Thesis Title Prediction of Disulfide Bond Formation Using Graphical Models

Author Mr. Watshara Shoombuatong

Degree Master of Science (Bioinformatics)

Thesis Advisory Committee Asst. Prof. Dr. Piyarat Nimmanpipug Chairperson
Asst. Prof. Dr. Jeerayut Chaijaruwanich Member
Assoc. Prof. Dr. Chatchai Tayapiwatana Member
Asst. Prof. Dr. Sukon Prasitwattanaseree Member

ABSTRACT

The formation of disulfide bond between cysteines plays a major role in protein folding, structure, function, and evolution. In addition, it conveys important information about the protein conformation and including helps toward the solution of the folding problem. Therefore, prediction of disulfide bond formation prediction is the primary key of proteins stability problem. As the determinations of bonding states of cysteines involve in the use of costly chemicals and equipments, computational approaches are very important to help in decreasing both costs and labors. Our assumptions are that the formation of disulfide bonds got influence from its

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

neighboring amino acids as well as their secondary structure and solvent accessibility. In this study, these factors have been accounted for the prediction of disulfide bond formation and disulfide bonding state of cysteines in proteins using the graphical models which are hidden Markov models (HMMs) and conditional random fields (CRFs). HMMs are used to predict disulfide bond formation from protein sequences. CRFs are used to predict disulfide bonding of cysteines. The classification of bonding state was successfully obtained at 84.4% accuracy, while the prediction of disulfide bond formation is close to 51.9% accuracy.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ชื่อเรื่องวิทยานิพนธ์ การทำนายการเกิดพันธะไดซัลไฟด์โดยใช้แบบจำลองเชิงกราฟ

ผู้เขียน นาย วัชระ ชุ่มบัวตอง

ปริญญา วิทยาศาสตรมหาบัณฑิต (ชีวสารสนเทศศาสตร์)

คณะกรรมการที่ปรึกษาวิทยานิพนธ์ ผศ .ดร. ปิยรัตน์ นิมมานพิภักดิ์ ประธานกรรมการ

ผศ.ดร. จีรยุทธ ไชยจารุวณิช กรรมการ

รศ.ดร. ชัชชัย ตะยาภิวัฒนา กรรมการ

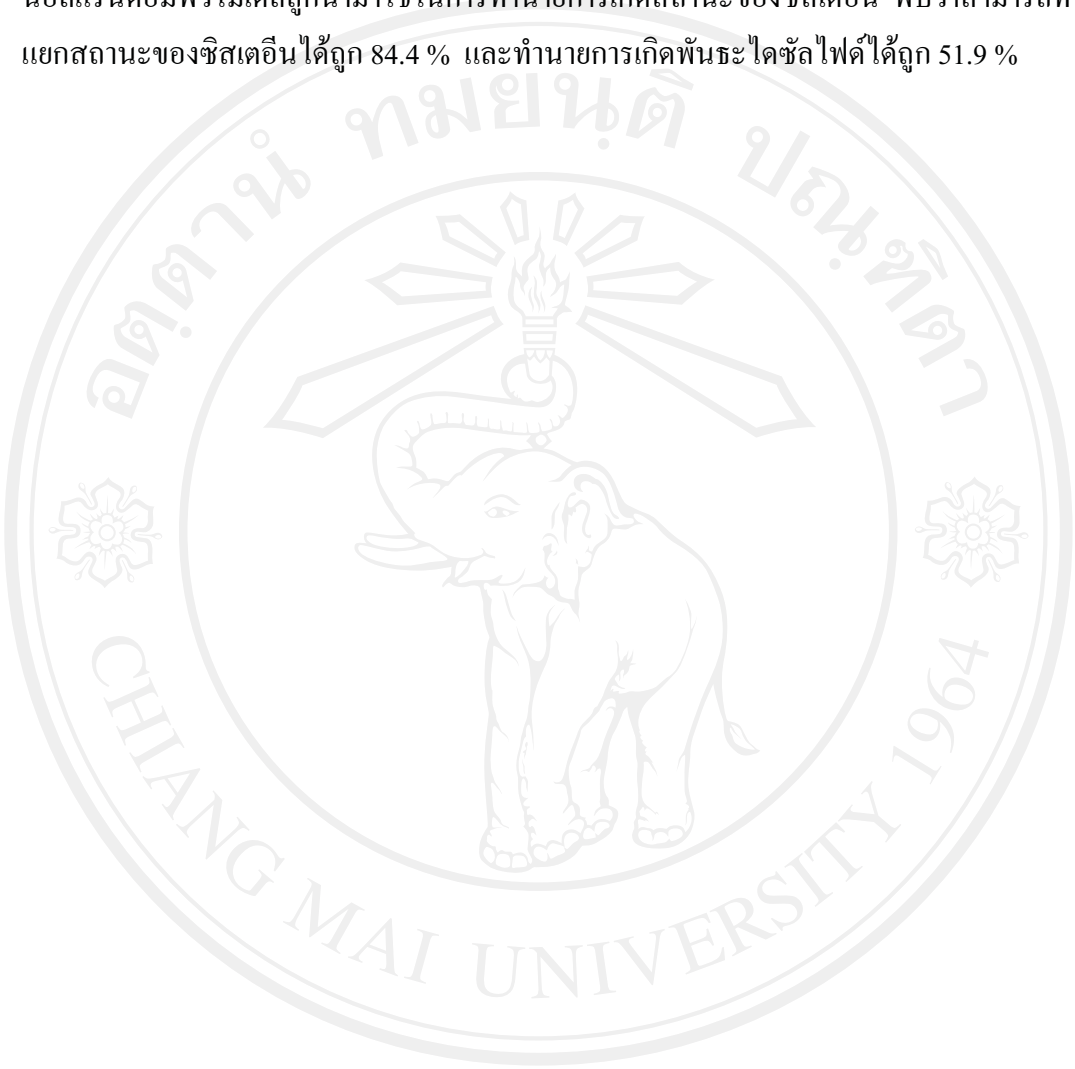
ผศ.ดร. สุคนธ์ ประสิทธิ์วัฒนเสรี กรรมการ

บทคัดย่อ

การสร้างพันธะไดซัลไฟด์ระหว่างซิสเตอีนมีบทบาทความสำคัญในการม้วนตัว , โครงสร้าง หน้าที่ และวิวัฒนาการของโปรตีน รวมถึงมีความสำคัญกับการแก้ปัญหาการม้วนตัวของโปรตีน ดังนั้นปัญหาการทำนายการสร้างพันธะไดซัลไฟด์เป็นกุญแจสำคัญสำหรับความเสถียรภาพของโครงสร้างโปรตีน สำหรับการทดลองเพื่อที่จะหาตำแหน่งของพันธะไดซัลไฟด์นั้นต้องใช้สารเคมีและอุปกรณ์ที่มีราคาแพง จากเหตุผลดังกล่าวการอาศัยวิธีทางการคำนวณจะช่วยลดค่าใช้จ่ายดังกล่าวได้มาก โดยผู้ทำการวิจัยมีสมมติฐานที่ว่า การสร้างพันธะไดซัลไฟด์ได้รับอิทธิพลจากกรดอะมิโนที่อยู่รอบข้างซิสเตอีน รวมไปถึงโครงสร้างทุติยภูมิและ ความสามารถของตัวทำละลายที่จะเข้าถึง หรือสัมผัสกับบริเวณหนึ่งๆ ของตัวถูกละลาย โดยปัจจัยดังกล่าวถูกนำมาใช้ในการทำนายการเกิดไดซัลไฟด์จากลำดับโปรตีน และ การทำนายการเกิดสถานะของซิสเตอินโดยใช้แบบจำลองเชิงกราฟมาประยุกต์ใช้กับปัญหาการทำนายการเกิดพันธะไดซัลไฟด์ โดยฮิดเดนมาร์คอฟโมเดลถูกนำมาใช้ในการทำนายการเกิดไดซัลไฟด์จากลำดับโปรตีน และคอนดิชั่น

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

นอลแรนดอมฟิวโมเดลถูกนำมาใช้ในการทำนายการเกิดสถานะของซิสเตอิน พบว่าสามารถที่จะแยกสถานะของซิสเตอีนได้ถูก 84.4 % และทำนายการเกิดพันธะไดซัลไฟด์ได้ถูก 51.9 %



ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved