REVIEW

Invertebrate trypsins: a review

**Adriana Muhlia-Almazán · Arturo Sánchez-Paz ·
Fernando L. García-Carreño**

Abstract Food protein hydrolysis, a crucial step in
digestion, is catalyzed by trypsin enzymes from the diges-
tive apparatus of invertebrates. Trypsin appeared early in
evolution and occurs in all phyla and, in the digestive
systems of invertebrates, it became the most abundant
proteinase. As in vertebrates, invertebrate trypsin is also
present in several forms (isoenzymes). Its physiological
importance in food protein digestion in several inverte-
brate species has emerged with compelling evidence; and
several other physiological functions, such as regulation of
digestive functions, are now settled. Recent advances in
the knowledge of invertebrate trypsin synthesis, regula-
tion, genetics, catalytic characteristics; structure, evolu-
tion, as well as inhibition, especially in non-Drosophilidae
insects and in some crustaceans are reviewed. Most of the
existing information is largely based on the use of several
tools, including molecular techniques, to answer many still
open questions and solve medical, agricultural, and food
quality problems.

A. Muhlia-Almazán
Centro de Investigación en Alimentación y Desarrollo,
Carretera a la Victoria Km. 0.6, Hermosillo,
Sonora 83000, Mexico

A. Sánchez-Paz
Centro de Investigaciones Biológicas del Noroeste (CIBNOR),
Unidad Hermosillo, Hermosillo, Sonora 83260, Mexico

F. L. García-Carreño (✉)
Centro de Investigaciones Biológicas del Noroeste,
Mar Bermejo 195, Col. Playa Palo de Santa Rita,
La Paz, Baja California Sur 23090, México
e-mail: fgarcia@cibnor.mx

Introduction

Included in the five classes of proteinases catalyzing hydro-
lysis of peptide bonds, serine endoproteinases are divided
into two superfamilies that evolved similar catalytic mecha-
nisms independently. The trypsin superfamily includes
trypsins and chymotrypsins, which are ubiquitous in ani-
mals, while the subtilysine superfamily is present in most
bacteria and fungi studied to date (Hu and Leger 2004).
Regardless of origin, trypsins are classified in clan SA,
family S1 (Webb 1993; Barrett et al. 1998). Trypsins pos-
sess a catalytic triad that characterize all serine proteinases,
consisting of His, Asp, and Ser amino acid residues and
typically show two characteristics that distinguish them
from other serine proteinases: (1) specificity for the peptide
bond formed by the carboxylic-side of Arg or Lys residues
and (2) ability to activate other pancreatic zymogens (de
Albuquerque et al. 2001).

The ability to digest protein must have emerged early in
evolution, as shown by the finding of trypsin coding genes
in eubacterial genomes (Rypniewski et al. 1994). Primitive
organisms selectively fixed the ability to use foreign protein
as a source of amino acids by synthesizing enzymes, now
termed proteinases, which were capable of recognizing and
hydrolyzing peptide bonds in a specific region of a poly-
peptide chain, based on specificity for the side chains of
amino acid residues in the carboxylic side of the site of
cleavage.

Invertebrate proteolytic enzymes have been widely stud-
ied and special attention has been directed to the trypsin
group (EC 3.4.21.4) because they take part in a number of

physiological processes, playing a major and better-understood key role in hydrolyzing food protein (Díaz-Mendoza et al. 2005; Delcroix et al. 2006). Among invertebrates, crustaceans and insects have been the traditional model, but an increasing list of species belonging to other taxa have contributed to understanding and characterizing physiological and biochemical processes in which these enzymes are involved (Noriega and Wells 1999; Maeda-Martínez et al. 2000; de Albuquerque et al. 2002). The first studies of proteinases, mainly trypsin and chymotrypsin in vertebrate species, started during the first part of the twentieth century, describing kinetic properties, isolation, some molecular characteristics, and evolution of trypsin and trypsinogen. The first reports of invertebrate trypsins were published in the 1960s, when secretion of digestive enzymes responsible for hydrolysis of polymeric nutrients in food were described in insects and crustaceans (Davie and Neurath 1955; Keil et al. 1968; Arnon and Neurath 1969; Pfleiderer et al. 1970; Yang and Davies 1971; Gooding 1975; Gibson and Barker 1979), showing that most terrestrial and aquatic invertebrates synthesize and secrete trypsin and other related enzymes. It is now well documented that invertebrates synthesize paralogue (homologue enzymes in the same organisms) and orthologue (homologue enzymes within a taxa) trypsins (Zhu et al. 2000; Heger and Ponting 2008). In 1876, Wilhelm Friederich Kuehne discovered a protein substance, in pancreatic juice of pig that degraded other protein substances and named it trypsin. Kuehne (1876) soon learnt that this pancreatic proteinase was initially in an inactive state, but spontaneously converts to the active form and is inactivated (denatured) by heat. One of the first proteinases isolated with sufficient purity and enough quantity for accurate biochemical analysis was trypsin from bovine pancreas (Northrup et al. 1948). Bovine trypsin was purified by crystallization in the early 1930s (Northrup and Kunitz 1931) and almost the entire amino acid sequence of the zymogen was known by protein sequencing in the 1960s (Walsh et al. 1964; Walsh and Wilcox 1970). With the advent of X-ray crystallography, the three-dimensional structure of bovine trypsin was resolved in the 1970s (Huber et al. 1974; Sweet et al. 1974; Kossiakoff et al. 1977; Rypniewski et al. 1993). Thus, over several decades, details of the amino acid sequence, structure, and mechanism of action of trypsin were revealed in a large number species, including invertebrates (Male et al. 1995; Noriega and Wells 1999; Heheman et al. 2007).

Given the importance of trypsin's functions and ubiquity, its evolution has been intensely studied. Using Fourier transformation of biological sequences, Baptista et al. (1998) found strong evidence for tandem gene duplication. A likely evolutionary path for the development of present-day trypsins involved an intrinsic extensive tandem gene duplication of a small DNA fragment of 15–18 nucleotides,

corresponding to five or six amino acids. This ancestral trypsin gene was subsequently duplicated, leading to the earliest version of a full-sized trypsin, from which contemporary trypsins have developed. Since invertebrates have a body temperature close to that of the environment, adaptations to temperature should have evolved. Comparative studies of structures from several trypsins revealed that several characteristics are unique to cold-adapted trypsins. An example of such characteristics is the increase in substrate affinity from a lower electrostatic potential of the S1-binding pocket provided by Glu221B and the absence of five hydrogen bonds adjacent to the catalytic triad, compared to mesophilic counterparts (Schrøder-Leiros et al. 2000).

In this review, we summarize the knowledge concerning invertebrate trypsins and make comparisons with vertebrates when a topic is poorly understood in invertebrates. The review encompasses molecular and kinetic characteristics, mechanisms regulating measurable activity in digestive organs, mechanisms dealing with unwanted autolysis of cell and tissue proteins, genetics, evolution, and function.

Characteristics of invertebrate trypsin

Trypsins show species-specific characteristics since there are significant differences in feeding habits, ingested food composition, and differences in the protein digestion process, according to the requirements of the species. These differences are adaptive responses to different life styles, environments, and mostly, different survival abilities among species. However, trypsins share some general characteristics, as observed in Tables 1 and 2, mainly related to their catalytic features.

Ca^{2+} dependence

To date, it is not conclusive whether invertebrate trypsins depend on a moderate concentration of Ca^{+2} for maximum activity and stability, as observed in mammalian trypsins. Trypsin from starfish *Asterias amurensis* and *Lysastrosoma anthosticta* and insects *Diatraea saccharalis*, *Periplaneta americana*, *Tenebrio mollitor*, and *Musca domestica* are not dependent on Ca^{+2} for activation or stability. In fact, the calcium-binding motif that is commonly observed in mammalian trypsins does not occur in most of these insect trypsins (Dionysius et al. 1993; Kishimura and Hayashi 2003; Lopes et al. 2006). However, trypsin structure from the crayfish *Astacus leptodactylus* has a high-affinity, calcium-binding site formed by the Glu$_{70}$–Glu$_{80}$ loop (Fodor et al. 2005). Some authors report activating and stabilizing effects of calcium on serine proteinases from invertebrates, such as the crab *Cancer pagurus* (Saborowski et al. 2004)

Table 1 General characteristics of various trypsins

Enzyme	Species	Molecular mass (kDa)	Tissue	Optimum pH	Inhibitor sensitivity	References
Trypsin	*Rhynchophorus ferrugineus*	–	Gut	9–10	SBTI, LBTI, PMSF	Alarcón et al. (2002)
Trypsin	*Asterias amurensis*	28	Pyloric ceca	8	–	Kishimura and Hayashi (2003)
Trypsin	*R. dominica*	23–24	Midgut gland	7.0–9.5	TLCK, STI, PMSF, E-64	Zhu and Baker (1999)
Trypsin	*Musca domestica*	26.6	Whole animal	8.5	–	Lopes et al. (2006)
Trypsin	*Diatraea saccharalis*	28.7	Midgut content	10.5	–	Lopes et al. (2006)
Trypsin	*Sesamia nonagrioides*	21–31	Midgut gland	8.5–10.5	E-64, HgCl₂, LBI, SBBI	Diaz-Mendoza et al. (2005); Novillo et al. (1999)
Trypsin	*Triatoma infestans*	40	Salivary glands	–	APMSF, SBTI	Amino et al. (2001)
Trypsin	*Manduca sexta*	25	Midgut gland	10.5	–	Peterson et al. (1994)
Trypsin	*Triops* sp.	22–31	Whole animal	8	TLCK	Maeda-Martínez et al. (2000)
Trypsin	*Astacus fluviatilis*	25	Midgut gland	7–8	PMSF, TLCK, DFP	Zwilling and Neurath (1981); Titani et al. (1983)
Trypsin	*Penaeus japonicus*	25	Midgut gland	8–8.3	PMSF, TLCK, SBTI	Galgani et al. (1985)
Trypsin	*Penaeus vannamei*	30–32	Midgut gland		–	Klein et al. (1996)
Trypsin	*Farfantepenaeus paulensis*	14.6–21.7	Midgut gland	–	TLCK, PMSF	Lemos et al. (1999)
Trypsin	*Cancer pagurus*	26	Gastric fluid	8	–	Saborowski et al. (2004); Heheman et al. (2007)

E-64 trans-epoxysuccinyl-L-leucylamido-(4-guanidino) butane, *PMSF* phenylmethansulfonyl fluoride, *TLCK* tosyl-L-lysyl-chloromethylketone, *SBTI* soybean trypsin inhibitor, *LBTI* lima bean trypsin inhibitor

and Australian redclaw crayfish *Cherax quadricarinatus* (Figueiredo et al. 2001).

Structure

The three-dimensional structure of enzymes in the S1 family is highly conserved, even though primary structures vary to a large extent. To date, the crystal structure of trypsin enzymes has been studied for only two crustaceans, the freshwater crayfish *A. leptodactylus* (Fodor et al. 2005) and the marine crab *C. pagurus* (Heheman et al. 2007). Active trypsin consists of a single polypeptide chain; the catalytic amino acid residues lie between two-stranded, β-barrel domains packed against each other. Ser_{195}, in the numbering system of bovine chymotrypsin (Rypniewski et al. 1994), acts as a nucleophile in hydrolysis and produces an acyl-enzyme intermediate with the substrate, while His_{57} acts as a general base and Asp_{102} stabilizes the correct tautomer of His_{57} and provides compensation for developing of a positive charge during the catalytic reaction. The substrate forms an anti-parallel β sheet with the protein-binding site and specificity is related to the Asp_{189} side chain that is located at the bottom of the binding pocket, which provides a negative charge for attracting the positively charged Lys or Arg amino acid residue in the substrate. Fodor et al. (2005) reported some important differences between crayfish trypsin structure and that of vertebrates. They found four loops (37, 60, 145, and 202) in crayfish trypsin that are different from vertebrates. Also, an extended hydrophobic region in crayfish trypsin includes a five-residue insertion at position 37, which is suspected to be involved with inhibitor interaction and a seven-residue insertion at position 60. These insertions affect the molecular mass (kDa) and pI values, yielding both cationic and anionic enzymes (Table 1; Barrett et al. 1998). Most relatives of invertebrate trypsins have an N-terminal signal peptide and are synthesized as precursors with an N-terminal extension acting as proenzymes that require proteolytic cleavage to yield the active enzyme (Amino et al. 2001; Sainz et al. 2004b).

Evolution

The two central elements of neo-Darwinian evolution are small random variations and natural selection. Comprehension of evolution requires the understanding of the details of random variation of DNA and the natural selection acting on gene products. The comparison of two or more sequences is the basis for evolutionary analysis, and it may reveal similarities that arise through random chance, sharing a common ancestor, or sharing a common selective pressure. In general, trypsin sequences in invertebrate species are still a small number, which provides an incomplete description of the long evolutionary process that these enzymes might have undergone.

