BJ www.biochemj.org

This is a data-enriched, interactive PDF that provides the gateway to a world of information when opened in Utopia Documents
Download FREE software now

BJ Signal

Biochem. J. (2011) 435, 167–174 (Printed in Great Britain) doi:10.1042/BJ20101613

online data

A new principle of oligomerization of plant DEG7 protease based on interactions of degenerated protease domains

Holger SCHUHMANN, Ulrike MOGG and Iwona ADAMSKA[1]
Department of Physiology and Plant Biochemistry, University of Konstanz, Universitätsstrasse 10, D-78457 Konstanz, Germany

Deg/HtrA proteases are a large group of ATP-independent serine endoproteases found in almost every organism. Their usual domain arrangement comprises a trypsin-type protease domain and one or more PDZ domains. All Deg/HtrA proteases form homo-oligomers with trimers as the basic unit, where the active protease domain mediates the interaction between individual monomers. Among the members of the Deg/HtrA protease family, the plant protease DEG7 is unique since it contains two protease domains (one active and one degenerated) and four PDZ domains. In the present study, we investigated the oligomerization behaviour of this unusual protease using yeast two-hybrid analysis *in vivo* and with recombinant protein *in vitro*. We show that DEG7 forms trimeric complexes, but in contrast with other known Deg/HtrA proteases, it shows a new principle of oligomerization, where trimerization is based on the interactions between degenerated protease domains. We propose that, during evolution, a duplicated active protease domain degenerated and specialized in protein–protein interaction and complex formation.

Key words: *Arabidopsis thaliana*, complex formation, degenerated protease domain, DEG7, serine protease, taxonomic distribution.

INTRODUCTION

Deg/HtrA (for degradation of periplasmic proteins/high temperature requirement A) proteases are ATP-independent serine proteases found in almost every organism [1,2]. They are generally defined by a trypsin-type protease domain (S1B, glutamylpeptidase I subfamily according to the nomenclature proposed by the MEROPS database [3], http://merops.sanger.ac.uk/) and one or more PDZ domains (originally identified in the postsynaptic density 95 protein, the *Drosophila* tumour-suppressor protein Discs large and the tight-junction protein zonula occludens 1) responsible for protein–protein interaction. Members of this family without a PDZ domain have also been described [4–6].

Deg/HtrA proteases form homo-oligomeric complexes, with trimers as the basic unit and oligomerization mediated by the protease domain [7–9]. Exceptions are plant DEG5 and DEG8 which form a heterohexameric complex in the thylakoid lumen of chloroplasts [6]. Biochemical and crystallographic analysis indicated that purified DegP from *Escherichia coli* [9,10] or HtrA from *Thermotoga maritima* [11] mainly exist as proteolytically inactive hexamers, where two homotrimers are stacked in a face-to-face manner. In solution, these hexamers assemble into large catalytically active spherical structures around their substrates forming 12- or 24-mers composed of four or eight homotrimers respectively [12,13].

Deg/HtrA proteases are involved in a variety of processes, including signalling [14,15], degradation of damaged proteins and housekeeping [16,17], apoptosis [18] and protein processing [4,5,19]. The *Arabidopsis thaliana* genome contains 16 genes encoding Deg/HtrA proteases [20]. Phylogenetic comparison of Deg/HtrA proteases from various organisms (including plants, animals, fungi and bacteria) revealed that this family is divided into four distinct groups [4]. DEG7 was the only protease from *A. thaliana* that clustered with Deg/HtrA proteases from fungi, forming a group of Deg/HtrA enzymes with an unusual domain arrangement. All members of this group are twice as long as other Deg/HtrA proteases, possess two protease domains (one degenerated) and up to four PDZ domains [2,4]. The best examined protease from this group is the Ynm3p protein (also called Nma111p, for nuclear mediator of apoptosis) from *Saccharomyces cerevisiae* [21–23]. It is a nuclear protein [21,22] interacting with the nuclear core complex [22] and long-chain acyl-CoA synthetases [23], as assayed by Y2H (yeast two-hybrid) screens. In contrast with the yeast protease, DEG7 from *A. thaliana* (At3g03380) was identified as a chloroplast protein involved in the degradation of photodamaged D1 protein, a core protein of the Photosystem II reaction centre [24].

In the present study, we investigated the taxonomic distribution of DEG7-like proteases and analysed how the unusual domain arrangement affects the oligomerization of DEG7 from *A. thaliana*. We demonstrated that DEG7 orthologues are restricted to fungi and plants, including various algae and land plants. Using Y2H assays and SEC (size-exclusion chromatography) with recombinant DEG7 purified from *E. coli*, we show that trimerization of DEG7 monomers is mediated by the second (degenerated) protease domain.

EXPERIMENTAL

Bioinformatics

We searched annotated protein databases for the presence of DEG7 orthologues using the BLAST algorithm [25] with default parameters as implemented on the web pages of the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/blast), The Gene Index Project (TGI, http://compbio.dfci.harvard.edu/tgi/), and the Joint Genome Institute (JGI, http://genomeportal.jgi-psf.org/). A list of DEG7 orthologues with their respective accession numbers is presented in Supplementary

Abbreviations used: AD, activation domain; BD, DNA-binding domain; 2ME, 2-mercaptoethanol; IPTG, isopropyl β-D-thiogalactopyranoside; NTA, nitrilotriacetic acid; SC, synthetic complete; SEC, size-exclusion chromatography; Y2H, yeast two-hybrid.

[1] To whom correspondence should be addressed (e-mail iwona.adamska@uni-konstanz.de).

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

Biochemical Journal

Table S1 (at http://www.BiochemJ.org/bj/435/bj4350167add.htm). Conserved protease and PDZ domains were identified using the InterProScan platform at the European Molecular Biology Laboratory/European Bioinformatics Institute (http://www.ebi.ac.uk/Tools/InterProScan) [26]. Secondary-structure prediction and comparison was performed using the HHpred platform (Max-Planck Institute for Developmental Biology, http://toolkit.tuebingen.mpg.de/hhpred/) [27]. Multiple sequence alignments of full-length and protease domain sequences were generated by the M-Coffee web server using standard settings [28]. Phylogenetic trees were constructed using the Phylogeny.fr platform (http://www.phylogeny.fr) [29]. Gaps in the alignment were removed with the built-in curation method, and phylogenetic trees were constructed using the Maximum-Likelihood, Parsimony or Neighbour-Joining method. All methods resulted in essentially the same tree. Human HtrA2, *E. coli* DegP and DEG1 from *A. thaliana*, representing the best examined Deg/HtrA proteases from animals, bacteria and plants respectively, were chosen as outgroups, since they do not belong to the DEG7 group [4]. Models of the first and second half of DEG7 were created using the MODELLER 9v6 program [30] with multiple template alignment, using the structure files 2pzd.pdb [31], 1lcy.pdb [7], 3cs0.pdb [12] and 1ky9.pdb [9] as templates. Manual assembly of the obtained structures into a DEG7 trimer was performed using the PyMOL program (http://www.pymol.org/).

Plasmid construction

General molecular biological procedures were conducted using the method described in [32]. Sequences of primers used in this study are given in Supplementary Table S2 (at http://www.BiochemJ.org/bj/435/bj4350167add.htm). A DNA fragment encoding *A. thaliana* DEG7 (At3g03380) was generated by PCR with SALK cDNA clone U21730 [33] as a template, using primers 0734 and 0740. The resulting cDNA fragment was cloned into the pET151-D/TOPO plasmid (Invitrogen), resulting in plasmid pHS36. This plasmid was mutagenized with primers 0724 and 0725 using the QuikChange® II site-directed mutagenesis kit (Stratagene) according to the manufacturer's instructions, resulting in plasmid pHS52 encoding $DEG7_{Ser206Ala}$. This mutation was introduced to prevent downstream applications (purification, expression in yeast) from negative effects of potential uncontrolled proteolysis. A change of the active-site serine residue to alanine does not influence the oligomerization behaviour of the protease domain, as was shown for human HtrA2 and *E. coli* DegP [7,9,12]. Amplification of the DEG7 cDNA with primers 0740 and 0799, using pHS52 as a template, and ligation of the respective DNA into pET151-D/TOPO, created a vector for overexpression of full-length His_6-tagged DEG7 in *E. coli* (pHS166). Plasmids for the overexpression of *DEG7* cDNA fragments were constructed by amplifying $\mathit{DEG7}_{Ser206Ala}$ with the following primers: first half, primers 0740 and 0771; second half, primers 0747 and 0777; active protease domain, primers 0765 and 0774; degenerated protease domain, primers 0793 and 0797) and cloning the fragment into pET151-D/TOPO, resulting in pHS183 (active protease domain), pHS184 (first half), pHS185 (second half) and pHS186 (degenerated protease domain).

For the Y2H assay, vectors were created using gateway technology (Invitrogen). Entry vectors were created by a TOPO reaction using cDNA fragments coding for $DEG7_{Ser206Ala}$ full-length (DEG7-fl, pHS73, primers 0740 and 0747), $DEG7_{Met1-Gln563}$ (first half, pHS81, primers 0740 and 0771), $DEG7_{Glu527-Gln1097}$ (second half, pHS80, primers 0777 and 0747), $DEG7_{Ser35-Lys256}$ (active protease domain, pHS77, primers 0765 and 0774), $DEG7_{Gly581-Gly840}$ (degenerated protease domain, pHS 172, primers 0793 and 0797), $DEG7_{Asp255-Ser373}$ (PDZ1, pHS78, primers 0775 and 0767) and $DEG7_{Ser868-Gln1097}$ (PDZ3 + 4, pHS88, primers 0776 and 0747), which were amplified by PCR using pHS52 as template. Final vectors for the assay were created by performing a gateway reaction with the entry vectors and modified pAD-GAL4–2.1 (Stratagene), introducing the GAL4 AD (activation domain) and modified pBD-Gal4 Cam (Stratagene), introducing the GAL4 BD (DNA-binding domain) respectively, resulting in pHS82 (AD–DEG7-fl), pHS83 (BD–DEG7-fl), pHS96 (AD–first half), pHS91 (BD–first half), pHS95 (AD–second half), pH-S84 (BD–second half), pHS93 (AD–active protease domain), pHS86 (BD–active protease domain), pHS 176 (AD–degenerated protease domain), pHS 179 (BD–degenerated protease domain), pHS90 (AD–PDZ1), pHS85 (BD–PDZ1), pHS97 (AD–PDZ3 + 4) and pHS89 (BD–PDZ3 + 4). All primers were obtained from Operon Biotechnologies. The inserts of all plasmids were sequenced (GATC Biotech), confirming the correct nucleotide sequence and insert orientation.

Protein overexpression, purification and SEC

Recombinant DEG7 containing an N-terminal His_6-tag was produced in *E. coli* strain BL21 Star™ (DE3) cells (Invitrogen) carrying plasmid pHS166. Cells were grown at 30 °C in 2 litres of 2 × YT medium [1.6 % (w/v) tryptone/1 % (w/v) yeast extract/0.5 % NaCl] [32] containing 2 % (w/v) glucose and 100 $\mu g \cdot l^{-1}$ ampicillin to a D_{600} of 0.8 and cooled down to 20 °C. *DEG7* overexpression was induced by adding 0.01 mM IPTG (isopropyl β-D-thiogalactopyranoside) (final concentration). Cells were harvested after 3 h by centrifugation at 6000 ***g*** for 10 min, resuspended in 2× PBS [32] supplemented with 20 mM 2ME (2-mercaptoethanol) and Complete™ EDTA-free protease inhibitor cocktail (Roche) according to the manufacturer's instructions, and lysed by ultrasonification for 20 s, followed by incubation on ice for 60 s, repeated ten times. After clarification of the lysate by centrifugation at 23 000 rev./min for 1 h using a Beckman Coulter JS13.1 rotor, the supernatant was applied to a Ni^{2+}-NTA (nitrilotriacetic acid) column using an ÄKTA FPLC system (GE Healthcare) and washed first with 2× PBS containing 20 mM 2ME, and then with 2× PBS containing 20 mM 2ME and 25 % (v/v) glycerol. His_6–DEG7 was eluted with 50 mM imidazole and concentrated using Vivaspin concentrator tubes (Vivascience), and subjected to SEC using a Superdex 200 16/300 GL column (GE Healthcare) and 2× PBS with 20 mM 2ME to remove any remaining impurities. For a more accurate identification of the elution volume, the main fraction containing DEG7 (as identified by immunoblotting) was concentrated as described above and subjected to a second run on a Superdex 200 16/300 GL column. To correlate elution volume and molecular mass, the column was calibrated with Blue Dextran, ferritin (440 kDa), aldolase (158 kDa) and ovalbumin (43 kDa) according to the manufacturer's instructions (GE Healthcare).

Truncated versions of DEG7 with an N-terminal His_6 tag were produced at 30 °C in *E. coli* strain BL21 Star™ (DE3) cells carrying plasmid pHS183, pHS184, pHS185 or pHS186. Cells were grown in 2 litres of 2 × YT medium [32] containing 2 % (w/v) glucose and 100 $\mu g \cdot l^{-1}$ ampicillin to a D_{600} of 0.8, and *DEG7* overexpression was induced by adding IPTG to a final concentration of 0.3 mM. After 3 h, cells were harvested by centrifugation at 6000 ***g*** for 10 min and resuspended in 50 mM Hepes/KOH (pH 8.0), 300 mM NaCl and 50 mM imidazole, supplemented with 10 mM 2ME and Complete™ EDTA-free protease inhibitor cocktail. Cells were lysed and centrifuged as described above. Proteins were purified with Ni^{2+}-NTA columns

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

on an ÄKTA FPLC system, using an imidazole gradient from 50 mM to 500 mM. The fraction containing the maximum amount of His_6-tagged proteins was subjected to SEC as described above, using a buffer containing 50 mM Hepes/KOH (pH 8.0), 300 mM NaCl and 10 mM 2ME.

Electrophoresis and immunoblotting

SDS/PAGE and immunoblotting were conducted using the method described in [32]. A DEG7-specific antibody was raised in rabbit against the peptide (Ac-)KGGSSGSPVIDWQGR(-COOH) (AgriSera AB) and used at 1:500 dilution.

Yeast two-hybrid assay

A Y2H assay was performed using pBD-Gal4 Cam as bait and pAD-Gal42.1 as the prey vector from the GAL4 phagemid vector kit (Stratagene). See the Plasmid construction section for details of cloning of the respective plasmids. pBD-WT and pAD-WT (Stratagene; both coding for fragment C of lambda *cl* repressor) were used as positive interaction control plasmids, pAD-WT and pLamin C (Stratagene; coding for human lamin C) were used as negative interaction control plasmids. *S. cerevisiae* strain YRG-2 was used for all Y2H assays. All experiments were performed at least in triplicate, starting each time from the transformation of the plasmids into the yeast cells. For composition of media and yeast transformation by the lithium acetate method, see http://home.cc.umanitoba.ca/∼gietz/ and [34]. For *HIS3* reporter gene assays, candidate clones were cultured overnight in liquid SC (synthetic complete) dropout medium lacking leucine, uracil and tryptophan, pelleted, and resuspended to give a D_{600} of 1.0. A 50 μl volume of the suspension, as well as 10×, 100× and 1000× dilutions were dropped on solid SC medium lacking either leucine, uracil and tryptophan (growth control) or histidine, leucine, uracil and tryptophan (reporter gene assay). Plates were grown for 2–3 days at 30 °C.

RESULTS AND DISCUSSION

Taxonomic distribution of DEG7 orthologues

We searched freely available genome databases for genes encoding DEG7-like proteases and retrieved hits for fungi and plant (including green, heterokont and haptophyte algae) genomes. All DEG7 orthologues share the unusual domain arrangement with two protease domains and up to four PDZ domains [2,4] (see Supplementary Figure S1 at http://www.BiochemJ.org/bj/435/bj4350167add.htm for sequence alignment data of the full-length proteins). The second protease domain is degenerated and lacks the active-site residues, such as histidine, aspartic acid and serine (results not shown), but is still recognized when scanned against the InterPro database [35] and the HHpred program [27] (results not shown). In contrast, four PDZ domains are not detected in every DEG7 orthologue, although the overall length of the proteins is comparable. This indicates that, provided the gene models retrieved from the databases are correct, some PDZ domains are not conserved enough to be detected by the domain prediction programs used in the present study. The domain architecture suggests that DEG7-like proteases evolved from a whole-gene duplication/fusion event of a DegP-like protease (containing a protease domain and two PDZ domains), followed by a subsequent degeneration of the second protease domain (Figure 1, and Supplementary Figure S2 at http://www.BiochemJ.org/bj/435/bj4350167add.htm). No DEG7 orthologues were found in animals, since all hits have much shorter sequences and no hit contained a second protease domain (query coverage <30 %; results not shown). Notably, genes coding for DEG7 orthologues were also absent from the genomes analysed of a primitive red alga (*Cyanidioschyzon merolae*), Cryptophyta (*Guillardia theta* and *Hemiselmis andersenii*) and diatoms (*Phaeodactylum tricornutum* and *Thalassiosira pseudonana*).

Phylogenetic comparison (see Supplementary Figure S3 at http://www.BiochemJ.org/bj/435/bj4350167add.htm for sequence alignment data) confirmed the close evolutionary relationship of DEG7 orthologues from various organisms (Figure 2). Two main clades of DEG7 proteases can be distinguished: clade A contains DEG7 proteases from fungi, whereas DEG7 proteins from green, heterokont and haptophyte algae, mosses and higher plants form clade B. Whereas most species investigated possess only one *DEG7*-like gene, the genomes of the yeasts *Schizosaccharomyces pombe* and *Schizosaccharomyces japonicus*, the ascomycete *Giberella zeae,* the oomycete *Phytophtora soja*, the haptophyte alga *Emiliania huxleyi* and the moss *Physcomitrella patens* contain two, whereas the higher plant *Populus trichocarpa* possesses three *DEG7* paralogues (Figure 2). Expressed sequence tag data from *S. pombe*, *P. patens* and *P. trichocarpa* indicate that all paralogues are expressed in these organisms (TGI; results not shown), but whether they have overlapping/complementary physiological roles is still unknown.

The distribution of DEG7 proteases to two clades may reflect the different intracellular localizations for DEG7-like proteins from plants and fungi: DEG7 from *A. thaliana* was reported to be a plastidial [24] protein, whereas Ynm3p from yeast was found in the nucleus [21,22]. However, one of the two DEG7 paralogues from *S. pombe* and *S. japonicus* (Figure 2, marked with asterisks) is also present in clade B, together with their plant orthologues. Although the intracellular localization of these proteases has not been examined so far, a lack of plastids in yeast points to other than a plastidial localization.

DEG7 forms homotrimers

Almost all Deg/HtrA proteases examined so far are present as homo-oligomers, ranging from trimers, e.g. DegS from *E. coli* [8,36] and mammalian HtrA2 [7], to 24-mers, e.g. DegP from *E. coli* [12,13]. To investigate the oligomerization state of DEG7, we performed SEC with recombinant tagged DEG7 purified from *E. coli* (Figure 3). Three absorption peaks at 280 nm were observed during the elution (Figure 3A). DEG7 was present in one peak, as analysed by immunoblotting with an anti-DEG7 antibody and SDS/PAGE with subsequent Coomassie Blue staining (Figures 3B and 3C). The elution volume of 11.84 ml corresponds to an apparent molecular mass of 352 kDa, as identified by calibrating the size-exclusion column with proteins of known molecular mass. Since the recombinant tagged DEG7 has a calculated molecular mass of 120 kDa per monomer, this result indicates the formation of a trimeric complex.

Identification of oligomerization site by Y2H screen and SEC

In all Deg/HtrA proteases examined so far, the formation of the trimer as the basic unit is mediated by residues of the individual active protease domains [1]. Since DEG7 seems to be the result of a *Deg/HtrA* gene-duplication and fusion event (Figure 1),

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

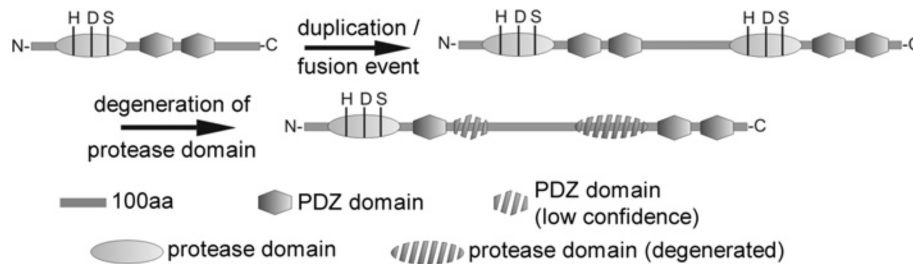

Figure 1 DEG7 in plant and fungi are the result of an internal whole-gene duplication/fusion event of an ancient *Deg/HtrA* gene

An ancient *Deg/HtrA* gene encoding one protease domain and two PDZ domains was duplicated and the gene copy generated was fused within the same open reading frame, resulting in a gene encoding a protein with two protease and four PDZ domains. Subsequently, the second protease domain, containing a catalytic triad of histidine (H), aspartate (D) and serine (S), residues, degenerated. 100aa, 100-amino-acid stretch.

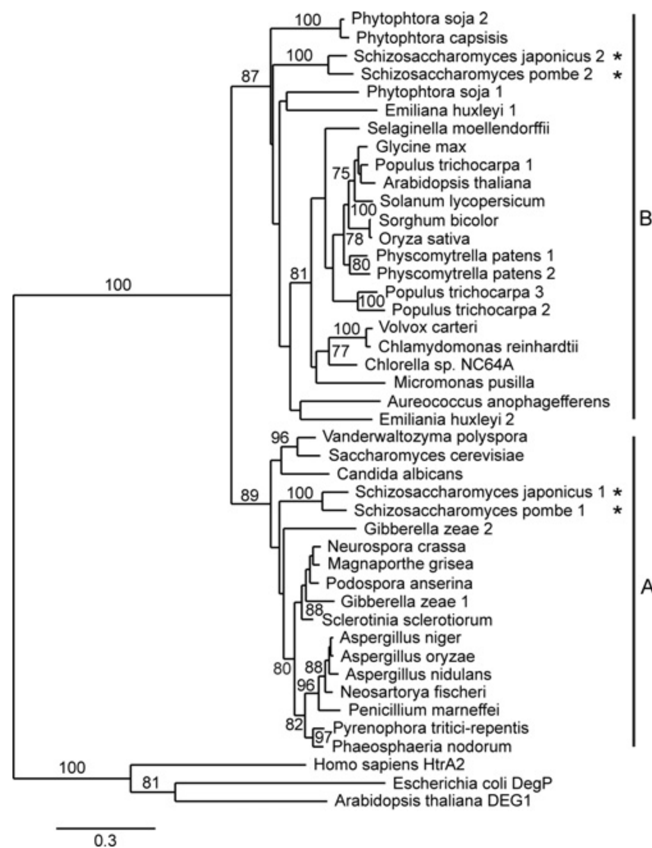

Figure 2 Evolutionary relationship of DEG7 proteases and selected members from the Deg/HtrA family

The Neighbour-Joining phylogenetic tree is based on the amino acid sequences of Deg/HtrA protease domains. The result of a bootstrap test (1000 replicates) is shown, and branches with confident values >75 % are indicated. Maximum-Likelihood and Parsimony methods resulted in essentially the same trees (not shown). Human HtrA2, *E. coli* DegP and DEG1 from *A. thaliana* were chosen as outgroups, since they do not belong to the DEG7 group [4]. Group A contains DEG7 orthologues from fungi, Group B contains mainly DEG7 proteases from algae, mosses and higher plants. Asterisks mark the distribution of two DEG7 paralogues from *S. japonicus* und *S. pombe* within groups A and B.

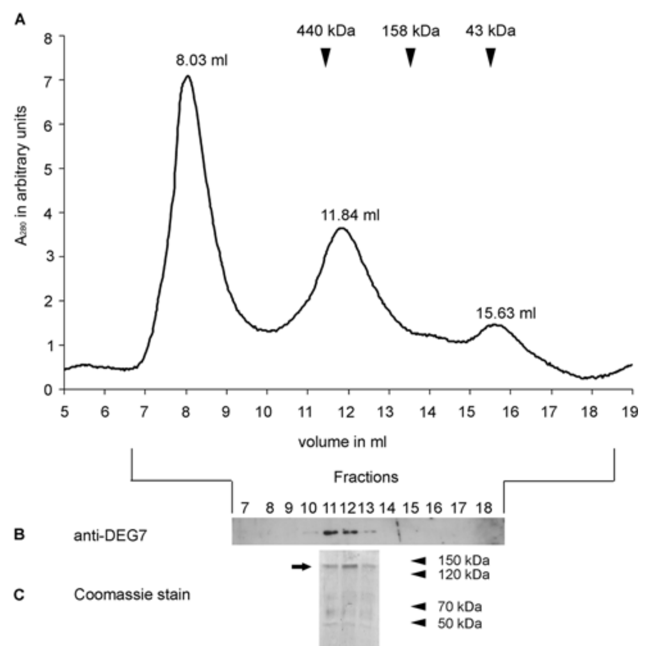

Figure 3 Recombinant DEG7 forms trimers

(**A**) Elution profile of purified recombinant DEG7. Three peaks were eluted from the SEC column with the indicated elution volumes. Approximate elution volumes of standard proteins are indicated by arrows. (**B**) Immunoblot analysis of proteins from marked fractions, using a DEG7 specific antibody, indicates the presence of DEG7 in the 11.84 ml elution peak. (**C**) Coomassie Blue staining of an SDS gel loaded with proteins from indicated elution fractions. DEG7 (arrow) is the main protein in the 11.84 ml elution peak. Molecular masses are indicated in kDa.

the DEG7 trimer may be compared with the hexameric state described for DegP from *E. coli* [9] rather than the trimers observed for DegS from *E. coli* [8,36] or human HtrA2 [7]. However, the unique domain arrangement of DEG7 raises the question of which domains mediate the observed trimerization (Figure 3). To address this question, a Y2H array was set up with protease and PDZ domains and combinations of these as baits and as preys: full-length inactive DEG7 with the catalytic serine residue mutated to alanine (DEG7-fl) was fused to the AD and BD domain of the yeast GAL4 protein respectively. Additionally, the first half (with the catalytic serine residue mutated to alanine) and the second half of this protein were also fused to the AD domain, as well as the BD domain. To analyse individual domains, the predicted active protease domain (with the catalytic serine residue mutated to alanine), the second (degenerated) protease domain, PDZ1 and PDZ3 + 4 domains were fused to the AD and the BD domain. Since the second PDZ domain in the first half of the protein was less defined (Figure 1), this construct was omitted. The resulting fusion constructs were co-expressed in *S. cerevisiae* in every possible AD/BD combination and tested for their ability to activate expression of the *HIS3* reporter gene that indicates physical interaction of the constructs (summarized in Table 1 and Figure 4). No interaction

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

Table 1 Interaction of DEG7 domains assayed by Y2H

Abbreviations: +/+, *HIS3* reporter gene activation with both possible AD- and BD-fusion combinations; +/−, *HIS3* reporter gene activation with one combination only, −/− no *HIS3* reporter gene activation with either possible combination. Interacting domains are marked in bold.

Constructs	Interaction
DEG7-fl/DEG7-fl	**+/+**
DEG7-fl/first half	+/−
DEG7-fl/second half	**+/+**
DEG7-fl/active protease domain	+/−
DEG7-fl/degenerated protease domain	+/−
DEG7-fl/PDZ1	−/−
DEG7-fl/PDZ3 + 4	−/−
DEG7-fl/empty	−/−
First half/first half	−/−
First half/second half	**+/+**
First half/active protease domain	−/−
First half/degenerated protease domain	+/−
First half/PDZ1	−/−
First half/PDZ3 + 4	−/−
First half/empty	−/−
Second half/second half	**+/+**
Second half/active protease domain	+/−
Second half/degenerated protease domain	**+/+**
Second half/PDZ1	−/−
Second half/PDZ3 + 4	−/−
Second half/empty	−/−
Active protease domain/active protease domain	−/−
Active protease domain/degenerated protease domain	+/−
Active protease domain/PDZ1	−/−
Active protease domain/PDZ3 + 4	−/−
Active protease domain/empty	−/−
Degenerated protease domain/degenerated protease domain	**+/+**
Degenerated protease domain/PDZ1	−/−
Degenerated protease domain/PDZ3 + 4	−/−
Degenerated protease domain/empty	−/−
PDZ1/PDZ1	−/−
PDZ1/PDZ3 + 4	−/−
PDZ1/empty	−/−
PDZ3 + 4/PDZ3 + 4	−/−
PDZ3 + 4/empty	−/−

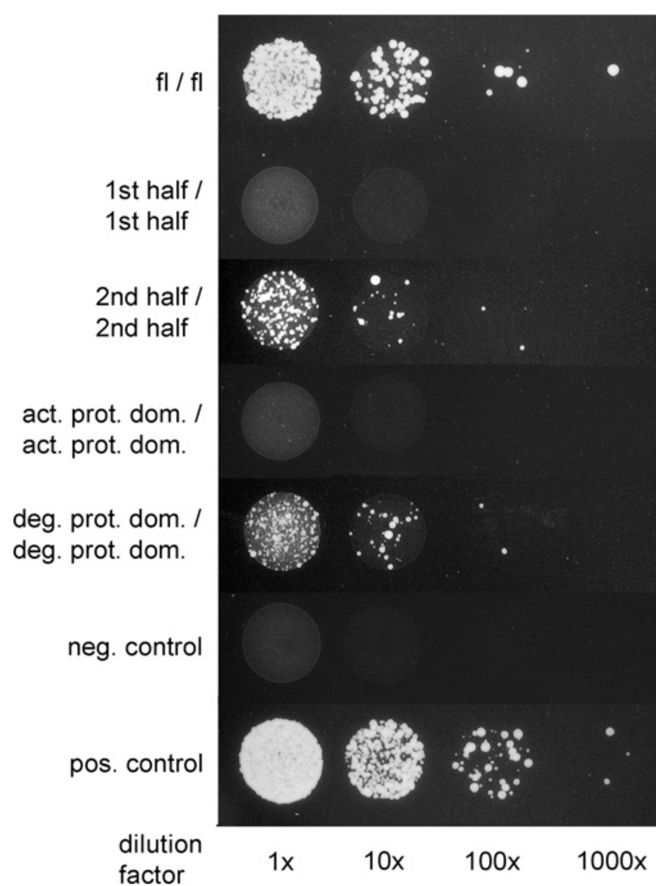

Figure 4 Y2H assay shows that oligomerization of DEG7 is mediated by the second (degenerated) protease domain

Growth of *S. cerevisiae* YRG2 cells on histidine-free medium (indicative for interaction) is only visible when the combination of interaction partners contains the second (degenerated) protease domain (deg. prot. dom.), which is part of the second half (2nd half) and the full-length protein (fl). No interaction occurs with constructs containing only the first half (1st half) or first (active) protease domain (act. prot. dom.). The columns show different dilutions of the yeast cultures. A dilution factor of 1× corresponds to a D_{600} of 1.0. Controls: neg., negative; pos., positive.

was detected when BD- and AD-fusion constructs were co-transformed with the empty AD or the BD vector respectively (Table 1), thus excluding autoactivation of the *HIS3* reporter gene. Similarly, no interaction was observed for constructs containing only the PDZ1 domain or PDZ3 + 4 domains, indicating that the PDZ domains are not essential for the assembly of the oligomeric complex. In all combinations that activated reporter gene expression, at least one of the interaction partners contained the second (degenerated) protease domain (Figure 4 and Table 1). No interaction was observed when this part was lacking in both fusion constructs (Figure 4 and Table 1). An additional interaction between the active and the degenerated protease domain cannot be ruled out, since some combinations (DEG7-fl/first half, DEG7-fl/active protease domain first half/second half, first half/degenerated protease domain, second half/active protease domain, active protease domain/degenerated protease domain, given as AD/BD combinations) also activated reporter gene expression (Table 1, see also Supplementary Figure S4 at http://www.BiochemJ.org/bj/435/bj4350167add.htm). This also indicates that the lack of reporter gene expression with the combinations first half/first half and active protease domain/active protease domain is not due to a lack of protein expression in the yeast, since some combinations including the AD–first half, the BD–first half, the AD–active protease domain, or the BD–active protease domain construct, promoted growth of the yeast strain on histidine-free medium (see Supplementary Figure S4). In summary, the Y2H data suggest that the second (degenerated) protease domain is the key mediator for the formation of the trimeric DEG7 complex in plants. The active protease domain does not seem to be sufficient for trimer formation, in contrast with all Deg/HtrA proteases investigated so far.

To confirm these findings by a method independent of the yeast system, the oligomerization status of the first and the second half of DEG7, as well as the first (active) and the second (degenerated) protease domain alone was analysed by SEC with recombinant proteins expressed in *E. coli*. Such truncated DEG7 constructs were affinity-purified as His_6-tagged recombinant soluble proteins, and subsequently analysed by SEC and immunoblotting (Figure 5). Although the constructs representing the first and second halves of DEG7 have similar calculated molecular masses (65.1 and 66.2 kDa respectively), they exhibit totally different elution patterns. The first half of DEG7 elutes as one peak in fractions between 14.0 and 14.5 ml (corresponding to an apparent molecular mass of approx. 75–100 kDa), whereas the second half of DEG7 is present in two distinct peaks (fractions

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

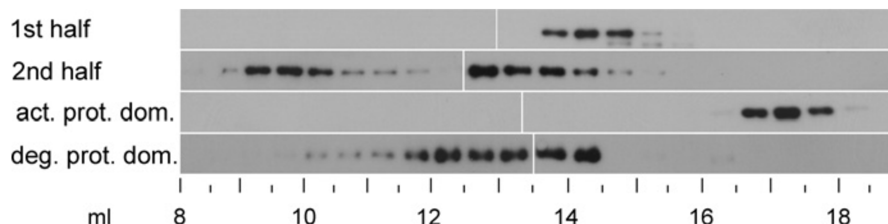

Figure 5 Only recombinant DEG7 containing the second (degenerated) protease domain forms oligomers

Immunoblot (anti-His$_5$-tag antibody) showing the distribution of truncated DEG7 after SEC. Fractions are indicated as vertical lines at the bottom, and numbers (in ml) indicate the respective elution volume. Truncated DEG7 proteins containing the second (degenerated) protease domain (2nd half, deg. prot. dom.) are eluted in two peaks at lower buffer volumes as compared with truncated proteins with the first (active) protease domain (1st half, act. prot. dom.), which are eluted in a single peak at the higher buffer volume.

between 9.5 and 10.0 ml, and 12.5 and 13.0 ml, corresponding to apparent molecular masses of 980–1300 and 170–240 kDa respectively). Similar differences in the elution patterns are also observed for the active protease domain as compared with the degenerated protease domain (Figure 5). Whereas the active protease domain (calculated molecular mass of 27.9 kDa) is present in the fractions between 17.0 and 17.5 ml (corresponding to an apparent molecular mass of 13–18 kDa) the degenerated protease domain (calculated molecular mass of 31.5 kDa) elutes as two poorly resolved peaks between 12.0 and 12.5 ml, and 14.0 and 14.5 ml, corresponding to apparent molecular masses of 230–315 and 75–100 kDa respectively). Although the estimation of apparent molecular masses of oligomers is too inaccurate to determine the exact stoichiometry of the complexes, these data suggest that the first half of DEG7 and the active protease domain are eluted as monomers, whereas the second half of DEG7 and the degenerated protease domains form at least two types of oligomer, possibly a trimer and a higher oligomeric complex. The inaccuracy in the estimation of molecular masses of monomers and oligomers could be due to sterical factors and the fact that DEG7 deletion constructs are probably far from a globular shape, which might be different for the intact protein. For the deletion construct containing the active protease domain, this might also explain the difference between the molecular mass according to the SEC and that estimated by SDS/PAGE, which matched the calculated values (results not shown). In summary, we can conclude from these data that only DEG7 deletion constructs with the degenerated protease domain are able to form oligomeric complexes, thus confirming the data obtained from the Y2H screen (Figure 4 and Table 1, and see Supplementary Figure S4).

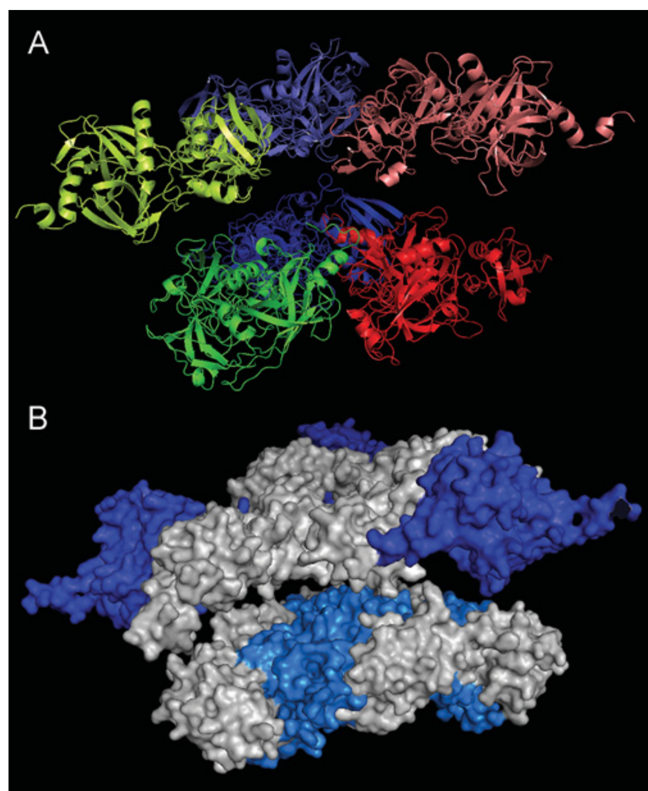

Figure 6 Homology modelling of a DEG7 trimer

The putative structure of a DEG7 trimer was modelled on the hexameric structure of *E. coli* DegP [9]. The approx. 100-amino-acid stretch between the second PDZ and the degenerated protease domain (Figure 1) was omitted for clarity. Relative orientations of the first half of the protein are not supported. (**A**) Cartoon model, side view. Monomeric subunits are coloured blue, green and red respectively. The second halves of the protein are at the bottom, coloured darker than the first halves. (**B**) Surface model, side view. The active protease domains are coloured blue, the second (degenerated) protease domains involved in oligomerization are coloured light blue. This model might represent a protease oligomer in an inactive conformation.

Modelling of DEG7 structure

To illustrate how DEG7 might form oligomers, we built a model based on our findings that DEG7 forms trimers mediated by interactions of second (degenerated) protease domains, and the available crystal structures of *E. coli* DegP [9,12] and human HtrA2 [7] (Figure 6). For clarity, the approx. 100-amino-acid stretch between the second PDZ domain and the degenerated protease domain (Figure 1) was omitted. The oligomerization via the second (degenerated) protease domain might enable the protein to keep the catalytic subunits in close proximity, without the actual formation of a proteolytic active complex. This might ensure a very rapid initiation of proteolysis after binding of a substrate. However, details in this model, especially the relative orientations of the active protease domains, need to be resolved further.

Conclusions

We have shown that the DEG7 protease from *A. thaliana* has orthologues in green, heterokont and haptophyte algae, in fungi, mosses and higher plants, but no orthologues can be found in animals. The unusual domain arrangement of DEG7 with two protease domains (one degenerated) and four PDZ domains requires a different mode of oligomerization than the one which was described for all other Deg/HtrA proteases. Although trimerization through an active protease domain is a common

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

feature of Deg/HtrA proteases in bacteria and mammals, the active protease domain of plant DEG7 is not essential for this process. Instead, the second (degenerated) protease domain is responsible for the trimerization of DEG7 in plants.

Inactive protease domains were reported from the proteolytic active oligomers of other proteases. Clp proteases in bacteria, for example, are oligomeric serine proteases with a proteolytic core constituted by two heptameric rings, with each heptamer composed of identical proteolytically active ClpP subunits [37]. In chloroplasts of plants and in cyanobacteria, however, the Clp protease core is a heteroheptamer, containing three proteolytically active ClpP subunits and four inactive ClpR variants, which show sequence similarity to ClpP on the amino acid level, but lack the catalytic triad of serine type proteases [38,39]. A high sequence similarity between the inactive ClpR and the active ClpP variants indicates a common ancestor [40]. This is supported by the fact that loss of the catalytic triad appears to be ongoing process since some cyanobacterial ClpR variants still retain one of the active-site residues [39].

A similar situation was reported for the 20S proteolytically active core unit of the 26S proteasome. This core unit is composed of four stacked heptameric rings. In Archaea, the two inner rings are homo-oligomers of proteolytically active β-subunits, whereas, in eukaryotes, the β-subunit has diverged into seven different β-proteins (reviewed in [41,42]). Of these seven β-subunits, only three contain the active-site residues and are proteolytically active. Despite the lack of enzymatic activity, the inactive β-subunits are essential for the correct assembly of the eukaryotic proteasome.

In the case of ClpP from photosynthetic organisms and the proteasome from eukaryotes, formerly active protease domains (or, in these cases, one-domain subunits) evolved into domains with tasks other than proteolysis, e.g. mediating subunit contacts in the oligomeric protease complex. The second (degenerated) protease domain of DEG7, which is described in the present paper, might be yet another example, where a formerly active enzymatic entity is now involved in the oligomerization of the protease complex. For Deg/HtrA proteases, inactive protease domains involved in oligomer formation have not been described so far. In contrast with ClpP and the proteasome, however, this inactive domain is not encoded by a separate gene, but is a part of the same polypeptide chain as the active protease domain.

AUTHOR CONTRIBUTION

Holger Schuhmann and Ulrike Mogg performed research, analysed data and wrote the draft of the paper. Iwona Adamska conceived the research and edited the paper before submission.

ACKNOWLEDGEMENTS

We are grateful to Silvia Kuhn for technical assistance in the Y2H assays, and Dr Dietmar Funck for helpful advice. We thank the Nottingham *Arabidopsis* Stock Centre (NASC) for providing the cDNA clones.

FUNDING

This work was supported by grants from the Deutsche Forschungsgemeinschaft [grant numbers AD92/8-2 and AD92/8-3] and Konstanz University (to I.A.).

REFERENCES

1 Kim, D. Y. and Kim, K. K. (2005) Structure and function of HtrA family proteins, the key players in protein quality control. J. Biochem. Mol. Biol. **38**, 266–274

2 Clausen, T., Southan, C. and Ehrmann, M. (2002) The HtrA family of proteases: implications for protein composition and cell fate. Mol. Cell **10**, 443–455

3 Rawlings, N. D., Morton, F. R., Kok, C. Y., Kong, J. and Barrett, A. J. (2008) MEROPS: the peptidase database. Nucleic Acids Res. **36**, D320–D325

4 Helm, M., Luck, C., Prestele, J., Hierl, G., Huesgen, P. F., Fröhlich, T., Arnold, G. J., Adamska, I., Gorg, A., Lottspeich, F. and Gietl, C. (2007) Dual specificities of the glyoxysomal/peroxisomal processing protease Deg15 in higher plants. Proc. Natl. Acad. Sci. U.S.A. **104**, 11501–1156

5 Schuhmann, H., Huesgen, P. F., Gietl, C. and Adamska, I. (2008) The DEG15 serine protease cleaves peroxisomal targeting signal 2-containing proteins in *Arabidopsis thaliana*. Plant Physiol. **148**, 1847–1856

6 Sun, X., Peng, L., Guo, J., Chi, W., Ma, J., Lu, C. and Zhang, L. (2007) Formation of DEG5 and DEG8 complexes and their involvement in the degradation of photodamaged photosystem II reaction center D1 protein in *Arabidopsis*. Plant Cell **19**, 1347–1361

7 Li, W., Srinivasula, S. M., Chai, J., Li, P., Wu, J. W., Zhang, Z., Alnemri, E. S. and Shi, Y. (2002) Structural insights into the pro-apoptotic function of mitochondrial serine protease HtrA2/Omi. Nat. Struct. Biol. **9**, 436–441

8 Wilken, C., Kitzing, K., Kurzbauer, R., Ehrmann, M. and Clausen, T. (2004) Crystal structure of the DegS stress sensor: how a PDZ domain recognizes misfolded protein and activates a protease. Cell **117**, 483–494

9 Krojer, T., Garrido-Franco, M., Huber, R., Ehrmann, M. and Clausen, T. (2002) Crystal structure of DegP (HtrA) reveals a new protease-chaperone machine. Nature **416**, 455–459

10 Jomaa, A., Damjanovic, D., Leong, V., Ghirlando, R., Iwanczyk, J. and Ortega, J. (2007) The inner cavity of *Escherichia coli* DegP protein is not essential for molecular chaperone and proteolytic activity. J. Bacteriol. **189**, 706–716

11 Kim, D. Y., Kim, D. R., Ha, S. C., Lokanath, N. K., Lee, C. J., Hwang, H. Y. and Kim, K. K. (2003) Crystal structure of the protease domain of a heat-shock protein HtrA from *Thermotoga maritima*. J. Biol. Chem. **278**, 6543–6551

12 Krojer, T., Sawa, J., Schäfer, E., Saibil, H. R., Ehrmann, M. and Clausen, T. (2008) Structural basis for the regulated protease and chaperone function of DegP. Nature **453**, 885–890

13 Jiang, J., Zhang, X., Chen, Y., Wu, Y., Zhou, Z. H., Chang, Z. and Sui, S. F. (2008) Activation of DegP chaperone-protease via formation of large cage-like oligomers upon binding to substrate proteins. Proc. Natl. Acad. Sci. U.S.A. **105**, 11939–11944

14 Ades, S. E., Connolly, L. E., Alba, B. M. and Gross, C. A. (1999) The *Escherichia coli* σ^E-dependent extracytoplasmic stress response is controlled by the regulated proteolysis of an anti-σ factor. Genes Dev. **13**, 2449–2461

15 Alba, B. M., Leeds, J. A., Onufryk, C., Lu, C. Z. and Gross, C. A. (2002) DegS and YaeL participate sequentially in the cleavage of RseA to activate the σ^E-dependent extracytoplasmic stress response. Genes Dev. **16**, 2156–2168

16 Lipinska, B., Fayet, O., Baird, L. and Georgopoulos, C. (1989) Identification, characterization, and mapping of the *Escherichia coli htrA* gene, whose product is essential for bacterial growth only at elevated temperatures. J. Bacteriol. **171**, 1574–1584

17 Strauch, K. L., Johnson, K. and Beckwith, J. (1989) Characterization of *degP*, a gene required for proteolysis in the cell envelope and essential for growth of *Escherichia coli* at high temperature. J. Bacteriol. **171**, 2689–2696

18 Verhagen, A. M., Silke, J., Ekert, P. G., Pakusch, M., Kaufmann, H., Connolly, L. M., Day, C. L., Tikoo, A., Burke, R., Wrobel, C. et al. (2002) HtrA2 promotes cell death through its serine protease activity and its ability to antagonize inhibitor of apoptosis proteins. J. Biol. Chem. **277**, 445–454

19 Kurochkin, I. V., Mizuno, Y., Konagaya, A., Sakaki, Y., Schonbach, C. and Okazaki, Y. (2007) Novel peroxisomal protease Tysnd1 processes PTS1- and PTS2-containing enzymes involved in β-oxidation of fatty acids. EMBO J. **26**, 835–845

20 Huesgen, P. F., Schuhmann, H. and Adamska, I. (2005) The family of Deg proteases in cyanobacteria and chloroplasts of higher plants. Physiol. Plant. **123**, 413–420

21 Padmanabhan, N., Fichtner, L., Dickmanns, A., Ficner, R., Schulz, J. B. and Braus, G. H. (2009) The yeast HtrA orthologue Ynm3 is a protease with chaperone activity that aids survival under heat stress. Mol. Biol. Cell **20**, 68–77

22 Fahrenkrog, B., Sauder, U. and Aebi, U. (2004) The *S. cerevisiae* HtrA-like protein Nma111p is a nuclear serine protease that mediates yeast apoptosis. J. Cell Sci. **117**, 115–126

23 Tong, F., Black, P. N., Bivins, L., Quackenbush, S., Ctrnacta, V. and DiRusso, C. C. (2006) Direct interaction of *Saccharomyces cerevisiae* Faa1p with the Omi/HtrA protease orthologue Ynm3p alters lipid homeostasis. Mol. Genet. Genomics **275**, 330–343

24 Sun, X., Fu, T., Chen, N., Guo, J., Ma, J., Zou, M., Lu, C. and Zhang, L. (2010) The stromal chloroplast Deg7 protease participates in the repair of photosystem II after photoinhibition in *Arabidopsis*. Plant Physiol. **152**, 1263–1273

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

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

26 Zdobnov, E. M. and Apweiler, R. (2001) InterProScan: an integration platform for the signature-recognition methods in InterPro. Bioinformatics **17**, 847–848
27 Soding, J., Biegert, A. and Lupas, A. N. (2005) The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res. **33**, W244–W248
28 Moretti, S., Armougom, F., Wallace, I. M., Higgins, D. G., Jongeneel, C. V. and Notredame, C. (2007) The M-Coffee web server: a meta-method for computing multiple sequence alignments by combining alternative alignment methods. Nucleic Acids Res. **35**, W645–W648
29 Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J. F., Guindon, S., Lefort, V., Lescot, M. et al. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. **36**, W465–W469
30 Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., Pieper, U. and Sali, A. (2007) Comparative protein structure modeling using MODELLER. Curr. Protoc. Protein Sci. **Chapter 2**, Unit 29
31 Zhang, Y., Appleton, B. A., Wu, P., Wiesmann, C. and Sidhu, S. S. (2007) Structural and functional analysis of the ligand specificity of the HtrA2/Omi PDZ domain. Protein Sci. **16**, 1738–1750
32 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: a Laboratory Manual, 2nd edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor
33 Yamada, K., Lim, J., Dale, J. M., Chen, H., Shinn, P., Palm, C. J., Southwick, A. M., Wu, H. C., Kim, C., Nguyen, M. et al. (2003) Empirical analysis of transcriptional activity in the *Arabidopsis* genome. Science **302**, 842–846
34 Gietz, R. D. and Woods, R. A. (2002) Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods Enzymol. **350**, 87–96
35 Hunter, S., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Das, U., Daugherty, L., Duquenne, L. et al. (2009) InterPro: the integrative protein signature database. Nucleic Acids Res. **37**, D211–D215
36 Zeth, K. (2004) Structural analysis of DegS, a stress sensor of the bacterial periplasm. FEBS Lett. **569**, 351–358
37 Wang, J., Hartling, J. A. and Flanagan, J. M. (1997) The structure of ClpP at 2.3 Å resolution suggests a model for ATP-dependent proteolysis. Cell **91**, 447–456
38 Peltier, J. B., Ripoll, D. R., Friso, G., Rudella, A., Cai, Y., Ytterberg, J., Giacomelli, L., Pillardy, J. and van Wijk, K. J. (2004) Clp protease complexes from photosynthetic and non-photosynthetic plastids and mitochondria of plants, their predicted three-dimensional structures, and functional implications. J. Biol. Chem. **279**, 4768–4781
39 Andersson, F. I., Tryggvesson, A., Sharon, M., Diemand, A. V., Classen, M., Best, C., Schmidt, R., Schelin, J., Stanne, T. M., Bukau, B. et al. (2009) Structure and function of a novel type of ATP-dependent Clp protease. J. Biol. Chem. **284**, 13519–13532
40 Garcia-Lorenzo, M., Sjodin, A., Jansson, S. and Funk, C. (2006) Protease gene families in *Populus* and *Arabidopsis*. BMC Plant Biol. **6**, 30
41 Gallastegui, N. and Groll, M. (2010) The 26S proteasome: assembly and function of a destructive machine. Trends Biochem. Sci. **35**, 634–642
42 Groll, M., Bochtler, M., Brandstetter, H., Clausen, T. and Huber, R. (2005) Molecular machines for protein degradation. ChemBioChem **6**, 222–256

Received 1 October 2010/19 January 2011; accepted 19 January 2011
Published as BJ Immediate Publication 19 January 2011, doi:10.1042/BJ20101613

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

Biochem. J. (2011) **435**, 167–174 (Printed in Great Britain) doi:10.1042/BJ20101613

online data

SUPPLEMENTARY ONLINE DATA

A new principle of oligomerization of plant DEG7 protease based on interactions of degenerated protease domains

Holger SCHUHMANN, Ulrike MOGG and Iwona ADAMSKA[1]

Department of Physiology and Plant Biochemistry, University of Konstanz, Universitätsstrasse 10, D-78457 Konstanz, Germany

Table S1 DEG7 orthologues and their accession numbers (SwissProt Protein Database)

*JGI protein ID, no entry in SwissProt Database; **GenBank® accession number, no entry in SwissProt Database.

Organism	Accession number
Arabidopsis thaliana	Q8RY22
Aspergillus nidulans	Q5B1Z4
Aspergillus niger	A5AB13
Aspergillus oryzae	Q2TYB1
Aureococcus anophagefferens	322*
Candida albicans	Q5A946
Chlamydomonas reinhardtii	A8JH35
Chlorella sp. NC64A	37665*
Emiliania huxleyi (1)	452178*
Emiliania huxleyi (2)	245118*
Gibberella zeae (1)	UPI000023D1E1
Gibberella zeae (2)	UPI000023F481
Glycine max	AK287315**
Magnaporthe grisea	A4RJH4
Micromonas pusilla	35996*
Neosartorya fischeri	A1DP85
Neurospora crassa	Q7S9D2
Oryza sativa	B9F2C1
Penicillium marneffei	B6QAL6
Phaeosphaeria nodorum	Q0UY70
Physcomitrella patens (1)	A9TIB2
Physcomitrella patens (2)	A9RQ61
Phytophtora capsici	27218*
Phytophtora soja (1)	133655*
Phytophtora soja (2)	199402*
Podospora anserina	B2ASP9
Populus trichocarpa (1)	B9GV35
Populus trichocarpa (2)	B9H390
Populus trichocarpa (3)	B9H391
Pyrenophora tritici-repentis	B2WNT3
Saccharomyces cerevisiae	P53920
Schizosaccharomyces japonicus (1)	B6K3R7
Schizosaccharomyces japonicus (2)	B6JWG1
Schizosaccharomyces pombe (1)	Q9P7S1
Schizosaccharomyces pombe (2)	O74325
Sclerotinia sclerotiorum	A7E9G4
Selaginella moellendorffii	165477*
Solanum lycopersicum	AK321684**
Sorghum bicolor	5004613*
Vanderwaltozyma polyspora	A7TGI3
Volvox carteri	79278*

Table S2 List of oligonucleotides used as PCR primers

Name	Sequence (5′→3′)
0724	GGTACTAAAGGTGGTTCAGCTGGTTCTCCCGTCATTG
0725	CAATGACGGGAGAACCAGCTGAACCACCTTTAGTACC
0734	TATGTCGACTTACTGCAAGGCTTTC
0740	CACCATGGGAGATCCGTTGGA
0747	GCGCCGCGGTTACTGCAAGGCTTTC
0765	GCGTTACTTATCCGTCCGCGAGTC
0767	GCGTTATGAGTGTAAATCTTGTACTGATA
0771	GCGTTATTGTTTCTTTGCTTCTGAGCC
0774	CACCTCCGTTGCCACCGCTGAAGATTG
0775	CACCGATAAGCCAAAAGCAGTTCATATTC
0776	CACCTCAAAAGCCCGGAGTTTTGGTC
0777	CACCGAACCCATGCATGAAGTGAATG
0793	CACCGGCAGTGAATTTAAATCTGATG
0797	CGCTTATCCATTTCCACCGGTTATGATT
0799	CGCTTAATCTTGTTTCAGAGTCAATACTC

[1] To whom correspondence should be addressed (e-mail iwona.adamska@uni-konstanz.de).

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

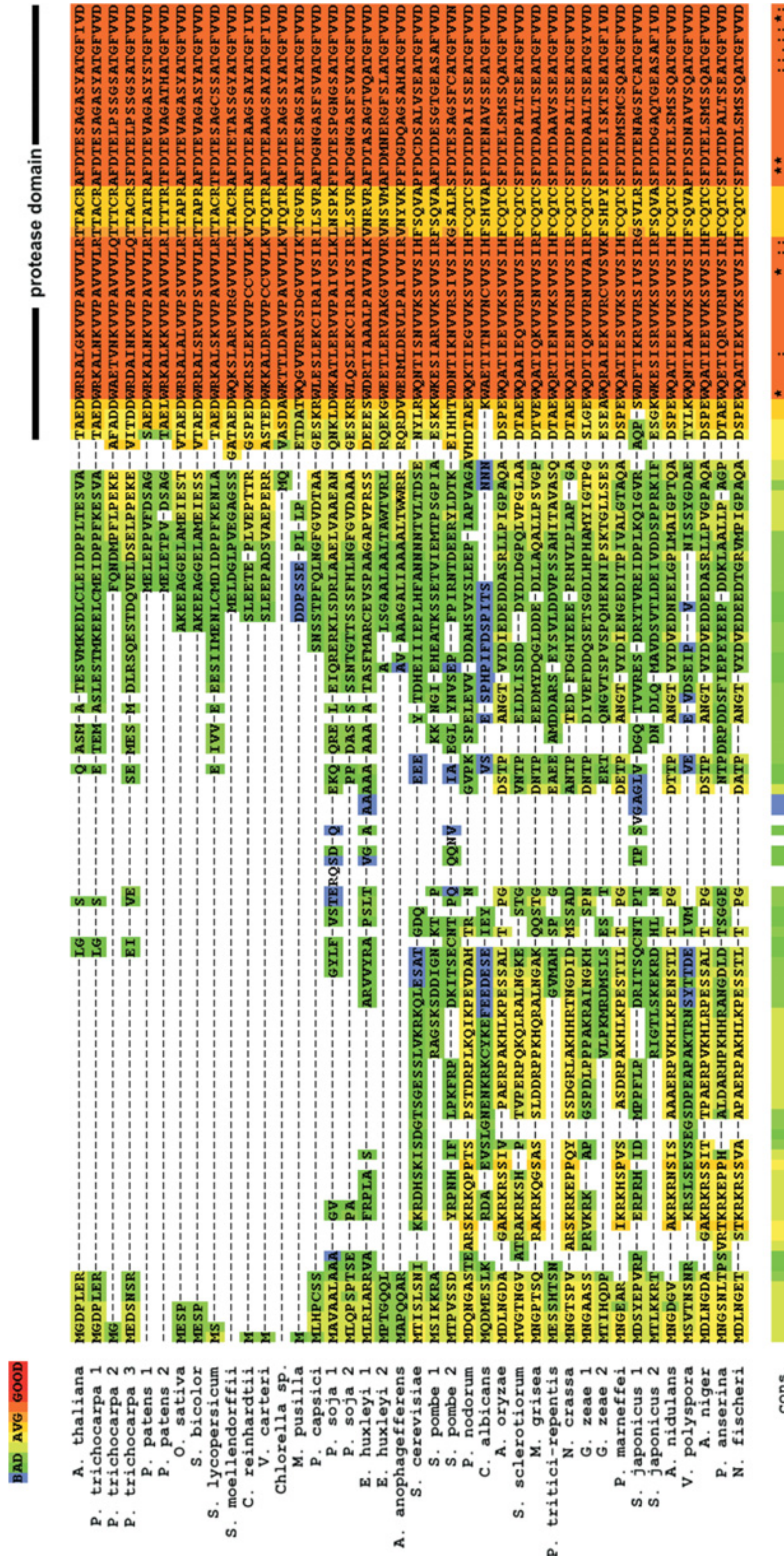

Figure S1 M-Coffee multiple sequence alignment of the amino acid sequences of the DEG7 orthologues used in the present study

The approximate position of the domains (as determined by using the InterProScan and HHpred prediction servers) are indicated by black bars on top of the alignment. Conserved residues of catalytic triad are shown as blue letters above the alignment. For full names of organisms, see Supplementary Table S1. cons., consensus; *, identity; :, conserved replacement; ., non-conserved replacement.

Continues...

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

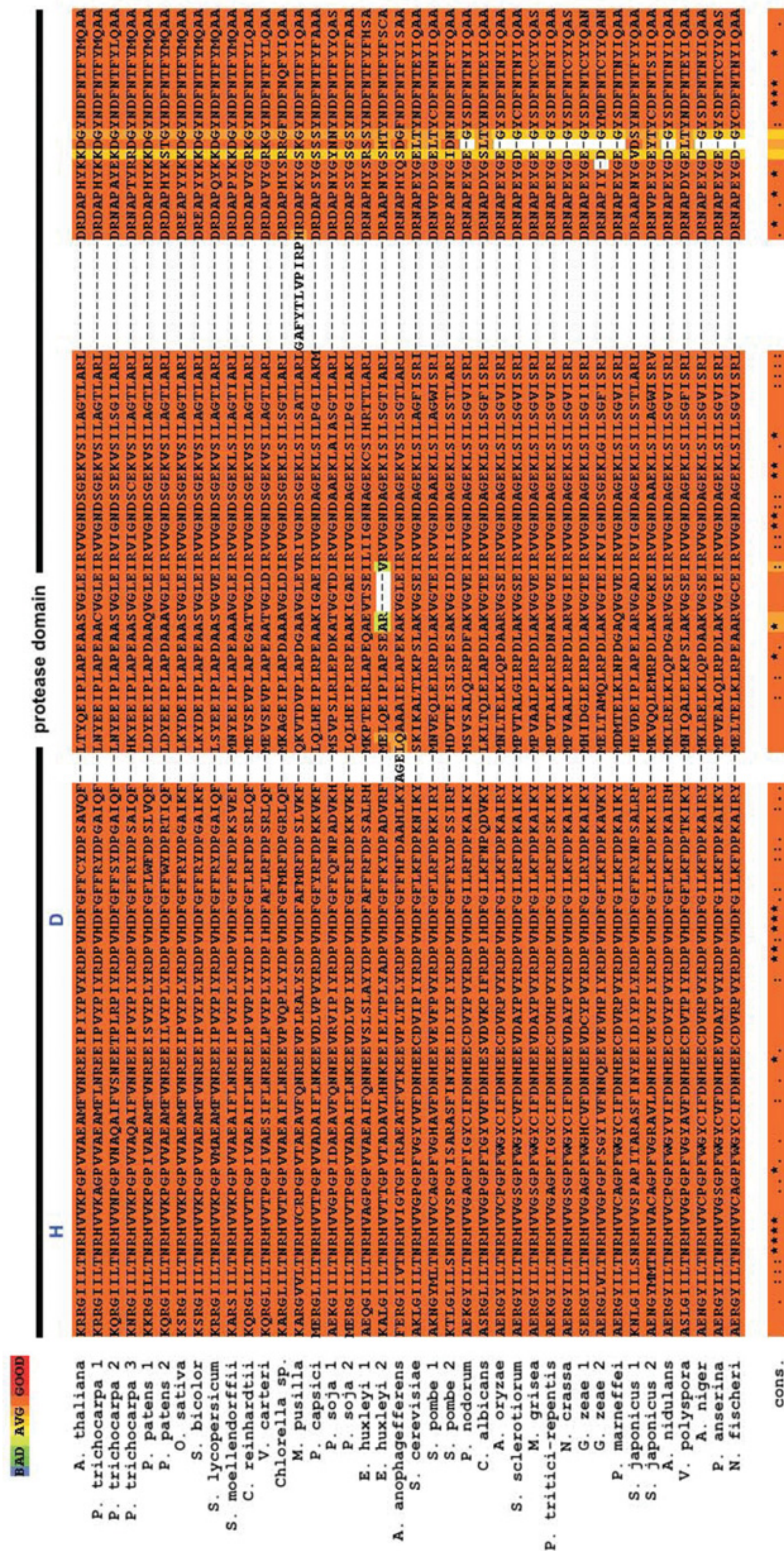

Figure S1 Continued

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

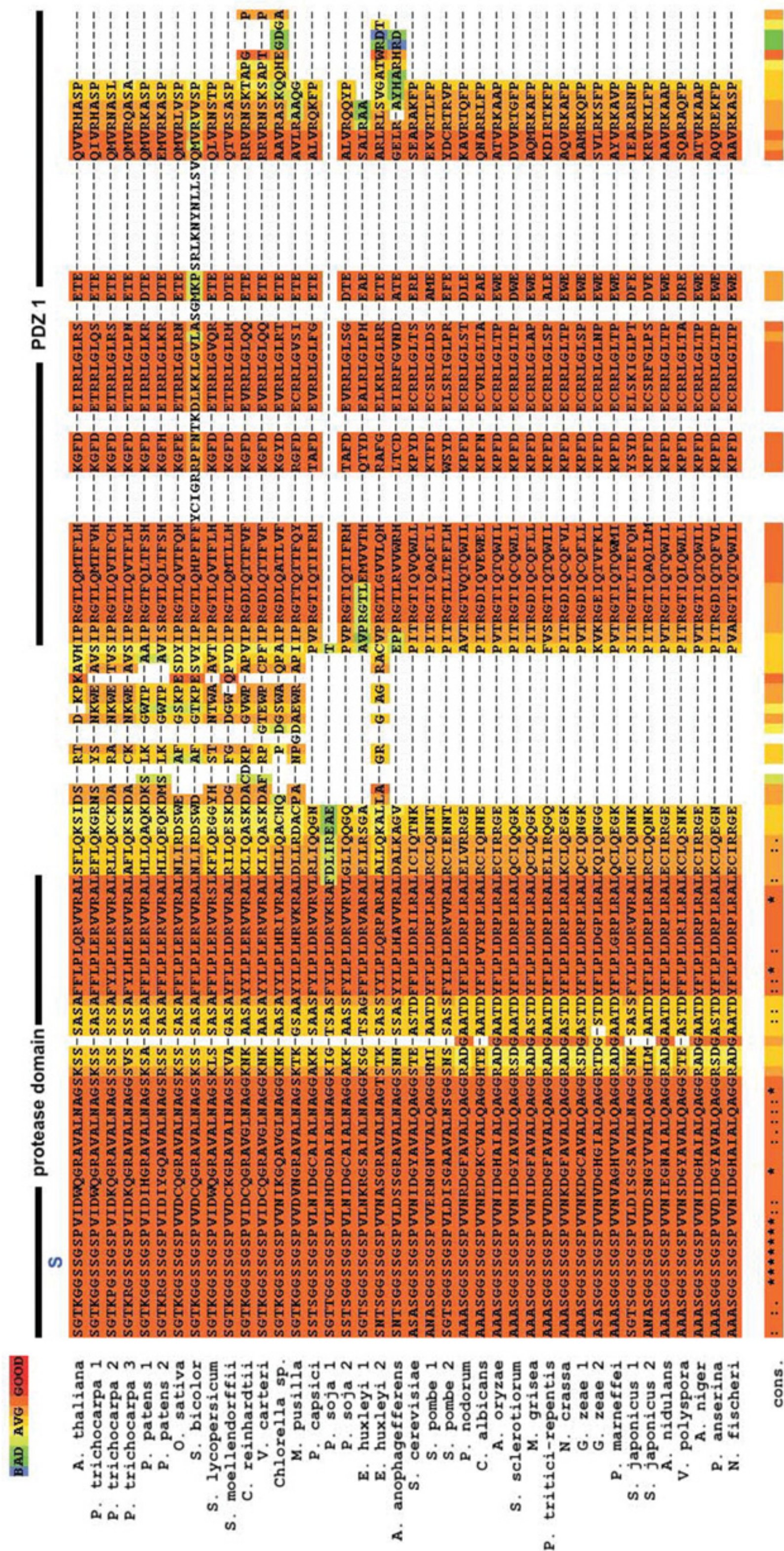

Figure S1 Continued

© 2011 The Author(s)

The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

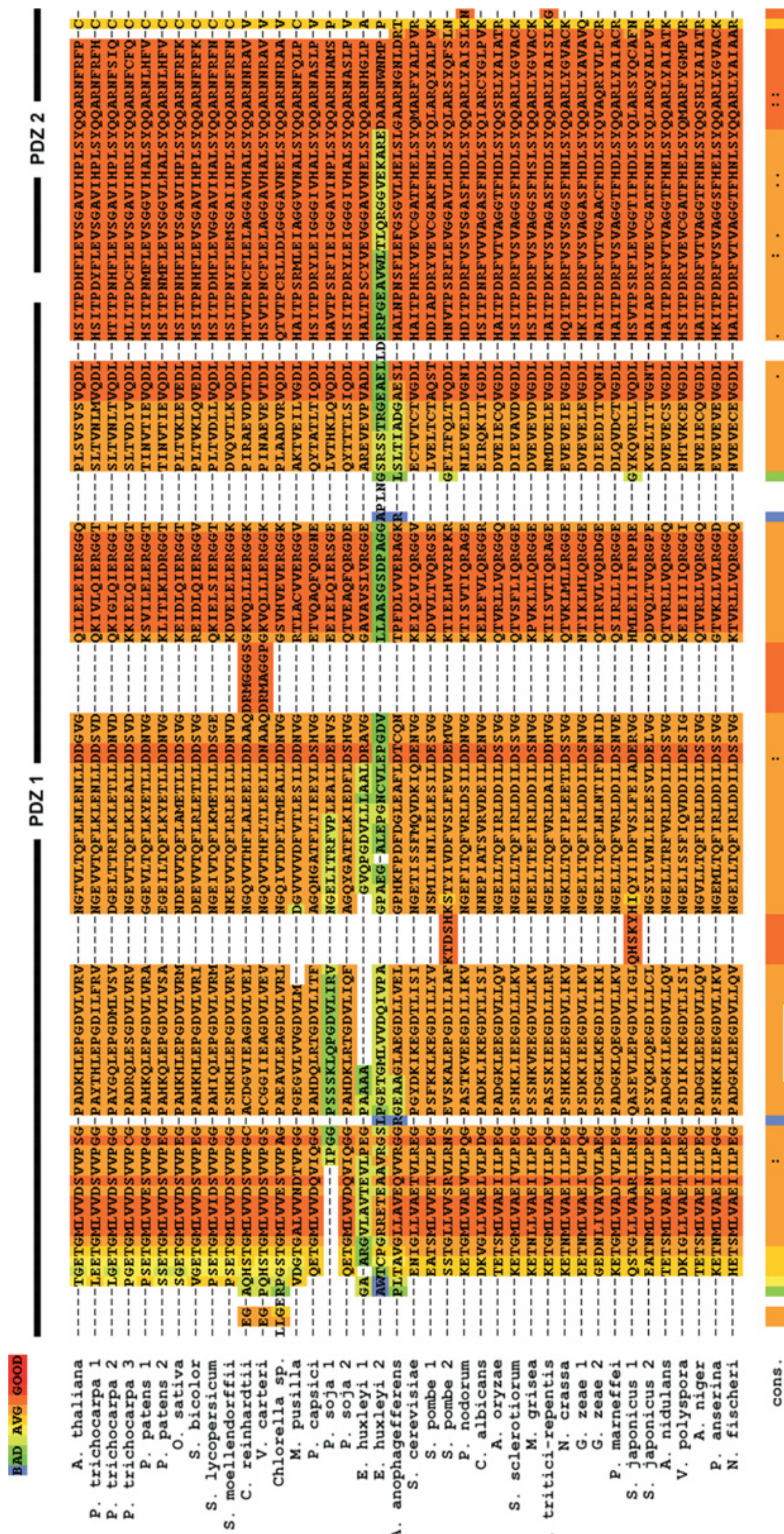

Figure S1 Continued

© 2011 The Author(s)

The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

Figure S1 Continued

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

Figure S1 Continued

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

Figure S1 Continued

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

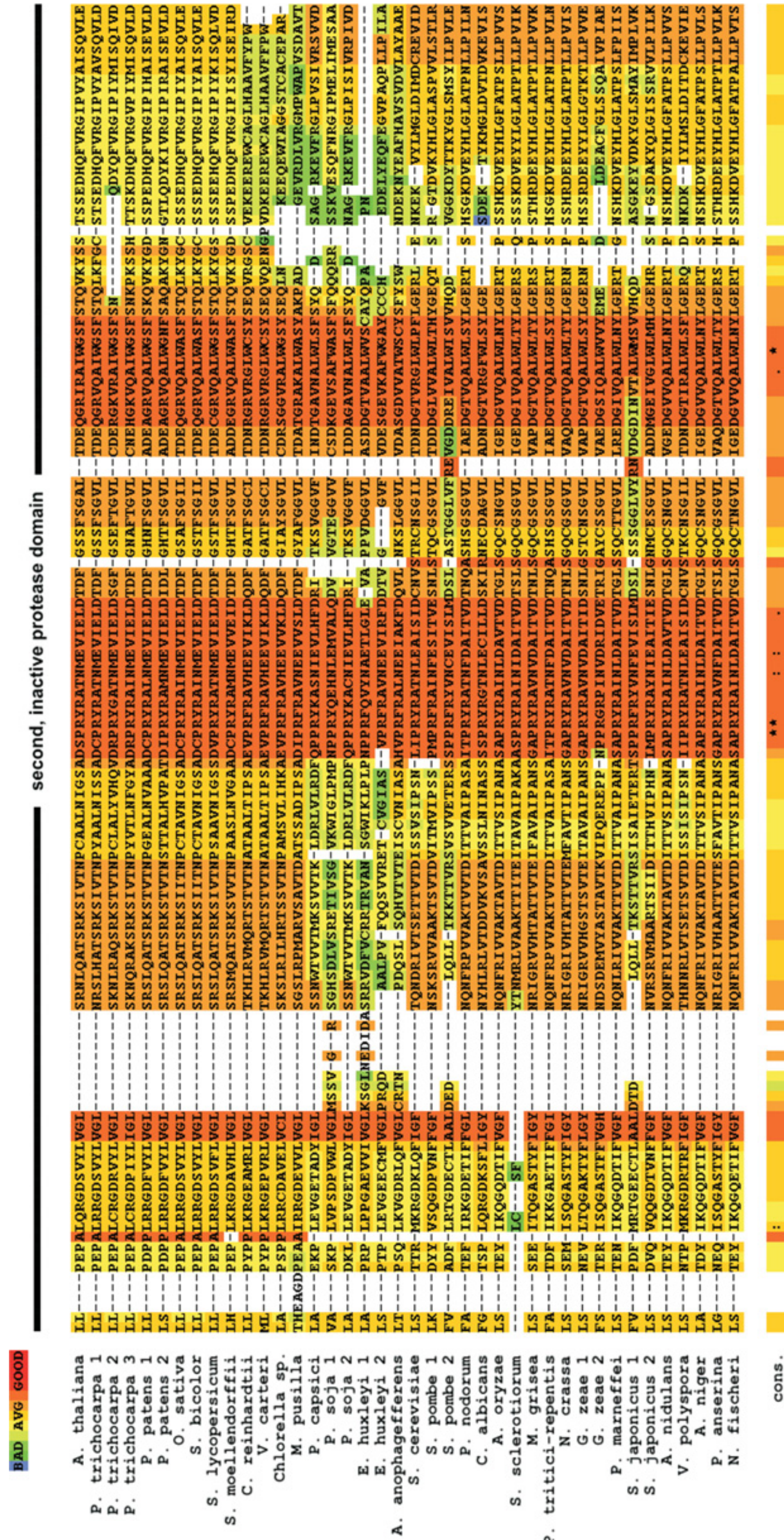

Figure S1 Continued

© 2011 The Author(s)

The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

Figure S1 Continued

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

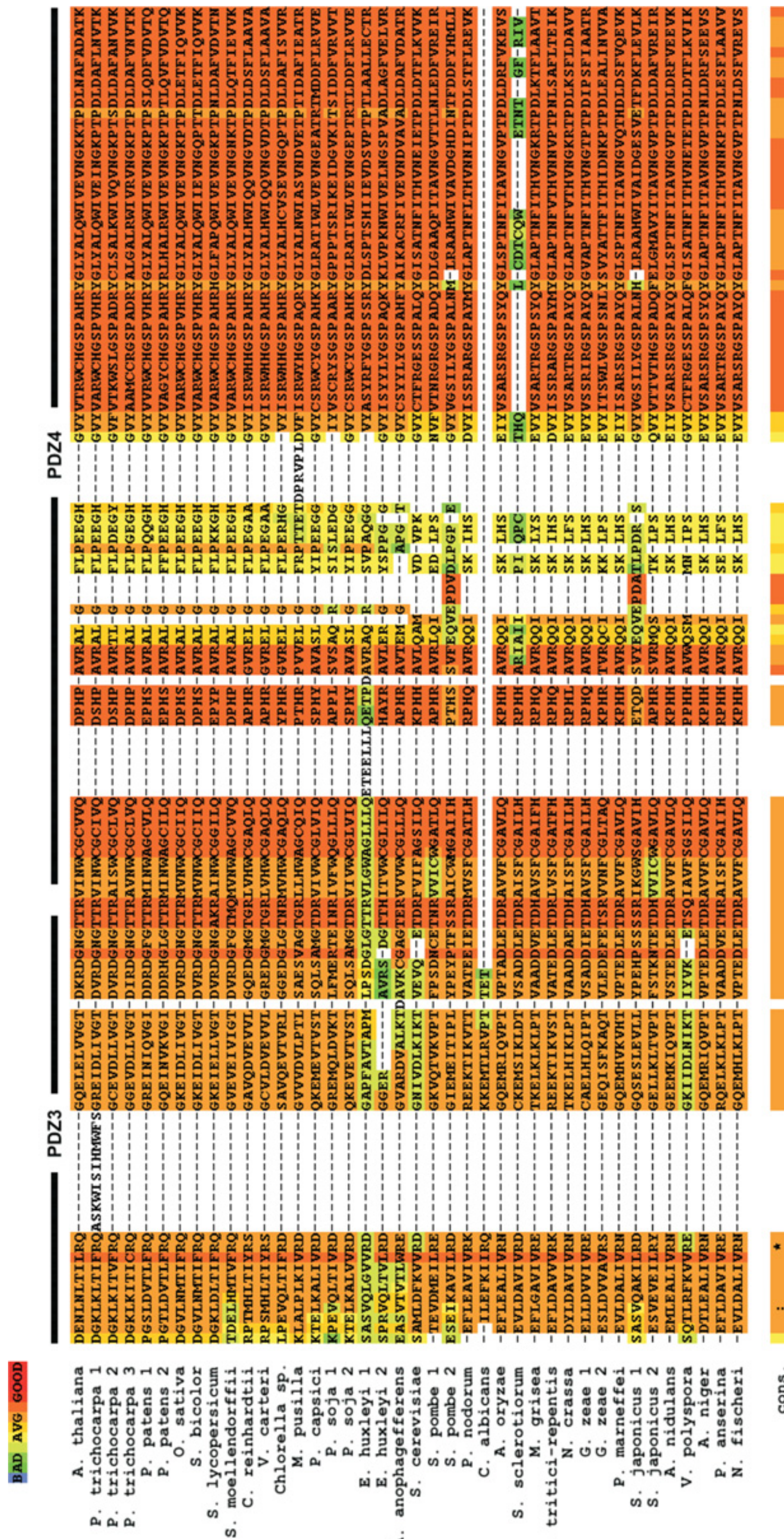

Figure S1 Continued

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

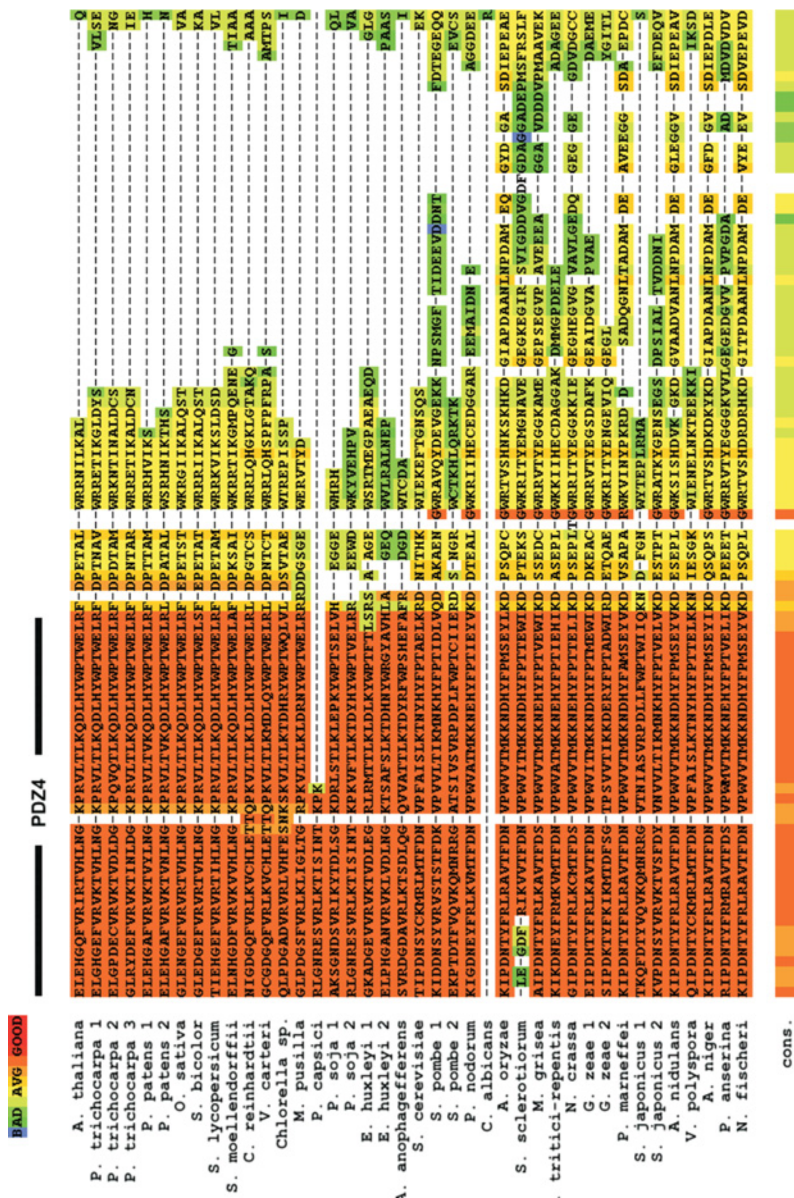

Figure S1 Continued

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

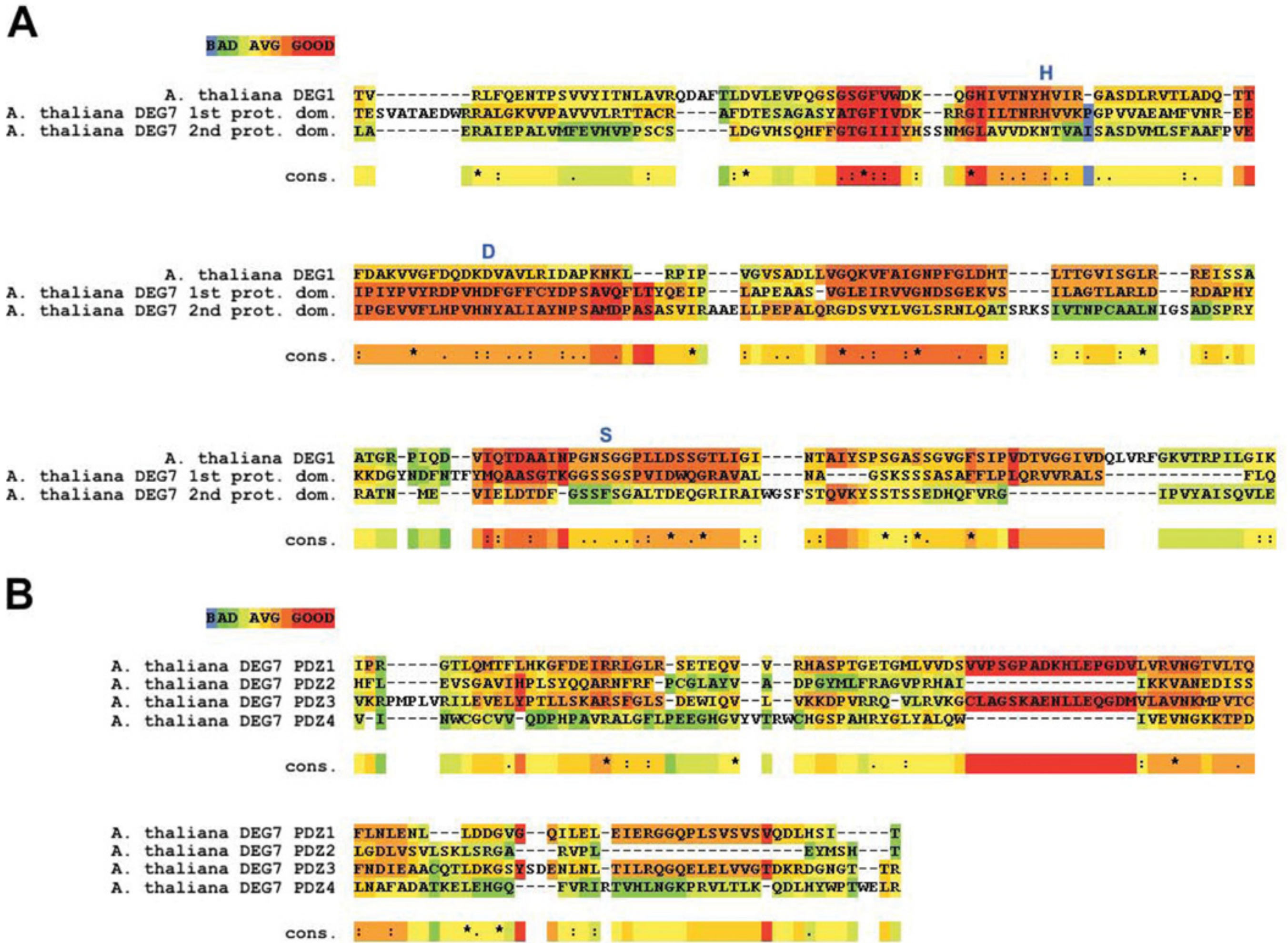

Figure S2 M-Coffee multiple sequence alignment of the amino acid sequences of protease domains and PDZ domains of *A. thaliana* DEG7

(**A**) Alignment of the first (active) protease domain (*A. thaliana* DEG7 1st prot. dom.) and the second (degenerate) protease domain (*A. thaliana* 2nd prot. dom.). The sequence of the protease domain of *A. thaliana* DEG1 is included to facilitate identification of catalytic side chains (shown in blue above the alignment). (**B**) Alignment of the PDZ domains. PDZ1 and PDZ3 both contain a stretch of additional amino acids not present in PDZ2 and PDZ4, indicating a whole-gene-duplication event as presented in Figure 1 of the main text. cons., consensus: *, identity; :, conservative replacement; ., non-conservative replacement.

© 2011 The Author(s)

The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

```
A. thaliana            AEDWR-RALGKVVPAVVVLRTTACRAFDTESAGASYATGFIVDKRRGIILTN
G. max                 AEDWR-KALNRVVPAVVVLRTTATRSFDTESAAASYATGFIVDKRRGIILTN
P. trichocarpa 1       AEDWR-KALNKVVPAVVVLRTTACRAFDTESAGASYATGFVVDKRRGIILTN
P. trichocarpa 2       ADDWA-ETVNKVVPAVVVLQTTTCRAFDTELPSSGSATGFVVDKQRGIILTN
P. trichocarpa 3       TDDWR-DAINKVVPAVVVLQTTACRSFDTELPSSGSATGFVVDKNRGIILTN
S. lycopersicum        AEDWR-KALSKVVPAVVVLRTTACRTFDTESAGCSSATGFVVDKRRGIILTN
P. patens 1            AEDWR-KALNKVVPAVVVLRTTATRAFDTEVAGASYSTGFVVDKKRGILLTN
P. patens 2            AELWR-KALKKVVPAVVVLRITTTRTFDTEVAGATHATGFVVDKQRGIILTN
O. sativa              AEDWR-RALALVVPSVVVLRTTAPRAFDTEVAGASYATGFVVDKSRGIILTN
S. bicolor             AEDWR-RALSRVVPSVVVLRTTAPRAFDTEVAGASYATGFVVDKSRGIILTN
S. moellendorffii      AEDWQ-KSLARVVRGVVVLRTTACRAFDTETASSGYATGFVVDKARSIILTN
C. reinhardtii         PEDWK-RSLEKVVPCCVVLKVTQTRAFDTEAAGSAYATGFIVDKQRGLILTN
V. carteri             TEDWK-KALDRVVPCCVVLKVTQTRAFDTEAAGSAYATGFIVDKQRGLILTN
M. pusilla             DATWQ-GVVRRVSDGVVVIKTTGVRAFDTESAGSAYATGFVVDKARGVVLTN
Chlorella sp. NC64A    SDAWK-TTLDAVVPAVVVLKVTQTRAFDTESAGSSYATGFVVDKARGLILTN
P. capsici             SKRWL-ESLEKCIRAIVSIRILSVRAFDGNGASFSVATGFVVDMERGLILTN
P. soja 1              KLDWK-ATLERVVPAIVSLKLNSPKFFDTESPGNGSATGFVVDAEKGIILTN
P. soja 2              SKRWL-QSLEKCIRAIVSIRLLSVRAFDGNGASFSVATGFVVDMERGIILTN
E. huxleyi 1           EESWD-RTIAAALPAVVAIKVNRVRAFDTASAGTVQATGFVVDAEQGIILTN
E. huxleyi 2           EKGWE-ETLERVAKGVVVVRVNSVMAFDMNERGFSLATGFVVDKALGIILTN
A. anophagefferens     RDVWE-RMLDRVLPAIVVIRVNYVKPFDGDQAGSAHATGFVVDFERGIIVTN
S. cerevisiae          YLRWQ-NTISNVVKSVVSIHFSQVAPFDCDSALVSEATGFVVDAKLGIILTN
S. pombe 1             SKKWK-ESIARVVKSVVSIRFSQVAAFDTDESGTGEASAFVVDAKNGYMLSN
S. pombe 2             HHTWD-NTIKNVVRSIVSIKGSALRSFDTESAGSFCATGFVVNKTLGLILSN
S. japonicus 1         QPSWD-FTIKRVVRSIVSIRGSVLRSFDTENAGSFCATGFVVDKNLGIILTN
S. japonicus 2         SGKWK-ESISRVVKSVVSIRFSQVASFDTDGAQTGEASAFIVDAENGYMMTN
P. nodorum             TAEWQ-KTIEGVVKSVVSIHFCQTCSFDTDPAISSEATGFVVDAEKGYILTN
C. albicans            NNKWA-ETITNVVNCVVSIHFSHVAPFDTENAVSSEATGFVVDASRGLILTN
A. oryzae              SPEWQ-ATIEEVVKSVVSIHFCQTCSFDTELSMSSQATGFVVDAERGYILTN
A. nidulans            SPEWQ-ATIEEVVKSVVSIHFCQTCSFDTELSMSSQATGFVVDAERGYILTN
A. niger               SPEWQ-ATIEEVVKSVVSIHFCQTCSFDTELSMSSQATGFVVDAENGYILTN
S. sclerotiorum        TAEWQ-AAIEQVVRNVVSIRFCQTCSFDTDPALTSEATGFVVDAERGYILTN
M. grisea              TVEWQ-ATIQKVVSNVVSIRFCQTCSFDTDAALTSEATGFVVDAERGYILTN
P. tritici-repentis    TAEWQ-RTIENVVKSVVSIHFCQTCSFDTDAAVSSEATGFVVDAEKGYILTN
N. crassa              TAEWQ-ATIENVVRNVVSIRFCQTCSFDTDPALTSEATGFVVDAERGYILTN
G. zeae 1              LGEWQ-DTIQKVVRNVVAIRFCQTCSFDTDAALTSEATGYVVDSERGYILTN
G. zeae 2              SEAWQ-RAIEKVVRCVVSVKFSHPYSFDTEISKTSEATGFIVDAERGLVLTN
P. anserina            TAEWQ-ETIQRVVRNVVSIRFCQTCSFDTDPALTSEATGFVVDAERGYILTN
N. fischeri            SPEWQ-ATIEKVVKSVVSIHFCQTCSFDTDLSMSSQATGFVVDAERGYILTN
P. marneffei           SPEWQ-ATIESVVKSVVSIHFCQTCSFDTDMSMCSQATGFVVDAERGYILTN
V. polyspora           YLKWQ-NTIAKVVKSVVSIHFSQVAPFDSDNAVVSQATGFVVDASLGIILTN
EcDegP                 DSPFCQEG-SPFQSSPFCQGGQGGNGGGQQQKFMALGSGVIIDADKGYVVTN
HsHtrA2                QYNFIADVVEKTAPAVVYIEILDRHPFLGREVPISNGSGFVVAAD-GLIVTN
AtDEG1                 TVRLFQENTPSVVYITNLAVRQDAFTLDVLEVPQGSGSGFVWDKQ-GHIVTN
                                                           .:. :     . :::*
```

Figure S3 Alignment of the amino acid sequences of the active protease domains of DEG7 proteins from higher plants, mosses, algae and fungi

Amino acid sequences of DEG7 active protease domains were aligned with those of the protease domains of DegP from *Escherichia coli* (EcDegP), human HtrA2 (HsHtrA2) and Deg1 from *Arabidopsis thaliana* (AtDEG1) as described in the Experimental section of the main text. For full names of organisms, see Supplementary Table S1. Identical residues are indicated by asterisks. Active-site residues of the catalytic triad are in bold. Conservative replacements are indicated by : symbols; non-conservative replacements are indicated by . symbols.

Continues...

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

```
A. thaliana            RHVVKPGPVVAEAMFVNREEIPIYPVYRDPVHDFGFFCYDPSAVQF---LTY
G. max                 RHVVKPGPVVAEAMFLNREEVPVHPIYRDPVHDFGFFRYDPGAIQF---LNY
P. trichocarpa 1       RHVVKAGPVVAEAMFLNREEIPVYPIYRDPVHDFGFFRYDPGAIQF---LNY
P. trichocarpa 2       RHVVNPGPVNAQAIFVSNEETPLRPIYRDPVHDFGFFSYDPGAIQF---LNY
P. trichocarpa 3       RHVVKPGPVVAQAIFVNNEEIPVYPIYRDPVHDFGFFRYDPSAIQF---HKY
S. lycopersicum        RHVVKPGPVMAEAMFVNREEIPVYPIYRDPVHDFGFFRYDPGAIQF---LSY
P. patens 1            RHVVKPGPIVAEAMFVNREEISVYPLYRDPVHDFGFLWFDPSLVQF---LDY
P. patens 2            RHVVKPGPVVAEAMFVNREEILVYPLYRDPVHDFGFFWYDPRTIQF---LDY
O. sativa              RHVVKPGPVVAEAMFVNREEIPVYPLYRDPVHDFGFFRYDPGAIKF---LKY
S. bicolor             RHVVKPGPVVAEAMFVNREEIPVYPLYRDPVHDFGFFRYDPGAIKF---LKY
S. moellendorffii      RHVVKPGPVVAEAIFLNREEIPVYPLYRDPVHDFGFFRFDPKSVEF---MNY
C. reinhardtii         RHVVTPGPIVAEAIFLNREELPVVPLYYDPIHDFGFLRFDPSRLQF---MEV
V. carteri             RHVVTPGPIVAESIFLNREELPVYPLYYDPIHDFAFLRFDPSRLQF---MEV
M. pusilla             RHVCRPGPVTAEAVFQNREEVPLRALYSDPVHDFAFMRFDPSLVKF---QKV
Chlorella sp NC64A     RHVVTPGPVVAEAIFLNREEVPVQPLYYDPVHDFGFMRFDPGRLQF---MKA
P. capsici             RHVVTPGPVVADAIFLNKEEVDLVPVYRDPVHDFGFYRFDPKKVKF---LQL
P. soja 1              RHVVGPGPIDAEAVFQNNEEVRVIPIYRDPVHDFGFFQFNPADVKH---MSV
P. soja 2              RHVVTPGPVVADAIFLNKEEVDLVPIYRDPVHDFGFFRFDPKKVKF---LQL
E. huxleyi 1           RHVAGPGPVVAEAIFQNNEEVSLSLAYYDPVHDFAFFRFDPSALRH---MKP
E. huxleyi 2           RHVVTTGPVTADAVLHNKEEIELTPLYADPVHDFGFFKYDPADVRF---MEL
A. anophagefferens     RHVIGTGPIRAEATFVTKEEVPLTPLYRDPVHDFGFFHFDAAHLKYAGELQA
S. cerevisiae          RHVVGPGPFVGYVVFDNHEECDVIPIYRDPVHDFGFLKFDPKNIKY---SKI
S. pombe 1             RHVVCAGPFVGHAVFDNHEEVEVFPVYRDPVHDFGFLRFDPKKIRY---MNV
S. pombe 2             RHVVSPGPISARASFINYEEIDIYPIYRDPVHDFGFFRYDPSSIRF---HDV
S. japonicus 1         RHVVSPAPITARASFINYEEIDIYPLYRDPVHDFGFFRYNPSALRF---HEV
S. japonicus 2         RHVACAGPFVGRAVLDNHEEVEVYPIYRDPVHDFGILKFDPKKIRY---MKV
P. nodorum             RHVVGAGPFIGYCIFDNHEECDVYPVYRDPVHDFGILRFDPKAIKY---MSV
C. albicans            RHVVGPGPFTGYVVFDNHESVDVKPIFRDPIHDFGILKFNPQDVKY---LKL
A. oryzae              RHVVCPGPFWGYCIFDNHEECDVRPVYRDPVHDFGILKFDPKAIRY---MNL
A. nidulans            RHVVCPGPFWGYVIFDNHEECDVYPVYRDPVHDFGFLKFDPKAIRH---MKL
A. niger               RHVVCPGPFWGYCIFDNHEECDVRPVYRDPVHDFGILKFDPKAIRY---MKL
S. sclerotiorum        RHVVGSGPFWGYCVFDNHEEVDAYPVYRDPVHDFGILRFDPKAIKY---MPV
M. grisea              RHVVGSGPFWGYCIFDNHEEVDAYPVYRDPVHDFGILKFDPKAIKY---MPV
P. tritici-repentis    RHVVGAGPFIGYCIFDNHEECDVHPVYRDPVHDFGILRFDPSKIKY---MPV
N. crassa              RHVVGSGPFWGYCIFDNHEEVDAYPVYRDPVHDFGILKFDPKAIKY---MPV
G. zeae 1              RHVVGAGPFWGHCVFDNHEEVDCYPVYRDPVHDFGILRYDPKAIKY---MHI
G. zeae 2              RHVVGPGPFSGYIVFNNQEEVEVHPIYRDPVHDFGFLKFDPKAVKY---MEL
P. anserina            RHVVGSGPFWGYCVFDNHEEVDAYPVYRDPVHDFGILKFDPKAIKY---MPV
N. fischeri            RHVVCAGPFWGYCIFDNHEECDVRPVYRDPVHDFGILKFDPKAIRY---MEL
P. marneffei           RHVVCAGPFWGYCIFDNHEECDVRPVYRDPVHDFGILKFDPKAIKY---MDM
V. polyspora           RHVVGPGPFVGYAVFDNHEECDVTPIYRDPVHDFGFLKFDPTKIKY---MNI
EcDegP                 NHVV-DNATVIKVQLSDGRKFDAKMVGKDPRSDIALIQIQ--NPKN-----L
HsHtrA2                AHVV-ADRRRVRVRLLSGDTYEAVVTAVDPVADIATLRIQ--TKEP-----L
AtDEG1                 YHVI-RGASDLRVTLADQTTFDAKVVGFDQDKDVAVLRID--APKN----KL
                        **           :             *   *..    :    .
```

Figure S3 Continued

© 2011 The Author(s)

The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

```
A. thaliana         QEIPLAPEAA-SVGLEIRVVGNDSGEKVSILAGTLARL-----------DRD
G. max              EEIPLAPEAA-CVGLEIRVVGNDSGEKVSILAGTLARL-----------DRD
P. trichocarpa 1    EEIPLAPEAA-CVGLEIRVVGNDSGEKVSILAGTLARL-----------DRD
P. trichocarpa 2    EEIPLAPEAA-SVGLEIRVIGNDSSEKVSILSGILARL-----------DRN
P. trichocarpa 3    EEIPLAPEAA-SVGLEIRVIGNDSCEKVSILAGTLARL-----------DRN
S. lycopersicum     EEIPLAPDAA-SVGVEIRVVGNDSGEKVSILAGTLARL-----------DRD
P. patens 1         EEIPLAPDAA-QVGLEIRVVGNDSGEKVSILAGTLARL-----------DRD
P. patens 2         EEIPLAPDAA-AVGLEIRVVGNDSGEKVSILAGTLARL-----------DRD
O. sativa           DEIPLAPEAA-SVGLEIRVVGNDSGEKVSILAGTLARL-----------DRE
S. bicolor          DEIPLAPEAA-SVGLEIRVVGNDSGEKVSILAGTLARL-----------DRE
S. moellendorffii   EEIPLAPEAA-AVGLEIRVVGNDSGEKLSILAGTIARL-----------DRD
C. reinhardtii      SEVPLAPEGA-TVGLDIRVVGNDSGEKVSILAGTLARL-----------DRD
V. carteri          SEVPLAPEAA-TVGLDIRVVGNDSGEKVSILAGTLARL-----------DRD
M. pusilla          TDVPLAPDGA-AVGLEVRIVGNDSGEKLSILSATLARLGAFYTLVPIRPHRD
Chlorella sp. NC64A GEIPLAPDAA-AVGLDIRVVGNDSGEKLSILSGTLARL-----------DRD
P. capsici          HEIPLRPEAA-KIGAEIRVVGNDAGEKLSILPGILAKM-----------DRD
P. soja 1           PSLRLEPDKA-TVGTDIRVVGNDAAEKLAIASGTLARL-----------DRD
P. soja 2           HEIPLRPEAA-KIGAEIRVVGNDAGEKLSILPGILAKL-----------DRD
E. huxleyi 1        TELRLAPEQA-EVTSEILIIGNNAGEKCSIHRTTLARL-----------DRN
E. huxleyi 2        QEIPLAPSEA-RV----RVVGNDAGEKISILSGTIARL-----------DRA
A. anophagefferens  AAIELAPEKA-KVGLEIRVVGNDAGEKVSILSGTLARL-----------DRN
S. cerevisiae       KALTLKPSLA-KVGSEIRVVGNDAGEKLSILAGFISRI-----------DRN
S. pombe 1          EQLELRPDLA-KVGTEIRVVGNDAAEKLSILAGWISRI-----------DRN
S. pombe 2          TEISLSPESA-KVGIDIRIIGNDAGEKLSILSSTLARL-----------DRP
S. japonicus 1      DEIPLAPELA-RVGADIRVIGNDAGEKLSILSSTLARL-----------DRA
S. japonicus 2      QQLEMRPDLA-KVGKEIRVVGNDAAEKLSILAGWISRV-----------DRN
P. nodorum          SALQLRPDFA-KVGVEIRVVGNDAGEKLSILSGVISRL-----------DRN
C. albicans         TQLELAPDLA-KVGTEIRVVGNDAGEKLSILSGFISRL-----------DRN
A. oryzae           TELKLQPDAA-RVGSEIRVVGNDAGEKLSILSGVISRL-----------DRN
A. nidulans         RELKLQPDGA-RVGSEIRVVGNDAGEKLSILSGVISRL-----------DRN
A. niger            RELKLQPDAA-KVGSEIRVVGNDAGEKLSILSGVISRL-----------DRN
S. sclerotiorum     TALGLRPDLA-KVGSEIRVVGNDAGEKLSILSGVISRL-----------DRN
M. grisea           AALPLRPDLA-KVGVEIRVVGNDAGEKLSILSGVISRL-----------DRN
P. tritici-repentis TALKLRPDNA-KVGVEIRVVGNDAGEKLSILSGVISRL-----------DRN
N. crassa           AALPLRPDLA-RVGIEIRVVGNDAGEKLSILSGVISRL-----------DRN
G. zeae 1          DGLELRPDLA-KVGTEIRVVGNDAGEKLSILSGIISRL-----------DRN
G. zeae 2          TAMQLRPDLA-KVGTEIKVIGNDSGEKLGILSGFISRL-----------DRN
P. anserina         EALQLRPDLA-KVGIEIRVVGNDAGEKLSILSGVISRL-----------DRN
N. fischeri         TELKLRPEAA-RVGCEIRVVGNDAGEKLSILSGVISRL-----------DRN
P. marneffei        TELKLNPDGA-QVGVEIRVVGNDAGEKLSILSGVISRL-----------DRN
V. polyspora        QALELKPSLA-KVGSEIRVVGNDAGEKLSILSGFISRL-----------DRN
EcDegP              TAIKMADSDALRVGDYTVAIGNPFGLGETVTSGIVSAL-----------GRS
HsHtrA2             PTLPLGRSADVRQGEFVVAMGSPFALQNTITSGIVSSA-----------QRP
AtDEG1              RPIPVGVSADLLVGQKVFAIGNPFGLDHTLTTGVISGL-----------RRE
                      : :  .           :*.       :    ::              *
```

