CMLS, Cell. Mol. Life Sci. 60 (2003) 2210–2218
1420-682X/03/102210-9
DOI 10.1007/s00018-003-3252-z

CMLS Cellular and Molecular Life Sciences

Research Article

Urochordates carry multiple genes for goose-type lysozyme and no genes for chicken- or invertebrate-type lysozymes

I. W. Nilsen [a, *], B. Myrnes [a], R. B. Edvardsen [b] and D. Chourrout [b]

[a] Marine Biotechnology and Fish Health, Norwegian Institute of Fisheries and Aquaculture Research, Postbox 6122, 9291 Tromsø (Norway), e-mail: inge-w.nilsen@fiskeriforskning.no
[b] Sars Centre for Marine Molecular Biology, Thormøhlensgt. 55, 5020 Bergen (Norway)

Received 29 June 2003; received after revision 24 July 2003; accepted 29 July 2003

Abstract. Genome clones and expressed sequence tags (ESTs) from the ascidian *Ciona intestinalis* and from the larvacean *Oikopleura dioica* were analysed for the presence of lysozyme-encoding genes. Two genes were found to potentially code for goose-type lysozymes in *Oikopleura*, while three or possibly more g-type proteins form the lysozyme complement of *C. intestinalis,* and at least one of these genes from each species is expressed based on EST data. No genes for chicken- or invertebrate-type lysozymes were found in either urochordate species. Consistent with this finding, extracts of *Oikopleura* animals possessed hydrolysing activity on bacterial cell walls, and this activity was not inhibited in the presence of a known inhibitor of chicken-type lysozyme. A wide range of isoelectric points for the predicted lysozymes from *Ciona* (pI 4.4, 6.4 and 9.9) and from *Oikopleura* (pI 5.0 and 8.0) suggests tissue-specific adaptations as well as specific functional roles of the lysozymes. Comparisons of gene structures, encoded sequences, cysteine residue content and their positions in the proteins indicate that the *g*-type lysozymes of *Ciona intestinalis* are more closely related to those of vertebrates than are the *g*-type lysozymes of *Oikopleura.* Multiple genes from each species may result from separate and lineage-specific duplications followed by functional specialisation.

Key words. Urochordate; lysozyme; goose-type; innate immunity.

Introduction

Bacterolytic enzymes such as lysozymes are important molecules produced for the purposes of defence and digestion. Three subclasses of lysozymes have been established in animals. The most extensively studied is the chicken (c-) lysozyme [see reviews in refs 1, 2] found in vertebrates and insects. The two other subclasses of animal lysozymes are the invertebrate (i-) type found in marine and terrestrial invertebrates including arthropods [3–6], and the goose (g-) type so far found only in vertebrates like birds, fish and mammals [7–10]. Multiple *c*-type genes have been reported for both mammals and insects [11–14], probably reflecting tissue-specialised functions and roles through adaptation similar to proposals for the multiple i-type lysozymes in shellfish [15]. Irwin and Gong [10] very recently reported that mammals and zebrafish carry two g-type lysozyme genes and that the pairs of genes are results of early lineage-specific gene duplication. As indicated, vertebrates and insects harbour two lysozyme subclasses, the c- plus g-type and the c- plus i-type, respectively. Whether or why two parallel lysozyme systems are needed is, however, not clear.

Many g-type lysozymes are recognised by the presence of a soluble lytic transglycosylase (SLT) domain of about 90 residues in the C-terminal part of transglycosylases

* Corresponding author.

in bacteria, which is shared by many prokaryotic and phage proteins [16–18]. Lysozymes and SLTs catalyse the hydrolysis of the β-1,4-glycoside bonds between N-acetylmuramic acid (NAM) and N-acetylglucosamine (NAG) in peptidoglycans of bacterial cell walls. Although the similar hydrolysing activity of various lysozymes obviously points to their analogous role in antibacterial defence, the three types of animal lysozymes have no close sequence relatedness. In fact, c-type lysozyme is a close homologue in primary as well as in tertiary structure to α-lactalbumin which possesses lactose synthetase activity produced in lactating mammals [see review in ref. 19] and the two enzymes are traditionally believed to have a common ancestor [20, 21]. Despite low primary sequence relationships among the different types of lysozyme, various theories of common gene precursors for these enzymes and others have been launched. In 1981, Matthews et al. [22] concluded, based on similar backbone conformations, modes of substrate binding, specific protein-substrate interactions and presumed mode of action, that the hen c-type lysozyme and the bacteriophage T4 lysozyme diverged from a common ancestor. This was further elaborated by the finding of a similar goose lysozyme structure, and g-type lysozymes were suggested to represent an evolutionary link between chicken and bacteriophage lysozymes with all three arising from the same predecessor gene [23]. Later, although not supportive of the idea, the authors also discussed a possible convergent evolution of lysozymes [24]. The findings of related overall tertiary structures and a low primary structure relationships between the bacterial SLT and goose lysozyme led to the notion that g-type lysozymes take up a central position in the lysozyme superfamily and that they are structurally closest to the lysozyme ancestors [18]. Holm and Sander [25] again used three-dimensional structural similarities, together with overlapping substrate specificities, to point to an evolutionary connection between plant chitinases and the animal lysozymes including also α-lactalbumin. Only recently were invertebrate lysozymes considered in an evolutionary context, and Bachali et al. [5] proposed, based on a common short domain in the encoded proteins of i-type and c-type lysozymes, that parts of their genes have evolved from the same ancestor exon.