In the late 1970s, immunological methods were used in studies on the evolution of peptidases. Studies of trypsins

Table 2 Trypsin genes from invertebrate species and translated/ isolated protein characteristics

Enzyme GenBank acc. #	Species	Number of genes/cDNA sequences	Size (aa)		Active site triad position	Cysteine pair residues	Residues conferring trypsin-like specificity	Mature protein size* (AA)	References
			Signal peptide	Activation peptide					
AF130840–2	*R. dominica*	Three cDNAs	16–18	12–15	His109, Asp156, Ser257	Three	Asp251, Gly274, Gly284	219, 261, 226	Zhu and Baker (1999)
AF173495, AF173496	*Plodia interpunctella*	Two cDNAs	16–18	12–15	His81, Asp133, Ser233	Three	Asp227, Gly250, Gly260	232	Zhu et al. (2000)
Z22930	*Anopheles gambiae*	Seven genes Two cDNAs	18–19	29–30	His91, Asp136, Ser232	Three	Asp226, Gly251, Gly259	225, 227	Müller et al. (1993, 1995)
AF072690	*Diaprepes abbreviatus*	Two genes	11	11	His41, Asp92, Ser185	Four	Asp 189, –, –	239	Yan et al. (1999)
	Drosophila melanogaster	Four genes	22	8	His57, Asp102, Ser195	Three	–	226	Davies et al. (1985); Wang et al. (1999)
AF487426	*Aedes aegypti*	Three genes One cDNA	15	10	His45, Asp100, Ser203	Three	Asp196, Gly229, Gly237	232	Barillas-Mury et al. (1991); Kalhok et al. (1993)
AY308482	*Pediculus humanus*	One gene One cDNA	16	12	His69, Asp114, Ser206	Three	Asp69, Gly114, Gly206	225	Kollien et al. (2004)
AF074955, AF074956	*Stomoxys calcitrans*	Four genes Two cDNAs	16–17	10–15	His68, Asp113, Ser208	Three	Asp211, Gly242, Gly250	239, 228	Lehane et al. (1998)
P35045	*Manduca sexta*	Three cDNAs	17	7	His46, Asp91, Ser194	Three	Asp188, Gly211, Gly221	232	Peterson et al. (1994)
AJ222658	*Pacifastacus leniusculus*	One cDNA	15	13	His57, Asp102, Ser195	Three	Asp189, Gly216, Gly226	237	Hernández-Cortés et al. (1999b)
Y15039 Y15040 Y15041	*Penaeus vannamei*	14 genes Five cDNAs	14	14	His57, Asp102, Ser195	Four	–	227	Klein et al. (1996, 1998)
AY294257 AY522438 to AY522441	*Lepeophtheirus salmonis*	Five cDNAs	15–17	23–24	His57, Asp102, Ser195	Three	Asp189, Gly216, Gly226	223	Kvamme et al. (2004, 2005)

* Some species include the deduced protein size of the mature protein

of vertebrates suggested an evolutionary relationship among these enzymes, finding that mutations affecting the structure of enzymes were extensive, although catalytic activity and specificity were barely affected (Pfleiderer et al. 1967; Arnon and Neurath 1969; Pfleiderer et al. 1970). The high degree of cross-reactivity among different trypsins, including those in invertebrates, indicated that several features on the surface of the molecule are common in these proteins. These data, in turn, suggested a high-degree of similarity in the three-dimensional structure of vertebrate and invertebrate trypsins, thus establishing that trypsins appeared before invertebrates and vertebrates diverged (Arnon and Neurath 1969).

Studies of the phylogeny of trypsins and trypsinogens of vertebrates and invertebrates provided some clues to the evolution of these protein families (Roach et al. 1997). Trypsin of the midgut gland of the crayfish *Astacus fluviatilis*, a species on the evolutionary pathway when decapods and vertebrates diverged 700 mya, shows that some essential structural properties are more closely shared with bacteria trypsin than with bovine trypsin, suggesting that crayfish trypsins should be placed closer to bacteria trypsins than to vertebrate trypsins. Vertebrate trypsins evolved new characteristics, such as a larger number of disulfide bridges (Zwilling and Neurath 1981; Titani et al. 1983; Neurath 1984). However, the main and more conserved sequence characteristics of vertebrate trypsins have also been observed in invertebrate trypsins, like the key amino acid residues in the specificity pocket Asp, Gly, Gly (Roach et al. 1997) and the three amino acid residues at the active site, His, Asp, and Ser, that are involved in the catalytic mechanism of serine proteinases (Neurath 1984; Pancer et al. 1996). Also, in the invertebrate trypsinogens, the presence of a signal peptide with a typical cleavage site, Arg or Lys/Ile, in vertebrates is observed (Table 2; De Haën et al. 1975; Klein et al. 1996; Lehane et al. 1998). However, important differences between vertebrate and invertebrate trypsinogens have been found. The shrimp and crayfish signal peptides in trypsinogens appear to be longer than those of vertebrates and other invertebrates, according to Peterson et al. (1994), but it seems that such results are not in agreement with new information (Fig. 1); however, we suggest that crustacean trypsins are more similar in size than those in insects.

Other important difference found between vertebrate and invertebrate trypsins is the number of Cys residues and, as a consequence, in the number and location of disulfide bridges. Vertebrate species commonly have six disulfide bridges, and some insect and crustacean trypsins have only three at conserved positions, which seems to be conserved bonds because their position is closer to the active site of the enzyme, as seen in the crayfish *Pacifastacus leniusculus* and *A. leptodactylus* and insects *Anopheles gambiae* and *Pediculus humanus* (Table 2; Fig. 1; Fletcher et al. 1987; Müller et al. 1993; Hernandez-Cortés et al. 1999a; Kollien et al. 2004; Fodor et al. 2005). In the shrimp *P. vannamei* (Klein et al. 1996) and the citrus weevil *Diaprepes abbreviatus*, eight Cys residues and four disulfide bonds are present. Differences in the number of Cys residues have been linked to the early evolutionary separation of vertebrates and invertebrates (Titani et al. 1983). De Haën et al. (1975) suggested that their common ancestor had the same number of disulfide bonds at the same positions as all vertebrate trypsins.

Rypniewsky et al. (1994) described a phylogenetic tree with a continuous evolution of trypsins from a single ancestral gene. A few years later, Roach et al. (1997) suggested that vertebrate trypsinogen evolution was dynamic and multi-modal, including gross duplication of a whole gene or part of the trypsinogen gene locus after elasmobranches diverged, giving rise to two groups of trypsinogen genes: anionic and cationic trypsinogens. This idea supports the hypothesis that a single, gene-duplication occurred prior to the divergence of mammals (Rypniewsky et al. 1994). The time needed for the divergence of trypsin into cationic and anionic forms has been subject to speculation. Fletcher et al. (1987) stated that, because of the high homology among the amino acid sequences of cationic trypsinogens in different species, divergence of a primordial trypsinogen gene should occur before the divergence of rodents and ungulates. The structure of anionic and cationic trypsinogens has been studied in vertebrates; however, the manner in which the sequences of these proteins dictate their structures is not understood (Pasternak et al. 1999).

Anionic and cationic trypsins have been observed in invertebrates. Ward (1975) reported the purification and properties of anionic trypsin in larvae of the webbing clothes moth *Tineola bisselliella* and Pancer et al. (1996) reported the sequences of two cDNAs coding for putative anionic trypsinogens from the colonial ascidian *Botryllus scholosseri*. In the insects *T. mollitor* and *Locusta migratoria*, two or three forms of anionic trypsinogen have been found in the midgut, together with a single cationic trypsinogen (Lam et al. 2000; Lopes and Terra 2003). Thus, besides differences in the overall net electric charge between cationic and anionic trypsinogens, the presence of trypsins, either cationic or anionic, is species-specific and the electric charge could be functionally important; however, little is known about the practical significance of different trypsins and its conservation during evolution (Fletcher et al. 1987; Pancer et al. 1996).

