Structure prediction and network analysis of chitinases from the Cape sundew, Drosera capensis.

Megha H. Unhelkar,[†,‡] Vy T. Duong,[†,¶,‡] Kaosoluchi N. Enendu,[¶] John E. Kelly,[†] Seemal Tahir,[†] Carter T. Butts,[*,§,‖,⊥] and Rachel W. Martin[*,†,¶]

†*Department of Chemistry, University of California, Irvine, Irvine, CA 92697*

‡

¶*Department of Molecular Biology and Biochemistry, University of California, Irvine, Irvine CA 92697*

§*Department of Sociology, UC Irvine, Irvine, CA, 92697 USA*

‖*Department of Electrical Engineering and Computer Science, UC Irvine, Irvine, CA, 92697 USA*

⊥*Department of Statistics, UC Irvine, Irvine, CA, 92697 USA*

E-mail: buttsc@uci.edu; rwmartin@uci.edu

Phone: 1(949)-824-7959

Sequence Alignments

The catalytic action of family 18 chitinases, which retains the β-anomeric carbon stereochemistry from the substrate to the product, is based on substrate-assisted hydrolysis of the glycosidic bond[1–3]. Catalysis is initiated by distorting the -1 sugar ring subsite adjacent to the glycosidic bond. Next, Asp 123 rotates to form hydrogen bonds with both Glu 127 and the N-acetyl group of the +1 sugar. This step protonates Glu 127. Then, the anomeric carbon is subjected to a nucleophilic attack by the oxygen from the N-acetyl group, forming an oxazolinium ion as an intermediate, followed by cleavage of the glycosidic bond by hydrolysis to generate smaller fragments. The DXDXE motif is essential for activity, hence fragments that were lacking this sequence due to truncation were excluded from our protein set.

A sequence alignment for Family 18 chitinases from Caryophylalles carnivorous plants is shown in Supplementary Figure 1. The figure is annotated to highlight specific amino acid properties and important sequence features. The chemical properties of amino acids are color-coded as follows: cysteines are yellow, positively charged residues are blue, negatively charged residues are red, hydrophobic residues are green, and all others are black. Highly conserved residues are indicated with a dot above the sequence position. Cysteine residues involved in structure-stabilizing disulfide bonds are indicated with yellow asterisks, while the active amino acid residues are marked with colored arrows. SignalP 4.1 is used to predict the signal peptide cleavage site, which is specified by underlining the residues on either of the cleavage point. The signal peptide itself is highlighted in light orange. Strikethrough text indicates sequence regions that are absent in the active enzyme, in this case the N-terminal signal peptide that is expressed but removed during maturation. Annotations were performed by homology to a well-characterized acidic endochitinase from *Vitis vinifera* (CHIT3_VITVI, Uniprot ID-P51614).

Family 19 contains Class I, II, and IV chitinases, all of which are characterized by an anomeric inverting mechanism.[4,5] The N-terminal chitin-binding domain is present in Class I and absent in Class II, which are otherwise similar in sequence. Family 19 chitinases from

plants have in common a catalytic domain with an active glutamic acid residue. The active site motif surrounding the active E is either HETT (type I and II) or HETG (type IV),[6] both of which are observed in this set of proteins. Annotations for the Family 19 chitinases are shown in Supplementary Figure 2. Amino acid and sequence features are indicated as in Supplementary Figure 1, with the following additions, when present: the C-rich domain is highlighted in light green, the P-rich hinge in light blue, and the C-terminal extension (CTE) in light gray. Both the C-rich domain and the P-rich hinge are highly variable in length and are absent in some sequences. Only three chitinases in this set contain the CTE, which targets those sequences to the vacuole. The reference sequences for this cluster are CHI3_CASSA (*Castanea sativa*), CHI2_BRANA (*Brassica napus*), and HORV2 (*Hordeum vulgare*).

Figure 1: Sequence alignment for Family 18 chitinases, annotated by homology to the reference sequence CHIT3_VITVI. The "DXDXE" motif, in which the acidic residues are marked with red arrows, is imperative for the enzyme activity. Orange arrows indicate residues implicated in substrate binding.