The progress of genomics in a variety of animal phyla allows a consideration of lysozyme evolution through whole-genome sequence examinations which in turn permits conclusions about the presence or absence of a given type of lysozyme. Of special interest are basal chordates such as urochordates and cephalochordates, since their common ancestors with other invertebrates and vertebrates were near the invertebrate-vertebrate transitions. Like many other marine invertebrates, the tunicates feed by filtering seawater, implying exposure to relatively high concentrations of microorganisms such as bacteria and algae. For nutritional and defensive purposes, these animals are therefore expected to have well-developed systems for innate immunity and for hydrolysis of microorganisms. Since the c-type lysozymes are known to be present in both invertebrates (insects) and vertebrates, one would expect them to be present also in urochordates. Predictions about the presence of g- and i-type lysozymes are per se more uncertain since they (so far) are restricted to vertebrates and to invertebrates, respectively. A very recent study on host defence genes expressed in hemocytes of *Ciona intestinalis* did not reveal the involvement in defence by, or even the presence of, any lysozyme genes in this tunicate [26].

Urochordates are divided into three classes: ascidians, having an adult benthic form, and larvaceans and thaliaceans that are pelagic throughout their life. The genome of the ascidian *C. intestinalis* has been almost entirely sequenced [27], although the sequence has not been finally assembled. The larvacean *Oikopleura dioica* has a very short life cycle that permits genetic analysis, and can be permanently cultivated. We have made important progress in sequencing its exceptionally compact genome [28]. Here we report the finding of only g-type lysozymes in both species of urochordate after examining genome data from both species and performing biochemical assays in *O. dioica*.

Materials and methods

Oikopleura resources

Methods for permanent culture of the species have been reported elsewhere [29]. Sequences of *Oikopleura* genes were identified using the TBLASTN program [30] on a dataset of genomic sequences produced by whole-genome shotgun sequencing [28], with the following query sequences: the c-type (accession no. P00698, [8]) and g-type (accession no. P27042, [31]) lysozymes from chicken, and the i-type chlamysin lysozyme from scallop (accession no. CAB63451, [4]). The genomic dataset consisted of 44,797 contigs representing a total of 40.98 megabases of non-redundant sequences. Introns were easily identified as a sequence with the GT-AG consensus ends interrupting the alignments. One of the genes also matched one of 1155 non-redundant contigs of expressed sequence tags (ESTs), obtained through 5′-end sequencing of a directional cDNA prepared at the adult stage. The EST-gene alignment allowed us to confirm the intron sequence. The *Oikopleura* lysozyme gene sequences were submitted to the EMBL Nucleotide Sequence Database and given the following gene names/accession numbers: lysg1/AJ564629 for genomic clone jq4f05, lysg2/AJ564 630 for genomic clone qy3c11, and AJ564631 for EST clone 004h09 of lysg2.

Ciona resources

We also retrieved all lysozyme genes from *C. intestinalis* using the TBLASTN program on the genome sequence recently made available [27] (http://genome.jgi-psf.org/ciona4/ciona4.home.html), using the above-mentioned query sequences. ESTs were also searched using the same program and the same website as well as the *Ciona* cDNA project (http://ghost.zool.kyoto-u.ac.jp/indexr1.html). The coding sequences of lysozyme genes were predicted through EST-gene alignments as well as using gene models also provided together with the genome data. The retrieved lysozyme gene sequences were Ciona1 from gene ci0100144442 (scaffold 194) and from EST BW231984 (identical to BW124765, BW231963), Ciona2 from gene ci0100140238 (scaffold 1222) and Ciona3 from gene ci0100132897 (scaffold 934).

Determination of lysozyme activity in Oikopleura

Bodies of approximately 1000 animals were thawed on ice followed by homogenisation in 1 ml 0.05 M sodium acetate buffer (pH 4.8) using a Downs glass homogeniser equipped with pestle B and finally centrifuged at 14,000 g for 6 min. The resulting supernatant was used as crude extract in enzyme assays. Lysozyme activity was measured as previously described [32] using lyophilised *Micrococcus luteus* (Sigma) substrate at a concentration of 0.2 mg/ml buffer and recording the continuous change in absorbance at 450 nm in the reaction carried out at room temperature. One unit of enzyme activity is defined as the amount of enzyme that catalyses a decrease in absorbance of 0.001 min^{-1}. For determination of pH for optimal activity, assay buffers of 10 mM sodium acetate and 5 mM sodium phosphate were used for pH ranges 3.6–5.6 and 6.0–8.0, respectively, after adjusting the buffer ionic strength to 0.03 using NaCl. The c-type lysozyme inhibitor Ivy was a kind gift from Dr. V. Monchois [33], the hen egg white c-type lysozyme was purchased from Sigma and the scallop i-type chlamysin lysozyme used in this study was isolated earlier [4].

Other sequence analyses

Protein signal peptide predictions were facilitated by the two programs Signal P [34] and PSORT II [35] hosted at http://www.cbs.dtu.dk/services/SignalP and at http://psort.nibb.ac.jp, respectively. Lysozyme sequences from *Oikopleura* and *Ciona* were also compared to the following g-type sequences with accession numbers P00719 (ostrich), P00718 (goose), P00717 (swan), BAB62406 (flounder), AAL08021 (grouper), BAB91437 (carp), Q9D7Q0 (mouse), Q8N1E2 (human1) and AAO32945 (human2). In addition, two genes were obtained from the *Fugu ruibripes* genome database (http://genome.jgi-psf.org/fugu6/fugu6.home.html) [protein id FRUP00000140020 (fugu1), and FRUP00000164529 (fugu2)] and a further two zebrafish genes were identified in sequence ctg9497.1 (zebrafish1) and sequence ctg 10816.1 (zebrafish2) using the Ensembl database (http://www.ensembl.org/Danio_rerio/). Multiple sequence alignments were obtained by running ClustalX [36]. The alignment was manually edited prior to phylogenetic analyses and subsequent phylogenies of protein sequences were estimated by the maximum likelihood (ML; TREE-PUZZLE [37]), neighbour-joining (NJ; ClustalX) and maximum parsimony (MP; PAUP) [38] methods.

