



รายงานวิจัยฉบับสมบูรณ์

โครงการ ศึกษาโครงสร้างของโปรตีน
และการทำงานของเอนไซม์ในระดับโมเลกุล
Protein Structure and Molecular Enzymology

โดย

ศ. ม.ร.ว. ชินุสรณ์ สวัสดิวัตน์
Professor M.R. Jisnuson Svasti

กรกฎาคม 2547

รายงานวิจัยฉบับสมบูรณ์

โครงการ ศึกษาโครงสร้างของโปรตีน และการทำงานของเอนไซม์ในระดับโมเลกุล Protein Structure and Molecular Enzymology

ศ. ม.ร.ว. ชินุสรณ์ สวัสดิวัตน์
Professor M.R. Jisnuson Svasti

สังกัด

ภาควิชาชีวเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล
ถนนพระราม 6 พญาไท ราชเทวี กทม. 10400

และ

ห้องปฏิบัติการชีวเคมี สถาบันวิจัยจุฬาภรณ์
ถนนวิภาวดีรังสิต ดอนเมือง กทม. 10210

สนับสนุนโดยสำนักงานกองทุนสนับสนุนการวิจัย

(ความเห็นในรายงานนี้เป็นของผู้วิจัย สกว. ไม่จำเป็นต้องเห็นด้วยเสมอไป)

Principal Investigator:

Professor M.R. Jisnuson Svasti

ศ. ม.ร.ว. ชัยนุสรณ์ สวัสดิวัตน์

Professor (C-11 Level), Department of Biochemistry, Faculty of Science,
Mahidol University, Rama VI Road, Bangkok 10400, Thailand.

Tel: 02 2015845, 2015840; Fax: 02 2015843, 3547174; E-mail: scjsv@mahidol.ac.th
and

Head, Laboratory of Biochemistry, Chulabhorn Research Institute,
Vibhavadee Rangsit Road, Bangkok 10210, Thailand

Tel: 02 5740622 to 31; ext. 1305; Fax: 02 5742027; E-mail: jisnuson@tubtim.cri.or.th

Co-Investigators:

1. Rudee Surarit

รศ. ดร. ฤดี สุราฤทธิ

Associate Professor, Department of Physiology and Biochemistry,
Faculty of Dentistry, Mahidol University,
Yothi Street, Bangkok 10400, Thailand.

Tel: 02 6448638; 02 6448644-46, ext 1415-16; Fax: 02 2466910

E-mail: dtrsr@mahidol.ac.th

2. James R. Ketudat Cairns

ผศ. ดร. เจมส์ เกตุทัต คาร์นส์

Assistant Professor, School of Chemistry,
Institute of Science, Suranaree University of Technology,
111 University Avenue, Muang District,
Nakhon Ratchasima 30000, Thailand.

Tel:(044) 224 304; Fax: (044) 224 185

E-mail: cairns@ccs.sut.ac.th, cairns@sura2.sut.ac.th, jrkcairns@yahoo.com

3. Jirundon Yuvaniyama

อ. ดร. จิรันดร ยูวะนิยม

Lecturer, Department of Biochemistry,
Faculty of Science, Mahidol University,
Rama VI Road, Bangkok 10400, Thailand.

Tel: 02 2015603; Fax: 02 3547174

e-mail: scjyv@mahidol.ac.th

4. **Palangpon Kongsaree** **รศ. ดร.พลังพล คงเสรี**

Associate Professor, Department of Chemistry,
Faculty of Science, Mahidol University,
Rama VI Road, Bangkok 10400, Thailand
Tel : 02 2015190, 2015848 Fax : 02 2458332
E-mail: scpks@mahidol.ac.th

5. **Pimchai Chaiyen** **ผศ. ดร. พิมพ์ใจ ใจเย็น**

Assistant Professor, Department of Biochemistry,
Faculty of Science, Mahidol University,
Rama VI Road, Bangkok 10400, Thailand.
Tel: 02 2015607; Fax: 02 3547174
e-mail: scpky@mahidol.ac.th

6. **Patjaraporn Wongvithoonyaporn** **อ.ดร. กัชรพร วงศ์วิฑูรยาพร**

Lecturer, Department of Biochemistry,
Faculty of Liberal Arts and Sciences,
Kasetsart University, Kamphangsaeen Campus,
Nakhorn Pathom 73140, Thailand.

Tel: 034 281106 ext. 462; Fax: 034 351894
E-mail: faaspw@nontri.ku.ac.th; patjaraporn@yahoo.com

7. **Apinya Chaivisuthangkura (Buranaprapuk)**
อ. ดร. อภิญญา ชัยวิสุทธางกูร (บุรณประพุกษ์)

Lecturer, Department of Chemistry,
Faculty of Science, Srinakharinwirot University Prasarnmitr,
Sukhumvit 23, Bangkok 10110, Thailand.
Phone: 02 6641000 ext. 8452; 02 2592097; Fax: 02 2592097
Email: apinyac@swu.ac.th

8. **Sujint Anguravirutt** **อ.ดร. สุจินต์ อังกราวีรุต**

Lecturer, Department of Chemistry,
Faculty of Science, Mahasarakham University
Mahasarakham 44150, Thailand
Tel & Fax: 043 754246
E-mail: sujint.a@msu.ac.th

9. Wipa Suginta

อ. ดร. วิภา สุจินต์

Lecturer, School of Chemistry,
Institute of Science, Suranaree University of Technology,
111 University Avenue, Muang District,
Nakhon Ratchasima 30000, Thailand
Tel 044 224169; Fax: 044 224185
E-mail: wipa@ccs.sut.ac.th

10. Pramvadee Y. Wongsangchantra

อ. ดร. เปรมวดี วงษ์แสงจันทร์

Lecturer, Department of Biotechnology,
Faculty of Science, Mahidol University,
Rama VI Road, Bangkok 10400, Thailand.
Tel: 02 2015316; Fax: 02 2463026
E-mail: scpyv@mahidol.ac.th

11. Prachumporn Toonkool

อ. ดร. ประชุมพร ทุนกุล

Lecturer, Department of Biochemistry,
Faculty of Science, Kasetsart University,
50 Paholyothin Rd., Bangkok 10900, Thailand.
Tel: 02 9428526-8, ext. 129; Fax: 02 5614627
E-mail: fscippt@ku

Acknowledgment

The financial support of the Thailand Research Fund, through the award of a Senior Research Scholar grant # RTA/02/2544 to Professor M.R. Jisnuson Svasti between August 2001 to July 2004, is gratefully acknowledged.

Abstract

This project had the aim of increasing Thailand's capability to undertake protein research, through establishment of modern research facilities, development of human resources, performance of high quality research publishable in international journals, and promotion of protein research. In terms of research facilities, our laboratory at the Chulabhorn Research Institute has equipment and expertise for determining the primary structure of a protein and for proteomic studies of gene expression, while the Center for Protein Structure and Function at Mahidol University is well equipped for x-ray crystallographic study of three-dimensional structure and for study of enzymes, including under the pre-steady state conditions. In terms of human resource development, the group consisted of 11 Ph.D. level scientists, mostly young investigators, from 5 universities. In addition, five other young Ph.D.s were linked as advisees to the principal investigator. Over the three year period of the grant, four Ph.D. and nine M.Sc. students graduated, and another 16 Ph.D. students and 15 M.Sc. students are presently enrolled.

Research covered three main themes: a) Protein structure-function relationships; b) Protein changes in disease; and c) Applications of enzymes in biotechnology. Understanding the relationship between the three-dimensional structure and the function of a protein not only helps us to understand the protein's mechanism of action, but can also lead to the design of drugs to inhibit the protein or the engineering of proteins to improve various properties. Proteins studied include enzymes involved in the synthesis of penicillin derivatives with potential uses in the pharmaceutical industry, glycosidase enzymes with potential uses in enzymatic synthesis of glycosides for pharmaceutical or cosmetic industries, and aromatic hydroxylases with potential applications in environmental remediation. In addition, chemical models for enzyme action, such as the cleavage of proteins, will provide better understanding of catalysis. Studies on protein changes in disease include the characterization of mutations in genetic diseases, such as the abnormal hemoglobins and inborn errors of metabolism, which will provide better understanding of these diseases and lead to improved diagnosis. Proteomic studies of cancer, involving both surgical specimens and cancer cell lines provide information on possible biomarkers for early detection and possible targets for chemotherapy. Finally, in terms of biotechnology, the isolation of a sericin-specific protease for degumming silk, if successful, will lower costs, lessen environmental problems and may potentially improve the quality of Thai silk.

Results from the research described yielded 22 publications in international journals, and two articles in international proceedings volumes, with some 5 papers were in high impact journals with impact factor greater than 4.0. In addition, 27 abstracts were presented at international meetings and another 54 abstracts presented at national meetings during the grant period.

Activities in promoting protein science and molecular enzymology included the extended visit of Nobel Laureate William Lipscomb as Stang Mongkolsuk Distinguished Professor, which had a broad impact not only on the group and in the academic community, but also on the government and the general public. An informal Protein Research Network was started to disseminate news and events to members by email. In addition, the research group organized various symposia, meetings and seminars, most notably a two-day Protein Research Network Symposium, attended by over 200 researchers from some 23 universities and institutions throughout Thailand.

Finally, the activities of our research group were well recognised by the scientific community through several awards. This included the Outstanding Scientist of Thailand Award from the Foundation for the Promotion of Science and Technology and the Outstanding National Researcher Award, Chemical Sciences and Pharmacy section, National Research Council of Thailand. Other awards included two Young Scientist of Thailand Awards from the Foundation for the Promotion of Science and Technology, a UNESCO-L'Oreal Women in Science Fellowship, and an Outstanding Thesis Award from the National Research Council of Thailand, Chemical Science and Pharmacy Section. Other honours included an Outstanding Lecturer Award from the Faculty Club, Faculty of Science, Mahidol University, and an Exemplary Lecturer Award from the Faculty Club, Mahidol University. These various awards confirm the quality of work performed by researchers in this Senior Research Scholar grant.

KEYWORDS: Protein / Enzyme/ Three-dimensional structure/ Kinetics/ Diseases

บทคัดย่อ

โครงการวิจัยนี้มีวัตถุประสงค์ที่จะเพิ่มความสามารถในการทำการวิจัยเรื่องโปรตีนของประเทศไทยโดยการจัดเตรียมเครื่องมือสมัยใหม่ต่างๆ การพัฒนาทรัพยากรมนุษย์ การทำงานวิจัยคุณภาพสูงที่สามารถตีพิมพ์ได้ในวารสารนานาชาติ และการส่งเสริมงานวิจัยต่างๆที่เกี่ยวข้องกับการศึกษาโปรตีนในด้านการจัดหาเครื่องมือสมัยใหม่ต่างๆ ห้องปฏิบัติการชีวเคมี สถาบันวิจัยจุฬาภรณ์ ได้มีเครื่องมือที่ทันสมัยในการวิเคราะห์โครงสร้างปฐมภูมิของโปรตีนและการศึกษาโปรตีโอมิกเพื่อการแสดงออกของยีนต่างๆ ในขณะที่ทางห้องปฏิบัติการของศูนย์ความเป็นเลิศทางวิชาการ โครงสร้างและหน้าที่การทำงานของโปรตีน คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดลได้มีการจัดหาเครื่องมือสำหรับการศึกษการหักเหของรังสีเอกซ์ เพื่อใช้ในการศึกษาโครงสร้างสามมิติของโปรตีนต่างๆ และสำหรับการศึกษาเรื่องเอนไซม์ได้มีการจัดเตรียมเครื่องมือที่ใช้ในการศึกษาจลนศาสตร์ของเอนไซม์ต่างๆรวมทั้งการศึกษายาได้สภาวะก่อนสมดุล ในด้านการพัฒนากำลังคน กลุ่มวิจัยประกอบด้วยนักวิจัยระดับปริญญาเอก 11 คน ส่วนใหญ่เป็นนักวิจัยรุ่นใหม่ จาก 5 มหาวิทยาลัย นอกจากนี้ยังมีนักวิจัยระดับปริญญาเอกรุ่นใหม่อีก 5 คนที่เป็นนักวิจัยภายใต้การดูแล ให้คำปรึกษาของหัวหน้าโครงการวิจัยนี้ ในระยะเวลา 3 ปีของการรับทุนเมธีวิจัยอาวุโสนี้ได้มีการผลิตดุษฎีบัณฑิต 4 คน และมหาบัณฑิต 9 คน และในปัจจุบันโครงการมีนักศึกษาระดับปริญญาเอก 16 คนและนักศึกษาระดับปริญญาโท 15 คน

งานวิจัยของโครงการนี้ครอบคลุม 3 เรื่องใหญ่ๆได้แก่ ก) การศึกษาความสัมพันธ์ของโครงสร้างและหน้าที่การทำงานของโปรตีน ข) การเปลี่ยนแปลงของโปรตีนในโรคต่างๆ และ ค) การประยุกต์ใช้เอนไซม์ทางเทคโนโลยีชีวภาพ การเข้าใจถึงความสัมพันธ์ระหว่างโครงสร้างสามมิติและการทำงานของโปรตีนไม่เพียงแต่จะช่วยให้เข้าใจกลไกการทำงานของโปรตีนเท่านั้นแต่ยังนำไปสู่การออกแบบยาต่างๆที่ยับยั้งการทำงานของโปรตีนหรือเป็นข้อมูลพื้นฐานที่นำไปใช้ในการออกแบบและการทำวิศวกรรมโปรตีนเพื่อพัฒนาคุณภาพของโปรตีนต่อไป สำหรับโปรตีนจำพวกเอนไซม์ได้ศึกษาเอนไซม์ที่เกี่ยวข้องในการสังเคราะห์อนุพันธ์ของเพนนิซิลินที่มีศักยภาพสูงที่จะนำไปใช้ในอุตสาหกรรมยา มีการศึกษาเอนไซม์ไกลโคซิเดสที่มีศักยภาพในการนำไปใช้ในการสังเคราะห์สารจำพวกไกลโคไซด์ที่ใช้ในอุตสาหกรรมยาและอุตสาหกรรมเครื่องสำอาง และศึกษาเอนไซม์อะโรมาติก ไฮโดรเลสซึ่งมีศักยภาพในการจัดการมลพิษในสิ่งแวดล้อม นอกจากนี้ยังได้ศึกษาตัวอย่างโมเดลทางเคมีสำหรับการทำงานของเอนไซม์ เช่นการย่อยโปรตีน ซึ่งจะช่วยให้เข้าใจกลไกการเร่งปฏิกิริยาของเอนไซม์สำหรับการศึกษการเปลี่ยนแปลงของโปรตีนในโรคต่างๆ ได้รวมถึงการศึกษวิเคราะห์การผ่าเหล่าของยีนในโรคพันธุกรรม เช่นโรคความผิดปกติของฮีโมโกลบินและในโรคพันธุกรรมบกพร่องของเมตาบอลิซึม ซึ่งจะช่วยให้เข้าใจกลไกการเกิดโรคและอาจนำไปสู่การพัฒนาวิธีการตรวจ วินิจฉัยที่มีประสิทธิภาพ การศึกษาโปรตีโอมิกของโรคมะเร็งซึ่งทำการศึกษาทั้งในเนื้อเยื่อจากผู้ป่วยโรคมะเร็งและในเซลล์เพาะเลี้ยงสายพันธุ์ต่างๆให้ข้อมูลพื้นฐานในการหาตัวบ่งชี้ทางชีวภาพของโรคมะเร็งเพื่อพัฒนาวิธี

การตรวจวิเคราะห์เรื่องต่างๆในระยะเริ่มต้น และอาจเป็นเป้าหมายในการรักษาทางเคมีต่อไป สุดท้ายในด้านของเทคโนโลยีชีวภาพมีการศึกษาเรื่องการแยกเอนไซม์ย่อยโปรตีนจำเพาะในเส้นไหม ซึ่งหากสำเร็จจะเป็นประโยชน์อย่างยิ่งในการลดค่าใช้จ่าย ลดมลภาวะและอาจนำไปสู่การปรับปรุงคุณภาพไหมไทยให้ดียิ่งขึ้น

ผลการวิจัยจากโครงการนี้สามารถตีพิมพ์ในวารสารระดับนานาชาติได้ทั้งสิ้น 22 เรื่อง และมี 2 เรื่องตีพิมพ์ในรายงานการประชุมวิชาการระดับนานาชาติ โดยบทความ 5 เรื่องตีพิมพ์ในวารสารที่มีแฟกเตอร์ผลกระทบ (impact factor) สูงกว่า 4.0 นอกจากนี้ในช่วงเวลาที่รับทุนนี้ ยังได้มีบทความนำเสนอผลงานในการประชุมวิชาการระดับนานาชาติ 27 เรื่อง และ บทความแนะนำเสนอผลงานในการประชุมวิชาการระดับชาติอีก 54 เรื่อง

มีกิจกรรมที่ส่งเสริมการศึกษาเรื่องวิทยาศาสตร์โปรตีนและการศึกษาเอนไซม์ระดับโมเลกุล ได้แก่ การเชิญศาสตราจารย์ วิลเลียม ลิปสคอม นักวิทยาศาสตร์ผู้ได้รับรางวัลโนเบลในเรื่องของการศึกษาโครงสร้างโปรตีนมาเป็นศาสตราจารย์รับเชิญเกียรติคุณสตาจ์ มงคลสุข ของคณะวิทยาศาสตร์ มหาวิทยาลัยมหิดลเป็นระยะเวลา 3 เดือน ซึ่งมีประโยชน์อย่างยิ่งต่อทั้งทางกลุ่มวิจัยและต่อสังคมวิชาการทั้งภาครัฐและสาธารณชนทั่วไป ได้เริ่มมีการก่อตั้งเครือข่ายการวิจัยโปรตีนอย่างไม่เป็นทางการโดยมีการเผยแพร่ข่าวสารเกี่ยวกับการประชุมและการบรรยายทางวิชาการต่างๆทางจดหมายอิเล็กทรอนิกส์แก่สมาชิก นอกจากนี้กลุ่มวิจัยโปรตีนนี้ยังได้จัดการประชุมทางวิชาการและการสัมมนาต่างๆหลายครั้งที่สำคัญคือการจัดการประชุมเครือข่ายการวิจัยโปรตีน (Protein Research Network Symposium) เป็นเวลา 2 วัน ซึ่งมีนักวิจัยเข้าร่วมการประชุมกว่า 200 คนโดยผู้เข้าร่วมประชุมมาจากมหาวิทยาลัยและสถาบันการศึกษาต่างๆทั่วประเทศไทยรวมทั้งสิ้น 23 แห่ง

กลุ่มการวิจัยนี้ได้รับการยอมรับจากสังคมวิชาการของประเทศไทยในระดับสูงโดยมีนักวิจัยในกลุ่มวิจัยได้รับรางวัลต่างๆ หลายรางวัล ได้แก่ รางวัลนักวิทยาศาสตร์ดีเด่นแห่งชาติ และรางวัลนักวิทยาศาสตร์รุ่นใหม่ จากมูลนิธิส่งเสริมวิทยาศาสตร์และเทคโนโลยีแห่งประเทศไทย รางวัลนักวิจัยดีเด่นแห่งชาติ และรางวัลผลงานวิทยานิพนธ์ดีเด่น สาขาวิทยาศาสตร์เคมีและเภสัช จากสำนักงานคณะกรรมการวิจัยแห่งชาติ รางวัล UNESCO-L'Oreal Women in Science Fellowship รางวัลอาจารย์ดีเด่นจากสภาคณาจารย์ คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล และรางวัลอาจารย์ตัวอย่างจากสภาคณาจารย์ มหาวิทยาลัยมหิดล ซึ่งรางวัลต่างๆ ที่นักวิจัยในกลุ่มได้รับสะท้อนให้เห็นถึงคุณภาพงานวิจัยของนักวิจัยในทุนเมธีวิจัยอาวุโส

KEYWORDS: โปรตีน / เอนไซม์ / โครงสร้างสามมิติของโปรตีน / จลนศาสตร์ของเอนไซม์ / โรคพันธุกรรมและโรคมะเร็ง

I. EXECUTIVE SUMMARY

The major aim of this project is to develop research on Protein Science in Thailand. This involves both establishment of research facilities and development of human resources. In terms of research facilities, the two laboratories at the Mahidol University and Chulabhorn Research Institute (CRI) have complementary equipment, which allows the group to study diverse aspects of protein structure and function. Thus, the Laboratory of Biochemistry at CRI has, for several years, been one of the few laboratories with a routinely functional amino acid analyzer and protein sequencer for study of primary structure. More recently, our laboratory at CRI was the first laboratory in Thailand to use the proteomics approach. This capability has recently been strengthened by the acquisition of a new Q/TOF micro mass spectrometer for identification of samples. In parallel to this, our group at Mahidol University has been designated as the Center for Excellence in Protein Structure and Function of the Faculty of Science. Major equipment items, amounting to approximately 30 million baht, have been provided by the faculty and are fully functional. This includes the first macromolecular X-ray crystallographic apparatus in Thailand, allowing study of the three-dimensional structures of proteins in-house. In addition, a stopped-flow spectrophotometer has been purchased allowing study of enzyme reactions at the pre-steady state stage, so that reaction intermediates of the reaction may be determined.

However, no matter how sophisticated the instrumentation available, few research advances can be made without qualified personnel. Thus, our research team consists of 11 Ph.D.-level co-investigators, four from within the Faculty of Science, one from another faculty at Mahidol University, and 6 from other universities, namely Suranaree University of Technology, Kasetsart University, Srinakharinwirot University, and Mahasarakham University. Most of whom are young investigators seeking to establish themselves. In addition, Professor Jisnuson Svasti became mentor to five other young Ph.D.s from Khonkaen University, Mae Fah Luang University, Mahidol University and Naresuan University, three of whom came from the Researcher Encouragement Project. Thus overall, staff linked to the project come from 7 other universities apart from Mahidol University, extending the reach of the grant to universities, which are less developed in research. Most of these staff have managed to find their own funding and to publish their work in international journals.

Research progress has been steady, and advances were made in all projects. Much of the work concerns the glycosidase enzymes, in particular β -glucosidases. β -Glucosidase enzymes with various substrate specificities are now being studied from 6 or more species, with the primary aim of studying structure-function relationships. This includes study of hydrolysis, reverse hydrolysis and transglucosylation reactions. Cassava linamarase has been found to show unique capability for transglucosylation of alcohols, including tertiary alcohols, and is being used for synthesis of chiral and other glycosides. Kinetic studies of the reactivation of inactivated enzyme by various alcohols are

providing better insight into the effectiveness of the cassava enzyme. In addition, cloning, expression and site-directed mutagenesis, and crystallization are being performed with both the Thai rosewood and cassava enzymes, but problems in protein expression still need to be solved. Additionally purification and characterisation of glucosidases and their substrates are being carried out on *Dalbergia nigrescens*, *Solanum torvum*, and *Plumeria obtusa*. In particular, preliminary data suggest that the *Solanum torvum* enzyme is family 3 glycosidase, despite having a steroid glucoside as substrate. Studies of rice β -glucosidases, including the detailed characterization of one enzyme, are opening new avenues of research, since there appear to be several β -glucosidases, which may use different substrates and perform different functions. Other work includes screening for novel glycosidases, and purification of α -mannosidases, which appears to be difficult, but recent results show some promise. In addition, the sequence and catalytic properties of *Vibrio* chitinase have been determined, including the development of novel mass spectroscopic techniques the hydrolysis reaction.

With the enzymes involved in synthesis of penicillin derivatives, X-ray diffraction data obtained from crystals of *Bacillus megaterium* penicillin G acylase (PGA) are being used for structure determination of the enzyme. The electron-density map is being interpreted and model building of the structure is in progress. The structure of *B. megaterium* PGA appears to be quite different from that of *Escherichia coli* PGA, which was used as a template for determination of initial phases, causing difficulties in structure determination. However, various approaches are being employed to improve map calculations and thus interpretation. The availability of three-dimensional structure of PGA will eliminate the errors associated with the currently used homology model and greatly facilitate the engineering of PGA specificity to recognize cephalosporin C substrate. For the X-ray crystal structure determination of the stereoinverting D-PhgAT from *Pseudomonas stutzeri*, the structure has been reinvestigated in a lower space group, P3 (1), with two molecules of D-PhgAT in the asymmetric unit. The structures of the two crystallographically dependent molecules differing around the catalytic region revealed that the structure of D-PhgAT is an asymmetric dimer. The current model has over 90% of the structure. The pyridoxal cofactor was clearly visible and covalently linked to the catalytic Lys-269. Further studies on the structure of the enzyme complexed with the inhibitor are in progress.

With the flavoprotein oxygenases, the mechanism of MHPCO was studied by replacing the natural cofactor FAD with FAD-analogues. Redox potential values of the reconstituted enzymes correlate well with the electronic effects of the 8-substituents of the FAD analogue. Pre-steady state kinetics of MHPCO was investigated using stopped-flow spectrophotometry, and showed that the rates of hydroxylation were 2.0, 1.5, 1.3, and 0.9 s⁻¹ for MHPCO, reconstituted with 8-CN, 8-Cl, 8-OCH₃, 8-CH₃ -FAD respectively. Studies of the oxidative half-reaction of MHPCO revealed that the oxygenation reaction of MHPCO occurs via an electrophilic aromatic substitution mechanism analogous to the mechanisms for parahydroxybenzoate and phenol

hydroxylases. On the other hand, the enzyme HPAH appears to consist of two protein components. C_1 is a reductase enzyme with subunit size of 32 kDa, catalyzing HPA-stimulated NADH oxidation without hydroxylation of HPA. C_2 is a tetrameric enzyme with subunit of 50 kDa, lacking a redox center, which hydroxylates HPA in the presence of C_1 . The complete genes of both C_1 and C_2 fragments were expressed in an *E. coli* system with good yield. The K_d for binding of FMN to C_1 was determined to be $0.02 \mu\text{M}$, and results from stopped-flow spectrophotometry indicated that $k_{\text{on}} = 1.7 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ and $k_{\text{off}} = 0.014 \text{ s}^{-1}$. Investigation of flavin specificity showed that C_2 is the only oxygenase component of the enzyme among the two-protein component aromatic hydroxylases capable of using three forms of flavin for hydroxylation. Thus, HPAH from *A. baumannii* is a novel prototype enzyme among the two-protein component aromatic hydroxylases.

In the study of chemical proteinases, two new bifunctional pyrenyl probes, L-phenylalanine-4(1-pyrenyl)butyramine chloride (Phe-L-Py) and L-phenylalanine-4(1-pyrenyl)methyramine chloride (PMA-L-Phe) were synthesized. Cleavage of both BSA and lysozyme is negligible when Phe-L-Py is used. However, PMA-L-Phe carrying a free amino terminus, photocleaves lysozyme with high efficiency and specificity, but shows negligible cleavage of BSA. Photocleavage of lysozyme results in at least two new fragments of molecular weights 11,000, with N-terminal sequence KVFG (identical to the N-terminus of lysozyme), and a 3,000 fragment of N-terminal sequence VAWRN, with a modification of the tryptophan. This indicates that cleavage by PMA-L-Phe occurs at Trp108-Val109, identical to that reported for Py-L-Phe (the probe carrying a free carboxyl terminus).

Some progress has been made in searching for proteinases, which are specific for sericin for use in degumming of silk. Several strains of microorganisms have been shown to produce proteinases that degrade sericin, and novel screening procedures have been developed and tested in an attempt to find enzymes that degrade sericin, but do not cleave fibroin.

Finally, research on human diseases at the Chulabhorn Research Institute continues to show novel abnormal hemoglobins, not previously found in Thailand. Work on the diagnostic aspects of inborn errors of metabolism has stimulated local hospitals to establish their own facilities. Accordingly, our own research can focus more on the molecular aspects, and in this connection we have focused on two diseases. In the case of methylmalonic acidemia, mutations were found both in the methylmalonyl CoA mutase enzyme itself, and in the enzymes involved in metabolism of the cobalamin coenzyme. The other area of interest involves mucopolysaccharidosis, since this complements our work on the glycosidase enzymes. So far, we have studied cases of Hurler's syndrome, and found various mutations in the α -iduronidase gene. Cancer research utilises mainly the proteomic approach, which has been successfully used in detecting the proteins over-expressed in thyroid cancer. Proteomics is being also being used to study other cancers,

such as cholangiocarcinoma, as well as cultured cancer cell lines, to detect possible biomarkers. Other studies involve testing extracts and pure compounds from medicinal plants for anti-cancer properties by cytotoxicity tests and *in vitro* invasion studies.

The output from this grant has generally been able to match or exceed the targets in the initial proposal. Over the three year period of the grant, research yielded 22 publications in international journals, and two articles in international proceedings volumes, similar to the 23 publications in international journal projected in the initial proposal. Moreover, some 5 papers were in high impact journals with impact factor greater than 4.0. It is gratifying to see that the young scientists in the team can publish papers in good international journals, since this is essential in enabling them to establish their reputations. Many presentations were also made at meetings anticipated, with 27 abstracts in international meetings and 54 abstracts at national meetings over the three year period. Some of these were invited lectures by staff, but many were poster presentations by students, thus providing the opportunity for young researchers to gain experience.

Apart from the development of the young Ph.D researchers in the team, training of additional personnel with interest and expertise was an important element of the grant. Over the three year period of the grant, four Ph.D. students graduated and another 16 students are currently enrolled in the Ph.D. program. Nine M.Sc. students graduated, and another 15 students are currently enrolled in the M.Sc. program. In addition, 52 B.Sc. students performed their Senior Project research with staff in the grant. These new Ph.D.s will be valuable resources for expanding research activities in protein science and enzymology in Thailand, while the M.Sc. and B.Sc. graduates can provide a supportive role in this endeavour, and some may receive further training to Ph.D level in the future.

Various activities were undertaken in terms of academic services and promotion of protein science and molecular enzymology. In particular, Nobel Laureate William Lipscomb's extended visit as Stang Mongkolsuk Distinguished Professor had broad impact in stimulating research at CPSF, in promoting protein research in the academic community, and encouraging the government and the general public to have a greater appreciation of the importance of science. An informal Protein Research Network was started to disseminate news and events to members by email. In addition, the research group organized various symposia, meetings and seminars, most notably a two-day Protein Research Network Symposium, attended by over 200 researchers from some 23 universities and institutions throughout Thailand.

Finally, the activities of our research group and the Center for Protein Structure and Function have been well recognised by the local scientific community through several awards from various agencies. Thus Professor Jisnuson Svasti received the Outstanding Researcher of Thailand Award from the Foundation for the Promotion of Science and Technology under the Royal Patronage of H.M. The King and the

Outstanding National Researcher Award, Chemical Sciences and Pharmacy section, National Research Council of Thailand. He was also named Outstanding Lecturer, Faculty Club, Faculty of Science, Mahidol University, and Exemplary Lecturer, Faculty Club, Mahidol University. Dr. Jirundon Yuvaniyama and Dr. Palangpon Kongsaree were awarded the Young Scientist of Thailand Award, Foundation for the Promotion of Science and Technology under Royal Patronage of H.M. The King. Then Dr. Pimchai Chaiyen was awarded UNESCO-L'Oreal Women in Science Fellowship and Dr. Apinya Buranaprapuk was awarded an Outstanding Thesis Award by the National Research Council of Thailand, Chemical Science and Pharmacy Section. These various awards are an assurance of the quality of work being carried out by our research group.

II. RESEARCH PROGRESS

1. Glycosidase Enzymes

Glycosidase enzymes are enzymes that hydrolyze the glycosidic bond between a sugar and an aglycone or another sugar, and have potential applications for oligosaccharide and glycoside synthesis. Over the years, we discovered several glycosidases from Thai plants by both enzymatic screening and DNA-based screening of plants, as well as detected glycosides in many species of Thai plants. We have already studied some of these glycosidases in more detail, in particular the β -glucosidases, in terms of purification, properties, cloning and sequencing, as well as use for glycoside synthesis. These studies are being continued in the laboratories of several members of our group, as outlined below. Progress in the study of glycosidases includes:

1.1 Screening for Novel Glycosidases

Work on screening for novel glycosidase enzymes in Thai plants was carried out in Dr. Sujint Anguravirutt's laboratory at Mahasarakham University. Six glycosidases: N-acetyl- β -D-glucosaminidase, α -D-mannosidase, α -D-galactosidase, β -D-galactosidase, α -D-glucosidase and β -D-glucosidase were screened from plants in Northeast Thailand. The results showed that among 41 types of seeds and 17 types of leaves studied significant amounts ($\geq 0.1 \mu\text{mole}/\text{min}/\text{g}$ sample) of the enzymes were found in 9 seeds namely: α -D-mannosidase ($0.23 \mu\text{mole}/\text{min}/\text{g}$ seed) in *Cleome viscosa* Linn. (เลื่อยนผี); α -D-mannosidase ($0.18 \mu\text{mole}/\text{min}/\text{g}$ seed) in *Ricinus communis* L. (ละหุ่ง); α -D-mannosidase ($0.25 \mu\text{mole}/\text{min}/\text{g}$ seed) and β -D-galactosidase ($0.11 \mu\text{mole}/\text{min}/\text{g}$ seed) in *Pueraria phaseoloides* (Roxb.) Benth. (ถั่วเลื่อยนป่า); α -D-mannosidase ($0.13 \mu\text{mole}/\text{min}/\text{g}$ seed) and α -D-galactosidase ($0.15 \mu\text{mole}/\text{min}/\text{g}$ seed) in *Bauhinia purpurea* Linn. (ชงโค); α -D-mannosidase ($0.29 \mu\text{mole}/\text{min}/\text{g}$ seed), α -D-galactosidase ($0.12 \mu\text{mole}/\text{min}/\text{g}$ seed) and N-acetyl- β -D-glucosaminidase ($0.16 \mu\text{mole}/\text{min}/\text{g}$ seed) in *Samanea saman* (Jaeq.) Merr. (จำปา); α -D-mannosidase ($0.44 \mu\text{mole}/\text{min}/\text{g}$ seed), α -D-galactosidase ($0.26 \mu\text{mole}/\text{min}/\text{g}$ seed) and β -D-galactosidase ($0.11 \mu\text{mole}/\text{min}/\text{g}$ seed) in *Moringa oleifera* Lamk. (มะรุม); α -D-mannosidase ($0.76 \mu\text{mole}/\text{min}/\text{g}$ seed), β -D-galactosidase ($0.11 \mu\text{mole}/\text{min}/\text{g}$ seed) and N-acetyl- β -D-glucosaminidase ($0.24 \mu\text{mole}/\text{min}/\text{g}$ seed) in *Ceiba pentandra* Gaertn. (จิง, ทุเรียน); α -D-mannosidase ($0.18 \mu\text{mole}/\text{min}/\text{g}$ seed), β -D-galactosidase ($0.22 \mu\text{mole}/\text{min}/\text{g}$ seed) and N-acetyl- β -D-glucosaminidase ($0.10 \mu\text{mole}/\text{min}/\text{g}$ seed) in *Jatropha curcas* Linn. (สบู่ดำ); α -D-mannosidase ($0.52 \mu\text{mole}/\text{min}/\text{g}$ seed), α -D-galactosidase ($1.09 \mu\text{mole}/\text{min}/\text{g}$ seed) and N-acetyl- β -D-glucosaminidase ($0.90 \mu\text{mole}/\text{min}/\text{g}$ seed) in *Albizia lebeckoides* (DC.) Benth. (คาง). The results indicated that α -D-mannosidase was found in all of the above samples while the highest enzyme activity was α -D-galactosidase ($1.09 \mu\text{mole}/\text{min}/\text{g}$ seed) in *Albizia lebeckoides* (DC.) Benth. However, overall the levels of enzymes found in plant sources were not high, so additional screening studies were performed on soil microorganisms, since soil in Mahasarakham Province has high salt content, and might contain unique

glycohydrolases. Preliminary experiments were performed on screening three glycohydrolases namely β -galactosidase, β -glucuronidase and β -glucosidase, from local soil bacteria, but no enzyme activity was detected.

Dr. Patjraporn Wongvithoonyaporn has screened for α -mannosidase and naringinase activities from approximately 70 fungal hosts collected from China and various Thai provinces, e.g. Bangkok, Chiangrai, Nakhorn Pathom, Supanburi, Saraburi, Nakhorn Ratchaseema, Amnart Chareon, Sreesaket, etc. Approximately 215 pure fungi were first isolated from these hosts using modified Sabouraud's dextrose agar supplemented with chloramphenicol. These fungal isolates are being used to screen for α -mannosidase using synthetic minimal medium supplemented with potato dextrose broth and mannan at 28°C and assaying enzyme with p-nitrophenyl- α -D-mannopyranoside. In contrast, for naringinase, about 166 fungal isolates producing naringinase with optimum acidic pH were directly screened from these hosts by using synthetic minimal medium supplemented with naringin and chloramphenicol at 28°C. Secondary screening of naringinase from the selected fungi is also being undertaken.

1.2 Structure Function Relationships in Plant β -Glucosidases

Study of structure-function relationships in plant β -glucosidases is a major focus of our work. These enzymes may differ in substrate specificity for the hydrolysis, transglucosylation and reverse hydrolysis reactions, and our objective is to correlate these differences in functional properties with differences in their structure. Various enzymes are being studied, including Thai Rosewood dalcochinase and cassava linamarase, which hydrolyze isoflavonoid glucoside and cyanogenic glucoside respectively. Work on plant β -glucosidases is being done in the laboratories of Dr. Jisnuson Svasti at Mahidol University, Dr. James Ketudat-Cairns at Suranaree University of Technology, and Dr. Prachumporn Toonkool at Kasetsart University.

1.2.1 Transglucosylation of alcohols using plant β -glucosidases

We have compared the transglucosylation activity of Thai Rosewood dalcochinase and cassava linamarase, purified in our laboratory, with commercial almond glucosidase. Initially, short chain C_1 - C_4 alkyl alcohols were used using a single phase buffer system, and the results showed that almond glucosidase and dalcochinase, gave poor yields of alkyl glucoside with secondary butanol and no product with tertiary butanol. However, with cassava linamarase, good yields (approaching 100%) were obtained with both secondary and tertiary butanol.

Later, studies were performed with longer chain alcohols, using buffer-saturated C_5 - C_8 primary, secondary and tertiary alcohols. The results (Table 1) showed that, in general, the three enzymes gave good yields (> 60%) of alkyl glucosides with both linear and branched C_6 primary alcohols. However, cassava

linamarase again gave a higher yield with octyl glucoside (85%) than the other two enzymes (25%-31%). Most interestingly, cassava linamarase also gave much better yields (approaching 100 %) with secondary alcohols than almond β -glucosidase (30-40%) and Thai rosewood β -glucosidase (<20%). In addition, cassava linamarase was the only enzyme capable of forming alkyl glucosides with C₅ and C₆ tertiary alcohols. Yields initially were rather low, but this could be optimized by reducing the concentration of enzyme and varying the time (Figure 1). This time course shows that with increasing amount of enzyme, there is more rapid and more complete cleavage of pNP-glucose. At the same time, the C₅-tertiary alkyl glucoside is also being formed with a rate that depends on the amount of enzyme. The alkyl glucoside, so formed, also becomes digested at high levels of enzyme. So the amount of enzyme used and the time of reaction have to be selected to optimise yields of tertiary alkyl glucoside. As a result, the maximum alkyl glucoside yields of 82% and 56% could be obtained with 2-methyl-butan-2-ol and 2-methyl-pentan-2-ol respectively. A 1% yield of glucoside was found with 2-methyl-butan-2-ol using *Pyrococcus furiosus* β -glucosidase, but there have been no previous reports of enzymatic glucosylation of 2-methyl-pentan-2-ol.

This high yield of alkyl glucoside with tertiary alcohols is exceptional for enzymatic transglucosylation. The enzymatic glucosylation of 2-methyl-pentan-2-ol has never been reported, and the only previous report with 2-methyl-butan-2-ol was the 1% glucoside yield found using *Pyrococcus furiosus* β -glucosidase. We have also started to explore the potential applications of this unique capability of cassava β -glucosidase in transglucosylating secondary and tertiary alcohols in the synthesis of chiral and other alkyl glucosides. So far, experiments performed in collaboration with Dr. Palangpon Kongsaree, indicate that alkyl glucosides can be obtained with butan-2-ol and 3-methyl-2-butanol, with structures being characterised by ¹H-NMR, ¹³C-NMR, and mass spectrometry. In both cases, product yields approached 80%, and 52% enantiomeric excess was obtained. In addition, some natural aglycones were also tested as acceptors, including linalool and quercetin, which gave glycoside yields of about 50%.

The possibility of using other glycosyl donors apart from Glc β -O-*p*NP was also explored using hexan-1-ol as glycosyl acceptor. The results showed that all both enzymes could also use Fuc β -O-*p*NP as donor for synthesis of hexyl glycosides. Thai rosewood β -glucosidase could only use gentiobiose but not cellobiose as donor, but cassava β -glucosidase could use neither. In terms of reverse hydrolysis using monosaccharide donors, Thai rosewood β -glucosidase could use glucose, but not fucose, but cassava linamarase could use neither free glucose nor free fucose as glycosyl donor for hexyl glycoside synthesis. The unique capability of cassava in transglucosylation, especially with secondary and tertiary alcohols will be further studied in terms of its potential applications in the synthesis of chiral and other alkyl glucosides.

Table 1: Transglucosylation of Longer Chain Alcohols by Cassava, Thai Rosewood, and Almond Glucosidases^a

Alcohol Acceptor	Mole % Alkyl Glucoside Obtained ^b		
	Cassava β -glucosidase	Thai rosewood β -glucosidase	Almond β -glucosidase
<i>Normal Primary alcohols</i>			
Hexan-1-ol	80	74	77
Octan-1-ol	85	25	31
<i>Branched Primary alcohols</i>			
2-Methyl-pentan-1-ol	100	100	74
3-Methyl-pentan-1-ol	81	60	75
4-Methyl-pentan-1-ol	100	74	74
<i>Secondary alcohols</i>			
Hexan-2-ol	100	3	37
3-Methyl-pentan-2-ol	100	2	39
4-Methyl-pentan-2-ol	100	16	30
<i>Tertiary alcohols</i>			
2-Methyl-butan-2-ol	82 ^c	0	0
2-Methyl-pentan-2-ol	56 ^d	0	0

^aEnzyme (17 nkat/ml) was incubated in buffer-saturated alcohol containing 15 mM Glc β -O-*p*NP for 3 days at 40°C; ^bMole % alkyl glucoside represents the mole percent of the alkyl glucoside spot relative to the total moles present in the Glc β -O-*p*NP, alkyl glucoside, and glucose spots: each value is an average from duplicate TLC plates; ^cMole % alkyl glucoside obtained from incubating with 8.3 nkat/ml enzyme for 4 days; ^dMole % alkyl glucoside obtained with 17 nkat/ml enzyme for 1 day.

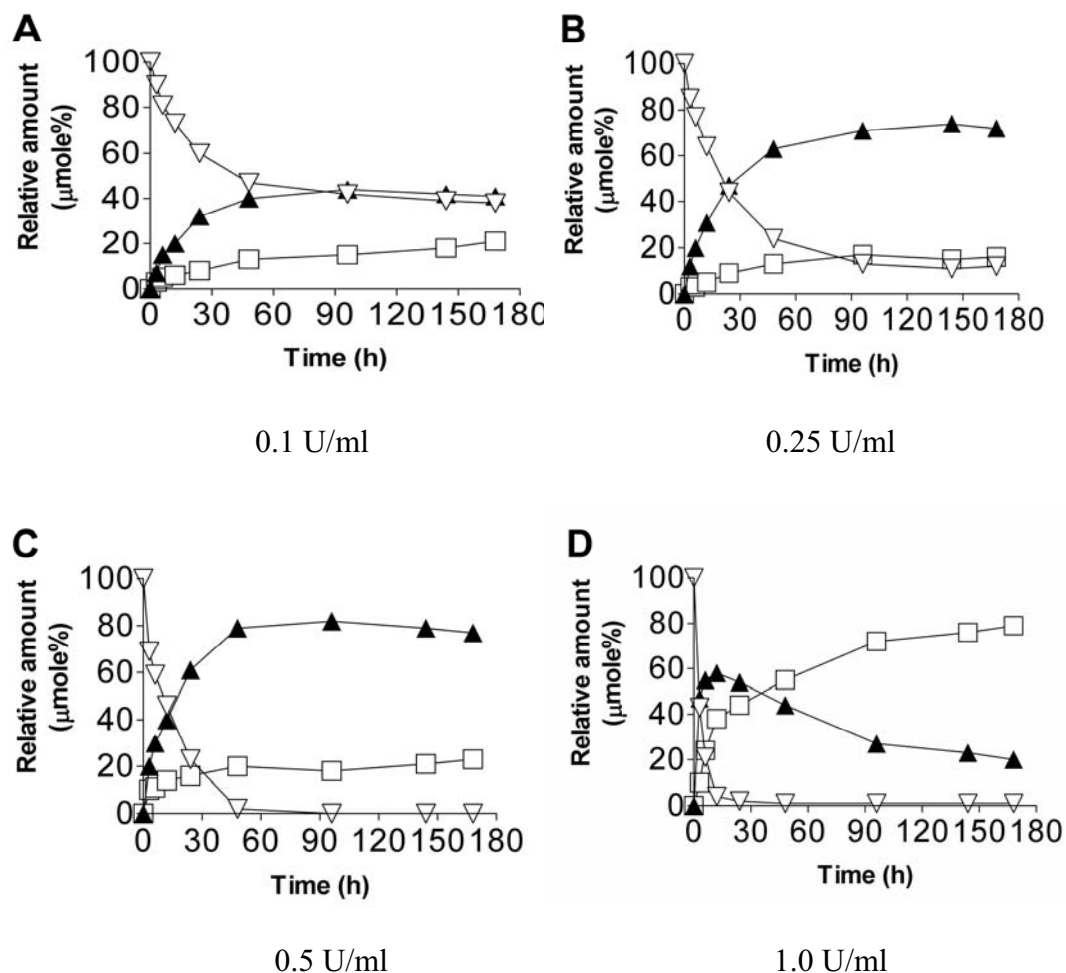


Figure 1: Transglucosylation of 2-methyl-butan-2-ol by cassava β -glucosidase. Various concentrations of enzyme were incubated with 15 mM pNP- β -Glc in 2-methyl-butan-2-ol, saturated with 0.1 M McIlvaine buffer, pH 5.5 at 40°C for various times (h). A: 1.7 nkat/ml enzyme; B: 4.2 nkat/ml enzyme; C: 8.3 nkat/ml enzyme; D: 17 nkat/ml enzyme. Products analyzed by TLC and quantitated by scanning. Relative amount (mole %) represents the moles of pNP- β -Glc (∇), alkyl glucoside ($\blacktriangle$) or free glucose ($\square$) as a percent of the total moles of glucosyl groups.

1.2.2 Kinetic Studies of Transglucosylation using Thai Plant β -Glucosidases

To study transglucosylation specificity of Thai plant β -glucosidases, enzymes were inactivated by 2FDNPG (2, 4-dinitrophenyl-2-fluoro-2-deoxyglucose), and various alcohols were used as acceptors for reactivation (Figure 2). Three enzymes were selected for study, cassava linamarase, Thai rosewood dalcochinase, and rice β -glucosidases BGlu1, since these three enzymes have different substrate specificity in hydrolysis, and two have them have also been shown to differ in transglucosylation specificity. The results so far show that the inactivation of all three enzymes by 2FDNPG resulted in time-dependent inactivation with pseudo-first-order kinetics. All three inactivated-enzymes can also be reactivated by addition of various acceptors. The reactivation rate was found to be dependent upon the concentration of alcohols in a saturable fashion.

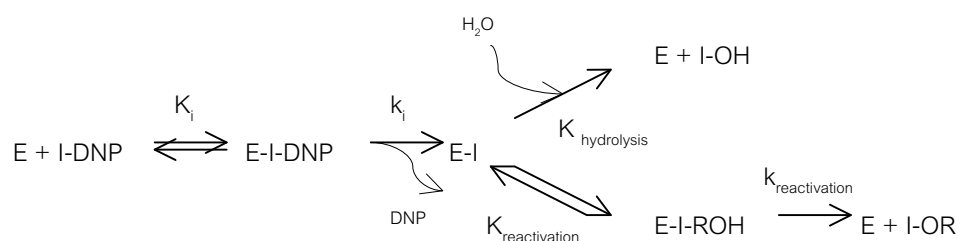


Figure 2: Scheme for reactivation study of β -glucosidases. E, free enzyme, I-DNP, 2FDNPGlc; DNP, 2,4-dinitrophenolate; E-I, glycosyl-enzyme intermediate; I-OH, 2FGlc; ROH, acceptor.

Inactivated cassava β -glucosidase can generally be reactivated by all alcohol acceptors with highest reactivation rate compared to the rate of reactivation seen with buffer alone (control). Thus, for example, transglucosylation of inactivated cassava β -glucosidases by n-butanol occurs approximately 100 fold faster than simple hydrolysis, while Thai rosewood dalcochinase occurred only 3-5 fold faster (Table 2). This indicates that cassava linamarase, in the presence of alcohols, has an unusually large preference for transglucosylation compared to hydrolysis. This is consistent with the high yields of alkyl glucosides obtained with cassava linamarase, when the enzyme is incubated with pNP-glucose and various alcohols (Results Section 1.2.1). In addition, cassava linamarase can use tert-butyl alcohol as acceptor in reactivation, as found in the synthesis reaction, but the relative rate of reactivation is much lower than that of n-butanol or secondary butanol. In contrast, reactivation of rice cassava β -glucosidases by alcohols occurs very slowly.

Table 2: Rate of Reactivation (Kobs) of Inactivated Enzyme by Various Alcohol Acceptors, with Cassava Linamarase and Thai Rosewood Dalcochinase.

0.5M Acceptors	Real Kobs (min ⁻¹) (X10 ⁻⁴)	
	Thai Rosewood	Cassava
Control	2.36	5.95
Methanol	2.69	61.7
Ethanol	3.57	137.3
n-Propanol	6.82	599.8
n-Butanol	15.0	444.1
Iso-Butanol	8.67	957.9
2-Butanol	3.70	243.5
3-Butanol	3.07	46.8
Glucose	2.82	6.99
Fucose	3.64	8.56

1.2.3 β -Glucosidase Enzyme from *Solanum torvum*

Screening studies using Thai Rosewood dalcochinase to digest ethanol extracts of more than 200 Thai plants, followed by t.l.c. analysis revealed the presence of several glycosides. A search for glucosidase enzymes in these plants has also been made by using the ethanol extract as the substrate. From these studies, *Solanum torvum* (makua puang) was chosen for further study. However, study of this system was difficult since both enzyme and substrate are unknown. Therefore, it was decided to purify the natural substrate first, for later use in enzyme isolation.

So, in collaboration with Dr. Prasart Kittakoop, air-dried fruits of *S. torvum* Sw. (2.5 kg wet weight) were sequentially extracted with MeOH, hexane and ethyl acetate, followed by LH-20 column chromatography (MeOH as eluent), then purified with MPLC on reversed phase C₁₈ column to yield pure compound torvoside A, and further fractionated by silica gel column chromatography to yield pure compound Torvoside H. Analyses of ¹H and ¹³C NMR spectral data reveal that compound torvoside A (**1**) was a known steroidal glycoside. But pure compound torvoside H (**2**) was a novel compound, derived from torvoside A. Both compounds were hydrolyzed by Thai rosewood dalcochinase, by extracts of leaves and stem from solanum, and also by commercial β -glucosidase from almond (Figure 3). The purification and characterization of β -glucosidase from *S. torvum* Sw. will be further studied.

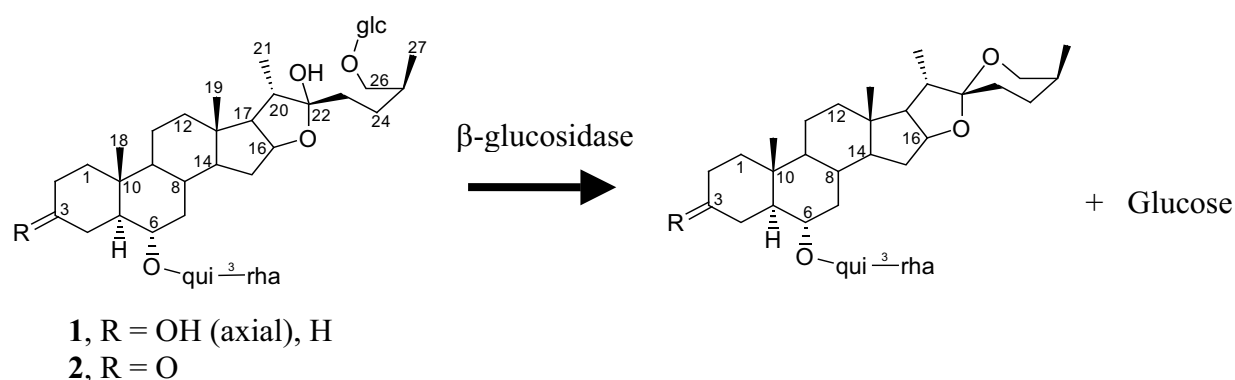


Figure 3: Natural Substrates of *Solanum torvum* and Their Hydrolysis by β -Glucosidase

We have shown that the fruits of *Solanum torvum* (makua puang) contained two steroidal glycosides, a known compound torvoside A and a novel compound torvoside H. Last year, *Solanum torvum* β -glucosidase enzyme (torvosidase) was

purified from young leaves in Professor Jisnuson Svasti's laboratory, using four steps of chromatography, including Butyl-Toyopearl, Con A-Sepharose, Sephacryl S-300 and Butyl-Sepharose chromatography. The purified enzyme showed protein band coincident with one activity band with 4-MU-glucoside. Native enzyme had MW 86 kD by gel filtration, similar to the 80 kD found on SDS-PAGE, and had pI of 9.3. Torvosidase had lower K_m values for the natural substrates torvoside A (63 μ M) and torvoside H (68 μ M) than for synthetic substrates pNP- β -glucoside (1.0 mM) and 4-MU- β -glucoside (0.74 mM). It showed little hydrolysis of dalcochin- β -glucoside, but could hydrolyze C₅-C₁₀ alkyl- β -glucosides (e.g. pentyl- β -glucoside, hexyl- β -glucoside, heptyl- β -glucoside, octyl- β -glucoside, decyl- β -glucoside).

Preliminary studies of the internal amino acid sequence of torvosidase have also been performed by trypsin digestion and sequence analysis by LC-MS/MS. Although, LC-MS/MS cannot separate leucine and isoleucine because of their identical mass, alignments of the sequence were made with the data available in the protein and nucleic acid database (<http://www.ncbi.nlm.nih.gov/BLAST/>). Seven peptides covering some 60 residues in *Solanum* β -glucosidase showed sequence similarity with family 3 β -glucosidase from plants, including β -D-glucan exohydrolases from *Hordeum vulgare* (barley), *Zea mays* (maize), *Arabidopsis thaliana*, *Nicotina tabacum* (tobacco) and *Oryza sativa* (rice). Further studies will be required to confirm this result, but from the data available, *Solanum* β -glucosidase, which shows high specificity for furostanol glycosides, should be classified in the family 3 of the glycosyl hydrolases according to the Henrissat classification. To our knowledge, this is the first report that a plant family 3 β -glucosidase specifically hydrolyzes the furostanol glycosides.

1.2.4 Other novel glucosides and glucosidases from Thai plants

In previous years, screening of Thai plant ethanolic extracts by digestion with Thai Rosewood dalcochinase showed other interesting glucosides in other plants, e.g. *Plumeria spp.* and *Nerium indicum*. These natural glucosides could be then be used to screen for novel glucosidase enzymes. Last year, a glucoside substrate has been purified from the flowers of *Plumeria obtusa*, and the structure of the compound was established by NMR and mass spectroscopic data. The ¹H and ¹³C NMR showed signals corresponding to a plumieride coumarate glucoside (Figure 4), previously found in the bark of *Plumeria rubra* Linn. This data shows two glucosyl groups attached at C-1 and C-23.

In addition, β -glucosidase was purified from the flowers of *Plumeria obtusa* Linn by DEAE-Cellulose, Con A-Sepharose, Sephacryl S-300 and Butyl-Sepharose chromatography. The cumulative yield and fold purification of the β -glucosidase were 2.3 % and 143 fold, respectively. The enzyme has a molecular

weight of 180,000 in the native state, and shows 2 bands of 59,000 Da and 67,000 Da on denaturing SDS-polyacrylamide gel electrophoresis. One fluorescent band of β -glucosidase was found with pI about 5.0 by agarose gel electrophoresis. The optimum pH of the purified β -glucosidase detected by pNP-Glc and its natural substrate was 5.5. The K_m values for pNP-Glc and *Plumeria* β -glucoside were 2.9 mM and 0.33 mM, respectively. The enzyme had hydrolytic activity towards pNP- β -Glc, pNP- β -Fuc, pNP- β -Gal and pNP- β -Man, and had lower activity for other pNP-glycosides, dalcochinin- β -glucoside (the natural substrate of Thai Rosewood β -glucosidase) and esculin, The enzyme could not hydrolyze cyanogenic glucosides (prunasin, linamarin and amygdalin), the natural substrate of *Solanum torvum* (torvoside A), disaccharides (gentiobiose), aromatic glucosides (arbutin and salicin) or alkyl glucosides (methyl-glucosides and hexyl-glucosides).

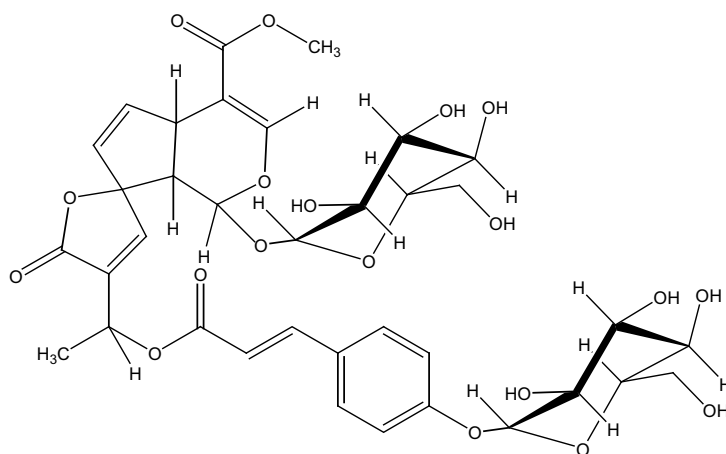


Figure 4: Natural glucoside substrate found in *Plumeria obtusa*

Since cassava linamarase showed exceptional ability to transglucosylate alkyl alcohols, we have also begun to search for linamarases from other plants. So far, we have found high activities of β -glucosidase in rubber (*Hevea brasiliensis*) leaves, which are known to have linamarase enzyme. It will be interesting to see whether rubber linamarase also shows high activity in transglucosylation, as well to test other possible sources of linamarase, such as linseed.

1.2.5 Expression and Purification of *Dalbergia* β -glucosidases

Work on the *Dalbergia* β -glucosidases at Dr. James Ketudat-Cairns' laboratory focused on cloning and purification of the *D. nigrescens* enzyme, characterization of its substrate, and recombinant production of both *D. nigrescens* and *D. cochinchinensis* enzymes. Additional work on expression and purification of

recombinant *D. cochinchinensis* enzyme is carried in the laboratory of Dr. Prachumporn Toonkool. Furthermore, work on the cloning and expression of cassava linamarase is also being started in Dr. Prachumporn Toonkool's laboratory, since this enzyme has exceptional capability in transglucosylation of secondary and tertiary alcohols.

In the first year of the project worked on developing better methods to purify *D. nigrescens* β -glucosidase, which tended to be contaminated with black phenolic compounds. By the second year, we purified β -glucosidase from *D. nigrescens* seeds with an improved procedure and characterized its activity and some peptide sequences. This purification procedure involved incorporating reducing agents and EDTA in the initial extraction buffer and extracting phenolic compounds with 70% ammonium sulfate before extracting the protein and purifying it by ammonium sulfate precipitation, DEAE chromatography, and gel-filtration chromatography. The resulting enzyme was about 90% pure. Greater purity could be achieved by passing this enzyme over a QAE column on an FPLC. The purified enzyme had a pH optimum of 5-6 and a temperature optimum (in a 10 min pNP- β -D-glucoside hydrolysis assay) of 60-70°C. Enzyme assays of this protein suggested that the K_m of the protein for pNP-glucoside and pNP-fucoside are both approximately 3-fold higher than with *D. cochinchinensis* dalcochinase. During this period, we sequenced the N-terminus of the protein and several peptides from a tryptic digest of the protein.

The purified enzyme was used to identify natural substrates in extracts of *D. nigrescens* seeds. By comparing TLC profiles of extracts added to assays with and without enzymes, several fluorescent spots were identified that appeared to be substrates for the *D. nigrescens* β -glucosidase. One of these substrates was not hydrolyzed by the *D. cochinchinensis* β -glucosidase (dalcochinase) and the other was hydrolyzed at a much lower rate by this enzyme than the *D. nigrescens* β -glucosidase, showing that, despite greater than 80% sequence identity between the two enzymes, their substrate-specificity appears to be different. Two *D. nigrescens* substrates have been purified and studied by NMR and mass-spectrometric analysis. They were found to be isoflavonoid glycosides containing two sugar residues, a hexose and a pentose, both of which are removed from the aglycone by the enzyme. The sugars appeared to be β -D-apiose and β -D-glucose from the NMR data. Bioactivity testing by the BioService Unit of the National Center for Genetic Engineering and Biotechnology (BIOTEC) indicated that substrates and their aglycones had little activity as antibacterials and antivirals, as well as no detectable antifungal or mammalian cell cytotoxic activity.

During the grant period, we worked to express β -glucosidase from a *D. nigrescens* cDNA cloned just before the grant started, which had nearly 90% identity with *D. cochinchinensis* dalcochinase. After several attempts to express the

protein in *Pichia pastoris* and *E. coli*, no significant β -glucosidase activity could be expressed and purified. The apparent mutation of a conserved Gly to Asp was corrected by site-directed mutagenesis, but still no activity was observed. In an attempt to verify the sequence of this clone, a short segment of a second β -glucosidase homologue cDNA from germinating *D. nigrescens* seeds was amplified and extended it by 3' and 5' RACE to attain the full length coding sequence. Now, we have cloned the rest of the sequence and found it to have 82% identity to *D. cochinchinensis* and 83% identity to *D. nigrescens* isozyme 1 β -glucosidases. We have cloned it into pPICZ α and shown it to have activity similar to the enzyme purified from the plant, i.e. it hydrolyzed pNP-glucoside and pNP-fucoside at similar levels and could hydrolyze the substrates isolated from the seeds. However, several tryptic peptides from the plant enzyme had sequences matching both isozymes of *D. nigrescens* β -glucosidase (Table 3). Then, both enzymes were cloned into the new *Pichia pastoris* expression vector described below, in the hope that they could be produced in an easily purifiable form. However, this has not worked for purification yet, and only the second isozyme was found to be active.

D. nigrescens β -glucosidase isozyme 1 was also produced in *E. coli*, but it was largely inactive and insoluble. The protein was gel purified on SDS-PAGE, and was used to produce antibodies in rabbits to help with further expression experiments in yeast.

Table 3: Comparison of two *D. nigrescens* cDNA with Edman peptide sequences of purified *D. nigrescens* β -glucosidase

Peptide	Sequence source	Sequence
N-terminus	DnBG gene 1 DnBG gene 2 Peptide	ATITEVPPF ATITEVPPF ATITEVPPF
Tryp 1	DnBG gene 1 DnBG gene 2 Peptide	YMNLDAYR YMNLDAYR YMNLDAYR
Tryp 2	DnBG gene 1 DnBG gene 2 Peptide	ASGGINSTGVDYYNR ASGGINSTGVDYYNR <i>likely</i> ASGGI <u>I</u> STGVD <i>glycosylation site</i>
Tryp 3	DnBG gene 1 DnBG gene 2 Peptide	LINELLANDITP... LINETLHNGITP... LINET L ANGI H
Tryp 4	DnBG gene 1 DnBG gene 2 Peptide	HWITVNEPSIFTMNGYAYGIFAPGR HWITINEPQVFTTNGYTYGMFAPGR HWITVNEP S I F T M NGY A Y G I F APGR Q M

Previously, a yeast expression system was tried for *Dalbergia* enzymes. Our collaborators in Dr. Mariena Ketudat-Cairns' group have developed a new fusion protein expression system for *D. cochinchinensis* β -glucosidase in *P. pastoris*. This system (pPICZ α), which incorporates a myc-epitope/His₆ tag at the C-terminus, was shown to produce active β -glucosidase in *P. pastoris*. However, the protein could not be purified by immobilized metal affinity chromatography. We have removed a protease cleavage site from the C-terminus of the protein to aid in purification from the C-terminal tag, however attempts by Dr. Mariena's, Dr. James' and Dr. Prachumporn's groups to purify the protein by immobilized metal affinity chromatography (IMAC) still failed to produce a significant amount of protein. Further attempts will be made to remove the last 36 amino acids from the C-terminus of the *D. cochinchinensis* dalcocinase, since these are not seen in the crystal structure of cyanogenic β -glucosidase from white clover (1CBG), but no active enzyme was produced from this construct. After consultation with Prof. Asim Esen of Virginia Polytechnic Institute and State University, U.S.A., we postulate that the yeast cell naturally cleaves off the C-terminus of the enzyme, so our enzyme has been made without the C-terminal His-tag for IMAC purification. A new construct of N-terminal His-tagged enzyme has been made, with active enzyme produced in *P. pastoris* cultures. In addition, we have generated a new plasmid, pPICZ α NH, which contains an N-terminal His₆ tag before the cloning site for the protein cDNA. This will allow *Dalbergia* and other enzymes to be expressed with the N-terminal His₆-tag for purification with the same restriction sites previously used for pPICZ α . Unfortunately, we are still unable to detect any significant binding between the new N-terminal His-tagged enzyme and the immobilized metal column (Ni²⁺-bound resin) for any of the *Dalbergia* β -glucosidases.

While purification via IMAC has not been successful, the laboratory of Prof. Svasti has used classical purification procedures to purify the recombinant *D. cochinchinensis* β -glucosidase (constructed with C-terminal His-tag) from *P. pastoris* culture media (grown under fermentation conditions). We are planning to perform amino acid sequencing to verify the N- and C-terminal sequences of this enzyme. The sequencing results will indicate the extent of proteolysis and suggest a suitable site for introducing the His-tag onto the protein.

During the past period, the *D. cochinchinensis* β -glucosidase was also cloned into a variety of expression vectors, namely pET32a, pGEX-4T-1 and pET15b, for expression by the bacterial systems. The pET32a plasmid was chosen as it was previously successful for rice β -glucosidase. Small amounts of soluble enzyme were expressed and purified from Origami *E. coli* culture, but the activity was very low. When the construct was cloned into the pGEX-4T-1 for expression as a GST-fusion protein in *E. coli* BL21(DE3), this has yielded high levels of β -glucosidase expression, but most proteins were insoluble. While it is

possible to solubilise protein with guanidine-hydrochloride, the refolding procedure has proved to be very difficult and an active enzyme has yet to be obtained. The enzyme was also cloned into the pET15b plasmid as an N-terminal His-tagged protein. Soluble enzyme can be obtained in substantial amounts when it is co-expressed with a chaperonin operon in *E. coli* BL21(DE3). The enzyme can be purified via IMAC, but it does not exhibit any activity. We are currently cloning the construct in pET15b into Origami *E. coli* that harbour a chaperonin operon, which may together help in the production of soluble active enzyme.

In addition to these microbial constructs. Two constructs have been made for expression of Thai rosewood β -glucosidase with an N-terminal His-tag in plant cells. The constructs include one with a His₆-tag after the signal sequence and before the mature protein, and one with a thrombin site for removal of the tag. The constructs were constructed in pCambia1301. In the future, they can be put into *Agrobacterium* and transferred into plant cells (eg. tobacco suspension cells) to see if the protein can be expressed and purified successfully from the plant.

Several attempts have also been made to crystallize the native enzyme purified from Thai rosewood seeds. So far, results have not been promising, and this may be due to the inherent heterogeneity of the natural enzyme, for example due to glycosylation. Once the enzyme can be expressed in microorganisms, more homogeneous enzyme may be obtained for crystallisation.

1.2.5 Rice β -glycosidases

Initially, two rice β -glucosidase cDNAs, *bglu1* and *bglu2* were cloned and their proteins, BGlu1 and BGlu2, were expressed in *E. coli* and characterized (Opassiri *et al.*, 2003). The activity of BGlu1, expressed in *E. coli* was high enough to allow purification by IMAC using the N-terminal thioredoxin-His-tag in the fusion produced from the pET32a vector. Levels of BGlu2 activity were too low for efficient purification and characterization, though hydrolysis of pNP- β -D-glucoside and pNP- β -D-fucoside were significantly above background. Comparison of BGlu1 activity against various synthetic glycosides was done to determine the glycone specificity as shown in Table 4. To clarify this further, kinetic parameters were determined for some of the glycosides for which hydrolysis was seen, as shown in Table 5.

As shown, Bglu1 has been estimated to have a K_m of approx. 0.23 mM and k_{cat} of 4.1 s⁻¹ for both pNP- β -fucoside and pNP- β -glucoside. It also hydrolyzes pNP- β -D-galactoside, β -D-mannoside, and pNP- α -L-arabinoside, but neither pNP- β -L-glucoside and -fucoside, pNP- α -D-glucoside and -fucoside, nor

pNP- β -D-arabinoside and pNP- β -D-thioglucoside. Of the natural substrates tested, Bglu1 seemed to hydrolyze 1,3- and 1,4- β -linked gluco-oligosaccharides best with highest activity toward laminaribiose and cellohexaose, as shown in Table 6. These studies indicated that these oligosaccharide substrates caused substrate inhibition and also served as substrates for transglycosylation, so low substrate concentrations, were used to determine kinetic constants for these substrates (Table 6).

Table 4: Substrate specificity of the purified rice BGlu1

<u>Substrate</u>	<u>Relative activity^a (%)</u>
pNP- β -D-glucopyranoside	100
pNP- β -D-fucopyranoside	100
pNP- β -D-galactopyranoside	15
pNP- β -D-mannoside	5.6
pNP- β -D-cellobioside	12
pNP- α -D-glucopyranoside	0
pNP- β -L-fucopyranoside	0
pNP- β -D-thioglucopyranoside	0
pNP- β -D-thiofucopyranoside	0
oNP- β -D-glucopyranoside	97
oNP- β -D-fucopyranoside	16
Methyl-umberriferyl β -D-glucoside	33
Methyl- β -D-glucopyranoside	0.1
n-Heptyl- β -D-glucopyranoside	3.8
n-Octyl- β -D-glucopyranoside	3.6
Phenyl- β -D-glucoside	0
Amygdalin	0.4
Prunasin	19
Dhurrin	0.3
DIMBOA-glucoside	0
Linamarin	0
Pyridoxin-5'-O- β -D-glucoside	+ (by TLC)

Table 5: Michaelis-Menten constants for hydrolysis of glycosides by rice BGlu1

Substrate	K_m (mM)	k_{cat} (s⁻¹)
pNP-β-D-glucopyranoside	0.23±0.01	4.07±0.04
pNP-β-D-fucopyranoside	0.23±0.01	4.08±0.04
pNP-β-D-galactopyranoside	3.03±0.22	2.10±0.13
pNP-β-D-mannoside	1.78±0.20	0.51±0.03
pNP-β-D-cellobioside	0.77±0.02	0.76±0.01
oNP-β-D-glucopyranoside	0.37±0.01	4.55±0.03
oNP-β-D-fucopyranoside	1.66±0.10	6.94±0.34

Table 6: Michaelis-Menten constants for hydrolysis of oligosaccharides by rice BGlu1

Substrate	K_m (mM)	k_{cat} (s⁻¹)	k_{cat}/K_m (s⁻¹ mM⁻¹)
Cello-oligosaccharides (DP)			
2	31.5 ± 1.6	1.52 ± 0.13	0.05 ± 0.002
3	0.72 ± 0.02	18.13 ± 0.35	25.36 ± 0.37
4	0.28 ± 0.01	17.34 ± 0.63	61.06 ± 0.37
5	0.24 ± 0.01	16.90 ± 0.06	71.5 ± 2.2
6	0.11 ± 0.01	16.93 ± 0.32	152.9 ± 0.5
Laminari-oligosaccharides (DP)			
2	2.05 ± 0.1	31.9 ± 3.1	15.7 ± 1.9
3	1.92 ± 0.04	21.2 ± 0.17	11.04 ± 0.16
4	N. D. ^a	N. D. ^a	N. D. ^a
5	N. D. ^a	N. D. ^a	N. D. ^a
Sophorose	13.89 ± 0.92	5.87 ± 0.21	0.42 ± 0.02
Gentiobiose	38.3 ± 4.1	0.99 ± 0.08	0.03 ± 0.003

The BGlu1 kinetic data was used to calculate the subsite binding affinities of the enzyme. As shown in Figure 5 below, the enzyme has significant differences with barley β -glucosidase, despite having high amino acid sequence identity (68%) and similar substrates.

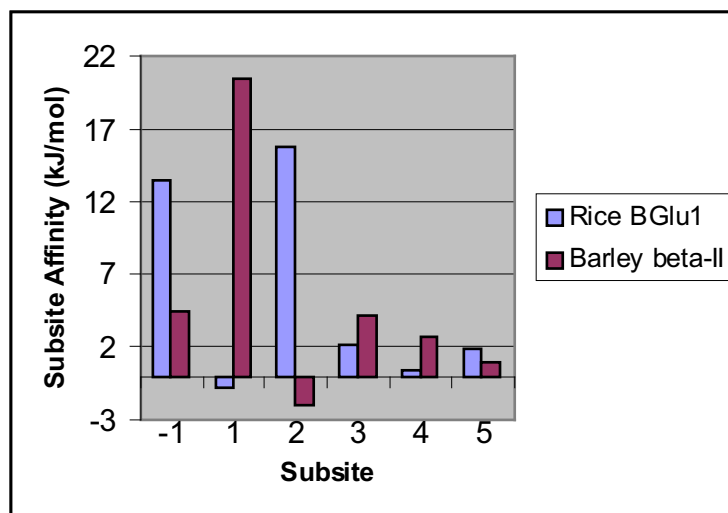


Figure 5: Subsite affinities of rice BGlu1 and barley β II β -glucosidase for cello- oligosaccharides

An HPLC assay was developed to measure pyridoxine-glucoside hydrolysis. It was found to be the best natural glycoside substrate hydrolysed with a k_{cat}/K_m ratio of $2.3 \text{ s}^{-1}\text{M}^{-1}$, although this was much worse than that for the cello-oligosaccharides DP 3-6 and laminari-oligosaccharides DP 2-3. We further showed that the enzyme can transglycosylate pyridoxine using pNP- β -D-glucoside as a donor to specifically produce pyridoxine-5'O- β -D-glucoside.

In addition to this functional characterization, we also set-up crystallization trials on rice BGlu1 thioredoxin- β -glucosidase fusion protein to attempt structural characterization. The microbatch crystallization trials resulted in small hexagonal crystals under some conditions, but larger crystals could not be obtained. The tag was removed with enterokinase and the protein was further purified by gel-filtration chromatography on a Sephadex S200 column. Peak fractions were used for crystallization trials and again resulted in small hexagonal crystals in three of the conditions tested, though these were different than the conditions that resulted in crystallization of the fusion protein. Again, the crystals were small, so further optimization will be needed to produce crystals big enough for data collection.

In addition to crystallization, structure-function aspects of BGlu1 are being studied by site-directed mutagenesis. We have mutated the catalytic acid/base and catalytic nucleophile glutamates to aspartate and glutamine and are currently sequencing the mutant cDNA before expressing the proteins. In addition, we have collaborated with Prof. Svasti's laboratory to mutate the catalytic nucleophile to alanine and serine, in an attempt to create a glycosyl synthase enzyme. These constructs still need to be confirmed and characterized.

It has been apparent that plants have multiple genes encoding β -glycosidase families. The opportunity to look at the plant wide role of these gene families has been opened by the sequencing of the rice genome, including the partial draft from Monsanto, the >90% complete drafts by the Chinese led by the Beijing Genomics Institute (indica rice) and the Syngenta Corp. Torrey Mesa Research Institute (japonica rice) and the ongoing efforts for a complete draft by the Rice Genome Project. We analyzed the glycosyl hydrolase family 1 and family 35 genes in the rice genome, and found approximately 47 family 1 genes and 15 family 35 genes, of which at least 34 family 1 and 13 family 35 genes appear to be expressed in rice, based on the presence of cDNA sequences, expressed sequence tags (ESTs) or full-length cDNA sequences in the database. The Japanese rice cDNA sequencing project, which deposited 28,000 cDNA sequences in the public databases, was particularly helpful to this analysis, along with the many EST projects around the globe.

Of the 47 genes identified for family 1, seven have no introns and are not expressed, indicating they may not be functional rice genes. Six of the seven intron-less genes, however, are more closely related to bacterial β -glucosidases, indicating they likely came from endophyte DNA in the indica rice genomic database. The large number of rice genes will provide a useful resource for evaluating structure and function relationships.

We have compared 12 of the family 35 genes with genes in this family from arabidopsis and found that both these groups contain some genes with C-terminal putative carbohydrate binding domains and some without. The phylogenetic analysis indicates that this domain seems to have been acquired early in plant evolution, and subsequently lost from some genes. We have isolated cDNA for 2 family 35 genes, and several other family 1 genes, and have begun to complete sequencing of these clones to confirm gene predictions. We also plan to amplify full-length coding regions of cDNA for other rice β -glycosidases.

In order to characterize the activity of a representative number of glycosyl hydrolase family 1 and 35 enzymes, we have obtained full-length cDNAs from twenty-five of these enzymes by RT-PCR or by request from the Rice Genome Resource Center to put into expression vectors. We have developed a set of

Gateway expression vectors to allow rapid screening for the expression of the enzymes in different systems. These include Gateway versions of the pET32 expression vector used to express the BGlu1 and BGlu2 proteins, pET32/DEST, the pPICZ α NH vector for secreted expression in *Pichia pastoris*, with an N-terminal His-tag, the pBAD/DEST an *Ara* promoter N-terminal His-patch-thioredoxin fusion from Invitrogen, and pMalc/DEST, pMalp/DEST and pThx/DEST Gateway vectors developed at the Salk Institute, San Diego, CA, USA. For overexpression in plants, we have obtained a set of Gateway plant expression vectors developed by Mark Curtis from the University of Zurich. We have currently cloned around 4 new cDNA into pENTR4 or pENTR/TOPO for introduction into this system. In addition to the new genes we have cloned into this system, we plan to put the *bglu2* gene into the system, since previous expression levels were too low for characterization.

Initially, one rice β -glucosidase cDNA, designated cg445-1 based on the corresponding gene in the Chinese indica rice database, was amplified by RT-PCR and cloned into this system (pENTR/TOPO). The sequence of this gene was nearly identical to the N-terminal protein sequence of the cell-wall-associated β -glucosidase purified from germinating rice by Akiyama and colleagues (1998). The protein with β -glucosidase activity was produced from the pBAD/DEST plasmid in *E. coli* strain Top Ten, from the pET32/DEST plasmid in both Origami and OrigamiB strains of *E. coli* and from the pPICZ α NH/DEST protein in GS115 *Pichia pastoris*. The highest levels of activity were obtained from the pET32/DEST plasmid in OrigamiB and the pPICZ α NH/DEST system. Both of these systems allowed a portion of the protein to be purified by IMAC purification, but the other two systems did not allow significant purification due to poor binding of the proteins to the resin. In all cases, it has been speculated that proteolysis that removes the tag may be responsible for the poor binding. Enough protein was produced from the two systems to allow some characterization. It was found that the enzyme hydrolyzed pNP- β -D-glucoside best, followed by pNP- β -D-galactoside, then pNP- β -D-fucoside, but did not hydrolyze pNP- β -D-mannoside. This was surprising, since the enzyme purified by Akiyama and colleagues did not hydrolyze the β -galactoside, but did hydrolyze the β -mannoside. The protein expressed in *E. coli* and *P. pastoris* had the same specificity.

For in plant analysis, we have grown up around 50 mutant lines with Tos17/2 insertions in glycosyl hydrolase family 1 and 35 genes. However, so far no obvious phenotypes have been identified and we have as yet only confirmed insertion in the gene of interest in five of these rice lines. This will require that a student be trained to do this kind of work, which is currently the slowest aspect. We have also obtained vectors for RNAi knock down and plan to try this with various genes of these families in the future.