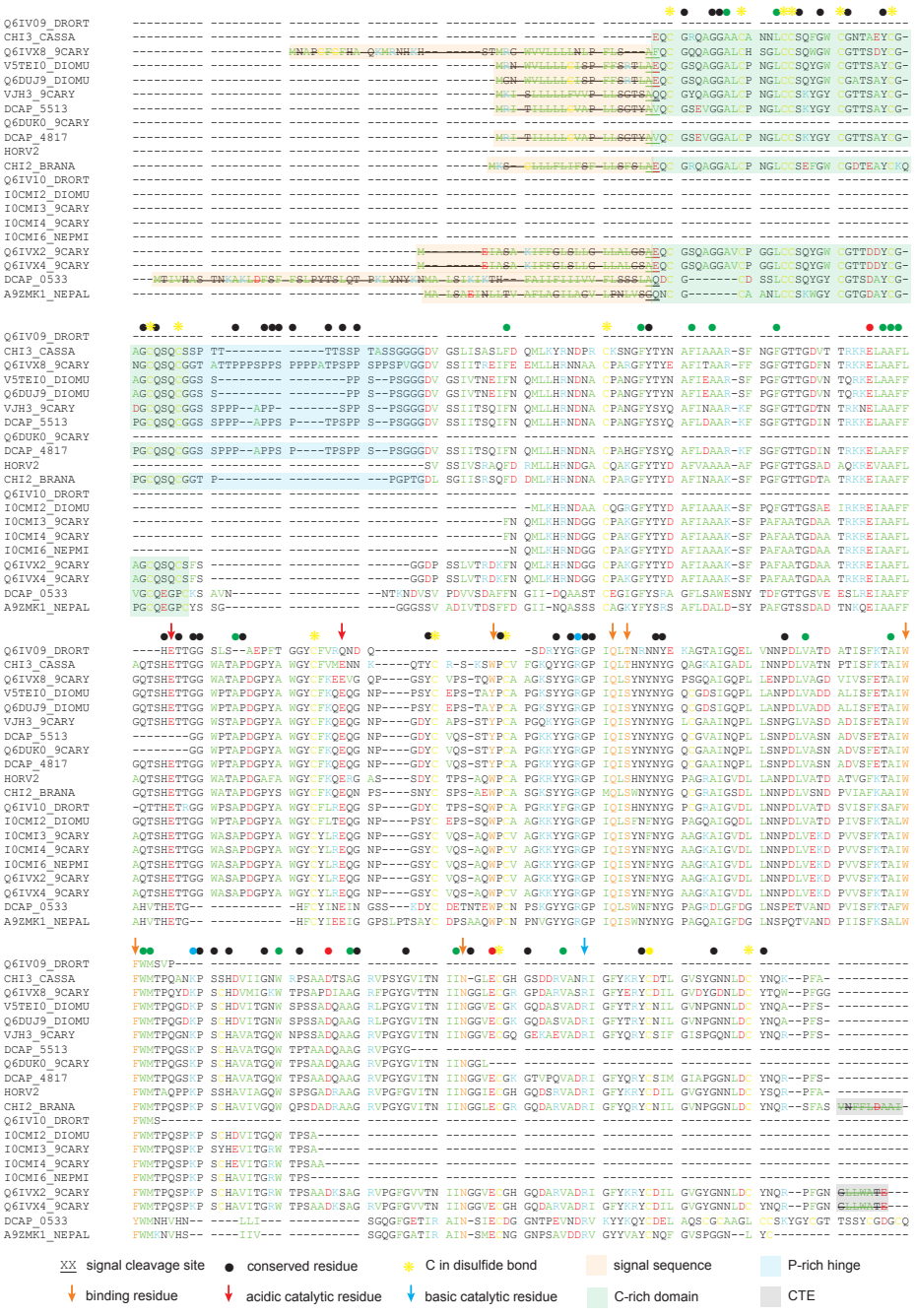

Figure 2: Sequence alignment and annotation for Family 19 chitinases. Many sequences in this cluster contain a chitin-binding C-rich domain (light green) that is connected to the active region by a P-rich hinge (light blue). Three sequences in this cluster contain a C-terminal extension (CTE) that causes the proteins to be targeted to the vacuole.

Four Family 19 chitinase fragments were identified from the *D. capensis* genome by performing a BLAST search for DcChit_1, a chitinase fragment previously identified from genomic DNA of the same organism.[6] Their sequences range from 41%-100% identity to DcChit1_1. These fragments contain part of the N-terminal region, including the C-rich domain and the P-rich hinge, neither of which was observed in the original fragment, along with part of the catalytic domain (Supplementary Figure 3). However, these sequences are all truncated before the catalytic residues. Sequencing of the *D. capensis* transcriptome will clarify whether these are fragments of active genes containing one or more introns, or inactive pseudogenes, which are relatively common in gene families undergoing rapid evolution[7] (as is the case for many proteins associated with pathogen defense[8]).

Figure 3: Chitinase 1 fragments discovered using a BLAST search of the *D. capensis* genome against the DcChitI_1 fragment previously identified by Renner and Specht from *D. capensis* genomic DNA.

