Thesis Title Development of Selecting Criteria for 3-Dimensional Poses of DARPin and CD4 Complex by ZDOCK Molecular Docking Histogram Analysis

Author Miss Tanchanok Wisitponchai

Degree Master of Engineering (Biomedical Engineering)

Thesis Advisory Committee

Assoc. Prof. Dr. Chatchai Tayapiwatana Advisor

Asst. Prof. Dr. Piyarat Nimmanpipug Co-advisor

ABSTRACT

Human immunodeficiency virus (HIV) inherits the active mutation, which promotes the viral survival in human and is the main hurdle of vaccine development. The mutations on HIV surface i.e. CD4 binding site on HIV gp120, cause the sneaking from the immune-surveillance but still retain the infectivity for CD4+ T cells. Formerly, Designed Ankyrin Repeat Protein (DARPin) technology has been innovated to inhibit the HIV infection. The CD4-specific DARPin specifically excludes the adhesion step of HIV to CD4 molecule and efficiently prevents HIV infection *in vitro* by competing with gp120. Moreover, the basis T cell functions are not disturbed, although DARPin can block binding of CD4 to MHCII. However, HIV still infects to CD4+ T cells in animal model maybe because of weak interaction

between CD4 and DARPin. Therefore, insight the interaction between CD4 and DARPin molecules may provide the information to improve the interaction affinity of DARPin for *in vivo* purpose. In this study, the hot spots on CD4 bound to DARPin 23.2 were identified for clarifying the interaction of them. The computational docking simulation, ZDOCK protocal, was available for generating the CD4-DARPin 23.2 complex structures. The 11 rational complex acheivement was convinced from the docking scores and the overlapping area on CD4 between DARPin 23.2 and gp120 as well as MHCII. All of 11 poses were carried out to identify the hot spots using computer programming to manage the complicated data. The desigining algorithm manipulated ZDOCK databases into histogram values based on five criteria. The combination of five criterion was available to identify the hot spots. The results showed that Lys35, Gln33, Gln40, and Arg59 are hot spots of CD4 to DARPin 23.2. Especially, Lys35 was the most important hot spot because (1) it located on critical residue on CD4 recognizing gp120 and also MHCII studied by biochemical mutagenesis; and (2) it had maximum histogram value of criteria combination in two out of three good complexes (p16, p26, and p1513). Our algorithm was confirmedly reliable for identifying hot spots because our prediction accorded to two servers which exhibiting different method with ours. The accordance was that hot spots of CD4 bound to DARPin predicted by three methods were somewhat hydrophilic residues. Moreover, our prediction has high identity residue with both methods by calculating the $PID^{predict}$. The advantage of this study is not only getting hot spots of CD4 to DARPin but also exhibiting the procedure to filter the rational complex and creating a new algorithm for identifying hot spots.

ชื่อเรื่องวิทยานิพนธ์ การพัฒนาเกณฑ์การคัดเลือกโครงสร้างสามมิติของดาร์ปินที่จับกับซีดีโฟร์ โดยการวิเคราะห์ข้อมูลฮีสโตแกรมจากซีต็อกโมเลกูลาร์ต็อกกิง

ผู้เขียน นางสาวธันย์ชนก วิศิษฐ์พลชัย

ปริญญา วิศวกรรมศาสตรมหาบัณฑิต (วิศวกรรมชีวการแพทย์)

คณะกรรมการที่ปรึกษาวิทยานิพนธ์

รศ.ดร.ชัชชัย ตะยาภิวัฒนา	อาจารย์ที่ปรึกษาหลัก
ผศ.ดร.ปิยรัตน์ นิมมานพิภักดิ์	อาจารย์ที่ปรึกษาร่วม

บทคัดย่อ

เชื้อไวรัสเอชไอวีมีความสามารถในการกลายพันธุ์ตลอดเวลาทำให้เชื้อดำรงอยู่ในร่างกายมนุษย์ได้และเป็นอุปสรรคสำคัญต่อการพัฒนาวัคซีนต่อเชื้อเอชไอวี การกลายพันธุ์ที่ผิวเซลล์ของเอชไอวี เช่น บริเวณที่ใช้จับกับโมเลกุลซีดีโฟร์บนจีพีหนึ่งร้อยยี่สิบของเอชไอวี จะทำให้เชื้อหลบหลีกการตรวจจับของระบบภูมิคุ้มกันและยังคงสามารถติดเชื้อในทีเซลล์ที่มีโมเลกูลของซีดีโฟร์บนผิวเซลล์ได้ ในช่วงเวลาที่ผ่านมาได้มีการใช้โปรตีนดาร์ปินซึ่งเป็นวิทยาการผลิตใหม่เพื่อที่จะยับยั้งการเข้าสู่เซลล์เป้าหมายของเชื้อเอชไอวี สำหรับโปรตีนดาร์ปินที่จำเพาะต่อโมเลกุลซีดีโฟร์นั้นจะเข้าไปจับกับโมเลกุลซีดีโฟร์โดยเข้าไปแย่งจับกับจีพีหนึ่งร้อยยี่สิบของตัวเชื้อเอชไอวี ส่งผลให้เกิดการขัดขวางการจับกันของเอชไอวีกับโมเลกุลซีดีโฟร์บนผิวของทีเซลล์ซึ่งการทดลองนี้เกิดขึ้นในหลอดทดลอง นอกจากนี้ทีเซลล์ยังคงทำงานได้เป็นปกติแม้ว่าการจับกันของดาร์ปินกับซีดีโฟร์จะไปขัดขวางการจับกันของซีดีโฟร์กับเอ็มเอชซีชนิดสอง แต่อย่างไรก็ตามเมื่อนำไปทดสอบในสัตว์ทดลองพบว่าดาร์ปินไม่สามารถป้องกันการติดเชื้อในทีเซลล์ที่มีซีดีโฟร์ได้ สาเหตุอาจเป็นเพราะซีดีโฟร์และดาร์ปินจับกันไม่แข็งแรง ดังนั้นความเข้าใจลักษณะการจับกันระหว่างโปรตีนดาร์ปินและโมเลกุลซีดีโฟร์อย่างถ่องแท้จะเป็นข้อมูลสำคัญเพื่อช่วยเพิ่มความชอบในการจับของโปรตีนแองไครินรีพีทต่อโมเลกุลซีดีโฟร์ให้ดีขึ้นสำหรับการทดลองในสัตว์ งานวิจัยนี้จะทำการค้นหากรดอะมิโนที่สำคัญบนซีดีโฟร์ที่ใช้จับกับดาร์ปินยี่สิบสามจุดสองเพื่อใช้ในการอธิบายการจับ

กันของซีดีโฟร์และดาร์ปินให้กระจ่างชัดขึ้น วิธีการทางคอมพิวเตอร์จะใช้การจำลองด็อกกิงด้วยซีด็อกโปรโตคอลในการทำนายลักษณะการจับกันของโครงสร้างซีดีโฟร์และโปรตีนดาร์ปิน จากการทดลองจะได้ลักษณะการจับกันของสองโปรตีนจำนวนสิบเอ็ดโครงสร้าง ซึ่งทั้งสิบเอ็ดโครงสร้างนี้มีความน่าเชื่อถือเนื่องจากมีบริเวณพื้นที่ทับซ้อนบนซีดีโฟร์ระหว่างดาร์ปินยี่สิบสามจุดสองกับจีพีหนึ่งร้อยยี่สิบและดาร์ปินกับเอ็มเอชซีชนิดสอง จากนั้นโครงสร้างทั้งสิบเอ็ดโครงสร้างนี้จะถูกนำไปหากรดอะมิโนที่สำคัญบนซีดีโฟร์โดยการใช้ศาสตร์ที่เกี่ยวกับการเขียนโปรแกรมเข้ามาจัดการข้อมูลที่ซับซ้อน แอลกอริทึมที่ออกแบบขึ้นมานี้จะเป็นการนำข้อมูลของซีด็อกมาใช้ในการสร้างเกณฑ์โดยให้อยู่ในรูปของฮีสโตแกรม จากนั้นเกณฑ์ทั้งห้าที่ถูกสร้างขึ้นจะถูกนำมาจัดกลุ่มใหม่เพื่อใช้ในการหากรดอะมิโนที่สำคัญของซีดีโฟร์ ผลจากการทดลองพบว่ากรดอะมิโนชนิดไลซีนตำแหน่งที่สามสิบห้า กลูตามีนตำแหน่งที่สามสิบสาม กลูตามีนตำแหน่งที่สี่สิบ และอาร์จินีนตำแหน่งที่ห้าสิบเก้าเป็นกรดอะมิโนที่สำคัญบนซีดีโฟร์ที่ใช้จับกับดาร์ปินยี่สิบสามจุดสองโดยเฉพาะอย่างยิ่งกรดอะมิโนไลซีนตำแหน่งสามสิบห้าถือว่าเป็นกรดอะมิโนที่สำคัญที่สุดเนื่องจาก (1) กรดอะมิโนนี้เป็นกรดอะมิโนที่สำคัญบนซีดีโฟร์ที่ใช้จับกับจีพีหนึ่งร้อยยี่สิบและยังใช้จับกับเอ็มเอชซีชนิดสองซึ่งเป็นการศึกษาจากการทำมิวเตชัน (2) กรดอะมิโนนี้มีค่าฮีสโตแกรมที่มากที่สุดของกลุ่มของเกณฑ์ทั้งห้าในจำนวนสองในสามของโครงสร้างที่ดี (พี16, พี26, และพี1513) แอลกอริทึมของเราได้รับการยืนยันว่ามีความน่าเชื่อถือในการหากรดอะมิโนที่สำคัญเนื่องจากผลการทำนายของเราสอดคล้องกับอีกสองโปรแกรมที่มีวิธีการที่ต่างจากแอลกอริทึมนี้ สิ่งที่สอดคล้องของสามวิธีการคือกรดอะมิโนที่สำคัญบนซีดีโฟร์ที่จับกับดาร์ปินถูกทำนายว่าเป็นกรดอะมิโนชนิดไฮโดรฟิลิก นับได้ว่าผลการทดลองของเรามีความน่าเชื่อถือในการหากรดอะมิโนที่สำคัญเนื่องจากมีความสอดคล้องกับอีกสองโปรแกรมค่อนข้างสูง จากการศึกษาครั้งนี้นอกจากจะได้ทราบกรดอะมิโนที่สำคัญที่ซีดีโฟร์ใช้ในการจับกับดาร์ปินแล้ว ยังเป็นการนำเสนอวิธีการที่ใช้ในการคัดเลือกหาโครงสร้างที่สมเหตุสมผล และได้แอลกอริทึมใหม่สำหรับการหากรดอะมิโนที่สำคัญ