Figure S3 Continued

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

```
A. thaliana          APHYKKDGYNDFNTFYMQAASGTKGGSSGSPVIDWQGRAVALNAGSKSS-SA
G. max               APHYKKDGYNDFNTFYMQAASGTKGGSSGSPVIDWQGRAVALNAGSKSS-SA
P. trichocarpa 1     APHYKKDGYNDFNTFYMQAASGTKGGSSGSPVIDWQGRAVALNAGSKSS-SA
P. trichocarpa 2     APAYEKDGYNDFNTFYLQAASGTKPGSSGSPVIDKQGRAVALNAGSSSS-SS
P. trichocarpa 3     APTYRRDGYNDFNTFYMQAASGTKRGSSGSPVIDKQGRAVALNAGGSVS-SS
S. lycopersicum      APQYKKDGYNDFNTFYMQAASGTKGGSSGSPVIDWQGRAVALNAGSKLS-SA
P. patens 1          APHYKKDGYNDFNTFYMQAASGTKGGSSGSPVIDIHGRAVALNAGSKSA-SA
P. patens 2          APHYKSTGYNDFNTFYMQAASGTKRGSSGSPVIDIYGQAVALNAGSRSS-SA
O. sativa            APYYKKDGYNDFNTFYMQAASGTKGGSSGSPVVDCQGRAVALNAGSKSS-SA
S. bicolor           APYYKKDGYNDFNTFYMQAASGTKGGSSGSPVVDCQGRAVALNAGSKSS-SA
S. moellendorffii    APPYKKDGYNDFNTFYMQAASGTKGGSSGSPVVDCKGRAVAINAGSKVA-GA
C. reinhardtii       APVYGRKGYNDFNTFYLQAASGTKGGSSGSPVIDCQGRAVGLNAGGKNK-AA
V. carteri           APVYGRKGYNDFNTFYLQAASGTKGGSSGSPVIDCQGRAVGLNAGGKNK-AA
M. pusilla           APKYGSKGYNDFNTFYIQAASGTKGGSSGSPVVDVNGRAVALNAGSKTK-GS
Chlorella sp. NC64A  APHYSRRGFNDFNQFYIQAASGTKGGSSGSPVVNIKGQAVGLNAGGKNK-AA
P. capsici           APSYGSSSYNDFNTFYFAAASSTSGGSSGSPVLNIDGCAIALNAGGAKK-SA
P. soja 1            APNYGYNNYNDFNTFYYQASSGTTGGSSGSPVLNHDGDAIALNAGGKIG-TS
P. soja 2            APSYGSGSYNDFNTFYFAAASSTSGGSSGSPVLNIDGCAIALNAGGAKK-AA
E. huxleyi 1         APHYSRSSYNDFNTFYFHSASGTSGGSSGSPVLNKRGSAIALNAGGKSG-TS
E. huxleyi 2         APNYGSHTYNDFNTFYFSCASNTSGGSSGSPVVNASGRAVALNAGTSTK-SA
A. anophagefferens   APHYQSDGFNDFNTFYISAASNTSGGSSGSPVLDSSGRAVALNAGGSNN-SS
S. cerevisiae        APEYGELTYNDFNTEYIQAAASASGGSSGSPVVNIDGYAVALQAGGSTE-AS
S. pombe 1           VPDYGELTYCDFNTNYIQAAANASGGSSGSPVVERNGNVVALQAGGHMI-AA
S. pombe 2           APNYGIDNYNDFNTFYYQAASGTSGGSSGSPVLDISGAAVALNSGGSNS-SA
S. japonicus 1       APNYGVDSYNDFNTFYYQAASGTSGGSSGSPVLDISGSAVALNAGGSNK-SA
S. japonicus 2       VPEYGEYTYCDFNTSYIQAAANASGGSSGSPVVDSNGYVVALQAGGHLM-AA
P. nodorum           APEYGE-GYSDFNTNYIQAAAAASGGSSGSPVVNRDGFAVALQAGGRADGAA
C. albicans          APDYGSLTYNDFNTEYIQAAAAASGGSSGSPVVNEDGKCVALQAGGHTE-AA
A. oryzae            APEYGE-GYSDFNTNYIQAAAAASGGSSGSPVVNIDGHAIALQAGGRADGAA
A. nidulans          APEYGD-GYSDFNTNYIQAAAAASGGSSGSPVVNIEGNAIALQAGGRADGAA
A. niger             APEYGD-GYSDFNTNYIQAAAAASGGSSGSPVVNIDGHAIALQAGGRADGAA
S. sclerotiorum      APEYGD-GYCDFNTNYIQAAAAASGGSSGSPVVNIDGYAVALQAGGRSDGAA
M. grisea            APEYGE-GYSDFNTCYYQASAAASGGSSGSPVVNIDGFAVALQAGGRADGAS
P. tritici-repentis  APEYGE-GYSDFNTNYIQAAAAASGGSSGSPVVDRDGFAVALQAGGRADGAA
N. crassa            APEYGD-GYSDFNTCYYQASAAASGGSSGSPVVNKDGFAVALQAGGRADGAS
G. zeae 1            APEYGE-GYSDFNTCYYQANAAASGGSSGSPVVNKDGCAVALQAGGRSDGAS
G. zeae 2            APIY-D-GYMDFNTCYYQANASASGGSSGSPVVNVDGHGIALQAGGRTD-GS
P. anserina          APEYGE-GYSDFNTCYYQASAAASGGSSGSPVVDIDGYAVALQAGGRSDGAS
N. fischeri          APEYGD-GYCDFNTNYIQAAAAASGGSSGSPVVNIDGHAIALQAGGRADGAA
P. marneffei         APEYGE-GYSDFNTNYIQAAAAASGGSSGSPVVNVAGHVVALQAGGRADGAA
V. polyspora         APDYGELTYNDFNTEYIQAAAAASGGSSGSPVVNSDGYAVALQAGGSTE-AS
EcDegP               GLN-AEN-Y----ENFIQTDAAINRGNSGGALVNLNGELIGINTAILAPDGG
HsHtrA2              ARDLGLPQT---NVEYIQTDAAIDFGNSGGPLVNLDGEVIGVNTMKVT----
AtDEG1               ISSAATGRP---IQDVIQTDAAINPGNSGGPLLDSSGTLIGINTAIYSPSGA
                                         :    *.**..:::  *  :.:::
```

Figure S3 Continued

© 2011 The Author(s)

The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

```
A. thaliana         S---AFFLPLQRVVRALSFLQKS----------------------I
G. max              S---AFFLPLERVVRALRFLQKGS---------------------E
P. trichocarpa 1    S---AFFLPLERVVRALEFLQK------------------------
P. trichocarpa 2    S---AFYLPLERVVRALRLLQKC----------------------K
P. trichocarpa 3    S---AFYLHLERVVRALAFLQKS----------------------K
S. lycopersicum     S---AFFLPLERVVRSLKFLQE------------------------
P. patens 1         S---AFFLPLERVVRALHLLQAQ----------------------K
P. patens 2         S---AFFLPLERVVRALHLLQEQ----------------------K
O. sativa           S---AFFLPLERVVRALNLIRD------------------------
S. bicolor          S---AFFLPLERVVRALNLIRDS----------------------W
S. moellendorffii   S---AYFLPLDRVVRALRILQES----------------------K
C. reinhardtii      S---AYYLPLERVVRALKLIQAS----------------------K
V. carteri          S---AYYLPLERVVRALKLIQASKDAFRPGTEWP-C----PF--IP
M. pusilla          A---AYYLPLHRVKRALDLLRDACP----------------A--NP
Chlorella sp. NC64A S---AYYLPLHRIVRALHLLQACHQ---PDGSWA-Q----PA--IP
P. capsici          A---SFYLPLDRVVRVLRLIQQGNP-------------------VP
P. soja 1           A---SFYLPLDRVKRAFDLIREAET-------------------IP
P. soja 2           S---SFYLPLDRVVRVLGLIQQGQP-------------------VP
E. huxleyi 1        A---GFFLPLDRVARALELLRSGAA-------------------VP
E. huxleyi 2        S---SYYLPLQRPARALALLQKALLAGRGAG----R----AC--VP
A. anophagefferens  A---SYYLPLHAVVRALDALKAG----------------------V
S. cerevisiae       T---DFFLPLDRILRALICIQTNKP-------------------IT
S. pombe 1          T---DYFLPLDRPLRALRCLQNNTP-------------------IT
S. pombe 2          S---SFYLPLDRVVRALRCIENNTP-------------------IT
S. japonicus 1      S---SFYLPLDRVVRALHCIQNNKS-------------------IT
S. japonicus 2      T---DYFLPLDRPLRALRCLQQNKP-------------------IT
P. nodorum          T---DYFLPLDRPLRALELVRRGEA-------------------VT
C. albicans         T---DYFLPVYRPLRALRCIQNNEP-------------------IT
A. oryzae           T---DYFLPLDRPLRALECIRRGEP-------------------VT
A. nidulans         T---DYFLPLDRPLRALECIRRGEP-------------------VT
A. niger            T---DYFLPLDRPLRALECIRRGEP-------------------VT
S. sclerotiorum     T---DYFLPLDRPLRALQCIQQGKP-------------------IT
M. grisea           T---DYFLPLDRPLRALQCLQQGKP-------------------IT
P. tritici-repentis T---DYFLPLDRPLRALELIRQGQF-------------------VS
N. crassa           T---DYFLPLDRPLRALKCLQEGKP-------------------IT
G. zeae 1           T---DYFLPLDRPLRALQCIQNGKP-------------------VT
G. zeae 2           T---DYFLPLDGPLRALKQIQNGGK-------------------VK
P. anserina         T---DYFLPLDRPLRALKCLQEGNP-------------------IT
N. fischeri         T---DYFLPLDRPLRALECIRRGEP-------------------VA
P. marneffei        T---DYFLPLGRPLRALQCLQEGKP-------------------VT
V. polyspora        T---DFFLPLDRILRALKCLQSNKP-------------------IT
EcDegP              NIGIGFAIPSNMVKNLTSQMVEYGQVK-----RG-E----LGIMGT
HsHtrA2             -AGISFAIPSDRLREFLHRGEKKNSSG-----ISGSQRRYIG--VM
AtDEG1              SSGVGFSIPVDTVGGIVDQLVRFGKVT-----RP-I----LG--IK
                         : :
```

Figure S3 Continued

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

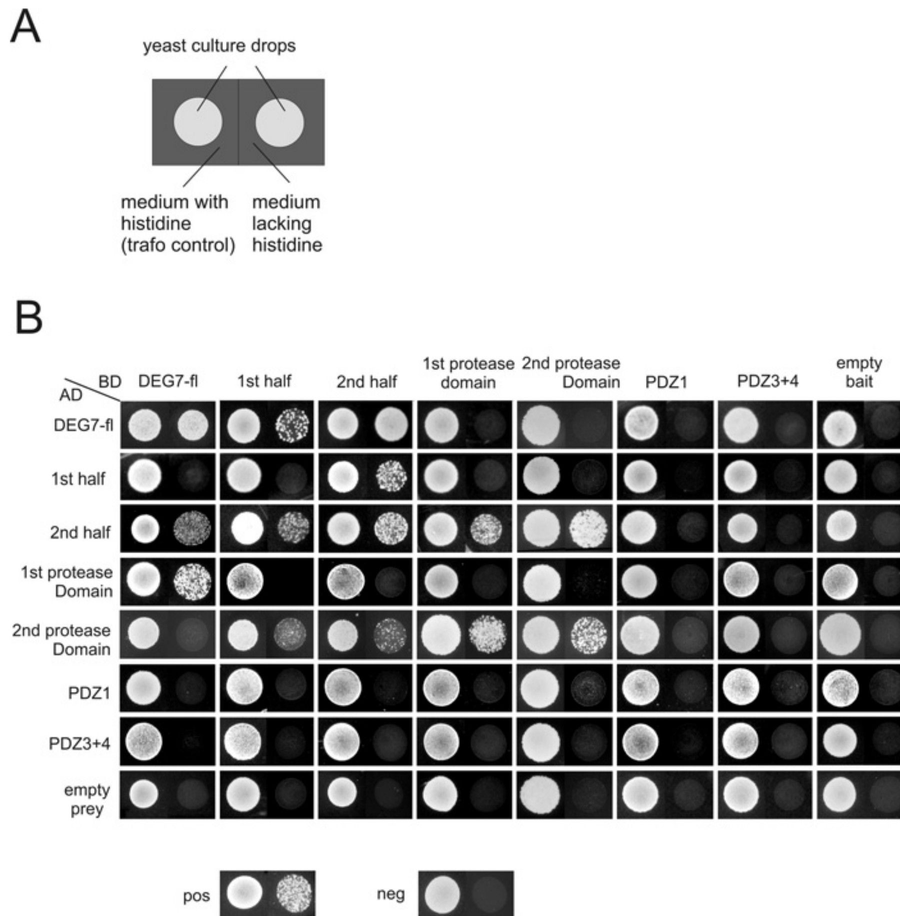

Figure S4 Y2H assay shows that oligomerization of DEG7 is mediated by the second (degenerated) protease domain (extended version)

(**A**) Schematic outline of the experiment. Yeast drop cultures on medium containing histidine, but lacking tryptophan and leucine were used as transformation controls, indicating that the strain contains both the AD and the BD vector. (**B**) Y2H assay as summarized in Table 1 and Figure 4 of the main text. Empty AD and BD vectors were used as autoactivation controls. Every yeast strain growing on histidine-free medium contains the second protease domain in at least one construct. The AD–first half, BD–first half, AD–active protease domain, and BD–active protease domain constructs respectively mediated growth on histidine-free medium in at least some combinations, indicating that lack of gene expression in the combinations first half/first half and active protease domain/active protease domain is not due to a lack of protein expression. pos, positive control; neg, negative control.

Received 1 October 2010/19 January 2011; accepted 19 January 2011
Published as BJ Immediate Publication 19 January 2011, doi:10.1042/BJ20101613

© 2011 The Author(s)
The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.