Preliminary Structural Models and *In silico* Maturation

Preliminary models for both Family 18 and Family 19 chitinases were produced using Rosetta[9], implemented in the online Robetta server[10]. The Rosetta structures contain the full sequences, including the N-terminal signal peptides, and in some cases, C-terminal targeting peptides that are also cleaved during maturation. The *in silico* maturation process, which we have previously described for cysteine proteases,[11] is illustrated in Supplementary

Figure 4 for a representative family 18 chitinase, DCAP_2209. The initial Rosetta sequence, including the signal peptide and lacking post-translational modifications, is shown in Supplementary Figure 4. In order to generate the equilibrated structure Supplementary Figure 4b, which more closely approximates the active form of the enzyme in solution, the signal sequence is removed, disulfide bonds are added using homology to a reference sequence (in this case CHIT3_VITVI), and the structure is equilibrated in explicit solvent. Many Family 18 chitinases from plants contain three disulfide bonds,[12,13] although examples without any disulfide bonds also exist.[14] Three are found in all the Family 18 chitinases in this set, as in CHIT3_VITVI,[15] and hevamine from *Hevea brasiliensis* (PDB ID: 2HVM).[16] The functionally important cis peptide bonds are captured by the molecular models for all the Family 18 chitinases examined here except for DCAP_7323, which unlikely to be active in any case because it is truncated at the N-terminal end.

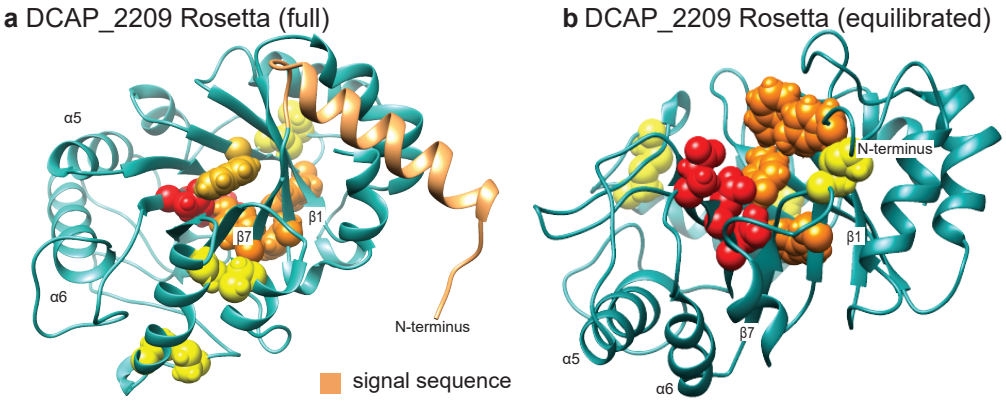

Figure 4: DCAP_2209 (a) before and (b) after *in silico* maturation. The light orange helix in part a is the N-terminal signal sequence. Important residues are color-coded as follows: Red: catalytically active residues of the "DXDXE" motif. Orange: aromatic substrate-binding residues. Yellow: Cysteines in disulfide bonds.

Supplementary Figure 5 shows full-length structures for Q6IVX8_9CARY and Q6IVX2_9CARY from *Drosera spatulata*. The N-terminal and C-terminal targeting sequences are exposed on the surface of the protein, as expected. The P-rich hinge in these proteins is variable in length, and highly flexible, as illustrated by the different relative conformations of of the

catalytic and C-rich chitin binding domains observed here.

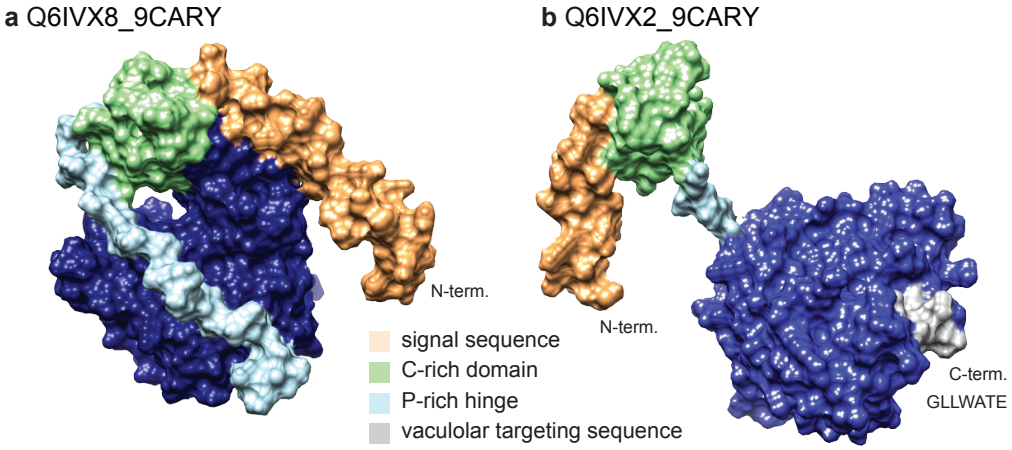

Figure 5: Initial Rosetta structures for two class I chitinases from *Drosera spatulata*, Q6IVX8_9CARY and Q6IVX2_9CARY, illustrating positioning of the N-terminal and C-terminal targeting sequences and the variability in length and conformation for the P-rich hinge.

Description of a Novel Two-Domain Class IV Chitinase

Class IV chitinases exhibit an amino acid substitution in the first active site region relative to Class I chitinases, resulting in a HETG/I motif instead of the HETT motif.[6] A deletion of four amino acids in the Cys-rich binding domain is also observed in class IV chitinases, as shown for a class IV chitinase from *Nepenthes alata* (A9ZMK1_NEPAL)[17] and DCAP_0533 in 2. Supplementary Figure 6 shows a sequence alignment of the N- and C-terminal domains of the Class IV chitinase DCAP_0533 with single domain class IV chitinases from *Picea abies* (Q6WSR8_PICAB), *Zea mays* (CHIA_MAIZE), and *Sorghum bicolor* (C5YBE7_SORBI). The two domains of DCAP_0533 were aligned with the most closely related annotated class IV chitinases, those from *Picea abies* (EC: 3.2.1.14, Uniprot: Q6WSR8_PICAB), *Zea mays* (EC: 3.2.1.14, Uniprot: CHIA_MAIZE), and *Sorghum bicolor* (Uniprot: C5YBE7_SORBI)[6,18,19] (Supplementary Figure 7).

Structurally, each domain consists of two lobes with eight helices each, separated by a large active site cleft (Supplementary Figure 7(a). In Supplementary Figure 7(b), the two domains of this protein are shown overlaid with the crystal structures of class IV chitinases from *Zea mays* (PDBID: 4MCK, 60% identity with the NTD) and *Picea abies* (PDBID: 3HBE, 64% identity with the CTD). The NTD Supplementary Figure 7(c) has an N-terminal signal peptide, a conserved C-rich binding domain, and a catalytic domain that appears to be functional. In its homolog CHIA_MAIZE, Chaudet et. al. characterized four catalytic residues (E62, E71, E165, and R171),[19] all of which have counterparts in the NTD of DCAP_0533 (E173, E182, E278, R290) (Supplementary Figures 6, 7. Previous modeling studies of well-characterized class I chitinases from barley, mustard, and chestnut seed homologs (barley: E67, mustard: E212, chestnut: E124) suggest the necessity of E62 in CHIA_MAIZE and E173 in the NTD of DCAP_0533 as a proton donor.[20–22] Overall, mutagenesis studies highlight the significance of E62 as an essential residue of the catalytic triad (E62, E165, R171 in CHIA_MAIZE) which we use to infer an equivalent catalytic triad in the NTD of DCAP_0533 (E173, E278, and R290). It has also been hypothesized that purpose of the triad is to alter the surrounding environment to induce activation of the glutamic acid in the HETG/I (class IV) or HETT (class I/II) motif by changing its pKa.[22]

Linked to the NTD by a cysteine and glycine-rich linker sequence, the CTD of DCAP_0533 (Supplementary Figure 7(d) potentially houses a second catalytic domain or binding domain whose closest structural homolog is Q6WSR8_PICAB from Norway spruce (*Picea abies*) (Supplementary Figure 6). Binding site residues and cysteines involved in disulfide bond formation are conserved in both chitinases. Comparing this sequence with the catalytic triad of Q6WSR8_PICAB (E113, R230, E218), we observe a potentially equivalent triad in the CTD (E407, E507, R519) (Supplementary Figure 7). Ubhayasekera et. al. describe the flexibility of E113 and demonstrate two conformations that it can adopt during catalysis.[18]

Although E407 is not located in the equivalent sequence position to E113, the flexibility of this residue in Q6WSR8_PICAB suggests that Glu407 may be at an appropriate distance to function as part of the CTD triad. Alternatively, the CTD may lack catalytic activity and act as a binding domain as in multidomain chitinases from archaea and bacteria.

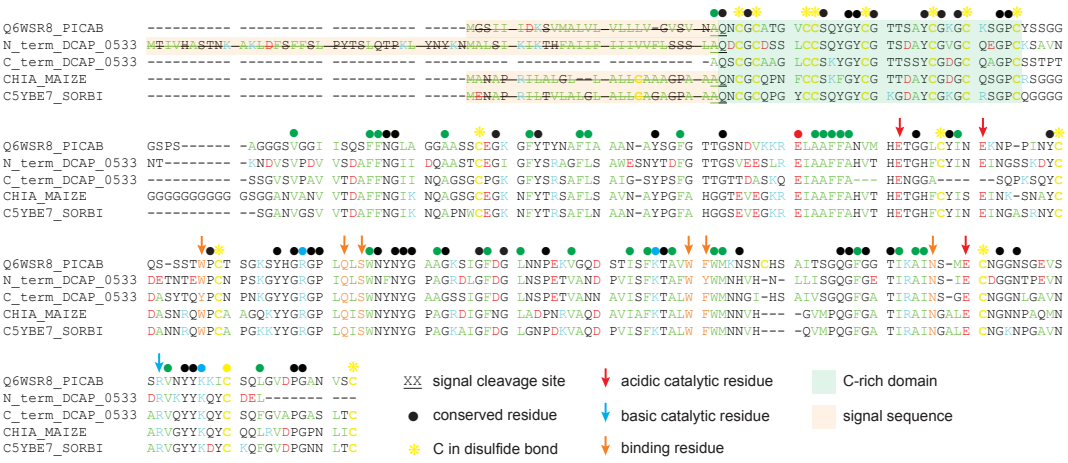

Figure 6: Sequence alignment and annotation of Q6WSR8_PICAB, CHIA_MAIZE, and the N-terminal domain (NTD) and C-terminal domain (CTD) of DCAP_0533. For the purpose of comparison, the sequence is manually separated above. We observe high sequence conservation regarding: the signal cleavage site, C-rich domain length and location, cysteines composing disulfide bonds, other binding site residues surrounding the main binding site residues (orange arrows), and catalytic residues except Glu407 of the CTD which is unaligned with Glu113 of Q6WSR8_PICAB

All initial and equilibrated structures are available for download as PDB files. The available structures for Families 18 and 19 are tabulated in Supplementary Tables 1 and 2, respectively.

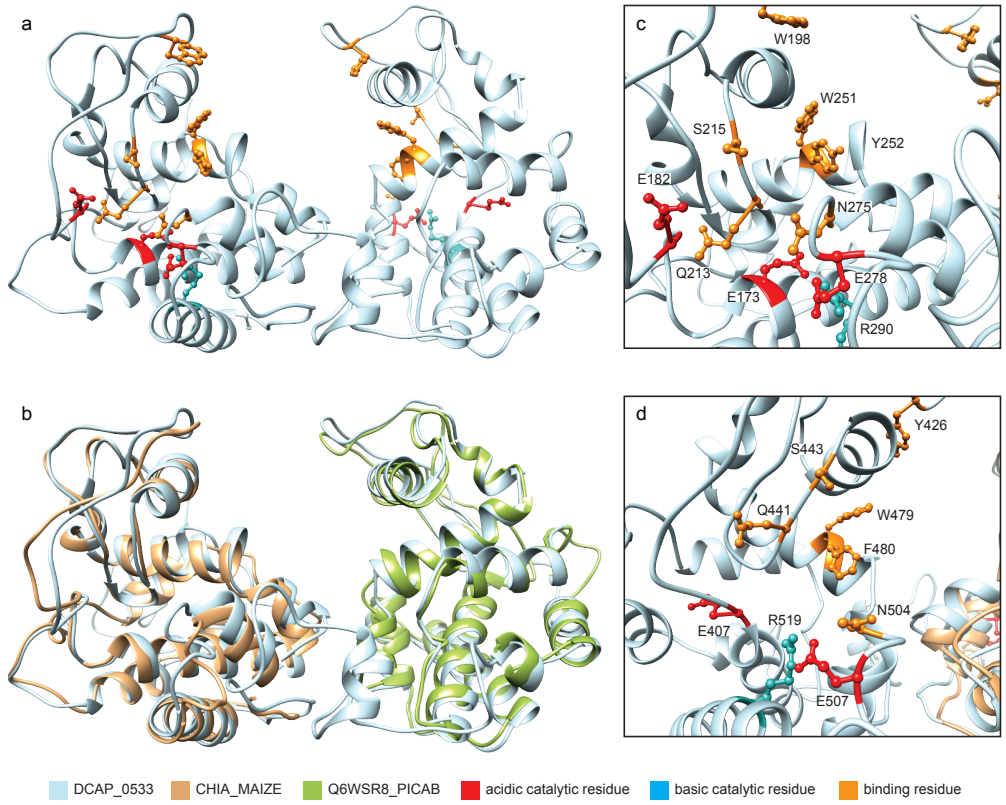

Figure 7: DCAP_0533 comparison with CHIA_MAIZE (4MCK) and Q6WSR8_PICAB (3HBE) and close up of catalytic residues and binding residues: (a) Robetta generated predicted structure with highlighted catalytic residues and binding residues. (b) Superimposition of CHIA_MAIZE and Q6WSR8_PICAB against DCAP_0533. (c) Catalytic site of NTD with 1-letter residue code and specifier. Catalytic triad consists of E173, E278, R290. (d) Catalytic site of CTD with 1-letter residue code and specifier. Catalytic triad consists of E407, E507, R519.

Supplementary Table 1: Rosetta and equilibrated structures for Family 18 Chitinases PDB files available for download

Protein	Organism	Sequence Elements included	File Name
CHIT3_VITVI	*Vitis vinifera*	signal, active region	CHIT3_VITVI_m1.pdb
CHIT3_VITVI	*Vitis vinifera*	active region	CHIT3_VITVI_mature_m1.pdb
DCAP_7323	*D. capensis*	active region	DCAP_7323_m1.pdb
DCAP_7323	*D. capensis*	active region	DCAP_7323_mature_m1.pdb
DCAP_0106	*D. capensis*	signal, active region	DCAP_0106_m1.pdb
DCAP_0106	*D. capensis*	active region	DCAP_0106_mature_m1.pdb
DCAP_7544	*D. capensis*	signal, active region	DCAP_7544_m1.pdb
DCAP_7544	*D. capensis*	active region	DCAP_7544_mature_m1.pdb
DCAP_2209	*D. capensis*	signal, active region	DCAP_2209_m1.pdb
DCAP_2209	*D. capensis*	active region	DCAP_2209_mature_m1.pdb
C7F821_NEPMI	*N. mirabilis*	signal, active region	C7F821_NEPMI_m1.pdb
C7F821_NEPMI	*N. mirabilis*	active region	C7F821_NEPMI_mature_m1.pdb
C7F817_9CARY	*D. spatulata*	signal, active region	C7F817_9CARY_m1.pdb
C7F817_9CARY	*D. spatulata*	active region	C7F817_9CARY_mature_m1.pdb
I7HCY7_NEPAL	*N. alata*	signal, active region	I7HCY7_NEPAL_m1.pdb
I7HCY7_NEPAL	*N. alata*	active region	I7HCY7_NEPAL_mature_m1.pdb
C7F818_9CARY	*D. spatulata*	signal, active region	C7F818_9CARY_m1.pdb
C7F818_9CARY	*D. spatulata*	active region	C7F818_9CARY_mature_m1.pdb
Q06SN0_9CARY	*D. spatulata*	signal, active region	Q06SN0_9CARY_m1.pdb
Q06SN0_9CARY	*D. spatulata*	active region	Q06SN0_9CARY_mature_m1.pdb
C7F824_9CARY	*D. spatulata*	signal, active region	C7F824_9CARY_m1.pdb
C7F824_9CARY	*D. spatulata*	active region	C7F824_9CARY_mature_m1.pdb
C7F822_9CARY	*D. spatulata*	signal, active region	C7F822_9CARY_m1.pdb
C7F822_9CARY	*D. spatulata*	active region	C7F822_9CARY_mature_m1.pdb
C7F819_9CARY	*D. spatulata*	signal, active region	C7F819_9CARY_m1.pdb
C7F819_9CARY	*D. spatulata*	active region	C7F819_9CARY_mature_m1.pdb
C7F823_NEPGR	*N. gracilis*	signal, active region	C7F823_NEPGR_m1.pdb
C7F823_NEPGR	*N. gracilis*	active region	C7F823_NEPGR_mature_m1.pdb
DCAP_5455	*D. capensis*	signal, active region	DCAP_5455_m1.pdb
DCAP_5455	*D. capensis*	active region	DCAP_5455_mature_m1.pdb
DCAP_2879	*D. capensis*	signal, active region	DCAP_2879_m1.pdb
DCAP_2879	*D. capensis*	active region	DCAP_2879_mature_m1.pdb
DCAP_4799	*D. capensis*	signal, active region	DCAP_4799_m1.pdb
DCAP_4799	*D. capensis*	active region	DCAP_4799_mature_m1.pdb
DCAP_2737	*D. capensis*	signal, active region	DCAP_2737_m1.pdb
DCAP_2737	*D. capensis*	active region	DCAP_2737_mature_m1.pdb

Supplementary Table 2: Rosetta and equilibrated structures for Family 19 Chitinases PDB files available for download

