การเกิดเอนไซม์คอมเพล็กซ์ระหว่างออโรเทต ฟอสฟอไรโบซิลทรานสเฟอเรส และออโรทิดีน 5'-โมโนฟอสเฟต ดีคาร์บอกซีลเลสของพลาสโมเดียม ฟัลซิพารัม

นางสาววรัญญา อิ่มประสิทธิชัย

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาชีวเวชศาสตร์ (สหสาขาวิชา)
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2556

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ ที่ส่งผ่านทางบัณฑิตวิทยาลัย
The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)
are the thesis authors' files submitted through the University Graduate School.

FORMATION OF ENZYME COMPLEX BETWEEN OROTATE PHOSPHORIBOSYLTRANSFERASE AND OROTIDINE 5'-MONOPHOSPATE DECARBOXYLASE OF PLASMODIUM FALCIPARUM.

Miss Waranya Imprasittichai

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Biomedical Sciences
(Interdisciplinary Program)
Graduate School
Chulalongkorn University
Academic Year 2013
Copyright of Chulalongkorn University

Thesis Title	FORMATION OF ENZYME COMPLEX BETWEEN OROTATE PHOSPHORIBOSYLTRANSFERASE AND OROTIDINE 5'-MONOPHOSPATE DECARBOXYLASE OF PLASMODIUM FALCIPARUM.
By	Miss Waranya Imprasittichai
Field of Study	Biomedical Sciences
Thesis Advisor	Professor Jerapan Krungkrai, Ph.D.
Thesis Co-Advisor	Sudaratana Krungkrai, D.Eng.
	Sittiruk Roytrakul, Ph.D.

Accepted by the Graduate School, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctoral Degree

.....Dean of the Graduate School

(Associate Professor Amon Petsom, Ph.D.)

THESIS COMMITTEE

.....Chairman

(Professor Apiwat Mutirangura, M.D., Ph.D.)

.....Thesis Advisor

(Professor Jerapan Krungkrai, Ph.D.)

.....Thesis Co-Advisor

(Sudaratana Krungkrai, D.Eng.)

.....Thesis Co-Advisor

(Sittiruk Roytrakul, Ph.D.)

.....Examiner

(Professor Sittisak Honsawek, M.D., Ph.D.)

.....Examiner

(Assistant Professor Nutchanat Wutipraditkul, Ph.D.)

.....External Examiner

(Assistant Professor Sumet Wajanarogana, Ph.D.)

วรัญญา อิ่มประสิทธิชัย : การเกิดเอนไซม์คอมเพล็กซ์ระหว่างออโรเทต ฟอสฟอไรโบซิลทรานสเฟอเรส และออโรทิดีน 5'-โมโนฟอสเฟต ดีคาร์บอกซีลเลสของพลาสโมเดียมฟัลซิพารัม. (FORMATION OF ENZYME COMPLEX BETWEEN OROTATE PHOSPHORIBOSYLTRANSFERASE AND OROTIDINE 5'-MONOPHOSPATE DECARBOXYLASE OF PLASMODIUM FALCIPARUM.) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: ศ. ดร.จิระพันธ์ กรึงไกร, อ.ที่ปรึกษาวิทยานิพนธ์ร่วม: ดร.สุดารัตน์ กรึงไกร, ดร.สิทธิรักษ์ รอยตระกูล, 96 หน้า.

พลาสโมเดียม ฟัลซิพารัม เป็นเชื้อสาเหตุหลักที่ทำให้เกิดการตายในผู้ป่วยที่เป็นโรคมาลาเรีย โดยอาศัยกระบวนการสังเคราะห์ไพริมิดีนแนววีถีดีโนโวอย่างเดียวในการเจริญเติบโต ในงานวิจัยนี้สนใจเอนไซม์ ออโรเทต ฟอสฟอไรโบซิลทรานเฟอเรส และ ออโรทิดีน 5´-โมโนฟอสเฟต ดีคาร์บอกซีเลส ซึ่งเป็นเอนไซม์ในลำดับที่ 5 และ 6 ของกระบวนการนี้ และถือเป็นตำแหน่งที่น่าสนใจในการพัฒนายารักษาโรคมาลาเรีย จากการศึกษาก่อนหน้านี้แสดงให้เห็นว่าเอนไซม์ทั้ง 2 ชนิด อยู่ในรูปของเฮเทโรเตตระเมอริก คอมเพล็กซ์ โดยประกอบด้วย ออโรเทต ฟอสฟอไรโบซิลทรานเฟอเรส และ ออโรทิดีน 5´-โมโนฟอสเฟต ดีคาร์บอกซีเลส อย่างละ 2 หน่วยย่อย การทำงานในรูปแบบเอนไซม์คอมเพล็กซ์จะมีประสิทธิภาพมากกว่าการทำงานในรูปแบบเอนไซม์เดี่ยว ในการศึกษานี้ได้นำเสนอถึง ปฏิสัมพันธ์ระหว่างกันของเอนไซม์ ในรูปแบบเอนไซม์คอมเพล็กซ์ ได้ทำการพิสูจน์โดยการเชื่อมต่อทั้งสองเอนไซม์เข้าด้วยกันโดยใช้สารเคมีเชื่อมต่อกรดอะมิโน จากนั้นวิเคราะห์หากรดอะมิโนที่เกี่ยวข้องโดยใช้เครื่องลิควิด โครมาโตกราฟฟี่-แมส สเปคโตรมิทรี และได้สร้างโครงสร้างจำลองของเอนไซม์ขึ้นมา ผลจากการทดลองปรากฏว่า บริเวณที่เป็นกรดอะมิโนที่เติมเข้ามาและมีลักษณะเป็นกรดอะมิโนที่มีความซ้ำกันโดยไม่พบในเอนไซม์เดียวกันในสิ่งมีชีวิตอื่นที่อยู่บริเวณเกลียวแอลฟา 2 และ 5 ของออโรทิดีน 5´-โมโนฟอสเฟต ดีคาร์บอกซีเลส เป็นบริเวณที่เกิดปฏิสัมพันธ์กันระหว่างออโรเทต ฟอสฟอไรโบซิลทรานเฟอเรส และ ออโรทิดีน 5´-โมโนฟอสเฟต ดีคาร์บอกซีเลส นำไปสู่การสร้างโครงสร้างจำลองของ เฮเทโรเตตระเมอริก ออโรเทต ฟอสฟอไรโบซิลทรานเฟอเรส และ ออโรทิดีน 5´-โมโนฟอสเฟต ดีคาร์บอกซีเลส มัลติเอนไซม์คอมเพล็กซ์ โดยการใช้ข้อมูลทางด้านโปรตี-โอมิกและการสร้างแบบจำลองโครงสร้าง ทำให้สรุปได้ว่าเชื้อมาลาเรียในมนุษย์ใช้บริเวณที่เป็นกรดอะมิโนซ้ำกันในการมีปฏิสัมพันธ์ระหว่าง ออโรเทต ฟอสฟอไรโบซิลทรานเฟอเรส และ ออโรทิดีน 5´-โมโนฟอสเฟต ดีคาร์บอกซีเลส และเกิดโครงสร้างคอมเพล็กซ์ ซึ่งมีผลทำให้ประสิทธิภาพการทำงานของเอนไซม์ดีขึ้น

สาขาวิชา ชีวเวชศาสตร์	ลายมือชื่อนิสิต .....
ปีการศึกษา 2556	ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก .....
	ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม .....
	ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม .....

5287817420 : MAJOR BIOMEDICAL SCIENCES
KEYWORDS: OROTATE PHOSPHORIBOSYLTRANSFERASE / OROTIDINE 5'-MONOPHOSPHATE DECARBOXYLASE / PROTEIN-PROTEIN INTERACTION / LOW COMPLEXITY REGION / PYRIMIDINE BIOSYNTHESIS / PLASMODIUM FALCIPARUM

WARANYA IMPRASITTICHAI: FORMATION OF ENZYME COMPLEX BETWEEN OROTATE PHOSPHORIBOSYLTRANSFERASE AND OROTIDINE 5'-MONOPHOSPATE DECARBOXYLASE OF PLASMODIUM FALCIPARUM.. ADVISOR: PROF. JERAPAN KRUNGKRAI, Ph.D., CO-ADVISOR: SUDARATANA KRUNGKRAI, D.Eng., SITTIRUK ROYTRAKUL, Ph.D., 96 pp.

Plasmodium falciparum, the causative agent of the most lethal form of human malaria, relies on de novo pyrimidime biosynthetic pathway. The parasite orotate phosphoribosyltransferase (OPRT) and orotidine 5'-monophosphate decarboxylase (OMPDC), the fifth and sixth enzyme of the pathway, is attractive target for antimalarial development. Previously, it was clearly established that the two enzymes in the malaria parasite exist physically as a heterotetrameric $(OPRT)_2(OMPDC)_2$ complex containing two subunits each of OPRT and OMPDC, and that the complex have catalytic kinetic advantages over the monofunctional enzyme. In this study, the protein-protein interaction in enzyme complex was identified using chemical cross-linker, liquid chromatography-mass spectrometric analysis and homology modeling. Interestingly, the unique insertions of low complexity region at the α2 and α5 helices of the parasite OMPDC, characterized by single amino acid repeat sequence which was not found in homologous proteins from other organisms, was located on the OPRT-OMPDC interface. The structural models for the protein-protein interaction of the heterotetrameric $(OPRT)_2(OMPDC)_2$ multienzyme complex were proposed. Based on the proteomic data and structural modeling, it is summarized that the low complexity region of human malaria parasite is responsible for the OPRT-OMPDC interaction. The structural complex of the parasite enzymes, thus, represents an efficient functional kinetic advantage.

Field of Study: Biomedical Sciences

Academic Year: 2013

Student's Signature .....

Advisor's Signature .....

Co-Advisor's Signature .....

Co-Advisor's Signature .....

ACKNOWLEDGEMENTS

I would like to express my appreciation and sincere thanks to my thesis advisor, Professor Dr. Jerapan Krungkrai for his invaluable help, excellent advice and constant encouragement throughout the course of this research. I am most grateful for his teaching and advice, not only the research methodologies but also alot of expertise in life.

In addition, I am extremely grateful to Dr. Sudaratana Krungkrai for her kind help and instruction in laboratory techniques. I am extremely grateful to Dr. Sittiruk Roytrakul for his suggestion in Mass Spectrometry techniques. I am extremely grateful to my supervisory committee, Professor Dr. Apiwat Mutirangura, Professor Dr. Sittisak Honsawek, Dr. Nutchanat Wutipraditkul and Assistant Professor Dr.Sumet Wajanarogana for their suggestions and criticism.

I would like to thanks all staffs in biomedical science unit and staffs in Department of Biochemistry, Faculty of Medicine, Chulalongkorn University for their help and advice during my study. And grateful to Miss Atchara Paemanee, staff at BIOTEC for her advice and take care of my laboratory at BIOTEC.

I am particularly grateful to the Graduate School of Chulalongkorn University for funding support of this thesis.

Finally, I most gratefully acknowledge my parents and my friends for all their support throughout the period of this research.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

CONTENTS

Page

THAI ABSTRACT iv

ENGLISH ABSTRACT v

ACKNOWLEDGEMENTS vi

CONTENTS vii

LIST OF TABLES x

LIST OF FIGURES xi

LIST OF ABBREVIATIONS xiii

CHAPTER I INTRODUCTION 1

1.1. Background and Rationale 1

1.2. Research question 3

1.3. Hypothesis 3

1.4. Objective 3

1.5. Conceptual framework 4

1.6. Key words 5

1.7. Expected Benefits 5

CHAPTER II LITERATURE REVIEWS 6

2.1 Human malaria parasite life cycle 6

2.2 Genomic of malaria parasite 10

2.3 Pyrimidine biosynthesis in human 11

2.3.1 Synthesis of pyrimidine nucleotides by *de novo* pathway 11

2.3.2 Synthesis of pyrimidine nucleotide by salvage pathway 13

2.4 Pyrimidine biosynthesis in *Plasmodium* parasite 15

2.5 Orotate phosphoribosyltransferase (OPRT) 17

2.6 Orotidine 5´-monophosphate decarboxylase (OMPDC) 21

2.7 *Pf*OPRT and *Pf*OMPDC enzyme complex 24

2.8 Mass spectrometry for protein complex interaction 29

2.9 Computational protein structure modeling 33

Page

CHAPTER III MATERIALS AND METHODS 37
3.1 Materials 37
3.2 Equipments 38
3.3 Reagents 40
3.3.1 General reagents 40
3.3.2 Reagent kits 43
3.4 Methods 44
3.4.1 Construction of plasmid in *E. coli* for protein expression 44
3.4.2 Protein expression and purification of recombinant *Pf*OPRT and *Pf*OMPDC. 44
3.4.3 Protein and enzymatic assay 46
3.4.4 Chemical cross-linking of *Pf*OPRT and *Pf*OMPDC 46
3.4.5 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE) 47
3.4.6 In gel digestion 47
3.4.7 Liquid chromatography-mass spectrometry (LC-MS/MS) and proteomic data analysis. 48
3.4.8 Homology modeling of *Pf*OPRT and *Pf*OMPDC in monomer, dimer and heterotetramer 50
CHAPTER IV RESULTS 52
4.1 Construction of recombinant plasmid *Pf*OPRT and *Pf*OMPDC in *E. coli* for protein expression 52
4.2 Recombinant enzyme *Pf*OPRT and *Pf*OMPDC preparation 54
4.3 Chemical cross-linking of *Pf*OPRT and *Pf*OMPDC 55
4.4 Liquid chromatography/mass spectrometry of chemical cross-linking products of dimeric forms and heterotetrameric $(Pf\text{OPRT})_2(Pf\text{OMPDC})_2$ 60
4.5 Homology models and proteomic data analysis of *Pf*OPRT and *Pf*OMPDC in monomeric, dimeric and heterotetrameric forms 64
CHAPTER V DISCUSSION AND CONCLUSION 76
REFERENCES 88

Page

VITA..... 96

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

LIST OF TABLES

Table		Page
1	The list of protein modeling web servers	35
2	LC-MS/MS data analysis of *Pf*OPRT in monomeric, dimeric forms and in heterotetrameric complexed with *Pf*OMPDC	62
3	LC-MS/MS data analysis of *Pf*OMPDC in monomeric, dimeric forms and in heterotetrameric complexed with *Pf*OPRT	63

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

LIST OF FIGURES

Figure		Page
1	The world map showing the area of malaria	7
2	Life cycle of human malaria	8
3	Pyrimidine *de novo* biosynthesis in human	14
4	The *de novo* pyrimidine biosynthesis of malaria parasites	16
5	Sequence alignment of OPRTs	19
6	The 3D structure of *S.cerevisiae* OPRTs	20
7	Sequence alignment and secondary structure assingment of OMPDCs	22
8	Crystal structure of *P. falciparum* OMPDC	23
9	Model of multienzyme $(OPRT)_2(OMPDC)_2$ complex formation in *P. falciparum*	25
10	Model of UMPS oligomerizationin *L. donovani*	28
11	Schematic workflow of cross-link residue for analysis for 3D structure information	31
12	Cross-link product of calmodulin-melittin complex	32
13	Agarose gel analysis of *Pf*OPRT-pQE30 Xa and *Pf*OMPDC-pTrc His A plasmid	53
14	SDS-PAGE analysis of monomeric and dimeric forms of *P. falciparum* OPRT	57
15	SDS-PAGE analysis of monomeric and dimeric forms of *P.*	58

falciparum OMPDC

16 SDS-PAGE analysis of heterotetrameric (*Pf*OPRT)$_2$(*Pf*OMPDC)$_2$ 59

17 Sequence alignment and secondary structure of OPRTs 68

18 Sequence alignment and secondary structure of OMPDCs 70

19 Phyre homology model of monomeric *P. falciparum* OPRT and OMPDC 72

20 Proposed homology model of monomeric *P. falciparum* OPRT in dimeric form 73

21 Proposed homology model of monomeric *P. falciparum* OMPDC in dimeric form 74

22 Proposed homology model of heterotetrameric (*Pf*OPRT)$_2$(*Pf*OMPDC)$_2$ 75

23 The sequential steps for dimer and tetramer formation of *P. falciparum* OPRT and OMPDC. 82

24 Comparative structural models among *P. falciparum, L. donovani* and human OPRT and OMPDC. 83

25 Homology models of tetramer (OPRT)$_2$(OMPDC)$_2$ of *P. falciparum* and dimer (OPRT-OMPDC)$_2$ of human. 84

LIST OF ABBREVIATIONS

µg	=	microgram
µl	=	microlitre
µM	=	micromolar
µm	=	micrometer
µmol	=	micromole
Arg	=	arginine
ATC	=	aspartate transcarbamoylase
ATP	=	adenosine triphosphate
Asp	=	aspartate
bp	=	base pair
BSA	=	bovine serum albumin
°C	=	degree Celsius

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

CAD	=	carbamoylphosphate synthetase-aspartate transcarbamoylase -dihydroorotase enzyme complex
CPS II	=	carbamoylphosphate synthetase II
CDP	=	cytidine diphosphate
CTP	=	cytidine triphosphate
CTPS	=	cytidine triphosphate synthase
Da	=	dalton
DHO	=	dihydroorotase
DHOD	=	dihydroorotate dehydrogenase
DMS	=	3, 3´-dimethylsuberimidate
DNA	=	deoxyribonucleic acid
dCDP	=	deoxycytidine diphosphate
dCTP	=	deoxycytidine triphosphate
dTDP	=	deoxythymidine diphosphate
dTMP	=	deoxythymidine monophosphate
dTTP	=	dexythymidine triphosphate

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

dUDP	=	deoxyuridine diphosphate
dUMP	=	deoxyuridine monophosphate
DTT	=	dithiothreitol
EC	=	enzyme commission
EDTA	=	ethylene diamine tertraacetic acid
ESI	=	electrospray ionization
FPLC	=	fast protein liquid chromatography
FTICRMS	=	fourier transform ion cyclotron resonance mass spectrometry
FT-MS	=	fourier transform ion cyclotron
Gln	=	glutamine
h	=	hour
HCl	=	hydrochloric acid
HCO_3^-	=	bicarbonate ion
HPLC	=	high performance liquid chromatography

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

IPTG	=	isopropyl β-D-thiogalactopyranoside
k_{cat}	=	catalytic number
k_{cat}/K_m	=	catalytic efficiency
K_m	=	Michaelis constant
Kb	=	kilo base pair
kDa	=	kilo dalton
LB media	=	Luria-Bertani media
LC	=	liquid chromatography
LCR	=	low complexity region
Ld	=	*leishmania donovani*
Lys	=	lysine
M	=	molar
MALDI	=	matrix-assisted laser desorption
mg	=	milligram
Mg^{2+}	=	magnesium

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

$MgCl_2$	=	magnesium chloride
min	=	minute
ml	=	millilitre
mM	=	millimolar
mm	=	millimeter
MS	=	mass spectrometry
m/z	=	mass to charge ratio
NaCl	=	sodium chloride
ng	=	nanogram
NH_4HCO_3	=	ammonium bicarbonate
nl	=	nanolitre
nM	=	nanomolar
Ni^{2+}-NTA	=	nickel nitrilotriacetate
nmol	=	nanomole
NMR	=	nuclear magnetic resonance

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

OA	=	orotic acid
OD	=	optical density
OMP	=	orotidine 5’-monophosphate
OMPDC	=	orotidine 5’-monophosphate decarboxylase
OPRT	=	orotate phosphoribosyltransferase
ORF	=	open reading frame
PCR	=	polymerase chain rection
PDB	=	protein data bank
Pf	=	*Plasmodium falciparum*
pH	=	potential of hydrogen ion
Phe	=	phenylalanine
PP_i	=	pyrosphosphate
PRPP	=	5’-phosphoribosyl-1-pyrophosphate
Q-TOF	=	quadrupole-time of fight
RNA	=	ribonucleic acid

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

rpm	=	round per minute
SDS-PAGE	=	sodium dodecyl sulfate polyacrylamine gel electrophoresis
Sc	=	*Saccharomyces cerevisiae*
SCOP	=	structure classification protein
sec	=	second
spp.	=	species
$Temp_{1/2}$	=	temperature half
Thr	=	threonine
TOF	=	time-of-fight
UMP	=	uridine 5’-monophosphate
UMPS	=	uridine monophosphate synthase
UDP	=	uridine 5´-diphosphate
UTP	=	uridine 5’-triphosphate
UV	=	ultraviolet
v	=	volt

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

v/v	=	volume per volume
V_{max}	=	maximum velocity
w/v	=	weight per volume

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

CHAPTER I

INTRODUCTION

1.1. Background and Rationale

Malaria is a mosquito-borne infectious disease of humans in tropical zone and subtropical zone caused by four species of *Plasmodium*, including *Plasmodium falciparum*, *Plasmodium ovale, Plasmodium malariae* and *Plasmodium vivax.* New retrospective studies have also shown hundreds of human cases infected with simian *Plasmodium knowlesi* (1).

Malaria remains a major and growing threat to the public health of the population living in the endemic areas (2). One of the five species that infect humans, *Plasmodium falciparum* is the most lethal parasite (3). It is estimated that the disease afflicts 450 million and kills 1.5-2.7 million people each year (4). Chemotherapy of malaria is available but is complicated for the spread of the disease by both antimalarial drugs resistance, including the first-line drug artemisinin (5), and their toxicity (6). This highlights the priority development for more effective

and less toxic new drugs with different mechanism of action, against malaria infection (7, 8).

Pyrimidine nucleotides are essential metabolites. Unlike human and other mammalian cells (9), *P. falciparum* cannot salvage preformed pyrimidine bases or nucleosides from human host, but relies solely on nucleotides synthesized through the *de novo* pathway (10). The biosynthetic pathway involves six sequential enzymes catalyzing the conversion of following precursors: HCO_3^-, ATP, Gln, Asp, and 5-phosphoribosyl-1-pyrophosphate (PRPP), to form uridine 5'-monophosphate (UMP). There are some key differences of the enzymes in the pathway between malarial parasites and higher eukaryotes, including human host, for instance, the parasite enzyme inhibitors are found to have strong antimalarial activity for *P. falciparum* grown *in vitro* and *Plasmodium berghei* propagated in mice (11-20). The final two steps of UMP synthesis require the addition of ribose 5´-phosphate from PRPP to orotate by orotate phosphoribosyltransferase (EC 2.4.2.10, OPRT) to form orotidine 5'-monophosphate (OMP) and the subsequent decarboxylation of OMP to form UMP by OMP decarboxylase (EC 4.1.1.23, OMPDC). In human, both enzymes exist as a bifunctional UMP synthase (UMPS), and in bacteria they express separately as monofunctional enzyme. Interestingly, the enzymes in human malaria parasite

express as monofunctional forms but catalytically function as a bienzyme complex in the form of heterotetrameric (*Pf*OPRT)$_2$(*Pf*OMPDC)$_2$ (21, 22). The enzymes in the parasite contain unique insertion sequences at N-terminus and internal regions and have specific characteristic Low Complexity Region (LCR), repeat of single amino acid or group of amino acid, which are not found in homologous proteins from other organisms. At present, we have not yet known that amino acid sequence plays a role interact in heterotetrameric (*Pf*OPRT)$_2$(*Pf*OMPDC)$_2$.