Mechanisms regulating trypsin activity

Trypsins are paramount effectors in the network of enzymatic functions, since its activation is the first step in a
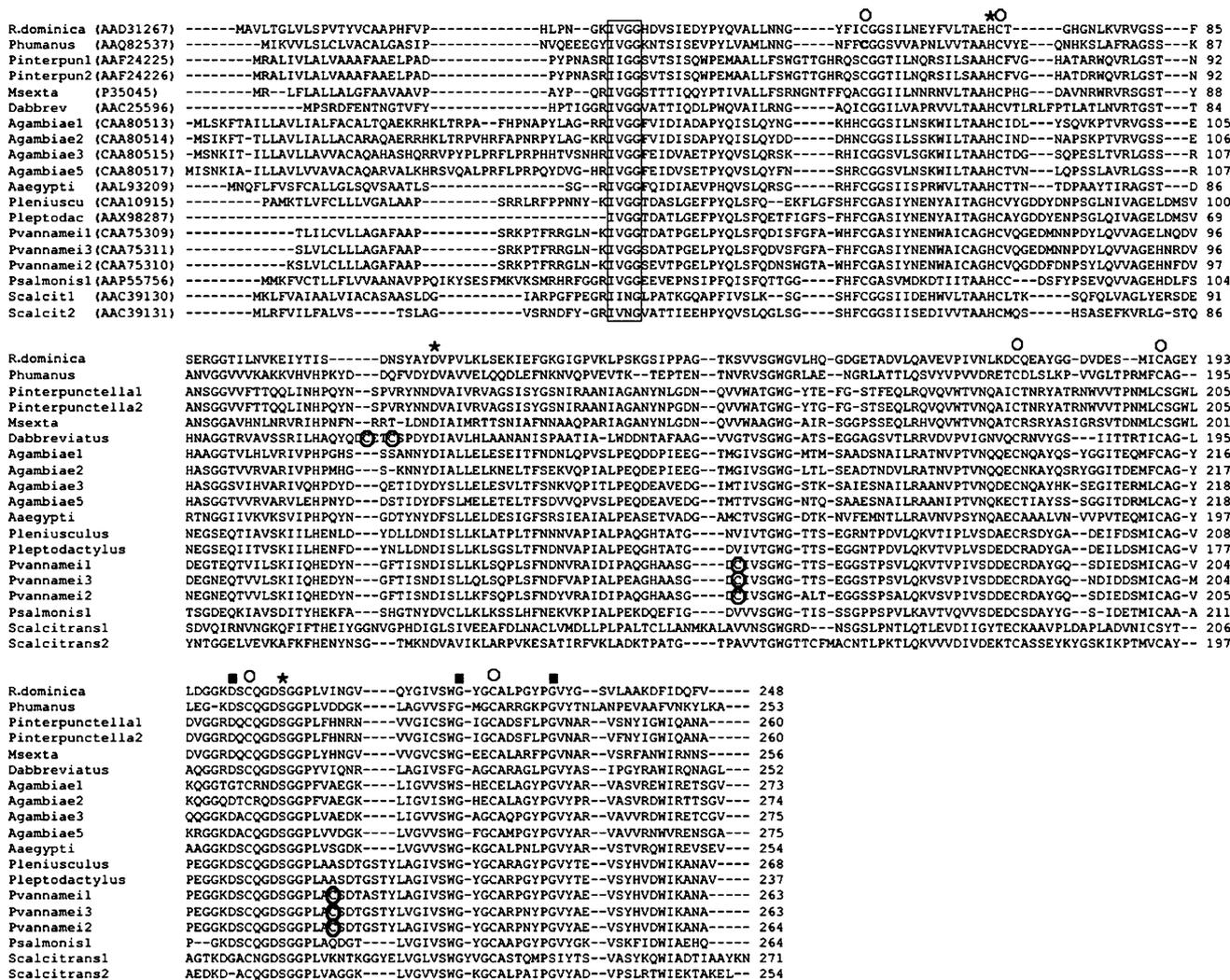
```
                                                                              O                   *O
R.dominica (AAD31267)  ------MAVLTGLVLSPVTYVCAAPHFVP------------HLPN--GKIVGGHDVSIEDYPYQVALLNNG-----YFICGGSILNEYFVLTAEHCT------GHGNLKVRVGSS---F  85
Phumanus   (AAQ82537)  ---------MIKVVLSLCLVACALGASIP--------------NVQEEEGYIVGGKNTSISEVPYLVAMLNNG-----NFFCGGSVVAPNLVVTAAHCVYE---QNHKSLAFRAGSS---K  87
Pinterpun1 (AAF24225)  --------MRALIVLALVAAAFAAELPAD--------------PYPNASRIIGGSVTSISQWPEMAALLFSWGTTGHRQSCGGTILNQRSILSAAHCFVG---HATARWQVRLGST---N  92
Pinterpun2 (AAF24226)  --------MRALIVLALVAAAFAAELPAD--------------PYPNASRIIGGSVTSISQWPEMAALLFSWGTTGHRQSCGGTILNQRSILSAAHCFVG---HATDRWQVRLGST---N  92
Msexta     (P35045)   --------MR--LFLLALLALGFAAVAAVP--------------AYP--QRIVGGSTTTIQQYPTIVALLFSRNGNTFFQACGGILNNRRNVLTAAHCPHG---DAVNRWRVRSGST---Y  88
Dabbrev    (AAC25596)  --------------MPSRDFENTNGTVFY--------------HPTIGGRIVGGVATTIQDLPWQVAILRNG-----AQICGGILVAPRVVLTAAHCVTLRLFPTLATLNVRTGST---T  84
Agambiae1  (CAA80513)  -MLSKFTAILLAVLIALFACALTQAEKRHKLTRPA--FHPNAPYLAG-KRIVGGFVIDIADAPYQISLQYNG-----KHHCGGSILNSKWILTAAHCIDL---YSQVKPTVRVGSS---E 105
Agambiae2  (CAA80514)  -MSIKFT-TLLAVLIALLACARAQAERRHKLTRPVHRFAPNRPYLAG-KRIVGGFVIDISDAPYQISLQYDD-----DHNCGGSILSSKWILTAAHCIND---NAPSKPTVRVGSS---E 106
Agambiae3  (CAA80515)  -MSNKIT-ILLAVLLAVVACAQAHASHQRRVPYPLPRFLPRPHHTVSNHRIVGGFEIDVAETPYQVSLQRSK-----RHICCGSVLSGKWILTAAHCTDG---SQPESLTVRLGSS---N  92
Agambiae5  (CAA80517)  MISNKIA-ILLALVVAVACAQARVALKHRSVQALPRFLPRPQYDVG-HRIVGGFEIDVSETPYQVSLQYFN-----SHRCGGSVLNSKWILTAAHCTVN---LQPSSLAVRLGSS---R 107
Aaegypti   (AAL93209)  -----MNQFLFVSFCALLGLSQVSAATLS----------------SG--RIVGGFQIDIAEVPHQVSLQRSG-----RHFCGGSIISPRWVLTAAHCTTN---TDPAAYTIRAGST---D  86
Pleniuscu  (CAA10915)  ---------PAMKTLVFCLLLVGALAAP---------SRRLRFPPNNY-KIVGGTDASLGEFPYQLSFQ--EKFLGFSHFCGASIYNENYAITAGHCVYGDDYDNPSGLNIVAGELDMSV 100
Pleptodac  (AAX98287)  ---------------------------------------------IVGGTDATLGEFPYQLSFQETFIGFS-FHFCGASIYNENYAITAGHCAYGDDYENPSGLQIVAGELDMSV  69
Pvannamei1 (CAA75309)  -------------TLILCVLLAGAFAAP---------SRKPTFRRGLN-KIVGGTDATPGELPYQLSFQDISFGFA-WHFCGASIYNENWAICAGHCVQGEDMNNPDYLQVVAGELNQDV  96
Pvannamei3 (CAA75311)  -------------SLVLCLLLAGAFAAP---------SRKPTFRRGLN-KIVGGSDATPGELPYQLSFQDVSFGFA-FHFCGASIYNENWAICAGHCVQGEDMNNPDYLQVVAGEHNRDV  96
Pvannamei2 (CAA75310)  -------------KSLVLCLLLAGAFAAP---------SRKPTFRRGLN-KIVGGSEVTPGELPYQLSFQDNSWGTA-WHFCGASIYNENWAICAGHCVQGDDFDNPSYLQVVAGEHNFDV  97
Psalmonis1 (AAP55756)  ---------MMKFVCTLLFLVVAANAVPPQIKYSESFMKVKSMRHRFGGRIVGGEEVEPNSIPFQISFQTTGG-----FHFCGASVMDKDTIITAAHCC---DSFYPSEVQVVAGEHDLFS 104
Scalcit1   (AAC39130)  --------MKLFVAIAALVIACASAASLDG----------IARPGFPEGRIINGLPATKGQAPFIVSLK--SG----SHFCGSIIDEHWVLTAAHCLTK-----SQFQLVAGLYERSDE  91
Scalcit2   (AAC39131)  --------MLRFVILFALVS-----TSLAG----------VSRNDFY-GRIVNGVATTIEEHPYQVSLQGLSG----SHFCGGSIISEDIVVTAAHCMQS-----HSASEFKVRLG-STQ  86

                              *                                                       O                   O
R.dominica        SERGGTILNVKEIYTIS------DNSYAYDVPVLKLSEKIEFGKGIGPVKLPSKGSIPPAG--TKSVVSGWGVLHQ-GDGETADVLQAVEVPIVNLKDCQEAYGG-DVDES--MICAGEY 193
Phumanus          ANVGGVVVKAKKVHVHPKYD----DQFVDYDVAVVELQQDLEFNKNVQPVEVTK--TEPTEN--TNVRVSGWGRLAE--NGRLATTLQSVYVPVVDRETCDLSLKP-VVGLTPRMFCAG-- 195
Pinterpunctella1  ANSGGVVFTTQQLINHPQYN--SPVRYNNDVAIVRVAGSISYGSNIRAANIAGANYNLGDN--QVVWATGWG-YTE-FG-STFEQLRQVQVWTVNQAICTNRYATRNWVVTPNMLCSGWL 205
Pinterpunctella2  ANSGGVVFTTQQLINHPQYN--SPVRYNNDVAIVRVAGSISYGSNIRAANIAGANYNPGDN--QVVWATGWG-YTG-FG-STSEQLRQVQVWTVNQAICTNRYATRNWVVTPNMLCSGWL 205
Msexta            ANSGGAVHNLNRVRIHPNFN--RRT-LDNDIAIMRTTSNIAFNNAAQPARIAGANYNLGDN--QVVWAAGWG-AIR-SGGPSSEQLRHVQVWTVNQATCRSRYASIGRSVTDNMLCSGWL 201
Dabbreviatus      HNAGGTRVAVSSRILHAQYQQTCSPDYDIAVLHLAANANISPAATIA-LWDDNTAFAAG--VVGTVSGWG-ATS-EGGAGSVTLRRVDVPVIGNVQCRNVYGS---IITTRTICAG-L 195
Agambiae1         HAAGGTVLHLVRIVPHPGHS---SSANNYDIALLELESEITFNDNLQPVSLPEQDDPIEEG--TMGIVSGWG-MTM-SAADSNAILRATNVPTVNQQECNQAYGS-YGGITEQMFCAG-Y 216
Agambiae2         HASGGTVVRVARIVPHPMHG---S-KNNYDIALLELKNELTFSEKVQPIALPEQDEPIEEG--TMGIVSGWG-LTL-SEADTNDVLRATNVPTVNQQECNKAYGSRYGGITDEMFCAG-Y 217
Agambiae3         HASGGSVIHVARIVQHPDYD----QETIDYDYSLLELESVLTFSNKVQPITLPEQDEAVEDG--IMTIVSGWG-STK-SAIESNAILRAANVPTVNQDECNQAYHK-SEGITERMLCAG-Y 218
Agambiae5         HASGGTVVRVARVLEHPNYD----DSTIDYDFSLMELETELTFSDVVQPVSLPEQDEAVEDG--TMTTVSGWG-NTQ-SAAESNAILRAANIPTVNQKECTIAYSS-SGGITDRMLCAG-Y 218
Aaegypti          RTNGGIIVKVKSVIPHPQYN----GDTYNYDFSLLELDSIGFSRSIEAIALPEAETVADG--AMCTVSGWG-DTK-NVFEMNTLLRAVNVPSYNQAECAAALVN-VVPVTEQMICAG-Y 197
Pleniusculus      NEGSEQTIAVSKIILHENLD----YDLLDNDISLLKLATPLTFNNNVAPIALPAQGHTATG----NVIVTGWG-TTS-EGRNTPDVLQKVTIPLVSDAECRSDYGA--DEIFDSMICAG-V 208
Pleptodactylus    NEGSEQIITVSKIILHENFD----YNLLDNDISLLKLSGSLTFNDNVAPIALPEQGHTATG----DVIVTGWG-TTS-EGGNTPDVLQKVTVPLVSDEDCRADYGA--DEILDSMICAG-V 177
Pvannamei1        DEGTEQTVILSKIIQHEDYN---GFTISNDISLLKLSQPLSFNDNVRAIDIPAQGHAASG----DQVSGWG-TTS-EGGSTPSVLQKVTVPIVSDDECRDAYGQ--SDIEDSMICAG-V 204
Pvannamei3        DEGNEQTVVLSKIIQHEDYN---GFTISNDISLLQLSQPLSFNDFVAPIALPEAGHAASG----DQVSGWG-TTS-EGGSTPSVLQKVSVPIVSDDECRDAYGQ--NDIDDSMICAG-M 204
Pvannamei2        NEGNEQTVVLSKIIQHEDYN---GFTISNDISLLKFSQPLSFNDYVRAIDIPAQGHAASG----DQVSGWG-ALT-EGGSSPSALQKVSVPIVSDDECRDAYGQ--SDIEDSMICAG-V 205
Psalmonis1        TSGDEQKIAVSDITYHEKFA----SHGTNYDVCLLKLKSSLHFNEKVKPIALPEKDQEFIG----DVVVSGWG-TIS-SSGPPSPVLKAVTVQVVSDEDCSDAYYG--S-IDETMICAA-A 211
Scalcitrans1      SDVQIRNVNGKQFIFTHEIYGGNVGPHDIGLSIVEEAFDLNACLVMDLLPLPALTCLLANMKALAVVNSGWGRD---NSGSLPNTLQTLEVDIIGYTECKAAVPLDAPLADVNICSYT-- 206
Scalcitrans2      YNTGGELVEVKAFKFHENYNSG---TMKNDVAVIKLARPVKESATIRFVKLADKTPATG----TPAVVTGWGTTCFMACNTLPKTLQKVVVDIVDEKTCASSEYKYGSKIKPTMVCAY-- 197

                       ■ O    *                                 ■  O
R.dominica        LDGGKDSCQGDSGGPLVINGV----QYGIVSWG-YGCALPGYPGVYG--SVLAAKDFIDQFV----- 248
Phumanus          LEG-KDSCQGDSGGPLVDDGK----LAGVVSFG-MGCARRGKPGVYTNLANPEVAAFVNKYLKA--- 253
Pinterpunctella1  DVGGRDQCQGDSGGPLFHNRN----VVGICSWG-IGCADSFLPGVNAR--VSNYIGWIQANA----- 260
Pinterpunctella2  DVGGRDQCQGDSGGPLFHNRN----VVGICSWG-IGCADSFLPGVNAR--VFNYIGWIQANA----- 260
Msexta            DVGGRDQCQGDSGGPLYHNGV----VVGVCSWG-EECALARFPGVNAR--VSRFANWIRNNS----- 256
Dabbreviatus      AQGGRDSCQGDSGGPVIQNR----LAGIVSFG-AGCARAGLPGVYAS--IPGYRAWIRQNAGL--- 252
Agambiae1         KQGGTGTCRNDSGGPFVAEGK----LIGVVSWS-HECELAGYPGVYAR--VASVREWIRETSGV--- 273
Agambiae2         KQGGQDTCRQDSGGPFVAEGK----LIGVISWG-HECALAGYPGVYPR--VASVRDWIRTTSGV--- 274
Agambiae3         QQGGKDACQGDSGGPVAEDK----LIGVVSWG-AGCAQPGYPGVYAR--VAVVRDWIRETCGV--- 275
Agambiae5         KRGGKDACQGDSGGPLVVDGK----LVGVVSWG-FGCAMPGYPGVYAR--VAVVRNWVRENSGA--- 275
Aaegypti          AAGGKDSCQGDSGGPLVSGDK----LVGVVSWG-KGCALPNLPGVYAR--VSTVRQWIREVSEV--- 254
Pleniusculus      PEGGKDSCQGDSGGPLAASDTGSTYLAGIVSWG-YGCARAGYPGVYTE--VSYHVDWIKANAV---- 268
Pleptodactylus    PEGGKDSCQGDSGGPLAASDTGSTYLAGIVSWG-YGCARPGYPGVYTE--VSYHVDWIKANAV---- 237
Pvannamei1        PEGGKDSCQGDSGGPLACDTASTYLAGIVSWG-YGCARPGYPGVYAE--VSYHVDWIKANA----- 263
Pvannamei3        PEGGKDSCQGDSGGPLACDTSYLVGIVSWG-YGCARPGYPGVYAE--VSYHVDWIKANA----- 263
Pvannamei2        PEGGKDSCQGDSGGPLACDTGSTYLAGIVSWG-YGCARPNYPGVYAE--VSYHVDWIKANA----- 264
Psalmonis1        P--GKDSCQGDSGGPLAQDGT----LVGIVSWG-YGCAAPGYPGVYGK--VSKFIDWIAEHQ----- 264
Scalcitrans1      AGTKDGACNGDSGGPLVKNTKGGYELVGLVSWGYVGCASTQMPSIYTS--VASYKQWIADTIAAYKN 271
Scalcitrans2      AEDKD-ACQGDSGGPLVAGGK----LVGVVSWG-KGCALPAIPGVYAD--VPSLRTWIEKTAKEL-- 254
```

Fig. 1 Amino acid sequences of purified proteins or deduced trypsinogen proteins from invertebrates. *IVGG* indicates the N-terminal residues of the active enzymes starting at Ile 16 (as numbered in vertebrate trypsinogens). Six conserved cysteines are marked with *circles*; two additional cysteines are marked in *Diaprepes abbreviatus* and *Penaeus vannamei* sequences. Active-site, triad-position amino acids are indicated with an *asterisk* on the top, and amino acid residues in the specific binding pocket are indicated with a *black square* on the top

series of consecutive reactions, where the function of multiple proteinases of different classes is part of a coordinated physiological process. Digestion of food protein is a prime example (Neurath 1984; Ehrmann and Clausen 2004; Delcroix et al. 2006).

Because synthesis and storage of active proteinases impose a risk of tissue damage by hydrolysis of autologous protein, organisms must possess a mechanism controlling activation until needed for a precise function. This is achieved by a complex system that involves control of synthesis; transcription and translation, storage, secretion, and activation of zymogens. Several mechanisms of regulation of proteolytic enzyme activity and function are known in vertebrates, and recently some of those mechanisms have been found to share similarities to invertebrate models, such as mosquitoes, other hematophagous insects, and crustaceans (Carreira et al. 1996; Lehane et al. 1998; Sahin-Tóth 2000). Different mechanisms of regulation at transcriptional, post-transcriptional, and translational levels have been suggested to exist in invertebrates, either alone or combined, and that they can be organized spatially and/or sequentially (Xiong and Jacobs-Lorena 1995; Noriega and Wells 1999).

Trypsin genes and transcriptional regulation

Transcription of genes encoding trypsin in vertebrates and invertebrates has been suggested as a major regulatory mechanism. Even though a few invertebrate trypsinogen nucleotide sequences have been reported, trypsin genes appear to be commonly represented as families in many invertebrates (Table 2; Zwilling and Neurath 1981; Zhu

et al. 2000; Diaz-Mendoza et al. 2005). In the mosquito *A. gambiae*, trypsin genes are arranged as a tightly clustered gene family consisting of seven related coding sequences (Müller et al. 1995). The sheep blowfly *Lucilia cuprina* has a four-member multi-gene family of trypsins (Casu et al. 1994). Wang et al. (1995) suggested that the ancestors of dipterans and lepidopterans had only one trypsinogen gene, and that extra copies were gained by gene duplication.

Kollien et al. (2004) found only one trypsin gene in the hematophagous louse *P. humanus*. Incubation of the translation product in the presence of chymotrypsin yielded bands at 33 and 26 kDa, suggesting a novel and uncommon mechanism of trypsin activation in insects. The significance of this finding remains to be investigated. Within a species, the way an individual regulates the presence and activity of a digestive enzyme is varied, and sometimes, opposing. Lehane et al. (1998) reported two trypsins produced as preproenzymes in the opaque zone of the midgut of the horsefly *Stomoxys calcitrans*, which are activated after a blood meal. In contrast, in the yellow fever mosquito *Aedes aegypti*, the absence of trypsin zymogens allows the organism to develop a particular regulation mechanism (Noriega and Wells 1999).

How trypsin multi-gene families provides adaptive value is not clear; however, Bown et al. (1997) and Zhu and Baker (1999) suggested that the complex provides an efficient mechanisms for protein degradation and an adaptive advantage for insects feeding on plants containing trypsin inhibitors ("Inhibition"). Various regulatory mechanisms of trypsins have been suggested for different invertebrates, especially during gene transcription in the cell nucleus and during translation or post-transcription. The female yellow fever mosquito *A. aegypti* is an example, possessing several regulatory mechanisms, since trypsin expression is a complex biphasic mechanism dependent on tissue and gender (Kalhok et al. 1993); transcription of early trypsin, preceded by juvenile hormone and ingestion of a blood meal, induces an increase of proteolytic activity by facilitating its translation. Early trypsin shows a negligible endopeptidase activity in the midgut. Late trypsin transcription reaches a maximum 18–24 h after the blood meal (Barrilas-Mury et al. 1991; Barrilas-Mury and Wells 1993; Lu et al. 2006). The presence and amount of each isoenzyme depends on a refined control mechanism after ingestion of blood; the amount of mRNA of early and late trypsins is also found to vary (Noriega et al. 1994). The mRNA for early trypsin is present in suitably small concentrations, affecting translation of late trypsin after ingestion of blood, then the concentration of the early trypsin mRNA drops. This sequence of events is repeated in the next gonadotropic cycle, when mRNA returns to its original level before the blood meal. The mature, early trypsin is

synthesized as an active form without a trypsinogen stage (Noriega et al. 1996; Noriega and Wells 1999). It seems that this regulation allows the mosquito to assess the quality and amount of the blood protein by early trypsin proteolysis. For some species, trypsin genes have been studied to understand its regulatory mechanisms. Some of these mechanisms, as promoter sequences in malaria mosquito group *A. gambiae*, have been described and partially elucidated (Müller et al. 1993; Skavdis et al. 1996; Lehane et al. 1998).

Four genes, belonging to a clustered gene family coding for trypsin enzymes, were isolated from the Mediterranean fruit fly *Drosophila melanogaster* (Davis et al. 1985). The four genes in the family are transcribed in alternating orientations and the activation peptide sequence of the α gene is not similar to that of most trypsins reported, indicating a modified activation mechanism. In *Drosophila*, the promoter function was highly conserved during evolution and *Drosophila* transcription factors recognize promoter DNA sequences from other distantly related species, such as the black fly *Simulium vittatum* and the mosquito *A. gambiae* (Shen and Jacobs-Lorena 1998). In fact, some *cis*-acting DNA sequences, such as the TATA box, the initiator region TCAGT, and other upstream promoter elements, have been identified immediately upstream of non-Drosophila insect trypsin genes (Xiong and Jacobs-Lorena 1995; Harshman and James 1998). This means that transcriptional regulatory elements do exist in invertebrates, perhaps distributed in a ubiquitous manner, and that these elements, according to the authors' suggestions, are evolutionarily conserved.

In contrast to the *A. aegypti* mosquito mechanism, where trypsin gene transcription seems to be inducible, in the midgut epithelium of the *S. calcitrans* fly, a large amount of trypsin mRNA is stored even in unfed individuals. This supports evidence that trypsin genes may be constitutively transcribed and that control over production of digestive trypsin can be regulated during post-transcription (Lehane et al. 1998). In other species, studies have demonstrated that a series of constitutive and blood meal-induced trypsin genes are present in the gut of both genders during the stages of the life cycle, as in the *A. gambiae* mosquitoes (Müller et al. 1995).

Studies of trypsins in marine invertebrates, especially crustaceans, have shown various trypsin genes dispersed in different multi-gene families, such as the whiteleg shrimp *Penaeus vannamei*, where the number of introns in the three trypsin genes varies from zero to three, showing relatively conserved positions and, commonly, shorter sequences than those in vertebrate species (Klein et al. 1998). Comparisons of DNA sequences and cDNAs show three exons and two introns, while vertebrates commonly have five exons separated by four introns (Craik et al. 1984). In *P. vannamei*, three isotrypsins, A, B, and C,

found in the digestive gland were purified and characterized by molecular, biochemical, and kinetic parameters (Sainz et al. 2004a), confirming that digestive trypsin is a polymorphic enzyme, as anticipated by Klein et al. (1996) in a study of cDNA. *P. vannamei* isotrypsins are codified at two loci: Locus *β*, which is homozygous and yields isoenzyme C and Locus *α*, which is heterozygous, yielding isoenzymes A and B (Sainz et al. 2005). In shrimp species, Al-Mohanna et al. (1985) and Lehnert and Johnson (2002) proposed that fibrillar cells (F) of the midgut gland tubules are the only sites for the secretion of digestive enzymes, which must be synthesized as zymogen granules and stored in supranuclear vacuoles similar to the formation of zymogens in the pancreatic exocrine cells of vertebrates. Al-Mohanna and Nott (1989) and Icely and Nott (1992) found that digestive enzymes in the vacuoles are liberated periodically into the gut, depending on the species and nutritional conditions. To date, the published information concerning isolation of invertebrate trypsinogens is scarce. Hernández-Cortés et al. (1999b) reported a trypsinogen in the midgut gland of the crayfish *P. leniusculus*. Lehnert and Johnson (2002) and Sainz et al. (2004b) used immunological approaches to find digestive trypsins in the giant tiger prawn *Penaeus monodon* and the whiteleg shrimp *P. vannamei*, respectively, stored as zymogens, which is a common strategy of frequent-feeder species. This information is relevant because studies of the organization and sequence of genes alone may never anticipate information about post-transcriptional modifications of trypsinogens.

We hypothesize that animals having isotrypsins have adaptive advantages; for example, Atlantic salmon with isotrypsin TRP-2*92 are larger than salmon without it (Rungruangsak-Torrisen et al. 1999). These authors believe that the effect is a consequence of better food conversion and protein efficiency, at least under the conditions of the experiment. Some isoenzymes may be vestigial, which is not rare, according to Pils and Shultz (2004) who suggested, when analyzing complete metazoan genomes, that up to 15% of the members of all encoded enzymes families may have lost their catalytic activity. In pig, Barret et al. (1998) found that pepsin B has a weak, general proteolytic activity, about 4% of what occurs with pepsin A, and it does not contribute much to digestion.

Limited information concerning mechanisms of trypsin activity regulation in marine invertebrates has been obtained. It is generally accepted that proteinolytic digestive enzymes are transcriptionally regulated in mammals, with the quantity and quality of the protein in food being one of the major inducers (Lhoste et al. 1994). Studies of insect species also suggest the influence of food in trypsin activity as in the larvae of the red palm weevil *Rhynchoporus ferrugineus* (Alarcon et al. 2002) and in the mosquito *A. aegypti* (Noriega et al. 1994, 1996) indicating that one of

the trypsin genes is mainly regulated for transcription by the quality and quantity of meal. Similar mechanisms have been described for some crustaceans. Le Moullac et al. (1994) suggested that food protein quality induces synthesis of digestive enzymes (trypsin, chymotrypsin, and amylase) in the larvae of whiteleg shrimp *P. vannamei* and some interesting observations of this species have been made in our lab. After starvation for 24 h, the amount of trypsin mRNA increased significantly, followed by a steep decline after feeding, compared to shrimp that had not been starved (Sánchez-Paz et al. 2003). Starvation affected midgut gland trypsin activity in a similar way (Muhlia-Almazán et al. 2002). Thus, a decrease of both trypsin activity and mRNA by starvation suggests a strategy to reduce enzyme synthesis, and consequently, to save energy required for protein synthesis, as well as preventing potential autolysis. Such results suggest a mechanism of regulation during transcription by external factors. Additionally, this could be an effective strategy for handling the effects of medium-to-long-term starvation and avoiding deleterious effects on fitness. This knowledge opens a door for management of the digestive system for greater efficiency, at least in pond-raised shrimp.

Zymogens

Vertebrates synthesize trypsin as a non-active precursor called trypsinogen, which is commonly stored in intracellular organelles until it is activated-secreted into the intestine lumen for food protein digestion. The active trypsin form is achieved by the hydrolytic post-translational removal of a short, highly charged peptide in the amino terminus of the zymogen, the trypsinogen activation peptide (TAP), by specific cleavage between lysine/arginine and isoleucine residues (Male et al. 1995). The transition from the inactive to an active state is achieved by reorientation of the so-called activation domain: the new N-terminus (Ile 16) folds into a pocket and makes a salt bridge with Asp 194, which results in an ordered, conformational change in the activation domain that creates the substrate binding site (Fig. 1; Pasternak et al. 1999; Ehrmann and Clausen 2004). Activation of trypsin zymogen and cleavage of signal peptides are two of various post-translational controls that digestive proteinases undergo. During the 1950s, bovine trypsinogen was shown to be activated through the cleavage of a peptide Val-Asp-Asp-Asp-Asp-Lys-Ile located on the N-terminus of the protein. The heptapetide was cleaved autocatalytically and by its physiological activator, enteropeptidase (Chen et al. 2003). The Asp4 sequence was found conserved in several mammal species, including human (Guy et al. 1978; Emi et al. 1986), cow (Davie and Neurath 1955; Le Huerou et al. 1990), goat (Bricteux-Gregoire et al. 1971),

sheep (Schyns et al. 1969), pig (Louvard and Puigserver 1974), dog (Pinsky et al. 1985), cat (Steiner et al. 1997), and rat (MacDonald et al. 1982). However, the number of consecutive aspartyl residues seems to reach a maximum in TAPs of mammalian vertebrates, while in other vertebrates this number decreases. In this way, none of the TAP sequences found in bacteria (Kim et al. 1991), fungi (Rypniewsky et al. 1993), crustaceans (Hernández-Cortes et al. 1999b; Johnson et al. 2002) and insects (Müller et al. 1993; Wang et al. 1999; Kollien et al. 2004) possess an acidic residue immediately before the Lys or Arg prior to the conserved sequence Ile-Val-Gly-Gly of the amino terminus of mature trypsins (Zhu and Baker 1999). In the tunicates (Roach et al. 1997) and the ascidian *Botryllus schlosseri* (Pancer et al. 1996), TAP sequences have an Asp before the Lys23-Ile24 cleavage bond. Additional Asp residues were added during the course of vertebrate evolution: one in sea lamprey (Roach et al. 1997); two, three, or four in fish, amphibians, and birds (Shi and Brown 1990; Gudmundsdottir et al. 1993; Male et al. 1995; Wang et al. 1995; Douglas and Gallant 1998; Kurokawa et al. 2002); and four in mammals. This fact suggests two features: (1) progressive increase in selective pressure for such acidic residues during the course of vertebrate evolution (Chen et al. 2003), and (2) increasing number of Asp residues preceding the Lys23-Ile24, associated with a progressive decline in the tendency for autoactivation. Studies using small, synthetic peptides indicate that Asp4 significantly slowed hydrolysis of the Lys23-Ile24 bond by bovine trypsin, but enhanced the action of enteropeptidase. This suggests that, although Asp4 might play a protective role against spontaneous activation of the zymogen within the pancreas, it must constitute the signal for specific activation of trypsinogen by enteropeptidase in the duodenum (Chen et al. 2003). Peterson et al. (1994) found that activation peptides in the midgut gland of insects, such as *Manduca sexta*, lack the enteropeptidase consensus sequence (DDDK), which is specific for cleavage in most vertebrate trypsinogens. This observation was confirmed for the human body louse *P. humanus* (Kollien et al. 2004) and fruit fly *D. melanogaster* (Davis et al. 1985). Recent reports show evidence suggesting the existence of reversible zymogen activation in bacteria and vertebrates, this on/off switch of proteinase activity may be regulated by concentrated salts (Huang et al. 2001; Friedrich et al. 2003) or temperature (Spiess et al. 1999; Ehrmann and Clausen 2004). However, no reports concerning this activity have been published for invertebrate trypsins. Undoubtedly, the activation peptide is located at the amino acid end of the molecule, which, by selective pressure, is a useful feature because zymogen is synthesized beginning with the activation peptide for the emerging proteinase born inactive.

Endocrinology

One of the distinctive features in studies of insect gene regulation is the support provided by the rich tradition of insect endocrinology. Insect endocrinology is based on a well-established experimental logic that uses extirpation, ligation, transplantation, and administration of hormonally active compounds (Harshman and James 1998). Only a few studies about hormone control of genes expression or protein activation have been reported in invertebrates (Harshman and James 1998). Hence, endocrine control of trypsin secretion in invertebrates is far from being fully understood.

During the late 1970s, researchers found that median neurosecretory cells of the yellow fever mosquito *A. aegypti* contain factors that enhance the secretion of trypsin twofold (Briegel and Lea 1979), and that ecdysteroids secreted by the ovaries were capable of enhancing the secretion of proteases (Briegel and Lea 1977). Later, Graf et al. (1998) showed that trypsin synthesis was reduced to less than half of its normal output when mosquitoes were decapitated or ovariectomized, indicating that hormonal factors released from the brain and ovaries influenced the rate of trypsin synthesis. Besides, researchers found that juvenile hormone regulates early trypsin synthesis by controlling transcription in the same species (Noriega and Wells 1999). Early trypsin mRNA is absent in larvae, pupae, and newly emerged females. The advent of the juvenile hormone is the signal for the transcription of the early trypsin gene when the organism reaches maturation. Thus, protein in the blood meal is responsible for translation (Noriega et al. 1996).

