An ancient lysozyme in placozoans

Henry Berndt[1], Igor Duarte[1], Urska Repnik[2], Michel Struwe[1], Mohammad Abukhalaf[3], Axel Scheidig[1], Andreas Tholey[3], Harald Gruber-Vodicka[1,*] & Matthias Leippe[1,*]

1 Zoological Institute, Kiel University, Kiel, Germany

2 Central Microscopy, Kiel University, Kiel Germany

3 Proteomics & Bioanalytics, Institute for Experimental Medicine, Kiel University, Kiel, Germany

Lysozymes are an essential part of nutrition and antibacterial immunity in metazoans, executing the breakdown of bacterial cell walls via the hydrolysis of peptidoglycan. Although various lysozymes have been reported for several bilaterian phyla, the origin of metazoan lysozymes remains elusive as they seem to be lacking in non-bilaterian animals. In this study, we investigated a putative goose-type lysozyme (PLys, glycoside hydrolase family 23, GH23) of the placozoan *Trichoplax* sp. H2 which we localized to gland cells of the ventral epithelium. N-terminal of the conserved GH23 lysozyme domain, PLys contains a non-conserved cysteine-rich domain. We could show a truncation of this N-terminal domain in the maturation process of PLys and a drastic increase in enzymatic activity at the cost of stability using recombinantly expressed physiological proteoforms of PLys. Phylogenetic analysis of GH23 lysozymes from all domains of life revealed a monophyletic radiation in animals. Based on structural comparisons and their distribution in the animal tree of life, metazoan g-type GH23 lysozymes appear to have originated from a horizontal gene transfer event from bacteria to an early pre-bilaterian ancestor. GH23 lysozymes have then been retained and expanded in many phyla, including Porifera, Cnidaria, Placozoa and chordates, acting as key component in the antibacterial arsenal since early metazoan evolution.

*: Corresponding authors e-mail addresses – mleippe@zoologie.uni-kiel.de , hgrubervodicka@zoologie.uni-kiel.de

Main:

Animals are always associated and interacting with bacteria. These interactions can range from beneficial mutualism between host and microbe to one side taking advantage, e. g. bacterial pathogens or bacterivorous animals[1]. For killing and digestion of bacteria, animals feature a variety of molecular processes and effector proteins. A key group of such antibacterial proteins are lysozymes, glycoside hydrolases (GHs) that attack the peptidoglycan layer by cleaving the β-1,4-glycosidic bond between its main building blocks N-acetylmuramic acid (MurNAc) and N-acetylglucosamine (GlcNAc). Lysozymes have been traditionally grouped into different classes, named after the organisms in which they were first discovered[2]. A more consistent naming convention for enzymes with a glycoside hydrolase activity categorizes them into families based on their fold and substrate specificity. For example, goose-type (g-type) lysozymes, first discovered in the egg white of the graylag goose *Anser anser* and found later in many vertebrates, , belong to the glycoside hydrolase family 23 (GH23), together with bacterial lytic transglycosylases[3].

Lysozymes appear to be very unevenly distributed in the animal tree of life, with most early branching clades such as Ctenophora and Cnidaria apparently lacking *bona fide* lysozymes[2]. While protostomian invertebrates as well as early branching deuterostomes such as echinoderms predominantly feature invertebrate-type (i-type) lysozymes of GH22i, chordates primarily possess g-type (GH23) and chicken-type (c-type) lysozymes (GH22c)[2,4]. Sparse reports of c-type lysozymes in Arthropoda and g-type lysozymes in mollusks, however, challenge the evolutionary argument that these two lysozyme families are limited to Chordata[2]. In addition, a short report of a putative g-type in placozoans, called PLys, suggests that the evolution of major lysozyme groups in animals was not a relatively late innovation in the lineages leading to bilaterians and then chordates[5,6]. The existence of a potential g-type lysozyme in some of arguably the simplest metazoan animals points instead to a much earlier occurrence and diversification of lysozymes in the animal tree of life (ATL). The position of Placozoa in the ATL is under ongoing discussion, but recent bioinformatic analyses consistently

place placozoans among the pre-bilaterian lineages and as the sister taxon to the clade formed by Cnidaria and Bilateria[7,8]. With such an interesting position in animal phylogeny, placozoans are increasingly used as a model organism to study the origin of multicellularity, developmental biology and neurophysiology[5,9,10]. As millimeter-sized and disk-shaped grazers of marine biofilms, placozoans continually interact with bacteria, and also harbor endosymbiotic bacteria in a cell-type specific symbiosis[11,12]. The basic body plan of the placozoans is very simple and consists of three layers. The ventral columnar epithelium has a mainly digestive function, whereas a thin dorsal epithelium has primarily a barrier function. A densely packed mesenchyme-like layer consists of fiber-cells, which form many cell-to-cell connections with other cell types in both epithelia[13]. In addition, fiber cells contain endosymbiotic bacteria, and have been highlighted as central immune cells in placozoa, as they are the only cells capable of phagocytosis and apoptosis[11,14]. During feeding, placozoa form a temporary gastric cavity between the ventral epithelium and the substrate, into which gland cells and lipophilic cells secrete digestive enzymes[15]. Comparative single-cell transcriptomics data of diverse placozoan lineages showed *plys* to be almost exclusively transcribed in gland cells of the ventral epithelium, from where PLys could be secreted into the gastric cavity (Fig. S1)[7].

To elucidate the role of PLys in *Trichoplax* sp., we identified the mature natural proteoforms of the lysozyme, produced the two lysozyme variants recombinantly, characterized them *in vitro* and determine their cellular distribution in whole animals. Using evolutionary analyses, we show that the highly active lysozyme PLys represents a broad radiation of G-type lysozymes across the animal tree of life that arguably were among the earliest antibacterial effectors in metazoan evolution.

Results and Discussion:

Identification and localization of native placozoan lysozyme

As a putative g-type lysozyme has been reported for placozoans (PLys)[6] and the homolog is transcribed across placozoan diversity, we screened for the activity of lysozymes using placozoan total protein extract from *Trichoplax* sp. H2. On a lysoplate, both placozoan total protein extracts in two concentrations (0.24 mg/mL and 1.24 mg/mL) produced zones of lysis after 24 h (Fig. 1A). Following this first indication of endogenous lysozyme activity, we validated the identity of the potential lysozymes by mass spectrometry and Western blotting of a reverse-phase chromatography fractionated *Trichoplax* proteome. The Western blot revealed two protein double bands, which correspond to molecular masses of about 30 kDa and 20 kDa (Fig. S2)[6]. We therefore used fractions with the strongest signals in the western blot for mass spectrometry to characterize these two potential proteoforms in a combination of proteomic analyses of intact proteins without digestion (top-down) and of digested proteins (peptide-based; bottom-up). In top-down experiments, the most abundant proteoform was lacking the N-terminus and started at amino acid position S100 (Fig. 1B). Based on the structural model of the full length PLys predicted by ColabFold, the N-terminal domain that is missing in this most abundant protein form consists of two repeated beta-sheets, is stabilized by 6 disulphide bonds and is connected to the active g-type lysozyme domain by a loop. N-terminomics (Bottom-up proteomics) experiments detected peptides specific to this N-terminal domain of the full-length PLys. The N-terminus (S100) of the shorter form of PLys is located in the connective loop region between the N-terminal domain and the lysozyme domain (Fig. 1C), which likely constitutes a cleavage site for specific proteolytic processing during protein maturation (limited proteolysis). We therefore differentiate between the truncated and likely mature PLys (mPLys) and the full-length precursor PLys (pPLys) in the later experiments.

Following the identification of the two forms of PLys in the *Trichoplax* sp. H2, we applied immunohistochemical labeling of PLys for cellular localization of the enzyme. Lysozymes are usually either secreted for extracellular lysis of bacteria or act on phagocytosed

124 bacteria in lysosomes[2]. As fiber cells have been indicated as the only cell type in
125 *Trichoplax* capable of engulfing bacteria[14], we predicted that PLys would be detectable
126 either in lysosomes of fiber cells or in exocytotic vesicles of secretory cells of the ventral
127 epithelium (gland cells or lipophilic cells). To be able to differentiate these cell types by
128 ultrastructural features, we opted for optimal tissue and subcellular preservation using
129 high pressure freezing (HPF) followed by freeze substitution (FS) and resin embedding.
130 The ventral epithelium features several groups of mucus-secreting gland cells, which we
131 visualized using rhodamine-coupled wheat germ agglutinin (WGA) to label mucus and
132 glycocalyces[16]. The signal of the specific PLys antibody was colocalized with WGA-
133 labeled mucus or glycocalyces and largely confined to vesicles. Most of these vesicles
134 were in cells of the ventral epithelium and to minor extent in cells of the interepithelial
135 layer, most likely corresponding to fiber cells (Fig. 2A-D). Our antibody staining is
136 consistent with the single-cell expression data of Najle *et al.* who reported *plys*
137 expression predominantly in gland cells[7]. We observed signals in digestive vacuoles of
138 fiber cells only occasionally, and the single cell transcriptomics data likewise shows no
139 substantial expression of *plys* in fiber cells[7]. It remains unclear, whether fiber cells are
140 able to import gland-cell secreted lysozyme, or if expression of lysozyme in fiber cells is
141 inducible by bacterial components. While the genome of the closely related *Trichoplax*
142 *adhaerens* showed a high variety of scavenger receptors and parts of Toll-like and NOD-
143 like signaling pathways that may detect molecular bacterial patterns, complete signaling
144 pathways from receptor to transcription factor that induce and regulate the synthesis of
145 innate immunity factors such as PLys have yet to be identified[17].
146