1.2. Research question

What are the amino acid sequences play a role in interaction between *Pf*OPRT and *Pf*OMPDC in the parasite?

1.3. Hypothesis

The unique insertion of the *Pf*OPRT and *Pf*OMPDC is responsible for the *Pf*OPRT-*Pf*OMPDC interaction in the parasite.

1.4. Objective

To identify protein-protein interaction in the heterotetrameric (*Pf* OPRT)$_2$(*Pf* OMPDC)$_2$ complex of the parasite.

1.5. Conceptual framework

Identification amino acid sequence, that play a role in interaction between molecule in dimeric *Pf*OPRT, dimeric *Pf*OMPDC and tetrameric $(PfOPRT)_2(PfOMPDC)_2$

↓

Amplification, expression and purification of recombinant proteins

↓

Cross-link between *Pf*OPRT and *Pf*OMPDC with 3, 3'-dimethylsuberimidate (DMS) to generate dimeric *Pf*OPRT, dimeric *Pf*OMPDC and tetrameric $(PfOPRT)_2(PfOMPDC)_2$ and separate cross-link product by SDS-polyacrylamind gel electrophoresis

↓

Identification of amino acid sequence used to interact between molecules by liquid chromatography-mass spectrometer

↓

Homology modeling of *Pf*OPRT and *Pf*OMPDC to generate dimeric and tetrameric enzyme

1.6. Key words: Orotate phosphoribosyltransferase; Orotidine 5'-monophosphate decarboxylase; Multienzyme complex; Pyrimidine biosynthesis; *Plasmodium falciparum*; Malaria

1.7. Expected Benefits

1.7.1 Understanding of the role of N-terminal insertion, internal insertion, and LCRs in *Pf*OPRT and *Pf*OMPDC will gain new information the protein-protein interaction.

1.7.2 Understanding of dimeric *Pf*OPRT, dimeric *Pf*OMPDC and tetrameric $(PfOPRT)_2(PfOMPDC)_2$ interaction will lead to design a new antimalarial drugs in the future.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

CHAPTER II

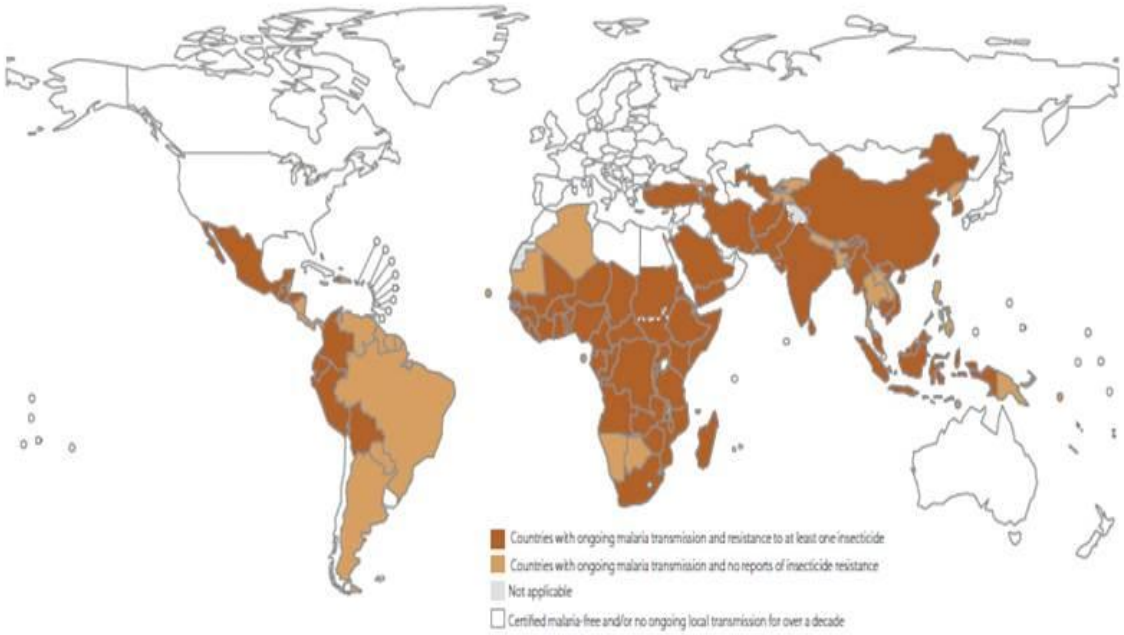
LITERATURE REVIEWS

The majority of malarial related mortality occurs in children below the age of 5 years, mostly in sub-Saharan region of Africa, due to their lack of naturally acquired malarial immunity. The majority of human malaria infection is caused by either *P. falciparum* or *P. vivax* and most malaria-associated deaths are due to *P. falciparum*. Of the two most common species of human malaria, *P. falciparum* is the most pathogenic, with approximately 1% of human infections giving rise to severe (life-threatening) disease. Severe disease encompasses a range of presentation including severe anemia, cerebral malaria, hypoglycemia and a systemic syndrome analogous to toxic shock (21). The malaria transmission area in the world distributes mostly in tropical zone and sub-tropical zone, such as Africa, South America and Asia (23) (Figure 1).

2.1 Human malaria parasite life cycle

The life cycle of the *Plasmodium* have sexual and asexual phases in the two hosts (Figure 2). The asexual phase, in human the parasites grow and multiply first in

the liver cell and then in the red blood cell. The sexual phase, in female *Anopheles* mosquito carrying the malaria parasites from human blood grow and then multiply in different stages in lumen of stomach and salivary glands. When the *Anopheles* mosquito takes a blood meal on another human, the sporozoites are injected with the mosquito's saliva and this will start another human infection.



Source: Adapted from Global Plan for Insecticide Resistance Management in malaria vectors, WHO, Geneva, 2012. From WHO regional entomologists in WHO Regional Offices and literature review by the Global Malaria Programme. Map production: Global Malaria Programme (GMP), World Health Organization

CHULALONGKORN UNIVERSITY

Figure 1 The world map showing the area of malaria. Malaria is presently endemic in a broad band around the equator, in areas of the Americas, many parts of Asia, and much of Africa; in Sub-Saharan Africa. This map is taken from World malaria report 2012 by World Health Organization (23).

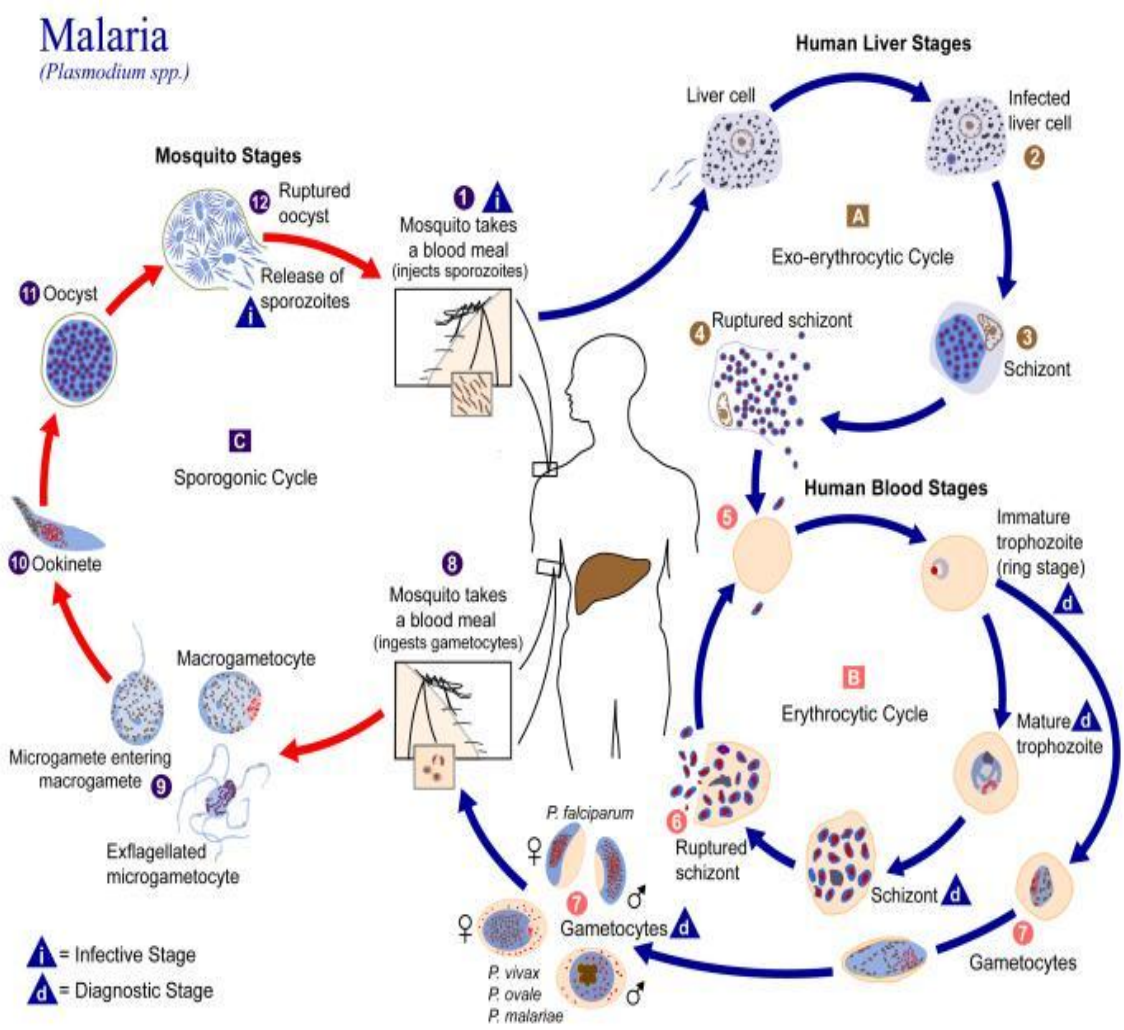


Figure 2 Life cycle of human malaria. Malaria parasites have sexual and asexual phases in two hosts, sexual phase in female *Anopheles* mosquito and asexual phase in human. (From http://www.dpd.cdc.gov/dpdx/HTML/Malaria.htm)

In mosquito's saliva gland, if it has sporozoites, during feeding, it secretes sporozoites into human bloodstream. After approximately one hour the sporozoites disappears from the circulation, 24 to 48 hours later in the parenchymal cell of the liver, sporozoites mature into schizonts, where the exo-erythrocytic schizogonic phase begins. The schizonts rupture out of liver cell in term merozoites. The merozoites enter the blood circulation and invade red blood cell, the merozoites grown to the tropozoites, the tropozoites mature into schizonts, initiating the erythrocytic schizogonic phase, which rupture releasing merozoites. Immature tropozoites differentiate into sexual erythrocytic phase (gametocytes), male (microgametocyte) and female (macrogametocyte).

The sexual phase occurs in the female *Anopheles* mosquito and initiate when the mosquito takes a blood meal that contains microgametocyte and macrogametocyte. In the lumen of mosquito's stomach, lysis of the erythrocytic material releases gametocytes. There, microgametocytes undergo a maturation process known as exflagellation, during including spindle formation and its nucleus undergoes three mitotic divisions, producing eight nuclei that migrate to the periphery of gametocytes. In cytoplasm, centriolar division generates eight filaments, following which one portion join each nucleus. This is completing the formation of

microgametes. During this period, the macrogametocytes undergo reducing nucleus division to become macrogamete. Microgametes swim to macrogametes, when the haploid microgamete fuses a haploid macrogamete, produces a diploid zygote. After 12 to 48 hours fertilization the zygotes, known as ookinetes, penetrate to the gut wall, the area between epithelium and basal lamina, where develop into oocyst. The oocysts repeat nucleus division, in which rates depend on *Plasmodium* species, in the part of proliferation of haploid cell, in term of sporoblasts, producing thousands of sporozoites. As membrane rupture, sporozoites carrying to the salivary gland and are ready to be inject to next human person (24-26).

2.2 Genomic of malaria parasite

Since 1996, the malaria genome project was initiated to determine the nucleotide sequence of *P. falciparum* genome by the Institute for Genomic Research, the Wellcome Trust Sanger Institute and Stanford Genome Technology Center and completed in 2002 (27). In 2005, the genome of *P. yoelii* was published (28) and the genome sequence of *P. vivax* and *P. knowlesi* were then completed (29, 30).

The genome of *P. falciparum* has 14 chromosomes and the genes encoding proteins are approximately 5,300 genes (27). The *P. falciparum* genome contains 23

million bases, they are rich in low complexity region (LCR) since nucleic acid level correlates with LCR at protein level (31). LCR are characterized by single amino acid or group of amino acids repeat, almost 50% of *Plasmodium* protein are longer than their yeast homologues, contain the insert segments, these inserts correspond to LCR (32). The genome of *P. falciparum* is extremely AT-rich (82% A+T) (33), *P. vivax* genome shows the lower level of AT (67.7% A+T), however, the richness AT in both species has probably not too much to do the disease (34).

2.3 Pyrimidine biosynthesis in human

In human pyrimidine nucleotides are synthesized by two routes: *de novo* pathway and salvage pathway**.** It is the basis of three of the bases (thymine, uracil and cytosine) found in DNA and RNA.

2.3.1 Synthesis of pyrimidine nucleotides by *de novo* pathway

De novo pyrimidine synthesis requires aspartate (Asp), glutamine (Gln) and HCO_3^- as precursors. Pyrimidines are assembling the single pyrimidine ring before being attached to 5- phosphoribosyl-1-pyrophosphate (PRPP). The first step in *de novo* pyrimidine biosynthesis is the formation of carbamoyl phosphate by carbamoyl

phosphate synthetase II (CPS II). Carbamoyl phosphate is condensed with aspartate in a reaction catalyzed by the aspartate carbamoyl transferase (ATC), which is then cyclized to dihydroorotate by dihydroorotase (DHO). The activities of CPS II, ATC and DHO are found on a trifunctional protein (namely, CAD). Dihydroorotate enters the mitochondria, and then converts into orotate by dihydroorotate dehydrogenase (DHOD). DHOD is located on the outer surface of the inner mitochondrial membrane. This is the only mitochondrial step in the nucleotide ring biosynthesis. Orotate is then converted to orotidine 5´-monophosphate (OMP) by orotate phosphoribosyltransferase (OPRT). Then OMP is decarboxylated to uridine 5´-monophosphate (UMP) by orotidine 5´-monophosphate decarboxylase (OMPDC). The final two steps in the pathway, OPRT and OMPDC activities are found on a bifunctional protein, also known as UMP synthetase (UMPS). Five of six reactions occur in the cytosol of the cell, whereas the DHOD reaction occurs in mitochondria.

Nucleotide kinases convert UMP to uridine diphosphate (UDP) and triphosphate (UTP). UTP is the direct substrate for CTP synthetase (CTPS), formation of cytidine triphosphate (CTP) requires Gln and ATP (35, 36). CDP is converted to its deoxyribonucleotide derivative by ribonucleotide reductase. UDP is s precursor for synthesis of thymidine nucleotide. UDP is also converted to dUDP by ribonucleotide

reductase. dUDP is dephosphorylated to dUMP. dUMP is then converted to deoxythymidine monophosphate (dTMP) by thymidylate synthase, requiring transfer of a single carbon from 5,10-methylenetetrahydrofolate (37) (Figure 3).

2.3.2 Synthesis of pyrimidine nucleotide by salvage pathway

Salvage pathway is used to recover bases and nucleotides that are formed during degradation of RNA and DNA, beginnig with deoxycytidine, deoxyuridine, or deoxythymidine nucleosides, which are each converted to nucleoside monophosphates in the first step by appropriate kinases. Uridine phosphorylase has an activity to inter-convert uracil, uridine and deoxyuridine, whereas uridine kinase phosphorylates the nucleoside into UMP. Deoxythymidine phosphorylase adds deoxyribose-1-phosphate to thymine to generate dTMP, and thymidine kinase then phosphorylates this nucleoside into dTDP and dTTP.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

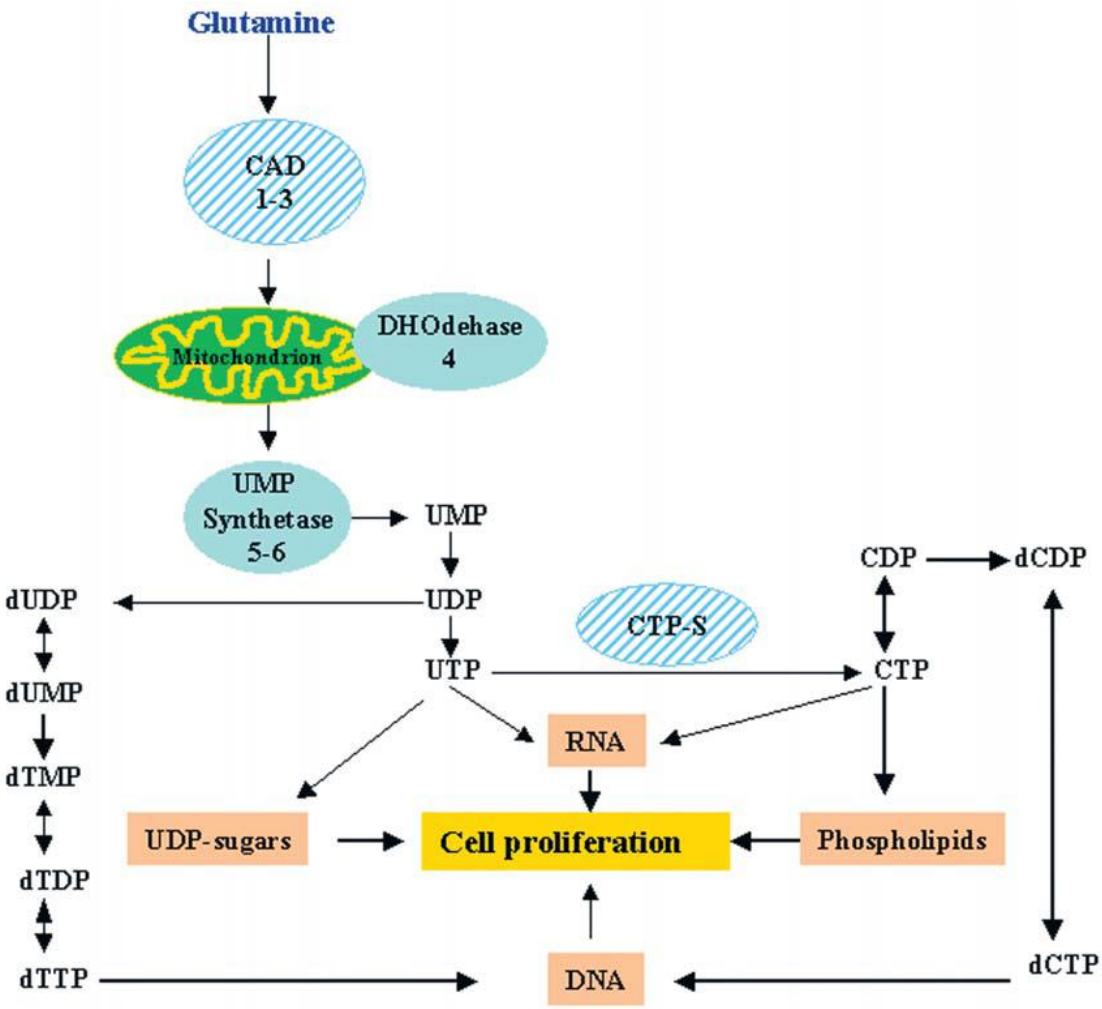


Figure 3 Pyrimidine *de novo* biosynthesis in human. The diagram shows the *de novo* pathway of pyrimidine biosynthesis in human. The *de novo* pyrimidine biosynthesis pathway requires six enzymes. The first three enzymes are catalyzed on trifunctional protein, a cytosolic enzyme known as CAD. The fourth enzyme is DHOD or DHOdehase, a mitochondrial enzyme. The last two enzymes are catalyzed by a bifunctional protein, known as UMPS, a cytosolic enzyme. CTP-S stands for CTP synthetase (37).

2.4 Pyrimidine biosynthesis in *Plasmodium* parasite

Human and malaria parasites have a different route of pyrimidine synthesis, human obtain their pyrimidines from 2 routes: salvage of the preexisting compounds and *de novo* synthesis, but the parasites have only *de novo* pathway. Because it lacks enzymes in the salvage pathway (9, 10, 38). The *de novo* pathway contains the six enzymes, using HCO_3^-, ATP, Gln, Asp, and PRPP as precursors. All enzymes in the *de novo* pathway of *P. falciparum* are encoded from the separate genes, in order from the first to the sixth sequential step; these are CPS II, ATC, DHO, DHOD, OPRT and OMPDC (39) (Figure 4).

The final two steps, OPRT and OMPDC in the malaria parasites exist as multienzyme complex containing two subunits of OPRT and two subunits of OMPDC (21). So its organization is different from human, in that both enzymes express in bifunctional UMPS and in most higher or multicellular eukaryotes both enzymes also exist as found in human (40-42). In many prokaryotes and a single cell eukaryote yeast, OPRT and OMPDC are encoded by two independent genes (40, 42, 43).

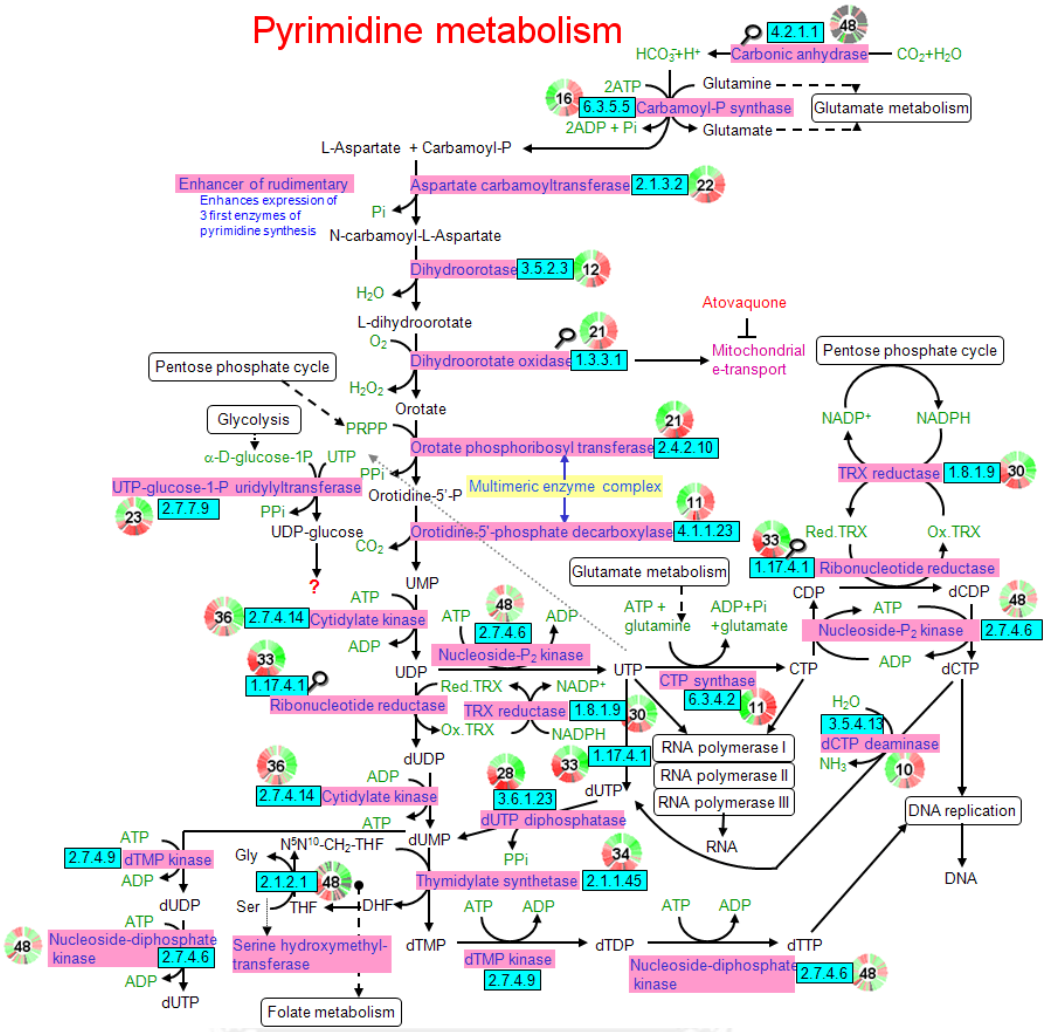


Figure 4 The *de novo* pyrimidine biosynthesis of malaria parasites. HCO_3^-, ATP, Gln, Asp, and 5-phosphoribosyl-1-pyrophosphate (PRPP) as precursors of *de novo* pyrimidine biosynthetic pathway. The fifth and the sixth enzymes in pyrimidine biosynthetic pathway in *Plasmodium* are active in multimeric enzyme complex. The *de novo* pyrimidine pathway is linked to other metabolic pathways, such as glycolysis, pentose phosphate pathway, folate and glutamate metabolism. (From http://sites.huji.ac.il/malaria/maps/pyrimidinemetpath.html).

2.5 Orotate phosphoribosyltransferase (OPRT)

Orotate phosphoribosyltransferase (OPRT) is the fifith enzyme in the *de novo* pyrimidine biosynthesis pathway. The novel form of the OPRT in *P. falciparum* (*Pf*OPRT) is a monofunctional protein. The open reading frame of *Pf*OPRT is located on chromosome 5 locus PFE0630C. The *Pf*OPRT gene has 846 bp and only one exon, encoding for 281 amino acids, with the molecular mass of about 33 kDa (16). The active form is dimeric enzyme with the molecular mass of about 66 kDa (16). The malaria parasite enzyme contains an extension of 66 amino acids at the N-terminus and unique insertion of 19 amino acids at 178-196 with high hydrophobic index of +1.0- +3.0 (16).

The *Pf*OPRT is classified in type I phosphoribosyltransferase, binding to PRPP, orotate and Mg^{2+}. The EC number of OPRT is EC 2.4.2.10. It catalyzes orotate + PRPP $\rightleftharpoons$ OMP +PP_i. The active site amino acids consist of Lys89, Phe97, Phe98, Thr212 and Arg241, which interact by extended hydrogen bonds among orotate, PRPP, surrounding amino acid residues and water molecules (44).

Similarity of sequence alignment of *Pf*OPRT comparing among *P. berghei*, *P. knowlesi*, *P. vivax*, *S. cerevisiae*, *E. coli and H. sapiens* are 65%, 62%, 56%, 30%, 22% and 18%, respectively, according to ClustalW2 calculation (45) (Figure 5).

The 3D structure of OPRT in protein data bank (PDB) is now deposited for 19 structures, containing three of *Salmonella enterica*, three of *S. cerevisiae* (*Sc*OPRT), two of *H. sapiens*, two of *Bacillus anthracis*, each of *Leishmania donovani*, *L. infuntum*, *E. coli* (*Ec*OPRT), *Vibrio cholerae*, *Streptococus mutans*, *S. pyogenes*, *Francisella tularensis*, *Aeropyrum pernix*, and *Corynebeacterium diphtheriae*. All of them are searched by using the EC number of the enzyme. The *Pf*OPRT shows the most similar to the *Sc*OPRT by having 30% identity. So the 3D structure of *Sc*OPRT, as template, is important to demonstrate a homology model of *Pf*OPRT.

The 3D structure of *Sc*OPRT (PDB ID: 2PS1), which is also a typical type I phosphoribosyltransferase family and exists as a homodimeric form, contains 226 residues and having secondary structures of seven α-helices and 10 β-sheets (Figure 6) (46). In addition, the dimer interface found in both *Ec*OPRT and *Sc*OPRT contains flexible loop, as a part of the active site (47).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

```
P.falciparum      -------MTTIKENEFLCDEEIYKSFVHLKDKICEERKKKELVNNNIDNVNFNDDDDNNY 53
P.knowlesi        ----MEEEKSTKQTNYLSDEDLHKKYIHLRECIELE----------------KDDMD--- 37
P.vivax           MLNRMEGEKSINQTNYLSDEDLHKKYTQLRECIELE----------------KDDQ---- 40
P.berghei         ----MD-ENNKETKN--IDEELHKKYNELCKRIELG----------------NDNKN--- 34
S.cerevisiae      ------------------------------------------------------------
E.coli            ------------------------------------------------------------
H.sapiens         ------------------------------------------------------------

P.falciparum      DDDGNSYSSYIKEMKKILKVVLLKYKALKFGEFILKSKRKSNYFFSS-GVLNNIVSSNII 112
P.knowlesi        -------NCHVKEMKKLLITALLKYKAIKFGDFILKSKRKSKYFFSS-GVLNNIVSAHII 89
P.vivax           -------NSHVKEMKSLLITALLKYNAIKFGDFILKSKRKSKYFFSS-GVLNNIVSAHII 92
P.berghei         -------CDDIRKMKKLLIDALIKYEAIKFGNFILKSKRKSKYFVST-GFLNNAISANIV 86
S.cerevisiae      ------MPIMLEDYQKNFLELAIECQALRFGSFKLKSGRESPYFFNL-GLFNTGKLLSNL 53
E.coli            ----------MKPYQRQFIEFALGKQVLKFGEFTLKSGRKSPYFFNA-GLFNTGRDLALL 49
H.sapiens         -------MAVARAAALGFLVTGLYDVQAFKFGDFVLKSGLSSPIYIDLRGIVSRPRLLSQV 53

P.falciparum      CFLLSELILKNKLSFDYLLGASYKGIPMVSLTSHFLFESKK--YSNIFYLYDRKEKKEYG 170
P.knowlesi        SFLISHLILKEKIPFDYLLGASYKGIPIATLTSHFLFQSNK--FSNVFYLYDRKEKKDYG 147
P.vivax           SFLISHLILKEKIPFDYLLGASYKGIPIATLTSHFLFRSNK--FANVFYLYDRKEKKDYG 150
P.berghei         SFLISNLILSKNIAFDYLFGASYKGIPIVSLTSHFLLNTNK--FHNIFYLYDRKEKKEYG 144
S.cerevisiae      ATAYAIAIIQSDLKFDVIFGPAYKGIPLAAIVCVKLAEIGGSKFQNIQYAFNRKEAKDHG 113
E.coli            GRFYAEALVDSGLEFDLLFGPAYKGIPIATTTAVALAEHHD---LDLPYCFNRKEAKDHG 106
H.sapiens         ADILFQTAQNAGISFDTVCGVEYTALPLATVICSTN---------QIPMLIRRKETKDYG 104

P.falciparum      DKNVIVGNLDDDDKDILNLKKKTKNNQDEEKKNIIIIDDVFTCGTALTEILAKLKTYEHL 230
P.knowlesi        DKTLIVGNLDEEFNGDV---HNAKDEKNEKK--VIIIDDVFTCGTALTEIMNKLKSYPNL 202
P.vivax           DKTLIVGNLDEAVNGDV---HNAKEAESEKK--VIIIDDVFTCGTALTEIINKLKSYPNL 205
P.berghei         DKTIIVGNIKESSQDCV---INSCNPQFEKKKKVIIIDDVFTCGTALTEIFNKMKAYEYL 201
S.cerevisiae      EGGIIVG--------------------SALENKRILIIDDVMTAGTAINEAFEIISNAK-G 153
E.coli            EGGNLVG--------------------SALQG-RVMLVDDVITAGTAIRESMEIIQANG-A 145
H.sapiens         TKRLVEG--------------------TINPGETCLIIEDVVTSGSSVLETVEVLQKEG-L 144

P.falciparum      KVVAFIVLLNRNEYEINENNQKIYFKDIFEKRVGIPLYSILSYKDDIQSMI--------- 281
P.knowlesi        KVVAFIVLLNRNEYELNEHNEKVYFKDLFEQKLKVPLYSILSYNEDLEPLMG-------- 254
P.vivax           RVVALIVLLNRNEYELNEQNEKVYFKDLFEQKLNVPLYSILSYHEDLEPLMG-------- 257
P.berghei         QVVACIVLLNRNEHEINENNEKVYFKDLFEQKYNIPIYSILNYEEDISHFIK-------- 253
S.cerevisiae      QVVGSIIALDRQEVVSTDDKEGLSATQTVSKKYGIPVLSIVSLIHITTYLEGR-ITAEEK 212
E.coli            TLAGVLISLDRQERGRGE----ISAIQEVERDYNCKVISIITLKDLIAYLEEKPEMAEHL 201
H.sapiens         KVTDAIVLLDREQ--GGKDKLQAHGIRLHSVCTLSKMLEILEQQKKVDAETVGRVKRFIQ 202

P.falciparum      ----------------
P.knowlesi        ----------------
P.vivax           ----------------
P.berghei         ----------------
S.cerevisiae      SKIEQYLQTYGASA-- 226
E.coli            AAVKAYREEFGV---- 213
H.sapiens         ENVFVAANHNGSPLSI 218
```

Figure 5 Sequence alignment of OPRTs. A sequence alignment of seven OPRTs was performed using ClustalW2 (45). Species name and accession number are as follows: *P. falciparum* OPRT (PFE0630C), *P. knowlesi* (XP_002259686.1), *P. vivax* (XP_001613829), *P. berghei* (PBANKA_111240), *S. cerevisiae* (NP_013601), *E. coli* (X00781) and *H. sapiens* (NP_000364).The completed conserved residues are shown with a yellow background while the similar residues are shown with a pink box.

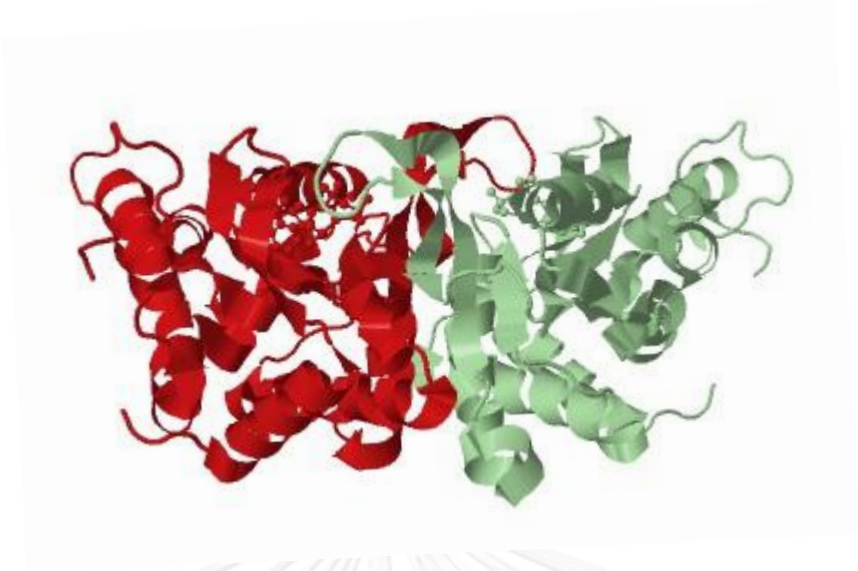


Figure 6 The 3D structure of *S. cerevisiae* OPRT. Homodimeric OPRT of *S. cerevisiae* has subunit A (red) and subunit B (green), PDB ID: 2PS1 (46).

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

2.6 Orotidine 5´-monophosphate decarboxylase (OMPDC)

Orotidine 5´-monophosphate decarboxylase (OMPDC) is the sixth enzyme in the *de novo* pyrimidine biosynthesis pathway, which catalyzes decarboxylation of OMP to UMP. The EC number of the enzyme is EC 4.1.1.23. The novel form of the OMPDC in *P. falciparum* (*Pf*OMPDC) is a monofunctional protein. The open reading frame of *Pf*OMPDC is located on chromosome 10 locus PF10_0225. The *Pf*OMPDC gene contains 969 bp and only one exon, encoding for 323 amino acids with the molecular mass of about 38 kDa. The active form is dimeric enzyme with the molecular mass of about 76 kD (22). In addition, the malaria parasite enzyme has also an extension of 32 amino acids at the N-terminus (Figure 7).

The alignment of *Pf*OMPDC with 6 known crystal structures (Figure 7) shows a large insertion at the N-terminal region in all *Plasmodium*. It has been proposed that these regions are responsible for binding between *Pf*OPRT and *Pf*OMPDC (21, 22, 48).

The 2.7 A° resolution crystal structure of *Pf*OMPDC has been recently identified (Figure 8) (48). The *Pf*OMPDC 3D structure folds as an $(\alpha/\beta)_8$ barrel, with eight β-sheets and 13 α-helices. The dimer interface consists a network of 10 hydrogen bonds. The active site is located at the open end of the $(\alpha/\beta)_8$ barrel (48).

```
                                                α1          β1                α2
                                                ^^^^^^^^^^^^^^  ---->  ^^^^^^^^^^^^^^
P.falciparum     ------------------MGFKVKLEKRRNAINTCLCIGLDPDEKDIENFMKNEKENNYN 42
P.knowlesi       ------------------MNLKTKLQKRRDEVNTCLCIGLDPDEADIKSFMQSEKQNGYQ 42
P.vivax          MGSSHHHHHHSSGLVPRGSNLKIKLQKRRDEVNTCLCIGLDPDEADIKSFMQSEKQNGYQ 60
P.berghei        MGSSHHHHHHSSGLVPRGSHFKTKLKNRRNEVNTCLCIGLDPDEDDIKNFMRNEEKNGYK 60
S.cerevisiae     ---------------------------------------------------MHKATYKER-- 9
E.coli           ------------------------------------------------------------
H.sapiens        ----------------------DAETVGRVKRFIQENVFVAANHNGSPLSIKEAPKELSF 38
                  α3                    η1        α4                  α5
                 ^^^^^^^                ^^^^^^   ^^^^^^^            ^^^^^^^^^^^^^^^
P.falciparum     NIKKNLK--------------EKYINNVSIKKDILLKAPDNIIREEKSEEFFYFFNHFCF 88
P.knowlesi       SIKKNLSNHGSSS--------QGGLFTPQVGGTMLLAENPPKE-VQEKDEFFYFFNHFCF 93
P.vivax          SVKKNLSNSGSSSSSSSNSSSGKGELFAPQMGGQMLLAETPPKE-AQEKDEFFYFFNHFCF 119
P.berghei        NVKNNMNS-----------NNNRIENVIKIGKEILLTDEENIENLSEEDKFFYFFNHFCF 109
S.cerevisiae     --AATHP------------------SPVAAKLFNIMHEKQTNLCASLDVRTTKELLELVE 49
E.coli           ------------------------MTLTASSSSRAVTN--SPVVVALDYHNRDDALAFVD 34
H.sapiens        GARAELPR----------------IHPVASKLLRLMQKKETNLCLSADVSLARELLQLAD 82
                                  β2    η2          α6                  β3
                                 ^^^^^^^  ---->  ^^^^  ^^^^^^^^^^^^^^^^^^^^^^^^  ---->  ^^^
P.falciparum     YIINETNKYALTFKMNFAFYIPYGSVGIDVLKNVFDYLYELNIPTILDMKINDIGNTVK- 147
P.knowlesi       YIINETKQYALSYKMNFAFYLPYGSLGVDVLKNVFDYLHYLDVPTILDIKMNDIGNTVK- 152
P.vivax          YIINETKEYALAYKMNFAFYLPYGSLGVDVLKNVFDYLHHLNVPTILDIKMNDIGNTVK- 178
P.berghei        YIINNTKEYALIYKMNFAFYIPYGSVGINALKNVFDYLNSMNIPTMLDMKINDIGNTVK- 168
S.cerevisiae     ALGPKICLLKTHVDILTDFSMEGTVKPLKALSAKYNFLLFE------DRKFADIGNTVKL 103
E.coli           KIDPRDCRLKVGKEMFTLFGPQFVRELQQRGFDIFLDLKFHDIP-------NTAAHAVAA 87
H.sapiens        ALGPSICMLKTHVDILNDFTLD-VMKELITLAKCHEFLIFE------DRKFADIGNTVKK 135
                  α7                β4          α8                  β5          α9
                 ^^^^^^^  ---->  ^^^^^^^^^^^^^^^^^^^^^^  ---->  ^^^^^^^^^^^^^^^^^^^
```

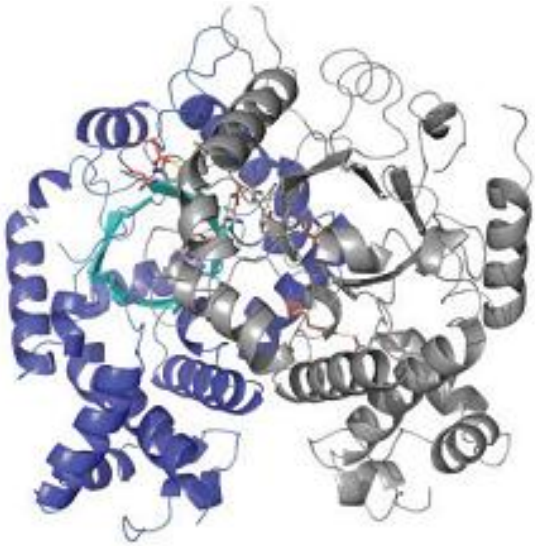


Figure 8 The 3D structure of *P. falciparum* OMPDC. Homodimeric *Pf*OMPDC contains subunit A (blue) and subunit B (gray), PDB ID: 2ZA2 (48).

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

2.7 *Pf*OPRT and *Pf*OMPDC enzyme complex

The single open reading frames of *PfOPRT* and *PfOMPDC* genes encode proteins with 281 and 323 amino acid residues, respectively (22). Both enzymes from *P. falciparum* have been purified and found to be a multienzyme complex with the molecular mass of 140 kDa, containing two molecules of each *Pf*OPRT (33 kDa) and *Pf*OMPDC (38 kDa) monofunctional form (22). This enzyme complex forms $\alpha_2\beta_2$ conformation, where α and β are *Pf*OPRT and *Pf*OMPDC respectively (Figure 9) (22). The heterotetrameric form is then represented as (*Pf*OPRT)$_2$(*Pf*OMPDC)$_2$.

Indeed, *Pf*OPRT and *Pf*OMPDC are encoded as stand-alone genes from chromosome 5 and 10, respectively. The *OPRT-OMPDC* gene fusion is present in the majority of eukaryotic groups, i.e. metazoa, amoebozoa, plantae, heterolobosea and euglenoids. On the other hand, they exit as separate genes in fungi, malaria parasite and most bateria (41). Interestingly, the inversely linked *OMPDC-OPRT* gene fusion is demonstrated in the kinetoplastids of the parasitic protists, e.g. *Leishmania spp.*, *Trypanosoma spp*. (41), representing an independent fusion event of both genes during evolution. The *OMPDC-OPRT* gene fusion is also found in some stramenopiles (diatom) and cyanobacteria as well (41).

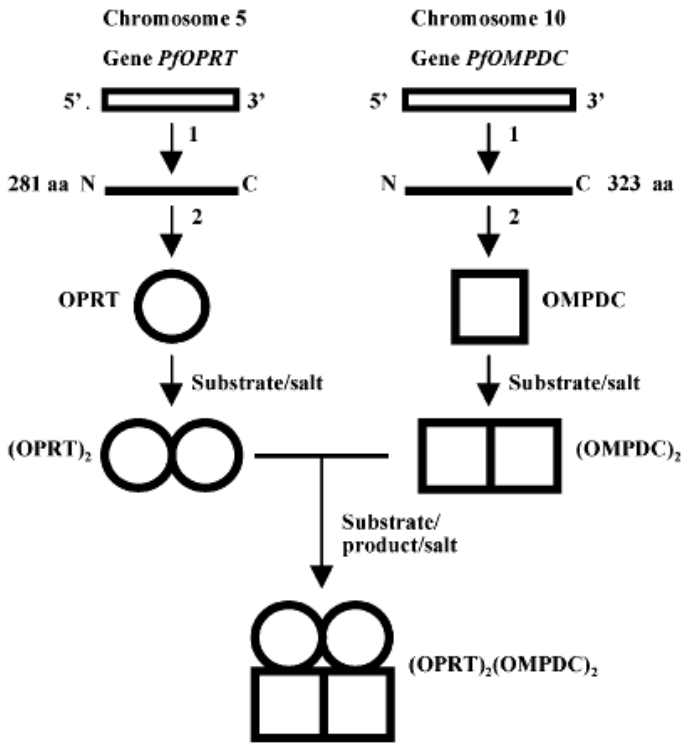


Figure 9 Model of multienzyme $(OPRT)_2(OMPDC)_2$ complex formation in *P. falciparum*. The OPRT and OMPDC are synthesized separately as inactive monomer, each individual enzyme forms a less active homodimer. Both enzymes are more active functioning in a tightly associated heterotetrameric $(OPRT)_2(OMPDC)_2$, as $\alpha_2\beta_2$ complex (22).

In the previous study, co-expression of *Pf*OPRT-*Pf*OMPDC as enzyme complex was carried out in *E. coli*, showing that the molecular mass of active *Pf*OPRT and *Pf*OMPDC enzyme complex was found to be 130.34 kDa in good agreement with the calculated mass of the $(Pf\text{OPRT})_2(Pf\text{OMPDC})_2$ tetramer. K_m, k_{cat}, and catalytic efficiency (k_{cat}/K_m) for orotate and PRPP of *Pf*OPRT in the enzyme complex were higher than the monofunctional form. In case of *Pf*OMPDC, K_m, k_{cat}, and k_{cat}/K_m for OMP was not significantly different from its monofunctional form (49, 50). In the part of enzyme inhibition, 6-aza UMP inhibited both the monofunctional *Pf*OMPDC form and the *Pf*OMPDC component in the enzyme complex but it likewise inhibited *Pf*OPRT in the enzyme complex, but did not inhibit the monofunctional *Pf*OPRT form. PP_i, a product of OPRT reaction, was found to be a good competitive inhibitor of *Pf*OMPDC in the enzyme complex but had not inhibitory effect in the monofunctional *Pf*OMPDC form and exhibited greater inhibitory effect to *Pf*OPRT in the enzyme complex than the monofunctional form. In addition, thermal stability characteristic was determined by calculating $temp_{1/2}$ (temperature at 50% activity remaining), the $temp_{1/2}$ of both components in the enzyme complex were higher than monofunctional forms (49). In deep, this study of the properties of the enzyme complex indicates that (1) it has kinetic benefit in both catalysis and inhibiton, (2) it

increases thermostability, and (3) the formation of enzyme complex stabilizes both components.

Nevertheless to the heterotetrameric formation of *Pf*OPRT and *Pf*OMPDC was reported for the first time in *P. falciparum* (22, 49). Most recently, in *L. donovani* it was shown that the *Ld*OPRT and *Ld*OMPDC were encoding from one gene, leading to a bifunctional protein *Ld*UMPS (41). The *Ld*OMPDC component was at the N-terminus, as what found in *Trypanosoma spp.*, whereas OMPDC in mammalian UMPS was at the C-terminus (41, 51, 52). After expression, the first step of oligomerization was a formation of tight dimer at the *Ld*OMPDC component (Figure 10B). This dimer was stable but did not catalyze the *Ld*OPRT reaction because the *Ld*OPRT had not yet to be actively dimer form. In the present of a ligand (e.g. OMP, UMP) a conformational change was occurred, and promoted dimerization of *Ld*OPRT component, leading to complete and active *Ld*UMPS tetramer (41) (Figure 10C).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

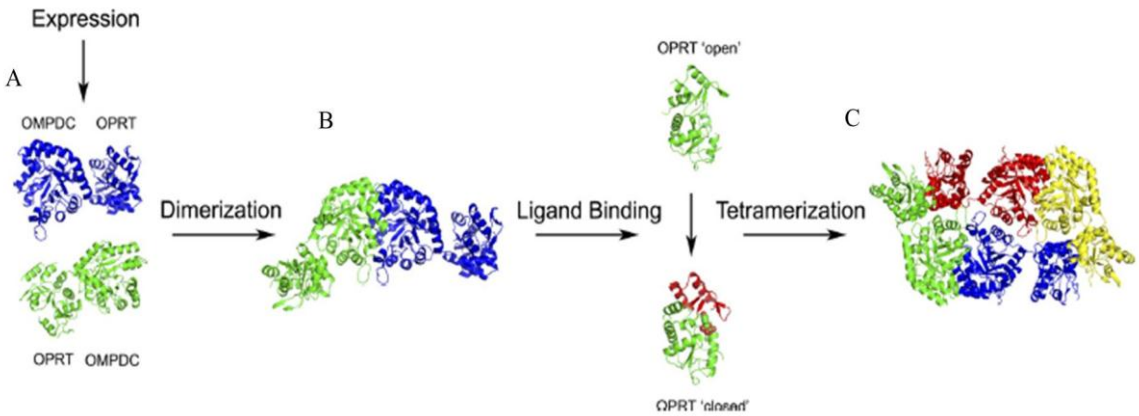


Figure 10 Model of UMPS oligomerization in *L. donovani*. A: the *Ld*UMPS monomer with the two components, OPRT and OMPDC. B: the asymmetric unit contains two monomers, formation of a dimer at the OMPDC component. C: tetramer formation after ligand binding of *Ld*UMPS (41).

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

2.8 Mass spectrometry for protein complex interaction

Mass spectrometry (MS) is an analytical technique for the determining element composition of a sample or molecule, which works by ionization chemical compounds to generate charged molecules or molecule fragment, and measuring their mass-to-charge ratio. A mass spectrometer consists of three components: an ion source (e.g. electrospray ionization (ESI) and matrix-assisted laser desorption (MALDI)), a mass analyzer (e.g. ion trap, time-of-fight (TOF), quadrupole, fourier transform ion cyclotron (FT-MS) analyzer) and a detector (e.g. photomultiplier, microchannel plate, electro multiplier) (53).

In general, MS based methods are applicable to soluble protein or protein complex. It is a powerful tool to identify protein in various ways, such as amino acid residues, protein structure and interaction in protein complex (54).

The powerful methods for prove the location of protein-protein interaction are X-ray crystallography, nuclear magnetic resonance (NMR) spectroscope and MS (55). MS method is the high sensitivity and speed for the analysis. It applies to soluble protein or protein complex, especially when sample is limited (55, 56). Application of MS in combination with chemical cross-link has been used to identify the cross-link residues (55).

Chemical cross-link is a process of covalent joining between molecules by a cross-link reagent (56). This method is established for structure information of protein complex and followed by MS analysis (54, 56) (Figure 11).

More recent studies, the 3D structure of calmodulin-melittin complex was performed by chemical cross-link and high resolution fourier transform ion cyclotron resonance mass spectrometry (FTICRMS). The cross-link products on SDS-PAGE, and mass spectra of non-digestion were analyzed after incubation with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride combination with *N*-hydroxysulfosuccinimide (EDC/sulfo-NHS) (Figure 12). The SDS-PAGE gel is digested with trypsin, the cross-link peptides are resistant to the trypsin digestion. The cross-link peptides is further determined by nano-ESI-FTICRMS and then identified by General Protein Mass Analysis for Windows (GPMAW). The expasy proteomic tools in the Swiss-Port database and Automatic Spectrum Assignment Program (ASAP) elucidated 12 cross-link peptides observed for the calmodulin-melittin complex (57).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

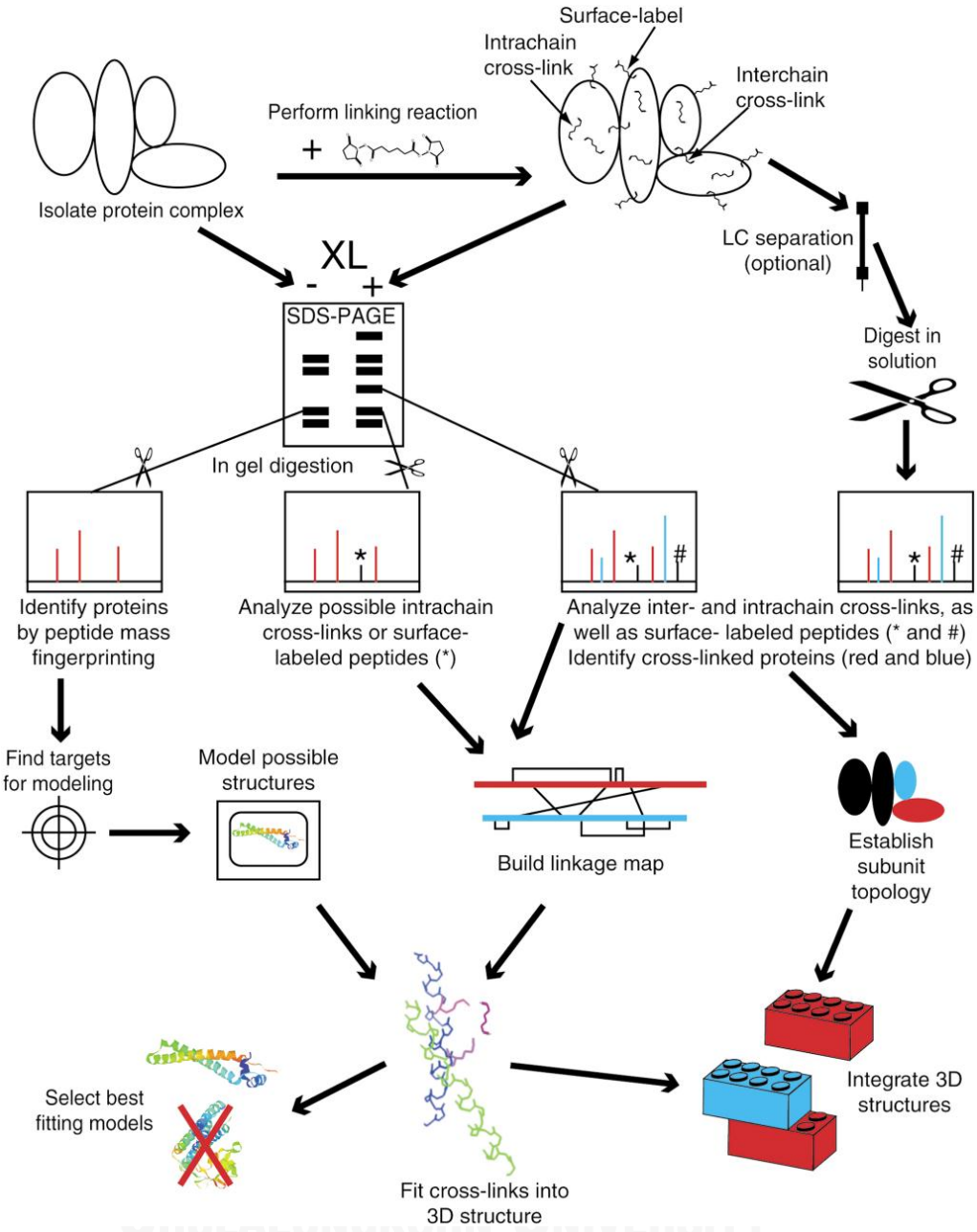


Figure 11 Schematic workflow of cross-link residue analysis for 3D structure information. Experimental steps are demonstrated to prove cross-link residue by MS method in combination with chemical cross-link, and the information for MS data are then used to generate 3D structure (56).

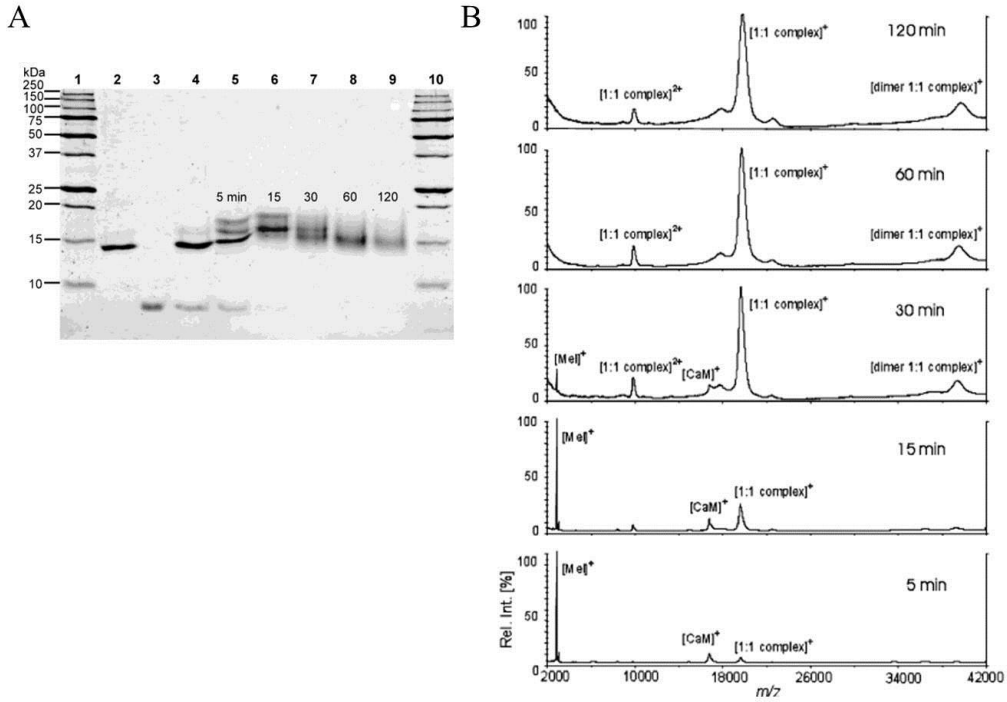


Figure 12 Cross-link product of calmodulin-melittin complex. A: calmodulin and melittin was cross-linked by EDC/sulfo-NHS and the cross-link was detected on the SDS-PAGE. B: mass spectra of non-digestion by MALDI-TOF of the complex at 5, 15, 30, 60 and 120 min incubation with chemical cross-link EDC/sulfo-NHS (57).

2.9 Computational protein structure modeling

If one wants to know the real 3D structure of a protein, great progress has been made using an experiment method of X-ray crystallography. This technique requires a high concentration, high purity and high stability of protein, moreover, condition for crystal growth is very important (58). So, the study of X-ray crystallography is relatively difficult and takes a lot of time. The 3D structure of protein can be predicted by the computation modeling. The prediction of 3D structure from known amino acid sequence of a protein have four different approaches: the first, comparative or homology modeling which is the most popular method for the protein in the same family as the template; the second, fold recognition and treading method aiming to assign folds of target amino acid sequence that have low or statistically insignificant sequence similarity to known protein structure; the third, *de novo* method aiming to predict the structure of protein from its amino acid sequence by principle of physic for protein folding; lastly, integrative or hybrid method that combines information from a varied set of computational and experimental sources, including all those described above (59).

Comparative or homology modeling is a template-based technique, exploiting the relationship between a target and a template protein with known 3D structure in

protein database. The identity between target and template shows the percentage of sequence correlating with model accuracy and quality. As a rule of thumb, homology model based on more than 50% sequence identity to template is considered as a high accurate model, 30%-50% sequence identity representing a medium accurate model, and less than 30% sequence identity considering as low accurate model which is the most substantial origin of errors in the building of homology model (59).

At the present, the computational protein structure modeling can be performed by having a lot of web server, each server uses a different method to predict 3D structure, and again the most favorite one is the homology modeling method. More recently, the Protein Homology/Analogy Recognition Engine or Phyre (pronounced as fire) is one of web server established for protein structure prediction, including protein complex structure, using the principle and technique of homology modeling (60). The Phyre server uses a library of known protein structure taken from the Structure Classification of Protein (SCOP) and Protein Data Bank (PDB), and has the Position-Specific Iterative Basic Local Alignment Search Tool (PSI-BLAST) to generate pairwise sequence alignment (60, 61). And also the server predicts secondary structure by using three independent secondary structure prediction programs of Psi-Pred, SSPro and JNet (60, 61). In addition, Jmol program, an open-

source Java viewer for chemical structures in 3D with features for chemicals, crystals, materials and biomolecules, is used to analyze interactive 3D viewing of the protein model.

Moreover, many web servers on the internet for protein structure prediction are freely available, as show in Table 1.

Table 1 The list of protein modeling web servers.

Server name	Web address
Phyre	http://www.imperial.ac.uk/phyre
I-TASSER	http://zhang.bioinformatics.ku.edu/I-TASSER
SAM-T06	http://www.soe.ucsc.edu/compbio/SAM_T06/T06-query.html
HHpred	http://toolkit.tuebingen.mpg.de/hhpred
Gen Threader	http://bioinf.cs.ucl.ac.uk/psipred/psiform.html
PCONS	http://pcons.net
Bioinfo	http://meta.bioinfo.pl

FFAS	http://ffas.ljcrf.edu
Robetta	http://robetta.bakerlab.org
SP^4	http://sparks.informatics.iupui.edu/SP4
SwissModel	http://swissmodel.expasy.org
ModWeb	http://salilab.org/modweb
Modeller	http://salilab.org/modeller
SCWRL3	http://dunbrack.fccc.edu/SCWRL3.php
WhatIf	http://swift.cmbi.ru.nl/whatif

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

CHAPTER III

MATERIALS AND METHODS

3.1 Materials

1) Alcohol lamp
2) Autopipet (P2, P10, P20, P100 and P1000)
3) Autopipet tip
4) Beaker
5) Cuvette (quartz, plastic)
6) Cylinder
7) Flask
8) Magnetic bar
9) Magnetic stirrer
10) Microcentrifuge tube
11) Microcentrifuge tube rack
12) Parafilm
13) Petridish

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

14) Pipette rack

15) Plastic wrap

16) Platinum loop

17) Polaroid film

18) Polypropylene conicol centrifuge tube

19) Screw cap tube

20) Spreader

21) Stirring magnetic bar

22) Test tube

23) Thermometer

3.2 Equipments

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1) Balance (Shimadzu)

2) Deep freezer -20°C,-80°C (Sharp, Revco)

3) DNA Thermal cycler (Thermo Hybaid)

4) Electrophoresis apparatus (Bio-Rad)

5) Fast Protein Liquid Chromatography (FPLC) system (Pharmacia)

6) High speed centrifuge (Sorvall)

7) High Performance Liquid Chromatography Quadrupole-Time Of Flight model SYNAPT™ High-Definition Mass Spectrometry (HPLC Q-TOF model SYNAPT™ HDMS) equipped with a Symmetry C_{18} 5 μm, 180-μm x 20-mm Trap column and a BEH130 C_{18} 1.7 μm, 100-μm x 100-mm analytical reversed-phase column (Waters corporation)

8) Incubator (Memmert)

9) Microcentrifuge (Heraeus)

10) pH meter (WPA)

11) Polaroid camera (Polaroid)

12) Power supply (Bio-Rad)

13) Refrigerator (Sanyo)

14) Shaking incubator (Kuhner)

15) Thermostat shaking water bath (GFL)

16) Ultrasonic homogenizer (Bandelin Sonoplus)

17) UV transilluminator (Spectroline)

18) UV-visible spectrophotometer model UV 1601 equipped with a temperature-controller (Shimadzu)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

19) Vertical electrophoresis apparatus (mini-gel) (Bio-Rad)

20) Vortex (Scientific industry)

21) Water bath (GFL)

3.3 Reagents

3.3.1 General reagents

1) 3,3'-Dimethylsuberimidate (DMS) (Fluka)

2) 5-Fluorouracil (Sigma)

3) 5-Phospho-D-ribose-1-diphosphate pentasodium salt (PRPP) (Sigma)

4) Acetic acid glacial (Merck)

5) Acetonitrile (Sigma)

6) Acrylamide (Sigma)

7) Agar (Becton, Dickinson and company)

8) Agarose (SeaKem)

9) Ammonium bicarbonate (NH_4HCO_3) (Sigma)

10) Ammonium persulfate (Sigma)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

11) Ampicillin sodium salt (Sigma)

12) Bromophenol blue (Fluka)

13) Bovine serum albumin (BSA) (Sigma)

14) Butanol (Merck)

15) Calcium chloride (Merck)

16) Comassie brilliant blue R250 (Sigma)

17) Disodium hydrogen phosphate (Merck)

18) Dithiothreitol (DTT) (Sigma)

19) DNA marker (1Kb DNA ladder) (Vivantis)

20) Ethanol, absolute (Merck)

21) Ethidium bromide (Sigma)

22) Ethylene diamine tetraacetic acid (disodium salt) (EDTA) (Fluka)

23) Formic acid (Fluka)

24) Glucose (Fluka)

25) Glycerol (Sigma)

26) Glycine (Vivantis)

27) Hydrochloric acid (Merck)

Chulalongkorn University

28) Imidazol (Sigma)

29) Iodoacetamide (Sigma)

30) Isopropanol (Merck)

31) Isopropyl-β-D-thiogalactopyranoside (IPTG) (Sigma)

32) Magnesium chloride (Merck)

33) Magnesium sulfate (Merck)

34) Methanol, absolute (Merck)

35) Mercaptoethanol (Sigma)

36) N,N-Methylene-bis-acrylamide (BIS) (Sigma)

37) N,N,N',N'-Tetramethylethylenediamine (TEMED) (Sigma)

38) Orotic acid monosodium (OA) (Sigma)

39) Orotidine-5'-monophosphate trisodium salt (OMP) (Sigma)

40) Phenol-chloroform (1:1, v/v) (Pierce)

41) Potassium chloride (Merck)

42) Potassium hydrogen phosphate (Merck)

43) Protease inhibitor cocktail (Roche)

44) Protein molecular mass marker for SDS-PAGE (low range) (Bio-Rad)

45) Sodium acetate (Merck)

46) Sodium dodecyl sulfate (SDS) (Sigma)

47) Sodium hydroxide (Merck)

48) Triton-X-100 (Sigma)

49) Trizma base (Sigma)

50) Trypsin (Promega)

51) Tryptone (Becton, Dickinson and company)

52) Uracil (Sigma)

53) Uridine (Sigma)

54) Yeast extract (Becton, Dickinson and company)

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

3.3.2 Reagent kits

1) Hi Trap Q HP anion-exchange column (Pharmacia BioSciences)

2) Protein assay (Biorad)

3) QIAGEN nickel-nitrilotriacetic acid (Ni^{2+}-NTA) agarose affinity resin

4) QIAGEN plasmid Midi kit

5) pQE30Xa vector (Qiagen)

6) pTrcHis TOPO vector (Invitrogen)

3.4 Methods

3.4.1 Construction of plasmid in *E. coli* for protein expression

*Pf*OPRT gene (846 bp) was cloned into pQE30Xa vector using competent *E. coli* M15 [pREP4] strain as host cell. The *E. coli* cell was grown in LB medium containing ampicillin (100 µg/ml) and kanamycin (10 µg/ml) (22). *Pf*OMPDC gene (969 bp) was cloned into pTrcHisA vector using *E. coli* TOP10 strain as host cell. The cell was propagated in the LB medium containing only ampicillin (100 µg/ml) (20).

3.4.2 Protein expression and purification of recombinant *Pf*OPRT and *Pf*OMPDC.

The *E. coli* cells harboring the constructed plasmids as described above was cultivated in their corresponding LB media at 37°C until the cell density at OD_{600nm} was 0.5. They were then induced with 1 mM IPTG for 18 h at 18°C. The recombinant *Pf*OPRT and *Pf*OMPDC proteins was purified by using three sequential steps of a Ni^{2+}-

NTA affinity chromatography, Hi-trap Q HP anion-exchange column and a Superose 12 gel filtration FPLC, as previously described (17, 22). In brief, cell pellets were suspened in lysis buffer (50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 10 mM imidazol, 10% glycerol, 1% triton x-100, 0.1%(w/v) lysozyme, 0.1% RNase and protease inhibitor cocktail) and then the mixture was sonicated with an ultrasonic homogenizer. The *E. coli* lysate was centrifuged for 30 min at 10,000 rpm. The supernate was load onto Ni^{2+}-NTA agarose affinity gel equilibrated with bufferA (50 mM Tris-HCl, pH 8.0, 300 mM NaCl, and 10 mM imidazole). The column was washed twice with 5 ml of bufferA and twice with 5 ml of bufferB (50 mM Tris-HCl, pH 8.0, 300 mM NaCl and 20 mM imidazole). The enzyme was eluted with 3 ml of bufferC (50 mM Tris-HCl, pH 8.0, 300 mM NaCl and 250 mM imidazole). The recombinant protein's eluted from Ni^{2+}-NTA affinity column was added into Hi Trap Q HP anion-exchange column equilibrated with 50 mM Tris-HCl, pH8.0 and then eluted with the 50 mM Tris-HCl, pH8.0, 250 mM NaCl. After that, the recombinate protein was purify by superpose 12 gel filtration FPLC, the column equilibrate with 50 mM Tris-HCl, pH 8.0, 250 mM NaCl. The flow rate was maintained at 0.5 ml/min and 0.5 ml fractions were collected.

3.4.3 Protein and enzymatic assay

Protein concentrations were determined by the Bradford assay (62) using bovine serum albumin as standard. *Pf*OPRT and *Pf*OMPDC activities were measured in quartz cuvette with a UV-visible spectrophotometer model UV 1601 equipped with a temperature-controller, according to previously described (62). In brief, the reaction of OPRT activity assay (1ml) contained 50mM Tris-HCl, pH 8.0, 5mM $MgCl_2$, 250 μM DTT, 250 μM OA and the enzyme (5-50 μl) were incubate for 1 min at 37°C. The reaction was started by add 250 μM PRPP, measuring decreasing OA concentration of absorbance change at 295 nm up to 3 min. The reaction of OMPDC activity assay (1 ml) contained 50 mM Tris-HCl, pH 8.0, 250 μM DTT and the enzyme (5-50 μl) were incubate for 1 min at 37°C. The reaction was started by add 250 μM OMP, measuring decreasing OMP concentration of absorbance change at 285 nm up to 3 min.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

3.4.4 Chemical cross-linking of *Pf*OPRT and *Pf*OMPDC

In vitro formation of heterotetrameric $(PfOPRT)_2(PfOMPDC)_2$ complex or dimeric forms of either $(PfOPRT)_2$ or $(PfOMPDC)_2$ was performed by using a bifunctional chemical cross-linking DMS which linked peptides having the nearest

neighbors of Lys residues at $-NH_2$ groups to yield amidine derivative of the peptides with a maximum link distance of 1.0 nm. The proteins and DMS ratios and the conditions used to generate dimeric (*Pf*OPRT)$_2$ or (*Pf*OMPDC)$_2$ and heterotetrameric (*Pf*OPRT)$_2$(*Pf*OMPDC)$_2$ complex were described (22). In brief, *Pf*OPRT and/or *Pf*OMPDC were incubated with DMS at 25°C, various time and stop reaction with 1 M glycine.

3.4.5 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was performed on a Bio-Rad minislab gel apparatus with 5% stacking and different % running gels by the Laemmli system (63). The molecular mass was determined compare the molecular mass standard. The gel was stained with Coomassie brilliant blue R 250 dye.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3.4.6 In gel digestion

Visualized protein bands in the SDS-PAGE gel was cut into small pieces and destained extensively with 50 mM NH_4HCO_3 followed by methanol, and then washed with deionized water. The chopped gels was dehydrated in acetonitrile, and reduced with 10 mM dithiothreitol in10 mM NH_4HCO_3 for 1 h at 56°C. The reduced gel was

then alkylated by 100 mM iodoacetamide in 10 mM NH_4HCO_3 for 1 h at 25 °C in dark. The alkylated gel was further digested with 10 ng/µl sequencing grade trypsin in 10 mM NH_4HCO_3 for 3 h at 37°C. The trypsin digested peptides was eluted from the gels with 0.1% formic acid/50% acetonitrile, evaporated at 40°C overnight, and then stored at -20°C until use (64).

3.4.7 Liquid chromatography-mass spectrometry (LC-MS/MS) and proteomic data analysis.

Nanoscale LC separation of tryptic peptides was performed with a NanoAcquity system. The samples was initially transferred with 0.1% formic acid to the trap column with a flow rate of 15 µl/min for 1 min. Mobile phase A was 0.1% formic acid in water, and mobile phase B was 0.1% formic acid in acetonitrile. The peptides were separated with a gradient of 15–50% mobile phase B over 15 min at a flow rate of 500 nl/min followed by a 3-min rinse with 80% of mobile phase B. The column temperature was maintained at 35 °C. The lock mass was delivered from the auxiliary pump of the NanoAcquity pump with a constant flow rate of 500 nl/min

at a concentration of 200 fmol/μl of [Glu^{1}] fibrinopeptide B to the reference sprayer of the NanoLockSpray source of the mass spectrometer.

Analysis of tryptic peptides was performed using a SYNAPT™ HDMS mass spectrometer. For all measurements, the mass spectrometer was operated in the V-mode of analysis with a resolution of at least 10,000 full-width half-maximum. All analyses were performed using positive nanoelectrospray ion mode. The time-of-flight (TOF) analyzer of the mass spectrometer was externally calibrated with [Glu^{1}] fibrinopeptide B from m/z 50 to 1600 with acquisition lock mass corrected using the monoisotopic mass of the doubly charged precursor of [Glu^{1}]fibrinopeptide B. The reference sprayer was sampled with a frequency of 20 sec. Accurate mass LC-MS data was acquired with data direct acquisition mode. The energy of trap was set at a collision energy of 6 V. In transfer collision energy control, low energy was set at 4 V. The quadrupole mass analyzer was be adjusted such that ions from m/z 300 to 1800 was efficiently transmitted. The MS/MS survey is over range 50 to 1990 Da and scan time was 0.5 sec.

The MassLynx version 2.3 software and Decyder MS were applied to analyze the LC-MS/MS data. The data was show peptide, which identify from monomeric,

dimeric and tetrameric enzyme. But some peptide can't identify in dimeric or tetrameric enzyme because this peptide was cross-link to the nearest molecule.

3.4.8 Homology modeling of *Pf*OPRT and *Pf*OMPDC in monomer, dimer and heterotetramer

The three-dimensional (3D) structures of *Pf*OPRT (residues 64-281) and *Pf*OMPDC (residues 1-323) was, respectively, determined by using crystal structures of *Saccharomyces cerevisiae* OPRT (PDB code 2PS1, chain A) and *P. falciparum* 3D7 OMPDC (PDB code 2Q8Z, chain A) as template. Homology models of the structures were generated by the Phyre program (Imperial College, London; a server at http://www.sbg.bio.ic.ac.uk/~phyre/index.cgi) (61). Homology search and multiple alignment of amino acid sequences was be performed using the BLAST (65) and the CLUSTALW2 (45) programs, respectively. After the binding sequences were identified by LC-MS/MS data, models for protein-protein interactions of dimeric $(Pf\text{OPRT})_2$ and $(Pf\text{OMPDC})_2$, and heterotetrameric $(Pf\text{OPRT})_2(Pf\text{OMPDC})_2$ was then generated. The constructed dimeric models was be stereo rotated with the amino acid sequences used for binding between chain A and chain B of *Pf*OPRT and *Pf*OMPDC. The dimeric models was then applied as the template to generate heterotetrameric model by

further stereo rotating with the amino acid sequences identified for *Pf*OPRT-*Pf*OMPDC interactions.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

CHAPTER IV

RESULTS

4.1 Construction of recombinant plasmid *Pf*OPRT and *Pf*OMPDC in *E. coli* for protein expression

The *Pf*OPRT (ORF, 846 bp) was inserted into pQE30 Xa plasmid (3.5 kb) and then transformed into *E. coli* M15. The clone containing *Pf*OPRT-pQE30 Xa vector was checked by plasmid extraction and restriction analysis. The size of *Pf*OPRT-pQE30 Xa vector was 4.3 kb, and the insert *Pf*OPRT was then confirmed by *Bam*HI and *Hind*III digestion. By agarose gel electrophoresis, it was shown that the size of the pQE30 Xa vector and *Pf*OPRT insert were 3.5 kb and 846 bp, respectively (Figure 13).

The *Pf*OMPDC (ORF, 969 bp) was cloned into pTrc His A plasmid (4.4 kb) and then transformed into *E. coli* TOP10. The clone having *Pf*OMPDC-pTrc His A vector, was checked by plasmid extraction and restriction analysis on agarose gel eletrophoresis. The size of *Pf*OMPDC-pTrc His A vector was 5.4 kb, and the insert *Pf*OMPDC was then confirmed by *Bam*HI and *Sal*I digestion. The result showed that

the size of pTrc His A vector and *Pf*OMPDC insert were 4.4 kb and 969 bp, respectively (Figure 13).

These results suggest the constructed plasmids have *Pf*OPRT and *Pf*OMPDC gene insert, corresponding to the design inserts with restriction sites for both PCR and cloning strategy.

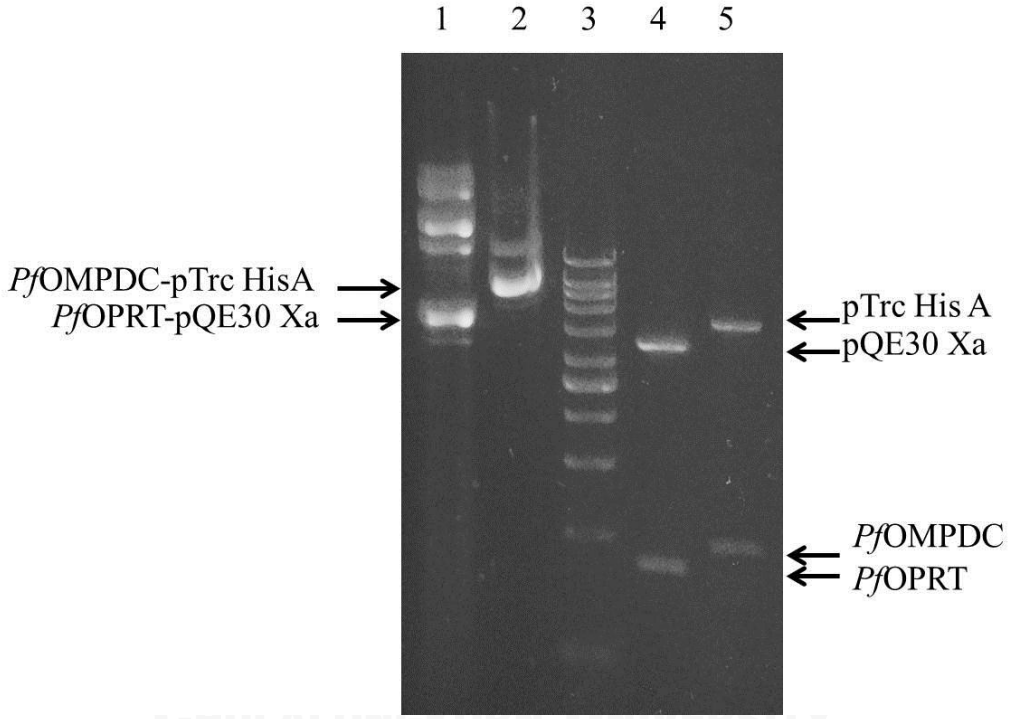


Figure 13 Agarose gel analysis of *Pf*OPRT-pQE30 Xa plasmid and *Pf*OMPDC-pTrc HisA plasmid. Lane1: *Pf*OPRT-pQE30 Xa plasmid, lane2: *Pf*OMPDC-pTrc HisA plasmid, lane3: 1kb DNA ladder marker, lane4: *Pf*OPRT-pQE30 Xa plasmid digested with *Bam* HI and *Hind* III, lane5: *Pf*OMPDC-pTrc HisA plasmid cut with *Bam* HI and *Sal* I. The agarose gel was 0.8%.

4.2 Recombinant enzyme *Pf*OPRT and *Pf*OMPDC preparation

The *Pf*OPRT and *Pf*OMPDC enzyme were expressed in *E. coli* M15 and *E. coli* TOP10, respectively under IPTG induction at 18°C for 18 h. The enzymes were purified by two sequential steps of a Ni^{2+}-NTA-affinity chromatography and Hi Trap Q chromatography. In some experiments, the enzymes were subjected into a superose 12 gel filtration column on FPLC system to analyze molecular mass of their native forms. The purity of enzyme preparation was analyzed by SDS-PAGE.

The recombinant *Pf*OPRT and *Pf*OMPDC were catalytically active in its dimeric form and had specific activities of 4.7 and 7.9 μmol/min/mg protein, respectively. Using this strategy, about 3-4 mg each of nearly 95% pure and active *Pf*OPRT and *Pf*OMPDC from 1 L of cell culture was obtained. By gel filtration FPLC and SDS-PAGE, the *Pf*OPRT was a 67-kDa homodimer of two 33-kDa $(Pf\text{OPRT})_2$ subunits (Figure 14), whereas the *Pf*OMPDC showed a monomeric subunit of 38 kDa (Figure 15) and a 76-kDa homodimeric $(Pf\text{OMPDC})_2$.

The above results suggest the recombinant protein of *Pf*OPRT and *Pf*OMPDC are enzymatically active with the native dimeric forms, whereas the monomers are also identified, in their denature forms.

4.3 Chemical cross-linking of *Pf*OPRT and *Pf*OMPDC

To prepare cross-linked products of either *Pf*OPRT or *Pf*OMPDC, the pure enzyme was incubated at 25 °C with DMS at a ratio of 1:2 protein to cross-linker, at various time intervals and the chemical cross-link products were then collected for SDS-PAGE analysis at 15, 30, 45 and 60 min (22). By using SDS-PAGE, monomers of *Pf*OPRT or *Pf*OMPDC and dimers of $(Pf\text{OPRT})_2$ or $(Pf\text{OMPDC})_2$ were detected (Figure 14, 15). At 30 min-incubation, the proteins were found to be cross-linked as dimeric $(Pf\text{OPRT})_2$ and dimeric $(Pf\text{OMPDC})_2$. There were no more oligomeric forms, such as trimer and tetramer, in each enzyme cross-link by this technique.

Chemical cross-linked heterotetrameric form of *Pf*OPRT and *Pf*OMPDC was achieved by DMS treatment at a ratio of 1:2 protein to cross-linker at 25 °C with various time intervals and the cross-linked products were then analyzed by SDS-PAGE at 10, 20, 30 and 60 min. At 30 min-incubation, more than 50% of the proteins were cross-linked as a tetrameric form (Figure 16). The molecular mass of each oligomer was determined as previously described (22). The time-dependent oligomerization of each enzyme corresponded to that expected for sequential cross-linking of monomer to form dimer, and each dimer to form heterotetramer. However, small

proteins of oligomeric forms such as octamer were observed after 30 min incubation of both enzymes with DMS treatment.

In summary, the chemical cross-linked products of dimeric (*Pf*OPRT)$_2$ and (*Pf*OMPDC)$_2$ and heterotetrameric (*Pf*OPRT)$_2$(*Pf*OMPDC)$_2$ were prepared and found to be the same results as established in the previously experiments (22) and then used for LC -MS/MS analysis.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

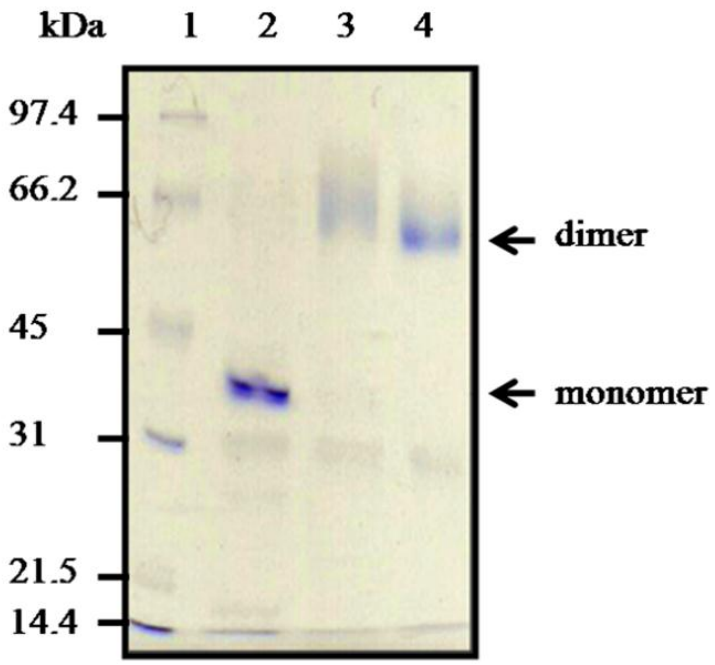


Figure 14 SDS-PAGE analysis of monomeric and dimeric forms of *P. falciparum* OPRT. The oligomers of *Pf*OPRT, cross-linked with 3,3'-dimethyl suberimidate (DMS), were identified on 12% polyacrylamide gels. Lane 1: molecular mass standard, lane 2: non-cross-linking monomeric form (10 μg protein), lanes 3, 4: cross-linking products of dimeric form incubated at 15 and 30 min-reactions, respectively. Gel of monomer on lane 2 and dimeric product on lane 4 were excised for further LC-MS/MS analysis.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

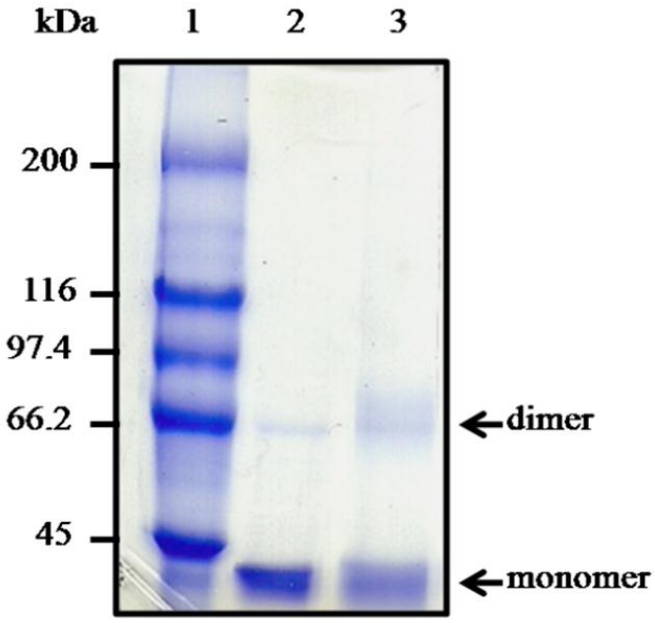


Figure 15 SDS-PAGE analysis of monomeric and dimeric forms of *P. falciparum* OMPDC. The oligomers of *Pf*OMPDC, cross-linked with 3,3'-dimethyl suberimidate (DMS), were identified on 7% polyacrylamide gels. Lane 1: molecular mass standard, lane 2: non-cross-linking monomeric form (20 μg protein), lane 3: cross-linking products of the dimeric form at 30 min-reaction. Gel of monomer on lane 2 and dimeric product on lane 3 were excised for further LC-MS/MS analysis.

