International Journal for Parasitology xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

International Journal for Parasitology

journal homepage: www.elsevier.com/locate/ijpara

Invited Review

Serine proteases in schistosomes and other trematodes

Jan Dvorak [a,*], Martin Horn [b,*]

[a] Department of Zoology and Fisheries, Faculty of Agrobiology, Food and Natural Resources, Czech University of Life Sciences in Prague, Kamycka 129, Prague CZ 165 21, Czech Republic
[b] Institute of Organic Chemistry and Biochemistry, The Czech Academy of Sciences, Prague CZ 166 10, Czech Republic

ARTICLE INFO

Article history:
Received 8 November 2017
Received in revised form 19 January 2018
Accepted 25 January 2018
Available online xxxx

Keywords:
Proteolytic enzyme
Serine protease
Peptidase
Trematoda
Platyhelminthes
Schistosoma
Fluke

ABSTRACT

Trematodes, also known as flukes, are phylogenetically ancient parasitic organisms. Due to their importance as human and veterinary parasites, their proteins have been investigated extensively as drug and vaccine targets. Among those, proteases, as crucial enzymes for parasite survival, are considered candidate molecules for anti-parasitic interventions. Surprisingly however, trematode serine proteases, in comparison with other groups of proteases, are largely neglected. Genes encoding serine proteases have been identified in trematode genomes in significant abundance, but the biological roles and biochemical functions of these proteases are poorly understood. However, increasing volumes of genomic and proteomic studies, and accumulated experimental evidence, indicate that this class of proteases plays a substantial role in host–parasite interactions and parasite survival. Here, we discuss in detail serine proteases at genomic and protein levels, and their known or hypothetical functions.

1. Introduction

Trematodes (Platyhelminthes, Trematoda, Digenea) represent a large group of parasites with more than 18,000 species recorded (Kostadinova and Perez-del-Olmo, 2014). They have complex and diverse life cycles that always involve mollusks as intermediate hosts, where they reproduce asexually, and a vertebrate definitive host, where sexual reproduction takes place. During their life cycles, parasites undergo dramatic changes in body organisation in different hosts and external environments. Importantly, many trematode species cause serious diseases and are of high medical and veterinary importance (Keiser and Utzinger, 2009; Mas-Coma et al., 2009; Lustigman et al., 2012). Thus most research has been focused on trematodes which impact human and livestock health; so-called "blood flukes" and "liver flukes".

Blood flukes are members of the genus *Schistosoma*, causing human and animal disease known as schistosomiasis (schistosomosis, bilharziasis) – chronic debilitating disease infecting more than 240 million people (Steinmann et al., 2006). There are several species responsible for human disease with five species being responsible for most of the disease cases recorded: *Schistosoma mansoni*, *Schistosoma intercalatum*, *Schistosoma japonicum*, *Schistosoma mekongi* and *Schistosoma haematobium* (Lockyer et al., 2003). Schistosome free-living larvae (cercariae), released into water from snail intermediate hosts, penetrate human skin and transform into schistosomula. They migrate to the vasculature where they mature into adult worms. Adult worms reside in the portal and mesenteric or bladder veins as male/female pairs, and survive for many years, producing hundreds of fertilised eggs per day. Chronic infections persist for years, or even decades, and severe morbidity results from host immune responses to eggs in host tissues (Pearce and MacDonald, 2002). Pathophysiological conditions such as spleno – and hepatomegalies due to immune-mediated entrapment (granulomas) of schistosome eggs, periportal fibrosis, portal hypertension, urinary obstruction, bladder carcinoma, sterility, malnutrition and developmental retardation are common with this parasitic infection. Existing treatment relies heavily on one drug, Praziquantel, and no vaccine has yet been developed (Pearce and MacDonald, 2002). During their complex life cycle, the parasites survive in various environments by presenting or releasing bioactive molecules that aid survival and modulate host physiology (Da'dara and Skelly, 2011). Disruption of these mechanisms by specific drugs/vaccines would provide therapeutic benefits.

Liver flukes are a group of polyphyletic trematodes residing in the host liver parenchyma, bile ducts or gallbladder, depending on the fluke species (Lotfy et al., 2008; Kostadinova and Perez-del-Olmo, 2014). Most research has focused on those species responsible for important human and livestock infections – fascioliasis (fasciolosis), clonorchiasis and opisthorchiasis. Fascioliasis is

* Corresponding authors.
E-mail addresses: hdvorak76@outlook.com (J. Dvorak), horn@uochb.cas.cz (M. Horn).

https://doi.org/10.1016/j.ijpara.2018.01.001
0020-7519/

a disease caused by *Fasciola hepatica*, *Fasciola gigantica* and several related species which affect livestock and free-living mammals worldwide (Mas-Coma et al., 2005). *Fasciola* larvae (cercariae) are released from the intermediate snail host and encyst as metacercariae on the surface of herbage; disease is transmitted by ingestion of these cysts. After ingestion, the parasites excyst and migrate via the intestine into the peritoneal cavity. The newly excysted juveniles (NEJs) migrate to the liver parenchyma and, after a feeding period that causes extensive damage to the liver tissue, they move to the bile duct, where they mature and produce eggs. The eggs are excreted with the bile via the intestine to faeces. Liver flukes cause serious economic losses in livestock production (in the European Union, more than US $1 billion/year and worldwide US $3 billion/year) (Mas-Coma et al., 2005; Piedrafita et al., 2010). The drug of choice is Triclabendazole (TCBZ), although resistance is spreading in Europe and other parts of the world (Fairweather, 2005). Pathophysiological conditions include anaemia at an early stage, bile duct inflammation and fibrosis, a decrease in bile production, liver atrophy, cirrhosis, weight loss and for some animals the disease is lethal (Mas-Coma et al., 2009). Human infection can occur in areas with poor farm management practices and the World Health Organization (WHO) recognises fascioliasis as a food-borne disease infecting more than 17 million people, with many more at risk (Mas-Coma et al., 2005). Clonorchiasis and opisthorchiasis are caused predominantly by *Clornorchis sinensis*, *Opisthorchis viverini* and *Opisthorchis felineus*, which are important food–borne trematodes with a high impact on human populations in endemic areas (Echaubard et al., 2016), with approximately 700 million people worldwide at risk of infection (Keiser and Utzinger, 2009; Sripa et al., 2010). Their life cycles involve three host species, a freshwater snail, a fish and a mammal, where they produce eggs after being ingested together with raw or partially cooked freshwater cyprinid fish (Echaubard et al., 2016). NEJs of these species enter the liver biliary system directly via the common bile duct (Moazeni and Ahmadi, 2016). Infection is linked to hepatobiliary pathology and can subsequently lead to cholangiocarcinoma, a malignant tumour of the biliary tract (Echaubard et al., 2016).

Proteolysis is an essential activity for numerous physiological processes (Rawlings et al., 2016). Proteolytic enzymes (proteases, peptidases) of trematodes and other parasite are attractive vaccine (Molina-Hernandez et al., 2015; Pearson et al., 2015; Tallima et al., 2017) and drug targets as they operate at the host–parasite interface (Abdulla et al., 2007; Jilkova et al., 2011). Proteases are generally crucial for parasitism as they facilitate important functions such as obtaining nutrients, successful parasite invasion, hatching, excystment (Sajid and McKerrow, 2002; Caffrey et al., 2004; McKerrow et al., 2006; Horn et al., 2014), as well as immune evasion and modulation of the host physiology (Carvalho et al., 1998; Cocude et al., 1999; Doenhoff et al., 2003; McKerrow et al., 2006; de Oliveira Fraga et al., 2010; Da'dara and Skelly, 2011). Most studies have been focused on cysteine proteases (CPs) and aspartic proteases (APs) from the digestive network, and cascades connected with the nutrient intake of flukes (Caffrey et al., 2004; Robinson et al., 2008; Kasny et al., 2009; McVeigh et al., 2012). On the other hand, higher organisms such as vertebrates employ serine proteases (SPs) for protein digestion instead of CPs (Delcroix et al., 2006), and this notable evolutionary switch is obvious in the digestive network of blood feeding arthropods. While ticks digest a blood meal using predominantly CPs and APs (Sojka et al., 2008; Horn et al., 2009), blood feeding insects such as mosquitoes, fleas, lice, sandflies or tsetse flies use similar SPs and APs as vertebrates (Sojka et al., 2016; Santiago et al., 2017). However some, such as Coleopteran insects, have retained the ability to utilise both CPs and SPs as a redundant or alternative mode of action (Petek et al., 2012; Zhu-Salzman and Zeng, 2015; Srp et al., 2016). This switch, employing the less reactive hydroxyl group of SPs instead of the thiol group of CPs, could have arisen due to the higher level of aerobic metabolism in insects.

In this review, we will focus on the involvement of trematode SPs in host–parasite interactions such as invasion processes involving cercarial elastases (CEs) (Ingram et al., 2012), modulation of host hemostasis and/or fluke physiology (Horn et al., 2014; Fajtova et al., 2015; Dvorak et al., 2016).

2. Serine proteases

SPs are so named due to the catalytic serine residue at the active site, where it is responsible for nucleophilic attack on the carbonyl carbon of a susceptible peptide bond. In addition to this serine residue, the catalytic mechanism involves additional residues that are proton donors. Based on the MEROPS database (Rawlings et al., 2016) (https://www.ebi.ac.uk/merops/), proteases are classified into clans and families. A clan contains proteases that have an evolutionary relationship and share similar tertiary structures. The clan names are formed from the letter for the catalytic type (S for serine proteases) followed by a serial second capital letter. Clans are divided into families based on sequence similarities and identities. Each protease family is named with a letter of the catalytic type followed by a sequentially assigned number. Alternatively, families are named by the exemplary protease, which is usually the protease that has been biochemically most studied (e.g. chymotrypsin for family S1).

SPs are an integral part of the proteolytic network in schistosomes, as shown by multiplex substrate profiling analysis of excretory/secretory products of developmental stages infecting the human host, where SP activities form a substantial part of all profiled activities (Dvorak et al., 2016). Genomes of schistosomes encode more than 300 predicted proteases (Berriman et al., 2009; Young et al., 2012), comprising approximately 2.5% of the whole genome. In the genomes of *S. mansoni* or *C. sinensis*, almost 25% or 30% of all proteases, respectively, in the database can be classified as SPs (Howe et al., 2017). The majority of them have been identified by prediction, and only a few of them have been biochemically characterised. Among these, proteases of clan PA family S1 and clan SC have been studied the most.

3. Clan PA, family S1

SPs from family S1, clan PA (P is used for clans containing families of more than one catalytic type), are generally endoproteases involved in many processes in which protein modification or degradation is required. They are the largest family of proteases recorded and their actions are essential in many critical events such as development, physiology, digestion, innate immunity, pathogen invasion etc. (Page and Di Cera, 2008; Rawlings et al., 2016). Family S1 proteases are characterised by the His-Asp-Ser order of catalytic triad in the primary sequence typical for all SPs from clan PA (Page and Di Cera, 2008). They are synthesised as inactive precursors (zymogens, proenzymes) and their activation is tightly controlled by proteolytic processing of the N-terminal extension that acts as a prodomain. This processing is required for structural rearrangement and generation of the active enzyme (Rawlings et al., 2016). The length of the N-terminal extension varies significantly from a few residues to hundreds of amino acids. The N-terminal extension is usually connected by disulfide bridges with the protease domain and may contain other non-proteolytic, usually binding, domains. An exemplary protease of family S1 is bovine chymotrypsin, with which other S1 proteases share similar features in their tertiary structures, although their particular substrate binding and cleavage specificities vary (Rawlings et al.,

Please cite this article in press as: Dvorak, J., Horn, M. Serine proteases in schistosomes and other trematodes. Int. J. Parasitol. (2018), https://doi.org/10.1016/j.ijpara.2018.01.001