Thesis Title Identification and Characterization of Cysteine and Serine Proteases from *Gnathostoma spinigerum*

Author Miss Natthawan Kongkerd

Degree Doctor of Philosophy (Microbiology)

Thesis Advisory Committee

Assoc. Prof. Dr. Pichart Uparanukraw, M.D.	Chairperson
Assoc. Prof. Dr. Nopporn Sittisombut, M.D.	Member
Asst. Prof. Dr. Sumalee Pruksakorn	Member

ABSTRACT

Gnathostoma spinigerum is a causative agent of human gnathostomiasis, a common parasitic disease involving skin and visceral organs including the central nervous system. Cysteine and serine proteases of many parasites have been demonstrated to play several roles in the mechanisms of parasitism. In this study, cDNA encoding a cathepsin L-like cysteine protease (GsCL1) and a serine protease (GsSP1) from *G. spinigerum* advanced third-stage larva (aL3) were identified and the biochemical properties of the recombinant enzymes characterized. The complete cDNA of the *GsCL1* gene of 1,484 bp encodes a protein of 398 amino acids consisting of the typical signal peptide sequence, the pro-domain and the mature domain of 23, 156 and 219 amino acids, respectively. The mature enzyme has the predicted molecular mass of 24 kDa. The amino acid sequence of GsCL1 contains the cysteine protease catalytic triad, Cys^{25}, His^{165}, and Asn^{186}, a highly conserved

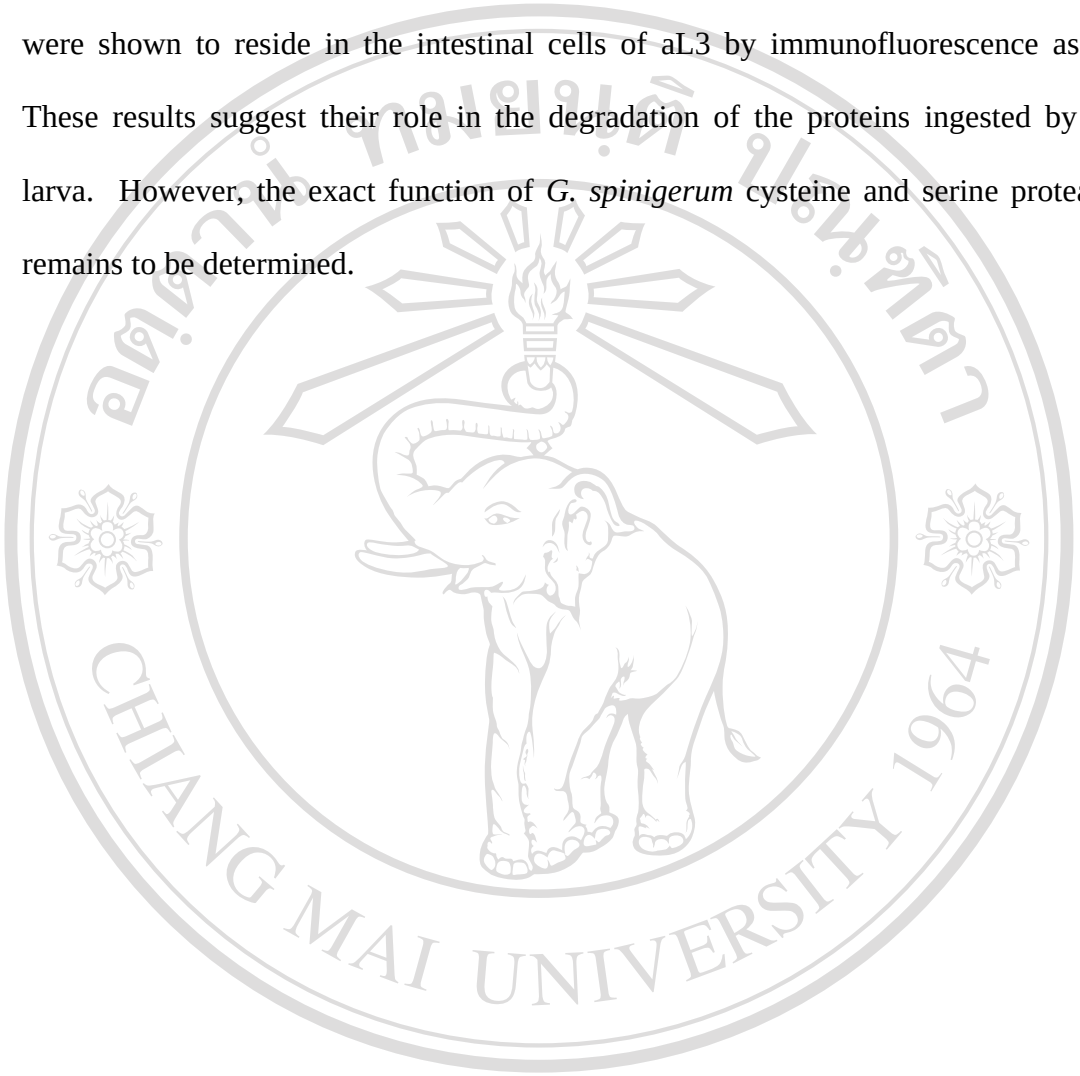
ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ERFNIN motif and a GNFD motif, which may be involved in intramolecular processing of the enzyme. The deduced amino acid sequence of *GsCL1* gene shows 53-64% identity to cathepsin L proteases of various organisms including a cathepsin L family member (cpl-1) of *Caenorhabditis elegans*. Phylogenetic analysis with 24 members of the papain superfamily has revealed that GsCL1 is closely related to cathepsin L proteases of nematodes including *C. elegans.* The recombinant pro-enzyme of GsCL1 expressed in *Pichia pastoris* displayed optimal protease activity toward Z-Phe-Arg-AMC substrate at pH 6.0. The activity was sensitive to cysteine protease inhibitors E-64 and K11777. The preference of GsCL1 for large hydrophobic and aromatic residues in the P2 position (Leu, Ile, Val, Phe, Trp, Tyr) was typical of cathepsin L proteases. Mouse anti-GST-proGsCL1 serum showed reactivities with 35-, 38- and 45-kDa proteins in the aL3 extracts. These proteins were shown to localize inside the intestinal cells of aL3 by immunofluorescence assay.

