

# SCIENTIFIC REPORTS



OPEN

## Expression and specificity of a chitin deacetylase from the nematophagous fungus *Pochonia chlamydosporia* potentially involved in pathogenicity

Almudena Aranda-Martinez<sup>1</sup>, Laia Grifoll-Romero<sup>2</sup>, Hugo Aragunde<sup>2</sup>, Enea Sancho-Vaello<sup>2</sup>, Xevi Biarnés<sup>2</sup>, Luis Vicente Lopez-Llorca<sup>1</sup> & Antoni Planas<sup>1</sup> 

Chitin deacetylases (CDAs) act on chitin polymers and low molecular weight oligomers producing chitosans and chitosan oligosaccharides. Structurally-defined, partially deacetylated chitooligosaccharides produced by enzymatic methods are of current interest as bioactive molecules for a variety of applications. Among *Pochonia chlamydosporia* (*Pc*) annotated CDAs, gene *pc\_2566* was predicted to encode for an extracellular CE4 deacetylase with two CBM18 chitin binding modules. Chitosan formation during nematode egg infection by this nematophagous fungus suggests a role for their CDAs in pathogenicity. The *P. chlamydosporia* CDA catalytic domain (*PcCDA*) was expressed in *E. coli* BL21, recovered from inclusion bodies, and purified by affinity chromatography. It displays deacetylase activity on chitooligosaccharides with a degree of polymerization (DP) larger than 3, generating mono- and di-deacetylated products with a pattern different from those of closely related fungal CDAs. This is the first report of a CDA from a nematophagous fungus. On a DP5 substrate, *PcCDA* gave a single mono-deacetylated product in the penultimate position from the non-reducing end (ADAAA) which was then transformed into a di-deacetylated product (ADDAA). This novel deacetylation pattern expands our toolbox of specific CDAs for biotechnological applications, and will provide further insights into the determinants of substrate specificity in this family of enzymes.

*Pochonia chlamydosporia* (Goddard) Zare and Gams is a nematophagous fungus which infects females and eggs of cyst or root-knot nematodes (RKN)<sup>1–3</sup>. It is a biocontrol agent against a number of plant parasitic nematodes in food-security crops such as tomato and barley<sup>4</sup>. *P. chlamydosporia* is also a soil saprophyte and a root endophyte<sup>5–7</sup>. Extracellular enzymes have been related with nematode egg infection<sup>8</sup>. Chitinases and especially proteases<sup>7,9,10</sup> are considered potential virulence factors for degradation of egg-shell components. Interestingly, the recently sequenced *P. chlamydosporia* genome<sup>11</sup> revealed a number of differentially expressed genes encoding for chitin modifying enzymes during the nematode infection process<sup>12</sup>.

Chitin, a linear polysaccharide of  $\beta$ -1,4-linked N-acetylglucosamine residues, is widely distributed in nature, being the major structural component of the exoskeletons of arthropods (including insects and crustaceans) and the fungal cell wall. Chitin is also present in the endoskeletons of mollusks, and in the cell wall of diatoms<sup>13,14</sup>. Chitin is depolymerized by chitinases, and deacetylated by the action of chitin deacetylases (CDAs) leading to chitosans and chitooligosaccharides, characterized by their degree of polymerization (DP), degree of acetylation (DA), and pattern of acetylation (PA). Chitin deacetylases (EC 3.5.1.41) belong to family 4 of carbohydrate esterases (CE4 in the Carbohydrate Active Enzyme classification, [www.cazy.org](http://www.cazy.org))<sup>15</sup> together with rhizobial NodB chitooligosaccharide deacetylases (EC 3.5.1.-), peptidoglycan N-acetylglucosamine deacetylases (EC 3.5.1.104), peptidoglycan N-acetylmuramic acid deacetylases (EC 3.5.1.-), acetyl xylan esterases (EC 3.1.1.72),

<sup>1</sup>Laboratory of Plant Pathology, Department of Marine Sciences and Applied Biology, Multidisciplinary Institute for Environmental Studies Ramón Margalef, University of Alicante, PO box 99, 03080, Alicante, Spain. <sup>2</sup>Laboratory of Biochemistry, Institut Químic de Sarrià, Universitat Ramon Llull, Via Augusta 390, 08017, Barcelona, Spain. Correspondence and requests for materials should be addressed to A.P. (email: [antoni.planas@iqs.edu](mailto:antoni.planas@iqs.edu))

and poly- $\beta$ -1,6-N-acetylglucosamine deacetylases (EC 3.5.1.-). All CE4 enzymes share the NodB homologous domain<sup>16</sup>, with a distorted ( $\beta/\alpha$ )<sub>8</sub> barrel structure<sup>17</sup> that contains the catalytic active site.

Chitin deacetylases (CDAs) play diverse biological functions. In bacteria, CDAs are involved in the catabolism of chitin (*i.e.* marine bacteria of the *Vibrionaceae* family for nitrogen recycling in chitinous debris<sup>18</sup>) or in signalling events (*i.e.* Rhizobia CDAs for Nod factors biosynthesis<sup>19,20</sup>). In fungi, they participate in cell wall morphogenesis and integrity, spore formation, germling adhesion, and fungal autolysis<sup>17,21–25</sup>. Fungal plant pathogens secrete CDAs during infection and early growth phase in the host to evade the plant defense mechanisms triggered by plant chitinases<sup>17</sup>. It has been hypothesized that partial deacetylation of their cell wall chitin or of the chitooligosaccharides (COS) produced by chitinases results in partially deacetylated oligomers that, unlike chitin oligosaccharides, are not well recognized by plant receptors reducing elicitation of plant defenses<sup>26</sup>. Few CDAs have been biochemically characterised with regard to substrate specificity: bacterial CDA such as *Rhizobium meliloti*<sup>27</sup> or *Vibrio cholerae*<sup>28,29</sup> and fungal CDAs such as *Mucor rouxii*<sup>30</sup>, *Aspergillus nidulans*<sup>31,32</sup>, *Colletotrichum lindemuthianum*<sup>33,34</sup>, *Puccinia graminis*<sup>35</sup>, *Pestalotiopsis* sp.<sup>26</sup>, and *Podospora anserina*<sup>36</sup>. They show different specificities on chitooligosaccharides leading to chitosan oligosaccharides with different patterns of acetylation as the result of random, sequential or processive mechanisms<sup>17</sup>. This rises the question on the role of the deacetylation pattern in the biological functions of CDAs.

Chitin is also a structural component of the eggshell of RKN which is the main barrier to pathogens<sup>37</sup> including the nematophagous fungus *P. chlamydosporia*. The fungus genome contains three putative CDA-encoding genes<sup>11,12</sup> of unknown function. Previous results detected chitosan formation in nematode eggs infected by *P. chlamydosporia*. Chitosan is associated with the sites of fungal penetration<sup>12</sup>, suggesting that *P. chlamydosporia* CDAs are involved in nematode infection. However, no studies have been carried out on the activity of these predicted CDAs.

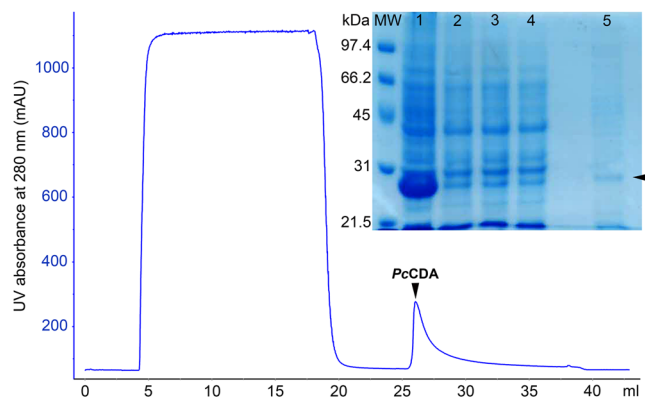
In this work, we report the analysis of *P. chlamydosporia* *pc\_2566*-encoded putative CDA protein (named *PcCDA* hereafter) in order to unravel structure-function relationships with regard to specificity and pattern of deacetylation. In a broader context, novel CDAs need to be characterized to: a) decipher the determinants of specificity leading to chitosan oligomers with different acetylation patterns, and b) enlarge the toolbox of CDAs to generate well-defined chitosan oligosaccharides with biological activities. After expressing the *PcCDA* catalytic domain in recombinant *E. coli*, the enzyme was purified from inclusion bodies by refolding and affinity chromatography. We demonstrate that the enzyme is indeed a deacetylase active on chitooligosaccharides with a novel deacetylation pattern compared to currently known CDAs.

## Results

***Pc\_2566* gene and encoded protein sequence.** Genome sequencing and annotation predicted gene *pc\_2566* (1368 bp ORF) as a putative chitin deacetylase<sup>11</sup> (*Supplementary Information*, Figure S1A). Augustus and GeneMark gene predictors were coincident in the first two introns but not conclusive on other potential introns at the 3'-end (Figure S2). Sequence alignment with the highly homologous (70% identity) *Metarhizium acridum* NW\_006916702.1 nucleotide sequence matched the first two introns in the sequences of both phylogenetically close fungi and identified the position of a third intron (Figure S3). The *PcCDA* translated protein sequence contains a N-terminal signal peptide from residues 1 to 18 (Figure S1). The mature protein (after signal peptide removal) is composed of 455 amino acids with a calculated molecular mass of 48.7 kDa, an isoelectric point of 7.7, and exhibits one potential N-linked glycosylation site and six potential O-glycosylation sites. The catalytic domain (CE4 domain, residues 107 to 303) is flanked by two (N- and C-terminal) CBM18 modules (residues 30 to 74 and 360 to 441, respectively). These family 18 carbohydrate binding modules are typically involved in chitin binding<sup>38</sup>. *PcCDA* full-length protein includes 25 cysteine residues, of which only two are located in the CE4 catalytic domain.

**Cloning, expression and purification of *PcCDA* catalytic domain.** The codon-optimized nucleotide sequence of the CE4 domain was subcloned into a pET22b vector for expression in *E. coli* (Figure S4). The expressed protein has a C-terminal Strep tag for purification by affinity chromatography (Figure S1B) and a predicted molecular mass of 26.8 kDa. All expression attempts varying temperature and time of induction rendered high protein expression but in the insoluble fraction after cell lysis. Solubilisation and refolding steps were necessary to obtain soluble and active protein. Inclusion bodies were solubilized in 7 M urea and soluble protein was recovered after refolding by dialysis (Figure S5) and purified by Strep tag affinity chromatography (Fig. 1). The eluted protein had an apparent molecular mass of 26.8 kDa in agreement with the expected mass. MALDI-TOF-MS analysis after in-gel trypsin digestion confirmed the identity of the *PcCDA* catalytic domain. The overall yield was low (0.43 mg per L of culture) but sufficient for enzyme characterization.

**Deacetylase activity and specificity on chitooligosaccharides.** *PcCDA* catalytic domain was assayed for deacetylase activity on GlcNAc<sub>3</sub> (A3) GlcNAc<sub>4</sub> (A4), and GlcNAc<sub>5</sub> (A5) substrates. Reaction mixtures were analysed by HPLC-MS at different incubation times to monitor products formation. Using the A5 substrate, *PcCDA* generated a mono-deacetylated product after 5 h of reaction (Fig. 2). A di-deacetylated product appeared after 24 h. A5 was almost completely consumed after 100 h reaction, when both mono-deacetylated and di-deacetylated products were present and no further deacetylation was observed. With the A4 substrate, the enzyme was also active but slower than with the A5 substrate. A mono-deacetylated product was formed after 16 h of reaction and a di-deacetylated product appeared only after 100 h reaction while A4 was still present in a significant amount (Figure S6). *PcCDA* had no activity on the A3 substrate (Figure S7), indicating that *PcCDA* is active on chitooligosaccharides with DP > 3.



**Figure 1.** Affinity chromatography purification of *PcCDA* catalytic domain. **(A)** Elution profile monitored by absorbance at 280 nm. **(B)** SDS PAGE analysis of fractions: Lane 1, sample after refolding by dialysis. Lanes 2 and 3, sample after centrifugation and filtered (0.45  $\mu\text{m}$ ) loaded into the column. Lane 4, flow-through from the column. Lane 5, eluted fraction with 2.5 mM d-Desthiobiotin after concentration by ultrafiltration. Arrow indicates column elution using d-Desthiobiotin. Arrowhead points to bands with the expected size (26.8 kDa) for *PcCDA* catalytic domain.

The deacetylation pattern of the products was determined by MALDI-TOF-MS/MS sequencing<sup>39,40</sup> using a preparative reaction with pentaacetylchitopentaose (A5) as substrate. A single mono-deacetylated product (ADAAA) was detected (Fig. 3). The di-deacetylated product (5% in this sample) was ADDAA. Therefore the enzyme starts deacetylating specifically the penultimate residue from the non-reducing end, and continues to the next residue towards the reducing end. No other products were detected under these experimental conditions.

**Sequence alignments and phylogenetic analysis.** *PcCDA* catalytic domain sequence was added to the multiple sequence alignment of CE4 enzymes active on chitooligosaccharides guided by the structural superimposition of available X-ray structures<sup>28</sup>. As seen in Fig. 4, the enzyme shows conservation of the active site motifs MT1–5 characteristic of CE4 enzymes<sup>41</sup>. Specifically, the TFDD (MT1) motif includes the general base aspartate (first D) and the metal-binding aspartate (second D), and motif MT2 (H(S/T)xxH) contains two histidines which, together with the Asp of MT1, form the so called His-His-Asp metal-binding triad of CE4 enzymes. Since the protein here studied was refolded in the presence of  $\text{Zn}^{2+}$ , the native metal of *PcCDA* is unknown. Finally, MT5 contains the general acid histidine for catalysis. The special disposition of the catalytic residues and metal-binding triad are presented below in the structural model of *PcCDA*.

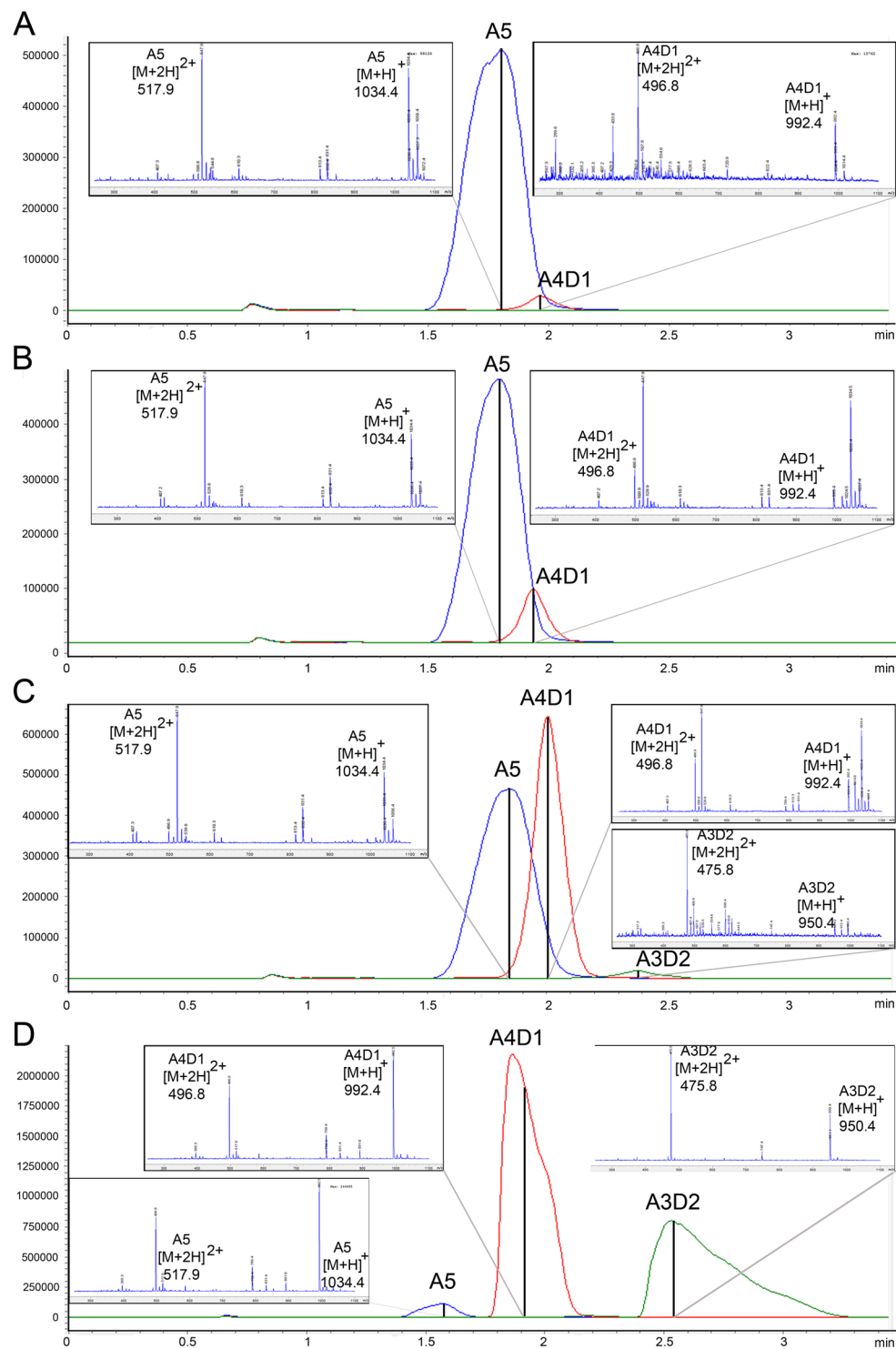
Using the protein sequences of characterized CAZymes from family CE4, the phylogenetic relationship (Fig. 5) showed that *Colletotrichum lindemuthianum* CDA (*CiCDA*) is *PcCDA* closest relative, sharing 43% sequence identity. Fungal CDAs seem to be grouped and segregated from the rest of the characterized CE4 enzyme members.

