Characterization of chitinases from the GH18 gene family in the myxomycete Physarum polycephalum

Stéphanie Renaud, Audrey Dussutour, Fayza Daboussi, Denis Pompon

BBA - General Subjects 1867 (2023) 130343

Contents lists available at ScienceDirect

BBA - General Subjects

journal homepage: www.elsevier.com/locate/bbagen

Characterization of chitinases from the GH18 gene family in the myxomycete *Physarum polycephalum*

Stéphanie Renaud [a], Audrey Dussutour [b], Fayza Daboussi [a], Denis Pompon [c,*]

[a] TWB, UMS INRAE / INSA / CNRS, Toulouse, France
[b] Centre de Recherche en Cognition Animale, UMR 5169 CNRS, Université Toulouse III, Toulouse, France
[c] Toulouse Biotechnology Institute, UMR CNRS / INRAE / INSA, Université de Toulouse, Toulouse, France

ARTICLE INFO

Keywords:
Glycoside hydrolase
Carbohydrate binding module
Chitinases
Physarum polycephalum
myxomycetes
GH18

ABSTRACT

Background: Physarum polycephalum is an unusual macroscopic myxomycete expressing a large range of glycosyl hydrolases. Among them, enzymes from the GH18 family can hydrolyze chitin, an important structural component of the cell walls in fungi and in the exoskeleton of insects and crustaceans.
Methods: Low stringency sequence signature search in transcriptomes was used to identify GH18 sequences related to chitinases. Identified sequences were expressed in *E. coli* and corresponding structures modelled. Synthetic substrates and in some cases colloidal chitin were used to characterize activities.
Results: Catalytically functional hits were sorted and their predicted structures compared. All share the TIM barrel structure of the GH18 chitinase catalytic domain, optionally fused to binding motifs, such as CBM50, CBM18, and CBM14, involved in sugar recognition. Assessment of the enzymatic activities following deletion of the C-terminal CBM14 domain of the most active clone evidenced a significant contribution of this extension to the chitinase activity. A classification based on module organization, functional and structural criteria of characterized enzymes was proposed.
Conclusions: Physarum polycephalum sequences encompassing a chitinase like GH18 signature share a modular structure involving a structurally conserved catalytic TIM barrels decorated or not by a chitin insertion domain and optionally surrounded by additional sugar binding domains. One of them plays a clear role in enhancing activities toward natural chitin.
General significance: Myxomycete enzymes are currently poorly characterized and constitute a potential source for new catalysts. Among them glycosyl hydrolases have a strong potential for valorization of industrial waste as well as in therapeutic field.

1. Introduction

Chitin is an important structural component of cell walls in fungi, including yeasts, and in the exoskeleton of insects and crustaceans. It is the second most abundant renewable polymer on earth after cellulose. The annual production and steady-state amount of this polymer are on the order of 10^{10} to 10^{11} tons per year [11,17]. A large part arises from fresh water and marine plankton [61–63]. This large and insoluble structural polysaccharide consists of repeats of *N*-acetyl-D-glucosamine (GlcNAc) monomers connected by β(1–4) bonds. There are three forms of chitin: α chitin (the most spread-out form), which consists of anti-parallel GlcNAc chains [61,62]; β-chitin, in which the chains are parallel; and γ chitin, which consists of a mix of parallel and antiparallel chains.

Despite its abundance, chitin does not accumulate in the environment due to the action of chitinolytic enzymes, which are widely spread throughout many organisms, even those that do not synthesize chitin themselves. Chitin depolymerization into GlcNAc requires the coordinated action of endochitinase, exochitinase, and chitobiosidase, which are mainly classified in the GH18 and GH19 families of the CAZY database [43]. In addition, *N*-acetylglucosaminidases from the GH20 family allow degradation of formed *N*, *N*′-diacetylchitobiose. They contribute to providing food sources, morphogenesis in fungi and insects, and defense against fungal diseases in plants. Over the last three decades, interest in chitinases has increased in various fields of biotechnology, such as the treatment of chitinous waste from the

* Corresponding author.
E-mail address: dpompon@insa-toulouse.fr (D. Pompon).

https://doi.org/10.1016/j.bbagen.2023.130343
Received 30 November 2022; Received in revised form 19 February 2023; Accepted 2 March 2023
Available online 17 March 2023

seafood industry, the biocontrol of phytopathogenic fungi, and their use as a biopesticide [43]. More recently, medical applications have involved the use of chitooligomers as antitumor agents, dietary fiber, antihypertensives agents, and asthma-related treatments [16,25,43].

Chitinases of the GH18 family are widely distributed in many organisms. GH18s are composed of several structural/functional elements: one or more catalytic domains, a signal peptide or transmembrane segment, zero to seven carbohydrate binding domains (CBMs), a serine/threonine domain, one or no type III fibronectin-like domain. Some CBMs are preferentially found in certain organisms, such as CBM5 and CBM12 in bacteria or fungi. CBM50 and CBM18 are generally found in the chitinases of subgroup C of fungi. CBM14 is thought to be found in the chitinases of mammals and insects. The functions of GH18 chitinases are very diverse: cell wall degradation, nutrition, autolysis, invasion and pathogenesis, moulting, immune defense, tolerance to infectious stress, nodulation.[8].

Because of their considerable potential for multiple applications, original chitinases have been extensively sought from cultivable and non-cultivable organisms [9,19,22,25,31,32,39,55,60]. However, few chitinases are currently commercially available and used for industrial applications, mainly due to the lack of efficient expression systems and simple purification methods [39,48]. The limited efficiency of heterologous chitinase expression in microbial hosts, such as *Escherichia coli* and *Pichia pastoris* [19,47,55,58] may result from the antibacterial and antifungal activities of chitinase, which interfere with the growth of the expression host [6].

In this context, we focused on the GH18 protein content of an unusual cell wall-free organism, the acellular slime mold *Physarum polycephalum* (PP). This macroscopic myxomycete is related to amoeba and phylogenetically distant from any major model organism. PP mainly feeds by phagocytosis of bacteria and fungi. However, it is able to degrade alternative energy sources, such as oat flakes, which requires the secretion of glycolytic enzymes for degradation [18,26,57]. PP features a particularly rich genome, exceeding 100 M-bases in size, that profits from its being at an evolutionary crossroad, bringing together genes frequently found in bacteria, plants, and animals [35,45]. We used a GH18 family signature encompassing the catalytic site in a low stringency search to scan a publicly available PP transcriptome [15]. This paper illustrates the presence of a large set of multidomain chitinases in PP that combine a well-conserved catalytic core with extra N-and C-terminal domains that modulate their cellular location and enzymatic properties.

2. Materials and methods

2.1. Sequence identification of GH18 chitinases

The gene search was performed at low stringency by scanning the pool of currently reported PP transcript sequences (NCBI-PRJNA295269 and footnote 1) *using* the nine amino-acid GH18 PROSITE signature of active site. The used proprietary search algorithm (MagicGene) surpassed the performance of classical PROSITE tools by allowing the identification of highly degenerate signatures up to the randomness statistical limit. The coding sequences of the resulting hits were extended on available transcripts upstream to the last available in-frame ATG codon preceding a stop codon and downstream to the first in-frame stop codon. Sequences without any detectable similarity with known chitinases or that were too short to encode most of the characteristic TIM Barrel structure of chitinases were discarded, as they were likely truncated in the library. This method proved to be more efficient and sensitive than a direct BLAST search in detecting putative chitinases poorly related to characterized references. Finally, selected hits were used to search available PP genomic sequences. However, the limited coverage and quality of available PP genome sequences did not allow us to significantly increase the set of hits or unambiguously extend the likely truncated transcripts.

2.2. Structural predictions of GH18 chitinases

3D-structural predictions were performed on hits using the TrRosetta algorithm [59] without the use of a specific template. Among the results, only structural predictions ranking with high or very high confidence levels (TM-score > 0.6) were retained for consideration. 3D-structures were visualized using chimeraX [40]. The search for putative signal peptides dedicated to protein exportation was conducted using SignalP-5.0 server [4]. Sequence segments corresponding to each structurally predicted domain were submitted separately to NCBI BlastP search to identify potential related domains of known function. The conserved domain database of the NCBI [33], SMART [30], HMMER [41] combined with Pfam [37] were used. For unsolved domain identification, protein sequences were submitted on Google Colaboratory to the DeepMind AlphaFold2 program, a deep-learning based algorithm, for structural prediction [23]. Resulting 3D-structures predicted with a high level of estimated confidence (pLDDT >70) were submitted to PDBefold [27] to identify structural domains with related folds and establish potential functions. The structures based alignment of the GH18 PP sequences was established using the T-coffee (Expresso) software [38].

2.3. Construction of a phylogenetic tree

Sequence alignments were performed using the Muscle algorithm. Phylogenic tree was established using the BioNJ software (datatype: protein, number of bootstraps: 1000, substitution model: John-Taylor-Thornton matrix or Dayhoff PAM matrix: the same results were obtained with both matrixes based on Muscle sequence alignments. Drawtree software was used as the tree viewer [10]. Phylogenetic trees in supplementary data have been built using IQtree web server and ITol web server (using 1000 bootstraps) [29,52] from Muscle alignments using the MPI bioinformatics toolkit [13].

2.4. Plasmid construction for Tr_10148PelB and Tr_10148PelB_DC expression

Plasmids expressing Tr_10148PelB and Tr_10148PelB_DC were constructed using the In-Fusion HD Cloning Plus kit (TakaraBio). Plasmid pET-29b(+) was prepared by restriction with *Nde*I and *Xho*I. The fragments cloned in the pET-29b (+) were amplified using the following matrix: pET-29b(+)-Tr_10,148 (built by Twist biosciences) and the Primers OBRT0062/OBRT0072 and OBRT0071/OBRT0072. These primers were used to add PelB sequence and to truncate the C-terminal domain to obtain plasmids expressing Tr_10148PelB and Tr_10148PelB_DC, respectively. The DNA sequence of Tr_10,148 and the primers are listed in supplementary Table S5. Both plasmids were validated by sequencing before being transformed into *E. coli* DE3 strain to express proteins.

2.5. Expression of PP GH18 chitinases

GH18 cDNA sequences encoding putative proteins were tagged with a 6-His-encoding sequence at their C-termini and synthesized by Twist Biosciences using the *E.coli* code bias and inserted into the expression vector pET-29b(+) (Novagen) between the *Nde*I and *Xho*I sites. Plasmids were transformed into an *E. coli* DE3 strain ("T7 Express"- Reference C2566 or "T7 express lysY"-reference C3010, from NEB) following the supplier's protocol. After transformation of sequence checked plasmids, LB medium (5 ml) containing 50 μg/ml kanamycin was inoculated with the strains and the strains were cultivated overnight at 37 °C and 150 rpm shaking (orbital 50 mm). These precultures were used to inoculate 50 ml of LB medium in 250 ml baffled erlenmeyers flasks. Cultures were incubated at 37 °C and 150 rpm shaking (orbital 50 mm). IPTG (0.1 or 1 mM) was added to induce expression when the OD_{600nm} reached 0.6 to 0.8. Different incubation *temperatures (16 °C or 37 °C) under* the same stirring conditions were used depending on the assay. An aliquot was

collected for each culture and pelleted. Cells were resuspended in 60 µl water before the addition of 20 µL 4× Laemmli buffer containing 1.4 M 2-mercaptoethanol and incubation for 10 min at 95 °C. Expressed proteins were monitored by SDS-PAGE using stain free gel imaging (Biorad). For enzyme production, cultures were harvested by centrifugation (4500 x*g*, 10 min, 4 °C) and suspended in 1× PBS (Abcam). Cells were disrupted by sonication with a microprobe (power 40%, 3 min, pulser 50%) and centrifuged at 15,000 *g* for 10 min at 4 °C to recover the supernatant, considered to be the soluble extract. Talon resin (HIS SPINTRAP TALON Ref: 29–0005-93 from Cytiva) was used to concentrate the proteins of the four candidates featuring the highest activity in cell extracts. For that, the supplier's recommended centrifugation protocol was applied by using 50 mM sodium phosphate, 300 mM NaCl, pH 7.4 as binding buffer, 50 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, pH 7.4 as wash buffer, and twice 200 µl of 50 mM sodium phosphate, 300 mM NaCl, 150 mM imidazole, pH 7.4 as elution buffer. Then, a concentration and buffer exchange with PBS were operated with Amicon centrifugal filter unit with a 10KDa cut off. The eluted proteins of interest were concentrated by a factor 4 and the final imidazole concentration estimated <1 mM.

2.6. Measurement of GH18 chitinases activity

The "fluorescent chitinases assay kit" (Sigma-Aldrich- CS1030) was used to measure GH18 activity. The assays were conducted as mentioned by the supplier with minor modifications. The assay buffer was substituted by McIlvaine's buffer pH 5.2 [20,36,54]. 4-methyl umbelliferone (0–100 ng) was used as a standard. One unit is defined as one µmole of 4-methylumbelliferone liberated per min with the substrate 4MU-$(NAcGlc)_3$. To determine the total activity in *E. coli* after overexpression, the activity was measured in the supernatant of the culture and the total activity was calculated using the volume of supernatant. The activity was also measured in soluble extract and the total activity calculated similarly. Then total activity in supernatant was reported as the sum of total activity in supernatant and in soluble extract. When needed, extracts were concentrated and buffer exchanged using Amicon units (Cut Off 10KDa).

2.7. Test of colloidal chitin degradation

Ten milliliters 37% hydrochloride acid was added to 1 g shrimp shell chitin (Sigma Aldrich-C9752) and mixed by magnetic stirring for 1 to 2 h before the very slow addition of 50 mL water. Colloidal chitin particles were collected by centrifugation (8000 x*g*, 20 min, 15 °C) and washed repeatedly with deionized water until the pH reached approximately 5. Finally, the pellet was suspended in deionized water to a final concentration of 50 mg/ml and autoclaved (20 min 121C). Colloidal chitin plates were prepared by mixing 10 ml 1% agarose, 8.4 mL 2× McIlvaine's buffer pH 5.2, and 1.6 ml 50 mg/ml colloidal chitin. Supernatants to be tested for chitinase hydrolysis were deposited on the plate surface. For volumes larger than 5 µl, deposition was performed in steps, with a drying time between each application. Commercial chitinases (Sigma-Aldrich) from *Trichoderma viride* or *Streptomyces griseus* were used as positive controls.

3. Results

3.1. Identification of encoded proteins exhibiting a GH18 signature in the PP transcriptome[1]

Available PP transcriptomes were scanned at low stringency using the GH18 PROSITE signature [LIVMFY]-[DN]-G-[LIVMF]-[DN]-[LIVMF]-[DN]-x-E. Translated sequence segments matching the signature were extended and filtered as described in Materials and Methods. This resulted in 19 distinct protein sequences (Supplementary Table S1) exhibiting the signature and matching at least one chitinase within the best 50 scores of a validation BLAST search. The best similarity (35 to 40%) between selected PP sequences and a member of the NCBI database was found for GH proteins (accession number: AYV84706.1; AYV82936.1; AYV76418.1) assigned to giant viruses of the *Hyperionvirus* and *Terrestrivirus* families. Giant viruses have been shown to act as transferable genetic elements between amoeba through phagocytosis [1,42,46]. A search to match the identified transcriptome sequences with published versions of the PP genome contigs [45] resulted in only a limited number of short matching genomic fragments. Such a result is consistent with the frequent presence of genomic introns that interrupt ORF frames and the relatively poor quality and incomplete nature of the available PP genome sequence. Alignment of selected PP GH18 sequences showed that all shared a well conserved common segment exhibiting high similarity with a part of the catalytic TIM barrel found in known GH18 chitinases (supplementary data: Fig. S2 and S13). Outside of this conserved sequence segment, similarity between the hits dropped off or disappeared within the N- and C-terminal extensions of various sizes.

3.2. 3D-structural predictions and sequence classification

The 19 selected sequences were run through the TrRosetta algorithm [59], resulting in 3D structural predictions rated with very high or high confidence levels (TM-score > 0.6) (Supplementary Fig. S1). Depending on the case, one to at least three independent folding domains were predicted for each sequence. All predicted PP structures included a similar $(\alpha/\beta)_8$ TIM barrel (in some cases incomplete) exhibiting a large central channel surrounded by the expected acidic catalytic residues of the search signature (Figs. 1, 2 and supplementary Fig. S2). Structural predictions revealed that five (Tr_09449, Tr_13,162, Tr_09402, Tr_08219, Tr_17,144) among selected hits lack parts of the barrel structure likely resulting from sequence truncation at mRNA levels or sequencing. The limits of the incomplete TIM barrels were confirmed on the basis of the structural alignment of Fig. 1. A search against the genome did not allow us to unambiguously assign the missing sequences. Interestingly, the presence of a hydrophobic N-terminal helix was predicted for all sequences, excepted for the four that were already expected to be truncated on the N-terminal side: Tr_09449, Tr_13,162, Tr_09402, Tr_08219. However, predicted structures evidenced that these helixes can form isolated domain as in Tr_29,722 or Tr_14,380 but also be part of domain interface as in Tr_13,810 or Tr_07972 (Supplementary Fig. S3). They thus could play different roles. Predictions (Supplementary Table S2) obtained with the "signal P" server [4] suggested that most could constitutes an extracellular export signal sequence, excepted for sequence Tr_29,722, which was better predicted as a lysosome localization signal based on Deeploc 1.0 software [3].