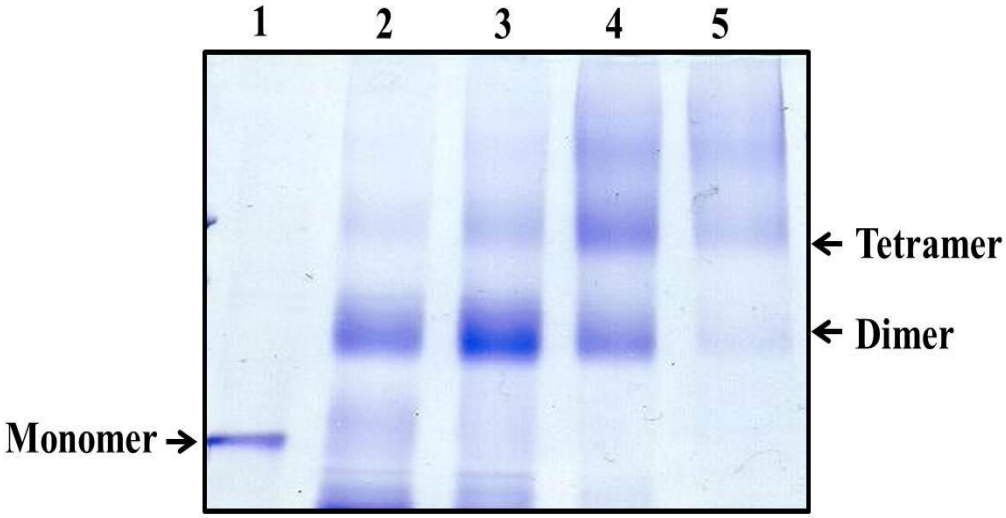


Figure 16 SDS-PAGE analysis of heterotetrameric (*Pf*OPRT)$_2$(*Pf*OMPDC)$_2$. The oligomers of *Pf*OPRT and *Pf*OMPDC, cross-linked with 3,3'-dimethyl suberimidate (DMS), were identified on 10% polyacrylamide gels. Lane 1: non-cross-linking monomeric form (10 µg protein each), lanes 2-5: cross-linking products of tetrameric form at 10, 20, 30 and 60 min-reactions, respectively. Gel of tetrameric product on lane 4 was excised for further LC-MS/MS analysis.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

4.4 Liquid chromatography/mass spectrometry of chemical cross-linking products of dimeric forms and heterotetrameric (*Pf*OPRT)$_2$(*Pf*OMPDC)$_2$

The non cross-linked monomeric forms, cross-linked dimeric and heterotetrameric forms of *Pf*OPRT and *Pf*OMPDC on the SDS-PAGE gel were subjected to in-gel trypsin digestion. The entire peptide samples were injected into the HPLC Q-TOF, and the LC-MS/MS data were analyzed by the MassLynx and Decyder MS software and summarized in Table 2 for *Pf*OPRT and Table 3 for *Pf*OMPDC. The results on the software analyses showed the amino acid sequences of *Pf*OPRT and *Pf*OMPDC obtained as monomer, dimer, and heterotetramer forms. Peptide score is the score for each peptide identification and MH^+(Da) is a peptide mass.

Combining the chemical cross-linking and LC-MS/MS led us to identify amino acid sequences that play role in the binding between domains and/or polypeptide chain of individual and both enzymes. The blank spaces in the Tables 2 and 3 represent the cross-linked peptides/sequences. Amino acid sequences that were present in the monomeric form but was not in the dimeric or heterotetarmeric form confirms that this sequence underlies domain-domain interface and contains amino acid residues used for binding between domains.

The results of LC-MS/MS analysis of the cross-link products of *Pf*OPRT and *Pf*OMPDC in their dimeric and tetrameric forms suggest that the identified amino acid sequences of both proteins play a role in protein-protein interaction of *Pf*OPRT and *Pf*OMPDC.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Table 2 LC-MS/MS data analysis of *Pf*OPRT in monomeric, dimeric forms and in heterotetrameric complexed with *Pf*OMPDC. The blank spaces represent the cross-linked peptides/sequences. The peptide matches OPRT of *P. falciparum* (3D7), protein ID is gi|124506173.

Peptide	Peptide score[a]	MH^+ (Da)[b]	Monomeric OPRT[a]	Dimeric OPRT[a]	Tetrameric OPRT-OMPDC[a]
ENEFLCDEEIYK	34.34	1032.6891	5.811603	10.67179	7.537772
EYGDKNVIVGNLDDDDKDILNLKK	15.6	2713.7146	8.755316		5.822202
GIPMVSLTSHFLFESK	85.93	1504.6726	7.166355	6.090981	6.881962
IYFKDIFEK	44.310001	1194.5951	7.241354	5.591504	6.965446
KNIIIIDDVFTCGTALTEILAK	71.029999	2461.8183	7.857266	4.154335	
KYSNIFYLYDR	68.019997	1089.6016	10.80939	7.890219	7.214506
LSFDYLLGASYK	60.650002	1375.9494	7.222945	3.474503	7.305166
NVIVGNLDDDDK	50.02	1154.3034	5.289937	5.497626	6.61994
NVIVGNLDDDDKDILNLK	103.83	2013.2179	14.65743	4.562683	5.403011
NVIVGNLDDDDKDILNLKK	35.790001	2122.6577	9.319596		6.336068
SFVHLK	8.3599997	499.33263	7.925779	6.753919	8.925819
VGIPLYSILSYK	61.27	1183.7367	4.989854	6.546392	6.391722
VVAFIVLLNR	47.689999	1144.1441	6.420709	9.93629	6.922265
VVAFIVLLNRNEYEINENNQK	33	2520.3911	9.863172	4.33497	4.282399
YSNIFYLYDR	73.480003	1353.232	13.4333	10.82894	9.282876
YSNIFYLYDRK	26.860001	1481.5222	4.393924	9.054119	6.417926

[a] The results were tested by the Decyder MS program with significance at $p<0.05$, whereas peptide score is the score for each peptide identification, high peptide score indicates high significant.

[b] MH^+(Da) is a peptide mass.

Table 3 LC-MS/MS data analysis of *Pf*OMPDC in monomeric, dimeric forms and in heterotetrameric complexed with *Pf*OPRT. The blank spaces in the represent the cross-linked peptides/sequences. The peptide matches OMPDC of *P. falciparum* (3D7); chain A, crystal structure of *P. Falciparum* OMPDC complexed with 6-amino-UMP, protein ID are gi|124802556 and gi|166235377, respectively..

Peptide	Peptibe Score[a]	MH^+ (Da)[b]	Monomeric OMPDC[a]	Dimeric OMPDC[a]	Tetrameric OPRT-OMPDC[a]
AAQMYYDQINAILK	58.610001	2210.5491	13.51126	13.3599	12.41457
APDNIIR	45.57	1000.6514	8.347193	7.84313	7.181398
APDNIIREEK	41.049999	1180.2404	9.541573	5.696492	4.269348
DICYDEEK	18.82	1065.6935	10.5044	4.767869	9.609203
DICYDEEKNK	21.639999	1313.5946	6.805417		3.950628
DILLKAPDNIIR	40.07	1374.5959	5.955823	7.277666	6.832736
DILLKAPDNIIREEK	9.0100002	1766.359	7.218944	8.964727	5.962379
FIFEYLK	22.33	4811.5051	10.32849	7.360944	
ILINIGR	26.360001	800.08878	7.875441	5.720076	5.588307
INDIGNTVK	36.799999	976.6666	8.848516	6.200231	3.874776
KFIFEYLK	25.84	1089.9767	9.965045	5.086093	6.27191
MNFAFYIPYGSVGIDVLK	46.419998	1422.1116	6.851215	5.939178	4.606118
NAINTCLCIGLDPDEK	80.889999	1833.0554	6.117382	9.577628	
NAINTCLCIGLDPDEKDIENFMK	24.33	2139.5581	6.240061	5.302535	5.325563
NVFDYLYELNIPTILDMK	62.799999	2217.121	14.35623	10.73046	
SDSCTVNIYMGTNMLK	41.169998	1388.5508	5.956773	4.115105	4.277855
SEEFFYFFNHFCFYIINETNK	29.08	2047.2428	6.717936	5.76970	
TTNPDSAIFQK	59.43	1221.3894	14.2892	13.02087	6.978079
TTNPDSAIFQKNLSLDNK	25.290001	1985.6077	8.036843	7.23289	6.672452
TYFPNCYILSPGIGAQNGDLHK	41.490002	2463.589	8.36222	11.827	5.371594
YALTFK	20.879999	3332.5879	12.87805	5.566563	8.408722
YINNVSIK	23.23	4797.7002	8.02208	6.701042	

[a] The results were tested by the Decyder MS program with significance at $p<0.05$, whereas peptide score is the score for each peptide in identification, high peptide score show high significant peptide.

[b] MH^+(Da) is a peptide mass.

4.5 Homology models and proteomic data analysis of *Pf*OPRT and *Pf*OMPDC in monomeric, dimeric and heterotetrameric forms

The Phyre program can predict the secondary structure and the homology model of three dimensional (3D) structure. As for the homology model, the Phyre program was utilized to construct models of monomeric *Pf*OPRT and *Pf*OMPDC. The model of *Pf*OPRT had no N-terminal domain (residues 1-63), which contains the unique insertion LCR amino acid sequence (Figure 17). The secondary structure prediction was showed, the N-terminal sequence contained one α-helix (residues 13-36). Sequence alignment between *Pf*OPRT and *S. cerevisiae* OPRT showed 30% identity. The most striking finding in all *Plasmodium* OPRTs were two large insertions; one, at the N-terminal region as described above and, the other, at the internal region (residues 178-196) with a high hydrophobic index of +1 to +3 (16).

Based on the sequence alignments of OMPDCs from available *Plasmodium* species, the model for *Pf*OMPDC using the full-length sequence contained a large unique insertion LCR in the α2, α3 and α4 helices at the N-terminal domain which were found in all *Plasmodium* OMPDCs (Figure 18) (22).

The monomeric *Pf*OPRT model consisted of 6 α-helices and 11 β-strands (Figure 19A), while the monomeric *Pf*OMPDC structure model exhibited 13 α-helices and 8 β-strands (Figure 19B). Both *Pf*OPRT and *Pf*OMPDC models highlight the LCR sequences that conserve amongst *Plasmodium* species, and participate in homodimeric and heterotetrameric formation.

The LC-MS/MS data of *Pf*OPRT (Table 2) identified the sequence 168-EYGDKNVIVGNLDDDDKDILNLKK-191 and 173-NVIVGNLDDDDKDILNLKK -191 containing amino acids responsible for β8 strand (residues 177- 199), which presumably locate at domain-domain interface and bind between chain A and chain B of *Pf*OPRT (Figure 20). Interestingly, this binding sequence is a main part of the internal unique insertion LCR (residues 178-196) of *Pf*OPRT (highlighted in Figure 17) (16). In contrast, the LC-MS/MS data of *Pf*OMPDC (Table 3) showed the sequence 175-DICYDEEKNK-184, which was not a LCR insertion sequence and located in α8 helix (residues 173-176)

(highlighted in Figure 18) (22), as the one for binding between chain A and chain B of *Pf*OMPDC (Figure 21).

Furthermore, the LC-MS/MS data of heterotetrameric (*Pf*OPRT)$_2$ (*Pf*OMPDC)$_2$ complex (Tables 2 and 3) indicated that the sequence 202-KNIIIIDDVFTCGTALTEILAK-223, spanning in α4 helix (residues 214-225) of *Pf*OPRT (Figure 20 and 22), was used to bind between the dimeric (*Pf*OPRT)$_2$ and the dimeric (*Pf*OMPDC)$_2$ at five positions as follows: 12-NAINTCLCIGLDPDEK-27 in α2 helix; 52-YINNVSIK-59 in β turn at α3-α4 position; 76-SEEFFYFFNHFCFYIINETNK-96 in α5 helix; 121-NVFDYLYELNIPTILDMK-138 in α6 helix; and 152-FIFEYLK-158, partially covering in α7 helix (Figure 21 and 22) of *Pf*OMPDC. All peptides involve in the tetrameric formation as described above are highlighted with yellow color on *P. falciparum* amino acid sequences (Figures 17 and 18). In addition, the peptide 12-NAINTCLCIGLDPDEK-27 of *Pf*OMPDC (Table 2) was cross-linked using K27 to the peptide 202-KNIIIIDDVFTCGTALTEILAK-223 of *Pf*OPRT (Table 2) to form heterotetrameric (*Pf*OPRT)$_2$(*Pf*OMPDC)$_2$, whereas the peptide 12-NAINTCLCIGLDPDEKDIENFMK-34 of *Pf*OMPDC (Table 3) was not cross-linked with DMS and it was then digested with trypsin when the enzyme existed in its dimeric form. The results suggest that the conformational change occurs during heterotetrameric association of dimeric (*Pf*OPRT)$_2$ and (*Pf*OMPDC)$_2$ (Figure 22).

It is noted here that, the sequence 202 -223 (α4 helix) of *Pf*OPRT and the sequence 12-27 (α2 helix), 52-59 (β turn of α3-α4 helix), 76-96 (α5 helix), 121-138 (α6 helix) and 152-158 (partial α7 helix) of *Pf*OMPDC were found to be located in the domain-domain interface of the heterotetrameric (*Pf*OPRT)$_2$(*Pf*OMPDC)$_2$. Indeed, the α2 and α5 helices, responsible for domain-domain interactions, are unique insertion LCR sequences located at the N-terminal extension of *Pf*OMPDC (22).

Based on the identified binding sequences by LC-MS/MS data, the structural models for protein-protein interactions of dimeric (*Pf*OPRT)$_2$ and (*Pf*OMPDC)$_2$, and heterotetrameric (*Pf*OPRT)$_2$(*Pf*OMPDC)$_2$ are then proposed and shown in Figure 20, 21 and 22.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

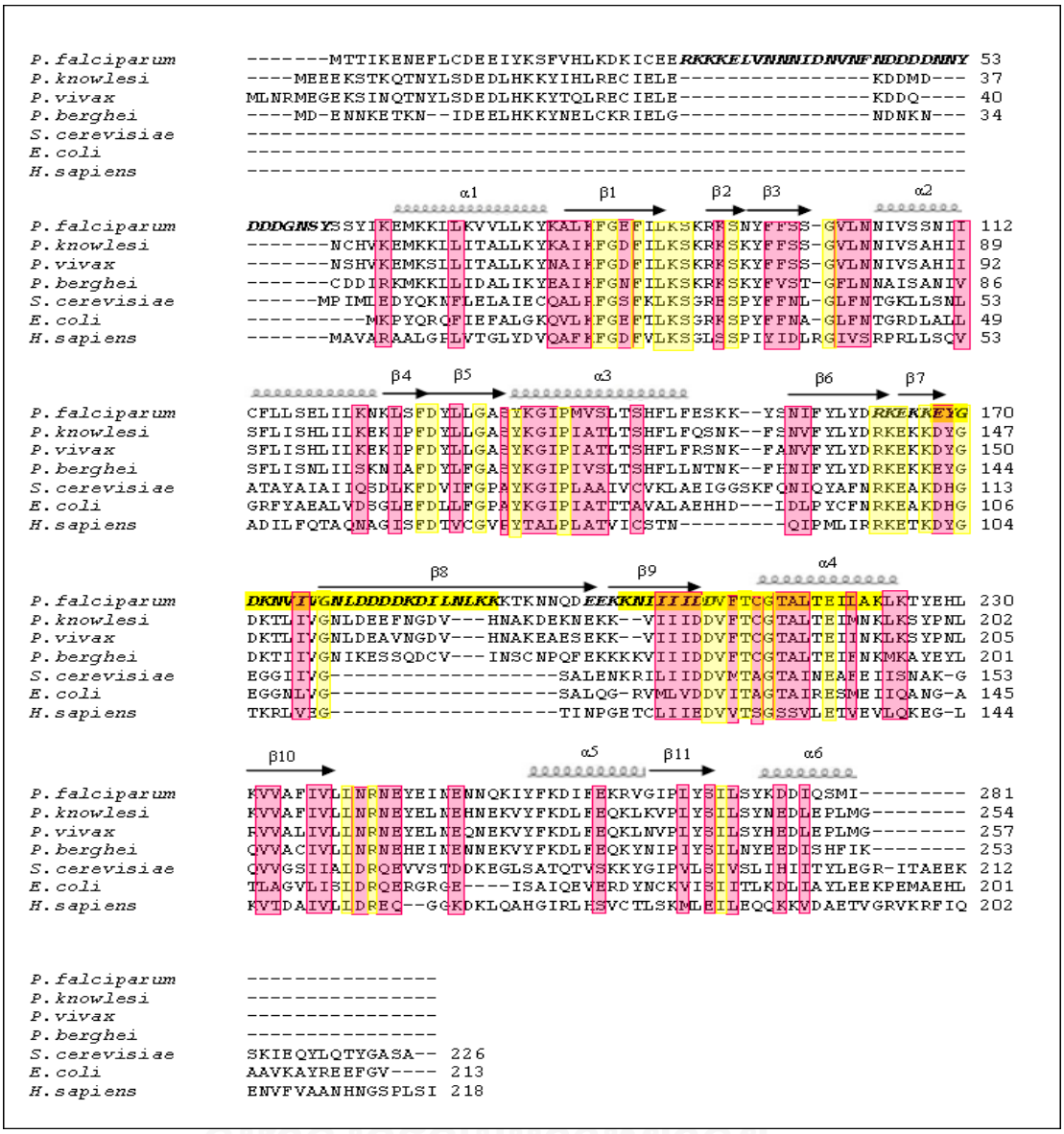


Figure 17 Sequence alignment and secondary structure of OPRTs. Sequence alignment of seven different OPRTs was performed using CLUSTALW (45). Species name and accession numbers are as follows: *P. falciparum* OPRT (PFE0630C), *P. knowlesi* (XP_002259686.1), *P. vivax* (XP_001613829), *P. berghei* (PBANKA_111240), *S. cerevisiae* (NP_013601), *E. coli* (X00781) and *H. sapiens* (NP_000364). Conserved

identical residues are highlighted in yellow, similar residues are shown in pink. In *P. falciparum* sequence, LCRs are bold and italic letters, cross-linking peptides are shown with yellow drag, sequences 168-191(β8) and 202-223 (α4 helix) are responsible for homodimeric and heterotetrameric formation, respectively. The Phyre program was used to predict α-helix and β-strand in the N-terminal sequence (residues 1-63), having one α-helix (residues 13-36) and no β-strand (61).

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

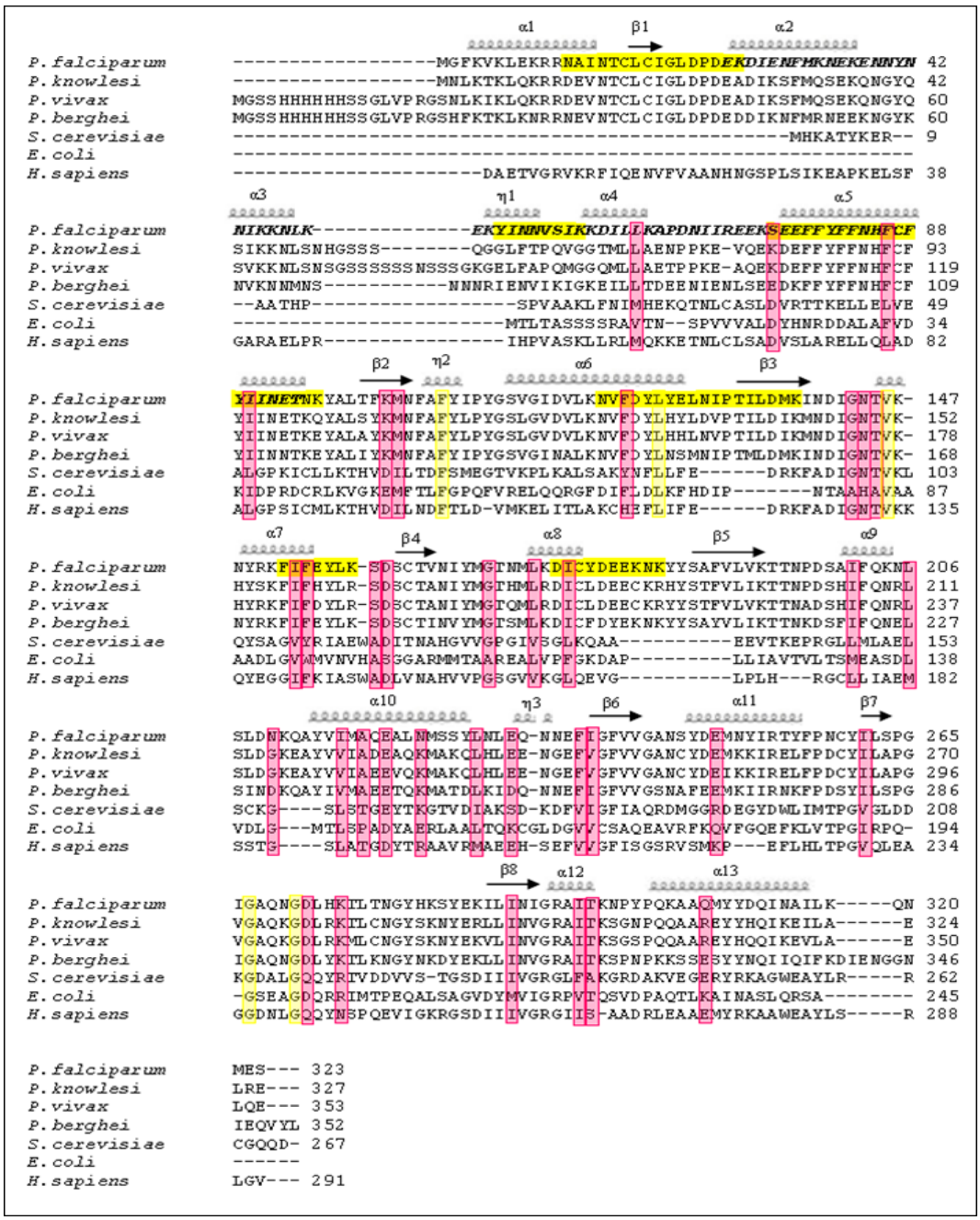


Figure 18 Sequence alignment and secondary structure of OMPDCs. Sequence alignment of seven different OMPDCs was performed using CLUSTALW (45). Species name, accession numbers and PDB ID are as follows: *P. falciparum* OMPDC (2ZA2), *P. knowlesi* (XP_002261749.1), *P. vivax* (2FFC), *P. berghei* (2FDS), *S. cerevisiae* (1DQW), *E.*

coli (1L2U) and *H. sapiens* (2EAW). Conserved identical residues are highlighted in yellow, similar residues are shown in pink. In *P. falciparum* sequence, LCRs are bold and italic letters, cross-linking peptides are shown with yellow drag, sequence 175-184 (α8) is used for homodimeric formation, sequences 12-27 (α2), 52-59 (η1, β turn of α3-α4), 76-96 (α5), 121-138 (α6) and 152-158 (α7) are responsible for heterotetrameric formation.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

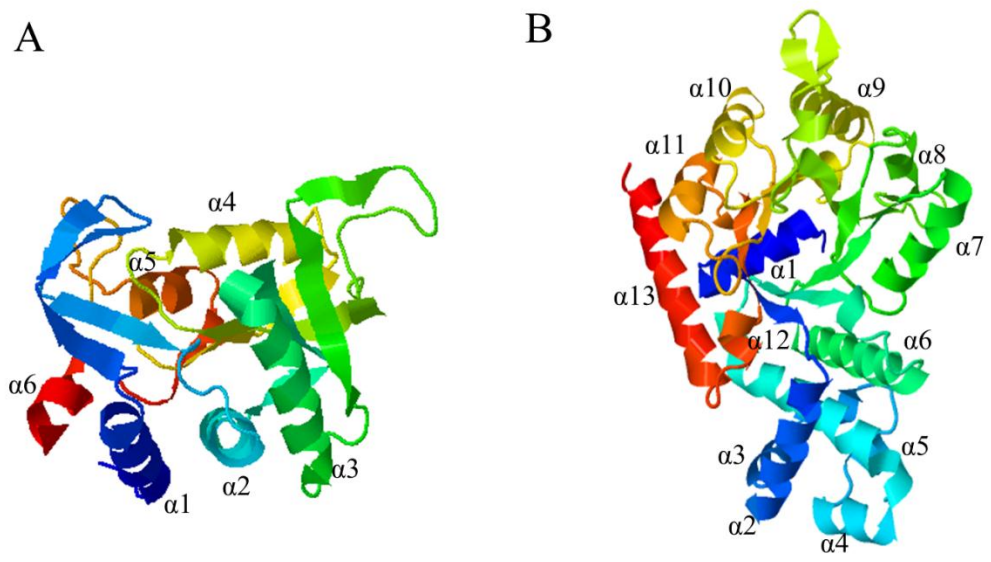


Figure 19 Phyre homology model of monomeric *P. falciparum* OPRT and OMPDC. The Phyre program was used to predict α-helix and β-strand (61). The α-helix numberings are shown in the built models, while the β-strands are not shown here but details can be found in the sequence alignments for *Pf*OPRT (Figure 17) and *Pf*OMPDC (Figure 18). (A) *Pf*OPRT model (64-281 residues) is depicted with the 1-63 residues omitted from the Phyre program due to its unique insertion sequence LCR with very low identity to other OPRTs. The model shows LCRs, α4 and β8 that used for dimeric and tetrameric formation. The active site was identified using Swiss model as previously described (44). (B) *Pf*OMPDC model (1-323 residues) is a completed full-length sequence in which its N-terminus harboring α2, α3, α4, and α5-helices as unique insertions (LCRs) to the enzyme and also α6,α7- helix. They play role in dimeric and tetrameric formation. The active site was identified based on our previously described crystal structure (48).

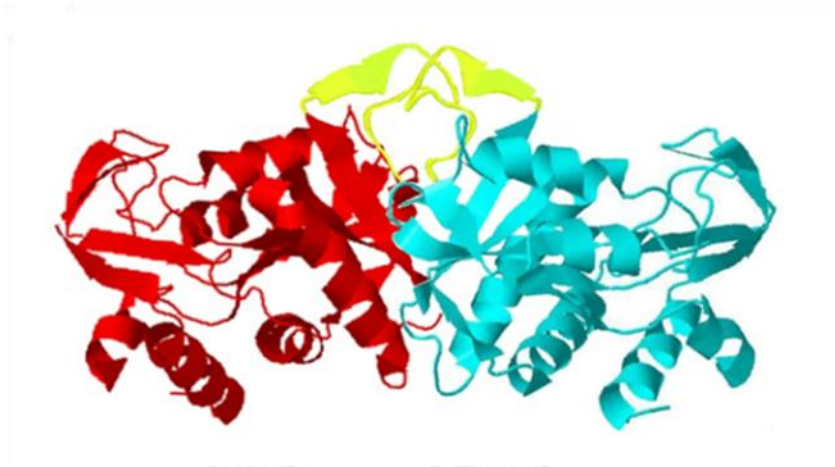


Figure 20 Proposed homology model of *P. falciparum* OPRT in dimeric form. The models were constructed by the Phyre program (61). Model of dimeric (*Pf*OPRT)$_2$, yellow ribbons represent the cross-linking and interacting sites of β8-strand for dimeric (*Pf*OPRT)$_2$.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

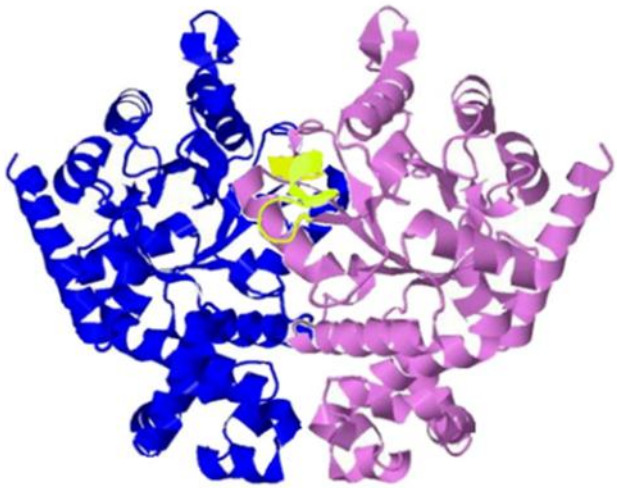


Figure 21 Proposed homology model of *P. falciparum* OMPDC in dimeric forms. The models were constructed by the Phyre program (61). Model of dimeric (*Pf*OMPDC)$_2$, yellow ribbons represent the cross-linking and interacting sites of α8-helix for dimeric (*Pf*OMPDC)$_2$.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

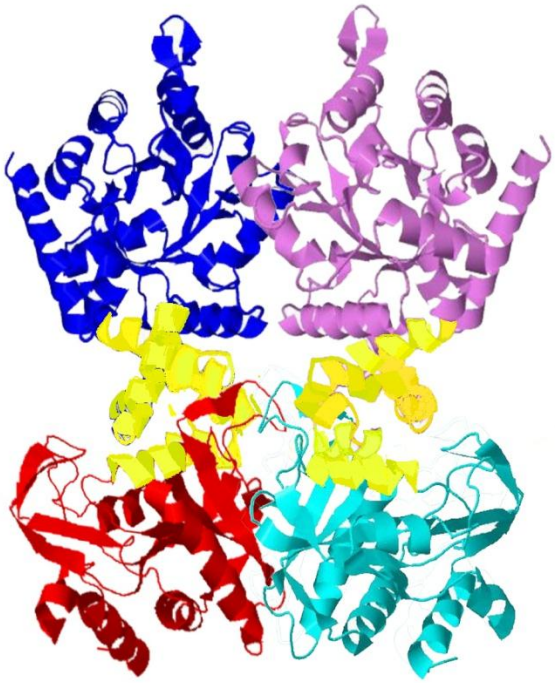


Figure 22 Proposed homology model of heterotetrameric (*Pf*OPRT)$_2$(*Pf*OMPDC)$_2$. The models were constructed by the Phyre program (61). Model of heterotetrameric (*Pf*OPRT)$_2$(*Pf*OMPDC)$_2$ complex, yellow ribbons indicate locations of protein-protein interface and cross-linking sequences for the heterotetrameric (*Pf*OPRT)$_2$(*Pf*OMPDC)$_2$ formation. The (*Pf*OPRT)$_2$ dimer uses α4-helix, where the (*Pf*OMPDC)$_2$ dimer show α2, α3-α4, α5, α6, α7 helices as binding motifs.

CHAPTER V

DISCUSSION AND CONCLUSION

The recent findings have highlighted the potential vulnerability of the human parasite toward agents which affect the pyrimidine biosynthetic pathway (12-20). However, at present only three enzymes have crystal structures and crystallographic analyses: the fourth (*Pf*DHOD) (17), fifth (*Pf*OPRT) (66), and sixth (*Pf*OMPDC) (48, 67) enzymes of the *P. falciparum* pathway. Much importance has been given to these enzymes since compounds presumed to affect the enzyme activities have been shown to be selectively toxic toward the human parasite (17, 18, 20). In humans and most eukaryotes, the genes encoding OPRT and OMPDC are fused into *OPRT-OMPDC* and expressed as bifunctional UMPS (40, 42), although an inversely linked *OMPDC-OPRT* gene, where the OMPDC is located at N-terminus and OPRT at C-terminus, have been reported in the trypanosomatid *Leishmania donovani* (41). With one exception, both genes in the human malarial parasites are acquired from eubacterial origins, are located on different chromosomes, and expressed as a heterotetrameric $(Pf\text{OPRT})_2(Pf\text{OMPDC})_2$ complex with *in vivo* concentration of 30-46 nM, relatively close to the UMPS concentration (17-32 nM) in human cells (16, 21, 22).

Moreover, there are key differences between malaria parasite and human, which are summaries as follows: 1) The malaria parasite genome has A+T rich of 82% (34), whereas the human genome contains A+T about 60% (68, 69). 2) Gene organization, *PfOPRT* and *PfOMPDC* are separated genes (22) but in human both genes are fused (42). 3) The amino acid sequences of *Pf*OPRT and *Pf*OMPDC show long amino acid sequences than human ones (16). 4) *Pf*OPRT and *Pf*OMPDC of the malaria parasite have unique insertions at the N-terminal and internal sequence (16). 5) The LCR is the special characteristic of both *Pf*OPRT and *Pf*OMPDC (31-33). 6) The *Pf*OPRT and *Pf*OMPDC can function as monofunctional enzyme and more active in multifunctional enzyme as heterotetrameric $(Pf\text{OPRT})_2(Pf\text{OMPDC})_2$ (22), whereas the human enzymes are active as dimeric bifunctional UMPS. The difference between the malaria parasite and human is a determining factor to inhibit OPRT and OMPDC in the malaria parasite, without affecting the human host enzyme.

Here, the *Pf*OPRT and *Pf*OMPDC expression systems using the constructed plasmids, *Pf*OPRT-pQE30 Xa vector and *Pf*OMPDC-pTrcHis A vector, are transformed into *E. coli* strain M15 and *E. coli* strain TOP10, respectively. The recombinant proteins are expressed under IPTG induction. Both enzymes are purified by Ni^{2+}-NTA agarose affinity and Hi Trap Q anion exchange chromatography. The activities of

monofunctional OPRT and OMPDC have high V_{max} of 4.7 and 7.9 μmol/min/mg protein, respectively. But the catalytic activity of monofunctional enzyme is 2-3-fold less than multienzyme complex (OPRT: V_{max} about 7.9-8.1 μmol/min/mg protein, k_{cat} 7.3-7.8 s^{-1}; OMPDC: V_{max} 25.1 μmol/min/mg protein, k_{cat} 30.7 s^{-1}) (49, 50). The purity of both *Pf*OPRT and *Pf*OMPDC checked on SDS-PAGE, about 3-4 mg each of nearly 95% from 1 L of *E. coli* cell culture was obtained. Using this strategy, the molecular masses of monomeric *Pf*OPRT and *Pf*OMPDC are 33 and 38 kDa respectively, as expected from the deduced amino acid sequences of the enzymes.

Indeed, the *Pf*OPRT and *Pf*OMPDC enzymes have specific charcters, comparing to the same homologs in other organisms, such as N-terminal insertion, internal insertion and LCRs. Moreover, the *Pf*OPRT and *Pf*OMPDC contribute their functions in a multienzyme complex, which are indicated by the chromatographic experiments of *Pf*OPRT and *Pf*OMPDC on the superpose 12 gel filtration FPLC column demonstrating a single peak with a molecular mass of 140 kDa and also by the SDS-PAGE showing the result of a monomeric form of *Pf*OPRT and *Pf*OMPDC with a molecular mass of 33 and 38 kDa, respectively. The monomers of both enzymes are inactive forms as shown by the previous study (21). The dimeric forms of *Pf*OPRT and *Pf*OMPDC are 67 and 76 kDa, respectively, as determined on the superpose 12 gel

filtration FPLC column (22). The structural pattern of the *Pf*OPRT and *Pf*OMPDC enzyme complex is the $\alpha_2\beta_2$ heterotetrameric form, where α and β represent *Pf*OPRT and *Pf*OMPDC respectively (Figure 23), similar to tryptophan synthase (70) and dihydroorofolate reductase-thymidylate synthase (71). Interestingly, many enzymes are found to be active in the multienzyme complex including the pyrimidine enzymes. For example, aspartate transcarbamoylase (ATC) and dihydroorotase (DHO), the second and the third enzymes of the pathway, form the complex in a eubacterium *Aquifex aeolicus*. Both ATC and DHO are active in trimeric forms and associate into a hexamer of $(ATC)_3(DHO)_3$. The hexamer is then in rapid equilibrium with the dodecamer (72, 73). Nevertheless, the crystal structure of the *A. aeolicus* enzyme complex between ATC and DHO is determined for the first time in the pyrimidine pathway (73).

The structural complex of the malaria parasite enzyme, thus, represents an efficient functional kinetic advantage, which is in line with co-localization principles for evolutional origin, that increases a random mutation in protein, possibly decreasing the free energy enough to create a tightly binding dimer resulted in increasing it concentration, which can amplify the effect of protein. And also with

allosteric property in the protein-protein interactions, in which protein networks or complex systems are regulated by allosteric control (74, 75).

We have, for the first time, examined the protein-protein interactions of the heterotetrameric (*Pf*OPRT)$_2$(*Pf*OMPDC)$_2$ complex using chemical cross-linking between the nearest Lys residues, and mapping the sites of the cross-link with the LC-MS/MS analysis, as previously described for protein structural model (56, 76, 77). Aside from establishing the domain-domain interactions between the *Pf*OPRT (α4 helix) and *Pf*OMPDC (α2, α5, α6, α7 helices, and β turn at α3-α4 position) complex, we have also shown that each *Pf*OPRT (β7 and β8 strands) and *Pf*OMPDC (α8 helix) form a relatively tight dimeric structure. Our Phyre structural model of *Pf*OPRT is comparable to the recent report of the Swiss model using the same template of *S. cerevisiae* OPRT (46), in which the unique insertion of 63 amino acids from N-terminus is omitted due to its low identity (44). Chemical cross-linking and LC-MS/MS results indicate that: 1) the dimer (*Pf*OPRT)$_2$ uses LCR insertion sequences for interaction between chain A and B, 2) the dimer (*Pf*OMPDC)$_2$ is homologous to all known monofunctional OMPDC structures (41, 48, 78) , and 3) the tetramer assembling each dimer of (*Pf*OPRT)$_2$ and (*Pf*OMPDC)$_2$ utilizes both the LCR sequences (α2 and α5 helices of *Pf*OMPDC) and the over-represented hydrophobic amino acid

sequences with some acidic and basic residues (α6 and α7 helices, β turn at α3-α4 position of *Pf*OMPDC; α4 helix of *Pf*OPRT) for the interactions (Figure 24A). The α2 and α5 helices are located at a protruding domain in the crystal structure of *Pf*OMPDC (48). In contrast, the *OPRT* and *OMPDC* genes and their enzymes in *Plasmodium* are different from *L. donovani* and human, but the *Pf*OPRT and *Pf*OMPDC are the most active in tetrameric form like *L. donovani*. Besides the bifunctional *L. donovani* UMPS structural models do not possess any significant domain-domain contacts between OPRT and OMPDC (41) (Figure 24B). The recent study of *L. donovani* enzyme have generated the superposition of human OPRT and OMPDC on *L. donovani* UMPS, showing that the OPRT C-terminus would conjugate with OMPDC N-terminus in view of missing approximately 20 residues of joining region (41) (Figure 24C). In addition, the previous study of bifunctional human UMPS has shown that it is active as a dimer (79), As far as the literature concerned, the homology model in space-filling representation of human enzyme UMPS enzyme follows the dimeric structure as $(OPRT\text{-}OMPDC)_2$, not tetrameric $(OPRT\text{-}OMPDC)_4$ as shown in Figure 24C. The homology model are constructed and then compared for the possible structure of tetrameric $(OPRT)_2(OMPDC)_2$ in the malaria parasite and

$(OPRT\text{-}OMPDC)_2$ in human (Figure 25). They are quite similar for the orientation of the enzyme. However, this needs to be further studied.

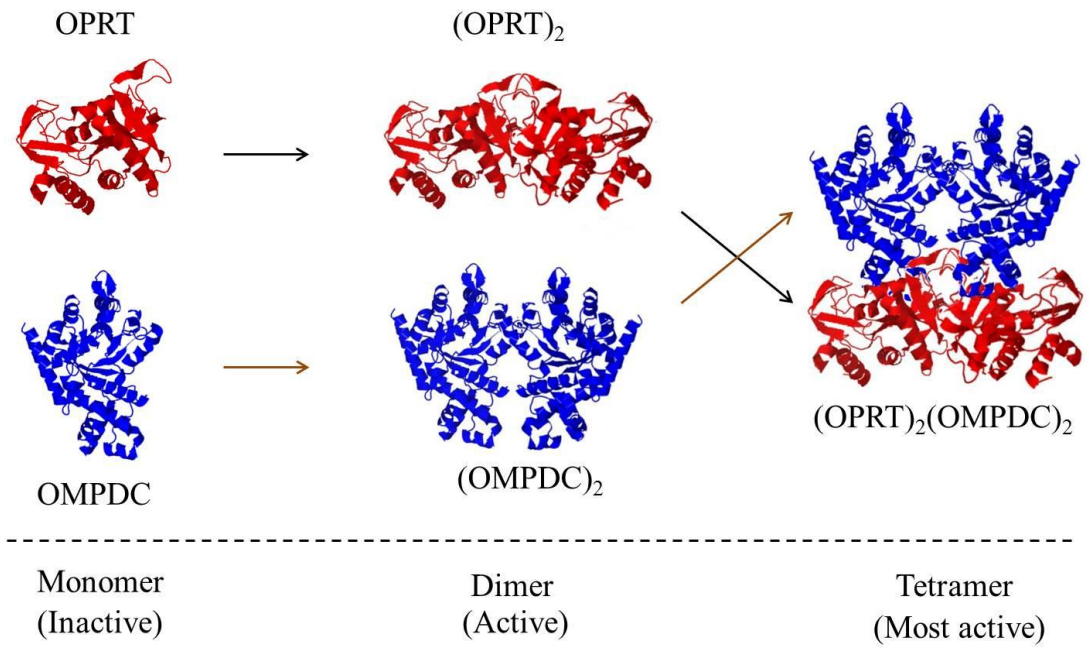


Figure 23 The sequential steps for dimer and tetramer formation of *P. falciparum* OPRT and OMPDC. The inactive monomer OPRT and OMPDC form homodimer $(OPRT)_2$ and $(OMPDC)_2$. Both homodimers are then assosiated into heterotetramer $(OPRT)_2(OMPDC)_2$. The OPRT and OMPDC homology models are shown in red and blue, respectively.

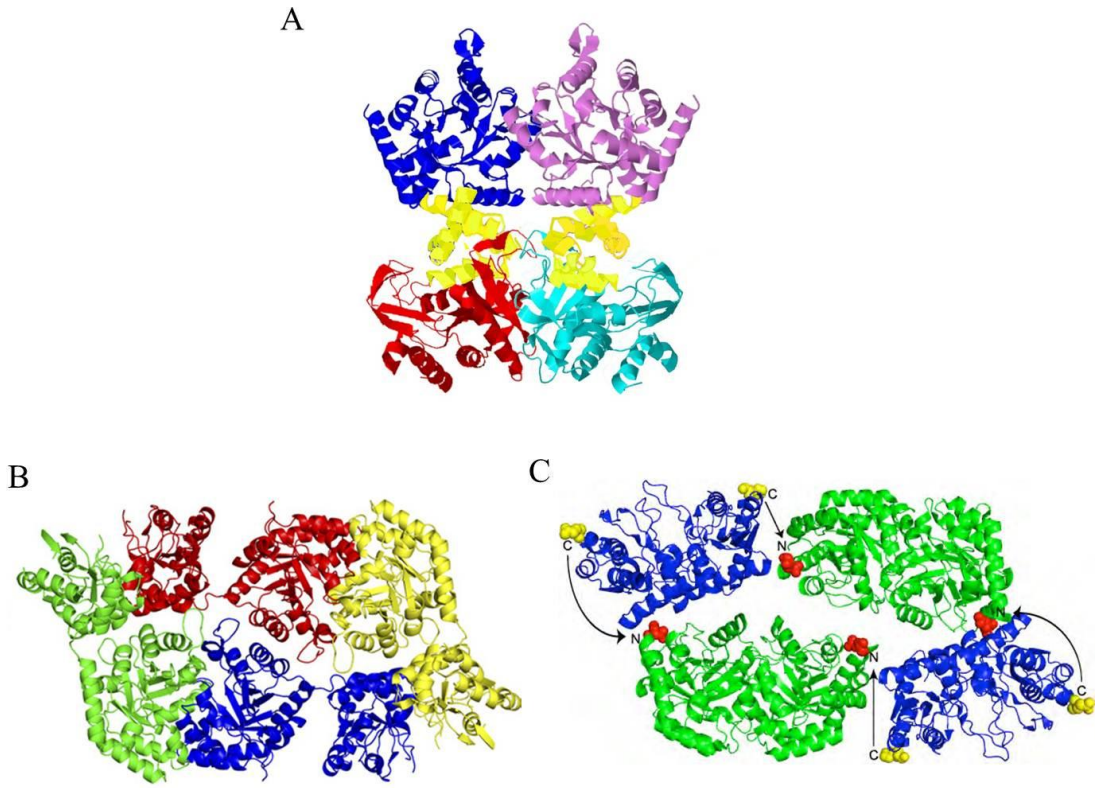


Figure 24 Comparative structural models among *P. falciparum*, *L. donovani* and human OPRT and OMPDC. A: the homology model of tetramer $(OPRT)_2(OMPDC)_2$ of *P. falciparum* (top, dimer $(OMPDC)_2$; bottom, dimer $(OPRT)_2$; yellow ribbons indicate the α helices responsible for binding the two dimers.), B: the crystal structure of bifunctional UMP synthase tetramer $(OPRT\text{-}OMPDC)_4$ of *L. donovani* (PDB ID: 3QW4) and C: the superposition of human OPRT and OMPDC on *L. donovani* UMP synthase (OPRT: blue, PDB ID: 2WNS and OMPDC: green, PDB ID: 2EAW). Arrows indicate the proposed path that joining region between the C-terminus of OPRT domain and N-terminus of OMPDC domain, as OPRT – OMPDC bifunctional enzyme in human.