**Structural model and substrate binding.** *C. lindemuthianum* CDA is the closest homologous protein with a solved 3D structure to the *PcCDA* catalytic domain, followed by the *Aspergillus nidulans* CDA. Both CDAs were used to build a first structural model by homology modelling. Loop 1 (amino acids 56 to 73, Fig. 4) is longer in *PcCDA* than in the templates and it was refined to the best empirical scoring energy, resulting in Model 1 shown in Fig. 6A. Since loop 1 has a similar length to the *VcCDA* protein, a second model using a combination of templates (*CiCDA* + *VcCDA*) was built to give Model 2, which does not leave any relevant part of the *PcCDA* sequence without templates (Fig. 6B). Both models are essentially identical along the protein structure, except for Loop 1, which appeared in two distinct conformations, extended and closed, respectively. The conformation of the loop could not be accurately determined in these calculations, being highly dependent on the template used. However, these different conformations are suggestive of intrinsic loop flexibility since no extensive interactions with core protein residues were observed in any of the models. The overall structure exhibits the canonical  $(\beta/\alpha)_8$  fold of CE4 enzymes and has the  $\text{Zn}^{2+}$  coordination and conserved catalytic residues properly oriented in the active site (Fig. 6D).

Ligand binding was simulated by computational docking of the A4 substrate to Model 1. The preferential binding mode of the ligand placed the substrate with the non-reducing end GlcNAc residue in subsite -1 (subsites numbering as previously defined)<sup>42</sup> (Fig. 6C), which is consistent with the experimentally observed first deacetylation event leading to ADAAA with the A5 substrate.

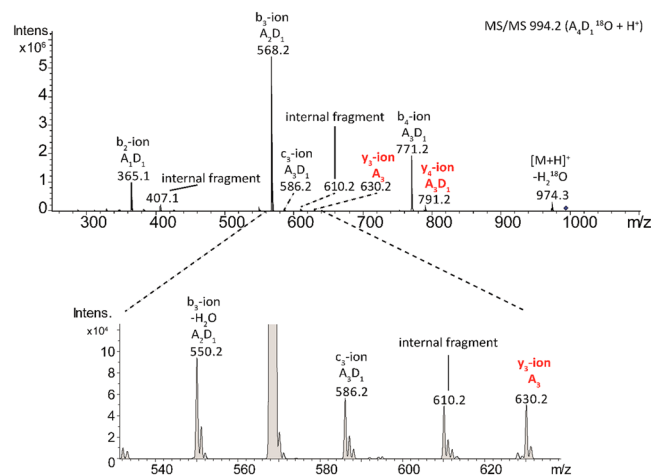
## Discussion

Few CDAs have been biochemically characterized and only four crystallographic structures have been reported, those from *Colletotrichum lindemuthianum* (first structure of a CDA)<sup>34</sup>, *Aspergillus nidulans*<sup>32</sup>, *Vibrio cholera*<sup>28</sup> (and its ortholog from *V. parahaemolyticus*)<sup>43</sup>, and just recently, from a marine *Arthrobacter* species<sup>44</sup>. Efforts are directed to characterize novel CDAs elucidating the determinants of their activity, specificity and deacetylation pattern, and to use them as biocatalysts for the preparation of pure chitosan oligosaccharides with defined structures. These pure compounds rather than mixtures are necessary to assay their bioactivities and implement



**Figure 2.** HPLC-MS monitoring of deacetylase activity by *PcCDA* catalytic domain on  $\text{GlcNAc}_5$ - substrate. Chromatograms show the presence of the substrate (A5) and the formation of mono-deacetylated (A4D1) and di-deacetylated (A3D2) products at different reaction times: (A) 10 min, (B) 5 h, (C) 24 h, and (D) 100 h. Reaction conditions: 0.2 mM substrate, 3.2 nM enzyme, 50 mM  $\text{K}_2\text{HPO}_4$ , 300 mM NaCl, pH 8.0, 37 °C.

applications in agriculture, medicine, cosmetics, and food sciences<sup>17,26,45</sup>. To our knowledge, this is the first report on a chitin deacetylase from nematophagous fungi. The *P. chlamydosporia* genome encodes for three putative CDAs classified in carbohydrate esterases family 4 (CE4)<sup>11</sup> based on their translated amino acid sequence. In this work, we report the protein purification, activity on different chitin oligomers, and de-*N*-acetylation pattern of the catalytic domain of *PcCDA* protein encoded by the *Pc\_2566* gene.



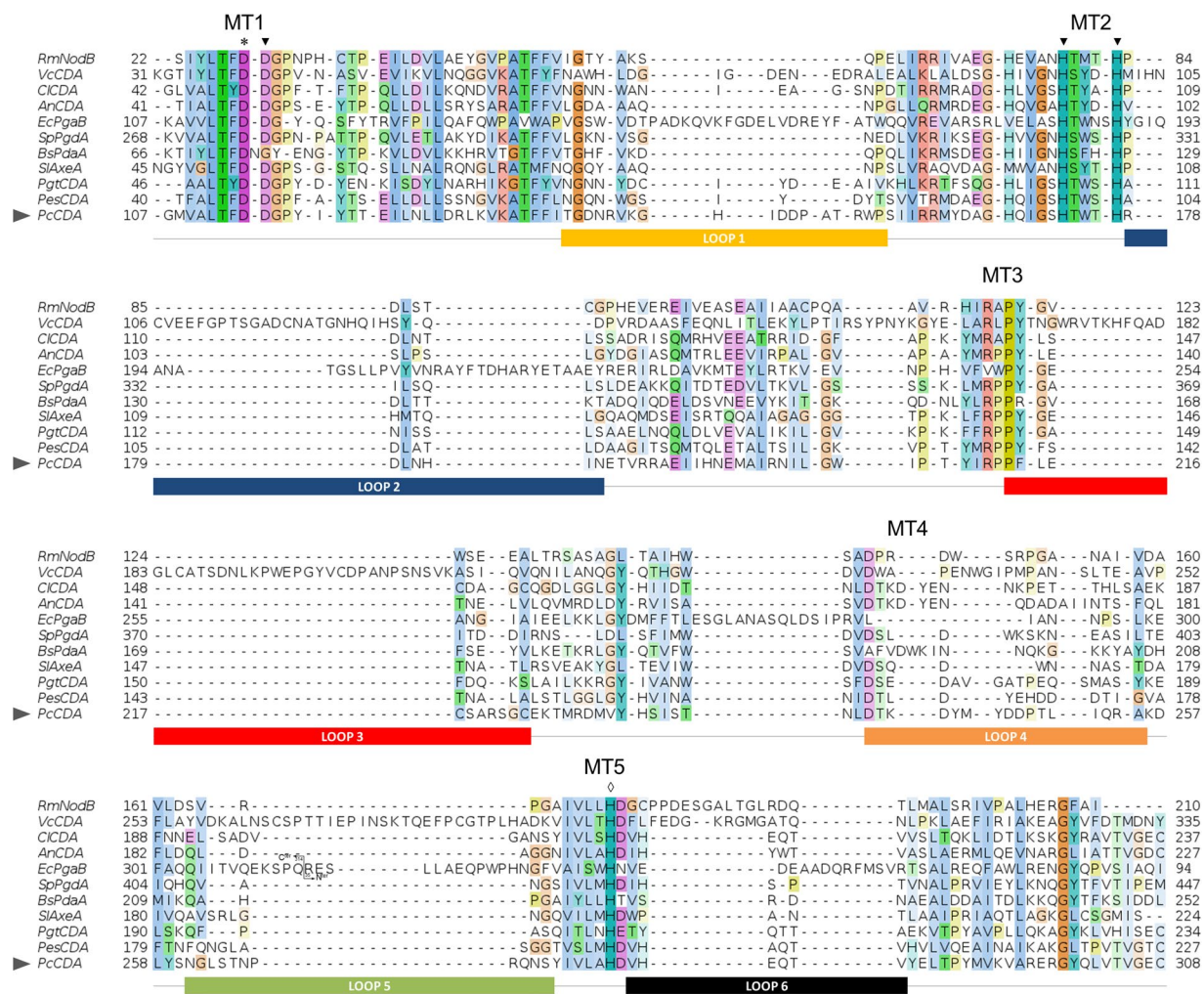
**Figure 3.** MS/MS spectrum of the mono-deacetylated product (ADAAA) from the PcCDA reaction with (GlcNAc)<sub>5</sub>. The reaction mixture contained mainly mono-deacetylated product and traces of di-deacetylated product. The mixture was subjected to the procedure reported in<sup>40</sup>. After reducing-end labelling with H<sub>2</sub>(<sup>18</sup>O), the sample was analyzed by UHPLC-ESI-MS<sup>2</sup>. Fragmentation spectrum of the mono-deacetylated product (A<sub>4</sub>D<sub>1</sub><sup>18</sup>O): b-ions are fragments from the non-reducing end, and y-ions are fragments from reducing end with the <sup>18</sup>O label.

**PcCDA was isolated by refolding from inclusion bodies.** Fungal proteins are difficult to express in bacterial hosts because they are often glycosylated proteins and/or contain many disulphide bonds<sup>35</sup>. Some fungal CDAs have been successfully expressed in *E. coli*. *Colletotrichum lindemuthianum* and *Aspergillus nidulans* CDAs were obtained after protein refolding<sup>46,47</sup> and CDAs from *Puccinia graminis* and *Pestalotiopsis* sp. overcame the difficult protein expression using protein fusions with maltose binding protein (MBP), which assisted folding and resulted in increased solubility and activity<sup>26,35</sup>. After codon optimization for expression in *E. coli*, we attempted, unsuccessfully, several protocols to express the isolated catalytic domain of PcCDA, including different induction times and temperatures, and heat shock before induction. Aggregates or inclusion bodies were always obtained. Best results were achieved by denaturation and refolding to recover soluble and active protein from inclusion bodies. Although the protein was prone to aggregation during the purification protocol, it remained soluble after the final chromatographic step. The final yield of active recombinant protein was low and may be improved through further optimization or using alternative expression systems.

**PcCDA is active on COS and exhibits a novel deacetylation pattern.** Whereas many fungal CDAs have been characterized on polymeric substrates (colloidal chitin, or soluble polymer glycol-chitin and CM-chitin), few CDAs have been analysed for substrate specificity and mode of action on low molecular weight COS, being the best characterized those from *Colletotrichum lindemuthianum*<sup>46,48</sup>, *Aspergillus nidulans*<sup>31,32</sup>, *Saccharomyces cerevisiae*<sup>49</sup>, and more recently from *Puccinia graminis*<sup>35</sup>, *Pestalotiopsis* sp.<sup>26</sup>, and *Podospira anserina*<sup>36</sup>. The closest homologue of PcCDA is the *C. lindemuthianum* ClCDA (Fig. 5) with 43% sequence identity. We show that PcCDA deacetylates COS with DP > 3. For A4 and A5, the mono-deacetylated product predominates during the early stage of the reaction and the di-deacetylated product appears after longer incubation periods. A similar enzymatic behaviour was described for *Pes*CDA<sup>26</sup> and *Ci*CDA<sup>50</sup> but with different specificities. As opposed to *Ci*CDA, which is active on COS as short as N,N'-diacetylchitobiose (A2), PcCDA does not deacetylate oligosaccharides shorter than DP4. PcCDA initially deacetylates A5 at the penultimate GlcNAc residue from the non-reducing end (ADAAA), and then deacetylates the next residue towards the reducing end (ADDAA). No further deacetylated products were observed under the assayed conditions. In contrast, *Ci*CDA starts deacetylating the third residue from the non-reducing end on COS with DP ≥ 3, and continues with full deacetylation of the substrates. This indicates that subtle differences in the active site topology are responsible for the binding specificity and deacetylation pattern exhibited by closely related CDAs.

**The modelled structure of PcCDA supports substrate specificity.** The recently proposed “Subsite Capping Model” suggests that the deacetylation pattern is dictated by critical loops that shape and differentially block accessible subsites in the binding cleft of CE4 enzymes<sup>28</sup>. Negative subsites accommodating GlcNAc units on the non-reducing end of the substrate are shaped by Loops 1, 2, and 6. PcCDA has a distribution of short loops, similar to its closest homologue *Ci*CDA, except for Loop 1, which is longer but of equivalent length than that of *Vibrio cholerae* CDA (*Vc*CDA) (Fig. 4). According to our model, long Loop 1 sequences partially block the accessibility of non-reducing end subsites. Accordingly, PcCDA and *Vc*CDA produce, as first deacetylated products, chitosan oligomers with the same deacetylation pattern on the penultimate GlcNAc residue from the non-reducing end (ADAAA). However, PcCDA continues by deacetylating a second residue while *Vc*CDA is highly specific for monodeacetylation. It may be speculated that, whereas Loop 1 in *Vc*CDA is fixed by a network of H-bond interactions with Loops 2 and 6, PcCDA Loop 1 might be more flexible as suggested by the structural models (Fig. 6A and B), since loop 1 establishes weaker interactions with neighbouring loops (in particular to

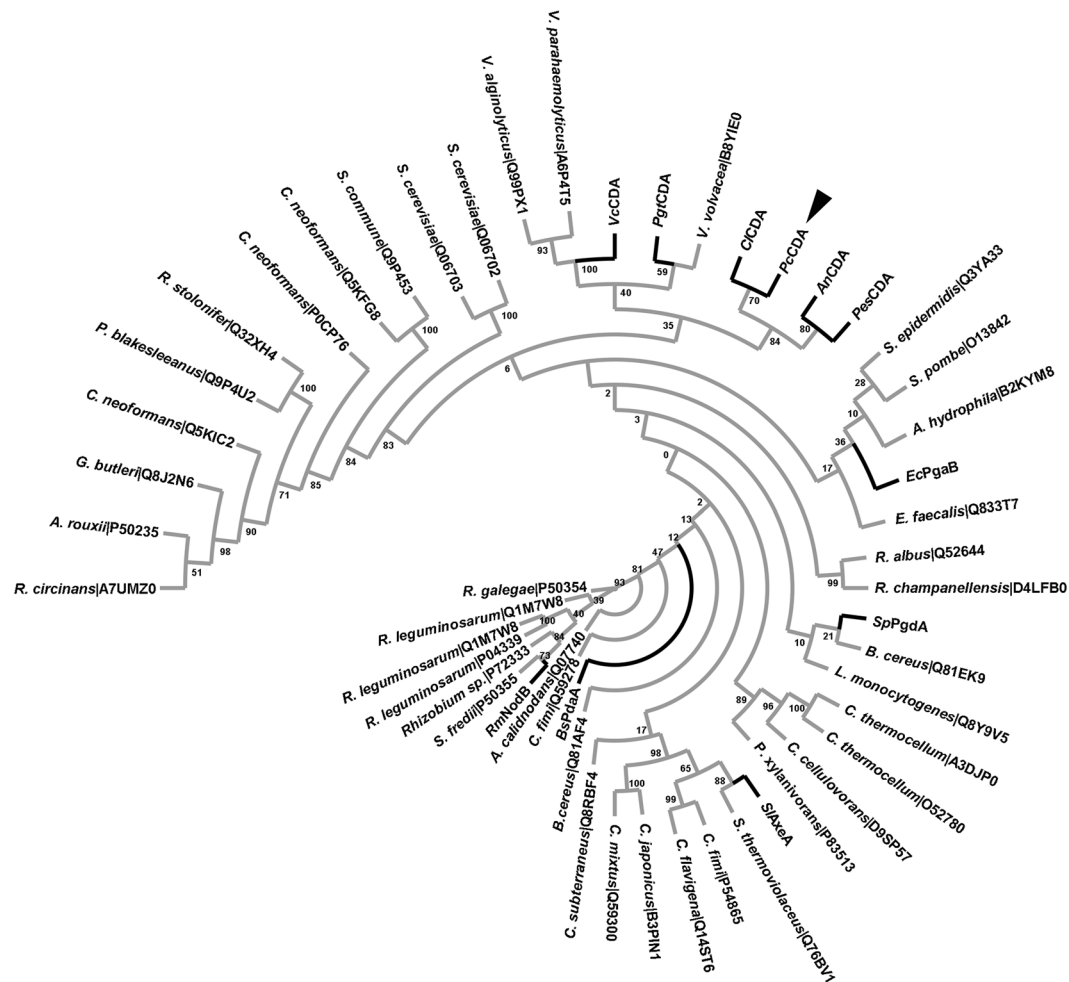




**Figure 4.** Multiple sequence alignment of chitin deacetylase (CDA) catalytic domains. Abbreviations: *CiCDA* (*Colletotrichum lindemuthianum*), *AnCDA* (*Aspergillus nidulans*), *PesCDA* (*Pestalotiopsis* sp.), *BsPdaA* (*Bacillus subtilis*), *PgtCDA* (*Puccinia graminis*), *VcCDA* (*Vibrio cholerae*), *RmNodB* (*Rhizobium meliloti*), *SlAxeA* (*Streptomyces lividans*), *SpPgdA* (*Streptococcus pneumoniae*), *EcPgaB* (*Escherichia coli*) and *PcCDA* (*Pochonia chlamydosporia*). Loops are highlighted with colored boxes according to<sup>28</sup>. The arrowhead indicates the sequence of *PcCDA* catalytic domain. Conserved catalytic motifs are labelled MT1-5. The 'His-His-Asp' metal binding triad (▼), catalytic base (\*), and catalytic acid (◇) are highlighted.

shorter Loops 2 and 6 in *PcCDA* as compared to *VcCDA* (Supplementary Information Figures S9 and S10). This may allow the exposure of an additional negative subsite for a second deacetylation to take place (ADDA) at a much slower rate. The role of loop dynamics is currently under study to rationalize (and be able to engineer) the deacetylation pattern exhibited by different CDAs.

**What is the biological function of *PcCDA*?** One of the most studied biological functions of fungal CDAs is the protection of plant pathogenic fungi from being lysed, avoiding plant immunity responses<sup>17,26,51</sup>. Partial deacetylation of the exposed chitin polymer (protecting it from the action of secreted plant chitinases) or of the elicitor-active chitin oligomers (preventing binding to receptors) are proposed mechanisms to evade the plant immune response<sup>52</sup>, as suggested in *Colletotrichum* spp.<sup>53</sup> and *Pestalotiopsis* sp.<sup>26</sup>. Fungal CDAs play also other physiological roles, participating in fungal nutrition, morphogenesis and development<sup>17,21-23</sup>, spore formation<sup>24</sup>, and germling adhesion<sup>25</sup>. CDAs have been described as virulence factors in animal pathogenic fungi. Cell wall chitin deacetylation seems essential for *Cryptococcus neoformans* virulence in lungs<sup>54</sup>. CDAs are considered putative virulence factors of entomopathogenic fungi since chitosan was detected in insect cuticle when infected by *Metarhizium anisopliae*<sup>55</sup>. In the case of *P. chlamydosporia*, the function of CDAs may be involved in nematode eggs infection rather than in a defense mechanism. Nematode eggshell contains chitin microfibrils<sup>56</sup>. *P. chlamydosporia* CDAs may act on the host nematode chitin because chitosan has been immunolocalised during infection of nematode eggs by the fungus<sup>12</sup>. Chitin deacetylation makes this polymer more elastic and soluble for fungal penetration using both hyphae and apressoria<sup>1</sup> with concomitant degradation of eggshell components by extracellular enzymes<sup>57</sup>. CDA activity from *P. chlamydosporia* would not only have relevance in biotechnology



**Figure 5.** Phylogram of CE4 catalytic domains. Amino acid sequences of characterized CAZymes from family CE4 were retrieved from Uniprot database. The enzymes reported experimentally to be active on chito oligosaccharides are highlighted using bold branches. Abbreviations: CLCDA (*Colletotrichum lindemuthianum*), PcCDA (*Pochonia chlamydosporia*), AnCDA (*Aspergillus nidulans*), PesCDA (*Pestalotiopsis* sp.), BsPdaA (*Bacillus subtilis*), PgtCDA (*Puccinia graminis*), VcCDA (*Vibrio cholerae*), RmNodB (*Rhizobium meliloti*), SLAxeA (*Streptomyces lividans*), SpPgDA (*Streptococcus pneumoniae*), and EcPgaB (*Escherichia coli*).

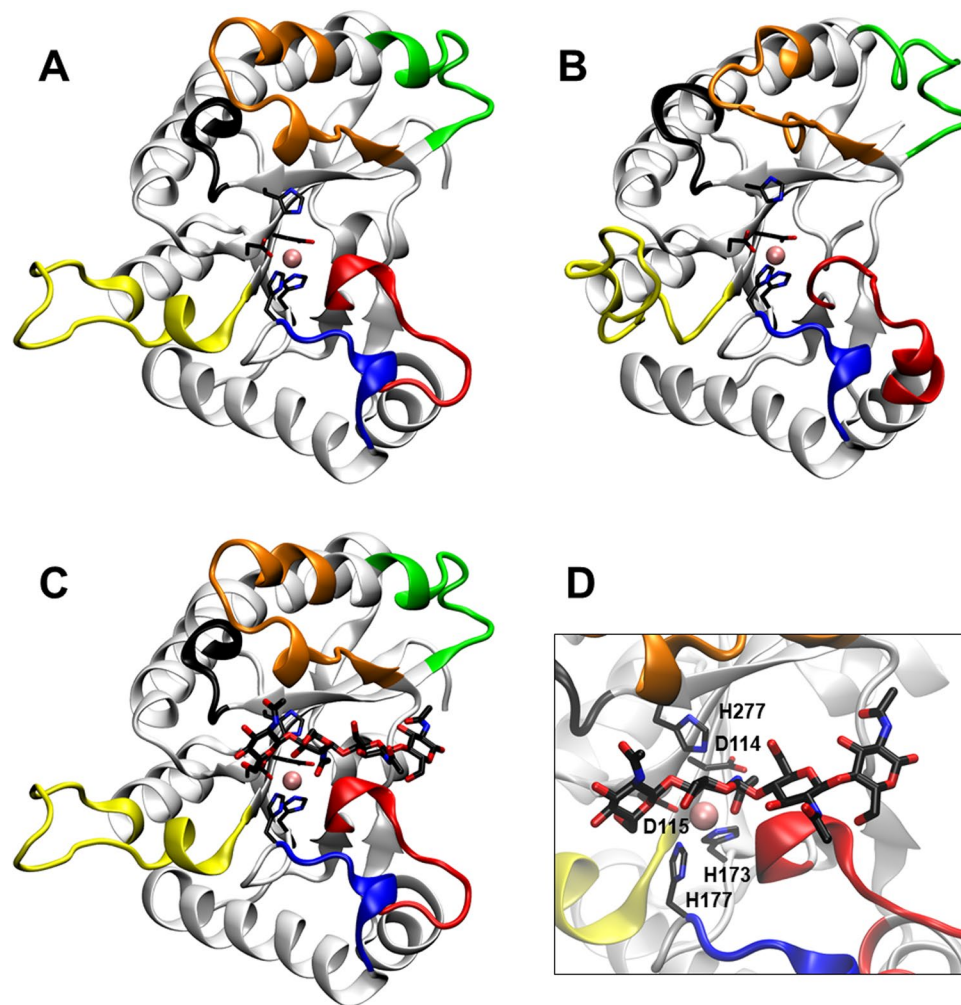
and agriculture but also in human and animal health because this fungus also infects eggs from animal parasitic nematodes<sup>58</sup>.

## Conclusion

*P. chlamydosporia* Pc\_2566 encodes for an active chitin deacetylase (PcCDA) potentially involved in nematode egg infection. The novel deacetylation pattern exhibited by the enzyme expands the repertoire of specific CDAs for biotechnological applications, where chitosan oligomers with defined pattern of acetylation are required to decipher structure-bioactivity relationships.

## Materials and Methods

**Analysis of Pc\_2566 gene sequence.** Pc\_2566 gene sequence of *Pochonia chlamydosporia* strain 123, predicted *in silico* as a chitin deacetylase<sup>11</sup>, was retrieved from [www.fungalinteractions.org](http://www.fungalinteractions.org) (accession date 5 February 2015, Genbank assembly accession GCA\_000411695.1). GeneMark and Augustus gene predictors were used for gene structure prediction. The presence of introns and their positions were verified using tBlastn searches (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) against RefSeq Representative Genome Database. The predicted protein-coding DNA sequence was then translated using Translate tool (<http://web.expasy.org/translate>). The multi-domain structure of PcCDA protein sequence was analysed using the online bioinformatic tools Prosite (<http://prosite.expasy.org/>) and Superfamily (<http://supfam.org/SUPERFAMILY/>). Signal peptide (<http://www.cbs.dtu.dk/services/SignalP/>), transmembrane helices (<http://www.cbs.dtu.dk/services/TMHMM/>), N-glycosylation (<http://www.cbs.dtu.dk/services/NetNGlyc/>) and O-glycosylation (mucin-type) (<http://www.cbs.dtu.dk/services/NetOGlyc/>) site predictors were also used.



**Figure 6.** Structural models of the *PcCDA* catalytic domain. (A) Model 1, using *ClCDA* (PDB 2IW0) and *AnCDA* (PDB 2Y8U) as templates. (B) Model 2, using *ClCDA* (PDB 2IW0) and *VcCDA* (PDB 4OUI) as templates. The loops are coloured as in Fig. 4 according to<sup>28</sup>. (C) Simulated docking of A4 ligand to Model 1, lowest energy binding mode which places the penultimate GlcNAc residue properly oriented for catalysis in subsite 0. (D) Magnification of the active site in Model 1, showing residues Asp<sup>115</sup>-His<sup>173</sup>-His<sup>177</sup> (metal binding triad), Asp<sup>114</sup> (general base), His<sup>277</sup> (general acid), and the Zn<sup>2+</sup> cation.

