



รายงานวิจัยฉบับสมบูรณ์

โครงการ “การใช้คอมพิวเตอร์ออกแบบและค้นหาสารต้านไวรัสเดงกี”

Computational design and discovery of dengue virus inhibitors

โดย ดร. เอกชัย เจนวิถีสุข

กุมภาพันธ์ 2554

สัญญาเลขที่ DBG5180013

รายงานวิจัยฉบับสมบูรณ์

โครงการ “การใช้คอมพิวเตอร์ออกแบบและค้นหาสารต้านไวรัสเดงกี”

Computational design and discovery of dengue virus inhibitors

ดร. เอกชัย เจนวิถีสุข

สังกัด ศูนย์พันธุวิศวกรรมและเทคโนโลยีชีวภาพแห่งชาติ

สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ

สนับสนุนโดยสำนักงานคณะกรรมการการอุดมศึกษา

และสำนักงานกองทุนสนับสนุนการวิจัย

(ความเห็นในรายงานนี้เป็นของผู้วิจัย สกอ. และ สกว. ไม่จำเป็นต้องเห็นด้วยเสมอไป)

กิตติกรรมประกาศ

คณะผู้วิจัยขอขอบคุณสำนักงานกองทุนสนับสนุนการวิจัย (สกว.) ที่ให้การสนับสนุนเงินงบประมาณ และ ศูนย์พันธุวิศวกรรมและเทคโนโลยีชีวภาพแห่งชาติ, สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติที่ให้การสนับสนุนทางด้านเครื่องมือและอุปกรณ์ในการวิจัยนี้

Project Code: DBG5180013

Project Title: การใช้คอมพิวเตอร์ออกแบบและค้นหาสารต้านไวรัสเดงกี

Investigator: น.สพ.ดร. เอกชัย เจนวิถีสุข

ศูนย์พันธุวิศวกรรมและเทคโนโลยีชีวภาพแห่งชาติ

สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ

113 อุทยานวิทยาศาสตร์ประเทศไทย

ถ.พหลโยธิน อ.คลองหลวง จ.ปทุมธานี 12120

E-mail Address: ekachai@biotec.or.th

Project Period: 30 มิถุนายน 2551 ถึง 30 มิถุนายน 2554

การติดเชื้อไวรัสโรคไขเลือดออกเป็นปัญหาด้านสาธารณสุขที่สำคัญของไทย เนื่องจากยังไม่มีวัคซีนและยารักษาต้านไวรัสที่ออกฤทธิ์ยับยั้งไวรัสเดงกีที่ได้ผล ทำให้มีจำนวนผู้ติดเชื้อและผู้เสียชีวิตในแต่ละปีเป็นจำนวนมาก งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาวิธีการทางคอมพิวเตอร์สำหรับ (1) ออกแบบเปปไทด์เพื่อยับยั้งการทำงานของโปรตีนเอนวิลอปของไวรัสเดงกี และ (2) ค้นหาหรือสารเคมีเพื่อยับยั้งการทำงานของโปรตีนเอนวิลอป, เอ็นเอสสามโปรตีนเอส, เฮลิเคสและเอ็นเอสห้าเมทิลทรานส์เฟอเรสของไวรัสเดงกีทั้งสี่ซีโรทัยป์ จากการศึกษาในโครงการนี้ผู้วิจัยได้พัฒนาวิธีการทางคอมพิวเตอร์โดยใช้ genetic algorithm พบว่าสามารถใช้คำนวณเพื่อค้นหาเปปไทด์สายสั้นได้เร็วกว่าวิธี monte carlo simulation ซึ่งเป็นวิธีดั้งเดิม นอกจากนี้ ผู้วิจัยได้พัฒนาวิธีการสกรีนหาสารเคมีเพื่อยับยั้งการทำงานของโปรตีนของไวรัสเดงกีโดยการใช้วิธีการกรองลิแกนด์ที่มีโครงสร้างที่ไม่เหมาะสมออกจาก screening ligand set ก่อนการสกรีนยาด้วยวิธี virtual screening ทำให้สามารถสกรีนยาจำนวนมากได้รวดเร็วขึ้น ผลการคำนวณพบเปปไทด์และลิแกนด์ที่คาดว่าจะมีฤทธิ์สามารถยับยั้งการทำงานของโปรตีนเอนวิลอป, เอ็นเอสห้าเมทิลทรานส์เฟอเรส, เอ็นเอสสามโปรตีนเอส, และเฮลิเคสของไวรัสเดงกี ซีโรทัยป์ 1, 2, 3 และ 4 ได้ ซึ่งควรได้รับการทดสอบฤทธิ์ดังกล่าวในห้องปฏิบัติการต่อไป

คำหลัก : ไวรัสเดงกี, ตัวยับยั้งชนิดเปปไทด์, genetic algorithm, virtual screening, สารต้านไวรัส

Project Code: DBG5180013

Project Title: Computational design and discovery of dengue virus inhibitors

Investigator: Ekachai Jenwitheesuk

National Center for Genetic Engineering and Biotechnology

National Science and Technology Development Agency

113 Thailand Science Park, Phaholyothin Road

Klong Luang, Pathumthani 12120

E-mail Address: ekachai@biotec.or.th

Project Period: 30 June 2008 to 30 June 2011

Dengue virus infection is an important public health problem in Thailand because there is no effective preventive vaccine and antiviral drugs available for dengue virus inhibition. This results in high rates of infection and death each year. This research aims to develop the computational methods (1) to design peptide inhibitors to inhibit protein functions of dengue envelop and (2) to identify drugs or chemical compounds to inhibit protein functions of envelop, NS3 protease, helicase and NS5 methyltransferase of four dengue serotypes. In this study, we applied genetic algorithm to identify short peptide inhibitors for envelop protein of four dengue serotypes. The result shows that our method is much faster in identifying peptides than the conventional method, i.e. monte carlo simulation. We also developed a new virtual screening method where the ligands with inappropriate structures were removed from the ligand library prior to the screening step. This method substantially shortens the calculation time and could be applied for screening of large ligand library. By using our peptide inhibitor design and virtual screening methods, we could identify potential peptides and chemical compounds that should be further tested against the envelop, NS3 protease, helicase and NS5 methyltransferase of dengue virus serotype 1, 2, 3 and 4 in vitro.

Keywords: dengue, genetic algorithm, virtual screening, peptide inhibitor, antiviral drugs

บทนำ

ไวรัสเดงกีเป็นไวรัสที่ก่อให้เกิดโรคไข้เลือดออก ซึ่งเป็นโรคสำคัญที่เป็นปัญหาด้านสาธารณสุข โดยเฉพาะ ประเทศในเขตร้อนอย่างประเทศไทยเนื่องจากมีผู้ติดเชื้อและเสียชีวิตในแต่ละปีเป็นจำนวนมาก ปัจจุบันยังไม่มีวัคซีนสำหรับป้องกันและยังไม่มียาต้านไวรัสที่ออกฤทธิ์ยับยั้งไวรัสเดงกีที่ได้ผล [1] วิธีการรักษาโรคติดเชื้อไวรัสเดงกีที่แพทย์ถือปฏิบัติในคลินิกในปัจจุบันเป็นเพียงการรักษาตามอาการ เช่นการให้สารน้ำหรือสารอื่นๆ เพื่อป้องกันการเกิดอาการช็อคหรือลดความรุนแรงของอาการช็อคและป้องกันอวัยวะของระบบต่างๆ ล้มเหลวซึ่งอาจนำไปสู่การเสียชีวิต [1-3] สารเหล่านี้ไม่ได้ออกฤทธิ์ยับยั้งไวรัสเดงกีโดยตรง แต่เป็นเพียงสารที่ช่วยพยุงไม่ให้ผู้ติดเชื้อมีอาการแยลงเพื่อรอให้ร่างกายสร้างภูมิคุ้มกันมากำจัดไวรัส ซึ่งบางครั้งการรักษาตามอาการนี้ไม่ได้ผล ผู้ติดเชื้อไม่ตอบสนองต่อการรักษาและเสียชีวิต [4-6] ดังนั้นยาหรือสารต้านไวรัสจึงมีความสำคัญในการรักษาโรคติดเชื้อไวรัสเดงกีเนื่องจากหากยาต้านไวรัสสามารถยับยั้งการเพิ่มจำนวนและกำจัดไวรัสเดงกีออกจากร่างกายได้อย่างรวดเร็ว โอกาสที่ผู้ติดเชื้อจะเกิดอาการช็อคและความเสี่ยงที่ผู้ติดเชื้อจะเสียชีวิตก็น่าจะลดลง การรักษาโดยใช้ยาต้านไวรัสที่จำเพาะกับไวรัสเดงกีจึงน่าจะมีประสิทธิภาพและประสบความสำเร็จสูงกว่าวิธีการรักษาตามอาการที่แพทย์ใช้อยู่ในปัจจุบัน