147 **PLys proteoforms show high lysozyme activity under acidic conditions**

148 Given the contrasting domain architecture of the two forms of PLys, we wanted to
149 understand the role of the cleaved N-terminal domain that characterizes pPLys. We
150 therefore heterologously expressed both pPLys and mPLys recombinantly, where
151 mPLys consisted of the g-type lysozyme domain starting at amino position 112 (M112)
152 of pPLys (mPLys_M112). A comparison of pPLys and mPLys_M112 lysozyme activity at
153 room temperature against freeze-dried and deactivated *Micrococcus lysodeicticus*

154 revealed high specific activities for both variants compared to hen egg white lysozyme
155 (HEWL) at pH 5.0 (21°C; pPLys = 9 U/pmol; mPLys_M112 = 11 U/pmol; HEWL
156 0.5 U/pmol) (Fig. 3B). Both variants had maximum activity under low ionic strength
157 (<50 mM) and at a pH between 5.0 and 5.5 (Fig. 3C-D). The specific activity of
158 mPLys_M112 at its pH optimum (pH 5.5) was 3-times higher than that of pPLys at its
159 respective optimum (pH 5.0) (Fig. 3E). At 21°C, which is the ambient temperature for our
160 long-term placozoan cultures, and their optimal pH, both proteoforms had a strong
161 bactericidal activity against viable gram-positive *Bacillus subtilis*. However, we found no
162 activity against the tested gram-negative *E. coli* strain (Fig. 3F and S3). The lack of
163 activity against gram-negative bacteria is contrary to the results of PLys by Bam and
164 Nakano[6] but consistent with most metazoan lysozymes being inactive towards gram-
165 negative bacteria[2].
166
167 **pPLys has a much higher optimal temperature and is active at a much broader**
168 **range of temperatures.**
169 Since the N-terminal domain of PLys is stabilized by six disulphide bonds, we
170 hypothesized that enzyme performance could strongly deviate under heat stress. Our
171 screens across a broad temperature range indeed revealed that the specific lysozyme
172 activity of pPLys increased until 60 °C and was still detectable at 70 °C, while
173 mPLys_M112 activity peaked at 30 °C and decreased proportional to the increase in
174 temperature (Fig. 3E). Following the strong difference in heat tolerance, we assessed the
175 denaturing dynamics (melting temperatures) for both variants. mPLys_M112 denatured
176 between 25 °C and 35 °C, whereas pPLys denatured between 50 °C and 60 °C,
177 consistent with their temperature-activity profiles (Fig. S4). Overall, the pPLys N-terminal
178 domain appears to have a strong stabilizing effect under physiologically relevant
179 temperatures, at the cost of activity. G-type lysozymes typically contain two disulphide
180 bonds, whereas the PLys lysozyme domain only contains one[18]. The increased structural
181 flexibility introduced by the loss of a disulphide bond could explain both increased
182 activity at low temperatures due to an induced-fit type of substrate binding, and
183 decreased stability at higher temperatures[19]. In pPLys, the cysteine-rich N-terminal

184 domain could compensate for the loss of the disulphide bond and stabilize the protein
185 until maturation to prevent a fast physiological turn-over. We hypothesize that *Trichoplax*
186 likely uses this trade-off between stability and activity by cleaving the N-terminus from
187 pPLys before secretion. The cysteine-rich repeats (CRRs) which constitute the N-
188 terminus are widely distributed across eukaryotes and were linked to horizontal gene
189 transfer (HGT) events from prokaryotes[20,21]. Those CRRs are abundant in proteins
190 associated with antimicrobial activity, where one or multiple CRRs constitute the N-
191 terminus. Temperature-dependent activity studies on full-length and truncated variants
192 of other proteins with this N-terminal CRR domain structure could elucidate if the
193 stabilizing effect is conserved and promotes the selection of this domain architecture.
194

195 **The puzzle of low pH and low ionic strength for optimal pLys activity**

196 Placozoans only form a temporary gastric cavity (feeding groove) when moving onto their
197 feeding substrate. As we primarily localized the lysozyme to vesicles of the ventral
198 epithelium, mPLys is presumably secreted to attack gram-positive bacteria and to
199 decompose accumulating peptidoglycan in this temporary feeding groove. Due to the
200 strong dependency of PLys activity on acidic conditions, we assume that placozoans
201 can acidify this gastric cavity despite its temporary nature. Previous studies on the
202 feeding behavior of *Trichoplax* sp. indicated gland cells and lipophilic cells as primary
203 secretory cells in the feeding process, in which gland cells largely contribute digestive
204 enzymes and lipophilic cells may contribute to the acidification with a secretory burst of
205 vesicles stainable with acidophilic dyes[15]. Release of the acidic content of those vesicles
206 in vast amounts might facilitate acidification of the gastric cavity, potentially down to the
207 physiologically optimal conditions for PLys activity at pH 5.0 to 5.5. However, how
208 placozoans lower ionic strength to less than 10% of the surrounding seawater medium
209 remains enigmatic. One possible mechanism could involve the mucus, which is secreted
210 by gland cells in the ventral epithelium. Mucins, the main component of metazoan
211 mucus, are complex proteoglycans which often carry negatively charged building blocks
212 such as carboxy groups, sulfate groups or sialic acid residues[22]. It has already been
213 shown that the negatively charged components of mucins can bind sodium ions and

214 their hydrate shells, causing swelling of the mucus and formation of a biogel[23]. Similarly,
215 in placozoans, the nascent mucus in the gastric cavity could potentially take up inorganic
216 cation species from the seawater medium, which is dominated by sodium ions, and lead
217 to a decrease of ionic strength in the aqueous phase inside the gastric cavity. Depending
218 on the pH inside the mucus-loaded exocytotic vesicles of the gland cells, carboxy-
219 groups in the mucus may be initially protonated and lead to proton-sodium exchange
220 between mucus and seawater after secretion, which would also cause a decrease of pH
221 in the gastric cavity. Experimentally, pH and conductivity microsensors with high
222 sensitivity and spatial resolution will possibly reveal the physicochemical dynamics
223 during the feeding process. Chemical analysis of placozoan mucus would be necessary
224 to characterize the molecular function of mucus in these early branching metazoans.
225
226 **Evolution of metazoan g-type lysozymes**
227 Placozoa are one of the earliest branching animal phyla, and their physiologically relevant
228 synthesis of g-type lysozymes with a two-step maturation suggested that related
229 lysozymes may be more widespread across the animal tree of life. To elucidate the
230 evolution of PLys and all other metazoan g-type lysozymes in the GH23 subfamily I, we
231 hence constructed a database of all GH23_I members including a set of closely related
232 bacterial lytic transglycosylases that possess the specific sequence motif of the GH23_I.
233 Amino acid sequences were obtained from the carbohydrate active enzyme (CAZY)
234 database, BLAST searches with PLys against nr, and the manual addition of a curated
235 group of GH23_I orthologs detected in metazoan genome assemblies without previously
236 annotated g-type lysozymes[24] (Fig. 4 and S5). In our phylogenetic analysis, out of the
237 131 metazoan g-type lysozymes included in the dataset, 129 clustered in a single clade
238 (clade support: 99.3) that is monophyletic except for a single bacterial sequence from
239 *Saprospira grandis* (acc. num. WBM73135.1). While the molecular evolution trajectory of
240 most g-type lysozyme orthologues was consistent with the phylogenetic relationships of
241 the respective organisms, we occasionally observed the clustering of orthologues from
242 phylogenetically distant species. This could be due to gene duplication and different
243 evolutionary trajectories of the resulting paralogues or the horizontal gene transfer

244 between different metazoan ancestors. The prevalence of g-type lysozymes in all early
245 branching animals (Porifera, Cnidaria and Placozoa) except Ctenophora suggests that
246 g-type lysozymes originated during early metazoan evolution[25]. The g-type lysozyme
247 appears to be an innovation of the lineage split from ctenophores and supports the
248 ctenophore first hypothesis of early animal evolution[26]. The gene was then lost in most
249 of the Protostomia except mollusks and brachiopods, while it was retained in most
250 chordate lineages (Fig. 5). All early branching animals lack homologues of any other
251 lysozyme subfamily, except a description of an i-type lysozyme in the sponge *Suberites*
252 *domuncula*[25]. In a previous phylogenetic analysis of i-type lysozymes, we showed the
253 *Suberites* lysozyme to be most similar to deuterostomian i-type lysozymes[27]. As all other
254 analyzed Porifera lack i-type lysozymes, this could either be a contamination or a
255 secondary acquisition of the i-type lysozyme by an ancestor of *Suberites*.
256 While all metazoan g-type lysozymes cluster together and diverge from bacterial lytic
257 transglycosylases, the exact relationship between these ancient and very similar protein
258 folds, which show low sequence but a substantial structural conservation, are difficult to
259 resolve from primary structure analyses. We therefore used a tertiary structure-based
260 phylogenetics approach to reconstruct their deeper relationships, and included c-type
261 and i-type lysozymes as well as the most closely related GH-families to elucidate the
262 start of the g-type lysozyme radiation[4,28]. The g-type lysozyme branch separates from all
263 other metazoan lysozyme families and branches within bacterial GH23 clades (Fig. 6 and
264 S6). This is further strengthened from previous structural biology studies, which showed
265 GH23 enzymes to be more similar to the putative common molecular ancestor of
266 chitinases and murein-active-glycosidases than c-type lysozymes (GH22c) and i-type
267 lysozymes (GH22i) in terms of active center organization, reaction mechanism and
268 fold[4,29]. As our sequence- and structure-based phylogenetic analyses indicate, a
269 bacterial lytic transglycosylase could have been horizontally transferred to a metazoan
270 or pre-metazoan ancestor in an ancient event and evolved into the metazoan g-type
271 lysozymes. To reconstruct the evolutionary trajectories of lysozymes in even higher
272 detail, future genomic analyses of early-branching metazoans and metazoan-related
273 holozoans may allow to provide more comprehensive databases and to identify novel

274 glycoside hydrolase orthologues. Moreover, detailed studies on the abundance and
275 mechanisms of horizontal-gene transfer events involving eukaryotes and prokaryotes are
276 certainly needed to account for the impact of horizontal gene transfer on metazoan
277 evolution.
278
279 **Materials and Methods:**
280 ***Trichoplax* sp. H2 culture and collection**
281 Organisms of *Trichoplax* sp. H2 strain "Panama" were cultivated at room temperature in
282 petri dishes in 34 PSU filtered sea water (FSW) and fed with *Rhodomonas* sp. Water was
283 changed every two weeks and every culture was discontinued after six weeks. For
284 protein extraction, individuals were picked and transferred into clean reaction tubes with
285 FSW. After transfer of 50 to 100 individuals, the tube was shortly centrifuged to sediment
286 the animals at the bottom. Following the removal of the supernatant, the pellets were
287 stored at -20 °C until use.
288
289 ***Trichoplax* sp. H2 lysis and protein detection**
290 Sample pellets comprising 500 individuals of *Trichoplax* sp. H2 in total were lysed using
291 a total volume of 500 µL of 8 M urea, 50 mM sodium bicarbonate buffer, 300 mM NaCl,
292 pH 7.4. The pellets were suspended and pipetted up and down until the lysate appeared
293 homogenous. The lysate was centrifuged at 15,000 x *g* at 4 °C for 15 min to remove
294 insoluble debris. The supernatant was diluted with 50 mM sodium bicarbonate buffer,
295 300 mM NaCl, pH 7.4 to a final urea concentration of 1.5 M. Subsequently, the sample
296 was subjected to reversed phase fast protein liquid chromatography (FPLC) using a GE
297 ÄKTA Purifier 100 FPLC system (Amersham Biosciences). The sample was loaded on a
298 Hamilton PRP-3 microbore column and subsequently eluted in a linear gradient from
299 0.1% trifluoroacetic acid (TFA) to 80% acetonitrile (ACN) in 0.1% TFA in 20 column
300 volumes (CV). The eluate was collected in fractions of 1 mL by an autosampler. The
301 fractions were split in two aliquots each and lyophilized. One aliquot of every fraction
302 was subjected to a Western blot analysis.

303 For protein separation using Tricine-SDS-PAGE, 50 µg of extracted protein were loaded
304 onto a polyacrylamide gel consisting of a 5% stacking gel and a 13% separation gel[30].
305 After electrophoresis, proteins were electroblotted onto a PVDF membrane following a
306 standard semi-dry blotting protocol. Successful blotting was verified by staining with
307 Ponceau S. The membrane was washed with Tris-buffered saline (TBS) and TBS with
308 0.05% Tween 20 (TBST) and afterwards blocked using 1% bovine serum albumin (BSA)
309 in TBST. The membrane was afterwards incubated overnight with a custom polyclonal
310 antibody mix (0.88 µg/mL) against PLys, generated in two rabbits (Eurogentec). One
311 rabbit was immunized with the peptide C-TKEQQMQGGVAAYNF (Res. 239-253), the
312 other one with the peptide C-PYSSTHISQATNILI (Res. 211-225) corresponding to two
313 putatively surface-exposed alpha helices in the lysozyme domain of PLys following
314 protein structure prediction. Antibodies were enriched and purified using affinity
315 chromatography after final bleeding of the animals. After incubation with the primary
316 antibodies, the membrane was washed three times with TBST for 15 min each and
317 incubated afterwards with a monoclonal antibody against rabbit IgG coupled to
318 horseradish peroxidase (HRP) (BioCat) for 1 h. Following three washes with TBST for 15
319 min, the membrane was developed using the Amersham ECL Advanced Western
320 Blotting Detection kit (GE Healthcare) and a luminescence detector (GE Healthcare) with
321 an exposure time of 5 s. Based on the results of the Western blot, the remaining aliquots
322 of fractions containing putative proteoforms of PLys were subjected to mass
323 spectrometry.
324
325 **Top-down proteomics**
326 Lyophilized fractions from the reversed-phase FPLC fractionation (A9, A10) were
327 dissolved in 50 µL 3% acetonitrile (ACN), 0.1% trifluoroacetic acid (TFA) and centrifuged
328 at 20,000 g for 25 min at 4°C. 10 µL of the supernatant was either reduced by 5 mM
329 TCEP for 1 hr at 25°C, or analyzed in its non-reduced form by LC-MS. Chromatographic
330 separation was performed on a Dionex U3000 nanoHPLC system equipped with an
331 Accucore C4 column (50 cm × 75 µm, 2.6 µm, 150 Å) coupled online to a Fusion Lumos
332 Tribrid mass spectrometer (Thermo Fisher Scientific) equipped with the High-Field-

333 Asymmetric Waveform Ion Mobility Spectrometry (FAIMS) Pro Interface. The eluents
334 used were: eluent A, 0.05% formic acid (FA); eluent B, 80% ACN, 0.04% FA. The
335 separation was performed over a programmed 90-min run. Initial chromatographic
336 conditions were 4% B for 5 min, then increased to 15% B over 2 min followed by a linear
337 gradient from 15% to 50% B over 60 min then 50 to 90% B over 2 min, and 11 min at
338 90% B. Following this, an inter-run equilibration of the column was achieved by 10
339 minutes at 4% B. A constant flow rate of 300 nl/min was employed. Wash runs were
340 performed between each sample injection. Full scan MS spectra were acquired (800-
341 1300 m/z, resolution 120,000, 4 microscans) and fragment scans of the target ions were
342 collected at CV of -15 or -10 (data dependent scans of 2, resolution 60,000) via CID
343 activation at NCE 25. Target ions were chosen after a survey scan that identified the
344 proteoform. Raw data were analyzed against a database containing a Uniprot *Trichoplax*
345 *sp. H2* proteome (13.02.2023) (12,150 sequences), Uniprot *Rhodomonas* sp. proteins
346 (374 sequences), and common contaminants (cRAP). The search was performed on
347 Proteome DiscovererTM 3.0 using a ProSightPD 4.2 search engine with a High/High
348 crawler Node (Max. RT difference 3 min and a precursor m/z tolerance of 0.5). Single
349 proteoform search node with the sequence of the proteoform was performed with a
350 precursor and fragment tolerances of 100 Da and 10 ppm, respectively.
351
352 **Bottom-up proteomics (N terminomics)**
353 Each 15 µL of the fractions 9 and 10 in 3% ACN, and 0.1% TFA were diluted to 100 µL
354 with TEAB then reduced with 10 mM DTT for 30 min at 56°C. Samples were alkylated
355 with 55 mM CAA at 25°C for 20 min in the dark. A first clean-up was performed using
356 SP3 beads (1:25 w/w protein to beads ratio) and ethanol (EtOH) added to 80% (v/v) with
357 mixing for 15 min at 25°C to induce protein binding. The supernatant was removed and
358 beads were washed three times with 80% EtOH. Dimethyl labelling (reductive
359 demethylation) of N-termini was performed by resuspending protein-bound beads in 200
360 mM HEPES (pH 7.0) and then mixing with formaldehyde (30 mM) and sodium
361 cyanoborohydride (15 mM) at 37°C for 2 h. Formaldehyde (30 mM) and sodium
362 cyanoborohydride (15 mM) was added again to each sample and kept mixing at 37°C

363 overnight. To quench the reaction, Tris buffer (pH 6.8) was added to 0.6 M for 5 h at
364 37°C. A second clean-up was done by adding EtOH to 80% for 15 min at 25°C. Then,
365 the supernatant was removed and beads were washed three times with 80% EtOH.
366 Beads were spun down to remove any remaining liquid and then reconstituted in 200
367 mM HEPES (pH 8). Digestion was performed with trypsin (1:50 enzyme to protein ratio)
368 overnight at 37°. Supernatant containing peptides was vacuum dried, then resuspended
369 in 3% ACN, 0.1% TFA and analysed by LC-MS. Chromatographic separation was
370 performed on a Dionex U3000 nanoHPLC system equipped with an Acclaim pepmap100
371 C18 column (2 µm particle size, 75 µm × 500 mm) coupled online to a mass
372 spectrometer. The same eluents as for the Top-down experiment were used. The
373 2separation was performed over a programmed 122-min run. After isocratic elution at
374 4% B for 2 min, a linear gradient from 4% to 50% B over 90 min then 50 to 90% B over
375 5 min, and 10 min at 90% B was utilized. A constant flow rate of 300 nl/min was
376 employed. Wash runs were performed between each sample injection. Data acquisition
377 following separation was performed on a QExactive Plus (Thermo). Full scan MS spectra
378 were acquired (300-1500 m/z, resolution 70,000) and subsequent data-dependent
379 MS/MS scans were collected for the 10 most intense ions (Top10) via HCD activation at
380 NCE 27 (resolution 17,500). Dynamic exclusion (20 s duration) and a lock mass
381 (445.120025) was enabled. Raw data were analyzed against a database containing a
382 Uniprot *Trichoplax sp. H2* proteome (13.02.2023) (12,150 sequences), Uniprot
383 *Rhodomonas* sp. proteins (374 sequences) and common contaminants (cRAP). The
384 search was performed on Proteome DiscovererTM 2.5 using a SequestHT search engine
385 with 10 ppm and 0.02 Da precursor and fragment ions tolerances, respectively. Digestion
386 with Trypsin_R (semi) with a max of two missed cleavages. Oxidation (15.995 Da) of
387 methionine and dimethylation (28.031 Da) and methylation (14.016 Da) at the peptide N-
388 terminus were set as dynamic modifications. Carbamidomethylation (57.02146 Da) on
389 cysteine and dimethylation on lysine was set as a static modification. Strict parsimony
390 criteria have been applied filtering peptides and proteins at 1% false discovery rate
391 estimation.
392

393 **Lysoplate assay**

394 For detection of lysozyme activity, 50 *Trichoplax* sp. H2 organisms were lysed with
395 500 µL of 50 mM MES buffer, 300 mM NaCl, 1% Triton X-100, pH 5.5. Following lysis,
396 the lysate was centrifuged for at 4 °C at 12,000 x *g* for 10 min to remove insoluble debris.
397 Protein concentration was determined using a Micro BCA assay (Thermo Fisher
398 Scientific). The lysate was concentrated using 10 kDa molecular weight cut-off (MWCO)
399 ultrafiltration inlets (Merck) and 10 µL were loaded onto a lysoplate with 20 mM sodium
400 acetate, pH 5.0, into 3 mm wells[31]. The raw extract was loaded onto the plate as well.
401 The lysoplate was incubated at room temperature for 24 h. Ten µL of 2 mg/mL hen egg
402 white lysozyme (Sigma-Aldrich) was used as positive control, lysis buffer as negative
403 control.