**Construction of a *PcCDA* expression plasmid.** A synthetic gene encoding for *Pochonia chlamydosporia* chitin deacetylase was designed using the predicted protein sequence and codon-optimised for *E. coli* expression (GeneOptimizer™ software, GeneArt® Gene Synthesis service, ThermoFisher)<sup>59</sup>. The signal peptide fragment was removed and restriction sites NdeI and SacI were added at the 5' and 3'-ends, respectively. The gene was sub-cloned in the vector pET22b(+)-StrepIIIC between NdeI and SacI restriction sites. The construction of pET22b(+)-StrepIIIC was described earlier<sup>39</sup>. Briefly, the vector pET22b(+) (Novagen) was used as a template to include a StrepII encoding sequence downstream of the multiple cloning site using PCR via a 50-mer phosphorylated primer pair. Electrocompetent *E. coli* DH5α cells were transformed with the pET22b(+)-StrepIIIC plasmid containing the *Pc\_2566* synthetic gene. Positive transformants were selected using LB medium supplemented with 0.1 mg mL<sup>-1</sup> ampicillin and cells were grown in 10 mL of the same media and incubated overnight at 37 °C (200 rpm). The plasmid was then extracted using QIA Prep Spin MiniPrep Kit (Qiagen), and *Pc\_2566* gene sequence was verified by DNA sequencing (Fig. S1). The sequence encoding the catalytic domain was amplified by PCR with specific primers containing NdeI and SacI restriction sites (Forward; 5'-TATGCATATGGTTCGATGGTCCGATGATTACC-3' and Reverse; 5'-TATCGAGCTCCAGACATTCACCAACGGTAACC-3') using iProof High Fidelity PCR Kit (Bio-Rad). The PCR product was purified (GenElute™ PCR Clean-Up Kit, Sigma-Aldrich) and digested with NdeI and SacI restriction enzymes (2.5 h, 37°). After purification (1% agarose gel electrophoresis and extraction with GenElute™ Gel Extraction Kit (Sigma-Aldrich)), the catalytic domain of *Pc\_2566* gene sequence was ligated (T4 DNA ligase, Bio-Rad) to pre-digested pET22b(+)-StrepIIIC vector with the same restriction enzymes at 16 °C overnight. *E. coli* DH5α cells were transformed with the ligation product, and positive transformants (ampicillin resistant) were verified by DNA sequencing. *E. coli* BL21(DE3) cells were transformed



with the resulting pET22b-*Pc\_2566\_CE4*-StrepIIC plasmid (encoding the *PcCDA* catalytic domain with C-terminal Step tag II sequence) for protein expression.

**Expression of *PcCDA* catalytic domain in *E. coli*.** *E. coli* BL21(DE3) cells containing the expression plasmid pET22b-*Pc\_2566\_CE*-StrepIIC were grown in LB medium supplemented with 0.1 mg mL<sup>-1</sup> ampicillin at 37 °C under agitation (200 rpm). When cell density ( $A_{600}$ ) reached 0.8, the culture was induced with 0.02 mM IPTG and 2% sterile ethanol (by filtration) at 16 °C and the culture was maintained with agitation at 16 °C for 18 h. Cells were harvested by centrifugation (20 min, 5000 rpm) and stored at -20 °C until processing.

**Protein purification by refolding and affinity chromatography.** Cell pellet was washed once with PBS buffer (50 mM K<sub>2</sub>HPO<sub>4</sub>, 300 mM NaCl, pH 8) containing 1 mM dithiothreitol (DTT) and then centrifuged (5000 rpm) for 20 min. The pellet was resuspended in 100 mL PBS buffer containing 1 mM DTT and 1 mM phenylmethylsulfonyl fluoride (PMSF) and lysed by sonication at 4 °C using a Soniprep 150 sonifier (7 min, 10 s ON/25 s OFF, 50% amplitude). Soluble and insoluble fractions were separated by centrifugation at 10000 rpm for 60 min. Both fractions were analysed by SDS-PAGE (14%) using Coomassie blue staining (Bio-Safe Coomassie G-250 Stain, Bio-Rad).

The insoluble fraction obtained from ca. 4 g of cells was washed twice with 20 mL of PBS buffer (50 mM K<sub>2</sub>HPO<sub>4</sub>, 300 mM NaCl, pH 8) containing 1 mM DTT and 1% Triton X-100 and then centrifuged at 15000 rpm for 20 min. The resulting pellet was resuspended, washed twice with PBS buffer containing 1 mM DTT followed by centrifugation (15000 rpm; 20 min). The final pellet was resuspended in 15 mL of a solution of 7 M urea, 1 mM DTT in PBS buffer and incubated at 4 °C for 30 min with shaking (150 rpm). After centrifugation (14800 rpm; 35 min), the solubilised inclusion bodies were refolded by dialysis against 1 L of PBS buffer with 1 mM DTT and 1 mM ZnCl<sub>2</sub> at 4 °C to remove urea (two buffer changes, 2 h each step) followed by dialysis against PBS buffer containing 1 mM ZnCl<sub>2</sub> overnight. Some protein precipitated and was removed by centrifugation (10 min at 15000 rpm, 4 °C). The final supernatant was sonicated for 1 min, 0.45 µm filtered and stored at 4 °C until used.

The refolded *PcCDA* catalytic domain was purified by affinity chromatography on an ÄKTA Protein Purification System (Amersham Biosciences) using a StrepTrap column (GE Healthcare). The protein was eluted with 2.5 mM d-Desthiobiotin in PBS buffer. The protein-containing fractions were combined, and the buffer was exchanged (PBS buffer) and the protein concentrated up to 2 mL using an Amicon Ultra-15 Centrifugal Filter, (Millipore). Protein concentration was determined with the BCA Protein Assay Kit (ThermoFisher).

**Analysis of chitin deacetylase activity of *PcCDA* catalytic domain.** Chitin deacetylase reactions were performed at final concentrations of 0.2 mM GlcNAc<sub>5</sub> (A5), GlcNAc<sub>4</sub> (A4) or GlcNAc<sub>3</sub> (A3) substrates, 3.2 nM protein, in 50 mM K<sub>2</sub>HPO<sub>4</sub>, 300 mM NaCl, buffer at pH 8 in a total volume of 200 µL at 37 °C. At different time intervals, 10 µL aliquots were withdrawn and mixed with 90 µL of H<sub>2</sub>O:propanol (1:1) in HPLC vials to stop the reaction. Samples were analysed by HPLC-MS (HPLC 1200, ESI-MS 6100 series SQ, Agilent Technologies) using a XBridge BEH Amide 2.5 µm 3.0 × 100 mm Column XP, (Waters) in combination with a XBridge Amide Guard Cartridge (2PK) pre-column (2.5 µm 4.6 × 20 mm; Waters), 5 µL injection, and isocratic elution at 60 °C with acetonitrile/water 65:35 v/v, 0.1% formic acid, at a flow rate of 0.4 ml/min. MS detection monitored (SIM mode) the following [M+H]<sup>+</sup> ion masses: m/z 628 (A3), 586 (A2D1), 544 (A1D2), 831 (A4), 789 (A3D1), 747 (A2D2), 1034 (A5), 992 (A4D1) and 950 (A3D2). Data was analysed with the ChemStation Software (Agilent Technologies).

**Pattern of deacetylation by *PcCDA* catalytic domain.** A preparative reaction was performed by incubating 0.3 mg of freshly prepared *PcCDA* catalytic domain protein and 5 mg pentaacetylchitopentaose (A5) substrate in PBS buffer (50 mM K<sub>2</sub>HPO<sub>4</sub>, 300 mM NaCl, pH 8) at 37 °C, in a final volume of 1 mL. After 48 h reaction time, the sample mostly contained mono-deacetylated product and traces of di-deacetylated product (as determined by HPLC-MS as above). The mixture was analysed following the procedure reported in<sup>40</sup>. Briefly, the freeze-dried sample was subjected to reducing-end labelling with H<sub>2</sub>(<sup>18</sup>O) and analysed by UHPLC-ESI-MS<sup>2</sup>, where each labelled product was sequenced by fragmentation in the MS/MS analyser. The fragmentation spectrum of the mono-deacetylated product is shown in Fig. 3.

**Multiple sequence alignment and phylogenetic analysis.** *PcCDA* catalytic domain protein sequence was incorporated to the multiple sequence alignment of CE4 deacetylases active on chitoooligosaccharides previously reported<sup>28</sup> by hidden Markov model comparisons using HMMER<sup>60</sup>. *PcCDA* was analysed together with *CiCDA* (*Colletotrichum lindemuthianum*, accession O87119), *AnCDA* (*Aspergillus nidulans*, accession Q5AAQ0), *PesCDA* (*Pestalotiopsis* sp., accession A0A1L3THR9), *BsPdaA* (*Bacillus subtilis*, accession O34928), *PgtCDA* (*Puccinia graminis*, accession E3K3D7), *VcCDA* (*Vibrio cholera*, accession Q9KSH6), *RmNodB* (*Rhizobium meliloti*, accession P02963), *SlAxeA* (*Streptomyces lividans*, accession Q54413), *SpPgdA* (*Streptococcus pneumoniae*, accession Q8DP63) and *EcPgaB* (*Escherichia coli*, accession P75906) CDA protein sequences. Sequence alignment is shown in Fig. 4.