Results

Detection of g-type lysozyme genes in urochordates

Protein sequences from three typical lysozyme members were used to tblastn-search the sequenced genome data from *Ciona* and *Oikopleura* for genes capable of encoding c-, i-, or g-type lysozymes. Query sequences of the Icelandic scallop i-type lysozyme [4] and the chicken c-type lysozyme [31] gave no hits of potential lysozyme genes in *Oikopleura* even at promiscuous cut-off values. When the chicken g-type lysozyme [8] was used to search for possible homologues in *Oikopleura,* sequences of two genomic genes and one EST were identified. The EST sequence appeared to be the transcript from one of the two genomic genes. Similar to *Oikopleura*, no genes for lysozymes of the c- or i-types were detected in *Ciona,* but again like *Oikopleura*, analysis of the *Ciona* genome sequences gave evidence for multiple g-type lysozyme genes. Out of three complete or nearly complete *Ciona* genes, one g-type lysozyme gene was also identified by several hits in a separate search for EST sequences. Assuming correctly assembled scaffolds in the *Ciona* sequence database, two or three additional g-type genes may be present in this tunicate genome. We were, however, not able to identify more than parts of their coding sequences and these potential genes and products are thus not further examined here.

Sequence similarities linked all five candidate lysozyme proteins of both species to family 23 of glycosyl hydrolases, and the predicted tunicate proteins LysG1 from *Oikopleura* and the *Ciona* genomic gene model clone ci0100140238 (Ciona2) carry the characteristic SLT domain (Pfam PF01464). The only urochordate lysozyme containing a putative secretion signal is the protein from the *Ciona* EST BW231984 (Ciona1) that is a transcript from the genomic sequence of *Ciona* scaffold 194 (position 115292–116815) and partly covered by the clone ci0100144442. The signal peptide is predicted to comprise the first 20 residues that are immediately followed by a predicted cleavage signal. However, one should keep in mind that two of the five urochordate putative protein sequences do not yet have complete N-termini. Similar signal predictions were run on those fish and terrestrial vertebrate g-type lysozymes that carry known complete

N termini and start with methionine, and we found that all the fish proteins lack an obvious signal for extracellular export. This lack of secretion signal in fish lysozyme *g* is consistent with a previous report for the Japanese flounder [9], which again is in contrast to findings in chicken and mammalian lysozymes [10].

The translation products of *Oikopleura* and *Ciona* lysozyme genes were aligned to known *g*-type lysozymes and substantial similarities among the urochordate and the vertebrate lysozymes were observed (fig. 1). Phylogenetic trees were constructed based on the multisequence alignments. The phylogeny of amino acid sequences of g-type lysozymes estimated by the quartet-puzzling ML method is shown in figure 2. Trees constructed by the NJ and MP methods confirmed the general topology of the ML tree (not shown). The different methods all supported the mammalian, bird and *Oikopleura* lysozymes as monophyletic groups. In addition, the NJ and ML methods divided the fish sequences into two groups, suggesting that all the fish proteins comprise one monophyletic group. The positioning of the *Ciona* sequences in the tree varies greatly between the three methods. The NJ tree places the Ciona1 sequence together with the two *Oikopleura* lysozyme *g* sequences with high bootstrap support. This could be due to long-branch attraction. By employing the MP method, two equally good trees were constructed. One tree clustered Ciona1 and Ciona2 with the *Oikopleura* LysG1 and LysG2, while the other tree positioned Ciona1 and Ciona3 together at the base of the fish sequences. The ML tree also clustered the Ciona1 and Ciona3 sequences, while Ciona2 is clustered with the *Oikopleura* lysozymes.

The lysg1 gene of *Oikopleura* is a full-length g-type lysozyme gene split in two exons, and except for a small 5′ truncation the lysg2 clone carries an almost complete gene for a similar lysozyme. Comparing the genomic lysg2 sequence to the *Oikopleura* EST clone shows that the lysg2 sequence starts within an intron implying that even this lysozyme gene has two exons. In striking contrast, two *Ciona* lysozyme genes contain four exons and the third lysozyme gene has three exons. Intron organisations of vertebrate and tunicate g-type lysozymes are illustrated in figure 3. As usual when comparing intron-exon organisations in genes from distantly related taxons, some intron positions are found to be shared and must correspond to old insertions, whereas others are taxon specific and may represent later insertions or lineage-specific sliding of old introns. The two *Ciona* lysozyme genes with four exons have the same three intron positions, two of which (intron 1 at position IV and intron 3 at position VI) are shared by vertebrate genes. The *Ciona* three-exon gene lacks the first intron of the other two *Ciona* genes. The two *Oikopleura* genes have their single introns at different positions, and these positions are not conserved at least in the species sample examined here. Predictably, *Oikopleura* introns are shorter than those of other species.

Properties of urochordate g-type lysozymes

We performed basic protein statistic analyses on goose lysozymes from vertebrates (bird, mammalian and fish,) and the results indicated a uniform group of basic proteins with net charges of plus 5–15 and isoelectric points (pIs) of 9–10. In contrast, putative lysozymes of urochordates appeared in two (*Oikopleura*) or three (*Ciona*) distinct classes based on pI. Two acidic g-type lysozymes encoded by *Oikopleura* lysg1 and by Ciona2 had theoretical pI values of 5.0 and 4.4, respectively. The Ciona1 EST clone was predicted to encode a weakly acidic lysozyme with a pI of 6.4, whereas moderately or highly basic proteins with pIs of 8.0 and 9.9 would be produced from the *Oikopleura* lysg2 and the Ciona3 gene, respectively.