The 1,032-bp cDNA sequence of the *GsSP1* gene encodes a protein of 343 amino acids comprising the putative signal peptide, the activation peptide and the mature enzyme of 19, 26 and 298 amino acids, respectively. The mature enzyme has the predicted molecular mass of 32.9 kDa and contains the serine protease catalytic triad, His^{46}, Asp^{95}, and Ser^{196}. The deduced amino acid sequence shows 30-35% homology to serine proteases from mammalians, insects and other eukaryotes. Phylogenetic analysis with selected members of serine proteases has revealed that GsSP1 is more closely related to those of insects and moths than those of mammalians. The recombinant proGsSP1 expressed in *P. pastoris* showed no serine protease activity against the fluorogenic substrates used in this study. Mouse

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

antiserum against GST-proGsSP1 fusion protein reacted with 35-, 38- and 39-kDa proteins in the aL3 extracts. These GsSP1-related proteins, just the same as GsCL1, were shown to reside in the intestinal cells of aL3 by immunofluorescence assay. These results suggest their role in the degradation of the proteins ingested by the larva. However, the exact function of *G. spinigerum* cysteine and serine proteases remains to be determined.



ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ชื่อเรื่องวิทยานิพนธ์ การค้นหาและศึกษาคุณลักษณะของซิสเทอีนและเซอรีน โพรทีเอสจากพยาธิตัวจี๊ด

ผู้เขียน นางสาวณัฐวรรณ คงเกิด

ปริญญา วิทยาศาสตรดุษฎีบัณฑิต (จุลชีววิทยา)

คณะกรรมการที่ปรึกษาวิทยานิพนธ์ รศ.ดร.นพ. พิชาติ อุปรานุเคราะห์ ประธานกรรมการ

รศ.ดร.นพ. นพพร สิทธิสมบัติ กรรมการ

ผศ.ดร. สุมาลี พฤกษากร กรรมการ

บทคัดย่อ

พยาธิตัวจี๊ด *Gnathostoma spinigerum* เป็นเชื้อที่ก่อให้เกิดโรคพยาธิตัวจี๊ดในคน (human gnathostomiasis) ซึ่งเป็นโรคติดเชื้อปรสิตที่เกิดขึ้นบริเวณผิวหนังและอวัยวะภายใน รวมถึงระบบประสาทส่วนกลาง cysteine และ serine proteases ในปรสิตหลายชนิด มีบทบาทอย่างมากต่อกลไกการอยู่ร่วมกันแบบปรสิต ในการศึกษาครั้งนี้ ได้ทำการแยก cDNA ที่กำหนดการสร้าง cathepsin L-like cysteine protease (GsCL1) และ serine protease (GsSP1) จากตัวอ่อนระยะที่ 3 ขั้นปลายของพยาธิตัวจี๊ด (aL3) และได้ศึกษาคุณลักษณะทางชีวเคมีของ recombinant enzymes ความยาวที่สมบูรณ์ของ cDNA ของยีน *GsCL1* คือ 1,484 bp ซึ่งกำหนดการสร้างโปรตีนที่มีขนาดกรดอะมิโนยาว 398 ตัว ซึ่งประกอบด้วย signal peptide, pro-domain และ mature domain ที่มีกรดอะมิโนจำนวน 23, 156 และ 219 ตัวตามลำดับ mature enzyme มีมวลโมเลกุล 24 kDa ลำดับกรดอะมิโนของ GsCL1 ประกอบด้วย cysteine protease catalytic triad (Cys^{25}, His^{165} และ Asn^{186}),