Amino acid sequences of characterized family CE4 CAZYmes ([www.CAZY.org](http://www.CAZY.org)) were retrieved from Uniprot database (August 2017). The phylogenetic relationships were inferred by using the Maximum Likelihood method based on the JTT matrix-based model<sup>61</sup>. Bootstrap analysis consisted of 500 replicates. The evolutionary analysis was conducted in MEGA7<sup>62</sup>, and the output dendrogram shown in Fig. 5.

**Molecular modelling of the *PcCDA* catalytic domain and ligand docking.** The three dimensional (3D) structure of *PcCDA* catalytic domain was modelled by means of homology modelling using the HHPRED server<sup>63</sup> and MODELLER software<sup>64,65</sup>. The X-ray structures of *CiCDA* (PDB: 2IW0), *AnCDA* (PDB: 2Y8U), and *VcCDA* (PDB: 4OUI, in complex with A3) were selected as templates for the threading. The model was refined

with a short simulated annealing protocol as implemented in MODELLER. Both 3D models, Model 1 using 2IW0 and 2Y8U as templates, and Model 2 using 2IW0 and 4OUI as templates, rendered an ensemble of conformations for Loop 1. The final structure models were assessed by means of empirical scoring energies with the DOPE score<sup>66</sup>.

The preferential binding modes of tetraacetylchitotetraose (A4) on to PcCDA Model 1 structure were evaluated by means of virtual docking with AutoDock VINA algorithm<sup>67</sup>. The structure of the A4 ligand was extracted from PDB 1LZC. Both the protein and ligand structures were first parametrized: polar hydrogens were added with AutoDockTools<sup>68</sup>, AutoDock4.2 atom typing was used, and Gaister partial charges were computed for each atom with AutoDockTools. All rotatable bonds of the ligand were considered free during the docking calculations, whereas the whole protein structure was kept fixed. A grid-box of  $34.5 \times 33.75 \times 24 \text{ \AA}^3$  centered at the active site was used as the search space for docking. The search for 20 different binding modes was requested with an exhaustiveness parameter set to 24. 3D structures were analyzed with the VMD visualization software<sup>69</sup>.

**Data Availability.** All data generated or analysed during this study are included in this published article (and its Supplementary Information file).

## References

- Lopez-Llorca, L. V. & Duncan, G. H. Effects of fungal parasites on cereal cyst nematode (*Heterodera avenae* Woll.) from naturally infested soil - a scanning electron microscopy study. *Can. J. Microbiol.* **37**, 218–225 (1991).
- Olivares-Bernabeu, C. M. & Lopez-Llorca, L. V. Fungal egg-parasites of plant-parasitic nematodes from Spanish soils. *Rev. Iberoam. Micol.* **19**, 104–110 (2002).
- Escudero, N. & Lopez-Llorca, L. V. Effects on plant growth and root-knot nematode infection of an endophytic GFP transformant of the nematophagous fungus *Pochonia chlamydosporia*. *Symbiosis*. **57**, 33–42 (2012).
- Manzanilla-Lopez, R. *et al.* *Pochonia chlamydosporia*: Advances and challenges to improve its performance as a biological control agent of sedentary endoparasitic nematodes. *J. Nematol.* **45**, 1–7 (2013).
- Domsch, K. H., Gams, W. & Anderson, T. H. *Compendium Of Soil Fungi* (IHW-Verlag: Eching (1993).
- Bordallo, J. J. *et al.* Colonization of plant roots by egg-parasitic and nematode-trapping fungi. *New Phytol.* **154**, 491–499 (2002).
- Lopez-Llorca, L. V. & Robertson, W. Immunocytochemical localization of a 32 kDa protease from the nematophagous fungus *Verticillium suchlasporium* in infected nematode eggs. *Exp. Mycol.* **16**, 261–267 (1992).
- Huang, X., Zhao, N. & Zhang, K. Extracellular enzymes serving as virulence factors in nematophagous fungi involved in infection of the host. *Res. Microbiol.* **155**, 811–816 (2004).
- Segers, R., Butt, T., Keen, J., Kerry, B. R. & Peberdy, J. The subtilisins of the invertebrate mycopathogens *Verticillium chlamydosporium* and *Metarhizium anisopliae* are serologically and functionally related. *FEMS Microbiol. Lett.* **126**, 227–231 (1995).
- Tikhonov, V. E., Lopez-Llorca, L. V., Salinas, J. & Jansson, H. B. Purification and characterization of chitinases from the nematophagous fungi *Verticillium chlamydosporium* and *V. suchlasporium*. *Fungal Genet. Biol.* **35**, 67–78 (2002).
- Larriba, E. *et al.* Sequencing and functional analysis of the genome of a nematode egg-parasitic fungus, *Pochonia chlamydosporia*. *Fungal Genet. Biol.* **65**, 69–80 (2014).
- Aranda-Martinez, A. *et al.* CAZyme content of *Pochonia chlamydosporia* reflects that chitin and chitosan modification are involved in nematode parasitism. *Environ. Microbiol.* **18**, 4200–4215 (2016).
- Brück, W. M., Slater, J. W. & Carney, B. F. *Chitin, Chitosan, Oligosaccharides and Their Derivatives: Biological Activities and Applications* (ed. Kim, S. K.) 11–12. CRC Press (2010).
- Peniche, C., Argüelles-Monal, W. & Goycoolea, F. M. Chitin and chitosan: Major sources, properties and applications. In *Monomers, Polymers and Composites from Renewable Resources* (eds Belgacem, M. N. & Gandini, A.) 517–542. Elsevier (2008).
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZY) in 2013. *Nucleic Acids Res.* **42**, 490–495 (2014).
- Kafetzopoulos, D., Thireos, G., Vournakis, J. N. & Bouriotis, V. The primary structure of a fungal chitin deacetylase reveals the function for two bacterial gene products. *Proc. Natl. Acad. Sci. USA* **90**, 8005–8008 (1993).
- Zhao, Y., Park, R. D. & Muzzarelli, R. A. A. Chitin deacetylases: Properties and applications. *Mar. Drugs*. **8**, 24–46 (2010).
- Ferguson, M. J. L. & Gooday, G. W. *Chitin Enzymology* (ed. Muzzarelli, R. A. A.) 393–396 (Atec; Grottammare, 1996).
- Gough, C. & Cullimore, J. Lipo-chitooligosaccharide signaling in endosymbiotic plant-microbe interactions. *Mol Plant-Microbe Interact.* **24**, 867–878 (2011).
- Hamel, L. P. & Beaudoin, N. Chitooligosaccharide sensing and downstream signaling: Contrasted outcomes in pathogenic and beneficial plant-microbe interactions. *Planta* **232**, 787–806 (2010).
- Hamid, R. *et al.* Chitinases: An update. *J. Pharm. Bioallied. Sci.* **5**, 21–29 (2013).
- Davis, L. L. & Bartnicki-Garcia, S. The co-ordination of chitosan and chitin synthesis in *Mucor rouxii*. *J. Gen. Microbiol.* **130**, 2095–2102 (1984).
- Baker, L. G., Specht, C. A., Donlin, M. J. & Lodge, J. K. Chitosan, the deacetylated form of chitin, is necessary for cell wall integrity in *Cryptococcus neoformans*. *Eukaryot. Cell*. **6**, 855–867 (2007).
- Matsuo, Y., Tanaka, K., Matsuda, H. & Kawamukai, M. *cdal+*, encoding chitin deacetylase is required for proper spore formation in *Schizosaccharomyces pombe*. *FEBS Lett.* **579**, 2737–2743 (2005).
- Geoghegan, I. A. & Gurr, S. J. Chitosan mediates germling adhesion in *Magnaporthe oryzae* and is required for surface sensing and germling morphogenesis. *PLoS Pathog.* **12**, e1005703, <https://doi.org/10.1371/journal.ppat.1005703> (2016).
- Cord-Landwehr, S., Melcher, R. L., Kolkenbrock, S. & Moerschbacher, B. M. A chitin deacetylase from the endophytic fungus *Pestalotiopsis* sp. efficiently inactivates the elicitor activity of chitin oligomers in rice cells. *Sci Rep.* **6**, 38018, <https://doi.org/10.1038/srep38018> (2016).
- John, M., Rohrig, H., Schmidt, J., Wieneke, U. & Schell, J. *Rhizobium* NodB protein involved in nodulation signal synthesis is a chitooligosaccharide deacetylase. *Proc. Natl. Acad. Sci. USA* **90**, 625–629 (1993).
- Andrés, E. *et al.* Structural basis of chitin oligosaccharide deacetylation. *Angew. Chem Int Ed.* **53**, 6882–6887 (2014).
- Li, X., Wang, L., Wang, X. & Roseman, S. The chitin catabolic cascade in the marine bacterium *Vibrio cholerae*: Characterization of a unique chitin oligosaccharide deacetylase. *Glycobiology*. **17**, 1377–1387 (2007).
- Kafetzopoulos, D., Martinou, A. & Bouriotis, V. Bioconversion of chitin to chitosan: purification and characterization of chitin deacetylase from *Mucor rouxii*. *Proc. Natl. Acad. Sci. USA* **90**, 2564–2568 (1993).
- Alfonso, C., Nuero, O. M., Santamaria, F. & Reyes, F. Purification of a heat-stable chitin deacetylase from *Aspergillus nidulans* and its role in cell wall degradation. *Curr. Microbiol.* **30**, 49–54 (1995).
- Liu, Z. *et al.* Structure and function of a broad-specificity chitin deacetylase from *Aspergillus nidulans* FGSC A4. *Sci Rep.* **7**, 1746, <https://doi.org/10.1038/s41598-017-02043-1> (2017).
- Tsigos, I. & Bouriotis, V. Purification and characterization of chitin deacetylase from *Colletotrichum lindemuthianum*. *J. Biol. Chem.* **270**, 26286–26291 (1995).