In mosquitoes, flies, fleas, and biting midges, the trypsin-modulating oostatic factor, TMOF, a decapeptide that functions as a physiological signal that terminates trypsin biosynthesis, was added to feed the citrus weevil *D. abbreviatus*. This stopped trypsin activity in the midgut gland, as well as significantly decreasing larval weight, growth, and survival. This suggests that trypsin synthesis decreases during translation, probably under the influence of TMOF-like hormones (Yan et al. 1999; Borovsky 2003).

In decapod crustaceans, a neuropeptide, the crustacean hyperglycemic hormone (CHH), is synthesized by neurosecretory somata, grouped together in the X-organ. CHH is a pleiotropic hormone that plays a central role in regulation of glycemia (Ollivaux and Soyes 2000). Sedlemeier (1988) showed that CHH exhibits a secretagogue effect on digestive enzymes, particularly controlling amylase release by the midgut gland of the crayfish *Orconectes limosus*. To date, a possible regulatory effect of CHH over trypsin synthesis remains unknown.

In penaeids, neuroendocrine regulation of trypsin activity was suggested by Gibson and Barker (1979) and

Sedlmeier (1988). Release of pancreatic enzymes requires participation of gastrointestinal hormones, such as cholecystokinin (CCK), gastrin, secretin, and bombesin. In mammals, regulation by hormones plays a role in appetite satiation, pain perception, and neuronal transmission (Gibbs et al. 1979). In insects, some substances of the CCK-gastrin family have been identified, but efforts to purify gastrointestinal hormones from crustaceans have failed and knowledge of these substances is still limited. Since Larson and Vigna (1983) proposed that CCK-gastrin evolved in invertebrates as a neural peptide and was subsequently used by vertebrates as a regulatory peptide in the nervous and gastrointestinal endocrine systems, we might expect these gastrointestinal materials in crustaceans would be widespread. These authors found that the crab *Cancer magister* had structurally similar peptides to the bioactive C-terminal amino acid sequence common to gastrin and CCK. Also, strong recognition for CCK-like peptides was detected when using antibodies generated against mammalian CCK8 in rock crabs *Cancer borealis* (Christie et al. 1995). Sulfakinins were recently discovered as multi-functional peptides with homologies to G-CCK peptides, which have, so far, been found only in insects and crustaceans; they inhibit food intake and stimulate release of enzymes, such as α-amylase, among others (Nichols et al. 1988; Torfs et al. 2002; Meyering-Vos and Müller 2007). Johnsen et al. (2000) purified three novel members of the sulfakinin family from an extract of the central nervous system of the giant tiger prawn *P. monodon*. Immuno-cytochemical studies with sulfakinin antisera showed a sparse neuronal distribution pattern, similar to that of insects. This suggests a role in neurotransmission or neuromodulation, although a direct role in digestive system physiology remains to be clarified, as has been shown by in vitro assays of the sulfakinins in orthopteran and blattodean insects (Johnsen et al. 2000; Meyering-Vos and Müller 2007). In the American crayfish *O. limosus*, Resch-Sedlmeier and Sedlmeier (1999) detected proteinase and amylase activity in the midgut gland after incubation with CCK-8, gastrin, secretin, and bombesin. These results may support the assumption that Crustacea may possess endogenous factors similar to vertebrate hormones, at least, that the pertinent receptors have the ability to recognize some hormones.

Inhibition

One of the most important mechanisms regulating proteolysis of autologous proteins is the synthesis of endogenous, cognate inhibitors that maintain inactivity in spontaneously activated enzymes. Inhibitors play a major role in regulating enzyme activity in plants, animals, and microorganisms. Avoiding unwanted autohydrolysis is so essential to the organism that considerable energy must be invested in producing proteinase inhibitors. Some data reveal the importance of inhibitors in a broad range of processes. Inhibitors hamper proteinase activity by forming inhibitor–enzyme complexes. The resulting inhibitor–proteinase complex is inactive because the catalytic site of the enzyme is blocked. According to Krowarsch et al. (2003), three mechanisms evolved for inhibitors to block access of the substrate to the enzyme. (1) Canonical inhibitors bind to the enzyme through an exposed, convex-binding loop, which is complementary to the active site of the enzyme. The mechanism of inhibition in this group is always very similar and resembles that of the ideal substrate. (2) Non-canonical inhibitors interact through their N-terminal segment. There are also extensive secondary interactions outside the active site, contributing significantly to the strength, speed, and specificity of recognition. (3) Serpins, similar to canonical inhibitors, interact with their target proteinases in a substrate-like manner; however, cleavage of a single peptide bond in the binding loop leads to dramatic structural changes.

Given the importance of proteinase inhibitors in modulating biological processes, some invertebrate trypsin inhibitors are intended for practical medical applications. Examples are the recombinant leech-derived tryptase inhibitor-2PL (LDTI-2PL) or the tick *Boophilus microplus* trypsin inhibitors (BmTIs) that are intended to study the effects of serine proteinase inhibitors on the inflammatory process induced by administration of carrageen into the rat pleural cavity and on release of kinins in pleural exudates (Malavazi-Piza et al. 2004). Proteinase inhibitors are extensively used in food technology. An example is surimi production, where hampering the muscle proteinases affects gel formation. (Comprehensive reviews of proteinase inhibitors in food technology can be found in García-Carreño et al. (2000) and García-Carreño and Hernández-Cortés (2000)). Because some invertebrate trypsin inhibitors are short polypeptides, ~35 amino acid residues, they are studied to modify proteinases of practical interest, either depressing or enhancing function, for noxious or useful proteinases (Boigegrain et al. 1994). Some insect species frequently feed on plants, synthesizing and storing proteinase inhibitors to disrupt digestive enzymes of pests and predators; thus proteinase inhibitors play a significant role in the regulation of proteolysis, whether the target enzymes are of exogenous or endogenous origin; hence, some researchers suggested that insect digestive systems adapt to the presence of inhibitors (Cherqui et al. 2001; De Leo et al. 2001; Diaz-Mendoza et al. 2005).

Inhibitors of proteinases are exceptional molecules because, in spite of being proteins, they are negligibly hydrolyzed by the specific cognate enzyme. In many arthropod species, serine proteinase inhibitors have been identified (Cherqui et al. 2001). The serine proteinase inhibitors

produced by plants and specific for invertebrate trypsins include inhibitors belonging to the Kunitz family as soybean trypsin inhibitor (SBTI) and soybean Kunitz trypsin inhibitor (SKTI) and the Kazal family in *Schistocerca gregaria*, *A. leptodactylus*, *A. fluviatilis*, *Pascifastacus leniusculus*, and *Helicoverpa zea* (Volpicella et al. 2003; Fodor et al. 2005) in arthropods. Such inhibitors are highly specific, possessing, at their surface, one or more peptide bonds made with the α-carboxylic side of Arg known as reactive sites. The reactive site specifically interacts with the active site of the cognate enzyme. Both natural and synthetic inhibitors inhibit trypsin. The extent of the inhibition (K_i = molar) for trypsin goes from mM (millimolar) to fM (femtomolar), indicating that quite different affinities between them are possible (Zollner 1993).

As proteases, inhibitors for digestive enzymes seem to have arisen early in evolution in invertebrates. The structure of inhibitors has been described and categorized. Inhibitors belong to ten families comprised of molecules with homologies, both in sequence, but more importantly, in the number and array of disulfide bridges that yield a particular topology. Each family has a number and position of conserved disulfide bridges. Several inhibitors for trypsin have been described in a sea anemone *Anemonia sulcata* (Tschesche and Kolkenbrock 1984). For a comprehensive review of inhibitors, see Laskowski and Kato (1980).

The presence of proteinaceous inhibitors for digestive proteinases gives some clues about mechanisms of regulation of activity of digestive enzymes. Some inhibitors have been described in mammals and are thought to control the enzyme activity in spontaneous activation of the matching proteinase before secretion to avoid the unwanted hydrolysis of proteinaceous intracellular components. In crustaceans, an inhibitor for trypsin isolated from the midgut gland of the whiteleg shrimp *P. vannamei* was described (García-Carreño et al. 1998). Inhibitors for trypsin have been found in four species of this penaeid genus (de Albuquerque et al. 2002). These inhibitors constrain, to different degrees, trypsins in the same species and, at some lesser rate, trypsins in related species. In penaeid species, the presence of inhibitors synthesized in the same organ as trypsin recall the mechanism of regulation of digestive enzymes in mammals.

Trypsin inhibitors in penaeids reduced the activity of paralogous and orthologous trypsins. In this review, we used the nomenclature of Southan (2000), where paralogues are homologue enzymes or inhibitors co-synthesized in the same organisms and orthologues are homologue enzymes or inhibitors found within other taxa. Inhibitors in four penaeid species are low molecular weight proteins and inhibition of autologous and orthologous trypsins are proportional to the concentration of the inhibitor (de Albuquerque et al. 2002). Although the four closely related shrimp species have similar total protein and total proteinase activity, the composition and molecular masses of the enzymes were distinct. The differences were more conspicuous when comparing protein composition in the midgut gland.

Trypsins as team players

When studying digestive physiology, new questions arise. Why does an organism need several different proteinases? Why is one class or specificity not enough? Why do animals evolve with several classes and specificities? Of course, the simple answer is: to increase its biological fitness. The formal cause is that organisms having multiple classes of proteinases, such as trypsin and chymotrypsin, can enhance hydrolysis of food protein, exposing a variety of essential amino acids to the action of carboxypeptidases, which will eventually be released for absorption. Thus, the mere fact of having a wider spectrum of proteinases increases the efficient use of food protein.

The advantage of having isotrypsins was already discussed. But, a deeper analysis of the presence of several classes of proteinases is required. Teschke and Saborowski (2005) found that in two caridean shrimp, contrary to expectations, trypsin activity represented a small portion of proteinases activity, where maximum activity was reached at acid pH, far from the pH for serine proteinases. Moreover, a small percentage of the individuals had high trypsin activity in the midgut gland and a large percentage had low trypsin activity. In individuals with low trypsin activity, a cysteine proteinase was the substitute contributor to total proteinase activity. Researchers concluded that eucarids may have different traits for using food protein. For example, the California spiny lobster *Panulirus interruptus* contained isotrypsins and isochymotrypins as major proteinases, but another band of activity when analyzing activity by gel electrophoresis at acid pH (Celis-Guerrero et al. 2004). At that time, this seemed unusual, but was later supported by evidence obtained from the European lobster *Homarus gammarus*, where most of the proteinolytic activity was an enzyme working at acid pH that was completely inhibited by pepstatin A (Navarrete del Toro et al. 2006).

Astacin family members contain proteases that are widely distributed in nature, in bacteria, crustaceans, nematodes, cnidaries, mollusks, and several vertebrates). The family received its name after purification, biochemical characterization, and sequencing of a protease from the noble crayfish *Astacus astacus* L. (EC 3.4.24.21) (Titani et al. 1987). This enzyme is synthesized in the crayfish midgut gland (Vogt et al. 1989). It is stored extracellularly as an active proteinase in the stomach-like cardia, and is thought to be a digestive enzyme in a variety of

crustaceans. Because of the fact that astacins are present in a tissue-specific manner in mature animals and are temporally and spatially present in developing systems, it has being proposed that its role is in mature and developing animals. Exploration of the structure, function, and regulation of these enzymes is at an early stage. A similar situation is found with other enzymes, such as cathepsins, which are distributed in several phyla. However, to date there is no evidences of a close link between trypsin functions and astacin or cathepsin functions, and if such a connection exists, it remains to be understood.

An interesting example of a coordinated network between proteinases occurs in the flatworm *Schistosoma mansoni*. Delcroix et al. (2006) used class-specific inhibitors and RNA interference techniques and determined "…that initial degradation of host blood proteins is ordered, occasionally redundant, and substrate specific…" in a proteinase network in which major proteinases function. They suggest that the protease network is conserved throughout invertebrate evolution.

Digestive proteinases do not work alone, with some exceptions, such as porcine pepsin B. Proteinases participate in food protein digestion in a coordinated manner, at least for endopeptidases, followed by exopeptidases. Other types or coordination among proteinases remain to be fully understood.

Trypsins from extremophiles

Temperature is one of the most important environmental factors sustaining life on Earth. Low temperatures do not favor chemical reaction rates catalyzed by enzymes (D'Amico et al. 2002). Organisms living under cold conditions require a variety of fundamental adaptations, including expression of enzymatic activities at appropriate levels, which must work under increased viscosity of the medium induced by low temperatures to allow the organism metabolic fluxes more or less comparable to those exhibited by closely related mesophilic species (Zecchinnon et al. 2001). Several taxa are devoid of temperature regulation, living close to 0 °C, referred as psychrophiles, such as bacteria, yeast algae, plants, insects, marine and terrestrial invertebrates, and fish (Feller and Gerday 2003). In terms of species diversity, microorganisms are the most abundant psychrophiles.

Even though the largest proportion of biomass on Earth is generated at cold temperatures, to our knowledge, there are only some reports related to invertebrate psychrophilic enzymes, for example: lysozyme from the tobacco hornworm moth *M. sexta* (Sotelo-Mundo et al. 2007), DNAse II from the Iceland scallop *Chlamys islamica* (Øverbø and Myrnes 2006), citrate synthase from the isopods *Idotea bal-*

tica and *Idotea emarginata* (Salomon and Buchholz 2000), alkaline phosphatase from the Arctic shrimp *Pandalus borealis* (de Backer et al. 2002), and glutathion *S*-transferase from the Icelandic scallop *Chlamys islandica* (Myrnes and Nielsen 2007).

With the interest for novel industrial applications in recent years, attention has focused on cold-adapted trypsins. However, the molecular basis of enzyme adaptation to low temperatures is still poorly understood, probably because of the limited number of available amino acid sequences and structural data. Leiros et al. (2000) reported an alignment of 27 vertebrate trypsin amino acid sequences and suggested that the increased substrate affinity of the psychrophilic trypsins compared to mesophilic forms is probably achieved by the optimization of electrostatic complementarity between the binding site and substrate. The generally reduced electrostatic surface potential of all cold trypsins, especially in the C-terminal domain, is also suggested as a factor that increases the substrate affinity by a more efficient guidance and orientation of the substrate prior to binding (Smalas et al. 1994).

Additionally, brachyurins (serine proteases isolated from several decapod crustaceans) show chymotrypsin-like, trypsin-like, and elastase-like activities towards synthetic substrates (Kristjáansdóttir and Gudmundsdóttir 2000) plus all the characteristics of cold-adapted enzymes (Kristjáansdóttir 1999). These enzymes are members of Clan SA of the S1 family of serine endopeptidases and include the euphaulysin from krill (Turkiewicz et al. 1991). Two distinguishing features of euphaulysins are their ability to cleave collagen and its sensitivity to autolysis (Gudmundsdóttir 2002). Considering, the recent finding of this enzyme and its commercial interest, it would not be surprising if new reports on the isolation and characterization of trypsins from invertebrates inhabiting cold environments are published in the near future.

After the discovery of a new family of crabs and dense aggregations of bresilioid shrimp at hydrothermal vents in the eastern Pacific Ocean (Williams 1980; Williams and Chace 1982) the number of publications dealing with taxonomy, ecological physiology, and distribution of invertebrates at these sites in the deep sea has increased markedly (Martin and Haney 2005). It is now known that the invertebrate fauna, distinguished by its extreme temperatures, low oxygen, and toxic hydrogen sulfide and heavy metals, is extraordinarily complex. However, to our knowledge only a few studies have been devoted to understanding the biochemical characteristics of enzymes from invertebrates living in such remote and extreme biotopes. References to enzymes of organisms in extreme environments have largely focused on hyperthermophilic microorganisms. As in the case of cold-adapted enzymes, in the near future, we would expect publications related to the isolation, purification,

and biochemical characterization of enzymes from invertebrates adapted to severe water environments.

Discussion

According to Neurath (2001), "Trypsin has fascinated biochemists, enzymologists, and biologists because it was one of first proteases to be identified and it plays an important digestive and regulatory role in a variety of physiological processes". Trypsin possesses two main functions, hydrolysis of food protein and activation of zymogens of trypsin and other digestive proteases. Understanding trypsin functions allowed us to learn about active site mechanisms, domain organization, evolution, phylogeny, and irreversible signal transduction, as well as zymogen activation. Learning about trypsins is possible with the development of new techniques. The earliest approach was biochemical characterization. Mammal trypsin was the first protease crystallized for understanding molecular characterization. Also, characterization of biochemical specificities was done at that moment by synthesizing specific substrates and inhibitors. With the advent of electrophoresis, researchers discovered that trypsins were synthesized as isoforms, which lead to a paradigm shift that several genes were involved in isotrypsins synthesis. Also, evolution and phylogeny of trypsin could be assessed by a combination of electrophoresis techniques and specific antibodies that elicited reactions to one enzyme and compared cross-reactions with others. Of course, several questions have arisen about the function of isoenzymes. During the era of gene cloning, another paradigm shift occurred in our understanding the genetics of trypsin when the number of genes expanded beyond the ones inferred from isoenzyme studies. Development of molecular biology techniques also opened the possibility of understanding the multi-gene families of proteases. Techniques, such as screening genomic libraries, cloning genes, and analyzing complete genomes, helped make important contributions in finding new functions for proteases genes (Heger and Ponting 2008).

Since the dawn of genomics, the picture of molecular genetics of trypsin continued to change as researched realized that the number of trypsin genes was even greater and that some transcribe in the digestive system cells while others transcribe in other tissues. This increased the opportunity of finding and studying specific domains in genes and their deduced proteins, and establishing integrative studies, such as that developed by Shah et al. (2008) for structural-functional analyses of three dimensional structures. Microarray gene expression profiles, protein–protein interactions, protein and genes databases, and bioinformatic tools allowed the introduction of functional residue clustering (FRC) as a new tool for transferring, in vivo, substrate specificity in a quantitative manner that is highly sensitive and specific. This novel method assigns putative inhibitor relationships for proteases, providing evidence that explains evolution of new functions for serine-protease homologues in immune response by using high-throughput data. Trypsin studies moved from crude extracts to pure proteins, to characterization of molecules and isoforms, to characterization of genes. In doing this, the vision changed and knowledge increased, each time supported by more analytical and explicative methods.

Proteinases evolved to fulfill advanced physiological functions; we now know that trypsins are involved in a plethora of activities in invertebrates, such as activation of prophenoloxidase synthesized in hemocytes and as a defense response to invading microorganisms in the so-called innate immunity of invertebrates that uses melanin to encapsulate invaders.

Obviously, the key role that trypsin plays in living organisms is not restricted to the more studied species; in fact, there are incalculable insect and crustacean species that show special proteins. For example, insects are agents of agricultural or medical problems and targets of several studies, including gene regulation. For crustaceans, many species can synthesize highly efficient proteinases adapted to extreme conditions. Still, several problems in studying proteinases in invertebrates must be addressed. It is difficult to obtain information about these enzymes because of the small size of invertebrates, the molting process, and the lack of species-specific substrates suitable for some evaluations of enzyme activity. A limitation in genetic studies is the absence of adequate in vivo expression systems to evaluate crustacean proteinases. The importance of several insect species in commercial production, plague control, and human diseases and crustaceans in ecology and as food for human consumption has increased interest in the science community and government organizations to find a deeper comprehension of factors involved in digestion.

In summary, digestive trypsins are members of a network of proteinases that act in a coordinated manner. They serve to activate zymogens, start digestion with endopeptidases and exopeptidases; some work where others are inhibited by specific inhibitors in food and some specifically hydrolyze some protein substrates. More complex networks and coordination types are expected to be found in the near future. Perhaps, Hans Neurath (1994) was right when he announced that the study of proteinases was in a second golden era.

References

Alarcón FJ, Martínez TF, Barranco P, Cabello T, Díaz M, Moyano FJ (2002) Digestive proteinases during development of larvae of red palm weevil, *Rhynchophorus ferrugineus* (Olivier, 1790) (Coleoptera: Cruculionidae). Insect Biochem Mol Biol 32:265–274

Al-Mohanna SY, Nott JA (1989) Functional cytology of the hepatopancreas of *Penaeus semisulcatus* (Crustacea: Decapoda) during the molt cycle. Mar Biol 101:535–544

Al-Mohanna SY, Nott JA, Lane D (1985) Mitotic E- and secretory F-cells in the hepatopancreas of the shrimp *Penaeus semisulcatus* (Crustacea: Decapoda). J Mar Biol Assoc UK 65:901–910

Amino R, Tanaka AS, Schenkman S (2001) Triapsin, an unusual activatable serine proteinase from the saliva of the hematophagous vector of Chagas disease *Triatoma infestans* (Hemiptera: Reduviidae). Insect Biochem Mol Biol 31:465–472

Arnon R, Neurath H (1969) An immunological approach to the study of evolution of trypsins. Proc Natl Acad Sci USA 64:1623–1628

Baptista A, Jonson P, Hough E, Petersen S (1998) The origin of trypsin: evidence for multiple gene duplications in trypsins. J Mol Evol 47:353–362

Barillas-Mury CV, Graf R, Hagedorn HH, Wells M (1991) cDNA and deduced amino acid sequence of a blood meal-induced trypsin from the mosquito *Aedes aegypti*. Insect Biochem 21:825–831

Barillas-Mury CV, Wells MA (1993) Cloning and sequencing of the blood meal-induced late trypsin gene from the mosquito *Aedes aegypti* and characterization of the upstream regulatory region. Insect Mol Biol 2:7–12

Barret A, Rawlings N, Woessner J (1998) Handbook of proteolytic enzymes. Academic Press, San Diego

Boigegrain R, Mattras H, Brehelin M, Coletti-Previero M (1994) Invertebrate proteinase inhibitors. Pure Appl Chem 66:1–7

Borovsky D (2003) Trypsin-modulating oostatic factor: a potential new larvicide for mosquito control. J Exp Biol 206:3869–3875

Bown DP, Wilkinson HS, Gatehouse JA (1997) Differentially regulated inhibitor-sensitive and insensitive proteinase genes from phytophagous insect pest *Helicoverpa armigera*, are members of complex multigene families. Insect Biochem Mol Biol 27:625–638

Bricteux-Gregoire S, Schyns R, Florkin M (1971) Purification, properties and N-terminal sequence of goat trypsinogen. Biochim Biophys Acta 229:123–135

Briegel H, Lea AO (1977) Ecdysone, the ovarian hormone and intestinal proteases in mosquitoes. Experientia 33:813

Briegel H, Lea AO (1979) Influence of the endocrine system on tryptic activity in female *Aedes aegypti*. J Insect Physiol 25:227–230

Carreira S, Fueri C, Chaix JC, Puigserver A (1996) Dietary modulation of the mRNA stability of trypsin isozymes and the two forms of secretory trypsin inhibitor in the rat pancreas. Eur J Biochem 239:117–123

Casu RE, Jarmey JM, Elvin CM, Eisemann CH (1994) Isolation of a trypsin-like serine proteinase gene family from the sheep blowfly *Lucilia cuprina*. Insect Mol Biol 3:159–170

Celis-Guerrero L, García-Carreño FL, Navarrete del Toro MA (2004) Characterization of proteases in the digestive system of spiny lobster (*Panulirus interruptus*). Mar Biotechnol 6:262–269

Chen JM, Kukor Z, Le Marechal C, Toth M, Tsakiris L, Raguenes O, Ferec C, Sahin-Toth M (2003) Evolution of trypsinogen activation peptides. Mol Biol Evol 20:1767–1777

Cherqui A, Cruz N, Simoes N (2001) Purification and characterization of two serine proteinase inhibitors from the hemolymph of *Mythimna unipuncta*. Insect Biochem Mol Biol 31:761–769

Christie AE, Baldwin D, Turrigiano G, Graubard K, Marder E (1995) Immunocytochemical localization of multiple cholecystokinin-like peptides in the stomatogastric system of the crab *Cancer borealis*. J Exp Biol 198:263–271

Craik CS, Choo QL, Swift GH, Quinto C, MacDonald RJ, Rutter WJ (1984) Structure of two related rat pancreatic trypsin genes. J Biol Chem 259:14255–14264

D'Amico S, Claverie P, Collins T, Georlette D, Gratia E, Hoyoux A, Meuwis MA, Feller G, Gerday C (2002) Molecular basis of cold adaptation. Phil Trans R Soc Lond 357B:917–925

Davie EW, Neurath H (1955) Identification of a peptide released during autocatalytic activation of trypsinogen. J Biol Chem 212:515–529

Davis CA, Riddell DC, Higgins MJ, Holden JJ, White BN (1985) A gene family in *Drosophila melanogaster* coding for trypsin-like enzymes. Nuuleic Acids Res 13:6605–6619

de Albuquerque C, Muhlia-Almazán A, Hernández-Cortes P, Garcia-Carreño FL (2001) Proteinases from marine organisms. In: Fingerman M, Nagabhushanam R (eds) Recent advances in marine biotechnology. Science Publishers, Plymouth, pp 209–238

de Albuquerque C, García-Carreño FL, Navarrete del Toro MA (2002) Trypsin and trypsin inhibitors from Penaeid shrimp. J Food Biochem 26:233–251

de Backer M, McSweeney S, Rasmussen HB, Riise BW, Lindley P, Hough E (2002) The 1.9 Å crystal structure of heat-labile shrimp alkaline phosphatase. J Mol Biol 318:1265–1274

De Haën C, Neurath H, Teller DC (1975) The phylogeny of trypsin-related serine proteinases and their zymogens. New methods for the investigation of distant evolutionary relationships. J Mol Biol 92:225–259

De Leo F, Bonadé-Bottino M, Ceci LR, Gallerani R, Jouanin L (2001) Effect of a mustard trypsin inhibitor expressed in different plants on three lepidopteran pests. Insect Biochem Mol Biol 31:593–602

Delcroix M, Sajid M, Caffrey CR, Lim K-C, Dvorak J, Hsieh I, Bahgat M, Dissous C, McKerrow JH (2006) A multienzyme network functions in intestinal protein digestion by a platyhelminth parasite. J Biol Chem 284:39316–39329

Diaz-Mendoza M, Ortego F, García de Lacoba M, Magaña C, de la Poza M, Farinós GP, Castañera P, Hernández-Crespo P (2005) Diversity of trypsins in the Mediterranean corn borer *Sesamia nonagrioides* (Lepidoptera: Noctuidae), revealed by nucleic acid sequences and enzyme purification. Insect Biochem Mol Biol 35:1005–1020

Dionysius DA, Hoek KS, Milne JM, Slattery SL (1993) Trypsin-like enzyme from sand crab (*Portunus pelagicus*): purification and characterization. J Food Sci 58:780–784

Douglas SE, Gallant JW (1998) Isolation of cDNAs for trypsinogen from the winter flounder, *Pleuronectes americanus*. J Mar Biotechnol 6:214–219

Ehrmann M, Clausen T (2004) Proteolysis as a regulatory mechanism. Ann Rev Genet 38:709–724

Emi M, Nakamura Y, Ogawa M, Yamamoto T, Nishide T, Mori T, Matsubara K (1986) Cloning, characterization and nucleotide sequences of two cDNAs encoding human pancreatic trypsinogens. Gene 41:305–310

Feller G, Gerday C (2003) Psychrophilic enzymes: hot topics in cold adaptation. Nat Rev Microbiol 1:200–208

Figueiredo MSRB, Kricker JA, Anderson AJ (2001) Digestive enzyme activities in the alimentary tract of redclaw crayfish, *Cherax quadricarinatus* (Decapoda, Parastacidae). J Crust Biol 21:334–344