A well-folded N-terminal domain between the putative export signal peptide and the TIM barrel structure was predicted for three sequences (Tr_08891, Tr_29,722, and Tr_08969). Two contained an already known structure: for Tr_08969, a double LysM pattern (also known as carbohydrate binding module (CBM) 50), and for Tr_08891, a double chitin-binding domain type1 (ChitBD1) pattern (also known as CBM18), as identified using the conserved domain database (CDD) from NCBI. The N-terminal domain of Tr_08891 featured 45% similarity with the corresponding region of the chitinase of *Fusarium bulbicola* and that of Tr_08969 showed 61% similarity with the peptidoglycan-binding domain (LysM) of *Eubacteriaceae bacterium* [33]. No known pattern was found for the third N-terminal domain of Tr_29,722. However, this domain was predicted by the TrRosetta and Alphafold algorithms to be a coiled coil domain attached by flexible polypeptide linkers composed of glycine and serine to the catalytic domain. Four of the selected GH18 hits (Tr_10,148, Tr_09446, Tr_07972, and Tr_07540) include various

[1] https://www.regulationsbiologie.ovgu.de/Downloads/Physarum+polycephalum+Genome+Resource.htm

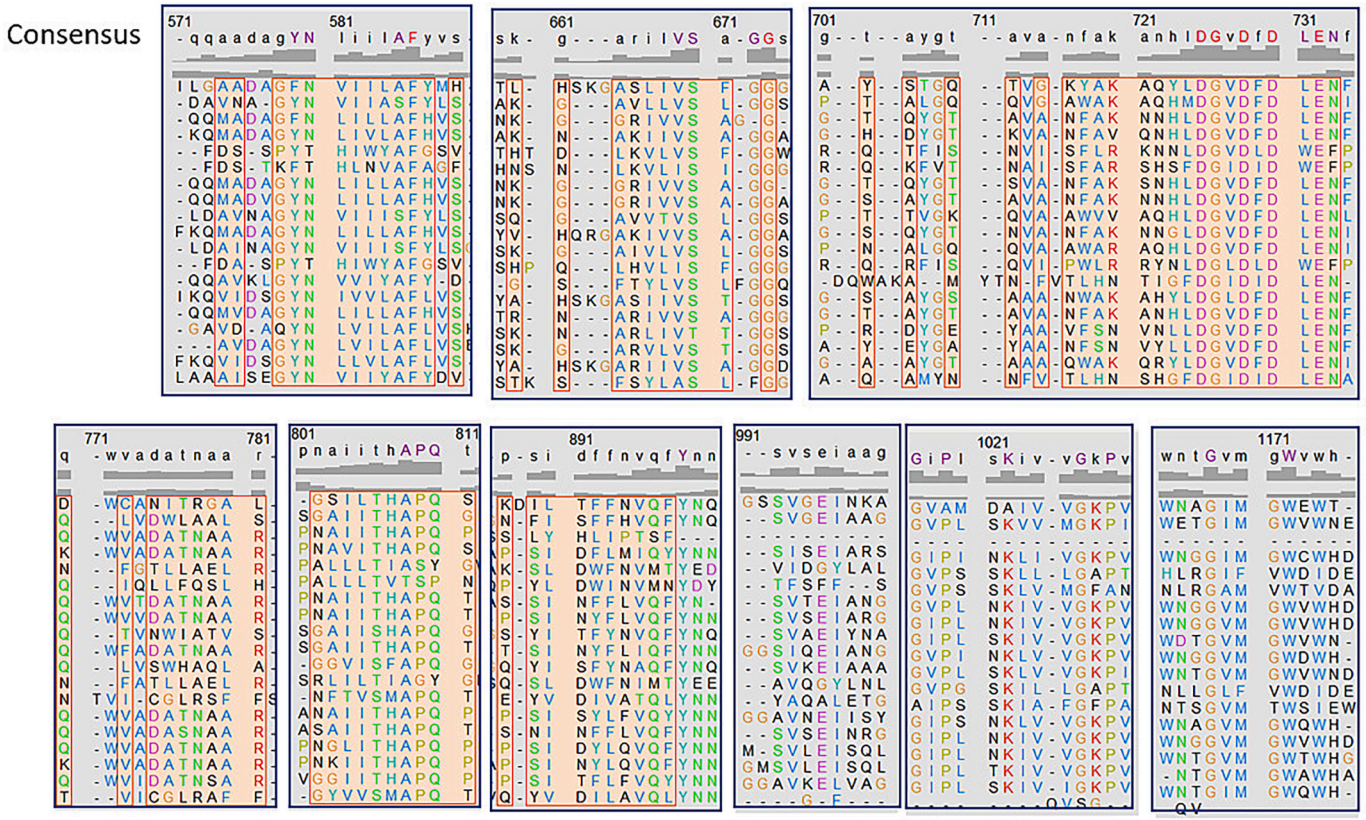

Fig. 1. Structurally conserved sequence segments of PP GH18. Alphafold predicted structures were RMSD minimalized using the structure of Tr-10,148 sequence as reference and the UCSF Chimera software. A structure-based sequence alignment was generated from matched structures using the software "Match- > Align" plug-in and a threshold distance of 5 Å. Resulting sequence segments structurally conserved between *Physarum* clones are listed on the figure and graphically represented on Fig. 2.

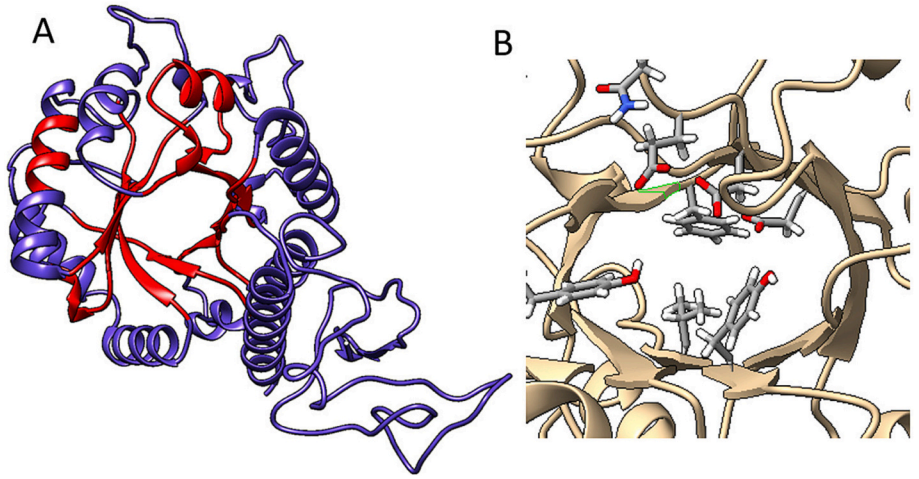

Fig. 2. Common structural feature of Physarum chitinases. A. Predicted structure of sequence Tr_10,148 exhibition the central hole of the Tim barrel. The red colored parts correspond to the structurally conserved elements (RMSD <5 Å) between all cloned PP GH18 as defined in Fig. 1. B. Close view of the conserved active site including the catalytic acidic residues (red colored). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

additional domains on the C-terminal side of the TIM barrel domain (Fig. 3A). A Blast search using the NCBI database revealed no similar sequences but a search by fold similarity with PDBefold using the structure predicted by the Alphafold program allowed us to determine that a part of this domain is likely chitin-binding domain type 2, also known as CBM14 (Supplementary Fig. S4). Matching of CBM14 in the CAZY database with the putative corresponding portion Tr_07972, Tr_09446 and Tr_10,148 was performed and confirmed the structural similarity. A chitin-binding function has been demonstrated in several cases for such module of approx. 70 residues that is found attached to several chitinase catalytic domains and non-catalytic proteins as well. Although the predicted C-terminal domain structure of Tr_09446 was of lower modelling quality, sequence alignments showed the characteristic expected pattern: six cysteines completed by several aromatic amino acids [14]. The C-terminal domain of Tr_07540 was probably incomplete however, the available part corresponded to the serine- and threonine-rich linker that is present in the three other C-terminal domains between the TIM barrel and CBM14.

