Comprehensive *in silico* structural-functional analysis of *Enterobacter* GH19 class I chitinase (*chiRAM*) gene: cloning and heterologous expression

Shahinaz M. Abady
Faculty of Science, Botany and Microbiology, Alexandria University, Egypt
Khaled M. Ghanem
Faculty of Science, Botany and Microbiology department, Alexandria University, Egypt
Nevine B. Ghanem
Faculty of Science, Botany and Microbiology department , Alexandria University, Egypt
Amira M. Embaby (amira.embaby@alexu.edu.eg)
Institute of Graduate Studies and Research, Alexandria University https://orcid.org/0000-0001-8068-4200

Research Article

Keywords: chiRAM Chitinase, Class I GH19 family, cloning, Structural modelling, Enterobacter sp. strain EGY1

Posted Date: April 19th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-420093/v1

Version of Record: A version of this preprint was published at Molecular Biology Reports on November 13th, 2021. See the published version at https://doi.org/10.1007/s11033-021-06914-9.

Abstract

I. Background. Present study aims to clone and express the gene-encoding chitinase / GH19 family from *Enterobacter* sp. in *E.coli* with *in silico* sequence analyses..

II. Methods and results. The putative open reading frame of GH19 chitinase from *Enterobacter* sp. strain EGY1 was cloned and expressed into pGEM-T and pET-28a + vectors, respectively using a degenerate primer. The isolated nucleotide sequence (1821 bp, Genbank accession no.: MK533791.2) was translated to chiRAM protein (606 amino acids, UniProt accession no.: A0A4D6J2L9). chiRAM *in silico* protein sequence analysis revealed GH19 class I chitinase: N-terminus signal peptide (Met^{1}-Ala^{23}), catalytic domain (Val^{83}-Glu^{347} & catalytic triad Glu^{149}, Glu^{171}, Ser^{218}), proline-rich hinge (Pro^{414} -Pro^{450}), (polycystic kidney disease protein motif (Gly ^{465}-Ser ^{533}), C-terminus chitin-binding domain (Ala^{553}-Glu^{593}), and class I conserved motifs (NYNY and AQETGG). Three dimensional model was constructed by LOMETS MODELLER, PDB template: 2dkvA (*Oryza sativa* L. japonica class I chitinase). Recombinant chiRAM was overexpressed as inclusion bodies (IBs) (~ 72kDa; SDS-PAGE) in 1.0 mM IPTG induced *E.coli* BL21 (DE3) Rosetta at room temperature, 18 hrs post induction. Optimized expression yielded active chiRAM with 1.974 U/mL ± 0.0002, on shrimp colloidal chitin (SCC), in induced *E.coli* BL21 (DE3) Rosetta cells growing in SB medium. LC-MS/MS identified the 72 kDa band in the soluble fraction with 52.3% coverage sequence exclusive to *Enterobacter cloacae* chitinase/GH19 (WP_063869339.1).

III. Conclusions. Despite the successful cloning and expression of chiRAM of *Enterobacter* sp. in *E.coli* with an appreciable chitinase activity, prospective studies would focus on minimizing IBs to facilitate chiRAM purification and characterization.

Introduction

Chitinases (E.C.3.2.1.14) are a group of glycoside hydrolases (GH) that catalyze the hydrolysis of the chitin polymer into chitooligosaccharides [(GlcNAC)n] [1, 2]. Chitinolytic enzymes display a wide range of applications in different sectors like pharmaceutical, medica, agricultural, and industrial fields [3]. Their diverse properties along with their broad spectrum of applications gained these enzymes much attention. Considering the amino acid sequences and catalytic mechanisms of chitinases [4], they are grouped into two main GH families: GH18 and GH19 as detailed in the CAZy database (Carbohydrate-Active EnZymes) (http://www.cazy.org). A large number of chitinases from bacteria [5], fungi [6], and plants [7] are highly distributed within GH18 with well-determined three-dimensional structures experimentally deduced from X-ray diffraction analysis. Mostly, five major classes (I, II, IV, VI, VII) are found in members of GH19 chitinases of plant origin [8]. Each class has its unique structural characteristics. For instance, the cysteine-rich domain at the N-terminus, involved in chitin binding, is a signature feature in members of GH19 chitinases of class I and IV. However, the small size of Class IV chitinases is primarily attributed to deletions in the catalytic domain (CatDom) and the cysteine-rich domain. Class II chitinases lack the cysteine-rich domain [9] and have only CatDom [8]. Class VI chitinases are characterized by the presence

of a heavily truncated chitin-binding domain (CBD) along with a proline-rich spacer [10]. However, class VII chitinases are devoid of the CBD but only possess CatDom [11].

Until now, the majority of discovered GH19 chitinases are distributed in viruses [12], nematodes [13], and plants like chitinases from *Brassica juncea* [14], *Canavalia ensiformis* [15], *Hevea brasiliensis* [8], *Oryza sativa* [16], *Bryum coronatum* [9], and *Picea abies* [17]. On the contrary, a small number of GH19 chitinases from bacteria such as *Aeromonas* [18], *Streptomyces sp*. [19], *Vibrio proteolyticus* [20], *Pseudoalteromonas* sp. [21], and *Bacillus circulans* [22] has been discovered to date. Generally, the structural-functional relationship of bacterial GH19 chitinases is not studied profoundly yet. Few reports highlighting the cloning and the expression of GH19 chitinases from bacteria are available so far [18, 23, 24]. Hence, rigorous researches should be conducted for unraveling and better understanding the features of bacterial GH19 chitinases.

Enterobacter spp. are Gram-negative, facultatively anaerobic bacteria belong to family *Enterobacteriaceae*. A well-known member is *Enterobacter cloacae* with serious human implications such as nosocomial infections, wound, respiratory and urinary tract infections [25]. The sequence of the whole genome of different species of *Enterobacter* is currently deposited in the GenBank database and available for the public. Retrieval of the nucleotide sequences of *Enterobacter* chitinases genes from GenBank revealed the presence of four annotated chitinase genes; two of GH18 (AAY67797.1 and ADF62010.1) and two of GH19 (WP_013096312.1 and WP_058672411.1). The chitinases with accession numbers AAY67797.1 and ADF62010.1 were well studied as native and recombinant enzymes [26, 27]. In contrast, the multispecies GH19 chitinases WP_013096312.1 and WP_058672411.1 among *Enterobacter* spp. have not been isolated yet.

In the context of unveiling the nature of GH19 chitinase from *Enterobacter* spp., the objective of the present study is to clone, express, and *in silico* analyse the retrieved open reading frame encoding a novel member of GH19 chitinases from *Enterobacter* sp. strain EGY1 for the first time ever.

Material And Methods

Bacterial strains and vectors

Enterobacter sp. strain EGY1, isolated previously from fish wastes (e.g., particles of flesh, skin, bones, entrails, and shells) and deposited in EMCCN-NRC under number 3074, was the source of the GH19 chitinase open reading frame (ORF) in the current study. The 16S rRNA was previously amplified by PCR according to a protocol previously published [28]. The 16S rRNA nucleotide sequence of this strain was previously deposited in GenBank under the accession number MK123383.1. The capability of the bacterial strain to produce chitinase was detected previously. *E. coli* DH5α (Promega Co., USA) was used as a propagation host. While four *E.coli* BL21 (DE3) strains (Novagen Co., USA) namely Rosetta, Tuner, C43, and ArcticExpress were used as expression hosts. pGEM®-T- Easy vector (Promega Co., USA) and pET-28a (+) vector (Novagen Co., USA) were used as cloning and expression vectors, respectively.

Genomic DNA isolation

The isolation of genomic DNA of *Enterobacter* sp. strain EGY1 was performed using ZR Fungal/Bacterial DNA Miniprep™ (Zymo Co., USA) according to the manufacturer's instructions. The concentration of genomic DNA was determined spectrophotometrically using Nano-drop ^{TM} 2000/2000c spectrophotometer.

PCR

The putative open reading frame (ORF) sequences encoding the multispecies GH19 chitinase gene were retrieved from the full genome sequences of *Enterobacter* spp; recently published and deposited in the international nucleotide sequence databases collaboration (INSDC). Then, multiple sequence alignment using the CLUSTALW program for the retrieved putative ORF sequences was performed through the BioEdit program 7.2 to infer the conserved regions among the aligned sequences. The degenerate primer set

F- *chiGH19-full length* (5'-*GAATTC*ATGATGAAYAAAAGGACVTTACTGAGYGT-3'and R- *chiGH19-full length* (5'-*AAGCTT*YTAGHTGCCGTCCAGGCVGTCAGACCA-3' was designed based on the inferred conserved regions at the boundaries of the aligned sequences to isolate the full length of GH19 chitinase from *Enterobacter* sp. strain EGY1. Degenerate positions are nominated by using the IUPAC (International Union of Pure and Applied Chemistry) code. The italic sequences at the 5' end of each primer are the recognition sites of *Eco*RI and *Hind* III, respectively. PCR was carried out in a 50 µL total volume containing genomic DNA (50 ng), 25 µL 2X PCR Master mix Solution (i-Taq, (iNTRON, Korea)), forward and reverse primers (0.5µM each) and nuclease-free water. PCR conditions were set as follows: 1 cycle at 95°C for 5 min, followed by 30 cycles each at 94°C for 1 min, 55°C for 1 min and 72°C for 110 sec and then 1 cycle for a final extension at 72°C for 10 min.

Cloning and sequencing of Enterobacter sp. strain EGY1 GH19 chitinase gene

The amplified PCR product (GH19 chitinase from *Enterobacter* sp. strain EGY1) was designated as chiRAM gene. It was purified using Zymoclean™ Gel DNA Recovery Kit (Zymo Co., USA) according to the instructions of the manufacturer. pGEM®–T Easy vector was used to clone the chiRAM fragment. The ligation reaction was performed in a total volume of 10µL: 5µL 2x ligation buffer, 1µL (50 ng) pGEM–T Easy vector, 1 µL (2Units) T4-DNA ligase (Promega Co., USA), and 3µL (150ng) purified PCR product. The reaction was incubated overnight at 4°C. The recombinant pGEM-T/ chiRAM construct was transformed into *E. coli* DH5α competent cells [29]. Blue-white screening of the recombinant clones was performed on LB/ampicillin/IPTG/X-gal plates. GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific, USA) was used to prepare plasmids from some recombinant white clones. Sequencing was carried out using ABI PRISM BigDye™ Terminator Cycle with the universal primer set Sp6 /T7.

Subcloning of chiRAM into pET28a (+) vector

The pGEM-T/ chiRAM construct and the pET-28a (+) vector were subjected to a double digestion reaction targeting *EcoR*I and *Hind*III restriction sites. The digestion reaction was executed in a total volume of 50 µL containing 30 µL of pGEM-T/ chiRAM or pET-28a (+) vector, 5µL of 10X buffer, 1µL (2units) of each

restriction enzyme and 13 µL of nuclease-free water. The reaction was incubated overnight at 37°C. The excised chiRAM fragment and the linearized pET-28a (+) vector were purified from the agarose gel using Zymoclean™ Gel DNA Recovery Kit (Zymo Co., USA). The ligation reaction was executed in a total volume of 10µL: 2 µL (100ng) of linearized pET-28a (+) vector, 3µL (150ng) purified chiRAM fragment, 1µL of 10X ligation buffer, 1 µL of T4 DNA ligase, and 3µL of nuclease-free water. The ligation reaction was carried out at 4°C overnight. The ligation mixture (3µL) was used to transform *E. coli* BL21 (DE3) Rosetta competent cells[29]. Transformants were grown on LB/ /kanamycin plates. Plasmids, isolated from some clones, were used as DNA templates in PCR using the universal primer set T7 promoter/ T7 terminator to confirm the presence of the chiRAM insert in the pET-28a (+)/chiRAM construct.

Transformation of pET-28a (+) / chiRAM into E.coli BL21 (DE3) non- Rosetta strains

The pET-28a (+)/chiRAM construct was further transformed into *E.coli* BL21 (DE3) Tuner, C43, and ArcticExpress competent cells [29].

Expression of recombinant chiRAM in E.coli BL21 (DE3) strains

A seed culture (16 hrs) of each *E. coli* BL21 (DE3) strain (Rosetta, Tuner, C43, and ArcticExpress) harbouring the pET28 (a) +/chiRAM construct was used inoculate 200 mL of LB supplemented with the appropriate antibiotics: 34 µg/mL kanamycin for all strains and 20 µg/mL gentamycin for the ArcticExpress strain. The inoculated broth was incubated at 180 rpm at 37^{O}C for 3hrs. Isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 1mM was added to the culture of 0.6–0.8 optical density at 600 nm. The induced cultures were incubated at 37^{O}C and 180 rpm for further 18 hrs for all strains except for the ArcticExpress induced culture was incubated at 10^{O}C and 180 rpm for further 18 hrs. After the indicated time, cells were subjected to sonication according to a procedure previously reported.

Optimized expression of recombinant chiRAM

In order to maximize the amount of recombinant chiRAM protein in a soluble active form, the following strategies were employed: a) expression and induction with 1 mM IPTG from three recombinant *E.coli* BL21 (DE3) strains harbouring the pET28 (a)+ /chiRAM construct in the following *E.coli* (DE3) strains: ArcticExpress, C43, and Tuner at room temperature, b) expression and induction with four different IPTG concentrations (0.4, 0.6, 0.8, 1.0mM) IPTG from recombinant *E.coli* BL21 (DE3) Rosetta harbouring the pET28 (a)+ /chiRAM construct, c) expression and induction with 1 mM IPTG under three different temperatures (room temperature, 30^{O}C, and 37^{O}C) from recombinant *E.coli* BL21(DE3) Rosetta harbouring the pET28 (a)+ /chiRAM construct, d) expression and induction with 1 mM IPTG under room temperature using different growth media (TB, 2xTY, LB, M9, 5xLB, and SB), e) treating the insoluble recombinant expressed chiRAM with four solubilisation buffers (B, C, E, and F) according to a procedure previously mentioned [30], B) 50 mM Tris-HCl, 5 mM EDTA, 1mM PMSF, 6M Guanidine-HCl, pH 8.5, C) 50 mM Tris-HCl, 5 mM EDTA, 1mM PMSF, 2M Urea, pH 12, E) 50 mM Tris-HCl, 5 mM EDTA, 1mM PMSF, 6M *n*-propanol, 2M Urea, pH 8.5, and F) 50 mM Tris-HCl, 5 mM EDTA, 1mM PMSF, 6M *B*-mercaptoethanol, 2M

Urea, pH 8.5 and the resultant solubilized chiRAM protein fractions were dialyzed separately using a dialysis membrane of 14 kDa against 50 mM Tris-HCl, pH 8.5 with four times buffer change along 48 hrs, f) subcloning of the truncated chiRAM (chiRAM-CatDom) on pET28a (+) vector using the primer set: F-*chiRAM-CatDom* (5'-*GAATTC*ATGATGAAAGCCAGCGACTGGGAATA-3' and R- *chiRAM-CatDom* (5'-*AAGCTT*CTTCATGTTGGCGCAACCCAAC-3', and g) checking the chiRAM activity in the cell free supernatant of *E.coli* (BL21) DE3 pET28 (a)+ /chiRAM.

Recombinant chiRAM activity

Shrimp colloidal chitin (SCC), prepared according to a previously mentioned procedure [31], was used as the substrate for enzyme assay. Briefly, the enzyme assay included 1 mL (2.5% SCC in 50 mM citrate phosphate buffer, pH 5.0) and 300 µL crude sonicated cell lysate. The reaction was incubated at 40 ^{o}C for 1 h in a shaker water bath. After that, the reaction tubes were boiled for 10 min after addition of 1 mL 3,5-Dinitrosalicylic acid [32]. Then the absorbance of the developed color was measured spectrophotometrically at 540 nm. A standard curve of N- acetylglucoseamine was established. One unit of enzyme activity is defined as the amount of enzyme that catalyzes the release of 1 µmole N-acetylglucoseamine from SCC per 30 min at 40^{o}C.

Sodium dodecyl sulphate gel electrophoresis (SDS-PAGE)

Expression of the recombinant expressed chiRAM protein was monitored on 10% SDS-PAGE according to Laemmli method [33]. The gel was stained with Coomassie brilliant blue R250. The molecular mass of chiRAM was determined using GangNam-STAIN™ Prestained Protein Ladder (iNTRON CO., Korea).

Scanning electron microscope (SEM)

The expressed recombinant chiRAM in insoluble form (inclusion bodies) was visualized under SEM (JSM-IT200, JEOL, Tokyo, Japan). The sample was prepared according to a previously mentioned procedure [34].

LC-MS/MS

Liquid chromatography electrospray-tandem mass spectrometry (LC-MS/MS) experiment was performed in TAMU Protein Chemistry Lab, Department of Biochemistry and Biophysics, Texas A &M University, USA). The full procedure was performed according to the protocol mentioned in Supplementary file 1. Peptide probabilities from MASCOT were assigned by the Peptide Prophet algorithm [35] with Scaffold delta mass correction. Protein identifications were accepted if they could be established greater than 99.0% and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm [36]

In silico chiRAM sequence analysis

BLAST algorithm (https://blast.ncbi.nlm.nih.gov/Blast.cgi) of NCBI was used to search for homologous nucleotide and protein sequences from INSCD to the generated nucleotide sequence and deduced amino acid sequence of chiRAM. ExPASY server (https://web.expasy.org/translate/) was used to obtain the

translated amino acid sequence of chiRAM. The secondary structure of chiRAM protein was predicted using SAS

(https://www.ebi.ac.uk/thornton-srv/databases/sas/). The presence of a signal peptide in chiRAM sequence was predicted using the SignalP-5.0 server (http://www.cbs.dtu.dk/services/SignalP/) based on neural networks (NN) and Hidden Markov Models (HMM). Multiple sequence alignment using the CLUSTALW program between the deduced amino acid sequence of chiRAM and those of GH19 chitinases from other species was carried out using CLC Sequence Viewer 8.0 program. The database CAZY (Carbohydrate-Active enzymes Database; available at http://www.cazy.org/GH5_19.html) was used to retrieve the deduced amino acids sequences of GH19 chitinases members (E.C. 3.2.1.14) from different organisms. Additionally, a phylogenetic tree was constructed by CLC Sequence Viewer 8.0 program to portray the evolutionary relationship of GH19 chitinases members among different species. The existence and delineation of protein domains were achieved using CDD (Conserved Domain Database); available at https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml. The three dimensional structure (3D) of chiRAM protein was built by submitting the protein sequence to the LOMETS online server (https://zhanglab.ccmb.med.umich.edu/LOMETS/). The LOMETS online programs normalize the selected predicted 3D structure of chiRAM with Z-score of 2.61 to guarantee the quality of the selected template used to build up the 3D structure. However, the evaluation of the 3D model quality was carried out by the Ramachandran plot. The online server http://mordred.bioc.cam.ac.uk/~rapper/rampage.php was used to build up a Ramachandran plot according to Paul de Bakker and Simon Lovell available on the server. The online server https://web.expasy.org/protparam/ was used to predict the theoretical PI and molecular mass of chiRAM. The three programs SOSUI, TMHMM2.0, and PHOBIUS available at http://harrier.nagahama-i-bio.ac.jp/sosui/cgi-bin/adv_sosui.cgi, http://www.cbs.dtu.dk/services/TMHMM/, and https://www.ebi.ac.uk/Tools/pfa/phobius/, respectively were used to predict the presence of any transmembrane helices in chiRAM. The presence of motifs was detected with Motif scan, available on the online server: https://myhits.isb-sib.ch/cgi-bin/motif_scan.

