Microbial Pathogenesis 110 (2017) 519–526

Contents lists available at ScienceDirect

Microbial Pathogenesis

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

In silico analysis of ChtBD3 domain to find its role in bacterial pathogenesis and beyond

Seema Patel [a, *], Abdur Rauf [b, **], Biswa Ranjan Meher [c]

[a] Bioinformatics and Medical Informatics Research Center, San Diego State University, San Diego, 92182, USA
[b] Department of Chemistry, University of Swabi, Anbar, 23561, Khyber Pakhtunkhwa, Pakistan
[c] Centre for Life Sciences, Central University of Jharkhand, Brambe, Ranchi, 835205, Jharkhand, India

ARTICLE INFO

Article history:
Received 26 December 2016
Received in revised form
24 July 2017
Accepted 27 July 2017
Available online 29 July 2017

Keywords:
Domain
ChtBD3
Pathogenesis
Glycosyl hydrolase
Chitinase
Peptidase

ABSTRACT

Chitin binding domain 3, known by the acronym ChtBD3, is a domain in the enzymes and proteins of several pathogenic virus, bacteria and fungi. As this domain is evolutionarily-conserved in virulence factors of these infectious agents, its detailed investigation is of clinical interest. In this regard, the current *in silico* study analyzed ChtBD3 domain distribution in bacterial proteins present in publicly-available SMART (simple modular architecture research tool) database. Also, the co-occurring domains of ChtBD3 in the studied proteins were mapped to understand positional rearrangement of the domain and consequent functional diversity. Custom-made scripts were used to interpret the data and to derive patterns. As expected, interesting results were obtained. ChtBD3 domain co-occurred with other critical domains like peptidase, glycol_hydrolase, kinase, hemagglutinin-acting, collagen-binding, among others. The findings are expected to be of clinical relevance.

1. Introduction

ChtBD3, a chitin-binding domain (ChtBD), occurring in pathogen and allergen proteins has been consistently associated with pathogenesis [1]. This domain occurs in chitinase, glycosidase (glucanase), related depolymerizing enzymes [2,3], or virulence proteins (as adhesins, lectins etc.) [4]. Chitinases (EC 3.2.1.14) are glycoside hydrolases (GH), belonging to the family GH18 and GH19 [5]. Chitin, the insoluble, linear β-1, 4-linked polymer of N-acetyl-glucosamine (GlcNAc) monosaccharide units, is the substrate of the enzyme chitinase. Chitin is the essential component of fungal (*Trichoderma, Cryptococcus* etc.) cell walls, nematodes, molluscs and insect exoskeletons (cuticle) and peritrophic membranes [6,7]. Also, it is food source for bacteria such as *Pseudomonas aeruginosa*, *Serratia marcescens*, *Vibrio harveyi*, *Bacillus circulans*, among others [8]. Chitinases from *Erwinia*, *Serratia*, *Bacillus* [9,10], *Vibrio* [11], *Pyrococcus* [12], *Streptomyces*, *Alteromonas* sp., and *Nocardiopsis* [13,14] have been well-studied. A study on *Bacillus circulans* chitinase (ChiA1) revealed that ChtBD is not indispensable for hydrolysis of chitin; however it enhances the degradation efficacy [9,15]. The bacterium *Francisella tularensis*, causative agent of tularemia, secretes chitinase for infecting the arthropod (ticks, mosquitoes etc.) vectors for disease transmission [16]. Malaria parasite (*Plasmodium* sp.) exploits this enzyme to traverse the chitin-containing peritrophic matrix of host mosquito, to gain access to its mid gut [17]. The cuticular proteins of malaria vector, the mosquito *Anopheles sinensis* has ChtBDs as well [18]. Plants such as rubber (*Hevea brasiliensis*) and tobacco (*Nicotiana tabacum*) elaborate this enzyme for defense purposes [19]. Active human chitinase includes chitotriosidase and acidic mammalian chitinase (AMCase) [5,20]. Human macrophage chitinase facilitates tissue remodeling, via the interaction with polysaccharides or extracellular matrix glycoproteins (Ujita et al., 2003). This enzyme has been detected in atherosclerotic plaques and serum of fungal-infected animal models. The degradation of the fungal chitin wall generates fragments, which elicit immune response, causing allergic reactions in atopic individuals [22]. The chitinolytic activity in human has defense purpose, as it lyses pathogen cell wall, and attempts to eliminate them. The dual role of chitinase in human pathogenesis is convincingly proved [7]. A recently-published literature review

* Corresponding author. Bioinformatics and Medical Informatics Research Center, San Diego State University, 5500 Campanile Dr, San Diego, CA 92182, USA.
** Corresponding author.
E-mail addresses: seemabiotech83@gmail.com (S. Patel), mashaljcs@yahoo.com (A. Rauf).

http://dx.doi.org/10.1016/j.micpath.2017.07.047
0882-4010/

discusses the pathogenic functions of chitinase and chitin [23].

ChtBD3 is present in virulent strains Ebola, HCV (Hepatitis C), and Flavivirus dengue [24]. Also, ChtBD3 is present in house dust mite and cockroach allergen [25]. Human macrophage-expressed chitinase has also ChtBD [21]. Though the domain location varies in proteins, the general topology of domains in chitinase is N terminal catalytic domain ChtBD, a fibronectin type III like domain (FN3) in the middle, and a C terminal CBM (carbohydrate-binding module) [10,26]. The general layout has been presented in Fig. 1. The conserved tryptophan residue in the ChtBD is assumed to be involved in hydrophobic interaction with the substrate [5,27]. Also, the conserved cysteine residues have been found responsible for interaction with the substrate [5]. ChtBD domain can be of several types, generally well-studied types are ChtBD1, ChtBD2 and ChtBD3. Some of the ChtBD3-containing proteins include ChiA, ChiB, ChiC, ChiD, Cbp1, and AprIV [28].

As chitinase enzyme is at the center of several infectious and allergic diseases, and the domain ChtBD3 is critical for its activity, this *in silico* work was pursued to garner insights. The findings are expected to contribute to the comprehension of pathogenesis and the manipulation of host immunity by ChtBD-containing proteins.

2. Material and methods

The *in silico* analysis conducted here, for the analysis of ChtBD3, can be fragmented into several steps. The details have been outlined below.

2.1. Protein retrieval from SMART platform

This investigation used ChtBD3 domain-containing bacterial proteins from the public platform SMART (simple modular architecture research tool) [29]. SMART identifies and annotates domains using HMMer (for alignment) and BLAST (for bit score) [29].

2.2. Custom scripts development for domain characterization

The downloaded data file was manipulated using custom-made Shell scripts. The distribution of ChtBD3 across bacteria species and the type of proteins containing this domain were determined. The scripts were constructed using the commands like awk, sort, grep, comm, and while loop. Based on the results of the output files, the critical co-occurring domains were discussed, and hypotheses were formulated, which is likely to be of relevance in better management of bacterial pathogen infections.

3. Results

About 670 ChtBD3-domain containing bacterial proteins were found in the SMART database, as of 15th May 2016. This domain occurs in repeats of 1–7. Other domains co-occurring with this domain included Cellulase, ChiC, Glyco_18, PKD, Glyco_hydro_19, Chitin_bind_3, Polysacc_deac_1, FN3, Tryp_SPc, Peptidase_S8, ChitinaseA_N, Mucin_bdg, Haemagg_act, Dioxygenase_C, Pro_Al_protease, SLT, Peptidase_M64, Peptidase_M66, Polysacc_deac_1, TagA, IG_like, Trypsin, Pro_Al_protease, PepSY, Peptidase_M4, Polysacc_deac_1, TSP_3, IG, CBM_6, BID_2, CBM_4_9, Phage_GPD, SKN1, Glyco_hydro_16, Fn3_assoc, CHB_HEX_C, 5_nucleotid_C, LTD, Hepar_II_III, He_PIG, CBM_3, CBM_10, PA14, GSDH, Peptidase_M28, FTP, XTALbg (Beta/gamma crystallins), Pkinase, PASTA, Exo_endo_phos, Alginate_lyase, ChitinaseA_N, CBM_5_12, CBM_X2, Peptidase_M66, Glyco_hydro_53, Glyco_10, Polysacc_deac_1, Flg_bbr_C, Glyco_hydro_46, Big_3, Collagen, and Metallophos.

Among the domains, PASTA and CBM_X2 occurred in repeats. Cellulose binding domain included CBM_3, CBM_4_9, CBM_5_12, CBM_6, CBM_10, and CBM_X2. Glyco_hydrolase (glycosyl hydrolase/GH) family domains included Glyco_10, Glyco_18, Glyco_hydro_19, Glyco_hydro_46, and Glyco_hydro_53. Peptidase domain families included Peptidase, M4, M28, M64, M66, and S8. These proteins are pathogenically-crucial and they contain critical domains like glycosyl hydrolase, protease, mucin binding, hemagglutinin binding, immunoglobin-like, heparin binding, and collagen binding, among others.