Concerning the catalytic TIM barrel domain, all characterized PP GH18 shared a very similar core structure (RMSD lower than 5 Å)

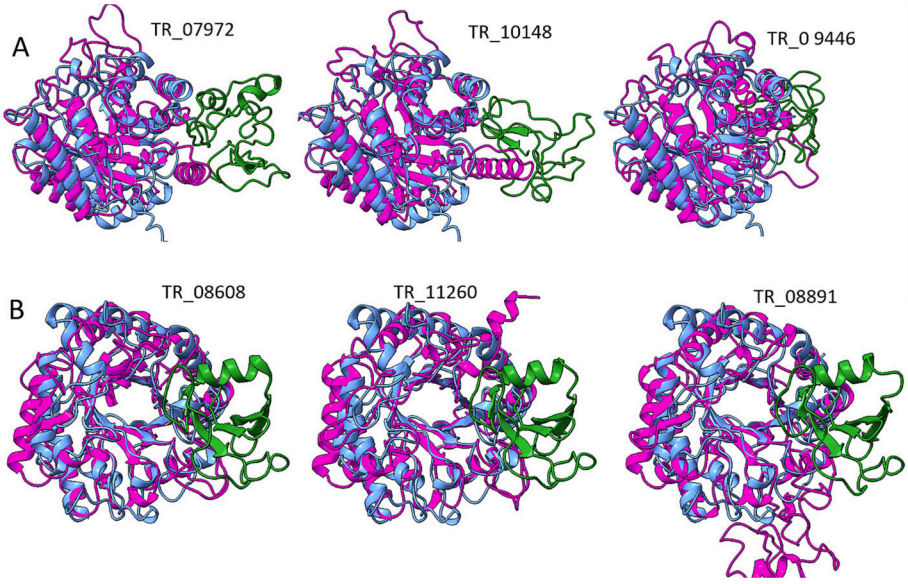

Fig. 3. Alphafold predicted structures of the additional C-terminal domains and CID loop insertions in the catalytic domains of PP-chitinases. Predicted structure for clone Tr_14,380 (in blue) was used as structural reference as its structure does not contain additional domain inserted in the TIM barrel or N- or C-extension. Top: chitinases (in pink colour) featuring a CBM14 like additional C-terminal chitin binding domain (green colored). Bottom additional CID domain (green colored) forming a subdomain inserted in the catalytic TIM barrel structures (pink colored). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

forming a large central channel where acidic catalytic residues are located (Fig. 2A). However, similar large extra-loops were identified in the Tr_08891, Tr_08608, and Tr_11,260 structures (Fig. 3B). This insertion, which occurred between the 7th and the 8th β-strand of the TIM barrel, was similar to the previously described (α + β) Chitin Insertion Domain (CID) present in subfamily A of some bacterial chitinases for which a classical model is the smChiA chitinase from *S. marcescens*. Tr_08608 predicted structure and X-ray structure PDB-7FD6 of smchiA chitinase exhibited a RMSD of 1.04 Å for AA-158-564. Such a very low value signed almost identical structures both for the TIM and CID parts. The extra-loop in the Tr_08608 and Tr_11,260 core sequences exhibited also 65% and 71% similarities, respectively, with lectin Ym1 (also named chitinase-like 3) from *Mus musculus*. Similarity with the same lectin module in Tr_08891 was lower (48%). The TIM barrel of Tr_12,536 that does not contain CID insertion was found more sequence related to bacterial enzymes that the previous ones and was particularly structurally related to structure PDB-3OAR of *Yersinia entomophaga* (RMSD of 1.27 Å for 78 of 269 atom pairs). Phylogenetic, modular structure and structural relationships of PP GH18 chitinases were summarized in Table 1.

The described modular structures were classified considering only the 13 apparently complete sequences (Fig. 4 and Table 1), resulting in five groups: group 1, containing only the catalytic domain barrel (CaD); group 2, including an additional C-terminal domain; group 3, including an N-terminal and catalytic domain; group 4, containing the catalytic domain, including an extra-loop; and group 5, including an additional N-

Table 1
Summary of the classification of chitinase like GH18 from *Physarum polycephalum*.

Group[1]	Clusters[2]	Tr_ Id	N-term. helix	N-term. CBM	Catalytic domain	C-term CBM	Activity[3] **Nat. lysY**	Classification	Related species
1A	C2	13,533	Yes	No	TB	No	**yes /yes**	See [5]	Giant virus
	C2	12,881	Yes	No	TB	No	ND /ND	See [5]	Giant virus
	C2	14,380	Yes	No	TB	No	ND/ yes	See [5]	Giant virus
	C2	13,810	Yes	No	TB	No	**yes/yes**	See [5]	Giant virus
1B	C5	12,536	Yes	No	TB	No	ND/yes	Bact. B type[6]	Bacteria
2	C3	09446	Yes	No	TB	CBM14	yes/yes	See [5]	Giant virus
	C3	10,148	Yes	No	TB	CBM14	**yes/yes**	See [5]	Giant virus
	C3	07972	Yes	No	TB	CBM14	**yes/yes**	See [5]	Giant virus
	C3	07540	Yes	No	TB	Trc CBM	ND/yes	See [5]	Giant virus
3	NC	29,722	Yes	Yes UNK	TB	No	yes/yes	See [5]	Giant virus
	C1	08969	Yes	2× CBM50	TB	No	**yes/yes**	See [5]	Giant virus
4	C4	08608	Yes	No	TB+ CID	No	**yes/yes**	See [4]	Plant
	C4	11,260	Yes	No	TB+ CID	No	**yes**/ND	See [4]	Plant
5		08891	Yes	2× CBM18	TB+ CID	No	ND/ND	See [4]	Plant parasite
6	C1	08219	No	No	trc TB	ND	ND		Giant virus
	C1	09402	No	No	trc. TB	No	ND		Giant virus
	C1	13,162	No	No	trc. TB	No	ND		Giant virus
	C1	09449	No	No	trc. TB	No	ND		Giant virus
	ND	17,144	Yes	No	C-trc. TB	No	ND		ND

[1]Groups as defined in Fig. 4. [2]Clusters from phylogenetic trees of supplementary figs. S5 A, B, C. [3]Activities as defined in Figs. S7 A and B toward synthetic substrates (Nat: T7 express cells, lysY: T7 express lysY), ND: not detected, yes: active (in bold higher values)). [4]Plant class V according to phylogenetic analysis but structurally related to CID (α + β) containing bacterial TIM barrels of subfamily A (structural RMSD of 1.04 Å (Tr_08608) with *Serratia marcescens* SmChia X-ray structure PDB-7FD6 AA 158–564). [5]Phylogenetically related to the independent branch of giant virus but structurally related to Plant and bacterial chitinases more or less decorated by N-or C-terminal CBMs.[6]Phylogenetically and structurally related for the TIM barrel part to bacterial chitinases (structural RMSD of 1.27 Å (78/269 pairs) to PDB-3OAR of *Yersinia entomophaga*). NC: sequence rejected during phylogenetic tree construction. CBM: carbohydrate binding module, TB: TIM barrel, TB + CID: TB containing a chitin insertion domain. N-trc or C-trc: N or C-terminal truncations. Trc-CBM: truncated CBM. Plant parasite refers to nematodes of *Heterodera* types and Giant virus to virus related to *Terrestrivirus*.

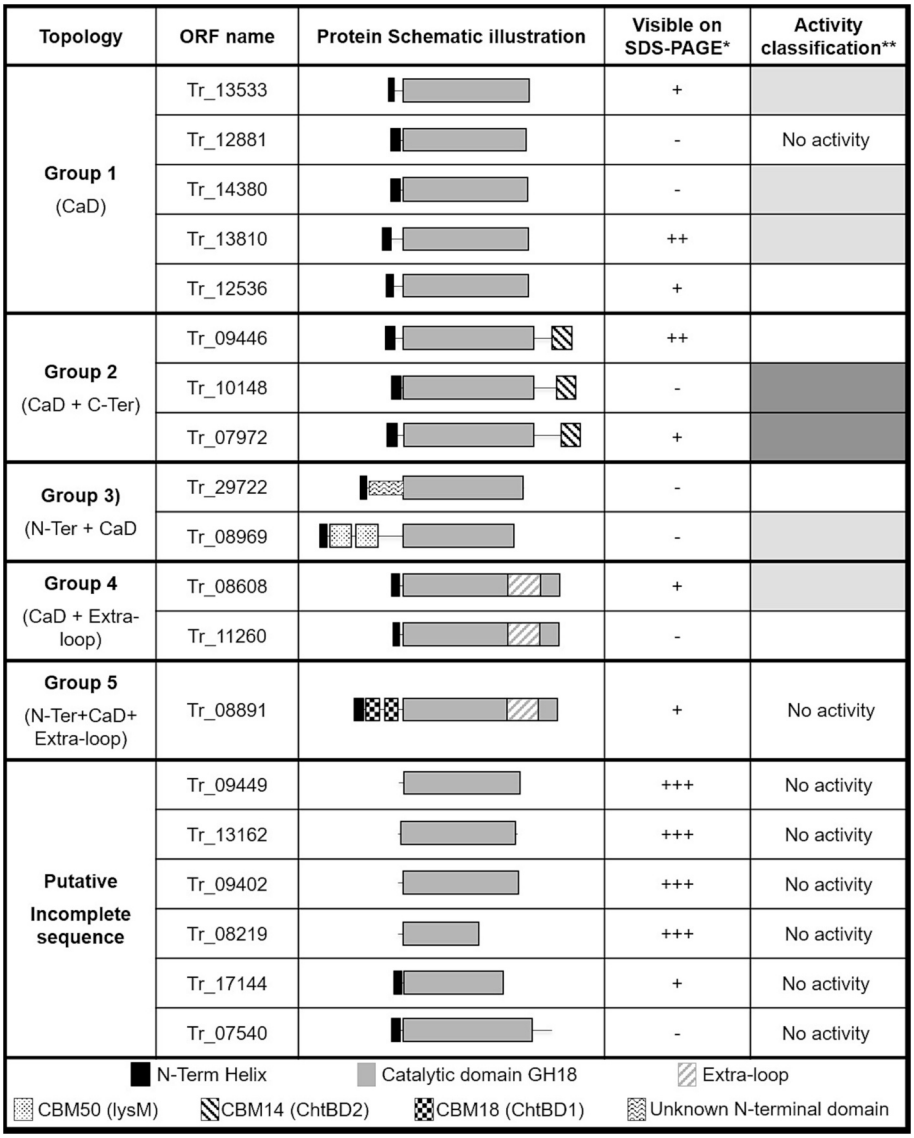

Fig. 4. PP GH18 classification based on domain topologies. Domain identification was based on predicted 3D-structures (TrRosetta and Alphafold). GH18-containing sequences with the same profile were gathered. Likely sequences featuring apparently incomplete Tim barrels were are grouped at the bottom of the table. *This column indicates that additional bands of the expected molecular weight were detectable (+, ++, +++) or not (−) in total extracts of "T7 express" transformed cells following induction with 1 mM IPTG for 2 h at 37 °C. Results were based on visual inspection of SDS-PAGE (supplementary, Fig. S6).**Expressed sequences were classified by relative activities: highest (dark grey), low but reliably detected (light grey), and negative or close to background (white) following measurements in soluble extracts prepared from cells expressing the sequences cultivated under various conditions (Supplementary Fig. S8). CaD: catalytic domain; C-ter: C-terminal domain; N-ter: N-terminal domain.

terminal domain relative to group 4. The relations between this proposed classification and the sequence conservation were evaluated. The catalytic domain sequences, including extra-loop (when present) were aligned and trees constructed as mentioned in the experimental section. The proteins belonging to the same classes in Fig. 4 also clustered in the same group of sequence similarity (Fig. 5). Only clone Tr_12,536 escape to the rule, suggesting that proteins of group 1 may potentially have different evolutive origins. These findings were confirmed by extending sequence classification to a global phylogenetic tree involving others species (Supplementary Figs. S5 A, B, C). Interestingly, all sequences lacking N-terminal helixes and considered as N-truncated belong to sequence cluster C1 (Fig. S5 C). Provided to exclude a clustering artifact, this may question a potential role of their truncated TIM barrel. The truncated sequence Tr_17,144 that was hardly classified could be a subpart of group 1 with Tr_12,536. Together, results suggested a possible evolutionary relationship between the capture of extra C- and N- domains surrounding the catalytic domain and the evolution of the catalytic domain itself. Most of the PP GH18 (Tr_12,881, Tr_14,380, Tr_13,533, Tr_13,810, Tr_09449, Tr_29,722, Tr_13,162, Tr_08219, Tr_08969, Tr_09402, Tr_10,148, Tr_07972, Tr_09446, Tr_07540) belonged to a unique part of the phylogenetic tree related to giant virus. This part is also related to bacterium of the genus *Francisella* and *Pseudoalteromonas*. A second branch including Tr_12,536 and possibly Tr_17,144 also hold GH18 from fungi (*Trichoderma*, *Beauveria*) and from bacteria (*Vibrio*, *Bacillus*, *Chromobacterium*). PP GH18 that included an extra-loop in the catalytic domain were in contrast related to chitinases from plants like *Medicago, cycas, Arabidopsis, nicotiana* (Tr_11,260 and Tr_08608) or from parasitic plant organisms like *Leishmania* and *Heterodera* (Tr_08891).

3.3. Expression and characterization of selected PP GH18 members in *E. coli*

The ORFs encoding the identified GH18 (full and truncated) were recoded using the standard *E. coli* codon bias, synthesized, and cloned into expression vectors. Bacteria-mediated expression in total and soluble extracts was analyzed by SDS-PAGE following IPTG induction. Additional protein bands at the expected size (Supplementary Table S3) were visualized in total extracts for a portion of the candidates (supplementary Fig. S6). By contrast, such additional bands were not identified in soluble extracts, suggesting that a large part of expressed proteins can be trapped into inclusion bodies.

Chitinase activities were tested on soluble extracts using model substrates: 4-methylumbelliferyl *N,N′*-diacetyl-β-D-chitobioside, a substrate suitable for fluorescent detection of chitobiosidase activity, 4-methylumbelliferyl *N*-acetyl-β-D-glucosaminide to monitor of β-*N*-

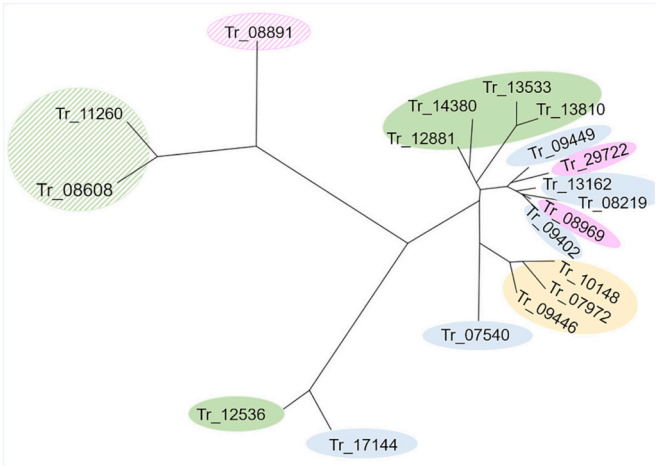

Fig. 5. Phylogenetic tree for the isolated catalytic domain of the GH18 family of *P. polycephalum*.
Sequences of the catalytic Tim barrel domains of PP GH18 were aligned using the Muscle algorithm. A phylogenetic tree was created using BioNJ and Drawtree software [10]. The colors and textures refer to the classification mentioned in Fig. 4. Green, yellow, pink, hatched green, and hatched pink represents groups 1 to 5, respectively. Blue arrays represent sequences featuring incomplete Tin barrel domains. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

acetylglucosaminidase activity, and 4-methylumbelliferyl β-D-*N*,*N*′,*N*′′-triacetylchitotriose (4MU-$(GlcNAc)_3$) for the detection of endochitinase activity. However, only the latter substrate allowed detection of significant activities, indicating that expressed proteins exhibited endochitinase but not detectable exochitinase activities. Prepared soluble extracts from transformants in strain *E. coli* T7 express, and in the corresponding lysY cells were cultivated and activities tested as described in Supplementary Fig. S7 legends. The best repressed lysY cells were tested to try limiting the *E. coli* growth inhibition (even without induction) observed with some of the GH18 candidates, inhibition that could have influenced activity ranking. In later experiments, constructs encompassing a lysS export sequence were also used to try limiting intracellular GH18 accumulation and toxicity. Based on the four expression conditions tested, activities of the described GH18 were ranked (Supplementary Fig. S7 and Table 1): Tr_10,148 and Tr_07972 performed the best; Tr_13,533, Tr_14,380, Tr_13,810, Tr_08969, and Tr_08608 showed lower activity; whereas the activity of Tr_12,881, Tr_12,536, Tr_09446, Tr_07540, Tr_29,722, Tr_11,260, Tr_08891, and Tr_17,144 were close to the limit of detection and found only under some of the conditions tested. Clones Tr_09449, Tr_13,162, Tr_09402, and Tr_08219, suspected to be truncated, exhibited no activity. The catalytic activities of Tr_10,148, Tr_07972, Tr_13,810, and Tr_08969 were also evaluated on colloidal chitin (Supplementary Fig. S8) using the formation of hydrolysis halos on agarose plates [44,49]. The results confirmed that the Tr_10,148, Tr_07972, and Tr_13,810 proteins efficiently hydrolyzed colloidal chitin but Tr_08969 did not exhibited detectable activity in such test. However, considering the limited sensitivity of halo formation test and the use of unpurified extracts, only positive results can be considered as robust. Attempts were performed to purify some of the expressed chitinases using affinity tags (Supplementary Fig. S12). However, such approach was found insufficient to reach a level of purity usable for more precise catalytic constant characterizations.

3.4. Analysis of the roles of the N- and C-terminal extensions surrounding the core domain

Functional contributions of the N- and C-terminal extensions present in the most active clone, Tr_10,148 were evaluated (Supplementary Fig. S9). A putative N-terminal export sequence signal is present on this chitinase but eukaryotic export signal might be poorly recognized and processed in *E.coli* [24]. When not recognized, such hydrophobic extension could cause protein aggregation. To avoid such limitation, a construct in which the eukaryotic signal of Tr_10,148 was substituted by a standard export signal used in *E. coli*: pelB, was designed (Supplementary Table S5). The limits of the sequence substitution was defined based on signalP 5.0 software predictions [4]. The activities of resulting constructs (Fig. 6) measured using similarly concentrated culture supernatants were higher for chitinase Tr_10148PelB than for Tr_10,148 (432 and 74 nmol/min/protein respectively), suggesting a better protein export with pelB fusions. However, activity level released in the culture medium still represented a small fraction (4%) of activity in cell lysates, indicating improved but still insufficient export. Concerning the CBM14-like C-terminal domain of Tr_10148PelB, corresponding amino acids 325 to 407 were deleted, giving construct Tr_10148PelB_DC. This resulted in an approximately two-fold decrease of activity toward the model substrate (Supplementary Fig. S10) suggesting that the C-terminal extension of Tr_10,148 (constituting a CBM14 domain) significantly enhanced activity by facilitating substrate recognition. Similar results were obtained with the C-terminal deletion on two versions of Tr_10,148 in which the 28-N-terminal residues were deleted and replaced or not by the export pelB sequence) (Supplementary Fig. S10). The activity of chitinase Tr_10148PelB and Tr_10148PelB_DC from concentrated supernatants of cultures on colloidal chitin were also tested (Supplementary Fig. S8). The same level of activity based on 4MU-$(GlcNAc)_3$ hydrolysis was deposited for all chitinases, including commercial ones from *S. griseus* and *T. viride*. Tr_10148PelB appeared to be at least as efficient as the *S. griseus* enzyme and exhibited better time dependence of hydrolysis than that the one of *T. viride*. The C-terminal domain truncation of Tr_10148PelB significantly decreased the efficiency of colloidal chitin hydrolysis, as already observed with the model substrate. This effect is not attributable to a loss of stability, as Tr_10148PelB, Tr_10148PelB_DC, and control activities evaluated similarly during prolongated incubation (Supplementary Fig. S11). Thus, the C-terminal domain in Tr_10,148 has a clear role in improving chitinase activity without affecting its stability.

4. Discussion

We identified nineteen PP sequences including a GH18 signature that are distantly sequence related to already characterized chitinases. Seven of them exhibited some endochitinase activities on the model substrate (4MU-$(GlcNAc)_3$ and for the most actives, hydrolysis of colloidal chitin was demonstrated. Eight others for which activities were close to background could correspond to partially truncated sequences. Sequence similarity analysis evidenced a modular structure, including various N- and C-terminal extensions surrounding the catalytic domain. Similarity of the core catalytic domain was maximal with related sequences found in giant viruses that infect amoebae. Most of the PP GH18 enzymes clustered together within a distinct phylogenetic branch (Supplementary Figs. S5 A, B, C) and share a highly structurally conserved TIM barrel catalytic structure. Notably, this branch included chitinase Chi23 from *P. aurantia,* an uncommon bacterial endochitinase [56]. The total number of GH18 chitinases found in PP was quite elevated for a chitin-free organisms, a feature also presents in fungi. This number was likely underestimated on the basis of transcriptome analysis as gene expression in PP is strongly dependent on the phase of development and on a range of environmental conditions. Attempts to correlate or extend the results from transcript screening with the available genomic sequences were mostly unsuccessful due to poor genome sequence quality, fragmentation and limited coverage.

A classification of PP GH18 chitinases was proposed mostly based on topological considerations (domain type and organization), sequence and structural conservations. The modular organization associating various carbohydrate binding domains to a common catalytic core could significantly contribute in diversifying activities. Some of the extensions

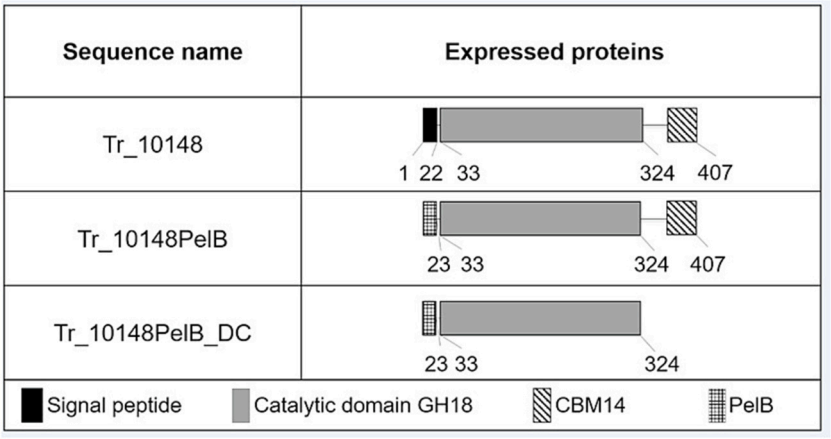

Fig. 6. Schematic structure of engineered chitinase derived from clone Tr_10,148.
The putative signal peptide (22 amino acids) of Tr_10,148 identified in the PP transcriptome, identified using Signal P 5.0, was replaced by the sequence of pelB, giving rise to Tr_10148PelB. Subsequent truncation of the C-terminal domain (amino acids 325 to 407) gave Tr_10148PelB_DC. Protein sequence of_ Tr_10148PelB and Tr_10148PelB_DC are available in supplementary Table S4.

show limited sequence similarity but strong predicted structure conservation with known domains involved in carbohydrate binding, such as CBM50 (LysM), which can bind (GlcNAc)-containing glycans [5], and CBM18 (ChitBD1) and CBM14 (chitBD2) for chitin binding. A BLAST search against bacteria, fungi, and amoebae taxids for sequences related to the PP lysM modules optimally matched with bacterial sequences. However, the presence of the four-cysteine residue motif in the lysM unit and the asparagine residue at position 33 or 34 from the domain start could favor a fungal origin [2]. This characteristic was also found in two plant chitinases (PrChiA and EaChiA) but not in bacteria [51]. LysM modules can exert strong antifungal activity, independently of the catalytic domain [51]. Interestingly, expression of Tr_08969 in *E. coli* had a strong negative impact on *E. coli* growth (data not shown). Bacteria do not contain chitin but some GlcNAc in peptidoglycans that can act as a target for lysM modules. Some chitinases can also have a lysozyme like effect [6], but this was not true for Tr_08969 under the tested conditions (data not shown).

Concerning the chitin binding modules, CBM18 is mainly associated with the GH19 chitinase catalytic domain and, exceptionally, with GH18 chitinase in plants [28,50]. This organization is also found in filamentous fungi (subgroup C1) as killer toxin-like chitinase [53]. CBM14 is also known as a peritrophin-A domain found in proteins strongly bound to the peritrophic matrix. Such matrices constitute acellular linings in the gut of most insects that is formed of proteins and sugar polymers, including chitin. The similarity between CBM14 domains is generally low but they remain identifiable due to the conserved positioning of the six cysteine residues and a high percentage of aromatic residues [14]. CBM14 is also found in pathogenic fungi as an effector to counteract host immunity during infection [21]. In addition to the N- and C- domain extensions, some of the characterized chitinases include a CID-like domain inserted within the catalytic TIM barrel domain. Such domain, that could be involved in chitin hydrolysis processivity, have already been observed in several chitinases: for example, CAD1 from *Ostrinia furnacalis* (5WUP), chitinase 40 from *Streptomyces thermoviolaceus* (4W5U), chitinase B from *Serratia marcescens* (7C34), and human chitotriosidase-1 (4WKA) [7,12,31,34].

5. Conclusions

PP expresses a large number of GH18 chitinases featuring large variations in their modular organization but surrounding a catalytic domain with a common and structurally conserved TIM barrel. Despite their modular and structural proximities with Plant and bacterial chitinases, the PP sequences mostly form an independent branch in phylogenetic trees much closer to giant virus than to any other characterized organisms. Such giant virus have hypothetical roles in genetic information transfers between myxomycetes. Furthermore, in the same organism, PP GH18 combines optional CID and several types of CBM found either in insects, fungi or bacteria but usually not in the same organism. Considering their N-terminal extension resembling export sequences, these enzymes could have extracellular location and be of interest in the development of biotechnical applications exploiting the PP or recombinant expression contexts. Tr_10,148 that contains a CBM14 domain, was characterized in more details and was found at least as efficient in hydrolyzing chitin than currently commercially available enzymes.

Author contributions

S. Renaud performed all experimental work and substantially contributed to data interpretation and preparation of the manuscript. A. Dussutour was involved as expert of the acellular slime mold *Physarum polycephalum*. F. Daboussi and D. Pompon acted as project supervisors, selected and interpreted the data, and wrote the manuscript.

Funding

This study was supported by funds received from Fonds Européen de Développement Régional (FEDER).

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Stephanie RENAUD reports financial support, administrative support, article publishing charges, and equipment, drugs, or supplies were provided by Fonds Européen de Développement Régional (FEDER).

Data availability

Data will be made available on request.

Acknowledgments

We acknowledge Sebastien Santini (CNRS/AMU IGS UMR7256) and the PACA Bioinfo platform (supported by IBISA) for the availability and management of the phylogeny.fr website used to build the phylogenetic tree of the catalytic domain of GH18 chitinases of PP.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbagen.2023.130343.

References

[1] S. Aherfi, P. Colson, B. La Scola, D. Raoult, Giant viruses of amoebas: an update, Front. Microbiol. 7 (2016), https://doi.org/10.3389/fmicb.2016.00349.