Results

In silico Sequence analysis and phylogenetic relationship of chiRAM

In this study, a set of retrieved putative sequences of the multispecies GH 19 family chitinase gene of *Enterobacter* spp., available in international nucleotide sequence databases collaboration INSDC, were collected. The degenerate primer set (F- *chi* -full length/R-*chi* –full length) mentioned in the section of materials and methods was designed in order to clone the gene encoding GH19 family chitinase gene from *Enterobacter* sp. strain EGY1. The full-length of GH19 chitinase gene from *Enterobacter* sp. strain EGY1 was cloned: 1821 bp (Additional file 1: Figure S1) with 606 deduced amino acid residues. The cloned chitinase gene was designated as chiRAM. The nucleotide sequence of chiRAM was deposited in GenBank under the accession number: MK533791.2. However, the translated amino acid sequence of chiRAM was deposited in UniProtKB database under the accession number A0A4D6J2L9. BLASTp search verified that the deduced amino acid sequence of chiRAM exhibited a high level of homology (99.34–

91.84%) to the multispecies GH19 chitinase sequences in protein databases from several *Enterobacter* spp. like *Ent. cloacae* (VAC58927, 99.34%), *Ent. cloacae* (VAC36707.1, 95.72%), *Ent. hormaechei* (WP_047354344.1, 94.59%), *Ent. cancerogenus* (WP_080327729.1, 94.55%), *Ent. chuandaensis* (WP_119914984.1, 94.22%), *Ent. bugandensis* (WP_059358028.1, 93.42%), and *Ent. ludwigii* (WP_044857960.1, 91.84%). The low levels of homology were observed with chitinase sequences from other species such as *Salmonella enterica* (EBB0882654, 70.01%), *Hevea brasiliensis* (4MST_A, 35.07%), *Secale cereale* (4J0L_A, 35.012%), and *Streptomyces griseus* (1WVU_A, 28.33%) (Table S1). The relationship between chiRAM amino acid sequences and other bacterial and plant GH19 chitinases was depicted in a phylogenetic tree (Fig. 1A). The predicted pI and molecular mass of chiRAM are 4.84 and 66.08 kDa, respectively. Analysis of the deduced amino acid sequence of chiRAM using CDD database revealed that chiRAM is a modular protein. The modular structure of chiRAM was depicted in Fig. 1B. The N-terminal sequence of the 23 amino acid residues (Met^{1} to Ala^{23}) is the signal peptide predicted by the analysis of chiRAM amino acid sequence on the Signal IP-5.0 Server. The signal peptide (MNKR*TLLSVLIAGACVA*PFMAQA) consisted of three segments: short positively charged segment (MNKR), central hydrophobic segment (*TLLSVLIAGACVA*), and cleavage segment (PFMAQA). The cleavage segment has the sequence Ala-X-Ala preceding the cleavage peptidase site. The remaining amino acid sequence encompasses catalytic domain (CatDom; residues Val^{83} to Glu^{347}), polycystic kidney disease protein motif (PKD; residues Gly^{465} to Ser^{533}), and chitin binding domain (ChiA1_BD; residues Ala^{553} to Glu^{593}). Analysis of the deduced amino acid residues of chiRAM with CDD database revealed that the chitin binding domain of chiRAM (ChiA1_BD) is belonging to chtBD3 superfamily which assigns to family 12 carbohydrate binding module (CBM12); in accordance with CAZY categorization. BLASTp search showed that chiRAM (ChiA1_BD) did exhibit moderate percent of identity (%) with diverse ChtBD(s) from other species (Fig. 2A) like chitin binding protein domain of protease precursor (*Pseudomonas aeruginosa* PAO1: Q9I589.1, 48.65%), full chitin binding protein domain (*P.aeruginosa* UCBPP-PA14: Q02I11.1, 48.65%), GlcNAc-binding protein A (*Vibrio parahaemolyticus* RIMD 2210633: Q87FT0.1, 47.5%), ChiC_BD and ChiA1_BD (*Aeromonas* sp. 10S-24: BAA09627.1, 41.67% & 39.13%, respectively), ChtBD3 (*Streptomyces cyaneus*: BAG55920.1, 36%), ChiA1_BD (*Str. olivaceoviridis*: CAB37321.1, 41.67%), ChiA1_BD (*Lysobacter enzymogenes*: AAT77163.1, 40.9%), ChiA1_BD (*Vibrio alginolyticus*: BAB21759.1, 36.36%), ChtBD3 (*Chitinophaga pinensis* DSM 2588: ACU62980.1, 30.4%), and ChtBD3 (*Micromonospora chokoriensis*: SCF20675.1, 31.82%).

The superfamily chtBD3 members are characterized by the presence of six conserved aromatic amino acid residues and three hydrophobic side chains as well. This signature feature of chtBD3 superfamily in ChiA1_BD of chiRAM is verified by the presence of six aromatic amino acid residues at Trp^{554}, Tyr^{560}, Tyr^{580}, Trp^{583}, Tyr^{590}, and Tyr^{591}; including two hydrophobic amino acid residues are allocated at Val^{565}, Cys^{574}, Pro^{587}(Fig. 2A) .

The proline-rich hinge region, a signature feature of class I GH19 chitinase separating the catalytic domain from the chitin binding domain [37], spanned from Pro^{414} to Pro^{450} in chiRAM inferred from sequence analysis by Motif Scan server.

Multiple sequence alignment among the amino acid sequences of chiRAM and other hits of class I plant and bacterial GH19 chitinases (Fig. 2B), gathered from CAZY, PDB, and UniprotKB databases, revealed the presence of the class I conserved motif NYNY among all aligned GH19 chitinase amino acid sequences. The evolutionary relationship and divergence of the conserved class I plant GH 19 chitinase SHETTG motif among class I bacterial GH19 chitinase was implied from Fig. 2B. The class I plant GH19 chitinase SHETTG signature motif was localized in the form of *AQ*ET*G*G in chiRAM amino acid sequence (Fig. 2B). Another, class I conserved motif, WPCAPGRKYFGRGP among class I plant GH19 chitinases [38], was traced partially in chiRAM as WPC*GKDKDGFLS*YFGRG*A* (Fig. 2B).

The secondary structure of chiRAM was predicted by the online SAS program (Fig. 3A) using Carica papaya chitinase (PDB: 3cql:A) as a template. The catalytic residues of chiRAM Glu^{149}, Glu^{171}, and Ser^{218} were localized on helix, coil, and coil, respectively. The topological features of chiRAM proved that it has 10 α-helical structures in the catalytic domain. The residue contacts to ligands or metals were presented in additional file 2: Table S2. The three dimensional structure of chiRAM was depicted in Fig. 4. The full-length three dimensional model for chiRAM was constructed by MODELLER of LOMETS as shown in Fig. 3B. The three dimensional built model for chiRAM was constructed based on the top template in PDB: 2dkvA with identity percent 40%. The normalized Z-score ranged from 1.09–7.69 for the selected template as deduced from the output of the ten individual threading programs. This in turn reflected that the selected templates were of good quality. The selected template 2dkvA in PDB was assigned to crystal structure of class I chitinase from *Oryza sativa* L. japonica. The quality of the built model was judged to be good as inferred from the Ramachandran plot depicted in Fig. 3C. The analysis of the Ramachandran plot for chiRAM model revealed that the number of residues in the favored region (~ 98.0% expected) was 558 (92.43%). Whilst, the number of residues in the allowed region (~ 2.0% expected) was 24 (4.0 %). The number of residues in the outlier region was 22 (3.6%).

Expression of recombinant chiRAM

The recombinant chiRAM was overexpressed from 1mM IPTG induced *E.coli* BL21 (DE3) Rosetta harboring the construct pET-28 (a) +/chiRAM after 18 h induction at room temperature. The chiRAM was overexpressed in the form of insoluble protein (inclusion bodies) deduced from SDS-PAGE (Fig. 4A) with an appreciable level of the expressed chiRAM in a soluble active form with 0.150 ± 0.0003 U/mL (Table 1). The chiRAM exhibited an apparent molecular weight of approximately 72 kDa (Fig. 4A). Examination of the insoluble fraction of the cell lysate of induced *E.coli* BL21(DE3) Rosetta harboring the construct pET-28 (a) +/chiRAM under scanning electron microscope did reveal the amorphous inclusion bodies of chiRAM as big, thick, and seemingly oval with an approximate diameter of 750 nm (Fig. 4B).

Table 1
Optimized expression of recombinant chiRAM using different parameters

Parameter	**chiRAM activity (U/mL)***
Uninduced *E.coli* BL21(DE3) pET-28a (+)/chiRAM Rosetta strain in LB	N.D
E.coli BL21(DE3) pET-28a (+)/chiRAM Rosetta strain in LB under different incubation temp., 1mM IPTG	
37^{o}C	N.D
30^{o}C	0.1809 ± 0.0003
RT*	0.301 ± 0.0003
E.coli BL21(DE3) pET-28a (+)/chiRAM Rosetta strain in LB, RT, at different 1mM IPTG conc.	
0.4	N.D
0.6	0.282 ± 0.0004
0.8	0.286 ± 0.0003
1.0	0.301 ± 0.0003
E.coli BL21(DE3) pET-28a (+)/chiRAM Rosetta strain in different culture media, 1mM IPTG, RT	
LB	0.301 ± 0.0003
M9	1.853 ± 0.0002
TB	0.561 ± 0.0001
5xLB	1.416 ± 0.0002
2xTY	1.238 ± 0.0002
SB	1.974 ± 0.0002
Cell free supernatant of *E.coli* BL21(DE3) pET-28a (+)/chiRAM Rosetta strain in LB, 1mM IPTG, RT	N.D
E.coli BL21(DE3) pET-28a (+)/chiRAM C43 strain in different culture media, 1mM IPTG, RT	
LB	0.112 ± 0.0004
M9	0.052 ± 0.0001
TB	0.066 ± 0.0002

N.D: not detected, RT: room temperature

* chitinase activity was estimated using shrimp colloidal chitin as a substrate

Parameter	chiRAM activity (U/mL)*
5xLB	0.242 ± 0.0002
2xTY	1.186 ± 0.0004
SB	0.718 ± 0.0003
E.coli BL21(DE3) pET-28a (+)/chiRAM ArcticExpress strain in LB medium, 1mM IPTG, 4^{o}C	N.D
E.coli BL21(DE3) pET-28a (+)/chiRAM Tuner strain in LB medium, 1mM IPTG, RT	
LB	0.090 ± 0.0002
M9	0.260 ± 0.0002
SB	0.050 ± 0.0002
2xTY	0.270 ± 0.0003
chiRAM IBs treated with different solubilization buffers	
Buffer B	N.D
Buffer C	N.D
Buffer E	N.D
Buffer F	N.D
N.D: not detected, RT: room temperature	
* chitinase activity was estimated using shrimp colloidal chitin as a substrate	

Optimized expression and solubilization of chiRAM

In an attempt to maximize the yield of the recombinant chiRMA activity in the soluble fraction, a detailed protocol was applied as mentioned in the section of materials and methods. Neither soluble or insoluble chiRAM could be detected upon using the *E.coli* BL21(DE3) ArcticExpress strain (Additional file 1: Figure S2 and Table 1). The recmbinanat *E.coli* BL21(DE3) Rosetta strain gave the highest chiRAM activity (0.300 ± 0.0003 U/mL) compared to the levels obtained from the recombinanat strains Tuner and C43 growing in LB medium, at room temperature, and 1 mM IPTG (Additional file 1: Figure S2 and Table 1). Conducting the induction at room temperature did result in a higher chiRAM activity (0.300 ± 0.0003 U/mL) than that obtained upon applying 30^{o}C as the induction temperature (Additional file 1: Figure S3 and Table 1). No chiRAM activity could be detected upon applying 37^{o}C as the induction temperature. The usage of 0.2% lactose as a chiRAM inducer did elicit low soluble and insoluble chiRAM levels compared to those obtained upon using the conventional inducer IPTG (Additional file 1: Figure S4). The

concentration of 0.4 mM IPTG did not elicit of the expression of chiRAM in a soluble active form. No significant difference could be obtained upon using three different concentrations of IPTG (0.6, 0.8, and 1 mM) using the recombinant rosetta strain (Additional file 1: Figure S5 and Table 1). For the recombinant rosettta strain, using the growth media M9 and SB did significantly impose the highest yield of chiRAM activity 1.852 ± 0.0002 and 1.974 ± 0.0002 U/mL, respectively when compared to the yield obtained upon using LB, 5xLB, and 2xTY as growth media (Additional file 1: Figure S6 and Table 1). The fold enhancement of 6.17 and 6.58 in chiRAM activity from the recombinant rosetta strain was obtained upon applying M9 and SB as growth media compared to the yield obtained upon using LB medium. For the recombinant C43 strain, 2xTY and SB did elicit the highest level of chiRMA activity 1.186 ± 0.0004 and 0.718 ± 0.0003 U/mL, respectively compared to the influence of the other growth media LB, M9, TB, and 5xLB (Table 1). For the Tuner strain, only M9 and 2xTY did elicit the highest level of chiRMA activity 0.260 ± 0.0002 and 0.270 ± 0.0003 U/mL, respectively compared to the influence of LB and SB growth media (Table 1). The four solubilization buffers could successfully solubilize the IBs of chiRAM (Additional file 1: Figure S7) without any detectable activity (Table 1). No chiRAM activity could be detected in the cell free supernatant of *E.coli* BL21(DE3) pET28 (a)+ /chiRAM Rosetta strain. Neither soluble nor insoluble fractions of the cell lysate of *E.coli* BL21(DE3) pET28 (a)+ /chiRAM-CatDom Rosetta strain display expression for the truncated chiRAM (chiRAM-CatDom).

Analysis of chiRAM using LC-MS-MS

LC-MS-MS did confirm the identity of chiRAM band of ~ 72 kDa in the soluble fraction as chitinase. LC-MS –MS deduced 31 unique peptides with 100% protein identification probability, and 52.3% coverage sequence exclusive to the multispecies *Enterobacter cloacae* chitinase/GH19 (WP_063869339.1 & 66.07 kDa) (Fig. 4C).

Discussion

Despite the plethora of reports to clarify the nature of GH19 chitinases from a panel of organisms including plants, viruses, nematodes, etc., still more research is mandatory to elucidate the main structural differences between the unexplored bacterial GH19 chitinases and the well-studied GH19 chitinases of non-bacterial members. Recently, the huge number of microbial genome projects does introduce a hill of complete microbial genome sequences in the international repository nucleotide databases (e.g., GenBank, EMBL, and DDJA) with putative and annotated GH19 chitinases from numerous bacteria. Consequently, the above mentioned does address the indispensable need for the heterologous expression of the putative bacterial GH19 chitinases sequences to better understand the nature and the mode of action of these enzymes.

In our study, the multispecies class I GH19 chitinase namely chiRAM was cloned and expressed in *E. coli* BL21 (DE3) Rosetta for the first time ever. Blastp analysis conferred a high sequence identitiy ranged from 99.34–91.84 % among protein sequences of GH19 chitinase of different *Enterobacter* spp. and chiRAM (Table 1). This would in turn explore the rationale behind the nomenclature of the multispecies GH19 chitinase from *Enterobacter* spp. in Genbank. Likewise GH18 chitinases, the number and

configuration of domains of plant GH19 chitinases differ widely in the five classes I, II, IV, VI, and VII [8]. In spite of this discrepancy, each domain has its own unique function that does support the overall chitinase activity or not. The domains of GH19 chitinases include the N-terminal signal peptide region, catalytic domain, chitin binding domain, and C-terminal extension region inferred from the literature of review [9]. A typical CBD in chiRAM belonged to chtBD3 superfamily was evidenced. The chtBD3 superfamily does encompass modules of 40–60 amino acid residues with six conserved aromatic amino acids residues and three hydrophobic amino acids side chains, that help bind chitin and /or cellulose. The present data are in a partial agreement with that of ChBD $_{chiA1}$ from *Bacillus circulans* WL-12; with three hydrophobic amino acid residues localized at Val^{668}, Cys^{677}, and Leu^{695} [39] and ChBD $_{CvChi45}$ of *Chromobacterium violaceium* with three hydrophobic amino acid residues at Val^{41}, Ala^{50}, and Leu^{70} [40]. Moreover, the presence and the position of the CBD region vary widely among different classes (I, II, IV, VI, and VII) of GH 19 chitinases members concluded from the CDD database. The CBD at the N-terminus region is considered one signature feature of class I GH 19 chitinases from plants [41]. In accordance, the absence of expression from the pET-28a(+)/chiRAM-CatDom greatly verified that chiRAM is belonging to class I GH19 chitinases, well known with their CBD as an essential domain for the overall enzyme activity and enzyme structure.

Conversely, members of class I GH19 chitinases from bacterial origin showed a wide divergence in the position of the CBD evidenced from the analysis of the collected sequences from CAZY database. Likewise C-terminus CBD of chiRAM, the C-terminus ChtBD(s) were annotated in the following class I GH19 chitinase sequences: BAE86996.1 chitinase (ChiC_BD: 437–474) [*V. proteolyticus*], SPM16686.1 putative chitinase (ChiC_BD: 439–479) [*V.cholerae*], and ACU62980.1 Chitinase (ChtBD3 superfamily: 488–528 & ChiC_BD: 554–595) [*Chitinophaga pinensis* DSM 2588]. Unlike C-terminus CBD of chiRAM, the N-terminus ChtBD(s) were annotated in the following putative class I GH19 chitinase sequences: APJ14276.1 chitinase ChiC_BD: 19–53 [*Aeromonas hydrophila*], ACH66624.1 basic endochitinase (ChiC_BD: 24–63) [*Aliivibrio fischeri* MJ11], ARU51128.1 chitinase (ChiC_BD: 58–100) [*Cellulosimicrobium cellulans*], AVG16867.1 chitinase (ChiC_BD: 20–59) [*Chromobacterium vaccinii*], ALN86380.1 secreted chitinase domain protein (ChtBD3 superfamily: 66–113) [*Lysobacter capsici*], SCF20675.1 chitinase (ChiC_BD: 97–138 & ChtBD3 superfamily: 42–76) [*Micromonospora chokoriensis*], AAQ24634.1 chitinase (ChiC_BD: 35–74) [*Streptomyces griseobrunneus*], 1WVU_A chitinase (ChtBD3 superfamily: 6–45) [*Str. griseus*], ABZ86730.1 chitinase (ChiC_BD: 153–193) [*Francisella philomiragia* subsp. philomiragia ATCC 25017], and ATP28490.1 chitinase (ChiC_BD: 21–59) [*Chromobacterium violaceum*]. The deviating sequences and assemblies among diverse groups suggest that they may have evolved from unlike ancestral genes.

Class I plant GH19 chitinases have signature motifs like SHETTG and NYNY. The signature motif SHETTG was found in many three dimensional structural models of plant chitinases (e.g., PDB 3cql endochitinase *Carica papaya*, PDB 2baa endochitinase *Hordeum vulgare*, PDB 3iwr chitinase *Oryza sativa* subsp. Japonica, PDB 4j01 basic endochitinase *Secale cereale*, PDB 2z37 chitinase *Brassica juncea*, PDB 4MST chain A, class I chitinase *Hevea brasiliensis*, and P24091.1 endochitinase B Precursor

Nicotiana tabacum). Upon tracing the SHETTG motif in class I GH19 chitinases among bacteria, a modification was noticed in this consensus sequence (Fig. 2B). For instance, the signature motif SHET*G*G was found in the following sequences: BAA92252.1 chitinase class I *Burkholderia gladioli*, BAA88833.1 chi35 *Str. thermoviolaceus*, PDB1WVU Chain A chitinase C *Str. griseus*, QAT82978.1 class I chitinase *Corallococcus coralloides*, ABL92611.1 chitinase *Mycobacterium* sp. KMS and QCW26020.1 chitinase *Lysobacter enzymogenes*. However, the signature motif *AQ*ET*G*G was observed in the next sequences: APJ14276.1 chitinase *A. hydrophila*, AWE26318.1 chitinase *Salmonella enterica*, and CCG31634.1 chitinase *Klebsiella aerogenes* EA1509E. While the signature motifs AQETGH, SQETTG, and AQETGG were found in BAD92016.1 chitinase I *B. circulans*, ACU62980.1 chitinase *Chitinophaga pinensis* DSM 2588, and QCC87378.1 chiRAM chitinase, respectively. In contrast, the NYNY signature motif was fully conserved among class I chitinases from plant and bacteria (Fig. 2B).

The CBD of chiRAM is linked to the catalytic domain by a hing region of 56 residues; rich in proline (Fig .1B). Simialrly, bacterial glycoside hydrolases (GH) like celulases, xylanases and chitinases are chacarerized by the presence of proline or threonine rich spacers connecting the catalytic domain to the non –catalytic binding domain[42]. These spacers or linkers were previously proved to fine tune the binding and the catalytic effieiciny of these enzymes to their insoluble substrates [43]. The modular structure of chiRAM revealed the presence of the unique PKD domain that is completely absent in plant GH18/GH19 chitinases and bacterial GH18 chitinases. Conversely, the PKD domain was annotated in the putative GH19 chitinase sequences of bacteria like (AWE30884.1) of *Salmonella enterica*, (AXY29321.1) of *Klebsiella aerogenes*, (ALB71769.1) of *Cronobacter muytjensii* ATCC 51329, and (ABK38513.1) of *Aeromonas hydrophila* subsp. hydrophila ATCC 7966. The PKD refers to a protein that is associated with the polycystic kidney syndrome. The function of the PKD domain in these putative GH19 chitinase sequences is still unraveled as these chitinases have not been studied yet. However, its evolution in *Enterobacter* spp. GH19 chitinases might be attributed to the colonization of *Enterobacter* in the intestine of human and animals.

Like all secreted proteins, chiRAM has a signal peptide that possibly does initiate the transport across the cytoplasmic membrane. Moreover, the fine structure of chiRAM signal peptide was in a great accordance with those of GH19 chitinases from *Aeromonas* sp., *Str. griseus*, and GH18 chitinase of *Str. lividance* [19]. The signal peptide in the wild type chitinase-producing bacteria plays an essential role to transfer the expressed chitinases from inside the cell to outside the cells as the chitin polymer is too big to be imported into the cell.