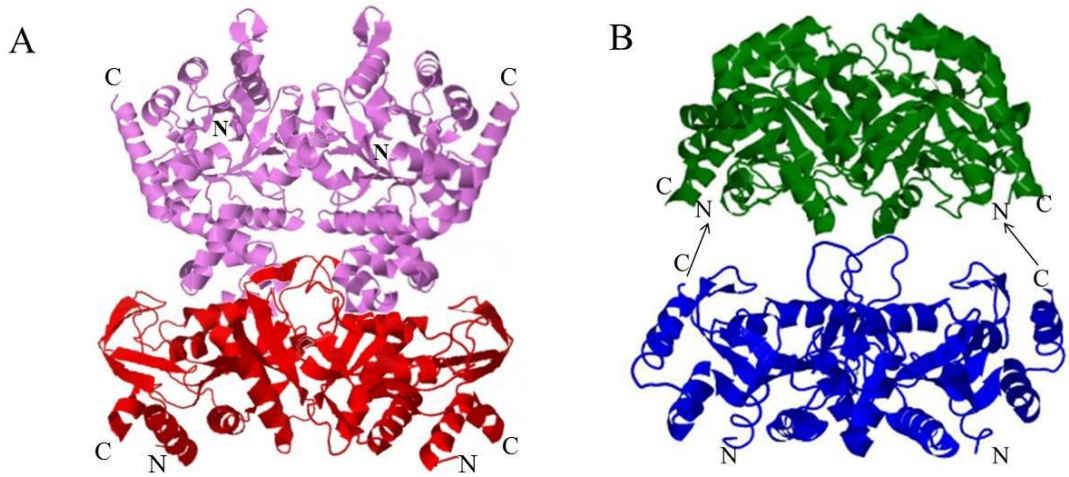


Figure 25 Homology models of tetramer $(OPRT)_2(OMPDC)_2$ of *P. falciparum* and dimer $(OPRT\text{-}OMPDC)_2$ of human. A: the homology model of tetramer $(OPRT)_2(OMPDC)_2$ of *P. falciparum*, where the OPRT (violet) and OMPDC (red) interact each other. B: the homology model of dimer $(OPRT\text{-}OMPDC)_2$ of human (OPRT: blue; OMPDC: green). Arrows indicate the proposed path that joining region between the C-terminus of OPRT domain and N-terminus of OMPDC domain, as described in Figure 24.

Hence, the formation of tetrameric (*Pf*OPRT)$_2$(*Pf*OMPDC)$_2$ is unique in the case of the insert LCR sequence in the PfOMPDC that found in all *Plasmodium* species plays a functional role for domain-domain interactions. The present investigation clearly supports our recent hypothesis that the unique insertions are responsible for PfOPRT-PfOMPDC assembly formation in vivo (22, 48).

The origin of LCR in the malarial parasite proteins remains unknown. The insertion LCR sequences and binding sequences identified in both PfOPRT and PfOMPDC contain acid, basic and hydrophobic amino acids, particular aromatic residues. Taken together, it is plausible that the interactions between PfOPRT and PfOMPDC domains apparently favored in the interfaces include mainly the hydrophobic pairs and some opposite charged pairs or salt bridges, which are supported by computational prediction of protein-protein interfaces (80-83). This conclusion appears to agree well with the recent results on hyperthermostability of the (PfOPRT)2(PfOMPDC)2 complex, comparing to the monofunctional form (49).

In addition to OPRT and OMPDC, insertion region and LCR in *Plasmodium* can be found in many bifunctional enzymes, such as bifunctional S-adenosylmethionine decarboxylase/ornithine decarboxylase (AdoMetDC/ODC), bifunctional dihydrofolate reductase-thymidylate synthase and glucose-6-phosphate dehydrogenase/6-

phosphogluconolactonase (84). For example, *Pf*AdoMetDC/ODC can be separated in three domains: the N-terminal AdoMetDC domain, joining region and the C-terminal ODC domain, and in all domains has an insertion region that play a role in enzyme function and folding (85). The insertion regions have rich charge amino acid residues, predominantly Asn, Asp, Lys, Ser and Glu and in the joining region is Asn-rich (84). Furthermore, these insertions can be existed in other monofunctional enzymes, e.g., protein kinase 7 has a long LCR insert in the amino acid sequence comparing to the same enzyme in the eukaryotic cells (86).

Since the actual protein interfaces are inferred from X-ray crystallographic structures (87, 88), further studies on the crystal structural analyses of the enzyme complex with full length proteins will provide additional information on the binding mechanisms including solvent-inaccessible area, hydrogen bonds, salt bridges, interface residues, and solvation free-energy between the two domains. Use of the interface peptides or inhibitors to disrupt the complex formation can be exploited for a rational drug design approach for more potential antimalarials (89, 90).

In conclusion, the protein-protein interactions between OPRT and OMPDC in homodimeric and heterotetrameric forms of the human malaria parasite are performed using chemical cross-linking in combination with LC-MS/MS and 3D

structure modeling. It is surmised that the human malaria parasite LCR of OPRT and the N-terminal insertion of OMPDC is responsible for heterotetrameric $(OPRT)_2(OMPDC)_2$. The structural complex of the parasite enzymes, thus, represents an efficient functional kinetic advantage, which in line with co-localization principle of evolutional origin, and allosteric regulation in the protein-protein interaction.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

REFERENCES

1. Guerin PJ, Olliaro P, Nosten F, et al. Malaria: current status of control, diagnosis, treatment, and a proposed agenda for research and development. *Lancet Infect Dis*. 2002;2:564-73.
2. Snow RW, Guerra CA, Noor AM, et al. The global distribution of clinical episodes of *Plasmodium falciparum* malaria. *Nature*. 2005;434:214-7.
3. Hay SI, Okiro EA, Gething PW, et al. Estimating the global clinical burden of *Plasmodium falciparum* malaria in 2007. *PloS Med*. 2010;7(e10000290).
4. Dondorp AM, Nosten F, Yi P, et al. Artemisinin resistance in *Plasmodium falciparum* malaria. *N Engl J Med*. 2009;361:455-67.
5. White NJ. Antimalarial drug resistance. *J Clin Invest*. 2004;113:1084-92.
6. Ridley RG. Medical need, scientific opportunity and the drive for antimalarial drugs. *Nature*. 2002;415:686-93.
7. Krungkrai SR, Krungkrai J. Malaria parasite carbonic anhydrase: inhibition of aromatic/heterocyclic sulfonamides and its therapeutic potential. *Asian Paci J Trop Biomed*. 2011;1:233-42.
8. Jones ME. Pyrimidine nucleotide biosynthesis in animals: genes, enzymes, and regulation of UMP biosynthesis. *Annu Rev Biochem*. 1980;49:253-79.
9. Gero AW, O'sullivan WJ. Purines and pyrimidines in malarial parasites. *Blood Cells*. 1990;16:467-84.
10. Krungkrai J, Cerami A, Henderson GB. Pyrimidine biosynthesis in parasitic protozoa: purification of a monofunctional dihydroorotase from *Plasmodium berghei* and *Crithidia fasciculata*. *Biochemistry*. 1990;29:6270-5.
11. Krungkrai J, Cerami A, Henderson GB. Purification and characterization of dihydroorotate dehydrogenase from the rodent malaria parasite *Plasmodium berghei*. *Biochemistry*. 1991;30:1934-9.
12. Krungkrai J, Krungkrai SR, Phakanont K. Antimalarial activity of orotate analogs that inhibit dihydroorotase and dihydroorotate dehydrogenase. *Biochem Pharmacol*. 1992;43:1295-301.
13. Seymour KK, Lyons SD, Phillips L, et al. Cytotoxic effects of inhibitors of *de novo* pyrimidine biosynthesis upon *Plasmodium falciparum*. *Biochemistry*. 1994;33:5268-74.
14. Krungkrai J. Purification, characterization and localization of mitochondrial dihydroorotate dehydrogenase in *Plasmodium falciparum*, human malaria parasite. *Biochim Biophys Acta*. 1995;1243:351-60.

15.Flores MVC, Atkins D, Wade D, et al. Inhibition of *Plasmodium falciparum* proliferation *in vitro* by ribozymes. *J Biol Chem*. 1997;272:16940-5.
16.Krungkrai SR, Aoki S, Palacpac NM, et al. Human malaria parasite orotate phosphoribosyltransferase: functional expression, characterization of kinetic reaction mechanism and inhibition profile. *Mol Biochem Parasitol*. 2004;134:245-55.
17.Hurt DE, Widom J, Clardy J. Structure of *Plasmodium falciparum* dihydroorotate dehydrogenase with a bound inhibitor. *Acta Cryst*. 2006;D62:312-23.
18.Meza-Avina Me, Wei L, Buhendwa Mg, et al. Inhibition of orotidine 5'-monophosphate decarboxylase and its therapeutic potential. *Mini Rev Med Chem*. 2008;8:239-47.
19.Krungkrai SR, Wutipraditkul N, Krungkrai J. Dihydroorotase of human malarial parasite *Plasmodium falciparum* differs from host enzyme. *Biochem Biophys Res Commun*. 2008;366:821-6.
20.Abdo M, Zhang Y, Schramm Vl, et al. Electrophilic aromatic selenylation: new OPRT inhibitors. *Org Lett*. 2010;12:2982-5.
21.Krungkrai SR, Prapunwattana P, Horii T, et al. Orotate phosphoribosyltransferase and orotidine 5'-monophosphate decarboxylase exist as multienzyme complex in human malaria parasite *Plasmodium falciparum*. *Biochem Biophys Res Commun*. 2004;318:1012-8.
22.Krungkrai SR, Delfraino BJ, Smiley JA, et al. A novel enzyme complex of orotate phosphoribosyltransferase and orotidine 5'-monophosphate decarboxylase in human malaria parasite *Plasmodium falciparum*: physical association, kinetics, and inhibition characterization. *Biochemistry*. 2005;44:1643-52.
23.World malaria report 2012: World Health Organization; 2012.
24.Bogitsh BJ, Carter CE, Oeltmann TN. Human parasitology: Elsevier Academic press; 2005.
25.Bray RS, Garnham PCC. The life cycle of malaria parasites. *Br Med Bull*. 1982;38:117-22.
26.Markell EK, Voge M. Medical parasitology. 4 ed: W.B Saunders company; 1976.
27.Gardner MJ, Hall N, Fung E, et al. Genome sequnence of human malaria parasite *Plasmodium falciparum Nature*. 2002;419:498-511.
28.Hall N, Karras M, Raine JD, et al. A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic and proteomic analyses. *Science*. 2005;307:82-6.
29.Carlton JM, Adams JH, Sliva JC, et al. Comparative genomic of the neglected human malaria parasite *Plasmodium vivax*. *Nature*. 2008;455:757-63.
30.Pain A, Bohme U, Berry AE, et al. The genome of the simian and human malaria parasite *Plasmodium knowlesi*. *Nature*. 2008;455:799-803.

31.Xue HY, Forsdyke DR. Low complexity segments in *Plasmodium falciparum* protein are primarily nucleic acid level adaptation. *Mol Biochem Parasitol*. 2003;128:21-32.
32.Frugier M, Bour T, Ayach M, et al. Low complexity regions behave as tRNA sponges to help co-translational folding of plasmodial proteins. *FEBS Lett*. 2011;584:448-54.
33.Pizzi E, Frontali C. Low-complexity regions in *Plasmodium falciparum* proteins. *Genome Res*. 2001;11:218-29.
34.Winzeler EA. Malaria research in the post-genomic era. *Nature*. 2008;455:751-6.
35.Devlin TM. Text book of biochemistry with clinical correlations. 7 ed: John Wiley & Sons, Inc; 2011.
36.Roskoski R. Biochemistry: W.B Saunders company; 1996.
37.Huang M, Graves LM. *De novo* synthesis of pyrimidine nucleotide; emerging interface with signal transduction pathway. *Cell Mol Life Sci*. 2002;60:321-36.
38.Reyes P, Rathod PK, Sanchez DJ, et al. Enzymes of purine and pyrimidine metabolism from human malaria parasite, *Plasmodium falciparum*. *Mol Biochem Parasitol*. 1982;5:275-90.
39.Krungkrai J, Prapunwatana P, Wichitkul C, et al. Molecular biology and biochemistry of malarial parasite pyrimidine biosynthetic pathway. *Southeast Asian J Trop Med Public Health*. 2003;34 Suppl 2:32-43.
40.Nara T, Hashimoto T, Aoki T. Evolutionary implications of the mosaic pyrimidine-biosynthetic pathway in eukaryotes. *Gene*. 2000;257:209-22.
41.French JB, Soysa DR, Yates PA, et al. The *Leishmania donovani* UMP synthase is essential for promastigote viability and has an unusual tetrameric structure that exhibits substrate-controlled oligomerization. *J Biol Chem*. 2011;286:20930-41.
42.Makiuchi T, Nara T, Annoura T, et al. Occurrence of multiple, independent gene fusion events for the fifth and sixth enzymes of pyrimidine biosynthesis in different eukaryotic groups. *Gene*. 2007;394:78-86.
43.De Montigny J, Belarbi A, Hubert JC, et al. Structure and expression of the URA5 gene of *Saccharomyces cerevisiae*. *MOl Gen Genet*. 1989;215:455-62.
44.Zhang Y, Deng H, Schramm V. Leaving group activation and pyrophosphate ionic state at the catalytic site of *Plasmodium falciparum* orotate phosphoribosyltransferase. *J Am Chem Soc*. 2010;132:17023-31.
45.Thompson JD, Higgins DG, Gibson TJ. CLUSTALW: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 1994;22:4673-80.

46.Gonzalez-Segura L, Witte Jf, Mcclard Rw, et al. Ternary complex formation and induced asymmetry in orotate phosphoribosyltransfrase. *Biochemistry*. 2007;46:14075-86.
47.Henriksen A, Aghajari N, Jensen KF, et al. A flexible loop at the dimer interface is a part of the active site of the adjacent monomer of *Escherichia coli* orotate phophoribosyltransferase. *Biochemistry*. 1996;35:3803-9.
48.Tokuoka K, Kusakari Y, Krungkrai SR, et al. Structural basis for the decarboxylation of orotidine 5'-monophosphate (OMP) by *Plasmodium falciparum* OMP decarboxylase. *J Bio Chem*. 2008;143:69-78.
49.Kanchanaphum P, Krungkrai J. Kinetic benefits and thermal stability of orotate phosphoribosyltransferase and orotidine 5'-monophosphate decarboxylase enzyme complex in human malaria parasite *Plasmodium falciparum*. *Biochem Biophys Res Commun*. 2009;390:337-41.
50.Kanchanaphum P, Krungkrai J. Co-expression of human malaria parasite *Plasmodium falciparum* orotate phosphoribosyltransferase and orotidine 5'-monophosphate decarboxylase as enzyme complex in *Escherichia coli*: a novel strategy for drug development. *Asian Biomed*. 2010;4:297-306.
51.Carter NS, Yates P, Arendt CS, et al. Purine and Pyrimidine Metabolism in *Leishmania*. *Adv Exp Med Biol*. 2008;625:141-54.
52.Gao G, Nara T, Nakajima-Shimada J, et al. Novel organization and sequences of five genes encoding all six enzymes for *de novo* pyrimidine biosynthesis in *Trypanosoma cruzi*. *J Mol Biol*. 1999;285:149-61.
53.Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature*. 2003;422:198-207.
54.Borch J, Jorgensen TJ, Roepstorff P. Mass spectrometric analysis of protein interactions. *Chem Biol*. 2005;9:509-16.
55.Walzthoeni T, Leitner A, Stengel F, et al. Mass spectrometry supported determination of protein complex structure. *Curr Opin Struct Biol*. 2013;23:252-60.
56.Back JW, De Jong L, Muijsers AO, et al. Chemical cross-linking and mass spectrometry for protein structural modeling. *J Mol Biol*. 2003;331:303-13.
57.Schulz DM, Ihling C, Clore GM, et al. Mapping the topology and determination of a low resolution three-dimensional structure of the calmodulin-melittin complex by chemical cross-linking and high resolution FTICRMS: direct demonstration of multiple binding modes. *Biochemistry*. 2004;43:4703-15.
58.Smyth MS, Martin JHJ. X-Ray crystallography. *J Clin Pathol:Mol Pathol*. 2000;53:8-14.

59.Schwede T. Computational structural biology: methods and applications: World scientific publishing company; 2008.
60.Bennett-Lovsey RM, Herbert AD, Sternberg MJE, et al. Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. *Protein*. 2007;70:611-25.
61.Kelley LA, Sternberg MJE. Protein structure prediction on the web: a case study using the Phyre server. *Nature Protocols*. 2009;4:363-71.
62.Bradford MM. A rapid and sensitive method for the quantitation of micrograms quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*. 1976;72:248-54.
63.Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. 1970;227:680-5.
64.Shevchenko A, Tomas H, Havlis J, et al. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols*. 2006;1:2856-60.
65.Altschul SF, Madden TL, Schaffer AA, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25:3389-402.
66.Takashima Y, Mizohata E, Tokuoka K, et al. Crystallization and preliminary X-ray diffraction analysis of orotate phosphoribosyltransferase from the human malaria parasite *Plasmodium falciparum*. *Acta Cryst*. 2012;F68:244-6.
67.Krungkrai SR, Tokuoka K, Kusakari Y, et al. Crystallization and preliminary crystallographic analysis of orotidine 5'-monophosphate decarboxylase from the human malaria parasite *Plasmodium falciparum*. *Acta Cryst*. 2006;F62:542-5.
68.I.H.G.S.C. (International Human Genome Sequencing Consortium). Initial sequencing and analysis of the human genome. *Nature*. 2001;409:860-921.
69.Brown TA. Genomes: Oxford: Wiley-Liss; 2002.
70.Mile EW. Structural basis for catalysis by tryptophan synthase. *Adv Enzymol Relat Areas Mol Biol*. 1991;64:93-172.
71.Knighton DR, Ken CC, Howland E, et al. struture of and kinetic channelling in bifunctional dihydrofolate reductase-thymidylate synthase. *Nat Struct Biol*. 1994;1:186-94.
72.Ahuja A, Purcarea C, Ebert R, et al. *Aquifex aeolicus* dihydroorotase: association with aspartate transcarbamoylase switches on catalytic activity. *J Bio Chem*. 2004;279:53136-44.
73.Zhang P, Martin PD, Purcarea C, et al. Dihydroorotase from the hyperthermophile *Aquifex aeolicus* is activated by stoichiometric association with aspartate

trancarbamoylase and forms a one-pot reactor for pyrimidine biosynthesis. *Biochemistry*. 2009;48:766-78.

74.Marcotte ED, Pellegrini M, Ng H-L, et al. Detecting protein function and protein-protein interactions from genome sequences. *Nature*. 1999;284:751-3.

75.Kuriyan J, Eisenberg D. The origin of protein interactions and allostery in colocalization. *Nature*. 2007;450:983-90.

76.Lee YJ. Mass spectrometric analysis of cross-linking sites for the structure of proteins and protein complexes. *Mol Biosyst*. 2008;4:816-23.

77.Leitner A, Walzthoeni T, Kahraman A, et al. Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol Cell Proteomics*. 2010;9:4634-49.

78.Bello AM, Poduch E, Liu Y, et al. Structure-activity relationships of C6-uridine derivatives targeting *Plasmodia* orotidine monophosphate decarboxylase. *J Med Chem*. 2008;51:439-48.

79.Yablonski MJ, Pasek DA, Han B-D, et al. Intrinsic activity and stability of bifunctional human UMP synthase and its two separate catalytic domain, orotate phosphoribisyltransferase and orotidine-5'-phosphate decarboxylase. *J Biol Chem*. 1996;271:10704-8.

80.Karlin S, Brocchieri L, Bergman A, et al. Amino acid runs in eukaryotic proteomes and disease associations. *Proc Natl Acad Sci USA*. 2002;99:333-8.

81.Ofran Y, Rost B. Analysing six types of protein-protein interfaces. *J Mol Biol*. 2003;325:377-87.

82.Skrabanek L, Saini HP, Bader GD, et al. Computational prediction of protein-protein interactions. *Mol Biotechnol*. 2008;38:1-17.

83.Yan C, Wu F, Jernigan Rj, et al. Characterization of protein-protein interfaces. *Protein J*. 2008;27:59-70.

84.Birkholtz L-M, Wrenger C, Joubert F, et al. Parasite-specific inserts in the bifunctional S-adenosylmethionine decarboxylase/ornithine decarboxylase of *Plasmodium falciparum* modulate catalytic activities and domain interactions. *Biochem J*. 2004;377:439-48.

85.Krause T, Luersen K, Wrenger C, et al. The oronithine decaboxylase domain of the bifunctional ornithine decarboxylase/*S*-adenosylmethionine decarboxylase of *Plasmodium falciparum*: recombinant expression and catalytic properties of two different constructs. *Biochem J*. 2000;352:287-92.

86.Merckx A, Echalier A, Langford K, et al. Structure of *P. falciparum* protein kinase 7 identify an activation motif and leads for inhibitor design. *structure*. 2008;16:228-38.

87.Dafforn TR. So how do you know you have a macromolecular complex? *Acta Cryst*. 2007;D63:17-25.
88.Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol*. 2007;372:774-97.
89.Singh SK, Maithal K, Balaram H, et al. Synthetic peptides as inactivators of multimeric enzymes: inhibition of *Plasmodium falciparum* triosephosphate isomerase by interface peptides. *FEBS Lett*. 2001;501:19-23.
90.Wells JA, Mcclendon CL. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature*. 2007;450:1001-9.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

APPENDIX

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

VITA

Name: Waranya Imprasittichai

Student ID: 5287817420

Sex: Female

Date of birth: July 29, 1986

Place of birth: Nakorn Pathom, Thailand

Nationality: Thai

Education: 2005-2008 Bachelor of Science in Biomedical sciences, Rangsit University, Thailand

Publication:

Waranya Imprasittichai, Sittiruk Roytrakul, Sudaratana R. Krungkrai, Jerapan Krungkrai. A unique insertion of low complexity amino acid sequence underlies protein-protein interaction in human malaria parasite orotate phosphoribosyltransferase and orotidine 5'-monophosphate decarboxylase. Asian Pac J Trop Med.2014;7:184-192 .

Oral presentation:

At The 17th Biological Science Graudate Congress, Hosted by Faculty of Science, Chulalongkorn University, in topic A unique insertion of low complexity amino acid sequence underlies protein-protein interaction in human malaria parasite orotate phosphoribosyltransferase and orotidine 5' monophosphate decarboxylase. In collaboration with Departmant of Biological Science, National University of Singapore, Singapore and Institute of Biological Sciences, Faculty of Science, University of Malaya, Kuala Lumpur, Malasia. December 8th-10th,2012.