[2] G.B. Akcapinar, L. Kappel, O.U. Sezerman, V. Seidl-Seiboth, Molecular diversity of LysM carbohydrate-binding motifs in fungi, Curr. Genet. 61 (2015) 103–113, https://doi.org/10.1007/s00294-014-0471-9.

[3] J.J. Almagro Armenteros, C.K. Sønderby, S.K. Sønderby, H. Nielsen, O. Winther, DeepLoc: prediction of protein subcellular localization using deep learning, Bioinformatics 33 (2017) 3387–3395, https://doi.org/10.1093/bioinformatics/btx431.

[4] J.J. Almagro Armenteros, K.D. Tsirigos, C.K. Sønderby, T.N. Petersen, O. Winther, S. Brunak, et al., SignalP 5.0 improves signal peptide predictions using deep neural networks, Nat. Biotechnol. 37 (2019) 420–423, https://doi.org/10.1038/s41587-019-0036-z.

[5] G. Buist, A. Steen, J. Kok, O.P. Kuipers, LysM, a widely distributed protein motif for binding to (peptido)glycans, Mol. Microbiol. 68 (2008) 838–847, https://doi.org/10.1111/j.1365-2958.2008.06211.x.

[6] S. Carneiro de Medeiros, J. Edvar Monteiro Júnior, G. Weyne Passos Sales, T. Barbosa Grangeiro, N. Accioly Pinto Nogueira, Chitinases as antibacterial proteins: A systematic review, J. Young Pharm. 10 (2018) 144–148, https://doi.org/10.5530/jyp.2018.10.33.

[7] L. Chen, L. Zhu, J. Chen, W. Chen, X. Qian, Q. Yang, Crystal structure-guided design of berberine-based novel chitinase inhibitors, J. Enzyme Inhib. Med. Chem. 35 (2020) 1937–1943, https://doi.org/10.1080/14756366.2020.1837123.

[8] W. Chen, X. Jiang, Q. Yang, Glycoside hydrolase family 18 chitinases: the known and the unknown, Biotechnol. Adv. 43 (2020), 107553, https://doi.org/10.1016/j.biotechadv.2020.107553.

[9] J.-J. Deng, D. Shi, H.-H. Mao, Z.-W. Li, S. Liang, Y. Ke, et al., Heterologous expression and characterization of an antifungal chitinase (Chit46) from Trichoderma harzianum GIM 3.442 and its application in colloidal chitin conversion, Int. J. Biol. Macromol. 134 (2019) 113–121, https://doi.org/10.1016/j.ijbiomac.2019.04.177.

[10] A. Dereeper, V. Guignon, G. Blanc, S. Audic, S. Buffet, F. Chevenet, et al., Phylogeny.fr: robust phylogenetic analysis for the non-specialist, Nucleic Acids Res. 36 (2008) W465–W469, https://doi.org/10.1093/nar/gkn180.

[11] G.S. Dhillon, S. Kaur, S.K. Brar, M. Verma, Green synthesis approach: extraction of chitosan from fungus mycelia, Crit. Rev. Biotechnol. 33 (2013) 379–403, https://doi.org/10.3109/07388551.2012.717217.

[12] F. Fadel, Y. Zhao, R. Cachau, A. Cousido-Siah, F.X. Ruiz, K. Harlos, et al., New insights into the enzymatic mechanism of human chitotriosidase (CHIT1) catalytic domain by atomic resolution X-ray diffraction and hybrid QM/MM, Acta Crystallogr. D Biol. Crystallogr. 71 (2015) 1455–1470, https://doi.org/10.1107/S139900471500783X.

[13] F. Gabler, S.-Z. Nam, S. Till, M. Mirdita, M. Steinegger, J. Söding, et al., Protein sequence analysis using the MPI bioinformatics toolkit, Curr. Protoc. Bioinformatics 72 (2020), e108, https://doi.org/10.1002/cpbi.108.

[14] P.J. Gaines, S.J. Walmsley, N. Wisnewski, Cloning and characterization of five cDNAs encoding peritrophin-A domains from the cat flea, Ctenocephalides felis, Insect Biochem. Mol. Biol. 33 (2003) 1061–1073, https://doi.org/10.1016/s0965-1748(03)00096-1.

[15] G. Glöckner, W. Marwan, Transcriptome reprogramming during developmental switching in Physarum polycephalum involves extensive remodeling of intracellular signaling networks, Sci. Rep. 7 (2017) 12304, https://doi.org/10.1038/s41598-017-12250-5.

[16] E.Z. Gomaa, Microbial chitinases: properties, enhancement and potential applications, Protoplasma. (2021), https://doi.org/10.1007/s00709-021-01612-6.

[17] Graham W. Gooday, The ecology of chitin degradation, in: Advances in microbial ecology, 1990, https://doi.org/10.1007/978-1-4684-7612-5. Available at:. [Accessed May 28, 2021].

[18] A. Haars, C.H.R. McCullough, A. Hütterman, J. Dee, Regulation of proteolytic enzymes in axenically grown myxamoebae of Physarum polycephalum, Arch. Microbiol. (1978) 55–60.

[19] P. Han, C. Yang, X. Liang, L. Li, Identification and characterization of a novel chitinase with antifungal activity from “Baozhu” pear (Pyrus ussuriensis maxim.), Food Chem. 196 (2016) 808–814, https://doi.org/10.1016/j.foodchem.2015.10.006.

[20] C.E. Hollak, S. van Weely, M.H. van Oers, J.M. Aerts, Marked elevation of plasma chitotriosidase activity. A novel hallmark of Gaucher disease, J. Clin. Invest. 93 (1994) 1288–1292, https://doi.org/10.1172/JCI117084.

[21] N.K. Hurlburt, L.-H. Chen, I. Stergiopoulos, A.J. Fisher, Structure of the Cladosporium fulvum Avr4 effector in complex with (GlcNAc)6 reveals the ligand-binding mechanism and uncouples its intrinsic function from recognition by the Cf-4 resistance protein, PLoS Pathog. 14 (2018), e1007263, https://doi.org/10.1371/journal.ppat.1007263.

[22] T. Itoh, T. Hibi, Y. Fujii, I. Sugimoto, A. Fujiwara, F. Suzuki, et al., Cooperative degradation of chitin by extracellular and cell surface-expressed chitinases from Paenibacillus sp. strain FPU-7, Appl. Environ. Microbiol. 79 (2013) 7482–7490, https://doi.org/10.1128/AEM.02483-13.

[23] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, et al., Highly accurate protein structure prediction with AlphaFold, Nature 596 (2021) 583–589, https://doi.org/10.1038/s41586-021-03819-2.

[24] A. Karyolaimos, H. Ampah-Korsah, T. Hillenaar, A. Mestre Borras, K.M. Dolata, S. Sievers, et al., Enhancing recombinant protein yields in the *E. coli* periplasm by combining signal peptide and production rate screening, Front. Microbiol. 10 (2019) 1511, https://doi.org/10.3389/fmicb.2019.01511.

[25] P.E. Kidibule, P. Santos-Moriano, E. Jiménez-Ortega, M. Ramírez-Escudero, M. C. Limón, M. Remacha, et al., Use of chitin and chitosan to produce new chitooligosaccharides by chitinase Chit42: enzymatic activity and structural basis of protein specificity, Microb. Cell Factories 17 (2018) 47, https://doi.org/10.1186/s12934-018-0895-x.

[26] D.C. Kilpatrick, J.L. Stirling, Glycosidases from the culture medium of Physarum polycephalum, Biochem. J. 161 (1977) 149–157, https://doi.org/10.1042/bj1610149.

[27] E. Krissinel, K. Henrick, Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 2256–2268, https://doi.org/10.1107/S0907444904026460.

[28] Y. Kuba, T. Takashima, K. Uechi, T. Taira, Purification, cDNA cloning, and characterization of plant chitinase with a novel domain combination from lycophyte Selaginella doederleinii, Biosci. Biotechnol. Biochem. 82 (2018) 1742–1752, https://doi.org/10.1080/09168451.2018.1491285.

[29] I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation, Nucleic Acids Res. 49 (2021) W293–W296, https://doi.org/10.1093/nar/gkab301.

[30] I. Letunic, S. Khedkar, P. Bork, SMART: recent updates, new developments and status in 2020, Nucleic Acids Res. 49 (2021) D458–D460, https://doi.org/10.1093/nar/gkaa937.

[31] C. Liu, N. Shen, J. Wu, M. Jiang, S. Shi, J. Wang, et al., Cloning, expression and characterization of a chitinase from Paenibacillus chitinolyticus strain UMBR 0002, PeerJ 8 (2020), e8964, https://doi.org/10.7717/peerj.8964.

[32] N.H. Loc, N.D. Huy, H.T. Quang, T.T. Lan, T.T. Thu Ha, Characterisation and antifungal activity of extracellular chitinase from a biocontrol fungus, Trichoderma asperellum PQ34, Mycology 11 (2020) 38–48, https://doi.org/10.1080/21501203.2019.1703839.

