

บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

ในรอบทศวรรษที่ผ่านมาข้อมูลทางชีววิทยามีการเพิ่มจำนวนขึ้นอย่างรวดเร็ว จนในปัจจุบันมีข้อมูลอยู่ในฐานข้อมูลขนาดใหญ่ทั่วโลกมากมาย ไม่ว่าจะเป็นข้อมูลลำดับนิวคลีโอไทด์หรือข้อมูลลำดับกรดอะมิโน ซึ่งส่วนใหญ่เป็นข้อมูลที่ได้จากการเก็บข้อมูลรหัสพันธุกรรมจากโครงการหาลำดับนิวคลีโอไทด์ โครงการรหัสพันธุกรรมขนาดใหญ่ที่รู้จักกันดีคือโครงการรหัสพันธุกรรมของมนุษย์ (Human Genome Project) นอกจากนี้ยังมีโครงการรหัสพันธุกรรมจากสิ่งมีชีวิตอื่นอีกมากมาย ข้อมูลลำดับเหล่านี้จะเก็บอยู่ในฐานข้อมูลรหัสพันธุกรรมที่มีขนาดใหญ่ เช่น ฐานข้อมูล GenBank [1] และฐานข้อมูล PIR [2-5] ข้อมูลรหัสพันธุกรรมในฐานข้อมูลเป็นข้อมูลดิบที่ยังไม่ได้รับการประมวลผล นอกจากนี้ยังมีกลุ่มนักวิจัยได้ทำการประมวลผลและจัดเก็บผลไว้ในฐานข้อมูลเฉพาะกลุ่มเช่น UniProt [6-8] SISYPHUS [9] หรือ HOMSTRAD [10] นอกจากนี้ยังมีงานวิจัยอีกหลายด้านที่ศึกษาเกี่ยวกับโปรตีน ไม่ว่าจะเป็นการวิเคราะห์หาความเหมือนของลำดับ (Sequence Similarity) การวิเคราะห์หาโครงสร้างปฐมภูมิ (Primary Structure) โครงสร้างทุติยภูมิ (Secondary Structure) โครงสร้างตติยภูมิ (Tertiary Structure) และโครงสร้างจตุรภูมิ (Quaternary Structure) รวมทั้งวิเคราะห์หาหน้าที่การทำงานของโปรตีน (Protein Function)

อย่างไรก็ตาม การวิเคราะห์หาหน้าที่ของโปรตีนที่ใช้กันอยู่ในปัจจุบันไม่สามารถวิเคราะห์หาหน้าที่การทำงานของโปรตีนได้อย่างมีประสิทธิภาพ ยังต้องอาศัยวิธีการวิเคราะห์จากห้องปฏิบัติการทางเคมีหรือห้องปฏิบัติการทางชีววิทยา ซึ่งการวิเคราะห์จากห้องปฏิบัติการจำเป็นต้องใช้งบประมาณสูงและใช้เวลานานเป็นอย่างมาก ดังนั้นเพื่อลดค่าใช้จ่ายของการหาหน้าที่การทำงานของโปรตีน นักวิทยาศาสตร์จึงใช้วิธีการวิเคราะห์หาหน้าที่การทำงานของลำดับกรดอะมิโนแทน โดยใช้เครื่องมือจับคู่เพื่อการเปรียบเทียบความเหมือนของกรดอะมิโนกับลำดับกรดอะมิโนในฐานข้อมูลที่ทราบหน้าที่การทำงาน

ในปัจจุบันเครื่องมือจับคู่ลำดับนิวคลีโอไทด์หรือลำดับกรดอะมิโนยังไม่มีความสามารถเพียงพอสำหรับการหาความเหมือนของโปรตีนที่ทำหน้าที่เหมือนกันในสิ่งมีชีวิตที่

แตกต่างกัน เนื่องจากวิวัฒนาการของสิ่งมีชีวิตรุ่นลูกอาจมีการเปลี่ยนแปลงลำดับพันธุกรรมไปจากพ่อแม่ แล้วทำให้เกิดช่องว่าง (Gap) ของลำดับกรดอะมิโน และเมื่อนำลำดับกรดอะมิโน จากสิ่งมีชีวิต 2 ชนิดที่มีหน้าที่เดียวกันมาเปรียบเทียบกันแบบ 1 มิติ พบว่าผลการทดลองให้ค่าความเหมือนของลำดับที่มีค่าต่ำ (Low Sequence Identity) ทำให้ไม่สามารถสรุปได้ว่าเป็นโปรตีนที่ทำหน้าที่เหมือนกัน ตัวอย่างเช่น การนำลำดับกรดอะมิโนฮีโมโกลบินของแมลงกับลำดับกรดอะมิโนฮีโมโกลบินของหนอนตัวกลมมาจับคู่กัน ซึ่งปรากฏว่าได้ค่าความเหมือนของลำดับเพียง 19 เปอร์เซ็นต์เท่านั้น จึงไม่สามารถสรุปได้ว่าลำดับกรดอะมิโนทั้ง 2 สายลำดับนี้ทำหน้าที่เหมือนกัน

ต่อมากลุ่มผู้วิจัยบางกลุ่มได้นำการวิเคราะห์กลุ่มที่ไม่ชอบน้ำมาใช้เปรียบเทียบความเหมือนของโปรตีน การวิเคราะห์นี้เป็นการวิเคราะห์กลุ่มของกรดอะมิโนที่ไม่ชอบน้ำที่อยู่ในรูปแบบแผนภูมิ 2 มิติ โดยการแสดงรูปร่างลักษณะของกลุ่มที่ไม่ชอบน้ำให้เห็นได้ชัดเจนยิ่งขึ้น ซึ่งใช้สำหรับหาโครงสร้างทุติยภูมิของโปรตีน [11, 12] แต่ผู้วิจัยบางกลุ่มได้นำการวิเคราะห์นี้ไปใช้เพื่อจับคู่ลำดับกรดอะมิโน [13-16] พบว่าผลการหาความเหมือนของลำดับให้ค่าที่ดีกว่าการจับคู่สายอักขระ แต่การเปรียบเทียบกลุ่มที่ไม่ชอบน้ำเหล่านี้ยังต้องให้การจับคู่กลุ่มโดยผู้เชี่ยวชาญ ซึ่งต้องอาศัยความชำนาญและทักษะความรู้เป็นอย่างมาก ในขณะนี้ยังไม่มีเครื่องมือใดที่ช่วยให้สามารถจับคู่กลุ่มเหล่านี้ได้อย่างอัตโนมัติ ดังนั้นวิธีการจับคู่ลำดับกรดอะมิโนโดยอาศัยการวิเคราะห์กลุ่มที่ไม่ชอบน้ำจึงยังไม่ได้ได้รับความนิยม

ในงานวิจัยฉบับนี้ ผู้วิจัยได้พัฒนาและเสนอเครื่องมือแสดงผลแผนภูมิ 2 มิติ เพื่อใช้สำหรับการศึกษากลุ่มที่ไม่ชอบน้ำ วิธีการแทนข้อมูลจากลำดับกรดอะมิโนใน 1 มิติเป็นรูปแบบแผนภูมิ 2 มิติที่สามารถนำไปใช้จับคู่กลุ่ม และเครื่องมือจับคู่กลุ่มที่ไม่ชอบน้ำแบบ 2 มิติโดยอัตโนมัติเพื่อใช้สำหรับการเปรียบเทียบหาโปรตีนที่ทำหน้าที่เหมือนกัน

1.2 วัตถุประสงค์ของการวิจัย

1. เพื่อสร้างรูปแบบของข้อมูลก่อนการจับคู่กลุ่มที่ไม่ชอบน้ำ
2. เพื่อนำเสนอแนวทางและสร้างเครื่องมือหาหน้าที่การทำงานของโปรตีนแบบอัตโนมัติ

1.3 ขอบเขตของการวิจัย

1. สร้างเครื่องมือจับคู่กลุ่มที่ไม่ชอบน้ำแบบ 2 มิติโดยอัตโนมัติ โดยอาศัยวิธีกำหนดการพลวัตในการหาคู่กลุ่มที่เหมือนกัน
2. ทดสอบความถูกต้องระหว่างวิธีการจับคู่กลุ่มโดยอัตโนมัติที่สร้างขึ้นกับวิธีการจับคู่กลุ่มที่ไม่ชอบน้ำโดยผู้เชี่ยวชาญ โดยใช้ชุดข้อมูลเดียวกัน
3. ทดสอบความถูกต้องของการเลือกลำดับกรดอะมิโนที่เหมือนมากที่สุดจากวิธีการเปรียบเทียบแบบ 2 มิติ เทียบกับการเลือกลำดับกรดอะมิโนที่เหมือนมากที่สุดจากวิธีการเปรียบเทียบแบบ 1 มิติ โดยใช้ชุดข้อมูลที่ได้จาก PIR และ HOMSTRAD ที่มีการระบุหน้าที่การทำงานแล้ว

1.4 ขั้นตอนของการวิจัย

1. ศึกษาวิธีการจัดโครงสร้างของโปรตีน
2. ศึกษาวิธีการวิเคราะห์กลุ่มที่ไม่ชอบน้ำ
3. ศึกษาวิธีการจับคู่แบบกำหนดการพลวัต
4. รวบรวมชุดข้อมูลเพื่อใช้ในการทดสอบความถูกต้องของเครื่องมือ
5. พัฒนาเครื่องมือแสดงผลแผนภูมิ 2 มิติ
6. พัฒนาการแทนข้อมูลเพื่อใช้สำหรับการจับคู่กลุ่มที่ไม่ชอบน้ำ
7. พัฒนาเครื่องมือจับคู่กลุ่มที่ไม่ชอบน้ำแบบ 2 มิติโดยอัตโนมัติ
8. ทดสอบความถูกต้องของเครื่องมือจับคู่กลุ่มที่ไม่ชอบน้ำแบบ 2 มิติโดยอัตโนมัติ
9. วิเคราะห์และสรุปผลผลการทดลอง
10. เรียบเรียงวิทยานิพนธ์

1.5 ประโยชน์ที่ได้รับ

1. ได้เครื่องมือจับคู่กลุ่มที่ไม่ชอบน้ำแบบ 2 มิติโดยอัตโนมัติ เพื่อใช้สำหรับหาความเหมือนของโปรตีนที่มีหน้าที่การทำงานเหมือนกันในสิ่งมีชีวิตที่แตกต่างกัน
2. ได้การแทนข้อมูลสำหรับการแทนกลุ่มที่ไม่ชอบน้ำ เพื่อใช้ในเครื่องมือจับคู่กลุ่มที่ไม่ชอบน้ำแบบ 2 มิติโดยอัตโนมัติ
3. ได้เครื่องมือแสดงผลแผนภูมิ 2 มิติ ที่สร้างตามหลักการวิเคราะห์กลุ่มที่ไม่ชอบน้ำ

1.6 โครงสร้างของวิทยานิพนธ์

เนื้อหาของวิทยานิพนธ์ฉบับนี้แบ่งออกเป็น 5 บท ดังนี้คือ บทที่ 1 เป็นบทนำ บทที่ 2 กล่าวถึงทฤษฎีและงานวิจัยที่เกี่ยวข้องกับงานวิจัยชิ้นนี้ บทที่ 3 กล่าวถึงการดำเนินงานวิจัย โดยอธิบายเป็นขั้นตอนต่างๆ อย่างละเอียด ส่วนในบทที่ 4 เป็นการทดลองและผลที่ได้จากการทดลองตามชุดข้อมูลต่างๆ และบทที่ 5 เป็นการสรุปผลการทดลองและข้อเสนอแนะของงานวิจัย ซึ่งอาจเป็นประโยชน์ต่องานวิจัยอื่นๆ ในอนาคต

1.7 ผลงานตีพิมพ์จากวิทยานิพนธ์

ส่วนหนึ่งของงานวิทยานิพนธ์นี้ได้รับการตีพิมพ์เป็นบทความวิชาการในหัวข้อเรื่อง "Towards 2-D Automatic Hydrophobic Cluster Alignment" โดย ภิสัทธี กรรณสูต รัฐ พิชญางกูร และ โชติรัตน์ รัตนามัทธนะ ในงานประชุมวิชาการ "ประชุมวิชาการวิศวกรรมชีวการแพทย์แห่งชาติครั้งที่ 5" ซึ่งจัดขึ้น ณ โรงแรมทวินทาวเวอร์ กรุงเทพมหานคร ประเทศไทย ในวันที่ 8 กรกฎาคม 2550 ดังภาคผนวก ข หน้า 79 และส่วนหนึ่งของวิทยานิพนธ์นี้ได้รับการตอบรับการตีพิมพ์วารสารทางวิชาการระดับนานาชาติในหัวข้อเรื่อง "Automatic 2-D Hydrophobic Cluster Alignment" โดย ภิสัทธี กรรณสูต รัฐ พิชญางกูร และ โชติรัตน์ รัตนามัทธนะ ในวารสารทางวิชาการระดับนานาชาติ "International Journal of Biomedical Engineering and Technology (IJ BET)" สำนักพิมพ์ Inderscience ได้รับการตอบรับการตีพิมพ์วันที่ 19 กันยายน 2550 ดังภาคผนวก ข หน้า 85