การศึกษาคุณลักษณะเฉพาะเชิงหน้าที่และโครงสร้างของ YKL-39 ซึ่งเป็นโปรตีนเหมือนไคติเนสจากมนุษย์

นางสาวอารยา รานอก
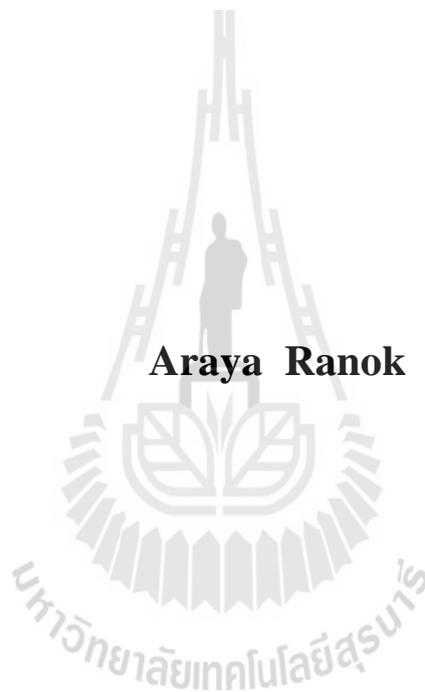
วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาชีวเคมี
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2556

FUNCTIONAL AND STRUCTURAL

CHARACTERIZATION OF YKL-39, A HUMAN

CHITINASE-LIKE PROTEIN

Araya Ranok

A Thesis Submitted in Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy in Biochemistry

Suranaree University of Technology

Academic Year 2013

FUNCTIONAL AND STRUCTURAL CHARACTERIZATION OF YKL-39, A HUMAN CHITINASE-LIKE PROTEIN

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

Thesis Examining Committee

_____

(Assoc. Prof. Dr. Jatuporn Witayakun)

Chairperson

_____

(Assoc. Prof. Dr. Wipa Suginta)

Member (Thesis Advisor)

_____

(Asst. Prof. Dr. Panida Khunkaewla)

Member

_____

(Dr. Buabarn Kuaprasert)

Member

_____

(Prof. Dr. James R Ketudat-Cairns)

Member

_____
(Prof. Dr. Sukit Limpijumnong)

Vice Rector for Academic Affairs

and Innovation

_____
(Assoc. Prof. Dr. Prapun Manyum)

Dean of Institute of Science

อารยา รานอก : การศึกษาคุณลักษณะเฉพาะเชิงหน้าที่และโครงสร้างของ YKL-39
ซึ่งเป็นโปรตีนเหมือนไคติเนสจากมนุษย์ (FUNCTIONAL AND STRUCTURAL
CHARACTERIZATION OF YKL-39, A HUMAN CHITINASE-LIKE PROTEIN)
อาจารย์ที่ปรึกษา : รองศาสตราจารย์ ดร.วิภา สุจินต์, 194 หน้า

โปรตีน YKL-39 ซึ่งเป็นโปรตีนเหมือนไคติเนสจากมนุษย์ จัดอยู่ในแฟมิลี glycosyl hydrolase 18 (GH-18) ที่ไม่สามารถย่อยไคตินได้ โปรตีนนี้ถูกจัดเป็นตัวบ่งชี้ทางชีวภาพการกระตุ้นของเซลล์กระดูกอ่อนและการติดตามความก้าวหน้าของโรคข้อเสื่อม แต่กลไกการทำงานยังไม่ทราบแน่ชัด วิทยานิพนธ์ฉบับนี้ได้แบ่งการศึกษาออกเป็น 2 ส่วน โดยส่วนแรกทำการศึกษาเกี่ยวกับการโคลน และผลิตรีคอมบิแนนท์โปรตีน YKL-39 ในแบคทีเรีย เพื่อใช้เป็นอิมมูนูเจนสำหรับผลิตโพลีโคลนอลและโมโนโคลนอลแอนติบอดีที่จำเพาะต่อโปรตีน YKL-39 จากการศึกษาพบว่า ทั้งโพลิโคนอลและโมโนโคลนอลแอนติบอดีมีความจำเพาะและทำปฏิกิริยาอย่างสูงต่อโปรตีน YKL-39 ผู้ทำวิจัยสามารถคัดเลือกโมโนโคลนที่สามารถสร้างโมโนโคลนอลแอนติบอดีจำเพาะต่อโปรตีน YKL-39 ได้จำนวน 2 โคลนให้ชื่อว่า 6H11 และ 8H3 ซึ่งโคลนทั้งสองสามารถผลิตแอนติบอดีชนิด IgM จากการศึกษาความจำเพาะของแอนติบอดีที่ผลิตได้ต่อโปรตีน YKL-39 ด้วยเทคนิค Dot blot พบว่าแอนติบอดีมีความจำเพาะสูงต่อโปรตีน YKL-39 ในตัวอย่างน้ำไขข้อของผู้ป่วยโรคข้อเสื่อม และในเซลล์เม็ดเลือดขาวชนิดโมโนไซต์ และทีลิมโฟไซต์ จากการค้นหาโปรตีนคู่จับของโปรตีน YKL-39 ในระบบฐานข้อมูล พบว่ามีไกลโคโปรตีนหลายชนิดที่เกี่ยวข้องกับโปรตีน YKL-39 ซึ่งมีบทบาทในเนื้อเยื่อกระดูกอ่อน การสร้างเนื้อเยื่อเกี่ยวพัน และการปฏิสัมพันธ์ระหว่างเซลล์ สามารถสรุปได้ว่าโพลิโคลนอลและโมโนโคลนอลแอนติบอดีต่อโปรตีน YKL-39 ที่ผลิตได้มีความเหมาะสมในการนำไปใช้ประโยชน์ในการศึกษาด้านภูมิคุ้มกันวิทยา เช่น การศึกษาวิถีการควบคุมของโปรตีน YKL-39 และการพัฒนาอุปกรณ์ในการตรวจวัดโปรตีนด้วยวิธีอิมมิวเซนเซอร์เพื่อติดตามความเสียหายของเนื้อเยื่อกระดูกอ่อน

ส่วนที่ 2 เป็นการศึกษาโครงสร้างสามมิติจากผลึกโปรตีนด้วยวิธีการหักเหของรังสีเอกซ์และศึกษาการจับของโปรตีน YKL-39 ต่อไคตินสายสั้นโดยวิธีอุณหพลศาสตร์ ผลการทดลองที่ได้พบว่าผลึกของโปรตีนสภาพธรรมชาติและที่จับกับไคตินสายสั้นขนาด 2-6 หน่วย สามารถหักเหรังสีเอกซ์ได้ความละเอียดถึง 1.53 ถึง 2.48 อังสตรอม มีโครงสร้างโดยรวมประกอบด้วยสองส่วนคือ โดเมนหลักมีโครงสร้างแบบ $(\beta/\alpha)8$-TIM-barrel และโดเมนเล็กแทรกระหว่างโดเมนหลักมีโครงสร้างเป็น $\alpha+\beta$ การวิเคราะห์ทางโครงสร้างพิสูจน์ให้เห็นว่าโปรตีน YKL-39 จับกับน้ำตาลสาย

สั้นด้วยแรงไฮโดรโฟบิคและด้วยเครือข่ายพันธะไฮโดรเจน ซึ่งการจับของน้ำตาลสายสั้นส่งผลให้โปรตีนเกิดการเปลี่ยนแปลงโครงรูปเพื่อให้การจับแน่นขึ้น ผลการวิเคราะห์โครงสร้างแสดงให้เห็นว่าโปรตีน YKL-39 มีบริเวณจับกับน้ำตาล 5 บริเวณย่อย ซึ่งครอบคลุมตำแหน่ง (-3)(-2)(-1)(+1)(+2) โดยมีกรดอะมิโนทริปโตแฟนตำแหน่ง 360 ทำหน้าที่หลักในการจับกับน้ำตาลด้วยแรงไฮโดรโฟบิคที่บริเวณศูนย์กลางของร่องจับของโปรตีน สำหรับการศึกษาอุณหพลศาสตร์ของโปรตีนจับกับไคตินสายสั้นด้วยวิธี isothermal titration calorimetry (ITC) และ intrinsic fluorescence spectroscopy ชี้ให้เห็นว่าโปรตีน YKL-39 จับกับน้ำตาลสายสั้นด้วยค่าคงที่ในการแตกตัว (K_d) ที่ความเข้มข้นระดับไมโครโมลาร์ และค่าพลังงานของการจับจะเพิ่มขึ้นตามจำนวนหน่วยของน้ำตาลที่ยาวขึ้น โดยค่าที่พบไม่มีความแตกต่างอย่างมีนัยสำคัญระหว่างน้ำตาลไคติน 5 หน่วย และ 6 หน่วย จากผลการทดลองนี้ได้สนับสนุนข้อมูลทางโครงสร้างสามมิติของโปรตีนที่แสดงบริเวณการจับต่อน้ำตาล 5 หน่วย และการวิเคราะห์ทางอุณหพลศาสตร์ชี้ให้เห็นว่าการจับของน้ำตาลต่อโปรตีน YKL-39 ถูกขับเคลื่อนโดยเอนทัลปี จากผลการศึกษาทั้งหมดเสนอให้เห็นว่าโปรตีน YKL-39 จับกับน้ำตาลอย่างไร และอาจส่งผลให้เกิดการกระตุ้นสัญญาณต่อระบบภูมิคุ้มกันอัตโนมัติและการปรับรูปร่างของเนื้อเยื่อ

สาขาวิชาชีวเคมี   ลายมือชื่อนักศึกษา_____

ปีการศึกษา 2556   ลายมือชื่ออาจารย์ที่ปรึกษา_____

                            ลายมือชื่ออาจารย์ที่ปรึกษาร่วม_____

ARAYA RANOK : FUNCTIONAL AND STRUCTURAL CHARACTERIZATION OF YKL-39, A HUMAN CHITINASE-LIKE PROTEIN. THESIS ADVISOR : ASSOC. PROF. WIPA SUGINTA, Ph.D. 194 PP.

FUNCTIONAL AND STRUCTURAL CHARACTERIZATION OF YKL-39, A HUMAN CHITINASE-LIKE PROTEIN

Human cartilage chitinase 3-like protein 2 (CHI3L2 or YKL-39) is a member of family-18 glycoside hydrolases that lacks chitinase activity. YKL-39 is known as a potential marker for the activation of chondrocytes and the progression of osteoarthritis. This thesis is divided into two parts. The first part involves cloning and expression of YKL-39 in the bacterial system that was used as for production of anti-YKL39 polyclonal and monoclonal antibodies. Both antibody types were highly selective, reacting only with YKL-39. Isotype mapping identified two generated hybridoma clones (so called clones 6H11 and 8H3) as the IgM isotype. Dot blot assay showed that the monoclonal antibody was strongly active with the synovial fluid of an osteoarthritis patient, and human monocyte and T lymphocyte cell lines. A database search for protein binding partners gave high hits with several glycoproteins that play particular roles in cartilage tissue scaffolding, connective tissue formation, and cell-cell interactions. In conclusion, anti-YKL39 polyclonal and monoclonal antibodies were raised and tested to be suitable for immunological application, such as the investigation of the YKL-39 regulating pathway and the development of an immunosensing tool for sensitive detection of cartilage tissue destruction.

In the second part, the binding of chitooligosaccharides to YKL-39 was investigated by protein crystallography and isothermal titration calorimetry. Four crystal structures of human YKL-39 were solved in the absence and presence of chitooligosaccharides. The overall structure of YKL-39 comprises a major $(\beta/\alpha)_8$ TIM barrel domain and a small $\alpha+\beta$ insertion domain. YKL-39 interacts with chitooligosaccharides through hydrophobic interactions, as well as a hydrogen bonding network. Detailed structural analysis revealed that the binding of chitin fragments induces local conformational changes that facilitate the tight binding and YKL-39 has the least extended chitin-binding cleft, containing five subsites for sugars, namely (-3)(-2)(-1)(+1)(+2), with Trp360 playing a prominent role in the sugar-protein interactions at the centre of the chitin binding cleft. Evaluation of binding affinities obtained from isothermal titration calorimetry and intrinsic fluorescence spectroscopy suggests that YKL-39 binds to chitooligosaccharides with K_d values in the micromolar concentration range and that the binding energies increase with the chain length. There were no significant differences between the K_d values of chitopentaose and chitohexaose, supporting the structural evidence for the five-binding subsite topology. Thermodynamic analysis indicates that binding of chitooligosaccharide to YKL-39 is mainly driven by enthalpy. In conclusion, our data suggest how YKL-39 could possibly interact with endogenous GlcNAc containing ligands that may stimulate the signaling cascades triggering autoimmune response and tissue remodeling.

School of Biochemistry   Student's Signature_____

Academic Year 2013   Advisor's Signature_____

Co-adviser's Signature_____

ACKNOWLEDGEMENTS

Araya Ranok

CONTENTS

CHAPTER

CONTENTS (Continued)

CONTENTS (Continued)

CONTENTS (Continued)

CONTENTS (Continued)

CONTENTS (Continued)

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF ABBREVIATIONS

(m, μ) L	(Milli, micro) liter
(m, μ) M	(Milli, micro) molar
(m, μ, n) g	(Milli, micro, nano) gram
A	Absorbance
Å	Angstrom(s)
aa	Amino acid(s)
Ab	Antibody
Ag	Antigen
B cell	Bursa produced cell
Bis Tris	Bis-(2-hydroxyethyl)imino-tris(hydroxymethyl)methane
bp	Base pair(s)
BSA	Bovine serum albumin
BSA	Bovine serum albumin
cDNA	Complementary deoxyribonucleic acid
CLPs	Chitinase-like proteins
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
EDTA	Ethylenediamine tetraacetic acid
ELISA	Enzyme linked-immunosorbent assay
FCA	Freund's complete adjuvant

LIST OF ABBREVIATIONS (Continued)

FCS	Fetal Calf Serum
FIA	Freund's incomplete adjuvant
GlcNAc	*N*-acetylglucosamine
GST	Glutathione S-transferase
h	Hour
HAT	Hypoxanthine Aminopetrin Thymidine
HRP	horseradish peroxidase
HT	Hypoxanthine-Thymidine
IMDM	Iscove's Modified Dulbeco's Medium
IPTG	Isopropyl-β-D-thiogalactopyranoside
ITC	Isothermal titration calorimetry
kDa	Kilodalton(s)
LB	Luria Bertani (broth)
mAb	Monoclonal antibody
min	Minute
Mr	Relative molecular mass
mRNA	Messenger ribonucleic acid
MW	Molecular weight
Ni-NTA	Ni-nitrilotriacetic acid
ºC	Degree Celsius
OD	Optical density
pAbs	Polyclonal antibody
PAGE	Polyacrylamide gel electrophoresis

LIST OF ABBREVIATIONS (Continued)

PBS	Phosphate buffered saline
PBS-T	Phosphate buffered saline containing Tween-20
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
pH	Negative logarithm of hydrogen ion activity
RMSD	root mean-square derivation
rpm	Round(s) per minute
s	Second
SDS	Sodium dodecyl sulfate
TEMED	N, N, N',N',-tetramethylenediamine
TMB	3, 3', 5, 5'-Tetramethylbenzidine substrate
Tris	Tris-(hydroxymethyl)-aminoethane
UV	Ultraviolet
v/v	Volume/volume
w/v	Weight/volume

CHAPTER I

INTRODUCTION

1.1 Chitin and chitinases

Chitin is a homopolysaccharide consisting of *N*-acetyl-glucosamine (GlcNAc) units joined together with *β*-1, 4-glycosidic linkages. Chitin is a highly abundant biopolymer in nature, second only to cellulose, and it serves as a major component of bacterial and fungal cell walls, exoskeletons of insects, microfilaria sheaths of parasitic nematodes, mollusks, protozoa, and crustaceans (Bussink, Speijer, Aerts, and Boot, 2007; Lee, 2009) . Although chitin is highly abundant, it does not accumulate in ocean floors, since chitin is degraded efficiently by chitinolytic enzymes from marine bacteria, especially in family *Vibrios* (Kawada, Hachiya, Arihiro, and Mizoguchi, 2007). Chitinases are hydrolytic enzymes that are responsible for the degradation of chitin into small oligosaccharide fragments. Chitin degradation is essential in a variety of biological processes. For example, bacteria produce chitinases to degrade chitin as their carbon and nitrogen sources (Cottrell, Moore, and Kirchman, 1999). Fungal chitinases have multiple functions, such as in nutrition, as well as in morphogenesis of filamentous fungi (Kuranda and Robbins, 1991). Plants are known to express chitinases for a defense mechanism against chitin-containing pathogens (Sahai and Manocha, 1993). Although humans do not contain endogenous chitin, chitinases expressed in humans have been reported to be associated with pathological conditions of certain infectious and inflammatory diseases (Elias, Homer, Hamid, and Lee, 2005;

Kawada *et al*., 2007; Kzhyshkowska, Gratchev, and Goerdt, 2007; Lee, 2009). In the Carbohydrate Active EnZymes (CAZy) database (http://www.cazy.org/), glycoside hydrolase families are described based on structurally-related catalytic and carbohydrate binding modules (or functional domains). Chitinases are grouped into family-18 and family-19 glycoside hydrolases (also referred to as GH-18 and GH-19 chitinases) (Henrissat and Bairoch, 1993). The two enzyme families do not share sequence identity and differ in both structures and mechanisms (Davies and Henrissat, 1995). GH-18 chitinases are mainly produced by bacteria, fungi, viruses, plants and mammals. The catalytic domain of GH-18 chitinases consists of an $(\alpha/\beta)_8$-TIM barrel fold, with a chitin binding domain (ChBD) attached either on the *N*-terminus or the *C*-terminus (Figures 1.1A and 1.1B). All GH-18 chitinases contain the catalytic domain with the DXXDXDXE conserved sequence motif (van Aalten *et al*., 2000), in which a glutamic residue at the end of strand $\beta 4$ is identified as the proton donor in the hydrolytic reaction (Brameld and Goddard, 1998; Kuranda and Robbins, 1991; Tews, Terwisscha van Scheltinga, Perrakis, Wilson, and Dijkstra, 1997; Watanabe *et al*., 1997). The catalytic mechanism with retention of the anomeric configuration of GH-18 chitinases is achieved by "substrate-assisted catalysis", by which the carbonyl oxygen of the *N*-acetyl group of the -1 sugar acts as a nucleophile (Cohen-Kupiec and Chet, 1998). In contrast, GH-19 chitinases are found mostly in plants (Iseli, Armand, Boller, Neuhaus, and Henrissat, 1996) and some Gram-positive bacteria, such as *Streptomyces* species (Iseli *et al*., 1996; Ohno *et al*., 1996). The catalytic domain of GH-19 chitinases consists of high α-helical content (Figure 1.1C) (Hart, Pfluger, Monzingo, Hollis, and Robertus, 1995; Monzingo, Marcotte, Hart, and Robertus, 1996). The 3D-structure of GH-19 chitinases shares a common fold with family

GH-46 chitosanases and lysozymes from families GH-22, GH-23, and GH-24 (Monzingo *et al*., 1996). Their catalytic mechanism is a general acid-base mechanism that inverts the anomeric configuration of the glycone moiety (Cohen-Kupiec and Chet, 1998).
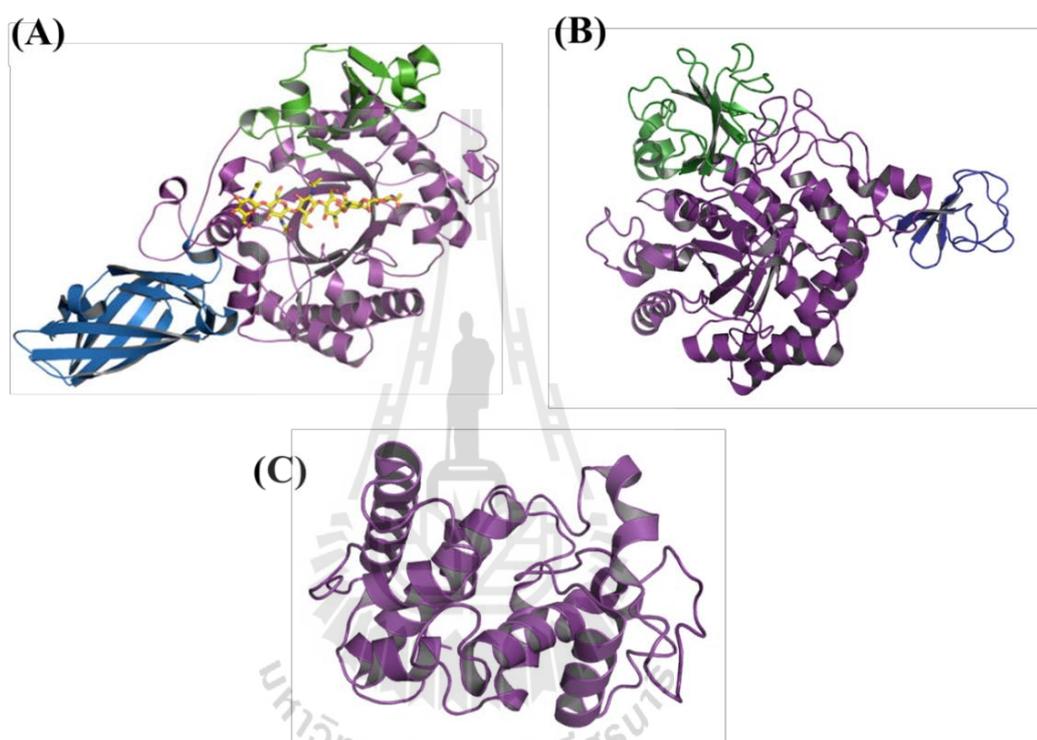
Figure 1.1 The 3D structures of GH-18 and GH-19 chitinases.

(A) GH-18 chitinase A with the ChBD attached on the *N*-terminus (blue) (Songsiriritthigul, Pantoom, Aguda, Robinson, and Suginta, 2008); (B) GH-18 chitinase B with the ChBD attached on the *C*-terminus (blue) (van Aalten *et al*., 2001). Both chitinase A and chitinase B contain the TIM barrel catalytic domain; (C) GH-19 chitinases with the bilobal catalytic domain (Hart *et al*., 1995; Monzingo *et al*., 1996).

1.2 Human chitinases

As mentioned earlier, there is no endogenous chitin in mammals. However, it is common for humans to be exposed to chitin, which is a major component of allergy-triggering antigens, such as cockroaches (Chew *et al*., 2008), house dust mites (Arlian and Platts-Mills, 2001) and helminthes, in daily life (Lee, Da Silva, Lee, Hartl, and Elias, 2008). To date, there are two human chitinases have been identified, Chitotriosidase and Acidic mammalian chitinase. They have been suggested to play a role in the defense mechanism against chitin-containing pathogens in the innate immunity of human (van Eijk *et al*., 2005).

Human macrophage chitinase (later known as chitotriosidase, CHIT1) was originally identified from the spleen of a symptomatic patient with Gaucher disease (GD, glucosylceramidosis). The source of the secreted chitotriosidase was found to be abnormal lipid-laden macrophages formed in tissues of Gaucher disease patients (Gaucher cells). The level of chitotriosidase in the plasma of Gaucher disease patients was found to be increased more than 600 times when compared with healthy people (Hollak, van Weely, van Oers, and Aerts, 1994). Thus, it has been proposed as a biomarker for diagnosis and monitoring the effectiveness of therapeutic approaches for treatment of GD (de Fost *et al*., 2006; Hollak *et al*., 1994; Pastores and Barnett, 2005). The expression of chitotriosidase at the mRNA level was also detected in lymph nodes, lung, and bone marrow that are involved in immune system (Boot, Bussink, and Aerts, 2005). Moreover, macrophages and neutrophils were also discovered to be a source of chitotriosidase in human (van Eijk *et al*., 2005). *In vitro*, monocyte-derived macrophages can be indeed to express chitotriosidase by stimulus

with proinflammatory cytokines, GM-CSF, TNF-α, and lipopolysaccharide (LPS), while IFN-γ and IL-4 inhibit the expression of chitotriosidase (Di Rosa *et al*., 2006; Malaguarnera, Musumeci, Di Rosa, Scuto, and Musumeci, 2005). *In vivo*, overexpression of chitotriosidase was found in several of diseases including atherosclerosis and sarcoidosis that are involved in chronic inflammation (Boot *et al*., 1999). Expression of chitotriosidase mRNA was also identified in Alzheimer's disease (AD), Ischemic cerebrovascular dementia (CvD), nonalcoholic fatty liver disease steatohepatitis (NASH) (Di Rosa *et al*. 2006; Malaguarnera *et al*. 2005). Moreover, chitotriosidase is the only biomarker for the monitoring of lipid-laden macrophages and Fabry disease; an X-linked globotriaosylceramidosis caused by the deficiency in the lysosomal α-galactosidase A (AGA) (Vedder *et al*., 2006). Treatment with recombinant AGA results in normalization of plasma chitotriosidase levels (Vedder *et al*., 2006). There are two major enzymatic isoforms of chitotriosidase, including 50 kDa and 39 kDa forms (Renkema, Boot, Muijsers, Donker-Koopman, and Aerts, 1995). The 39 kDa isoform which lacks a *C*-terminal chitin-binding domain is made during the posttranslational processing, as showed in Figure 1.3 (Renkema *et al*., 1997). Chitotriosidase has been classified into GH-18 chitinases based on its amino acid sequence, structure, and ability to hydrolyze natural chitin in the cells wall of the fungal pathogen *Candida albicans* (Boot *et al*., 2001; Boot, Renkema, Strijland, van Zonneveld, and Aerts, 1995; Boot *et al*., 1998; Hollak *et al*., 1994). Increased chitotriosidase activity was identified in plasma of neonates with fungal and bacterial infections, suggesting that it can be used as an indicator of macrophage activity during the progression of infectious diseases (Labadaridis *et al*., 2005). In addition, enzymatically chitinase activity to hydrolyzed chitin was found in

both *in vitro* and in an animal model, suggesting that it may play a role in defense mechanism against chitin-containing pathogens in innate immunity (van Eijk *et al*., 2005).

The structure of CHIT1, with the *N*-terminal catalytic domain adopts the $(\alpha/\beta)_8$ TIM-barrel structure that has a chitin-binding domain at its *C*-terminus. Like other GH-18 chitinases, the β_4 strand within the TIM barrel domain of CHIT1 (also AMCase) contains a conserved sequence motif $D_{133}XXD_{136}XD_{138}XE_{140}$ that forms the active site of the enzyme with the last glutamic acid acting as the proton donor required for hydrolyzing the β (1, 4) glycosidic bond in a chitin chain as showed in Figure 1.3 (Brameld and Goddard, 1998; Kuranda and Robbins, 1991; Tews *et al*., 1997; Watanabe *et al*., 1997). Although no endogenous substrate for these true chitinases has been identified in mammals, recent studies suggest that they have significant roles in inflammation, tissue injury responses, and pathogenesis of human diseases (Lee *et al*., 2011).
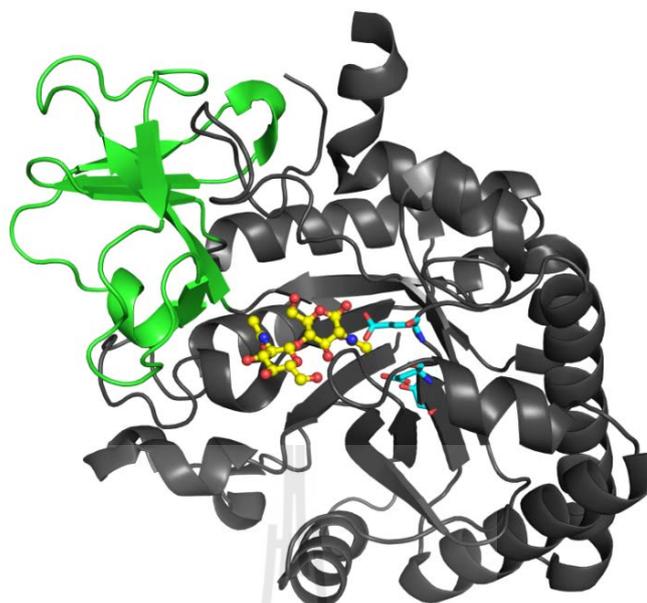
Figure 1.2 The 3D structure of 39 kDa isoform of human chitotriosidase that lacks ChBD at *C*-terminus.

The TIM barrel domain is made up of a cylinder of 8 parallel β-strands surrounded by 8 parallel α-helices (dark gray color) with small insertion domain (green clolor). The critical aspartic acid (D) and catalytic glutamic acid residue (E) at conserved sequence motif $D_{133}XXD_{136}XD_{138}XE_{140}$ are indicated in blue (Fusetti *et al*., 2002). The substrate $GlcNAc_2$ is showed as ball-stick yellow color.

Acidic mammalian chitinase (AMCase) is human form of GH-18 chitinase, which was initially identified in human stomach. AMCase is stable at extremely low pH and has an optimum pH of approximately 2 (Boot *et al*., 2001). AMCase has been suggested to be particularly involved in progression of asthma (Elias *et al*., 2005, a chronic disease characterized by exaggerated Th2 airway inflammation (Ray and Cohn, 1999). AMCase was reported to be highly up-regulated in the airways of asthmatic patients (Zhu *et al*., 2004) and in patients with pulmonary T helper 2 (Th2)

inflammations, but not with T helper 1 (Th1) inflammation (Robert *et al*., 2006). Another study in a mouse model of ovalbumin-induced bronchial asthma suggested that the AMCase was involved in the pathophysiology of asthma and acted downstream of interleukin-13 (Zhu *et al*., 2004). Administration of anti-AMCase antibody led to a significant decrease of Th2-inflammation, tissue eosinophilia and lymphocyte accumulation. Similar effects were detected after the application of a potent chitinase inhibitor, allosamidin (Zhu *et al*., 2004). Although the mechanism underlying the human chitinase activation is unknown, human chitinases may function as host defense against pathogenic infection and inflammatory responses (Shuhui, Mok, and Wong, 2009). Chitin-containing pathogens, such as dust mites, cockroaches or fungi, may be introduced into human's airway via inhalation. These allergens may also be introduced into the lungs, the pulmonary circulation and the gastrointestinal tract. These pathogens stimulate macrophages, which recruit eosinophils and basophils and inflammatory cells of the innate immune system to produce the type 2 helper T-cell (Th2 cell) cytokines interleukin-4 (IL-4) and interleukin-13 (IL-13). Such cytokines are known to induce the asthmatic phenotype. Secretion of the Th2 cytokines leads to an increase in the production of macrophages chitinase (chitotriosidase) and AMCase, both of which function enzymatically to kill chitin-producing pathogens and reduce inflammatory signaling by chitin (Figure 1.2).
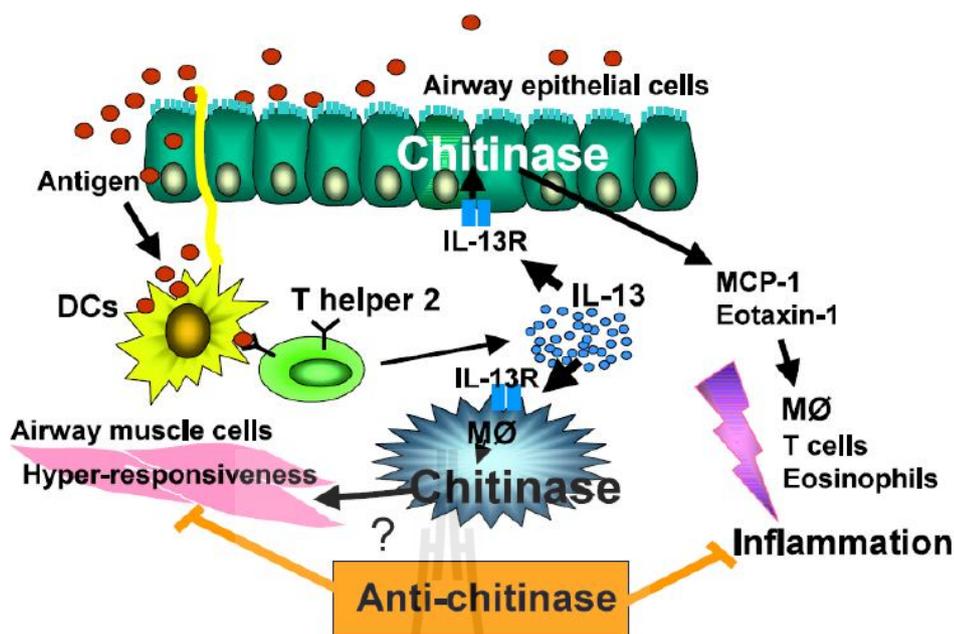
Figure 1.3 The role of human chitinases in host defense against pathogen infection and inflammatory responses of allergic inflammation in the airway of the asthmatic lungs (Kawada *et al*., 2007).

1.3 Human chitinase-like proteins

Recently, a number of chitinase-like proteins (CLPs) were identified in secretions of epithelial tissues and blood circulation (Chupp *et al*., 2007; Hakala, White, and Recklies, 1993; Johansen *et al*., 1996; Knorr, Obermayr, Bartnik, Zien, and Aigner, 2003). Expression of chitinase-like proteins is associated directly with pathogenesis of certain types of inflammatory diseases (Kzhyshkowska *et al*., 2007. CLPs have amino acid sequences homologous to GH-18 chitinases but lack chitinase activity due to substitution of the essential catalytic residue (glutamic acid) at the conserved sequence motif with leucine, isoleucine or tryptophan, as showed in Figure 1.4 (Kawada *et al*., 2007. Despite the lack of enzymatic activity, all CLPs contain the

$(\alpha/\beta)_8$ TIM-barrel domain and retain the ability to bind chitin and chitin oligosaccharides with high affinity (Figure 1.5) (Houston, Recklies, Krupa, and van Aalten, 2003). To date, four members of CLPs have been identified, including chitinase 3-like 1 protein (CHI3L1 or YKL-40 or HC *gp*-39) (Johansen *et al*., 1996), chitinase 3-like 2 protein (YKL-39) (Hu, Trinh, Figueira, and Price, 1996), oviduct-specific glycoprotein (oviductin or, Mucin9) (Arias, Verhage, and Jaffe, 1994), and stabilin-1 interacting chitinase-like protein (SI-CLP) (Kzhyshkowska *et al*., 2006).
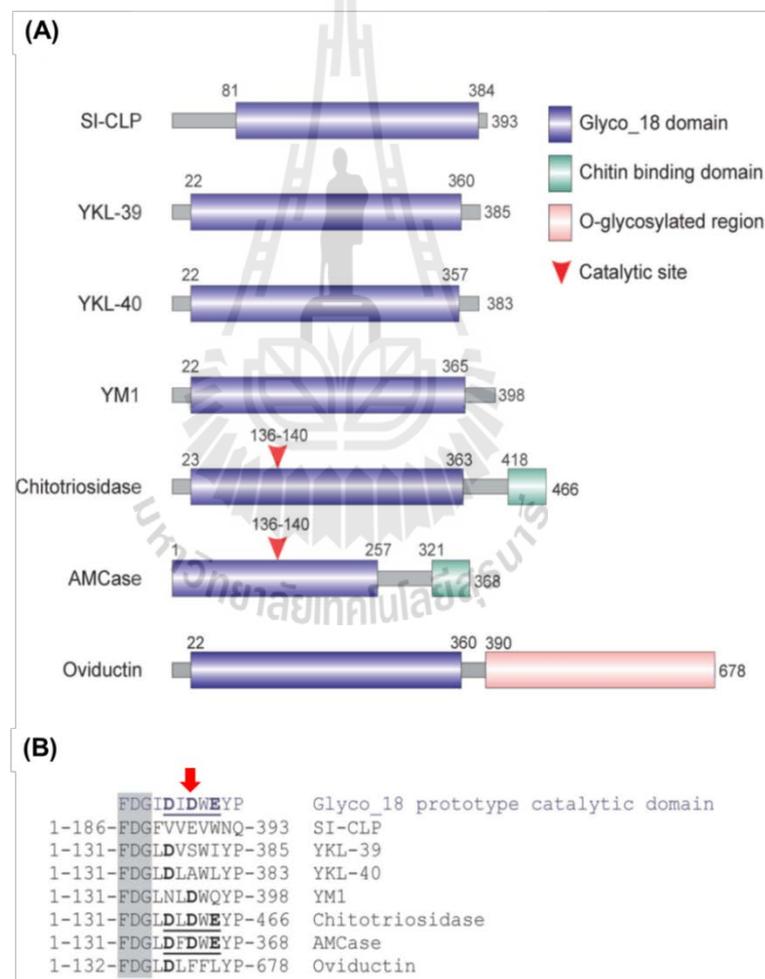
Figure 1.4 Human chitinases and chitinase-like proteins as members of GH-18. (A) A schematic representation of domain organization of GH-18 members, which include the TIM barrel domain, the ChBD, the *O*-glycosylated region, and the

cleavage site; (B) The conserved sequence motif within the GH-18 family. Chitotriosidase (CHIT) and AMCase contain Glu140, which acts as proton donor. For CLPs, the glutamic acid is substituted to Leu, Ile and Trp, making them lack catalytic activity (Kzhyshkowska *et al*., 2006).

YKL-40 (also called human cartilage glycoprotein-39 or HC *gp*-39, CHI3L1) was originally isolated as one of the major proteins secreted from cultured human osteosarcoma cell line MG63 (Johansen, Williamson, Rice, and Price, 1992). YKL-40 was named based on its three specific *N*-terminal amino acids: tyrosine (Y), lysine (K) and leucine (L), and its apparent molecular weight on SDS-PAGE analysis (about 40 kDa) (Johansen *et al*., 1992).

YKL-40 is secreted by macrophages and neutrophils (Volck *et al*., 1998), presumably in response to primary infection (Renkema *et al*., 1998). Because YKL-40 was highly expressed during pathological conditions associated with tissue inflammation, and development of fibrosis, it has been proposed that YKL-40 may involve in inflammation and tissue remodeling (Johansen *et al*., 2003; Kirkpatrick *et al*., 1997; Malinda, Ponce, Kleinman, Shackelton, and Millis, 1999; Shackelton, Mann, and Millis, 1995). Previous reports showed high levels of YKL-40 mRNA and protein expression in differentiated macrophages (Rehli, Krause, and Andreesen, 1997), and in a subset of macrophages under pathological conditions, such as rheumatoid arthritis (Kirkpatrick *et al*., 1997), giant cell arteritis (Johansen *et al*., 1999), arteriosclerosis (Boot *et al*., 1999), osteoarthritis (Volck *et al*., 2001) and small cell lung cancer (Junker, Johansen, Hansen, Lund, and Kristjansen, 2005). *In vitro* studies also confirmed high expression of YKL-40 in numerous human cancer cell

lines, including glioblastoma, colon cancer, ovarian cancer, prostate cancer, osteosarcoma, and malignant melanoma (Johansen, Jensen, Roslind, Nielsen, and Price, 2006). In addition, *in vivo* studies also reported elevated levels of YKL-40 in the circulating serum in number of solid tumors, such as breast cancer (Johansen *et al*., 2003), colorectal cancer (Cintin *et al*., 2002), ovarian cancer (Dupont *et al*., 2004, gastrointestinal cancer (Johansen *et al*., 2009), glioblastoma (Francescone *et al*., 2011, metastatic renal and prostate cancer (Brasso *et al*., 2006), malignant melanoma (Schmidt *et al*., 2006), small cell lung cancer (Johansen, Drivsholm, Price, and Christensen, 2004), and cervical adenocarcinoma (Mitsuhashi *et al*., 2009). However, the physiological roles of YKL-40 in cancer development are not yet known. Although serum concentrations of YKL-40 are not specific to a primary stage of solid tumors, YKL-40 has been proposed as a novel biomarker for the detection of advanced stages of various types of cancer (Johansen *et al*., 2006). Also, several studies demonstrated that high levels of YKL-40 expression are especially found in cancer patients with poor prognosis (Dehn *et al*., 2003; Jensen, Johansen, and Price, 2003; Johansen *et al*., 2006; Yamac *et al*., 2008).

The 3D structure of YKL-40 was solved by X-ray crystallography in a native form in complex with chitin oligosaccharides (Fusetti, Pijning, Kalk, Bos, and Dijkstra, 2003). The structural complexes showed that YKL-40 contained the $(\alpha/\beta)_8$ TIM-barrel structure with a small insertion domain in a fashion identical to GH-18 chitinases. In contrast to human chitinases, YKL-40 lacks the ChBD domain (Figure 1.5).
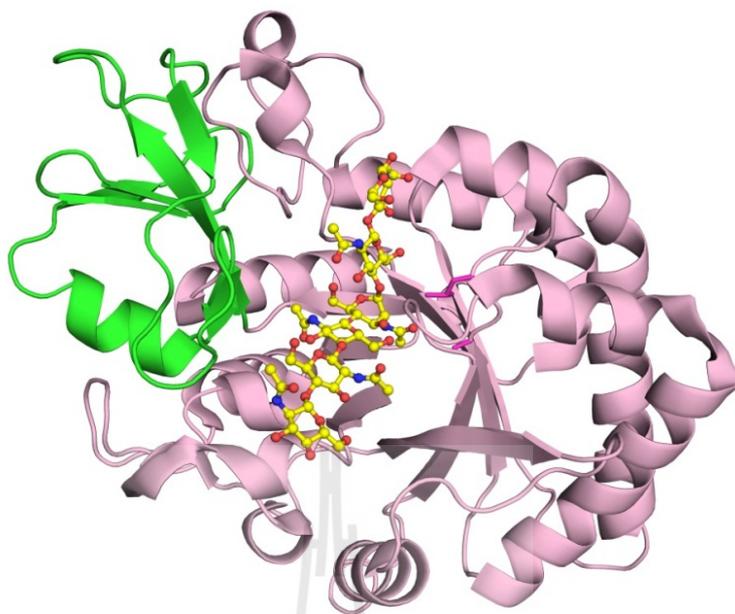
Figure 1.5 The ribbon representation of YKL-40 in complex with chitin oligosaccharide GlcNAc$_8$.

The $(\alpha/\beta)_8$ TIM barrel domain shown as purple color, while a small insertion domain shown as green color. The side chain of Leu140 that substitutes the catalytic Glu in GH-18 chitinases is shown as purple sticks. Chitin oligosaccharide is represented in ball and stick with carbons of GlcNAc$_8$ in yellow, nitrogen in blue and oxygen in red (Houston *et al*., 2003).

The crystal structures of YKL-40 complexed with various chitin oligosaccharides identified nine GlcNAc-binding subsites that are formed by a number of conserved aromatic residues (Trp[71], Trp[99], Tyr[141], Phe[208], Trp[212], and Trp[352]) that provide hydrophobic interactions with the hydrophobic faces of the sugar rings (Figure 1.6). Of all, Trp99 and Trp352 participate in ligand recognition by making hydrogen bonds with the C3-OH of the (-1)GlcNAc and the C7-O of the (-2)GlcNAc, respectively. Moreover, the crystal structure of YKL-40 with GlcNAc$_8$

showed that binding of the chitooligosaccharide to YKL-40 induced a large conformational change in the protein structure that is in contrast to chitinases. Even through, YKL-40 lacks enzymatically chitinase activity, it still retains the ability to bind chitin and chitin oligosaccharides with high affinity and the binding affinity depends on the length of chitin oligosaccharides (Fusetti *et al*., 2003; Houston *et al*., 2003). YKL-40 is also suggested as a lectin, which binds GlcNAc-containing ligands and possibly plays a role in innate immune responses to chitin-containing pathogens (Houston *et al*., 2003).
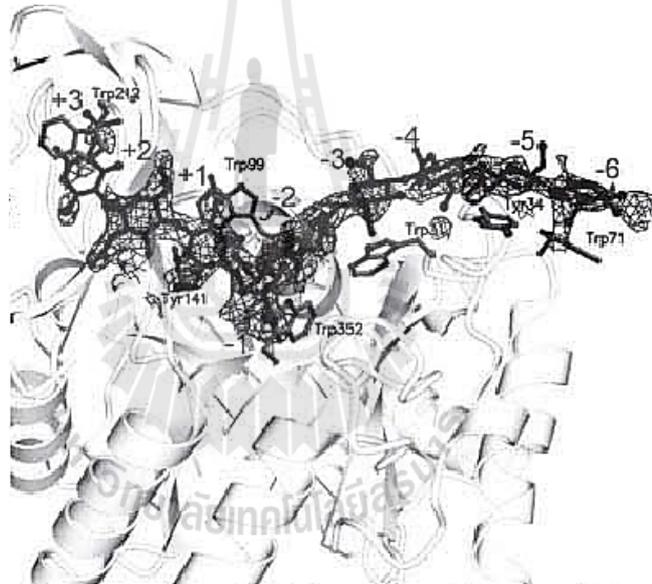
Figure 1.6   The structure of YKL-40 bound to chitin oligosaccharide.

The linings of aromatic amino acids that form the GlcNAc-binding cleft of YKL-40 are displayed in sticks (Houston *et al*., 2003).

Oviductin is a mammalian CLPs expressed and secreted by oviductal epithelium (Arias *et al*., 1994; Malette, Filion, St-Jacques, Kan, and Bleau, 1995). The protein was detected at highest level during time of ovulation. Its structure is

predicted to comprise an *O*-glycosylated mucin-like region, following the TIM-barrel domain. Oviductin has been suggested to play a role in regulating fertilization and early embryo development (Lapensee, Paquette, and Bleau, 1997; Lok, Briton-Jones, Yuen, and Haines, 2002). The oligosaccharide moieties in the *Zona pellucida* may bind to the TIM-barrel domain of oviductin, where by oviductin molecules might form a protective shield around the oocytes (Malette *et al*., 1995). Oviductin mRNA in fertile women was stimulated by estradiol and luteinizing hormones, while progesterone was shown to have a clear inhibitory effect (Briton-Jones *et al*., 2001). However a specific function of oviductin in human is not yet known.

SI-CLP is an interacting partner of the sorting receptor stabilin-1, which has been induced by Th2 cytokine IL-4 and glucocorticoid dexamethasone. *In vivo* studies showed that SI-CLP was highly expressed in macrophages from bronchoalveolar lavage of patients with chronic airway inflammation and in patients with sarcoidosis undergoing corticoid therapy (Kzhyshkowska *et al*., 2006). SI-CLP protein expression is not restricted to primary macrophages, but is also found in Raji cells, Jurkat cells and various tumor cell lines (Kzhyshkowska *et al*., 2006). Determination of the structure of SI-CLP by X-ray crystallography showed that it also displayed a TIM-barrel fold with four oligosaccharide binding subsites. Interestingly, SI-CLP could also bind a lipopolysaccharide (LPS) *in vitro* and neutralized its endotoxin effect on macrophages. The results suggested the possible role of SI-CLP as pathogen sensing and endotoxin neutralization (Meng *et al*., 2010).

Table 1.1 Cell type specificity and propose roles of CLPs.

CLPs	Substituted residue	Cell types	Functions
Oviductin	Leu	Oviduct epitheliums	Involved in regulating fertilization and early embryo development.
SI-CLP	Trp	MF	Associated with pathogen sensing and endotoxin neutralization.
YKL-40	Leu	MF, neutrophill, tumor cell, cancer cell line	Involved in inflammation and tissue remodeling.
YKL-39	Ile	articular chondrocyte, synovial cells	Involved in Osteoarthritis

1.4 YKL-39

YKL-39 (also known as CHI3L2) was originally discovered as a novel secreted protein in primary culture of human articular cartilage chondrocytes that was co-purified with YKL-40 (Hu *et al*., 1996). YKL-39 is currently recognized as a biochemical marker for chondrocyte activation and the progression of osteoarthritis in

humans (Knorr *et al.*, 2003. However, the exact biological function of YKL-39 is still unknown. YKL-39 is more closely related in size and sequence to YKL-40 than other members of the GH-18 family. Like YKL-40, the name of YKL-39 was given based on its three *N*-terminal amino acids: tyrosine (Y), lysine (K) and leucine (L) and its apparent molecular weight of 39 kDa. YKL-39 was co-purified with YKL-40 (Hu *et al.*, 1996). When the secreted protein YKL-40 from human articular chondrocyte culture was harvested and then fractionated by Sephacryl S-300 HR chromatography, the *N*-terminal sequencing data displayed the presence of two amino acid sequences, one of which had the sequence of YKLV*X*Y**YT S**WSQ**Y**R (YKL-40), whereas the other had the sequence of YKLV*X*Y**FTN**WSQ**D**R (later being identified as YKL-39) (Hu *et al.*, 1996). Functional characterization showed that YKL-40 is different from YKL-39, in that YKL-40 contains the heparin binding sequence at the *C*-terminus. Therefore, YKL-40 can bind to heparin with high affinity, whereas YKL-39 does not contain such a sequence. SDS-PAGE analysis showed that YKL-39 had a molecular mass slightly smaller than YKL-40 (Figure 1.7, 1a and 2a). Staining of the SDS-PAGE gel for carbohydrate showed that YKL-39 is not a glycoprotein (Figure 1.7, 2b), but YKL-40 is (Figure 1.7, 2b).
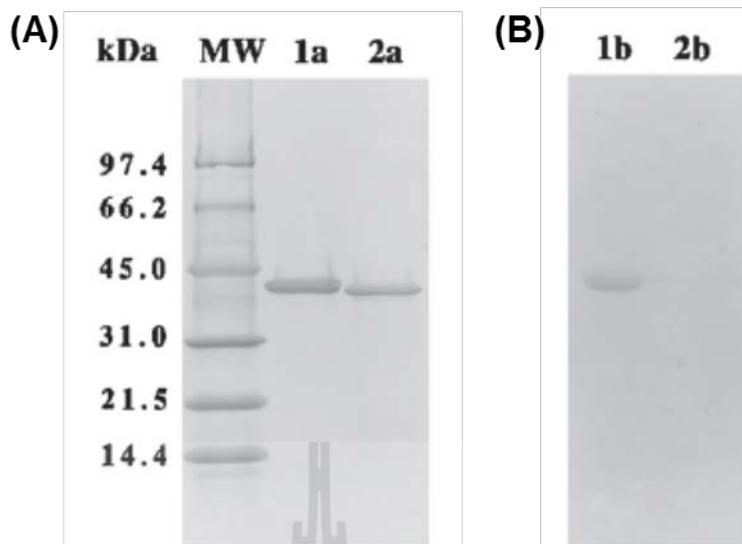
Figure 1.7 SDS-PAGE analysis of YKL-39 and YKL-40.

(A) Coomassie Brilliant Blue stain; (B) Glycoprotein stains using the Sigma glycoprotein detection. Lanes MW: molecular mass standards; lanes 1a and 1b: YKL-40; lanes 2a and 2b: YKL-39 (Hu *et al*., 1996).

A full length YKL-39 cDNA encodes a single polypeptide chain of 390 amino acids (Figure 1.8), which contains the 26-amino acid signal peptide (Hu *et al*., 1996). The mature protein, after removal of the signal peptide, has a length of 364 amino acids and a calculated molecular mass of 40,825 Da. Within the conserved sequence motif region, the glutamic acid, which is essential for catalytic activity of human chitinases, is substituted with isoleucine, thus causing YKL-39 to lack chitinase activity (Hu *et al*., 1996).

```
        10          20          30          40          50          60
MGATTMDQKS  LWAGVVVLLL LQGGSAYKLV CYFTNWSQDR QEPGKFTPEN IDPFLCSHLI

        70          80          90          100         110         120
YSFASIENNK VIIKDKSEVM LYQTINSLKT KNPKLKILLS IGGYLFGSKG FHPMVDSSTS

        130         140         150         160         170         180
RLEFINSIIL FLRNHNFDGL DVSWIYPDQK ENTHFTVLIH ELAEAFQKDF TKSTKERLLL

        190         200         210         220         230         240
TVGVSAGRQM IDNSYQVEKL AKDLDFINLL SFDFHGSWEK PLITGHNSPL SKGWQDRGPS

        250         260         270         280         290         300
SYYNVEYAVG YWIHKGMPSE KVVMGIPTYG HSFTLASAET TVGAPASGPG AAGPITESSG

        310         320         330         340         350         360
FLAYYEICQF LKGAKITRLQ DQQVPYAVKG NQWVGYDDVK SMETKVQFLK NLNLGGAMIW

        370         380         390
SIDMDDFTGK SCNQGPYPLV QAVKRSLGSL
```

Figure 1.8   Amino acid sequence of YKL-39 containing 390 amino acids with the 26 amino acids signal peptide sequence (gray box). The first three *N*-terminal amino acids of the mature protein are tyrosine (Y), lysine (K) and leucine (L) (blue box). The conserved sequence motif shows Ile substitution instead of Glu (the red letter in green shade).

YKL-39 accounts for 4% and YKL-40 for 33% of the total proteins secreted in chondrocyte-conditioned medium. YKL-39 and YKL-40, which are closely related in size and sequence, show more than 50% identical with each other (Figure 1.9).

```
                 *          20          *          40          *
YKL-39 : MGATTMDQKSLWAGVVVLLLLQGGSAYKLVCYFTNWSQDRQEPGKFTPEN :  50
YKL-40 : -----MGVKASQTGFVVLVLLQCCSAYKLVCYYTSWSQYREGDGSCFPDA :  45


                 60         *          80         *         100
YKL-39 : IDPFLCSHLIYSFASIENNKVIIKDKSEVMLYQTINSLKTKNPKLKILLS : 100
YKL-40 : LDRFLCTHIIYSFANISNDHIDTWEWNDVTLYGMLNTLKNRNPNLKTLLS :  95


                 *         120         *         140         *
YKL-39 : IGGYLFGSKGFHPMVDSSTSRLEFINSIILFLRNHNFDGLDVSWIYPDQK : 150
YKL-40 : VGGWNFGSQRFSKIASNTQSRRTFIKSVPPFLRTHGFDGLDLAWLYPGRR : 145


                 160        *         180         *         200
YKL-39 : ENTHFTVLIHELAEAFQKDFTKSTKERLLLTVGVSAGRQMIDNSYQVEKL : 200
YKL-40 : DKQHFTTLIKEMKAEFIKEAQPGKKQ-LLLSAALSAGKVTIDSSYDIAKI : 194


                 *         220         *         240         *
YKL-39 : AKDLDFINLLSFDFHGSWEKPLITGHNSPLSKGWQDRGPSSYYNVEYAVG : 250
YKL-40 : SQHLDFISIMTYDFHGAWRG--TTGHHSPLFRGQEDASPDRFSNTDYAVG : 242


                 260        *         280         *         300
YKL-39 : YWIHKGMPSEKVVMGIPTYGHSFTLASAETTVGAPASGPGAAGPITESSG : 300
YKL-40 : YMLRLGAPASKLVMGIPTFGRSFTLASSETGVGAPISGPGIPGRFTKEAG : 292


                 *         320         *         340         *
YKL-39 : FLAYYEICQFLKGAKITRLQDQQVPYAVKGNQWVGYDDVKSMETKVQFLK : 350
YKL-40 : TLAYYEICDFLRGATVHRTLGQQVPYATKGNQWVGYDDQESVKSKVQYLK : 342


                 360        *         380         *
YKL-39 : NLNLGGAMIWSIDMDDFTGKSCNQGP-YPIVQAVKRSLGSL : 390
YKL-40 : DRQLAGAMVWALDLDDFQGSFCGQDLRFPLTNAIKDALAAT : 383
```

Figure 1.9 Amino acids alignment of YKL-39 and YKL-40.

The alignment was created in ClustalW. Identity of their amino acid sequences is shown by gray shading. The three identical *N*-terminal amino acids (YKL) are shown in a black box.

```
                       *        20        *        40        *        60
YKL-39     : MGATTMDQKSLWAGVVVLLLLQGGSAYKLVCYFTNWSQDRQEPGKFTPENIDPFLCSHLI :  60
YKL-40     : -----MGVKASQTGFVVLVLLQCCSAYKLVCYYTSWSQYREGDGSCFPDALDRFLCTHII :  55
CHIT1      : -----MVRSVAWAGFMVLLMTPWGSAKLVCYFTNWAQYRQGEARFLPKDLDPSLCTHLI :  55
AMCase     : -----MTKLILLTGLVLIINLQLGSAYQLTCYFTNWAQYRPGLGRFMPDNIDPCLCTHLI :  55
Oviductin  : ------------MWKWVGVVKHHDGAAHKCY-LNWAHSRGASHDDCTHAASMNNNVAKD :  47
SI-CLP     : -----------MRTLFNLLWTALACSPVHTTLSKSDAKKAASKTLLEKSQFSDKFVQDRG :  49

                       *        80        *        100       *        120
YKL-39     : YSEASIENNKVIIKDKSEVMLYQTINSLKTKNPKLKILLSIGGYLFGSKGFHPMVDSSTS : 120
YKL-40     : YSEANISNDHLDTWEWNDVTLYGMLNTLKNRNPNLKTLLSVGGWNFGSQRFSKIASNTQS : 115
CHIT1      : YAFAGMTNHQISTTEWNDETLYQEFNGLKKMNPKLKTLLAIGGWNFGTQKFTDMVATANN : 115
AMCase     : YAFAGRQNNEITTIEWNDVTLYQAFNGLKNKNSQLKTLLAIGGWNFGTAPFTAMVSTPEN : 115
Oviductin  : DKYNKKRNRKTSGG-WN--GTSRTTMSTANRKASVSRTHDDCDYGRGSMHDRWTARKATM : 104
SI-CLP     : LVVTDLKAESVVLEHRSYCSAKARDRHFAGDVLGYVTPWNSHCYDVTKVFGSKFTQISPV : 109

                       *        140        *        160       *        180
YKL-39     : RLEFINSIILFLRNHNFDGLDVSWIYPDQK-----ENTHFTVLIHELAEAFQKDFTKSTK : 175
YKL-40     : RRTFIKSVPPFLRTHGFDGLDLAWLYPGRR-----DKQHFTTLIKEMKAEFIKEAQP-GK : 169
CHIT1      : RQTFVNSAIRFLRKYSFDGLDLDWEYPGSQGSPAVDKERFTLVQDLANAFQQEAQTSGK : 175
AMCase     : RQTFITSVIKFLRQYEFDGLDFDWEYPGSRGSPPQDKHLFTVLVQEMREAFEQEAKQINK : 175
Oviductin  : RRSAAVSGVHVT-SYDVRCRDNVSYDHGSWRTGHNSSDKSLAYAMNYWRKGASKMGTYGR : 163
SI-CLP     : WLQLKRRGREMFEVTGLHDVDQGAMRAVRKHAKGLHIVPRLLFEDWTYDDERNVLDSEDE : 169

                       *        200        *        220       *        240
YKL-39     : ERLLLTVGVSAGRQMIDNSYQVEKLAKDLDFINLLSEDFHGSWEKPLITGHNSPLSKGWQ : 235
YKL-40     : KQLLLSAASAGKVTIDSSYDIAKISQHLDFISIMTYDFHGSWEKPLITGHNSPLFRGQE : 227
CHIT1      : ERLLLSAAVPAGQTYVDAGYEVDKIAQNLDFVNIMAYDFHGSWEK--VTGHNSPLYRQE : 233
AMCase     : PRLLVTAAVAAGISNIQSGYEIPQLSQYLDYIHVMTYDLHGSWEG--YTGENSPLYRYPT : 233
Oviductin  : TRKASKNGARASGKYTKGAYCSVWGAKKHWDYYVLANKCKVVG-----YNASSYKAWR : 218
SI-CLP     : IEELSKTVVQVAKNQHFDGEVVEVWNQLLSQKRVGLIHMLTHLAEALHQARLLATLVIPP : 229

                       *        260        *        280       *        300
YKL-39     : DRGPSSYYNVEYAVGYWIHKCMPSEKVVMGIPTYGHSFTLA-S-AELTVGAPASGPGAAG : 293
YKL-40     : DASPDRFSNTDYAVGYMLRLGAPASKIVMGIPTFGRSFTLA-S-SELGVGAPISGPGIPG : 285
CHIT1      : ESGAAASLNVDAAVQQWLQKGTPASKIILGMPTYGRSFTLA-SSSDLRVGAPATGSGTPG : 292
AMCase     : DTGSNAYLNVDYVVNYWKDNGAPAEKIVGFPTYGHNFILS-NPSNLGIGAPTSGAGPAG : 292
Oviductin  : RHGGAMVWTDMDDVRGTCGTGVYVNDVRASSTSWSSAVNSSSTDRAVTTAWTTDSKGGAG : 278
SI-CLP     : AITPGTDQLGMFTHKEFEQLAPVLDGFSLMTYDYSTAHQPGPNAPLSWVRACVQVLDPKS : 289

                       *        320        *        340       *        360
YKL-39     : PITESSGFLAYYEICQELKG--------AKITRLQDQCVPYAVKGNQWVGYDLVKSMETK : 345
YKL-40     : RFTKEAGTLAYYEICDELRG--------ATVHRTLGQQVPYATKGNQWVGYDQESVKSK : 337
CHIT1      : PFTKEGGMLAYYEVCSWKG--------ATKQRIQDQKVPYIFRDNQWVGFDLVESFKTK : 343
AMCase     : PYAKESGIWAYYEICTELKNG--------ATQGWDAPQEVPYAYQGNVWVGYDNIKSFDIK : 345
Oviductin  : VTHGKCNMTTRGTTVTTKTVSGKHTVAGKTTGATTVTSVGHSMTGKATVGHSVTTGKTTS : 338
SI-CLP     : KWRSKILLGLNEYGMDYATSK-------DAREPVVGARYIQTLKDHRPRMVWLSQASEHF : 342

                       *        380        *        400       *        420
YKL-39     : VQELENLNLGGAMTWSIDMDDFTGKSCNQGP-YPLVQAVKRSLGSL-------------- : 390
YKL-40     : VQYLKDRQLAGAMVWALDLDDFQGSFCNQGP-YPLIQTLRQELSLPYLPSGTPELEVPKP : 383
CHIT1      : VSYLKQKGLGGAMVWAIDLDDFAGFSCNQGR-YPLIQTLRQELSLPYLPSGTPELEVPKP : 402
AMCase     : AQWLKHNKFGGAMVWAIDLDDFTGTFCNQGK-FPLISTLKKALGLQSASCTAPAQPIEPI : 404
Oviductin  : VGYSVTGKTTVGHSVTVSHSVSGGTTMTVHTTRNTVARRKAVARKVTVSRNS-------- : 390
SI-CLP     : FEYKKSRSGRHVVFYPTLKSLQVRLELARELGVGVSIWELGQGLDYFYDLL--------- : 393

                       *        440        *        460       *        480
YKL-39     : ------------------------------------------------------------ :  -
YKL-40     : ------------------------------------------------------------ :  -
CHIT1      : GQP--------SEPEHGPSPGQDTFCQGKADGLYPNPRERSSFYSCAAGRLFQQSCPTGL : 454
AMCase     : TAAPSGSGNGSGSSSSGGSSGGSGFCAVRANGLYPVANNRNAFWHCVNGVTYQQNCQAGL : 464
Oviductin  : --------------VTGTMRGNTSVGTHRMGNGMANRMMSSSVTADNRVYGNHSSVNSVT : 436
SI-CLP     : ------------------------------------------------------------ :  -

                       *
YKL-39     : ------------ :  -
YKL-40     : ------------ :  -
CHIT1      : VFSNSCKCCTWN : 466
AMCase     : VFDTSCDCCNWA : 476
Oviductin  : TSSKKNSAVDA- : 447
SI-CLP     : ------------ :  -
```

Figure 1.10 Human chitinases and chitinase-like proteins as members of GH-18. The

amino acid sequence alignment of human chitinases and CLPs is shown with the

conserved sequence motif in a black box. Chitotriosidase (CHIT) and AMCase contain Glu140 that acts as H^+ donor (arrow). For CLPs, this glutamic acid is substituted to Leu, Ile or Met.

1.4.1 Cell type specificity of YKL-39

YKL-39 was reported to be predominantly expressed in chondrocytes and synoviocytes (Hu *et al*., 1996). In addition, YKL-39 was also reported in glioblastoma cells (Kavsan *et al*., 2008). YKL39 mRNA was also detected in lung and heart, but not in brain, spleen, pancreas, and liver (Hu *et al*., 1996). Unlike human chitinases, YKL-39 was expressed in macrophages only under a strong stimulation by a combination of Th2 cytokine IL-4 and TGF-β growth factor (Gratchev *et al*., 2008). There is only one report presenting that upon the alternative activation macrophages can express YKL-39. They found that YKL-39 and YKL-40 mRNA were strongly overexpressed in the alternative activation of microglial macrophage in the brain of patients with Alzheimer's disease during the course of the disease (Colton *et al*., 2006).

1.4.2 Biological function of YKL-39

YKL-39 is currently recognized as a biological marker for the activation of chondrocytes and the progress of osteoarthritis (OA) (Knorr *et al*., 2003). Real time PCR and DNA microarray analysis showed that YKL-39 mRNA was significantly up-regulated in cartilage of patients with osteoarthritis (OA). While the YKL-39 mRNA expression was correlated with collagen type 2 up regulation, YKL-40 mRNA showed no significant up-regulation in osteoarthritic cartilage (Steck, Breit, Breusch,

Axt, and Richter, 2002). Another study reported that the mRNA expression of YKL-39 was elevated in both early degenerative and late stages of osteoarthritis, while that of YKL-40 was significantly suppressed during the progression of osteoarthritis, instead (Knorr *et al*., 2003). Proteomic analysis further identified YKL-39, but not YKL-40, to be secreted by human osteoarthritic cartilage culture (De Ceuninck *et al*., 2005). Recently, the biological role of YKL-39 in the proliferation or differentiation in chondrocytes was suggested (Miyatake *et al*., 2013). An overexpression of YKL-39 in chondrocytes enhanced cell proliferation approximately 3-fold, resulting in a significant increase in the percentage of the population in S-phase of the cell cycle. These results indicated that YKL-39 enhanced cell proliferation by stimulating G1/S transition. Moreover, overexpression of YKL-39 was shown to enhance the colony-forming of chondrocytes due to its increasing the expression of the cell adhesion molecule DDR1, and it also enhanced the expression of type II collagen. These data suggested that YKL-39 act as a novel growth/differentiation factor for articular cartilage chondrocytes that controls joint homeostasis in adults (Miyatake *et al*., 2013). YKL-39 was also found to induce autoimmune response of rheumatoid arthritis (RA) (Du *et al*., 2005; Sekine *et al*., 2001). ELISA and Western blot analysis showed that YKL-39 antibodies were detected in up to 8% of human serum samples of patients with RA, while only 1% of patients with RA had elevated level of YKL-40 antibodies. These results indicated that the pathway regulating immune response to YKL-39 was independent from that to YKL-40 (Sekine *et al*., 2001). When *BALB/c* mice were immunized with recombinant YKL-39, both anti-YKL-39 antibodies and the anti-type II collagen antibodies were detected, which suggested the spread of autoimmune reactions

(Sakata *et al*., 2002). Moreover, histological examination revealed synovial proliferation and irregularity of the cartilage surface in *BALB/c* mice (Figure 1.11). The results obtained from such studies suggested that YKL-39 acted as an inducer of autoimmune processes related to arthritis (Sakata *et al*., 2002).
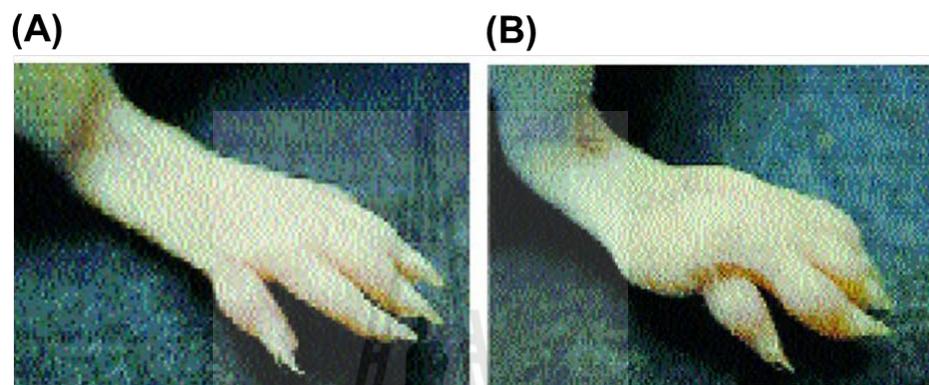
Figure 1.11 YKL-39 induced arthritis in *BAL /c* mice.

(A) The right foot pad of a control mouse and (B) the right foot pad of a YKL-39 immunized *BAL /c* mouse 30 days after the first immunization (Sakata *et al*., 2002).

Involvement of YKL-39 in tissue remodeling was suggested on the basis of the high level of the protein detected in chondrocyte cultures and it was shown to induce cell proliferation and cell migration. A recent report in human embryonic kidney (HEK293) and human glioblastoma (U87 MG) cells showed that YKL-39 activated signal transduction regulated by the phosphorylation of ERK1/ERK2 kinases, which led to either cell proliferation or differentiation (Areshkov and Kavsan, 2010). Even though, YKL-39 and YKL-40 are closely related in size and sequence, their effects in activating the transduction of MAPK signaling and ERK1/2 phosphorylation in the U393 cell line are different. The activation of ERK1/2 phosphorylation by YKL-40

leads to a proliferative signal, while the activation of ERK1/2 phosphorylation by YKL-39 inhibits cell mitogenesis and proliferation (Areshkov, Avdieiev, Balynska, Leroith, and Kavsan, 2012). However, a report on the YKL-39 function in the ATDC5 mouse chondrocytes cell line, showed that overexpression of YKL-39 enhanced cell proliferation of chondrocytes (Miyatake *et al*., 2013). One probable explanation for that difference is that sustainable activation of YKL-39 signal pathways may alter the cellular responses, which are observed in the transient activation (Miyatake *et al*., 2013).
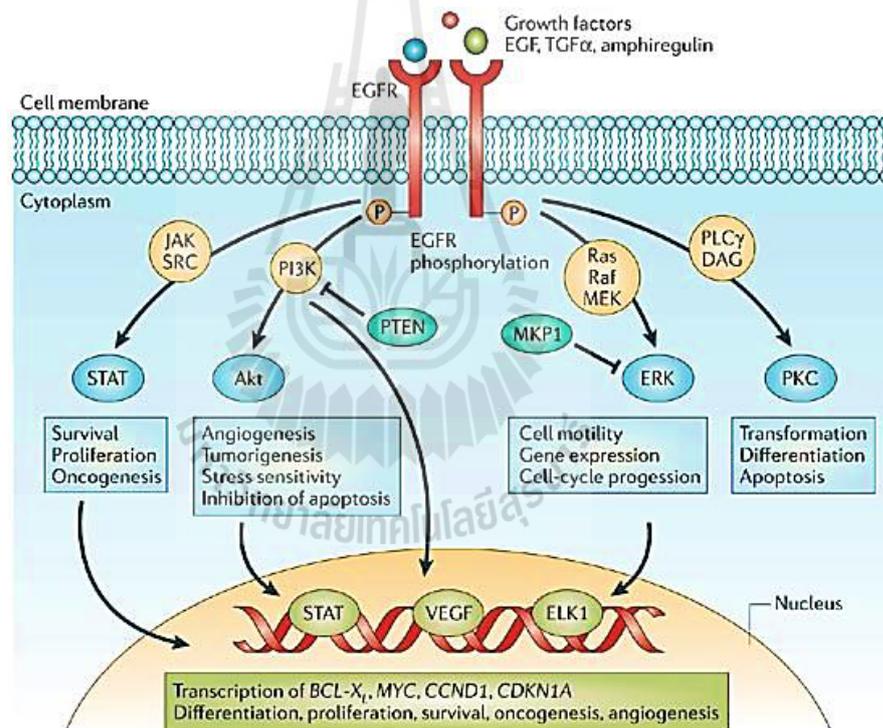
Figure 1.12 Overview of MAPK/ERK pathways in mammals.

Protein phosphorylation is involved in the regulation of a variety of growth and differentiation pathways through several phosphorylation cascades that communicate a signal from a receptor on the surface of the cell to the DNA in the nucleus of the cell to initiate cellular processes, such as proliferation, differentiation, and development.

The signaling cascades consist of many proteins, Raf, MEK1/2 and ERK act as the core components of the pathway (Nyati, Morgan, Feng, and Lawrence, 2006).

1.4.3   The 3D structure of YKL-39

To date, there is only one report on the 3D structure of YKL-39 in complex with $GlcNAc_6$ (Schimpl *et al*., 2012). The structure of this protein comprised a $(\beta/\alpha)_8$ TIM barrel domain with a small insertion domain similar to the YKL-40 structure. In the Schimpl *et al*. structure, the electron density map of the $GlcNAc_6$ chitooligosaccharide was well-defined only for four sugars in the -2 to +2 subsites (Figure 1.14A). The binding groove of YKL-39 was lined by a few aromatic residues that are conserved in the GH-18 chitinases, except that a Tyr at position 104 in YKL-39 was found in place of Trp in other GH-18 members (Figure 1.14B). Binding of YKL-39 with the chitooligosaccharide was achieved through a combination of hydrophobic stacking interactions between the sugar rings and aromatic amino acids and hydrogen-bonding, in a similar fashion to that found in chitinase-oligosaccharide complexes (Figure 1.14C). The 3D structure of YKL-39 confirmed that the essential catalytic Glu residue at the conserved sequence motif DxxDxDx**E** was substituted with Ile145, making it lack chitinase activity, as mentioned earlier (Figure 1.14B). Shimpl's YKL-39 sequence also presented the substitutions of D143 and E145 within the $D_{138}xxD_{141}x\mathbf{D_{143}}x\mathbf{E_{145}}$ conserved sequence motif to Ser143 and Ile145. However, reverse mutations of these two residues to aspartate and glutamate, generating the S143D/I145E double mutant that revealed significant chitinase activity with nanomolar range (k_{cat}/K_m = 2 s^{-1} nM^{-1}) that was lower than human chitinase CHIT1 (k_{cat}/K_m = 296 s^{-1} nM^{-1}). These results indicated that the active chitinase activity of

YKL-39 was recovered by returning the substitutions of essential catalytic acid residues in the binding site to the residues present in the active enzymes, suggesting that YKL-39 is a pseudo-chitinase with reservation of the chitin-binding function (Schimpl *et al.*, 2012).
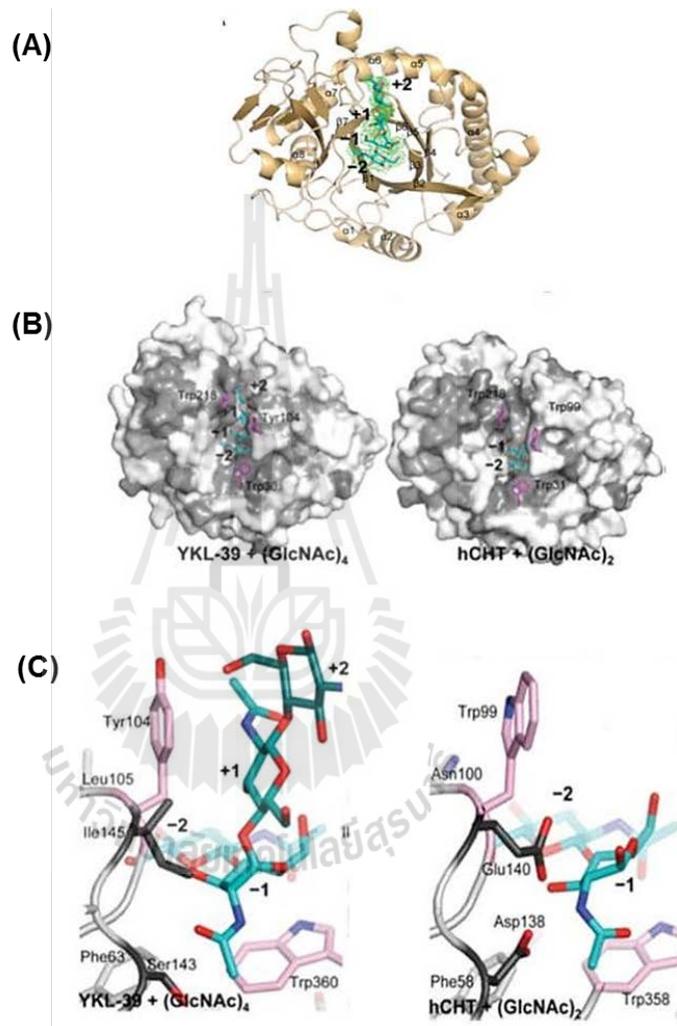
Figure 1.13 Structure of the YKL-39 in complex with chitohexaose.

(A) Stereo image of a cartoon representation of the structure of YKL-39 in complex with GlcNAc$_6$, showing four *N*-acetylglucosamine units of the ligand as cyan carbon atoms with electron density map of the chitinsugar contoured at 2.25 σ that shown as green color. (B) Surface representation of the YKL-39 in complex with chitooligosaccharide, with the conserved with chitotriosidase (PDB code 1LQ0)

colored by light grey, and identical residues colored by grey. The side chains of aromatic residues that are important to binding activity are shown as sticks with pink carbon atoms. The ligand $GlcNAc_6$ is shown as a stick model with cyan carbon atoms. (**C**) Stereo images of the ligand binding site of YKL-39 and the active site of chitotriosidase. The substitutions of catalytic residues in YKL-39 and catalytic residues in chitotriosidase are displayed with dark grey carbon atoms. The side chains of aromatic residues that are involved in binding property are presented as sticks with pink carbon atoms.

1.5 Research objectives

To address the functional and structural roles of YKL-39, this thesis has been divided into two parts. The first part involves cloning, and expression of YKL-39 for production of anti-YKL39 antibodies. In the second part, the structures of YKL-39 and its ligand complexes were solved using protein X-ray crystallization technique. The binding kineteics of YKL-39 against various chitoologosaccharides were investigated by Isothermal titration calorimetry (ITC) and fluorescence spectroscopy.

The objectives of this research are

1. To clone, express and purify human chitinase-like protein YKL-39 in *E. coli* BL21(DE3) cells.

2. To produce polyclonal and monoclonal antibodies specific for human YKL-39.

3. To solve the X-ray structures of YKL-39 protein alone and in the presence of various chitooligosaccharides.

4. To determine binding affinities of YKL-39 towards its specific sugar ligands.

CHAPTER II

MATERIALS AND METHODS

2.1 Materials

2.1.1 Bacterial strains and plasmids

Escherichia coli strain DH5α was used for routine cloning and plasmid preparation. The expression vector pET32a(+), which is a thioredoxin fusion system, was used for cloning and expression of the full-length *YKL-39* cDNA. *E. coli* type strain BL21(DE3) cells were used for high-level expression of recombinant thioredoxin-His$_6$/YKL-39 fusion protein.

2.1.2 Molecular cloning and plasmid purification

Chemicals and reagents used for molecular cloning were of molecular biology grade. All oligonucleotide primers (Table 2.1) used for amplification and DNA sequencing of *YKL-39* gene were ordered from BioDesign (Bangkok, Thailand). Human cDNA was purchased from GenScript Corporation (Piscataway, NJ, USA). *Pfu* DNA polymerase, *Pfu* polymerase 10x buffer, dNTP mix, *Bam*H1 and *Xho*I restriction enzymes, T4 DNA ligase, ligase 10x buffer and pGEM-T easy vector system were purchased from Promega (Madison, WI, USA). High-Speed Plasmid Mini Kit was product of Geneid (Taipei, Taiwan) and QIAquick Gel Extraction Kit of QIAGEN (QIAGEN GmbH, Hilden, Germany). HyperLadder^{Tm} I DNA marker was

bought from Bioline (Tacerton MA, USA) and SYBRÒ Green I nucleic acid gel stain from Sigma-Aldrich (St. Louis, MO, USA).

2.1.3 Protein expression and purification

Chemicals and reagents used for protein expression and purification were of analytical grade. Acrylamide, ammonium sulfate, ammonium persulfate, bromophenol blue, bis-N, N"-methylenebisacrylamide, calcium chloride, coomassie blue R250, coomassie blue G250, ethylenediamine tetra-acetic acid (EDTA), isopropyl β-D-thiogalactoside (IPTG), 2-mercaptoethanol, magnesium chloride, glycerol, glycine, sodium azide, sodium dodecyl sulfate (SDS) and Tris (hydroxymethylamine), *N, N', N", N"'*-tetramethylethylenediamine (TEMED) were products of Sigma-Aldrich. Ammonium hydroxide, glacial acetic acid, hydrochloric acid, methanol, n-butanol, phosphoric acid, potassium chloride, potassium hydroxide, sodium acetate, sodium bicarbonate, sodium chloride, sodium hydroxide, sodium dihydrogen phosphate and sodium hydrogen phosphate were obtained from Carlo ERBA (Rodano, Milano, Italy). Ampicillin, phenyl methylsulfonyl fluoride (PMSF), imidazole, isopropyl β-D-thiogalactoside (IPTG), hen egg white lysozyme, Triton X-100, Bacto tryptone, yeast extract, bacto agar and Luria-Bertani broth (LB) were product of USB Corporation (Cleaveland, OH, USA). Enterokinase was purchased from GenScript. Vivaspin-20 ultrafiltration membrane concentrators (10 kDa molecular-weight cut-off) were product of Vivascience (Hanover, Germany). 0.2 μm pore size filter membrane was bought from Nalgene (Rochester NY, USA). MF-Millipore Membrane Filters (0.22 μm and 0.45 μm pore size) were purchased from Millipore Corporation (Beverly, MA, USA). Ni-nitrilotriacetic acid (Ni-NTA) agarose

resin was a product of QIAGEN. Chromatographic media and pre-packed columns for automated protein purification were products of Amersham Pharmacia (Amersham Biosciences Biotech, Piscataway, NJ, USA). These pre-packed columns included HisTrap HP 1-mL, HisTrap HP 5-mL, Hiprep 26/10 Desalting column, HiTrap Q FF 5-mL, and HiLoad 16/60 Superdex 200 prep grade (1.6×60 cm) gel filtration columns.

2.1.4 Polyclonal antibody and monoclonal antibody production

Chemicals and reagents used for polyclonal and monoclonal antibody production were of analytical grade. Freund's complete adjuvant, polyethylene glycol (PEG), IsoQuick™ Kits for Mouse Monoclonal Isotyping, Hypoxanthine Aminopetrin Thymidine (HAT), and Hypoxanthine-Thymidine (HT) were purchased from Sigma-Aldrich. Horseradish peroxidase-conjugated goat anti-rabbit immunoglobulins (goat-anti-Igs-HRP) were purchased from Merck Millipore (Billerica, MA, USA). Horseradish peroxidase-conjugated rabbit anti-mouse immunoglobulins (Rab-anti-mIgs-HRP) were purchased from Dako (Glostrup, Denmark). Skimmed milk powder was product of Himedia (Himedia Laboratories Pvt. Ltd., Mumbai, India), 3, 3', 5, 5'-Tetramethylbenzidine substrate (TMB) was obtained from Invitrogen (Carlsbad, CA, USA). Protein A agarose column was a product of GenScript Corporation (CA, USA). BM condimed H1 (10x Hybridoma cloning supplement) was purchased from Roche (Roche Diagnostics GmbH, Mannheim, Germany). Fetal Calf Serum (FCS), Hybridoma-Serum free media, Iscove's Modified Dulbeco's Medium (IMDM) medium, and Dulbecco's Modified Eagle's Medium (DMEM) were obtained from Gibco® (Grand Island, NY, USA).

2.1.5 Protein crystallization

Crystallization screening kits used for finding crystallization conditions included JB Screen HTS I and HTS II that were purchased from Jena Bioscience GmbH (Jena, Germany), Wizard I and II were purchased Emerald (Emerald BioSystems, Bainbridge Island, WA, USA), and Crystal Screen High Throughput, HR2-110, HR2-112 and HR2-130 from Hampton Research (Aliso Viejo, CA, USA). Microbatch sitting drop 60 well plates and hanging drop vapor diffusion 24 well plates were purchased from Hampton Research or Nunc (Nunc GmbH, Wiesbaden, Germany). Al's oil, Nylon loop 0.05-0.5 mm, 20 mm Siliconized glass cover, Crystal Clear sealing tape were purchased from Hampton Research**.** High vacuum grease was purchased from Dow Corning Corporation (Midland, MI, USA). PEG3350 was a product of USB Corporation. 2, 2-bis (hydroxymethyl)-2, 2', 2"-nitrilotriethanol (Bis-Tris) were purchased from Sigma-Aldrich.

N-acetyl-chitooligosaccharides used for protein crystallization and binding studies, including di-*N*-acetyl-chitobiose, tri-*N*-acetylchitotriose, tetra-*N*-acetyl-chitotetraose, penta-*N*-acetyl-chitopentaose, and hexa-*N*-acetylchitohexaose, were purchased from Megazyme (Megazyme International Ireland Co., Ltd., Wicklow, Ireland).

2.2 Analytical programs

An amino acid sequence alignment of human chitinase and chitinase-like proteins was constructed using the "CLUSTALWW" algorithm in a GCG package (Thompson, Higgins, and Gibson, 1994) and displayed in Genedoc (http://www.psc.edu/biomed/genedoc/) and the DNASTAR package (DNASTAR,

Inc., Madison, WI, USA). The crystallographic were as follows. The HKL-2000 program (Otwinowski and Minor, 1997) was used for data indexing, refinement, and scaling. The *MOLREP* program (Vagin and Teplyakov, 1997) was used for molecular replacement method, the Coot program (Emsley and Cowtan, 2004) for model building, REFMAC5 in the CCP4 suit (Murshudov, Vagin, Lebedev, Wilson, and Dodson, 1999) for refinement, and PROCHECK (Laskowski, MacArthur, Moss, and Thornton, 1993) for validation of the stereochemistry of the final model.

2.3 Gene isolation and cloning of human YKL-39

A set of specific oligonucleotides was designed to amplify the full-length *CHI3L2* gene encoding YKL-39 (Genbank accession number NM_004000) by PCR technique using human uterus cDNA (GenScript Corporation) as DNA template. The forward oligonucleotide included a *Bam*HI cloning site and the reverse oligonucleotide included an *Xho*I cloning site. The sequences of both primers are shown in Table 2.1.

Table 2.1 Primers for synthesis of human YKL-39 cDNA. Underlined sequences indicate the restriction sites.

YKL-39	Primers
YKL-39BamHI	Forward 5'-GC<u>GGATCC</u>TACAAACTGGTTTGCTAC-3'
YKL-39XhoI	Reverse 5'-A<u>CTCGAG</u>TTACAGGGAGCCAAGGCTTCT-3'

The construct was designed to include six histidine residues tagged to the *C*-terminus of thioredoxin connected with the full-length human YKL-39 at the

C-terminus. The DNA fragment encoding YKL-39 was amplified by PCR performed under the following conditions: 35 cycles, denaturation at 97°C for 30 sec, annealing at 60°C for 1 min and extension at 72°C for 1 min, except for the final cycle where extension proceeded for 4 min. The amplified PCR product was then examined by 1% agarose gel electrophoresis that contained SyberGreen for detection of PCR products. After separating on agarose gel, the PCR product was observed under UV light using The Gel Doc 2000 system (Bio-Rad, USA). The PCR product of expected size (1.1 kb) was cut to further purify using the QIAquick Gel Extraction Kit (QIAGEN), following the manufacture's instructions. The purified DNA fragment encoding YKL-39 was cloned into the digested pET32a(+) vector, to generate the recombinant plasmid, called pET32a(+)/*CHI3L2*.In order to identify the presence of the insert *CHI3L2* gene in successfully constructed pET32a(+)/*CHI3L2*, colony PCR, restriction digestion analysis and DNA sequencing were carried out. Both sense and anti-sense strands of the *CHI3L2* gene were confirmed by automated DNA sequencing (First Base Laboratories, Selangor, Malaysia).

2.4 Recombinant expression and purification

2.4.1 Transformation of recombinant plasmid into *E. coli* BL21 (DE3) host cells

The recombinant pET32a(+)/*CHI3L2* was transformed into *E. coli* BL21(DE3) expression host cells as described by Sambrook *et al*. (1989). The competent cells of *E. coli* BL21(DE3) host were prepared as described in Appendix A. One microliter (50-100 ng) of plasmid was added to 100 μL of thawed *E. coli* stain BL21(DE3) competent cells was mixed gently by swirling and continued to store the

tube on ice for 30 min. The mixed cells were heat shocked at 42°C for 45 s, and then rapidly transferred to ice for 10 min. The transformed cells were added with 900 μL of LB broth, and then grown at 37°C for 45 min. Afterwards, 200 μL of the cell suspension was spread on LB agar plates, containing 100 μg/mL of ampicillin, and incubated overnight at 37°C.

2.4.2 Expression of fusion Trx-His$_6$/YKL-39

Single colonies of the *E. coli* BL21 (DE3) cells harboring the recombinant plasmid pET32a(+)/*CHI3L2* were grown overnight in LB medium containing 100 mg/mL ampicillin (LB/Amp) at 37°C with gentle shaking at 200 rpm. After overnight incubation, the bacterial culture was diluted to a ratio of 1:100 with LB/Amp broth, and further grown at 37°C until OD$_{600}$ reached 0.6-0.8. The culture was cooled down to 25°C for 30 min, then IPTG was added to a final concentration of 0.5 mM, and the culture was shaken at 200 rpm at 25°C for an additional 18 h. The induced YKL-39 expression cells were harvested by centrifugation at 4,500 rpm at 4°C for 30 min, and the cell pellet kept at -30°C until use. To obtain highest level of protein expression, different conditions for YKL-39 induction were optimized, including IPTG concentration, temperature, and induction time. The concentrations of IPTG were varied from 0-1.0 mM. Three induced temperatures, including at 15°C, 25°C and 37°C were tried with induction times of 2, 4, 8, 12 and 18 h. The condition, which gave highest level of protein expression, was chosen for further studies. To examine protein expression level, total protein profiles of *E. coli* BL21(DE3) cells, containing the recombinant pET32a(+)/*CHI3L2* plasmid with and without DNA insert were compared. After IPTG induction, one milliliter of the induced cells was

harvested by centrifugation, and then the cell pellets were resuspended in 100 µL of 1×SDS sample buffer (50 mM Tris, pH 6.8, 0.03% bromophenol blue, 2% SDS, 10% glycerol and 1.6% of 2-mercaptoethanol). The cell suspension was boiled for 5 min, then centrifuged at 10,000 rpm for 5 min and 5 µL of supernatant was analyzed on 12% SDS-PAGE. To obtain crude supernatant, containing soluble fusion protein, the induced cell pellet was re-suspended in 30 mL of lysis buffer (50 mM Tris-HCl, pH 8.0 containing 50 mM NaCl, 1 mM PMSF, 1 mg/mL lysozyme and 1% (v/v) triton X-100), and lysed on ice using a Sonopuls Ultrasonic homogenizer with a 6-mm-diameter probe (Sigma-Aldrich, St Louis, MO, USA). Sonication was carried out 15 times on ice with 30% power output, for 15 s, and with 1 min for cooling interval. Crude supernatant was transferred to a new vial after centrifugation at 12,000 rpm for 45 min.

2.4.3   Purification of recombinant YKL-39

2.4.3.1   Purification of YKL-39 for antibody production

Crude supernatant containing soluble fusion Trx-His$_6$/YKL-39 protein that obtained from Section 2.5.2 was applied onto a Ni-NTA agarose (1x1 mL, QIAGEN) column that was pre-equilibrated with five column volumes (CV) of equilibration buffer (50 mM NaCl, 50 mM Tris-HCl, pH 8.0). The column was operated manually under gravity flow. After sample loading, the column was washed thoroughly with 10 CV of 5 mM imidazole, followed by 10 CV of 20 mM imidazole in the equilibration buffer to remove unbound proteins. Subsequently, the bound fractions were eluted with 5 CV of 250 mM imidazole in equilibration buffer. These bound fractions that contained mainly the Trx-His$_6$/YKL-39 fusion protein were

pooled, and then thoroughly dialyzed against 50 mM NaCl in 50 mM Tris-HCl, pH 8.0, for a complete removal of imidazole. The protein solution was concentrated using Vivaspin-20 membrane concentrator (Mr cut-off 10,000, Vivascience AG). Further purification was performed by anion exchange chromatography on a 5x1 cm^2 HiTrap Q FF pre-packed column connected to an ÄKTAprime^{TM} plus system (GE Healthcare). The column was equilibrated with 5 CV of 20 mM Tris-HCl, pH 8.0, at a flow rate of 2.50 mL/min. One milliliter of concentrated protein was filtered through 0.45 μm pore size membrane filter (Millipore) before applying to the column, and then 1.0 mL fractions were collected. Unbound proteins were removed with 2 CV of 20 mM Tris-HCl, pH 8.0. Afterward, bound proteins were eluted with 10 CV of a linear gradient of 0-1 M NaCl in the same buffer. The containing Trx-His$_6$/YKL-39 fractions were pooled and concentrated, then were further cleaved by enterokinase for removal of the fusion tag thioredoxin. The cleavage reaction was carried out by incubating 0.2 unit of enterokinase per 100 μg of the fusion protein at 15°C for 18 h. Afterwards, the thioredoxin segment, which still connects with the His$_6$ at the *C*-terminus, was separated from the YKL-39 protein by a Ni-NTA agarose affinity column as described in the first step. Unbound fractions containing purified YKL-39 were pooled, exchanged into 20 mM Tris-HCl (pH 8.0), and concentrated to 10-20 mg/mL using the Vivaspin membrane concentrator. The purity of YKL-39 was verified by SDS-PAGE. Protein concentration was determined by the Pierce BCA assay (Novagen, Darmtadt, Germany). Aliquots of the purified YKL-39 were stored at -30°C until use.

2.4.3.2 Purification for crystallization and protein-ligand binding study

Crude supernatant containing soluble fusion Trx-His$_6$/YKL-39 obtained from section 2.5.2 was centrifuged at 19,000 rpm for 1 h, then filtered through a 0.45 μm pore size membrane filter (Millipore). Further purification was performed using two HisTrap HP (1x5 mL) pre-packed columns connected in series with Hiprep 26/10 Desalting column (Amersham Pharmacia), which was operated by an ÄKTAprime^{TM} plus FPLC system. The columns were equilibrated with 10 CV of equilibration buffer (50 mM NaCl, 50 mM Tris-HCl, pH 8.0), a constant flow rate of 1 mL/min was maintained and protein fractions of 2 mL were collected. After sample loading, the bound protein resin was washed with 25 CV of equilibration buffer to remove unbound proteins. Finally, the bound protein was eluted with 10 CV of 250 mM imidazole in equilibration buffer, then passed through a Hiprep 26/10 Desalting column (Amersham Pharmacia) to completely remove imidazole. The Trx-His$_6$/YKL-39 containing fractions were pooled and concentrated, then further cleaved by enterokinase to eliminate the Trx-His$_6$ fragment. The cleavage reaction was carried out by incubating with 0.2 unit of enterokinase per 100 μg of the fusion protein at 15°C for 18 h. Afterwards, the Trx-His$_6$ segment was separated from the YKL-39 protein by adsorption to a (1x1 mL) pre-packed HisTrap HP column (Amersham Pharmacia), connected to an ÄKTAprime^{TM} plus system. Unbound fractions containing purified YKL-39 were pooled, concentrated, and then further purified through a HiLoad 16/60 Superdex 200 prep grade gel filtration (Amersham Biosciences) to obtain the homogeneous protein solution. The column was equilibrated and eluted with 50 mM NaCl, 50 mM Tris-HCl (pH 8.0) with a flow rate

of 0.50 mL/min. The protein sample was filtered through 0.45 μm pore size membrane filter (Millipore), then 1.0 mL was injected to the system, and 1.0 mL fractions were collected. YKL-39-containing fractions were combined, exchanged into 20 mM Tris–HCl (pH 8.0), and concentrated to 10-20 mg/mL in the Vivaspin membrane concentrator. All purification steps were carried out at 4°C. The purity and protein concentration were determined as described in Section 2.4.3.1.

2.5 Polyclonal antibody production

2.5.1 Immunization

Two male New Zealand rabbits (eight week old) were immunized subcutaneously with the purified antigen YKL-39 emulsified in PBS with 50% Freund's complete adjuvant (Sigma-Aldrich). Immunizations were repeated twice every second week with the same antigen. Immunization and bleeding were scheduled as follows:

Week 0   Collection of pre-immune serum (10 mL)

            1st immunization with 500 μg YKL-39

Week 2   2nd immunization with 200 μg YKL-39

Week 4   3rd immunization with 200 μg YKL-39

Week 5   Collection of immunized serum (10-20 mL) and verification of anti-

            YKL39 production by Western blot technique

            4th immunization with 200 μg YKL-39

Week 6   3rd bleed 10 mL

Week 7   4th bleed 10 mL

Week 8   5th bleed 10 mL

Five days after the final boost, the immunized serum was collected from the central ear artery using a 19-gauge needle. The immunized serum was allowed to clot and retract at 4°C overnight, and then retrieved by centrifugation at 4,500 rpm for 15 min. The titer of anti-YKL39 antibodies in the serum was determined by indirect ELISA and Western blotting.

2.5.2 Purification of anti-YKL39 polyclonal antibodies

The immunized rabbit serum was further purified by affinity chromatography on a protein A agarose (1x1 mL) column (GenScript Corporation, USA). Serum sample was diluted with binding buffer (0.15 M NaCl, 20 mM sodium phosphate, pH 8.0) at a ratio of 1:1 to ensure that proper ionic strength and pH were maintained for optimal binding. The diluted serum sample was filtered through a 0.45 μm filter, and then applied onto protein A resin column, which was equilibrated with 10 CV of binding buffer. Bound IgG resin was washed with 30 CV of binding buffer. Bound IgG resin was washed with 30 CV of binding buffer. The bound antibodies were eluted with 10 CV of elution buffer (0.1 M glycine, pH 2.5), and were immediately neutralized to pH 7.4 with neutralization buffer (1 M Tris-HCl, pH 8.5) at a ratio of 1:10 volume of total eluted volume. The eluted fraction was then applied onto Hiprep 26/10 Desalting column (Amersham Pharmacia) connected with an ÄKTAprime^{TM} plus system for the exchanging buffer to PBS. The purity of the purified antibodies was verified by SDS-PAGE. Antibody concentration was determined by the Pierce BCA assay. Aliquots of the purified pAbs were flash-frozen in liquid N_2, followed by storage at -30°C until use.

2.5.3 Characterization of anti-YKL-39 antisera

2.5.3.1 Determination of the antibody titers and specificity by Western blotting

To determine the sensitivity of the raised antibodies, titers and cross-reactivity of the purified pAbs was evaluated by Western blotting. The purified YKL-39 (5 μg) was resolved on 12% polyacrylamide gel, and then electrophoretically transferred onto a nitrocellulose membrane (Amersham Pharmacia) by the semi-dry transferring technique at 15 volts for 1 h. After blotting, non-specific bindings were blocked by incubation with 5% (w/v) skimmed milk in 1x PBS for 1 h at room temperature. The blocked membrane was rinsed once with 1x PBS, then incubated for additional 1 h with anti-YKL39 antisera at various concentrations of the purified pAbs. Then, the membrane was washed five times with 0.1% Tween 20 in PBS, pH 7.2 (washing buffer, PBS-T) for 5 min of each washing, followed by incubation with horseradish peroxidase-conjugated goat anti-rabbit immunoglobulins (goat-anti-Igs-HRP) at a dilution of 1:10,000. After 1 h of incubation, followed by three more washes, signals representing antibody-antigen interactions were detected by the chemiluminescence using enhanced chemiluminescence reagent (Amersham Pharmacia), and then the membrane was exposed to X-ray film (Kodak, Belgium). Cross-reactivity of the purified pAbs (10 ng) was tested against three different YKL-39 homologues, including human YKL-40, human acidic mammalian chitinase (AMCase), and bacterial chitinase.

2.5.3.2 Determination of the antibody titers and specificity by indirect ELISA

Each well of the ELISA plate was coated with 50 μL of 10 μg/mL YKL-39 protein in carbonate/bicarbonate buffer, pH 9.6, at 4°C for 18 h. The coated plate was blocked with 2% (w/v) skim milk in PBS, pH 7.2 (blocking buffer) by incubation for 1 h at room temperature. Then, the plate was washed four times with 0.05% Tween 20 in PBS, pH 7.2 (washing buffer, PBS-T). Fifty microliters of various 2-fold serial dilutions of antiserum (1:10-1:20,480) or purified pAbs in the blocking buffer were added to the plates in triplicate. For specificity determination of pAbs, 96-well of ELISA plates were coated with varying protein concentrations with each protein concentration added to each well of the ELISA plate in triplicate. After 1 h of incubation at room temperature, the plate was washed for additional four times with the washing buffer, followed by adding 50 μL horseradish peroxidase-conjugated goat anti-rabbit immunoglobulins (goat-anti-Igs-HRP) at dilution of 1:2,000. After 1 h of incubation, and then three more washes, 100 μL of 3, 3', 5, 5'-tetramethylbenzidine substrate (TMB) (Invitrogen) was added and incubated at room temperature in the dark. The reactions were stopped by adding 100 μL of 1 N HCl. The intensity of the developed signals was determined spectrophotometrically with a Biochrom Anthos MultiRead Microplate Reader (Biochrome Ltd.) at a wavelength of 450 nm (A_{450}). The antibody-antigen binding curve was evaluated by a nonlinear regression function available in Prism version 5.0 (GraphPad Software, La Jolla, CA, USA) using the following single-site binding model.

2.6 Monoclonal antibody production

2.6.1 Immunization

Three eight-weeks-old female BALB/c mice were immunized intraperitoneally with 70 μg of the purified YKL-39 protein in sterile PBS at two weeks intervals. Serum was collected a week before each injection. The immunization and bleeding schedule was as follows:

Week 0   Collection of pre-immune serum, 1 mL

                1st Immunization with 70 μg of YKL39

Week 1   Collection of immunized bleed, 1 mL

Week 2   2nd Immunization with 70 μg of YKL39

Week 3   Collection of immunized bleed, 1 mL

Week 4   3rd Immunization with 70 μg of YKL39

Week 5   Verification of monoclonal antibody by indirect ELISA

Indirect ELISA was employed to determine the immune response of the mice prior to their use for cell fusion. Three days prior to cell fusion, mouse was boosted with 70 μg of YKL-39 protein in PBS.

2.6.2 Determination of antibody response in the immunized mice by indirect ELISA

To determine the antibody response in the immunized mice, indirect ELISA was employed. Each well of the ELISA plate was coated with 50 μL of 10 μg/mL YKL-39 in carbonate/bicarbonate buffer, pH 9.6, at 4°C for 18 h. The coated plate was blocked with 2% (w/v) skim milk in PBS, pH 7.2 (blocking buffer) by incubation for 1 h at room temperature. Then, the plate was washed four times with

0.05% (v/v) Tween 20 in PBS, pH 7.2 (washing buffer). Fifty microliters of 5 fold serial dilutions (1:5-1:10x10^6) of mice antiserum in the blocking buffer were added in triplicate. For determination of the specificity of the mAb, 96-well of ELISA plates were coated with varying protein concentrations with each protein concentration added to each well of the ELISA plate in triplicate. After 1 h of incubation at room temperature, the plate was washed an additional four times with the washing buffer, followed by adding 50 μL horseradish peroxidase-conjugated rabbit anti-mouse immunoglobulins (Rab-anti-mIgs-HRP) at dilution of 1:2,000. After 1 h of incubation and then three more washes, 100 μL of TMB substrate was added and incubated at room temperature in the dark. The reactions were stopped by adding 100 μL of 1 N HCl. The intensity of the developed signal was determined by measuring the absorbance in an ELISA plate reader at the wavelength 450 nm.

2.6.3 Hybridoma production

2.6.3.1 Preparation of splenocytes

The immunized mouse was sacrificed and its spleen was removed and placed in tissue culture dish containing 5 mL of Iscove's Modified Dulbeco's Medium (IMDM) medium (Gibco®). After rinsing, the spleen was transferred to a new tissue culture dish containing 5 mL of IMDM medium. Splenocytes were collected by crushing the spleen in IMDM medium. The cell suspension was transferred to a 50 mL falcon tube and then allowed to settle for 10 min. The cells were collected by centrifugation at 2000 rpm, 4°C for 5 min, while supernatant was discarded. Red blood cells (RBC) were lysed in 0.83% (w/v) NH$_4$Cl hypotonic solution. The RBC cell suspension was diluted by adding 20 mL of the

IMDM medium. Then, the splenocytes were washed two times with serum-free IMDM. The number and viability of the cells were counted in 0.2% (v/v) tryphan blue on a hemacytometer.

2.6.3.2 Preparation of mouse myeloma cell lines

Approximately two weeks prior to the fusion, the P3X63-Ag8.653 mouse myeloma cell line was thaw and cultured in a petri-dish plate containing complete IMDM media. They were subcultured daily at the splitting ratio 1:2 in order to maintain log phase of growth before the fusion process. It is crucial that the myeloma cells should be in log phase and healthy in order for the fusion process be successful. On the day of fusion, the myeloma cells were collected, and then washed two times with serum-free IMDM by centrifugation at 1,500 rpm, 4°C for 5 min. The number and viability of the cells were counted in 0.2% (v/v) tryphan blue on a hemacytometer.

2.6.3.3 Cells fusion

Splenocytes and mouse myeloma cells were mixed together at a ratio of 2:1, and then centrifuged at 1,500 rpm for 10 min. Cell pellets were incubated at 37°C for 5 min before being fused together by a standard hybridoma technique (Kohler and Milstein, 1975) using polyethylene glycol (PEG). Addition of reagents is as follows: 1.5 mL of 50% (v/v) PEG 1450, which was incubated at 37°C for 5 min before being used, was added in a drop wise fashion with gentle shaking for over 1 min, followed by continuously shaking for another minute. One milliliter of warm IMDM was added within 1 min. Then, 3 mL of warm IMDM was added into the tube

for over 1 min. Finally, 16 mL of the IMDM was added into the tube within 2 min. After centrifugation at 400 g for 10 minutes, cell pellets were incubated at 37°C for 5 min in a water bath before removing the supernatant. Then, the fused cells were re-suspended in 100 mL of Hypoxanthine Aminopetrin Thymidine (HAT) medium. One hundred microliters of the cell suspension were added in each well of a 96-well plate and cultivated at 37°C in a 5% CO_2 incubator. After five days of cultivation, 150 μL of Hypoxanthine-Thymidine (HT) medium was added into each well. The plates were incubated at 37°C in a 5% CO_2 incubator to allow the hybridoma cells to proliferate.

2.6.3.4 Hybridoma screening by ELISA

Supernatants taken from the culture medium, which contained hybridoma cells, were used to determine the production of monoclonal antibody against YKL-39 antigen by indirect ELISA, as described in Section 2.8.2.

2.6.3.5 Limiting dilutions

To obtain a single hybridoma clone, limiting dilutions were carried out. The hybridomas from the positive wells were counted and cell concentrations were adjusted to 4, 2 and 1 cell per 150 μL of the IMDM supplemented with 10% (v/v) BM condimed H1 and 10% (v/v) fetal calf serum (FCS) (Gibco®). Then, the cells were expanded by cultivation at 37°C in a 5% CO_2 incubator. After 7-10 days of cultivation, wells which contained single hybridoma clones were marked. Supernatant from the single clones was collected and checked by indirect ELISA for the presence of monoclonal antibodies capable of binding specifically to YKL-39. The single clones obtained from the limiting dilution

procedure that produced the highest levels of anti-YKL39 antibody were selected and further expanded.

2.6.4 Isotype determination of monoclonal antibodies

The isotypes of the newly-generated monoclonal antibodies (mAb) were determined by captured ELISA technique using IsoQuick™ Kits for Mouse Monoclonal Isotyping (Sigma-Aldrich). Each well of ELISA plate was coated with 50 μL of different goat anti-mouse immunoglobulin isotypes in carbonate/bicarbonate buffer, pH 9.6, at 4°C for 18 h. The coated plate was blocked with 2% (w/v) skim milk in PBS, pH 7.2, (blocking buffer) by incubation for 1 h at room temperature. Then the plate was washed four times with 0.05% (v/v) Tween 20 in PBS, pH 7.2 (washing buffer). Fifty microliters of 10 μg/mL of the produced mAb were added. After 1 hour of incubation at room temperature, the plate was washed four times with the washing buffer, followed by addition of 50 μL of rabbit-anti-mouse Igs-conjugated with HRP at dilution 1:2,000. After 1 h of incubation, following three washes, 100 μL of TMB substrate was added, and the plate was incubated at room temperature in the dark. The reactions were stopped by addition of 100 μL of 1 N HCl. The intensity of the developed color was determined by measuring the absorbance on an ELISA plate reader at the wavelength 450 nm.

2.6.5 Monoclonal production in serum free media

A hybridoma cell line capable of producing monoclonal antibody against human YKL-39 with IgM isotype was cultured in IMDM supplemented with 10%

(w/v) heat inactivated FCS, 40 μg/mL gentamycin and 2.5 μg/mL amphotericin B (Biolab, Smutprakarn, Thailand) in a humidified atmosphere of 5% (v/v) CO_2 at 37°C and was further subcultured in serum free media IMDM, containing 1% (w/v) penicillin/streptomycin. The antibody-containing culture media was collected, sterile filtered and stored at -30°C until use.

2.6.6 Monoclonal antibody purification

The culture media were collected and the presence of anti-YKL39 mAb was checked by indirect ELISA. The mAb was then purified by Hitrap IgM purification HP (Amersham Biosciences) connected to an ÄKTA purifier system. The culture media containing anti-YKL39 mAb was adjusted to the same concentration of ammonium sulfate as in the binding buffer. Small amounts of solid ammonium sulfate were added to the sample until the final concentration was 1.0 M, while stirring slowly and continuously during this procedure. Crude antibody was filtered through a 0.45 μm membrane filter immediately before loading on the column. Sample was applied to a Hitrap IgM column which was pre-equilibrated with 5 CV of binding buffer (20 mM sodium phosphate, 1.0 M $(NH_4)_2SO_4$, pH 7.5). Unbound materials were washed with 15 CV of the binding buffer, while bound IgM antibody was eluted with 12 CV of elution buffer (20 mM sodium phosphate, pH 7.5). The purity of the purified antibody was verified by SDS-PAGE. Antibody concentration was determined by the Pierce BCA assay. Aliquots of the purified mAb were stored at -30°C until use.

2.6.7 Characterization of anti-YKL39 monoclonal antibody

2.6.7.1 Specificity of anti-YKL39 mAb

To assess cross-reactivity of the purified mAb, Western blotting was carried out against three different YKL-39 homologues. These proteins included human YKL-40, human acidic mammalian chitinase (AMCase), and bacterial chitinase at various concentrations, as described in the polyclonal antibody production Section 2.6.3.1. The ability of the mAb to bind to the native form of human YKL-39 expressed by various human cell lines was also evaluated by immunoblotting.

2.6.7.2 Sensitivity of anti-YKL39 mAb

The sensitivity of generated anti-YKL39 mAb was also examined by Western blot analysis. Two-fold serial dilutions of anti-YKL-39 antibody were used, ranging from 0.65-10 µg, and YKL-39 was fixed at 2.5 µg. Detail is as described in the polyclonal antibody production section (2.6.3.1).

2.6.7.3 Cell culture

All human cell lines were obtained from the American Cell Culture Collection. Human T cell lines (Jurkat) and myeloid cell lines (U937 and THP-1) were cultured in Roswell Park Memorial Institute (RPMI-1640) medium, supplemented with 10% (w/v) heat inactivated FCS, 40 µg/mL gentamycin and 2.5 µg/mL amphotericin B in a humidified atmosphere of 5% (v/v) CO_2 at 37°C. Hela and COS-7 (monkey, kidney cell line) cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) under the conditions mentioned earlier. After cell passages of three times a week, the cells were collected, and then washed three times with PBS by

centrifugation at 1,500 rpm for 5 min. Then, 1×10^7 cells were lysed on ice for 30 min in 1 mL of lysis buffer (50 mM Tris-HCl (pH 8.2), 100 mM NaCl, 2 mM EDTA, 0.02% (w/v) NaN$_3$) containing 1% (v/v) Triton X-100, 1 mM phenylmethylsulphonylfluoride (PMSF), 5 mM iodoacetamide, and 10 µg/mL aprotinin) at 4°C for 30 min. Clear lysate was collected by centrifugation at 13,000 rpm, 4°C for 30 min, and used for functional assays.

2.6.7.4 Dot blot analysis

The purified YKL-39 or whole cell lysate prepared from various cell lines was spotted onto a nitrocellulose membrane, and then allowed to dry. Nonspecific binding was blocked by 5% (w/v) skim milk in 1% (v/v) Tween 20 in PBS (PBS-T) for 1 h at 28°C. The membrane was incubated for further 1 h with the mAb from clone 6H11. The membrane was washed five times in PBS-T, 5 min each, and incubated with rabbit-anti-mouse Igs-conjugated with HRP at a dilution of 1:5,000 for 1 h. After thoroughly washing, the signals were detected by the enhanced chemiluminescence method.

2.7 Protein crystallization

2.7.1 Initial screening by microbatch screening crystallization

2.7.1.1 Protein sample preparation

The freshly purified YKL-39 was concentrated in a Vivaspin-20 membrane concentrator (M_r cut-off 10,000) at 4,500 rpm, to 20 mg/mL stock concentration in 20 mM Tris-HCl, pH 8.0, and the protein was kept at 4°C until crystallization set up. To eliminate microbial contamination, dust, and aggregation,

the protein solution was filtered through a 0.2 μm centrifugal filter at 8,000 rpm for 5 min before crystallization setup. For initial screening, the microbatch under oil and sitting drop vapor diffusion methods were employed to obtain suitable conditions for reproducible single crystals for X-ray diffraction experiments.

2.7.1.2 Initial screening by microbatch under oil screening crystallization

Initial crystallization experiments were performed manually by the microbatch under oil method (Chayen, 1999). The initial screening was performed in 60 well plates (Nunc, Wiesbaden, Germany). Firstly, each well of the crystallization plate was filled with 10 μL of 100% paraffin oil to prevent evaporation of the aqueous drops. For each crystallization drop, 1 μL of each aqueous precipitant cocktail solution from Jena screen HTS I and HTS II (Jena Bioscience), Crystals Screen High Throughput HR2-110, or HR2-112 kit (Hampton Research) were added under oil into each well, followed by 1 μL of 10 mg/mL YKL-39 solution (Figure 2.1). To complete a single crystallization drop, each well was carefully examined under a Stemi 2000-C stereo microscope (Zeiss Corp, NJ, USA). The crystallization plate was covered with a plate cover to prevent dust from outside. The evaporation rate of water in the crystallization drop through the oil was minimized by placing it on a moist sponge in a plastic box, followed by incubation at 15°C or 25°C to allow protein crystals to grow. The crystal growth was examined daily for one week, and then continued once a week afterwards.
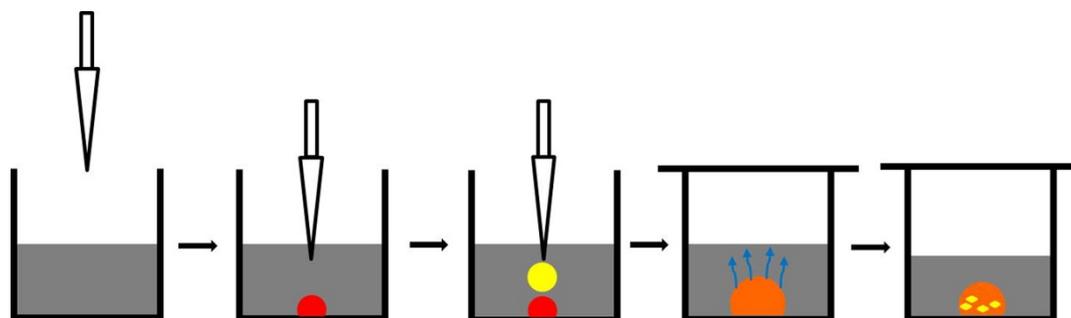
Figure 2.1 Schematic representation of Microbatch under oil crystallization technique**.**

2.7.1.3 Initial screening by robotic screening crystallization

From this section, crystallization experiments and protein-ligand binding by ITC technique were performed at the Institute of Molecular and Cell Biology (IMCB), Singapore, under advising by Prof. Dr. Robert C. Robinson, collaboration. Initial crystallization of YKL-39 was further set up using a Screenmaker 96+8TM Robot (Innovadyne Technologies, Inc., Santa Rosa, CA, USA) with sitting drop vapor diffusion technique in CrystalQuickTM plates (Greiner bio-one). The appropriate protein concentration was determined, and crystallization screening was performed by hanging drop technique using Pre-Crystallization Test (PCT, Hampton) as precipitant. Initially, 1 µL of the purified protein was mixed with 1 µL of the PCT reagents on a 20 mm siliconized cover slip, then the cover slip, containing the mixing drop was inverted and put on to the top of the sealed grease well of a 24 well plate containing 0.5 mL of PCT reagent. After incubation for 30 min at 25°C, the drop was carefully examined under a Stemi 2000-C stereo microscope. If the protein concentration for crystallization screening was appropriate,

microcrystallines or light granular precipitates would be observed throughout the drop. The protein concentration, which gave light granular precipitates, was then chosen for further optimization. For screening with the sitting drop vapor diffusion technique, the crystallization drop included 20 nL of the freshly prepared YKL-39 (12.75 mg/mL in 20 mM Tris–HCl, pH 8.0) and 20 nL of each precipitating agent from the screening kits, including JCSG-*plus* HT96 (Molecular Dimensions), Wizard I and II (Emerald) and Crystal Screen High Throughput HR2-130 (Hampton). Then, 170 nL of the same precipitant was applied into the well as reservoir (Figure 2.2). The crystallization plates were incubated at two different temperatures of 15°C and 25°C to allow crystal growth. The protein drops that were equilibrated over the reservoir of the precipitating agent were examined under a stereomicroscope periodically, as mentioned for the Microbatch under oil method.
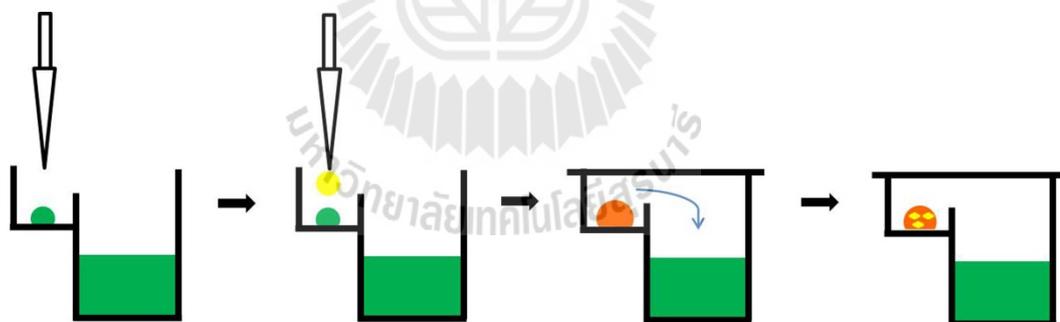
Figure 2.2   Schematic representation of Sitting Drop Vapor Diffusion technique.

2.7.2   **Optimization of crystallization conditions**

The initial screening conditions that yielded small crystals were further optimized by Hanging Drop Vapor Diffusion technique (Weber, 1997) to obtain high quality crystals. Factor variables, including concentrations of PEG and

salts, were adjusted at fixed pH and protein concentration during optimization. A grid of 24 conditions was set up as shown in the Figure 2.3. The crystallization droplet contained equal volumes of protein and mother liquor, for which 1 µL of YKL-39 (12.75 mg/mL) was pipetted onto the center of a 22 mm siliconizd cover slip, followed by an equal volume of the reservoir solution. The prepared cover slip, containing the crystaliization drop, was subsequently inverted and placed on the top edge of the wells, containing 0.5 mL of the reservoir (Figure 2.4), along the top edge of the well containing reservoir was firstly sealed with high vacuum grease. The prepared crystallization plates were subsequently incubated at 15°C and 25°C.
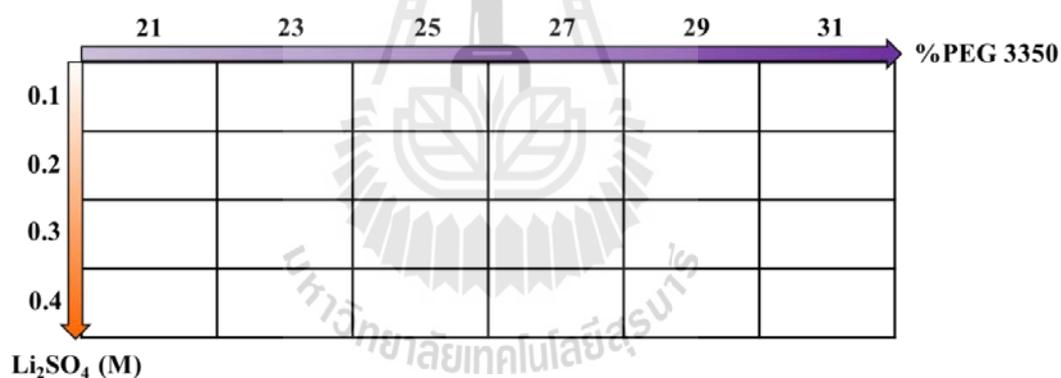
Figure 2.3 Grid screen of variable concentrations of PEG and salts.
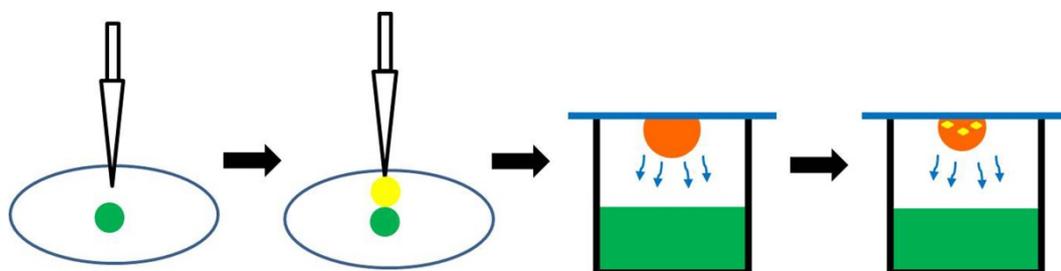
Figure 2.4 Schematic representation of Hanging Drop Vapor Diffusion technique.

2.7.3 Crystallization of protein/ligand complexes

To determine the crystal structures of YKL-39 in complex with its ligands, the YKL-39 crystals were grown under condition 27% (w/v) PEG 3350, 0.2 M Li_2SO_4, 0.1 M Bis Tris, pH 5.5 by the hanging drop vapor diffusion method. The YKL-39 crystals were soaked with various chitooligosaccharides ($GlcNAc_2$ to $GlcNAc_6$). Initially, a single X-ray quality crystal of YKL-39 was transferred to a new drop of precipitant containing each chitooligosaccharide at a specified concentration: 0.1 mM for $GlcNAc_5$ or $GlcNAc_6$, 5 mM for $GlcNAc_4$, and 10 mM for $GlcNAc_3$ or $GlcNAc_2$. The prepared glass cover slip, containing the soaked crystals, was subsequently inverted and placed on the top edge of the wells containing 0.5 mL of that precipitant as described for the hanging drop vapor diffusion method (Section 2.8.2). The crystal plates were then incubated at 25°C for 16 h. The survival of the crystals was observed every 4 h by stereomicroscope. The soaked crystals were mounted in a nylon loop, and flash-cooled in liquid nitrogen for X-ray diffraction.

2.7.4 Data collection and processing

2.7.4.1 Crystal transportation by dry shipper

The crystals were fished from the drop by a nylon loop connected with the wand plug into the bottom of the caps, then rapidly immersed into liquid nitrogen in a dewar. Afterwards, the loop containing a crystal was capped into the associated cryo vial (Hampton Research) immersed and kept in liquid nitrogen for the entire shipping time.

2.7.4.2 In-house X-ray diffraction

Initial X-ray diffraction data sets were collected in-house X-ray generator located at the Institute of Molecular and Cell Biology, Proteos, Singapore using the X-ray source from an RU-H3R rotating anode. The X-ray generator (Rigaku/*M*SC) was operated at 50 kV and 100 mA, with an incident wavelength setting at 1.54 Å. During diffraction, the crystals were maintained at 100K with a nitrogen cold stream that was generated by an X-Stream 2000 low temperature system. The X-ray diffraction was recorded on an R-AXIS IV++ image plate area detector (Rigaku/MSC). The crystal in its nylon-loop was placed onto the goniometer head, which allowed it to be positioned accurately within the X-ray beam and rotated between the X-ray beam and the detector, and the position of the mounted crystal was adjusted to center with assistance by a digital camera. The distance between the crystal and detector was varied from 60 to 70 mm in order to maximize the resolution of diffracted spots. The exposure time was set at 1 min. Once an X-ray diffraction image was observed, the first image was preliminarily analyzed by indexing to check crystal quality. Initial space group, lattice type, and unit cell parameters were then determined by a single oscillation image in the *HKL-2000* program (Otwinowski and Minor, 1997). For data collection, X-ray diffraction data were collected with 0.50° oscillations for 180° rotation or until the data were complete. All the diffraction images were also indexed, integrated, and scaled with the same program.

2.7.4.3 Synchrotron X-ray diffraction

All diffraction data of the ligand-free YKL-39 and the crystal complexes were collected on beamline BL13B1 at the National Synchrotron

Radiation Research Center (NSRRC, Hsinchu, Taiwan). The mounted crystal was kept with a nitrogen cold stream at 105K during the diffraction process. X-ray diffractions were recorded on the ADSC Quantum 315 CCD detector. The wavelength was set at 1.00 Å. The distance between the crystal and detector was varied from 200 mm to 350 mm in order to maximize the resolution of diffracted spots. The exposure time was set between 4 to 10 s and collected for 180° rotations with 0.50-1.00º oscillations per frame around the omega axis until the data were complete. All the diffraction images were also indexed, integrated and scaled with program *HKL-2000*. The maximum resolution for each data set was determined based on the criteria that the average ratio of measured intensity to its standard deviation ($I/\sigma(I)$) was more than 2-fold in the outer shell. The most widely used global quality indicator is the R_{merge} (Rhodes, 2000). It measures the spread of individual intensity measurements around the average value for the group of equivalent reflections. It has the formula

$$R_{merge} = [\Sigma_{hkl}\Sigma_i \ |I_i(hkl) - \langle I(hkl)\rangle|]/[\ \Sigma_{hkl}\Sigma_i I_i(hkl)] \tag{2.1}$$

Where I_i is the intensity for the *i*th measurement of an equivalent reflection with indices *hkl*.

R_{merge} at the highest resolution shell that strongly depends on redundancy was considered for evaluating the quality of data set, and was kept at lower than 40%. Moreover, the completeness was more than 90%, and redundancy was at least 4-fold in all shells. After the X-ray data collections were completed, the merged data of each data set were then converted into an *mtz* format using the program Scalepack2mtz from the CCP4 suite (CCP4, 1994). Scaling, merging and statistical analysis of all the data sets were further examined with programs from the CCP4 suite (CCP4, 1994).

2.7.5 Phase determination by molecular replacement method

To determine the orientation and position of the molecules within the unit cell, molecular replacement method (Navaza, 1994) was used to solve the phase problem. In this study, all the YKL-39 structures were solved by the molecular replacement using the program *MOLREP* from the CCP4 suite (Vagin and Teplyakov, 1997). The first data set to be processed was that of the YKL-39/GlcNAc$_2$ complex, which was processed to highest of resolution of 1.53 Å using the previously known structure of YKL-39 in complex with GlcNAc$_6$ (PDB code 4AY1) as a search model (Schimpl *et al*., 2012). The rotation and translation function solutions were solved for one protein molecule per asymmetric unit with correlation coefficient of 71.0% and R_{factor} of 46.7% in resolution range of 19.60-1.74 Å. The other data sets, including the ligand-free YKL-39, in complex YKL-39/GlcNAc$_4$ and YKL-39/GlcNAc$_6$ crystals were solved by using the final model of the YKL-39/GlcNAc$_2$ complex as a search model.

2.7.6 Model building and refinement

After obtaining initial phases from the search model, an initial model was manually rebuilt in Coot (Emsley and Cowtan, 2004). The electron density map of $|F_{obs}|$-$|F_{cal}|$ and $2|F_{obs}|$-$|F_{cal}|$ was calculated from the observed structure factor amplitudes and the calculated structure factor amplitudes. A new model was then fit to the new electron density map and further rounds of structural refinement were performed with Refmac5 in the CCP4 suit (Murshudov *et al*., 1999) and Phenix (Adams *et al.*, 2002). The resultant electron density map showed the regions of disagreement between the search model and the new structure, which was further

determined through rebuilding and refinement processes. These processes were continued until the correlation between the diffraction data and the model were maximized. The agreement was measured by an R_{factor} defined as

$$R = \frac{\Sigma\text{all reflections } |F_{\text{obs}} - F_{\text{cal}}|}{\Sigma\text{all reflections } |F_{\text{obs}}|} \qquad (2.2)$$

A similar quality criterion was R_{free}, which was calculated from a subset (~5%) of reflections that were not included in the structure refinement. The water molecule positions were added into the model in the refinement stages by inspecting the $\left||2F_{\text{obs}}|-|F_{\text{cal}}|\right|$ electron density map for peaks at 1.0σ.

2.7.7 Validation of model quality

The quality of geometry of the final model was validated by PROCHECK (Laskowski *et al.*, 1993). The Ramachandran plot (Ramachandran and Sasisekharan, 1968) of each model was calculated by PROCHECK. The number of Ramachandran outliers was found to be less than 0.2%. In addition, the root-mean-square (R.M.S.) deviations from ideality of bond length and angles were also determined with PROCHECK. The criteria of a well-refined model exhibits R.M.S. deviations were less than 0.014 Å for bond lengths and 3° for bond angles. All the structures and electron density maps were created and displayed with Pymol (www.pymol.org).

2.8 Protein-ligand interaction

2.8.1 Isothermal titration calorimetry (ITC)

ITC experiments were carried out using the ITC-200 system (Microcal Inc.). All solutions were degassed to avoid air bubbles in the calorimeter. All of measurements were carried out at 25°C with stirring speed of 260 rpm after the baseline was set with distilled water. Protein and chitooligosaccharide were prepared in 20 mM potassium phosphate buffer, pH 8.0. For the titration with $GlcNAc_6$ and $GlcNAc_5$, 10 µM of YKL-39 was placed in the reaction cell with a volume of 0.4 mL, and 0.25 mM ligand was placed in the syringe. Aliquots of 4 µL of ligands were injected into the reaction cell. The injections were repeated 20 times over 140 s intervals. The background was measured by injecting the ligand into the cell containing only buffer. ITC experiments with $GlcNAc_4$, $GlcNAc_3$ and $GlcNAc_2$ were performed as described for $GlcNAc_6$ but with higher concentrations of the reaction components. For the experiment with $GlcNAc_4$, 30 µM YKL-39 in 20 potassium phosphate buffer, pH 8.0, was added in the reaction cells and 0.45 µM $GlcNAc_4$ was placed in the syringe. For experiments with $GlcNAc_3$ and $GlcNAc_2$, 15 µM YKL-39 was added in the reaction cells and 3 µM $GlcNAc_3$ and 4 µM $GlcNAc_2$, respectively, were added into the syringe. The ITC data were collected by using the Microcal Origin v.7.0 software. All data were subtracted from the heat releasing profile obtained by injections of the corresponding ligand into the reaction cell after the saturated binding reaction. The data were fitted by a nonlinear least-squares algorithm fitted to a single-site binding model. After background subtraction, all data were fit to yield the stoichiometry (n), the equilibrium binding association constant (K_a), and the enthalpy change (ΔH) of the reaction. The Gibb's free energy (ΔG) and the entropy

change (ΔS) were calculated from the enthalpy change (ΔH) and binding constant with the following equation.

$$\Delta G = \Delta H - T\Delta S = -RT\ln(K_a) \qquad (2.3)$$

Where R represents the gas constant (1.98 cal K^{-1} mol^{-1}) and T the absolute temperature in Kelvin (K). The enthalpic term represents the change in heat that is produced by breaking and forming noncovalent bonded interactions, whereas the entropic term reflects changes in order. In these experiments, all data from the binding reactions fitted well to a single-site binding model with a calculated n (stoichiometry) of 0.9-1.1. The shape of the ITC binding curve is dependent on the association constant (K_a) and the molar concentration of macromolecule [M] being titrated which can be determined by the so-called Wiseman c-value (Wiseman, Williston, Brandts, and Lin, 1989). The relation is shown as:

$$c = nK_a[M] \qquad (2.4)$$

Where n is the stoichiometry of the reaction, K_a is the equilibrium binding association constant, and [M] is the molar protein concentration. For the binding of GlcNAc$_4$, GlcNAc$_5$, and GlcNAc$_6$, they could be optimized to give a c-value to be in the range of $10 < c < 1000$. This ensures that K_a can be determined from the Wiseman binding isotherm. For the binding of GlcNAc$_2$ and GlcNAc$_3$ to YKL-39 yielded a c-value of 0.1. Binding thermodynamics can be obtained by ITC even if c is in the range of $0.01 < c < 10$ if a sufficient portion of the binding isotherm is used for analysis (Turnbull and Daranas, 2003).

2.8.2 Intrinsic fluorescence studies of protein-ligand binding

The purified YKL-39 was titrated with different concentrations of the chitooligosaccharides in 20 mM Tris-HCl, pH 8.0 at 25°C. Changes in intrinsic tryptophan fluorescence were monitored directly in a LS-50 fluorescence spectrometer (Perkin-Elmer Limited). The excitation wavelength was 295 nm and emission intensities were collected over 300-450 nm with the excitation and emission slit widths being kept at 5 nm. Each protein spectrum was corrected for the buffer spectrum. Binding curves of the ligands, which were detected as increasing fluorescence intensities after adding the ligand, were evaluated by a nonlinear regression function available in Prism version 5.0 (GraphPad Software, California, USA) via the following single-site binding model:

$$F\text{-}F_0 = \frac{(F_{max} - F_0) \times (L_0)}{K_d + (L_0)} \qquad (2.5)$$

F and F_0 refer to the fluorescence intensity in the presence and absence of ligand, respectively; F_{max} refers to the maximum fluorescence signal of the protein-ligand complex; L_0 is the initial ligand concentration (µM) and K_d is the equilibrium dissociation constant (µM).

CHAPTER III

RESULTS

3.1 Molecular Cloning, Expression, and Purification of YKL-39

3.1.1 Molecular cloning

A database search showed that the *CHI3L2* gene encoding YKL-39 is located in human chromosome 1 at locus 1p13.3. In this study, the nucleotide sequence of *CHI3L2* cDNA in the GenBank database (www.ncbi.nlm.nih.gov/genbank) was used as the template to design a set of oligonucleotides, which was designed to generate in-frame protein expression. The full-length *CHI3L2* gene encoding YKL-39 (Genbank accession number NM_004000) was successfully amplified by PCR technique with human uterus cDNA (GeneScript) as the DNA template. The pairs of oligonucleotide primers could produce the expected PCR products (~1.1 kb), which were analyzed by agarose gel electrophoresis (Figure 3.1.1).
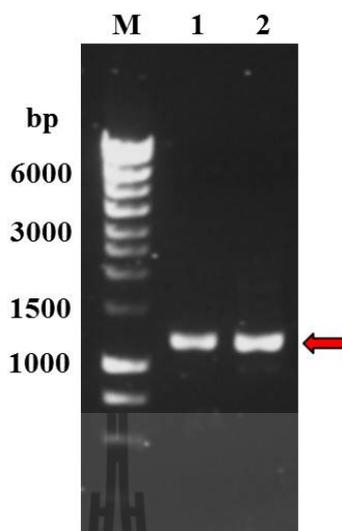
Figure 3.1.1 Identification of the amplified PCR product of the *CHI3L2* gene.

The PCR products of *CHI3L2* gene were analyzed on 1% (w/v) agarose containing syber green. Lane M: 1 Kb DNA marker; lanes 1 and 2: ~1,100 bp of the expected PCR products (arrow).

The PCR product of the expected size (~1.1 kb) was purified and directly cloned into the pET32a(+) expression vector, generating the recombinant plasmid, so called pET32a(+)/*CHI3L2*. The fusion protein expressed from the designed construct (Figure 3.1.2) consisted of five components of functional importance to this study: a thioredoxin fusion tag, a sixhistidine tag (His$_6$), S-tag, an enterokinase cleavage site, and a YKL-39 polypeptide. The Trx tag was fused to the protein to enhance protein solubility, while the His$_6$ tag was induced to aid protein purification. Enterokinase cleavage site is used for fusion tag removal that contains a five amino acid sequence, DDDDK.
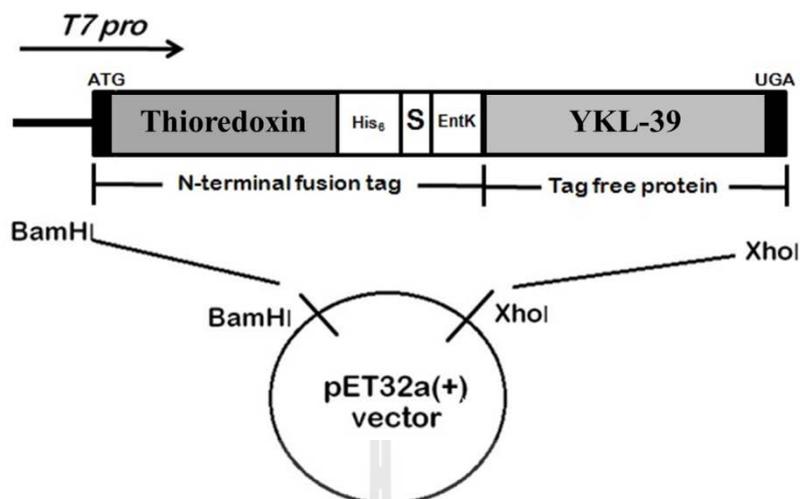
Figure 3.1.2 The gene organization of the pET32a(+)/*CHI3L2* construct.

The recombinant plasmid pET32a(+)/*CHI3L2* clones were examined by colony PCR, restriction enzyme digestion, and DNA sequencing to ensure the success of this construction. The recombinant plasmids were isolated from the positive *E. coli* colonies, which were grown on LB/Amp medium, and were subsequently determined for the *CHI3L2* gene insertion by the colony PCR technique with specific oligonucleotide primers for the *CHI3L2* gene. Agarose gel revealed the PCR product of expected size (1.1 kbp) from clones 1 and 2 (Figure 3.1.3A). The restriction endonuclease digestion assay was also performed to confirm the presence of the *CHI3L2* gene in the pET32a(+)/*CHI3L2* construct. After digestion the isolated plasmid obtained from positive colonies with *BamH*I and *Xho*I, two DNA bands were seen on agarose gel to the smaller size DNA fragment of about 1.1 kbp corresponded to the size of the *CHI3L2* gene obtained from PCR amplification, whereas the bigger DNA band of about 6.0 kbp was corresponded to the pET32a(+) expression vector

(Figure 3.1.3B). Therefore, both colony PCR and restriction enzyme digestion suggested that the pET32a(+)/*CHI3L2* constructed was generated successfully.

(A)   **(B)**
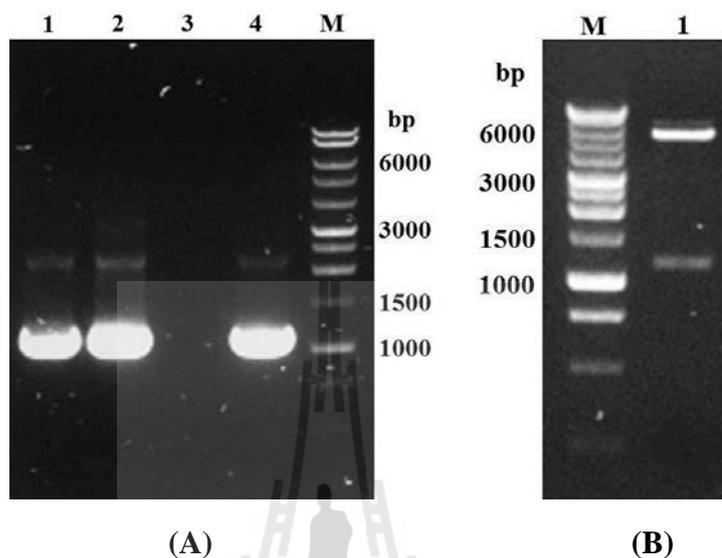
Figure 3.1.3 Confirmation of recombinant plasmid pET32a(+)/*CHI3L2* by colony PCR and restriction enzyme digestion.

(A) Colony PCR, The PCR products of the selected *E. coli* colonies were analyzed on 1% agarose gel electrophoresis. Lane M: 1 kb DNA marker; lanes 1-2: PCR products obtained from clones 1 and 2; lane 3: negative control; lane 4: positive control. (B) Restriction enzyme digestion: the restriction fragment of the pET32a(+) expression vector and the *CHI3L2* gene obtained from restriction endonuclease digestion with *BamH*I and *Xho*I. Lane M: 1 kb DNA marker; lane 1: 1.1 kb band of the *CHI3L2* gene and 6.0 kbp band of pET32a(+) vector after the recombinant plasmid pET32a(+)/*CHI3L2* was digested with *BamH*I and *Xho*I.

Furthermore, both sense and anti-sense strands of the *CHI3L2* gene were confirmed by automated DNA sequencing. The DNA sequencing results indicated

that the recombinant plasmid pET32a(+)/*CHI3L2* had an open reading frame (ORF)
that encoded the *N*-terminal thioredoxin, followed by the His$_6$ tag connected with the
N-terminus of the mature YKL-39. Overall, this recombinant plasmid
pET32a(+)/*CHI3L2* encoded 364 amino acids of the mature YKL-39 lacking a 26-aa
signal sequence. Figure 3.1.4 represents an open reading frame of the construct, which
expressed as a fusion protein that contains a thioredoxin fragment, followed by the
His$_6$ residues (purple color) attached at the *N*-terminus of the YKL-39 polypeptide,
via an enterokinase cleavage site (orange color) for removal of the *N*-terminus fusion
tag. In this construct, there are an extra seven amino acids A-M-A-D-I-G-S flanking
between the enterokinase recognition site and the first three residues: Y (Tyrosine), K
(Lysine) and L (Leucine) as shown in blue color of the YKL-39 polypeptide. The
deduced amino acid sequence of YKL-39 was found to be identical to the 390-aa
sequence of CHI3L2 isoform 1 (identifier: Q15782-4) reported in the
UniProtKB/Swiss-Prot database. The exception is that a single amino acid residue in
our sequence at position 318 had Trp instead of Arg as show in red color with arrow
(Figure 3.1.4). The discrepancy of this amino acid has been suggested previously to
be a natural variation that occurs due to a single nucleotide polymorphism (The MGC
Project Team, 2004).

	10	20	30	40	50	60
	MSDKIITKTD	DSFDTDVLKA	DXAILVDFWA	EWCGPCKMIA	PILDEIADEY	QGKLTVAKLN

	70	80	90	100	110	120
	IDQNPGTAPK	YGIRGIPTLL	LFKNGEVAAT	KVGALSKGQL	KEFLDANLAG	SGSGHMHHHH

	130	140	150	160	170	180
	HHSSGLVPRG	SGMKETAAAK	FERQHMDSPD	LGTDDDDKAM	ADIGSYKLVC	YFTNWSQDRQ

	190	200	210	220	230	240
	EPGKFTPENI	DPFLCSHLIY	SFASIENNKV	IIKDKSEVML	YQTINSLKTK	NPKLKILLSI

	250	260	270	280	290	300
	GGYLFGSKGF	HPMVDSSTSR	LEFINSIILF	LRNHNFDGLD	VSWIYPDQKE	NTHFTVLIHE

	310	320	330	340	350	360
	LAEAFQKDFT	KSTKERLLLT	VGVSAGRQMI	DNSYQVEKLA	KDLDFINLLS	FDFHGSWEKP

	370	380	390	400	410	420
	LITGHNSPLS	KGWQDRGPSS	YYNVEYAVGY	WIHKGMPSEK	VVMGIPTYGH	SFTLASAETT

	430	440	450	460	470	480
	VGAPASGPGA	AGPITESSGF	LAYYEICQFL	KGAKITWLQD	QQVPYAVKGN	QWVGYDDVKS

	490	500	510	520	529	
	METKVQFLKN	LNLGGAMIWS	IDMDDFTGKS	CNQGPYPLVQ	AVKRSLGSL	

Figure 3.1.4 The deduced amino acid sequences of the constructed recombinant plasmid pET32a(+)/*CHI3L2*.

The constructed recombinant plasmid contained the thioredoxin fusion tag, followed by the His$_6$ (purple) connected with the enterokinase cleavage site (orange), attached at the *N*-terminus of the mature YKL-39 polypeptide. The single amino acid residue exception to the sequence in the UniProtKB/Swiss-Prot database (identifier: Q15782-4), Arg318 substituted with Trp, is shown in red with an arrow.

The amino acid sequence of human YKL-39 was highly homologous to many of the recently identified members of the human chitinase and chitinase-like protein family. Sequence alignment of the human YKL-39 with these proteins, including YKL-40 (CHI3L1, P36222), acidic mammalian chitinase (AMCase, Q9BZP6-1), human chitotriosidase (CHIT1, Q13231), human oviduction (oviduction, Q86YN0),

and stabilin-1 isoform 1(SI-CLP, Q9NY15) by CLUSTAL W is shown in Figure. 3.1.5. The YKL-39 sequence is closest to CHIT1 (55.2% identity), followed by YKL-40 (51.3%). On the other hand, YKL-39 had lowest identity to SI-CLP (17.5%). The amino acid residues S143 and I145 located at the ends of the DxxDxSxE motif (Figure 3.1.5, black asterisks) are reported to suppress chitinase activity. In true chitinases like AMCase and CHIT1, these two residues are found as Asp and Glu instead. Double mutations of S143 to D143 and I145 to E145 in YKL-39 created pseudo-chitinase activity for this protein (Schimpl *et al*., 2012).
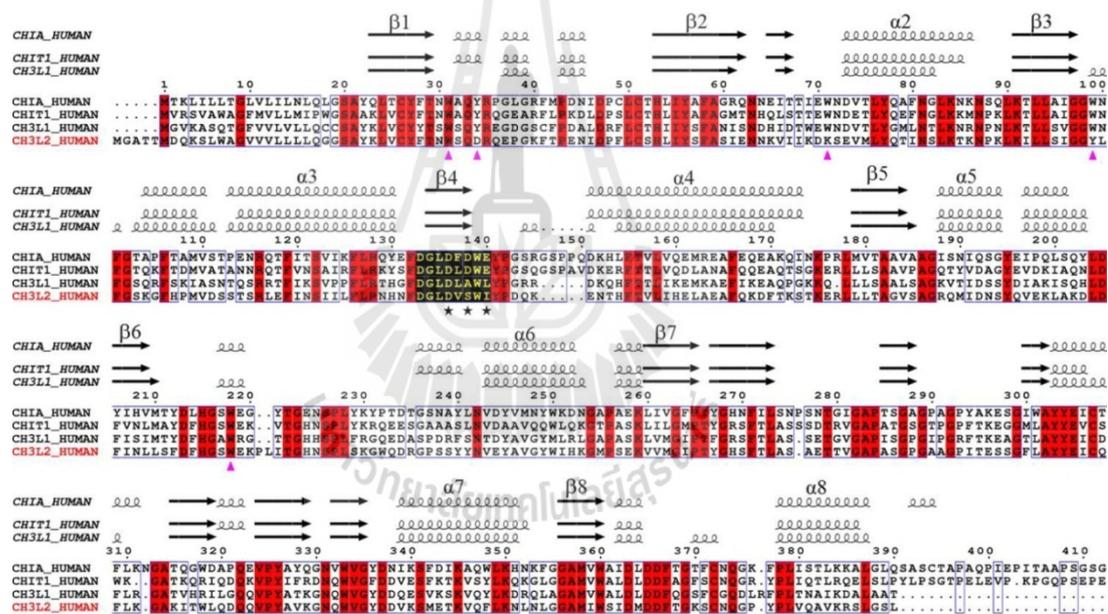
Figure 3.1.5 Structure-based alignments of human YKL-39 and its homologues.

The sequence alignment was generated by CLUSTAL W. The YKL-39 sequence (SwissProt: Q15782) was aligned with the amino acid sequence of AMCase (Swiss-Prot: Q9BZP6), chitotriosidase (Swiss-Prot: Q13231), and (SwissProt: P36222). The secondary structural components of AMCase (pdb 3FXY), Chitotriosidase (pdb 1GUV), and YKL-40 (pdb 1HJW) were identified, edited, and displayed as a

3.1.6B revealed that the majority of the soluble protein was found primarily in the supernatant after centrifugation to remove insoluble material. SDS-PAGE analysis of the induced cells in the soluble and insoluble fractions exhibited the expected protein band of Trx-His$_6$/YKL-39 at 55 kDa, however, the protein band was mainly seen in the soluble fraction (lane 3), as compared to the protein band in the insoluble fraction (lane 2) and the control cells (lane 1). The results showed that YKL-39 was expressed in *E. coli* in soluble form.
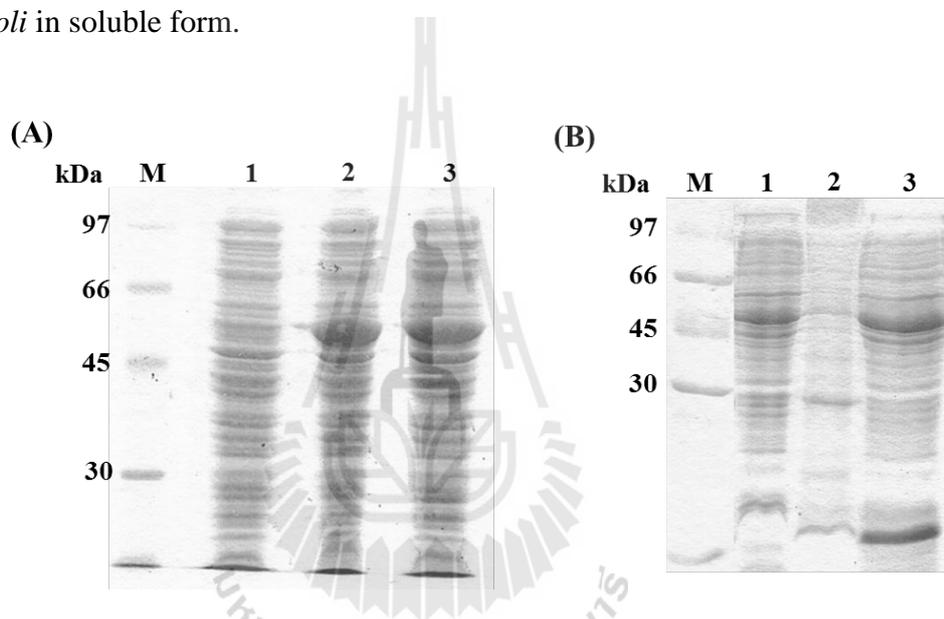
Figure 3.1.6 SDS-PAGE of the expression of the Trx-His$_6$/YKL-39 fusion protein. (A) Protein expression analysis: Lane M: protein molecular weight marker; lane 1: whole cell lysate of un-induced BL21(DE3) host containing pET32a(+)/YKL-39; lanes 2-3: whole cell lysate of induced BL21(DE3) host containing pET32a(+)/YKL-39 with IPTG. (B) Solubility determination. Lane M: protein molecular weight marker; lane 1: total proteins of induced BL21(DE3) host containing pET32a(+)/YKL-39 with IPTG; lane 2: pellet of cell lysate after induction with 0.5 mM IPTG for 16 h; lane 3: supernatant of cell lysate after induction.

3.1.2.2   Optimization of recombinant YKL-39 expression from *E. coli* BL21(DE3) cells

In this study, expression trials were carried out in order to obtain optimal conditions that gave highest yield, with highest solubility of human YKL-39. Different factors, including IPTG concentration, temperature, and time of induction, that generally affect yield and solubility of the recombinant expression were investigated.   Initially, IPTG concentration was analyzed. SDS-PAGE analysis showed that the concentration of 0.5 mM IPTG yielded the highest level of soluble protein. Further optimization was carried out at a fixed concentration of 0.5 mM IPTG for 16 h with various temperatures (15°C, 25°C, 30°C and 37°C). SDS-PAGE analysis of total proteins obtained in the cell pellet and crude supernatant revealed high expression level of the soluble YKL-39 was observed when induced at 15°C and 25°C. Thus, all the following experiments were performed under the induction temperature at 25°C. For further optimization, the expression of Trx-His$_6$/YKL-39 in *E. coli* BL21(DE3) was examined at different times of induction (0, 4, 6, 9, 12, and 16 h). As shown in Figure 3.1.7, SDS-PAGE gel analysis revealed no protein band of 55 kDa at 0 h of induction. The soluble form of Trx-His$_6$/YKL-39 fusion protein was clearly observed even at 4 h. However, protein expression reached highest level after 6 h of IPTG induction and maintained a similar level up to 16 h. Thus, all the following experiments were performed at the induction time of over 6 h. For the convenience of experimental handling, an overnight induction was preferred for YKL-39 expression. In conclusion, the expression level of the soluble YKL-39 was dependent on IPTG concentration, temperature and time of induction, with the optimal conditions set as 0.5 mM IPTG concentration at 25°C for 16 h. This condition

was further utilized for high expression level of YKL-39 functional and structural characterization.
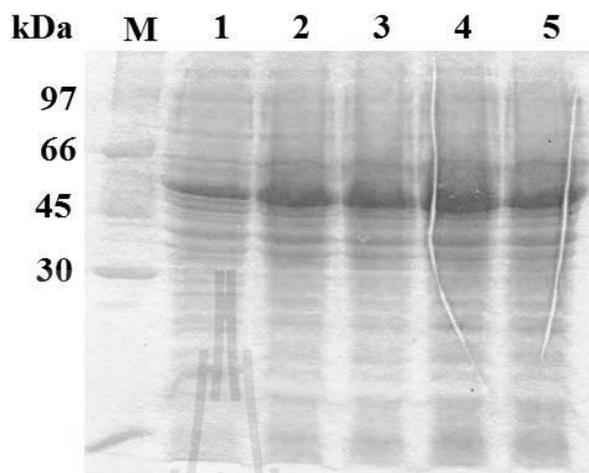
Figure 3.1.7 SDS-PAGE of Trx-His$_6$/YKL-39 fusion protein at different induction times with 0.5 mM IPTG, and 25°C.

Lane M: protein molecular weight marker; lanes 1-5: supernatant of whole cell lysate of induced BL21(DE3) host containing pET32a(+)/*CHI3L2* with IPTG for 4, 6, 9, 12, and 16 h, respectively.

3.1.2.3 Protein purification for polyclonal and monoclonal antibody production

Recombinant Trx-His$_6$/YKL-39 was expressed as an *N*-terminal thioredoxin and His$_6$ tag fusion protein in the *E. coli* BL21(DE3) under the optimized condition as described in Section 3.1.2.2. The presence of the His$_6$ tag facilitated the protein purification by Ni-NTA resin affinity chromatography. The Trx-His$_6$/YKL-39 fusion protein contained the cleavage site for entorokinase at the *N*-terminus of the mature YKL-39 protein. Initially, the recombinant fusion protein containing His$_6$ tag

was purified by gravity Ni-NTA resin affinity column. To minimize nonspecific binding and reduce the amount of contaminating proteins, a two-step wash (5 mM, followed by 20 mM imidazole) was carried out before the column was eluted with 250 mM imidazole. SDS-PAGE analysis showed the majority of the 55 kDa protein band, corresponding to the Trx-His$_6$/YKL-39 fusion protein in the Elute I and Elute II fractions, as showed in Figure 3.1.8. However, several contamination proteins were also observed in the elution I fraction. Approximately 80-90% purity of the YKL-39 protein was obtained by first Ni-NTA resin affinity chromatography atep, with the overall yield of about 6.8 mg protein per liter of *E. coli* culture. After elution, the buffer of the eluted sample was immediately changed to 50 mM Tris-HCl, pH 8.0, containing 50 mM NaCl to remove imidazole.

Figure 3.1.8 SDS-PAGE of the Trx-His$_6$/YKL-39 fusion protein purified by gravity Ni-NTA resin affinity chromatography.

Lane M: molecular weight protein marker; lane 1: flow-through fraction; lanes 2-3: wash fraction with 5 mM imidazole; lane 4: wash fraction with 20 mM imidazole; lanes 5-6: eluted fractions with 250 mM imidazole.

For further purification, Sepharose Q HP anion exchange chromatography was applied to eliminate the contaminating proteins, so as to obtain highly purified fusion protein. The protein purified from gravity Ni-NTA resin was filtered, and then applied onto a Sepharose Q HP column. SDS-PAGE analysis displayed the majority of the fusion protein band in fractions 39-41 (Figure 3.1.9). However, several low molecular weight contaminated proteins were still observed. These eluted fractions were then pooled and concentrated. After this purification step, approximately 5.1 mg purified protein was obtained from 1 liter of *E. coli* culture.
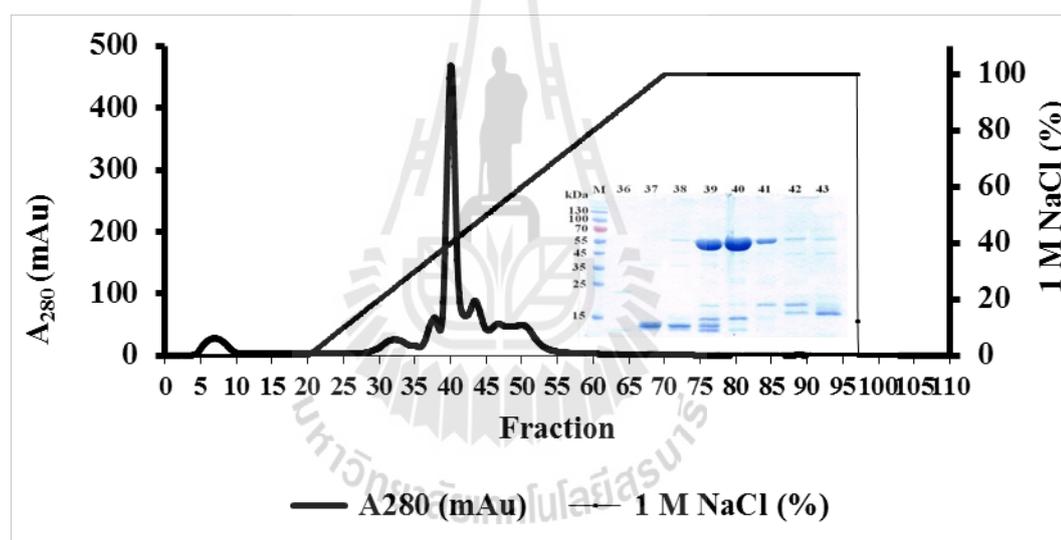
Figure 3.1.9 An elution profile of the fusion Trx-His$_6$/YKL-39 purified from an ÄKTA purifier system with a Sepharose Q HP anion exchange chromatography.

The *E. coli* expressed YKL-39 fusion protein was applied onto a Ni-NTA resin column, followed by a Sepharose Q HP pre-packed column. The bound protein was eluted with a linear gradient of 1 M NaCl in 10 mM Tris-HCl, pH 8.0. The inset displays SDS-PAGE analysis of the eluted fractions containing the Trx-His$_6$/YKL-39 fusion protein in fractions 39-41.

To remove the Trx-His$_6$ fusion tag, the purified protein obtained from Sepharose Q HP anion exchange chromatography was cleaved with enterokinase to generate the fusion tag free YKL-39. After digestion overnight, the Trx-His$_6$ fusion tag was completely removed from the protein solution by the second gravity Ni-NTA resin step. The purification profile of the free fusion-tag YKL-39 protein obtained from the second Ni-NTA resin was evaluated by SDS-PAGE, as shown in Figure 3.1.10.
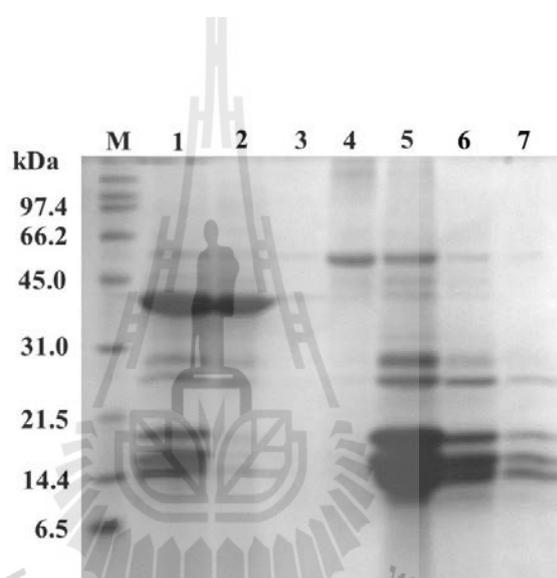
Figure 3.1.10 SDS-PAGE of the free-tag YKL-39 protein purified by second Ni-NTA agarose.

Lane M: molecular weight protein marker; lane 1: digested fusion Trx-His$_6$/YKL-39 solution before purification; lane 2: flow-through fraction contained the tag-free YKL-39 protein; lanes 3: fraction of wash with 5 mM imidazole; lane 4: 20 mM imidazole wash fraction; lanes 5-7: fractions of uncleav Trx-His$_6$/YKL-39 and Trx-His$_6$ tag fragment eluted with 250 mM imidazole.

The flow-through fractions containing nonfusion YKL-39 protein were detected at approximately 39 kDa, which corresponds to the theoretically expected molecular weight of YKL-39. The uncleav Trx-His$_6$/YKL-39 and Trx-His$_6$ tag fragment were observed in the Wash II and elute fractions as expected. After the second Ni-NTA resin step exchange of the buffer and concentration, the purity of YKL-39 was evaluated to be >95%, with a yield of about 3.4 mg per liter of culture (Table 3.1). The protein from the second Ni-NTA resin was used for polyclonal and monoclonal antibody production. Improvement of the purity of YKL-39 after purification is displayed in Figure 3.1.11.
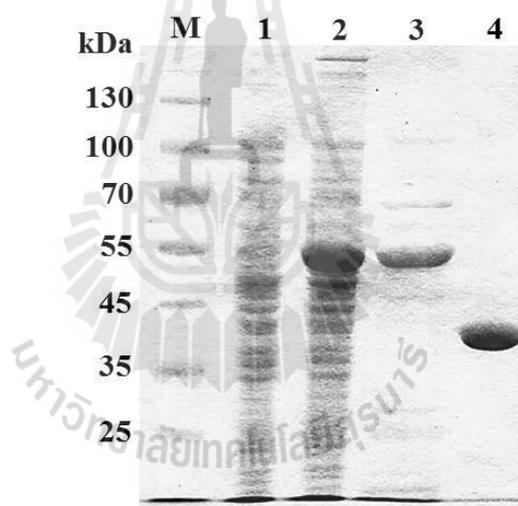
Figure 3.1.11 SDS-PAGE analysis of the purified YKL-39 after each purification step.

Lane M: protein molecular weight marker; lane 1: cell lysate of un-induced BL21(DE3) host containing pET32a(+)/*CHI3L2*; lane 2: cell lysate of induced with IPTG; lane 3: elution of Trx-His$_6$/YKL-39 from Ni-NTA affinity column and anion exchange chromatography, lane 4: the purified YKL-39 after thioredoxin removal and further purification by the second Ni-NTA affinity column.

Table 3.1.1 Complete purification of human YKL-39.

Purification step	Total volume (mL)	Total protein (mg)	Yield (%)	Protein form
Crude extract	30	105	100	Trx-His$_6$/YKL-39
First Ni-NTA affinity column	15	6.8	6.5	Trx-His$_6$/YKL-39
Anion exchange column	5	5.1	4.8	Trx-His$_6$/YKL-39
Second Ni-NTA affinity column	5	3.4	3.3	YKL-39

Table 3.1.1 shows that the purification of the recombinant Trx-His$_6$/YKL-39 yielded approximately 80-90% purified YKL-39 after first Ni-NTA resin affinity chromatography step. After enterokinase cleavage, followed by second Ni-NTA affinity chromatography, 3.4 mg of the purified YKL-39 was obtained from one liter of bacterial culture.

3.1.2.4 Protein purification for crystallization

To obtain high quality and quantity of YKL-39, the purification protocol was further developed using His Trap HP (1-mL). As shown in the elution profile in Figure 3.1.12 present the elution profile of the protein purification by His Trap HP, which showed that most contaminating proteins were washed out in the unbound and wash fractions. Bound proteins were found in fractions eluted with 5 to 250 mM imidazole. This result indicated that fusion protein was not strongly bound to

the Ni^{2+} resin. After removal of the Trx-His$_6$ tag by enterokinase cleavage overnight and further purification by a second His trap column, SDS-PAGE showed the 39 kDa free-tag YKL-39 in flow-through fractions, while noncleav Trx-His$_6$/YKL-39 and Trx-His$_6$ tag fragments were observed in the eluent (Figure 3.1.13). As shown in Figure 3.1.14, SDS-PAGE analysis of eluted fractions showed a single band of the 39 kDa protein. The purity of homogeneous protein of >95% was obtained after final purification step. Therefore, it was used for crystallization and protein-ligand binding studies.
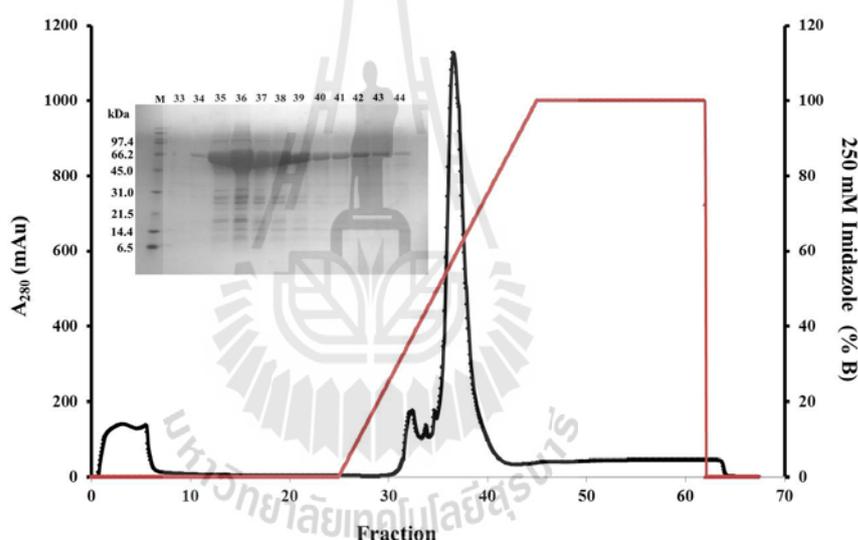
Figure 3.1.12 An elution profile of the recombinant Trx-His$_6$/YKL-39 recombinant obtained from an ÄKTA purifier system with a His trap column.

Crude supernatant was applied onto a His trap (1-mL) column. Bound proteins were eluted with a linear gradient of 0-250 mM imidazole in 50 mM Tris-HCl, pH 8.0. Inset displays SDS-PAGE gel of the eluted fractions containing the Trx-His$_6$/YKL-39 fusion protein in fractions 35-44.
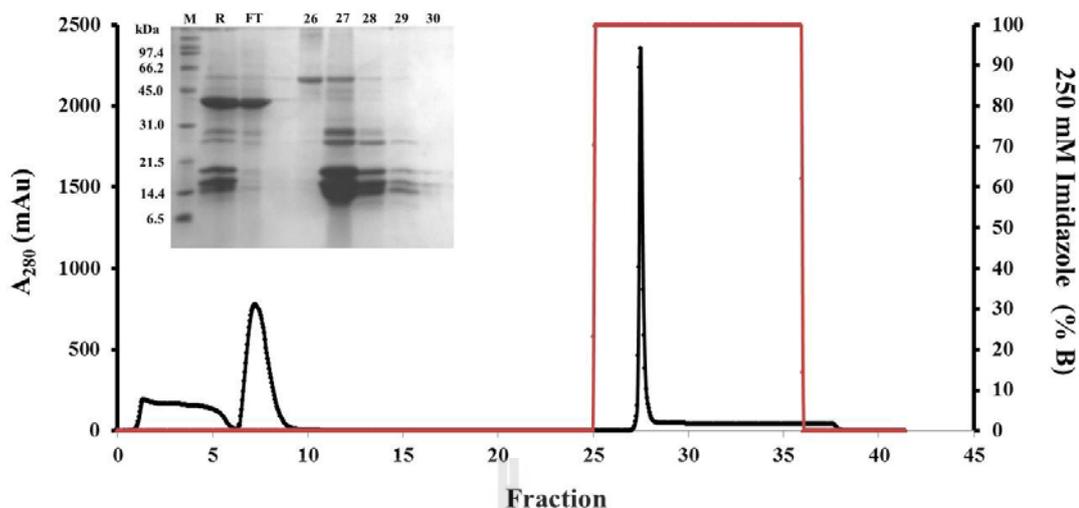
Figure 3.1.13 Elution profile of the HisTrap HP column of Trx-His$_6$/YKL-39 fusion protein after cleavage with enterokinase.

Inset represents SDS-PAGE gel of the eluted fractions. Lane M: low molecular weight protein marker; lane R: the overnight digestion of fusion protein Trx-His$_6$/YKL-39 by enterokinase; lane FT: the free-tag YKL-39 eluted in the flow-through; lanes 26-30: the un-cut fusion protein and *N*-terminus Trx-His$_6$ tag fractions eluted with 250 mM imidazole.
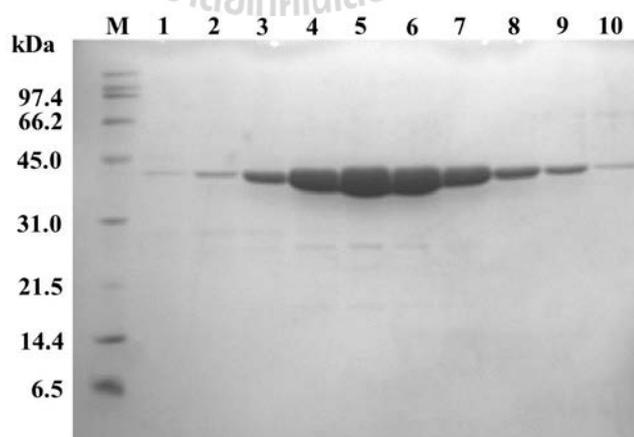
Figure 3.1.14 SDS-PAGE analysis of the purified YKL-39 obtained after purification with a Hiload 16/60 Superdex 200 prep grade gel filtration column.

The protein sample obtained from the second Ni-NTA affinity step was applied onto a Hiload 16/60 Superdex 200 prep grade gel filtration column under the conditions given in Section 2.5.3.2. Lane M: molecular weight protein maker; lanes 1-10: eluted fractions with 50 mM NaCl and 50 mM Tris buffer, pH 8.0.

3.2 Production of polyclonal and monoclonal antibodies

YKL-39 is recently known as an osteoarthritis marker, since expression of YKL-39 was highly up-regulated at both mRNA and protein levels in cartilage from patients with osteoarthritis (Steak *et al*., 2002 and Knorr *et al*., 2003). Therefore, the availability of in-house polyclonal and monoclonal antibodies will open up an opportunity for the development of immunological tools (i.e., specific YKL-39 ELISA kit and YKL-39 immunosensor) for a sensitive detection of the disease progression. In this study, we describe the production of anti-YKL-39 polyclonal and monoclonal antibodies for investigation of protein-protein or protein-ligand interactions, as well as for development of a sensitive enzyme-linked immunoassay that requires specific, well characterized, and uniform antibodies against YKL-39.

3.2.1 Polyclonal antibody production

3.2.1.1 Determination of immune response to produce anti-YKL39 antibodies

The purified YKL-39 protein was used as immunogen for production of anti-YKL-39 polyclonal antibodies (pAbs). After the final immunization, an average of 5-10 mL of anti-serum was obtained after each bleeding. To determine the antibody response in the immunized rabbit, antibody titers were

determined by indirect ELISA assay. Figure 3.2.1A, shows the Ag:Ab binding curve, presenting concentration dependence of anti-YKL39 antisera towards the recombinant YKL-39. The hyperbolic feature, mimicking the Michaelis-Menten plot, between the interaction signals versus YKL-39 concentrations indicated that YKL-39 protein was able to stimulate a specific immune response in rabbits to produce high-specific pAbs. The binding curve demonstrated saturated and specific binding of the anti-YKL39 antiserum to the YKL-39 antigen (10 µg/mL) within a range of 1:10-1:327,680 dilutions of the antisera used. Pre-immune serum gave no reactivity to YKL-39 at the given ratios of Ag:Ab concentrations. Secondary antibody treatment alone (1:2,000 dilution) also demonstrated no reactivity (data not shown). These results showed that the anti-YKL39 antisera were specifically reactive to the purified YKL-39 protein. The polyclonal antibodies reacted strongly with the YKL-39 at dilution up to 1:40,000. In addition, the signals, representing specific interactions was seen with the linear range of detection from 1:2,560 to 1:20,480 of anti-serum dilutions ($R^2 =$ 0.9934), as shown in Figure 3.2.1B.
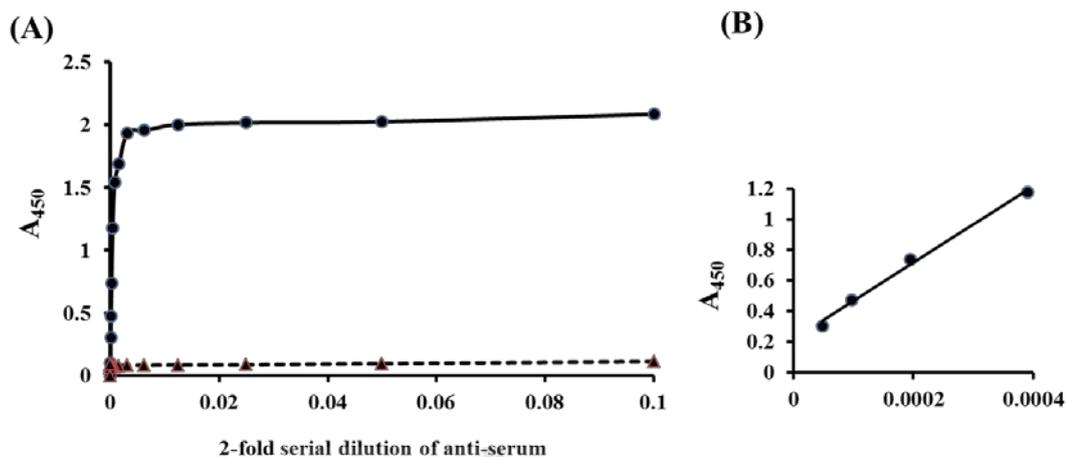
Figure 3.2.1 Antibody titers after final immunization were determined by Indirect ELISA assay.

(A) The binding curve of Ag:Ab. 96 well plates were coated with 10 µg/mL YKL-39 and 2-fold serial dilutions of the antisera from 1:10 to 1:327,680 were added in to the wells in triplicate (closed-circle), while the pre-immune rabbit serum was coated as a negative control (closed-triangle). A HRP-conjugated goat anti-rabbit antibody was used to detect protein-antibody complex and the absorbance measured at 450 nm with TMB as substrate. The binding curve representing antigen-antibody interactions were obtained from a nonlinear regression fit available in Prism v5.0 (GraphPad Software) using the following single-site binding model. (B) The linear range of detection was found to be from 1:2,560 to 1:20,480 of anti-serum dilutions.

3.2.1.2 Affinity purification of anti-YKL39 polyclonal antibodies

The antisera were further purified by affinity chromatography on a protein A agarose column (GenScript Corporation, USA). Immunized serum was diluted with binding buffer (0.15 M NaCl, 20 mM Na_2HPO_4, pH 8.0) at a ratio of 1:1 to ensure that proper ionic strength and pH are maintained for optimal binding, as

described in Section 2.7.2. Bound antibodies were eluted with 10 CV of elution buffer (0.1 M glycine, pH 2.5), and were immediately neutralized to pH 7.4 with neutralization buffer. Figure 3.2.2 is SDS-PAGE analysis of the anti-YKL39 antisera purification. Panel A represents the elution profile of each purification step. To evaluate the quality of the purified antibodies, the eluted fractions were further analyzed by SDS-PAGE. Under non-reducing conditions, the immunoglobulin migrated behind the 97 kDa marker (panel B, lanes 1-2). Under reducing conditions, two protein bands from the eluted fractions were observed, one of which had a molecular weight of 50 kDa and the other one a molecular weight of 25 kDa that belonged to heavy and light chains, respectively (panel B, lanes 3-4). The antibodies bound with protein A column, indicating that it was a protein A-specific immunoglobulin G (IgG) isotype.
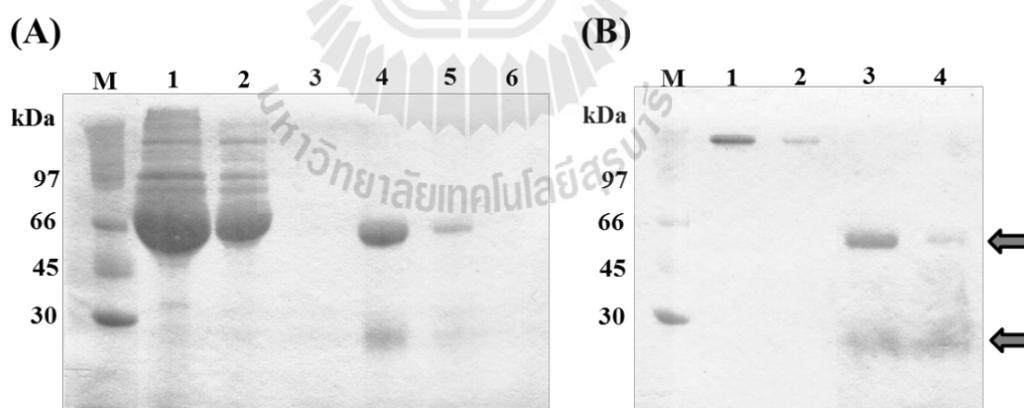
Figure 3.2.2 Purification of anti-YKL39 polyclonal antibodies by protein A-agarose chromatography.

(A) SDS-PAGE of purified anti-YKL39 under reducing conditions: Lane M, molecular weight protein marker. Lane 1, crude antiserum. Lane 2, flow-through fraction. Lane 3, wash fraction with binding buffer. Lanes 4-5, eluted fractions with

elution buffer (0.1 M glycine, pH 2.5). (B) SDS-PAGE of eluted fractions. Lane M, molecular weight protein marker. Lanes 1-2, eluted fractions under non-reducing condition, while Lanes 3-4 eluted fractions under reducing condition.

3.2.1.3 Characterization of purified anti-YKL39 polyclonal antibodies

3.2.1.3.1 Specificity of purified anti-YKL39 polyclonal antibodies

Indirect ELISA and Western blot immunoassay were carried out to determine specificity and sensitivity to YKL-39. Titers of the purified antibodies were further determined by indirect ELISA. Our raised antibodies were found to be highly sensitive, reacting strongly with the YKL-39 protein even when only 5 ng of the purified anti-YKL39 pAbs was used. For further biochemical characterization, 10 ng of anti-YKL39 pAbs was selected as the optimum concentration. Figure 3.2.3A displays the Ag:Ab binding curves, showing the concentration dependence of anti-YKL39 antiserum towards the recombinant YKL-39 and other chitinase and CLP homologs. The hyperbolic feature between the interaction signals versus YKL-39 concentrations indicated specific interactions between the anti-YKL39 pAbs and antigen (Figure 3.2.3A, blue curve). When YKL-40 (purple), AMCase (pink), and bacterial chitinase A (orange) were used as immunogens in range of 0.24-500 ng, no significant increase in the signal was observed. In addition, the pre-immune serum, which was used as control did not react with protein even at the maximum amount of YKL-39 used (yellow line). The results indicate that the raised antibodies were highly specific to the YKL-39 antigen. The

binding intensity of each antigen was further tested at a fixed concentration of 500 ng/well and the anti-YKL39 antibodies of 10 ng/well (Figure 3.2.3A). The absorbance obtained for YKL-39 was 15–19 folds greater than that for human YKL-40, AMCase, and bacterial chitinase. This further confirmed a high specificity of the purified antibodies towards YKL-39. To ensure the specificity of purified pAbs to YKL-39, Western blot analysis was carried out with various GH-18 homologs. As shown in Figure 3.2.3C, the antibodies recognized both fused (lanes 2 and 3) and bare (lane 4) forms of human YKL-39, but did not recognize YKL-40 (lane 5), AMCase (lane 6), and bacterial chitinase (lane 7). The antibodies also did not cross-react with proteins in human serum (lane 8). The results suggested that the purified pAbs were highly specific for YKL-39.
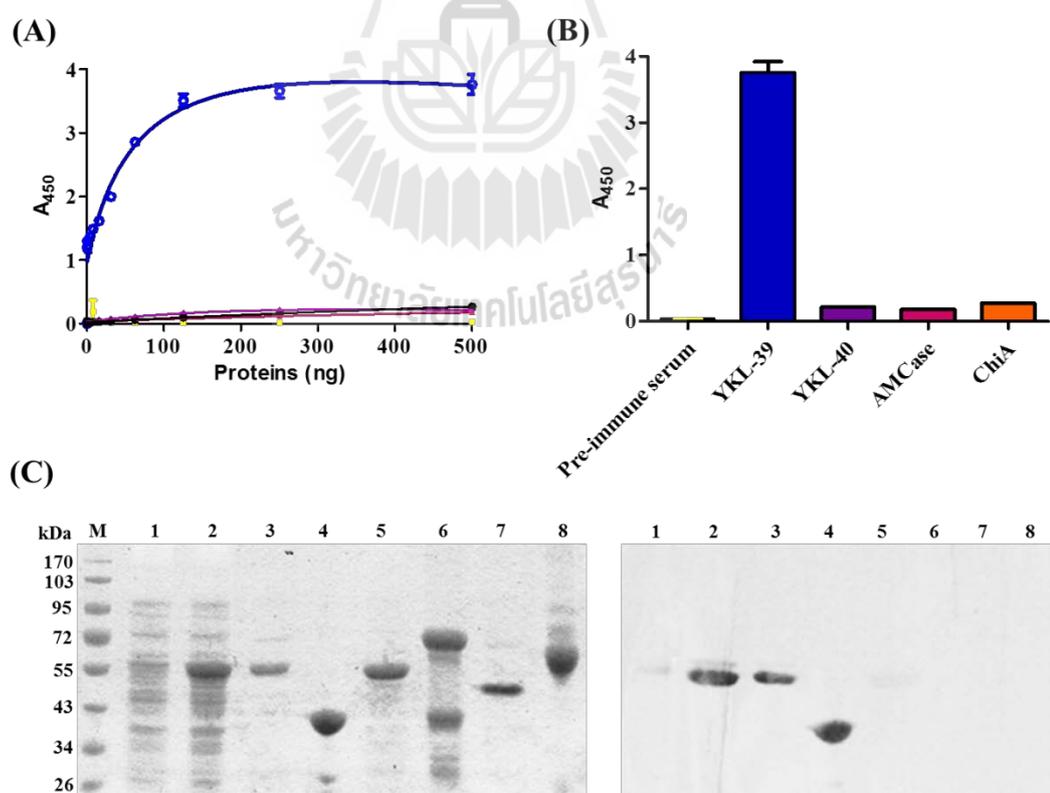
Figure 3.2.3 Specificity of purified anti-YKL39 antisera by indirect ELISA and Western blot analysis.

(A) Anti-YKL39 antiserum was raised against the purified YKL-39. Specificity of the purified anti-YKL39 antibodies (10 ng/well) was tested towards different concentrations (0–500 ng/well) of the CLPs by indirect ELISA assay. Blue, YKL-39; purple, YKL-40; pink, AMCase; orange, bacterial chitinase. Pre-immune serum against YKL-39 was used as control (yellow line). (B) Bar chart showing immunological interaction between a fixed concentration (500 ng/well) of YKL-39 homologs and purified YKL-39 antisera (10 ng/well). Values are averaged from the data of triplicate trials. (C) Cross-reactivity of anti-YKL39 pAbs with other human chitinases and CLPs proteins was analyzed by SDS-PAGE (left panel) in comparison with Western blot analysis (right panel). Lane M, protein molecular weight marker; lane 1, cell lysate of uninduced *E. coli* BL21(DE3) host containing pET32a(+)/*CHI3L2*; lane 2, cell lysate of cells induced with IPTG; lane 3, elution of Trx-His$_6$/YKL-39 from first Ni-NTA affinity column; lane 4, purified YKL-39 after ion exchange chromatography, after removal of the Trx-His$_6$ tag, and further purified by second Ni-NTA affinity column; lane 5, YKL-40; lane 6, AMCase; lane 7, chitinaseA; lane 8, healthy human serum.

3.2.1.3.2 Sensitivity of purified anti-YKL39 polyclonal antibodies

The sensitivity of the raised antibodies was carried out by indirect ELISA and Western blot analysis as described in Section 2.7.3.2 and 2.7.3.1. Indirect ELISA assay was performed to determine the optimal coating concentration of YKL-39. Concentrations of YKL-39 antigen coated on each well of ELISA plate were varied from 0.05 to 10 µg/mL, followed by the addition of 10 ng of

the purified anti-YKL39 pAbs. Figure 3.2.3A is the binding curve, showing saturation of the purified anti-YKL39 to the YKL-39 antigen above 10 µg/mL (left panel). The binding affinity (K_a) of YKL-39 protein for the detection by anti-YKL39 antibody was estimated to be 2.4×10^7 M^{-1}. In addition, the signal of interaction with the linear range of detection was found to be from 0.16 to 1.25 µg/mL of YKL-39 with $R^2 =$ 0.9999, as shown in Figure 3.2.3A (right panel). The limit of detection, which was estimated as the concentration of YKL-39 giving 10% of the maximum A_{450}, for the assay was found to be 0.04 µg/mL of YKL-39. The optimal YKL-39 concentration, which gave half the maximum value of A_{450} or B_{max} of the binding curve, was determined to be 1.7 µg/mL. In addition, the titer of the optimal YKL-39 concentration determination was also carried out by Western blot analysis. Different concentrations of YKL-39 were used in a range of 0.08-10 µg while the anti-YKL39 pAbs were fixed at 10 ng. After the membrane was exposed to an X-ray film, it was observed that at 10 ng of anti-YKL39 antibody generated a strong signal at 10 µg, and became fainted when the amount of YKL-39 was decreased down to 0.31 µg of YKL-39 (Figure 3.2.3B). At amounts of YKL-39 antigen lower than 0.16 µg, no signal of Ag:Ab reaction was observed.
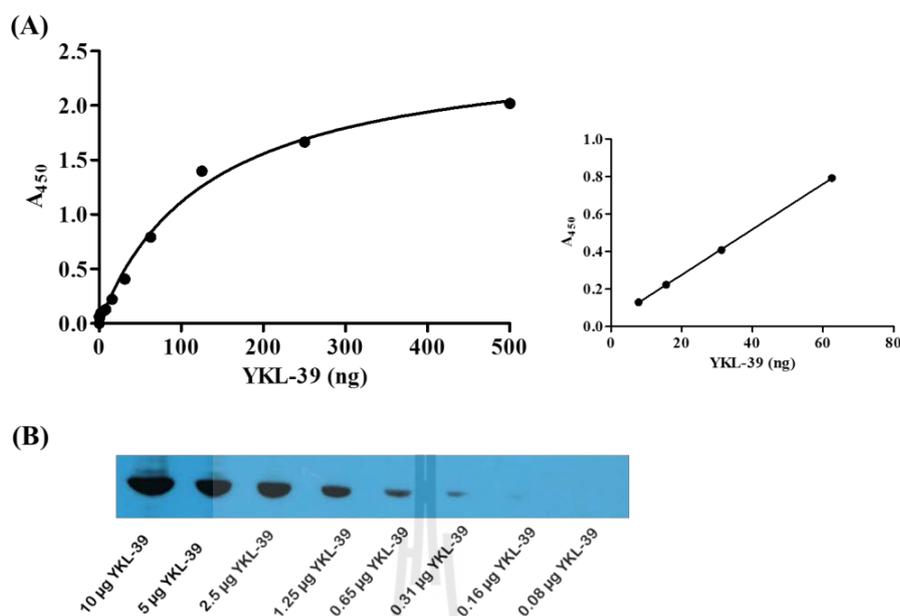
Figure 3.2.4 Sensitivity of the purified anti-YKL39 pAbs by indirect ELISA and Western blot analysis.

(A) Assessment of the binding affinity of the purified anti-YKL39 antibodies by indirect ELISA. The left panel is the binding curve of the Ag:Ab interaction. The purified anti-YKL39 antibodies (10 ng/well) were tested with different concentrations (0–10 $\mu g.mL^{-1}$/well) of YKL-39. Binding affinity was analyzed by nonlinear regression of data to a specific one-site binding function (Prism5). The right panel shows a linear response range at 0.16-1.25 µg/mL of YKL-39. (B) Sensitivity of the purified anti-YKL39 pAbs with YKL-39 protein was analyzed by Western blot analysis. YKL-39 amount was varied from 0.08-10 µg was used and the purified pAbs were fixed at 10 ng.

The optimal concentrations of anti-YKL39 antibodies used for indirect ELISA and Western blot assays were further investigated. Firstly, an indirect ELISA was

carried out. Different concentrations of anti-YKL39 antibodies in a range of 0.05 to 10 µg/mL were tested against 2.5 µg/mL of YKL-39. Signals representing antibody-protein interactions were detected with the HRP-conjugated IgG by the enhanced chemiluminescence method. The interaction of Ag:Ab gave the binding curve, which demonstrated saturation of the purified anti-YKL39 at above 2.5 µg/mL of the YKL-39 antigen as shown in Figure 3.2.3.4A (left panel). The binding affinity (K_a) of anti-YKL39 antibody for YKL-39 protein was estimated to be $1.3 \times 10^7 \, M^{-1}$. The right panel represents the linear range of detection, which was found to be from 0.4 to 1.25 µg/mL of anti-YKL39 ($R^2 = 0.978$). The limit of detection, which was calculated as the concentration of anti-YKL39 giving a 10% of the maximum A_{450}, was found to be 0.8 µg/mL. Moreover, the titer of the optimal anti-YKL39 concentration determination was also carried out by Western blot analysis. Two-fold serial dilutions of anti-YKL39 ranging from 0.04 to 10 µg were used to estimate the immunoactivity toward 2.5 µg of YKL-39. After the membrane exposed to the X-ray film, it was observed that at 10 ng of anti-YKL39 antibody generates a strong signal while 0.31 µg of anti-YKL39 that was able to give the weakest of signal.
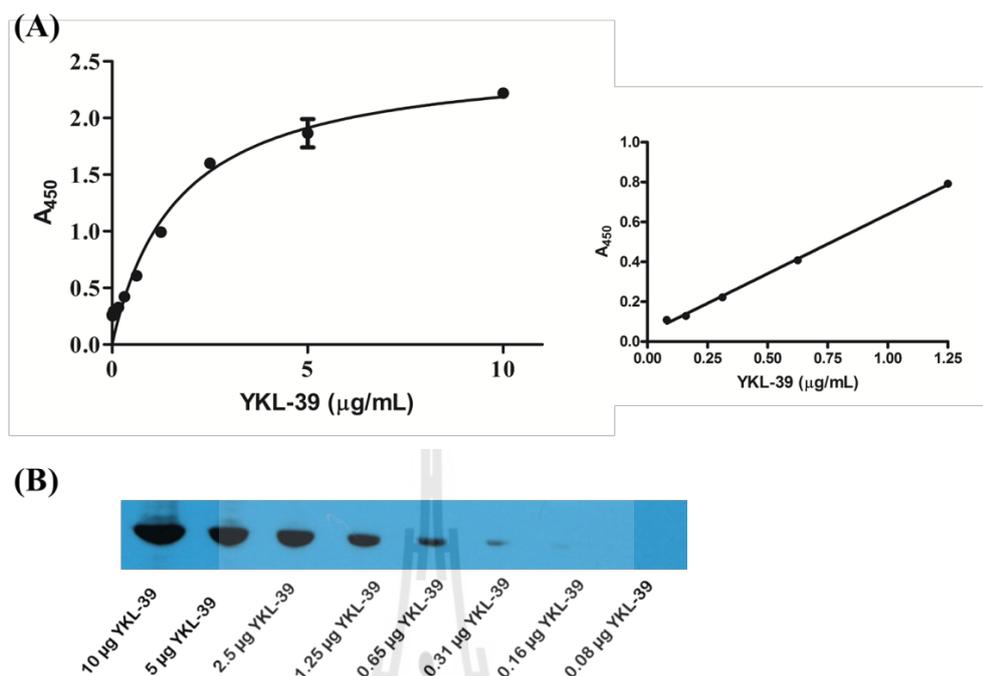
Figure 3.2.5 Titer of the anti-YKL39 by indirect ELISA and Western blot analysis.

(A) Assessment of the binding affinity of the purified anti-YKL39 antibody to YKL-39 protein by indirect ELISA. The left panel is binding curve of the Ag:Ab interaction. The purified anti-YKL39 antibody concentrations were varied from 0.5-100 µg/mL at fixed 2.5 µg/mL YKL-39. Binding affinity was analyzed by nonlinear regression, specific one-site binding (Prism5). The right panel show linear range at 0.4-12.5 µg/mL of anti-YKL39 antibody. (B) Sensitivity of purified anti-YKL39 pAbs for YKL-39 protein was analyzed by Western blot analysis. Anti-YKL39 antibody was used ranging from 0.005-10 µg and YKL-39 was fixed at 2.5 µg.

3.2.2   Monoclonal antibody production

3.2.2.1 Determination of immune response to YKL-39

To produce monoclonal antibody (mAb) specific for YKL-39, three female BALB/c mice were immunized intraperitoneally with 70 µg of the

purified YKL-39 protein in sterile PBS at two week intervals. Immunized mice sera were collected after 7 days of each injection as described in Section 2.8.1. The immune responses against YKL-39 antigen were determined by indirect ELISA using YKL-39. As shown in Figure 3.2.5, high titers of antibody response in the immunized mice serum were observed, whereas no signal was observed in the pre-immunized sample. Titer of raised antibody in immunized mice serum was high at dilutions of 1/390,625. The results indicated the success of induction of antibody production against YKL-39 protein.
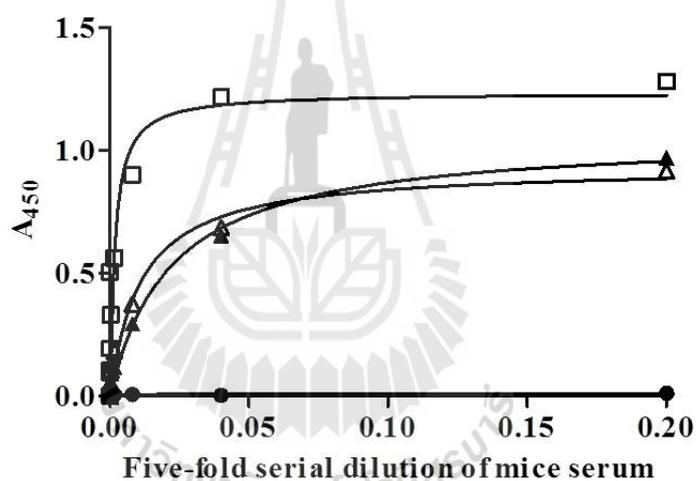
Figure 3.2.6 Determination of immune response against YKL-39 in immunized mice by indirect ELISA.

The binding curve of Ag:Ab, the 96 well plates were coated with a 10 µg/mL of YKL-39 and 5-fold serial dilutions of the immunized mice serum from 1/5 to 1/1,953,125 were performed in triplicate. The HRP-conjugated rabbit anti-mouse Igs antibody was used to detect the protein-antibody complex. Mouse 1, box; mouse 2, closed-triangle, mouse 3, open-triangle; pre-immunized mouse, closed-circle. Protein-

antibody binding curves were evaluated by a nonlinear regression function available in Prism version 5.0 fit to the single-site binding model.

3.2.2.2 Monoclonal antibody production by standard hybridoma technique

Spleen cells were further fused with P3X63-Ag8.653 myeloma cells by standard hybridoma technique using 50% (v/v) polyethylene glycol (PEG) as described in Section 2.8.3. After ten days of cultivation, cell growth was monitored under a microscope. Numbers and sizes of hybridoma containing wells were recorded. It was found that 477 wells contained healthy hybridomas, which was approximately 50% of the 960 wells observed. The presence of antibodies in positive wells was determined by indirect ELISA. The results showed 71 hybridomas containing wells (7.4%) were obtained that were strongly reacted with YKL-39. In order to obtain single hybridoma clones, ten positive wells that gave high responses to YKL-39 antigen were further chosen for dilution by limiting dilution as described in Section 2.8.4. After 7-10 days of cultivation, wells that contained single clones were marked and screened by indirect ELISA to determine the level of anti-YKL39 mAb production. From our data, five clones that gave highest signals were selected, and subsequently re-cloned by the 2nd round of limiting dilution until the stable single clone were obtained. As tested by indirect ELISA assay, two single clones gave high positive reactivity to YKL-39 protein were obtained. These clones were referred to as 8H3 and 6H11. The 8H3 monoclone was originated from the well of plate no. 8, row H, column 3 of the fusion step. For the 6H11clone, 94% of the wells from 192 wells in 1st limiting dilution showed strong reactivity with YKL-39 by indirect ELISA

screening. To confirm that these two clones were not the same, a 2nd limiting dilution was carried out. It was found that most positive wells still reacted strongly with YKL-39, indicating that 6H11 was originally derived from the same clone in the fusion experiment, prior to cloning by limiting dilution. Finally, they were further expanded, purified and characterized as described in the next section. Figure 3.2.6 represented the summarization of the anti-YKL39 mAb production.
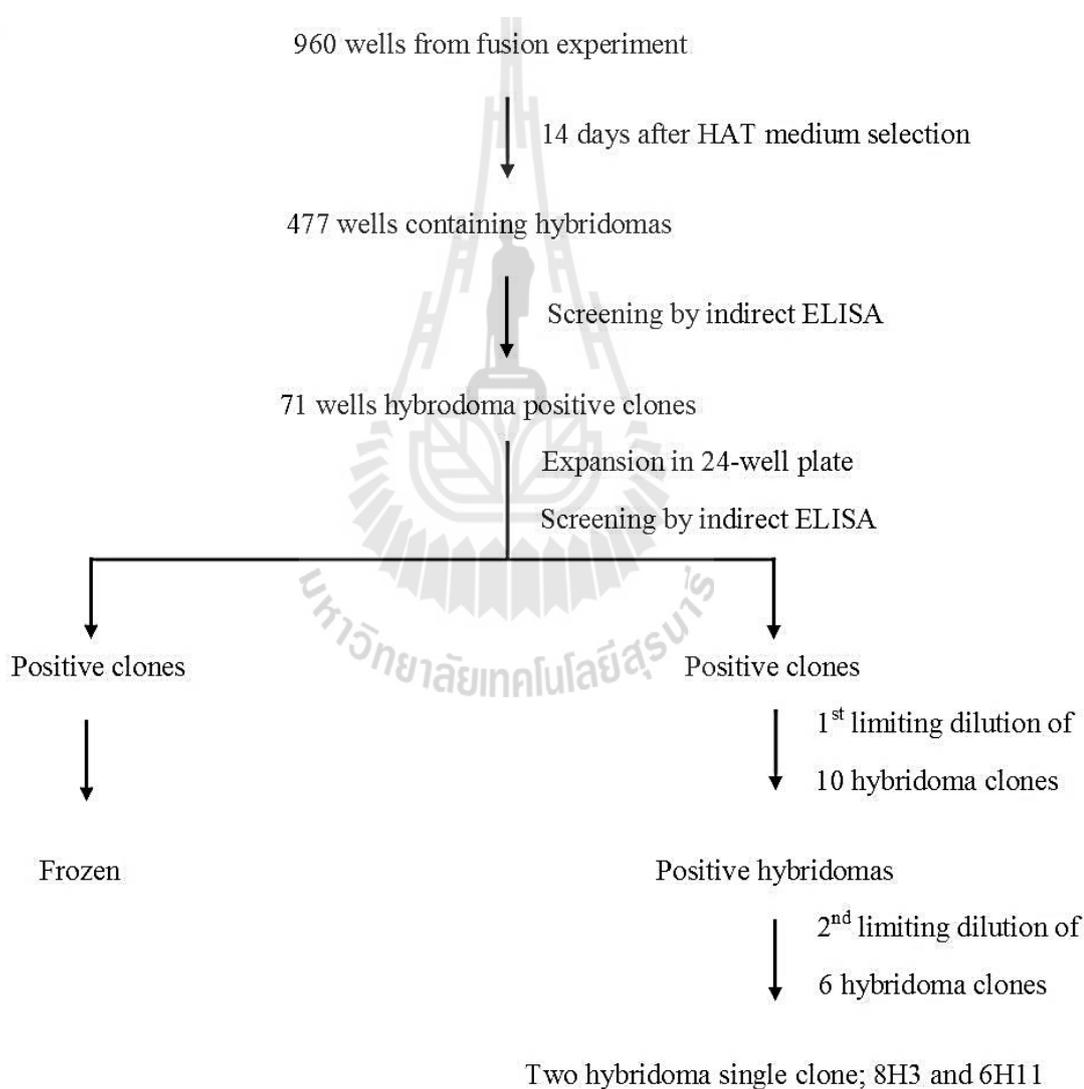
Figure 3.2.7 Summarization of the anti-YKL39 mAb production.

3.2.2.3 Determination of monoclonal antibody isotyping

The specific isotype of monoclones 6H11 and 8H3 was determined by captured ELISA technique using IsoQuick_Kits for Mouse Monoclonal Isotyping (Sigma-Aldrich) as described in Section 2.8.5. As shown in Figure 3.2.7, the isotype mapping identified two hybridoma clones, 6H11 and 8H3, to be IgM isotype. Because the 6H11 clone exhibited greater immunoactivity, this clone was selected for further purification and characterization.
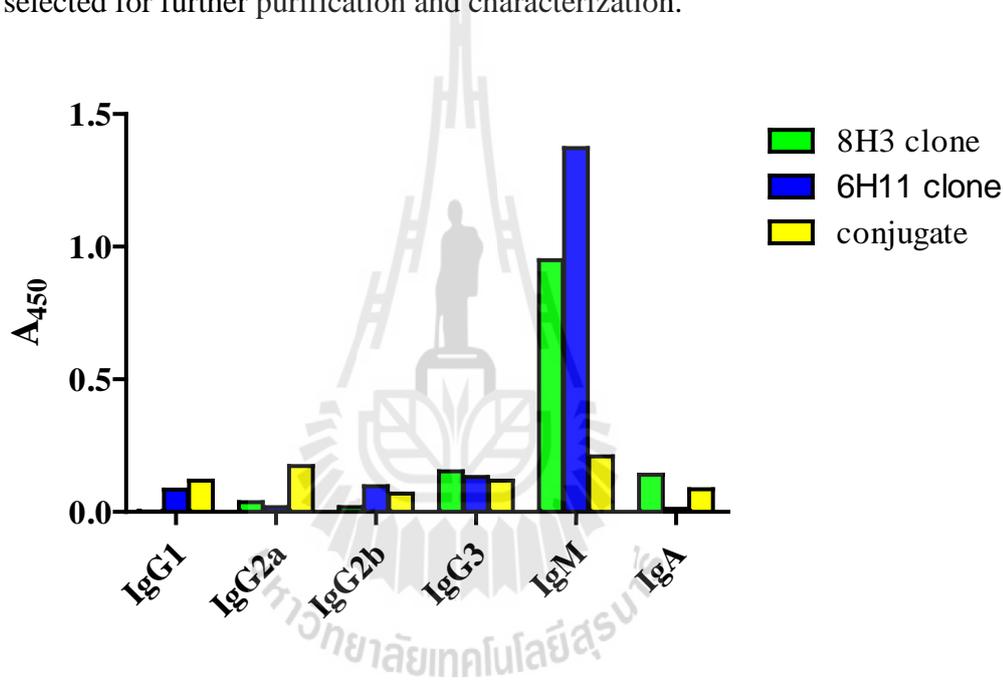
Figure 3.2.8 Determination of the isotype of the generated mAb of clones 6H11 and 8H3 clone by capture ELISA.

The culture supernatant was added into each well of a 96-well plate with isotype antibodies (goat anti-mouse IgG1, IgG2a, IgG2b, IgG3, IgM, and IgA), and antigen-antibody reaction was visualized by use of HRP-conjugated anti-mouse immunoglobulin and TMB was used as substrate. Absorbance was read at 450 nm. For conjugate reaction, the culture supernatant of a single clone is absent.

3.2.2.4 Large scale production of anti-YKL39 monoclonal antibody

The 6H11 monoclone was expanded in scale up. The reactivity of the stable 6H11 cell line was determined by indirect ELISA. The anti-YKL39 mAb-containing culture supernatant was collected, and then the titer of mAb was determined by indirect ELISA. Figure 3.2.11 shows the titer of anti-YKL-39 mAb from culture supernatant of the 6H11 hybridoma cell line. The mAb from culture supernatant strongly reacted with YKL-39 by giving A_{450} signal of more than 1.0 unit above A_{450} of the conjugate control. These results showed that crude anti-YKL39 medium strongly reacted with the purified YKL-39 protein.
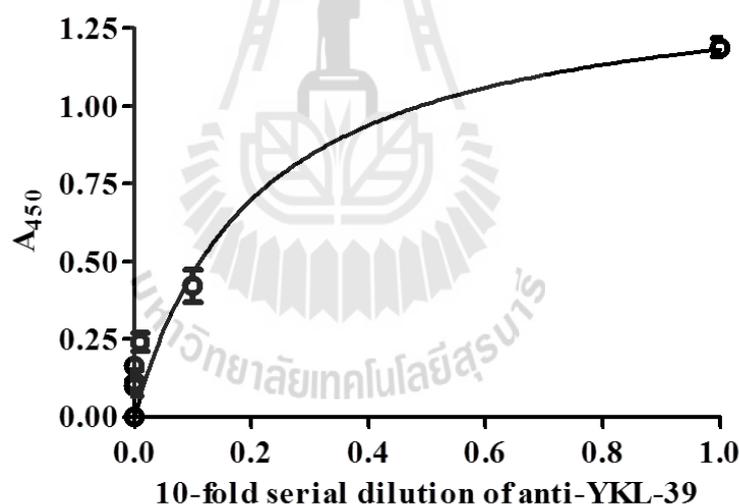
Figure 3.2.9 Determination of anti-YKL39 mAb titer of the 6H11 hybridoma cell line by indirect ELISA.

A 96-well plate was coated with 10 μg/mL YKL-39 and 10-fold serial dilutions of anti-YKL39 mAb-containing culture supernatant ranging from 1 to 1×10^{-5} were added in to the wells in triplicate. HRP-conjugated rabbit anti-mouse Igs antibody was used to detect protein-antibody complex and the absorbance measured at 450 nm with

TMB as substrate. Protein-antibody binding curve were evaluated by a nonlinear regression function available in Prism version 5.0 (GraphPad Software) fit to the single-site binding model.

3.2.2.5   Purification of anti-YKL39 mAb

The antibody-containing culture media were collected and checked for reactivity by indirect ELISA, as described in Section 3.2.2.4. Crude anti-YKL39 mAb was purified with a HiTrap IgM HP column (GE HealthCare) connected to an ÄKTA purifier system. The culture medium containing mAb was applied on a Hitrap IgM column, and unbound medium were washed with binding buffer, while bound IgM antibodies were eluted with elution buffer (20 mM sodium phosphate, pH 7.5). The purity of the mAb against YKL-39 was verified by SDS-PAGE. Figure 3.2.10 presents an elution profile of the purification, which showed a large peak in unbound fraction. Heavy and light chains of the antibody showed two bands migrating to approximately 92 kDa and 25 kDa. These protein bands corresponded to the molecular mass of IgM isotype that was seen in both unbound and eluted fractions as shown in Figure 3.2.11A. For Figure 3.2.11B, each purified fraction was also analyzed on a 8% PAGE gel under both reducing and non-reducing conditions. Under non-reducing condition, both unbound and eluted fractions were observed to contain a band on the top of the gel which corresponded to the high molecular weight of IgM, which consists of five subunit of approximately 180 kDa each (MW approximately 900 kDa). These results indicated that most of the monoclonal IgM did not bind to the column. However, it was found that there was no detectable contaminant in the eluted antibody, indicating that the mAb was successfully purified.
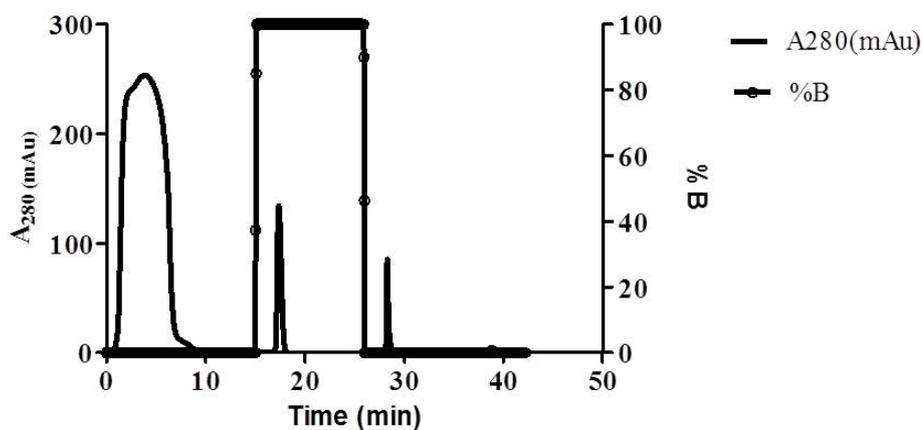
Figure 3.2.10 An elution profile of anti-YKL39 mAb purified by affinity chromatography on a HiTrap IgM HP column.

Crude supernatant containing anti-YKL39 mAb was adjusted to high salt concentration, in order to improve the binding ability. The sample was loaded onto a 1-mL Hitrap IgM column. Unbound samples were washed with binding buffer, while bound IgM antibodies were eluted with 20 mM sodium phosphate, pH 7.5. The purity of the purified mAb against YKL-39 was verified by SDS-PAGE.
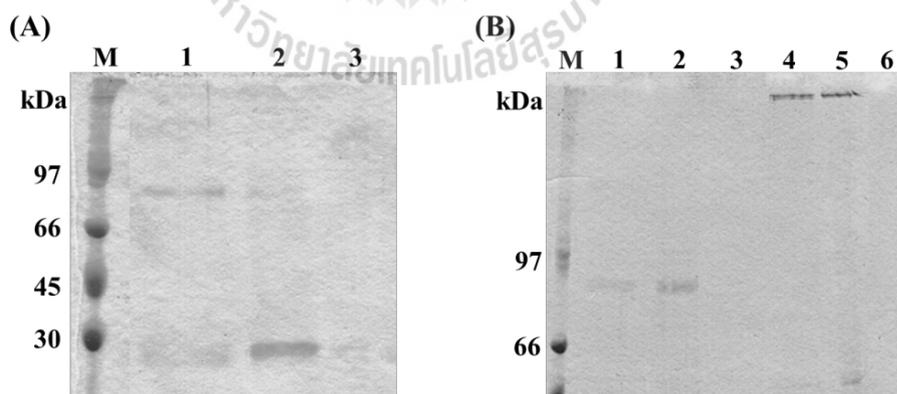
Figure 3.2.11 SDS-PAGE analysis of the purification of anti-YKL39 mAb using a HiTrap IgM HP column.

(A) 12% SDS-PAGE electrophoresis of purified anti-YKL39 mAb under reducing conditions: Lane M, molecular weight protein marker. Lane 1, unbound fraction. Lanes 2, eluted fraction. Lane 3, washed fraction with binding buffer. Lanes 4-5, the eluted fraction of IgG with elution buffer (0.1 M glycine, pH 2.5). (B) 8% PAGE of the eluted fractions: Lane M, molecular weight protein marker. Lane 1, unbound fraction. Lane 2, eluted fraction. Lane 3, washed fraction, while lanes 4-5, unbound and eluted fractions under non-reducing condition.

3.2.2.6 Characterization of anti-YKL39 monoclonal antibodies

To ensure the ability of anti-YKL39 mAb, the characterization of anti-YKL39 monoclonal antibody was characterized for specificity, sensitivity, and the ability of the monoclonal antibodies to bind to native human YKL-39 from various human cell lines.

3.2.2.6.1 Determination for specificity and sensitivity of anti-YKL39 mAb

To assess the specificity of the newly raised mAb against various GH-18 homologs, including YKL-40, acidic mammalian chitinase (AMCase) and bacterial chitinase, western blotting was carried out as described in polyclonal antibody production. As shown in Figure 3.2.12A, the antibody recognized both fused (Lanes 2 and 3) and bare (Lane 4) forms of human YKL-39, but did not recognize YKL-40 (Lane 5), AMCase (Lane 6), and bacterial chitinase (Lane 7). The antibody also did not cross-react with proteins in human serum (Lane 8). The results suggested that the mAb was highly specific for YKL-39. The sensitivity of generated anti-

YKL39 mAb was also examined by western blot analysis. Two-fold serial dilution of anti-YKL39 antibody was used ranging from 0.65-10 µg and YKL-39 was fixed at 2.5 µg. As seen in Figure 3.2.12B, the strongest titer was seen at a ratio of 10:5 (µg/µg) of mAb:Ag, while the slightest titer of the antibody that recognized the YKL-39 Ag was 2.5:5(µg/µg).
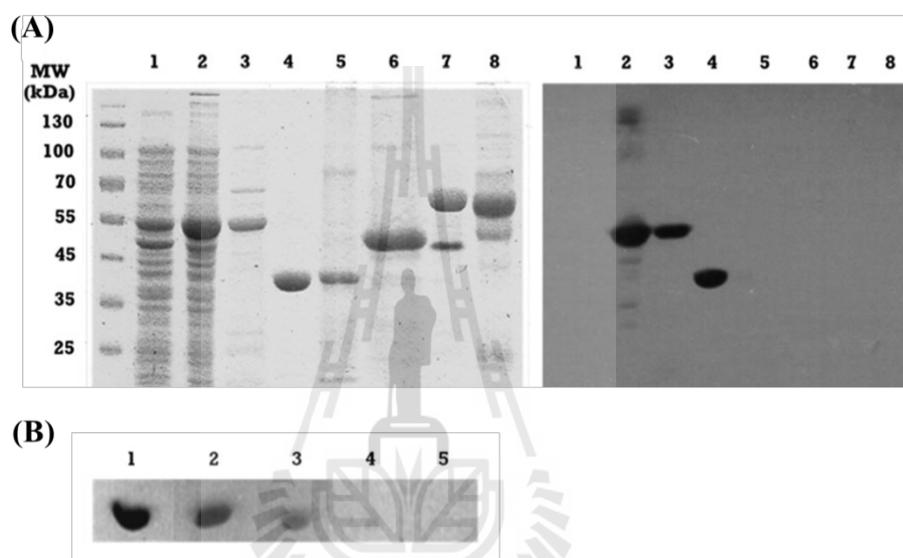
Figure 3.2.12 Cross-reactivity of anti-YKL39 mAb.

(A) Chitinases and chitinase-like proteins were analyzed by SDS-PAGE (left panel) in comparison with Western blot analysis (right panel). Lane M, protein molecular weight marker; Lane 1, cell lysate of uninduced *E. coli* BL21(DE3) host containing pET32a(+)/YKL-39; Lane 2, cell lysate of the cells induced with IPTG; Lane 3, elution of Trx-His$_6$/YKL-39 fusion protein from the first Ni-NTA affinity column; Lane 4, purified YKL-39 after anion exchange chromatography, followed by thioredoxin removal, and further purified by the second Ni-NTA affinity column; Lane 5, YKL-40; Lane 6, AMCase; Lane 7, bacterial *V. harveyi* chitinase A; Lane 8, human serum. (B) Sensitivity of purified anti-YKL39 mAb for YKL-39 protein was

analyzed by Western blot analysis. Two-fold serial dilution of the amount of anti-YKL39 antibody was used ranging from 0.65-10 µg and YKL-39 were fixed at 2.5 µg. Lane 1, 10 µg of anti-YKL39 mAb: Lane 2, 3, 4, and 5 are 5, 2.5, 1.25 and 0.62 µg of anti-YKL39 mAb, respectively.

3.2.2.6.2 Determination for reactivity of anti-YKL39 mAb against native YKL-39 in human cell lines and osteoarthritis sample

The ability of the monoclonal antibodies to bind to the native human YKL-39 form in various human cell lines, including Jurkat, U937, Hela, COS-7, and THP-1 and synovial fluid of osteoarthritic patients was also determined by immunoblotting. Purified YKL-39 or whole cell lysate prepared from various cell lines was spotted onto a nitrocellulose (NC) membrane, followed by incubation with the mAb from clone 6H11. The signals were detected by enhanced chemiluminescence using rabbit anti-mouse IgG antibody conjugated with horseradish peroxidase at 1:5000 dilution. Figure 3.2.13A (SDS-PAGE) and 3.2.13B (immunoblot) show that anti-YKL39 mAb reacted strongly with a protein band in the whole cell lysate of Jurkat cells (Lane 2). This protein migrated to the same position of recombinant YKL-39 (Lane 1) on SDS-PAGE. Jurkat cells have been reported previously to express SI-CLP, but not other CLPs (Kzhyshkowska *et al.*, 2007). However, Santa Cruz Biotechnology (Santa Cruz, CA) used the whole cell lysate of Jurkat cells to determine the specificity of their commercially available anti-YKL-39 polyclonal antibody (H-80) (Santa Cruz Biotechnology Inc.). Figure 3.2.13C is a dot blot assay showing strong cross-reactivity of anti-YKL39 mAb with YKL-39, the synovial fluid of a volunteer with osteoarthritis, and the whole cell lysate of U937

cells. However, the mAb did not recognize all proteins in the whole cell lysates of Hela cells, COS-7 cells, and THP-1 cells. The negative control, in which IgG secondary antibody was added in the absence anti-YKL39 mAb, did not produce a positive signal.
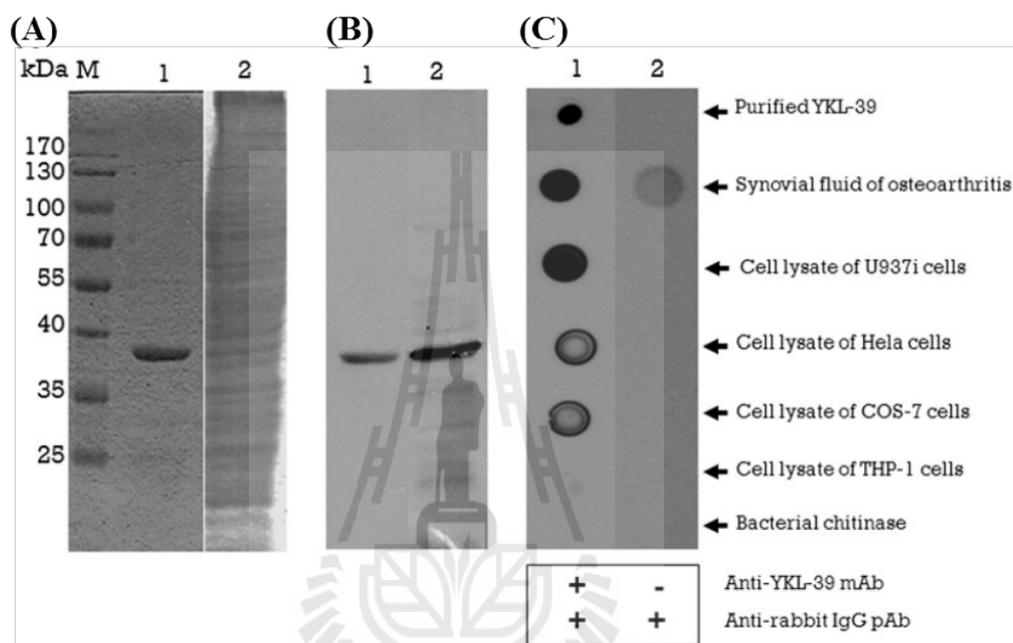
Figure 3.2.13 Detection of YKL-39 expression using anti-YKL39 mAb in human cell line and synovial fluid of osteoarthritic patients by dot blot analysis.

(A) SDS-PAGE analysis. (B) Western blot analysis. Lane 1, YKL-39; Lane 2, whole cell lysate of Jurkat cells. (C) Dot blot analysis showing antigen-antibody interactions with human sample or human cell lines. Purified YKL-39 or whole cell lysate was spotted onto nitrocellulose membrane, and then the membrane was allowed to dry. After blocking with 5% (w/v) skim milk, the membrane was incubated for 1 h with 6H11 mAb. The signals were detected by enhanced chemiluminescence using rabbit anti-mouse IgG antibody conjugated with horseradish peroxidase at 1:5000 dilution.

3.3 Determination of crystal structures of YKL-39

This section described determination of crystal structures of YKL-39 in the absence and presence chitooligosaccharide and details of sugar-protein interactions from data obtained from protein crystallization.

3.3.1   Crystallization of human YKL-39

3.3.1.1 Initial screening of crystallization of YKL-39

Initial crystallization experiments were firstly performed manually by the microbatch under oil method (Chayen *et al*., 1992). Equal volumes of protein and mother liquor were mixed under oil. After monitoring crystal growth for two months, YKL-39 crystals were not yet observed. The crystallized drops were mostly observed as clear drops and precipitates.

To solve the crystallization problem, the appropriate protein concentration for crystallization screening was firstly determined with the Pre-Crystallization Test (PCT ™, Hampton), using hanging drop vapor diffusion technique as described in Section 2.9.1.1. After incubation for 30 min, the reaction drops were carefully examined under a stereo microscope. The results revealed that the appropriate protein concentration for crystallization screening, which generated microcrystalline or light granular precipitate throughout the drop, was 12.75 mg/mL. So this protein concentration was chosen for further studies. Subsequently, initial crystallization of YKL-39 was performed by sitting drop vapor diffusion technique. After mixing equal volumes of each precipitating agent and YKL-39 solution, the crystal growth was monitored. The crystals with various sizes and needle clusters were observed under different conditions in the crystallization trials at both 15°C and 25°C. As listed in

Table 3.3.1, the YKL-39 crystals were obtained under conditions, including 25% (w/v) PEG 3350, 0.2 M Li_2SO_4, 0.1 M Bis Tris, pH 5.5, at 25°C on day 21 (Figure 3.3.1A); 30% (w/v) PEG 1500 at 15°C on day 5 (Figure 3.3.1B) and 20% (w/v) PEG 3000, 0.1 M citrate, pH 5.5 at 15°C on day 14 (Figure 3.3.1C). In addition, needle clusters were also observed under the condition containing 10% (w/v) PEG 8000, 0.2 M $Ca(OAc)_2$, 0.1 imidazole, pH 8.0, at 15°C on day 30 (Figure 3.3.1D). The precipitant, which was found to be appropriate for further crystallization of this protein, was PEG in a wide range of the molecular weights from 1,500 to 8,000. YKL-39 crystallization was successful under the mono-valent salt, Li_2SO_4, and di-valent salts, $Ca(OAc)_2$, while the buffers were Bis Tris, citrate and imidazole at pH 5.5 and 8.0.
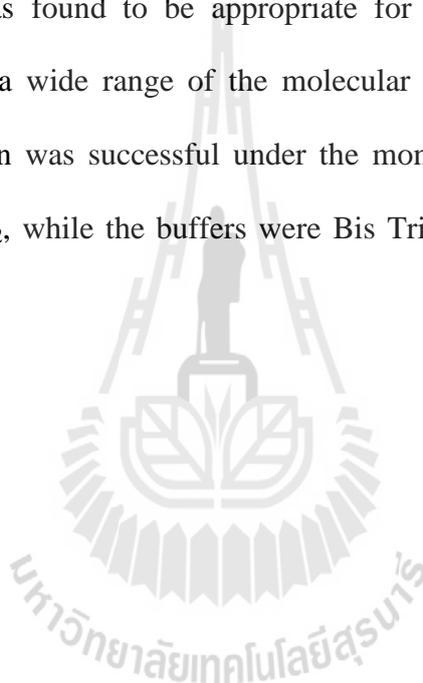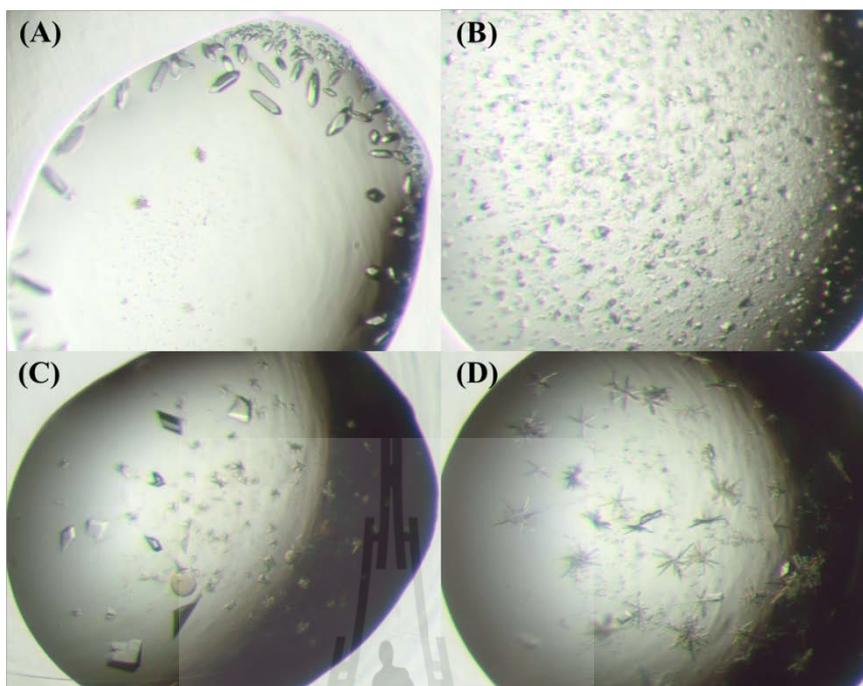
Figure 3.3.1 Crystals and needle clusters of ligand-free YKL-39 from initial screenings using sitting drop technique.

The conditions included (A) 25% (w/v) PEG 3350, 0.2 M Li_2SO_4, 0.1 M Bis Tris, pH 5.5 at 25°C; (B) 30% (w/v) PEG 1500 at 15°C; (C) 20% (w/v) PEG 3000, 0.1 M Citrate, pH 5.5 at 15°C; and (D) 10% (w/v) PEG 8000, 0.2 M $Ca(OAc)_2$, 0.1 M imidazole, pH 8.0 at 15°C.

Table 3.3.1 Crystallization conditions of ligand-freeYKL-39 from screening kits that yielded small crystals and needle clusters in sitting drop technique.

Condition no.	Precipitant	Salt	Buffer	Temperature (ºC)	Days of observation	Crystal morphology
H9 (JCSG-*plus* HT96)	25% (w/v) PEG 3350	0.2 M Li_2SO_4	0.1 M Bis Tris, pH 5.5	25	21	Crystals
D7 (*HR2-130)*	30% (w/v) PEG 1500	-	-	15	5	Small crystals
No. 6 (Wizard I)	20% (w/v) PEG 3000	-	0.1 M citrate, pH 5.5	15	14	Crystals
No. 46 (Wizard I)	10% (w/v) PEG 8000	0.2 M $Ca(OAc)_2$	0.1 imidazole, pH 8.0	15	30	Needle clusters

3.3.1.2 Optimization of YKL-39 crystallization

From all the positive screening conditions, the condition H9 from JCSG-*plus* HT96 gave the good quality hexagonal crystals under condition 25% (w/v) PEG 3350, 0.2 M Li$_2$SO$_4$, 0.1 M Bis Tris, pH 5.5. Thus, it was then selected for further optimization by the hanging drop method. The concentrations of PEG 3350 were varied in range of 21-31% (w/v) and lithium sulfate was varied ranging in 0.1-0.4 M in two dimension grid screening as shown in Figure 3.3.2. The high quality crystals of ligand-free YKL-39 were observed within one days at 25°C in the condition 27% (w/v) PEG 3350, 0.2 M lithium sulfate in 0.1 M Bis Tris, pH 8.0 and another condition of 31% (w/v) PEG 3350, 0.3 M lithium sulfate in 0.1 M Bis Tris, pH 8.0, while the condition of 29 and 31% (w/v) PEG 3350, 0.2 M lithium sulfate in 0.1 M Bis Tris, pH 8.0 generated single crystals, but smaller sizes with the cracked surface as shown in Figure 3.3.3. These crystals were used to perform preliminary X-ray diffraction by an in-house X-ray diffraction system. All crystals were diffracted, and gave diffraction patterns similar to a protein diffraction pattern (the data are not shown). However, they diffracted to low resolution at about 3.0-6.0 Å, which is not enough for collection and processing.

% (w/v) PEG 3350

	21	23	25	27	29	31
Li_2SO_4 **0.1 M**						
0.2 M						
0.3 M						
0.4 M						

Figure 3.3.2 Grid screen with variation of % PEG 3350 and lithium sulfate.
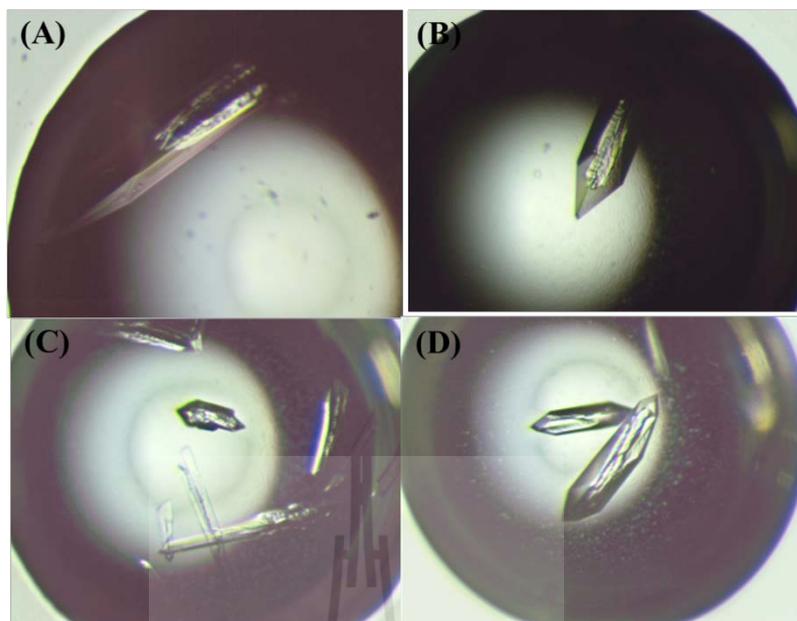
Figure 3.3.3 Photographs of YKL-39 crystals obtained from the crystal optimization using a variation of PEG 3350 and lithium sulfate concentrations.

The positive conditions, which yielded the crystals of YKL-39, are:

(A) 27% PEG 3350, 0.2 M Li_2SO_4 in 0.1 M Bis Tris, pH 8.0

(B) 29% PEG 3350, 0.2 M Li_2SO_4 in 0.1 M Bis Tris, pH 8.0

(C) 31% PEG 3350, 0.2 M Li_2SO_4 in 0.1 M Bis Tris, pH 8.0

(D) 31% PEG 3350, 0.3 M Li_2SO_4 in 0.1 M Bis Tris, pH 8.0

Subsequently, a refined grid screening optimization was performed by increasing concentrations of PEG 3350 from 27 to 37% (w/v) and lithium sulfate from 0.20 to 0.35 M, as shown in Figure 3.3.4. Big sizes of single crystals were observed within one day under the conditions: 27% PEG 3350, 0.2 M Li_2SO_4 in 0.1 M Bis Tris, pH 8.0, and 31% PEG 3350, 0.3 M Li_2SO_4 in 0.1 M Bis Tris, pH 8.0, (Figure 3.3.5). Under these two conditions, the fully-grown crystals were observed within seven days

after their first appearance. Thus, these conditions were chosen for crystallization of the ligand-free YKL-39 and for soaking with chitooligosaccharides in order to study protein-ligand binding. Many small crystals with cracked surface were observed under other conditions with highly percentages of PEG as seen in Figure 3.3.4. The result indicated that high concentrations of precipitant resulted in a decrease in crystal size.
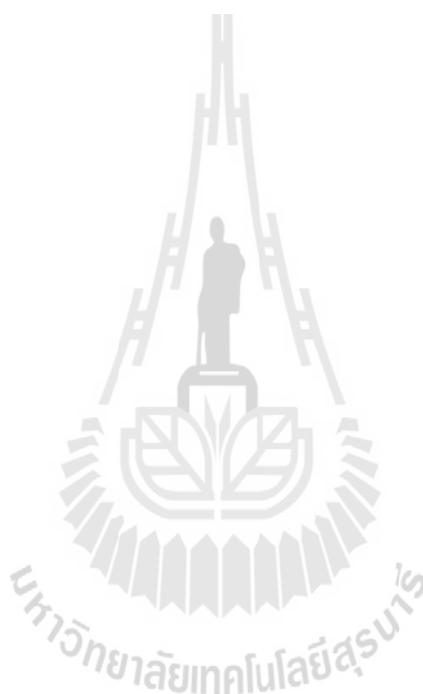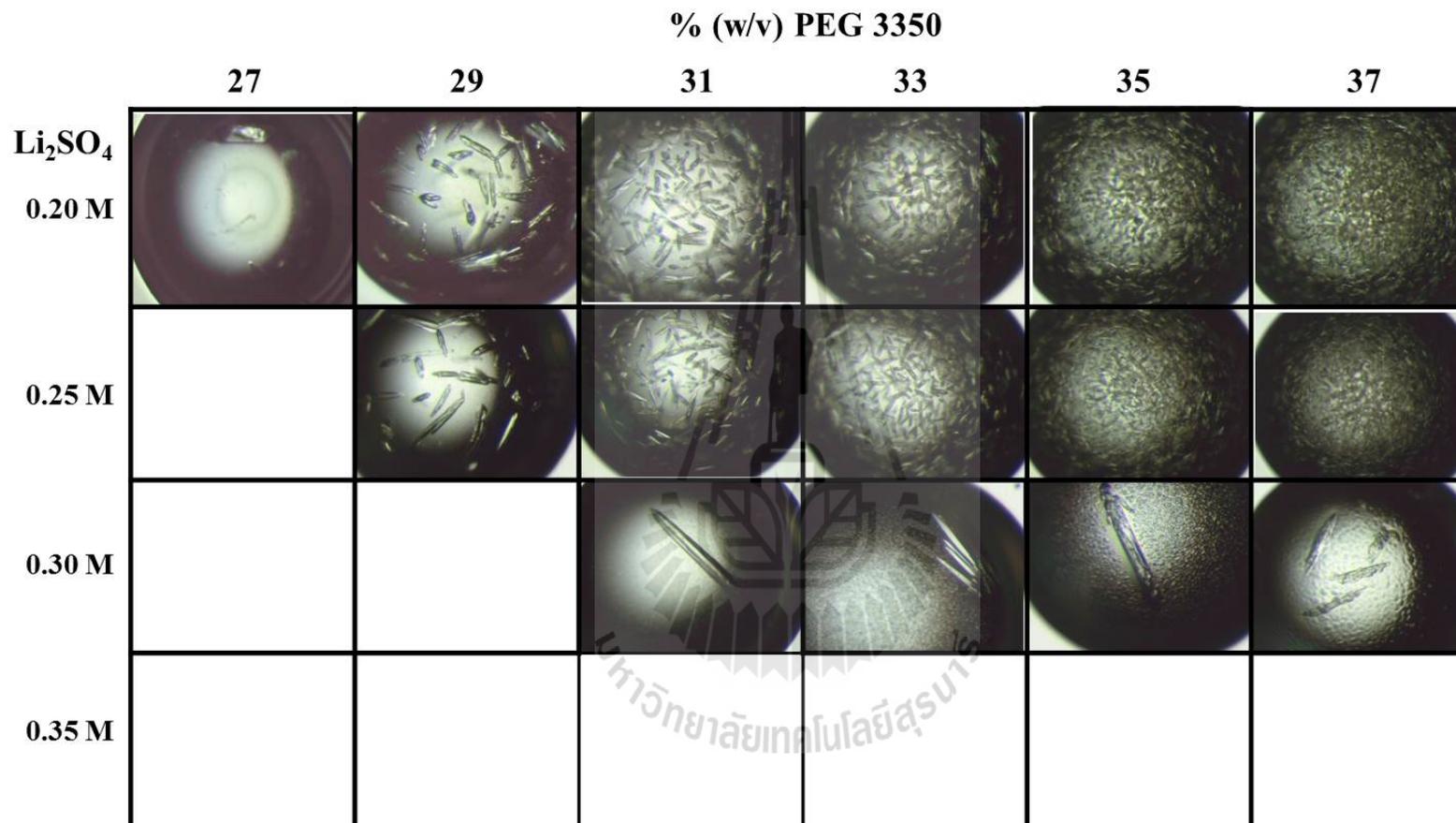
Figure 3.3.4 The refined grid screen with a variation of lithium sulfate and PEG 3350.
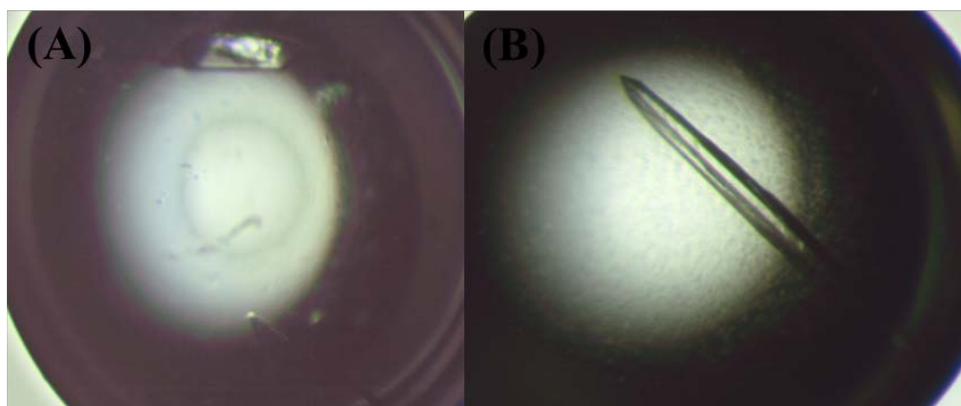
Figure 3.3.5   The positive conditions that gave large crystals within 1 day are:

(A) 27% PEG 3350, 0.2 M Li_2SO_4 in 0.1 M Bis Tris, pH 8.0

(B) 31% PEG 3350, 0.3 M Li_2SO_4 in 0.1 M Bis Tris, pH 8.0

3.3.1.3   Crystallizations of YKL-39 in complex with chitooligosaccharides

The suitable conditions for crystallization of ligand-free YKL-39 were used to produce high quality diffraction crystals as described in Section 3.3.2. Single crystals were allowed to fully grow for a week, and then soaked with various chains of chitin oligosaccharides, including $GlcNAc_2$ to $GlcNac_6$. The concentration of ligands was also optimized. It was found that the chitinsugar concentrations, which gave the X-ray quality crystals without being cracked or damaged, were 0.1 M $GlcNAc_5$ and $GlcNAc_6$, 5 M $GlcNAc_4$, 10 M $GlcNAc_3$ and $GlcNAc_2$.

3.3.2   X-ray diffraction analysis of ligand-free YKL-39 and its complex crystals.

Four X-ray data sets of ligand-free YKL-39 and its complexes with

GlcNAc$_2$, GlcNAc$_4$ and GlcNAc$_6$ crystals were successfully collected. They were diffracted at 100 K in a nitrogen stream from an Oxford Cryosystems Cryo-stream and data collected on an ADSC Quantum 315 CCD detector with the wavelength at 1.00 Å. All the diffraction data were indexed, integrated and scaled with the *HKL*-2000 program (Otwinowski and Minor, 1997). The ligand-free YKL-39 and YKL-39 soaked with GlcNAc$_2$, GlcNAc$_4$ and GlcNAc$_6$ crystals diffracted X-ray to 2.40 Å, 1.53 Å, 1.87 Å and 2.48 Å resolutions, respectively.
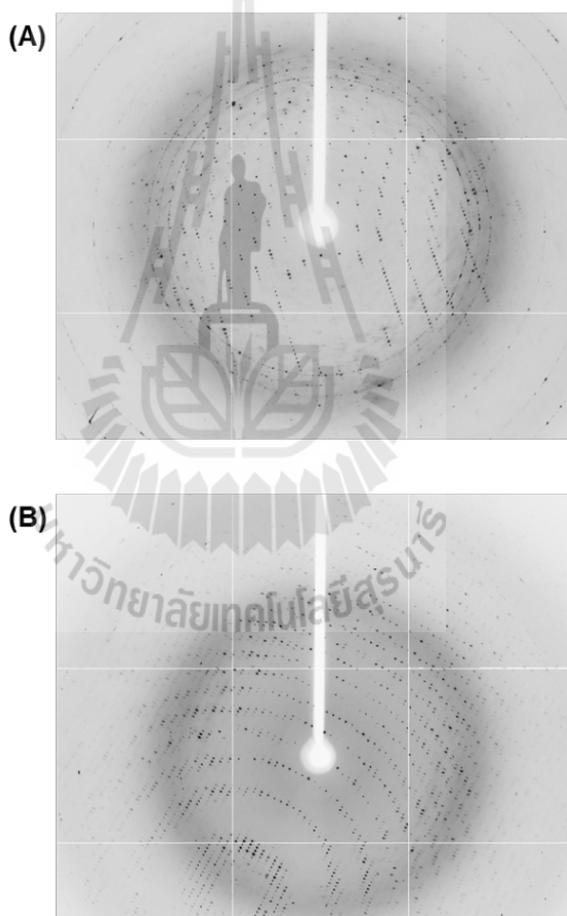
Figure 3.3.6 Diffraction images of ligand-free YKL-39 and YKL-39 complexed with GlcNAc$_2$.

The ligand-free YKL-39 (A) and YKL-39 complexed with GlcNAc$_2$ (B), crystals diffracted to 2.40 and 1.53 Å resolutions, respectively.

Table 3.3.2 Statistics of X-ray diffraction data collection of YKL-39.

	YKL-39	YKL39/ GlcNAc$_2$	YKL39/ GlcNAc$_4$	YKL39/ GlcNAc$_6$
Data collection statistics				
Beamline	BL13B1	BL13B1	BL13B1	BL13B1
Wavelength (Å)	1.00	1.00	1.00	1.00
Space group	$P4_12_12$	$P4_12_12$	$P4_12_12$	$P4_12_12$
Unit-cell parameter (Å)	a = 71.62	a = 70.68	a = 71.35	a = 71.21
	b = 71.62	b = 70.68	b = 71.35	b = 71.21
	c = 140.66	c = 140.26	c = 140.53	c = 141.02
Resolution range (Å)	26.27- 2.40	19.61 -1.53	19.8 - 1.87	24.5 - 2.48
Resolution outer shell (Å)	2.44 - 2.40	1.56 -1.53	1.97 - 1.90	2.54 - 2.50
No. unique reflections	11748	50934	26230	12681
No. observed reflections	181110	617279	225642	90749
Completeness (%)	99.9 (100.0)	99.3 (99.3)	99.0 (87.8)	99.6 (94.6)
Average redundancy per shell	15.3 (15.8)	11.5 (8.6)	7.4 (5.8)	6.8 (5.3)
$I/\sigma(I)$	101.0 (6.4)	26.3 (2.8)	45.1 (2.8)	11.0 (3.1)
$R_{(merge)}$ (%)	3.0 (46.9)	7.5 (75.3)	4.3 (45.5)	15.2 (34.7)

3.3.3 Phase determination by molecular replacement

The first data set belonging to YKL-39/GlcNAc$_2$ complex that was processed to 1.53 Å was solved by the molecular replacement with MOLREP (Vagin and Teplyakov, 1997, 2010) using the known structure of YKL-39 in complex with GlcNAc$_6$ (PDB access number 4AY1) as a model (Schimpl et al., 2012). The rotation, translation and subsequent fitting functions were performed by the same software. In the final fitting solution was performed in the P4$_1$2$_1$2 space group, and gave an R-factor of 47.25%.

3.3.4 Refinements and Structural determination of YKL-39 structures

3.3.4.1 The model quality of the YKL-39 structures

The other data sets of ligand-free YKL-39 and in complex YKL-39/GlcNAc$_4$ and YKL-39/GlcNAc$_6$ crystals were solved by using the final model structure of complex YKL-39/GlcNAc$_2$ as the starting model. The analysis of the F_o - F_c and 2F_o - F_c electron density map and model building was carried out in the COOT program (Emsley and Cowtan, 2004), restrained refinement in REFMAC5 from the CCP4 suite ("Collaborative Computing Project Number 4", 1994), and Phenix (Adams et al., 2002; Afonine et al., 2010). The geometry of each final model was validated by PROCHECK (Laskowski et al., 1993). All four diffraction data sets were determined to have tetragonal P4$_1$2$_1$2$_1$ space group symmetry with similar unit cell parameters, as summarized in the Table 3.3.2. All data sets were collected to 99.3-99.9% completeness over in the resolution range of 2.48-1.53 Å. These crystals contained one monomer per asymmetric unit, which corresponded to the same Matthew's coefficient (V_M) of 2.2 Å^3/Da. A solvent content was calculated to be

42.5% to 43.0%. The final model of the YKL-39/GlcNAc$_2$ complex, which was used as a model of the other data sets, was refined at 1.53 Å resolution with an R_{factor} of 17.4%, R_{free} of 20.7% with the R.M.S. deviations of bond length of 0.010 Å and bond angle of 1.48°. The 2,952 protein atoms (365 amino acid residues for one molecule in the asymmetric unit) exhibited a good fit with the electron density map with an average B-factor of 12.35 Å^2. The final $2F_o$-F_c map, contoured at 1.0 σ, showed a clear electron density map for GlcNAc$_2$ (29 ligand atoms) in subsites -2 to -1, with an average B-factor of 14.61 Å^2, while the 780 water molecules were refined with an average B-factor of 36.10 Å^2. In a good agreement with data, no residue was found in the outlier regions of the Ramachandran plots as summarized in Table 3.3.3. The ligand-free YKL-39 structure was refined at 2.40 Å resolution to an R_{factor} of 21.6% and R_{free} of 24.3% for 2,894 protein atoms (365 amino acid residues for a molecules in the asymmetric unit) and 168 water molecules with an average B-factor of 66.6 Å^2 and 79.4 Å^2, respectively. The final model of the YKL-39/GlcNAc$_4$ complex was refined to 1.87 Å with an R_{factor} of 18.6%, and an R_{free} of 23.6% with the R.M.S. deviations of bond length of 0.010 Å and bond angle of 1.37°. The 2,919 protein atoms (364 amino acid residues) showed good fit with the electron density map with an average B-factor of 23.7 Å^2. The final $2F_o$ - F_c map, contoured at 1.0 σ, shows an electron density map for GlcNAc$_4$ (57 ligand atoms) covering subsites -3 to +1 with an average B-factor of 31.46 Å^2, while 428 water molecules were refined with an average B-factor of 39.04 Å^2. The crystal structure of the YKL-39/GlcNAc$_6$ complex was refined at 2.48 Å with an R_{factor} of 18.7% and R_{free} of 23.5% with the R.M.S. deviations of bond length of 0.008 Å and bond angle of 1.28°. The 2,894 protein atoms (365 amino acid residues) exhibited a good fit with the electron density map

(an average B-factor of 48.6 Å^2). The final $2F_o$ - F_c map, contoured at 1.0 σ, showed clearly electron density map for GlcNAc$_6$, which appeared to cover subsites -3 to +3. The average B-factor was found to be of 53.1 Å^2 (85 ligand atoms), while 263 water molecules were refined with an average B-factor of 49.8 Å^2. In summary, the fitted model was in a good agreement with the calculated model, with no residue in the outlier regions of the Ramachandran plots, as shown in Table 3.3.3. Atomic coordinates and structure factors of the final models of YKL-39 have been deposited in the Protein Data Bank with PDB accession numbers: 4P8U for apo YKL-39, 4P8V for the YKL-39/GlcNAc$_2$ complex, 4P8W for the YKL-39/GlcNAc$_4$ complex, and 4P8X for the YKL-39/GlcNAc$_6$ complex.
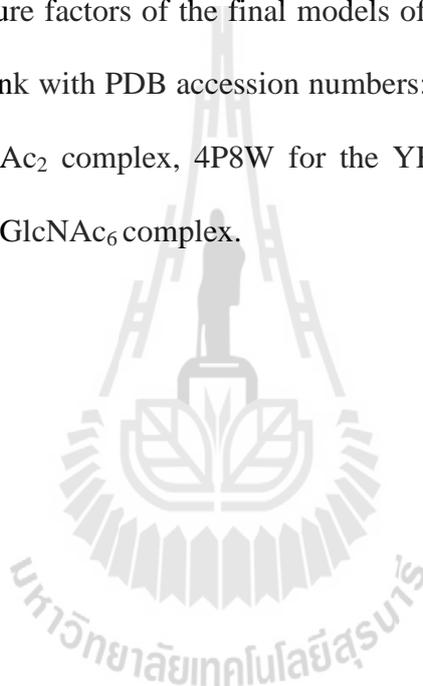
Table 3.3.3 Statistics of structural refinement.

	YKL-39	YKL39/ GlcNAc$_2$	YKL39/ GlcNAc$_4$	YKL39/ GlcNAc$_6$
PDB access code	4P8U	4P8V	4P8W	4P8X
Refinement statistics				
Resolution range (Å)	26.27-2.40	19.61-1.53	19.8- 1.87	24.5-2.48
[a]R$_{factor}$ (%)	21.6	17.4	18.6	18.7
[b]R$_{free}$ (%)	24.3	20.7	23.6	23.5
No. of residue in protein	365	365	364	365
No. protein atoms	2896	2952	2919	2894
No. ligand atoms	-	29	57	85
No. waters	168	780	428	263
Mean B-factor				
Protein	66.64	16.35	23.72	48.55
Ligand	-	14.61	28.81	53.11
Water	79.45	36.10	34.63	49.85
R.M.S. deviations				
Bond length	0.013	0.010	0.010	0.008
Bond angle	1.39	1.48	1.37	1.28
Ramanchandan plot				
Residue in favoured region (%)	96.1	98.6	97.8	97.8
Residue in allowed region (%)	3.9	1.4	2.2	2.2
Residue in outlier (%)	0.0	0.0	0.0	0.0

Numbers in parentheses are outer shell parameters.

[a]$R_{factor} = (\Sigma|F_{obs}|-|F_{cal}|/\Sigma|F_{obs}|)$.

[b]Based on 5% of the data.

[c]Calulated with PROCHECK (version 1.00.0)

3.3.4.2 The overall structure of YKL-39 in the presence and absence chitin fragments

The overall structure of YKL-39 contains two conserved domains (Figure 3.3.7A) that are found in all GH-18 chitinases and chitinase-like proteins. The major $(\beta/\alpha)_8$ TIM barrel domain (Domain I) comprises eight parallel strands B1-B8 (purple), which run in alternation with eight helices A1-A8 (cyan) as shown in Figure 3.3.7B. It is noticeable that helices A1, A3, and A6 are broken and contain short helices: G1-1 in helix A1, G3-1 and G3-2 in helix A3, and G6-1 and G6-2 in helix A6. The TIM barrel domains of GH-18 glycosyl hydrolases are known to interact specifically with chitin oligosaccharides (Aalten *et al.*, 2001, Fusetti *et al.*, 2002, Fusetti *et al.*, 2003, Meng *et al.*, 2010, Songsiriritthigul *et al.*, 2008, and Schimpl *et al.*, 2012,). The second domain is termed the small $\alpha+\beta$ insertion domain (Domain II) (Figure 3.3.7A). This domain is inserted between the tail of strand B7 and the start of helix A7. It is made up of six anti-parallel strands connected by a short helix that forms a typical Greek key motif. The exact function of this domain is unknown, but certain amino acid residues in this domain help to stabilize the sugar-protein complex. The substituted catalytic residue Glu145 to Ile and Asp143 to Ser in the loop of strand β_4, which is part of a DxxDxDxE conserved sequence motif, makes it lack chitinase hydrolytic activity. In contrast to this subfamily, YKL-39 is without a chitin binding domain. Detailed structural analysis revealed that YKL-39 contains two disulfide bonds, the first bond links Cys31 and Cys56 that are located within the TIM barrel domain. The second bond forms between Cys308 of the TIM barrel domain and Cys372 of the small insertion domain. Both bonds help maintain the structural integrity of the protein. In addition, six *cis* peptide bonds are observed, including

Glu42-Pro43, Ser62-Phe63, His112-Pro113, Ile145-Tyr146, Lys220-Pro221, and Trp360-Ser361, two of which (Ile145-Tyr146 and Trp360-Ser361) directly participate in ligand binding. The structures of YKL-39 in the presence of chitin fragments detail the binding surface of the TIM barrel domain, consisting a long, deep groove with approximate dimensions of 35 Å (long) x 17 Å (wide) x 7 Å (deep). This groove has been reported to interact with chitin fragments in other GH-18 homologues (Fusetti *et al*., 2003; Fusetti *et al*., 2002; Meng *et al*., 2010; Songsiriritthigul *et al*., 2008; van Aalten *et al*., 2001).. The shape of this groove is a crevice in which both ends are open, permitting chitooligosaccharides of different sizes to be accommodated and extended out from the groove in either direction. The narrowest part of the crevice points towards the center of the TIM barrel domain, a feature that favors a bent sugar chain.
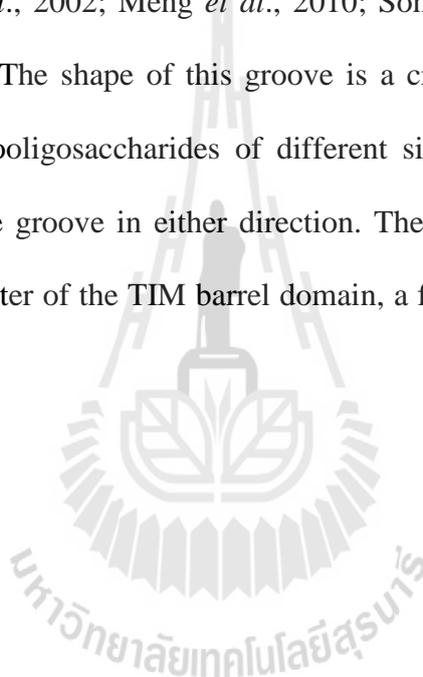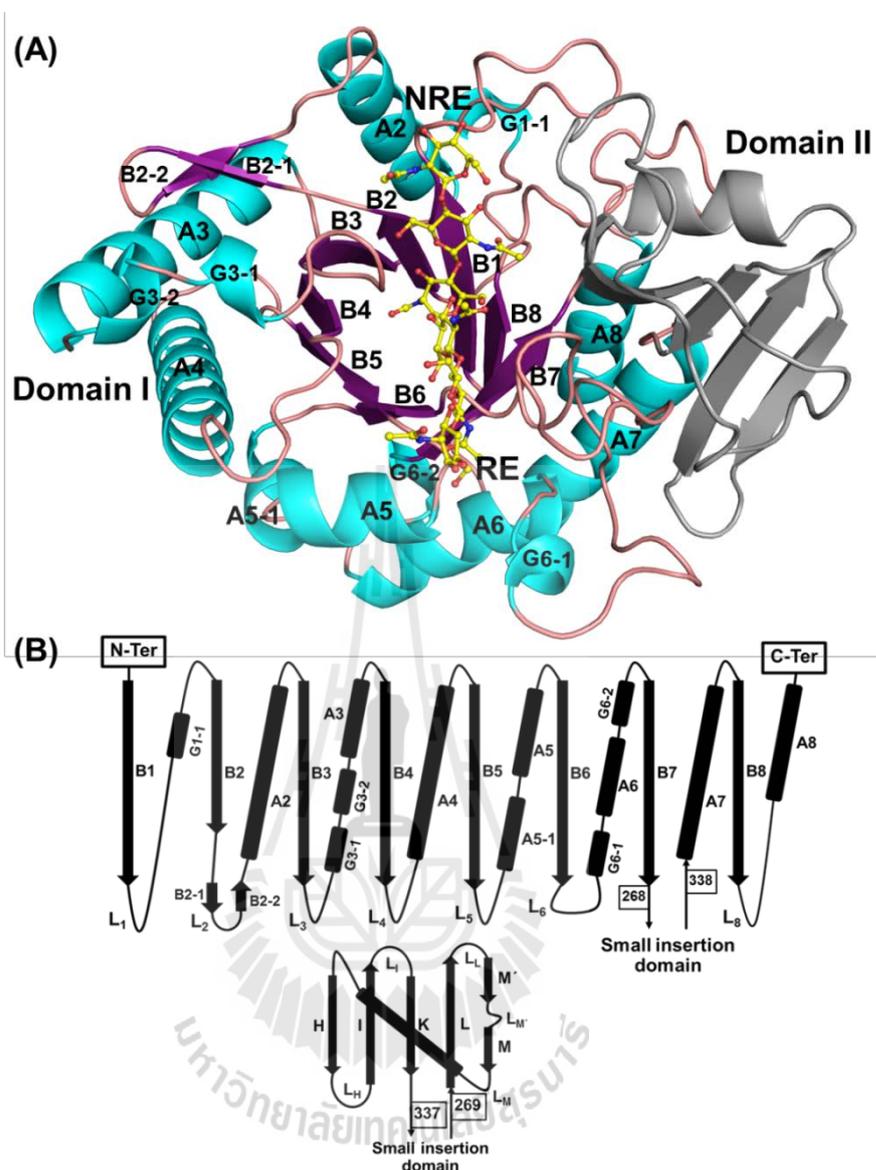
Figure 3.3.7 The overall structure of YKL-39.

(A) The cartoon model of YKL-39 in complex with $GlcNAc_6$. The chitin-binding cleft is located on the surface of the $(\beta/\alpha)_8$ TIM-barrel domain (domain I), which is presented in cyan for helices and purple for strands, while the $\alpha+\beta$ insertion domain (domain II) flanked by the strand B7 and the helix A7 is shown in gray. The chitooligosaccharide $GlcNAc_6$ is shown as ball and sticks and colored by atoms with yellow for carbons, blue for nitrogen, and red for oxygen atoms. RE and NRE denote

the "reducing" and "non-reducing" ends of the sugar chain. (B) Structural topology of YKL-39. The eight strands are designated as B1 to B8, loops as L, and eight α-helices connected between β-strands as A1 to A8. The helices are often segmented. Short helices are referred to by the letter G, for instance, G1-1 indicates the short helix within the region of the A1 helix.

The structures of YKL-39 in the presence of $GlcNAc_2$, $GlcNAc_4$, and $GlcNAc_6$ were successfully determined. The overall structures of YKL-39 in the presence chitin fragment were considered. The longest chitin fragment, $GlcNAc_6$, complex was found within the long deep groove through the central of protein molecule covering subsites (-3)(-2)(-1)(+1)(+2)(+3) as seen in Figure 3.3.8. From the subsite definition as assigned by Davies and Henrissat (1995), the non-reducing end is designated as subsite -3, followed by the subsites located towards the reducing end are designated as subsites -2, -1, +1, +2 and +3, respectively. The binding site was lined with a few invariant aromatic acid residues. Three of them, Trp36, Trp360 and Trp218 were found in suitable positions to form stacking interactions with the bound chitin fragment. A conserved residue, Asp213, which was observed in GH-18 chitinases and CLPs, was in the position to form H-bond with the subsite -1GlcNAc (Figure 3.3.8A). Figure 3.3.8B represents the electrostatic analysis, which revealed the surface of binding cleft of YKL-39. A negatively charged region was observed in the deep, central part of binding sites.
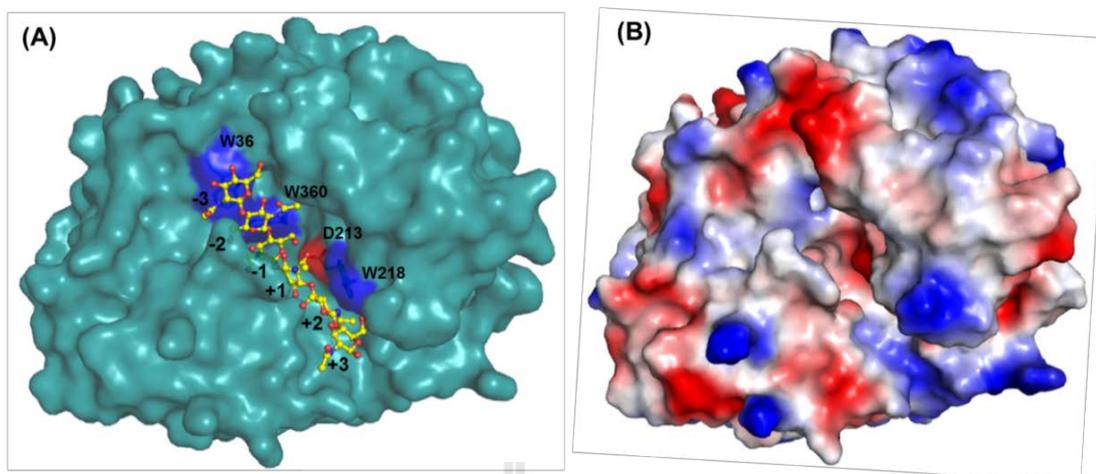
Figure 3.3.8 Electrostatic surface of YKL-39 in complex with $GlcNAc_6$.

(A) The important amino acids interacting with $GlcNAc_6$ by hydrophobic interaction are shown in blue and H-bonds in red. The sugar chain is represented as sticks, carbon in yellow, oxygen in red and nitrogen in blue. (B) Electrostatic surface representation of YKL-39 in complex with $GlcNAc_6$. The positively charged regions are shown in blue, the neutral regions in white, and the negatively charged regions in red. The figures were created by Pymol (DeLano, 2002).

3.3.4.3 Chitooligosaccharide binding induces structural movements

Superimposition of the apo YKL-39 structure with structures in complex with $GlcNAc_2$, $GlcNAc_4$, and $GlcNAc_6$ gives C_α RMSDs for 326 residues of 0.46, 0.38, and 0.39 Å, respectively, as shown in Table 3.3.4. These values reflect some structural non-identity and that may be induced upon sugar binding. Figure 3.3.9A presents the $(\beta/\alpha)_8$ TIM barrel domains of the ligand-free YKL-39 and the sugar-bound YKL-39, showing the small structural differences that were observed in the loop regions surrounding the surface of the chitin binding cleft. Close inspection of the GlcNAc-binding subsites of the apo protein (Figure 3.3.9A, blue) in

comparison with that of the protein in complex with the longest chitooligosaccharide GlcNAc$_6$ (Figure 3.3.9, cyan) reveals considerable movements of loop L$_1$ on the surface of subsite -3, loops L$_3$ and L$_{M'}$ (see Figure 3.3.7B for loop assignment) near subsites -2 and -1, and L$_6$ near subsites +2 and +3 in the structure with GlcNAc$_6$. As compared to the exterior of the native protein (Figure 3.3.9B), four key residues: Tyr104 and Leu105 (part of the bottom loop L$_3$), and Met364 and Phe301 (part of the top loop L$_{M'}$) were found to shift significantly towards the center of the chitin-binding cleft. This narrows the cleft of the sugar bound protein around subsites -2/-1 by 1.3 Å (Figure 3.3.9C). Such local movements engender close contacts between the GlcNAc rings and the binding residues around the corresponding subsites. In contrast, Tyr243 (part of loop L$_6$) rotates away from its original position, making the cleft at this particular subsite (+3) to widen by 1.9 Å in comparison to that of the non-ligand-bound structure (Figure 3.3.9D). This orientation of Tyr243 increases the access to subsite +2 from the subsite +3 direction, suggesting that longer sugars may be accommodated in this conformation. Taken together, these structural differences provide evidence that the binding of chitooligosaccharides induces local structural changes that help strengthen the interactions between the YKL-39 and the chitosugars.

Table 3.3.4 Root-mean-square deviations between the structures of ligand-free YKL-39 and YKL-39 in complex with chitooligosaccharides.

The complex crystal structures	RMSD	The number of C_α
YKL-39/GlcNAc$_2$	0.46	327
YKL-39/GlcNAc$_4$	0.38	321
YKL-39/GlcNAc$_6$	0.39	334
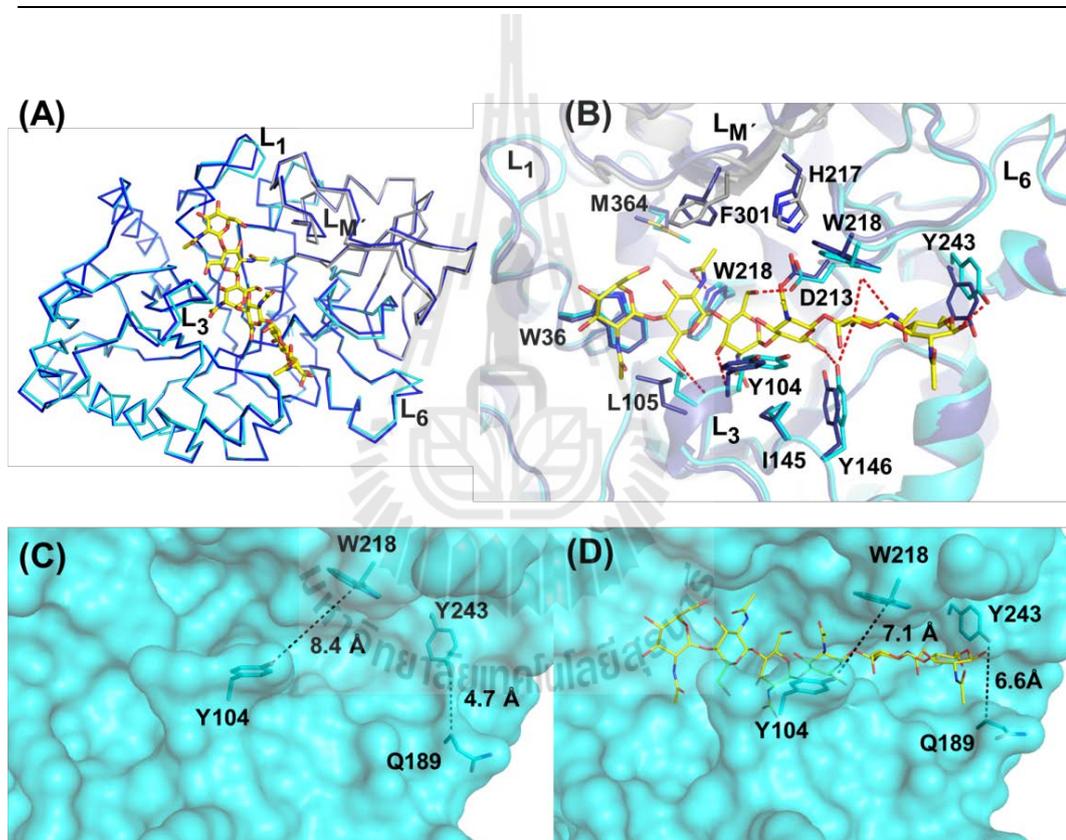
Figure 3.3.9 Structural comparison of the YKL-39 structures in the absence and presence of GlcNAc$_6$.

(A) The structure of ligand-free YKL-39 (royal blue) was superimposed with the structure of YKL-39 in complex with GlcNAc$_6$ (cyan and grey). (B) A close up of the GlcNAc$_6$ binding pocket from (A). The binding residues around the chitin binding cleft that have different conformations in the two structures are included (Leu105,

Tyr104, Met364, Tyr243 and Phe30), which are displayed as sticks. (C) Surface representation of the sugar-binding groove of the apo form of YKL-39. (D) Surface representation of the sugar binding groove of the YKL-39/GlcNAc$_6$ complex, showing local changes causing the closure of the surface area around subsites -2 and -1, and the widening of the area beyond subsite +2 as compared to the binding groove of the apo protein.

3.3.4.4 Specific interactions of YKL-39 with chitooligosaccharides

Like all GH-18 members, the sugar binding cleft of YKL-39 comprises multiple subsites for GlcNAc units. The critical binding features of these subsites are aromatic residues, which interact with the GlcNAc rings mainly through hydrophobic interactions, and charged or polar side chain residues, which form hydrogen bonding interactions with the saccharide chain. Inspection of the substrate binding cleft of YKL-39 in apo form (Figure 3.3.10A) and YKL-39 in complexes with GlcNAc$_2$ (Figure 3.3.10B), GlcNAc$_4$ (Figure 3.3.10C), and GlcNAc$_6$ (Figure 3.3.10D) reveals the binding behaviors of individual chitooligosaccharides. GlcNAc$_2$ was found at the center of the binding cleft between subsites -2 and -1 (Figure 3.3.10B). As summarized in Table 3.3.5 (H-bonding interaction) and Table 3.3.6 (binding residues), subsites -2 and -1 comprise a large number of binding residues. Residues Phe63, Leu105, Phe301, Trp360, Met364 interact with the GlcNAc ring at subsite -2, while Tyr32, Ser143, Tyr104, Ile145, Asp213, Tyr269, and Trp360 are found at subsite -1. The contacts around the center of the binding cleft make highly ordered interactions, which suggest that an arriving chitin chain would initially and preferentially interact at the center of the binding groove. For a longer chain

chitooligosaccharide, the interaction may become extended through occupation of the neighboring weaker-affinity subsites, where the four sugar rings $GlcNAc_4$ were extended from subsites -3 to +1 (Figure 3.3.10C), and by the complex with $GlcNAc_6$ from subsites -3 to +3 (Figure 3.3.10D). This affinity gradient is well supported by lowest B-factor for the subsite -2GlcNAc (43.63 $Å^2$), indicating high rigidity due to tight interactions at the internal site. Increases in B-factor for bound sugar at the extended subsites (-1GlcNAc, 44.76, +1GlcNAc, 44.75, +2GlcNAc, 52.48, +3GlcNAc, 81.64, and -3GlcNAc, 64.28 $Å^2$) apparently reveal high flexibility of the sugar rings forming weak interactions with the binding residues at such subsites. Comparing the three complex structures, they displayed essentially identical. The superimposition of all the structures with the ligand gave an RMSD value of 0.12 Å for the 325 Cα residues. Taken together, these results demonstrated that only $GlcNAc_2$ is sufficient to induce the local conformation change at the binding site.
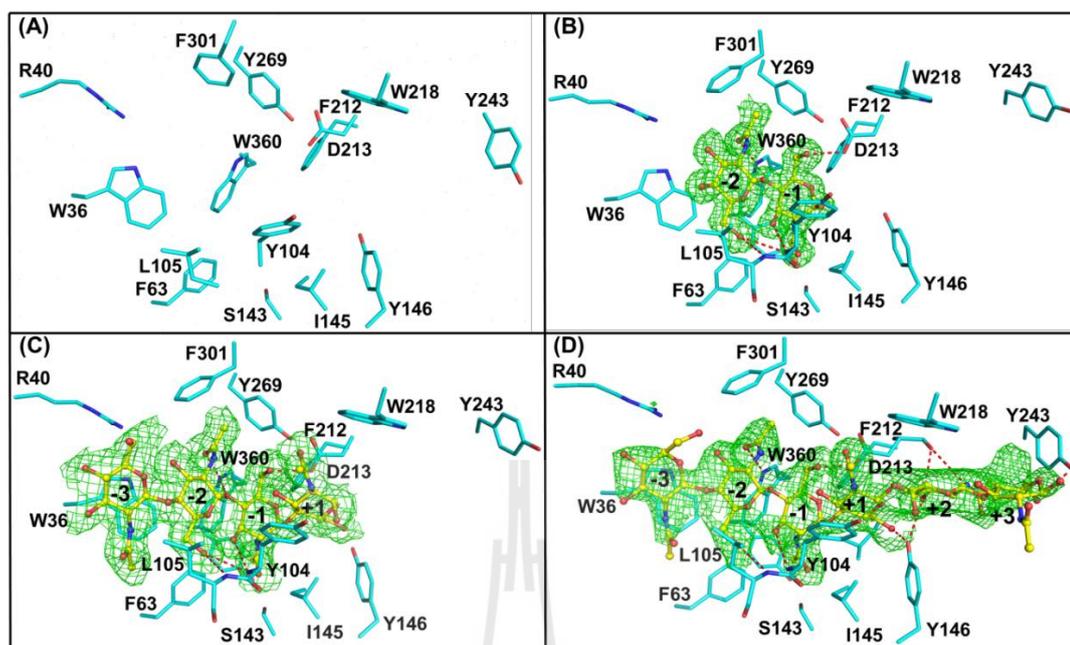
Figure 3.3.10 The residues involved in chitooligosaccharide/YKL-39 binding.

The residues contributing hydrogen bonding and hydrophobic interactions are displayed as sticks. (A) the unliganded form of YKL-39; (B) YKL-39 in complex with GlcNAc$_2$; (C) YKL-39 in complex with GlcNAc$_4$; (D) YKL39 in complex with GlcNAc$_6$. The sugar residues are shown as sticks and colored by atom, with yellow for carbon, blue for nitrogen and red for oxygen. Hydrogen bonds are shown as red dashed lines. The $2F_o$-F_c is contoured at 1σ.

Hydrophobic interactions between chitohexaose (GlcNAc$_6$) and aromatic/hydrophobic residues surroundings subsites -3 to +3 of the YKL-39 binding groove are analyzed and shown in Figure 3.3.11A. Three key residues, Trp36 at subsite -3, Trp360 at subsite -1, and Trp218 at +2, are found to stack directly against the plane of the pyranose rings of the GlcNAc units. These interactions are expected to contribute significantly to binding at these subsites. The first two residues (Trp36

and Trp360) are completely conserved in other GH-18 homologues and act as key binding residues (Fusetti *et al.*, 2003; Fusetti *et al.*, 2002; Meng *et al.*, 2010; Songsiriritthigul *et al.*, 2008; van Aalten *et al.*, 2001). The conserved aromatic Trp36, located at the edges of the binding cleft, marks the non-reducing end of the long chain chitin fragment by stack interaction to the planes of the bound sugar at subsite -3GlcNAc, while Trp218 marks the reducing end of the chitin sugar, and is in the position suitable for binding with the subsite +2GlcNAc. For the subsite +3GlcNAc, no hydrophobic interaction was found, since the crucial binding residue equivalent to W212, which is seen in the YKL-40 homologues, is missing in YKL-39. It is presumed that there is a weak binding at this subsite.

In spite of lacking chtinase activity in YKL-39, chitooligasaccharides still bind to YKL-39 in a fashion similar to that found in chitinase. A long chain sugar, such as GlcNAc$_6$ was found in the chitin binding cleft of YKL-39 in a bent conformation, where the sugar ring at subsite -1 still adopting a boat conformation like the sugar chain in true chitinase (Songsiriritthigul *et al.*, 2008; van Aalten *et al.*, 2001). The electron density of YKL-39/GlcNAc$_6$ complex suggested that the sugar units at the central subsites are held by a larger number of weak interactions, including salt bridges, H-bond, and hydrophobic interactions. Trp360 seem to play a particular role in ligand recognition by making hydrogen bonds with the C7-O of the subsite -2GlcNAc. This residue is conserved in GH-18 chitinases and CLPs (Fusetti *et al.*, 2003; Fusetti *et al.*, 2002; Meng *et al.*, 2010; Schimpl *et al.*, 2012; Songsiriritthigul *et al.*, 2008; van Aalten *et al.*, 2001). The binding of -2GlcNAc was stabilized by a hydrogen-bonding network, involving Trp360, and Leu105, while the conserved Asp213 and Tyr104 residues formed hydrogen bonds with C6-O and C3-O of the

subsite -1GlcNAc, respectively. Figure 3.3.11B displays a number of hydrogen bonds that help to stabilize the sugar/protein complexes. Such hydrogen bonds are formed either directly between the sugar rings and the binding residues or mediated by water molecules. As summarized in Tables 3.3.5 and 3.3.6, high densities of interactions are seen at subsites -2, -1, and +1, which are likely to determine the preferential binding at these sites.
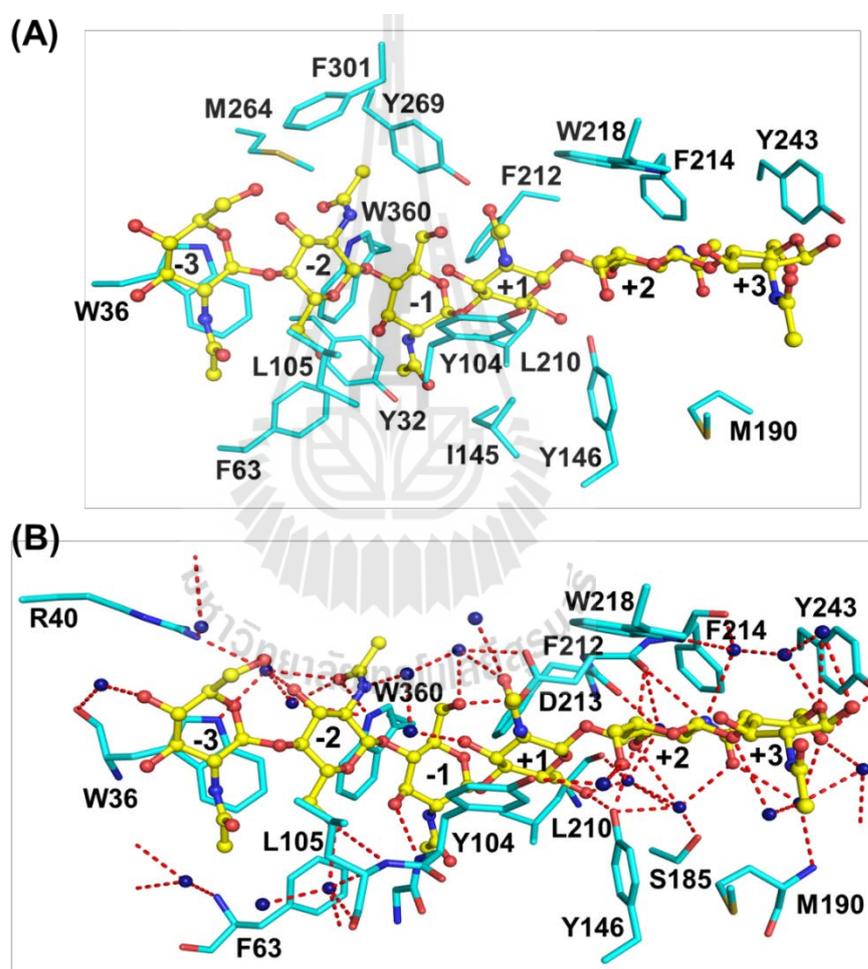
Figure 3.3.11 Specific interactions within the chitin binding cleft of YKL-39.

(A) The aromatic residues that are involved hydrophobic interactions with the sugar molecule at each subsite. (B) The amino acid residues and water molecules that

mediate the formation of the hydrogen bonding network. Water molecules are shown as blue diamonds and hydrogen bonds are shown as red dashed lines.

Table 3.3.5 A summary of the H-bonding interactions between GlcNAc$_6$ and the binding residues in the chitin binding cleft of YKL-39.

GlcNAc subsite	GlcNAc atom	Binding residues/atom	Distance (Å)
+3	O6	Y243/OH	3.1
+2	O3	Y146/OH	3.1
	N2	D213/O	3.3
+1	O6	Y146/OH	2.5
-1	O3	Y104/N	2.9
	O6	D213/OD2	2.7
-2	O6	L105/N	3.3
	O7	W360/NE1	2.9
-3		Not observed	

Table 3.3.6 A summary of the interactions between GlcNAc₆ with the binding residues in the chitin binding cleft of YKL-39.

GlcNAc subsites	Binding residues
+3	Y243, E219
+2	Y146, S185, M190, D213, F214, W218
+1	Y104, Y146, H271
-1	Y32, Y104, S143, I145, L210, F212, D213 , Y269, W360
-2	F63, F65, Y104, L105, M264, F301, W360
-3	N35, R40, W36, Y269

3.5.3 Structural comparison with other human GH-18 members

Superimposition of the wild type YKL39/GlcNAc$_6$ complex obtained in this study with the complex of YKL39/GlcNAc$_6$ complex (4AY1) reported by Schimpl *et al*., 2013, is characterized by an RMSD of 0.22 for Cα positions over 334 atoms (Figure 3.3.12). Thus the two complexes are essentially identical, even though they are derived from crystals of different space groups. The differences between the two protein sequences are at residues 35 and 318. In Schimpl's structure Asn35 was mutated to Gln, generating N35Q mutant, to prevent glycosylation. In this study, YKL-39 was functionally expressed in *E. coli* system as a native, non-glycosylated protein. In our structure, Asn35 makes three salt bridges, two of which link with Tyr61 and Asp75 and the third one with a neighboring water molecule (Figure

3.3.12A, left panel and inset). Gln35 in Schimpl's structure flips to a vertical direction and forms salt bridges with Tyr61, Lys74, and also a neighboring water molecule. Although the residue 35 in both structures lies in subsite -3, none makes contact with the GlcNAc ring at the corresponding subsite. Residue 318, which is seen as Trp in our structure, lies in a region of the protein distant to the chitin-binding groove. In the present structure, Trp318 forms hydrophobic interactions with Pro325 and the stalk of Lys340, while in structure 4AY1 Arg318 forms a salt bridge with Asp338 (Figure 3.3.12A, right panel and inset). These interactions are standard residue:residue interactions, in line with the view that these substitutions arise from naturally occurring polymorphisms. In our structure, all six GlcNAc rings could be fitted into subsites -3 to +3 (Figure 3.3.12B, left panel), whereas only four GlcNAc rings of chitohexaose $GlcNAc_6$ were fitted into subsites -2 to +2 in the 4AY1 structure (Figure 3.3.12B, right panel).
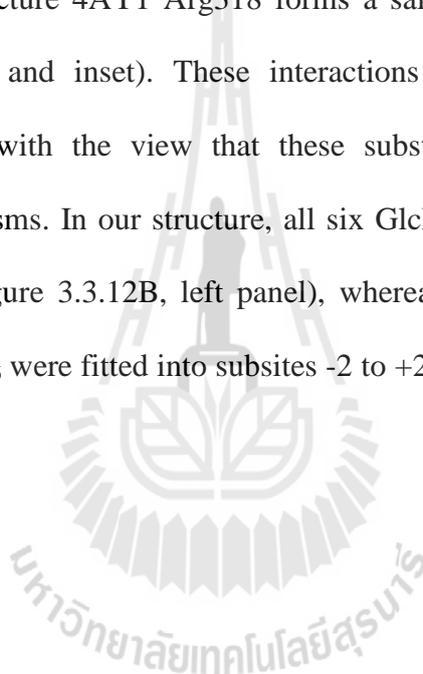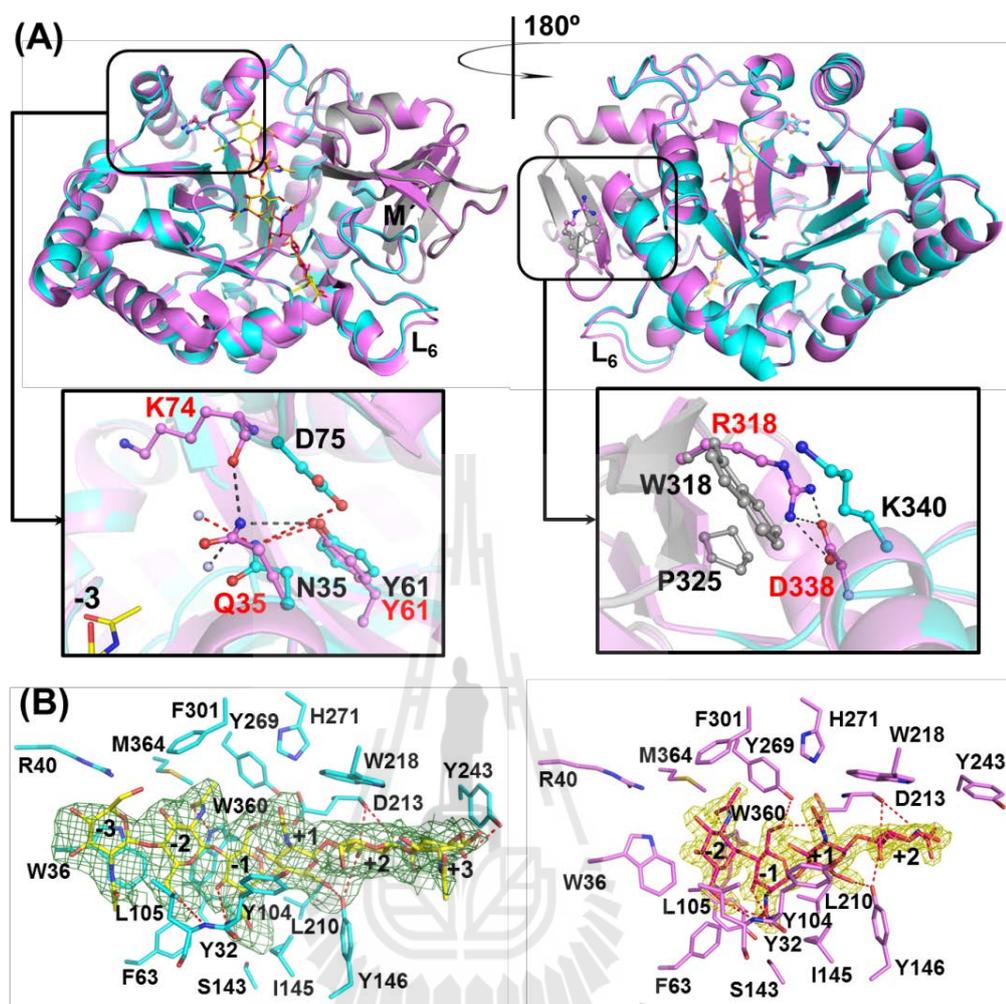
Figure 3.3.12 Structural comparison of two YKL-39 structures.

(A) Superposition of the currently-studied YKL-39 (PDB: 4P8X) and the previously-reported YKL-39 (PDB: 4AY1, magenta) in complex with $GlcNAc_6$. The differences between the two protein sequences at residues N35Q (left, inset) and W318R (right, inset) are displayed as ball and sticks. (B) $GlcNAc_6$ with map in the binding cleft of the currently-studied YKL-39 (green), while $GlcNAc_6$ with map in the binding cleft of the previously-reported YKL-39 (yellow). Color: carbon in the 4P8X structure is displayed in cyan, in the 4AY1 structure in magenta. For both structures, nitrogen atoms are in blue and oxygen atoms in red. Both $2F_o$-F_c omit maps are contoured at 1σ.

The superimposition of the YKL-39 structure to the available the crystal structure of human chitinase (CHIT1 and AMCase) and chitinase-like protein YKL-40 gave RMSD of 0.59, 0.58 and 0.69 Å, respectively, as list in Table 3.3.7. Figure 3.3.12 shows the structural superimposition of the structures of these proteins. Their overall the structures are high similarity in part of the $(\beta/\alpha)_8$ TIM barrel domain structures. Nevertheless, a few significant differences may explain divergent binding features between YKL-39 and the other two homologues. The most obvious variance is at subsite -4. In CHIT1 and YKL-40, this subsite has Tyr 34 as the key residue forming a hydrophobic stacking interaction against -4GlcNAc. This Tyr residue is missing in YKL-39, and the closest residue, Asp39, would be unlikely to make productive contact with a GlcNAc ring (Table 3.3.8). As a result, no affinity of binding at subsite -4 is likely for YKL-39. Another considerable difference is observed at subsite +3. In YKL-40, this subsite comprises Trp212 that directly stacks against the plane of the subsite +3GlcNAc. In contrast, Tyr243 in YKL-39 forms a hydrogen bond of a distance of 3.1 Å with the equatorial C6-OH group of the subsite +3GlcNAc, suggesting that the interaction at this position is weaker than that in YKL-40. The subsite -3GlcNAc unit at the non-reducing end of the GlcNAc$_6$ in YKL-39 (Figure 3.3.13B) is well defined and interacts with Trp36, whereas the subsite +3GlcNAc unit is more flexible, adopting a more open position than the GlcNAc ring that forms stacking with Trp212 at subsite +3 in the YKL-40 structure (Figure 3.3.13C). Weak affinity at the reducing end of the sugar chain is well characterized by high B-factor for the subsite +3GlcNAc (81.64 Å^2), reflecting high flexibility due to loose fitting at this location. Within the YKL-40 structure, the chitosugar occupies two positions between subsites -4 to +2 (Figure 3.3.13C, green) and subsites -3 to +3

(Figure 3.3.13C, grey) due to its more extended binding cleft. The residue at the homologous position to the catalytic residue Glu140 in the human chitinase (CHIT1) at bottom of subsite -1 is substituted by Ile145 in YKL-39 and Leu140 in YKL-40 (Figure 3.3.13A). Thus, neither YKL-39 nor YKL-40 have catalytic activity, although the overall feature of the bent conformation of the interacting sugar, suitable for being cleaved by normal chitinases, is maintained in the CLPs (Fusetti *et al.*, 2003; Fusetti *et al.*, 2002; Meng *et al.*, 2010; Songsiriritthigul *et al.*, 2008; van Aalten *et al.*, 2001).
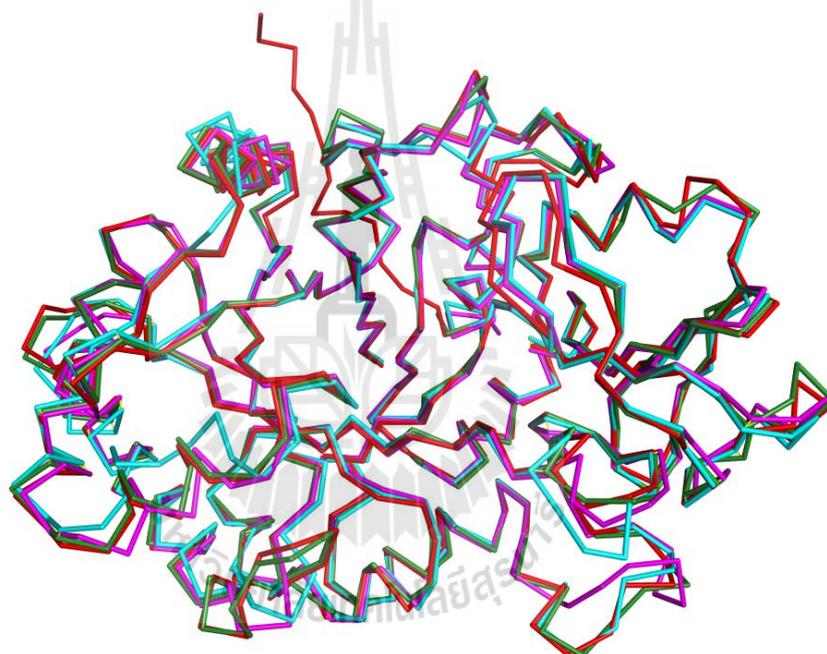
Figure 3.3.13 Ribbon representation of YKL-39, superimposed with other GH-18 human chitinases and CLPs

The crystal structure of YKL-39 is colored in cyan, while human chitinase, CHIT1 (PDB code: 1LG1) and AMCase (PDB code: 3FXY) are colored green and red, respectively. Chitinase-like protein YKL-40 (PDB code: 1HJW) is indicated by pink color.

Table 3.3.7 Root-mean-square deviation between the structures of complex YKL-39/GlcNAc$_6$ and other human chitinases and chitinase-like proteins.

Human chitinases and CLPs (PDB code)	RMSD	Sequence identity (%)	Number of C$_\alpha$
CHIT1 (1LG1)	0.59	55.2	296
AMCase (3FXY)	0.58	49	288
YKL-40 (1HJW)	0.69	51.3	292

YKL-39 and YKL-40 are closely related in size and sequence. Structural comparison also shows that the overall of the structures of both proteins are similar with a few significant differences in their chitin binding grooves (Figure 3.3.16A). For binding of GlcNAc$_4$ to YKL-39, the $2F_o$ - F_c Fourier maps obtained from the refined model revealed the electron density map covering subsites -3 to +1. This is in contrast to binding features observed with YKL-40 and chitinase B (ChiB), where GlcNAc$_4$ preferably occupies the central part of the binding cleft spanning from subsites -2 to +2 (Fusetti *et al*., 2003 and Gennady *et al*., 2004).
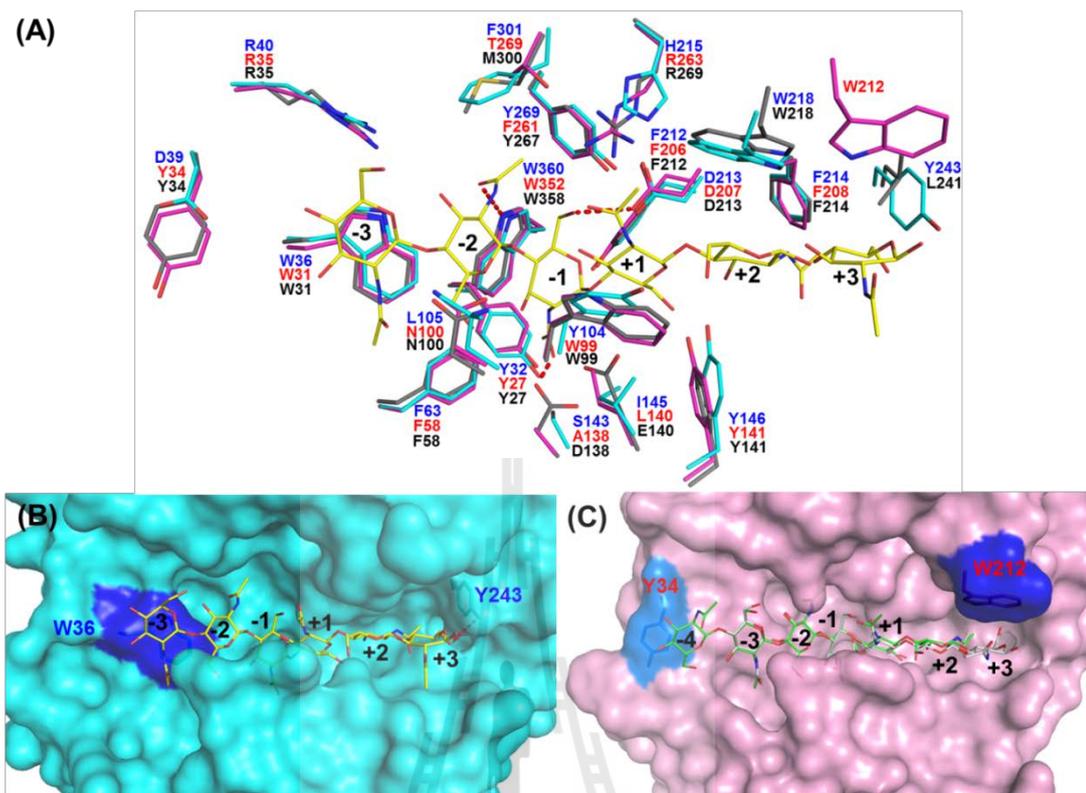
Figure 3.3.14 Superimposition of YKL-39 with its human GH-18 homologues.

(A) The amino acid residues of YKL-39 (cyan), YKL-40 (margenta), and CHIT1 (gray) that participate in sugar binding are shown. For the sake of clarity, only GlcNAc$_6$ from the YKL-39 structure is presented (yellow). (B) Surface representation and of the sugar-binding cleft of YKL-39, showing chitohexaose occupying subsites -3 to +3. However, the interaction at subsite +3 appears weak. Trp36 is highlighted in dark blue and Tyr234 is labeled. (C) Surface representation of the sugar-binding cleft of YKL-40, showing six GlcNAc units of the GlcNAc$_8$ occupying subsites -3 to +3 (PDB: 1NWU, green) or subsites -4 to +2 (PDB: 1HJW, gray). Tyr34 is highlighted in cyan and Trp212 in dark blue.

Table 3.3.8 A summary of the interactions between GlcNAc$_6$ and the binding residues in each subsites of YKL-39, YKL-40 and CHIT1.

Subsite / Protein	-4	-3	-2	-1	+1	+2	+3
YKL-39 (PDB: 4P8X)		**W36**	***W360***	**W360**		**W218**	
			L105	*D213,Y104*	*Y146*	*Y146, D213*	*Y243*
		R40	F63, F301, M364	Y32, S143, I145, Y269	Y104, L210, F212, H217	F214	
YKL-40 (PDB:1HJW, 1NWT)	**Y34**	**W31**	***W352***	**W352**	**W99**		**W212**
	E70	*N100, E290*	*N100, E290*	*D207, R263*	*Y141, R263*	*Y141*	
		R35, W69	Y27, F58, T269	Y27, A138, L140, Y206, F261		F208	
CHIT1 (PDB: 1LG1)	**Y34**	**W31**	***W358***	**W358**	**W99**	**W218**	
			N100, E297	*D138, E140*	*Y141, R269*		
				D213, R269,Y212			
		R35	Y27, F58, M300	Y267		F214	

Bold: stacking interaction residues

Bold-and-underline: recognition residues

Bold-and-italic: hydrogen bonding residues

3.4 Protein-ligand binding study

Human YKL-39 is known as a member of GH-18 glycosyl hydrolases that acts as a chitooligosaccharide binding protein (Schimpl *et al.*, 2012). A previous study by Schimpl *et al*. (Schimpl *et al.*, 2012) showed the 3D structure of YKL-39 in complex with chitin fragment. However, there are no direct experimental data showing the free energy changes associated with chitooligosaccharide binding to this protein. This section described binding thermodynamics of YKL-39 with its sugar ligand determined by intrinsic tryptophan fluorescence titration assay and isothermal titration calorimetry (ITC). ITC particularly provides a direct measure of thermodynamics parameters, including the stoichiometry of the interaction (*n*), the association constant (K_a), the free energy change (ΔG), the enthalpy change (ΔH) and the entropy change (ΔS). The data obtained from binding study and structural investigation provide the mechanistic ideas about how various chitooligosaccharides interacts at atomic levels with their protein partner and the energetic contributions to binding.

3.4.1   Protein-ligand binding study by intrinsic tryptophan fluorescence spectroscopy

Intrinsic fluorescence titration technique was used to evaluate the binding affinity of various chitin fragments in range of $GlcNAc_2$ to $GlcNAc_6$ towards the human YKL-39 protein. In agreement with the data obtained from the dose-response curves, the titration of chitin oligosaccharides ($GlcNAc_2$-$GlcNAc_6$) using various concentrations to 10 µg of YKL-39 resulted in a progressive increases in the fluorescence intensity, with no shift in the λ max and present changes in fluorescence intensities over a range of emission wavelengths of 330-400 nm as shown in Figure

3.4.1 to 3.4.5. Such changes reflect interaction of tryptophan residues with chitooligosaccharides of various lengths: $GlcNAc_2$ (Figure 3.4.1); $GlcNAc_3$ (Figure 3.4.2); $GlcNAc_4$ (Figure 3.4.3); $GlcNAc_5$ (Figure 3.4.4), and $GlcNAc_6$ (Figure 3.4.5), respectively. As shown in an inset of the figures, transformation of the fluorescence data as described in Section 2.9.2, yielding a plot of relative intensities versus sugar concentrations gave a good fit with a single-site binding model. The relative intensities also reached saturation at certain concentrations of sugar used for titration, giving an idea that sugar reached specifically to the protein with a ratio of 1:1. From the plot, the dissociation constant (K_d) values for each corresponding sugar could be obtained from the concentration of ligand (µM) at half saturation and the change in the free energy (ΔG, Kcal.mol^{-1}) was estimated and both values are shown in Table 3.4.1. The K_d values indicated that the long chains of chitooligosaccharides, including $GlcNAc_5$ and $GlcNAc_6$, reacted with YKL-39 with the highest affinity with K_d of 0.7 and 0.4 µM, respectively, followed by $GlcNAc_4$ (4.4 µM), $GlcNAc_3$ (183 µM), and $GlcNAc_2$ (216 µM), respectively. A decrease in K_d values with the increasing lengths of the chitin oligosaccharides suggest that the binding affinity depends on the length of the chitooligosaccharide. These results agree well with to the binding energy, which increased as the chain lengths of the chitosugar increased.
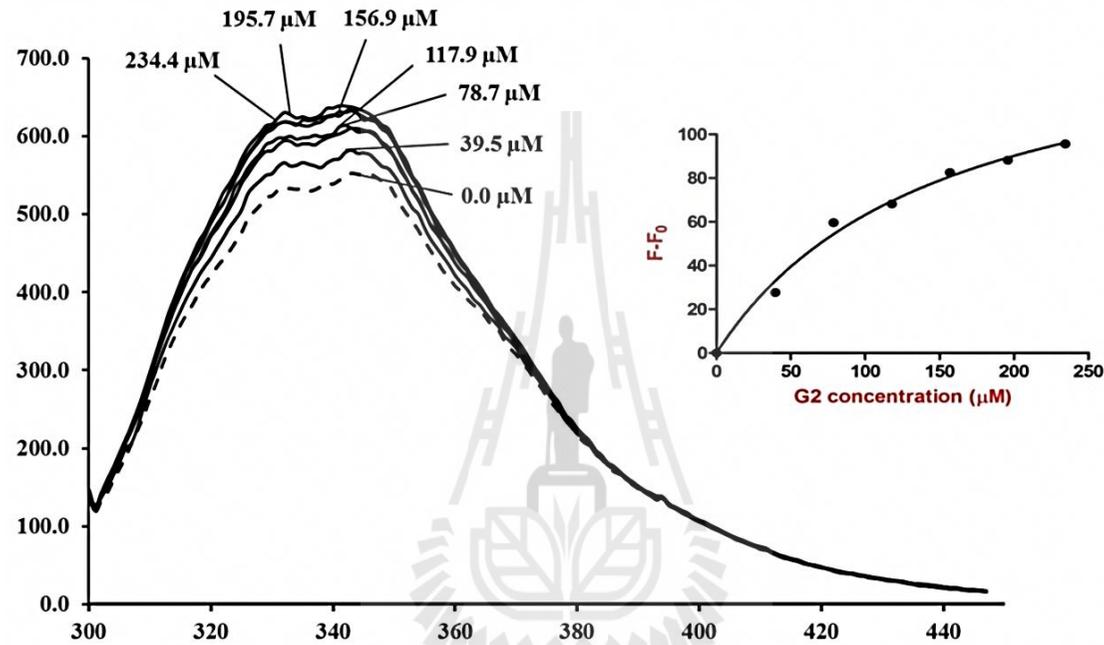
Figure 3.4.1 Binding of GlcNAc$_2$ to YKL-39 as studies by intrinsic fluorescence spectroscopy.

The purified YKL-39 (10 µg) was titrated with GlcNAc$_2$. An aliquot of 2.5 µL of GlcNAc$_2$ prepared as a 10 mM stock solution was added to the protein. The emission spectra were collected from 300-450 nm with an excitation wavelength of 295 nm. The dashed line indicates the protein solution without ligand added. An inset presents the relative fluorescence intensities (F-F_0) plotted against sugar concentrations. The binding curve was fitted to a single-binding site model.
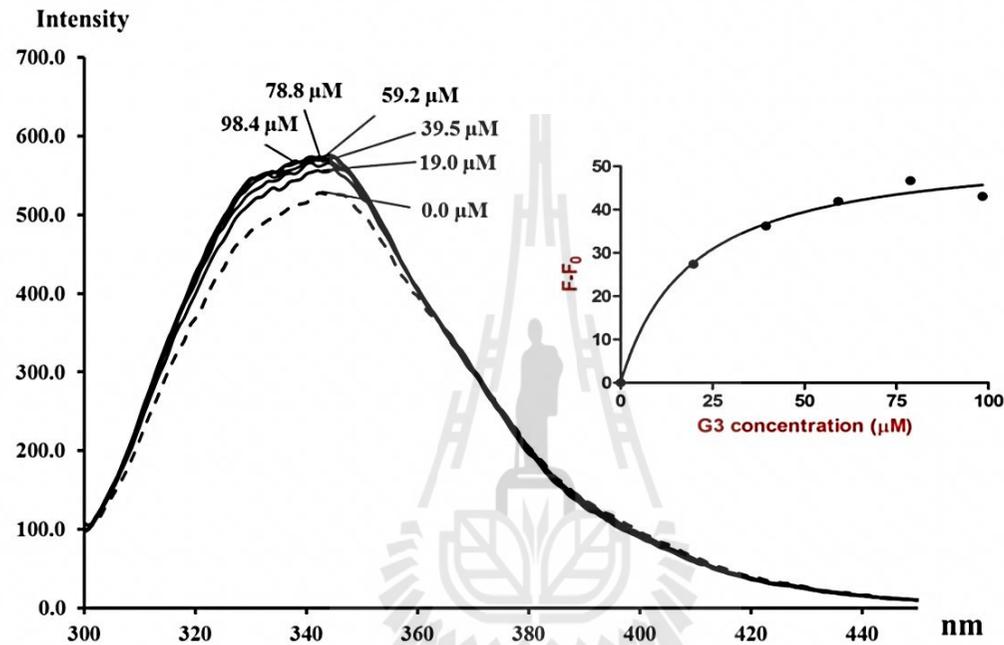
Figure 3.4.2 Binding of GlcNAc$_3$ to YKL-39 as studies by intrinsic fluorescence spectroscopy.

The purified YKL-39 (10 µg) was titrated with GlcNAc$_3$. An aliquot of 1µL of GlcNAc$_3$ prepared as a 10 mM stock solution was added to the protein. The emission spectra were collected from 300-450 nm with an excitation wavelength of 295 nm. The dashed line indicates the protein solution without ligand added. An inset presents the relative fluorescence intensities (F-F_0) plotted against sugar concentrations. The binding curve was fitted to a single-binding site model.
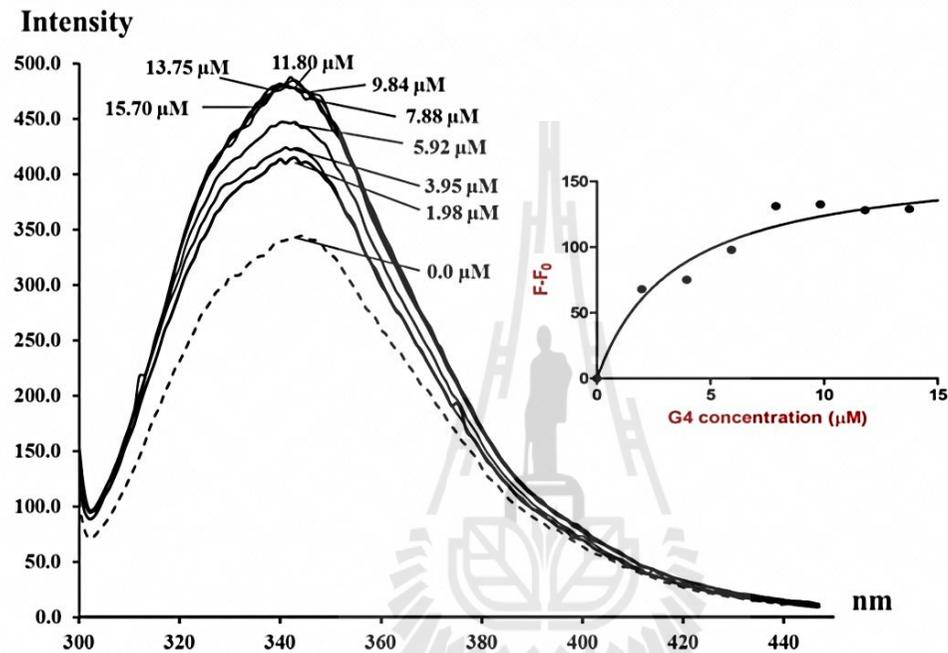
Figure 3.4.3   Binding of GlcNAc$_4$ to YKL-39 as studies by intrinsic fluorescence spectroscopy.

The purified YKL-39 (10 μg) was titrated with GlcNAc$_4$. An aliquot of 1 μL of GlcNAc$_4$ prepared as a 10 mM stock solution was added to the protein. The emission spectra were collected from 300-450 nm with an excitation wavelength of 295 nm. The dashed line indicates the protein solution without ligand added. An inset presents the relative fluorescence intensities (F-F_0) plotted against sugar concentrations. The binding curve was fitted to a single-binding site model.
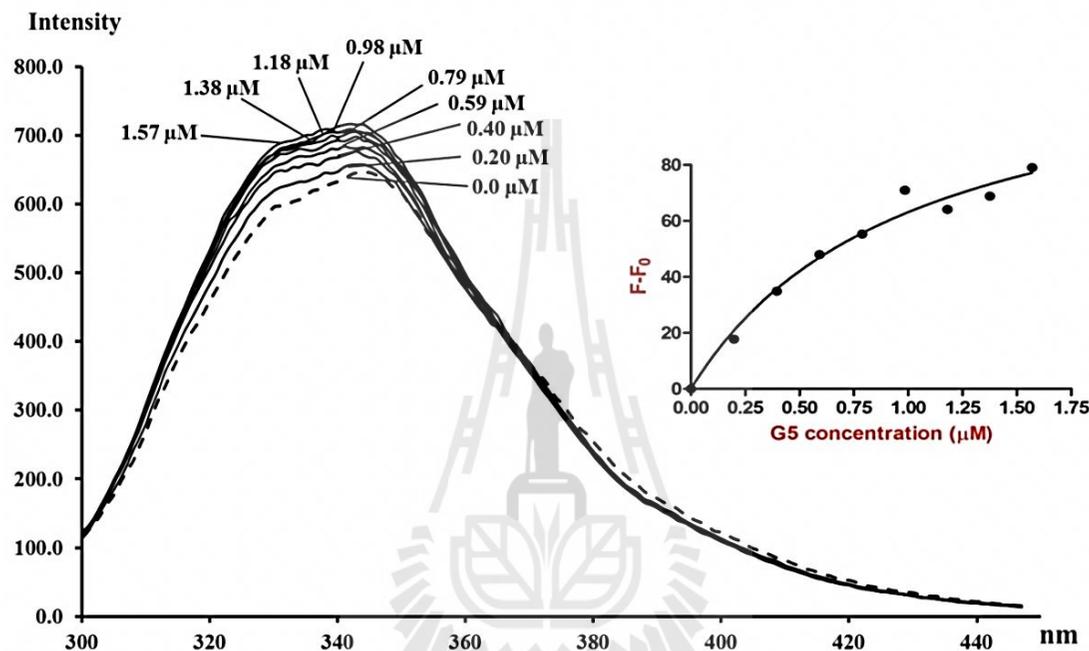
Figure 3.4.4 Binding of GlcNAc$_5$ to YKL-39 as studies by intrinsic fluorescence spectroscopy.

The purified YKL-39 (10 µg) was titrated with GlcNAc$_5$. An aliquot of 1 µL of GlcNAc$_5$ prepared as a 0.1 mM stock solution was added to the protein. The emission spectra were collected from 300-450 nm with an excitation wavelength of 295 nm. The dashed line indicates the protein solution without ligand added. An inset presents the relative fluorescence intensities (F-F_0) plotted against sugar concentrations. The binding curve was fitted to a single-binding site model.
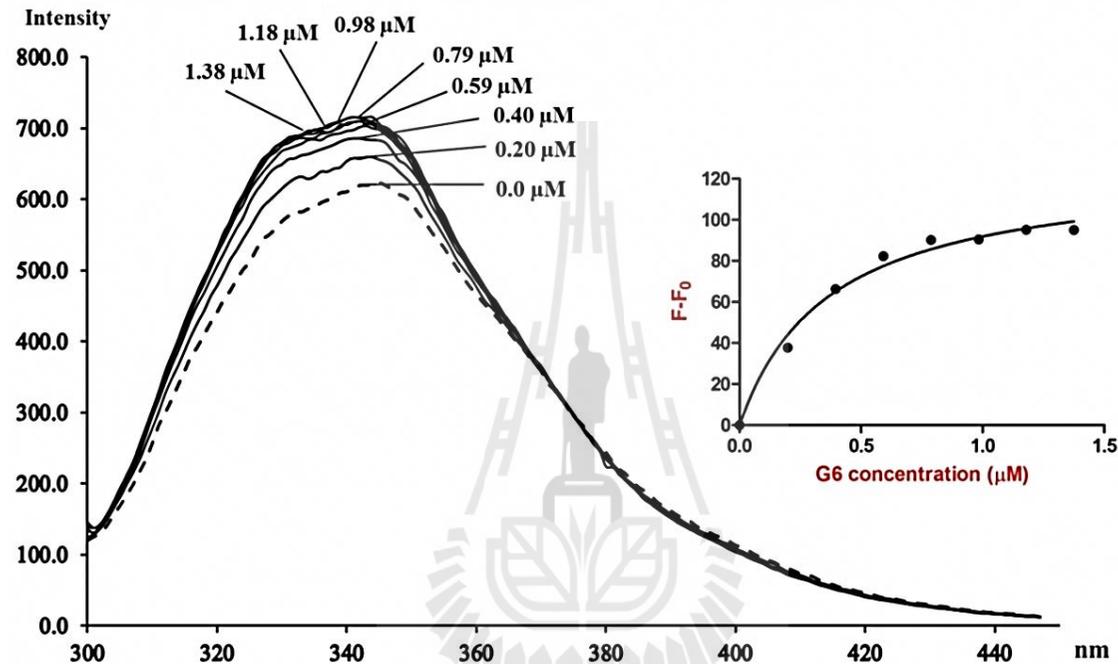
Figure 3.4.5 Binding of GlcNAc$_6$ to YKL-39 as studies by intrinsic fluorescence spectroscopy.

The purified YKL-39 (10 µg) was titrated with GlcNAc$_6$. An aliquot of 1 µL of GlcNAc$_6$ prepared as a 0.1 mM stock solution was added to the protein. The emission spectra were collected from 300-450 nm with an excitation wavelength of 295 nm. The dashed line indicates the protein solution without ligand added. An inset presents the relative fluorescence intensities (F-F_0) plotted against sugar concentrations. The binding curve was fitted to a single-binding site model.

Table 3.4.1 The equilibrium dissociation constant (K_d) and $\Delta G_{binding}$ values derived from the binding curves obtained from intrinsic fluorescence study.

Chitooligosaccharides	K_d (µM)	$\Delta G_{binding}$ (kcal/mol)
GlcNAc$_2$	216 ± 42	-5.0
GlcNAc$_3$	183 ± 47	-5.1
GlcNAc$_4$	4.4 ± 1.9	-7.3
GlcNAc$_5$	0.7 ± 0.3	-8.4
GlcNAc$_6$	0.4 ± 0.0	-8.8

3.4.2 Thermodynamics of binding by Isothermal titration calorimetry (ITC)

To gain an insight into the mechanism of binding, the binding kinetics of YKL-39 against various long chains of chitin oligosaccharides (GlcNAc$_2$ to GlcNAc$_6$) was investigated by ITC. This technique measures heat released or absorbed during binding events, providing information about binding in a single experiment (Agnieszka, 2011; Bissantz, Kuhn, and Stahl, 2010; Leavitt and Freire, 2001; Perozzo, Folkers, and Scapozza, 2004; Pierce, Raman, and Nall, 1999).

3.4.2.1 Binding of GlcNAc$_2$ to YKL-39

For ITC experiments to study binding of GlcNAc$_2$ to YKL-39, the 4 µM stock solution of GlcNAc$_2$ was added in discrete injections into 15 µM YKL-39 in 20 mM potassium phosphate buffer, pH 8.0, at 25°C. These conditions gave the ITC thermograms (Figure 3.4.6A, upper panels) and binding isotherms with theoretical single-site binding curve fit (Figure 3.4.6A, lower panels). These data were

further analyzed, which yielded ΔH of -13.1±4.4 kcal/mol, -TΔS of 7.6±3.8 kcal/mol, ΔG of -5.0±0.01 kcal/mol, and K_d of 204±3 µM. These binding thermodynamic parameters are plotted against the number of sugar ring and presented in Figure 3.4.7B. It shows that the overall binding of GlcNAc$_2$ to YKL-39 protein is an exothermic and enthalpy driven reaction, due to the negative values of ΔH of −13.1±4.4 kcal/mol, and an unfavorable entropic term (-TΔS = 7.6±3.8 kcal/mol/K and ΔS = -27±14 cal/K.mol).

3.4.2.2 Binding of GlcNAc$_3$ to YKL-39

For binding of GlcNAc$_3$ to YKL-39, 15 µM YKL-39 was titrated with GlcNAc$_3$. Under the conditions given in the Material and Methods, the titration thermogram was obtained as shown in Figure 3.4.6B (upper panel), and the data fitted to a single-site binding model (Figure 3.4.6B, lower panel) yielded ΔH of -11.3±1.3 kcal/mol, -TΔS of 6.1±1.3 kcal/mol, ΔG of -5.3±0.01 kcal/mol, and K_d of 142±8 µM. Figure 3.4.7B show that binding of GlcNAc$_3$ to YKL-39 was also driven by enthalpy, with an unfavorable entropic term (-TΔS = 6.1±1.3 kcal/mol and ΔS = -20.4±4.5 cal/K.mol.

3.4.2.3 Binding of GlcNAc$_4$ to YKL-39

For binding of GlcNAc$_4$ to YKL-39, 30 µM YKL-39 was titrated with GlcNAc$_4$. The titration thermogram was obtained as shown in Figure 3.4.6C (upper panel), and the data fitted to a single-site binding model (Figure 3.4.6C, lower panel) yielded ΔH of -17.0±2.0 kcal/mol, -TΔS of 9.1±1.8 kcal/mol, ΔG of -7.7±0.02

kcal/mol, and K_d of 1.7±0.7 µM. The binding of GlcNAc$_4$ to YKL-39 was also driven by enthalpy, with an unfavorable entropic term (-TΔS = 9.1±1.8 kcal/mol and ΔS = -30.5±5.9 cal/K.mol.

3.4.2.4 Binding of GlcNAc$_5$ to YKL-39

For binding of GlcNAc$_5$ to YKL-39, 10 µM YKL-39 was titrated with GlcNAc$_5$. The titration thermogram was obtained as shown in Figure 3.4.6D (upper panel), and the data fitted to a single-site binding model (Figure 3.4.6D, lower panel) yielded ΔH of -17.2±3.7 kcal/mol, -TΔS of 7.2±3.9 kcal/mol, ΔG of -10.0±0.4 kcal/mol, and K_d of 0.06±0.04 µM. The binding of GlcNAc$_3$ to YKL-39 was also driven by enthalpy, with an unfavorable entropic term (-TΔS = 6.1±1.3 kcal/mol and ΔS = -24.3±13.3 cal/K.mol.

3.4.2.5 Binding of GlcNAc$_6$ to YKL-39

For binding of GlcNAc$_6$ to YKL-39, 10 µM YKL-39 was titrated with GlcNAc$_6$. The titration thermogram was obtained as shown in Figure 3.4.6E (upper panel), and the data fitted to a single-site binding model (Figure 3.4.6E, lower panel) yielded ΔH of -12.2±1.5 kcal/mol, -TΔS of 2.0±1.5 kcal/mol, ΔG of -10.2±0.3 kcal/mol, and K_d of 0.04±0.02 µM. The binding of GlcNAc$_6$ to YKL-39 was also driven by enthalpy, with an unfavorable entropic term (-TΔS = 2.0±1.5 kcal/mol and ΔS = -6.9±5.2 cal/K.mol.
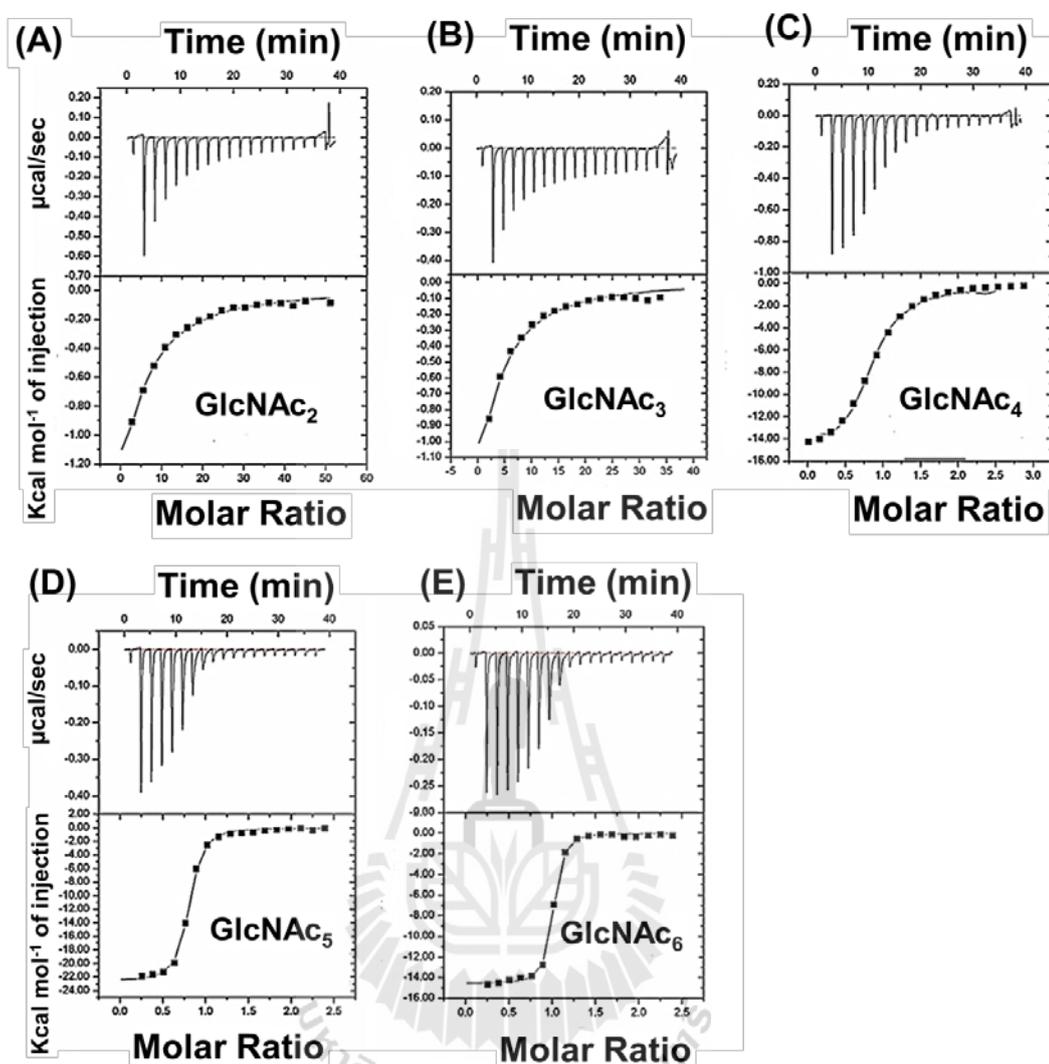
Figure 3.4.6 YKL-39 binding to various chitooligosaccharides.

Microcalorimetric titrations for the binding of chitooligosaccharide to YKL39 (upper panels), and binding isotherms with theoretical fittings (lower panels) obtained for the binding of GlcNAc$_2$ (A), GlcNAc$_3$ (B), GlcNAc$_4$ (C), GlcNAc$_5$ (D), and GlcNAc$_6$ (E) to YKL-39. Each ligand was added to protein solution in 20 mM potassium phosphate buffer, pH 8.0, at 25°C.

Table 3.4.2 Thermodynamic parameters of chitooligosaccharides binding to YKL-39.

Chitooligosaccharide	ITC						Wiseman c-value
	$K_d{}^a$	ΔH^b	ΔS^c	- $T\Delta S^b$	ΔG^b	n	
GlcNAc$_2$	204 ± 3	-13.1 ± 4.4	-27.0 ± 14.8	8.0 ± 4.4	-5.1 ± 0.01	1.0±0	0.1
GlcNAc$_3$	142 ± 8	-11.3 ± 1.3	-20.4 ± 4.5	6.1 ± 1.3	-5.3 ± 0.01	1.0±0	0.1
GlcNAc$_4$	1.7 ± 0.7	-17.0 ± 2.0	-30.5 ± 5.9	9.1 ± 1.8	-7.8 ± 0.2	0.9±0.1	18
GlcNAc$_5$	0.06 ± 0.04	-17.2 ± 3.7	-24.3 ± 13.3	7.2 ± 3.9	-10.0 ± 0.4	0.9±0.1	167
GlcNAc$_6$	0.04 ± 0.02	-12.2 ± 1.5	-6.9 ± 5.2	2.0 ± 1.5	-10.2 ± 0.3	1.1±0.1	250

a µM.

b kcal/mol.

c cal/K mol.
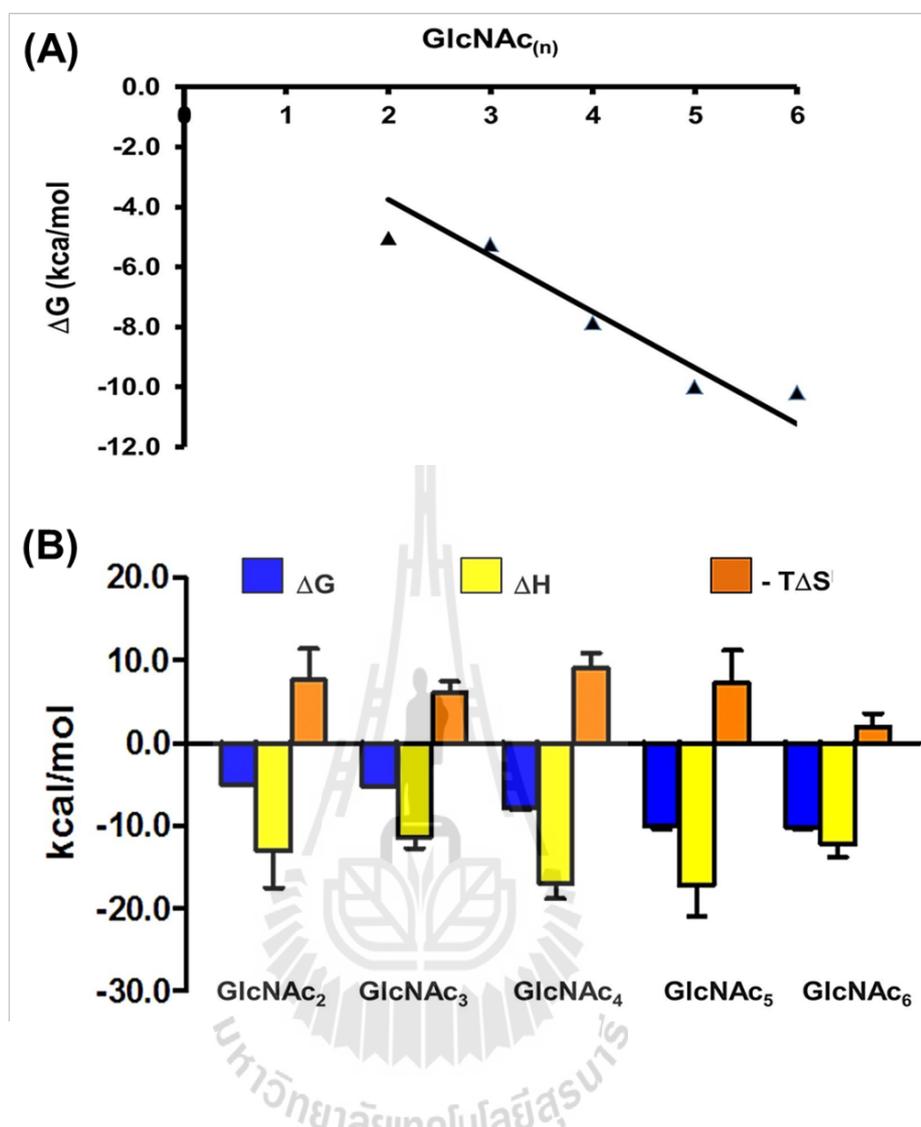
Figure 3.4.7 Thermodynamic parameters plot.

(A) Free energy of binding of chitooligosaccharide to YKL-39 increased with the length. Free energy of binding of GlcNAc$_5$ and GlcNAc$_6$ are not significantly different, indicating that binding cleft of YKL-39 has five subsites. (B) The binding signature (free energy, binding enthalpy, and entropy factor) plotted for ligand binding events.

Table 3.4.3 Thermodynamic parameters of chitooligosaccharides binding to YKL-39.

Chitooligosaccharide	Intrinsic fluorescence	ITC				
	$K_d{}^a$	$K_d{}^a$	ΔH^b	ΔS^c	$- T\Delta S^b$	ΔG^b
GlcNAc$_2$	216 ± 42	204 ± 3	-13.1 ± 4.4	-27.0 ± 14.8	8.0 ± 4.4	-5.1 ± 0.0
GlcNAc$_3$	183 ± 47	142 ± 8	-11.3 ± 1.3	-20.4 ± 4.5	6.1 ± 1.3	-5.3 ± 0.01
GlcNAc$_4$	4.4 ± 1.9	1.7 ± 0.7	-17.0 ± 2.0	-30.5 ± 5.9	9.1 ± 1.8	-7.8 ± 0.2
GlcNAc$_5$	0.7 ± 0.3	0.06 ± 0.04	-17.2 ± 3.7	-24.3 ± 13.3	7.2 ± 3.9	-10.0 ± 0.4
GlcNAc$_6$	0.4 ± 0.0	0.04 ± 0.02	-12.2 ± 1.5	-6.9 ± 5.2	2.0 ± 1.5	-10.2 ± 0.3

a µM.

b kcal/mol.

c cal/K mol.

CHAPTER IV

DISCUSSION

4.1 Molecular cloning, expression and purification of human YKL-39

The aim of the numerous protein characterization studies is improvement of techniques for molecular cloning, protein expression, purification, and protein functional and characterization methods. Generally, the most widely used expression host is *E. coli* due to it being a relatively simple and well characterized system capable of producing large quantities of soluble protein in a short period of time. In this study, we employed the nucleotide sequence of *CHI3L2* cDNA in the GenBank database (www.ncbi.nlm.nih.gov/genbank) to design a set of oligonucleotides, and then amplify the full-length *CHI3L2* gene. The PCR fragment of the *CHI3L2* gene was directly cloned into different expression vectors, and then expression from each vector was tested with the attempt to obtain a soluble YKL-39 in an *E. coli* system. Initially, the *CHI3L2* gene encoding YKL-39 protein was successfully cloned into pGEX-6P-1 expression vector, generating a construct called pGEX-6P-1/*CHI3L2*. This vector was designed to express a fusion protein with Glutathione S-Transferase (GST) at the *N*-terminus, which contains a thrombin cleavage site for deletion of the fusion protein tag. Unfortunately, the fusion proteins were mainly expressed as inclusion bodies from all the conditions tested. An alternative system was tried. Expression of YKL-39 as a thioredoxin fusion protein from the pET32a(+) vector was carried out. A construct called pET32a(+)/*CHI3L2* was made suitable for protein

expression in *E. coli* BL21 (DE3) host cells. The presence of a fusion tag at the upstream of the YKL-39 insert contains three desired characteristics: a thioredoxin (Trx) fusion tag, a hexahistidine tag, and an enterokinase cleavage site. The Trx fragment was fused to the protein to enhance protein solubility, whereas the hexahistidine residues facilitated protein purification, and the enterokinase cleavage site that contains a five amino acid recognition sequence, DDDDK, helped to remove the fusion tag. In this construct, there are an extra seven amino acids A-M-A-D-I-G-S between the Trx removal site and the mature YKL-39. Both sense and anti-sense strands of the *CHI3L2* gene were confirmed by automated DNA sequencing. All of identification assay confirmed that the pET32a(+)/*CHI3L2* construct had a correct structure, the insert *CHI3L2* gene did not introduce any reading frame shifting. The deduced amino acid sequence of YKL-39 was found to be almost identical to the 390-aa sequence of *CHI3L2* isoform1 (identifier: Q15782-4) reported in the UniProtKB/Swiss-Prot database. The exception is that a single amino acid residue in our sequence at position 318 had Trp instead of Arg. The discrepancy of this amino acid has been suggested previously to be a natural variation that occurs due to a single nucleotide polymorphism (The MGC Project Team, 2004). The deduced amino acid sequence of human YKL-39 was compared with that of YKL-40 (CHI3L1, P36222), acidic mammalian chitinase (AMCase, Q9BZP6-1), human chitotriosidase (CHIT1, Q13231), human oviduction (oviduction, Q86YN0), and stabilin-1 isoform 1(SI-CLP, Q9NY15). The YKL-39 sequence is closest to CHIT1 (55.2% identity), followed by YKL-40 (51.3%). On the other hand, YKL-39 showed lowest identity to SI-CLP (17.5%). The amino acid residues S143 and I145 located at the end of DxxDxDxE motif are reported to suppress chitinase activity. In true chitinases like AMCase

(Olland *et al*., 2009) and CHIT1 (Fusetti *et al*., 2002), these two residues are found as Asp and Glu instead. Double mutations of S143 to D143 and I145 to E145 in YKL-39 created pseudo-chitinase activity for this protein (Schimpl *et al*., 2012).

Optimization for high level of YKL-39 expression was investigated. The effects of various IPTG concentrations, induction temperatures, and times were tested. High expression level of the soluble YKL-39 was found to be under the following conditions: 0.5 mM IPTG concentration, 25°C and 16 h of induction temperature and time. The fusion protein containing His$_6$ tag was initially purified by Ni-NTA resin affinity column, followed by Sepharose Q HP anion exchange chromatography to eliminate the contaminating proteins. After digestion by enterokinase, the Trx-His$_6$ fusion tag was completely removed from the protein solution by the second Ni-NTA resin step, followed by HiLoad 16/60 Superdex 200 prep grade gel filtration to improve the purity and homogeneous of protein solution. The high quality of homogeneous protein with >95% purity was obtained after purification and was appropriate to be used for production of polyclonal and monoclonal antibodies, and used for crystallization and protein-ligand binding studies.

4.2 Production of anti-YKL39 polyclonal and monoclonal antibodies

4.2.1 Production of anti-YKL39 polyclonal antibodies

The Ag:Ab binding curves indicated concentration dependence of anti-YKL39 antisera towards the recombinant YKL-39, whereas no signal was detected in the reaction of pre-immune serum. Secondary antibody treatment alone also demonstrated no reactivity to protein. This indicated that YKL-39 protein was able to stimulate immune response in rabbits to produce highly specific and strongly reactive

polyclonal antibodies against the protein of interest. Subsequently, the immunized serum was further purified with protein-A column. SDS-PAGE analysis showed that the purified anti-YKL39 antisera migrated behind the 97 kDa protein marker under non-reducing condition, while two proteins of about 55 kDa and 25 kDa were observed under denaturing condition. Together with the achievement of antibody purification by Protein-A agarose, the results confirmed that anti-YKL39 pAbs had the specific IgG isotype. Indirect ELISA assay showed that the raised antibodies were highly sensitive, reacting strongly with the YKL-39 protein even when only 5 ng of the purified antibody was used. The Ag:Ab binding curves showed concentration dependence of anti-YKL39 pAbs towards the recombinant YKL-39, but not reaction with other GH-18 human chitinases and CLPs, human YKL-40, AMCase, and bacterial chitinaseA, indicating a high specificity of the purified antibodies towards YKL-39 antigen. Western blot analysis showed that the anti-YKL39 pAbs recognized both fusion Trx tag and native form of YKL39, but did not recognize other proteins even proteins in human serum. These results indicated that the purified pAbs were highly selective for YKL-39.

The binding curve demonstrated hyperbolic saturation feature, yielding the binding affinity (K_a) of YKL-39 protein for the detected by anti-YKL39 antibody to be 2.4×10^7 M^{-1}. In addition, the signal of interaction with the linear range was found to be from 0.16 to 1.25 µg/mL of YKL-39 with $R^2 = 0.9999$. By Western blot analysis, the anti-YKL39 antibody generated a strong signal at 10 µg, and became faint when the amount of YKL-39 was decreased down to 0.31 µg of YKL-39. These results suggested that 10 ng of anti-YKL39 pAbs was able to detect YKL-39 antigen in range of 0.16 to 1.25 µg/mL by indirect ELISA and 0.3-10 µg by Western blot assay.

Optimization of the antigen coating and antibody concentration of these conditions can be used to develop a solid-phase immunoassay for the detection of YKL-39 in real sample of patients with progressive osteoarthritis.

4.2.2 Production of anti-YKL39 monoclonal antibody

Hybridoma technique is the most widely used to produce monoclonal antibodies. The myeloma tumor cells are fused with spleen cells that challenge to produce antibody against the specific antigen. The fusion cells, namely hybrinoma, can grow indefinitely and rapidly produce a large amount of antibodies in culture. In this study, we further employed the hybridoma technique (Kohler and Milstein, 1975; Tomita and Tsumoto, 2011) to raise monoclonal antibody specific for the *E. coli-*expressed YKL-39. The immunized mice with YKL-39 protein showed high titers of antibody response, indicating the success of induction of antibody production against YKL-39 protein. Spleen cells from the highest of immune response mouse were further fused with myeloma cells, following hybridoma technique, as described in Section 2.8.3. Hybrodomas were selected by cultivation fused cells in HAT medium (hypoxanthine-aminopterin-thymidine medium). This medium combined aminopterin, a drug that acts as a powerful block to the de novo nucleotide synthesis pathway, with hypoxanthine (a purine derivative) and thymidine (a deoxynucleoside), which are intermediates in DNA synthesis. Thus, unfused myeloma cells die, as they cannot produce nucleotides by the de novo or salvage pathways. Unfused spleen cells die as they have a short life span. In this way, only the hybridomas can survive (Holliday and Ho, 1998).

After ten days of cultivation, healthy hybridomas were observed in 477 wells. Screening for the positive hybridomas using indirect ELISA identified 71 hybridomas that strongly reacted with YKL-39. To obtain a single clone of hybrodomas, limiting dilution was investigated by separating hybridomas into the individual clones. Unfortunately, most of hybridomas lost their activity during cultivation. However, a positive single clones from the 1st limiting dilution were re-cloned by the 2nd round of limiting dilution until the stable single clone of hybridoma were obtained. Finally, two positive clones, namely 6H11 and 8H3, were obtained. Isotyping mapping clearly suggested that both monoclones belong to the IgM isotype. Because the 6H11 clone exhibited greater immunoactivity, it was then selected for large scale production.

Crude anti-YKL39 mAb was purified with a HiTrap IgM HP column (GE Health-Care). This column is made of 2-mercaptopyridine coupled to sepharose, thiophilic affinity medium, the binding affinity to ligand is promoted by water-structuring salts. The antibody: coupled ligand column interaction has been suggested to result from a combined electron donating and accepting action of the 2-mercaptopyridine ligand (Hutchens and Porath, 1987; Lihme and Heegaard, 1991).

Western blot analysis demonstrated that the antibody titer against recombinant YKL-39 had the strongest titer at a ratio of 10:5 (µg/µg) of mAb:Ag. Our newly raised mAb recognized both fused and bare forms of human YKL-39, but did not recognize other GH-18 homologs, suggesting that the anti-YKL39 mAb was highly specific for YKL-39. Detection of YKL-39 in Jurkat cells (a model for differentiation of T lymphocytes) and U937 cells (a model for the differentiation of monocytes and macrophages) by our mAb indicated that YKL-39 may be involved in both innate and adaptive immune systems in response to the likelihood of infection or tissue

inflammation. Dot blot assay showed that the monoclonal antibody was strongly active with the synovial fluid of an osteoarthritis patient, and human monocyte and T lymphocyte cell lines. Such results suggested that our monoclonal antibody was functionally active and exhibited cell type specificity. Taken together, both polyclonal and monoclonal antibodies were highly specific and active with various types of immunological assays, making them suitable for functional studies such as the identification of the YKl-39 regulated pathways involving cartilage tissue degeneration and development of an immunosensing tool for monitoring the progress of degenerative joint disease.

4.3 The structural determination of YKL-39 in the absence and presence of chitin fragments

4.3.1 The Crystallization of YKL-39

The appropriate conditions to crystallize YKL-39 were found under conditions, containing 27 to 31% (w/v) PEG 3350 and 0.2 to 0.3 M lithium sulfate in 0.1 M Bis Tris, pH 8.0, 25°C. The optimized conditions showed that high concentrations of precipitant resulted in a decrease in crystal size. To obtain structures in the presence chitin fragments, crystals of apo YKL-39 were soaked with the optimal concentration of each chitin fragment. Finally, four crystal structures of YKL-39 in the absence and presence of chitooligosaccharides, including ligand-free YKL-39 and in complex with $GlcNAc_2$, $GlcNAc_4$ and $GlcNAc_6$ were solved to resolutions of 2.48-1.53 Å.

4.3.2 The overall structure of YKL-39 in the presence and absence chitin fragments

Like GH-18 chitinases and CLPs, the overall structure of YKL-39 displayed two conserved domains, including the major $(\beta/\alpha)_8$ TIM barrel domain (Domain I) and a small insertion domain. The TIM barrel domains interact specifically with chitin oligosaccharides, while the exact function of the insertion domain is unknown (Fusetti *et al*., 2003; Fusetti *et al*., 2002; Meng *et al*., 2010; Schimpl *et al*., 2012; Songsiriritthigul *et al*., 2008; van Aalten *et al*., 2001). The structures of YKL-39 in the presence of chitin fragments revealed the binding surface of the TIM barrel domain to be a long, deep groove that is open at both ends to facilitate the accommodation of chitin fragments that extend out from the groove in both directions.

4.3.3 Chitooligosaccharide binding induces structural movements

Superimposition of the apo YKL-39 structure on the structures in complex with GlcNAc$_2$, GlcNAc$_4$, and GlcNAc$_6$, showed that the presence of the saccharide chain in the chitin binding cleft induces local conformation changes that are mainly observed in the loop regions surrounding the surface of the chitin binding cleft. The most observations were found at the center of the chitin-binding cleft. The movement of key residues moved the binding cleft boundary residues of the sugar bound protein around subsites -2 and -1, closer by 1.3 Å, to close contact with the GlcNAc rings. In contrast, the rotation of Tyr243 widened the binding cleft at this particular subsite (+3) by 1.9 Å, resulting in the subsites to be extended from +2 to +3. Taken together, these structural differences provide evidence that the binding of

chitooligosaccharides induces local structural changes that strengthen the interactions between the YKL-39 and that facilitated the accommodation of incoming ligand.

4.3.4 Specific interactions of YKL-39 with chitooligosaccharides

Examination of the substrate binding cleft of ligand-free YKL-39 and all sugar-bound YKL-39 structures reveals the binding behaviors of individual chitooligosaccharides. The strongly binding interaction was found at the center of the binding cleft and extended in both directions with weaker-affinity interactions. This was supported by $GlcNAc_2$ model that bound at the center covering subsite -2/-1. Two or three molecules of $GlcNAc_2$ could bind within the chitin-binding site of YKL-39 based on its size, but only a single bound GlcNAc2 was observed, suggesting that an incoming chitin chain would preferentially interact at the center of the binding groove. Where a longer chain chitooligosaccharide, four sugar rings $GlcNAc_4$ extended over subsites -3 to +1, and $GlcNAc_6$ bound at subsites -3 to +3. This affinity gradient is well supported by lower B-factors for -2GlcNAc (43.63 $Å^2$), -1GlcNAc (44.76 $Å^2$) and +1GlcNAc (44.75 $Å^2$) that are located at the center of the binding sites, indicating more rigid and tight interactions at the internal sites. Increases in B-factor for bound sugar were found to increase with the extending subsites. Especially, +3GlcNAc, showed the highest of B-factor, reflecting high flexibility of the sugar rings forming weak interactions with the binding residues at such subsites.

The protein-ligand binding interaction was considered with the structure of YKL-39 in complex with the longest chitin chain, $GlcNAc_6$, as model. The critical binding features are aromatic residues, which interact with the GlcNAc rings mainly

through hydrophobic interactions, and charged or polar side chain residues that form hydrogen bonding interactions with the saccharide chain. For hydrophobic interactions, three key residues, Trp36 at subsite -3, Trp360 at subsite -1, and Trp218 at +2, are found to stack with the plane of GlcNAc rings, resulting in the increasing strength of its binding affinity. The first two residues (Trp36 and Trp360) are completely conserved in other GH-18 homologues and act as key binding residues (Fusetti *et al*., 2003; Fusetti *et al*., 2002; Meng *et al*., 2010; Schimpl *et al*., 2012; Songsiriritthigul *et al*., 2008; van Aalten *et al*., 2001). The conserved aromatics Trp36, located at the edges of the binding cleft, marks the non-reducing end of long chain chitin fragment by stacking against the planes of the bound sugar at -3GlcNAc, while Trp218 marks the reducing end of chitin sugar, and is in the position suitable for binding with the subsite +2GlcNAc. For the subsite +3GlcNAc, no hydrophobic interaction was found. It is presumed that there is a weak binding at this subsite. This observation was confirmed by ITC results, which showed the insignificantly different values of binding affinity and binding energy for $GlcNAc_5$ and $GlcNAc_6$, suggesting insignificant differences in the binding strength for these two chitooligosaccharides. The electron density of YKL-39/$GlcNAc_6$ complex suggested that the sugar units at the central subsites are held by a larger number of weak interaction, including salt bridges, H-bond, and hydrophobic interactions. Trp360 seems to play a particular role in ligand recognition by making hydrogen bonds with the C7-O of -2GlcNAc. This residue is conserved in GH-18 chitinases and CLPs (Fusetti *et al*., 2003; Fusetti *et al*., 2002; Meng *et al*., 2010; Schimpl *et al*., 2012; Songsiriritthigul *et al*., 2008; van Aalten *et al*., 2001). Figure 4.1B displays a number of hydrogen bonds that help to stabilize the sugar/protein complexes. Such hydrogen bonds are formed either directly

between the sugar rings and the binding residues or mediated by water molecules. High densities of interactions are seen at subsites -2, -1, and +1, which are likely to determine the preferential binding at these sites.

4.3.5 Structural comparison with other human GH-18 members

Superimposition of the structure of YKL39:GlcNAc$_6$ complex obtained in this study on YKL39:GlcNAc$_6$ complex (4AY1) reported by Schimpl et al. (Schimpl *et al*., 2012) was found to be essentially identical, even though they are derived from crystals with different space groups. The two proteins differed in sequences at residues 35 and 318, resulting in dissimilarity of interactions around these residues. Although residue 35 in both structures is located near subsite -3, it makes no contact with the GlcNAc ring. Residue 318 lies in a region distant from binding site, Trp318 in native YKL-39 forms hydrophobic interactions with Pro325 and the stalk of Lys340, while Arg318 in structure 4AY1 forms a salt bridge with Asp338. In addition, all six GlcNAc rings could be fitted into subsites -3 to +3 in our structure, whereas only four GlcNAc rings of GlcNAc$_6$ were fitted into subsites -2 to +2 in the 4AY1 structure.

The sugar-binding grooves of YKL-39, CHIT1 (Fusetti *et al*., 2002), and YKL-40 (Fusetti *et al*., 2002; Houston *et al*., 2003) were compared and found to be very similar with a few significant differences in the chitin binding subsites. These explain divergent binding features between YKL-39 and other two homologues. The most different variation is at subsite -4. In CHIT1 and YKL-40, the sugar ring formed a hydrophobic stack against the key residue Tyr34, but this residue is missing in YKL-39, causing no affinity at subsite -4 in the YKL-39 structure (see Figure 3.3.13).

Another considerable difference was observed at subsite +3. In YKL-40, this subsite sugar stacks directly against the Trp212 residue. In contrast, Tyr243 in YKL-39 forms a hydrogen bond with the subsite +3GlcNAc, suggesting that the interaction at this position is weaker than that in YKL-40. Weak affinity at the reducing end of the sugar chain is suggested by a high B-factor for +3GlcNAc (81.64 $Å^2$), reflecting high flexibility due to loose fitting at this location. Within the YKL-40 structure, the chitosugar occupies two positions between subsites -4 to +2 and subsites -3 to +3 due to its more extended binding cleft. The catalytic residue Glu140 in CHIT1 at bottom of subsite -1 is clearly substituted by Ile145 in YKL-39 and Leu140 in YKL-40, making them lack chitinase activity.

The chitin-protein complex suggested that GlcNAc containing oligosaccharide could act as the physiological ligand for YKL-39. It has been explained recently that YKL-39 enhances colony forming activity, cell proliferation, and type II collagen expression in mouse chondrogenic cells. These data suggest that YKL-39 may be a novel growth and differentiation factor involved in regulating cartilage homeostasis (Miyatake *et al*., 2013). This report demonstrates that YKL-39 could participate in specific signaling processes and could function as a lectin able to recognize specific glycan structures present in mammalian tissues.
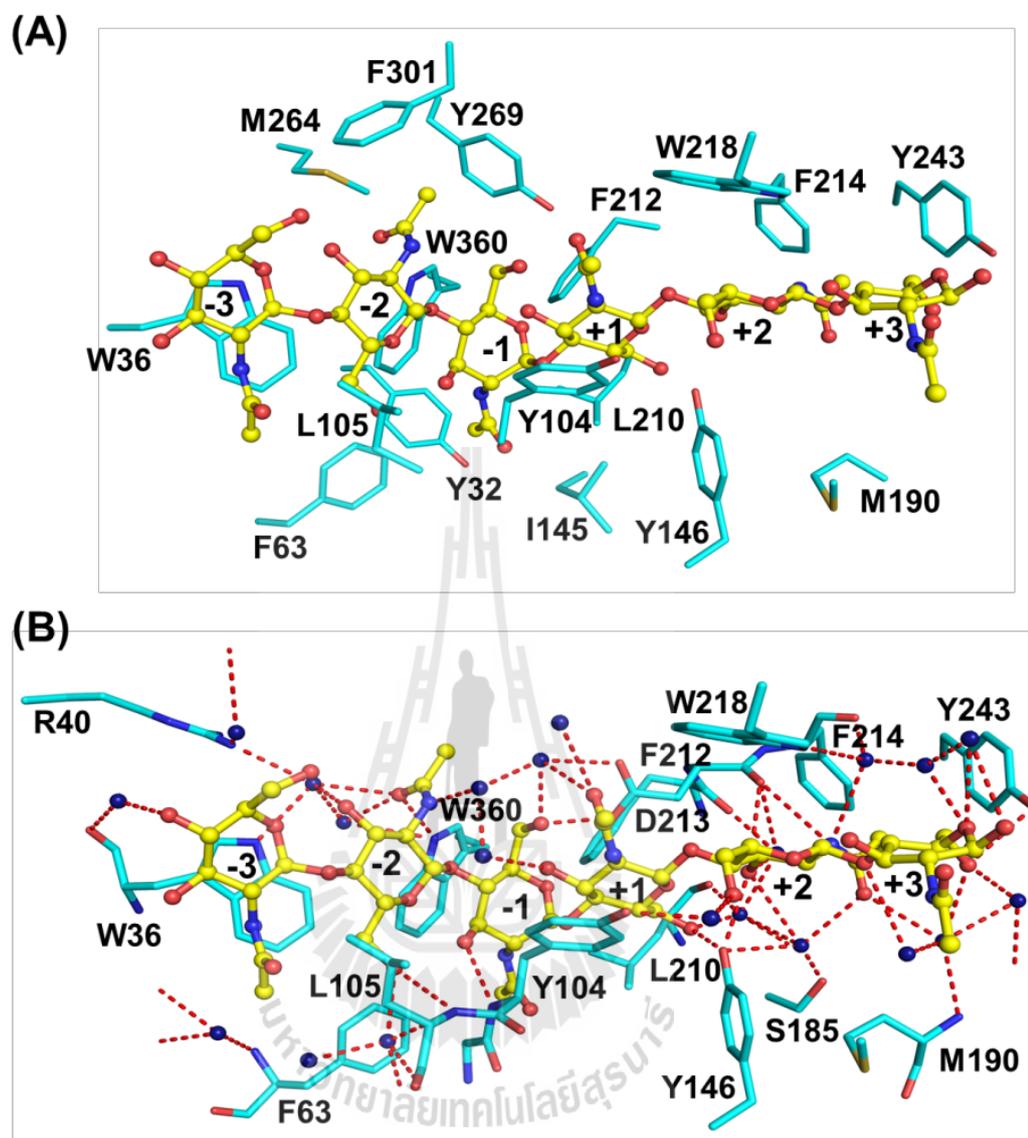
Figure 4.1 Specific interactions within the chitin binding cleft of YKL-39.

(A) The aromatic residues that are involved hydrophobic interactions with the sugar molecule at each subsite. (B) The amino acid residues and water molecules that mediate the formation of hydrogen bonding network. Water molecules are shown as blue diamonds and hydrogen bonds are shown as red dashed lines.

4.4 The binding thermodynamics of YKL-39 with chitin oligosaccharides

Isothermal titration calorimetry (ITC) and intrinsic tryptophan fluorescence were used to determine binding thermodynamics of YKL-39 towards various chitin fragments. For ITC measurements, YKL-39 was titrated with chitooligosaccharides until the binding isotherms were saturated, allowing the determination of binding affinities and thermodynamic parameters for YKL-39 (Agnieszka, 2011; Baranauskiene *et al*., 2009; Pierce *et al*., 1999). All data obtained from the binding reactions were fitted using a single-site binding model with calculated stoichiometry (n) of approximately 1.0, indicating that one molecule of the sugar interacts within the binding cleft of YKL-39. This is consistent with that found in the crystal structures, each of which shows only a single chitin chain, even $GlcNAc_2$, as described in Results (Section 4.3.4). Analysis of the ITC data gave the equilibrium binding constants (K_d), which indicated the binding affinities that clearly increased with increasing chitooligosaccharide length. These binding kinetics are consistent with the structural observations that $GlcNAc_2$ occupies the center of the protein binding cleft at subsites -2 and -1, whereas $GlcNAc_4$ and $GlcNAc_6$ stretches along the entire binding groove in both direction, having the most stable binding of the chitooligosaccharides studied. The binding affinity of $GlcNAc_5$ (0.06 μM) and $GlcNAc_6$ (0.04 μM) are insignificantly difference. These correspond to the structure of the YKL-39/$GlcNAc_6$ complex that suggests strong interactions from subsites -3 to +2. It shows hydrophobic stacking interaction by Trp36 at -3GlcNAc and Trp218 at +2GlcNAc, respectively, but no interaction was observed at +3GlcNAc. These binding characteristics also correlated with fluorescence spectroscopy measurements in which

the data fit well to the single-site binding model of a non-linear regression function as described in Section 3.4.1. The estimated K_d values are in the following order: $GlcNAc_6 \cong GlcNAc_5 > GlcNAc_4 > GlcNAc_3 > GlcNAc_2$.

Comparing the data from both the fluorescence measurements and ITC experiments found that the values of K_d obtained from fluorescence measurements are in general higher than those from ITC measurements, reflecting the different binding concepts of the two binding methods. ITC assay account for all interactions, while in the fluorescence assay, quenching tryptophan residues that are located within binding cleft is measured. Nevertheless, both show the difference in K_d values is small for short chain chitooligosaccharides, and becomes greater with increasing chain length and tighter binding. Both binding methods showed insignificantly different K_d for $GlcNAc_2$ (see Table 3.4.3). The results suggested that important key residue Trp360 that stacks against the face of two GlcNAc rings at subsites -2/-1 plays an exclusive role in the sugarprotein interactions at such central subsites. The enthalpy change (ΔH), the entropy change ($-T\Delta S$) and the free energy change (ΔG) obtained from the ITC experiments clearly demonstrate that all chitooligosaccharides bind to YKL-39 in spontaneous, exothermic reactions (see Table 3.4.2). As shown in Figure 3.4.7A, similar values of the binding energies estimated for $GlcNAc_5$ and $GlcNAc_6$ suggest insignificant differences in the binding strengths for these two chitooligosaccharides. The results of these kinetic analyses are consistent with the crystal structure showed strong binding covering subsites $(-3)(-2)(-1)(+1)(+2)$, instead of six subsites, as observed for its closely-related homologs CHIT1 and YKL-40 (Fusetti *et al*., 2003; Fusetti *et al*., 2002; Meng *et al*., 2010). Thermodynamic parameters revealed that all ligands bind to YKL-39 (see Figure 3.4.7B), were mainly driven by enthalpy, as

indicated by the dominant negative enthalpy change with unfavorable entropy. The dominant -ΔH reflected the strength of binding that was caused from hydrogen bonds and electrostatic interactions (Richard, 2012).

CHAPTER V

CONCLUSIONS

The *CHI3L2* gene encoding YKL-39 was successfully cloned into the pET32a(+) expression vector, yielding the construct named pET32a(+)/*CHI3L2*. Employing this vector system, the recombinant YKL-39 was expressed as a fusion protein that contains the thioredoxin (Trx) fragment, followed by six histidine residues (His$_6$) attached at the *N*-terminus of the YKL-39 polypeptide. Colony PCR, restriction enzyme digestion and DNA sequencing confirmed that the newly generated construct pET32a(+)/*CHI3L2* had a correct structure, and the *CHI3L2* gene insertion did not introduce any reading frame shift.

The recombinant YKL-39 polypeptide lacking its 26-aa signal sequence was successfully expressed in a soluble form with *E. coli* strain BL21(DE3) as host. The protein was expressed as a Trx-His$_6$/YKL-39 fusion protein with the total molecular weight of approximately 55 kDa. The expression level of the soluble YKL-39 was dependent on the IPTG concentration, temperature and time of induction, with the optimal expression conditions being 0.5 mM IPTG concentration at 25°C for 16 h.

For polyclonal and monoclonal antibody production, the Trx-His$_6$/YKL-39 fusion protein was purified by gravity flow Ni-NTA agarose chromatography, followed by Sepharose Q HP anion exchange chromatography, to eliminate the contaminating proteins for high purity fusion protein. After enterokinase digestion to remove the Trx/His$_6$ *N*-terminal tag fragment, the free tag YKL-39 was further

purified by a second round of Ni-NTA agarose column chromatography. The entire YKL-39 polypeptide had seven extra residues (A-M-A-D-I-G-S) preceding the YKL sequence. The protein was seen on SDS-PAGE gel at the molecular weight between 35 and 45 kDa, which corresponded to the predicted mass of 41.5 kDa. The YKL-39 was evaluated to be about 95% pure, with the overall yield of about 3.4 mg per liter of culture. For crystallization of YKL-39, a large amount of homogeneous protein is required. Therefore, after the thioredoxin fusion tag removal, the tag-free protein was purified by Ni-NTA agarose, followed by S200 gel filtration chromatography. The purified YKL-39 appeared to be a single protein with MW of about 40 kDa on SDS-PAGE and was evaluated to be >95% pure. This highly purified protein was considered appropriate for crystallization trials and protein-ligand binding studies.

For anti-YKL39 polyclonal antibody production, data obtained from indirect ELISA assay showed that YKL-39 protein was able to stimulate a response in the immunized rabbits to produce highly specific anti-YKL39 antisera. The Ag:Ab binding curves showed hyperbolic concentration dependence of anti-YKL39 antiserum towards the recombinant YKL-39, while pre-immune serum showed no reactivity to YKL-39. The results suggested that anti-YKL39 antiserum was specifically and strongly reactive towards the purified YKL-39 protein. The raised antisera were further purified via a protein A column. Two bands with molecular weights of 55 kDa and 25 kDa, representing the heavy and light chains of IgG, respectively, appeared on SDS-PAGE analysis, as expected. The specificity and sensitivity of the purified anti-YKL39 antisera were further characterized by indirect ELISA and Western blot immunoassay. Our raised polyclonal antibodies were found to be highly sensitive, reacting strongly with the YKL-39 protein even when only 5 ng

of the purified antibody was used. The Ag:Ab binding curves which gave the saturated hyperbolic feature also showed concentration dependence of anti-YKL39 antisera towards the recombinant YKL-39. To determine the specificity of anti-YKL39 pAbs, other GH-18 homologs YKL-40, AMCase, and chitinase A were used as immunogens. It was found that no significant increase in the signal was observed. The binding intensity of each antigen was further tested. It was found that the absorbance obtained for YKL-39 was 15-19 folds greater than that for human YKL-40, AMCase, and chitinase A. These further confirmed a high specificity of the purified antibodies towards YKL-39.

For anti-YKL39 monoclonal antibodies were generated by the hybridoma method. Spleen cells, which produced antibody against YKL-39, were fused with myeloma cells using 50% (v/v) PEG. Finally, two positive clones, namely 6H11 and 8H3, were obtained after a limiting dilution screen of about 200 hybridoma clones. Isotype mapping of these clones were determined by capture ELISA, which was found that both a single clones belonged to the IgM isotype. Crude anti-YKL39 mAb, which was obtained from cultured serum free media were further purified via a specific HiTrap IgM HP column. The purified IgM titers were further tested against recombinant YKL-39 by Western blot analysis. It revealed that the strongest titer was seen at a ratio of 10:5 (µg/µg) of mAb:Ag, while the slightest titer of the antibody that recognized the YKL-39 Ag was 2.5:5 (µg/µg). The antibodies were found to recognize both fused and free forms of human YKL-39, but did not recognize YKL-40, AMCase, and bacterial chitinase. The antibodies also did not cross-react with proteins in human serum. Based on these findings, it has been suggested that the anti-YKL39 mAb was highly specific for YKL-39. The anti-YKL39 mAb could react

with YKL-39 in synovial fluid of a patient with osteoarthritis and various human cell lines, as well as in the whole cell lysate of Jurkat cells. Such results suggested that our monoclonal antibody was functionally active and exhibited cell type specificity.

For structural determination of YKL-39 in the presence and absence chitin fragments, the suitable conditions for crystallization of YKL-39 were screened by the sitting drop technique. It was found that the precipitants, which were suitable to crystallize this protein, were PEG derivatives of the molecular weight from 1500 to 8000, containing the Li_2SO_4 and $Ca(OAc)_2$. The most suitable buffers were Bis Tris, citrate and imidazole at pH 5.5 and 8.0. After optimization using the hanging drop method with various concentrations of precipitants PEG 3350 and lithium sulfate in 0.1 M Bis Tris pH 8.0, the X-ray diffracting quality crystals of ligand- free YKL-39 were obtained under a condition of 27% (w/v) PEG 3350, 0.2 M lithium sulfate in 0.1 M Bis Tris, pH 8.0. Ligand-free crystals were soaked with chitin fragments of $GlcNAc_2$ to $GlcNAc_6$. Finally, four crystal structures of YKL-39 in the free form and in complex with $GlcNAc_2$, $GlcNAc_4$ and $GlcNAc_6$ were solved to resolutions of 2.48-1.53 Å. All the crystals belong to space group $P4_12_12$ with one molecule in an asymmetric unit and unit cell dimensions of approximately 71x71x40 Å.

The overall structure of YKL-39 contains the major $(\beta/\alpha)_8$ TIM barrel domain with the small insertion domain between stand B_7 and helix A_7 as in GH-18 chitinase and CLPs. In contrast to other subfamilies, YKL-39 contains no chitin binding domain. The catalytic residue Glu145, which is part of a conserved DxxDxDxE sequence motif is substituted to Ile, so YKL-39 lacks chitinase hydrolytic activity. In the structures of YKL-39 in the presence of chitin fragments, the ligand-binding cleft is observed to be long and open at both ends in order to facilitate the accessibility to

the sugar ligand. Superimposition of the free form YKL-39 to the YKL-39/chitooligosaccharide structures revealed that their TIM barrel domains are identical with only a few differences in the binding cleft. In addition, the presence of chitin fragment in the binding site induces local conformational changes, which potentially reflect the structural movement upon the ligand binding event. The $2F_o–F_c$ Fourier maps showed that $GlcNAc_2$ bound to protein covers subsites -2 to -1, $GlcNAc_4$ bound from subsites -3 to +1, while $GlcNAc_6$ bound from subsites -3 to +3. These findings suggested that the short chitin fragment bound primarily at the center of the binding groove. For the YKL-39/$GlcNAc_6$ structure, it was found that YKL-39 interacted with its sugar counterpart mainly through a combination of hydrogen bonding network and hydrophobic interactions. A larger number of contacts at the non-reducing end binding sites probably reflect stronger interactions. When the YKL-39 structure was compared with those of CHIT1, AMCase, and YKL-40, it was observed that their overall structures had high similarity with a few differences, within the chitin binding cleft.

The thermodynamics of binding of YKL-39 with chitooligosaccharides ($GlcNAc_2$ to $GlcNAc_6$) was investigated by intrinsic tryptophan fluorescence and ITC. All data from the binding reactions fitted well to a single-site binding model with a calculated *n* (stoichiometry) of 0.9-1.1. ITC experimental data for binding demonstrated that the binding affinity of YKL-39 for ligands increased with the increasing length of ligands in the micromolar range for K_d. The chitin binding cleft of YKL-39 contains a maximum of five subsites, namely (-3)(-2)(-1)(+1)(+2). Thermodynamic parameter determination clearly demonstrated that all ligands bind to protein with exothermic

and spontaneous reactions, which were mainly driven by enthalpy, as indicated by the dominant negative enthalpy change with unfavorable entropy.
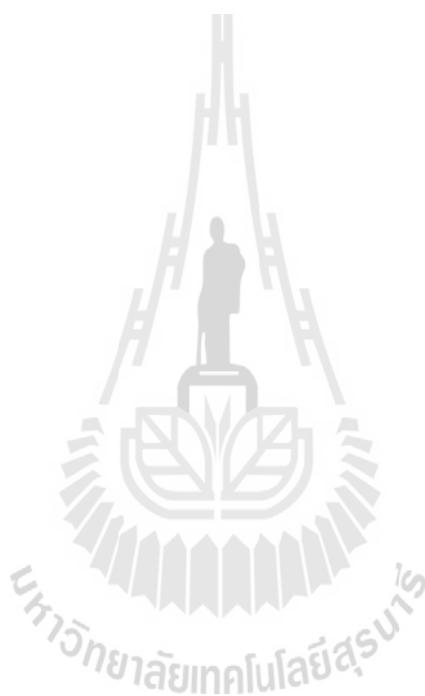
REFERENCES

REFERENCES

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002). PHENIX: building new software for automated crystallographic structure determination. **Acta Crystallogr. Sect. D-Biol. Crystallogr.** 58(11): 1948-1954.

Afonine, P. V., Grosse-Kunstleve, R. W., Chen, V. B., Headd, J. J., Moriarty, N. W., Richardson, J. S., Richardson, D. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2010). phenix.model_vs_data: a high-level tool for the calculation of crystallographic model and data statistics. **J. Appl. Crystallogr.** 43(4): 669-676.

Agnieszka, K. B. (2011). Thermodynamics of Ligand-Protein Interactions: Implications for Molecular Design.

Areshkov, P. A. and Kavsan, V. M. (2010). Chitinase 3-like protein 2 (CHI3L2, YKL-39) activates phosphorylation of extracellular signal-regulated kinases ERK1/ERK2 in human embryonic kidney (HEK293) and human glioblastoma (U87 MG) cells. **Tsitol. Genet.** 44(1): 3-9.

Areshkov, P. O., Avdieiev, S. S., Balynska, O. V., Leroith, D., and Kavsan, V. M. (2012). Two closely related human members of chitinase-like family, CHI3L1 and CHI3L2, activate ERK1/2 in 293 and U373 cells but have the different influence on cell proliferation. **Int. J. Biol. Sci.** 8(1): 39-48.

Arias, E. B., Verhage, H. G., and Jaffe, R. C. (1994). Complementary deoxyribonucleic acid cloning and molecular characterization of an estrogen-dependent human oviductal glycoprotein. **Biol. Reprod.** 51(4): 685-694.

Arlian, L. G. and Platts-Mills, T. A. (2001). The biology of dust mites and the remediation of mite allergens in allergic disease. **J. Allergy Clin. Immunol.** 107(3 Suppl): 406-413.

Avdieiev, S., Savinska, L., Filonenko, V., and Kavsan, V. (2012). Chitinase 3-like 2 protein monoclonal antibodies. **Hybridoma (Larchmt).** 31(1): 32-39.

Bissantz, C., Kuhn, B., and Stahl, M. (2010). A medicinal chemist's guide to molecular interactions. **J. Med. Chem.** 53(14): 5061-5084.

Boot, R. G., Blommaart, E. F., Swart, E., Ghauharali-van der Vlugt, K., Bijl, N., Moe, C., Place, A., and Aerts, J. M. (2001). Identification of a novel acidic mammalian chitinase distinct from chitotriosidase. **J. Biol. Chem.** 276(9): 6770-6778.

Boot, R. G., Bussink, A. P., and Aerts, J. M. (2005). Human acidic mammalian chitinase erroneously known as eosinophil chemotactic cytokine is not the ortholog of mouse YM1. **J. Immunol.** 175(4): 2041-2042.

Boot, R. G., Renkema, G. H., Strijland, A., van Zonneveld, A. J., and Aerts, J. M. (1995). Cloning of a cDNA encoding chitotriosidase, a human chitinase produced by macrophages. **J. Biol. Chem.** 270(44): 26252-26256.

Boot, R. G., Renkema, G. H., Verhoek, M., Strijland, A., Bliek, J., de Meulemeester, T. M., Mannens, M. M., and Aerts, J. M. (1998). The human chitotriosidase gene. Nature of inherited enzyme deficiency. **J. Biol. Chem.** 273(40): 25680-25685.

Boot, R. G., van Achterberg, T. A., van Aken, B. E., Renkema, G. H., Jacobs, M. J., Aerts, J. M., and de Vries, C. J. (1999). Strong induction of members of the chitinase family of proteins in atherosclerosis: chitotriosidase and human cartilage gp-39 expressed in lesion macrophages. **Arterioscler Thromb. Vasc. Biol.** 19(3): 687-694.

Brameld, K. A. and Goddard, W. A. (1998). Substrate Distortion to a Boat Conformation at Subsite -1 Is Critical in the Mechanism of Family 18 Chitinases. **J. Am. Chem. Soc.** 120(15): 3571-3580.

Brasso, K., Christensen, I. J., Johansen, J. S., Teisner, B., Garnero, P., Price, P. A., and Iversen, P. (2006). Prognostic value of PINP, bone alkaline phosphatase, CTX-I, and YKL-40 in patients with metastatic prostate carcinoma. **Prostate.** 66(5): 503-513.

Briton-Jones, C., Lok, I. H., Yuen, P. M., Chiu, T. T., Cheung, L. P., and Haines, C. (2001). Regulation of human oviductin mRNA expression in *vivo*. **Fertil. Steril.** 75(5): 942-946.

Bussink, A. P., Speijer, D., Aerts, J. M., and Boot, R. G. (2007). Evolution of mammalian chitinase(-like) members of family 18 glycosyl hydrolases. **Genetics.** 177(2): 959-970.

Chayen, N. E. (1999). Crystallization with oils: a new dimension in macromolecular crystal growth. **J. Cryst. Growth.** 196(2-4): 434-441.

Chew, G. L., Perzanowski, M. S., Canfield, S. M., Goldstein, I. F., Mellins, R. B., Hoepner, L. A., Ashby-Thompson, M., and Jacobson, J. S. (2008). Cockroach allergen levels and associations with cockroach-specific IgE. **J. Allergy Clin. Immunol.** 121(1): 240-245.

Chupp, G. L., Lee, C. G., Jarjour, N., Shim, Y. M., Holm, C. T., He, S., Dziura, J. D., Reed, J., Coyle, A. J., Kiener, P., Cullen, M., Grandsaigne, M., Dombret, M. C., Aubier, M., Pretolani, M., and Elias, J. A. (2007). A chitinase-like protein in the lung and circulation of patients with severe asthma. **N. Engl. J. Med.** 357(20): 2016-2027.

Cintin, C., Johansen, J. S., Christensen, I. J., Price, P. A., Sorensen, S., and Nielsen, H. J. (2002). High serum YKL-40 level after surgery for colorectal carcinoma is related to short survival. **Cancer.** 95(2): 267-274.

Cohen-Kupiec, R. and Chet, I. (1998). The molecular biology of chitin digestion. **Curr. Opin. Biotechnol.** 9(3): 270-277.

Collaborative Computing Project Number 4. (1994). The CCP4 suite: programs for protein crystallography. **Acta Crystallogr. Sect. D-Biol. Crystallogr.** 50(Pt 5): 760-763.

Colton, C. A., Mott, R. T., Sharpe, H., Xu, Q., Van Nostrand, W. E., and Vitek, M. P. (2006). Expression profiles for macrophage alternative activation genes in AD and in mouse models of AD. **J. Neuroinflammation**. 3: 27.

Cottrell, M. T., Moore, J. A., and Kirchman, D. L. (1999). Chitinases from uncultured marine microorganisms. **Appl. Environ. Microbiol.** 65(6): 2553-2557.

Davies, G. and Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. **Structure.** 3(9): 853-859.

De Ceuninck, F., Marcheteau, E., Berger, S., Caliez, A., Dumont, V., Raes, M., Anract, P., Leclerc, G., Boutin, J. A., and Ferry, G. (2005). Assessment of some tools for the characterization of the human osteoarthritic cartilage proteome. **J. Biomol. Tech.** 16(3): 256-265.

De Fost, M., Hollak, C. E., Groener, J. E., Aerts, J. M., Maas, M., Poll, L. W., Wiersma, M. G., Haussinger, D., Brett, S., Brill, N., and vom Dahl, S. (2006). Superior effects of high-dose enzyme replacement therapy in type 1 Gaucher disease on bone marrow involvement and chitotriosidase levels: a 2-center retrospective analysis. **Blood.** 108(3): 830-835.

Dehn, H., Hogdall, E. V., Johansen, J. S., Jorgensen, M., Price, P. A., Engelholm, S. A., and Hogdall, C. K. (2003). Plasma YKL-40, as a prognostic tumor marker in recurrent ovarian cancer. **Acta. Obstet. Gynecol. Scand.** 82(3): 287-293.

Di Rosa, M., Dell'Ombra, N., Zambito, A. M., Malaguarnera, M., Nicoletti, F., and Malaguarnera, L. (2006). Chitotriosidase and inflammatory mediator levels in Alzheimer's disease and cerebrovascular dementia. **Eur. J. Neurosci.** 23(10): 2648-2656.

Du, H., Masuko-Hongo, K., Nakamura, H., Xiang, Y., Bao, C. D., Wang, X. D., Chen, S. L., Nishioka, K., and Kato, T. (2005). The prevalence of autoantibodies against cartilage intermediate layer protein, YKL-39, osteopontin, and cyclic citrullinated peptide in patients with early-stage knee osteoarthritis: evidence of a variety of autoimmune processes. **Rheumatol. Int.** 26(1): 35-41.

Dupont, J., Tanwar, M. K., Thaler, H. T., Fleisher, M., Kauff, N., Hensley, M. L., Sabbatini, P., Anderson, S., Aghajanian, C., Holland, E. C., and Spriggs, D. R. (2004). Early detection and prognosis of ovarian cancer using serum YKL-40. **J. Clin. Oncol.** 22(16): 3330-3339.

Elias, J. A., Homer, R. J., Hamid, Q., and Lee, C. G. (2005). Chitinases and chitinase-like proteins in T(H)2 inflammation and asthma. **J. Allergy Clin. Immunol.** 116(3): 497-500.

Emsley, P. and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. **Acta Crystallogr. Sect. D-Biol. Crystallogr.** 60(12 Part 1): 2126-2132.

Francescone, R. A., Scully, S., Faibish, M., Taylor, S. L., Oh, D., Moral, L., Yan, W., Bentley, B., and Shao, R. (2011). Role of YKL-40 in the angiogenesis, radioresistance, and progression of glioblastoma. **J. Biol. Chem.** 286(17): 15332-15343.

Fusetti, F., Pijning, T., Kalk, K. H., Bos, E., and Dijkstra, B. W. (2003). Crystal structure and carbohydrate-binding properties of the human cartilage glycoprotein-39. **J. Biol. Chem**. 278(39): 37753-37760.

Fusetti, F., von Moeller, H., Houston, D., Rozeboom, H. J., Dijkstra, B. W., Boot, R. G., Aerts, J. M., and van Aalten, D. M. (2002). Structure of human chitotriosidase. Implications for specific inhibitor design and function of mammalian chitinase-like lectins. **J. Biol. Chem.** 277(28): 25537-25544.

Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. **Bioinformatics.** 15(4): 305-308.

Gratchev, A., Schmuttermaier, C., Mamidi, S., Gooi, L., Goerdt, S., and Kzhyshkowska, J. (2008). Expression of Osteoarthritis Marker YKL-39 is Stimulated by Transforming Growth Factor Beta (TGF-beta) and IL-4 in Differentiating Macrophages. **Biomark. Insights.** 3: 39-44.

Hakala, B. E., White, C., and Recklies, A. D. (1993). Human cartilage gp-39, a major secretory product of articular chondrocytes and synovial cells, is a mammalian member of a chitinase protein family. **J. Biol. Chem.** 268(34): 25803-25810.

Hart, P. J., Pfluger, H. D., Monzingo, A. F., Hollis, T., and Robertus, J. D. (1995). The refined crystal structure of an endochitinase from Hordeum vulgare L. seeds at 1.8 A resolution. **J. Mol. Biol.** 248(2): 402-413.

Henrissat, B. and Bairoch, A. (1993). New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. **Biochem. J.** 293(Pt 3): 781-788.

Henrissat, B. and Davies, G. J. (2000). Glycoside hydrolases and glycosyltransferases. Families, modules, and implications for genomics. **Plant Physiol.** 124(4): 1515-1519.

Hollak, C. E., van Weely, S., van Oers, M. H., and Aerts, J. M. (1994). Marked elevation of plasma chitotriosidase activity. A novel hallmark of Gaucher disease. **J. Clin. Invest.** 93(3): 1288-1292.

Holliday, R. and Ho, T. (1998). Evidence for gene silencing by endogenous DNA methylation. **Proc. Natl. Acad. Sci. USA.** 95(15): 8727-8732.

Houston, D. R., Recklies, A. D., Krupa, J. C., and van Aalten, D. M. (2003). Structure and ligand-induced conformational change of the 39-kDa glycoprotein from human articular chondrocytes. **J. Biol. Chem.** 278(32): 30206-30212.

Hu, B., Trinh, K., Figueira, W. F., and Price, P. A. (1996). Isolation and sequence of a novel human chondrocyte protein related to mammalian members of the chitinase protein family. **J. Biol. Chem.** 271(32): 19415-19420.

Iseli, B., Armand, S., Boller, T., Neuhaus, J. M., and Henrissat, B. (1996). Plant chitinases use two different hydrolytic mechanisms. **FEBS Lett.** 382(1-2): 186-188.

Jensen, B. V., Johansen, J. S., and Price, P. A. (2003). High levels of serum HER-2/neu and YKL-40 independently reflect aggressiveness of metastatic breast cancer. **Clin. Cancer Res.** 9(12): 4423-4434.

Johansen, J. S., Williamson, M. K., Rice, J. S., and Price, P. A. (1992). Identification of proteins secreted by human osteoblastic cells in culture. **J. Bone. Miner. Re**s. 7(5): 501-512.

Johansen, J. S., Baslund, B., Garbarsch, C., Hansen, M., Stoltenberg, M., Lorenzen, I., and Price, P. A. (1999). YKL-40 in giant cells and macrophages from patients with giant cell arteritis. **Arthritis Rheum.** 42(12): 2624-2630.

Johansen, J. S., Bojesen, S. E., Mylin, A. K., Frikke-Schmidt, R., Price, P. A., and Nordestgaard, B. G. (2009). Elevated plasma YKL-40 predicts increased risk of gastrointestinal cancer and decreased survival after any cancer diagnosis in the general population. **J. Clin. Oncol.** 27(4): 572-578.

Johansen, J. S., Christensen, I. J., Riisbro, R., Greenall, M., Han, C., Price, P. A., Smith, K., Brunner, N., and Harris, A. L. (2003). High serum YKL-40 levels in patients with primary breast cancer is related to short recurrence free survival. **Breast Cancer Res. Treat.** 80(1): 15-21.

Johansen, J. S., Drivsholm, L., Price, P. A., and Christensen, I. J. (2004). High serum YKL-40 level in patients with small cell lung cancer is related to early death. **Lung Cancer.** 46(3): 333-340.

Johansen, J. S., Hvolris, J., Hansen, M., Backer, V., Lorenzen, I., and Price, P. A. (1996). Serum YKL-40 levels in healthy children and adults. Comparison with serum and synovial fluid levels of YKL-40 in patients with osteoarthritis or trauma of the knee joint. **Br. J. Rheumatol.** 35(6): 553-559.

Johansen, J. S., Jensen, B. V., Roslind, A., Nielsen, D., and Price, P. A. (2006). Serum YKL-40, a new prognostic biomarker in cancer patients? **Cancer Epidemiol Biomarkers Prev.** 15(2): 194-202.

Junker, N., Johansen, J. S., Hansen, L. T., Lund, E. L., and Kristjansen, P. E. (2005). Regulation of YKL-40 expression during genotoxic or microenvironmental stress in human glioblastoma cells. **Cancer Sci.** 96(3): 183-190.

Kavsan, V., Dmitrenko, V., Boyko, O., Filonenko, V., Avdeev, S., Areshkov, P., Marusyk, A., Malisheva, T., Rozumenko, V., and Zozulya, Y. (2008). Overexpression of YKL-39 Gene in Glial Brain Tumors. **Scholarly Research Exchange.** 2008.

Kawada, M., Hachiya, Y., Arihiro, A., and Mizoguchi, E. (2007). Role of mammalian chitinases in inflammatory conditions. **Keio. J. Med.** 56(1): 21-27.

Kirkpatrick, R. B., Emery, J. G., Connor, J. R., Dodds, R., Lysko, P. G., and Rosenberg, M. (1997). Induction and expression of human cartilage glycoprotein 39 in rheumatoid inflammatory and peripheral blood monocyte-derived macrophages. **Exp Cell Res.** 237(1): 46-54.

Knorr, T., Obermayr, F., Bartnik, E., Zien, A., and Aigner, T. (2003). YKL-39 (chitinase 3-like protein 2), but not YKL-40 (chitinase 3-like protein 1), is up regulated in osteoarthritic chondrocytes. **Ann. Rheum. Dis.** 62(10): 995-998.

Kohler, G. and Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. **Nature.** 256(5517): 495-497.

Kuranda, M. J. and Robbins, P. W. (1991). Chitinase is required for cell separation during growth of Saccharomyces cerevisiae. **J. Biol. Chem.** 266(29): 19758-19767.

Kzhyshkowska, J., Gratchev, A., and Goerdt, S. (2007). Human chitinases and chitinase-like proteins as indicators for inflammation and cancer. **Biomark. Insights.** 2: 128-146.

Kzhyshkowska, J., Mamidi, S., Gratchev, A., Kremmer, E., Schmuttermaier, C., Krusell, L., Haus, G., Utikal, J., Schledzewski, K., Scholtze, J., and Goerdt, S. (2006). Novel stabilin-1 interacting chitinase-like protein (SI-CLP) is up-regulated in alternatively activated macrophages and secreted via lysosomal pathway. **Blood.** 107(8): 3221-3228.

Labadaridis, I., Dimitriou, E., Theodorakis, M., Kafalidis, G., Velegraki, A., and Michelakakis, H. (2005). Chitotriosidase in neonates with fungal and bacterial infections. **Arch. Dis. Child Fetal Neonatal Ed.** 90(6): F531-532.

Lapensee, L., Paquette, Y., and Bleau, G. (1997). Allelic polymorphism and chromosomal localization of the human oviductin gene (MUC9). **Fertil Steril.** 68(4): 702-708.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. **J. Appl. Crystallogr.** 26(2): 283-291.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. **J. Appl. Crystallogr.** 26(2): 283-291.

Leavitt, S. and Freire, E. (2001). Direct measurement of protein binding energetics by isothermal titration calorimetry. **Curr. Opin. Struct. Biol.** 11(5): 560-566.

Lee, C. G. (2009). Chitin, chitinases and chitinase-like proteins in allergic inflammation and tissue remodeling. **Yonsei. Med. J.** 50(1): 22-30.

Lee, C. G., Da Silva, C. A., Dela Cruz, C. S., Ahangari, F., Ma, B., Kang, M. J., He, C. H., Takyar, S., and Elias, J. A. (2011). Role of chitin and chitinase/chitinase-like proteins in inflammation, tissue remodeling, and injury. **Annu. Rev. Physiol.** 73: 479-501.

Lee, C. G., Da Silva, C. A., Lee, J. Y., Hartl, D., and Elias, J. A. (2008). Chitin regulation of immune responses: an old molecule with new roles. **Curr. Opin. Immunol.** 20(6): 684-689.

Lok, I. H., Briton-Jones, C. M., Yuen, P. M., and Haines, C. J. (2002). Variable expression of oviductin mRNA at different stages of human reproductive cycle. **J. Assist. Reprod. Genet.** 19(12): 569-576.

Malaguarnera, L., Musumeci, M., Di Rosa, M., Scuto, A., and Musumeci, S. (2005). Interferon-gamma, tumor necrosis factor-alpha, and lipopolysaccharide promote chitotriosidase gene expression in human macrophages. **J. Clin. Lab. Anal.** 19(3): 128-132.

Malette, B., Filion, B., St-Jacques, S., Kan, F. W., and Bleau, G. (1995). Hormonal control of the biosynthesis of hamster oviductin. **Microsc. Res. Tech.** 31(6): 470-477.

Malinda, K. M., Ponce, L., Kleinman, H. K., Shackelton, L. M., and Millis, A. J. (1999). Gp38k, a protein synthesized by vascular smooth muscle cells, stimulates directional migration of human umbilical vein endothelial cells. **Exp. Cell Res.** 250(1): 168-173.

Meng, G., Zhao, Y., Bai, X., Liu, Y., Green, T. J., Luo, M., and Zheng, X. (2010). Structure of human stabilin-1 interacting chitinase-like protein (SI-CLP) reveals a saccharide-binding cleft with lower sugar-binding selectivity. **J. Biol. Chem**. 285(51): 39898-39904.

Mitsuhashi, A., Matsui, H., Usui, H., Nagai, Y., Tate, S., Unno, Y., Hirashiki, K., Seki, K., and Shozu, M. (2009). Serum YKL-40 as a marker for cervical adenocarcinoma. **Annals of Oncology.** 20(1): 71-77.

Miyatake, K., Tsuji, K., Yamaga, M., Yamada, J., Matsukura, Y., Abula, K., Sekiya, I., and Muneta, T. (2013). Human YKL39 (chitinase 3-like protein 2), an osteoarthritis-associated gene, enhances proliferation and type II collagen expression in ATDC5 cells. **Biochem. Biophys. Res. Commun.** 431(1): 52-57.

Monzingo, A. F., Marcotte, E. M., Hart, P. J., and Robertus, J. D. (1996). Chitinases, chitosanases, and lysozymes can be divided into procaryotic and eucaryotic families sharing a conserved core. **Nat. Struct. Biol.** 3(2): 133-140.

Nyati, M. K., Morgan, M. A., Feng, F. Y., and Lawrence, T. S. (2006). Integration of EGFR inhibitors with radiochemotherapy. **Nat. Rev. Cancer**. 6(11): 876-885.

Ohno, T., Armand, S., Hata, T., Nikaidou, N., Henrissat, B., Mitsutomi, M., and Watanabe, T. (1996). A modular family 19 chitinase found in the prokaryotic organism Streptomyces griseus HUT 6037. **J. Bacteriol.** 178(17): 5065-5070.

Olland, A. M., Strand, J., Presman, E., Czerwinski, R., Joseph-McCarthy, D., Krykbaev, R., Schlingmann, G., Chopra, R., Lin, L., Fleming, M., Kriz, R., Stahl, M., Somers, W., Fitz, L., and Mosyak, L. (2009). Triad of polar residues implicated in pH specificity of acidic mammalian chitinase. **Protein Sci.** 18(3): 569-578.

Otwinowski, Z. and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode (Vol. 276, pp. 307-326): Elsevier.

Pastores, G. M. and Barnett, N. L. (2005). Current and emerging therapies for the lysosomal storage disorders. **Expert Opin. Emerg. Drugs.** 10(4): 891-902.

Perozzo, R., Folkers, G., and Scapozza, L. (2004). Thermodynamics of protein-ligand interactions: history, presence, and future aspects. **J. Recept. Signal Transduct Res.** 24(1-2): 1-52.

Ramachandran, G. N. and Sasisekharan, V. (1968). Conformation of polypeptides and proteins. **Adv. Protein Chem.** 23: 283-438.

Ray, A. and Cohn, L. (1999). Th2 cells and GATA-3 in asthma: new insights into the regulation of airway inflammation. **J. Clin. Invest.** 104(8): 985-993.

Rehli, M., Krause, S. W., and Andreesen, R. (1997). Molecular characterization of the gene for human cartilage gp-39 (CHI3L1), a member of the chitinase protein family and marker for late stages of macrophage differentiation. **Genomics.** 43(2): 221-225.

Renkema, G. H., Boot, R. G., Au, F. L., Donker-Koopman, W. E., Strijland, A., Muijsers, A. O., Hrebicek, M., and Aerts, J. M. (1998). Chitotriosidase, a chitinase, and the 39-kDa human cartilage glycoprotein, a chitin-binding lectin,

are homologues of family 18 glycosyl hydrolases secreted by human macrophages. **Eur. J. Biochem.** 251(1-2): 504-509.

Renkema, G. H., Boot, R. G., Muijsers, A. O., Donker-Koopman, W. E., and Aerts, J. M. (1995). Purification and characterization of human chitotriosidase, a novel member of the chitinase family of proteins. **J. Biol. Chem.** 270(5): 2198-2202.

Renkema, G. H., Boot, R. G., Strijland, A., Donker-Koopman, W. E., van den Berg, M., Muijsers, A. O., and Aerts, J. M. (1997). Synthesis, sorting, and processing into distinct isoforms of human macrophage chitotriosidase. **Eur. J. Biochem.** 244(2): 279-285.

Sahai, A. S. and Manocha, M. S. (1993). Chitinases of fungi and plants: their involvement in morphogenesis and host-parasite interaction. **FEMS Microbiology Reviews.** 11(4): 317-338.

Sakata, M., Masuko-Hongo, K., Tsuruha, J., Sekine, T., Nakamura, H., Takigawa, M., Nishioka, K., and Kato, T. (2002). YKL-39, a human cartilage-related protein, induces arthritis in mice. **Clin. Exp. Rheumatol.** 20(3): 343-350.

Schimpl, M., Rush, C. L., Betou, M., Eggleston, I. M., Recklies, A. D., and van Aalten, D. M. (2012). Human YKL-39 is a pseudo-chitinase with retained chitooligosaccharide-binding properties. **Biochem. J.** 446(1): 149-157.

Schmidt, H., Johansen, J. S., Sjoegren, P., Christensen, I. J., Sorensen, B. S., Fode, K., Larsen, J., and von der Maase, H. (2006). Serum YKL-40 predicts relapse-free and overall survival in patients with American Joint Committee on Cancer stage I and II melanoma. **J. Clin. Oncol.** 24(5): 798-804.

Sekine, T., Masuko-Hongo, K., Matsui, T., Asahara, H., Takigawa, M., Nishioka, K., and Kato, T. (2001). Recognition of YKL-39, a human cartilage related protein, as a target antigen in patients with rheumatoid arthritis. **Ann. Rheum. Dis.** 60(1): 49-54.

Shackelton, L. M., Mann, D. M., and Millis, A. J. (1995). Identification of a 38-kDa heparin-binding glycoprotein (gp38k) in differentiating vascular smooth muscle cells as a member of a group of proteins associated with tissue remodeling. **J. Biol. Chem.** 270(22): 13076-13083.

Shuhui, L., Mok, Y. K., and Wong, W. S. (2009). Role of mammalian chitinases in asthma. **Int. Arch. Allergy Immunol.** 149(4): 369-377.

Songsiriritthigul, C., Pantoom, S., Aguda, A. H., Robinson, R. C., and Suginta, W. (2008). Crystal structures of Vibrio harveyi chitinase A complexed with chitooligosaccharides: implications for the catalytic mechanism. **J. Struct. Biol.** 162(3): 491-499.

Steck, E., Breit, S., Breusch, S. J., Axt, M., and Richter, W. (2002). Enhanced expression of the human chitinase 3-like 2 gene (YKL-39) but not chitinase 3-like 1 gene (YKL-40) in osteoarthritic cartilage. **Biochem. Biophys. Res. Commun.** 299(1): 109-115.

Terpe, K. (2003). Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. **Appl. Microbiol. Biotechnol.** 60(5), 523-533.

Tews, I., Terwisscha van Scheltinga, A. C., Perrakis, A., Wilson, K. S., and Dijkstra, B. W. (1997). Substrate-Assisted Catalysis Unifies Two Families of Chitinolytic Enzymes. **J. Am. Chem. Soc.** 119(34): 7954-7959.

Tolia, N. H. and Joshua-Tor, L. (2006). Strategies for protein coexpression in Escherichia coli. **Nat. Methods.** 3(1): 55-64.

Turnbull, W. B. and Daranas, A. H. (2003). On the value of c: can low affinity systems be studied by isothermal titration calorimetry? **J Am Chem Soc**. 125(48): 14859-14866.

Vagin, A. and Teplyakov, A. (1997). MOLREP: an Automated Program for Molecular Replacement. **J. Appl. Crystallogr.** 30(6): 1022-1025.

Vagin, A. and Teplyakov, A. (2010). Molecular replacement with MOLREP. **Acta Crystallogr. Sect. D-Biol. Crystallogr.** 66(Pt 1): 22-25.

van Aalten, D. M., Komander, D., Synstad, B., Gaseidnes, S., Peter, M. G., and Eijsink, V. G. (2001). Structural insights into the catalytic mechanism of a family 18 exo-chitinase. **Proc. Natl. Acad. Sci. USA.** 98(16): 8979-8984.

van Eijk, M., van Roomen, C. P., Renkema, G. H., Bussink, A. P., Andrews, L., Blommaart, E. F., Sugar, A., Verhoeven, A. J., Boot, R. G., and Aerts, J. M. (2005). Characterization of human phagocyte-derived chitotriosidase, a component of innate immunity. **Int. Immunol.** 17(11): 1505-1512.

Vedder, A. C., Cox-Brinkman, J., Hollak, C. E., Linthorst, G. E., Groener, J. E., Helmond, M. T., Scheij, S., and Aerts, J. M. (2006). Plasma chitotriosidase in male Fabry patients: a marker for monitoring lipid-laden macrophages and their correction by enzyme replacement therapy. **Mol. Genet. Metab.** 89(3): 239-244.

Volck, B., Johansen, J. S., Stoltenberg, M., Garbarsch, C., Price, P. A., Ostergaard, M., Ostergaard, K., Lovgreen-Nielsen, P., Sonne-Holm, S., and Lorenzen, I. (2001). Studies on YKL-40 in knee joints of patients with rheumatoid arthritis

and osteoarthritis. Involvement of YKL-40 in the joint pathology. **Osteoarthritis Cartilage.** 9(3): 203-214.

Volck, B., Price, P. A., Johansen, J. S., Sorensen, O., Benfield, T. L., Nielsen, H. J., Calafat, J., and Borregaard, N. (1998). YKL-40, a mammalian member of the chitinase family, is a matrix protein of specific granules in human neutrophils. **Proc. Assoc. Am. Physicians.** 110(4): 351-360.

Watanabe, T., Kimura, K., Sumiya, T., Nikaidou, N., Suzuki, K., Suzuki, M., Taiyoji, M., Ferrer, S., and Regue, M. (1997). Genetic analysis of the chitinase system of Serratia marcescens 2170. **J. Bacteriol.** 179(22): 7111-7117.

Weber, P.C. (1997). **Methods Enzymol**. 276: 13-23.

Wiseman, T., Williston, S., Brandts, J. F., and Lin, L. N. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. **Anal Biochem**. 179(1): 131-137.

Yamac, D., Ozturk, B., Coskun, U., Tekin, E., Sancak, B., Yildiz, R., and Atalay, C. (2008). Serum YKL-40 levels as a prognostic factor in patients with locally advanced breast cancer. **Adv. Ther.** 25(8): 801-809.

CURRICULUM VITAE

Name: Miss Araya Ranok

Date of Birth: October 10th, 1977

Place of Birth: Nakhon Ratchasima, Thailand

Education:

2008- 2013:   School of Biochemistry, Suranaree University of Technology,

                       Nakhon Ratchasima, Thailand

                       Ph.D. in Biochemistry

2006-2008:   School of Health Science, Sukhothai Thammathirat Open

                       University, Bangkok, Thailand

                       B.Sc. in Occupation Health and Safety

2004-2006   Public Health Department, Khon Kaen University, Khon Kaen,

                       Thailand

                       M. Public Health in Environmental Health.

1995-1999:   Biochemistry Department, Khon Kaen University, Khon Kaen,

                       Thailand

                       B.Sc. in Biochemistry

Fellowships: The Office of the Higher Education Commission, Thailand