Protein	Organism	Sequence Elements included	File Name
HORV2	*H. vulgare*	active region	HORV2 PDBID: 2BAA
HORV2	*H. vulgare*	active region	HORV2_crystal_struc_mature_m1.pdb
Q6IV09_DRORT	*D. rotundifolia*	active region	Q6IV09_DRORT_m1.pdb
Q6IV09_DRORT	*D. rotundifolia*	active region	Q6IV09_DRORT_mature_m1.pdb
CHI3_CASSA	*Castanea sativa*	C-rich domain, P-rich hinge, active region	CHI3_CASSA_m1.pdb
CHI3_CASSA	*Castanea sativa*	C-rich domain, P-rich hinge, active region	CHI3_CASSA_mature_m1.pdb
Q6IVX8_9CARY	*D. spatulata*	signal, C-rich domain, P-rich hinge, active region	Q6IVX8_9CARY_m1.pdb
Q6IVX8_9CARY	*D. spatulata*	C-rich domain, P-rich hinge, active region	Q6IVX8_9CARY_mature_m1.pdb
V5TEI0_DIOMU	*D. muscipula*	signal, C-rich domain, P-rich hinge, active region	V5TEI0_DIOMU_m1.pdb
V5TEI0_DIOMU	*D. muscipula*	C-rich domain, P-rich hinge, active region	V5TEI0_DIOMU_mature_m1.pdb
Q6DUJ9_DIOMU	*D. muscipula*	signal, C-rich domain, P-rich hinge, active region	Q6DUJ9_DIOMU_m1.pdb
Q6DUJ9_DIOMU	*D. muscipula*	C-rich domain, P-rich hinge, active region	6DUJ9_DIOMU_mature_m1.pdb
VJH3_9CARY	*D. spatulata*	signal, C-rich domain, P-rich hinge, active region	VJH3_9CARY_m1.pdb
VJH3_9CARY	*D. spatulata*	C-rich domain, P-rich hinge, active region	VJH3_9CARY_mature_m1.pdb
DCAP_5513	*D. capensis*	signal, C-rich domain, P-rich hinge, active region	DCAP_5513_m1.pdb
DCAP_5513	*D. capensis*	C-rich domain, P-rich hinge, active region	DCAP_5513_mature_m1.pdb
Q6DUKO_9CARY	*D. spatulata*	active region	Q6DUKO_9CARY_m1.pdb
Q6DUKO_9CARY	*D. spatulata*	active region	Q6DUKO_9CARY_mature_m1.pdb
DCAP_4817	*D. capensis*	signal, C-rich domain, P-rich hinge, active region	DCAP_4817_m1.pdb
DCAP_4817	*D. capensis*	C-rich domain, P-rich hinge, active region	DCAP_4817_mature_m1.pdb
CHI2_BRANA	*B. napus*	signal, C-rich domain, P-rich hinge, active region, CTE	CHI2_BRANA_m1.pdb
CHI2_BRANA	*B. napus*	C-rich domain, P-rich hinge, active region	CHI2_BRANA_mature_m1.pdb
Q6IV10_DRORT	*D. rotundifolia*	active region	Q6IV10_DRORT_m1.pdb
Q6IV10_DRORT	*D. rotundifolia*	active region	Q6IV10_DRORT_mature_m1.pdb
I0CMI2_DIOMU	*D. muscipula*	active region	I0CMI2_DIOMU_m1.pdb
I0CMI2_DIOMU	*D. muscipula*	active region	I0CMI2_DIOMU_mature_m1.pdb
I0CMI3_9CARY	*D. spatulata*	active region	I0CMI3_9CARY_m1.pdb
I0CMI3_9CARY	*D. spatulata*	active region	I0CMI3_9CARY_mature_m1.pdb
I0CMI4_9CARY	*D. spatulata*	active region	I0CMI4_9CARY_m1.pdb
I0CMI4_9CARY	*D. spatulata*	active region	I0CMI4_9CARY_mature_m1.pdb
I0CMI6_NEPMI	*N. mirabilis*	active region	I0CMI6_NEPMI_m1.pdb
I0CMI6_NEPMI	*N. mirabilis*	active region	I0CMI6_NEPMI_mature_m1.pdb
Q6IVX2_9CARY	*D. spatulata*	signal, C-rich domain, P-rich hinge, active region, CTE	Q6IVX2_9CARY_m1.pdb
Q6IVX2_9CARY	*D. spatulata*	C-rich domain, P-rich hinge, active region	Q6IVX2_9CARY_mature_m1.pdb
Q6IVX4_9CARY	*D. spatulata*	signal, C-rich domain, P-rich hinge, active region, CTE	Q6IVX4_9CARY_m1.pdb
Q6IVX4_9CARY	*D. spatulata*	C-rich domain, P-rich hinge, active region	Q6IVX4_9CARY_mature_m1.pdb
DCAP_0533	*D. capensis*	signal, C-rich domain, P-rich hinge, active region, C-terminal domain	DCAP_0533_m1.pdb
DCAP_0533	*D. capensis*	C-rich domain, P-rich hinge, active region, C-terminal domain	DCAP_0533_mature_m1.pdb
A9ZMK1_NEPAL	*N. alata*	signal, C-rich domain, P-rich hinge, active region	A9ZMK1_NEPAL_m1.pdb
A9ZMK1_NEPAL	*N. alata*	C-rich domain, P-rich hinge, active region	A9ZMK1_NEPAL_mature_m1.pdb