404

405 **Immunofluorescence of *Trichoplax* sp. H2**

406 Trichoplax sp. H2 individuals were cryofixed using an EM ICE high pressure freezer
407 (Leica) as described previously[11]. The animals were first transferred into planchets,
408 covered with hexadecane as filler, and subsequently subjected to high pressure freezing.
409 The vitrified samples were subjected to freeze substitution using a freeze substitution
410 (FS) machine (Leica), following an adapted protocol from Mayorova et al.[16] Briefly,
411 samples were freeze substituted in acetone with 0.5% uranyl acetate at -90 °C for 8 h,
412 followed by incubation at -60 °C for 8 h and at -40 °C for 8 h. After washing with ethanol,
413 samples were progressively infiltrated with medium hardness LOWICRYL HM20 resin at
414 -40 °C (33%, 66%, 100%, each for 2 h), followed by 100% resin overnight. Samples
415 were transferred to fresh resin in 0.2 ml PCR tubes and polymerized under indirect UV-
416 radiation at -40 °C for 48 h, followed by direct UV irradiation for 12 h at room temperature
417 in a nail polish lamp. The hardened samples were remounted and sectioned using a
418 Leica UC7 ultramicrotome and a histo diamond knife (Diatome). Transverse 300 nm thin
419 sections were transferred onto water drops on glow-discharged cover slips and heated
420 until the water evaporated. Prior to immunofluorescence labelling, the sections were
421 etched with 10% sodium ethoxide in ethanol for 5 min to increase antigen accessibility[32]
422 For immunofluorescence labelling, the cover slips were first blocked with 0.5% fish skin

423 gelatin (Sigma-Aldrich) in phosphate-buffered saline (PBS) for 30 min, followed by
424 overnight incubation with the primary antibody against PLys in a concentration of
425 25 µg/mL in blocking solution at 4 °C in a wet chamber. A negative control was incubated
426 in the blocking solution only. Afterwards, the cover slips were washed with multiple
427 changes of PBS and incubated with a FITC-coupled anti-Rabbit IgG, diluted 1:100 in
428 blocking buffer, together with Rhodamine-coupled wheat germ agglutinin (WGA), diluted
429 1:300, at room temperature in a wet chamber for 3 h. Subsequently, the cover slips were
430 washed with several changes of PBS and incubated in 1 µg/ml DAPI in ddH_2O, for
431 10 min. The cover slips were then washed in two quick changes of ddH_2O. Excessive
432 water was removed with a filter paper and the cover slips were mounted onto an 8-µL
433 drop of Mowiol (Carl Roth). The cover slips were stored for one day at room temperature
434 at a dark place to ensure hardening of Mowiol and subsequently stored at 4 °C until
435 image acquisition. Imaging was conducted with an LSM 900 with Airyscan confocal
436 microscope (Carl Zeiss) using a Plan-Apochromat 40x/1.4 Oil DIC M27 objective and
437 three laser excitation wavelengths (405/488/561). Images were acquired in the Airyscan
438 Multiplex mode with tiling and subsequently Airyscan processed and stitched using the
439 Zeiss ZEN v. 3.2 software.
440

441 **Recombinant protein production and purification**

442 cDNA of mPLys_M112 and pPLys, without a signal peptide predicted using SignalP v5.0,
443 was synthesized and inserted into pET-30-a(+) plasmids, codon optimized for *E. coli*
444 (BioCat)[33]. The sequence of mPLys_M112 was generated from Uniprot entry B3RIZ1,
445 while the sequence of pPLys was obtained from Uniprot entry A0A369RX22. The amino
446 acid sequence of the pPLys fusion protein (pPlys-FP) was
447 MHHHHHHSSGLVPRGSGMKETAAAKFERQHMDSPDLGTDDDDKAMKSVILLSAFVAVA
448 FAALNDDCNAGQGSCQYDSQCYTGTPASGLCPYDPSNVKCCPKVGQSCKSSTGNCML
449 TTHCGGTTYSGYCPGPSSVRCCVSSSGSGGSYPTKYGDFMRINPSGASSATARQDGLS
450 YSGVAASNKLASNDYNRCLRYKSQFQSASSSTQIPVGLITAIASRESRCGGALDSNGYGD
451 HGNGYGLMQVDKRYHSLQGGPYSSTHISQATNILISSINGVANNHRSWTKEQQMQGGV

452 AAYNFGVGNVQSIGGMDIGTTGNDYSNDVVARAQWFHNNGFN. The amino acid
453 sequence of the mPLys_M112-FP was
454 MHHHHHHSSGLVPRGSGMKETAAAKFERQHMDSPDLGTDDDDKAMGRINPSGASSAT
455 ARQDGLSYSGVAASNKLASNDYNRCLRYKSQFQSASSSTQIPVGLITAIASRESRCGGAL
456 DSNGYGDHGNGYGLMQVDKRYHSLQGGPYSSTHISQATNILISSINGVANNHRSWTKEQ
457 QMQGGVAAYNFGVSNVQSIGGMDIGTTGNDYSNDVVARAQWFHNNGFN.
458 Besides the lacking N-terminal domain, mPLys_M112 differed from pPLys only in a
459 single amino acid substitution from glycine to serine in amino acid position 256. As this
460 residue is located inside a loop, oriented towards the solvent, and does not form the
461 conserved substrate binding site or active site in the predicted protein structure, we
462 considered this amino acid substitution not to affect either lysozyme activity or protein
463 stability. A chemically competent *E. coli* Shuffle *LysY* strain, which is engineered for
464 cytoplasmic expression of disulphide-bond containing proteins, was transformed with
465 the respective plasmids and grown in LB medium (50 µg/mL Kanamycin) at 37 °C and
466 150 rpm overnight. The overnight culture was split into 500 µL aliquots and mixed 1:1
467 (V/V) with glycerol. The resulting glycerol stocks were stored at -80 °C until further use.
468 Prior to recombinant protein expression, 100 mL LB medium was inoculated with 100 µL
469 glycerol stock and grown over night as described above. At the next day, 1 L TB medium
470 (50 µg/mL kanamycin) was inoculated with overnight culture to an OD600 of 0.1. When
471 the culture reached an OD600 of 0.4 to 0.6, protein expression was induced using 0.4
472 mM IPTG. For protein overexpression, the culture was incubated at 18 °C and 150 rpm
473 for 16 h. Bacteria were harvested by centrifuging at 14,500 x *g* at 4 °C for 30 min. The
474 resulting pellets were resuspended in 50 mM sodium phosphate buffer, 300 mM NaCl,
475 pH 7.4. Bacterial cells were lysed using an MS-73 sonotrode (Bandelin) following the
476 manufacturer's instructions. The lysate was cleared by centrifuging at 35,000 *x g* at 4 °C
477 for 30 min. The pellet was resuspended in buffer and the lysing and centrifuging step
478 was repeated. The resulting supernatants were collected and subjected to immobilized
479 metal-ion affinity chromatography using HiTrap Talon Crude columns with a CV of 5 mL
480 (Cytiva) and a peristaltic pump (Pharmacia). After sample loading, the column was
481 washed with 10 CV sample buffer. Overexpressed fusion protein was eluted with sample

482 buffer containing 500 mM imidazole. All further chromatographic purification steps were
483 performed on an ÄKTA Purifier FPLC. The eluted fusion proteins were desalted into a 20
484 mM Tris-HCl buffer containing 50 mM NaCl, 20 % glycerol (w/V) and 2 mM $CaCl_2$, pH
485 6.8, using a HiPrep 26/10 desalting column (Cytiva). Enterokinase, light chain (New
486 England Biolabs) was added in a ratio of 15 ng Enterokinase per milligram fusion protein.
487 After two days of incubation at room temperature, the cleaved protein of interest (POI)
488 was purified using size exclusion chromatography (SEC). For SEC, a Superdex PG 75
489 16/600 column (Cytiva) was used. After sample loading, the sample was eluted with a
490 20 mM MES buffer containing 300 mM NaCl and 20% glycerol, pH 5.5, using a flow rate
491 of 0.75 mL/min. Eluted proteins were collected in 1-mL fractions. The purified proteins
492 were snap frozen in liquid nitrogen and stored at -80 °C until use. Protein concentration
493 was determined measuring the absorbance at 280 nm and applying the calculated
494 extinction coefficient at this wavelength.
495
496 **Lysozyme assay**
497 For quantification of lysozyme activity, a turbidimetric assay after Shugar was used.[34]
498 Lyophilized *Micrococcus lysodeikticus* ATCC No. 4698 (Sigma-Aldrich) (0.2 mg/mL)
499 were suspended in reaction buffer. For standard measurements a 20 mM sodium acetate
500 buffer with varying pH values was used. Different ionic strengths were adjusted with
501 NaCl. To determine the optimal pH value for lysozyme activity, a modified buffer after
502 Ellis was used, containing 5 mM sodium carbonate, 5 mM 2-amino-2-methylpropan-1,3-
503 diol, 5 mM sodium dihydrogen phosphate and 5 mM citric acid[35]. For every
504 measurement, 20 µL of lysozyme solution were added to 730 µL *Micrococcus*
505 suspension. The decrease in extinction at a wavelength of 450 nm was measured
506 photometrically. For measurements at room temperature, an EasySpec UV-Vis
507 spectrophotometer (Safas) was used. Measurements at different temperatures were
508 performed with an UV2600i dual-beam absorption spectrophotometer equipped with a
509 CPS-100 temperature-controlled cuvette holder (Shimadzu). One U of lysozyme activity
510 was defined as a decrease of extinction at 450 nm of 0.001per minute[34].
511

512 **Bacterial lysis assay**

513 For assaying the lysing potential of pPLys and mPLys_M112, a membrane-
514 permeabilization assay was used, described in detail by Bruhn et al.[36] . *Escherichia coli*
515 K12 D31 YpCr[37] and *Bacillus subtilis* (ATCC 6051) were used as target cells in this assay.
516 In brief, $1 \cdot 10^5$ CFU/well bacteria in mid-logarithmic growth phase were added to a
517 dilution series of the protein of interest in a black flat-bottom 96-well plate (Sarstedt),
518 together with the DNA binding dye SYTOX Green (Invitrogen). Fluorescence was
519 measured with a Tecan Infinite M200 plate reader (Tecan) measuring emission at
520 538 nm, excited at 495 nm. Relative lysis was calculated by setting the offset
521 fluorescence of unlysed bacteria as 0% and the fluorescence of totally lysed bacteria as
522 100%. For *E. coli*, total lysis was correlated to the fluorescence caused by 625 nM
523 Melittin, for *B. subtilis*, total lysis was interpreted as the maximum fluorescence
524 generated by mPLys_M112, as the fluorescence after Melittin treatment was drastically
525 below the fluorescence plateau of the lysozyme treatment (Fig. S3). If possible, sigmoidal
526 curves were fitted to the data in ORIGIN v2023 using the Boltzmann function.

527

528 **Thermal shift assay**

529 Stability of recombinant mPLys_M112 and pPLys in different buffers from the Durham
530 pH Screen (Molecular Dimensions #MD1-101) was assessed by thermal shift assays[38].
531 Four µL of 5000x SYPRO Orange dye (Thermo Fisher) were added to 1 mL of protein
532 solution in 100 mM sodium phosphate, 500 mM NaCl, pH 6.5. Ten µL of this mixture
533 were distributed into a 96-well PCR plate (Sarstedt). Ten µL buffer concentrate was
534 added and the plate sealed with a transparent foil (Carl Roth). Melting curves were
535 recorded in an ABI7300 qPCR thermocycler (Thermo Fisher) through the TAMRA
536 channel. The plate was first equilibrated at 21 °C for 2 min and then the fluorescence
537 was measured for 1 min. Subsequently, the temperature was increased in 1-°C
538 increments (0.25 min equilibration followed by 1 min measurement) up to 96 °C. Curves
539 were inspected visually and then processed using *Excel* 16.69.1 (Microsoft) and *Prism*
540 9.4.1 (Graphpad) as previously described[38].

541

542 **Phylogenetic analyses of g-type lysozyme amino acid sequences**

543 The evolutionary relationships between PLys and other members of the GH23 subfamily
544 I were inferred through Maximum Likelihood (ML) phylogenetics, using the amino acid
545 alignment of the conserved g-type lysozyme domain. From the CAZY database[39], we
546 obtained the list of all 120,967 GH23 annotated proteins
547 (http://www.cazy.org/IMG/cazy_data/GH23.txt; accessed on the 1st of September
548 2023), and used the accession numbers to retrieve public sequences from the NCBI
549 protein database. The dataset was filtered to keep sequences with a length between 100
550 and 1,000 residues and that contained both the motives Glu-Ser and Glu-Leu-Met-Gln
551 that are conserved in GH23 subfamily I. Further downsampling was carried out in two
552 steps, first by randomly selecting one sequence from each host species, and then by
553 selecting one sequence per cluster at a 50% identity threshold using CD-HIT v. 4.8.1[40].
554 Additional metazoan g-type lysozyme sequences were added through rounds of
555 BLASTp searches against the non-redundant sequences database (nr) to improve the
556 representation of the group. Finally, lysozyme sequences from Porifera and Cnidaria
557 from a g-type lysozyme orthogroup[24] were manually added to the dataset. A chitosanase
558 sequence from *Bacillus subtilis* was included in the dataset as outgroup (acc. num.
559 QHE14361.1). Multiple sequence alignment was carried out with Muscle v.5.1[41] and the
560 resulting alignment was trimmed to keep residues only that aligned against the g-type
561 lysozyme conserved domain which was determined for PLys sequence through the
562 InterProScan web application[42]. The ML tree of the resulting alignment was calculated
563 with IQ-TREE v.2.0.3[43], based on the general variable time matrix (VT), a FreeRate
564 heterogeneity model with 10 categories (+R10) as selected by ModelFinder[44] and using
565 the SH-aLRT test with 1,000 bootstrap replicates for branch support. Tree representation
566 was obtained with iTOL[45]. To account for the simplification generated by the
567 downsampling steps, all eukaryotic sequences removed from the initial dataset, which
568 represented 0.1% of the complete dataset, were placed in the tree following an
569 Evolutionary Placement Analysis (EPA) approach using MAFFT v.7.520[46] to add the
570 sequences to the alignment and epa-ng v.0.3.8[47] to build the final tree.
571

572 **Protein structure prediction and structure-based phylogenetics of the GH**
573 **lysozyme superfamily**
574 For protein structure prediction, we used ColabFold v1.5.5 which combines MMseqs2
575 with Alphafold 2, as previously described[48]. Out of five generated models, we chose the
576 highest ranked model as predicted protein structure. Chimera v1.16 was used for
577 molecular visualization[49]. Crystal structures of proteins from all GH families within the
578 lysozyme superfamily were obtained from PDB database and included in an unrooted
579 ML tree with the predicted protein structure of PLys. The software FoldTree[28] was used
580 to both generate pairwise alignments between the three-dimensional structures and to
581 derive an ML tree based on a local structural alphabets approach. We randomly
582 subsampled the generated structure models three times and calculated a tree for every
583 subsample. One tree is presented in the main text, the additional two trees in Fig. S7.
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601

References

1. McFall-Ngai, M. *et al.* Animals in a bacterial world, a new imperative for the life sciences. *Proc. Natl. Acad. Sci.* **110**, 3229–3236 (2013).
2. Callewaert, L. & Michiels, C. W. Lysozymes in the animal kingdom. *J. Biosci.* **35**, 127–160 (2010).
3. Blackburn, N. T. & Clarke, A. J. Identification of Four Families of Peptidoglycan Lytic Transglycosylases. *J. Mol. Evol.* **52**, 78–84 (2001).
4. Wohlkönig, A., Huet, J., Looze, Y. & Wintjens, R. Structural relationships in the lysozyme superfamily: Significant evidence for glycoside hydrolase signature motifs. *PLoS ONE* **5**, e15388 (2010).
5. Ringrose, J. H. *et al.* Deep proteome profiling of *Trichoplax adhaerens* reveals remarkable features at the origin of metazoan multicellularity. *Nat. Commun.* **4**, 1408 (2013).
6. M. Bam, & H. Nakano. Lysozyme from placozoa, an early diverging metazoan, lyses both gram-positive and negative bacteria. *JSFS 85th Anniv.-Commem. Int. Symp. Symp. Proc. 06006* (2017).
7. Najle, S. R. *et al.* Stepwise emergence of the neuronal gene expression program in early animal evolution. *Cell* **186**, 4676-4693.e29 (2023).
8. Laumer, C. E. *et al.* Revisiting metazoan phylogeny with genomic sampling of all phyla. *Proc. R. Soc. B Biol. Sci.* **286**, 20190831 (2019).
9. DuBuc, T. Q., Ryan, J. F. & Martindale, M. Q. “Dorsal–Ventral” genes are part of an ancient axial patterning system: Evidence from *Trichoplax adhaerens* (Placozoa). *Mol. Biol. Evol.* **36**, 966–973 (2019).
10. Romanova, D. Y. *et al.* Glycine as a signaling molecule and chemoattractant in *Trichoplax* (Placozoa): insights into the early evolution of neurotransmitters. *NeuroReport* **31**, 490–497 (2020).
11. Gruber-Vodicka, H. R. *et al.* Two intracellular and cell type-specific bacterial symbionts in the placozoan *Trichoplax* H2. *Nat. Microbiol.* **4**, 1465–1474 (2019).
12. Kamm, K., Osigus, H.-J., Stadler, P. F., DeSalle, R. & Schierwater, B. Genome analyses of a placozoan rickettsial endosymbiont show a combination of mutualistic and parasitic traits. *Sci. Rep.* **9**, 17561 (2019).
13. Smith, C. L. *et al.* Novel cell types, neurosecretory cells, and body plan of the early-diverging metazoan *Trichoplax adhaerens*. *Curr. Biol.* **24**, 1565–1572 (2014).
14. Mayorova, T. D. *et al.* Placozoan fiber cells: mediators of innate immunity and participants in wound healing. *Sci. Rep.* **11**, 23343 (2021).
15. Smith, C. L., Pivovarova, N. & Reese, T. S. Coordinated feeding behavior in *Trichoplax*, an animal without synapses. *PLOS ONE* **10**, e0136098 (2015).
16. Mayorova, T. D., Hammar, K., Winters, C. A., Reese, T. S. & Smith, C. L. The ventral epithelium of *Trichoplax adhaerens* deploys in distinct patterns cells that secrete digestive enzymes, mucus or diverse neuropeptides. *Biol. Open* bio.045674 (2019) doi:10.1242/bio.045674.
17. Kamm, K., Schierwater, B. & DeSalle, R. Innate immunity in the simplest animals – placozoans. *BMC Genomics* **20**, 5 (2019).
18. Kawamura, S. *et al.* Role of disulfide bonds in goose-type lysozyme. *FEBS J.* **275**, 2818–2830 (2008).
19. Zavodszky, M. *et al.* Disulfide bond effects on protein stability: Designed variants of *Cucurbita maxima* trypsin inhibitor-V. *Protein Sci.* **10**, 149–160 (2001).
20. Daugavet, M. A. *et al.* Features of a novel protein, rusticalin, from the ascidian *Styela rustica* reveal ancestral horizontal gene transfer event. *Mob. DNA* **10**, 4 (2019).

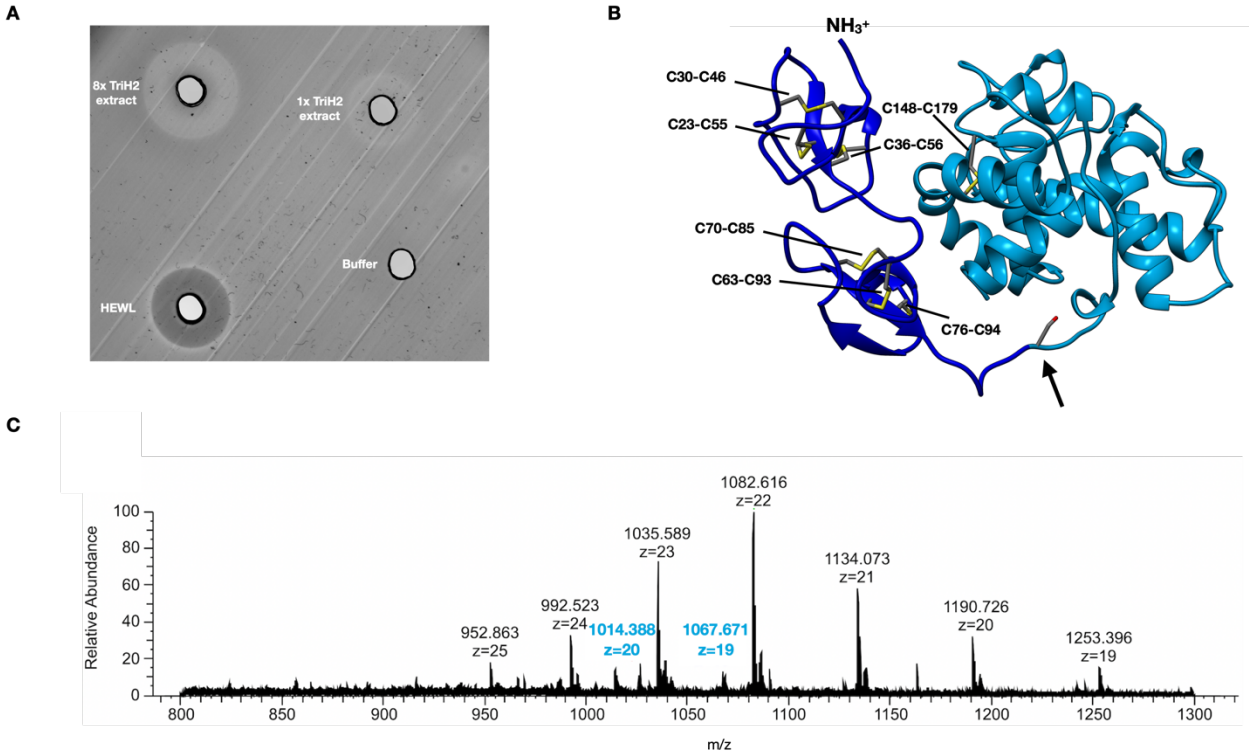
650 21. Daugavet, M. A., Shabelnikov, S. V. & Podgornaya, O. I. Amino acid sequence
651 associated with bacteriophage recombination site helps to reveal genes potentially acquired
652 through horizontal gene transfer. *BMC Bioinformatics* **21**, 305 (2020).
653 22. McShane, A. *et al.* Mucus. *Curr. Biol.* **31**, R938–R945 (2021).
654 23. Weston, A. *et al.* The influence of ions on the lubricative abilities of mucin and the role of
655 sialic acids. *Colloids Surf. B Biointerfaces* **227**, 113327 (2023).
656 24. Laumer, C. E. *et al.* Support for a clade of Placozoa and Cnidaria in genes with minimal
657 compositional bias. *eLife* **7**, e36278 (2018).
658 25. Wiens, M. *et al.* Innate immune defense of the sponge *Suberites domuncula* against
659 bacteria Involves a MyD88-dependent signaling pathway. *J. Biol. Chem.* **280**, 27949–27959
660 (2005).
661 26. Schultz, D. T. *et al.* Ancient gene linkages support ctenophores as sister to other
662 animals. *Nature* **618**, 110–117 (2023).
663 27. Berndt, H. *et al.* Molecular and functional characterization of ILYS-5, a major
664 invertebrate lysozyme of *Caenorhabditis elegans*. *Dev. Comp. Immunol.* **159**, 105220 (2024).
665 28. Moi, D. *et al.* Structural phylogenetics unravels the evolutionary diversification of
666 communication systems in gram-positive bacteria and their viruses. Preprint at
667 https://doi.org/10.1101/2023.09.19.558401 (2023).
668 29. Monzingo, A. F., Marcotte, E. M., Hart, P. J. & Robertas, J. D. Chitinases, chitosanases,
669 and lysozymes can be divided into procaryotic and eucaryotic families sharing a conserved
670 core. *Nat. Struct. Biol.* **3**, 133–140 (1996).
671 30. Schägger, H. & Von Jagow, G. Tricine-sodium dodecyl sulfate-polyacrylamide gel
672 electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.*
673 **166**, 368–379 (1987).
674 31. Osserman, E. F. & Lawlor, D. P. Serum and urinary lysozyme (muramidase) in monocytic
675 and monomyelocytic leukemia. *J. Exp. Med.* **124**, 921–952 (1966).
676 32. Holderith, N., Heredi, J., Kis, V. & Nusser, Z. A High-Resolution Method for Quantitative
677 Molecular Analysis of Functionally Characterized Individual Synapses. *Cell Rep.* **32**, 107968
678 (2020).
679 33. Almagro Armenteros, J. J. *et al.* SignalP 5.0 improves signal peptide predictions using
680 deep neural networks. *Nat. Biotechnol.* **37**, 420–423 (2019).
681 34. Shugar, D. The measurement of lysozyme activity and the ultra-violet inactivation of
682 lysozyme. *Biochim. Biophys. Acta* **8**, 302–309 (1952).
683 35. Ellis, D. A. A new universal buffer system. *Nature* **191**, 1099–1100 (1961).
684 36. Bruhn, H., Riekens, B., Berninghausen, O. & Leippe, M. Amoebapores and NK-lysin,
685 members of a class of structurally distinct antimicrobial and cytolytic peptides from protozoa
686 and mammals: a comparative functional analysis. *Biochem. J.* **375**, 737–744 (2003).
687 37. Boman, H. G. & Monner, D. A. Characterization of lipopolysaccharides from
688 *Escherichia coli* K-12 mutants. *J. Bacteriol.* **121**, 455–464 (1975).
689 38. Huynh, K. & Partch, C. L. Analysis of protein stability and ligand interactions by thermal
690 shift assay. *Curr. Protoc. Protein Sci.* **79**, (2015).
691 39. Drula, E. *et al.* The carbohydrate-active enzyme database: functions and literature.
692 *Nucleic Acids Res.* **50**, D571–D577 (2022).
693 40. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-
694 generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).
695 41. Edgar, R. C. Muscle5: High-accuracy alignment ensembles enable unbiased
696 assessments of sequence homology and phylogeny. *Nat. Commun.* **13**, 6968 (2022).
697 42. Jones, P. *et al.* InterProScan 5: genome-scale protein function classification.

Bioinformatics **30**, 1236–1240 (2014).
43. Nguyen, L.-T., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. IQ-TREE: A fast and effective stochastic algorithm for estimating Maximum-Likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
44. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Von Haeseler, A. & Jermiin, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
45. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).
46. K. Kathoh & M. Frith. Adding unaligned sequences into an existing alignment using MAFFT and LAST. *Bioinformatics* **28**, 3144–3146 (2012).
47. P. Barbera *et al.* EPA-ng: Massively parallel evolutionary placement of genetic sequences. *Syst. Biol.* **68**, 365–369 (2019).
48. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).
49. Pettersen, E. F. *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

Acknowledgements

We thank Ina Kraus-Stojanowic and Antonia Stanislaus for their support during the recombinant expression and characterization of pPLys, Heidrun Ließegang for assistance of placozoan culturing and the Central Microscopy (CAU) for their support with HPF&FS, ultramicrotomy (Ulrike Voigt and Tanja Zabel) and for providing the Zeiss LSM900 confocal microscope (the German Research Foundation INST 257/650-1 FUGG). Part of this work was performed at the Multi-User CryoEM Facility at the Centre for Structural Systems Biology, Hamburg, supported by the Universität Hamburg and DFG grant numbers (INST 152/772-1|152/774-1|152/775-1|152/776-1|152/777-1 FUGG). We also thank Tim Laugks for technical assistance with high pressure freezing and Brandon Seah for valuable suggestions on the manuscript.

735 **Figures**



736

737 **Figure 1: PLys is an active lysozyme with a truncated mature proteoform.**
738 **A**): Lysoplate assay with 8x concentrate of raw *Trichoplax* sp. H2 8x, unconcentrated
739 *T*. sp. H2 extract extract, 20 µg HEWL and lysis buffer. **B)**: Structure model of PLys as
740 ribbon diagram. The backbone of the mature proteoform of PLys is colored in light
741 blue, while the cysteine-rich N-terminal domain is colored in dark blue. Cysteine
742 residues forming disulphide bonds are highlighted and annotated. An arrow marks
S100, the apparent N-terminus of the physiologically mature proteoform.
C): Representative mass spectrum of the top-down proteomics experiment with the
identified mature PLys proteoform with charge states z=19 and z=20.

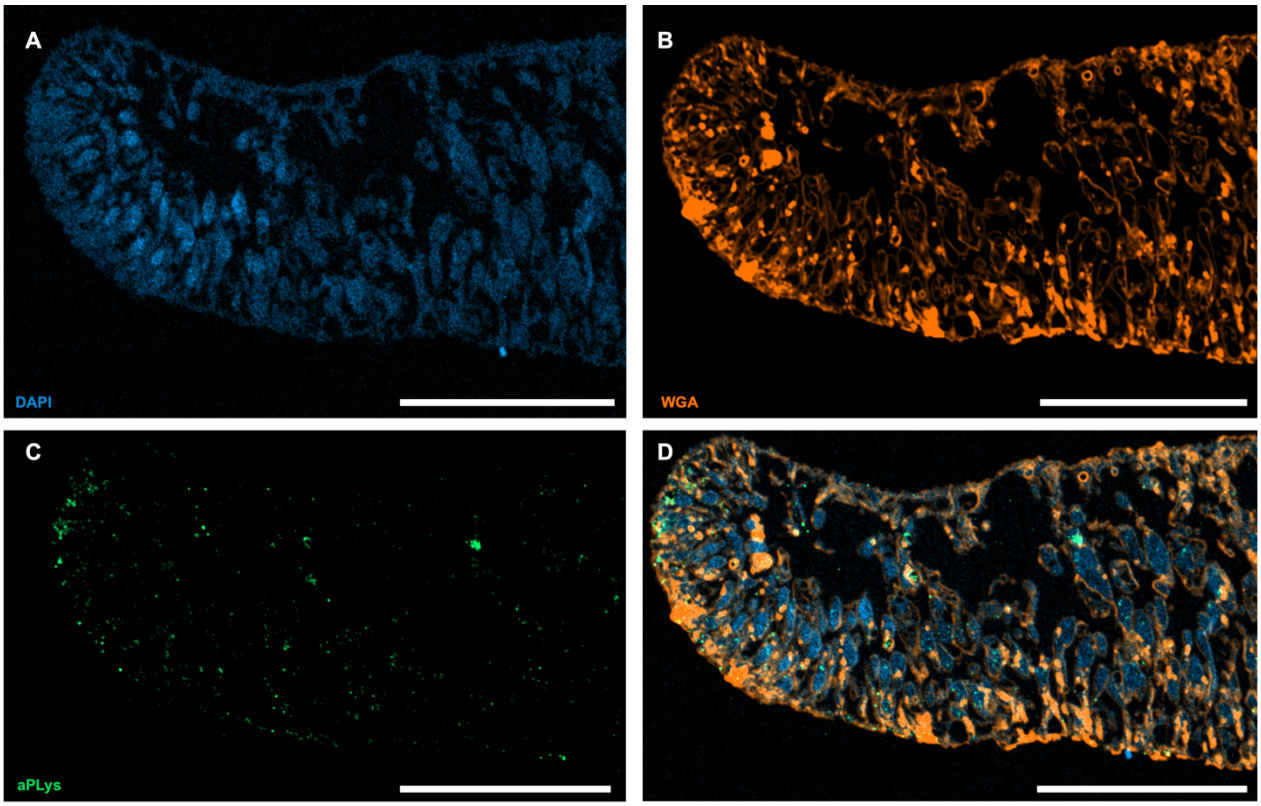


Figure 2: PLys is abundant across placozoan tissues with high local concentrations in peripheral gland cells. Immunofluorescence images of 300 nm sections of high pressure frozen and freeze substituted *Trichoplax* sp. H2 embedded in LOWICRYL HM20. Sections were etched with sodium ethoxide prior to labeling. Signals of different channels are shown in false colours. **A**): DAPI: blue. **B**): antiPLys: green **C**): WGA-Rhodamin: orange. **D**): Merged channels. Scale bars correspond to 20 µm.

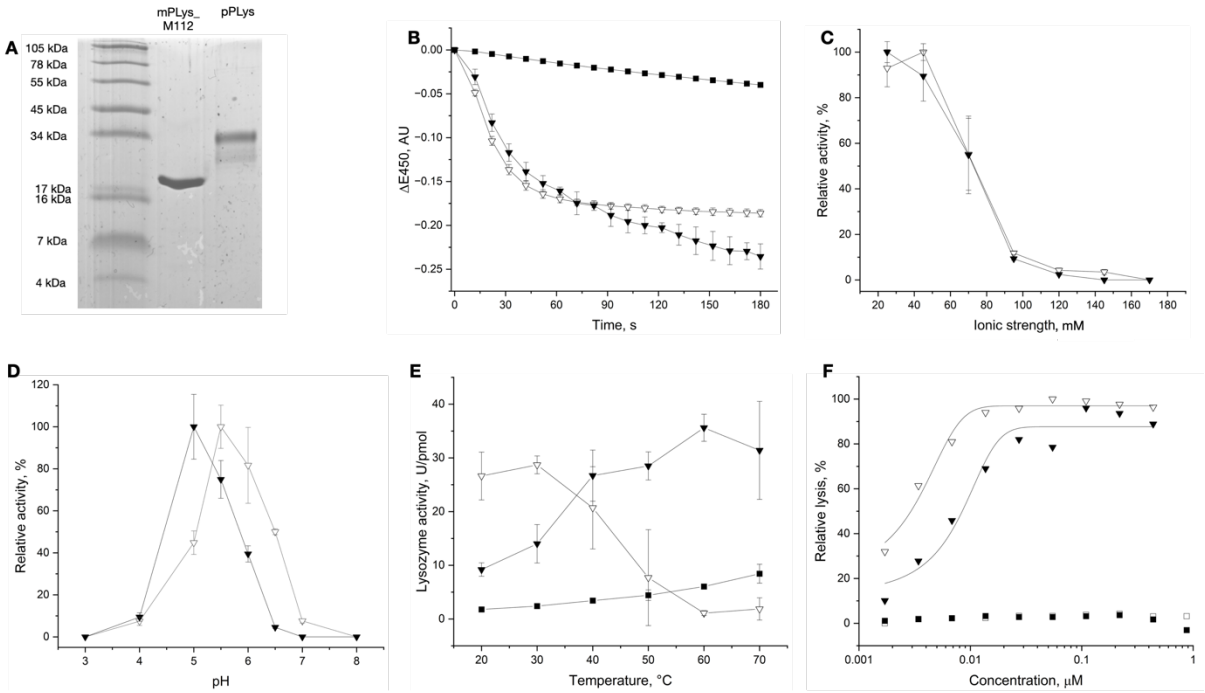


Figure 3: PLys is a potent lysozyme under acidic conditions. A): SDS-PAGE gel of purified recombinant mPLys_M112 and pPLys. The SeeBlue Plus2 protein standard (Invitrogen) was used as molecular mass reference. **B**): Lysozyme assay of 32 pmol pPLys (closed triangles), 32 pmol mPLys_M112 (open triangles) and 32 pmol HEWL (closed squares) at pH 5.0. Reaction buffer consisted of 20 mM sodium acetate, pH 5.0 (I = 13 mM). **C,D**): Dependency of lysozyme activity of pPLys (closed triangles) and mPLys_M112 (open triangles) on ionic strength (C) and pH value (D) at room temperature. **E**): Lysozyme activity of pPLys (closed triangles), mPLys_M112 (open triangles) and HEWL (closed squares) at different temperatures. All measurements were performed in triplicates. **F**): Relative lysis of *B. subtilis* (triangles) and *E. coli* K12 D31 cells (squares) after treatment with pPLys (filled symbols) and mPLys_M112 (open symbols). The depicted values are the average of two replicate measurements.

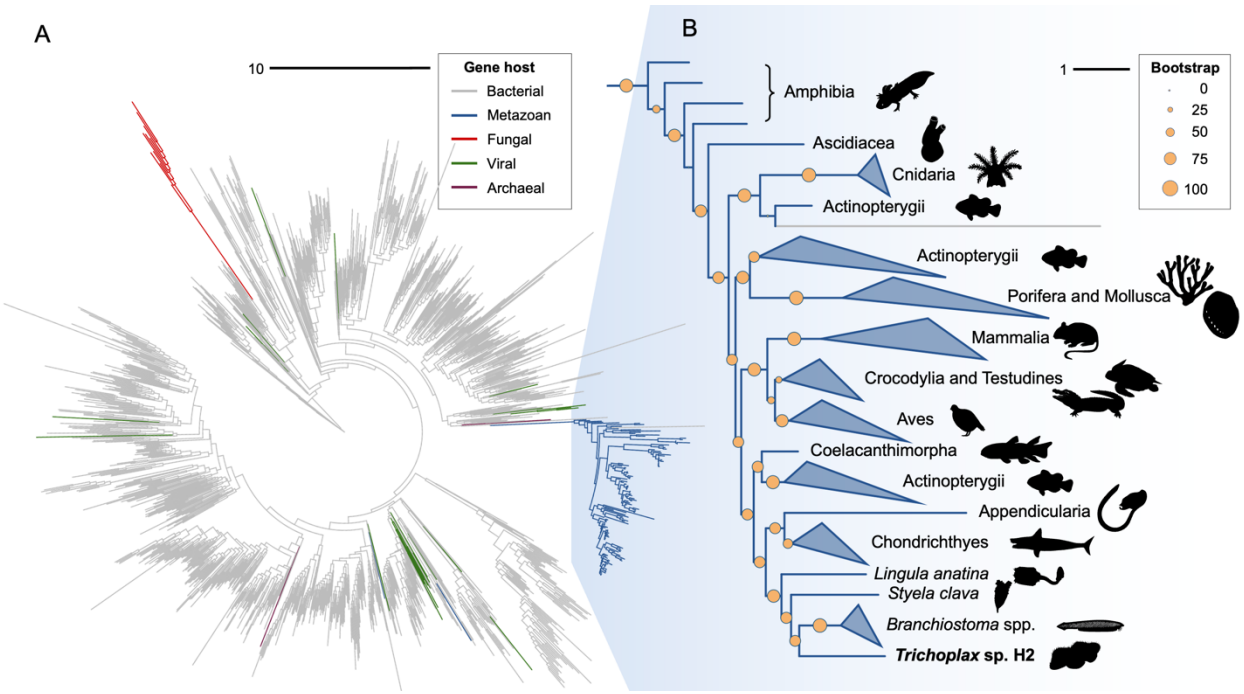


Figure 4: Metazoan g-type lysozymes cluster predominantly into a single clade and are present in six animal phyla. A) Phylogenetic tree of 1410 glycoside hydrolase family 23 sequences from all domains of life and **B**) expanded tree branch containing all metazoan sequences. Colors indicate host groups, orange dots show phylogenetic support (aLRT).

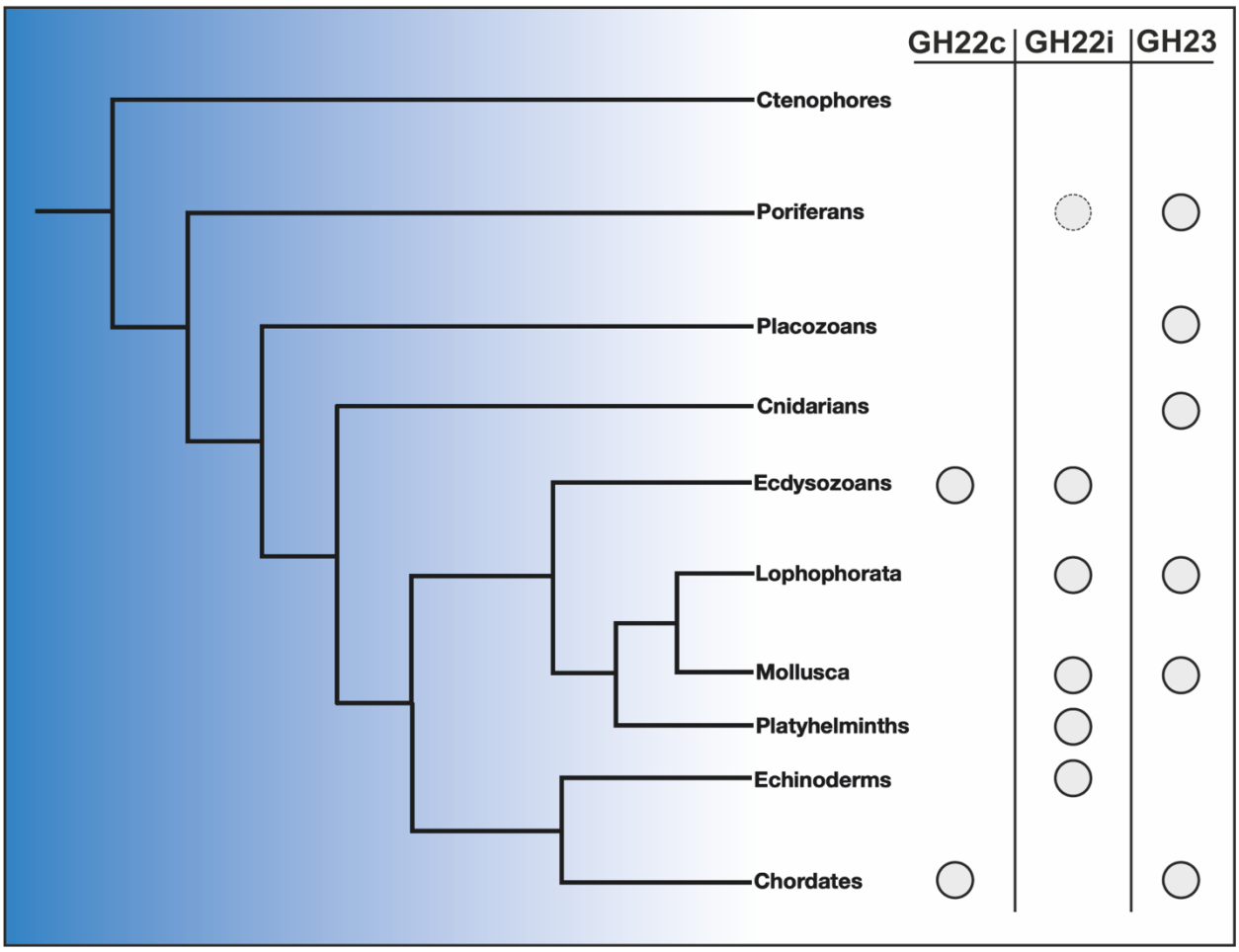


Figure 5: G-type lysozymes are widely distributed over the animal tree of life and the only lysozyme of early all early branching metazoans except Ctenophores. Overview phylogram of metazoans based on Laumer et al.[8], that shows the distribution of the most abundant lysozyme families (closed circles). The status of lysozyme abundance is based on Callewaert and Michiels[2], Berndt et al.[27] and this study. As only one GH22i lysozyme has been reported for poriferans, the respective circle is dotted.

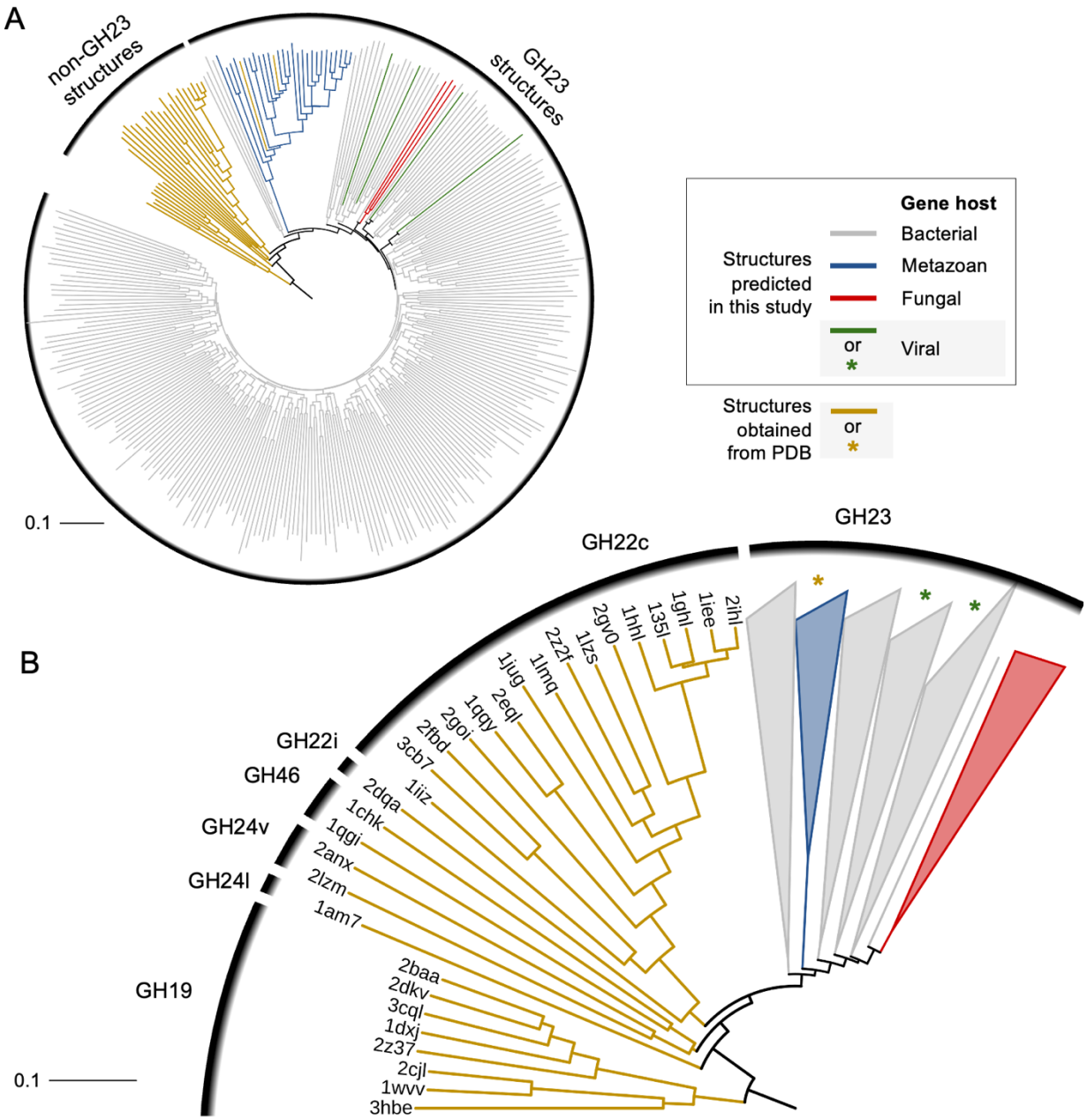


Figure 6: Structure-based phylogeny of GH families within the lysozyme superfamily. In **A**), the whole tree is depicted, in **B**), GH23 clades were collapsed. GH23 structures were randomly sampled from our working dataset (20%). Asteriscs indicate presence of structures from a phage host within the collapsed clade.