อนุภาคของไวรัสเดงกีประกอบไปด้วยโปรตีนสองกลุ่มหลัก [7] กลุ่มแรกเป็นโปรตีนที่เป็นส่วนประกอบของโครงสร้างของอนุภาคไวรัส (structural protein) ประกอบไปด้วยโปรตีนสามชนิดคือ นิวคลีโอแคปซิดโปรตีนซี (nucleocapsid protein C), โปรตีนเอ็ม (prM) หรือ non-glycosylated membrane protein และกลัยโคโปรตีนอี (glycoprotein E) ซึ่งเป็นเอ็นวีลอปโปรตีน (envelope protein) ส่วนโปรตีนกลุ่มที่สองเป็นโปรตีนที่ไม่ได้เป็นส่วนประกอบของโครงสร้างของอนุภาคไวรัส (non-structural protein) เรียกโดยย่อว่า เอ็นเอสโปรตีน (NS protein) ประกอบไปด้วยโปรตีนห้าประเภทคือ NS1, NS2, NS3, NS4 และ NS5 โดยโปรตีน NS2 และ NS4 สามารถแบ่งย่อยลงไปได้อีกประเภทละสองชนิดคือ NS2A, NS2B และ NS4A, NS4B ในขั้นตอนการแบ่งตัวของไวรัสเดงกี โปรตีนทั้งหมดของไวรัสจะถูก translate ออกมาเป็นสาย polypeptide สายเดี่ยว ซึ่งต่อมาจะถูกตัดด้วยโปรตีนโปรตีเอสจนได้เป็นโปรตีนแต่ละชนิดซึ่งทำหน้าที่ต่างๆ กัน ตามที่ได้บรรยายมาแล้วข้างต้น โปรตีนโปรตีเอสนี้มีทั้งโปรตีเอสที่เป็นของตัวไวรัสเองและเป็นโปรตีเอสที่มาจากเซลล์ที่ไวรัสอาศัยอยู่ โดยโปรตีนโปรตีเอส เช่น ฟิวรีน (furin) และ ซีครีเตส (secretase) ที่มาจากเซลล์ที่ไวรัสอาศัยอยู่ จะตัดพันธะระหว่างโปรตีน NS1 และ NS2A เป็นอันดับแรกใน endoplasmic reticulum ตามด้วยพันธะระหว่าง C-prM, prM-E, E-NS1 และ NS4A-NS4B ส่วนพันธะระหว่าง NS2A-NS2B, NS2B-NS3, NS4A-NS4A และ NS4B-NS5 จะถูกตัดโดย NS3 protease ซึ่งเป็นโปรตีนประเภทซีรีนโปรตีเอส (serine protease) ที่มาจากตัวไวรัสเดงกีเอง [7]

ตารางที่ 1 แสดงตำแหน่งและลำดับกรดอะมิโนของโปรตีนที่ถูกตัดอย่างจำเพาะโดย NS3 protease ของไวรัสเดงกีซีโรทัยปี 1-4 เครื่องหมาย | แสดงตำแหน่งของ scissile bond ที่ถูกตัดด้วย NS3 protease และสัญลักษณ์กรดอะมิโนที่แสดงด้วยตัวอักษรตัวหนาเป็นกรดอะมิโนตำแหน่ง P1 [7]

	Capsid-protC		NS2A-NS28		NS2B-NS3		NS3-NS4A		NS4B-NS5	
DENV1	97	104	1341	1348	1471	1478	2090	2097	2489	2496
	RRK R SVTM		WGR K SWPL		KKC R SGVL		AGR R SVSG		GGR R GTGA	
DENV2	97	104	1342	1349	1472	1479	2090	2097	2488	2495
	RRR R SAGV		SKK R SWPL		KKQ R AGVL		AGR K SLTL		NTR R GTGN	
DENV3	97	104	1342	1349	1472	1479	2090	2097	2488	2495
	KRK K TSLC		LKR R SWPL		QTQ R SGVL		AGR K SIAL		TGK R GTGS	
DENV4	96	103	1341	1348	1471	1478	2089	2096	2484	2491
	GRK R STIT		ASR R SWPL		KTC R SCAL		SGR K SITL		TPR R CTGT	

โปรตีน NS3 เป็นโปรตีนที่ทำหน้าที่หลายอย่าง บริเวณ N-terminal (กรดอะมิโนตำแหน่งที่ 1-180) เป็นโดเมนที่ทำหน้าที่เป็น trypsin-like serine protease ส่วนด้าน C-terminal ที่เหลือ เป็นโดเมนที่ทำหน้าที่เป็น nucleoside triphosphatase, 5'-RNA triphosphatase และ helicase การยับยั้งการทำงานที่เป็นโปรตีเอส ของโปรตีน NS3 ของไวรัสเดงกี พบว่าสามารถยับยั้งขบวนการแบ่งตัวเพิ่มจำนวนของไวรัสได้เป็นอย่างดี [8-11] เนื่องจากโปรตีนของไวรัสไม่ถูกตัดออกจากสาย polyprotein ได้

NS3 protease ของไวรัสเดงกียังมีความจำเพาะสูงในการจับและตัดสายเปปไทด์ของตัวไวรัสเดงกี สายเปปไทด์ที่มาจากเซลล์ที่ไวรัสอาศัยอยู่นั้น NS3 protease ของไวรัสก็สามารถเข้าไปจับได้แต่ประสิทธิภาพในการจับและย่อยสายเปปไทด์นั้นต่ำกว่ามาก [7-9] ดังนั้น NS3 protease ของไวรัสเดงกีจึงเป็นเป้าหมายที่สำคัญสำหรับการออกแบบและค้นหาสารต้านไวรัส เนื่องจาก NS3 protease เป็น essential protein ในขบวนการแบ่งตัวเพิ่มจำนวนและมีความจำเพาะกับลำดับกรดอะมิโนของสายเปปไทด์ของตัวไวรัสเดงกีเท่านั้น ดังนั้น peptide inhibitor ที่ออกแบบมาให้จับอย่างจำเพาะกับ NS3 protease ของไวรัส (ตารางที่ 1) จึงมีโอกาสน้อยที่จะไปยับยั้งการทำงานของโปรตีเอสในมนุษย์ซึ่งเป็นผลดีเนื่องจากคุณสมบัติดังกล่าวจะไปลดโอกาสการเกิดผลข้างเคียงที่ไม่พึงประสงค์ของยา

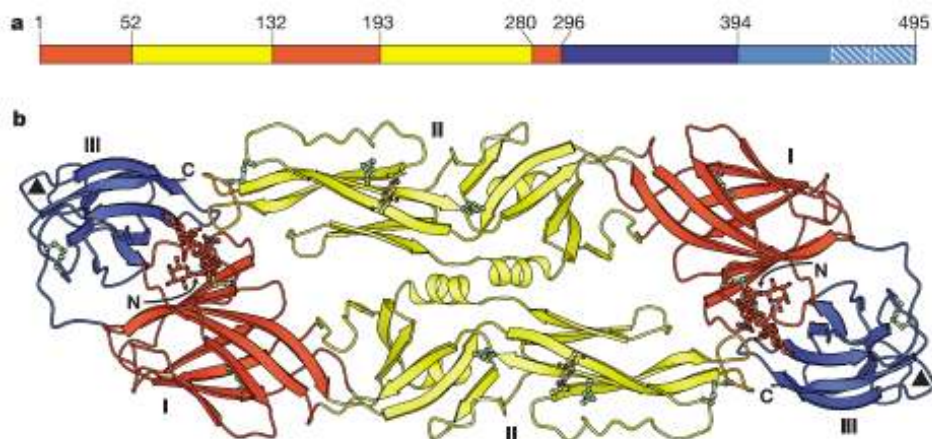
นักวิทยาศาสตร์ได้ใช้ข้อมูลในตารางที่ 1 ออกแบบและสังเคราะห์สารต้านไวรัสที่เป็น peptide mimic สายสั้นๆ ซึ่งสามารถเข้าไปแย่งสาย polyprotein เพื่อจับอย่างจำเพาะกับ NS3 protease และป้องกันการตัด scissile bond ของสาย polyprotein ทำให้ไวรัสไม่สามารถแบ่งตัวเพิ่มจำนวนได้ โดยสารต้านไวรัสเหล่านี้เป็นเปปไทด์ที่เป็นสารประกอบของสารหลายชนิด ได้แก่ alpha-keto amide [12, 13], amide [14], aldehyde [12, 15], trifluoromethyl ketone [13], boronic acid [13], และ cyclohexenyl chalcone derivative [16] เมื่อทดสอบประสิทธิภาพการยับยั้งการเพิ่มจำนวนของไวรัสเดงกีซีโรทัยปีสองพบว่าสารเหล่านี้มีค่า inhibitory constant (Ki) อยู่ในช่วง 1-50 uM [7]