References

(1) Van Aalten, D.; Komander, D.; Synstad, B.; Gåseidnes, S.; Peter, M.; Eijsink, V. *Proceedings of the National Academy of Sciences* **2001**, *98*, 8979–8984.

(2) Payne, C. M.; Baban, J.; Horn, S. J.; Backe, P. H.; Arvai, A. S.; Dalhus, B.; Bjørås, M.; Eijsink, V. G. H.; Sørlie, M.; Beckham, G. T.; Vaaje-Kolstad, G. *Journal of Biological Chemistry* **2012**, *287*, 36322–36330.

(3) Madhuprakash, J.; Singh, A.; Kumar, S.; Sinha, M.; Kaur, P.; Sharma, S.; Podile, A. R.; Singh, T. P. *International journal of biochemistry and molecular biology* **2013**, *4*, 166.

(4) T., F. *Current Protein and Peptide Science* **2000**, *1*, 105–124.

(5) Iseli, B.; Armand, S.; Boller, T.; Neuhaus, J.; Henrissat, B. *FEBS Letters* **1996**, *382*, 186–188.

(6) Renner, T.; Specht, C. D. *Molecular Biology and Evolution* **2012**, *29*, 2971–2985.

(7) Hua, Z.; Zou, C.; Shiu, S. H.; Vierstra, R. *PLoS One* **2011**, *6*, e16219.

(8) Kaschani, F.; Shabab, M.; Bozkurt, T.; Shindo, T.; Schornack, S.; Gu, C.; Ilyas, M.; Win, J.; Kamoun, S.; van der Hoorn, R. *Plant Physiology* **2010**, *154*, 1794–1804.

(9) D.E., K.; Chivian, D.; Baker, D. *Nucleic Acids Research* **2004**, *32*, W526–31.

(10) Raman, S. et al. *Proteins* **2009**, *77*, 89–99.

(11) Butts, C. T.; Zhang, X.; Kelly, J. E.; Roskamp, K. W.; Unhelkar, M. H.; Freites, J. A.; Tahir, S.; Martin, R. W. *Computational and Structural Biotechnology Journal* **2016**, *in press*.

(12) Beintema, J. J. *FEBS Letters* **1994**, *350*, 159–163.

(13) Kesari, P.; Patil, D. N.; Kumar, P.; Tomar, S.; Sharma, A. K.; Kumar, P. *PROTEOMICS* **2015**, *15*, 1693–1705.

(14) Kitaoku, Y.; Umemoto, N.; Ohnuma, T.; Numata, T.; Taira, T.; Sakuda, S.; Fukamizo, T. *Planta* **2015**, *242*, 895–907.

(15) Busam, G.; Kassemeyer, H.-H.; Matern, U. *Plant Physiol.* **1997**, *15*, 1029–1038.

(16) Terwisscha van Scheltinga, A.; Kalk, K.; Beintema, J.; B.W., D. *Structure* **1994**, *2*, 1181–1189.

(17) Ishisaki, K.; Honda, Y.; Taniguchi, H.; Hatano, N.; Hamada, T. *Glycobiology* **2012**, *22*, 345–351.

(18) Ubhayasekera, W.; Rawat, R.; Ho, S. W. T.; Wiweger, M.; Von Arnold, S.; Chye, M.-L.; Mowbray, S. L. *Plant Molecular Biology* **2009**, *71*, 277–289.

(19) Chaudet, M. M.; Naumann, T. A.; Price, N. P.; Rose, D. R. *Protein Science* **2014**, *23*, 586–593.

(20) Andersen, M.; Jensen, A.; Robertus, J.; Leah, R.; Skriver, K. *Biochemical Journal* **1997**, *822*, 815–822.

(21) Garcia-Casado, G.; Carmen, C.; Allona, I.; Casado, R.; Pacios, L.; Aragoncillo, C.; Gomez, L. *Glycobiology* **1998**, *8*, 1021–1028.

(22) Tang, C.; Chye, M.; Ramalingam, S.; Ouyang, S.; Zhao, K.; Ubhayasekera, W.; Mowbray, S. *Plant Molecular Biology* **2004**, *56*, 285–298.