Regarding to the structural configuration of chiRAM, the predicted secondary structure and the predicted three dimensional structure assumed its richness in α-helix chains (10 chains). This is quite compatible with the structural folding of plant GH19 chitinases [44] and bacterial GH19 chitinases (e.g., GH19 *Pseudomonas aeruginosa*) [45] with a conserved core region consisting of 10 α-helices. Conversely, the configuration of GH18 chitinases is in α/β barrel [46].

The estimated molecular weight of the recombinant chiRAM (~ 72 kDa) was slightly higher than than the theortical molecualr weight (66 kDa). The increase in the molecular weight of the recombinant chiRAM could be imposed by the presence of his-tags (~ 2.5kDa) [47] and the slow migration of chiRAM on SDS-PAGE with relation to the protein structure of GH19 chitinases [19].

Heterologous expression in the microbial cell factory *E.coli* is one of the well-established strategies for both small and large scale production of recombinant proteins because *E.coli* is easy to grow and the overall process is cost-effective. Notwithstanding, the formation of large, insoluble, mis-folded versions of the recombinant proteins in the cytoplasm namely inclusion bodies (IBs) is the main obstacle that hinders the consecution of the overall process, obtaining high levels of active soluble proteins. The exact mechanism controlling the formation of IBs is still not completely understood [48]. However, altered cellular homeostasis would trigger the formation of IBs [48]. The literature of review reported a number of strategies aimed at obtaining soluble, properly folded, active protein species like tailoring cultural conditions (temperature, inducer, and growth rate), genetic manipulation of the target (protein truncation and fusion to solubilization tags), specific codon usage optimization, and exposing the IBs to a panel of solubilization buffers [49]. Nevertheless, there is no a generic method fitting for the solubilization and re-folding of all proteins so far. For each case, there must be a strategy tailored based on trials and errors; custom–made strategies [50]. The initial level of recombinant chiRAM expressed in *E.coli* BL21 (DE3) Rosetta strain was 0.300 ± 0.0003 U/mL. The applied optimized strategy did successfully enhance the traced chiRAM activity to 1.852 ± 0.0002 and 1.974 ± 0.0002 U/mL in the cell lysate of the recombianant Rosetta strain upon substituting the LB growth medium with M9 and SB, respectively. Herein, a fold enhancement of 6.17 and 6.58 in chiRAM activity expressed by from the recombinant Rosetta strain was obtained upon applying M9 and SB as growth media compared to the yield obtained upon using the conventional LB medium. The compositions of the different media (LB, TB, M9, 2xTY, 5xLB, and SB) do likely trigger an influence on the cellular physiology of the recombinant *E.coli* (DE3) cells. Consequently, the differences in cellular physiology among different *E.coli* (DE3) growing on different media most probably impose influences on the nature and the folding status of the expressed recombinant enzymes from these strains. Interestingly, a further enhancement in chiRAM activity from the recombinant Rosetta strain would likely be achieved upon scaling up the process in a laboratory scale fermenter in prospective studies. Conversely, the four solubilzation buffers (B,C, E, and F) could successfully solubilize the chiRAM IBs without tracing any acivity. This could be attributed to the irreversible enzyme denaturation triggered by the denaturants Guanidine-HCl and urea included in these buffers.

Since a small amount of the recombinant chiRAM was expreesed in the soluble fraction, it was mandatory to get further confirmation about the identity of the band of ~ 72 kDa in the soluble fraction by mass spectrometry. LC-MS-MS confirmed the identity of this band in the soluble fraction as chitinase similar to that of *Enterobacter cloacae* chitinase/GH19 (WP_063869339.1 & 66.07 kDa) with 52% sequence coverage.

Conclusions

Conclusively, the present study did successfully clone, heterlogously express and *in silico* analyze the structure-functional properties the novel GH19 class I chitinase (chiRAM) from *Enterobacter* sp. Strain EGY1. Prospective studies would focus on minimization the formation of IBs of recombinant chiRAM by fusing chiRAM to a panel of solubilization tags. Further studies would also address purification and characterization of the recombinant chiRAM. Present data would open the avenue towards presenting a novel member of the rarely studied GH19 chitinase of bacterial origin.

Declarations

Authors' contributions

S.M. Abady performed all laboratory experiments. K.M.Ghanem put the main research idea, put the experimental designs, and revised the manuscript. N. B. Ghanem participated in the data analysis and revised the manuscript. A.M. Embaby participated in putting the research idea, setting the design of all experiments, analysed the whole laboratory data and wrote the whole draft manuscript. All authors approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article and its additional files.

Code availability

Not applicable.

Ethical approval

Not applicable.

Consent to publish

Not applicable.

Consent to participate

Not applicable.

Conflicts of interest

All authors declare that there is no any conflict of interest.

Authors' details

[1] Department of Botany and Microbiology, Faculty of Science, Alexandria University, Alexandria, Egypt.

[2] Department of Biotechnology, Institute of Graduate Studies and Research, Alexandria University, Alexandria, Egypt.

References

1. Synowiecki J, Al-Khateeb NA (2003) Production, Properties, and Some New Applications of Chitin and Its Derivatives. Crit Rev Food Sci Nutr. https://doi.org/10.1080/10408690390826473
2. Dahiya N, Tewari R, Hoondal GS (2006) Biotechnological aspects of chitinolytic enzymes: A review. Appl Microbiol Biotechnol 71:773–782. https://doi.org/10.1007/s00253-005-0183-7
3. Bhattacharya D, Nagpure A, Gupta RK (2007) Bacterial chitinases: Properties and potential. Crit Rev Biotechnol 27:21–28. https://doi.org/10.1080/07388550601168223
4. Robertus JD, Monzingo AF (1999) The structure and action of chitinases. EXS 87:125–135. https://doi.org/10.1007/978-3-0348-8757-1_9
5. Henrissat B, Bairoch A (1993) New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem J 293:781–788. https://doi.org/10.1042/bj2930781
6. Hartl L, Zach S, Seidl-Seiboth V (2012) Fungal chitinases: Diversity, mechanistic properties and biotechnological potential. Appl Microbiol Biotechnol 93:533–543. https://doi.org/10.1007/s00253-011-3723-3
7. Grover A (2012) Plant Chitinases: Genetic Diversity and Physiological Roles. CRC Crit Rev Plant Sci 31:57–73. https://doi.org/10.1080/07352689.2011.616043
8. Martínez-Caballero S, Cano-Sánchez P, Mares-Mejía I, et al (2014) Comparative study of two GH19 chitinase-like proteins from Hevea brasiliensis, one exhibiting a novel carbohydrate-binding domain
9. Taira T, Mahoe Y, Kawamoto N, et al (2011) Cloning and characterization of a small family 19 chitinase from moss (Bryum coronatum). Glycobiology 21:644–654. https://doi.org/10.1093/glycob/cwq212
10. Berglund L, Brunstedt J, Nielsen KK, et al (1995) A proline-rich chitinase from Beta vulgaris. Plant Mol Biol. https://doi.org/10.1007/BF00019193
11. Truong NH, Park SM, Nishizawa Y, et al (2003) Structure, heterologous expression, and properties of rice (oryza sativa L.) family 19 chitinases. Biosci Biotechnol Biochem. https://doi.org/10.1271/bbb.67.1063
12. Berini F, Katz C, Gruzdev N, et al (2018) Microbial and viral chitinases: Attractive biopesticides for integrated pest management. Biotechnol Adv 36:818–838. https://doi.org/10.1016/j.biotechadv.2018.01.002
13. Tachu B, Pillai S, Lucius R, Pogonka T (2008) Essential role of chitinase in the development of the filarial nematode Acanthocheilonema viteae. Infect Immun 76:221–228.

https://doi.org/10.1128/IAI.00701-07

14. Guan Y, Ramalingam S, Nagegowda D, et al (2008) Brassica juncea chitinase BjCHI1 inhibits growth of fungal phytopathogens and agglutinates Gram-negative bacteria. J Exp Bot 59:3475–3484. https://doi.org/10.1093/jxb/ern197
15. Schlesier B, Koch G, Horstmann C (1998) Characterization of a class II chitinase from jack bean (Canavalia ensiformis) seeds. Nahrung / Food 42:170–170. https://doi.org/10.1002/(sici)1521-3803(199808)42:03/04<170::aid-food170>3.3.co;2-1
16. Sasaki C, Itoh Y, Takehara H, et al (2003) Family 19 chitinase from rice (Oryza sativa L.): Substrate-binding subsites demonstrated by kinetic and molecular modeling studies. Plant Mol Biol 52:43–52. https://doi.org/10.1023/A:1023972007681
17. Wiweger M, Farbos I, Ingouff M, et al (2003) Expression of Chia4-Pa chitinase genes during somatic and zygotic embryo development in Norway spruce (Picea abies): Similarities and differences between gymnosperm and angiosperm class IV chitinases. J Exp Bot 54:2691–2699. https://doi.org/10.1093/jxb/erg299
18. Ueda M, Kojima M, Yoshikawa T, et al (2003) A novel type of family 19 chitinase from Aeromonas sp. No.10S-24: Cloning, sequence, expression, and the enzymatic properties. Eur J Biochem 270:2513–2520. https://doi.org/10.1046/j.1432-1033.2003.03624.x
19. Okazaki K, Yamashita Y, Noda M, et al (2004) Molecular cloning and expression of the gene encoding family 19 chitinase from Streptomyces sp. J-13-3. Biosci Biotechnol Biochem 68:341–351. https://doi.org/10.1271/bbb.68.341
20. Honda Y, Taniguchi H, Kitaoka M (2008) A reducing-end-acting chitinase from Vibrio proteolyticus belonging to glycoside hydrolase family 19. Appl Microbiol Biotechnol 78:627–634. https://doi.org/10.1007/s00253-008-1352-2
21. García-Fraga B, da Silva AF, López-Seijas J, Sieiro C (2015) A novel family 19 chitinase from the marine-derived Pseudoalteromonas tunicata CCUG 44952T: Heterologous expression, characterization and antifungal activity. Biochem Eng J 93:84–93. https://doi.org/10.1016/j.bej.2014.09.014
22. Yano S, Rattanakit N, Wakayama M, Tachiki T (2005) Cloning and expression of a Bacillus circulans KA-304 gene encoding chitinase I, which participates in protoplast formation of Schizophyllum commune. Biosci Biotechnol Biochem 69:602–609. https://doi.org/10.1271/bbb.69.602
23. Yano S, Kanno H, Tsuhako H, et al (2020) Cloning, expression, and characterization of a GH 19-type chitinase with antifungal activity from Lysobacter sp. MK9-1. J Biosci Bioeng xxx: https://doi.org/10.1016/j.jbiosc.2020.11.005
24. Li S, Zhang B, Zhu H, Zhu T (2018) Cloning and expression of the Chitinase Encoded by ChiKJ406136 from Streptomyces Sampsonii (Millard & Burr) Waksman KJ40 and its antifungal effect. Forests 9:1–18. https://doi.org/10.3390/f9110699
25. Mezzatesta ML, Gona F, Stefani S (2012) Enterobacter cloacae complex: Clinical impact and emerging antibiotic resistance. Future Microbiol.

26. Salam M, Dahiya N, Sharma R, et al (2008) Cloning, characterization and expression of the chitinase gene of Enterobacter sp. NRG4. Indian J Microbiol. https://doi.org/10.1007/s12088-008-0044-z
27. Mallakuntla MK, Vaikuntapu PR, Bhuvanachandra B, et al (2017) Transglycosylation by a chitinase from Enterobacter cloacae subsp. cloacae generates longer chitin oligosaccharides. Sci Rep. https://doi.org/10.1038/s41598-017-05140-3
28. Embaby AM, Heshmat Y, Hussein A, Marey HS (2014) A sequential statistical approach towards an optimized production of a broad spectrum bacteriocin substance from a soil bacterium bacillus sp. YAS 1 Strain. Sci World J 2014:. https://doi.org/10.1155/2014/396304
29. Chang AY, Chau VW, Landas JA, Pang Y (2017) Preparation of calcium competent Escherichia coli and heat-shock transformation. JEMI Methods
30. García-Fruitós E (2014) Insoluble proteins: Methods and protocols
31. ROBERTS WK, SELITRENNIKOFF CP (1988) Plant and Bacterial Chitinases Differ in Antifungal Activity. Microbiology. https://doi.org/10.1099/00221287-134-1-169
32. Miller GL (1959) Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. Anal Chem 31:426–428. https://doi.org/10.1021/ac60147a030
33. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. https://doi.org/10.1038/227680a0
34. Abby SS, Melcher M, Kerou M, et al (2018) Candidatus Nitrosocaldus cavascurensis, an ammonia oxidizing, extremely thermophilic archaeon with a highly mobile genome. Front Microbiol 9:1–19. https://doi.org/10.3389/fmicb.2018.00028
35. Keller A, Nesvizhskii AI, Kolker E, Aebersold R (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem 74:5383–5392. https://doi.org/10.1021/ac025747h
36. Nesvizhskii AI, Keller A, Kolker E, Aebersold R (2003) A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem 75:4646–4658. https://doi.org/10.1021/ac0341261
37. Renner T, Specht CD (2012) Molecular and functional evolution of class i chitinases for plant carnivory in the caryophyllales. Mol Biol Evol. https://doi.org/10.1093/molbev/mss106
38. Cao J, Tan X (2019) Comprehensive analysis of the chitinase family genes in tomato (Solanum Lycopersicum). Plants. https://doi.org/10.3390/plants8030052
39. Ikegami T, Okada T, Hashimoto M, et al (2000) Solution structure of the chitin-binding domain of Bacillus circulans WL-12 chitinase A1. J Biol Chem. https://doi.org/10.1074/jbc.275.18.13654
40. Lobo MDP, Silva FDA, Landim PG de C, et al (2013) Expression and efficient secretion of a functional chitinase from Chromobacterium violaceum in Escherichia coli. BMC Biotechnol. https://doi.org/10.1186/1472-6750-13-46
41. Neuhaus JM, Fritig B, Linthorst HJM, et al (1996) A revised nomenclature for chitinase genes. Plant Mol Biol Report 14:102–104. https://doi.org/10.1007/BF02684897

42. Anthony R, Warren J (1993) β-1,4-Glycanases and β-Glycosidases. Curr Opin Biotechnol 4:469–473. https://doi.org/10.1016/0958-1669(93)90014-N
43. Poon DKY, Withers SG, McIntosh LP (2007) Direct demonstration of the flexibility of the glycosylated proline-threonine linker in the Cellulomonas fimi xylanase Cex through NMR spectroscopic analysis. J Biol Chem 282:2091–2100. https://doi.org/10.1074/jbc.M609670200
44. Tanaka J, Fukamizo T, Ohnuma T (2017) Enzymatic properties of a GH19 chitinase isolated from rice lacking a major loop structure involved in chitin binding. Glycobiology. https://doi.org/10.1093/glycob/cwx016
45. Chen L, Chen J, Kumar A, Liu Z (2015) Effects of domains modification on the catalytic potential of chitinase from Pseudomonas aeruginosa. Int J Biol Macromol 78:266–272. https://doi.org/10.1016/j.ijbiomac.2015.04.017
46. Van Aalten DMF, Synstad B, Brurberg MB, et al (2000) Structure of a two-domain chitotriosidase from Serratia marcescens at 1.9-Å resolution. Proc Natl Acad Sci U S A. https://doi.org/10.1073/pnas.97.11.5842
47. Booth WT, Schlachter CR, Pote S, et al (2018) Impact of an N-terminal polyhistidine tag on protein thermal stability. ACS Omega. https://doi.org/10.1021/acsomega.7b01598
48. Ramón A, Señorale-Pose M, Marín M (2014) Inclusion bodies: Not that bad... Front. Microbiol.
49. Basu A, Li X, Leong SSJ (2011) Refolding of proteins from inclusion bodies: Rational design and recipes. Appl. Microbiol. Biotechnol.
50. Burgess RR (2009) Chapter 17 Refolding Solubilized Inclusion Body Proteins. In: Methods in Enzymology

Supplementary Tables

Table S1

Percent of identity of the deduced amino acid sequence of chiRAM from *Enterobacter* sp. Strain EGY1 and other species

Species	Accession number	% Identity
Enterobacter cloacae	VAC58927.1	99.34
Ent.cloacae	VAE90674.1	97.69
Ent.cloacae	VAC36707.1	95.72
Ent.hormaechei	WP_047354344.1	94.59
Ent.cancerogenus	WP_080327729.1	94.55
Ent.chuandaensis	WP_119914984.1	94.22
Ent. bugandensis	WP_059358028.1	93.42
Ent.cloacae	WP_151412814.1	93.40
Ent.kobei	WP_088126968.1	93.26
Ent.asburiae	WP_148242681.1	92.41
Ent.mori	WP_126815644.1	92.08
Ent.huaxiensis	WP_125914156.1	91.91
Ent.ludwigii	WP_044857960.1	91.84
Citrobacter koseri	WP_130028211.1	89.33
Salmonella enterica	EBB0882654.1	74.01
Hevea brasiliensis	4MST_A	35.07
Secale cereale	4J0L_A	35.12

Table S2
Residue contacts of chiRAM to ligands and metals

Residue	Contact to ligand
Tyr106	Sulfate ion: SO^{-}_{4} & GOL: Glycerol
Thr107, Tyr108	Sulfate ion: SO^{-}_{4} & GOL: Glycerol & CL: Chloride ion
Ser109	GOL: Glycerol & CL: Chloride ion
Val115, Phe118, Pro119, ALa120, Leu121, Arg129, Ser155, Glu161, Asp205, Phe206, Arg212, Cys304, Gly305, Asn315, Val329, Pro330, Pro332, Asn334, Gln338, Cys339, Ala340, Asn341, Met342	CL: Chloride ion
Asp130, Gly180, Gly181, Asp189, Trp191, Trp196, Ser208, Phe210, Tyr326	Sulfate ion: SO^{-}_{4}
Ala133, Gly306	Sulfate ion: SO^{-}_{4} & CL: Chloride ion
Lys137	Sulfate ion: SO^{-}_{4} & MDP: (4s)-2-Methyl-2-4-Pentanediol
Thr150, Gly151, Gln262, Pro264, Gly301, Val302, Ala313,	NAG-NAG-NAG-NAG-: N-Acetyl -D-glucoseamine
Gly152	NAG-NAG-NAG-NAG-: N-Acetyl -D-glucoseamine & MDP: (4s)-2-Methyl-2-4-Pentanediol & EDO: 1,2 Ethanediol
His153	NAG-NAG-NAG-NAG-: N-Acetyl -D-glucoseamine & EDO: 1,2 Ethanediol
Glu154	NAG : N-Acetyl -D-glucoseamine & NAG-NAG-NAG-NAG-: N-Acetyl -D-glucoseamine & EDO: 1, 2 Ethanediol
Trp156	NAG : N-Acetyl -D-glucoseamine & NAG-NAG-NAG-NAG-: N-Acetyl -D-glucoseamine & CL: Chloride ion
Arg157	EDO: 1,2 Ethanediol
Trp162	NAG : N-Acetyl -D-glucoseamine & CL: Chloride ion
Lys179	NAG-NAG-NAG-NAG-: N-Acetyl -D-glucoseamine & NDG: 2-(Acetylamino)-2-Deoxy-A-D-Glucopyranose
Glu185	NAG-NAG-NAG-NAG-: N-Acetyl -D-glucoseamine

Residue	Contact to ligand
Lys215, Leu217, Ser218, Asn222, Phe257	NAG : N-Acetyl -D-glucoseamine & NAG-NAG-NAG-NAG-: N-Acetyl -D-glucoseamine & MES: 2-(N-Morpholino)-Ethanesulfonic acid
Gln216, Ile298, Asn299	NAG : N-Acetyl -D-glucoseamine & NAG-NAG-NAG-NAG-: N-Acetyl -D-glucoseamine & MES: 2-(N-Morpholino)-Ethanesulfonic acid & NDG: 2-(Acetylamino)-2-Deoxy-A-D-Glucopyranose
Tyr219	NAG : N-Acetyl -D-glucoseamine & NAG-NAG-NAG-NAG-: N-Acetyl -D-glucoseamine & Zn: Zinc ion
Tyr221	NAG : N-Acetyl -D-glucoseamine & NAG-NAG-NAG-NAG-: N-Acetyl -D-glucoseamine & Sulfate ion: SO^{-}_{4}
Lys265	NAG : N-Acetyl -D-glucoseamine & NAG-NAG-NAG-NAG-: N-Acetyl -D-glucoseamine
Ile272, Pro289	MPD: (4s)-2-Methyl-2,4-Pentanediol
Glu303, Gln312	NAG-NAG-NAG-NAG-: N-Acetyl -D-glucoseamine & NDG: 2-(Acetylamino)-2-Deoxy-A-D-Glucopyranose & CL: Chloride ion
Arg316	NAG-NAG-NAG-NAG-: N-Acetyl -D-glucoseamine & NDG: 2-(Acetylamino)-2-Deoxy-A-D-Glucopyranose
Lys318	NDG: 2-(Acetylamino)-2-Deoxy-A-D-Glucopyranose & CL: Chloride ion
Lys312	NDG: 2-(Acetylamino)-2-Deoxy-A-D-Glucopyranose

Figures

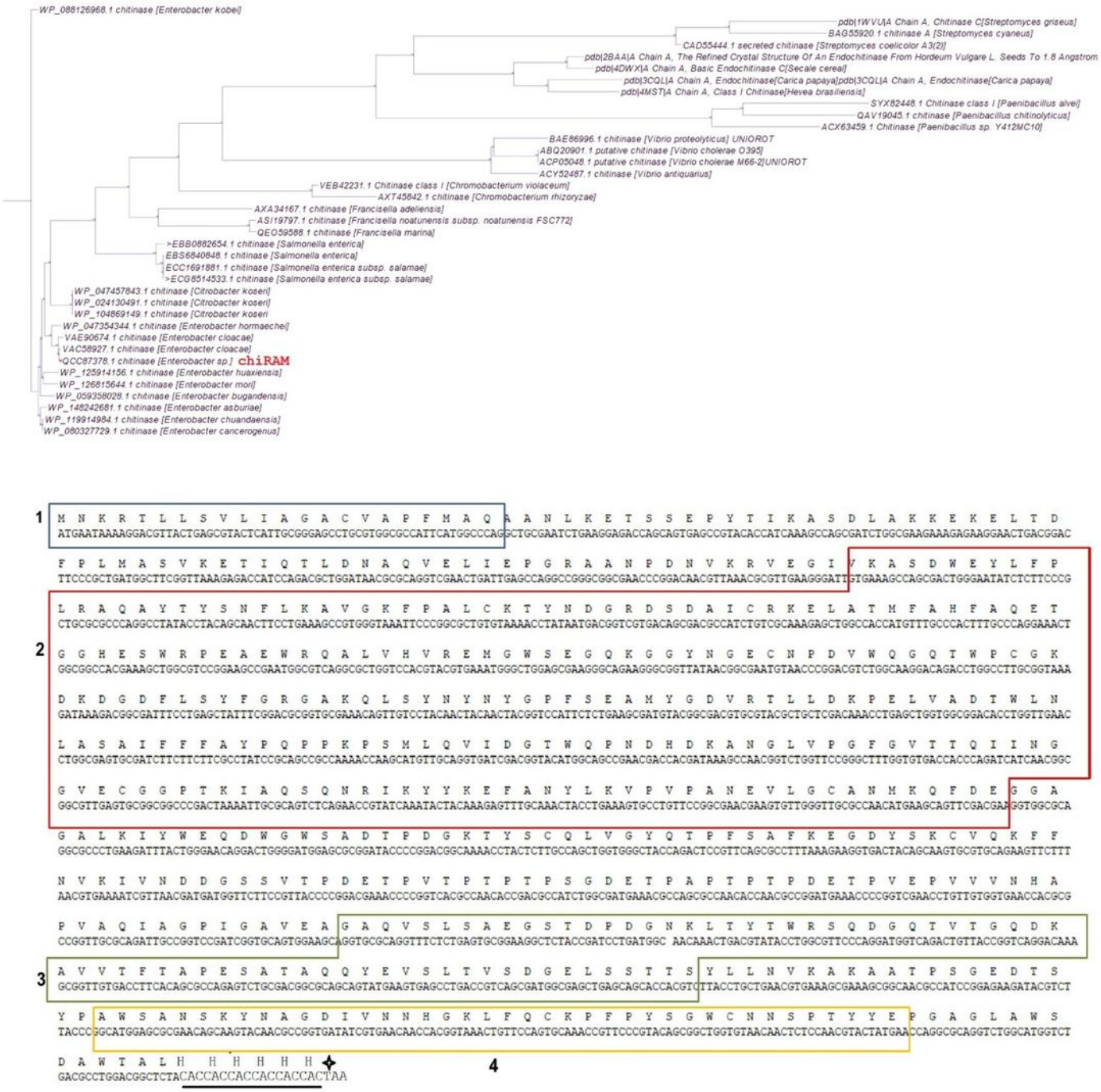


Figure 1

A Phylogenetic relationships of chiRAM in relation to GH19 chitinases from other species (bacteria and plants). The rooted phylogenetic tree was built up according to an alignment of full length amino acid sequences of all selected GH19 chitinases using the neighbor-joining method (Saitou and Nei 1987) employed in the ClustalW program, Jakes-Cantor model. Phylogenetic tree was built with CLC Sequence Viewer 8.0. The bootstrapping value was set to be 1000. The accession numbers of chitinases amino

acid sequences followed by the organisms' names were put in the tips of the branch. The bar indicates the branch length was 0.12. B The amino acid sequence (612 amino acids) of the recombinant chiRAM protein deduced from the translation of the amplified full length open reading frame (1821 bp) of chiRAM. Rectangle1: indicated the N-terminus signal peptide (Met1 to Ala23). Rectangle 2: indicated the catalytic domain (Val83 to Glu347). Rectangle 3: indicated the polycystic kidney disease protein motif (Gly 465 to Ser 533). Rectangle 4: indicated the C-terminus chitin-binding domain (Ala553 to Glu593). The underlined residues are the six histidine tags of the recombinant chiRAM. Star: indicated the position of the stop codon (TAA).

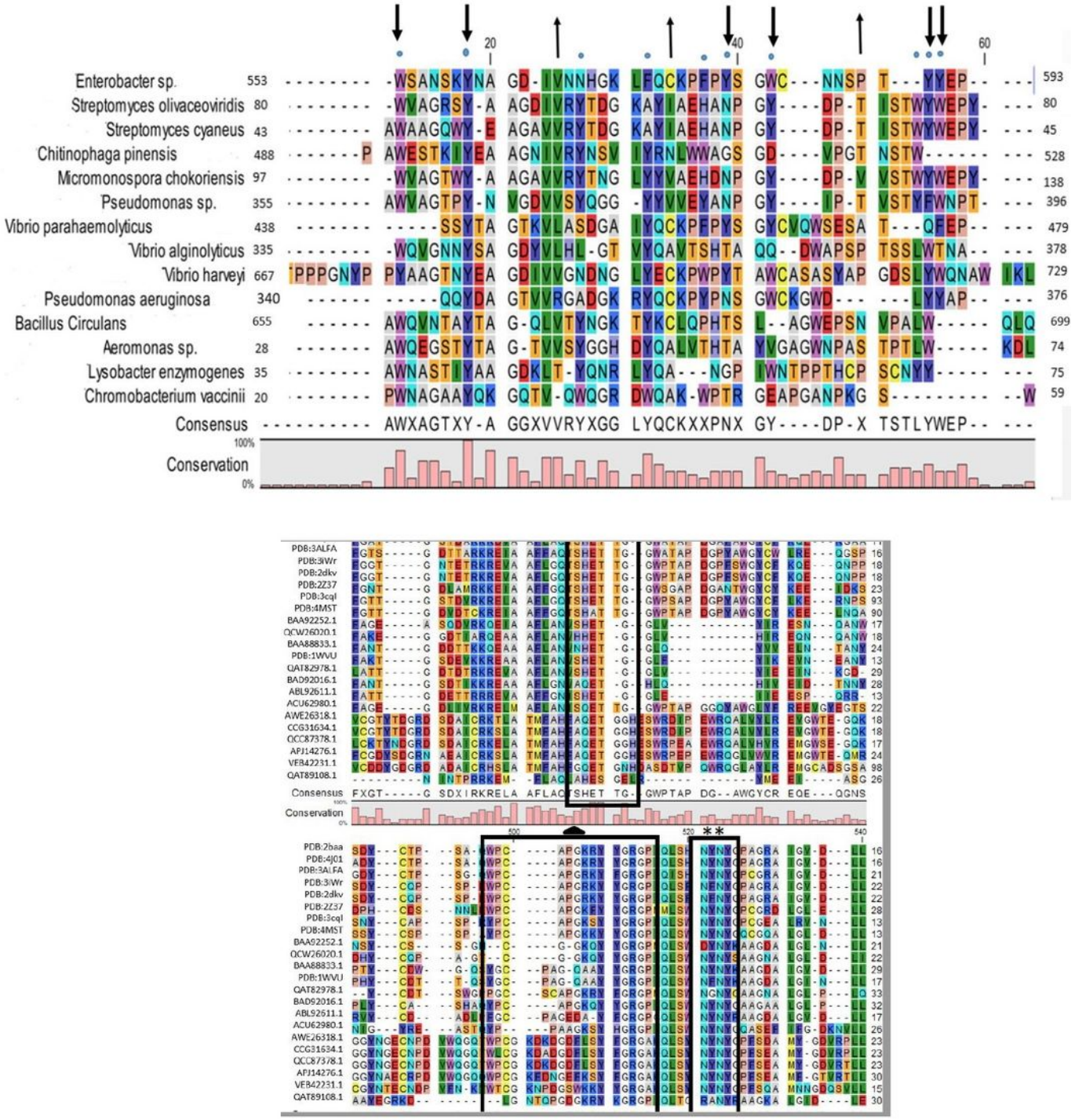


Figure 2

A Multiple sequence alignment of amino acid sequences for chitin binding domain of some members of GH19 chitinases showing the signature feature of chtBD3 super family in ChiA1_BD among amino acid sequences of different GH19 chitinases from bacteria performed by CLC Sequence Viewer 8.0 program. Filled black circles indicated the distribution of six aromatic amino acid residues among chitin binding domains of chitinases amino acid sequences of different bacterial species. Upward arrows indicated the

position of the six aromatic amino acid residues in the chitin binding domain of chiRAM: Trp554, Tyr560, Tyr580, Trp583, Try590, and Tyr591. Downward arrows indicated the position of three hydrophobic amino acids in the core of the chitin binding domain. B Multiple sequence alignment of amino acid sequences of GH19 chitinases from different species including chiRAM showing the evolutionary relationship and divergence in three conserved motifs NYNY, SHETTG , and WPCAPGRKYFGRGP, performed with CLC Sequence Viewer 8.0 program. The rectangles with symbols (*), (**), and (▲) indicated the conserved motifs NYNY, SHETTG, and WPCAPGRKYFGRGP, respectively. The following accession numbers: 2baa, 4j01, 3AFLA, 3iwr, 2kdv, 2Z37,3cq1, 4MST,BAA92252.1, QCW26020.1, BAA88833.1, 1WVU, QAT82978.1, BAD92016.1, ABL92611.1, ACU62980.1, AWE26318.1, CCG3164.1, QCC87378.1, APJ14276.1, VEB42241.1, and QAT89108.1 refer to chitinases from Hordeum vulgare, Secale cereale, Nicotiana tabacum, Oryza sativa subsp. Japonica, Oryza sativa subsp.Japonica, Brassica juncea, Carica papaya, Hevea brasilliensis, Burkhoderia gladioli, Lysobacter enzymogenes, Str.thermoviolaceus, Str.griseus, Corallococcus coralloides, Bacillus circulans, Mycobacterium sp.KMS, Chitinophaga pinensis DSM 2588, Salmonella enterica, Klebsiella aerogenes EA1509E, Enterobacter sp.strain EGY1, Aeromonas hydrophila, Chromobacterium violaceum, and Corallococcus corallaides, respectively.