[33] S. Lu, J. Wang, F. Chitsaz, M.K. Derbyshire, R.C. Geer, N.R. Gonzales, et al., CDD/SPARCLE: the conserved domain database in 2020, Nucleic Acids Res. 48 (2020) D265–D268, https://doi.org/10.1093/nar/gkz991.

[34] P.H. Malecki, M. Bejger, W. Rypniewski, C.E. Vorgias, The crystal structure of a streptomyces thermoviolaceus thermophilic Chitinase known for its refolding efficiency, Int. J. Mol. Sci. 21 (2020) 2892, https://doi.org/10.3390/ijms21082892.

[35] C.L. McGrath, L.A. Katz, Genome diversity in microbial eukaryotes, Trends Ecol. Evol. 19 (2004) 32–38, https://doi.org/10.1016/j.tree.2003.10.007.

[36] T.C. McIlvaine, A buffer solution for colorimetric comparison, J. Biol. Chem. 49 (1921) 183–186, https://doi.org/10.1016/S0021-9258(18)86000-8.

[37] J. Mistry, S. Chuguransky, L. Williams, M. Qureshi, G.A. Salazar, E.L. L. Sonnhammer, et al., Pfam: the protein families database in 2021, Nucleic Acids Res. 49 (2021) D412–D419, https://doi.org/10.1093/nar/gkaa913.

[38] C. Notredame, D.G. Higgins, J. Heringa, T-coffee: A novel method for fast and accurate multiple sequence alignment, J. Mol. Biol. 302 (2000) 205–217, https://doi.org/10.1006/jmbi.2000.4042.

[39] T.M. Pasin, T.B. de Oliveira, A.S.A. de Scarcella, M.L.T.M. de Polizeli, M.-E. Guazzaroni, Perspectives on expanding the repertoire of novel microbial chitinases for biological control, J. Agric. Food Chem. 69 (2021) 3284–3288, https://doi.org/10.1021/acs.jafc.1c00219.

[40] E.F. Pettersen, T.D. Goddard, C.C. Huang, E.C. Meng, G.S. Couch, T.I. Croll, et al., Structure visualization for researchers, educators, and developers, Protein Sci. 30 (2021) 70–82, https://doi.org/10.1002/pro.3943.

[41] S.C. Potter, A. Luciani, S.R. Eddy, Y. Park, R. Lopez, R.D. Finn, HMMER web server: 2018 update, Nucleic Acids Res. 46 (2018) W200–W204, https://doi.org/10.1093/nar/gky448.

[42] D. Raoult, M. Boyer, Amoebae as genitors and reservoirs of giant viruses, Intervirology 53 (2010) 321–329, https://doi.org/10.1159/000312917.

[43] A.S. Rathore, R.D. Gupta, Chitinases from bacteria to human: properties, applications, and future perspectives, Enzyme Res. 2015 (2015), 791907, https://doi.org/10.1155/2015/791907.

[44] R. Salas-Ovilla, D. Gálvez-López, A. Vázquez-Ovando, M. Salvador-Figueroa, R. Rosas-Quijano, Isolation and identification of marine strains of Stenotrophomona maltophilia with high chitinolytic activity, PeerJ 7 (2019), e6102, https://doi.org/10.7717/peerj.6102.

[45] P. Schaap, I. Barrantes, P. Minx, N. Sasaki, R.W. Anderson, M. Bénard, et al., The *Physarum polycephalum* genome reveals extensive use of prokaryotic two-component and metazoan-type tyrosine kinase signaling, Genome Biol. Evol. 8 (2016) 109–125, https://doi.org/10.1093/gbe/evv237.

[46] F. Schulz, L. Alteio, D. Goudeau, E.M. Ryan, F.B. Yu, R.R. Malmstrom, et al., Hidden diversity of soil giant viruses, Nat. Commun. 9 (2018) 4881, https://doi.org/10.1038/s41467-018-07335-2.

[47] W. Song, N. Zhang, M. Yang, Y. Zhou, N. He, G. Zhang, Multiple strategies to improve the yield of chitinase a from Bacillus licheniformis in Pichia pastoris to obtain plant growth enhancer and GlcNAc, Microb. Cell Factories 19 (2020) 181, https://doi.org/10.1186/s12934-020-01440-y.

[48] Y.M. Stoykov, A.I. Pavlov, A.I. Krastanov, Chitinase biotechnology: production, purification, and application, Eng. Life Sci. 15 (2015) 30–38, https://doi.org/10.1002/elsc.201400173.

[49] W. Suginta, P.A. Robertson, B. Austin, S.C. Fry, L.A. Fothergill-Gilmore, Chitinases from Vibrio: activity screening and purification of chiA from Vibrio carchariae, J. Appl. Microbiol. 89 (2000) 76–84, https://doi.org/10.1046/j.1365-2672.2000.01076.x.

[50] T. Taira, Structures and antifungal activity of plant chitinases, J. Appl. Glucosci. (2010) 167–176.

[51] T. Takashima, R. Sunagawa, K. Uechi, T. Taira, Antifungal activities of LysM-domain multimers and their fusion chitinases, Int. J. Biol. Macromol. 154 (2020) 1295–1302, https://doi.org/10.1016/j.ijbiomac.2019.11.005.
[52] J. Trifinopoulos, L.-T. Nguyen, A. von Haeseler, B.Q. Minh, W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis, Nucleic Acids Res. 44 (2016) W232–W235, https://doi.org/10.1093/nar/gkw256.
[53] G. Tzelepis, M. Karlsson, Killer toxin-like chitinases in filamentous fungi: structure, regulation and potential roles in fungal biology, Fungal Biol. Rev. 33 (2019) 123–132, https://doi.org/10.1016/j.fbr.2018.11.001.
[54] M. Uehara, E. Tabata, M. Okuda, Y. Maruyama, V. Matoska, P.O. Bauer, et al., Robust chitinolytic activity of crab-eating monkey (Macaca fascicularis) acidic chitinase under a broad pH and temperature range, Sci. Rep. 11 (2021) 15470, https://doi.org/10.1038/s41598-021-95010-w.
[55] S. Vincenzi, J. Bierma, S.I. Wickramasekara, A. Curioni, D. Gazzola, A. T. Bakalinsky, Characterization of a grape class IV chitinase, J. Agric. Food Chem. 62 (2014) 5660–5668, https://doi.org/10.1021/jf501225g.
[56] Y.-J. Wang, W.-X. Jiang, Y.-S. Zhang, H.-Y. Cao, Y. Zhang, X.-L. Chen, et al., Structural insight into chitin degradation and thermostability of a novel endochitinase from the glycoside hydrolase family 18, Front. Microbiol. 10 (2019) 2457, https://doi.org/10.3389/fmicb.2019.02457.
[57] J.H. Waterborg, M. Poot, C.M.A. Kuyper, Nucleic acids and related enzymes, J. Biochem. (Tokyo) 88 (1980) 1715–1721, https://doi.org/10.1093/oxfordjournals.jbchem.a133146.
[58] Q. Yan, S.S. Fong, Bacterial chitinase: nature and perspectives for sustainable bioproduction, Bioresour. Bioprocess. 2 (2015) 31, https://doi.org/10.1186/s40643-015-0057-5.
[59] J. Yang, I. Anishchenko, H. Park, Z. Peng, S. Ovchinnikov, D. Baker, Improved protein structure prediction using predicted interresidue orientations, Proc. Natl. Acad. Sci. U. S. A. 117 (2020) 1496–1503, https://doi.org/10.1073/pnas.1914677117.
[60] A. Zhang, Y. He, G. Wei, J. Zhou, W. Dong, K. Chen, et al., Molecular characterization of a novel chitinase CmChi1 from Chitinolyticbacter meiyuanensis SYBC-H1 and its use in N-acetyl-d-glucosamine production, Biotechnol. Biofuels 11 (2018) 179, https://doi.org/10.1186/s13068-018-1169-x.
[61] K.H. Gardner, J. Blackwell, Refinement of the structure of beta-chitin, Biopolymers. 14 (8) (1975) 1581–1595, https://doi.org/10.1002/bip.1975.360140804.
[62] M. Kaya, M. Mujtaba, H. Ehrlich, A.M. Salaberria, T. Baran, C.T. Amemiya, R. Galli, L. Akyuz, I. Sargin, J. Labidi, On chemistry of γ-chitin, Carbohydr Polym 176 (2017) 177–186, https://doi.org/10.1016/j.carbpol.2017.08.076.
[63] K.E. Köllner, D. Carstens, E. Keller, F. Vazquez, C.J. Schubert, J. Zeyer, H. Urgmann, Bacterial chitin hydrolysis in two lakes with contrasting trophic statuses, Appl. Environ. Microbiol. 78 (3) (2012) 695–704, https://doi.org/10.1128/AEM.06330-11.