34. Blair, D. E. *et al.* Structure and mechanism of chitin deacetylase from the fungal pathogen *Colletotrichum lindemuthianum*. *Biochemistry* **45**, 9416–9426 (2006).
35. Naqvi, D. *et al.* A recombinant fungal chitin deacetylase produces fully defined chitosan oligomers with novel patterns of acetylation. *Appl. Environ. Microbiol.* **82**, 6645–6655 (2016).
36. Hofbach, J. *et al.* A chitin deacetylase of *Podospora anserina* has two functional chitin binding domains and a unique mode of action. *Carbohydr. Polym.* <https://doi.org/10.1016/j.carbpol.2017.11.015> (2017).
37. Bird, A. F. & Bird, J. *The Structure of Nematodes*. 2<sup>nd</sup> edition (Academic Press Inc, 1991).
38. Suetake, T. *et al.* Chitin-binding proteins in invertebrates and plants comprise a common chitin-binding structural motif. *J. Biol. Chem.* **275**, 17929–17932 (2000).
39. Hamer, S. N. *et al.* Enzymatic production of defined chitosan oligomers with a specific pattern of acetylation using a combination of chitin oligosaccharide deacetylases. *Sci Rep.* **5**, 8716, <https://doi.org/10.1038/srep08716> (2015).
40. Cord-Landwehr, S. *et al.* Quantitative mass-spectrometric sequencing of chitosan oligomers revealing cleavage sites of chitosan hydrolases. *Anal. Chem.* **89**, 2893–2900 (2017).
41. Blair, D. E., Schuttelkopf, A. W., MacRae, J. I. & van Aalten, D. M. F. Structure and metal-dependent mechanism of peptidoglycan deacetylase, a streptococcal virulence factor. *Proc. Natl. Acad. Sci.* **102**, 15429–15434 (2005).
42. Davies, G. J., Wilson, K. S. & Henrissat, B. Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J.* **321**, 557–559 (1997).
43. Hirano, T. *et al.* Structure-based analysis of domain function of chitin oligosaccharide deacetylase from *Vibrio parahaemolyticus*. *FEBS Lett.* **589**, 145–151 (2015).
44. Tuving, T. R. *et al.* Structure and function of a CE4 deacetylase isolated from a marine environment. *PLoS ONE* **12**, e0187544, <https://doi.org/10.1371/journal.pone.0187544> (2017).
45. Tsigos, I., Martinou, A., Kafetzopoulos, D. & Bouriotis, V. Chitin deacetylases: New, versatile tools in biotechnology. *Trends Biotechnol.* **18**, 305–312 (2000).
46. Tokuyasu, K., Ohnishi-kameyama, M., Hayashi, K. & Mori, Y. Cloning and expression of chitin deacetylase gene from a Deuteromycete. *Colletotrichum lindemuthianum*. *J. Biosci. Bioeng.* **87**, 418–423 (1999).
47. Wang, Y. *et al.* Cloning of a heat-stable chitin deacetylase gene from *Aspergillus nidulans* and its functional expression in *Escherichia coli*. *Appl. Biochem. Biotechnol.* **162**, 843–854 (2010).
48. Shrestha, B., Blondeau, K., Stevens, W. F. & Hegarat, F. L. Expression of chitin deacetylase from *Colletotrichum lindemuthianum* in *Pichia pastoris*: Purification and characterization. *Protein Expr. Purif.* **38**, 196–204 (2004).
49. Martinou, A., Koutsoulis, D. & Bouriotis, V. Expression, purification, and characterization of a cobalt-activated chitin deacetylase (Cda2p) from *Saccharomyces cerevisiae*. *Protein Expr. Purif.* **24**, 111–116 (2002).
50. Tokuyasu, K., Mitsutomi, M., Yamaguchi, I., Hayashi, K. & Mori, Y. Recognition of chitoooligosaccharides and their N-acetyl groups by putative subsites of chitin deacetylase from a Deuteromycete. *Colletotrichum lindemuthianum*. *Biochemistry*. **39**, 8837–8843 (2000).
51. El Gueddari, N. E. E., Rauchhaus, U., Moerschbacher, B. M. & Deising, H. B. Developmentally regulated conversion of surface-exposed chitin to chitosan in cell walls of plant pathogenic fungi. *New Phytol.* **156**, 103–112 (2002).
52. Lenardon, M. D., Munro, C. A. & Gow, N. A. R. Chitin synthesis and fungal pathogenesis. *Curr. Opin. Microbiol.* **13**, 416–423 (2010).
53. Hekmat, O., Tokuyasu, K. & Withers, S. G. Subsite structure of the endo-type chitin deacetylase from a Deuteromycete, *Colletotrichum lindemuthianum*: an investigation using steady-state kinetic analysis and MS. *Biochem. J.* **374**, 369–380 (2003).
54. Baker, L. G., Specht, C. A. & Lodge, J. K. Cell wall chitosan is necessary for virulence in the opportunistic pathogen *Cryptococcus neoformans*. *Eukaryot. Cell.* **10**, 1264–1268 (2011).
55. Nahar, P., Ghormade, V. & Deshpande, M. V. The extracellular constitutive production of chitin deacetylase in *Metarhizium anisopliae*: Possible edge to entomopathogenic fungi in the biological control of insect pests. *J. Invertebr. Pathol.* **85**, 80–88 (2004).
56. Wharton, D. Nematode egg-shells. *Parasitology* **81**, 447–463 (1980).
57. Esteves, I., Peteira, B., Atkins, S. D., Magan, N. & Kerry, B. Production of extracellular enzymes by different isolates of *Pochonia chlamydosporia*. *Mycol. Res.* **113**, 867–876 (2009).
58. Ferreira, S. R. *et al.* Biological control of *Ascaris suum* eggs by *Pochonia chlamydosporia* fungus. *Vet. Res. Commun.* **35**, 553–558 (2011).
59. Fath, S. *et al.* Multiparameter RNA and Codon Optimization: A standardized tool to assess and enhance autologous mammalian gene expression. *PLoS ONE* **6**, e17596, <https://doi.org/10.1371/journal.pone.0017596> (2011).
60. Eddy, S. R. Profile hidden Markov models. *Bioinformatics*. **14**, 755–763 (1998).
61. Jones, D. T., Taylor, W. R. & Thornton, J. M. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**, 275–282 (1992).
62. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).
63. Söding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **1**, 33, W244–8 (2005).
64. Sali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
65. Fiser, A. & Šali, A. Modeller: generation and refinement of homology-based protein structure models. *Methods Enzymol.* **374**, 461–491 (2003).
66. Shen, M. Y. & Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **15**, 2507–2524 (2006).
67. Trott, O. & Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).
68. Morris, G. M. *et al.* Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comp. Chem.* **16**, 2785–91 (2009).
69. Humphrey, W., Dalke, A. & Schulten, K. VMD - Visual MolecularDynamics. *J. Molec. Graphics.* **14**, 33–38 (1996).

## Acknowledgements

The authors thank Prof. Bruno Moerschbacher and Dr. Stefan Cord-Landwehr, University of Münster, for the sequencing of deacetylated chitoooligosaccharides. This work was supported by the European Commission NANO3BIO project, grant agreement n° 613931 (to A.P.), and grants BFU2016–77427–C2–1–R (to A.P.) and AGL2015–66833–R (to L.L.) from MINECO, Spain. Pre-doctoral contracts are acknowledged to Generalitat Valenciana (to A.A.), Generalitat de Catalunya (to H.A.), and European Commission NANO3BIO project (to L.G.).

## Author Contributions

L.L., and A.P. designed and supervised the work, A.A., L.G., and H.A. performed the experiments and analysed data, E.S. and X.B. carried out the bioinformatics and modelling work. A.A., L.L., X.B., and A.P. wrote the paper with critical inputs from all other authors. All authors reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-19902-0>.

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018