Both *Oikopleura* enzymes contained more cysteines (5.6–6.3%) than did g-type lysozymes from *Ciona* or vertebrates (0–3.5%). The cysteine content of vertebrate g-type lysozymes was however fairly variable: fish enzymes have either no cysteine residue, as in flounder, one, as salmon, or two, as in zebrafish, while birds and mammals have four conserved cysteines in common, and another three are conserved from mouse to human (fig. 1). Notably, the human2 lysozyme lacks these latter three C-terminal cysteines. Ciona2 and -3 lysozymes would contain four and two cysteines, respectively, and all four in the former protein and one in the latter protein are at or very close to positions conserved in mammalian proteins. The expressed Ciona1 lysozyme has three cysteines of which two are near to conserved positions and the third is in the signal peptide. The respective numerous cysteines in the two *Oikopleura* lysozymes are at identical positions and only one or two of them may possibly be in common with Ciona or mammalian lysozymes.

Water-soluble extracts of *Oikopleura* were prepared from ~ 1000 cultured specimens and used to hydrolyse *M. luteus* cell walls. The extract contained a total of 2200 units when measured at pH 5.2 and ionic strength 0.03. To abolish potential c-type lysozyme activity, we added an excess of Ivy, the recently discovered *Escherichia coli* inhibitor of c-type lysozyme [33]. In a control batch, 66 nM Ivy totally suppressed the activity of 8 nM c-type lysozyme from hen egg white and the hydrolysis could be observed only after adding 0.9 nM of the i-type chlamysin (fig. 4A). The hydrolysing activity of *Oikopleura* extracts proved unaffected by co-incubation with 66 nM Ivy (fig. 4B). The in vitro activity of *Oikopleura* lysozyme was found to be strongly pH dependent with maximum activity at pH 5.2 (fig. 4C). Increasing the ionic strength from 0.03 to 0.1 resulted in a 50-fold reduction in *Oikopleura* lysozyme activity (not shown).

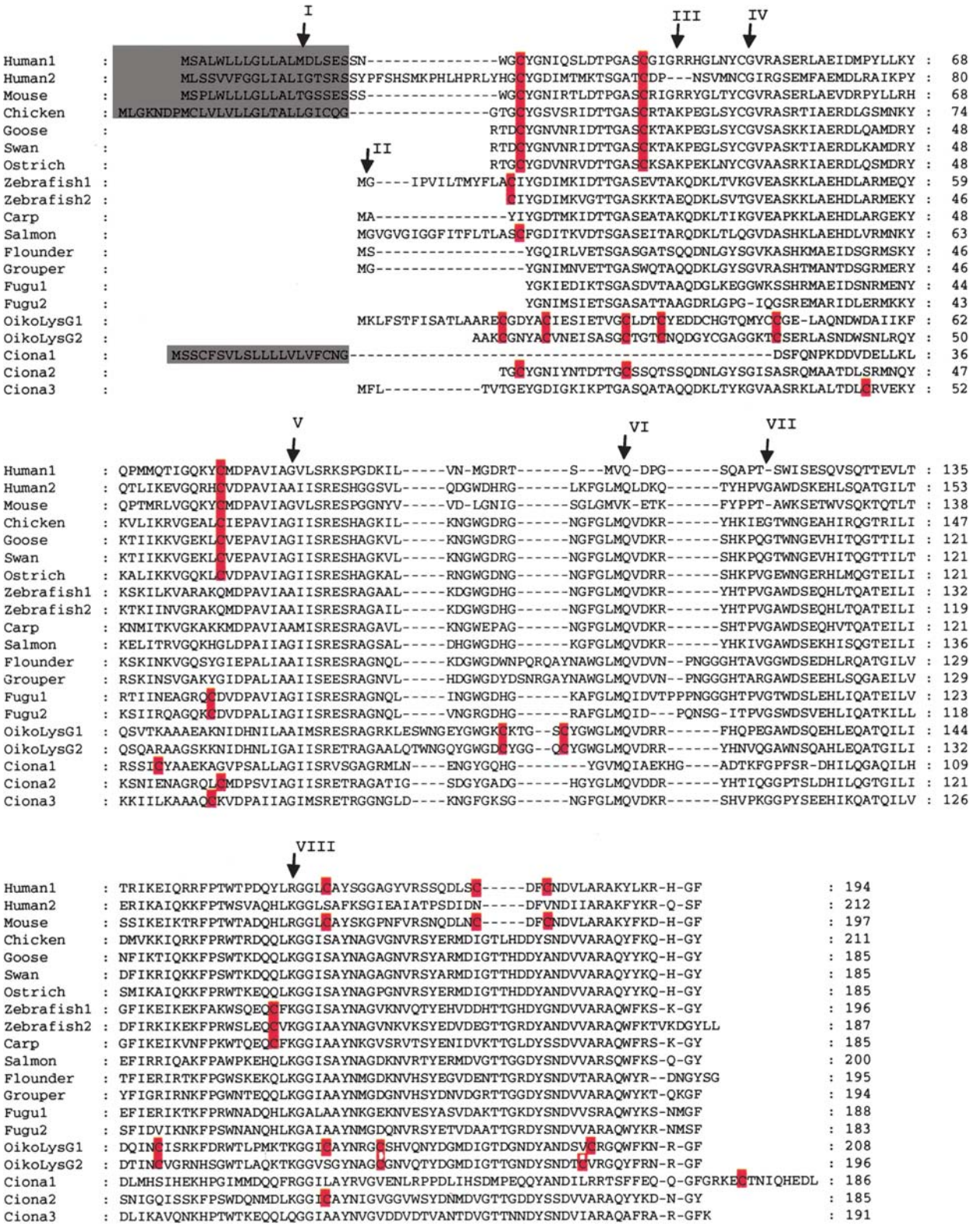

Figure 1. Amino acid sequence alignment of urochordate and vertebrate goose-type lysozymes. Multiple alignment of deduced and predicted lysozyme *g* sequences (see text for accession numbers and sources) was initially constructed using the ClustalX program and corrected by eye. Dashes (–) represent gaps inserted to maximise alignment. Predicted signal peptide sequences and cysteine residues are shaded in grey and red, respectively. The intron-exon organization of the g-type lysozyme genes has been investigated for the human, mouse, chicken, zebrafish, flounder, fugu, *C. intestinalis* and *O. dioica* sequences, and arrows point to positions representing encoded residues of exon-exon junctions of one or several lysozymes. See also figure 3 for details.

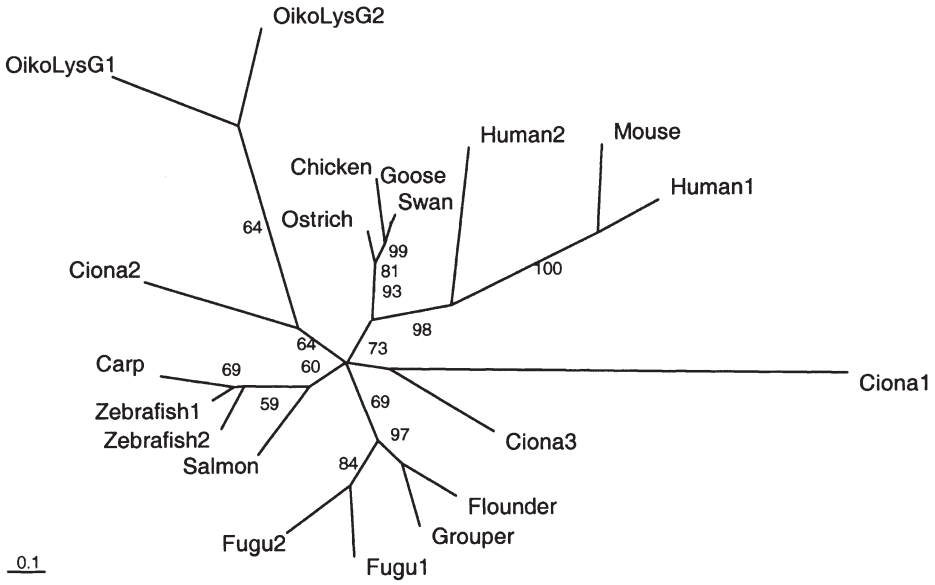

Figure 2. Phylogeny of *g*-type lysozymes in urochordates and vertebrates. The phylogeny of amino acid sequences of g-type lysozymes was estimated by the quartet-puzzling maximum likelihood method. The Dayhoff matrix was used as a model of substitution. The N-terminal part of the sequences was omitted due to ambiguous alignment. In the unrooted tree, internal branch support is shown for values of 50% and higher.

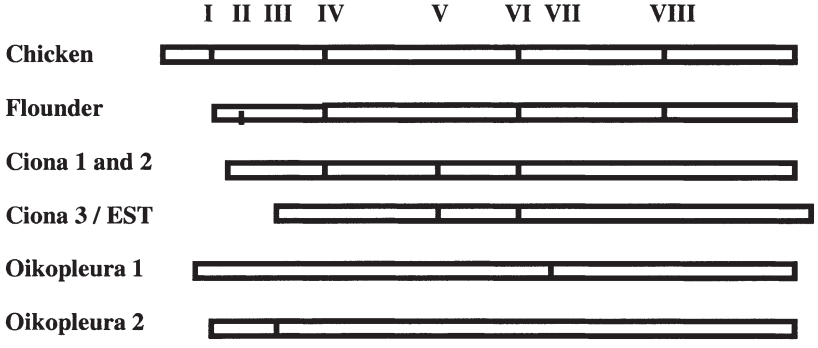

Figure 3. Exon structures of vertebrate and tunicate g-type lysozyme genes. Individual exons are separated in boxes, and intron positions are indicated following their order in one or several sequences as they appear when multiply aligned (see also fig. 1).

Search of lysozyme genes in other invertebrates

We searched for all lysozyme genes in the genome sequences of the fruit fly *Drosophila melanogaster* and the nematode *Caenorhabditis elegans* and in the EST databases of all invertebrates. Several recent sequence entries (accession numbers AAN86085, AY169675, ABO74894, AAN16375) revealed that shrimps have genes for lysozymes of the c-type in addition to the i-type. Genes encoding c- plus i-type lysozymes were found in the fly, and genes encoding i-type lysozyme were found in the nematode. No g-type lysozyme gene was detected in either species. Unexpectedly, one sequence encoding the carboxy-terminal half of a g-type lysozyme was found (accession No. AW739817) in an EST collection from hemocytes of the bloodfluke mollusc *Biomphalaria glabrata*.

Discussion

Based on sequence analyses, we have identified genes for the g-type but not for the c- or i-type lysozyme in the two marine urochordates *O. dioica* and *C. intestinalis*. The presence of g-type lysozyme ESTs from these two tunicate species demonstrates that lysozymes are expressed in these animals. The genes from *Oikopleura* have a unique intron organisation, while the *Ciona* lysozyme gene structures (see below) are clearly reminiscent of vertebrate genes and suggest that these genes have a common origin. Our study therefore shows that g-type lysozymes are not vertebrate specific. The phylogenetic analyses suggest that even though the exact positioning is unresolved, the *Ciona* g-type lysozymes are more closely related (or more similar) to the vertebrate sequences than

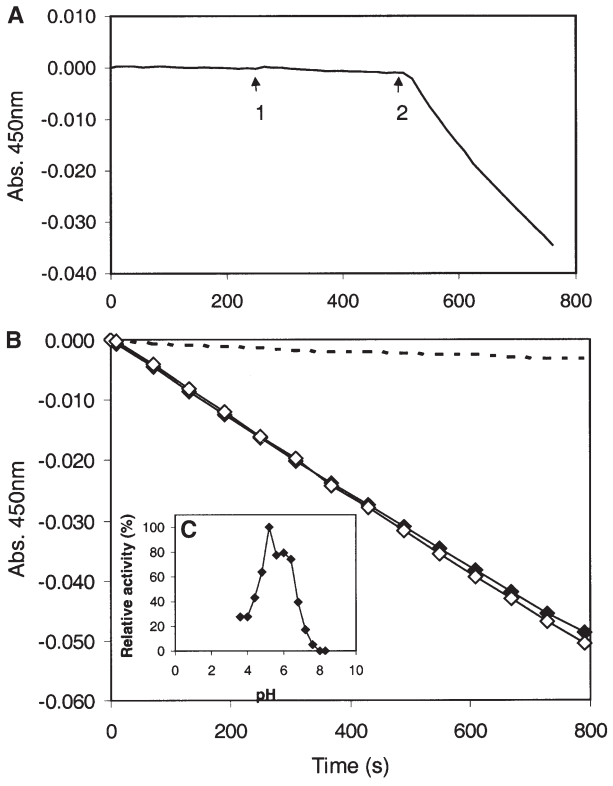

Figure 4. Lysozyme activity in *O. dioica*. (**A**) Control of the Ivy lysozyme inhibitor effect on c-type (HEWL) and i-type (chlamysin) lysozyme activity. Cell walls of *M. luteus* (0.2 mg/ml) in 10 mM sodium acetate buffer at pH 5.2 were incubated with 66 nM Ivy and 8 nM HEWL: (1) after 50 s HEWL was added to the incubation solution to a final concentration of 16 nM, and (2) after an additional 250 s, 0.9 nM chlamysin was included. (**B**) *Oikopleura* lysozyme activity on the substrate *M. lysodeikticus* (0.2 mg/ml) was measured in sodium acetate buffer at pH 5.2 and I = 0.03 (closed symbols) and compared to the activity in the presence of 66 nM Ivy (open symbols). Substrate alone (dashed line) represents blank in the assay. (**C**) The effect of pH on *Oikopleura* lysozyme activity. Activities from *Oikopleura* extracts were measured in sodium acetate buffers of pH 3.6–5.6 and in sodium phosphate buffers of pH 6.0–8.0 at ionic strength of 0.03. Enzyme activity is expressed as relative activity.

are the *Oikopleura* lysozymes. In the course of the present survey, we retrieved one typical g-type lysozyme sequence from a collection of mollusc ESTs. Confirmation will be needed to show that this sequence is encoded by the *Biomphalaria* genome, and to determine whether the origin of the g-type lysozyme preceded the emergence of chordates and even the emergence of deuterostomes, like the c-type lysozymes.

The fact that no other types of lysozyme were found in either the *Ciona* genome or the very rich collections of *Ciona* ESTs clearly suggests that c-type lysozymes have been lost in at least some chordate lineages. This also seems to be the case for larvaceans, since only g-type genes were detected in the shotgun sequencing data set or in the EST collection. Even though our *Oikopleura* genome dataset still comprises a very large number of relatively small contigs, we have indication from alignments with BAC insert sequences and with a sample of ESTs that it covers the genome and the majority of genes very well. Alignments of 800 non-redundant ESTs with this shotgun dataset showed an average coverage of 65%; 83% of the ESTs matched on more than one-quarter of their length and a further 48% covered at least three-quarters of it. Since lysozyme genes are well conserved throughout their full length, we estimate the chance that additional genes are present in the genome but undetected by BLAST is fairly low, and probably less than 10%. Therefore, the most probable situation is that both urochordate species only have g-type lysozymes, and that c-type genes were lost in a common ancestor. The i-type would also be absent and perhaps lost in all chordates. As a matter of fact, another invertebrate deuterostome, the echinoderm starfish *Asteria rubens,* was the first organism found to contain the i-type lysozyme [6]. Gene losses are one process that in combination with and in addition to gene genesis, gene duplication and horizontal gene transfer are important for shaping genomes in the course of evolution [39]. Loss of genes seems to be especially significant for evolution of small genomes and the proposed loss of c- as well as i-type lysozyme genes in urochordates may provide further examples of such genome evolutionary events.

In support of our gene data results, the enzyme activity assays strongly suggested the absence of c-type lysozymes. Extracts of *Oikopleura* exhibited *in vitro* hydrolysing activities on bacterial cell walls at acidic conditions under low ionic strength, and the activities were not inhibited by a known c-type lysozyme inhibitor. Furthermore, Ivy did not affect the activity of the i-type lysozyme from Icelandic scallop. The latter is a lysozyme with antibacterial activity against both Gram-positive and Gram-negative bacteria as opposed to the predominantly Gram-positive activity of c-type lysozyme [4, 33].

Contrasting the few known i-type gene structures, lysozyme genes of the c- and g-types vary little in their respective organisation although some introns may be lost or gained between phyla [5, 8, 10, 31, 40, 41]. While *Ciona* g-type lysozyme genes share intron positions with those of vertebrates, both genes of *Oikopleura* display divergent intron-exon organisations. This does not indicate a distant or independent origin of their lysozyme genes, since we have observed with a large number of other genes and gene families that a global remodelling of intron-exon organisations has occurred in the *Oikopleura* lineage [R. Edvardsen, E. Lerat, A. D. Maeland, M. Flåt, R. Tewari, M. J. Jensen et al., unpublished data]. This crisis involved an unusually high frequency of intron gains and losses, which may be related to the strong compaction of the genome and/or the short life cycle. Nonetheless, *Oikopleura* lysozymes display several divergent genetical and biochemical features. The moni-

tored activity in *Oikopleura* extracts depends on a low pH relative to the higher pH preference for the activity of vertebrate g-type lysozymes from fish or, in particular, chicken [9]. In the absence of available developmental- or tissue-specific expression profiles for lysozymes in *Oikopleura*, we do not know whether this acid-dependent activity arises from one or both (or all) lysozyme genes. The varying pIs of the *Oikopleura* and *Ciona* lysozymes indicate that tunicate lysozymes have undergone a specialisation process. Lysozyme specialisation or adaptation to certain tissues and conditions such as digestive organs and acidic environments is also known for the c-type lysozymes in vertebrates [11–14] and for the i-type lysozymes in bivalves [15]. One may fairly reasonably assume that the multiple g-type lysozymes in urochordates have similarly specialised, and the biochemically significant wide range of pI of the g-type lysozymes in these animals may support this assumption. Examples of g-type specialisation were recently reported in mammals, where two g-type genes are expressed at varying levels and in various tissues [10].

Of the five marine tunicate lysozymes, the complete N-terminal sequences of one enzyme from each of the two species are uncertain, and of the three apparently complete sequences from the tunicates, from signal peptide predictions, apparently only one protein from *Ciona* is secreted. The g lysozyme of Japanese flounder is expected to function at an intracellular level due to lack of a secretion signal [9], in contrast to the situation for similar enzymes from terrestrial vertebrates [10]. According to our signal predictions, the absence of secretion signals seems to be a common characteristic among the fish *g* lysozymes. Thus, tunicates seem to produce g-type lysozymes for intracellular use as fishes do, but the not fully characterised lysg2 gene of *Oikopleura* and a second gene of *Ciona* together with the EST-identified gene may produce secreted lysozymes. The two tunicate lysozymes predicted to lack secretion signals are the proteins that have a very low pI (4.4 and 5.0) and are thus probably well suited for storage in intracellular compartments.

The Japanese flounder *g*-type lysozyme was found to be devoid of cysteines, and the authors consequently suggested that disulphide bonds, and indirectly their effect on folding, are not necessary for lysozyme activity [9]. This is unexpected, since the divergent types of lysozymes from phages to mammals display similar enzyme activities believed to be contributed by their similar tertiary structure. Then again, disulphide bonds may function for stabilising purposes, and the lower stability of the fish lysozyme is reflected by the low temperature optimum of 25 °C for the flounder enzyme compared to 55 °C for the chicken enzyme [9]. Four cysteine residues are conserved in *g* lysozymes from birds to mammals, and these residues have been shown to be engaged in intramolecular disulphide binding in the bird proteins [21]. One of the *Ciona* lysozymes has two or possibly three of these four cysteines in addition to one cysteine conserved in mammals only. This demonstrates a rather high degree of conservation between mammals, birds and some tunicate lysozymes.

The close sequence relationship between the two *Oikopleura* lysozymes and the tight conservation of a uniquely high number of cysteines at specific positions in the two proteins are clear indications for an ancestor gene that duplicated and then evolved as a class-specific gene in the larvaceans. In contrast, the ascidian lysozymes share conserved cysteines as well as gene structures with vertebrate lysozymes. Thus, lineage-specific gene duplications seem to have happened in the tunicates as previously proposed for the two g-type lysozyme genes found in mammalian genomes [10].

Note added in proof. The salmon goose-type lysozyme sequence is derived from the DNA sequence (Accession No. CB517582), and during the proof-reading process the protein sequence was re-analysed and found to be the only presented fish sequence containing a secretion signal and which is comprised by 21 amino acid residues.

Acknowledgement. We would like to thank Drs. Hans Lehrach and Richard Reinhardt, Max-Planck Institute for Molecular Genetics, Berlin, Germany, for providing shotgun data, and Jean-Marie Bouquet at Sars Centre, Bergen, Norway, for the *Oikopleura* specimens.

1 Jollès P. and Jollès J. (1984) What's new in lysozyme research? Mol. Cell. Biochem. **63:** 165–189
2 Jollès P. (1996) Lysozymes: Model Enzymes in Biochemistry and Biology. Birkhäuser, Basel
3 Ito Y., Yoshikawa A., Hotani T., Fukuda S., Sugimura K. and Imoto T. (1999) Amino acid sequences of lysozymes newly purified from invertebrates imply wide distribution of a novel class in the lysozyme family. Eur. J. Biochem. **259:** 456–461
4 Nilsen I. W., Øverbø K., Sandsdalen E., Sandaker E., Sletten K. and Myrnes B. (1999) Protein purification and gene isolation of chlamysin, a cold-active lysozyme-like enzyme with antibacterial activity. FEBS Lett. **464:** 153–158
5 Bachali S., Jager M., Hassanin A., Schoentgen F., Jollès P., Fiala-Medioni A. et al. (2002) Phylogenetic analysis of invertebrate lysozymes and the evolution of lysozyme function. J. Mol. Evol. **54:** 652–664
6 Jollès J. and Jollès P. (1975) The lysozyme from *Asterias rubens*. Eur. J. Biochem. **54:** 19–23
7 Canfield R. E. and McMurry S. (1967) Purification and characterisation of a lysozyme from goose egg white. Biochem. Biophys. Res. Commun. **26:** 38–42
8 Nakano T. and Graf T. (1991) Goose-type lysozyme gene of the chicken: sequence, genomic organization and expression reveals major differences to chicken-type lysozyme gene. Biochem. Biophys. Acta **1090:** 273–276
9 Hikima J.-i., Minagawa S., Hirono I. and Aoki T. (2001) Molecular cloning, expression and evolution of the Japanese flounder goose-type lysozyme gene, and the lytic activity of its recombinant protein. Biochim. Biophys. Acta **1520:** 35–44
10 Irwin D. M. and Gong Z. (2003) Molecular evolution of vertebrate goose-type lysozyme genes. J. Mol. Evol. **56:** 234–242
11 Irwin D. M. and Wilson A. C. (1989) Multiple cDNA sequences and the evolution of bovine stomach lysozymes. J. Biol. Chem. **264:** 11387–11393

12 Stewart C. B., Schilling J. W. and Wilson A. C. (1987) Adaptive evolution in the stomach lysozymes of foregut fermenters. Nature **330:** 401–404
13 Kylsten P., Kimbrell D. A., Daffre S., Samakovlis C. and Hultmark D. (1992) The lysozyme locus in *Drosophila melanogaster*: different genes are expressed in midgut and salivary glands. Mol. Gen. Genet. **232:** 335–343
14 Takeuchi K., Irwin D. M., Gallup M., Shinbrot E., Kai H., Stewart C. B. et al. (1993) Multiple cDNA sequences of bovine tracheal lysozyme. J. Biol. Chem. **268:** 27440–27446
15 Olsen Ö., Nilsen I. W., Sletten K. and Myrnes B. (2003) Multiple invertebrate lysozymes in blue mussel (*Mytilus edulis*). Comp. Biochem. Physiol. **136:** 107–115
16 Keck W., Kazemier B. and Engel H. (1991) Murein-metabolizing enzymes from *Escherichia coli*: sequence analysis and overexpression of the *slt* gene, which encodes the soluble lytic transglycosylase. J. Bacteriol. **173:** 6773–6782
17 Koonin E. V. and Rudd K. E. (1994) A conserved domain in putative bacterial and bacteriophage transglycosylases. Trends Biochem. Sci. **19:** 106–107
18 Thunnissen A. M., Isaacs N. W. and Dijkstra B. W. (1995) The catalytic domain of a lytic transglycosylase defines a novel class of lysozymes. Proteins **22:** 245–258
19 Qasba P. K. and Kuman S. (1997) Molecular divergence of lysozymes and α-lactalbumin. Crit. Rev. Biochem. Mol. Biol. **32:** 255–306
20 Brew K., Castellino F. J., Vanaman T. C. and Hill R. L. (1970) The complete amino acid sequence of bovine α-lactalbumin J. Biol. Chem. **245:** 4570–4582
21 Qasba P. K. and Safaya S. K. (1984) Similarity of the nucleotide sequences of rat α-lactalbumin and the chicken lysozyme genes. Nature **308:** 377–380
22 Matthews B. W., Grütter M. G., Anderson W. F. and Remington S. J. (1981) Common precursor of lysozymes of hen egg-white and bacteriophage T4. Nature **290:** 334–335
23 Grütter M. G., Weaver L. H. and Matthews R. W. (1983) Goose lysozyme structure: an evolutionary link between hen and bacteriophage lysozymes? Nature **303:** 828–831.
24 Weaver L. H., Grütter M. G., Remington S. J., Dray T. M., Isaacs N. W. and Matthews B. W. (1985) Comparison of goose-type, chicken-type and phage-type lysozymes illustrates the changes that occur in both amino acid sequences and three-dimensional structures during evolution. J. Mol. Evol. **21:** 97–111
25 Holm L. and Sander C. (1994) Structural similarity of plant chitinase and lysozymes from animals and phage. FEBS Lett. **340:** 129–132
26 Shida K., Terajima D., Uchino R., Ikawa S., Ikeda M. and Asano K. (2003) Hemocytes of *Ciona intestinalis* express multiple genes involved in innate immune host defense. Biochem. Biophys. Res. Commun. **302:** 207–218
27 Dehal P., Satou Y., Campbell R. K., Chapman J., Degnan B., De Tomaso A. et al. (2002) The draft genome of *Ciona intestinalis:* insights into chordate and vertebrate origins. Science **298:** 2157–2167
28 Seo H. C., Kube M., Edvardsen R. B., Jensen M. F., Beck A., Spriet E. et al. (2001) Miniature genome in the marine chordate *Oikopleura dioica*. Science **294:** 2506
29 Spada F., Steen H., Troedsson C., Kallesoe T., Spriet E., Mann M. et al. (2001) Molecular patterning of the oikoplastic epithelium of the larvacean tunicate *Oikopleura dioica*. J. Biol. Chem. **276:** 20624–20632
30 Altschul S. F., Madden T. L., Schaffer A. A., Zhang J., Zhang Z., Miller W. et al. (1997) Gapped BLAST and PS*I*-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:** 3389–3402
31 Jung A., Sippel A. E., Grez M. and Schutz G. (1980) Exons encode functional and structural units of chicken lysozyme. Proc. Natl. Acad. Sci. USA **77:** 5759–5763
32 Myrnes B. and Johansen A. (1994) Recovery of lysozyme from scallop waste. Prep. Biochem. **24:** 69–80
33 Monchois V., Abergel C., Sturgis J., Jeudy S. and Claverie J. M. (2001) *Escherichia coli* ykfE ORFan gene encodes a potent inhibitor of *C*-type lysozyme. J. Biol. Chem. **276:** 18437–18441
34 Nielsen H., Engelbrecht J., Brunak S. and Heijne G. von (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. **10:** 1–6
35 Nakai K. and Kanehisa M. (1992) A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics **14:** 897–911
36 Thompson J. D., Gibson T. J., Plewniak F., Jeanmougin F. and Higgins D. G. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24:** 4876–4882
37 Schmidt H. A., Strimmer K., Vingron M. and Haeseler A. von (2002) TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics **18:** 502–504
38 Swoffor D. L. (2001) *PAUP* 4.0.* Sinauer, Sunderland, Mass.
39 Snel B., Bork P. and Huynen M. A. (2002) Genomes in flux: the evolution of archaeal and proteobacterial gene content. Genome Res. **12:** 17–25
40 Sun S. C., Åsling B. and Faye I. (1991) Organization and expression of the immunoresponsive factor in the giant silk moth *Hyalophora cecropia*. J. Biol. Chem. **266:** 6644–6649
41 Nilsen I. W. and Myrnes B. (2001) The gene of chlamysin, a marine invertebrate-type lysozyme, is organized similar to vertebrate but different from invertebrate chicken-type lysozyme genes. Gene **269:** 27–32