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ERFNIN motif ที่มีลักษณะคงไว้ และ GNFD motif ซึ่งอาจเกี่ยวข้องในกระบวนการภายในโมเลกุลของเอนไซม์ ลำดับกรดอะมิโนของ GsCL1 มีความเหมือน 53-64% กับ cathepsin L proteases ของสิ่งมีชีวิตอื่น ๆ รวมถึง cathepsin L (cpl-1) ของ *Caenorhabditis elegans* การวิเคราะห์ทาง phylogenetic กับสมาชิกใน papain superfamily 24 ชนิด แสดงให้เห็นว่า GsCL1 มีความเกี่ยวข้องอย่างใกล้ชิดกับ cathepsin L proteases ของพยาธิตัวกลม รวมถึง *C. elegans* ด้วย recombinant pro-enzyme ของ GsCL1 ที่ถูกสร้างขึ้นใน *Pichia pastoris* มีปฏิกิริยาต่อ Z-Phe-Arg-AMC substrate ได้ดีที่สุดที่ pH 6.0 ซึ่งปฏิกิริยานี้ถูกยับยั้งได้ด้วย cysteine protease inhibitors E-64 และ K11777 GsCL1 มีความชอบกรดอะมิโนที่เป็น hydrophobic ขนาดใหญ่ และ aromatic (Leu, Ile, Val, Phe, Trp, Tyr) ที่ตำแหน่ง P2 ซึ่งเป็นลักษณะเฉพาะของ cathepsin L proteases ซีรั่มหนูต่อ GST-proGsCL1 ทำปฏิกิริยากับโปรตีนขนาด 35, 38 และ 45 kDa ที่สกัดจาก aL3 และพบว่าโปรตีนเหล่านี้อยู่ในเซลล์ลำไส้ของ aL3 โดยวิธี immunofluorescence

cDNA ของยีน *GsSP1* ขนาด 1,032 bp กำหนดการสร้างโปรตีนที่มีกรดอะมิโนยาว 343 ตัว ซึ่งประกอบด้วย putative signal peptide, activation peptide และ mature enzyme ที่มีกรดอะมิโนจำนวน 19, 26 และ 298 ตัว ตามลำดับ mature enzyme มีมวลโมเลกุล 32.9 kDa และประกอบด้วย serine protease catalytic triad (His^{46}, Asp^{95} และ Ser^{196}) ลำดับกรดอะมิโนมีความเหมือน 30-35% กับ serine proteases ของสัตว์เลี้ยงลูกด้วยนม แมลง และ eukaryotes ชนิดอื่น ๆ การวิเคราะห์ทาง phylogenetic กับสมาชิกของ serine proteases ที่คัดเลือกมาแสดงให้เห็นว่า GsSP1 มีความเกี่ยวข้องอย่างใกล้ชิดกับ serine proteases ของแมลงและผีเสื้อกลางคืนมากกว่า serine proteases ของสัตว์เลี้ยงลูกด้วยนม recombinant pro-enzyme ของ GsSP1 ที่ถูกสร้างขึ้นใน *P. pastoris* ไม่มีปฏิกิริยาต่อ fluorogenic substrates ที่ใช้ในการศึกษาครั้งนี้ ซีรั่มหนูต่อ GST-proGsSP1 fusion protein ทำ

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ปฏิกิริยากับโปรตีนขนาด 35, 38 และ 39 kDa ที่สกัดจาก aL3 พบว่าโปรตีนที่เกี่ยวข้องกับ GsSP1 เหล่านี้อยู่ในเซลล์ลำไส้ของ aL3 เช่นเดียวกันกับ GsCL1 จากผลการทดลองนี้บอกเป็นนัยถึงบทบาทของเอนไซม์โพรทีเอสในการย่อยสลายโปรตีนที่ตัวอ่อนพยาธิกินเข้าไป อย่างไรก็ตามหน้าที่ที่แน่ชัดของ cysteine และ serine proteases ของพยาธิตัวจี๊ดยังคงต้องมีการศึกษาต่อไป

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved