การใช้วิธีชีวสารสนเทศวิเคราะห์โปรตีนก่อโรคของมัยโคพลาสมานิวมอเนียที่มีศักยภาพในการผ่านเข้าสู่สมองมนุษย์

นางสาว วสุนันท์ ชุ่มเชื้อ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาชีวเวชเคมี ภาควิชาชีวเคมี
คณะเภสัชศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2549

ISBN 974-14-3916-4

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

BIOINFORMATIC ANALYSIS OF *MYCOPLASMA PNEUMONIAE* VIRULENCE FACTORS WITH A POTENTIAL ROLE IN HUMAN BRAIN INVASION

Miss Vasunun Chumchua

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Biomedicinal Chemistry
Department of Biochemistry
Faculty of Pharmaceutical Sciences
Chulalongkorn University
Academic year 2006
ISBN 974-14-3916-4

490114

BIOINFORMATIC ANALYSIS OF *MYCOPLASMA PNEUMONIAE* VIRULENCE FACTORS WITH A POTENTIAL ROLE IN HUMAN BRAIN INVASION

Miss Vasunun Chumchua

A Dissertation Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy Program in Biomedicinal Chemistry

Department of Biochemistry

Faculty of Pharmaceutical Sciences

Chulalongkorn University

Academic year 2006

ISBN 974-14-3916-4

Thesis Title	BIOINFORMATIC ANALYSIS OF *MYCOPLASMA PNEUMONIAE* VIRULENCE FACTORS WITH A POTENTIAL ROLE IN HUMAN BRAIN INVASION
By	Miss Vasunun Chumchua
Field of study	Biomedicinal Chemistry
Thesis Advisor	Associate Professor Duangdeun Meksuriyen, Ph.D.
Thesis Co-Advisor	Chinae Thammarongtham, Ph.D.

Accepted by the Faculty of Pharmaceutical Sciences, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctor's Degree

.....Pornpen Pramyothin.....Dean of Faculty of Pharmaceutical Sciences
(Associate Professor Pornpen Pramyothin, Ph.D.)

THESIS COMMITTEE

.....Vimolmas Lipipun.....Chairman
(Associate Professor Vimolmas Lipipun, Ph.D.)

.....D. Meksuriyen.....Thesis Advisor
(Associate Professor Duangdeun Meksuriyen, Ph.D.)

.....Chinae Thammarongtham.....Thesis Co-advisor
(Chinae Thammarongtham, Ph.D.)

.....Boonsri Ongpipattanakul.....Member
(Assistant Professor Boonsri Ongpipattanakul, Ph.D.)

.....Member
(Sudsanguan Chusaculthanachai, Ph.D.)

.....Member
(Somchai Pongpattanakitchote, Ph.D.)

วสุนันท์ ชุ่มเชื้อ: การใช้วิธีชีวสารสนเทศวิเคราะห์โปรตีนก่อโรคของมัยโคพลาสมานิวมอเนียที่มีศักยภาพในการผ่านเข้าสู่สมองมนุษย์ (BIOINFORMATIC ANALYSIS OF *MYCOPLASMA PNEUMONIAE* VIRULENCE FACTORS WITH A POTENTIAL ROLE IN HUMAN BRAIN INVASION) อ.ที่ปรึกษา: รศ.ดร. ดวงเดือน เมฆสุริเยนทร์, อ.ที่ปรึกษาร่วม: ดร. ชิเน ธำมรงค์ธรรม, 100 หน้า, ISBN 974-14-3916-4

วัตถุประสงค์ของการศึกษาครั้งนี้เพื่อคัดกรองโปรตีนก่อโรคของมัยโคพลาสมานิวมอเนียที่ทำให้เชื้อสามารถผ่านจากเซลล์เยื่อบุโพรงหลอดเลือดเข้าสู่สมองมนุษย์ ด้วยวิธีทางชีวสารสนเทศ พบว่าเอนไซม์อินอลเลส (e) เป็นโปรตีนก่อโรคที่สำคัญ โดยโครงสร้างสามมิติของเอนไซม์นี้ประกอบด้วยสองโดเมนหลักคือ N-terminal domain และ C-terminal domain โดยทางด้าน N-terminal domain ประกอบด้วยสายแอลฟาและสายเบตาอย่างละสามสาย ส่วนด้าน C-terminal domain ประกอบด้วยสายแอลฟาและสายเบตาเรียงสลับกันเป็นแบบ TIM barrel และมีลูปขนาดเล็กยื่นออกไป ผลการศึกษาครั้งนี้พบว่าการเรียงตัวของ TIM barrel เป็นแบบ anti-parallel $(\alpha/\beta)_8$ barrel โดยมีไอออนแมกนีเซียมซึ่งเป็นโคแฟกเตอร์ของเอนไซม์อินอลเลสถูกล้อมรอบด้วยกรดอะมิโน D_{256}-E_{310}-D_{337} ทั้งนี้บริเวณกรดอะมิโนตำแหน่งที่ 264 ถึง 276 พบว่าพับตัวเป็นลูป ซึ่งเป็นบริเวณที่มีความแปรปรวนของลำดับกรดอะมิโนเมื่อเปรียบเทียบกับโครงสร้างของเอนไซม์ชนิดนี้ในเชื้อชนิดอื่นการศึกษาอันตรกริยาระหว่างเอนไซม์อินอลเลสและพลาสมิโนเจน (plg) พบว่าบริเวณที่โปรตีนทั้งสองชนิดสามารถจับกันได้ด้วยพันธะไฮโดรเจนระหว่างกันมีดังนี้คือ eR_{24}-$plgK_{48}$, eK_{70}-$plgY_{50}$, eN_{165}-$plgT_{66}$, eA_{168}-$plgE_{21}$, eD_{171}-$plgK_{70}$, และ eN_{213}-$plgP_{68}$/$plgN_{69}$ ซึ่งการศึกษาตำแหน่งที่เกิดอันตรกริยาระหว่างเอนไซม์อินอลเลสของมัยโคพลาสมานิวมอเนียและพลาสมิโนเจนครั้งนี้เป็นรายงานครั้งแรกที่แสดงให้เห็นถึงความสามารถในการจับกันระหว่างโปรตีนคู่นี้อาจเป็นปัจจัยส่งเสริมช่วยให้มัยโคพลาสมาผ่านเข้าสมองโดยอาศัยพลาสมิโนเจนที่ได้รับการกระตุ้นจาก tissue plasminogen activator (tPA) เพื่อเปลี่ยนไปเป็นพลาสมินซึ่งจะไปทำลายโปรตีนบนเซลล์เยื่อบุโพรงหลอดเลือดในสมองเป็นผลให้เซลล์หลุดลอกออก รวมถึงส่งผลให้มีการหลั่งของ cytokine และ chemokine เหนี่ยวนำให้เซลล์มีการจัดเรียงตัวใหม่ทำให้มัยโคพลาสมาฉวยโอกาสเข้าสู่สมองแบบเดียวกับเชื้อในกลุ่ม *Escherichia coli* และ *Streptococcus* ผ่านเข้าสู่เซลล์เจ้าบ้าน จากผลการศึกษาครั้งนี้ได้ให้ข้อมูลเบื้องต้นเกี่ยวกับโปรตีนที่มีบทบาทต่อการผ่านเข้าสมอง อันจะเป็นประโยชน์สำหรับออกแบบยาและการคัดเลือกโปรตีนสำหรับการทดลองต่อไปในห้องปฏิบัติการ

ภาควิชา.....ชีวเคมี.....ลายมือชื่อนิสิต.....วสุนันท์ ชุ่มเชื้อ.....

สาขาวิชา.....ชีวเวชเคมี.....ลายมือชื่ออาจารย์ที่ปรึกษา.....

ปีการศึกษา.....2549.....ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

4676960733: MAJOR BIOMEDICINAL CHEMISTRY

KEY WORDS: *MYCOPLASMA PNEUMONIAE* ENOLASE / HUMAN PLASMINOGEN KRINGLE DOMAIN 2 / VIRULENCE FACTOR / COMPARATIVE MODELING / MOLECULAR DOCKING / MOLECULAR DYNAMICS

VASUNUN CHUMCHUA: BIOINFORMATIC ANALYSIS OF *MYCOPLASMA PNEUMONIAE* VIRULENCE FACTORS WITH A POTENTIAL ROLE IN HUMAN BRAIN INVASION, THESIS ADVISOR: ASSOC. PROF. DUANGDEUN MEKSURIYEN, Ph.D., THESIS CO-ADVISOR: CHINAE THAMMARONGTHAM, Ph.D., 100 PAGES. ISBN 974-14-3916-4

This investigation was conducted to identify virulence factors of *Mycoplasma pneumoniae* from its genome by bioinformatic schemes and to determine their affinities on plasminogen which circulate in blood brain vessel, by molecular modeling. This study shows that enolase is a putative virulence factor. Therefore enolase was then taken to model its three-dimensional structure and to determine the interaction of *M. pneumoniae* enolase with plasminogen using molecular modeling and molecular docking, respectively. Subsequently the obtained docking model was taken to study in dynamics with implicit water under NVT ensemble to investigate their stabilities. The *M. pneumoniae* enolase model obtained shows a kidney shape consisted of two domains, the N-terminal domain containing three α helices and three anti-parallel β-sheets, the C-terminal domain forming a $(\beta/\alpha)_8$-barrel structure with a small protruding loop. The Mg^{2+} ion, a metal cofactor is surrounded by D_{256}-E_{310}-D_{337}. Although this model appears to share the same fold of overall structure with those of template and other microbial enolase, it has a particular variant loop of amino acid residue 264 to residue 276. Molecular docking shows the key residues involved in hydrogen bonding which plays an important role in the interaction between *M. pneumoniae* enolase (e) and human plasminogen (plg) as followed, eR_{24}-$plgK_{48}$, eK_{70}-$plgY_{50}$, eN_{165}-$plgT_{66}$, eA_{168}-$plgE_{21}$, eD_{171}-$plgK_{70}$, and eN_{213}-$plgP_{68}$/$plgN_{69}$. The interaction result obtained is consistent with experimental data on binding assay that enolase has a binding affinity to plasminogen. This information may propose the prospective of *M. pneumoniae* invasion across human brain either endothelial cell detachment or endothelial cell remodeling by the facilitating of plasmin, an active form of plasminogen, as well as the pathophysiology of invasion of *Escherichia coli* and *Streptococcus* species. Plasmin is activated by tissue plasminogen activator (tPA). Subsequently it would involve in endothelial cell detachment and remodeling by the authorizing of cytokine and chemokine induced by the increase in plasmin. Our result is the first report on molecular interaction, by computational simulation, of *M. pneumoniae* enolase-plasminogen which provides the structural basis of binding complex that suggests to further experiment and rationale drug design.

Department.....Biochemistry.....Student's signature.....Vasunun Chumchua

Field of study....Biomedicinal Chemistry.....Advisor's signature.....D. Meksuriyen

Academic year...2006.....Co-advisor's signature.....Chinae.....

ACKNOWLEDGEMENTS

I would like to acknowledge many people for helping me during my doctoral work. I would especially like to thank my advisor, Associate Professor Dr. Duangdeun Meksuriyen, for her generous time and commitment. Throughout my doctoral work she encouraged me to develop independent thinking and research skills. She continually stimulated my analytical thinking and greatly assisted me with scientific writing.

I wish to thank Dr. Chinae Thammarongtham, my co-advisor, for invaluable advice guidance and encouragement throughout this sophisticate study including allowing me to use facility for the Bioinformatics and Molecular dynamics simulation work at BIOTEC. I am also very grateful to Natapol Pornputtapong for initial teaching me the computational modelling method. His knowledge of computers was also greatly helpful.

I am extremely grateful for the assistance, generosity, and advice I received from Dr. Sudsanguan Chusakulthanachai at Duke University, Dr. Somchai Pongpattanakitchote at BIOTEC, Associate Professor Vimolmas Lipipun, and Assistant Professor Boonsri Ongpipattanakul at Chulalongkorn University. Their assistance was helpful to complete this dissertation.

I owe a special note of gratitude to Dr. Anuchai Theerareungchaisri at Chulalongkorn University and Tim Searles at Duke University for assisting me with miscellaneous tasks for video conference setting. This was especially helpful during my dissertation examination time.

I am grateful to many persons who shared their memories and experiences, especially, Supim Wongtongtair, Chompunuch Boonyarkart, and Pilaiwanwadee Hutamekalin. I need to express my gratitude and deep appreciation for their friendship, hospitality, and wisdom have supported, enlightened, and entertained me over the many years of our friendship. They have consistently helped me keep perspective on what is important in life and show me how to deal with reality.

I would like to thank Associate Professor Dr. Supapon Cheevadhanarak, Dr. Junthira Punya, and all members of combinatorial biosynthesis laboratory at BIOTEC for being support, room space, computer, printer and their helpful during I wrote my dissertation.

Finally, I would like to thank my family for their patience and for helping me keep my life in proper perspective and balance.

This research was partially funded by Thailand Graduate Institute of Science and Technology of National Science and Technology Development Agency (operating grant TGIST 01-48-039)

CONTENTS

		Page
ABSTRACT (Thai).....		iv
ABSTRACT (English).....		v
ACKNOWLEDGEMENTS.....		vi
CONTENTS.....		vii
LIST OF TABLES.....		viii
LIST OF FIGURES.....		ix
LIST OF ABBREVIATIONS.....		x
CHAPTER		
I.	INTRODUCTION.....	1
II.	LITERATURE REVIEW.....	7
	MYCOPLASMA PNEUMONIAE AND ITS ROLE EXRA-PULMONARY DISEASE.....	7
	BIOINFORMATICS APPROACH FOR IDENTIFYING VIRULENCE FACTOR.....	13
	BIOLOGICAL SEQUENCE ANALYSIS.....	14
	COMPARATIVE MODELING.....	15
	MOLECULAR DOCKING.....	20
	APPLICATION OF MOLECULAR DYNAMICS SIMULATION.....	21
III.	MATERIALS AND METHODS.....	25
IV.	RESULTS.....	37
V.	DISCUSSION AND CONCLUSION.....	64
REFERENCES.....		71
APPENDICES.....		81
VITA.....		100

LIST OF TABLES

Table		Page
1.	Selected structure annotation tools	26
2.	Selected sequence annotation tools	31
3.	Properties of molecular system used in investigation of protein complexes by MD simulation	35
4.	The protein property prediction of un-annotate gene that found in *PAIs* performed by sequence annotation tool	38
5.	Model evaluations the model quality between crude and refinement model	46
6.	The model refinement quality evaluated by PROCHECK	47
7.	The calculated energies (kcal/mol) of MpnE-Plasminogen interaction comparison between rigid-docking and MD simulation	56
8.	The distance between MpnE model and human plasminogen that makes contact by hydrogen bonding	57
9A.	Different Accessible Molecular Surface Accesibility (ΔASA) values of the amino acid of *M. pneumoniae* enolase model that makes contact or produces hydrogen bonding with human plasminogen kringle domain 2 that evaluated by *WHATIF*	61
9B.	Different Accessible Molecular Surface Accesibility (ASA) values of the amino acid of human plasminogen kringle domain 2 that makes contacts or produces hydrogen bonding with MpnE model that evaluated by *WHATIF*	62
10.	Amino acid characterized in the interface between MpnE and Human plasminogen kringle domain 2	63

LIST OF FIGURES

Figure		Page
1.	Primary structure of human plasminogen.....	12
2.	Summary of putative virulence factors from entire genome of *M.pneumoniae*.....	38
3.	The proteomics profile of entire *M. pneumoniae* protein.....	40
4.	The *M. pneumoniae* proteome overview.....	41
5.	Protein domain prediction.....	43
6.	Multiple sequence alignment of microbial enolases	45
7.	RMSD of MpnE with Mg^{2+} model from MD simulation.....	50
8.	MpnE model.....	51
9.	The structural alignment of enolase among 3 organisms.....	52
10.	Plausible interactions model.....	55
11.	PyMol representation of the electrostatic potential surface area of interaction.....	58
12.	PyMol represents electrostatic potential surface particular on 10 residues that make contact with human plasminogen kringle domain 2 has been defined by Δ ASA.....	59
13.	PyMol representation of the electrostatic potential surface of human plasminogen kringle domain 2 that make contact with MpnE model has been defined by Δ ASA.....	60
14.	Propose mechanism of M. pneumoniae involved in human brain invasion...	68
15.	Plasmin augments the releasing of cytokine.....	69

LIST OF ABBREVIATIONS

Å	angstrom
$Å^3$	cubic angstrom
α	alpha
ASA	accessible surface area
β	beta
BMEC	brain microvascular endothelial cells
BSGC	Berkeley Structural Genomics Center
CASP	critical assessment in structure prediction
3D	Three dimensional
EM	energy minimization
hBBB	human blood brain barrier
HBMEC	human brain microvascular endothelial cell
KNB	knob domain
LBS	lysyl binding site
MD	molecular dynamics
MpnE	*Mycoplasma pneumoniae* enolase
NCBI	National center of biotechnology institute
NMR	nuclear magnetic resonance
NVT	number, volume, and temperature
ns	nanosecond
PAIs	pathogenicity island
PDB	protein data bank
plg	plasminogen
ps	picosecond
RMSD	root mean square deviation
SCRs	structural conserved regions
SEN	streptococcal surface enolase
SVRs	structural variable regions
tPA	tissue-type plasminogen activator
uPA	urokinase plasminogen activator