Also, transmembrane helices, coiled coil region, and signal peptide occurred variably in those proteins containing ChtBD3. Some DUFs (domains of unknown function) harbored in the proteins included DUF922, DUF3472, DUF3739 and DUF5011, of which DUF3472 and DUF5011 occurred frequently.

The bacterial species containing the domain ChtBD3 belonged to the genus such as *Acholeplasma*, *Achromobacter*, *Acidobacterium*, *Acidothermus*, *Actinomyces*, *Actinoplanes*, *Actinosynnema*, *Aeromonas*, *Agrobacterium*, *Aliivibrio*, *Amycolatopsis*, *Arthrobacter*, *Azorhizobium*, *Bacillus*, Beutenbergia, Bilophila, *Bradyrhizobium*, *Brevibacillus*, *Burkholderia*, *Butyrivibrio*, *Catenulispora*, *Cellulophaga*, *Cellvibrio*, *Chitinophaga*, *Chromobacterium*, *Chthoniobacter*, *Citrobacter*, *Clostridium*, *Collinsella*, *Conexibacter*, *Cronobacter*, *Desulfitobacterium*, *Dickeya*, *Enterobacter*, *Erwinia*, *Escherichia*, *Eubacterium*, *Ferrimonas*, *Francisella*, *Frankia*, *Gramella*, *Granulicella*,

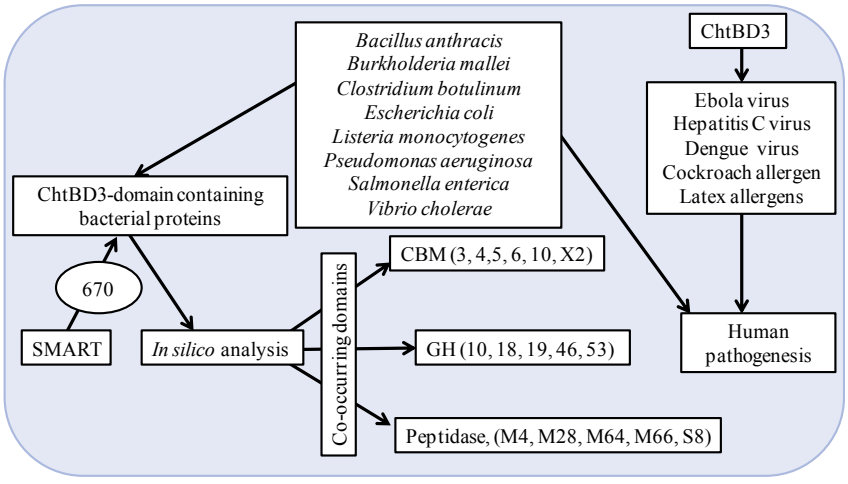

Fig. 1. ChtBD-containing pathogens, allergens and analysis-derived co-occurring domains.

Grimontia, *Hahella*, *Haliangium*, *Herpetosiphon*, *Isoptericola*, *Klebsiella*, *Lactobacillus*, *Listeria*, *Microbacterium*, *Micromonospora*, *Microscilla*, *Moritella*, *Mucilaginibacter*, *Mycobacterium*, *Myxococcus*, *Nakamurella*, *Niastella*, *Nocardiopsis*, *Oceanobacillus*, *Paenibacillus*, *Pantoea*, *Photobacterium*, Photorhabdus, *Pseudoalteromonas*, *Pseudomonas*, *Psychromonas*, *Rahnella*, *Reinekea*, *Rheinheimera*, *Ruminococcus*, *Saccharophagus*, *Salinispora*, *Salmonella*, *Sanguibacter*, *Serratia*, *Shewanella*, *Sodalis*, *Sorangium*, *Spirochaeta*, *Stackebrandtia*, *Stenotrophomonas*, *Stigmatella*, *Streptomyces*, *Streptosporangium*, *Teredinibacter*, *Thermoanaerobacterium*, *Thermobispora*, *Verrucosispora*, *Vibrio*, *Wolinella*, *Xanthobacter*, and *Xylanimonas*. Several members of important bacterial genus such as *Aeromonas*, *Bacillus*, *Listeria*, *Paenibacillus*, *Pseudomonas*, *Salmonella*, *Serratia*, *Shewanella*, *Streptomyces*, and *Vibrio* elaborate proteins containing this domain. Pathogenic species with this domain included *Bacillus anthracis* (anthrax), *Burkholderia mallei* (glanders), *Clostridium botulinum* (botulism/muscle paralysis), *Escherichia coli* (diarrhea and urinary tract infections), *Listeria monocytogenes* (listeriosis), *Pseudomonas aeruginosa* (bacteremia), *Salmonella enterica* (gastroenteritis), and *Vibrio cholerae* (cholera) etc.

Fig. 1 presents the list of ChtBD-containing pathogens and allergens. Also, the co-occurring domains of pathologic significance have been illustrated.

4. Discussion

The domains co-occurring with the ChtBD3 domain might or might not be essential for the latter's activity; however, they can enhance our understanding of ChtBD3 and its functionality. Hence, these domains have been briefly discussed with literature support. A huge majority of the co-occurring domains have been found in proteases, glycosyl hydrolases, though some are detected in other critical proteins and enzymes as well.

Peptidases can be of several types, predominant of which are chymotrypsin (PA), subtilisin (SB) and carboxypeptidase C (SC) [30]. These clans have the same catalytic triad of serine (S), aspartate (D) and histidine (H), which acts as nucleophile, electrophile, and base, respectively [31]. The order of amino acid residues of the catalytic triad in the chymotrypsin, subtilisin, and carboxypeptidase is HDS, DHS and SDH, respectively [31]. These proteases are secreted via secretory pathway or retained as vesicles in leukocyte (neutrophil) granules, for discharge during requirement (such as immune activation) [32]. Pro_Al_protease domain occurs in serine proteases as exopeptidase, endopeptidase, oligopeptidase, omega-peptidase, streptogrisin etc. [33]. Metalloproteases use metal ions (Zn, Ca, Co etc.) to bind to amino acids (such as His, Glu, Asp or Lys) arranged as diads or triads [34,35]. Common motif in metalloprotease includes an HExxH, a hallmark of Zincin superfamily of metalloprotease [36]. This motif has been detected in *Treponema pallidum* adhesin pallilysin [37] and *Listeria monocytogenes* thermolysin [36]. Peptidase_M4 is a family of metalloprotease class that includes thermolysin, bacillolysin, protealysin, aureolysin, pseudolysin and other bacterial endopeptidases [38]. PepSY domain has protease inhibitor function. It is found in propeptide (protein precursor) of family M4 peptidase, tethered to cell wall, secreted, or occurring as hypothetical proteins [39]. FTP (fungalysin/thermolysin propeptide) domain is present in bacterial M4 peptidase propeptide and the fungal M36 propeptide [40]. The propeptides are likely to show chaperone activity and prevent peptidase activation. Peptidase_M28 is a domain in aminopeptidase Y, belonging to clan MH [41]. Peptidase_M64 domain is characteristic of metallo-endopeptidases belonging to the MEROPS peptidase family M64, members of which include IgA peptidase. Peptidase_M66 is a zinc metalloprotease, detected in bacterial pathogens like enterohemorrhagic *Escherichia coli*, *Aeromonas hydrophila* and *Vibrio cholerae* [42]. Peptidase_M66 domain-containing metalloproteases include ADAMs (a disintegrin and metalloprotease), astacins, and matrixins [43]. Further information on protease hierarchy can be obtained from MEROPS database [44].

Enzymes hydrolyzing the glycosidic bonds constitute the enzyme group glycosyl hydrolases (GH) (EC 3.2.1). These enzyme families have been described in CAZy (CArbohydrate-Active EnZymes) database (http://www.cazy.org) [45]. Glycoside hydrolase family 5 (GH5) comprises many well-known enzymes such as endoglucanase, beta-mannanase, exo-1,3-glucanase, endo-1,6-glucanase, xylanase, endoglycoceramidase [46]. Glyco_10 domain occurs in members of glycoside hydrolase family 10 (GH10) which include xylanase, endo-1,3-beta-xylanase, cellobiohydrolase [47]. Glyco_hydro_16 domain occurs in the members of glycoside hydrolase family 16 (GH16), which includes lichenase, xyloglucan - xyloglucosyltransferase, agarase, kappa-carrageenase, endo-beta-1,3-glucanase, endo-beta-1,3-1,4-glucanase etc. [48]. Glyco_18 domain occurs in glycoside hydrolase family 18 (GH18), which comprises of chitinase, chitodextrinase and mycotoxins (as in *Kluyveromyces lactis*, *Candida sphaerica*) [49]. GH19 or Glyco_hydro_19 also comprises of chitinase, hydrolyzing the β-1, 4-N-acetyl-D-glucosamine linkages in chitin polymers and chitin oligosaccharides. Plants also elaborate it as a weapon to disrupt chitin wall of invaders such as fungi and insects. Glyco_hydro_46 domain is found in family 46 of the glycosyl hydrolase, characterized as chitosanase enzymes, which catalyze the endohydrolysis of β-1, 4-linkages between N-acetyl-D-glucosamine and D-glucosamine residues, in a partly acetylated chitosan. Glyco_hydro_53 is TIM (triosephosphateisomerase) barrel domain, found in family 53 of the glycosyl hydrolase, characterized as endo-1, 4- beta-galactanases. FN3 (Fibronectin type 3) domain is found in glycohydrolases (cellobiohydrolase) of bacterium like *Clostridium thermocellum* [50]. ChitinaseA_N domain is found in the bacterial chitinases and viral proteins. It is organized into a FN3 domain-like fold, comprising only of beta strands. This domain is predicted to interact with chitin. Alginate_lyase domain belongs to alginate lyases, which catalyze the depolymerization of alginates by β-elimination. Alginate is a 1-4-linked polysaccharide of β-D-mannuronic acid and α-L-guluronic acid, produced by some bacteria (*Azotobacter* and *Pseudomonas*) and algae (brown). Cellulase is a member of O-glycosyl hydrolases. Polysacc_deac_1 domain is found in polysaccharide deacetylase or chitooligosaccharide deacetylase (such as nodulation protein B from *Rhizobium*, chitin deacetylase of yeast) and endoxylanases.

CBMs (carbohydrate-binding module) are non-catalytic sites in the carbohydrate-active enzymes or glycosyl hydrolases (GHs), binding to cellulose or other related carbohydrates. CBMs comprise of contiguous amino acid sequence, with specific folds, occurring as modules within large enzymes [51]. Based on amino acid sequence configuration, CBMs have been classified into 71 families, as summarized in CAZy (Carbohydrate-Active enZymes) database (http://www.cazy.org) [45]. However, the repertoire is expanding. Some CBMs associated with ChtBD3 have been discussed here. CBM_3 is β sandwich domain in bacterial GHs and it is involved in cellulose binding [52,53]. CBM_4_9 is a module in 1, 4-β-glucanase, and arranged as two domains in N termini [54]. CBM_5_12 is found in proteins such as chitinase A1, chitinase B, and endoglucanase Z [55]. CBM_6 has a lectin-like β-jelly roll fold that binds to amorphous cellulose, xylan, mixed β-(1,3)(1,4) glucan, and β-1,3-glucan [56]. CBM_10 occurs in the proteins of aerobic bacteria and anaerobic fungi, where they bind to cellulose and proteins, respectively [57]. In the latter case, the domain is called dockerin which mediates multiprotein complex formation [58]. CBM_X2 has an immunoglobulin (Ig)-like fold and it binds to cellulose and to bacterial cell walls [59]. CBMs are involved in cellulosomes (multiprotein

glycosyl hydrolase complexes) formation as well [60].

IG domain occur in Ig (immunoglobulin), the highly modular proteins, with tetrameric structure made of two light chains (kappa and lambda) and two heavy chains (alpha, delta, epsilon, gamma and mu) linked by disulfide bonds [61]. IG-like (immunoglobulin-like) domains are involved in functions such as cell-cell recognition, cell-surface receptors, muscle structure and the immune system, among others [62]. LTD (lamin-tail domain) has an Ig fold which occurs in the proteins like nuclear lamins (a type of intermediate filaments forming nuclear lamina) and hydrolases. This domain interacts with actin, emerin, and SREBP1 (sterol regulatory element-binding protein 1), among others [63]. Mutation in lamins can lead to diseases like muscular dystrophy and Hutchinson-Gilford progeria syndrome [64–66]. BID_2 (Bacterial Ig-like domain 2) is an Ig-like domain in bacterial and phage surface proteins such as intimins, a bacterial cell-adhesion proteins with IG-like and C-type lectin-like domains [67]. CD73 occurs in C terminal of surface protein ecto-5′-nucleotidase, the glycosyl phosphatidylinositol-linked membrane-bound enzymes, catalyzing the dephosphorylation of purine and pyrimidine ribo- and deoxyribonucleoside monophosphates to their corresponding nucleosides [68]. This lymphocyte differentiation antigen CD73 protein is expressed on fibroblasts, and it regulates adenosine level (by controlling AMP to adenosine conversion) [69]. Thus it plays role as an immunomodulator, via the activation of adenosine receptors. Haemagg_act is a domain in an array of hemagglutinins, hemolysins, adhesins (TpsAs), and hemopexin-binding proteins [70]. Big_3 (bacterial Ig-like domain) is a domain with an Ig-like fold, occurring in a multitude of bacterial surface proteins, playing role in the interaction with host components [71]. He_PIG is a domain in haemagglutinins and cell surface proteins. This domain has Ig-like fold, so likely to be involved in immune response.

Collagens are extracellular structural proteins with repeats of G-X-Y (glycine, proline and hydroxyproline), and triple helix configurations [72]. This protein forms a component of connective tissue structure, abnormality of which causes several diseases, including osteogenesis imperfecta, Ehlers-Danlos syndrome, Alport syndrome, Bethlem myopathy, chondrodysplasias, epidermolysis bullosa, arterial aneurysms, and osteoporosis [73]. Hepar_II_III domain occurs in heparinases, the substrate-inducible heparin and heparan sulfate-degrading enzymes, elaborated by wide range of organisms [74]. TSP_3 is an aspartate residue-rich repeat in thrombospondin, an extracellular glycoprotein which binds to calcium ions and inhibits angiogenesis (by interfering with endothelial cell migration, proliferation, apoptosis etc.) [75,76]. This domain mediates interaction with glycosaminoglycans, calreticulin, integrins, and fibrinogen, during cellular adhesion, platelet aggregation etc. [77]. Flg_bbr_C (flagellar basal-body rod) is a C terminal domain in flagellar proteins [78]. XTALbg is a beta/gamma domain in crystallins, the water-soluble proteins in the cytoplasm of eye lens fiber cell. Crystallins are evolutionarily-related to stress proteins, and they can be of four types such as alpha-, beta-, gamma- (lacking in birds), and delta-(mostly in reptiles and birds) [79]. SKN1 is a domain in beta-glucan synthesis-associated protein [80]. Synthesis of β1,6-glucan starts in the endoplasmic reticulum (ER) and extends to Golgi complex via the glucosyltransferases. This protein is vital for cell wall stability, and it acts as virulence factors [81]. TagA is a bacterial lipoprotein [82]. Many bacterial lipoproteins (as in *Flavobacterium*) that serve in motility, contain chitin-utilizing domain [83]. SLT domain occurs in proteins from phages and secretion systems (type II- IV). Bacterial lytic transglucosylases (soluble (Slt) and membrane-bound type (Mlt)) degrade peptidoglycan, by cleaving the β-1, 4-glycosidic bond between N-acetylmuramic acid and N-acetylglucosamine [84]. Slt70 has N-terminal superhelical U-shaped domain, a superhelical linker L domain, and a C-terminal catalytic domain (has a lysosome-like fold). DUF3472, though sparsely-annotated, occurs with ChtBD3. This DUF has been reported in arthropod venoms [85].

GSDH domains occur in glucose/sorbosone dehydrogenases, a type of quinoproteins which oxidize glucose to gluconolactone. This enzyme is a calcium-dependent homodimer which uses PQQ (pyrroloquinoline quinone) as a cofactor [86]. 5_nucleotid_C occur in 5′-nucleotidases acting on nucleotides [87]. This enzyme is a membrane-bound glycoprotei,n which requires metal ions (Zn, Co) for activity. On activation, it hydrolyses UDP-glucose and other nucleotide diphosphate sugars. One of its homolog mosquito (*e.g. Aedes aegypti*) salivary gland apyrase (ATP-diphosphohydrolase) converts ATP into AMP, which-prevents platelet aggregation and facilitates hematophagy [88,89]. Polysacc_deac_1 domain is present in polysaccharide deacetylase. This enzyme deacetylates chitin and chito-oligosaccharide by acting on the carbon-nitrogen bond [90,91]. PA14, a β-barrel domain, occurs in an array of glycosidases, glycosyltransferases, proteases, amidases, yeast adhesins, bacterial toxins (anthrax protective antigen), amoeba proteins, and mammalian proteins (fibrocystin, whose mutation leads to polycystic kidney and hepatic disease) [92–94]. This domain is predicted to be involved in carbohydrate-binding. Pkinase is domain of protein kinases (such as PAKs), the enzymes that phosphorylate proteins (by transferring phosphate from ATP to amino acid), leading to their activation [95]. The phosphorylated protein has changed conformation, which affects protein function. The delicate balance between phosphorylated and dephosphorylated states is essential for survival. So, critical functions like cell division, proliferation, apoptosis, and differentiation rely on this enzyme. Protein kinases can be ramified as serine/threonine, tyrosine or dual specificity type (such as calcium-dependent protein kinases) [96]. PASTA (PBP and Serine/Threonine kinase Associated) is a sensory domain located at the C-terminal of Penicillin-binding proteins (PBP) and bacterial serine/threonine kinases [97,98]. A PASTA domain-rich protein includes PknB of *Mycobacterium tuberculosis* [99]. This domain binds to beta-lactam antibiotics and peptidoglycan [98]. Exo_endo_phos domain is present in magnesium-dependent endonucleases and phosphatases. It plays role in nucleic acid cleavage, protein inactivation, and intracellular signaling [100]. Metallophos domain is found in phosphoesterases *e.g.* bis (5′-nucleosyl)-tetraphosphatase (apaH), sphingomyelin phosphodiesterases, 2′-3′ cAMP phosphodiesterases, and nucleases [101]. Dioxygenase_C is the C-terminal domain of dioxygenases, the enzymes cleaving aromatic rings. The dioxygenases play key roles in the degradation of aromatic compounds [102]. Several of the dioxygenases form the intradiol family, to which bacterial proteins of clinical-significance such as adhesins, filamentous hemagglutinins, and hemopexin-binding protein belong [103]. Dioxygenase_C domain is likely to be a carbohydrate-dependent hemagglutination activity site, as observed in indoleamine 2,3-dioxygenase (IDO)-mediated immune modulation in dengue virus infected patient [104].

A number of glycosyl hydrolase domains co-occur together. Many clinically-relevant bacteria possess this domain. Previous analysis has shown the lethal viruses like HCV, dengue and some allergens to harbor this domain. It indicates the critical role of this domain in a broad form of pathogenesis. Those bacterial pathogens that do not occur in the databases might not have been analyzed for this domain yet, or their culture conditions did not enable the expression of these proteins. Many virulence factors are enzymes, which are produced in substrate-, temperature-, pH-, and cofactor-dependent manner. In fact, the lack of ambient conditions prevents elaboration of some proteins, which lead to the lack of experimental evidence, and the accumulation of hypothetical proteins in the *in silico*-predicted databases [105–107]. Also, it is likely that

some homolog of this domain might be playing the pathogenic role. Adhesins, the surface proteins of infectious agents provoke human platelet activation and host cell adhesion [108]. A large repertoire of inflammatory conditions has been associated with adhesins. We hypothesize that ChtBD or close homologs might be present in the adhesins.

Proteases and glycosyl hydrolases are not completely distinct but share several common domains. It makes the annotations of proteins more complex and the current enzymatic classification criteria defective.

Lytic polysaccharide monooxygenases are important enzymes for the decomposition of polymers such as chitin [109]. Deacetylase enzyme catalyzes the deacetylation of oligosaccharides [110]. It explains the presence of oxygenase and deacetylase domains co-occurring with the ChtBD3 domain.

Plants produce peptides (such as hevein) that contain ChtBDs [111]. Plant cells sense chitin oligosaccharides for defense signaling through a plasma membrane receptor CERK1 protein (with three LysM motifs in the extracellular domain, and an intracellular Ser/Thr kinase intracellular domain having autophosphorylation/myelin basic protein kinase activity) [112,113]. A study reports that the conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants [114]. As many bacterial pathogens have proteins with this domain, it is likely that they might be inhibiting chitin-mediated innate immune responses in human.

From the findings, it can be gathered that all living organisms have ChtBDs, though few might have lost them during the course of evolution or modified into other forms by gene rearrangement-driven domain swapping [115]. ChtBD-containing proteins are critical for pathogenesis. This domain enhances substrate binding and thus promotes protease and glycosidase activity of the enzyme. Within same species of a pathogen, virulence strengths of different strains vary, which can be explained by the presence or loss of the ChtBDs or their different positions in the protein. In fact, this claim has been proven by *in silico* analysis of several Ebola and HCV virus strains [116,117].

Like the translations step of protein formation, the post-translational modifications are critical for gene expression. A number of critical host proteins in their glycosylated form might be in inactive forms. The cleavage of the glycosyl moieties of these conjugated proteins, by the pathogen and allergen glycosyl hydrolases, is assumed to activate the proteins. The resultant perturbed protein cascade is likely to be leading to an array of diseases, including cancer and autoimmune diseases. So, it is suggested that ChtBD3-containing proteins, even if mutated to or annotated as hypothetical proteins due to various stresses (by drugs, or *in vitro* culture conditions), are hydrolases and thus potential drug targets. ChtBD3 as well as other ChtBD domains can be used as a motif to identify potential virulence factors.

5. Conclusion

This *in silico* study generated useful information on the pathogenically-critical domain ChtBD3. This domain occurring in a wide range of glycosyl hydrolases, binds to chitins and chito-oligosaccharides, thus promoting their hydrolysis by the enzyme catalytic site. The fierce evolutionary-retention of this domain and its clustering with some other virulence-associated domains can be considered as a clue to probe further into pathogenicity and strain evolutions.

Declaration

There is no conflict of interest in submission of this manuscript.

References

[1] H.T. Tran, N. Barnich, E. Mizoguchi, Potential role of chitinases and chitin-binding proteins in host-microbial interactions during the development of intestinal inflammation, Histol. Histopathol. 26 (2011) 1453–1464. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3181078&tool=pmcentrez&rendertype=abstract. (Accessed 25 May 2016).

[2] S. Javed, M. Ahmad, M. Ahmad, M. Abdin, R. Hamid, M. Khan, J. Musarrat, Chitinases: an update, J. Pharm. Bioallied Sci. 5 (2013) 21, http://dx.doi.org/10.4103/0975-7406.106559.

[3] R. Hamid, M.A. Khan, M. Ahmad, M.M. Ahmad, M.Z. Abdin, J. Musarrat, S. Javed, Chitinases: an update, J. Pharm. Bioallied Sci. 5 (2013) 21–29, http://dx.doi.org/10.4103/0975-7406.106559.

[4] Y. Itakura, S. Nakamura-Tsuruta, J. Kominami, H. Tateno, J. Hirabayashi, Sugar-binding profiles of chitin-binding lectins from the hevein family: a comprehensive study, Int. J. Mol. Sci. 18 (2017) 1160, http://dx.doi.org/10.3390/ijms18061160.

[5] F. Fadel, Y. Zhao, A. Cousido-Siah, F.X. Ruiz, A. Mitschler, A. Podjarny, X-Ray crystal structure of the full length human chitotriosidase (CHIT1) reveals features of its chitin binding domain, PLoS One 11 (2016) e0154190, http://dx.doi.org/10.1371/journal.pone.0154190.

[6] G. Tetreau, N.T. Dittmer, X. Cao, S. Agrawal, Y.-R. Chen, S. Muthukrishnan, J. Haobo, G.W. Blissard, M.R. Kanost, P. Wang, Analysis of chitin-binding proteins from Manduca sexta provides new insights into evolution of peritrophin A-type chitin-binding domains in insects, Insect Biochem. Mol. Biol. 62 (2015) 127–141, http://dx.doi.org/10.1016/j.ibmb.2014.12.002.

[7] L.W. Tjoelker, L. Gosting, S. Frey, C.L. Hunter, H. Le Trong, B. Steiner, H. Brammer, P.W. Gray, Structural and functional definition of the human chitinase chitin-binding domain, J. Biol. Chem. 275 (2000) 514–520, http://dx.doi.org/10.1074/jbc.275.1.514.

[8] J. Folders, J. Algra, M.S. Roelofs, L.C. van Loon, J. Tommassen, W. Bitter, Characterization of Pseudomonas aeruginosa chitinase, a gradually secreted protein, J. Bacteriol. 183 (2001) 7044–7052, http://dx.doi.org/10.1128/JB.183.24.7044-7052.2001.

[9] T. Watanabe, Y. Ito, T. Yamada, M. Hashimoto, S. Sekine, H. Tanaka, The roles of the C-terminal domain and type III domains of chitinase A1 from Bacillus circulans WL-12 in chitin degradation, J. Bacteriol. 176 (1994) 4465–4472. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=196264&tool=pmcentrez&rendertype=abstract. (Accessed 11 April 2016).

[10] Y. Zhan, S. Guo, Three-dimensional (3D) structure prediction and function analysis of the chitin-binding domain 3 protein HD73_3189 from Bacillus thuringiensis HD73, Biomed. Mater. Eng. 26 (Suppl 1) (2015) S2019–S2024, http://dx.doi.org/10.3233/BME-151506.

[11] A.L. Svitil, D.L. Kirchman, A chitin-binding domain in a marine bacterial chitinase and other microbial chitinases: implications for the ecology and evolution of 1,4-beta-glycanases, Microbiology 144 (Pt 5) (1998) 1299–1308, http://dx.doi.org/10.1099/00221287-144-5-1299.

[12] S. Mine, T. Nakamura, T. Sato, T. Ikegami, K. Uegaki, Solution structure of the chitin-binding domain 1 (ChBD1) of a hyperthermophilic chitinase from Pyrococcus furiosus, J. Biochem. 155 (2014) 115–122, http://dx.doi.org/10.1093/jb/mvt104.

[13] H. Tsujibo, H. Orikoshi, N. Baba, M. Miyahara, K. Miyamoto, M. Yasuda, Y. Inamori, Identification and characterization of the gene cluster involved in chitin degradation in a marine bacterium, Alteromonas sp strain O-7, Appl. Environ. Microbiol. 68 (2002) 263–270. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=126582&tool=pmcentrez&rendertype=abstract. (Accessed 16 May 2016).

[14] H. Tsujibo, T. Kubota, M. Yamamoto, K. Miyamoto, Y. Inamori, Characterization of chitinase genes from an alkaliphilic actinomycete, Nocardiopsis prasina OPC-131, Appl. Environ. Microbiol. 69 (2003) 894–900. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=143619&tool=pmcentrez&rendertype=abstract. (Accessed 16 May 2016).

[15] T. Ikegami, T. Okada, M. Hashimoto, S. Seino, T. Watanabe, M. Shirakawa, Solution structure of the chitin-binding domain of Bacillus circulans WL-12 chitinase A1, J. Biol. Chem. 275 (2000) 13654–13661. http://www.ncbi.nlm.nih.gov/pubmed/10788483. (Accessed 14 May 2016).

[16] A.J. Hager, D.L. Bolton, M.R. Pelletier, M.J. Brittnacher, L.A. Gallagher, R. Kaul, S.J. Skerrett, S.I. Miller, T. Guina, Type IV pili-mediated secretion modulates Francisella virulence, Mol. Microbiol. 62 (2006) 227–237, http://dx.doi.org/10.1111/j.1365-2958.2006.05365.x.

[17] S. Takeo, D. Hisamori, S. Matsuda, J. Vinetz, J. Sattabongkot, T. Tsuboi, Enzymatic characterization of the Plasmodium vivax chitinase, a potential malaria transmission-blocking target, Parasitol. Int. 58 (2009) 243–248, http://dx.doi.org/10.1016/j.parint.2009.05.002.

[18] D. Zhou, B. Duan, Y. Sun, L. Ma, C. Zhu, B. Shen, Preliminary characterization of putative structural cuticular proteins in the malaria vector *Anopheles sinensis*, Pest Manag. Sci. (2017), http://dx.doi.org/10.1002/ps.4649.

[19] S. Martínez-Caballero, P. Cano-Sánchez, I. Mares-Mejía, A.G. Díaz-Sánchez, M.L. Macías-Rubalcava, J.A. Hermoso, A. Rodríguez-Romero, Comparative study of two GH19 chitinase-like proteins from Hevea brasiliensis, one exhibiting a novel carbohydrate-binding domain, FEBS J. 281 (2014)

4535–4554, http://dx.doi.org/10.1111/febs.12962.
[20] O. Crasson, G. Courtade, R.R. Léonard, F.L. Aachmann, F. Legrand, R. Parente, D. Baurain, M. Galleni, M. Sørlie, M. Vandevenne, Human chitotriosidase: catalytic domain or carbohydrate binding module, Who's leading HCHT's biological function, Sci. Rep. 7 (2017) 2768, http://dx.doi.org/10.1038/s41598-017-02382-z.
[21] M. Ujita, K. Sakai, K. Hamazaki, M. Yoneda, S. Isomura, A. Hara, Carbohydrate binding specificity of the recombinant chitin-binding domain of human macrophage chitinase, Biosci. Biotechnol. Biochem. 67 (2003) 2402–2407. http://www.ncbi.nlm.nih.gov/pubmed/14646200. (Accessed 14 May 2016).
[22] S.M. Levitz, Innate recognition of fungal cell walls, PLoS Pathog. 6 (2010) e1000758, http://dx.doi.org/10.1371/journal.ppat.1000758.
[23] S. Patel, A. Goyal, Chitin and chitinase: role in pathogenicity, allergenicity and health, Int. J. Biol. Macromol. 97 (2017) 331–338, http://dx.doi.org/10.1016/j.ijbiomac.2017.01.042.
[24] J.P. Messina, O.J. Brady, T.W. Scott, C. Zou, D.M. Pigott, K.A. Duda, S. Bhatt, L. Katzelnick, R.E. Howes, K.E. Battle, C.P. Simmons, S.I. Hay, Global spread of dengue virus types: mapping the 70 year history, Trends Microbiol. 22 (2014) 138–146, http://dx.doi.org/10.1016/j.tim.2013.12.011.
[25] C.-R. Shen, H.-H. Juang, H.-S. Chen, C.-J. Yang, C.-J. Wu, M.-H. Lee, Y.-S. Hwang, M.-L. Kuo, Y.-S. Chen, J.-K. Chen, C.-L. Liu, The correlation between chitin and acidic mammalian chitinase in animal models of allergic asthma, Int. J. Mol. Sci. 16 (2015) 27371–27377, http://dx.doi.org/10.3390/ijms161126033.
[26] F. Uni, S. Lee, R. Yatsunami, T. Fukui, S. Nakamura, Mutational analysis of a CBM family 5 chitin-binding domain of an alkaline chitinase from Bacillus sp J813, Biosci. Biotechnol. Biochem. 76 (2012) 530–535, http://dx.doi.org/10.1271/bbb.110835.
[27] S. Ferrandon, T. Sterzenbach, F.B. Mersha, M.-Q. Xu, A single surface tryptophan in the chitin-binding domain from Bacillus circulans chitinase A1 plays a pivotal role in binding chitin and can be modified to create an elutable affinity tag, Biochim. Biophys. Acta 1621 (2003) 31–40. http://www.ncbi.nlm.nih.gov/pubmed/12667608. (Accessed 14 May 2016).
[28] H. Orikoshi, S. Nakayama, K. Miyamoto, C. Hanato, M. Yasuda, Y. Inamori, H. Tsujibo, Roles of four chitinases (chia, chib, chic, and chid) in the chitin degradation system of marine bacterium Alteromonas sp strain O-7, Appl. Environ. Microbiol. 71 (2005) 1811–1815, http://dx.doi.org/10.1128/AEM.71.4.1811-1815.2005.
[29] C.P. Ponting, J. Schultz, F. Milpetz, P. Bork, SMART: identification and annotation of domains from signalling and extracellular protein sequences, Nucleic Acids Res. 27 (1999) 229–232, http://dx.doi.org/10.1093/nar/27.1.229.
[30] M.B. Rao, A.M. Tanksale, M.S. Ghatge, V. V Deshpande, Molecular and biotechnological aspects of microbial proteases, Microbiol. Mol. Biol. Rev. 62 (1998) 597–635. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=98927&tool=pmcentrez&rendertype=abstract. (Accessed 28 March 2016).
[31] O.D. Ekici, M. Paetzel, R.E. Dalbey, Unconventional serine proteases: variations on the catalytic Ser/His/Asp triad configuration, Protein Sci. 17 (2008) 2023–2037, http://dx.doi.org/10.1110/ps.035436.108.
[32] S.M. Uriarte, D.W. Powell, G.C. Luerman, M.L. Merchant, T.D. Cummins, N.R. Jog, R.A. Ward, K.R. McLeish, Comparison of proteins expressed on secretory vesicle membranes and plasma membranes of human neutrophils, J. Immunol. 180 (2008) 5575–5581, http://dx.doi.org/10.4049/jimmunol.180.8.5575.
[33] N.N. Nickerson, L. Prasad, L. Jacob, L.T. Delbaere, M.J. McGavin, Activation of the SspA serine protease zymogen of Staphylococcus aureus proceeds through unique variations of a trypsinogen-like mechanism and is dependent on both autocatalytic and metalloprotease-specific processing, J. Biol. Chem. 282 (2007) 34129–34138, http://dx.doi.org/10.1074/jbc.M705672200.
[34] S. Ravaud, P. Gouet, R. Haser, N. Aghajari, Probing the role of divalent metal ions in a bacterial psychrophilic metalloprotease: binding studies of an enzyme in the crystalline state by x-ray crystallography, J. Bacteriol. 185 (2003) 4195–4203. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=164877&tool=pmcentrez&rendertype=abstract. (Accessed 16 May 2016).
[35] H. Yang, K. Makaroff, N. Paz, M. Aitha, M.W. Crowder, D.L. Tierney, Metal ion dependence of the matrix Metalloproteinase-1 mechanism, Biochemistry 54 (2015) 3631–3639, http://dx.doi.org/10.1021/acs.biochem.5b00379.
[36] A.P. Bitar, M. Cao, H. Marquis, The metalloprotease of Listeria monocytogenes is activated by intramolecular autocatalysis, J. Bacteriol. 190 (2008) 107–111, http://dx.doi.org/10.1128/JB.00852-07.
[37] S. Houston, R. Hof, L. Honeyman, J. Hassler, C.E. Cameron, Activation and proteolytic activity of the Treponema pallidum metalloprotease, pallilysin, PLoS Pathog. 8 (2012) e1002822, http://dx.doi.org/10.1371/journal.ppat.1002822.
[38] A. de Kreij, G. Venema, B. van den Burg, Substrate specificity in the highly heterogeneous M4 peptidase family is determined by a small subset of amino acids, J. Biol. Chem. 275 (2000) 31115–31120, http://dx.doi.org/10.1074/jbc.M003889200.
[39] C. Yeats, N.D. Rawlings, A. Bateman, The PepSY domain: a regulator of peptidase activity in the microbial environment? Trends Biochem. Sci. 29 (2004) 169–172, http://dx.doi.org/10.1016/j.tibs.2004.02.004.
[40] N.N. Nickerson, V. Joag, M.J. McGavin, Rapid autocatalytic activation of the M4 metalloprotease aureolysin is controlled by a conserved N-terminal fungalysin-thermolysin-propeptide domain, Mol. Microbiol. 69 (2008) 1530–1543, http://dx.doi.org/10.1111/j.1365-2958.2008.06384.x.
[41] M.B. Suárez, J.A. Vizcaíno, A. Llobell, E. Monte, Characterization of genes encoding novel peptidases in the biocontrol fungus Trichoderma harzianum CECT 2413 using the TrichoEST functional genomics approach, Curr. Genet. 51 (2007) 331–342, http://dx.doi.org/10.1007/s00294-007-0130-5.
[42] R.L. Szabady, J.H. Yanta, D.K. Halladin, M.J. Schofield, R.A. Welch, TagA is a secreted protease of Vibrio cholerae that specifically cleaves mucin glycoproteins, Microbiology 157 (2011) 516–525, http://dx.doi.org/10.1099/mic.0.044529-0.
[43] F.X. Gomis-Rüth, Catalytic domain architecture of metzincin metalloproteases, J. Biol. Chem. 284 (2009) 15353–15357, http://dx.doi.org/10.1074/jbc.R800069200.
[44] N.D. Rawlings, A.J. Barrett, R. Finn, Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors, Nucleic Acids Res. (2015), http://dx.doi.org/10.1093/nar/gkv1118.
[45] V. Lombard, H. Golaconda Ramulu, E. Drula, P.M. Coutinho, B. Henrissat, The carbohydrate-active enzymes database (CAZy) in 2013, Nucleic Acids Res. 42 (2014) D490–D495, http://dx.doi.org/10.1093/nar/gkt1178.
[46] H. Aspeborg, P.M. Coutinho, Y. Wang, H. Brumer, B. Henrissat, Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5), BMC Evol. Biol. 12 (2012) 186, http://dx.doi.org/10.1186/1471-2148-12-186.
[47] T.T. Do, D.T. Quyen, T.N. Nguyen, V.T. Nguyen, Molecular characterization of a glycosyl hydrolase family 10 xylanase from Aspergillus Niger, Protein Expr. Purif. 92 (2013) 196–202, http://dx.doi.org/10.1016/j.pep.2013.09.011.
[48] S.M. Linton, M.S. Cameron, M.C. Gray, J.A. Donald, R. Saborowski, M. von Bergen, J.M. Tomm, B.J. Allardyce, A glycosyl hydrolase family 16 gene is responsible for the endogenous production of β-1,3-glucanases within decapod crustaceans, Gene 569 (2015) 203–217, http://dx.doi.org/10.1016/j.gene.2015.05.056.
[49] G.D. Tzelepis, P. Melin, D.F. Jensen, J. Stenlid, M. Karlsson, Functional analysis of glycoside hydrolase family 18 and 20 genes in Neurospora crassa, Fungal Genet. Biol. 49 (2012) 717–730, http://dx.doi.org/10.1016/j.fgb.2012.06.013.
[50] T. Kraushaar, S. Brückner, M. Veelders, D. Rhinow, F. Schreiner, R. Birke, A. Pagenstecher, H.-U. Mösch, L.-O. Essen, Interactions by the fungal Flo11 adhesin depend on a fibronectin type III-like adhesin domain girdled by aromatic bands, Structure 23 (2015) 1005–1017, http://dx.doi.org/10.1016/j.str.2015.03.021.
[51] H. Hashimoto, Recent structural studies of carbohydrate-binding modules, Cell. Mol. Life Sci. 63 (2006) 2954–2967, http://dx.doi.org/10.1007/s00018-006-6195-3.
[52] P.-H. Ferdinand, R. Borne, V. Trotter, S. Pagès, C. Tardif, H.-P. Fierobe, S. Perret, Are cellulosome scaffolding protein CipC and CBM3-containing protein HycP, involved in adherence of Clostridium cellulolyticum to cellulose? PLoS One 8 (2013) e69360, http://dx.doi.org/10.1371/journal.pone.0069360.
[53] C.R. Santos, J.H. Paiva, M.L. Sforça, J.L. Neves, R.Z. Navarro, J. Cota, P.K. Akao, Z.B. Hoffmam, A.N. Meza, J.H. Smetana, M.L. Nogueira, I. Polikarpov, J. Xavier-Neto, F.M. Squina, R.J. Ward, R. Ruller, A.C. Zeri, M.T. Murakami, Dissecting structure-function-stability relationships of a thermostable GH5-CBM3 cellulase from Bacillus subtilis 168, Biochem. J. 441 (2012) 95–104, http://dx.doi.org/10.1042/BJ20110869.
[54] B.R. Urbanowicz, C. Catalá, D. Irwin, D.B. Wilson, D.R. Ripoll, J.K.C. Rose, A tomato endo-beta-1,4-glucanase, SlCel9C1, represents a distinct subclass with a new family of carbohydrate binding modules (CBM49), J. Biol. Chem. 282 (2007) 12066–12074, http://dx.doi.org/10.1074/jbc.M607925200.
[55] E. Garcia-Gonzalez, L. Poppinga, A. Fünfhaus, G. Hertlein, K. Hedtke, A. Jakubowska, E. Genersch, Paenibacillus larvae chitin-degrading protein PlCBP49 is a key virulence factor in American Foulbrood of honey bees, PLoS Pathog. 10 (2014) e1004284, http://dx.doi.org/10.1371/journal.ppat.1004284.
[56] M. Jam, E. Ficko-Blean, A. Labourel, R. Larocque, M. Czjzek, G. Michel, Unraveling the multivalent binding of a marine family 6 carbohydrate-binding module with its native laminarin ligand, FEBS J. (2016), http://dx.doi.org/10.1111/febs.13707.
[57] S. Raghothama, P.J. Simpson, L. Szabó, T. Nagy, H.J. Gilbert, M.P. Williamson, Solution structure of the CBM10 cellulose binding module from Pseudomonas xylanase A, Biochemistry 39 (2000) 978–984. http://www.ncbi.nlm.nih.gov/pubmed/10653641. (Accessed 16 May 2016).
[58] X. Dai, Y. Tian, J. Li, Y. Luo, D. Liu, H. Zheng, J. Wang, Z. Dong, S. Hu, L. Huang, Metatranscriptomic analyses of plant cell wall polysaccharide degradation by microorganisms in the cow rumen, Appl. Environ. Microbiol. 81 (2015) 1375–1386, http://dx.doi.org/10.1128/AEM.03682-14.
[59] Z. Yi, X. Su, V. Revindran, R.I. Mackie, I. Cann, Molecular and biochemical analyses of CbCel9A/Cel48A, a highly secreted multi-modular cellulase by Caldicellulosiruptor bescii during growth on crystalline cellulose, PLoS One 8 (2013) e84172, http://dx.doi.org/10.1371/journal.pone.0084172.
[60] O. Yaniv, Y. Halfon, L.J.W. Shimon, E.A. Bayer, R. Lamed, F. Frolow, Structure of CBM3b of the major cellulosomal scaffoldin subunit ScaA from Acetivibrio cellulolyticus, Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun. 68 (2012) 8–13, http://dx.doi.org/10.1107/S174430911104807X.
[61] H.W. Schroeder, L. Cavacini, Structure and function of immunoglobulins, J. Allergy Clin. Immunol. 125 (2010) S41–S52, http://dx.doi.org/10.1016/

j.jaci.2009.09.046.
[62] D.D. Chaplin, Overview of the immune response, J. Allergy Clin. Immunol. 125 (2010) S3–S23, http://dx.doi.org/10.1016/j.jaci.2009.12.980.
[63] D.N. Simon, T. Domaradzki, W.A. Hofmann, K.L. Wilson, Lamin A tail modification by SUMO1 is disrupted by familial partial lipodystrophy-causing mutations, Mol. Biol. Cell. 24 (2013) 342–350, http://dx.doi.org/10.1091/mbc.E12-07-0527.
[64] T.A. Dittmer, T. Misteli, The lamin protein family, Genome Biol. 12 (2011) 222, http://dx.doi.org/10.1186/gb-2011-12-5-222.
[65] Z. Qin, A. Kalinowski, K.N. Dahl, M.J. Buehler, Structure and stability of the lamin A tail domain and HGPS mutant, J. Struct. Biol. 175 (2011) 425–433, http://dx.doi.org/10.1016/j.jsb.2011.05.015.
[66] D.K. Shumaker, L. Solimando, K. Sengupta, T. Shimi, S.A. Adam, A. Grunwald, S. V Strelkov, U. Aebi, M.C. Cardoso, R.D. Goldman, The highly conserved nuclear lamin Ig-fold binds to PCNA: its role in DNA replication, J. Cell Biol. 181 (2008) 269–280, http://dx.doi.org/10.1083/jcb.200708155.
[67] C.D. Boyd, T.J. Smith, S. El-Kirat-Chatel, P.D. Newell, Y.F. Dufrêne, G.A. O'Toole, Structural features of the Pseudomonas fluorescens biofilm adhesin LapA required for LapG-dependent cleavage, biofilm formation, and cell surface localization, J. Bacteriol. 196 (2014) 2775–2788, http://dx.doi.org/10.1128/JB.01629-14.
[68] S.P. Colgan, H.K. Eltzschig, T. Eckle, L.F. Thompson, Physiological roles for ecto-5'-nucleotidase (CD73), Purinergic Signal 2 (2006) 351–360, http://dx.doi.org/10.1007/s11302-005-5302-5.
[69] T. Hashikawa, M. Takedachi, M. Terakura, T. Saho, S. Yamada, L.F. Thompson, Y. Shimabukuro, S. Murakami, Involvement of CD73 (ecto-5'-nucleotidase) in adenosine generation by human gingival fibroblasts, J. Dent. Res. 82 (2003) 888–892. http://www.ncbi.nlm.nih.gov/pubmed/14578500. (Accessed 15 May 2016).
[70] S. Ruer, G. Ball, A. Filloux, S. de Bentzmann, The "P-usher", a novel protein transporter involved in fimbrial assembly and TpsA secretion, EMBO J. 27 (2008) 2669–2680, http://dx.doi.org/10.1038/emboj.2008.197.
[71] H. Bierne, P. Cossart, Listeria monocytogenes surface proteins: from genome predictions to function, Microbiol. Mol. Biol. Rev. 71 (2007) 377–397, http://dx.doi.org/10.1128/MMBR.00039-06.
[72] J.P.R.O. Orgel, A. V Persikov, O. Antipova, Variation in the helical structure of native collagen, PLoS One 9 (2014) e89519, http://dx.doi.org/10.1371/journal.pone.0089519.
[73] J. Myllyharju, K.I. Kivirikko, Collagens and collagen-related diseases, Ann. Med. 33 (2001) 7–21. http://www.ncbi.nlm.nih.gov/pubmed/11310942. (Accessed 15 May 2016).
[74] Z. Wei, M. Lyon, J.T. Gallagher, Distinct substrate specificities of bacterial heparinases against N-unsubstituted glucosamine residues in heparan sulfate, J. Biol. Chem. 280 (2005) 15742–15748, http://dx.doi.org/10.1074/jbc.M501102200.
[75] M.L. Iruela-Arispe, A. Luque, N. Lee, Thrombospondin modules and angiogenesis, Int. J. Biochem. Cell Biol. 36 (2004) 1070–1078, http://dx.doi.org/10.1016/j.biocel.2004.01.025.
[76] P.R. Lawler, J. Lawler, Molecular basis for the regulation of angiogenesis by thrombospondin-1 and -2, Cold Spring Harb. Perspect. Med. 2 (2012) a006627, http://dx.doi.org/10.1101/cshperspect.a006627.
[77] K. Tan, M. Duquette, J.-H. Liu, R. Zhang, A. Joachimiak, J. Wang, J. Lawler, The structures of the thrombospondin-1 N-terminal domain and its complex with a synthetic pentameric heparin, Structure 14 (2006) 33–42, http://dx.doi.org/10.1016/j.str.2005.09.017.
[78] S. Kojima, N. Nonoyama, N. Takekawa, H. Fukuoka, M. Homma, Mutations targeting the C-terminal domain of FliG can disrupt motor assembly in the Na(+)-driven flagella of Vibrio alginolyticus, J. Mol. Biol. 414 (2011) 62–74, http://dx.doi.org/10.1016/j.jmb.2011.09.019.
[79] U.P. Andley, Crystallins in the eye: function and pathology, Prog. Retin. Eye Res. 26 (2007) 78–98, http://dx.doi.org/10.1016/j.preteyeres.2006.10.003.
[80] T. Roemer, S. Delaney, H. Bussey, SKN1 and KRE6 define a pair of functional homologs encoding putative membrane proteins involved in beta-glucan synthesis, Mol. Cell. Biol. 13 (1993) 4039–4048. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=359953&tool=pmcentrez&rendertype=abstract. (Accessed 12 May 2016).
[81] N.M. Gilbert, M.J. Donlin, K.J. Gerik, C.A. Specht, J.T. Djordjevic, C.F. Wilson, T.C. Sorrell, J.K. Lodge, KRE genes are required for beta-1,6-glucan synthesis, maintenance of capsule architecture and cell wall protein anchoring in Cryptococcus neoformans, Mol. Microbiol. 76 (2010) 517–534, http://dx.doi.org/10.1111/j.1365-2958.2010.07119.x.
[82] C.W. Harkey, K.D. Everiss, K.M. Peterson, Isolation and characterization of a Vibrio cholerae gene (tagA) that encodes a ToxR-regulated lipoprotein, Gene 153 (1995) 81–84. http://www.ncbi.nlm.nih.gov/pubmed/7883190. (Accessed 15 May 2016).
[83] M.J. McBride, T.F. Braun, J.L. Brust, Flavobacterium johnsoniae GldH is a lipoprotein that is required for gliding motility and chitin utilization, J. Bacteriol. 185 (2003) 6648–6657. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=262120&tool=pmcentrez&rendertype=abstract. (Accessed 15 May 2016).
[84] K. Arends, E.-K. Celik, I. Probst, N. Goessweiner-Mohr, C. Fercher, L. Grumet, C. Soellue, M.Y. Abajy, T. Sakinc, M. Broszat, K. Schiwon, G. Koraimann, W. Keller, E. Grohmann, TraG encoded by the pIP501 type IV secretion system is a two-domain peptidoglycan-degrading enzyme essential for conjugative transfer, J. Bacteriol. 195 (2013) 4436–4444, http://dx.doi.org/10.1128/JB.02263-12.
[85] E.A.B. Undheim, A. Jones, K.R. Clauser, J.W. Holland, S.S. Pineda, G.F. King, B.G. Fry, Clawing through evolution: toxin diversification and convergence in the ancient lineage Chilopoda (centipedes), Mol. Biol. Evol. 31 (2014) 2124–2148, http://dx.doi.org/10.1093/molbev/msu162.
[86] A. Oubrie, H.J. Rozeboom, K.H. Kalk, A.J. Olsthoorn, J.A. Duine, B.W. Dijkstra, Structure and mechanism of soluble quinoprotein glucose dehydrogenase, EMBO J. 18 (1999) 5187–5194, http://dx.doi.org/10.1093/emboj/18.19.5187.
[87] P.L. Ipata, F. Balestri, The functional logic of cytosolic 5'-nucleotidases, Curr. Med. Chem. 20 (2013) 4205–4216. http://www.ncbi.nlm.nih.gov/pubmed/23992316. (Accessed 15 May 2016).
[88] D.E. Champagne, C.T. Smartt, J.M. Ribeiro, A.A. James, The salivary gland-specific apyrase of the mosquito Aedes aegypti is a member of the 5'-nucleotidase family, Proc. Natl. Acad. Sci. U. S. A. 92 (1995) 694–698. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=42686&tool=pmcentrez&rendertype=abstract. (Accessed 15 May 2016).
[89] D. Sun, A. McNicol, A.A. James, Z. Peng, Expression of functional recombinant mosquito salivary apyrase: a potential therapeutic platelet aggregation inhibitor, Platelets 17 (2006) 178–184, http://dx.doi.org/10.1080/09537100500460234.
[90] J.E. Urch, R. Hurtado-Guerrero, D. Brosson, Z. Liu, V.G.H. Eijsink, C. Texier, D.M.F. van Aalten, Structural and functional characterization of a putative polysaccharide deacetylase of the human parasite Encephalitozoon cuniculi, Protein Sci. 18 (2009) 1197–1209, http://dx.doi.org/10.1002/pro.128.
[91] Y. Zhao, R.-D. Park, R.A.A. Muzzarelli, Chitin deacetylases: properties and applications, Mar. Drugs 8 (2010) 24–46, http://dx.doi.org/10.3390/md8010024.
[92] A. Kolbinger, T. Gao, D. Brock, R. Ammann, A. Kisters, J. Kellermann, D. Hatton, R.H. Gomer, B. Wetterauer, A cysteine-rich extracellular protein containing a PA14 domain mediates quorum sensing in Dictyostelium discoideum Eukaryot, Cell 4 (2005) 991–998, http://dx.doi.org/10.1128/EC.4.6.991-998.2005.
[93] D.J. Rigden, R.Y. Eberhardt, H.J. Gilbert, Q. Xu, Y. Chang, A. Godzik, Structure- and context-based analysis of the GxGYxYP family reveals a new putative class of glycoside hydrolase, BMC Bioinforma. 15 (2014) 196, http://dx.doi.org/10.1186/1471-2105-15-196.
[94] M.W. Zmudka, J.B. Thoden, H.M. Holden, The structure of DesR from Streptomyces venezuelae, a β-glucosidase involved in macrolide activation, Protein Sci. 22 (2013) 883–892, http://dx.doi.org/10.1002/pro.2204.
[95] M. Lei, M.A. Robinson, S.C. Harrison, The active conformation of the PAK1 kinase domain, Structure 13 (2005) 769–778, http://dx.doi.org/10.1016/j.str.2005.03.007.
[96] M.-H. Oh, X. Wu, H.S. Kim, J.F. Harper, R.E. Zielinski, S.D. Clouse, S.C. Huber, CDPKs are dual-specificity protein kinases and tyrosine autophosphorylation attenuates kinase activity, FEBS Lett. 586 (2012) 4070–4075, http://dx.doi.org/10.1016/j.febslet.2012.09.040.
[97] P. Maurer, K. Todorova, J. Sauerbier, R. Hakenbeck, Mutations in Streptococcus pneumoniae penicillin-binding protein 2x: importance of the C-terminal penicillin-binding protein and serine/threonine kinase-associated domains for beta-lactam binding, Microb. Drug Resist 18 (2012) 314–321, http://dx.doi.org/10.1089/mdr.2012.0022.
[98] C. Yeats, R.D. Finn, A. Bateman, The PASTA domain: a beta-lactam-binding domain, Trends Biochem. Sci. 27 (2002) 438. http://www.ncbi.nlm.nih.gov/pubmed/12217513. (Accessed 15 May 2016).
[99] Y. Chawla, S. Upadhyay, S. Khan, S.N. Nagarajan, F. Forti, V.K. Nandicoori, Protein kinase B (PknB) of Mycobacterium tuberculosis is essential for growth of the pathogen in vitro as well as for survival within the host, J. Biol. Chem. 289 (2014) 13858–13875, http://dx.doi.org/10.1074/jbc.M114.563536.
[100] C. Li, S.-Y. Sun, F.R. Khuri, R. Li, Pleiotropic functions of EAPII/TTRAP/TDP2: cancer development, chemoresistance and beyond, Cell Cycle 10 (2011) 3274–3283, http://dx.doi.org/10.4161/cc.10.19.17763.
[101] K. Pawłowski, A. Muszewska, A. Lenart, T. Szczepińska, A. Godzik, M. Grynberg, A widespread peroxiredoxin-like domain present in tumor suppression- and progression-implicated proteins, BMC Genomics 11 (2010) 590, http://dx.doi.org/10.1186/1471-2164-11-590.
[102] T. Iwasaki, K. Miyauchi, E. Masai, M. Fukuda, Multiple-subunit genes of the aromatic-ring-hydroxylating dioxygenase play an active role in biphenyl and polychlorinated biphenyl degradation in Rhodococcus sp strain RHA1, Appl. Environ. Microbiol. 72 (2006) 5396–5402, http://dx.doi.org/10.1128/AEM.00298-06.
[103] J. Schlosrich, K.L. Eley, P.J. Crowley, T.D.H. Bugg, Directed evolution of a non-heme-iron-dependent extradiol catechol dioxygenase: identification of mutants with intradiol oxidative cleavage activity, Chembiochem 7 (2006) 1899–1908, http://dx.doi.org/10.1002/cbic.200600296.
[104] A. Becerra, R. V Warke, K. Xhaja, B. Evans, J. Evans, K. Martin, N. de Bosch, A.L. Rothman, I. Bosch, Increased activity of indoleamine 2,3-dioxygenase in serum from acutely infected dengue patients linked to gamma interferon antiviral function, J. Gen. Virol. 90 (2009) 810–817, http://dx.doi.org/10.1099/vir.0.004416-0.
[105] S. Sivashankari, P. Shanmughavel, Functional annotation of hypothetical proteins - A review, Bioinformation 1 (2006) 335–338. http://www.pubmedcentral.nih.gov/articlerender.fcgi?

artid=1891709&tool=pmcentrez&rendertype=abstract. (Accessed 23 March 2015).
[106] R. Mohan, S. Venugopal, Computational structural and functional analysis of hypothetical proteins of Staphylococcus aureus, Bioinformation 8 (2012) 722–728, http://dx.doi.org/10.6026/97320630008722.
[107] S. Patel, Letter to the Editor on "The urgency of annotating pathogen hypothetical genes,", Gene Rep. 4 (2016) 233–234, http://dx.doi.org/10.1016/j.genrep.2016.07.004.
[108] S. Patel, N. Mathivanan, A. Goyal, Bacterial adhesins, the pathogenic weapons to trick host defense arsenal, Biomed. Pharmacother. 93 (2017) 763–771, http://dx.doi.org/10.1016/j.biopha.2017.06.102.
[109] P.K. Busk, L. Lange, Classification of fungal and bacterial lytic polysaccharide monooxygenases, BMC Genomics 16 (2015) 368, http://dx.doi.org/10.1186/s12864-015-1601-6.
[110] X. Li, L.-X. Wang, X. Wang, S. Roseman, The chitin catabolic cascade in the marine bacterium Vibrio cholerae: characterization of a unique chitin oligosaccharide deacetylase, Glycobiology 17 (2007) 1377–1387, http://dx.doi.org/10.1093/glycob/cwm096.
[111] S.G. Kini, P.Q.T. Nguyen, S. Weissbach, A. Mallagaray, J. Shin, H.S. Yoon, J.P. Tam, Studies on the Chitin Binding Property of Novel Cysteine-Rich Peptides from Alternanthera sessilis, Biochemistry 54 (2015) 6639–6649, http://dx.doi.org/10.1021/acs.biochem.5b00872.
[112] H. Kaku, Y. Nishizawa, N. Ishii-Minami, C. Akimoto-Tomiyama, N. Dohmae, K. Takio, E. Minami, N. Shibuya, Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor, Proc. Natl. Acad. Sci. U. S. A 103 (2006) 11086–11091, http://dx.doi.org/10.1073/pnas.0508882103.
[113] A. Miya, P. Albert, T. Shinya, Y. Desaki, K. Ichimura, K. Shirasu, Y. Narusaka, N. Kawakami, H. Kaku, N. Shibuya, CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in Arabidopsis, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 19613–19618, http://dx.doi.org/10.1073/pnas.0705147104.
[114] R. de Jonge, H.P. van Esse, A. Kombrink, T. Shinya, Y. Desaki, R. Bours, S. van der Krol, N. Shibuya, M.H.A.J. Joosten, B.P.H.J. Thomma, Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants, Science 329 (2010) 953–955, http://dx.doi.org/10.1126/science.1190859.
[115] S. Patel, Drivers of bacterial genomes plasticity and roles they play in pathogen virulence, persistence and drug resistance, Infect. Genet. Evol. 45 (2016) 151–164, http://dx.doi.org/10.1016/j.meegid.2016.08.030.
[116] S. Patel, S. Patel, Analysis of Ebola virus polymerase domains to find strain-specific differences and to gain insight on their pathogenicity, VirusDisease 27 (2016) 242–250, http://dx.doi.org/10.1007/s13337-016-0334-8.
[117] S. Patel, In silico analysis of Hepatitis C virus (HCV) polyprotein domains and their comparison with other pathogens and allergens to gain insight on pathogenicity mechanisms, Comput. Biol. Chem. 65 (2016) 91–102, http://dx.doi.org/10.1016/j.compbiolchem.2016.10.006.