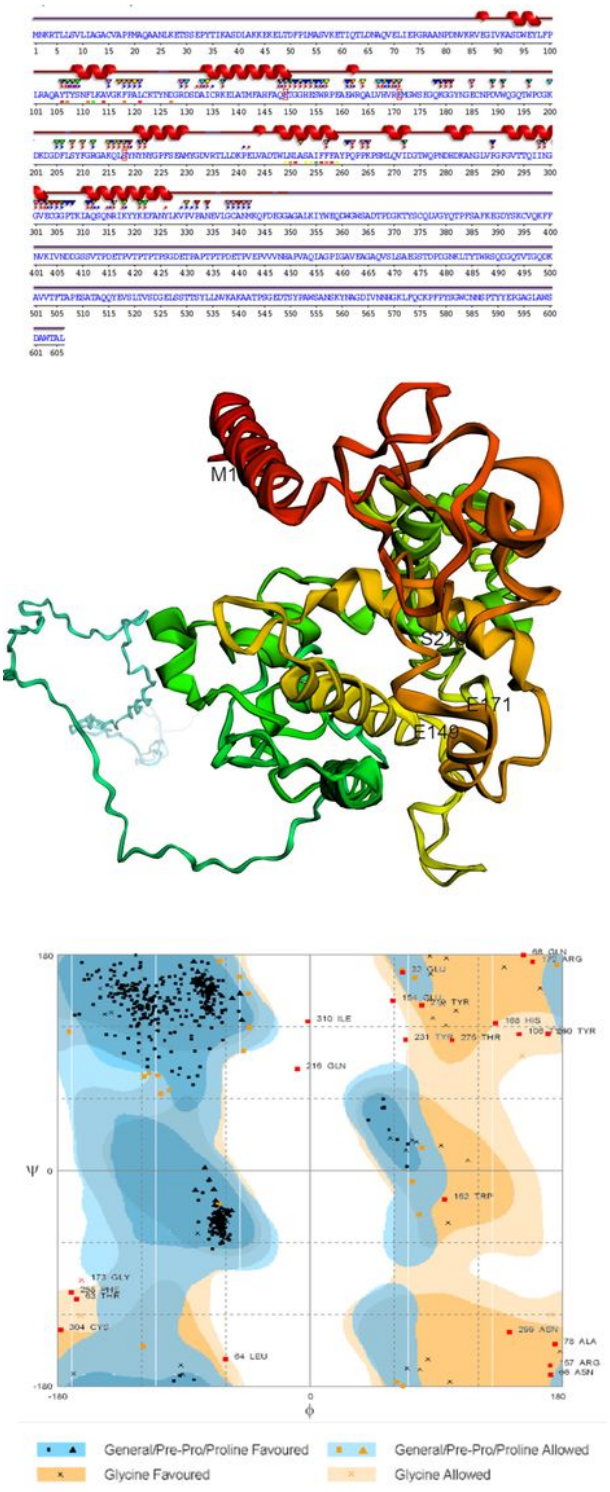
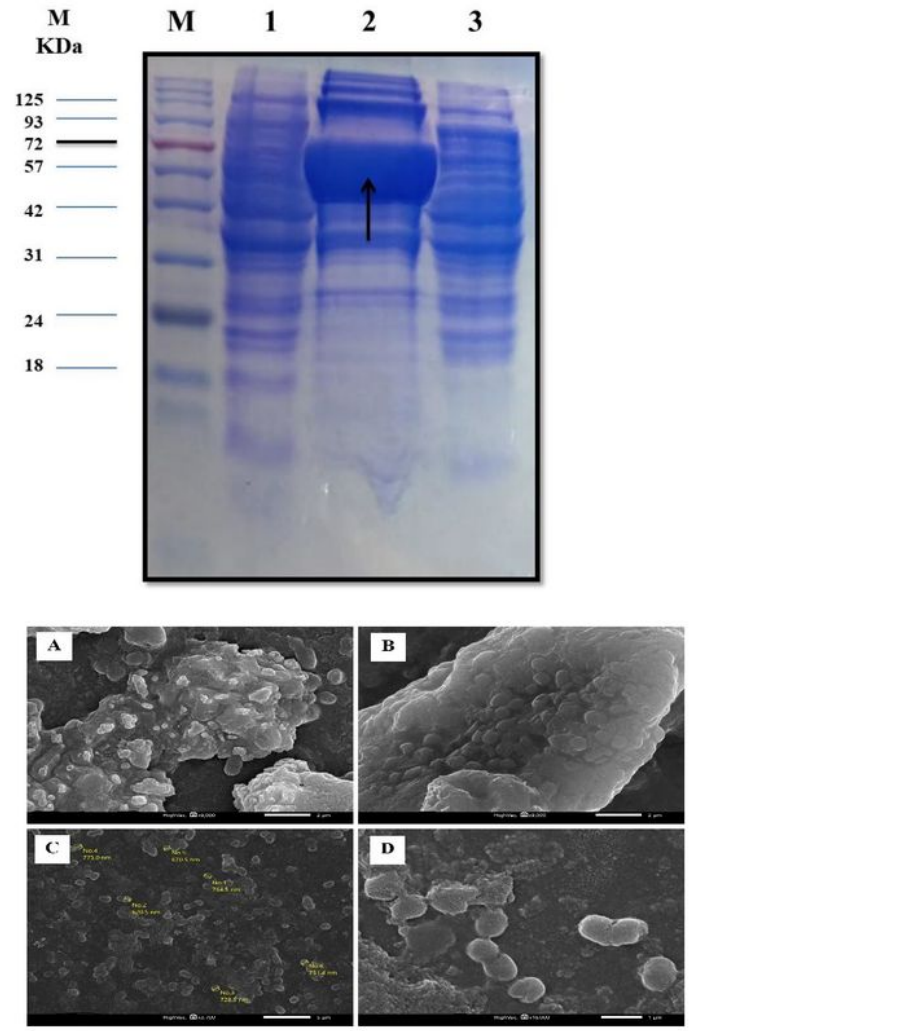


Figure 3

A Predicted secondary structure of chiRAM as determined by SAS online program using Carica papaya (PDB: 3cql:A) as a template. The red rectangle pointed out the position of amino acids included in the catalytic triad (Glu149- Glu171-Ser218) of chiRAM. Blue filled squares: residues contact to metals. Red filled squares: residues contact to ligands. Green triangles pointed out the active sites. B Cartoon representation of predicted 3D structure of chiRAM based on the template Oryza sativa Japonica Group

(PDB: 2dkvA) using the multiple threading approach (LOMETS) provided by I-TASSER server. The catalytic triad (Glu149, Glu171, and Ser218) was highlighted in black in both views. M1: refers to the first amino acid methionine in chiRAM amino acid sequence. C Ramachandran plot, for the predicted 3D chiRAM model. The number of residues in the favored region (~98.0% expected) was 558(92.4%), the number of residues in the allowed region (~2.0% expected) was 24 (4.0%) and the number of residues in the outlier region was 22(3.6%).



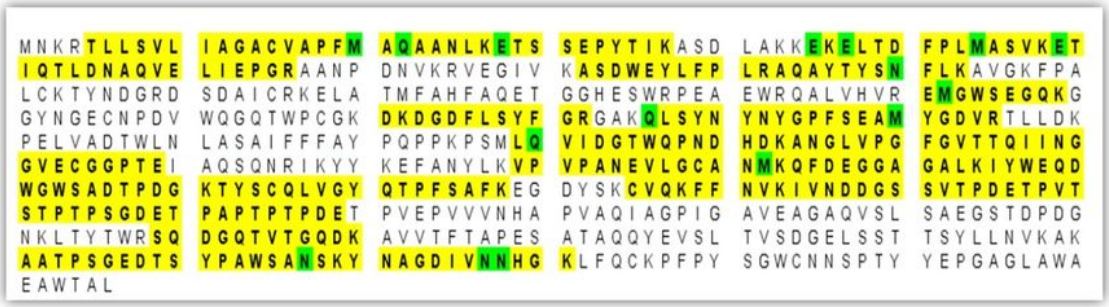


Figure 4

A Expression of chiRAM by E.coli BL21(DE3) Rosetta/ pET28a(+)/chiRAM with 1mM IPTG at room temperature. M: protein ladder. Lanes 1 and 2: soluble proteins and insoluble proteins of induced E.coli BL21 (DE3) Rosetta/ pET28a(+)/chiRAM, respectively. Arrow indicates the inclusion bodies of chiRAM. Lane 3: soluble proteins of uninduced E.coli BL21 (DE3) Rosetta / pET28a-chiRAM. Each lane contain 50µg protein of total protein. B Field scanning electron microscopy of chiRAM inclusion bodies. (A-D): Shows different pattern of inclusion bodies aggregation. A & B: scale bar of 2µm. C & D: scale bar of 5 and 1 µm, respectively. chiRAM inclusion bodies were spherical and oval in shape (A-D). C): Shows the diameter of chiRAM inclusion bodies ranged from 670.5 – 764.5 nm. C Amino acid sequence of the multispecies chitinase of Enterobacter cloacae complex (WP_063869339.1) with 52% sequence coverage with chiRAM amino acid sequence after tryptic digestion and separation with LC-MS-MS. The yellow color pointed out the location of the determined LC-MS-MS sequence peptides of chiRAM that were found in WP_063869339.1.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Fig.S1.png
- Fig.S2.png
- Fig.S3.png
- Fig.S4.png
- Fig.S5.png
- Fig.S6.png
- Fig.S7.png