รายงานผลการวิจัยฉบับสมบูรณ์
โครงการวิจัย (Project)
โครงการวิจัยทุนอุดหนุนวิจัย มก. ปีงบประมาณ 2552

***ส่วนที่ 1* สรุปผลการดำเนินงานโครงการวิจัย** (Project)

1.1 รหัส ว-ท(ด) 12.52 ชื่อโครงการ ระบบประมาณโครงสร้างสามมิติลำดับกรดอะมิโนของโปรตีนฮีแมกกลูตินินชนิด A สายพันธุ์ย่อย H5N1 บนระบบอินเทอร์เน็ต

1.2 ลักษณะโครงการ ● เป็นโครงการวิจัยเดี่ยว
○ เป็นโครงการย่อยในชุดโครงการวิจัย (ระบุชื่อชุดโครงการวิจัย) .....

1.3 ชื่อหัวหน้าโครงการ นายจักร แสงมา

1.4 หน่วยงานหลักรับผิดชอบ หน่วยปฏิบัติการวิจัยเคมีสารสนเทศ ภาควิชาเคมี คณะวิทยาศาสตร์

1.5 ประเภทโครงการ ● โครงการวิจัย 3 สาขา; ☐ เกษตรศาสตร์ ■ วิทยาศาสตร์และเทคโนโลยี ☐ สังคมศาสตร์ฯ
○ โครงการวิจัยสถาบันเพื่อพัฒนาคุณภาพ
○ โครงการวิจัยและถ่ายทอดงานวิจัยสู่ประชาชน
○ โครงการเสริมสร้างความเข้มแข็งด้านการวิจัย
○ โครงการวิจัยเพื่อพัฒนาหน่วยปฏิบัติการวิจัยเชี่ยวชาญเฉพาะ (SRU)
○ โครงการวิจัยและพัฒนาเพื่อเพิ่มศักยภาพเชิงบูรณาการเพื่อการแข่งขันฯ
○ โครงการวิจัยพัฒนาร่วมภาครัฐและเอกชน

1.6 ระยะเวลาดำเนินงานวิจัยตลอดโครงการ 1 ปี ตั้งแต่ปีงบประมาณ 2552 ถึงปีงบประมาณ 2552

1.7 สถานที่ดำเนินงานวิจัย/เก็บข้อมูล หน่วยปฏิบัติการวิจัยเคมีสารสนเทศ ภาควิชาเคมี คณะวิทยาศาสตร์

1.8 งบประมาณรวมตลอดโครงการ 200,000 บาท ประกอบด้วย

ปีงบประมาณ 2552 ได้รับ 190,000 บาท ปีงบประมาณ ..... ได้รับ ..... บาท
ปีงบประมาณ ..... ได้รับ ..... บาท ปีงบประมาณ ..... ได้รับ ..... บาท
ปีงบประมาณ ..... ได้รับ ..... บาท ปีงบประมาณ ..... ได้รับ ..... บาท

1.9 วัตถุประสงค์โครงการวิจัย

1.9.1 เพื่อสร้างระบบประมาณโครงสร้างสามมิติลำดับกรดอะมิโนของโปรตีนฮีแมกกลูตินินบนระบบอินเทอร์เน็ต

1.9.2 ใช้ระบบที่สร้างขึ้นในการตรวจสอบการกลายพันธุ์ของโปรตีนฮีแมกกลูตินินชนิด H5N1 เพื่อเฝ้าระวังการกลายพันธุ์ของเชื้อไข้หวัดนก

1.10 เป้าหมายผลงานวิจัยตลอดโครงการ

ปีที่	เดือนที่	ผลงานวิจัยที่คาดว่าจะได้
1.	1-6	- ได้ free ware ที่เหมาะสมในการพัฒนาระบบ เข้าใจการทำงานและผลลัพท์ของโปรแกรมที่เลือกใช้ - ได้ข้อมูล sequence ของโปรตีนฮีแมกกลูตินินที่พบในฐานข้อมูล NCBI เพื่อนำมาศึกษาการกลายพันธุ์

7-12 - ได้ระบบเว็บไซต์และฐานข้อมูล ที่สามารถจำลองโครงสร้าง HA ได้
-ได้ส่วนแสดงผลแบบ 2 มิติ และ 3 มิติ

1.11 สรุปการดำเนินงานวิจัยตลอดโครงการ

วัตถุประสงค์ (ตามแผน)	**เป้าหมาย / ผลที่คาด (ตามแผน)**	**ผลการดำเนินงาน (ปฏิบัติได้จริง)**
1.คัดเลือก free ware ที่สามารถทำงานบนระบบเครือข่ายได้	1.ได้โปรแกรมที่มีประสิทธิภาพเหมาะนำมาใช้พัฒนาระบบ	1.1 ได้คัดเลือกโปรแกรม Modeller มาใช้ในการประมาณโครงสร้างสามมิติของ HA 1.2 เลือก Jmol applet ในการแสดงผลแบบสามมิติบนระบบอินเทอร์เน็ต
2. รวบรวม sequence ของโปรตีนฮีแมกกลูตินินที่พบในฐานข้อมูล NCBI และเก็บในฐานข้อมูล	2. ข้อมูล sequence ของ H5N1 ที่สามารถนำมาศึกษาการกลายพันธุ์ได้	2. ได้ข้อมูล sequence ของ H5N1 จาก NCBI จำนวน 3329 sequence นำมาศึกษาตำแหน่งกลายพันธุ์ ด้วยการทำ sequence alignment
3. พัฒนาระบบฐานข้อมูลและเว็บไซต์	3. ได้ระบบฐานข้อมูลและเว็บไซต์เพื่อรองรับการทำงานของผู้ใช้	3. ได้ระบบฐานข้อมูลและเว็บไซต์
4. พัฒนาส่วนแสดงผลแบบ 2 มิติ และ 3 มิติ	4. ได้ส่วนแสดงผลแบบ 2 มิติ และ 3 มิติ	4. สามารแสดง alignment file ให้เห็นตำแหน่งกลายพันธุ์ได้ชัดเจน ใช่ Jmol แสดงโครงสร้างสามมิติได้

1.12 สรุปผลการดำเนินงานตามวัตถุประสงค์

● บรรลุ ได้ระบบประมาณโครงสร้างสามมิติลำดับกรดอะมิโนของโปรตีนฮีแมกกลูตินินชนิด A สายพันธุ์ย่อย H5N1 บนระบบอินเทอร์เน็ต

○ บรรลุบางส่วน (ร้อยละ……) เหตุผล

○ ไม่บรรลุ เหตุผล

1.13 ผลผลิต/ สิ่งที่ได้จากการวิจัย (Outputs) (โปรดระบุรายละเอียด)

- ○ องค์ความรู้/ข้อมูลพื้นฐาน .....
- ○ สายพันธุ์พืช/สัตว์/จุลินทรีย์ .....
- ○ ผลิตภัณฑ์ .....
- ○ สิ่งประดิษฐ์ .....
- ○ เทคโนโลยี/นวัตกรรม .....
- ● ฐานข้อมูล/ซอฟแวร์ ระบบประมาณโครงสร้างสามมิติลำดับกรดอะมิโนของโปรตีนฮีแมกกลูตินินชนิด A สายพันธุ์ย่อย H5N1
- ○ คู่มือ .....
- ○ วิดีทัศน์ .....
- ● การสร้างนักวิจัย/สนับสนุนนิสิตปริญญาตรี ..... คน ปริญญาโท 1 คน ปริญญาเอก ..... คน
- ○ สนับสนุนการศึกษาปัญหาพิเศษ ..... เรื่อง (ระบุ) ..... วิทยานิพนธ์ ..... เรื่อง (ระบุ) .....
- ○ อื่นๆ (ระบุ) .....

1.14 ผลลัพธ์/ผลสำเร็จที่ได้/หรือคาดว่าจะได้จากการนำผลการวิจัยไปใช้ประโยชน์ (Outcomes)

(1) เป้าหมายการนำไปใช้ประโยชน์ (ระบุกลุ่มเป้าหมายของงานวิจัยเชิงปริมาณ/คุณภาพ)

- ○ ด้านการศึกษา/เสริมการเรียนการสอน .....
- ○ ด้านการเกษตร .....
- ○ ด้านอุตสาหกรรม .....
- ○ ด้านทรัพยากรธรรมชาติ/สิ่งแวดล้อม .....
- ○ ด้านคุณภาพชีวิต สุขภาพอนามัย .....
- ○ ด้านเศรษฐกิจ .....
- ○ ด้านสังคม .....
- ○ ด้านการทำนุบำรุงศิลป ศาสนา วัฒนธรรม .....
- ○ ด้านการถ่ายทอดเทคโนโลยี/ฝึกอบรมแก่กลุ่มเป้าหมาย .....
- ○ เสนอภาครัฐ เพื่อใช้กำหนดแผน/นโยบาย ฯลฯ .....
- ● นำความรู้ไปวิจัย/พัฒนาขั้นต่อไป ระบบประมาณโครงสร้างสามมิติของโปรตีนฮีแมกกลูตินินชนิด A สายพันธุ์ย่อย H5N1ที่สร้างขึ้นจะเป็นประโยชน์ในการศึกษาและวิจัยเรื่องไข้หวัดนก ใช้ในการตรวจสอบการกลายพันธุ์เพื่อเฝ้าระวังการกลายพันธุ์ได้ นอกจากนี้ยังสามารถขยายระบบรองรับเพื่อศึกษาโปรตีนชนิดอื่นๆได้ .....
- ○ ก่อให้เกิดความร่วมมือระหว่างหน่วยงาน/การสร้างเครือข่าย .....
- ○ อื่นๆ (ระบุ) .....

(2) สรุปผลการนำผลการวิจัยไปเผยแพร่ / ถ่ายทอด **ตั้งแต่เริ่มต้นจนสิ้นสุดโครงการ** (ระบุรายละเอียดอยู่ระหว่างดำเนินการส่งตีพิมพ์/ตีพิมพ์แล้วในรูปแบบ**เอกสารอ้างอิง**และ**แนบสำเนาเป็นภาคผนวก**ของรายงาน)

◯ ตีพิมพ์ในวารสารวิชาการต่างประเทศ ..... เรื่อง (ระบุ) .....

◯ ตีพิมพ์ในวารสารวิชาการในประเทศ ..... เรื่อง (ระบุ) .....

◯ นำเสนอในการประชุม/สัมมนา ต่างประเทศ ..... เรื่อง (ระบุ) .....

◯ นำเสนอในการประชุม/สัมมนา ในประเทศ 1 เรื่อง (ระบุ) .....

Taweesak Poochai, Daungmanee Chuakhaew, Chak Sangma, Web-Based Application for Automated 3D Structure of Hemagglutinin (H5N1), PACCON 2008, Jan 30 – Feb 1, 2008, Sofitel Centara Grand Bangkok, Bangkok

◯ นำเสนอทางวิทยุ/โทรทัศน์/Website 1 เรื่อง/ครั้ง (ระบุ) .....