1.3 Over-expression of *Vibrio carchariae* chitinase A in *E. coli* System for Functional and Structural Studies

Since chitin is the second-most abundant hydrocarbon source after cellulose in nature, it has recently received attention as a target in the exploitation of the biomass. Chitin is a recalcitrant molecule, insoluble in water and organic solvents, hence it requires harsh conditions for chemical recycling. On the other hand, it is attractive to develop a gentler bioconversion process based on chitinases. Chitinases are a diverse of enzymes that mainly degrade chitin polymer, releasing smaller fragments of chito-oligosaccharide products. The enzymes are considered as an efficient source for treatment of chitin, which is produced in a multi-million ton quantity per annum as a waste by-product by seafood industry. Chitinase hydrolytic products have potentially useful applications in the areas of medicine and pharmaceutical industry. For example, chito-oligosaccharides have a remarkable compatibility with living tissue and have been proved to promote the wound-healing process. Some forms of chitin derivatives could be used as drug carriers or conjugated drug delivery systems. In addition, chitin metabolism is a potential drug target for the treatment of some forms of human parasite and fungal infections.

Dr. Wipa Suginta has previously purified chitinase A from a marine bacterium, *Vibrio carchariae*. The gene that encodes chitinase A has been isolated from a genomic library of *V. carchariae*, and then subsequently cloned into the pBluescript KS II (-) cloning vector. A clone, designated P3C1, carrying a 4.0-kb DNA fragment, which contains the full-length nucleotide sequence of chitinase A. This 2.7-kb DNA fragment included the putative regulatory region (*Chi A* promoter), ribosome binding sequence (Shine-Dalgarno sequence), 21-amino acid signal peptide, a 1.7-kb structural gene encoding the mature chitinase A, and the C-terminally processed fragment. This P3C1 clone exhibited high expression level in *E. coli* DH5a upon induction by swollen chitin and IPTG. However, the protein was found to be expressed as an inactive precursor with higher molecular weight (95 kDa) than the secreted, functional enzyme (63 kDa).

Dr. Wipa Suginta's research proposed to: i) establish a high expression system of *V. carchariae* chitinase A in *E. coli* to obtain the recombinant chitinase A in active form; ii) study the enzymatic properties of chitinase A by comparing the chitinolytic activity, as well as synthetic properties of the expressed enzyme with the native enzyme; iii) investigate the molecular structure of chitinase A based upon the available nucleotide sequence data and; iv) set up crystallization trials aiming to obtain the complete 3D-structure of the enzyme.

1.3.1 Peptide mass analysis and determination of C-terminal proteolytic cleavage site of *V. carchariae* chitinase A

The ensemble of peptides, which were eluted from in-gel digestion of chitinase A purified from *V. carchariae*, was analyzed by means of MALDI-TOF or

nanoESI mass spectrometry and subjected to a data-bank search. This process, commonly denoted as mass fingerprinting, resulted in an unambiguous match of this protein to the chitinase A gene (Table 7). The molecular mass of peptide T1 (2040.7) agreed well with the theoretical mass (2040.0) of the N-terminal peptide identified previously by microsequencing. In addition, peptide T39, identified as being nearest to the C-terminus, had a mass of 551.0, which matched the mass of the tryptic peptide sequence GNYAK.

The chitinase A precursor had a calculated M_r of 90,249, which was slightly less than indicated by SDS-PAGE (95,000). Because the chitinase precursor was inactive and its molecular mass was approx. 23 kDa larger than the native enzyme, the precursor must be cleaved by a proteinase in *V. carchariae* in order to form the active 63-kDa chitinase A. MALDI-TOF measurements yielded a peak of M_r 62,698, which corresponded to the mass of chitinase A predicted to end at Arg597 (calculated M_r 62,718.12).

Table 7: Mass identification of tryptic peptides of *V. carchariae* chitinase A by MALDI-TOF or nanoESI mass spectrometry

position in the sequence ^a	tryptic peptide	expected mass	observed mass	peptide AA sequence
22-40	T1	2040.0	2040.7	APTAPSIDMYGSNNLQFSK
41-56	T2	1800.8	1800.5	IELAMETTSGYNDMVK
57-62	T3	759.4	759.3	YHELAK
67-86	T6	2284.0	2283.8	FNQWSGTSGDTYNVYFDGVK
124-151	T8	2976.5	2976.6	SAPVEITIADTDGSHLKPLTMNVD PNNK
152-173	T9	2554.2	2553.6	SYNTDPSIVMGTYFVEWGIYGR
204-218	T11	1523.7	1523.6	SVGGNSFNALQTAC
219-236	T12	2103.0	2102.8	GVNDYEVVIHDPWAAYQK
237-250	T13	1560.8	1560.7	SFPQAGHEYSTPIK
251-262	T14+T15	1394.7	1394.5	GNYAMLMALKQR
268-288	T17+T18	2397.2	2397.0	IIPSIGGWTLSDPFYDFVDKK
289-298	T19+T20	1135.6	1135.6	NRDTFVASVK
303-326	T23+T24	2601.2	2601.2	TWKFYDGDVIDWEFPGGGGAAADK
327-341	T25	1587.8	1587.6	GDPVNDGPAYIALMR
345-356	T27	1361.7	1361.7	VMLDELEAETGR
454-463	T31	1079.6	1079.6	LVLGTAMYGR
464-487	T32	2460.1	2460.1	GWEGVTPDTLTDPNPMTGTATGK
488-505	T33+T34	1965.0	1964.8	LKGSTAQQVWEDGVIDYK
509-538	T36	3377.5	3378.5	SFMLGANNTGINGFEYGYDAQA EAPWVWNR
539-550	T37	1389.7	1389.6	STGELITFDDHR
551-555	T38	516.3	516.5	SVLAK
556-560	T39	551.3	551.3	GNYAK

^aUnidentified peptides are not included.

1.3.2 Sequence comparison and structural topology prediction

Since the gene encoding chitinase A from *V. carchariae* was isolated and its nucleotide sequences was determined previously by Suginta et al.(2000), the deduced amino acid sequence of chitinase A from *V. carchariae* was then compared with other bacterial Chi A sequences. The putative mature chitinase A gave highest identity with Chi A from *V. parahaemolyticus* (94%), followed by Chi A from *S. liquefaciens* (48%), Chi A from *Alteromonas* sp. (47%), and Chi A from *Enterobacter* sp. (47%), Chi A from *S. marcescens* (47%), and Chi A from *Pantoea agglomerans* (44%), respectively. *V. carchariae* chitinase A aligned with Chi A from *Bacillus circulans* with low identity (18%). The amino acid alignment revealed that bacterial chitinase A is highly conserved in the catalytic region (data not shown). Two completely conserved motifs are found within the predicted TIM barrel catalytic domain: SxGG (located at the β 3 strand) and DxxDxDxE (located at the β 4 strand) The function of the SxGG is still unknown. In the enzyme catalysis of family 18 chitinases, which requires a substrate-assisted mechanism, Glu315 located at the end of the DxxDxDxE motif has been strongly proved to serve as the essential catalytic residue of chitinase A. The aspartic acid residue Asp392 in the *V. carchariae* sequence has been suggested to help stabilise the transition states flanking the oxazoline intermediate and to assist the correct orientation of the 2-acetamido group during the enzyme catalysis.

The secondary structure of *V. carchariae* chitinase A was predicted to be similar to the published 3D-structure of SmChiA from *Serratia marcescens*. Basically, the structure contains three main domains: (i) the N-terminal chitin binding domain (Chi N domain) (residues 21-138) consisting only of β -strands, which are connected through a hinge region (residues 139-159) to (ii) the main $(\alpha/\beta)_8$ -TIM barrel domain (residues 160-458 and residues 549-591) and (iii) a small domain (residues 459-548), which has an $\alpha + \beta$ -fold structure. This small insertion domain is an excursion from the $(\alpha/\beta)_8$ barrel between strand β 7 and helix α 7. The predicted topology of the catalytic domain of *V. carchariae* chitinase A (Figure 6) is shown to be similar to the catalytic domain of SmChi A from *Serratia marcescens*. The *V. carchariae* catalytic domain covers Val160-The458 and His549-Val591 and comprises eight β -strands running parallel to one another throughout the barrel, together with α -helices that lie anti-parallel to the barrel. Interestingly, small α -helices (G1-1, G1-2, and G1-3) that have been identified as parts of helix A1 in the *S. marcescens* $(\alpha/\beta)_8$ barrel are missing in the *V. carchariae* barrel. As a result, the predicted catalytic domain of *V. carchariae* chitinase A would comprise seven helices instead of eight helices as commonly found for other $(\alpha/\beta)_8$ -barrel enzymes.

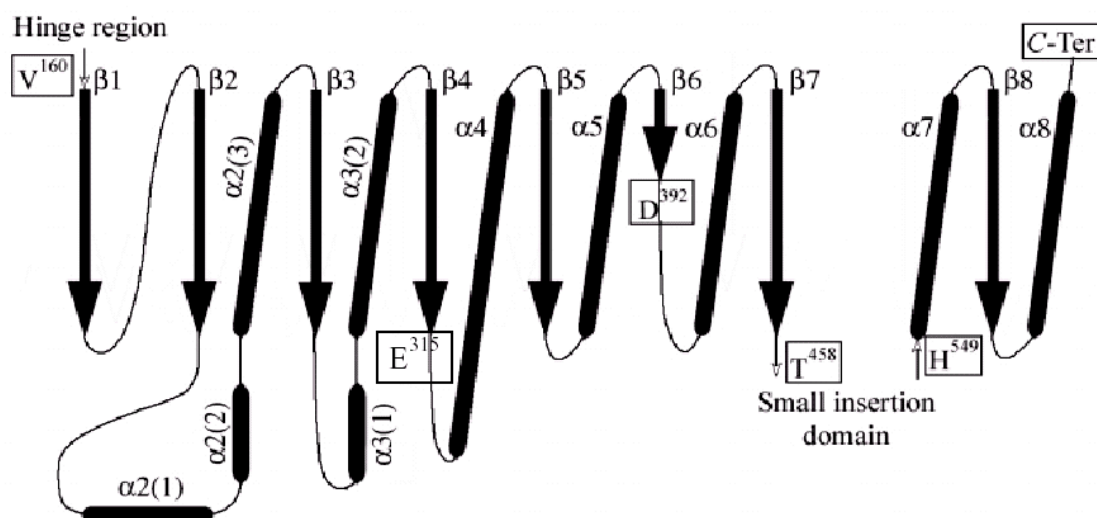


Figure 6: Predicted topology of the catalytic domain of *V. carchariae* chitinase A

1.3.3 Cloning and expression of chitinase A in *E. coli*

Chitinase A expressed in *E. coli* as the unprocessed precursor was much less active than the enzyme purified from *V. carchariae*. Taking advantage of the M_r of the native enzyme obtained from MALDI-TOF measurement, two oligonucleotides were designed to generate the protein without the 23-kDa C-terminal proteolytic peptide. A 1.7-kb DNA fragment encoding the C-terminally processed chitinase A was cloned into pDrive cloning vector, and later transferred to the pQE60 expression vector. The protein was expressed under the T5 promoter in *E. coli* M15 under optimized conditions with high yield (~ 10 mg/l culture). Using such a system, the expressed chitinase was a hybrid protein with six histidines tagged at the C-terminus. The protein was purified using Ni-NTA agarose affinity chromatography, followed by Superdex-S200 HR FPLC. Electrospray MS confirmed the M_r of the expressed protein to be 63,823 (± 15), corresponding to the calculated M_r of the mature chitinase (62,718.12 Da) plus two additional amino acids: arginine and serine (260.27 Da). These amino acids were encoded by the six nucleotides (AGATCT) corresponding to the *Bgl* II cloning site following the codon of Arg597 of the mature chitinase A, and formed a link to the C-terminal histidine tag residues (839.85 Da). The total calculated mass of the expressed protein is therefore 63,818.24 Da. The expressed protein exhibited chitinase activity using the gel activity assay with glycol chitin substrate.

As analyzed by HPLC/ESI-MS, chitinase A, expressed in *E. coli*, was able to hydrolyze colloidal chitin. Figure 7 shows a HPLC-MS chromatogram of chitooligosaccharide products acquired after 5 min of reaction time. The enzyme degraded chitin polymer releasing chitooligosaccharide products ranging from GlcNAc to [GlcNAc]₇ with [GlcNAc]₂ as the major product (> 80% of the total products). Although [GlcNAc]₅ and [GlcNAc]₇ were not clearly seen in the HPLC-MS chromatogram, their molecular masses were certainly observed in the MS spectrum. The release of chitooligosaccharide products with various sizes confirmed the endo characteristic of *V. carbariae* chitinase A. The hydrolytic activity determined for the expressed chitinase A was found to be identical to that of the native enzyme tested with the same method.

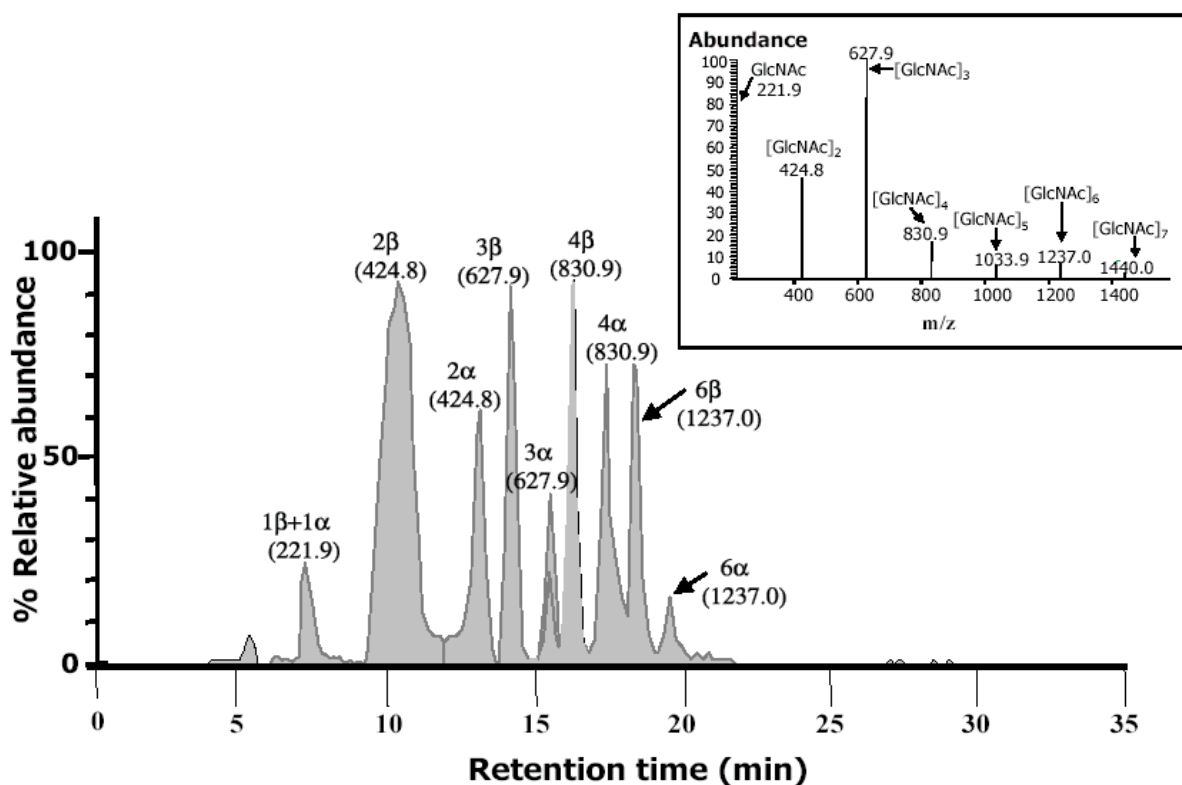


Figure 7: HPLC-MS chromatogram of hydrolytic products of chitinase A expressed in *E. coli*.

1.3.4 Enzymatic properties of chitinase A from *V. carbariae*

Enzymatic properties of the native and recombinant chitinase A were studied using quantitative HPLC–ESI mass spectrometry. Since initial data gave identical characteristics between the two enzymes, chitinase A from *V. carbariae* was later chosen for detailed investigation of substrate hydrolysis. In conclusion, a combination of HPLC and ESI-MS allowed separation of α and β anomers and simultaneous monitoring of all chitooligosaccharides produced. Chitinase A primarily produced β -anomeric products, indicating that it catalyzed hydrolysis

through a retaining mechanism. Notably, the enzyme did not hydrolyze dimers, but required trimers as shortest substrate, producing (GlcNAc)₂ as the major product, together with GlcNAc after complete hydrolysis. Hydrolytic activity of Chitinase A against short chitooligosaccharides and pNP glycosides was determined as function of time. The “classical” colorimetric assay employing pNP glycosides gave identical results to the analysis based on calibrated mass spectrograms. The kinetics of hydrolysis showed K_m and k_{cat}/K_m values to be: pNP-(GlcNAc)₂ (1.1 mM, $6.8 \times 10^3 \text{ s}^{-1}\text{M}^{-1}$), (GlcNAc)₃ (11.9 mM, $8.7 \times 10^2 \text{ s}^{-1}\text{M}^{-1}$), (GlcNAc)₄ (2.2 mM, $2.7 \times 10^2 \text{ s}^{-1}\text{M}^{-1}$), (GlcNAc)₆ (0.2 mM, $3.1 \times 10^4 \text{ s}^{-1}\text{M}^{-1}$), and chitin (0.2 mg/ml), respectively. Hence, Chitinase A had a greater affinity towards higher molecular weight chitooligosaccharides. Chitinase A exhibited endo-characteristics while cleaving either the second β -glycosidic linkage from the non-reducing end or the internal glycosidic linkages of the sugar chains. Relating our findings to the reported structures of other chitinases, we could confirm six GlcNAc subsites, two of which forming a predominant recognition site.

Using HPLC/MS to study the hydrolysis of pNP-(GlcNAc)_n substrates, we also found evidence for polymerisation. The amount of oligomers produced was in the order of 1-5% of the substrates. Apart from the much more prominent formation of the (GlcNAc)₂ hydrolysis product, trimer (GlcNAc)₃ and tetramer (GlcNAc)₄ were also primarily synthesized from pNP-(GlcNAc)₂ by chitinase A. Upon the availability of the standard sugars tested in the experiments, the synthetic products were observed to be (GlcNAc)₄ and (GlcNAc)₆ when pNP-(GlcNAc)₃ was used as the substrate. Figure 8 shows the synthesis of (GlcNAc)₃ and (GlcNAc)₄ from pNP-(GlcNAc)₂ within the first 60 minutes of reaction time. It is notable that the tetramer, once synthesized, is hydrolysed considerably faster than the trimer. Under given conditions (low temperature, short reaction time and relatively low substrate concentrations), oligosaccharide synthesis was likely to take place through transglycosylation rather than reversal of the hydrolytic reaction.

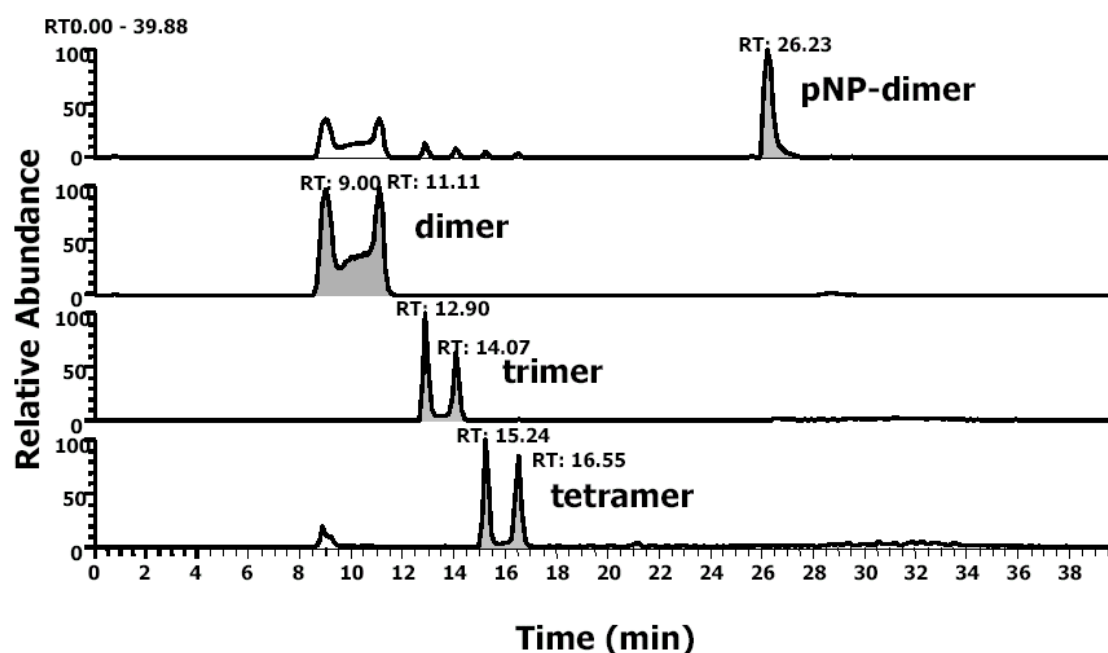


Figure 8: Oligosaccharide synthesis from pNP-(GlcNAc)₂ by chitinase A.

1.3.5 Preliminary study of protein crystallization

Since high-level expression of the active chitinase A has recently been established in the *E. coli* M15, rapid screening for chitinase crystals were carried out by a microbatch method using the Crystal Screen kits from Hampton Research, USA and the JBScreen HTS I and HTS II from Jena Bioscience, Germany. Crystallization of the purified chitinase A (10 mg/ml) was set up using the hanging-drop method. Protein crystals were grown in the crystallization buffer containing 0.1 M sodium acetate buffer, pH 4.6 containing 10% PEG 400, and 0.125 M CaCl₂ as precipitant, at 15°C. Crystals of chitinase were observed after 4 days of incubation (Figure 4). We found that 0.1 M sodium acetate buffer, pH 4.6 containing 20% Glycerol, 10% PEG 400 and 0.125 M CaCl₂ could be used as a cryoprotectant solution for crystal mounting in cold N₂-stream (−160°C).

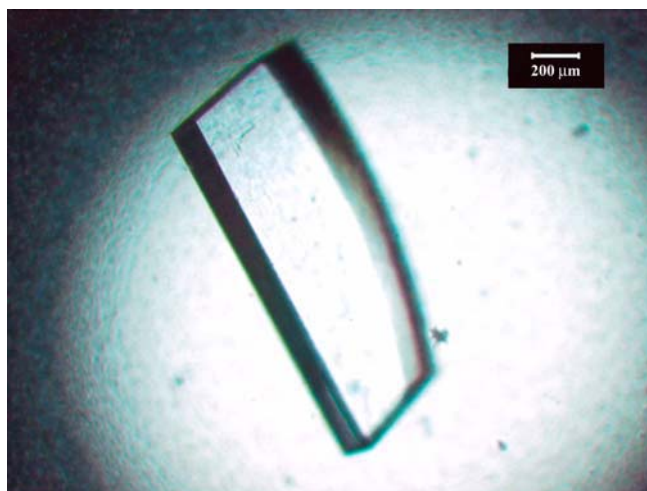


Figure 9: A crystal of chitinase A expressed in *E. coli*

Under a collaborative project with Dr. Jirundon Yuvaniyama, Mahidol University, diffraction data of the obtained crystals were collected using a Rigaku/MSR R-Axis IV⁺⁺ detector with a RU-H3R rotating-anode generator with Osmic Blue confocal focusing mirrors running at 50 kV and 100 mA. Diffraction data were recorded over 65° rotation of the crystal around phi axis in 260 diffraction images with a width of 0.25° per image. Using the program d*TREK, primarily structural analysis revealed that the crystal diffracted X-rays up to 2.14 Å resolution with 97.2% completeness and belonged to the tetragonal space group $P4_122$, or $P4_322$ with unit-cell parameters $a = 127.64$, $b = 127.64$, $c = 171.42$ Å. Currently, SmChiA from *Serratia marcescens* and a fungal chitinase (CiX1) from *Coccidioides immitis* are being used as phasing models to provide an initial Molecular Replacement (MREP) solution. Other chitinases from the family 18 may also be tried considering their moderate sequence identities with chitinase A. Alternatively, Multiple Anomalous Diffraction (MAD) technique performed with a synchrotron radiation source will be attempted to overcome the phase problem.

In conclusion, the enzymatic properties of chitinase A from *V. carchariae* were studied using HPLC-ESI MS. Under given conditions, chitinase A revealed both hydrolytic, as well as synthetic activities, giving an indication in utilisation of the enzyme in biotechnological and medical applications. Identification of the C-terminal proteolytic cleavage site using MALDI-TOF MS provided the possibility for the gene encoding the processed chitinase A to be cloned and functionally expressed in a high expression *E. coli* system. The recombinant protein exhibited chitinase activity and could form crystals with x-ray crystallization quality,

providing a good implication in 3D-structural determination. For future research, x-ray crystallization will be employed to obtain the 3D-structure of the enzyme both in free form and in complex with chitoooligosaccharide substrates. It is hoped that the structural information will provide an insight into the molecular mechanism of the enzyme in chitin hydrolysis. Understanding the mode of enzyme action will eventually lead to generation of a new enzyme molecule with particular properties that are applicable for biotechnological purposes.

1.4 Purification and Catalytic Properties of Plant α -Mannosidases

α -Mannosidases (EC.3.2.24) are glycoside hydrolases involved in both the maturation and the degradation of Asn-linked oligosaccharides in all eukaryotes. Traditionally, they have been grouped into two classes (I and II) based on functional characteristics, sequence homology and their cellular compartmentation. Class I α -mannosidases, classified to family 47 of glycosyl hydrolase, only hydrolyze α -1,2 mannose bonds and are all involved in the maturation of Asn-linked oligosaccharides, by trimming of $\text{Man}_9\text{GlcNAc}_2$ to $\text{Man}_5\text{GlcNAc}_2$. Class II α -mannosidases, classified to family 38 of glycosyl hydrolase, possess α -1,3, α -1,6, and α -1,2 hydrolytic activity, and catalyze the degradation of Asn-linked oligosaccharides. Class II enzymes can be localized to the cytosol, lysosomes, and Golgi complex, while class I enzymes are found in either the endoplasmic reticulum (ER) or the Golgi complex.

α -Mannosidases can be used in carbohydrate bioengineering by using their specific hydrolytic activity and in synthesis of various mannose-containing oligosaccharides both in natural and synthetic substrates. Of the plant enzymes, Jack Bean α -mannosidase is well known and is often used for reverse hydrolysis. α -Mannosidases have also been studied in our laboratory. Notably, we have purified α -mannosidases from red bean, and studied their specificity for hydrolysis of different mannose disaccharides, as well as specificity of bond formation in reverse hydrolysis. We have also found moderately high levels of α -mannosidases from *Hibiscus sabdariffa* var *sabdariffa* (roselle) and *Albizzia procera* Benth, and both of these enzymes have been shown to give good yields in reverse hydrolysis involving mannose alone and mixtures of mannose with raffinose, producing various oligosaccharide products including tetrasaccharides.

1.4.1 α -D-Mannosidase from *Albizzia procera* Benth.

α -Mannosidase is being purified from *Albizzia procera* Benth. in Dr. Sujint Anguravirutt's laboratory at Mahasarakham University (MSU). Seeds of *Albizzia procera* Benth. were surface-sterilized, soaked overnight in distilled water, and homogenized in 0.05 M sodium acetate buffer, pH 5.0, containing 1 mM phenylmethyl sulphonylfluoride and 5% (w/v) polyvinyl-polypyrrolidone. The

homogenate was centrifuged and treated with 25% Dowex 2-X8. The crude extract was fractionated by ammonium sulfate at 35-75% saturation, and the precipitate was resuspended in 0.1 M sodium acetate buffer, pH 5.0, dialyzed against 0.02 M potassium phosphate buffer, pH 6.5 overnight, and fractionated by anion-exchange chromatography on a Diethylaminoethyl-Sepharose column (ANX Sepharose 4 Fast Flow, Amersham Biosciences) equilibrated with 0.02 M potassium phosphate buffer, pH 6.5. The column was first washed with 2 volumes of the starting buffer to elute the unbound proteins. The bound proteins were eluted with 10 volumes of the same buffer containing a linear gradient of 0-0.2 M NaCl, followed by 10 volumes of a linear gradient of 0.2-0.5 M NaCl, 5 volumes of a linear gradient of 0.5-1.0 M NaCl, and 5 volumes of 1.0 M NaCl in the buffer, respectively. Fractions containing α -mannosidase activity, eluted at 0.25-0.4 M NaCl, 0.02 M potassium phosphate buffer, pH 6.5, were pooled.

The enzyme pool was concentrated and loaded onto a Sephacryl S-300 HR gel filtration column (HiPrep 16/60, Amersham Biosciences) equilibrated with 0.05 M potassium phosphate buffer, pH 6.5 containing 0.15 M NaCl. Proteins were eluted from the column with the same buffer. α -Mannosidase fractions had molecular weight of about 320 kDa, and were loaded on a Butyl Sepharose 4 Fast Flow column equilibrated with 1.5 M ammonium sulfate in 0.02 M potassium phosphate buffer, pH 6.5. After washing with 2 volumes of starting buffer to elute unbound proteins, bound proteins were eluted with 20 volumes of the same buffer containing a decreasing linear gradient of 1.5-0 M ammonium sulfate, followed by 10 volumes of 0.02 M potassium phosphate buffer, pH 6.5.

From 100 g of *Albizzia procera* Benth. seeds, 1.8 mg of purified α -mannosidase was obtained with specific activity 594 nkat/mg, representing a 160-fold purification with 12% yield. The resulting enzyme obtained is at least 95% pure by polyacrylamide gel electrophoresis, and of sufficient purity for use in study of kinetic properties. .

The enzymes properties of the α -D-mannosidase from *Albizzia procera* Benth., from the Butyl Sepharose 4 column were studied, as follows. Study of the dependence of activity on pH showed a rather broad pH optimum, ranging from pH 3.0-pH 5.5, with maximal activity at pH 4.0. When temperature was varied, highest enzyme activity was obtained at 60 °C. Study of activity with various *p*-NP-glycosides at 1 mM indicated that the enzyme showed by far the highest activity for *p*-NP- α -D-mannoside, with other *p*-NP-glycosides showing <2.5% of the activity compared to *p*-NP- α -D-mannoside. Kinetic studies showed that the K_m and V_{max} of α -D-mannosidase from *Albizzia procera* Benth. for *p*-NP- α -D-mannoside were 1.2 mM and 0.9 μ mol/min, respectively.

The effect of various compounds on the hydrolytic activity of the α -D-mannosidase from *Albizzia procera* Benth. was studied. Aliquots of the enzyme were incubated in buffer containing 1mM EDTA, 1mM ZnCl₂, 1mM CaCl₂, 1mM MgCl₂, 1mM KCN, respectively, and the enzyme activity was determined using *p*-NP- α -D-mannoside as substrate. The results indicated that none of the metal ions tested had significant effect on the activity of the enzyme.

The α -D-mannosidase from *Albizzia procera* Benth. was also tested for transmannosylation of alcohols to synthesize alkyl mannosides. D-mannose or *p*-nitrophenyl α -D-mannopyranoside was used as mannosyl-donor while methanol, ethanol, 1-propanol, 2-propanol, 1-butanol, 2-methyl-1-propanol, 2-butanol, 2-methyl-2-propanol, 1-pentanol, 1-hexanol, and 1-heptanol, respectively, were used as mannosyl-acceptor substrates. The products formed were analyzed by TLC. The results showed that the alkyl mannosides could be obtained from every alcohol tested, except from 2-methyl-2-propanol, but with different yields. Short-chain primary alcohols gave good yield of alkyl mannoside, while long-chain primary alcohol and secondary alcohols gave poorer yields.

1.4.2 α -D-Mannosidase from *Hibiscus* spp.

Since the content of glycosidase in *Hibiscus* spp. was variable, we checked the content of glycosidases, i.e., α -mannosidase, α -galactosidase, β -galactosidase, β -glucosidase, β -N-acetyl-glucosaminidase in *Hibiscus* spp. plant of 3 varieties (var. *sabdariffa*, CV. Kleebyao, and a hybrid variety) and okra (*Hibiscus esculentus* L. CV. OK#5) at different imbibing and germination time. The isozyme content was also studied by non-denaturing gel electrophoresis staining with 4-methylumbelliferyl- α -D-mannopyranoside. Compared to the other varieties, *Hibiscus* spp. plant var. *sabdariffa* shows highest activity of α -mannosidase, so this variety is used for further study. Identification of the variety of this plant was confirmed by taxonomic study (coordinated with Assist. Prof. Oraphan Sangkajantranon (Dept. of Botany, Faculty of Liberal Arts and Science, Kasetsart University, Kamphangsaeen Campus) and Dr. Supachitra Chadchawan (Dept. of Botany, Faculty of Science, Chulalongkorn University), indicating its species is *Hibiscus sabdariffa* L. var. *sabdariffa* or in common name, roselle.

As a result, two isozymes of α -mannosidase were purified from this roselle by 30-80% ammonium sulfate fractionation and various chromatographic steps, i.e., DEAE-cellulose, CM-cellulose, hydroxyapatite, Sephacryl HR S-200 and Sephacryl HR S-300. In all cases, buffers used were sodium phosphate buffers, pH 7.7, with appropriate concentration, except for CM-cellulose, which was sodium acetate buffer, pH 4.4 + 1 mM ZnCl₂. The enzyme, passing through hydroxyapatite chromatography, was free from β -mannosidase and the other major glycosidases, i.e., α -galactosidase, β -galactosidase, β -glucosidase, β -N-

acetyl-glucosaminidase and used for preliminary oligosaccharide synthesis study using the reverse hydrolysis reaction. Appropriate analysis systems for separation of mannose, xylose, glucose, galactose, arabinose, and various mannobioses linked with $\alpha(1\rightarrow2)$, $\alpha(1\rightarrow3)$, $\alpha(1\rightarrow4)$, and $\alpha(1\rightarrow6)$ were first studied by using thin-layer chromatographic (TLC) technique (silica gel) with various developing systems. Moderately good separation of the above standards was obtained from the combination of the 3 successive analysis from 5 developing systems. Reverse hydrolysis of α -mannosidase using mannose as sugar donor and acceptor gave various mannobioses as products. Highest product yield was obtained at pH 4.0, and at 60°C. The amount of products occurred was measured from the intensity of the spots. After 3 and 7 days incubation with 50% mannose at optimal pH and temperature, $\text{Man}\alpha(1\rightarrow6)\text{Man}$ (623 mM at 3 days and 783 mM at 7 days) was the major product compared to $\text{Man}\alpha(1\rightarrow2)\text{Man}$, $\text{Man}\alpha(1\rightarrow3)\text{Man}$, or $\text{Man}\alpha(1\rightarrow4)\text{Man}$ (total of 541 mM at 3 days and 662 mM at 7 days). Reverse hydrolysis of α -mannosidase was also performed using mannose as sugar donor and other monosaccharides as sugar acceptors. No products were detected by using 10% mannose with 40% xylose or 40% glucose. However, when enzyme was incubated with 10% mannose with 40% galactose or 40% arabinose, unknown mixed products were obtained in good yield.

Both isozymes from Sephacryl HR S-300 were recovered in approximately in 3% yield compared to crude extract. Two major bands from SDS-PAGE were obtained from Sephacryl HR S-300, showing molecular weight at 69.6 kDa and 56.6 kDa, with the 69.6 kDa-band having double intensity. The isozyme I, with pI value of 4.8 was separated from isozyme II, with pI value of 4.9-6.0, using CM-cellulose chromatography at pH 4.4 + 1 mM ZnCl_2 with a pH step pH gradient from pH 4.4-5.0 at intervals of 0.2 pH unit and then with salt linear gradient from 0-0.3 M NaCl at pH 5.0. Both isozymes are drastically lost their activities. However, a single band having the molecular weight of 69.6 kDa was obtained from SDS-PAGE with the isozyme I sample, whereas isozyme II could not be seen since the protein concentration was not sufficient to run SDS-PAGE with coomassie staining. The native molecular weight of the enzyme on Sephacryl HR S-300 column was in the range of 360-480 kDa, indicating that the enzyme should comprise of 6 subunits. Both isozyme I and II, seem to be unstable after prolong storage at 4°C in buffer with various pHs, even containing ZnCl_2 , and after running on hydroxyapatite. However, isozyme I seem to be more promising for study than isozyme II, due to the higher yield and purity, so work will focus on isozyme I. The kinetic values, K_m and V_{max} , of isozyme I towards p-nitrophenyl- α -D-mannopyranoside are 1.05 mM and 0.36 $\mu\text{mol}/\text{min}$, respectively. For hydrolysis, optimum pH and temperature were at pH 4.0-4.5 and 60-70°C, respectively, similar to the optimum pH and temperature for synthesis. When dialysing the partially purified enzyme in buffer at pH 5.5 and 7.0 with and without EDTA overnight, only the activity of enzyme dialysed in

buffer pH 5.5 with EDTA is lost by about 50% compared to the sample dialysed in the same buffer but without EDTA. But the activities of both enzymes dialysed in buffer pH 7.0 with and without EDTA are in the same range. Incubation of sample dialysed at pH 5.5 with EDTA in ZnCl_2 , CoCl_2 or CaCl_2 for 1 hr at 30°C shows that the activity of only the sample incubated in ZnCl_2 was fully recovered. However, addition of Zn^{2+} in assay buffer at pH 5.0 does not improve the activity of the purified enzyme. In addition, assay of the enzyme in 1 mM EDTA with and without addition of 1 mM ZnCl_2 does not seem to decrease or increase the activity. This indicates that the enzyme requires Zn^{2+} only for stabilising activity and possibly not for catalysis.

2. Enzymes Involved in Synthesis of Penicillin Derivatives

Penicillin is one of the most important antibiotics in the world, and many derivatives have been developed with increased efficacy. Use of enzymes for various steps in the synthesis of penicillin derivatives provides important alternatives for chemical synthesis, due to the high specificity and moderate conditions of enzyme reactions. The enzymes to be studied here are not only of potential value in terms of applications, but will provide interesting lessons in understanding protein structure and function relationships. Two projects are in progress.

2.1 Conversion of Bacterial Penicillin G Acylase to Cephalosporin C Acylase by Protein Engineering

The production of semisynthetic cephem antibiotics requires 7-amino-cephalosporanic acid (7ACA) as an important intermediate, which can be generated by hydrolysis of cephalosporin C (CC) as shown in Figure 10. Enzymatic hydrolysis of CC is preferable to chemical reactions due to the mild conditions used and advantages in terms of cost effectiveness as well as environmental safety. Unfortunately, CC acylases identified from various sources either have low activity toward CC or are actually specific for 7- β -(4-carboxy-butanamido)-cephalosporanic acid (glutaryl-7ACA or GL-7ACA), an intermediate in an alternative pathway (shown with dashed arrows in Figure 10). The high costs of running multiple bioreactors, and the low activity of CC acylase for a single-step reaction make enzymatic production of these antibiotics unfavorable. Penicillin G acylase (PGA; benzylpenicillin acylase; EC 3.5.1.11) hydrolyses penicillin G (PG) to phenylacetic acid (PA) and 6-aminopenicillanic acid (6APA). As the CC and PG structures are relatively similar (see Figure 1) and they can be hydrolysed by the same deamidation mechanism, changing substrate specificity of PGA to recognise CC is a possibility for developing the desired CC acylase activity. Work with this aim is being carried out in Dr. Jirundon Yuvaniyama's laboratory, and progress so far is as follows.

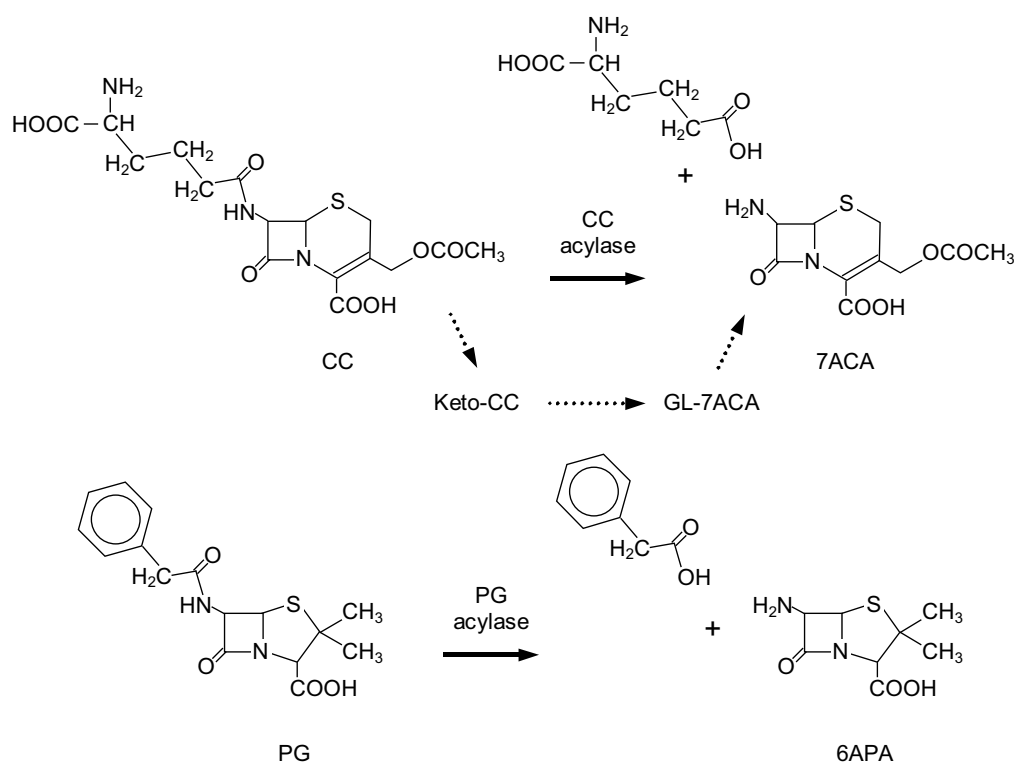


Figure 10: Deamidation of cephalosporin C (CC) to 7-aminocephalosporanic acid (7ACA) and penicillin G (PG) to 6-aminopenicillanic acid (6APA). Dashed arrows show an alternative pathway of CC hydrolysis currently used in industry that occurs through 7-β-(5-carboxy-5-oxopentanamido)-cephalosporanic acid (keto-CC) and 7-β-(4-carboxybutan-amido)-cephalosporanic acid (glutaryl-7ACA; GL-7ACA) intermediates.

2.1.1 Homology modeling of wild-type *Bacillus megaterium* penicillin G acylase (PGA)

Since three-dimensional structures of *Escherichia coli* (ATCC 11105) PGA are publicly available, we have used one of them (PDB: 1PNM) as a template structure in the building of homology model of *B. megaterium* PGA. An alignment of amino-acid sequences of both proteins shows 31% identity (49% homology) suggesting that the homology model can be built with substantial reliability. The alignments of chains A and B of both enzymes were manually checked for possible loop structures of the gap/insertion regions. Then, the structural models of *B. megaterium* PGA chains A and B were separately built based on the optimally aligned sequences.

2.1.2 Identification of amino-acid candidates for mutagenesis of *Bacillus megaterium* penicillin G acylase

A model structure of cephalosporin C (CC) was manually docked into the active site of the *B. megaterium* PGA based on the knowledge of protein–ligand interactions obtained from various structures of *E. coli* PGA complexes. This suggested a binding mode in that the side chain of CC would be buried in the deep hydrophobic pocket of the enzyme. It should be noted that the CC nucleus did not make extensive contacts with the enzyme in this binding mode. Since the

side chain of CC should contain both positive and negative charges in the working pH range of PGA, this binding mode predicted unfavorable interactions between the ligand and enzyme in agreement with the known data.

We have hypothesized that the protein–ligand interactions around the region of CC side chain are the major determinants for CC being a poor substrate for this PGA. Therefore, amino-acid mutations in this hydrophobic pocket which offer more favorable interactions to the CC side chain may contribute to an improved binding recognition of the mutant enzyme towards CC. Although flexibility of the CC side chain may allow for several possibilities in designing the mutations to improve the protein–ligand interactions, we have chosen only one possibility for mutagenesis. Our plan is to initially start with minimal mutation for simplicity in both design and experimental set up and possibly optimize or redesign the interactions later once the actual three-dimensional structure of *B. megaterium* enzyme becomes available. With this plan, we decided to incorporate a histidine and a glutamic acid to provide positive and negative charges in places of a leucine and a methionine, respectively.

2.1.3 Site-directed mutagenesis of *Bacillus megaterium* penicillin G acylase (PGA)

The plasmid pBA402 containing the *pac* gene of *B. megaterium* PGA enzyme kindly provided by Prof. Vithaya Meevothisom has been engineered to add a new *Bam*HI restriction site upstream of the *pac* gene. Together with the existing downstream *Xba*I site, this allowed for the *pac* gene transfer such that all mutagenesis works could be made in the *E. coli* system, which provide an easier control over the plasmid manipulations. The *pac* gene was inserted in the pET3a expression vector and the codons for leucine-56b and methionine-181b were then replaced with those of histidine and glutamate, respectively, as planned. The mutagenesis work was done using DyNazyme EXT (Finnzymes) high-fidelity polymerase in the PCR-based mutation with appropriate oligonucleotide primers. All mutations have been confirmed by DNA sequencing. The mutant gene will be inserted into the modified pBA402 vector, in order to replace the wild-type gene for further characterization in *B. megaterium* system.

2.1.4 Transformation of *B. megaterium* UN-CAT1 host cells

The mutated *pac* gene on the pET3a vector has been re-inserted into the modified pBA402 plasmid vector and then electroporated into *B. megaterium* UN-CAT1 host cells for expression of the engineered protein as planned. It was found later that the host cells obtained from our collaborator already contained some plasmid DNA and could not be used for our work, so the work needs to be repeated with proper host cells.

2.1.5 Preliminary characterization of *Bacillus megaterium* penicillin G acylase (PGA) in *Escherichia coli*

In addition to the planned transformation of *B. megaterium* UN-CAT1 cells with the mutated *pac* gene, the mutant and wild-type genes also have been inserted into the pET3a vector and put in *E. coli* BL21 (DE3) pLysS for expression study. Both the wild-type and mutant proteins could be expressed in comparable amounts upon IPTG induction. The proteins could auto-process resulting in subunits of the correct sizes as estimated with SDS-PAGE. Both types of cells were then tested for survival and growth in liquid media containing either penicillin G (PG) or cephalosporin C (CC) of various concentrations. The *E. coli* cells expressing the mutant PGA could grow in the presence of lower concentrations of either PG or CC, in comparison with those cells expressing the wild-type PGA. This preliminary finding indicates that the mutations of active-site residues have affected the catalytic ability of the enzyme although further investigation is required to understand whether this is due to the change on protein stability or enzyme specificity. However, with recent progress in crystallization of the wild-type *B. megaterium* PGA, the priority of research is being given to X-ray structure determination of the enzyme, since identification of candidate residues for protein engineering is more accurate with its actual structure than with a homology model.

2.1.6 Expression and purification of wild-type *Bacillus megaterium* penicillin G acylase

Wild-type PGA was expressed in *B. megaterium* UN-CAT1 containing pBA402 and then purified using SP-Sephadex C25 column at pH 6.0. The bound PGA was eluted with a 0–500 mM NaCl gradient. SDS-PAGE electrophoresis showed the denatured heterodimer bands of 23 kD and 61 kD (for chains A and B, respectively) with no other detectable impurities.

2.1.7 Crystallization trials of purified *B. megaterium* PGA

The concentrated, purified wild-type PGA obtained from *B. megaterium* UN-CAT1 cells was subjected to crystallization trials in order to assess its crystallizability. Plates and prism-shaped crystals were obtained from the preliminary crystallization screens using a number of buffer, salt, and precipitant combination in the modified microbatch set up. The positive conditions were optimized to grow large single crystals of PGA enzyme suitable for X-ray diffraction studies.

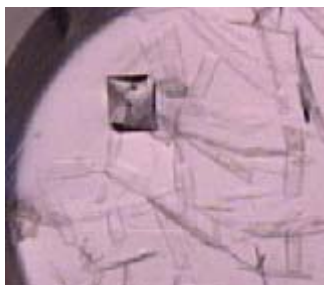


Figure 11: Crystals of wild-type *B. megaterium* penicillin G acylase (PGA) from crystallization screens.

One of the positive crystallization conditions of the wild-type *B. megaterium* PGA has been optimized to obtain moderate-sized crystals ($0.2 \times 0.05 \times 0.05 \text{ mm}^3$). A complete diffraction data set has been collected at cryo-temperature (-160°C) on the R-Axis IV⁺⁺ image-plate detector system mounted on Rigaku/MSR rotating anode generator running at 50 kW. The data were processed with the program CrystalClear to $2.8 \text{ \AA}$ resolution with an R-merge of 6.0%. Preliminary characterization showed the crystal belonged to the monoclinic $P2_1$ space group with cell parameters: $a=59.6 \text{ \AA}$, $b=76.9 \text{ \AA}$, $c=79.4 \text{ \AA}$, and $\beta=99.0^\circ$ and 42% estimated solvent content ($V_M=2.1 \text{ \AA}^3/\text{Da}$). The *Escherichia coli* PGA structures (PDB 1AI5 or related) were used as molecular templates in determining the initial phases using the techniques of Molecular Replacement under the AMORE program of the CCP4 suite. We found that the *E. coli* and *B. megaterium* PGA were quite different, making it difficult to obtain Molecular Replacement solutions. This could explain our earlier lack of success in protein engineering of *B. megaterium* PGA based on a homology model built on the *E. coli* structure. However, preliminary solutions have been obtained and are being further investigated.

Initial electron-density maps were calculated after rigid-body minimization calculation of the template against our X-ray diffraction data. We have been interpreting the electron-density maps and rebuilding the template PGA structural model to reflect the actual *B. megaterium* PGA density. The polypeptide backbone is being traced based on 2mFo-DFc and mFo-DFc Fourier syntheses with difficulty, probably due to significant differences from the *E. coli* PGA structural template. We are still continuing to pursue this with additional approaches in electron-density map calculation in order to simplify the interpretation. Among these, calculations of composite simulated-annealing omit map, as well as phase improvement using ARP/wARP procedure, are being attempted.

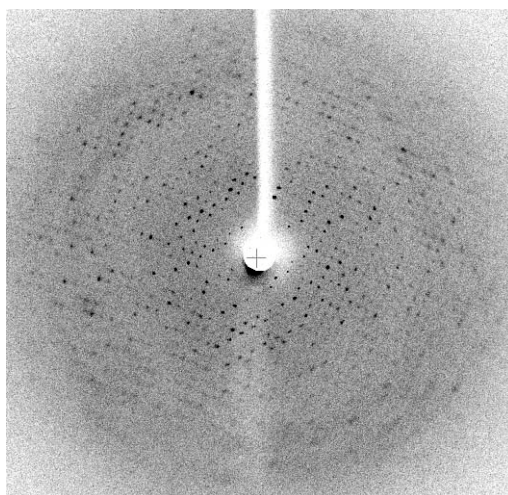


Figure 12: X-ray diffraction pattern of a crystal of wild-type *B. megaterium* PGA

2.2 Three Dimensional Structure of *Pseudomonas stutzeri* D-Phenylglycine Aminotransferase

The enzyme amino acid aminotransferase acts to transfer amino ($-\text{NH}_2$) groups, using pyridoxal-5'-phosphate (PLP) as a coenzyme, and in general does not change the stereochemistry of the substrate or product of the reaction. However, certain enzymes in this group have been found in plants and bacteria, that can change the stereochemistry to give compounds of opposite configuration.

D-phenylglycine aminotransferase (D-PhgAT) has been discovered in soil bacteria *Pseudomonas stutzeri* ST-201, and is a dimeric protein with molecular weight of 92 kDa. It uses D-phenylglycine or D-4-hydroxyphenylglycine as amino group donor and converts 2-oxoglutarate, amino group acceptor, to L-glutamic acid. The enzyme has high substrate specificity and cannot use D- or L-phenylalanine, tyrosine, alanine, valine, leucine, isoleucine or serine as substrate. Studies on the structure of L-amino acid aminotransferase and D-amino acid aminotransferase have been reported. However, the mechanism of the stereo-inversion catalysed by D-PhgAT is an interesting problem that has not been studied yet. Studies of the three-dimensional structure of D-PhgAT will not only provide information about substrate specificity and catalytic mechanism, but will also provide clues on the mechanism of the stereoinversion.

Work in Dr. Palangpon Kongsaree's laboratory, in collaboration with Prof. Vithaya Meevootisom, has shown that recombinant D-PhgAT can be expressed at a level of 20% of total protein using pET-17b as expression vector in *E. coli* BL21 (DE3). The enzyme was purified to 90% homogeneity on SDS-polyacrylamide gel electrophoresis using ammonium sulfate fractionation, Phenyl-agarose chromatography, and DEAE-cellulose chromatography. Initial studies of crystallization were performed at 10 mg/ml

protein concentration, with variation of ammonium sulfate concentration, pH and buffer.

Suitable conditions for crystallisation were found to be 30% saturated ammonium sulfate, pH 5.8-6.0 in 200 mM phosphate buffer. Using these conditions with the hanging drop method, crystals were seen in 5-7 days, and these grew to a size of 150-250 μm in 3-4 weeks. In 35% ammonium sulfate, 200 mM phosphate buffer, pH 6.0, 100 mM NaCl, these crystals were stable at 4°C for at least 1-2 months and at room temperature for at least 1-2 weeks.

For data collection, crystals of native D-PhgAT in cryoprotectant (10-30% glycerol) were frozen at 100 K. Since an X-ray diffractometer for macromolecules was not available at Mahidol University at the time, crystals of native D-PhgAT were studied at the National Synchrotron Light Source, Brookhaven, New York, U.S.A. The crystal has trigonal symmetry with possible space groups being P3(1)21 or P3(2)21, with unit cell parameters of $a = b = 75.155 \text{ \AA}$, $c = 147.559 \text{ \AA}$, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$ and $V = 721792 \text{ \AA}^3$. Each asymmetric should have 1 protein molecule, with solvent content of about 50%. With monochromatic X-ray of $\lambda = 0.97950 \text{ \AA}$, ADSC Quantum-4 CCD detector, crystal and detector distance of 200 mm, oscillation angle of 0.5° , 750 frames, the crystals could diffract to a resolution of 2.2 and 98.9% completeness.

The crystal structure of D-phenylglycine aminotransferase was determined by using the modified X-ray crystal structure coordinates of glutamate-1-semialdehyde aminomutase (Protein Data Bank code: 2GSA) as a search model. Molecular replacement was carried out using data from 8-4 $\text{\AA}$ resolution with a program CNS suite running on a Linux operating system. The cross-rotation function search yielded a clear solution of 4σ with 1.7σ higher than the next best solution. The subsequent PC-refinement and translation function search established the correct space group to be P3₁21 with packing coefficient of 0.54. The molecular replacement model showed an overall satisfactory packing diagram. Preliminary inspection of experimental electron density map revealed a continuous electron density in the core domain of the structure with unclear electron density in some regions. After removing regions with missing electron density, the model was subjected to 20 rounds of rigid-body refinement at 20-3.5 $\text{\AA}$ that yielded a slightly decreased R-factor to about 52.5%. The model was further refined using simulated-annealing refinement protocol at 5000 K with 25 K temperature decrease in each step using data from 20-3.5 $\text{\AA}$. As a result, the protein model was much improved and the electron density was noticeably superior to the previous electron density maps. The model was visually inspected on a graphics workstation with a model-editing program and then was submitted for a simulated-annealing refinement. This protocol was repeated many times between model inspection, model building and model refinement. The current model was refined to 2.2 $\text{\AA}$ with R-factor of 24% (28% R-free). Two regions are missing in the central polypeptide chains, and there are ten disordered residues near the C-terminus.

The disorder problem near the binding site of the current model of D-PhgAT at 2.4 Å has been investigated. One possibility was that the current structure was determined in the space group $P3_121$ with one molecule of D-PhgAT in the asymmetric unit. In crystal structures of aminotransferases, the enzyme exists either as an asymmetric or a symmetric homodimer. This raises the possibility that D-PhgAT could exist as an asymmetric dimer, so that the real space group would be lower and the two-fold axis would become a pseudo local two-fold axis. We then decided to pursue the investigation of the crystal structure in the lower space group, $P3(1)$ with two molecules of D-PhgAT in the asymmetric unit.

The structure of D-PhgAT in the space group $P3_1$ has been re-investigated using the model in the space group $P3_121$ as a search model. The two expected solutions were readily located. The electron density maps looked different and allowed the model to be constructed in each monomer individually. After several rounds of model editing and refinement, there are still some disordered residues. The pyridoxal cofactor binding sites were located in the electron density maps ($1\sigma 2F_o-F_c$ and $3\sigma F_o-F_c$) near the catalytic lysine-269. In both monomers, there was some additional electron density clearly observed. A closer look in this region suggested that the enzyme may have a pyridoxal cofactor existing in the imine complex, covalently linked to the lysine-269. Model building of the co-factor is in progress. Interestingly, the comparison with superimposed models of related aminotransferases suggested that the mode of binding in D-PhgAT may differ from others. The spatial arrangement and the conformation of the co-factor may explain the regiospecificity and stereospecificity of the transamination in D-PhgAT.

Work on the two possible space groups progressed satisfactorily. The current model has over 90% of the structure. The active site region is still not finished as the data showed some degrees of disorder in the crystal structure. So we are trying to collect better resolution data from newly prepared crystals, as well as initiating study of the complex. Much effort has been made to improve the resolution limit of the native crystal by growing large crystals and analyzing the freshly-prepared crystals on a newly installed rotating-anode generator equipped with an imaging-plate area detector R-AXIS IV++ at the Center for Protein Structure and Function. The best resolution obtained was about 2.0-2.1 Å, slightly higher than the original synchrotron data set. The higher resolution will be useful to help us understand the enzymatic mechanism.



Figure 13: Ribbon Diagram of D-Phenylglycine aminotransferase

In order to study the detailed mechanism of the D-PhgAT, the structure of D-PhgAT complexed with its inhibitor was started. The native crystal of D-PhgAT and PLP complex was soaked in a stabilizing solution containing different concentrations of gabaculine, an aminotransferase inhibitor, at various times. The soaked crystal was then transferred sequentially to a stabilizing solution containing 20-30% glycerol as a cryoprotectant before mounting in a nylon loop and flash-freezing under a cryogenic nitrogen stream at $-160\text{ }^{\circ}\text{C}$. The soaked crystal did not diffract X-rays as well as the native crystal with only visually observed reflections to about $3.3\text{-}3.5\text{ }\text{\AA}$ resolution. We expect that the resolution limit of the derivatized crystals can be improved to better than $2.8\text{ }\text{\AA}$, in order to see the mode of binding in the active site of the D-PhgAT enzyme. The preliminary X-ray crystallographic analysis of the gabaculine-soaked crystal revealed similar unit cell parameters with those of the native enzyme. At this stage, the presence of the inhibitor in the enzyme cannot yet be confirmed. The structure determination of the complex crystal will be carried out by difference Fourier synthesis in comparison to the native enzyme.

3. Flavin-Containing Oxygenase Enzymes

Dr. Pimchai Chaiyen's laboratory focuses on studying the mechanistic enzymology of some selected oxygenase enzymes, especially those having flavins as a cofactor. Flavoprotein oxygenases are present in all aerobic organisms and crucial for the endogenous metabolic pathways, as well as participating in the degradation of xenobiotic compounds. Many aromatic compounds in soil are degraded by these flavoprotein oxygenases, so these enzymes may be useful for environmental remediation. Two enzymes involved in degradation of aromatic compounds are being studied.

3.1 2-Methyl-3-hydroxypyridine-5-carboxylic acid (MHPC) oxygenase

2-Methyl-3-hydroxypyridine-5-carboxylic acid (MHPC) oxygenase (MHPCO) is a flavoprotein catalyzing the oxygenation of a substrate MHPC to α -(N-acetylaminomethylene) succinic acid. In our study, the mechanism of the reaction was investigated by replacing the natural cofactor FAD with FAD-analogues. FAD-analogues employed were those with various substituents at the 8-position of the flavin ring system. Substituents with electron withdrawing effects were 8-Cl- / 8-CN- groups and substituents with electron donating effects were 8-NH₂- / 8-CH₃-O- groups. Natural FAD was removed from the native holoenzyme, and the resulting apoenzyme was reconstituted with FAD-analogues. The thermodynamic and catalytic properties of the reconstituted enzymes were investigated and found to be similar to the native one. A substrate (MHPC) bound to the reconstituted enzyme less tightly than to native enzyme. Dissociation constants for binding of substrate analogue (5HN) and the reconstituted enzymes vary in the range of 1-20 μ M and these values indicate that the reconstituted enzymes still bind well with ligands. All of the reconstituted enzymes have hydroxylation efficiency higher than 85 %, with the 8-Cl- / 8-CN- MHPCO having higher hydroxylation efficiency than 8-NH₂ / 8-CH₃O- MHPCO. Redox potential values of the reconstituted enzymes were measured, and found to be more positive than the values of free FAD-analogues. These values also correlate well with the electronic effects of 8-substituents. The electronegativity of 8-substituents also affected the stability of anionic semiquinone during the enzyme reduction.

The pre-steady state kinetics of MHPCO was investigated using stopped-flow spectrophotometry. Results have shown that the rate of the hydroxylation was 2.0, 1.5, 1.3, and 0.9 s⁻¹ for MHPCO reconstituted with 8-CN, 8-Cl, 8-OCH₃, 8-CH₃ -FAD respectively. We have tested to see if MHPCO uses the electrophilic aromatic substitution mechanism, by correlating the rate of the hydroxylation step with the stability or the ability to stabilize the negative charge of FADHO⁻, the flavin leaving group if FADHOOH was the electrophile. The calculated values of pK_a of FADHOH, the HOMO of FADHO⁻ which represents another method of estimating pK_a of FADHOH, and the difference in heat of formation of FADHO⁻ and FADHOOH were used as the values to represent the relative stability of the FADHO⁻ leaving group. The result showed that the better stability of FADHO⁻, the better rate of hydroxylation.

In summary, correlation between the rate constant and electronic effect has shown that in the reductive half-reaction of MHPCO, the rates of flavin reduction by NADH could be described as a parabolic relationship with the redox potential values of the reconstituted enzymes, which is consistent with the Marcus electron transfer theory. Studies of the oxidative half-reaction of MHPCO revealed that the oxygenation reaction of MHPCO occurs via an electrophilic aromatic substitution mechanism analogous to the mechanisms for parahydroxybenzoate and phenol hydroxylases.

3.2 *p*-Hydroxyphenylacetate hydroxylase (HPAH)

p-Hydroxyphenylacetate (HPA) hydroxylase (HPAH) was purified from *Acinetobacter baumannii* and shown to be a two-protein component enzyme. The small component (C₁) is the reductase enzyme with a subunit molecular weight of 32 kDa. C₁ alone catalyzes the HPA-stimulated NADH oxidation without hydroxylation of HPA. C₁ is a flavoprotein with FMN as a native cofactor but can also bind to FAD. The large component (C₂) is the hydroxylase component that hydroxylates HPA in the presence of C₁. C₂ is a tetrameric enzyme with a subunit molecular weight of 50 kDa and apparently contains no redox center. FMN, FAD, or riboflavin could be used as coenzyme for hydroxylase activity with FMN showing the highest activity. Our data demonstrated that C₂ alone was capable of utilizing reduced FMN to form the product 3,4-dihydroxyphenyl-acetate. Mixing reduced flavin with C₂ also resulted in the formation of a flavin intermediate, that resembled a C(4a)-substituted flavin species indicating that the reaction mechanism of the enzyme proceeded via C(4a)-substituted flavin intermediates. Based on the available evidence, we conclude that the reaction mechanism of HPAH from *Acinetobacter baumannii* is similar to that of bacterial luciferase. The enzyme uses a luciferase-like mechanism and reduced flavin (FMNH₂, FADH₂, or reduced riboflavin) to catalyze the hydroxylation of aromatic compounds, which are usually catalyzed by FAD-associated aromatic hydroxylases.

The N-terminal sequences for each enzyme component were successfully determined. Although the data obtained from the initial enzyme preparation showed a mixed sequence, the problem was shown to be due to proteolytic cleavage. *Acinetobacter baumannii* appeared to have a high level of proteases that could digest the N-terminal part of C₁ and C₂ protein. With a low concentration of protease inhibitors included, as in the initial enzyme preparation procedure, the preparation yielded the enzyme with some partial cleavage by protease, causing a mixed sequence when analyzed by Edman degradation. With the new improved enzyme preparation method, the N-terminal sequences of C₁ and C₂ could be determined by Edman degradation, yielding a 40 residue sequence for C₂ and 26 residue sequence for C₁. Amino acid residues 1-34 of C₂ were used to design degenerate primers for use in PCR with the genomic DNA of *A. baumannii* as a template. A 102 bp product was obtained from the PCR and the nucleotide sequence of this PCR product was shown to correspond to the N-terminal sequence of C₂ enzyme. This 102 bp sequence will be used as a nucleotide probe to screen the *A. baumannii* library for the gene corresponding to enzyme C₂.

The N-terminal sequences of C₂ and C₁ were also used to design the degenerate primers to amplify part of the gene coding for C₂ and C₁. Using this approach, we were able to conclude that C₂ gene was located upstream of C₁ gene. The sequence obtained from such PCR product was used to design another nucleotide probe to screen for C₁ gene. Results have shown that we have obtained full-length of the genes for both C₂ and C₁ by using library screening method and two of nucleotide probes. Sequence analysis

has shown that the unique broad flavin specificity of C_2 is paralleled by the lack of homology to oxygenases in the same class with specificity for reduced FAD. Although C_1 presumably has a domain for binding flavin and NADH, similar to other reductases of the same class, it also contains an extra unique C-terminal half, which may be responsible for HPA-stimulation of NADH oxidation.

We have expressed both C_1 and C_2 genes in an *E. coli* system. Results have shown that most of the enzyme was expressed in the soluble form. Protein purification process was carried out and yielded the purified C_1 (about 150 mg) and the purified C_2 (about 100 mg) from 3.6 L cell culture. We have characterized steady-state kinetic properties of the purified recombinant HPAH and compared these properties to the native enzyme. Results have shown that the recombinant enzymes have similar K_m and k_{cat} values when compared to the native one. We have investigated the binding of FMN to C_1 by using static titration and pre-steady state kinetics. The K_d for binding of FMN to C_1 was determined to be 0.02 μ M, and results from stopped-flow spectrophotometry has indicated that $k_{on} = 1.7 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ and $k_{off} = 0.014 \text{ s}^{-1}$.

We investigated the ability of C_2 to use various reduced flavins for hydroxylation by measuring the amount of the product 3,4-dihydroxyphenylacetate (DHPA) formed when different limited amounts of reduced FMN, FAD, and riboflavin were provided to C_2 . A solution containing HPA and various concentrations of flavin was placed in the main body of an anaerobic cuvette and the recombinant C_2 was placed in the side arm. The cuvette was made anaerobic and the flavin was then photo-reduced by exposing the cuvette to high-intensity visible light. C_2 in the side arm was then mixed with the reduced flavin and the inlet was opened to admit air. The solution in the cuvette was analyzed for DHPA by using the HPLC method. Results showed that the amount of DHPA formed was directly dependent on the amount of reduced FMN, FAD and riboflavin provided, indicating the use of these flavins by C_2 . This also indicated that reduced FMN and FAD interacted with C_2 equally well since DHPA was produced in about 80% of reduced FMN and FAD provided while reduced riboflavin served as a substrate for C_2 in lesser extent. This result is consistent with our previous observation in native enzyme that all of reduced FMN, FAD, and riboflavin were used by C_2 to hydroxylate HPA (Chaiyen, et al., 2001). To our knowledge, C_2 is the only oxygenase component of the enzyme in the class of two-protein component aromatic hydroxylase capable of using three forms of flavin for hydroxylation. Previous studies have shown that the oxygenase components of HPAH from *E. coli*, of phenol hydroxylase from *B. thermoglucosias* A7, and of chlorophenol-4-monooxygenase from *B. cepacia* AC1100 specifically used reduced FAD for hydroxylation reaction.

In conclusion, HPAH from *A. baumannii* represents a novel prototype of enzyme in the class of two-protein component aromatic hydroxylases. The expression of C_1 and C_2 in *E. coli* system allows the enzymes to be obtained in reasonable amount for the future investigation on enzyme structure and function.

4. Design of Chemical Proteinases

Reagents that can specifically cleave the protein backbone (chemical proteases) can be useful in chemistry and biology. Chemical proteases can be useful for manipulation of proteins, for example, for structure-activity studies of proteins and their structural domains, for study of ligand-biomolecule interactions, for design of new therapeutic agents, and for converting large proteins into smaller fragments amenable for sequencing. Light has been proposed as a reagent for protein cleavage reactions, since it has advantages over thermal reactions, namely: 1) photoreactions provide a sharp control for initiation and termination of the reaction; 2) specific chromophores can be selectively activated by controlling the wavelength of excitation, minimizing side reactions; 3) visible light is one of the least toxic reagents; and 4) one less reagent needs to be removed, since light is dissipated during the reaction.

Site specific photocleavage of proteins with high efficiency, using probes based on a pyrenyl chromophore coupled to natural amino acids, or short peptides, has been achieved. Molecular modeling studies using the X-ray crystal structure of lysozyme suggested that the environment of the cleavage site is quite hydrophobic with some ionic residues located near the binding site. Salt bridge formation between the probe site chain and charge residues in the protein was proposed to be the binding mechanism. Such binding interaction is expected to have a strong influence on the binding selectivity and possibly on the cleavage properties. The role of the ionic group of the probe side chain on the binding properties and protein photocleavage is examined in Apinya Buranaprapuk's laboratory. Two new bifunctional pyrenyl probes, L-phenylalanine-4(1-pyrenyl)butyramine chloride (Phe-L-Py) and L-phenylalanine-4(1-pyrenyl)methyramine chloride (PMA-L-Phe) (Figure 14), have been synthesized and used to test this hypothesis.

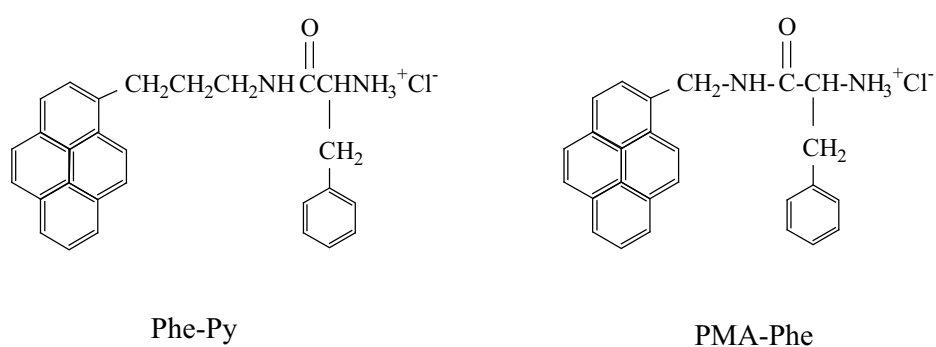


Figure 14: Structures of L-phenylalanine-4(1-pyrenyl)butyramine chloride (Phe-Py) and L-phenylalanine-4(1-pyrenyl)methyramine chloride (PMA-Phe)

4.1 Synthesis of new bifunctional pyrenyl probes

Synthesis of L-phenylalanine-4(1-pyrenyl)butyramine chloride (Phe-L-Py) was performed by dissolving 4(1-pyrenyl)butyramine chloride (PBAC) and N-t-Boc-L-phenylalanine (excess) in 50% acetonitrile/THF, 5% sodium bicarbonate. Dicyclohexylcarbodiimide (DCC) was then added into the solution, and the solution was stirred at room temperature for 1.15 hr. Water was added to stop the reaction. Then, the product was extracted with methylene chloride (CH_2Cl_2). Methylene chloride layer was separated and washed with 5% sodium bicarbonate and water twice. The purification was achieved by using column chromatography (silica gel). The N-t-Boc group of N-t-Boc-Phe-Py was removed by stirring N-t-Boc-Phe-Py in 50% TFA/ CH_2Cl_2 , 5% triisopropylsilane for 10 minutes. All solvents were evaporated to dryness. The product is light yellow in color with the yield of 90%.

Synthesis of L-phenylalanine-4(1-pyrenyl)methyramine chloride (PMA-L-Phe) was performed in the same direction, but 4(1-pyrenyl)methyramine chloride (PMAC) was used instead of PBAC. The products have light yellow color with the yield of ~90%.

The structures of these two pyrenyl probes were confirmed by UV, Fluorescence and NMR spectroscopy. UV spectra of Phe-L-Py and PMA-L-Phe show highest peak (in the range of 300-400 nm) at 344 nm. Both pyrenyl probes have emission spectrum at $\lambda_{\text{max}} = 378$ nm. ^1H NMR for Phe-L-Py (400 MHz, d_6 -acetone): 8.0-8.4 ppm (9H), 7.3 ppm (1H), 7.1-7.2 ppm (5H), 4.0 ppm (1H), 3.2-3.5 ppm (6H) 2.8 ppm (2H), 2.7 ppm (2H). ^1H NMR for PMA-L-Phe (400 MHz, CDCl_3): 7.9-8.3 ppm (9H), 7.2-7.3 ppm (5H), 5.2 ppm (2H), 4.0 ppm (1H), 3.7 ppm (1H), 3.3-3.4 ppm (2H), 2.1 ppm (2H).

4.2 Binding studies with proteins

4.2.1 With L-phenylalanine-4(1-pyrenyl)methyramine chloride (PMA-L-Phe)

Addition of lysozyme (0-10 μM) to a solution of PMA-L-Phe (2 μM) results in a small hypochromism (1.1% at 342 nm). The absorption peak positions are unchanged. In contrast, the binding of PMA-L-Phe to BSA results in red shifts of its absorption peak positions with isosbestic points at 338, 342.5 nm, and hyperchromism. The red shift of the absorption spectra suggests a hydrophobic environment surrounding the pyrenyl chromophore in BSA, and the hypochromism is attributed to partial stacking of the aromatic residues of the protein with that of the aromatic pyrenyl chromophore. Analysis of the absorption spectra data indicates a binding constant of $3.3 \pm 0.3 \times 10^5 \text{ M}^{-1}$ for PMA-L-Phe/lysozyme, and $7.8 \pm 0.2 \times 10^5 \text{ M}^{-1}$ for PMA-L-Phe/BSA.

The fluorescence spectra of PMA-L-Phe (0.7 μM) recorded in the presence of increasing lysozyme concentrations (0, 2, 4, 6, 8, 10 μM) indicate

weak quenching of the pyrenyl fluorescence by the protein. In contrast, the fluorescence intensity of PMA-L-Phe (0.7 μ M) is dramatically enhanced by the addition of BSA (0-10 μ M, excitation at 343 nm) with 1 nm-red shift.

The CD spectra of PMA-L-Phe, in the absence and in the presence of the proteins are examined. The spectra are expected to be sensitive to the chiral environment of the probes. However, the CD spectrum of PMA-L-Phe (30 μ M), in the presence of BSA (50 μ M) or lysozyme (50 μ M), is nearly unchanged. Only a slight increase in the peak intensity is observed.

4.2.2 With L-phenylalanine-4(1-pyrenyl)butyramine chloride (Phe-L-Py)

The Phe-L-Py absorption spectrum shows self dimerization at concentrations above 2 μ M. The absorption spectrum of Phe-L-Py (0.7 μ M) is red shifted in the presence of BSA resulting in an isosbestic point at 344 nm and hyperchromism as observed with PMA-L-Phe. Only hypochromism is observed when Phe-L-Py is bound to lysozyme. The fluorescence spectra of Phe-L-Py (0.7 μ M) is dramatically enhanced by the addition of BSA (0-10 μ M, excitation at 344 nm). The fluorescence spectrum of Phe-L-Py at higher concentration (6 μ M) shows excimer formation of the probe centered at $\sim$ 480 nm. Further binding studies of Phe-L-Py has not been performed due to the subtle spectroscopic data.

4.3 Fluorescence quenching studies:

Quenching of the fluorescence intensity of the free and bound probes by Cobalt (III) hexammine chloride (CoHA) (0-1 mM) has been carried out to examine the access of the pyrenyl chromophore to the aqueous solution. Quenching constants can be calculated using the Stern-Volmer equation: ($I_0/I = 1 + K_{SV} [Q]$). The fluorescence quenching of PMA-L-Phe (2 μ M) by CoHA, recorded in the absence and presence of BSA or lysozyme (10 μ M), indicates almost the same quenching constant of $1.05 \times 10^3 \text{ M}^{-1}$.

4.4 Photochemical protein cleavage

The new probe, L-phenylalanine-4(1-pyrenyl)methyramine chloride (PMA-L-Phe) carrying a free amino terminus photocleaves lysozyme with high efficiency and specificity, while the cleavage of BSA seems to be negligible.

The mixture of protein (15 μ M), PMA-L-Phe (15 μ M) and $\text{Co}(\text{NH}_3)_6\text{Cl}_3$ (CoHA, 1 mM) was irradiated at 342 nm (at the pyrenyl absorption band) using a 150 W xenon lamp attached to a PTI model A1010 monochromator (provided by Dr. C. V. Kumar, University of Connecticut). UV cut-off filter (WG-345) was used to remove stray UV light. The irradiation time was varied from 0 to 30 minutes, and the cleavage of proteins was visualized in gel electrophoresis experiments (SDS-PAGE; 12% gel for lysozyme and

8% gel for BSA). Dark control samples were prepared under the same conditions, as described above, except that the solutions were protected from light.

Lysozyme is efficiently cleaved by PMA-L-Phe with specificity and high efficiency (57% yield), as shown in Figure 15. Photocleavage of lysozyme results in at least two new fragments of molecular weights 11,000 and 3,000 (lanes 3, 4, and 5). No products are produced in the absence of light (lane 2), or CoHA (lane 6) or PMA-Phe (data not shown). The probe, CoHA, and light, therefore, are essential for the protein photocleavage.

Photocleavage of BSA by PMA-L-Phe results in two fragments (Mol. Wt. ~40,000 and ~28,000 (data not shown), but the photocleavage yields of the fragments obtained from PMA-L-Phe/BSA are quite low (less than 10%). Therefore, these results show dependency of the ionic group at the probe site chain on the cleavage efficiency and specificity.

The cleavage of both BSA and lysozyme seems to be negligible when Phe-L-Py is used for the photoreaction instead of PMA-L-Phe.

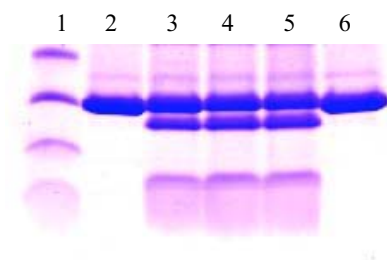


Figure 15: SDS-PAGE of lysozyme. Lane 1: M.W. protein makers; Lanes 2-5: lysozyme (15 μ M), PMA-L-Phe (15 μ M) and CoHA (1 mM). Lane 2 is the dark control while samples in lanes 3-5 were exposed to 342 nm radiation (10, 20, and 30 minutes, respectively). Lane 6 contained lysozyme and PMA-L-Phe (irradiated at 342 nm for 30 minutes).

4.5 Western blot transfer and amino acid sequencing of cleaved fragments

The separated protein fragments on 12% SDS-polyacrylamide gel were transferred to PVDF membrane with a current of 60 mA for 1.5 hours using the semi-dry system (BIORAD) using CAPS buffer, pH 10.5. The transferred protein fragments on PVDF membrane were stained with Coomassie brilliant blue (0.1% Coomassie brilliant blue R-250 in 40% methanol and 1% acetic acid). The desired bands were cut and sent for N-terminal amino acid composition analysis (Midwest Analytical, Inc., MO, USA).

N-terminal sequencing of the 11 kDa fragment indicated the residues KVFGR, the five N-terminal residues of lysozyme, so these residues have not been altered by the

cleavage chemistry. The newly formed N-terminus of the 3 kDa fragment was amenable to the sequencing chemistry, and it indicated the residues VAWRN, where the Trp was modified in some way with an altered retention time. The observed modification of the Trp residue provides additional evidence that the probe binds in this region of the protein, and induces protein cleavage at an adjacent site. The sequencing data indicated that the 11 kDa band was from the N-terminus of lysozyme, and that the 3 kDa band has been identified as the C-terminal fragment. KVFGR). In addition, the 3 kDa band also indicated a minor product (<5%) with the N-terminal residues KVFGR. The result indicated another minor cleavage site, but there was no indication of similar product in the 11 kDa band. This observed cleavage site with PMA-L-Phe (Trp108-Val109) is identical to that reported for Py-L-Phe (the probe carrying a free carboxyl terminus), and the structural changes introduced in PMA-L-Phe did not make a significant impact on the location of the cleavage site. These results clearly show that the cleavage of lysozyme is not site-specific but showed a high selectivity. This loss of specificity, but enhanced cleavage yield, clearly indicates the strong role of the side chain in the protein photocleavage. While the details of this role are yet to be established, the improved photocleavage yields will be of practical importance. The comparison of photocleavage yield is shown in Table 8.

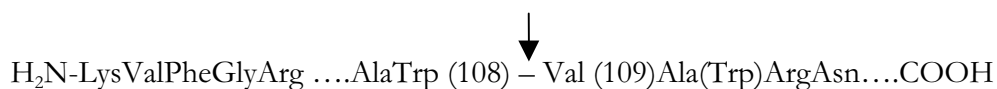


Figure 16: Cleavage pattern of lysozyme from the reaction with PMA-L-Phe.

Table 8: %Photocleavage yield of BSA and lysozyme.

Probe	Protein	Time of irradiation (min)	%Yield
Py-L-Phe	BSA	20	21
	Lysozyme	10	35
PMA-L-Phe	BSA	20	<10
	Lysozyme	10	57
Phe-L-Py	BSA	20	-
	Lysozyme	10	-

5. Isolation and Characterization of a Sericin-specific Protease for Use in Degumming of Silk

Silk consists mainly of fibroin, the silk fibers, and sericin, the gummy proteins holding the fibers together. Sericin causes hardness of silk-fiber texture and difficulty in the dyeing of silk. Therefore, it is necessary to remove sericin using the process called

silk degumming. Degumming of silk traditionally requires alkaline solutions and consumes much water and energy. In addition, quality control of the process is difficult, depending largely on the skills of the workers. A prolonged process increases the possibility of fibroin being damaged although it can remove more sericin. Dr. Pramvadee Y. Wongsangchantra is trying to isolate a protease that would specifically digest sericin, without digesting fibroin. This protease may form the basis of an alternative degumming of Thai silk using milder, less destructive conditions than the traditional one.

5.1 Soil enrichment and selection of bacteria

According to ecophysiological enrichment, we have used soil and wastewater from various places especially in the area of silk industry as the sources of microorganisms to be screened. To confine to the bacteria that can initiate the degradation of silk proteins at 37°C and pH 7, cycloheximide was used in our selective medium, of which small pieces of silk cocoon or raw silk (un-degummed silk thread) serve as a sole carbon and nitrogen source. In addition, successive subcultures were performed whenever changes in appearance of silk cocoon or raw silk were observed. The silk cocoon was changed from square pieces (or from hard threads if raw silk was used) into loose fibers and finally densely packed spherical forms, which were probably due to the effect of rotary shaker. Several batches of enrichment indicated significant changes in silk cocoon appearance even the last (9th) subculture (Figure 17, 18). From these mixed cultures, pure cultures have been isolated on the agar medium supplemented with 0.3 % skimmed milk as the sole carbon and nitrogen source. It was possible to differentiate different colony morphologies and the general activity of extracellular protease among the isolates. In summary, enrichment with consecutive subcultures was used as a strategy in primary screening of our screening program.

In secondary screening, individual cultures have been investigated using different approaches. One approach is cultivation of individual cultures in the same conditions as those of enrichment cultures to select whether they can cause a significant change in silk cocoon appearance. For example, pure bacterial strains KC001-009 showed different degrees in the ability to alter the appearance of silk cocoon. The order of the most effective strain to the least effective strain is KC007, 009, 008, 003, 004, 005, 006, 001, and 002, respectively. This result was observed in comparison with a control, within 4 to 6 days of cultivation to the end of experiment, which was 30 days (Figure 19).



Figure 17: Cocoon or raw silk buried in soil GR, H1C, and MR for 3 months (Dec 2003 – Feb 2004).

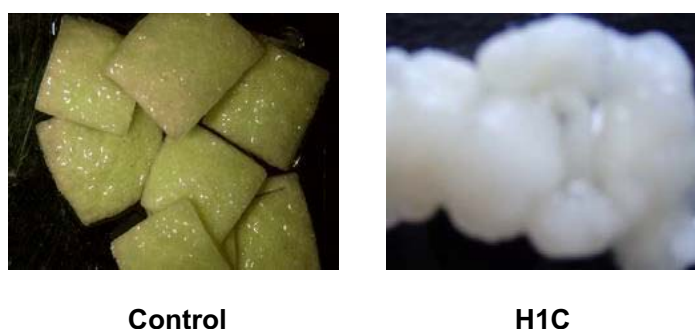


Figure 18: Appearance of 0.5x0.5 cm-whole shell of silk cocoon, after incubation at pH7, 37°C for 7 days in the 4th enrichment culture derived from the 3rd subculture of original H1C soil and its control (without inoculation)



Figure 19: Appearance of 0.5x0.5 cm-whole shell of silk cocoon after incubation at pH7, 37°C for 30 days without inoculation (control) and with KC003 or KC007, or KC008 inoculation.

5.2 Testing cell-free cultures of individual isolates for digestability

For another approach in secondary screening, individual cell-free supernatants (CFS) from cultures have been tested for protease activity at various cultivation times by using radial diffusion in agar assay (RD). It was possible to explore the digestability on 3 different substrates (sericin, fibroin, and skimmed milk) in such assay, since we have optimized conditions for extraction of sericin, and dissolution of fibroin. These

preparations provide a soluble form of either sericin or fibroin, which is more vulnerable to enzymatic degradation than the native or solid form. However, they are easier to be prepared in a homogeneous manner. A result of digestability on 3 different substrates (sericin, fibroin, and skim milk) evaluated by RD assay is showed in the following table. The digestability value (DV) is defined as the diameter of clear zone divided by the diameter of well, and 1DV means no digestability under the conditions evaluated. In addition both LC_1_1 and LC_1_6 could also cause a significant change in raw silk appearance, in which LC_1_6 made the raw silk whiter than LC_1_1. These results help to define the optimal cultivation time for protease production under the conditions used.

Table 9: Digestability of sericin, fibroin or skimmed milk by cell-free culture supernatants of bacteria

CFS	Digestability value on		
	sericin	fibroin	skim milk
9-day old LC_1_1 (with raw silk in culture)	3.7	3.3	1
9-day old LC_1_1	1	1	2.5
9-day old LC_1_6 (with raw silk in culture)	3.3	1.3	1
9-day old LC_1_6	2	1	2.1

5.3 Production, purification, and characterization of extracellular protease

In our secondary screening for a bacterial protease with preference for sericin from our culture collection, we have applied radial diffusion (RD) in agar gel plus thin-layer enzyme assay (TEA) technique. In this technique, enzyme produced by bacteria can diffuse through agar medium and then hydrolyze the coated protein at the bottom revealing the hydrophobic surface of polystyrene petri dish. The hydrophobic surface can be visualized by removal of agar and condensation of water vapor. According to Wikstrom, we can apply sericin or fibroin coated on the surface of polystyrene petri dish to serve as substrate while a solid-culture medium is placed on the top layer. A standard procedure was optimized for preparation of such plates, that includes the amount and the age of master culture, and incubation time. This procedure was focused to make sure that our method is reproducible.

Pure cultures, GR_4_4, MR_4_2, and MR_4_3, did not show a significant change in silk cocoon appearance. However, different media showed different results for the same bacterial strain examined. We found conditions that GR_4_4, MR_4_2, and MR_4_3 showed their clear zones on sericin coated plate but not on fibroin coated plate. In addition, a mix culture of GR_4_4 mutants (in the same medium used in the RD plus TEA assay) showed a significant change in the appearance of silk cocoon. This mutant will be isolated. Although KC007 could alter the appearance of silk cocoon, its cell-free culture showed minute amount of its digestability on casein in the RD assay (2 days for

the small clear zone appearance). This unexpected result was possibly due to the improper sampling time, small amount and unstableness of enzyme, as well as its substrate preference. Upon changing media in the RD plus TEA, KC007 showed a clear zone on sericin coated plate but not on fibroin coated plate.

In addition, during optimization of the RD plus TEA for our screening, we found that Alcalase[®], a commercial protease, could hydrolyze fibroin better than sericin under the condition examined. This condition showed sharp clear zones resulting from Alcalase[®] treatment. The area of the clear zone on the sericin coated surface (3 cm in diameter including well diameter) was less than that on the fibroin coated surface (4 cm in diameter including well diameter) when the same amount of Alcalase[®] was used for treatment (Figure 20).

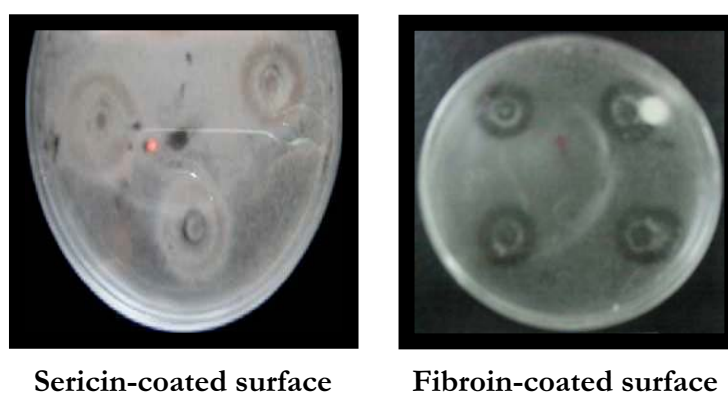


Figure 20: The enzyme-affected surfaces visualized by condensation of water vapor after agar removal. The surface of polystyrene Petri dish was coated with sericin or fibroin. Each petri dish was tested with 5.8 mU of Alcalase[®] per well (10 μ l) at pH 7.0, 37°C, 24 hrs.

From the results of secondary screening, we will further characterize the proteases to confirm this digestability using a zymogram assay, with the procedure which has been modified for our application. The zymogram can be used for purification of small amount of interesting enzyme for further N-terminal amino acid sequencing and then cloning.

When 3 approaches of secondary screening of our screening program are compared, the benefit of this RD plus TEA technique over the observation of significant changes in raw silk appearance (as in the enrichment technique) or the RD assay alone is its high throughput in screening. This benefit is considered under the basis of the following factors. Firstly, testing the substrate preference of the protease is independent, but continuous from its production. Secondly, the sericin or fibroin homogeneously coated on the surface of polystyrene petri dish is likely to have a structure more closely similar to that of raw silk thread. Finally, smaller amounts of substrate are needed in this technique. Therefore, by varying culture media, we can also use this technique for

preliminary optimization (e.g. selection of a suitable media) for bacterial growth and protease production.

In parallel to the optimization, the gene encoding for the extracellular proteases from the isolates will be cloned (as mentioned above) to allow production of the protein in quantity. Purification and characterization of the protease(s) will be attempted. This will lead to optimization of conditions and formulation of the enzymes for effective usage in silk-degumming process.

6. Protein Changes In Human Disease

Prof. Jisnuson Svasti's laboratory at the Chulabhorn Research Institute, together with Dr. Chantragan Srisomsap, Dr. Rudee Surarit, and Dr. James Ketudat-Cairns, is interested in study of protein changes in various human diseases. This is because many diseases result from abnormalities in protein structure or expression, and study of the relationship between the functional or physiological abnormality with the nature of the amino acid interchange can provide insight into structure-function relationships in proteins. Current areas of interest are the abnormal hemoglobins, inborn errors of metabolism and protein changes in cancer. Although many studies have been performed in these areas elsewhere in the world, studies in Thailand are still rather limited. Our studies provide information on the defects occurring in the Thai population, and offer medical practitioners the opportunity of collaborating to study diseases of their interest at the molecular level.

6.1. Characterization of Abnormal Hemoglobins Found In Thailand

Our present studies seek to characterize new abnormal Hb not previously found in Thailand, through cooperation with Professor Suthat Fucharoen at the Thalassemia Research Center, Institute of Science and Technology, Mahidol University. These studies are done at both the protein level using tryptic peptide mapping of abnormal chain by HPLC, amino acid analysis, and automated peptide sequencing, and at the DNA level, by amplification of globin genes from leukocytes, followed by automated DNA sequencing. Moreover, since Thailand has high incidence of Hb E [β 26 Glu-Lys], α -thalassemia and β -thalassemia, abnormal Hb can often be found in compound heterozygosity with these conditions in Thailand unlike in most countries, so that the effect of such associations on the hematological and clinical profile is of interest.

6.1.1 Hb Kodaira [β 146 (HC3) His→Gln (CAC→CA4)]

Hb Kodaira [β 146 (HC3) His→Gln (CAC→CA4)], was characterised for the first time in Thailand. This Hb variant shows increased oxygen affinity and was first described in a Japanese male in 1992. Recently, while our study was in progress, a second report described the same phenotypic variant Hb Kodaira II, but with a different

nucleotide change (CAC→CAG) at codon 146 of the β -globin gene. We have now found the first case of Hb Kodaira II in Thailand, in a 31-year old healthy Thai female (S.P), who came for a routine check up. She had the following hematological profile at the steady state: Hb 14.7 g/dL (normal 13.8 ± 0.7); Hct 47% (normal 42.5 ± 1.7); MCV 94 fL (normal 92.5 ± 3.7); MCH 29 pg (normal 30.6 ± 1.9); MCHC 31 g/dL (normal 32.6 ± 0.9); RBC $5.07 \times 10^9/L$ (normal 4.5 ± 0.3). Analysis of the blood sample by automated high performance HPLC (VariantTM System; Bio-Rad Laboratories, Hercules, CA, USA) revealed an abnormal Hb present at a level of 45.4%, eluted earlier than Hb A, with Hb F and Hb A₂ being in the normal range. The abnormal Hb could not be separated from Hb A by cellulose acetate electrophoresis, pH 8.6. However, the abnormal β -globin chain migrated slightly slowly toward the cathode than normal β -globin chain on cellulose acetate electrophoresis in acid-urea-mercaptoethanol, at pH 6.5. In addition, the abnormal Hb had slightly lower pI than Hb A on isoelectric focusing (IEF; pH range 6-8).

The characterization of the variant was performed at the protein and DNA level. Since denaturing cellulose acetate electrophoresis at pH 6.5 separated the β -globin chains, the abnormal Hb was purified by CM-cellulose (CM-52, Whatman Co.) chromatography in 0.03 Bis-Tris HCl, pH 6.1 with an NaCl gradient. The abnormal Hb was eluted out early, completely separated from Hb A. Globin chains, prepared by the acid-acetone method were separated by reversed phase HPLC (Vydac C₄) to yield abnormal β -chain and normal α^A chain. The $\beta^{\text{Kodaira II}}$ -chain and normal β^A -chain were compared by tryptic peptide mapping on an HPLC (510, Waters, Milford, MA, USA) using a C₁₈ column (Aquapore OD-300, 7 micron, 250 x 4.6 mm; Applied Biosystems, Foster City, CA, USA) in an ammonium acetate buffer, pH 5.7, and acetonitrile. Peptide maps were similar except that peptide $\beta^{\text{Kodaira II}}$ -T15 (6.13 min) was eluted out faster than peptide β^A -T15 (10.13 min). Peptide $\beta^{\text{Kodaira II}}$ -T15 gave the sequence Tyr-Gln on a protein sequencer (ABI 473A; Applied Biosystems), confirming the mutation [β 146 (HC3) His→Gln].

Since the amino acid substitution [β 146 (HC3) His→Gln] has been described both as a CAC→CAA mutation in Hb Kodaira and as a CAC→CAG mutation in Hb Kodaira II, DNA analysis was also performed. Genomic DNA was extracted from white blood cells of the proband by phenol-chloroform extraction. The whole β -globin gene was amplified using the primers 5'-CCTAAGCCAGTGCCAGAAG-3' and 5'-AACTGAGTGGAGTCAAGGCT-3'. PCR products were purified by using QIAquickTM Gel Extraction Kit (Qiagen Inc., Valencia, CA, USA). Direct DNA sequencing on an ABI 310 Prism DNA sequencer (Applied Biosystems) demonstrated that the β 146 codon in the proband was heterozygous for the normal CAC (His) and the abnormal CAG (Gln) sequence. This indicates that the proband was heterozygous for Hb Kodaira II, as found recently by So et al.

Hb Kodaira has a slightly lower P_{50} than Hb A and a significantly decreased Bohr effect, since the imidazole group of the C-terminal histidine in β -globin chain normally forms an internal salt-link with the carboxylate of $\beta 94$ Asp. In this case, increased oxygen affinity did not lead to erythrocytosis, since the patient had renal failure and was anemic (Hb = 7.9 g/dL). However, in the second case, Hb Kodaira II, the patient was reported to have mild polycythemia, and also had slightly elevated Hb F (1.6%), which was presumed to be related to pregnancy. The proband, described here as Hb Kodaira II, also showed mildly elevated RBC and Hb level, but was otherwise in normal health.

The present studies confirm that the Hb Kodaira [$\beta 146$ (HC3) His $\rightarrow$ Gln] phenotype can result from two mutations, namely CAC $\rightarrow$ CAA in Hb Kodaira and CAC $\rightarrow$ CAG in Hb Kodaira II. All three cases described were found in Asians, but the evolutionary origin of these cases has not been investigated. The mutation is not detectable by the traditional technique of electrophoresis at alkaline pH, where the imidazole group of His146 is expected to be uncharged. However, it is possible, that with improved techniques of detection by HPLC and IEF, Hb Kodaira may be found in non-Asiatic populations.

6.1.2 Hb Hekinan [$\alpha(27)$; Glu-Asp] with α -Thalassemia

In this phase, two cases of compound heterozygosity for Hb Hekinan [$\alpha(27)$; Glu-Asp] and α -thalassemia were characterised. This electrophoretically silent hemoglobin (Hb) variant, Hb Hekinan ($\alpha 27$; Glu $\rightarrow$ Asp) was first observed in 46-year-old Japanese male, later in a Chinese-Black female from French Guyana, and in three Chinese from Macau. Recently, the mutation has also been reported in Thailand by Supan Fucharoen at Khonkaen University, in association with HbE ($\beta 26$; Glu $\rightarrow$ Lys) and deletional α -thalassemia 1, South East Asian Type (SEA Type). The lack of a charge change makes Hb Hekinan difficult to separate from Hb A by standard chromatographic techniques or by HPLC, causing problems in quantitation of abnormal Hb and in obtaining purified protein for characterization. However, Hb Hekinan and Hb A may be well separated by isoelectric focusing (IEF). We have studied two new unrelated cases of Hb Hekinan in association with α -thalassemia-1, with one case also having Hb E.

Both probands were accidentally found by routine screening. The first case, S.J., a 20-year-old Burmese woman, had the following hematological profile at steady state: Hb 11.2 g/dL; Hct 35%; MCV 68 fL; MCH 22 pg; MCHC 32 g/dL; RBC 5.1×10^6 /ml, while the second case S.S., an adult woman of Thai origin had the profile: Hb 10.2 g/dL; Hct 35%; MCV 83 fL; MCH 24 pg; MCHC 29 g/dL; RBC 4.3×10^6 /ml. Hemoglobin typing by cation exchange HPLC on the VARIANTTM instrument (Bio-Rad Laboratories, Hercules, CA, USA) revealed a major peak of an abnormal Hb poorly separated from Hb A in both probands. With S.S., a minor peak was also observed at the HbA₂/HbE region, indicating the presence of Hb E. Separation of the abnormal Hb is much better on IEF in the pH range 6-8. A distinct band of Hb Hekinan migrated slightly more

anodal than Hb A in both probands, with minor bands of oxidized Hb Hekinan and Hb A also being detectable in proband S.J. With proband S.S., bands were also detected in the Hb A₂/ Hb E regions, due to Hb E ($\alpha^A_2\beta^E_2$), as well as another band Hb X₁ presumed to be $\alpha^{\text{Hek}}_2\beta^E_2$, together with a faint unknown band Hb X₂.

Hb Hekinan was purified by DEAE-cellulose column chromatography, and, although Hb A and Hb Hekinan were not completely separated, several pools were taken and analyzed by IEF, allowing Hb Hekinan to be obtained free of contamination by Hb A. The α^{Hek} -chain and normal α^A -chain were purified and compared by tryptic peptide mapping, and an abnormal peptide α^{Hek} -T4 (23.57 min) was isolated with amino acid composition (Asx 1.12, Glx 2.29, Gly 3.06, Ala 3.50, Val 1.03, Leu 0.88, Tyr 0.74, His 0.82, Arg 1.04). Protein sequencing confirmed the $\alpha 27$ Glu→Asp, indicating Hb Hekinan.

Direct DNA sequencing indicated that in both probands, $\alpha 1^{\text{Hek}}$ gene had T as the third base of codon 27, instead of the G found in the $\alpha 1^A$ gene. In addition, gap-PCR analysis showed that both probands were heterozygous for α -thalassemia-1 (SEA type), where both the $\alpha 1$ and $\alpha 2$ genes are deleted. In the case of proband S.S., direct DNA sequence analysis of the β globin gene also showed a heterozygous change from GAG (Glu) to AAG (Lys) at codon 26 in exon 1, resulting in the replacement of glutamic acid at position 26 by lysine, and indicating the Hb E ($\beta 26\text{Glu-Lys}$) mutation. The location of the α^{Hek} mutation in the $\alpha 1$ locus is similar to two earlier reports in Asiatic populations, but differs from the $\alpha 2^{\text{Hek}}$ mutation associated with α -thalassemia-2 (-3.7), reported in a survey of α globin gene variants.

The Hb Hekinan mutation does not cause functional abnormalities, since it conserves charge properties and occurs at the interface between α and β chains. Thus, simple heterozygotes showed normal hematology and had no pathological conditions. In addition, associations of α^{Hek} with other Hb mutations, including $\alpha 2^{\text{Hek}}\alpha 1^A/-^{3.7}\alpha 1^A$, as well as $--^{\text{SEA}}/\alpha 2^A\alpha 1^{\text{Hek}}$, β^A/β^E and $\alpha 2^A\alpha 1^A/\alpha 2^A\alpha 1^{\text{Hek}}$, β^A/β^E showed the hematological parameters expected for the equivalent genotypes lacking the α^{Hek} gene. In our study, both probands showed co-inheritance of the α -thalassemia-1 and $\alpha 1^{\text{Hek}}$ genes. Thus S.S. had the same $--^{\text{SEA}}/\alpha 2^A\alpha 1^{\text{Hek}}$, β^A/β^E genotype as in the previous report, while the unrelated proband, S.J. is the first case described with the genotype $--^{\text{SEA}}/\alpha 2^A\alpha 1^{\text{Hek}}$, β^A/β^A . Both also had hematological profiles similar to that expected for equivalent genotypes lacking the α^{Hek} gene. Our studies indicate that, while the α^{Hek} gene is a rare variant, it can be found associated in various combinations, with the β^E gene and with α -thalassemia in Thailand, due to their high frequency.

Early estimates of the content of Hb Hekinan in simple heterozygotes have ranged between 12.9-13.4% (Table 10), using IEF or HPLC systems providing good separation of Hb Hekinan and Hb A. Hb Hekinan level in the genotype $\alpha 2^{\text{Hek}}\alpha 1^A/-^{3.7}\alpha 1^A$, β^A/β^A with three α genes, was found to be 28%. However, the commercially

available HPLC instrument does not provide adequate separation of Hb Hekinan and Hb A, so Hb Hekinan levels in complex associations were not reliable when quantitated by HPLC (Table 10). Poor separation leads to underestimation of Hb Hekinan compared to Hb A, since Supan Fucharoen et al could not detect Hb Hekinan in the proband with the genotype $--^{SEA}/\alpha 2^A \alpha 1^{Hek}, \beta^A/\beta^E$. Thus, we have also quantitated Hb Hekinan levels by IEF followed by densitometry, obtaining higher levels than that found when quantitated by HPLC (Table 10). Proband S.J., with genotype $--^{SEA}/\alpha 2^A \alpha 1^{Hek}, \beta^A/\beta^A$, had 42.6% Hb Hekinan, while proband S.S., with genotype $--^{SEA}/\alpha 2^A \alpha 1^{Hek}, \beta^A/\beta^E$, had 32% Hb Hekinan and 8 % of $\alpha^{Hek}_2 \beta^E_2$ band, making a total of 40 % (Table 10). However, both probands had α -thalassemia-1, so that $\alpha 1^{Hek}$ is one of only two α globin genes present. While the level of $\alpha 2:\alpha 1$ mRNA appears to be in the range of 1.5-2.8, the ratio of $\alpha 2:\alpha 1$ globin chains has been previously estimated as being 2-3, but has more recently been estimated as being only 1.19 for stable variants. The latter ratio would lead to % Hb Hekinan levels of 43.7 % in compound association of $\alpha 1^{Hek}$ with α -thalassemia-1, similar to estimates reported here using IEF and densitometry for quantitation.

6.1.3 Hb Kurosaki [$\alpha 7$ (A5) Lys-Glu] Found In Thailand

Hemoglobin (Hb) Kurosaki [$\alpha 7$ (A5) Lys-Glu] was first discovered in 1995 by Harano et al in a 70-year-old Japanese woman. As far as we know, no other cases have been reported, until the present paper, which describes a Thai male carrying this variant.. Cellulose acetate electrophoresis at pH 8.6 revealed Hb variant as a fast-moving band. Isoelectric focusing (pH range 6-8) of the hemolysate showed the presence of the abnormal Hb at a more anodic position than Hb A. On cellulose acetate electrophoresis in 6 M urea and 2-mercaptoethanol, pH 6.5, the proband also showed an α chain variant with more anodic mobility than α^A .

The hemolysate was fractionated by DEAE-column chromatography, and an abnormal Hb peak (16.9%) was found, eluting out after HbA. Abnormal Hb was treated with acid-acetone and fractionated by HPLC (510; Waters, Milford, MA, USA) on a Vydac C4 (The Separations Group, Hesperia, CA, USA) column to separate the α and β globin chains. Peptide maps of normal and abnormal α chains were compared by HPLC on a C18 column (Aquapore OD-300, 7 micron, 250x4.6 mm; Applied Biosystems, Foster City, CA, USA) in ammonium acetate buffer, pH 5.7 and acetonitrile. The peptide map of abnormal $\alpha^{Kurosaki}$ chain was similar to that of normal α^A , except for the lack of peptide α^A -T-1, which was instead replaced by the later eluting $\alpha^{Kurosaki}$ -T-1.

Table 10. Amounts of Hb Hekinan in Simple and Compound Heterozygotes

Method	Our data	Our data	Our data	Our data	Fucharoen, 2003	Fucharoen, 2003	Molchanova, 1994	Harano, 1988
Sample	IEF	IEF	HPLC	HPLC	HPLC	HPLC	HPLC	IEF
α globin genotype	S.S.	S.S.	S.S.	S.J.	Proband	Mother	Proband	Proband
	$\frac{SEA}{--} \alpha^A_{Hek} \alpha^A_{Hek}$	$\frac{SEA}{--} \alpha^A_{Hek} \alpha^A_{Hek}$	$\frac{SEA}{--} \alpha^A_{Hek} \alpha^A_{Hek}$	$\frac{SEA}{--} \alpha^A_{Hek} \alpha^A_{Hek}$	$\alpha^A \alpha^A / \alpha^A \alpha^A_{Hek}$	$\alpha^A \alpha^A / \alpha^A \alpha^A_{Hek}$	$\alpha^{Hek} \alpha^A / \alpha^A_{-3.7} \alpha^A$	$\alpha^A \alpha^A / \alpha^A \alpha^A_{Hek}$
β globin genotype	β^A / β^E	β^A / β^E	β^A / β^E	β^A / β^A	β^A / β^E	β^A / β^E	β^A / β^A	β^A / β^A
Hb Hekinan	32	26.5	27.1	26.5	0	27.9	12.9	
Hb A	40	50.3	63	53.9	59.4	?	84	
Hb X ₁ (α^{Hek}) ₂ (β^E) ₂	8.2	1.6		0.8				
Hb X ₂	3.7							
Hb E + HbA2	12.2	9.3	1.1	9.5	28		3.1	
Hb F		1.4	+	0.9	5.2		0.5	

The amino acid composition (Asp 1.74, Glu 1.19, Ser 0.74, Thr 0.83, Ala 1.08, Pro 1.07, Val 2.19, Leu 1.00, Lys 1.05) of abnormal peptide $\alpha^{\text{Kurosaki T-1}}$, determined by the Pico Tag method (Waters) showed the presence an $\alpha^{\text{T-1,2}}$ peptide, due to substitution of Lys by Glu or Gln. Automated protein sequence analysis of the abnormal α^{X} chain showed that lysine at position 7 is indeed replaced by glutamic acid, so that tryptic cleavage did not occur, leading to formation of the $\alpha^{\text{Kurosaki T-1,2}}$ peptide extending from residue 1 to 11.

The earlier patient showed normal hematological parameters, and the variant showed normal function and stability. This is consistent with the fact that the seventh position of α chain is located in an internal position of the A helix of the Hb molecule and is not involved in contacts with heme or between chains. Another case of a mutation at the $\alpha 7$ position, namely Hb Tatra ($\alpha 7(\text{A5}) \text{Lys} \rightarrow \text{Asn}$) found in a 72-year-old woman born in Czechoslovakia, also showed normal hematological parameters in the heterozygote.

6.1.1 Compound Heterozygosity for Hb D Punjab/ HbE

The proband P.S. was a woman with normal hematological data at the steady state: Hb 12.0 g/dL; Hct 38.0 %; MCV 84.1 fL; MCH 26.5 pg. Hemoglobin typing by HPLC showed: Hb F 0.9 %, Hb A₂ 33.5 %; abnormal Hb 59.0 %. Cellulose acetate electrophoresis of the hemolysate indicated that no Hb A was present and instead showed two major bands in the position of Hb A₂ (Hb E) and an unknown abnormal Hb running intermediate between Hb A and Hb A₂. The abnormal Hb was purified by DEAE-cellulose chromatography and the abnormal β -chains were isolated by HPLC on a Vydac C4 column. Comparison of the tryptic peptide maps of normal β^{A} and abnormal β^{PS} chains by HPLC revealed the presence of a new peptide $\beta^{\text{PS}}\text{T-13}$ (Glx 2.28, Thr 0.84, Ala 1.89, Pro 1.72, Tyr 0.64, Val 1.24, Phe 1.08, Lys 1.00). Amino acid sequence analysis revealed the mutation $\beta 121\text{Glu-Gln}$, which has previously been reported as Hb D Punjab or Hb D Los Angeles: $\beta^{\text{PS}}\text{-Tp13}$: ¹²¹**Gln**-Phe-Thr-Pro-Pro-Val-Gln-Ala-Ala-Tyr-Gln-Lys. Since Hb E was found by HPLC and on cellulose acetate electrophoresis, the proband P.S. was a compound heterozygote for Hb D Punjab and Hb E. This combination does not cause any adverse symptoms.

6.1.2 Compound Heterozygosity for Hb E/ Hb Tak

The proband K.R. was a female with normal Hb level, but slightly low MCV and MCH at the steady state, namely: Hb 13.7 g/dl; Hct 40 %; MCV 69 fL; MCH 23.7 pg. Hemoglobin typing by HPLC showed Hb F 1.9 %; Hb A₂/ Hb E 47.3%; abnormal Hb 46.4%. Cellulose acetate electrophoresis and isoelectric focusing indicated that Hb A was absent, replaced by two major peaks at the position of Hb A₂, intermediate between Hb A and Hb A₂. The hemolysate was

fractionated by DEAE-cellulose chromatography to yield the Hb E and the unknown abnormal Hb. These Hb were separately treated with acid-acetone and further fractionated by HPLC on a Vydac C4 column to separate chains. The abnormal β^E and β^{KR} chains were compared to normal β^A chains by tryptic mapping on HPLC. β^E chains showed two new peptides β^E -T3a (Asx 1.51, Glx 0.91, Gly 1.80, Val 3.07, Lys 1.00) and β^E -T3b (Gly 1.01, Arg 1.00, Ala 1.14, Leu 1.03), consistent with the mutation $\beta 26\text{Glu}\rightarrow\text{Lys}$, characteristic of Hb E. β^{KR} chains showed two peptides β^{KR} -T15a (Thr 1.03, Tyr 0.83, His 0.98, Lys 1.00) and β^{KR} -T15b (Asx 1.18, Ser 1.00, Ala 1.11, Tyr 0.74, Leu 2.67, Phe 2.00). Protein sequencer analysis showed the following sequences: β^{KR} -Tp15a: Tyr-His-**Thr-Lys** and β^{KR} -Tp15b: **Leu-Ala-Phe-Leu-Leu-Ser-Asn-Phe-Tyr**.

The results therefore indicate that proband K.R. was a compound heterozygote for HbE [$\beta 26(\text{Glu}\rightarrow\text{Lys})$] and Hb Tak [$\beta 147(+\text{AC})$]: (147)Thr-Lys-Leu-Ala-Phe-Leu-Leu-Ser-Asn-Phe-(157)Tyr-COOH.

6.2 Inborn Errors of Metabolism

Inborn errors of metabolism can cause severe clinical manifestations, such as mental retardation or developmental abnormalities. In general, they arise from deficiencies in enzymes in various metabolic pathways, such as the urea cycle, pathways for synthesis of specific amino acids, or mucopolysaccharide degradation. Such disorders may be due to mutations leading to dysfunctional or poorly functioning enzyme, or may result from lowered expression. Enzyme deficiencies are typically detected by an accumulation of the substrate of the enzyme reaction and/or a decrease level of metabolites, which occur after the enzyme reaction. We have established various analytical techniques for determining the levels of metabolites, such as amino acids, homocysteine and sulfur amino acids, and organic acids. The debilitating effects of many disorders, such as phenylketonuria, may be abolished or lessened if the disease is detected early, and appropriate intervention is carried out with suitable dietary programs. Thus, early detection by using appropriate diagnostic procedures is of much importance.

We are collaborating primarily with Dr. Pornswan Wasant (Division of Genetics, Department of Pediatrics, Siriraj Hospital), Dr. Vorasak Shotelersuk (Faculty of Medicine, Chulalongkorn University), and Dr. Suthipong Pangkanon (Queen Sirikit National Institute of Child Health, Ministry of Public Health), and Dr. Duangrurdee Wattanasirichaigoon (Faculty of Medicine, Ramathibodi Hospital) and Dr. Pranoot Tanpaiboon (Faculty of Medicine, Chiangmai University). Apart from the analytical studies to assist diagnosis of disease, we also study selected cases in greater depth, by analyzing the levels of enzymes suspected of being deficient using leukocytes or cultured fibroblasts from patients. Then, in cases where the gene is known, we design primers for preparing cDNA by RT-PCR or PCR, and then perform automated sequence of the cDNA to determine the mutation. We are presently concentrating on two major

diseases, methylmalonic acidemia and Hurler syndrome, and we will continue with these studies.

6.2.1 Amino Acid Analysis of Normal Thai Children of Different Ages

Many disorders of amino acid metabolism, urea cycle disorders, and organic acid disorders may be detected by amino acid analysis. Since amino acid analysis facilities were not readily available in Thailand, we established a simple methodology for analysis of plasma free amino acid levels by reverse-phase HPLC and pre-column derivatization with phenylisothiocyanate. Amino acid levels were determined in plasma samples from 57 normal children of various ages, obtained from Dr. Pornswan Wasant and Dr. Suthipong Pangkanon. Data for different age groups were obtained (0-6 months, 6-12 months, 1-3 years, 3-6 years, and 6-12 years), but the different age groups did not show major differences. Thus, the plasma amino acid levels in normal children are summarized for the overall age range of 0-12 years of age (Table 11), providing baseline values for detection of abnormal amino acid levels. Later, amino acid analysis facilities were established in Siriraj Hospital by Dr. Wasant, using our methodology. We have also assisted Ramathibodi Hospital in standardization of their amino acid analysis facilities. Thus our research has had impact on the development of amino acid analysis facilities in Thailand. As a result, we perform fewer plasma amino acid analyses now, but we still perform occasional analyses for various physicians in case of emergency or special need. In addition, amino acid levels in children with genetic diseases were studied, as follows.

6.2.2 Plasma Amino Acid Analyses in Children with Abnormal Protein Metabolism

Some 221 plasma amino acid analyses have now been performed for children suspected of having inborn errors of protein metabolism. The most common inborn errors were phenylketonuria and maple syrup urine disease, but other diseases were also found, including urea cycle disorders, homocystinuria, and tyrosinemia.

6.2.2.1 *Phenylketonuria*: Eight cases showed abnormally high levels of plasma phenylalanine (1 with 544 nmole/ml, and 7 with > 1,000 nmol/ml). This suggested that the patients are deficient in the enzyme phenylalanine hydroxylase, leading to improper metabolism of phenylalanine. The high level of phenylalanine is toxic to the central nervous system and causes brain damage. However, if detected early enough, proper treatment with a diet containing low phenylalanine can prevent mental retardation. From recent screening studies, the frequency of phenylketonuria in Thailand appears to be lower than that reported in many other countries.

6.2.2.2 *Maple Syrup Urine Disease*: Amino acid analysis of plasma from thirteen patients showed increases in the level of valine (8 patients with >300 nmole/ml), isoleucine (5 patients with >300 nmole/ml) and leucine (2 patients with 400-650 nmole/ml and 11 patients in the range 1,350-4300 nmole/ml). Comparison of the levels of the three amino acids to the levels found in normal children of the same age indicated greatly increased levels of leucine and moderately increased levels of valine and isoleucine. These results are consistent with the diagnosis of maple syrup urine disease (MSUD). Such patients can be treated by feeding with special milk formulation designed for maple syrup urine disease.

6.2.2.3 *Argininosuccinate synthetase deficiency*: The HPLC chromatogram of this patient showed abnormally high levels of citrulline (2,960 nmole/ml) and glutamine (1,640 nmole/ml) in plasma, compared to the levels found in normal children of the same age group. Based on these results, the patient was diagnosed as having argininosuccinate synthetase deficiency, the first such case discovered in Thailand.

6.2.2.4 *Homocystinuria*: Homocystinuria is an autosomal recessive disorder that arises from deficiency in the metabolism of methionine due to lack of the enzyme cystathionine beta-synthetase (CBS). This causes accumulation of methionine, homocysteine, and metabolites of these amino acids in the blood and urine. Two cases were studied.. The first case, W.J., was a 9 year old female, having nephrotic syndrome from the age of 18 months, who, was referred to Siriraj Hospital due to eye conditions, iridodonesis, anterior lens dislocation, and glaucoma. She was tall and had arachnodactyly. Urine analysis showed cystine/homocystine and reducing substances. The second case, B.J., was the younger sister, aged 6 years, also with an eye disorder, namely myopia, but without iridodonesis. Cystine/homocystine and reducing substances were present in plasma. Both also showed elevated levels of methionine (>500 nmole/ml) in plasma and high levels of homocysteine (>400 nmole/ml) in urine, agreeing with clinical diagnosis of homocystinuria. Both patients were treated with vitamin B₆, folic acid, and betaine.

Table 11 Plasma Free Amino Levels (nmol/ml) in Normal Thai Children (0-12 yrs)

Amino Acid Compounds	Mean (n=57)	Standard Deviation	95% Confidence Interval	
			Lower	Upper
Phosphoserine	19.6	21.5	13.9	25.3
Aspartic acid	6.8	5.0	5.4	8.1
Glutamic acid	171.3	140.7	134.0	208.7
α -Aminoadipic acid	1.7	2.5	1.0	2.3
Hydroxyproline	18.1	13.2	14.6	21.6
Phosphoethanolamine	1.6	3.2	0.8	2.5
Serine	116.6	45.0	104.7	128.6
Asparagine	80.7	27.2	73.5	87.9
Glycine	179.1	69.6	160.6	197.6
Glutamine	353.5	166.4	309.3	397.6
β -Alanine	5.2	4.0	4.1	6.2
Taurine	112.9	77.4	92.3	133.4
Histidine	22.9	14.0	19.1	26.6
γ -Aminobutyric acid	< 10			
Citrulline	34.0	34.8	24.8	43.3
Threonine	87.5	71.2	68.4	106.5
Alanine	265.8	120.7	233.8	297.8
β -Amino-isobutyric acid	< 5			
Carnosine	< 5			
Arginine	70.8	35.4	61.2	80.3
Proline	172.6	86.8	149.6	195.7
1-Methylhistidine	< 5			
Anserine	< 5			
3- Methylhistidine	< 5			
α -Amino-n-butyric acid	15.6	10.7	12.7	18.5
Tyrosine	69.8	58.8	54.2	85.5
Valine	173.3	63.5	156.4	190.2
Methionine	25.3	11.5	22.2	28.4
Isoleucine	46.0	18.9	41.0	51.0
Leucine	96.9	34.5	97.8	106.1
Hydroxylysine	< 5			
Phenylalanine	57.1	14.7	53.2	61.0
Tryptophan	16.1	7.7	14.0	18.2
Ornithine	44.0	24.2	37.6	50.4
Lysine	110.2	50.4	96.7	123.7

6.2.2.5 *Tyrosinemia Type I*: This is an autosomal recessive disorder, due to deficiency of the enzyme fumarylacetoacetate hydrolase (FAH), which is the last enzyme in the pathway for tyrosine breakdown. This defect leads to accumulation of tyrosine and its metabolites in the liver, causing liver disease. Tyrosine may also accumulate in the kidney and central nervous system, leading to delayed development. The patient was a 2 month old boy, with hepatomegaly and prolonged coagulogram. Analysis of urinary organic acid showed elevated succinylacetone. Alpha-fetoprotein was elevated. Amino acid analysis showed elevated tyrosine (560 nmole/ml), methionine (600 nmole/ml) and phenylalanine (188 nmole/ml) in plasma. This data assisted and confirmed clinical diagnosis.

6.2.2.6 *Nonketotic hyperglycinemia (NKH)*: This is an autosomal recessive disorder in the degradation of the glycine, due to a defect in the glycine cleavage system. This leads to accumulation of glycine in the whole body, including the central nervous system. Diagnosis of NKH requires amino acid analysis of plasma and cerebrospinal fluid (CSF amino acids). NKH patients typically have a CSF glycine/plasma glycine ratio greater than 0.08. The two patients studied had CSF glycine/plasma glycine ratios of 0.095 and 0.125 respectively. Patients with this disorder show symptoms within a few days after birth and decline rapidly, leading to lethargy, convulsions and if not treated, possibly cessation of breathing. Treatment involves use of sodium benzoate as anit-convulsant, and reducing the level of glycine in plasma and CSF.

6.2.3 Mucopolysaccharide Disorders

Several mucopolysaccharide disorders have been found in Thai patients, including Hurler Syndrome, Hunter Syndrome, Maroteaux-Lamy Syndrome, Sly Syndrome, Morquio Syndrome, and Scheie Syndrome. We have studied a Thai patient diagnosed as having Hurler syndrome, who was referred to the Department of Pediatrics, Siriraj Hospital for further treatment. The patient had physical examination showing coarse facies, claw-hand deformities, cloudy cornea, pectus deformities, macrocephaly and hepatosplenomegaly. Laboratory investigation showed positive urinary test for MPS, film spine characteristic of dysostosis multiplex, cloudy cornea from eye examination, DQ of 73 and IQ of 61 with mild mental retardation. These are positive findings for diagnosis of mucopolysaccharidosis (Hurler syndrome).

Studies in our laboratory showed a high level of urinary GAGs by the DMB method, confirming mucopolysaccharidosis. TLC analysis of GAGs

showed heparan sulfate and dermatan sulfate, consistent of Hurler's syndrome. Leukocytes were collected and assayed for enzyme activity. Enzymes assayed are those deficient in various mucopolysaccharide disorders, namely α -L-Iduronidase (lacking in MPS I, Hurler), α -N-Acetyl glucosaminidase (lacking in MPS III B, Sanfilippo B), Galactose-6-sulfatase (lacking in MPS IV A, Morquio A), β -Galactosidase (lacking in MPS IV B, Morquio B), N-Acetyl galactosamine-4-sulfatase (lacking in MPS VI Maroteaux-Lamy), and β -Glucuronidase (lacking in MPS VII, Sly).

Measurement of enzyme levels showed that five mucopolysaccharide degrading enzymes (β -galactosidase, β -glucuronidase, galactose-6-sulfatase, arylsulfatase B, α -N-ac-glucosaminidase) were present at normal levels. However, α -L-iduronidase was totally absent in agreement with the clinical diagnosis of Hurler syndrome. Genomic DNA and cDNA sequencing showed that the patient, SP, had the following differences with the database sequence for the α -L-Iduronidase gene: Q33H (CAG>CAT), A75T (TTG>CTG), R105Q (CGG>CAG), L118L (TTG>CTG), N181N (AAT>AAC), A314A (GCG>GCC), A361T (GCG>ACG), T388T (ACG>ACC), T410T (ACC>ACG), V454I (GTC>ATC) and S633L (TCG>TTG). The mRNA message level appeared to be normal, based on the signal ratio of 1.51 for α -iduronidase/ β -actin in SP vs. the ratio of 1.54 in the control by northern blot analysis. SP's phenotypically normal mother carried all the same polymorphisms as SP, except for Q33H and S633L. The mutation A75T, carried by SP and her mother, is known to cause a severe defect in α -iduronidase, Q33H is a non-disease-causing polymorphism, while S633L is a recently described mutation associated with Scheie syndrome. Thus, the α -iduronidase deficiency in SP appears to stem from the heterozygous combination of the A75T and S633L mutations, although the modifying effects of Q33H and other SNPs cannot be ruled out.

Another Hurler patient has been characterized and found to be deficient in α -iduronidase enzyme. The patient's α -iduronidase gene was sequenced and two new mutations were observed, 247insert-C and E299X, both of which result in production of incomplete proteins. So, this patient is sick due to lack of complete α -iduronidase being produced. Subsequently, we have begun to characterize the frequency of the SNPs we detected that affect the protein structure in the Thai population. Interestingly, we see an allele Q105, T361, I454 (and likely Q33) at a frequency of 19%. The Q33/Q105/T361/I454 allele has previously been characterized in Taiwan and found to have higher than average α -iduronidase activity.

6.2.4 Methylmalonic acidemia

We have evaluated 3 patients with methylmalonic acidemia (MMA), which results from the functional deficiency in the methylmalonyl CoA mutase (MCM) gene (*mut* gene). This disorder may be caused by a deficiency in the gene itself (*mut^f* or *mut* mutations) or in genes responsible for transport and processing of the cobalamin cofactor (*cbl* mutations). Two of the *cbl* genes, those impaired in the *cblA* and *cblB* mutations which are thought to be involved in transport into the mitochondria and deoxyadenylation, have been identified in the last year.

One patient, P.S., from Dr. Vorasuk Shotelersuk, was found to have 2 heterozygous mutations, 1048delT and 1706_1707delGGinsTA (G544X), which were found to be inherited from the mother and father, respectively, and result in an absence of functional proteins due to a frame shift and a premature stop codon. One other polymorphism, A499T, was found the proband and her mother and was tested for its affect on recombinant MCM expressed in *E. coli*. The K_m for both the cofactor and the substrate for this mutant were within error of the wild-type enzyme, so the mutation seemed to have little effect. In addition, the allele was found in 8% of the chromosomes from a non-MMA Thai population. So, this polymorphism seems to be a normal variant without effect on enzyme activity and the defect in the patient is due to the 1048delT and G544X mutations (Champattachai et al., 2003, *Mol. Genet. Metab.* **79**, 300-302).

The other two patients did not show any defect in MCM enzyme, so studies were begun to see whether there are mutations of the *cbl* genes. We found that the two patients had the same mutation in the *cblB* gene, E152X (nonsense stop mutation). One patient was homozygous, while the other was heterozygous, but the other mutation has not been found.

6.3 **Protein Changes in Human Cancer**

We are collaborating with Dr. Phaibun Punyarit of the Department of Pathology, Phramongkutkloao Hospital, Bangkok to analyze changes in protein in human cancer tissues. Specimens of tumor tissue and normal tissue from the same patient have been collected from many types of cancer, and have been characterized in terms of pathology. These tissues have been solubilized and fractionated by two-dimensional electrophoresis, comparing the patterns of tumor tissue with normal tissue from the same patient. Protein patterns are compared to databases available on the Worldwide Web, and some proteins of interest are identified by peptide mass fingerprinting, ESI/MS/MS or automated protein sequence analysis.

6.3.1. Diagnosis of thyroid cancer and non-cancer diseases by proteomics

Thyroid nodules are relatively common and are found in approximately 6% of women and 2% of men. Nodular or multinodular goiter is the most common non-neoplastic thyroid disease, but differential diagnosis of nodular hyperplasia from true neoplastic thyroid nodule by microscopic criteria is often difficult. It is also difficult to distinguish between benign and malignant forms of follicular neoplasms at the pre-operative stage. Thyroid scan, ultrasonography and fine-needle aspiration cytology (FNAC) are well established techniques for primary diagnosis of benign and malignant thyroid diseases. But the FNAC has inherent limitations due to inadequate sampling and overlapping cytological features between benign and malignant follicular lesions.

We have performed 2-DE analysis of soluble proteins from normal, benign and malignant thyroid tissues, including multinodular goiter (G), diffuse hyperplasia or Graves'disease (Hy), follicular adenoma (FA), follicular carcinoma (FC) and papillary carcinoma (PC) tissues. First dimension was electrofocusing in IPG, pH 3-10, and second dimension was 12.5% SDS-polyacrylamide gels. Selected gel spots were sequenced using an 473A Protein Sequencer (Perkins-Elmer) or identified by identified by nanoelectrospray mass spectrometry after tryptic digestion and desalting, using a Q-ToF instrument in collaboration with Prof. Wittman-Liebold of the Max Delbrück Center for Molecular Medicine, Berlin. Protein was identified using the sequence tag program and this information was then applied to search in a non-redundant translated nucleotide database (http://pepsea.protana.com/PA_PeptidePatternForm.html).

The 2-DE pattern of papillary carcinoma thyroid tissue is shown in Figure 21; to ensure reproducibility, this PC sample was studied 5 times by cutting different areas of tissue, and the same protein pattern was obtained. Tissues from several patients were studied for each disease, namely 5 G patients, 7 Hy patients, 6 FA patients, 7 FC patients and 9 PC patients, and for each patient, samples of diseased and normal tissue were compared. Thirty-two spots were identified by N-terminal sequencing and ESI-MS/MS to serve as landmarks for comparisons between tissues. These included various groups of proteins, such as cytoskeletal elements (actin, vimentin, tropomyosin, and myosin), proteins concerned with folding and assembly related to molecular chaperones (HSP27 and GRP78), proteins involved in transcription and translation (nucleoside diphosphate kinase, elongation factor1 β) and proteins involved in sulfur metabolism (glutathione-S-transferase P, thioredoxin).

The 2-D electrophoretic patterns of different disease states N, G, Hy, FA, FC and PC showed unique protein patterns in the molecular weight range of 15-30 kDa and pI of 4.5-6.5 (Figure 22). More spots were found in the neoplastic tissues (FA, FC and PC) than the non-neoplastic tissues (G, Hy). Eleven protein spots showed differences in the level of expression in different diseases, namely glutathione-S-transferase P (GSTP), thioredoxin peroxidase 1

(TPX1), superoxide dismutase (SODC), heat shock protein 27 (HSP27), prohibitin (PHB), 4 spots of cathepsin B (CB), DJ-1 Protein, and ATP synthase D chain (ATPQ). These have previously been reported as tumor-related proteins. In addition, TCTP gave variable expression with different thyroid diseases. Non-neoplastic tissues (both nodular goiter and diffuse hyperplasia) gave similar patterns to normal tissues. The level of TCTP (spot 8) was found to be very low in the G and Hy disease states. The level of HSP27 (spot 11) was higher in Hy compared to normal. The expression of cathepsin B (CB) was found to be the most distinctive. This protein was up-regulated in neoplastic tissues (Figure 22), and the very high levels of CB in FA disease allow FA to be readily distinguished from G or Hy. Comparison of benign (FA) and malignant tumors (FC, PC), shows that prohibitin (spot 12) and ATPQ (spot1) were up-regulated in PC, while HSP27 (spot 11) was present at higher levels in FA than in normal but at lower levels in both FC and PC, compared to the normal state.

Our results therefore suggest that neoplastic and non-neoplastic changes in the thyroid may be distinguished by proteomic studies. The most distinctive protein found is cathepsin B (CB), which is up-regulated in neoplastic thyroid tissue. Furthermore, up-regulation of ATPQ and PHB in papillary carcinoma provides another choice of markers for diagnosis, allowing PC to be distinguished from FC. Interestingly, four spots (spot 4, 5, 6 and 17) of cathepsin B (CB) could be identified (Figure 22). Spot 5 (CB1) and 6 (CB2) had the same pI but slightly different MW. CB1 and CB2 were found to be glycosylated, since they were positive for staining with PAS, with CB1 being more intensely stained, so possibly having more sugar than CB2.

6.3.2. Diagnosis of thyroid cancer and non-cancer diseases by proteomics

Two-dimensional electrophoresis showed some characteristic differences in the protein patterns from tissues from different thyroid diseases, including normal thyroid, non-neoplastic diseases, such as multinodular goiter (G), diffuse hyperplasia or Graves' disease (Hy), and neoplastic diseases such as the benign follicular adenoma (FA), and the malignant follicular carcinoma (FC) and papillary carcinoma (PC). The most distinctive protein found was cathepsin B, which could be detected as four spots, with differential expression in different thyroid diseases. In particular, two of these cathepsin B spots CB2 and CB3 are strongly upregulated in neoplastic diseases, compared to non-neoplastic diseases. However, the proteomic approach is a useful diagnostic tool for studying diseases involving the thyroid, but this technique is rather time-consuming, so another approach has also been investigated employing immunological techniques.