Fletcher TS, Alhadeff M, Craik CS, Langerman C (1987) Isolation and characterization of a cDNA encoding rat cationic trypsinogen. Biochemistry 26:3081–3086

Fodor K, Harmat V, Hetenyi C, Kardos J, Antal J, Perczel A, Patthy A, Katona G, Graf L (2005) Extended intermolecular interactions in a serine proteinase–canonical inhibitor complex account for strong and highly specific inhibition. J Mol Biol 350:156–169

Friedrich R, Panizzi P, Fuentes-Prior P, Richter K, Verhamme I, Anderson PJ, Kawabata S-I, Huber R, Bode W, Bock PE (2003) Staphylocoagulase is a prototype for the mechanism of cofactor-induce zymogen activation. Nature 425:535–539

Galgani FG, Benyamin Y, Van Wormhoudt A (1985) Purification, properties and immunoassay of trypsin from the shrimp *Penaeus japonicus*. Comp Biochem Physiol 81B:447–452

García-Carreño FL, Carrillo O, Navarrete MA (1998) Control of digestive functions in shrimp. I. An inhibitor of the trypsin activity in the hepatopancreas. In: von Vaupel Klein JC, Schram FR (eds) Proceedings of fourth international crustacean congress, Brill, Leiden, pp 915–922

García-Carreño F, An H, Haard N (2000) Protease inhibitors in food processing. In: Michaud D (ed) Recombinant protease inhibitors in plants, pp 215–223. Landes Biosciences and http://Eurekah.com, Georgetown. TX, pp 214–220

García-Carreño F, Hernández-Cortés P (2000) Use of protease inhibitors in seafood products In: Naard N, Simpson B (eds) Seafood enzymes. Marcel Dekker, New York, pp 531–547

Gibbs J, Fauser DJ, Rowe EA, Rolls BJ, Rolls ET, Maddison SP (1979) Bombesin suppresses feeding in rats. Nature 282:208–210

Gibson R, Barker PL (1979) The decapod hepatopancreas. Oceanogr Mar Biol Ann Rev 17:285–346

Gooding RH (1975) Digestive enzymes and their control in haematophagus arthropods. Acta Trop 32:96–111

Graf R, Lea AO, Briegel H (1998) A temporal profile of the endocrine control of trypsin synthesis in the yellow fever mosquito, *Aedes aegypti*. J Insect Physiol 44:451–454

Gudmundsdóttir A (2002). Cold-adapted and mesophilic brachyurins. Biol Chem 383:1125–1131

Gudmundsdóttir A, Gudmundsdóttir E, Oskarsson S, Bjarnason JB, Eakin AK, Craik CS (1993) Isolation and characterization of cDNAs from Atlantic cod encoding two different forms of trypsinogen. Eur J Biochem 217:1091–1097

Guy O, Lombardo D, Bartlet DC, Amic J, Figarella C (1978) Two human trypsinogens. Purification, molecular properties, and N-terminal sequences. Biochemistry 17:1669–1675

Harshman LG, James AA (1998) Differential gene expression in insects: transcriptional control. Annu Rev Entomol 43:671–700

Heger A, Ponting CP (2008) Evolutionary rate analyses of orthologs and paralog from 12 *Drosophila* genomes. Genome Res 17:1837–1849

Hehemann JH, Redecke L, Perbandt M, Saborowski R, Betzel C (2007) Crystallization and preliminary X-ray diffraction studies of trypsin-like proteinases from the gastric fluid of the marine crab *Cancer pagurus*. Acta Crystallograph Sect F Struct Biol Cryst Commun 1:242–245

Hernández-Cortés P, Quadros W, Navarrete del Toro A, Portillo G, Colado G, García-Carreño FL (1999a) Rate of ingestion and proteolytic activity in digestive system during continuous feeding of juveniles shrimps. J Appl Aquacult 9:35–45

Hernández-Cortes P, Cerenius L, García-Carreño FL, Soderhall K (1999b) Trypsin from *Pacifastacus leniusculus* hepatopancreas: purification and cDNA cloning of the synthesized zymogen. Biol Chem 380:499–501

Hu G, Leger RJ (2004) A phylogenomic approach to reconstructing the diversification of serine proteinases in fungi. J Evol Biol 17:1204–1214

Huang X, Knoell C, Frey G, Hazegh-Azam M, Tashjian AH, Hedstrom L, Abeles RH, Timasheff SN (2001) Modulation of recombinant human prostate-specific antigen: activation by Hofmeister salts and inhibition by azapeptides. Biochemistry 40:11734–11741

Huber R, Kukla D, Bode W, Schwager P, Bartels K, Deisenhofer J, Steigemann W (1974) Structure of the complex formed by bovine trypsin and bovine pancreatic trypsin inhibitor II. Crystallographic refinement at 1.9 A resolution. J Mol Biol 89:73–101

Icely JD, Nott JA (1992) Digestion and absorption: digestive system and associated organs. In: Harrison FW (ed) Microscopic anatomy of invertebrates. Wiley-Liss, New York, pp 147–202

Johnsen AH, Duve H, Davey M, Hall M, Thorpe A (2000) Sulfakinin neuropeptides in a crustacean. Isolation, identification and tissue localization in the tiger prawn *Penaeus monodon*. Eur J Biochem 267:1153–1160

Johnson SC, Ewart KV, Osborne JA, Delage D, Ross NW, Murray HM (2002) Molecular cloning of trypsin cDNAs and trypsin gene expression in the salmon louse *Lepeophtheirus salmonis* (Copepoda: Caligidae). Parasitol Res 88:789–796

Kalhok SE, Tabak LM, Prosser DE, Brook W, Downe AER, White BN (1993) Isolation, sequencing and characterization of two DNA clones coding for trypsin-like enzymes from the midgut of *Aedes aegypti*. Insect Mol Biol 2:71–79

Keil B, Dlouha V, Holeysovsky V, Sorm F (1968) Hypothesis of three-dimensional arrangement of polypeptide chain in trypsin. Coll Czch Chem Comm 33:2307–2315

Kim JC, Cha SH, Jeong ST, Oh SK, Byun SM (1991) Molecular cloning and nucleotide sequence of *Streptomyces griseus* trypsin gene. Biochem Biophys Res Commun 181:707–713

Kishimura H, Hayashi K (2003) N-terminal amino acid sequence of trypsin from the pyloric ceca of starfish *Asterias amurensis*. Fisheries Sci 69:867–869

Klein B, Le Moullac G, Sellos D, Van Wormhoudt A (1996) Molecular cloning and sequencing of trypsin cDNAs from *Penaeus vannamei* (Crustacea, Decapoda): use in assessing gene expression during the moult cycle. Int J Biochem Cell Biol 28:551–563

Klein B, Sellos D, Van Wormhoudt A (1998) Genomic organization and polymorphism of a Crustacean trypsin multi-gene family. Gene 216:123–129

Kollien AH, Waniek PJ, Prols F, Habedank B, Schaub GA (2004) Cloning and characterization of a trypsin-encoding cDNA of the human body louse *Pediculus humanus*. Insect Mol Biol 13:9–18

Kossiakoff AA, Chambers JL, Kay LM, Stroud RM (1977) Structure of bovine trypsinogen at 1.9 Å resolution. Biochemistry 16:654–664

Kristjánsdóttir KS (1999) Isolation and characterization of a broad specificity collagenolytic protease from Antarctic krill. M.Sc thesis, University of Iceland, Iceland

Kristjánsdóttir S, Gudmundsdóttir A (2000) Propeptide dependent activation of the Antarctic krill euphauserase precursor produced in yeast. Eur J Biochem 267:2632–2639

Krowarsch D, Cierpicki T, Jelen F, Otlewski J (2003) Canonical protein inhibitors of serine proteinases. Cell Mol Life Sci (CMLS) 60:2427–2444

Kuehne W (1876) Ueber das trypsin (Enzyme des Pankreas). Heidelberg Nat Med Ver 1:194–198

Kurokawa T, Suzuki T, Ohta H, Kagawa H, Tanaka H, Unuma T (2002) Expression of pancreatic enzyme genes during the early larval stage of Japanese eel, *Anguilla japonica*. Fish Sci 68:736–744

Kvamme BO, Frost P, Nilsen F (2004) The cloning and characterization of full-length trypsins from the salmon louse *Lepeophtheirus salmoni*. Mol Biochem Parasitol 136:303–307

Kvamme BO, Kongshaug H, Nilsen F (2005) Organisation of trypsin genes in the salmon louse (*Lepeophtheirus salmonis*, Crustacea, copepoda) genome. Gene 352:63–74

Lam W, Coast GM, Rayne RC (2000) Characterisation of multiple tripsins from the midgut of *Locusta migratoria*. Insect Biochem Mol Biol 30:85–94

Larson BA, Vigna SR (1983) Species and tissue distribution of cholecystokinin/gastrin-like substances in some invertebrates. Comp Endocrinol 50:469–475

Laskowski M, Kato I (1980) Protein inhibitors of proteinases. Ann Rev Biochem 49:593–626

Le Huerou I, Wicker C, Guilloteau P, Toullec R, Puigserver A (1990) Isolation and nucleotide sequence of cDNA clone for bovine pancreatic anionic trypsinogen. Structural identity within the trypsin family. Eur J Biochem 193:767–773

Lehnert SA, Johnson SE (2002) Expression of hemocyanin and digestive enzyme messenger RNAs in the hepatopancreas of the black tiger shrimp *Penaus monodon*. Comp Biochem Phys B 133:63–171

Lehane SM, Assinder SJ, Lehane MJ (1998) Cloning, sequencing, temporal expression and tissue-specificity of two serine proteinases from the midgut of the blood-feeding fly *Stomoxys calcitrans*. Eur J Biochem 254:290–296

Leiros HKS, Willassen NP, Smalås AO (2000) Structural comparison of psychrophilic and mesophilic trypsins. Elucidating the molecular basis of cold-adaptation. Eur J Biochem 267:1039–1049

Lemos D, Hernández-Cortés MP, Navarrete del Toro MA, García-Carreño FL, Phan VN (1999) Ontogenetic variation in digestive proteinase activity of larvae and postlarvae of pink shrimp *Farfantepenaeus paulensis* (crustacea: Decapoda:Penaeidae). Mar Biol 135:653–662

Le Moullac G, van Wormhoudt A (1994) Adaptation of digestive enzymes to dietary protein, carbohydrate and fiber levels and influence of protein and carbohydrate quality in *Penaeus vannamei* larvae (Crustacea, Decapoda). Aquat Liv Res 7:203–210

Lhoste EF, Fiszlewicz M, Gueugneau AM, Tranchant T, Corring T (1994) Early adaptation of pancreas to protein-enriched feed: role of cholecystokinin and gastrin-releasing peptide. Pancreas 9:624–632

Lopes AR, Terra WR (2003) Purification, properties and substrate specificity of a digestive trypsin from *Periplaneta americana* (Dictyoptera) adults. Insect Biochem Mol Biol 33:407–415

Lopes AR, Juliano MA, Marana SR, Juliano L, Terra WR (2006) Substrate specificity of insect trypsins and the role of their subsites in catalysis. Insect Biochem Mol Biol 36:130–140

Louvard MN, Puigserver A (1974) On bovine and porcine anionic trypsinogens. Biochem Biophys Acta 371:177–185

Lu SJ, Pennington JE, Stonehouse AR, Mobula MM, Wells MA (2006) Reevaluation of the role of early trypsin activity in the transcriptional activation of the late trypsin gene in the mosquito *Aedes aegypti*. Insect Biochem Mol Biol 36:336–343

MacDonald RJ, Stary SJ, Swift GH (1982) Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide sequences of the cloned cDNA. J Biol Chem 257:9724–9732

Maeda-Martínez A, Obregon-Barboza V, Navarrete del Toro MA, Obregon-Barboza H, García-Carreño FL (2000) Trypsin-like enzymes from two morphotypes of the 'living fossil' Triops (Crustacea: Brachiopoda:Notostraca). Comp Biochem Physiol B 126:317–323

Malavazi-Piza K, Araujoa M, Godinhob R, Tanaka A (2004) Effect of invertebrate serine proteinase inhibitors on carrageenan-induced pleural exudation and bradykinin release. Int Immunopharmacol 4:1401–1408

Male R, Lorens JB, Smalas AO, Torrissen KR (1995) Molecular cloning and characterization of anionic and cationic variants of trypsin from Atlantic salmon. Eur J Biochem 232:677–685

Martin JW, Haney TA (2005) Decapod crustaceans from hydrothermal vents and cold seeps: a review through 2005. Zool J Linn Soc 145:445–522

Meyering-Vos M, Müller A (2007) Structure of the sulfakinin cDNA and gene expression from the Mediterranean field cricket *Gryllus bimaculatus*. Insect Mol Biol 16:445–454

Muhlia-Almazán A, García-Carreño FL (2002) Influence of molting and starvation on the synthesis of proteolytic enzymes in the midgut gland of the white shrimp *Penaeus vannamei*. Comp Biochem Physiol B 133:383–394

Müller HM, Crampton JM, della Torre A, Sinden R, Crisanti A (1993) Members of a trypsin gene family in *Anopheles gambiae* are induced in the gut by blood meal. EMBO J 12:2891–2900

Müller HM, Catteruccia F, Vizioli J, della Torre A, Crisanti A (1995) Constitutive and blood meal-induced trypsin genes in *Anopheles gambiae*. Exp Parasitol 81:371–385

Myrnes B, Nilsen IW (2007) Glutathione *S*-transferase from the Icelandic scallop (*Chlamys islandica*): Isolation and partial characterization. Comp Biochem Physiol 144C:403–407

Navarrete del Toro MA, García-Carreño FL, Díaz LM, Celis-Guerrero L, Saborowski R (2006) Aspartic proteinases in the digestive tract of marine decapod crustaceans. J Exp Zool 305A:645–654

Neurath H (1984) Evolution of proteolytic enzymes. Science 224:350–357

Neurath H (1994) Proteolytic enzymes past and present: the second golden era. Protein Sci 3:1734–1739

Neurath H (2001) Recollections: from proteases to proteomics. Protein Sci 10:892–904

Nichols R, Schneuwly SA, Dixon JE (1988) Identification and characterization of a *Drosophila* homologue to the vertebrate neuropeptide cholecystokinin. J Biol Chem 263:12167–12170

Noriega FG, Wells MA (1999) A molecular view of trypsin synthesis in the midgut of *Aedes aegypti*. J Insect Physiol 45:613–620

Noriega FG, Barillas-Mury CV, Wells MA (1994) Dietary control of late-trypsin gene transcription in *Aedes aegypti*. Insect Biochem Mol Biol 24:627–631

Noriega FG, Pennington JE, Barrillas-Mury CV, Wang XY, Wells MA (1996) *Aedes aegypti* midgut early trypsin is post-transcriptionally regulated by blood feeding. Insect Mol Biol 5:25–29

Northrup JH, Kunitz M (1931) Isolation of protein crystals possessing tryptic activity. Science 73:262–263

Northrup JH, Kunitz M, Herriott RM (1948) Crystalline enzymes. Columbia University Press, New York

Novillo C, Castañera P, Ortego F (1999) Isolation and characterization of two digestive trypsin-like proteinases from larvae of the stalk corn borer *Sesamia nonagrioides*. Insect Biochem Mol Biol 29:177–184

Ollivaux C, Soyes D (2000) Dynamics of biosynthesis and release of crustacean hyperglycemic hormone isoforms in the X-organ/sinus gland complex of the crayfish *Orconectes limosus*. Eur J Biochem 267:5106–5114

Øverbø K, Myrnes B (2006) Deoxyribonuclease II from the Icelandic scallop (*Chlamys islandica*): isolation and partial characterization. Comp Biochem Physiol 143B:315–318

Pancer Z, Lenck J, Ronkevick B, Steffen R, Muller I, Muller WE (1996) Molecular cloning and sequence analysis of two cDNAs coding for putative anionic trypsinogens from the colonial Urochordate *Botryllus schlosseri* (Ascidiacea). Mol Mar Biol Biotechnol 5:326–333

Pasternak A, Ringe D, Hedstrom L (1999) Comparison of anionic and cationic trypsinogens: the anionic activation domain is more flexible in solution and differs in its mode of BPTI binding in the crystal structure. Protein Sci 8:253–259

Peterson AM, Barillas-Mury CV, Wells MA (1994) Sequence of three cDNAs encoding an alkaline midgut trypsin from *Manduca sexta*. Insect Biochem Mol Biol 24:463–471

Pfleiderer G, Zwilling R, Sonneborn HH (1967) Eine proteinase vom molekulargewicht 11000 und eine trypsinahnliche fraktion aus *Astacus fluviatilis*. Hoppe-Seyler's Z Physiol Chem 348:1319–1331

Pfleiderer G, Linke R, Reinhardt G (1970) On the evolution of endopeptidases. 8. Cross-reactions of trypsins and chymotrypsins of different species. Comp Biochem Physiol 33:955–967

Pils B, Shultz J (2004) Inactive enzyme homologues find new function in regulatory processes. J Mol Biol 340:399–404

Pinsky SD, La Forge KS, Scheele G (1985) Differential regulation of trypsinogen mRNA translation: full-length mRNA sequences encoding two oppositely charged trypsinogen isoenzymes in the dog pancreas. Mol Cell Biol 5:2669–2676

Resch-Sedlmeier G, Sedlmeier D (1999) Release of digestive enzymes from the crustacean hepatopancreas: effect of vertebrate gastrointestinal hormones. Comp Biochem Physiol B 123:187–192

Roach JC, Wang K, Gan L, Hood L (1997) The molecular evolution of the vertebrate trypsinogens. J Mol Evol 45:640–652

Rungruangsak-Torrisen K, Carter CG, Sundby A, Berg A, Houlihan DF (1999) Maintenance ration, protein synthesis capacity, plasma

insulin and growth of Atlantic salmon (*Salmo salar* L.) with genetically different trypsin isoenzymes. Fish Physiol Biochem 21:223–233

Rypniewsky WR, Hastrup S, Betzel C, Dauter M, Dauter Z, Papendorf G, Branner S, Wilson KS (1993) The sequence and X-ray structure of trypsin from *Fusarium oxysporum*. Prot Eng Des Selec 6:341–348

Rypniewsky WR, Perrakis A, Vorgias CE, Wilson KS (1994) Evolutionary divergence and conservation of trypsin. Protein Eng 7:57–64

Saborowski R, Sahling G, Navarrete del Toro MA, Garcia-Carreño F (2004) Stability and effects of organic solvents on endopeptidases from the gastric fluid of the marine crab *Cancer pagurus*. J Mol Catal-B Enzym 30:109–118

Sahin-Tóth M (2000) Human cationic trypsinogen. J Biol Chem 275:22750–22755

Sahin-Toth M, Toth M (2000) Gain-of-function mutations associated with hereditary pancreatitis enhance autoactivation of human cationic trypsinogen. Biochem Bioph Res Co 278:286–289

Sainz JC, García-Carreño FL, Hernández-Cortés P (2004a) *Penaeus vannamei* isotrypsins: purification and characterization. Comp Biochem Physiol 138:155–162

Sainz JC, García-Carreño FL, Sierra-Beltrán A, Hernández-Cortés P (2004b) Trypsin synthesis and storage as zymogen in the midgut gland of the shrimp *Litopenaeus vannamei*. J Crustacean Biol 24:266–273

Sainz JC, García-Carreño FL, Cordova-Murueta J, Cruz-Hernandez P (2005) *Penaeus vannamei* (Boone, 1931) isotrypsins, genotype and modulation. J Exp Mar Biol Ecol 326:105–113

Salomon M, Buchholz F (2000) Effects of temperature on the respiration rates and the kinetics of citrate synthase in two species of *Idotea* (Isopoda, Crustacea). Comp Biochem Physiol 125B:71–81

Sánchez-Paz A, García-Carreño FL, Muhlia-Almazán A, Hernández-Saavedra NY, Yepiz-Plascencia G (2003) Differential expression of trypsin mRNA in the white shrimp (*Penaeus vannamei*) midgut gland under starvation conditions. J Exp Mar Biol Ecol 292:1–17

Schrøder-Leiros HK, Willassen N, Smalås A (2000) Structural comparison of psychrophilic and mesophilic trypsins: elucidating the molecular basis of cold-adaptation. Eur J Biochem 267:1039–1049

Schyns R, Bricteux-Gregoire S, Florkin M (1969) Purification, properties and N-terminal sequence of sheep trypsinogen. Biochim Biophys Acta 175:97–112

Sedlmeier D (1988) The crustacean hyperglycemic hormone (CHH) releases amylase from the crayfish midgut gland. Regul Peptides 20:91–98

Shah PK, Tripathi LP, Jensen LJ, Gahnim M, Mason C, Furlong EE, Rodrigues V, White KP, Bork P, Sowdhamini R (2008) Enhaced function annotations for *Drosophila* serine proteases: a case study for systematic annotation of multi-member gene families. Gene 407:199–215

Shen Z, Jacobs-Lorena M (1998) Nuclear factor recognition sites in the gut-specific enhancer region of an *Anopheles gambiae* trypsin gene. Insect Biochem Molec Biol 28:1007–1012

Shi YB, Brown DD (1990) Developmental and thyroid hormone-dependent regulation of pancreatic genes in *Xenopus laevis*. Genes Dev 4:1107–1113

Skavdis G, Siden-Kiamos I, Müller HM, Crisanti A, Louis C (1996) Conserved function of *Anopheles gambiae* midgut-specific promoters in the fruitfly. EMBO J 15:344–350

Smalås AO, Heimstad ES, Hordvik A, Willassen NP, Male R (1994) Cold adaption of enzymes: structural comparison between salmon and bovine trypsins. Proteins 20:149–166

Sotelo-Mundo RR, López-Zavala AA, Garcia-Orozco KD, Arvizu-Flores AA, Velázquez-Contreras EF, Valenzuela-Soto EM, Rojo-Dominguez A, Kanost MR (2007) The lysozyme from insect (*Manduca sexta*) is a cold-adapted enzyme. Protein Pept Lett 14:774–778

Southan C (2000) Assessing the proteinase and proteinase inhibitor content of the human genome. J Peptide Sci 6:453–458

Spiess C, Beil A, Ehrmann M (1999) A temperature-dependent switch from chaperone to proteinase in a widely conserved heat shock protein. Cell 97:333–347

Steiner JM, Medinger TL, Williams DA (1997) Purification and partial characterization of feline trypsin. Comp Biochem Physiol B 116:87–93

Sweet RM, Wright HT, Janin J, Chothia CH, Blow DM (1974) Crystal structure of the complex of porcine trypsin with soybean trypsin inhibitor (Kunitz) at 2.6-Å resolution. Biochemistry 13:4212–4228

Teschke M, Saborowski R (2005) Cysteine proteinases substitute for serine proteinases in the midgut glands of *Crangon crangon* and *Crangon allmani* (Decapoda: Caridea). J Exp Mar Biol Ecol 316:213–229

Titani K, Sasagawa T, Woodbury RG, Ericsson LH, Dorsam H, Kraemer M, Neurath H, Zwilling R (1983) Amino acid sequence of crayfish (*Astacus fluviatilis*) trypsin. Biochemistry 22:1459–1465

Titani K, Torff HJ, Hormel S, Kumar S, Walsh KA, Rodl J, Neurath H, Zwilling R (1987) Amino acid sequence of a unique protease from the crayfish *Astacus fluviatilis*. Biochemistry 26:222–226

Torfs P, Baggerman G, Meeusen T, Nieto J, Nachman RJ, Calderon J (2002) Isolation, identification, and synthesis of a disulfated sulfakinin from the central nervous system of an arthropod, the white shrimp *Litopenaeus vannamei*. Biochem Bioph Res Co 299:312–320

Tschesche H, Kolkenbrock H (1984) An inhibitor of elastase from *Anemonia sulcata*. Chem Peptide Protein 2:350–354

Turkiewicz M, Galas E, Kalinowska H (1991) Collagenolytic serine protease from *Euphausia superba* Dana (Antarctic krill). Comp Biochem Physiol 99B:359–371

Vogt G, Stocker W, Storch V, Zwilling R (1989) Biosynthesis of *Astacus* protease a digestive enzyme from crayfish. Histochemistry 91:373–381

Volpicella M, Ceci LR, Cordewener J, America T, Gallerani R, Bode W, Jongsma MA, Beekwilder J (2003) Properties of purified gut trypsin from *Helicoverpa zea*, adapted to proteinase inhibitors. Eur J Biochem 270:10–19

Walsh KA, Kauffman DL, Sampath-Kumar KSV, Neurath H (1964) On the structure and function of bovine trypsinogen and trypsin. Proc Natl Acad Sci USA 51:301–308

Walsh KA, Wilcox PE (1970) Serine proteinases. Meth Enzymol 19:31–41

Wang K, Gan L, Lee I, Hood L (1995) Isolation and characterization of the chicken trypsinogen gene family. Biochem J 307:471–479

Wang S, Magoulas C, Hickey D (1999) Concerted evolution within a trypsin gene cluster in *Drosophila*. Mol Biol Evol 16:1117–1124

Ward CW (1975) Properties and specificity of the major anionic trypsin-like enzyme in the keratinolytic larvae of the webbing clothes moth. Biochim Biophys Acta 391:201–211

Webb EC (1993) Enzyme nomenclature: a personal perspective. FASEB J 7:1192–1194

Williams AB (1980) A new crab family from the vicinity of submarine thermal vents on the Galapagos Rift (Crustacea: Decapoda: Brachyura). Proc Biol Soc Wash 93:443–472

Williams AB, Chace FA Jr (1982) A new caridean shrimp of the family Bresiliidae from thermal vents of the Galapagos Rift. J Crust Biol 2:136–147

Xiong B, Jacobs-Lorena M (1995) The black fly *Simulium vittatum* trypsin gene: characterization of the 5′-upstream region and induction by blood meal. Exp Parasitol 81:363–370

Yan XH, DeBondt HL, Powell CC, Bullock RC, Borovsky D (1999) Sequencing and characterization of the citrus weevil, *Diaprepes*

abbreviatus, trypsin cDNA. Effect of *Aedes trypsin* modulating oostatic factor on trypsin biosynthesis. Eur J Biochem 262:627–636

Yang YJ, Davies D (1971) Trypsin and chymotrypsin during metamorphosis in *Aedes aegypti* and properties of the chymotrypsin. J Insect Physiol 17:117–131

Zecchinon L, Claverie P, Collins T, D'Amico S, Delille D, Feller G, Georlette D, Gratia E, Hoyoux A, Meuwis MA, Sonan G, Gerday C (2001) Did psychrophilic enzymes really win the challenge? Extremophiles 5:313–321

Zhu YC, Baker JE (1999) Characterization of midgut trypsin-like enzymes and three trypsinogen cDNAs from the lesser grain borer,

R. dominica (Coleoptera: Bostrichidae). Insect Biochem Mol Biol 29:1053–1063

Zhu YC, Kramer KJ, Dowdy AK, Baker JE (2000) Trypsinogen-like cDNAs and quantitative analysis of mRNA levels from the Indian meal moth, *Plodia interpunctella*. Insect Biochem Mol Biol 30:1027–1035

Zollner H (1993) Handbook of enzyme inhibitors. VCH Verlagsgesellschaft, Weinheim

Zwilling R, Neurath H (1981) Invertebrate proteinases. Methods Enzymol 80:633–664