เว็บไซต์ฐานข้อมูล http://chemiebase.ku.ac.th/3dvis

◯ นำเสนอทางนิทรรศการ ..... เรื่อง/ครั้ง (ระบุ) .....

◯ บทความ/เอกสารสิ่งพิมพ์/วีดีทัศน์ ..... เรื่อง/ครั้ง (ระบุ) .....

◯ ถ่ายทอด/ฝึกอบรมแก่เกษตรกร/ผู้สนใจ ..... เรื่อง/ครั้ง (ระบุ) .....

◯ ถ่ายทอดสู่ภาคเอกชน/อุตสาหกรรม/ผู้ประกอบการ (ประโยชน์เชิงพาณิชย์) ..... เรื่อง/ครั้ง (ระบุ) .....

◯ ภาครัฐนำไปใช้กำหนดแผน/นโยบาย ฯลฯ (ระบุ) .....

◯ มีผู้นำผลงานวิจัยไปอ้างอิง (ระบุ) .....

◯ อื่นๆ (ระบุ) .....

1.15 การยื่นจด ◯ สิทธิบัตร ◯ อนุสิทธิบัตร ◯ ลิขสิทธิ์

◯ มีศักยภาพที่จะยื่นจด (ระบุ) ..... ◯ ยื่นจดแล้ว เมื่อ .....

1.16 ผลกระทบ (Impact) ที่เกิดจากการนำผลการวิจัยไปใช้ (ระบุว่าก่อให้เกิดผลกระทบอย่างไร)

◯ ด้านความมั่นคง อาทิ การเมืองการปกครอง กฎหมาย การต่างประเทศ โครงสร้างพื้นฐาน และบริการโทรคมนาคม ฯลฯ (ระบุ) .....

◯ ด้านการเศรษฐกิจ อาทิ การพาณิชยกรรม การเกษตรกรรม การอุตสาหกรรม การท่องเที่ยว วิทยาศาสตร์และเทคโนโลยี พลังงาน ฯลฯ (ระบุ) .....

◯ ด้านคุณภาพชีวิตและสังคม ศักยภาพของคนและการศึกษา การแพทย์และสาธารณสุข หลักประกันความมั่นคง สวัสดิการสังคม วัฒนธรรม จริยธรรมและค่านิยม ฯลฯ (ระบุ) .....

◯ ด้านทรัพยากรธรรมชาติและสิ่งแวดล้อม การบริการจัดการการใช้ทรัพยากรให้เกิดประโยชน์ การป้องกันการทำลาย ลดการสูญเสีย การฟื้นฟูทรัพยากรและสิ่งแวดล้อม ฯลฯ

◯ อื่นๆ (ระบุ) .....

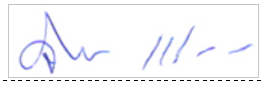
1.17 ผลการดำเนินงานวิจัยสอดคล้องกับยุทธศาสตร์ชาติ ในด้าน

- ◯ ยุทธศาสตร์การขจัดความยากจน
- ● ยุทธศาสตร์การพัฒนาคนและสังคมที่มีคุณภาพ
- ◯ ยุทธศาสตร์การปรับโครงสร้างเศรษฐกิจให้สมดุล และแข่งขันได้
- ◯ ยุทธศาสตร์การบริหารจัดการทรัพยากรธรรมชาติและสิ่งแวดล้อม
- ◯ ยุทธศาสตร์การต่างประเทศและเศรษฐกิจระหว่างประเทศ
- ◯ ยุทธศาสตร์การพัฒนากฎหมายและส่งเสริมการบริหารกิจการบ้านเมืองที่ดี
- ◯ ยุทธศาสตร์การส่งเสริมประชาธิปไตยและกระบวนการประชาสังคม
- ◯ ยุทธศาสตร์การรักษาความมั่งคงของรัฐ
- ◯ ยุทธศาสตร์การรองรับการเปลี่ยนแปลงและพลวัตโลก
- ◯ อื่นๆ โปรดระบุ .....

1.18 ปัญหา อุปสรรค ในการดำเนินงานวิจัยและแนวทางแก้ไข .....

1.19 งานที่จะทำต่อไป/คำชี้แจงเพิ่มเติม .....

1.20 ได้แนบ “**รายงานผลการวิจัยฉบับสมบูรณ์ของโครงการ** (Project)” ตามหัวข้อ **ในส่วนที่** 2 (หน้าถัดไป) มาด้วยแล้ว

ลงชื่อ  หัวหน้าโครงการ

(นายจักร แสงมา)

6 / พฤศจิกายน /2556 วัน/เดือน/ปี ที่รายงาน

ส่วนที่ 2 **รายงานผลการวิจัยฉบับสมบูรณ์**

โครงการวิจัยทุนอุดหนุนวิจัย มก. ปีงบประมาณ 2552

โครงการวิจัยรหัส ว-ท(ด) 12.52

(ชื่อโครงการภาษาไทย ระบบประมาณโครงสร้างสามมิติลำดับกรดอะมิโนของโปรตีนฮีแมกกลูตินินชนิด A สายพันธุ์ย่อย H5N1 บนระบบอินเทอร์เน็ต)

(ชื่อโครงการภาษาอังกฤษ Web-Based Application for Automated 3D Structure of Hemagglutinin (H5N1))

(ชื่อผู้วิจัยภาษาไทย นายจักร แสงจักร[1] นายศิษฎ์ศ ทองสิมา [2] นส.ดวงมณี เชื้อเขียว [3])

(ชื่อผู้วิจัยภาษาอังกฤษ Chak Sangma[1] Sissades Tongsima [2] Daungmanee Chuakheaw [3])

บทคัดย่อ

ฮีแมกกลูตินินเป็นโปรตีนที่หวัดนกใช้ในการจับเซลล์เจ้าบ้าน การกลายพันธุ์ของฮีแมกกลูตินินจึงอาจทำให้หวัดนกติดคนได้มากขึ้น มีการกลายพันธุ์ของฮีแมกลลูตินินเกิดขึ้นมากมายแต่เราไม่สามารถโคลนการกลายพันธุ์ทุกๆ แบบได้ จึงต้องอาศัยเครื่องมือที่สามารถจำลองโครงสร้างของโปรตีนเพื่อช่วยคัดกรองโครงสร้างที่มีแนวโน้มที่จะติดคนได้มาก เพื่อนำไปโคลนก่อน โครงการวิจัยนี้ได้สร้างเครื่องมือจำลองโครงสร้างโดยอาศัยโปรแกรมฟรีและโปรแกรมที่เปิดเผยโค๊ด เพื่อจำลองโครงสร้างของโปรตีนฮีแมกกลูตินิน ผลการทดสอบพบว่าโครงสร้างที่ได้มีคุณภาพดีเทียบเท่าโปรแกรมทางการค้า และเว็บแอพพลิเคชั่นที่นิยมใช้

คำสำคัญ: ฮีแมกกลูตินิน, จำลองโครงสร้าง

ABSTRACT

Influenza A hemagglutinin is a protein that birdflu virus use to bind with host cell receptor. Mutations on this protein could make the birdflu adapt to human by changing the protein structure. Several mutations have been found but cloning all of these proteins is not possible. Prediction tool that use sequence data to build approximation structures suitable for modeling the binding mechanism can help to select a portion of sequences that start to bind to human better to be cloned. This project was to build software from an open source that can produce an approximate hemagglutinin structure form protein sequence. The result was comparable to the commercial software based on the same prediction protocol.

Key words: hemagglutinin, comparative modeling, 3D model, homology modeling

บทนำ

โปรตีนฮีแมกกลูตินิน (hemagglutinin) เป็นส่วนสำคัญของเชื้อไข้หวัดนกที่ทำให้เกิดการระบาดของเชื้อ โดยโปรตีนฮีแมกกลูตินินมีบทบาทที่ทำให้เชื้อสามารถจับกับเซลล์เจ้าบ้าน (host cell) ทั้งที่เป็นคนและสัตว์ ทำ

ให้เชื้อเข้าสู่เซลล์และเกิดกลไกการของเชื้อและเกิดการระบาดต่อไปได้ การกลายพันธุ์ของโปรตีนฮีแมกกลูตินินหรือการเปลี่ยนแปลงชนิดของกรดอะมิโนของโปรตีน อาจมีผลทำให้เชื้อติดคนได้ง่ายขึ้น ปัจจุบันมีโครงสร้างของโปรตีนฮีแมกกลูตินินทั้งที่แยกได้จากคนและสัตว์แล้วหลายพันโครงสร้าง โดยโครงสร้างเหล่านี้จะมีชนิดของกรดอะมิโนแตกต่างกันออกไปในบางตำแหน่ง ซึ่งหากสามารถติดตามการเปลี่ยนแปลงชนิดของกรดอะมิโนเหล่านี้จะทำให้ทราบแนวโน้มและรูปแบบการกลายพันธุ์ใน host แต่ละชนิดได้

การศึกษาโปรตีนฮีแมกกลูตินินที่นิยมวิธีหนึ่งในปัจจุบัน คือ การใช้ระเบียบวิธีทางคอมพิวเตอร์ แต่ในการศึกษาการกลายพันธุ์โดยใช้ระเบียบวิธีทางคอมพิวเตอร์นั้น จำเป็นต้องมีโครงสร้างสามมิติของโปรตีนที่จะทำการศึกษา ซึ่งปัจจุบันได้จากฐานข้อมูล Protein Data Bank (PDB) [1] โดยเป็นข้อมูลโครงสร้างทาง X-ray แต่ในการได้มาของโครงสร้างนี้ต้องเสียเวลาและค่าใช้จ่ายมาก ทางหนึ่งที่เป็นไปได้ในการแก้ปัญหาดังกล่าวคือ การใช้โปรแกรมประยุกต์โดยอาศัยพื้นฐานของระเบียบวิธี comparative modeling เพื่อสร้างโครงสร้างสามมิติของโปรตีนจากลำดับกรดอะมิโน (sequence) แต่วิธีดังกล่าวมีข้อจำกัดหลายประการ ได้แก่ ความจำเป็นต้องใช้คอมพิวเตอร์สมรรถนะสูงในการคำนวณ ค่าลิขสิทธิ์ของโปรแกรม ค่าใช้จ่ายในการอบรมบุคลากรเพื่อให้ปฏิบัติงานกับโปรแกรมประยุกต์ ระยะเวลาในการสร้างข้อมูลตั้งต้นของแต่ละโครงสร้างโปรตีนเพื่อทำการคำนวณ และการตรวจสอบผลลัพธ์ที่ได้จากการคำนวณ ดังนั้น โครงการวิจัยจึงเกิดขึ้นเพื่อสร้างระบบประมาณโครงสร้างสามมิติลำดับกรดอะมิโนของโปรตีนฮีแมกกลูตินินบนระบบอินเทอร์เน็ต เพื่ออำนวยความสะดวก และลดค่าใช้จ่ายด้านต่างๆ แก่ผู้ใช้ในการศึกษาและติดตามการกลายพันธุ์ของโปรตีนฮีแมกกลูตินิน โดยระบบสามารถแสดงลำดับกรดอะมิโนที่เหมือนและต่างกันให้เห็นอย่างชัดเจน แสดงโครงสร้างสามมิติพร้อมทั้งระบุบริเวณของการกลายพันธุ์ที่สำคัญ เปรียบเทียบโครงสร้างที่จำลองขึ้นกับ template และสามารถนำไปประยุกต์ใช้ภายในหน่วยงานทำให้ไม่ต้องส่ง sequence ไปยัง server ภายนอก การพัฒนาระบบเพื่อใช้เองนี้ยังทำให้เราสามารถปรับเปลี่ยนการแสดงผลให้เหมาะตามความต้องการและเลือกใช้โปรแกรมที่มีประสิทธิภาพในการสร้างโครงสร้างสามมิติได้ โครงการนี้จึงถือเป็นโครงการนำร่องในส่วนการพัฒนาเว็บแอพพลิเคชันเพื่อจำลองโครงสร้างสามมิติของโปรตีน และส่งเสริมการศึกษาวิจัยทางเคมี/ชีวสารสนเทศในประเทศไทยอีกทางหนึ่ง

ระบบประมาณโครงสร้างสามมิติบนระบบอินเทอร์เน็ตที่นิยมใช้ในปัจจุบัน ได้แก่ Swiss-model [2] ซึ่ง ผู้ใช้สามารถเลือกได้ว่าจะสร้าง 2D alignment หรือ 3D model ของ sequence ที่สนใจ โดยเทียบโครงสร้างใน Protein databank การใช้งานระบบค่อนข้างง่าย ให้ข้อมูลของโครงสร้างที่สร้างขึ้น และสามารถเลือกบันทึกข้อมูลได้หลายแบบ ระบบที่ผู้วิจัยพัฒนาขึ้นจะมีลักษณะการทำงานคล้ายกับ Swiss-model คือ สามารถสร้าง 2D alignment และ 3D model แต่จะเพิ่มความสามารถในการนำโครงสร้างที่สร้างขึ้นไปคำนวณด้วยระเบียบวิธี Molecular dynamics simulation (MD) เพื่อความสะดวกในการศึกษาระบบเชื้อไข้หวัดด้วยระเบียบวิธีเคมีคอมพิวเตอร์มากขึ้น

วิธีวิจัย

การดำเนินงานวิจัยเริ่มจากการหาโปรแกรมที่เหมาะสมสำหรับการทำวิจัย โดยต้องเป็นโปรแกรมที่สามารถนำมาใช้งานได้โดยไม่เสียค่าใช้จ่าย หรือ ฟรีแวร์ (free ware) และต้องมีประสิทธิภาพในการประมาณ

โครงสร้างสามมิติของลำดับอะมิโน (sequence) นำมาทำการทดสอบการทำงานเพื่อยืนยันความถูกต้องก่อนนำไปพัฒนาต่อ จากนั้นจึงออกแบบการทำงานของระบบทั้งในส่วนของฐานข้อมูลและเว็บไซต์

เมื่อได้โปรแกรมที่เหมาะสมและได้ระบบฐานข้อมูลแล้วจึงรวบรวมข้อมูล sequence ของโปรตีนฮีแมกกลูตินินสายพันธุ์ย่อย H5N1 จากฐานข้อมูล NCBI [3] และรวบรวมข้อมูลโครงสร้างทางเอ็กเรย์จาก PDB เพื่อนำมาทดสอบการสร้างโครงสร้างสามมิติ เทียบโครงสร้างที่จำลองขึ้นกับโครงสร้างแม่แบบ (template) ผลจากการจำลองโครงสร้างจะให้ข้อมูล เปอร์เซ็นความคล้ายกับ template (%identity) ค่าพลังงานของโครงสร้างซึ่งจะทำให้ผู้ใช้ทราบความสมบูรณ์ของโครงสร้างที่ประมาณขึ้น

จากนั้นสร้างระบบการแสดงผลแบบสองมิติและสามมิติที่สามารถแสดงความแตกต่างของตำแหน่งการกลายพันธุ์ของโปรตีน ทดสอบระบบและปรับปรุงแก้ไขเพื่อให้ง่ายต่อการใช้งาน

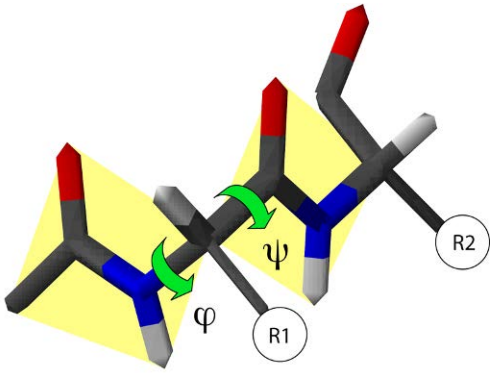
ผลและวิจารณ์

1. ผลการทดสอบโปรแกรม

การดำเนินงานช่วงหกเดือนแรกของการวิจัยเป็นการทดสอบโปรแกรมที่จะนำมาพัฒนาระบบประมาณโครงสร้างสามมิติ ซึ่งจะทำการศึกษาลักษณะการทำงาน ข้อดี ข้อด้อย ของแต่ละโปรแกรมเพื่อให้ได้โปรแกรมที่เหมาะสมและสามารถนำมาพัฒนาให้เกิดเป็นระบบที่มีประสิทธิภาพ

ผู้วิจัยได้เลือกโปรแกรม modeller [4] และ scwrl4 [5] ซึ่งเป็น free ware ที่เป็นที่ยอมรับที่นำมาใช้ในการประมาณโครงสร้างสามมิติลำดับอะมิโนของโปรตีน (protein sequence) มาทดสอบสร้างโครงสร้างสามมิติของโปรตีนฮีแมกกลูตินิน และเทียบผลกับ swiss-model server [6] ซึ่งเป็นระบบประมาณโครงสร้างสามมิติบนอินเทอร์เน็ตที่เป็นที่นิยม โดยจะนำมาวิเคราะห์ความถูกต้องของของโครงสร้างด้วยแผนภูมิ ramachandran (ramachandran plot หรือ ramachandran diagram) [7] และตัวแปรต่างๆ ที่คำนวณจากโปรแกรม procheck [8]

แผนภูมิ ramachandran พัฒนาโดย Gopalasamudram Narayana Iyer Ramachandran เพื่อใช้ในการตรวจสอบความถูกต้องของโครงสร้างโปรตีน เนื่องจากโครงสร้างหลัก หรือ backbone ของโปรตีนต้องมีคุณสมบัติ peptide planarity และมีลักษณะ trans-conformer เป็นส่วนใหญ่ การจัดเรียงอะตอมของ backbone จึงค่อนข้างจำเพาะ การตรวจโครงรูปของ backbone ของโปรตีนจะระบุด้วยมุมทอร์ชัน ϕ และ ψ ดังรูป 1



http://en.wikipedia.org

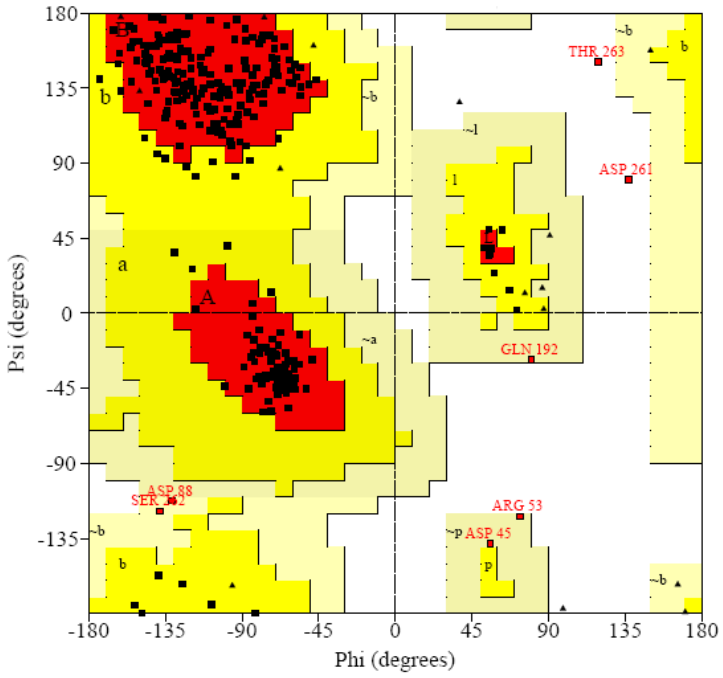
C

C_α

N

รูปที่ 1 แสดงโครงสร้าง backbone ของโปรตีน

มุม ϕ เป็นมุมจากการหมุนของ C_α-N และ ψ เป็นมุมจากการหมุนของ C_α-C ซึ่งทั้งสองเป็นพันธะที่สามารถหมุนได้ (rotatable bond) มีค่าตั้งแต่ -180° ถึง 180° จากการวิเคราะห์โครงสร้างของโปรตีนชนิดต่างๆ พบว่าอิทธิพลด้านความเกะกะของโครงสร้าง (steric effect) จะทำให้ ϕ และ ψ มีค่าจำเพาะและเป็นช่วงๆ ดังนั้นเมื่อนำมาสร้างกราฟจะได้กราฟที่บอกถึงความสัมพันธ์ของ ϕ และ ψ ดังรูป 2

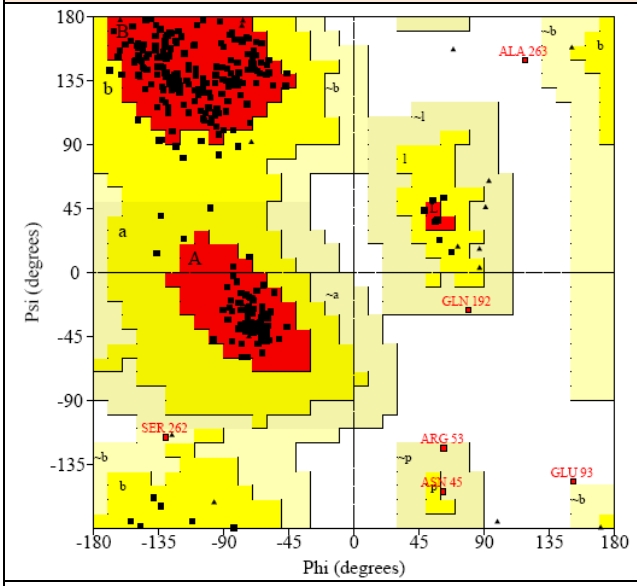
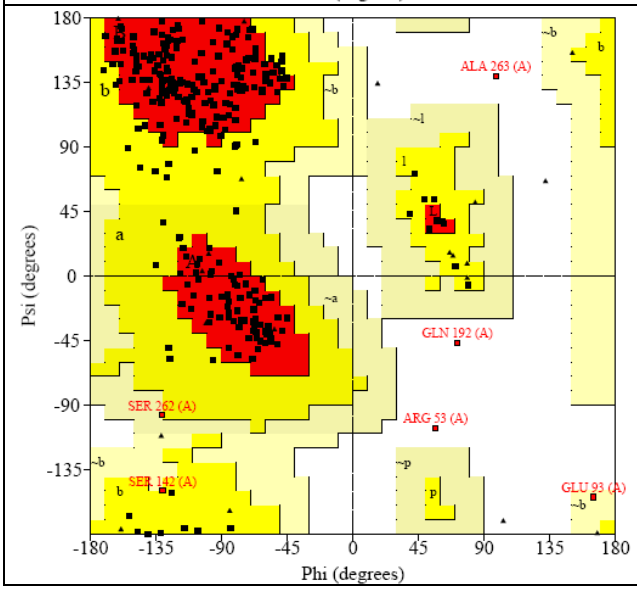


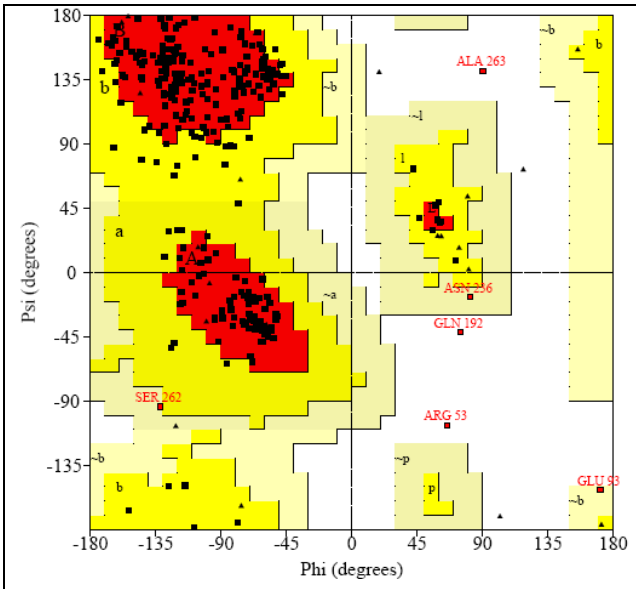
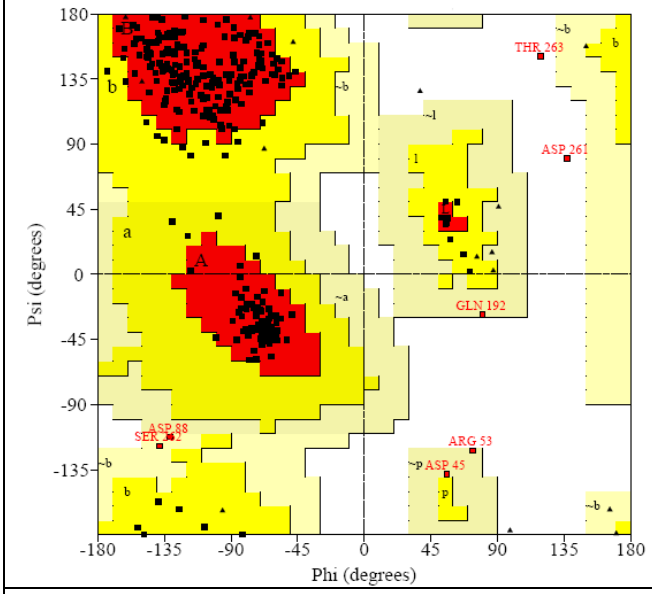
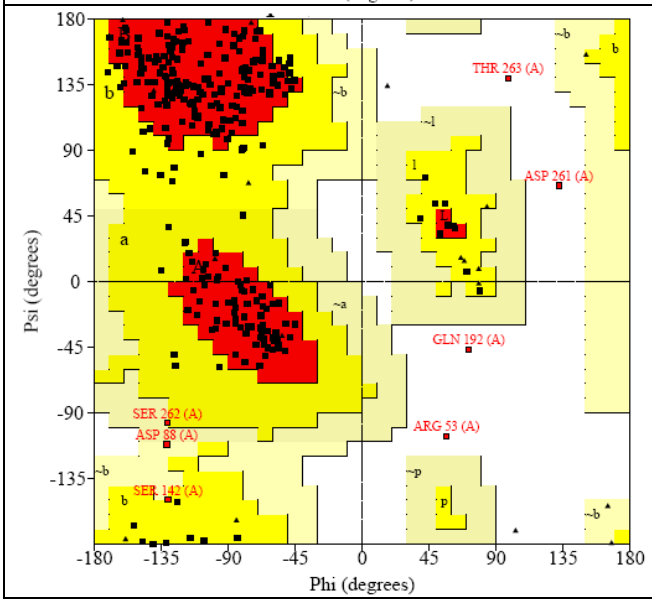
รูปที่ 2 แสดงแผนภูมิ ramachandran โดยบริเวณที่แรเงา แสดงถึง ϕ และ ψ ที่ยอมรับได้ ส่วนบริเวณพื้นสีขาวแสดงถึง ϕ และ ψ ที่ไม่ถูกต้องเป็นบริเวณที่โครงรูปของโปรตีนที่บิดเบี้ยวไป พื้นที่สีแดงเรียกว่า favored region หรือ core region โครงสร้างที่ถือว่าสมบูรณ์ต้องมีเรซิดิวส์ที่ให้ค่ามุมอยู่บริเวณนี้มากกว่า 90 % (ramachandran plot > 90 % core)

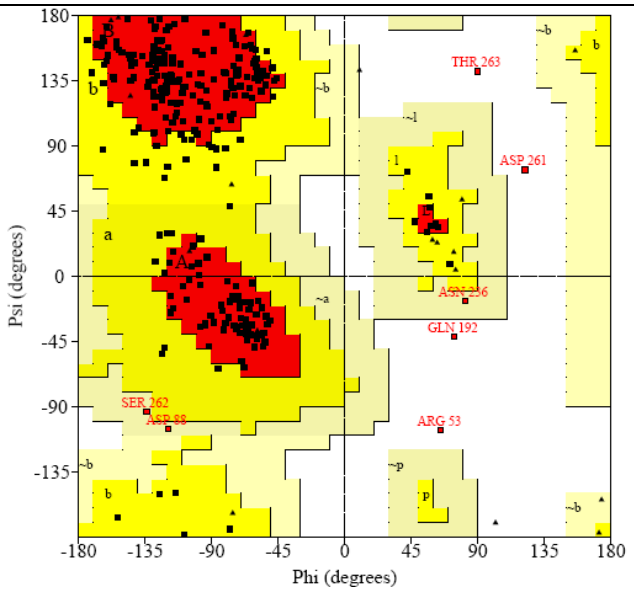
ระบบที่นำมาทดสอบการทำงานของโปรแกรมในเบื้องต้นนี้ได้แก่ A/Duck/Singapore/3/97 เรียกโดยย่อว่า Sing-97, A/Thailand/1(Kan-1)/2004 หรือ Kan-1 และ L129V/A134V Kan-1 โดย Sing-97 เป็น HA ที่มีโครงสร้างทางเอ๊กเรย์แล้ว และพบว่ามี ramachandran plot 92.5% core ซึ่งจะได้นำมาเป็นข้อมูลอ้างอิง

เพื่อเปรียบเทียบความถูกต้องของแต่ละโปรแกรม ส่วนการสร้างโครงสร้างสามมิติของ Kan-1 และ L129V/A134V Kan-1 นั้น จะใช้ Sing-97 เป็น Template ผลการทดสอบแสดงดังตาราง 1

ตารางที่ 1 แสดงผลการประมาณโครงสร้างสามมิติของโครงสร้าง Sing-97, Kan-1 และ L129V/A134V Kan-1 ด้วยโปรแกรม modeller, scwrl4 และ swiss-model server

ระบบ Sing-97	
	**ระบบ**: Sing-97 **จาก** modeller - ramachandran plot: 88.6% core : 9.3% allow : 1.4% gener : 0.7% disall - planar groups: 100.0% - average B-factor for buried atoms : 68.543
	**ระบบ**: Sing-97 **จาก** scwrl4 ramachandran plot: 82.1% core : 15.7% allow : 1.1% gener : 1.1% disall planar groups: 100.0% - average B-factor for buried atoms : 0.000

	**ระบบ: Sing-97 จาก swiss-model** ramachandran plot: 81.7% core : 16.1% allow : 1.1% gener : 1.1% disall planar groups: 100.0% - average B-factor for buried atoms : 50.000
ระบบ Kan-1	
	**ระบบ: Kan-1 จาก modeller** ramachandran plot: 88.0% core : 9.5% allow : 1.8% gener : 0.7% disall planar groups: 100.0% - average B-factor for buried atoms : 68.357
	**ระบบ: Kan-1 จาก scwrl4** ramachandran plot: 82.0% core : 15.5% allow : 1.1% gener : 1.4% disall planar groups: 100.0% - average B-factor for buried atoms : 0.000



ระบบ: Kan-1 จาก swiss-model

ramachandran plot: 81.6% core

: 15.9% allow

: 1.1% gener

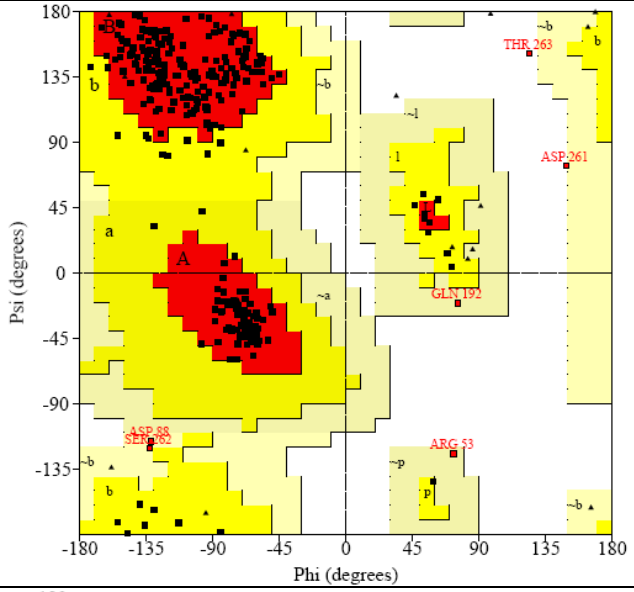
: 1.4% disall

planar groups: 83.5%

- average B-factor for buried atoms

: 50.000

ระบบ L129V/A134V Kan-1



ระบบ: L129V/A134V Kan-1

จาก modeller

ramachandran plot: 88.4% core

: 9.5% allow

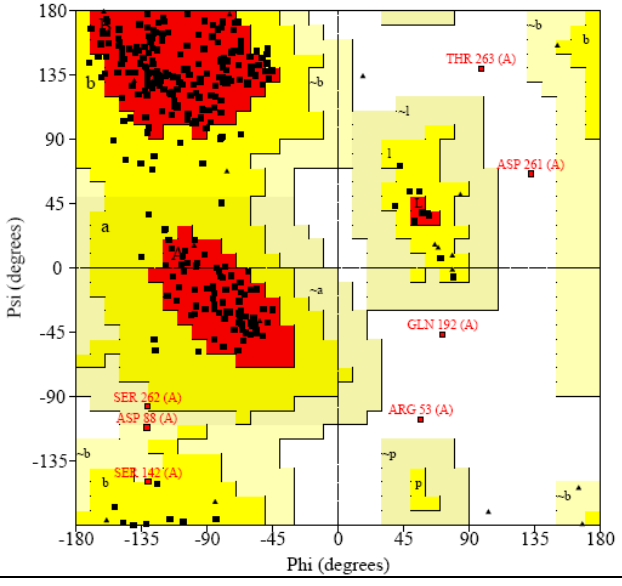
: 1.4% gener

: 0.7% disall

planar groups: 100.0%

- average B-factor for buried atoms

: 68.271



ระบบ: L129V/A134V Kan-1

จาก scwrl4

ramachandran plot: 82.0% core

: 15.5% allow

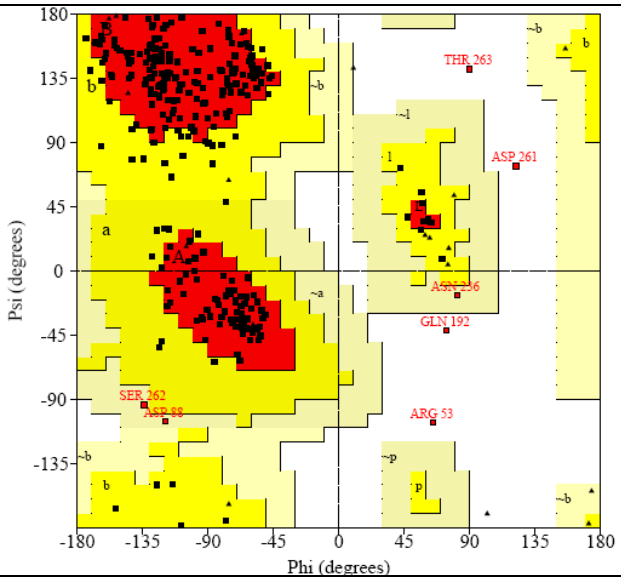
: 1.1% gener

: 1.4% disall

planar groups: 100.0%

- average B-factor for buried atoms

: 0.000



ระบบ: L129V/A134V Kan-1

จาก swiss-model

ramachandran plot: 81.3% core

: 16.3% allow

: 1.1% gener

: 1.4% disall

planar groups: 83.5%

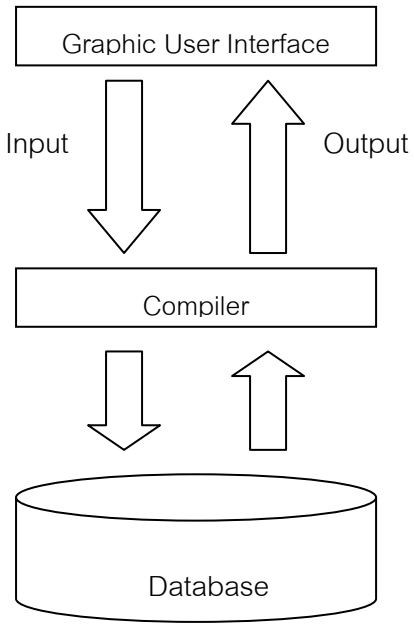
- average B-factor for buried atoms

: 50.000

จากผลการศึกษา ramachandran plot และตัวแปรอื่นๆ ที่คำนวณจากโปรแกรม procheck เพื่อตรวจสอบความถูกต้องของโครงสร้างโปรตีนที่สร้างขึ้น พบว่า โปรแกรม modeller สามารถประมาณโครงสร้างได้ดีกว่า scwrl4 และ swiss-model จึงเป็นเหตุผลหลักที่ทำให้ผู้วิจัยเลือกโปรแกรม modeller เพื่อใช้ในการพัฒนาระบบประมาณโครงสร้างสามมิติของ HA ที่จะดำเนินการในขั้นตอนต่อไป นอกจากนี้ modeller ยังมีโมดูลอื่นๆ ที่มีประสิทธิภาพที่สามารถนำมาพัฒนาต่อยอดและใช้งานร่วมกับโปรแกรมอื่นๆ เช่นโปรแกรมวิเคราะห์ผลและแสดงผลต่างๆ ในส่วนโปรแกรมแสดงผลแบบสามมิตินั้น ผู้วิจัยเลือกใช้ Jmol [9] เนื่องจากเป็น Java applet มีความยืดหยุ่นและใช้งานง่าย นอกจากนี้ยังเปิดเผยโค๊ดสามารถนำมาแก้ไขปรับปรุงให้เหมาะสมกับการทำงานของระบบได้

1. ผลการออกแบบและการทำงานของระบบ

การทำงานของระบบจะมีลักษณะเป็นเว็บแอพพลิเคชั่น โดยผู้ใช้ (client) รับ-ส่ง ข้อมูลมาประมวลผลยังฝั่งผู้ให้บริการ (server) ผ่านระบบอินเทอร์เน็ต ตามแบบฟอร์มที่ออกแบบในส่วน graphic user interface (GUI) ไว้ ลักษณะดังรูป 3



รูปที่ 3 แสดงการทำงานของระบบ

ส่วน GUI หรือ ส่วนการติดต่อกับผู้ใช้ พัฒนาขึ้นโดยใช้ภาษา PHP (PHP Hypertext Processor) ร่วมกับ HTML (Hyper Text Markup Language) โดย PHP จะเป็นคำสั่งที่ใช้ในการประมวลผลฝั่ง server แล้วจึงส่งมาแสดงผลด้วย HTML ฝั่ง client

ส่วนการประมวลผล (compiler) ของระบบจะเรียกใช้โปรแกรม modeler และ procheck ในการประมาณโครงสร้างสามมิติและคำนวณความถูกต้องของโครงสร้างที่ได้บน server ผ่านไพธอนสคริปต์ (python script)

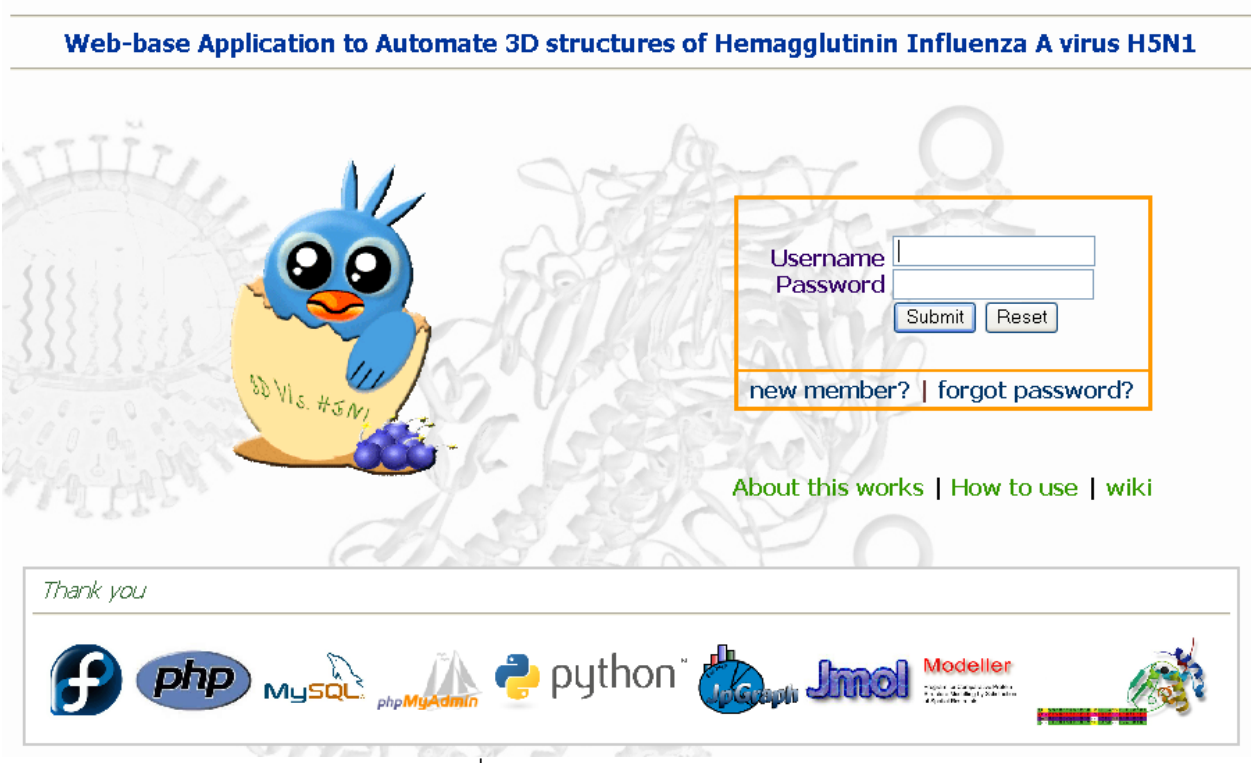
ส่วนฐานข้อมูล (database) ผู้วิจัยเลือกใช้ MySQL ซึ่งเป็นโปรแกรมจัดการฐานข้อมูลที่มีประสิทธิภาพทำงานร่วมกับ PHP ได้ดี รองรับข้อมูลขนาดใหญ่ และจัดการระบบได้ง่ายผ่าน phpmyadmin และเผยแพร่อยู่บนเครือข่ายนนทรีในชื่อ http://chemiebase.ku.ac.th/3dvis

การทำงานของระบบ

ลักษณะการทำงานของระบบที่พัฒนาขึ้นจะคล้ายกับระบบของ Swiss-model คือ ให้ผู้ใช้เลือกได้ว่าจะสร้าง 2D alignment หรือ 3D model ของ sequence ที่สนใจ แต่จะเพิ่มความสามารถในการส่งโครงสร้างที่สร้างขึ้นไปคำนวณ MD ได้ เพื่อความสะดวกในการศึกษาเชื้อไข้หวัดด้วยระเบียบวิธีทางเคมีคอมพิวเตอร์

การทำงานของระบบประกอบด้วยหน้า (form) ต่างๆ ดังนี้

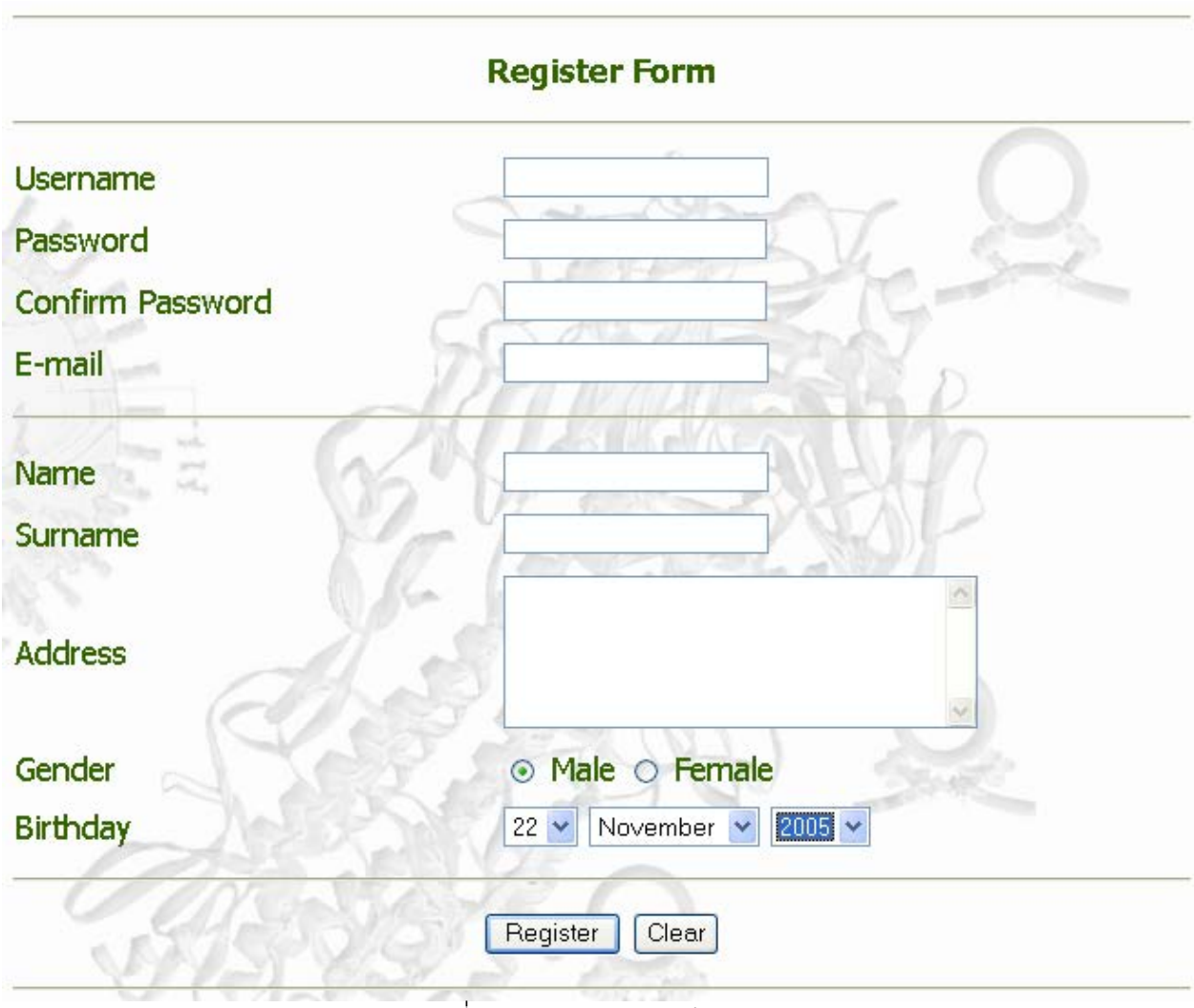
- home – เป็นหน้าแรกของเว็บไซต์ แสดงดังรูป 4



รูปที่ 4 แสดงหน้าแรกของเว็บไซต์

- register หน้าลงทะเบียนสมาชิกใหม่

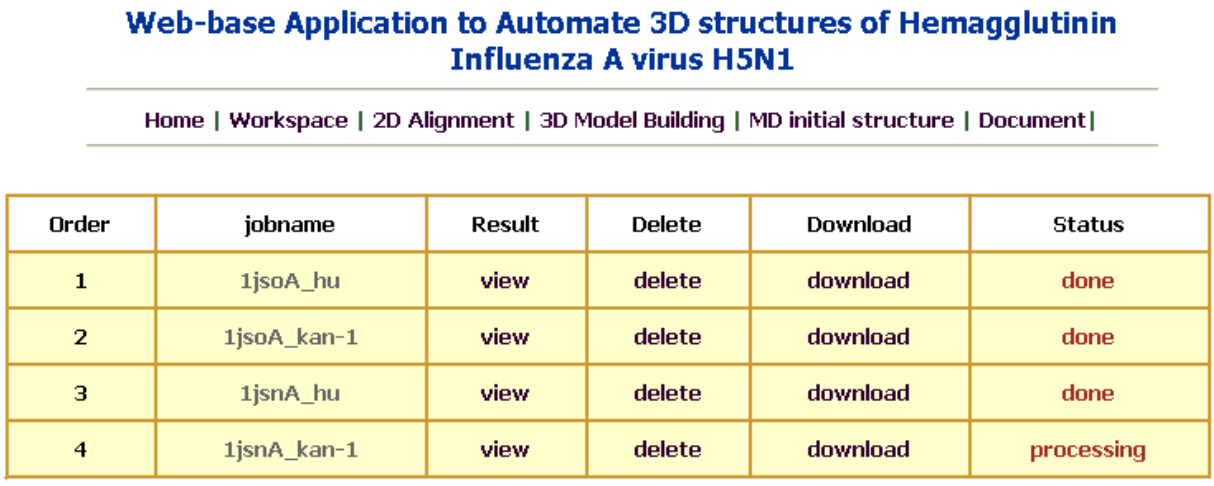
ในหน้าแรกจะมีส่วนสำหรับ login เข้าระบบ โดยผู้ใช้จะต้องลงทะเบียนก่อนการใช้งาน โดยคลิกที่ “new member” จากนั้นจะปรากฏหน้า register.php สำหรับลงทะเบียนดังรูป 5



รูปที่ 5 แสดงหน้าลงทะเบียนสมาชิกใหม่

- workspace เป็นหน้าแสดงผลการทำงานสำหรับผู้ใช้แต่ละคน

เมื่อผู้ใช้ login เข้าระบบแล้วจะพบหน้าแสดงการทำงาน ซึ่งหากได้เคยใช้งานระบบแล้วจะปรากฏรายละเอียดการทำงานดังรูป 6 ในส่วนนี้ผู้ใช้สามารถเรียกดู ลบ และ ดาวน์โหลดข้อมูลได้

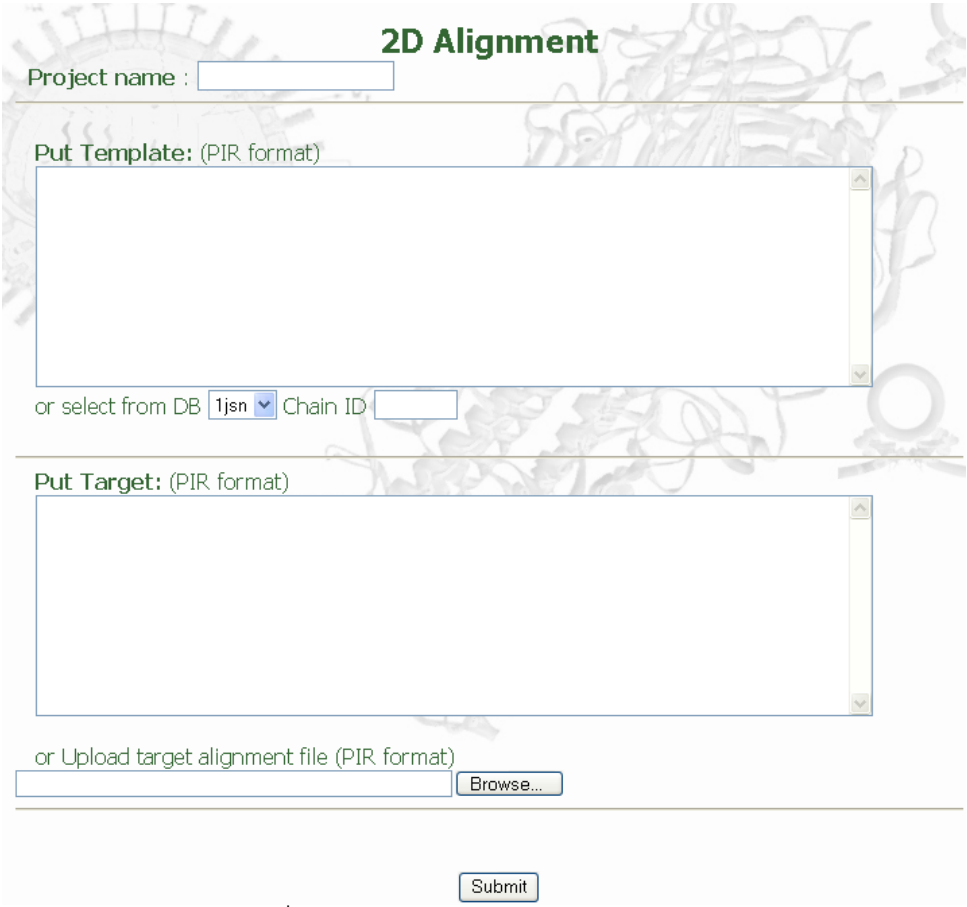


Order	jobname	Result	Delete	Download	Status
1	1jsoA_hu	view	delete	download	done
2	1jsoA_kan-1	view	delete	download	done
3	1jsnA_hu	view	delete	download	done
4	1jsnA_kan-1	view	delete	download	processing

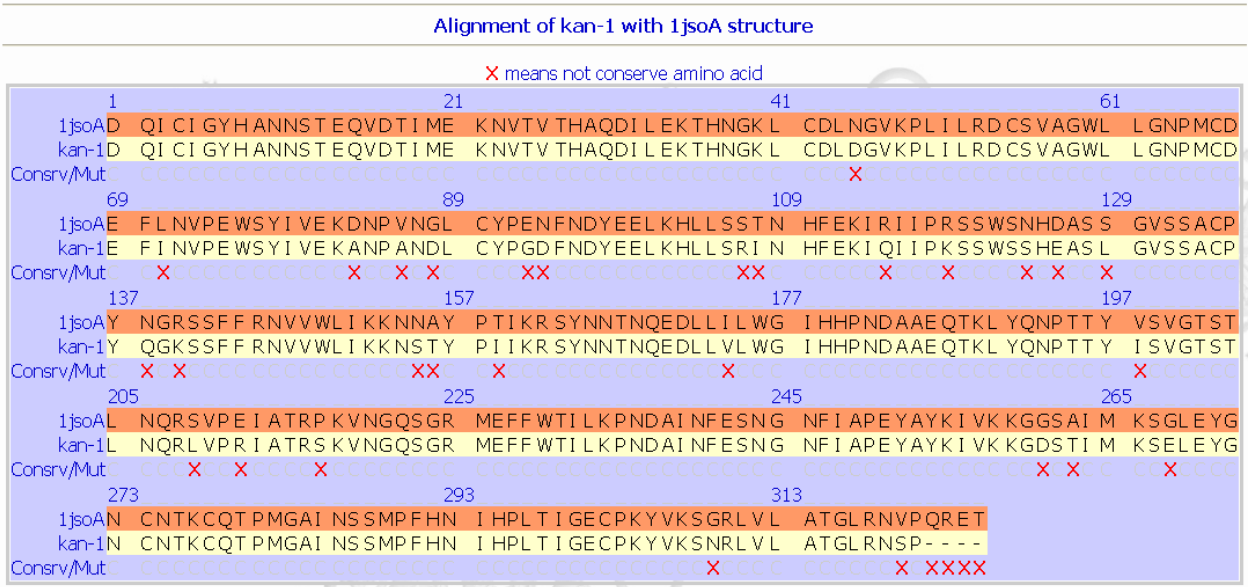
รูปที่ 6 แสดงรายละเอียดการทำงานของผู้ใช้

ในส่วนบนของหน้าต่างการทำงานนี้ มีเมนูการใช้งานหลัก ได้แก่ 2D alignment, 3D model building และ MD initial structure

- 2D alignment เป็นการเทียบลำดับกรดอะมิโนของโปรตีนที่ต้องการกับแม่แบบ (template) เพื่อดูว่ามีตำแหน่งเหมือนหรือต่างกันอย่างไร ซึ่งสามารถเทียบ 1 ต่อ 1 หรือ เทียบหลายๆ sequence (multiple sequence alignment) ก็ได้ หน้าต่างการทำงานแสดงดังรูป 7 และผลแสดงดังรูป 8



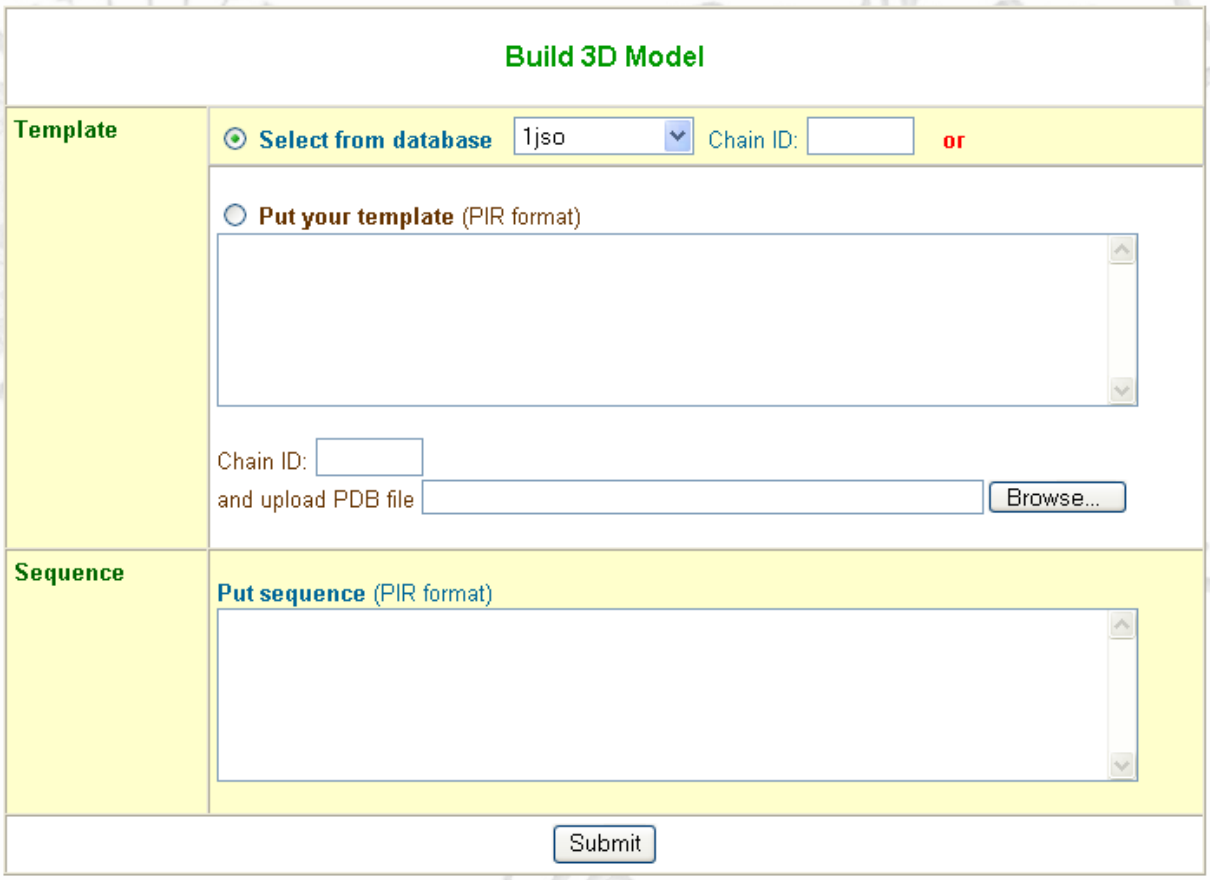
รูปที่ 7 แสดงหน้าจอการทำ 2D alignment



รูปที่ 8 แสดงผลการทำ 2D alignment ของ sequence kan-1 เทียบกับโครงสร้างจาก PDB คือ 1jso chain A

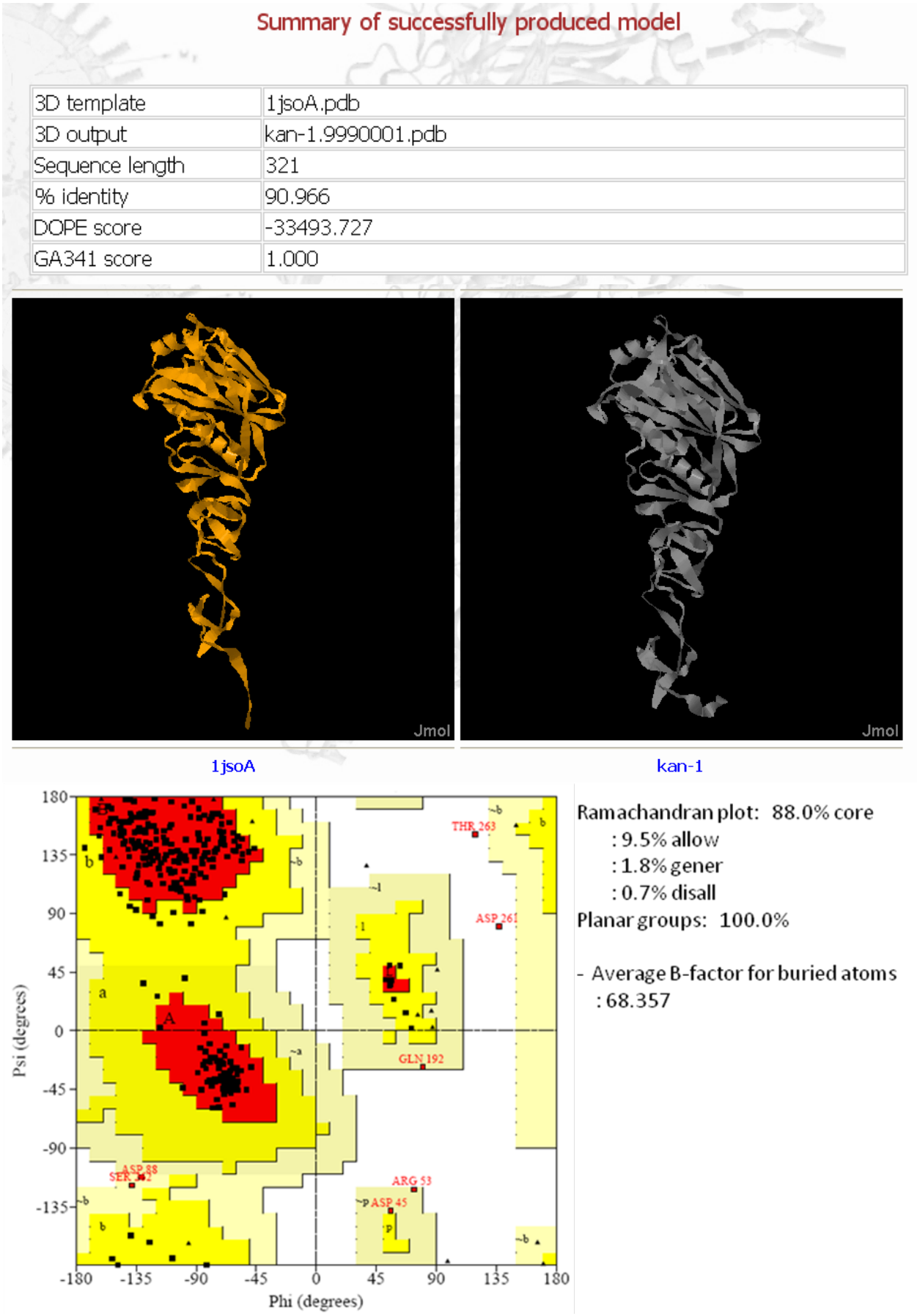
- 3D automate model เป็นการประมาณโครงสร้างสามมิติของ sequence เทียบกับ template โดยหาก sequence ที่สนใจมีความใกล้เคียงกับ template โครงสร้างที่ได้ก็จะมีความสมบูรณ์มาก ค่าพลังงานของระบบที่คำนวณได้จะเป็นลบมาก (DOPE score) ผู้ใช้สามารถเลือก template จากที่มีในฐานข้อมูล หรืออัพโหลดตัวที่สนใจได้ หากเลือก template เอง ผู้ใช้จะต้อง upload ไฟล์โครงสร้างสามมิติในรูปแบบของ pdb เข้าไปด้วยเพื่อใช้เป็น template ในการสร้างโครงสร้างสามมิติ ลักษณะการทำงานแสดงดังรูป 9

-



รูปที่ 9 แสดงหน้าจอการทำ 3D model building

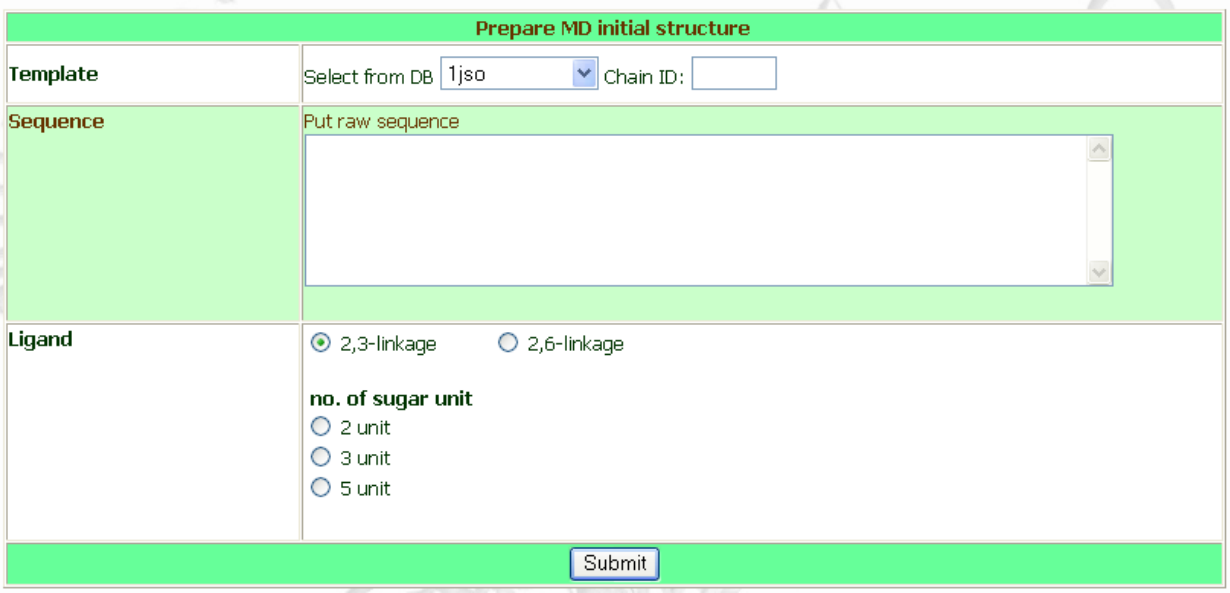
ผลที่ได้จากการประมาณโครงสร้างสามมิติจะแสดงเป็นรูปสามมิติ ด้วยโปรแกรม Jmol เทียบระหว่าง template กับ sequence และให้ข้อมูลด้านโครงสร้าง เช่น %identity, DOPE score [10], GA341 score, ramachandran plot เป็นต้น ผลแสดงดังรูป 10



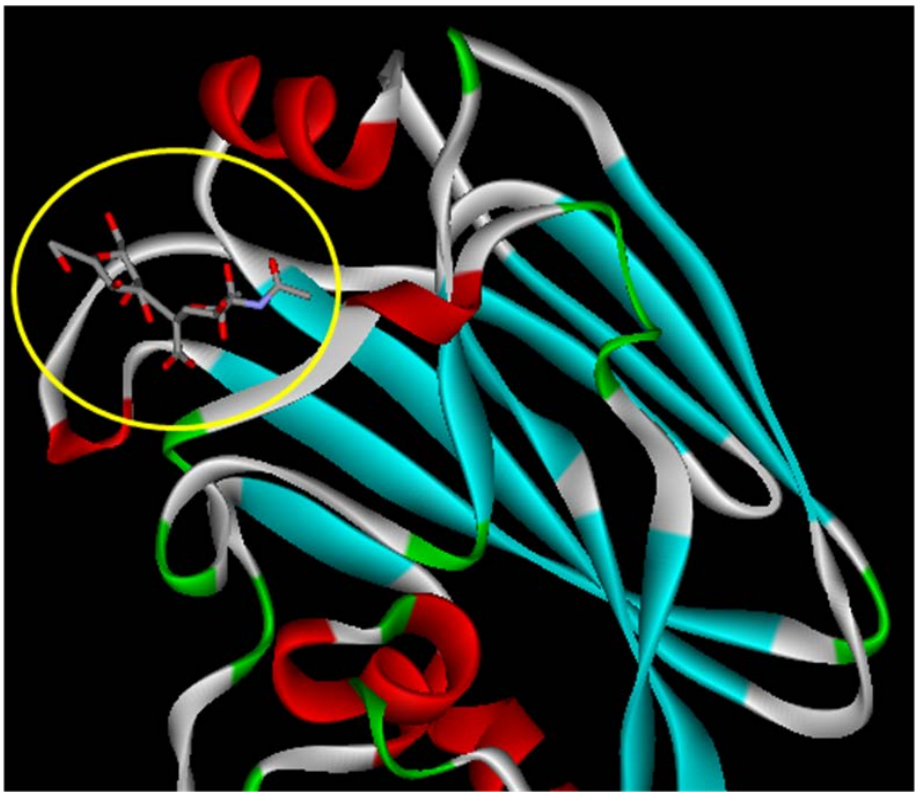
รูปที่ 10 แสดงผลการประมาณโครงสร้างสามมิติ ของ kan-1 จากโครงสร้าง 1jso chain A และข้อมูลทางโครงสร้างที่คำนวณจากโปรแกรม modeler และ procheck

- MD initial structure

เป็นเมนูสำหรับสร้างไฟล์เริ่มต้นในการทำ molecular dynamics simulation (MD) ซึ่งสามารถเลือก ligand แบบ 2,3-linkage หรือ 2,6-linkage จำนวนน้ำตาลแบบ 2 3 หรือ 5 unit ได้ ดังรูป 11 ซึ่งหากเลือกน้ำตาลแบบ 2 unit เมื่อเอาไฟล์ไปเปิดด้วยโปรแกรมแสดงผลจะเห็นว่ามี ligand แบบ 2 unit อยู่ในโครงสร้างบริเวณโพรงการจับ (binding site) ดังรูป 12



รูปที่ 11 หน้าต่างสำหรับสร้างไฟล์เริ่มต้นการทำ MD



รูปที่ 12 แสดงโครงสร้าง kan-1 ที่ใส่ ligand แบบ 2 unit เป็นโครงสร้างเริ่มต้นในการรัน MD

สรุปและเสนอแนะ

ระบบประมาณโครงสร้างสามมิติลำดับกรดอะมิโนของโปรตีนฮีแมกกลูตินินชนิด A สายพันธุ์ย่อย H5N1 บนระบบอินเทอร์เน็ตนี้ พัฒนาขึ้นโดยอาศัยการคำนวณด้วยโปรแกรม modeler ซึ่งทดสอบแล้วว่าให้ความสมบูรณ์ของโครงสร้างดีกว่าเว็บแอพพลิเคชั่นที่นิยมใช้ เช่น swiss model สามารถ upload template ที่สนใจเข้าไปในฐานข้อมูลได้ แม้ว่าในด้านข้อมูลประกอบหรือการแสดงผลอาจจะยังไม่ดีเทียบเท่าแต่ระบบยังสามารถพัฒนาต่อได้ ซึ่งคาดว่าระยะยาวจะมีการพัฒนาให้ระบบมีประสิทธิภาพ ใช้งานง่าย และให้ข้อมูลประกอบมากขึ้น

ในเบื้องต้นนี้ มีการใช้งานทั้งในลักษณะการทำ 2D alignment 3D model และ เตรียมไฟล์เริ่มต้นสำหรับรัน MD ระบบให้ข้อมูลเชิงโครงสร้างที่คำนวณได้จากโปรแกรม modeler และ procheck ที่บอกถึงคุณภาพของโครงสร้างที่สร้างขึ้น แสดงผลการ align ในแบบ 2 มิติ ทำให้เห็นตำแหน่งการกลายพันธุ์ได้ชัดเจน แสดงโครงสร้างสามมิติที่สร้างขึ้นได้หลายรูปแบบโดยอาศัยความสามารถของโปรแกรม Jmol หน่วยวิจัยได้ทดสอบแล้วพบว่าทำให้การดำเนินงานวิจัยด้าน MD มีความสะดวกมากขึ้น นอกจากการศึกษากับ HA แล้วยังสามารถใช้ในการเตรียมโครงสร้างสามมิติของ antibody เพื่อศึกษาการกลายพันธุ์ด้วยระเบียบวิธี MD ได้ด้วย จะเห็นว่าระบบมีความยืดหยุ่น สามารถปรับเปลี่ยนแก้ไข เพิ่มเติม ให้รองรับการดำเนินงานวิจัยในหลายๆ ด้าน ทั้งยังสามารถขยายให้ครอบคลุมโปรตีนชนิดอื่นๆ และสามารถพัฒนาให้มีประสิทธิภาพมากขึ้นได้ ซึ่งนับเป็นประโยชน์จากการที่เราเริ่มพัฒนาระบบขึ้นใช้เอง นำเทคโนโลยีที่มีอยู่มาประยุกต์ใช้ให้เกิดประโยชน์ และพึ่งพาตนเองมากขึ้น ระบบที่พัฒนาขึ้นนี้จะ เป็นประโยชน์ต่อการวิจัยไข้หวัดนกในประเทศ ทั้งสามารถขยายให้เกิดประโยชน์ในการศึกษาโปรตีนอื่นๆ ได้

เอกสารอ้างอิง

1. www.rcsb.org
2. swissmodel.expasy.org
3. www.ncbi.nlm.nih.gov
4. N. Eswar, M. A. Marti-Renom, B. Webb, M. S. Madhusudhan, D. Eramian, M. Shen, U. Pieper, A. Sali. Comparative Protein Structure Modeling With MODELLER. Current Protocols in Bioinformatics, John Wiley & Sons, Inc., Supplement 15, 5.6.1-5.6.30, 2006.
5. G. G. Krivov, M. V. Shapovalov, and R. L. Dunbrack, Jr. Improved prediction of protein side-chain conformations with SCWRL4. *Proteins* (2009)
6. www. swissmodel.expasy.org
7. Ramachandran, G.N., Ramakrishnan, C., and Sasisekharan, V. 1963. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**: 95–99.
8. Laskowski R A, MacArthur M W, Moss D, Thornton J M. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.,* **26**: 283-291.
9. jmol.sourceforge.net
10. M.-y. Shen and A. Sali. Statistical potential for assessment and prediction of protein structures. Protein Science 15, 2507–2524, 2006.

---

(1) *(ภาษาไทย) ภาควิชา / ศูนย์ / สถานี เคมี คณะ / สถาบัน / สำนัก วิทยาศาสตร์*
(ภาษาอังกฤษ) Department of Chemistry, Faculty of Science

หมายเหตุ : ให้ส่งรายงานผลการวิจัยฉบับสมบูรณ์ (ฉบับร่าง) จำนวน 3 ชุด ก่อน โดยสถาบันวิจัยและพัฒนาจะส่งให้ผู้ทรงคุณวุฒิ ประเมิน / วิจารณ์ หากไม่มีการแก้ไข จะแจ้งให้ส่งเพิ่ม แต่หากมีความเห็นข้อเสนอแนะจากผู้ทรงคุณวุฒิให้ปรับแก้ไข จะแจ้งให้ดำเนินการแก้ไข และให้ส่งรายงานผลงานวิจัยฉบับสมบูรณ์ (ฉบับจริง) **จำนวน 12 ชุด** พร้อม Diskette ต่อไป