อีกโดเมนหนึ่งของโปรตีน NS3 ซึ่งทำหน้าที่เป็นเอนไซม์ helicase ก็มีความสำคัญเช่นเดียวกัน โดย helicase ทำหน้าที่คลาย (unwinding) สายอาร์เอ็นเอที่ม้วนกันเป็น duplexes ในขณะที่ไวรัสทำการแบ่งตัวเพื่อเพิ่มจำนวน นักวิทยาศาสตร์ได้สร้างโครงสร้างสามมิติโดยใช้เทคนิค x-ray crystallography และได้ทำ alanine scanning mutagenesis เพื่อค้นหาตำแหน่งที่เป็น active site และศึกษาผลของเปลี่ยน

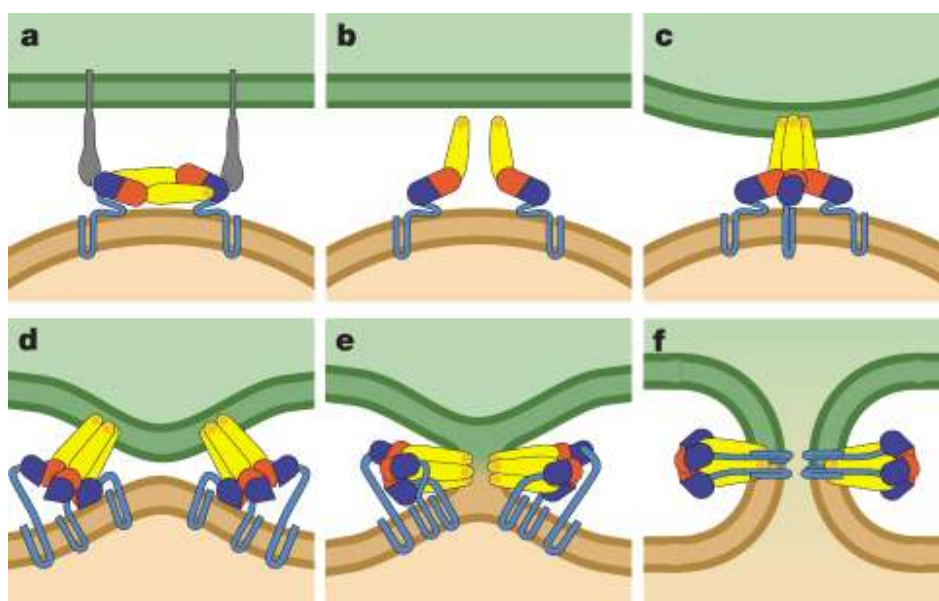
กรดอะมิโนใน active site ต่อ enzymatic activity พบว่าตำแหน่ง Ser-364, Ile-365, Arg-376 และ Lys-396 อยู่ในบริเวณ active site และมีความ conserve สูง การเปลี่ยนแปลงกรดอะมิโนในบริเวณนี้ส่งผลทำให้ helicase มี enzymatic activity ลดลงอย่างมาก ซึ่งเป็นตำแหน่งที่เหมาะสมสำหรับการออกแบบยาให้เข้ามามีจับอย่างจำเพาะ อย่างไรก็ตาม ยังไม่มีกลุ่มวิจัยกลุ่มใดออกแบบยาหรือสกรีนยาต่อ NS3 helicase ของไวรัสเดงกี

นอกจากโปรตีน NS3 แล้ว โปรตีน NS5 ของไวรัสเดงกีก็เป็นโปรตีนที่สำคัญสำหรับการออกแบบสารต้านไวรัส เนื่องจากโปรตีน NS5 นี้เป็นโปรตีนที่มีความ conserve สูง [17, 18] โดยโดเมนทางด้าน N-terminal ทำหน้าที่เป็น methyltransferase และปลายทางด้าน C-terminal ทำหน้าที่เป็น RNA-dependent RNA polymerase นักวิทยาศาสตร์พบว่า ribavirin ซึ่งเป็น guanosine analogue มีฤทธิ์สามารถยับยั้งไวรัสเดงกี [19] โดยเมื่อ ribavirin เข้าสู่เซลล์แล้วจะถูก phosphorylated โดยเอนไซม์ไคเนส (kinase) ซึ่งอยู่ภายในเซลล์ ทำให้มีการเติมฟอสเฟตเข้าที่ปลาย 5' ของโมเลกุลของ ribavirin จากนั้น ribavirin ที่มีกลุ่มฟอสเฟต จะเข้าไปแย่ง guanosine triphosphate จับกับ NS5 methyltransferase และยับยั้งการทำงานของ NS5 methyltransferase นอกจาก ribavirin แล้ว acyclovir และ 5-ethynyl-1-beta-D-ribofuranosylimidazole-4-carboxamide (EICAR) ก็มีฤทธิ์ในการยับยั้งไวรัสเดงกีได้ใกล้เคียงกับ ribavirin แต่ยาทั้งสามตัวนี้ยังมีประสิทธิภาพในการยับยั้งไวรัสเดงกีได้ไม่ดีนัก โดยมีค่า inhibitory concentration อยู่ที่ high micromolar level [19]

โปรตีนเอนวิลอปเป็นโปรตีนที่อยู่บนผิวของไวรัสทำหน้าที่ในขบวนการ membrane fusion และ cell entry โดยโปรตีนเอนวิลอปของไวรัสเดงกีมีโรทียีสสองมีขนาดความยาว 495 กรดอะมิโน แบ่งออกเป็นสามโดเมน (domain) ดังที่แสดงในรูปที่ 1 โดเมนที่หนึ่งสีแดง, โดเมนที่สองสีเหลือง และโดเมนที่สามสีน้ำเงิน นอกจากนี้มีส่วนที่เป็น trans-membrane region ยาวประมาณ 100 กรดอะมิโน (สีฟ้า) [20] ไวรัสสามารถเข้าสู่เซลล์เป้าหมายได้โดยสอดส่วนปลายของโดเมนที่สอง (สีเหลือง) เข้าไปในผิวของเซลล์เป้าหมาย จากนั้นจะม้วนพับโดเมนที่สาม (สีน้ำเงิน) เข้าหาโดเมนที่สองโดยใช้โดเมนที่หนึ่ง (สีแดง) เป็นจุดหมุน ทำให้ผิวของไวรัสเข้ามาใกล้กับผิวเซลล์เป้าหมายและเกิดการประสานตัวกันของผิวของไวรัสและผิวเซลล์เป้าหมาย ทำให้ไวรัสสามารถเข้าสู่เซลล์เป้าหมายได้ (รูปที่ 2) การออกแบบ peptide inhibitor หรือสารเคมีที่สามารถจับแบบจำเพาะกับโปรตีนเอนวิลอปของไวรัสเดงกีเพื่อยับยั้งการทำงานของโปรตีนดังกล่าวก็จะสามารถป้องกันไม่ให้ไวรัสเข้าสู่เซลล์เป้าหมายซึ่งจะส่งผลให้ไวรัสไม่สามารถแบ่งตัวเพิ่มจำนวนได้ [20]



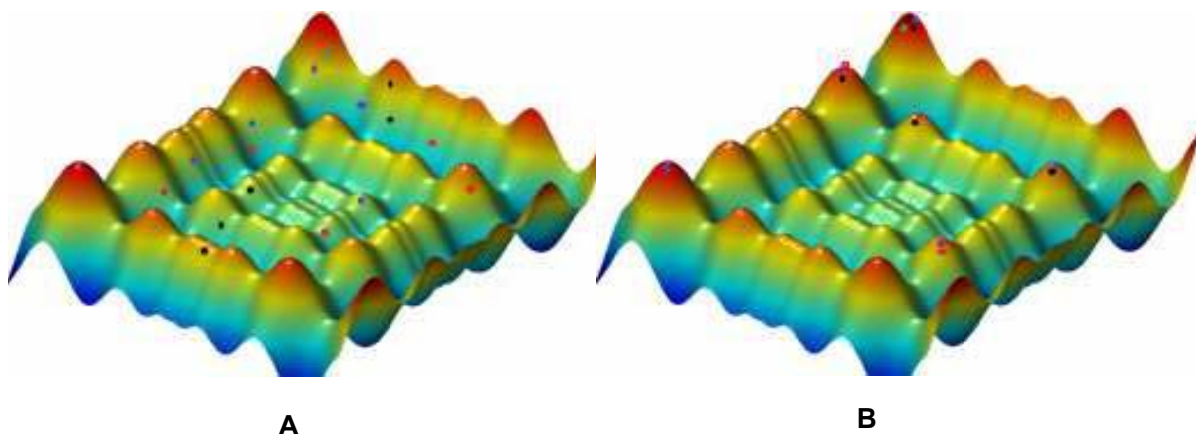
รูปที่ 1 แสดงโครงสร้าง X-ray crystallography ของโปรตีนเอนวิลอปของไวรัสเดงกีซีโรทัยปัสสอง (a) โปรตีนเอนวิลอปมีความยาว 495 กรดอะมิโน (b) มีลักษณะเป็น homodimer รูปร่างคล้ายตัวอักษร L ประกอบไปด้วยสามโดเมน คือ โดเมนที่หนึ่ง (สีแดง) อยู่ตรงกลางของ monomer ทำหน้าที่เป็นจุดหมุน โดเมนที่สอง (สีเหลือง) เป็นด้านปลายที่ไวรัสสอดเข้าไปในผิวของเซลล์เป้าหมาย และโดเมนที่สาม (สีน้ำเงิน) เป็นบริเวณที่อยู่ติดกับผิวของไวรัส ส่วนสีฟ้า (กรดอะมิโนที่ 394 ถึง 495) เป็นส่วน transmembrane region



รูปที่ 2 แสดงกลไกการเข้าสู่เซลล์เป้าหมายของไวรัสเดงกีซีโรทัยปัสสอง โดย (a-c) ไวรัสสอดส่วนปลายของโดเมนที่สองเข้าไปในผิวของเซลล์เป้าหมาย (สีเขียว), (d) จากนั้นเกิดการหมุนพับของโดเมนที่สามไปทางโดเมนที่สองโดยใช้โดเมนที่หนึ่งเป็นจุดหมุน, (e) ทำให้ผิวของไวรัส (สีน้ำตาล) และผิวเซลล์เป้าหมายเคลื่อนเข้ามาใกล้กัน, (f) และเกิดการประสานตัวกันของผิวของไวรัสและผิวเซลล์เป้าหมาย ทำให้ไวรัสสามารถเข้าสู่เซลล์เป้าหมายได้ การออกแบบ peptide inhibitor หรือสารเคมีที่สามารถจับแบบจำเพาะกับโปรตีนเอนวิลอปของไวรัสเดงกีซีโรทัยปัสสองเพื่อยับยั้งการทำงานของโปรตีนดังกล่าวก็จะสามารถป้องกันไม่ให้ไวรัสเข้าสู่เซลล์เป้าหมายซึ่งจะส่งผลให้ไวรัสไม่สามารถแบ่งตัวเพิ่มจำนวนได้

ในปัจจุบันวิธีการค้นหาเปปไทด์ที่มีลำดับกรดอะมิโนที่มีค่าคะแนนฟิตเนส (ค่าคะแนนในการประเมินสายเปปไทด์) ที่ดีที่สุดใช้วิธีการทางคอมพิวเตอร์ ทำโดยการพิจารณาในแต่ละตำแหน่งบนโปรตีนเป้าหมาย แล้วเลือกกลุ่มสายเปปไทด์ขึ้นมา จากนั้นหาค่าคะแนนฟิตเนสไปเรื่อยๆ ตามลำดับกรดอะมิโนโดยตรง ซึ่งการทำงานในลักษณะดังกล่าวยังมีข้อจำกัด คือ ต้องทำการพิจารณาค้นหาจำนวนเปปไทด์ที่มีค่าคะแนนฟิตเนสที่เหมาะสมจำนวนมากจนเกินความสามารถที่คอมพิวเตอร์จะคำนวณได้หมดในระยะเวลาอันสั้น ตัวอย่างเช่น ในกรณีที่จะออกแบบตัวยับยั้งชนิดเปปไทด์ที่มีความยาว 10 กรดอะมิโน แต่ละตำแหน่งของเปปไทด์สามารถเปลี่ยนกรดอะมิโนได้ 20 ชนิด ดังนั้นหากเปลี่ยนชนิดกรดอะมิโนโดยสุ่มในทุกๆ ตำแหน่งของเปปไทด์ ก็จะได้เปปไทด์ทั้งหมด จำนวน 20^{10} หรือประมาณสิบล้านล้านเปปไทด์ โดยใช้เวลาในการคำนวณค่าคะแนนฟิตเนสของเปปไทด์แต่ละสายประมาณ 5 นาทีต่อหนึ่งซีพียู จะใช้เวลาทั้งหมดประมาณ 96,500 ปี สำหรับการคำนวณทั้งสิบล้านล้านเปปไทด์ต่อหนึ่งซีพียู ดังนั้น จึงจำเป็นต้องใช้อัลกอริทึมค้นหา (search algorithm) เช่น มอนติคาร์โลซิมูเลชัน (Monte-Carlo simulation) ซึ่งเป็นวิธีการทางคอมพิวเตอร์อีกชนิดหนึ่งมาช่วยในการค้นหาเปปไทด์ที่มีค่าคะแนนฟิตเนสที่ดี ดังเช่นวิธีการออกแบบเปปไทด์ที่แสดงไว้ในคำขอสิทธิบัตร เลขที่ประกาศโฆษณา WO/2009/045596 Jenwitheesuk และคณะ เรื่อง “Optimized dengue virus entry inhibitory peptide (1OAN)” และคำขอสิทธิบัตร เลขที่ประกาศโฆษณา WO/2009/048658 Jenwitheesuk และคณะ เรื่อง “Optimized dengue virus entry inhibitory peptide (DN81)” อย่างไรก็ตามวิธีมอนติคาร์โลซิมูเลชันนี้ยังมีข้อจำกัดด้านความไวในการค้นหาเปปไทด์ที่มีค่าคะแนนฟิตเนสที่ดีและเปปไทด์ที่ค้นหาได้มักเป็นเปปไทด์ที่มีค่าคะแนนฟิตเนสที่ระดับจุดสูงสุดสัมพัทธ์ (local minima) ซึ่งไม่ใช่ค่าที่สูงสุดของฟังก์ชัน เพียงแต่เป็นค่าสูงสุดเมื่อเทียบกับจุดข้างเคียงเท่านั้น หรือกล่าวอีกนัยหนึ่ง คือ จุดดังกล่าวยังไม่ใช่ว่าจุดที่เหมาะสมที่สุดในการออกแบบเปปไทด์ และต้องใช้เวลาอย่างมากในการค้นหาเปปไทด์สายอื่นที่มีค่าคะแนนฟิตเนสที่ดีกว่าหรือหาสายเปปไทด์ที่มีค่าสูงสุดสมบูรณ์ (global minima)

ด้วยข้อจำกัดของวิธีการออกแบบตัวยับยั้งชนิดเปปไทด์ที่มีมาก่อนดังกล่าวข้างต้น ดังนั้นในโครงการนี้จึงเลือกใช้วิธีเจเนติกอัลกอริทึมที่ได้พัฒนาปรับปรุงให้เหมาะสมกับการออกแบบตัวยับยั้งชนิดเปปไทด์ร่วมกับการใช้สกอริงฟังก์ชัน 2 ชนิดในการหาค่าคะแนนฟิตเนส เนื่องจากเจเนติกอัลกอริทึมในรูปแบบดั้งเดิมเป็นขั้นตอนในการแก้ปัญหาด้วยการคำนวณเพื่อหาคำตอบที่ดีที่สุดของปัญหา โดยเลียนแบบกระบวนการการถ่ายทอดพันธุกรรมของยีนตามธรรมชาติ ซึ่งสิ่งมีชีวิตที่มีลักษณะทางพันธุกรรมที่เหมาะสมจะดำรงอยู่ได้ดีกว่าสิ่งมีชีวิตที่มีลักษณะทางพันธุกรรมไม่เหมาะสมกับสิ่งแวดล้อมนั้นๆ โดยในธรรมชาติพันธุกรรมของสิ่งมีชีวิตจะมีการพัฒนาการโดยการคัดเลือกสิ่งที่ดีที่สุดสายพันธุ์เพื่อสืบทอดไปยังรุ่นต่อไป เมื่อนำหลักการนี้มาใช้ในการคำนวณ จะเป็นการใช้ประสบการณ์ที่มีมาในขั้นตอนการคำนวณก่อนหน้านี้ เพื่อค้นหาคำตอบที่ดีกว่าในขั้นต่อไป ซึ่งจะสามารถช่วยแก้ไขปัญหาคำตอบที่ติดอยู่ที่ระดับจุดสูงสุดสัมพัทธ์ได้ ดังรูปที่ 3 ดังนั้นเจเนติกอัลกอริทึมจึงเหมาะสมในการประยุกต์ใช้กับการแก้ปัญหาที่ใช้เวลาในการค้นหาคำตอบเป็นเวลานาน



รูปที่ 3 ภาพแสดงตัวอย่างการแก้ไขปัญหาการติดอยู่ที่ค่าสูงสุดสัมพัทธ์ (จุดสูงสุดของบางพื้นที่) ของวิธีเจเนติกอัลกอริทึม โดยภาพ A แสดงการสุ่มค่าเริ่มต้นของคำตอบ โดยจะกระจายไปยังคำตอบที่ตำแหน่งต่างๆ ของพื้นที่ ภาพ B แสดงให้เห็นว่าคำตอบที่ได้จะอยู่ที่ตำแหน่งที่ใกล้เคียงหรือตรงกับค่าสูงสุดสัมบูรณ์ (ตำแหน่งที่สูงที่สุดของพื้นที่)

จากผลการวิจัยที่ผ่านมา ได้พบตำแหน่งที่เหมาะสมสำหรับการออกแบบ peptide inhibitor 2 ตำแหน่ง โดยใช้อัลกอริทึม Metropolis Monte Carlo ในการค้นหา และใช้วิธี RAPDF ในการหาค่าคะแนนเพื่อทำการออกแบบตัวยับยั้งชนิดเปปไทด์สำหรับไวรัสเดงกีซีโรทัยปสอง พบว่าตำแหน่งและเปปไทด์ดังกล่าว สามารถยับยั้งได้เกือบร้อยเปอร์เซ็นต์ โดยเปปไทด์นี้จะเข้าไปรบกวนการม้วนพับของโปรตีนเอนวิลอปทำให้ไวรัสไม่สามารถเข้าสู่เซลล์เป้าหมายได้ [21] นอกจากนี้ มีผู้ออกแบบ peptide inhibitor โดยเลียนแบบลำดับกรดอะมิโนของโดเมนที่สองและโดเมนที่สามของโปรตีนเอนวิลอปของไวรัสเดงกีซีโรทัยปสองและนำมาทดสอบหาค่า inhibitory activity แต่พบว่าเปปไทด์ยับยั้งไวรัสเดงกีได้ไม่ดีเท่าที่ควร [22] ดังนั้นในโครงการนี้ จึงใช้สองตำแหน่งบนไวรัสเดงกีซีโรทัยปสองมาใช้ในการออกแบบตัวยับยั้งชนิดเปปไทด์สำหรับซีโรทัยปอื่นๆ และพัฒนาวิธีการทางคอมพิวเตอร์เพิ่มเติมเพื่อให้สามารถออกแบบตัวยับยั้งเปปไทด์ที่มีประสิทธิภาพมากที่สุด

ส่วนงานวิจัยส่วนที่สองพบสารเคมีหลายชนิดในกลุ่ม β -glucosidase inhibitor สามารถยับยั้งการทำงานของโปรตีนเอนวิลอปได้อย่างมีประสิทธิภาพ เช่น castanospermine และ deoxynojirimycin สามารถยับยั้งไวรัสเดงกีได้มากกว่า 90% ที่ความเข้มข้น 50 μ M [23-25] โดย crystallographer ได้พบบริเวณที่สำคัญซึ่งเป็นช่องเล็กๆ อยู่บริเวณระหว่างโดเมนที่หนึ่งกับโดเมนที่สองของโปรตีนเอนวิลอป ในขณะที่ทำการตกผลึกโปรตีนเอนวิลอปโดยใช้ n-octyl- β -D-glucoside (β -OG) ซึ่งเป็น detergent ชนิดหนึ่ง พบว่า β -OG สามารถเข้าไปอยู่ในช่องเล็กๆ ดังกล่าวและสามารถขัดขวางการม้วนพับของโปรตีนเอนวิลอปได้ซึ่งบ่งชี้ว่าบริเวณดังกล่าวเป็นบริเวณที่สามารถใช้ในการออกแบบหรือสกรีนยา [23, 25] ถึงแม้จะมีความพยายามในการออกแบบ peptide inhibitor และค้นหาต้านไวรัส แต่เปปไทด์ที่ออกแบบมายังมีประสิทธิภาพที่ไม่เป็นที่น่าพอใจ และยังไม่มีผู้ทำวิจัยสกรีนยาในลักษณะที่เป็น large-scale virtual screening เพื่อค้นหาสารเคมีที่มีโมเลกุลขนาดเล็ก

วัตถุประสงค์

ผู้วิจัยแบ่งโครงการนี้ออกเป็นสองส่วน ดังนี้

วัตถุประสงค์ของโครงการส่วนที่ 1

พัฒนาวิธีการทางคอมพิวเตอร์โดยใช้เทคนิค genetic algorithm สำหรับคำนวณออกแบบ peptide inhibitor สองตำแหน่ง ที่มีขนาดความยาว 20 กรดอะมิโนและ 28 กรดอะมิโน เพื่อยับยั้งการทำงานของโปรตีนเอนวิลอป (envelope) ของไวรัสเดงกีทั้ง 4 ซีโรทัยป์

วัตถุประสงค์ของโครงการส่วนที่ 2

ค้นหาหรือสารต้านไวรัสเดงกี โดยใช้เทคนิค virtual screening สกรีนโครงสร้างสามมิติของ approved drugs และ drug-like compounds จำนวนหนึ่งล้านโครงสร้างต่อโครงสร้างสามมิติของโปรตีนเอนวิลอป (envelope), เอ็นเอสสามโปรตีเอส (NS3 protease), เฮลิเคส (helicase) และ เอ็นเอสห้าเมทิลทรานส์เฟอเรส (NS5 Methyltransferase) ของไวรัสเดงกีซีโรทัยป์ 1, 2, 3 และ 4

วิธีการทดลอง

ส่วนที่ 1 การออกแบบ peptide inhibitor

1. ดาวนโหลดโครงสร้างสามมิติของโปรตีนเอนวิลอป (envelope) ของไวรัสเดงกีซีโรทัยป์ 2 และ 3 ซึ่งได้จากการทำ crystallography จาก Protein Data Bank (PDB) โดยมี PDB ID: 1OAN และ 1UZG
2. ดาวนโหลดลำดับกรดอะมิโนของโปรตีนเอนวิลอปของไวรัสเดงกีซีโรทัยป์ 1 และ 4 จากเว็บ GenBank และนำมาสร้างโครงสร้างสามมิติโดยใช้โครงสร้างสามมิติจากข้อ 1. เป็นโครงสร้างต้นแบบโดยใช้โปรแกรม Modeller บนเซอร์เวอร์ (<http://www.salilab.org/modeller/>)
3. เทคนิค genetic algorithm ถูกนำมาใช้ในการควบคุมการสร้าง mutant structures เพื่อค้นหาลำดับกรดอะมิโนที่มีค่า scoring function ของ discrete optimized protein energy (DOPE) [26] และ residue-specific all-atom probability discriminatory function (RAPDF) [27] ที่ดีที่สุด เพื่อใช้ในการออกแบบตัวยับยั้งชนิดเปปไทด์สำหรับโปรตีนเอนวิลอปของไวรัสเดงกี ทั้งสี่ซีโรทัยป์ โดยใช้ตำแหน่งที่เหมาะสมสำหรับการออกแบบตัวยับยั้งชนิดเปปไทด์จำนวนสองตำแหน่ง โดยมีวิธีดังรูปที่ 4 มีรายละเอียด ดังนี้

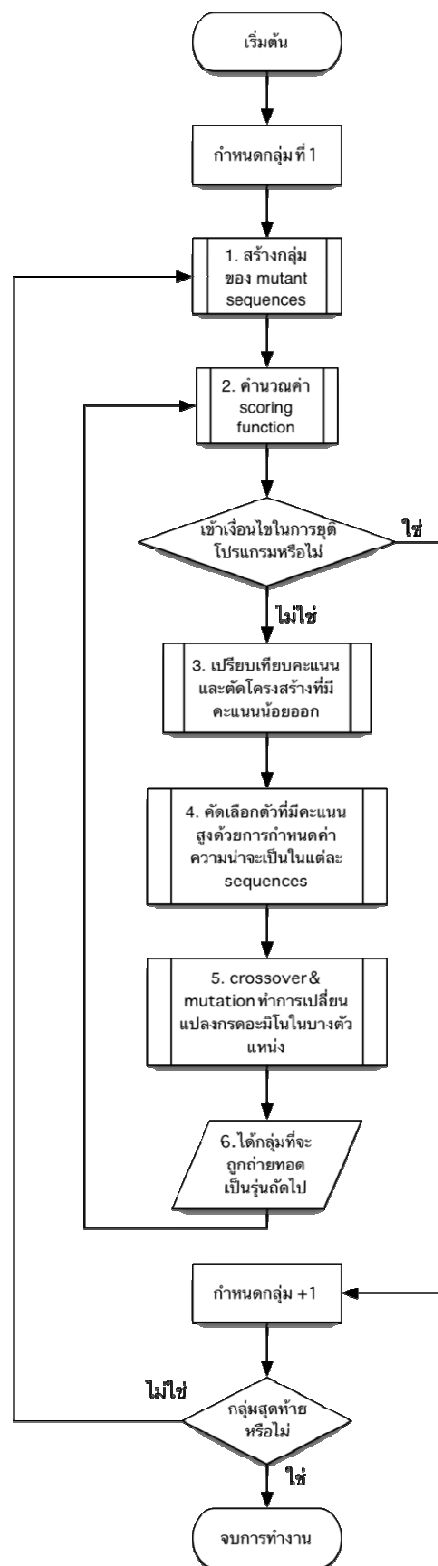
3.1 สร้างกลุ่มของ mutant sequences โดยกำหนดให้กลุ่มหนึ่งมีทั้งหมด 50 ลำดับ โดยแต่ละลำดับจะมีความยาวตามตำแหน่งที่สนใจ โดยในแต่ละตำแหน่งเกิดจากการสุ่มตัวกรดอะมิโนทั้ง 20 ตัว โดยเมื่อกำหนดเรียบร้อยแล้ว ก็จะเก็บไว้ในฐานข้อมูล เพื่อจะได้ไม่เกิดการสุ่มซ้ำตัวเดิม

3.2 ในแต่ละ mutant sequences ในกลุ่มจะถูกนำมาคำนวณหาค่าคะแนนฟิตเนสโดยจะทำการแทนที่ mutant sequences ลงในตำแหน่งที่เหมาะสมในลำดับกรดอะมิโนของแต่ละ ซีโรทัยป์ แล้วทำการสร้างโครงสร้างโปรตีนของลำดับกรดอะมิโนนั้นๆ โดยใช้โปรแกรม scwrl3 [28] เมื่อได้โครงสร้างโปรตีนแล้ว จะทำการ minimize energy ของโครงสร้างโดยใช้โปรแกรม NAMD [29] แล้วจึงนำไปคำนวณหาค่า scoring function โดยใช้ทั้งโปรแกรม DOPE และ RAPDF

3.3 เมื่อได้คะแนนของทุกๆ โครงสร้างมาแล้ว ในแต่ละรอบจะทำการตัดโครงสร้าง ที่มีคะแนนน้อยออก โดยจะเก็บโครงสร้างและลำดับเบสที่มีคะแนนดีที่สุด 50 ลำดับแรกไว้ใช้ในรอบถัดไป

3.4 วิธีคัดเลือก จะใช้วิธี Roulette-wheel โดยที่แต่ละลำดับเบส จะถูกเรียงตามคะแนนมากที่สุด ไปน้อยที่สุด แล้วกำหนดค่าความ น่าจะเป็นให้กับแต่ละลำดับเบส เพื่อตัวที่ดีที่สุดจะมีโอกาส ถ่ายทอดต่อไปยังรุ่นต่อไป โดยเอาคะแนนของโครงสร้างนั้นๆ หารด้วย คะแนนรวมทั้งหมดของ ทุกโครงสร้าง เมื่อได้ค่าความน่าจะเป็นแล้ว ต่อไปคือการสุ่มตัวเลขตั้งแต่ 0-1 เพื่อใช้ในการ ตัดสินอย่างสุ่มว่าลำดับเบสไหนควรจะถูกลำดับทอดไปยังรุ่นถัดไป โดยที่ลำดับเบสที่มีค่าความน่าจะเป็นมาก อาจจะถูกเลือกออกมาหลายครั้ง

3.5 เมื่อได้ลำดับเบสกลุ่มใหม่ออกมาแล้ว จะทำการจับคู่ของแต่ละลำดับเบสอย่างสุ่ม แล้วทำการ crossover สลับตำแหน่งระหว่างคู่ เพื่อให้เกิดการกระจายของกรดอะมิโนเมื่อทำการ crossover เรียบร้อยแล้ว ต่อไปคือการทำ mutation โดยการที่พิจารณาในแต่ละลำดับเบส ทำการเปลี่ยนบางตำแหน่งของกรดอะมิโนโดยขึ้นอยู่กับค่าการสุ่มตัวเลข และเปอร์เซ็นต์ที่ได้กำหนดไว้



รูปที่ 4 การออกแบบตัวยับยั้งชนิดเปปไทด์ ด้วยเทคนิค genetic algorithm

3.6. กลุ่มของลำดับเบสที่ได้ล่าสุดนี้ จะถูกนำไปคิดคะแนนใหม่ในข้อ 2.2 และ จะทำการวนซ้ำไปเรื่อยๆ โดยในแต่ละรอบจะเก็บตัว ที่มีคะแนนสูงสุดเอาไว้ เงื่อนไขในการยุติโปรแกรมคือ ถ้าเกิดไม่มีการ

เปลี่ยนแปลงตัวที่มีค่าสูงสุดเป็นจำนวนครั้งตามที่กำหนดแล้ว (ในโครงการนี้กำหนดไว้ 200 ครั้ง) ก็จะทำให้การยุติโปรแกรม แล้วเริ่มสุ่ม mutant sequences กลุ่มใหม่ในข้อ 2.1

4. ใช้เทคนิคจากข้อ 3 ในการออกแบบเปปไทด์ขนาดความยาว 20 กรดอะมิโนสำหรับตำแหน่งที่ 1 เพื่อยับยั้งการทำงานของโปรตีนเอนวีลอป (envelope) ของไวรัสเดงกีทั้งสี่ซีโรทัยป์

5. ใช้เทคนิคจากข้อ 3 ในการออกแบบเปปไทด์ขนาดความยาว 28 กรดอะมิโนสำหรับตำแหน่งที่ 2 เพื่อยับยั้งการทำงานของโปรตีนเอนวีลอป (envelope) ของไวรัสเดงกีทั้งสี่ซีโรทัยป์

6. วิเคราะห์ลักษณะเฉพาะของตัวยับยั้งชนิดเปปไทด์ที่ได้มาทั้งหมดของแต่ละซีโรทัยป์เพื่อให้มีความแม่นยำมากขึ้นและสามารถคัดเลือกตัวยับยั้งชนิดเปปไทด์ที่มีความน่าจะเป็นสูงสุด โดยมีขั้นตอนดังนี้

6.1 คัดเลือกสายเปปไทด์ที่มีคะแนนฟิตเนสดีกว่า wildtype

6.2 คัดเลือกสายเปปไทด์ที่ไม่มี poly-H และไม่มีกรดอะมิโนซ้ำกันอยู่ติดกันมากกว่าสามตัว

6.3 นำสายเปปไทด์ที่ผ่านเกณฑ์สองข้อแรก ไปหาค่าต่างๆ ดังนี้

6.3.1) % hydrophilic (ร้อยละของการละลายน้ำ) คำนวณโดยสายเปปไทด์ที่มีกรดอะมิโนเป็น D, E, R, K, หรือ H เป็นส่วนประกอบ โดยสายเปปไทด์ที่เหมาะสมจะต้องมีค่ามากกว่า 25 %

6.3.2) % hydrophobic (ร้อยละของการไม่ละลายน้ำ) คำนวณโดยสายเปปไทด์ที่มีกรดอะมิโนเป็น F, I, L, M, V, W, หรือ Y โดยสายเปปไทด์ที่เหมาะสมจะต้องมีค่าน้อยกว่า 25 %

6.3.3) % forming H bond (ร้อยละของพันธะไฮโดรเจน) คำนวณโดยสายเปปไทด์ที่มีกรดอะมิโนเป็น S, T, E, D, K, R, H, N, Q, หรือ Y โดยสายเปปไทด์ที่เหมาะสมจะต้องมีค่าน้อยกว่า 75 %

6.3.4) Isoelectric point (pI) (ค่าคะแนน Isoelectric) โดยสายเปปไทด์ที่เหมาะสมจะต้องมีค่าใกล้เคียงกับ 7

6.4 สายเปปไทด์ที่ผ่านเกณฑ์ทั้งหมด เป็นตัวยับยั้งชนิดเปปไทด์ที่อยู่ในเกณฑ์ดี และจะถูกนำไปพิจารณาความเหมาะสมและนำไปใช้งานต่อไป

ส่วนที่ 2 สกรีนหายาหรือสารต้านไวรัสเดงกี

งานวิจัยนี้ ใช้เทคนิค virtual screening ทำการสกรีนโครงสร้างสามมิติของ approved drugs และ drug-like compound จำนวนหนึ่งล้านโครงสร้างต่อโครงสร้างสามมิติของโปรตีนเอนวีลอป, เอ็นเอสสามโปรตีเอส, เฮลิเคส และเอ็นเอสห้าเมทริลทรานส์เฟอเรส ของไวรัสเดงกีสี่ซีโรทัยป์ เพื่อค้นหาหรือสารต้านไวรัสเดงกี โดยแสดงรายละเอียดการทดลอง ดังนี้

1. ขั้นตอนการเตรียมไฟล์โครงสร้างสามมิติของลิแกนด์และโปรตีนของไวรัสเดงกี (รูปที่ 5)

1.1 ดาวน์โหลดโครงสร้างสามมิติของลิแกนด์จำนวนหนึ่งล้านโครงสร้าง (รวมลิแกนด์ทั้งหมด 9 กลุ่ม) จากเว็บไซต์ฐานข้อมูล <http://ligand.info> ทำการแปลงโครงสร้างลิแกนด์ให้อยู่ในรูปแบบ PDB format และเซฟเป็นไฟล์เดี่ยว หนึ่งไฟล์ต่อหนึ่งโครงสร้าง แล้วทำการ assign atomic charges และ rotatable bond ให้ทุก single bond สามารถหมุนได้โดยอิสระ

1.2 ทำการคำนวณหา molecular volume ของลิแกนด์แต่ละโครงสร้าง

1.3 คำนวณโครงสร้างสามมิติของโปรตีนเอนวีลอป, เอ็นเอสสามโปรตีเอส, เฮลิเคส และ เอ็นเอสห้าเมทิลทรานส์เฟอเรส ของไวรัสแดงที่ type 2 ที่ได้จากการทำ crystallography จาก Protein Data Bank (PDB) ซึ่งมี PDB ID ดังต่อไปนี้ 1OKE (envelop), 1R6A (NS5 methyltransferase), 2FOM (NS2B/NS3 protease) และ 2BMF (helicase)

1.4 คำนวณหา volume ของ active site ของโปรตีนแต่ละโครงสร้างจากข้อ 1.3 และทำการเลือกใช้ชุดของลิแกนด์ที่มี molecular volume ที่เหมาะสมกับขนาดของ active site ของโปรตีนโดยภาพรวม ทำการจัดลำดับความสำคัญของชุดลิแกนด์แต่ละชุดว่าควรเลือกลิแกนด์ใดมาสกรีนก่อนหรือหลัง

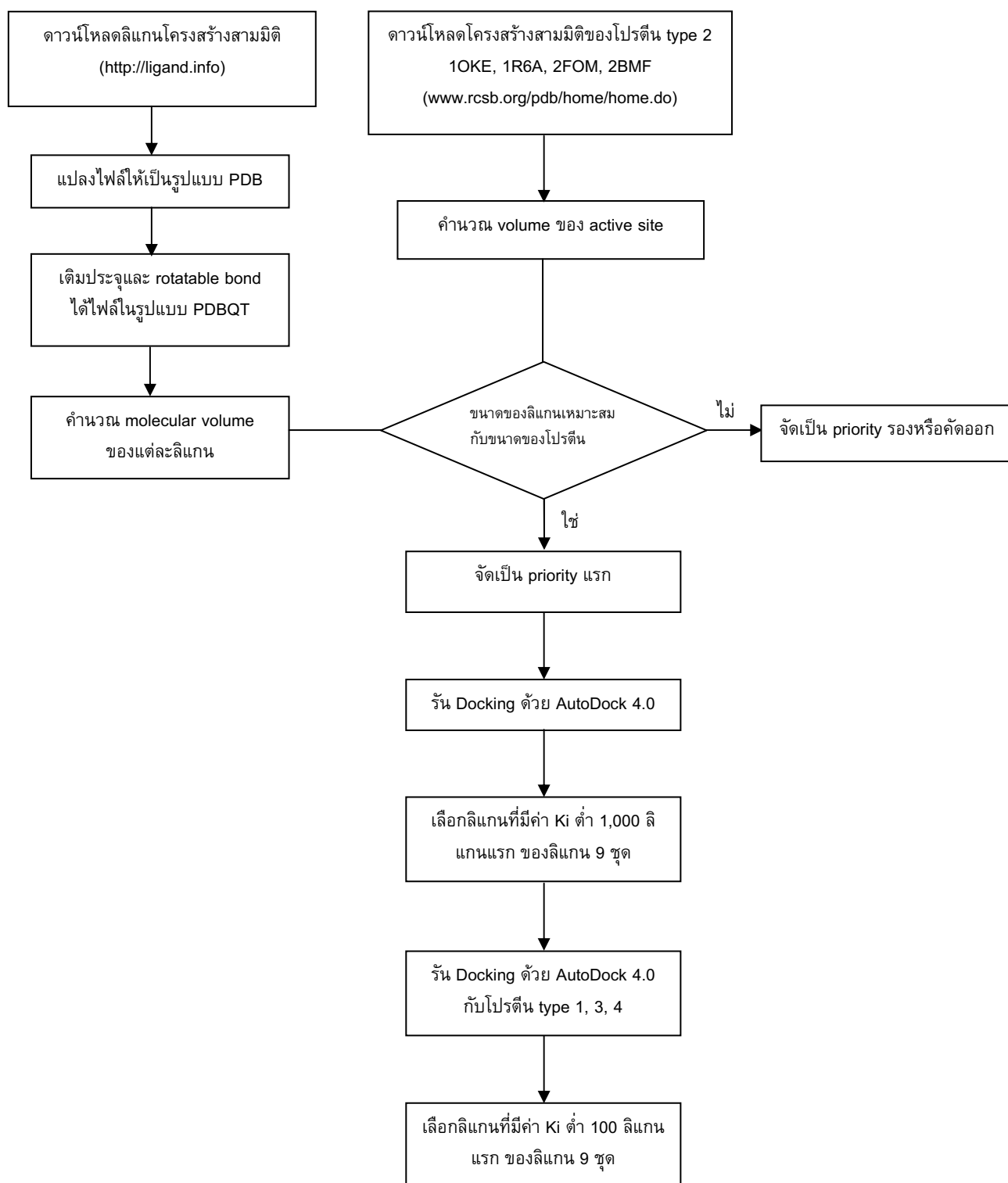
1.5 คำนวณหาค่า inhibitory constant (Ki) ระหว่างโปรตีนกับลิแกนด์โดยใช้โปรแกรม AutoDock version 4.0 และเลือกผลของลิแกนด์ในแต่ละชุดของลิแกนด์ที่ให้ค่า Ki ที่ต่ำที่สุดใน 1,000 ลิแกนด์แรก เพื่อนำไปใช้ในการหาค่า Ki ระหว่างโปรตีนของทั้งสี่โปรตีนของไวรัสแดงที่ type 1, 3, และ 4

1.6 นำลิแกนด์ที่ได้ 1,000 ลิแกนด์แรกของชุดลิแกนด์ทั้ง 9 กลุ่ม คำนวณหาค่า Ki ระหว่างโปรตีนของไวรัสแดงที่ type 1, 3, และ 4 ด้วยโปรแกรม AutoDock [30] อีกครั้ง จากนั้นเลือกลิแกนด์ในแต่ละชุดของลิแกนด์ที่ให้ค่า Ki ที่ต่ำที่สุดใน 100 ลิแกนด์แรกของโปรตีนทั้งสี่โปรตีนของไวรัสแดงที่ type 1, 3, และ 4

1.7 วิเคราะห์ผลของลิแกนด์ที่ได้ โดยเลือกลิแกนด์ใน top 100 ของลิแกนด์ทั้งหมด 9 กลุ่มในแต่ละโปรตีน

2. ขั้นตอนการสร้างโครงสร้างสามมิติของไวรัสแดงที่ type 1, 3 และ 4

โครงสร้างเริ่มต้นที่ทราบโครงสร้างสามมิติ คือ โปรตีนในซีโรทัยป์ 2 ดังนั้น ในการหาซีโรทัยป์ที่เหลือทำได้ โดยนำเอา amino acid sequence จาก NCBI ซึ่งมี accession number ของซีโรทัยป์ 1, 3 และ 4 คือ AAB70695.1, ACZ15977.1 และ NP_073286.1 ตามลำดับ จากนั้น นำ accession number ดังกล่าว ทำการ alignment ด้วยโปรแกรม clustalw กับลำดับเบสของโปรตีนในซีโรทัยป์ 2 ซึ่งจะได้ลำดับเบสของโปรตีนในซีโรทัยป์ที่เหลือ แล้วนำลำดับเบสที่ได้ดังกล่าว ทำการสร้างโครงสร้างสามมิติโดยใช้โปรแกรม CPH models 3.0 server บนเว็บไซต์ <http://www.cbs.dtu.dk/services/CPHmodels/> โครงสร้างสามมิติของโปรตีนซีโรทัยป์ 1, 2, 3 และ 4 แสดงดังรูปที่ 6-9 ตามลำดับ



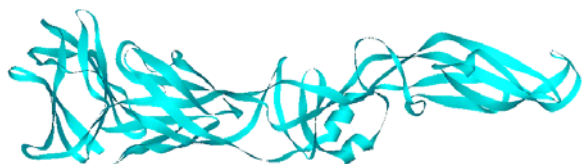
รูปที่ 5 ขั้นตอนการใช้เทคนิค virtual screening สกรีนโครงสร้างสามมิติของลิแกนด์ต่อโครงสร้างสามมิติของโปรตีน



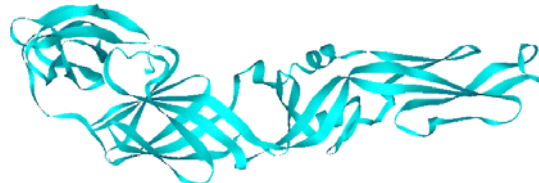
ซีโรทัยป์ 1



ซีโรทัยป์ 2



ซีโรทัยป์ 3



ซีโรทัยป์ 4

รูปที่ 6 โครงสร้างสามมิติของโปรตีนแอนวิลอป ซีโรทัยป์ 1, 2, 3 และ 4



ซีโรทัยป์ 1



ซีโรทัยป์ 2



ซีโรทัยป์ 3



ซีโรทัยป์ 4

รูปที่ 7 โครงสร้างสามมิติของโปรตีนเอ็นเอสห้าเมทิลทรานส์เฟอเรส ซีโรทัยป์ 1, 2, 3 และ 4



ชีโรทัยป์ 1



ชีโรทัยป์ 2

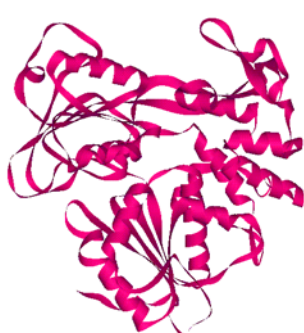


ชีโรทัยป์ 3



ชีโรทัยป์ 4

รูปที่ 8 โครงสร้างสามมิติของโปรตีนเอ็นเอสสามโปรตีนชีโรทัยป์ 1, 2, 3 และ 4



ชีโรทัยป์ 1



ชีโรทัยป์ 2



ชีโรทัยป์ 3



ชีโรทัยป์ 4

รูปที่ 9 โครงสร้างสามมิติของโปรตีนเฮลิเคส ชีโรทัยป์ 1, 2, 3 และ 4

3. วิธีการทำ Molecular docking โดยใช้ชุดซอฟต์แวร์ AutoDock version 4

3.1 เติม charges ให้แก่โครงสร้างของโปรตีนและลิแกนด์ และ assign rotatable bonds ของโครงสร้างของยาโดยใช้ซอฟต์แวร์ AutoDockTool (ADT) โดยใช้ค่า default ที่ซอฟต์แวร์กำหนดมาให้

3.2 ใช้โปรแกรม ADT สร้างไฟล์ grid parameter file (GPF) ขึ้นเพื่อกำหนดค่า parameters ต่างๆ เพื่อสั่งให้โปรแกรม AutoGrid4 (อยู่ในชุดซอฟต์แวร์ AutoDock version 4) สร้าง Grid-box ขนาด 60x60x60 angstroms โดยให้จุดศูนย์กลางของ Grid-box อยู่ที่จุดศูนย์กลางของ active site ของโปรตีน ซึ่ง Grid-box ขนาด 60x60x60 angstroms นี้สามารถครอบคลุม active site ของโปรตีนได้ทั้งหมดและโครงสร้างของลิแกนด์สามารถเคลื่อนไหวอยู่ใน Grid-box ได้อย่างอิสระ

3.3 ใช้โปรแกรม ADT สร้างไฟล์ dock parameter file (DPF) ขึ้นเพื่อกำหนดค่า parameters ต่างๆ เพื่อสั่งให้โปรแกรม AutoDock4 (อยู่ในชุดซอฟต์แวร์ AutoDock version 4) ทำการคำนวณค่าพลังงานระหว่างโครงสร้างของลิแกนด์และโปรตีน เมื่อเสร็จสิ้นการคำนวณ จะได้ลิสต์ของลิแกนด์เรียงตามลำดับความสามารถในการจับกับโปรตีนแต่ละตัวของไวรัสแดงก็

ผลการทดลอง

ส่วนที่ 1 การออกแบบ peptide inhibitor

1. ลำดับกรดอะมิโนในแต่ละซีโรทัยป์แต่ละตำแหน่ง และจำนวนของลำดับทั้งหมดที่หาได้ในแต่ละวิธี พร้อมทั้งค่าคะแนนฟิตเนสของตัว wildtype แสดงดังตารางที่ 2
2. ลำดับกรดอะมิโนของตัวยับยั้งเปปไทด์ที่ผ่านเกณฑ์ของทุกซีโรทัยป์ทุกตำแหน่ง รวมถึงผลที่ได้จากเกณฑ์ แสดงดังตารางที่ 3 ถึง 10

ตารางที่ 2 ลำดับกรดอะมิโนในแต่ละซีโรทัยป์แต่ละตำแหน่ง และจำนวนของลำดับทั้งหมดที่หาได้ในแต่ละวิธีพร้อมทั้งค่าคะแนนฟิตเนสของตัว wildtype

serotype	region	wildtype sequence	RAMP score	DOPE score
1	1	LDIELLKTEVTNPVLRKLC	-14982.60	-37586.10
	2	SWLVHKQWFLDLPLPWTSGASTSQETWN	-15142.11	-37773.16
2	1	LDFELIKTEAKQPATLRKYC	-32466.92	-81226.66
	2	AWLVHRQWFLDLPLPWLPGADTQGSNWI	-32519.60	-81313.80
3	1	LDIELQKTEATQLATLRKLC	-32020.57	-78558.65
	2	AWMVHRQWFFDLPLPWTSGATTKTPTWN	-32016.59	-78637.19
4	1	LDFELTKTTAKEVALLRTYC	-15419.25	-37338.92
	2	TWLVHKQWFLDLPLPWTAGADTSEVHWN	-15517.08	-37414.58

ตารางที่ 3 แสดงตัวบ่งชี้ที่ได้ที่ผ่านเกณฑ์ของที่ร้อยละ 1 region 1

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
FCFCHSHHCEKNERWSRFC	-16025.04	45	20	60	7.64
FCFCHDDHDDKNEREWRFC	-15948.03	60	20	65	6.05
FCFSHDDHHGERNEEWWRCC	-15918.61	50	20	60	6.05
FCFHHEHHECETERWARYC	-15910.76	55	20	65	6.48
FCFWCEHHRPERNERPERFC	-15909.72	50	20	55	6.96
ICFCHERHQIEDNRRSTRFC	-15876.68	45	20	65	7.74
FCFCHHAHCEKNERWWRDC	-15842.59	50	20	55	7.06
ICFCHEQHQNDRNRQSWRFC	-15823.74	35	20	65	7.74
FCFCSQRHEPHRGERGWSFC	-15795.87	35	20	50	7.74
MCFCHEHHQWEDNRRRKARFC	-15786.85	50	20	60	7.74
FCFGHSHHSQRNERWDRFC	-15783.11	45	20	65	7.91
FCFCNHCSHGEKNERQFRFC	-15772.68	35	20	55	7.61
FCFCHHNHHCDRNEQWWRRRC	-15768.42	45	20	60	7.64
ACFCHERHECEDTRRNWRFC	-15763.97	50	15	60	6.90
FCFCSQRHEPHHPRDAWSFC	-15753.59	35	20	50	7.05
ICFCHEHHESKDSSEKWRFC	-15747.99	50	20	65	6.32
FCFCHHPHCEKTERTWRFPC	-15747.65	45	20	55	7.75
FCFCQHDDHTDRNERWWRQC	-15738.70	45	20	65	7.05
TCFCHHEHHIEKNERWWRSC	-15736.89	50	20	65	7.11
FCGCHDHNPEKNERWWKFC	-15734.96	45	20	55	7.05
ICFHHEHHQPRDNRWWREC	-15712.49	55	20	65	7.91
FCGCSHAHCEKNERWWRFC	-15700.81	40	20	50	7.62
PCFCHHGHDDPSNERWWRFC	-15699.45	40	20	50	7.11
FCFCHHPHHEEREERRWRFC	-15681.29	60	20	60	7.12

ตารางที่ 3 แสดงตัวบ่งชี้ชนิดเปปไทด์ที่ได้ผ่านเกณฑ์ของที่ร้อยละ 1 region 1 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
ACFCHCEHHWRDNRREIRFC	-15677.86	50	20	55	7.62
QCFCHPHHPERPERWWRFC	-15666.05	45	20	50	7.75
FCCHHTHDHLEKPERWWGCC	-15664.28	50	20	55	7.06
FEFCTHTHHPERNQRRWRGC	-15661.94	40	20	60	7.91
KCFCHEHHQSEDNRRYIRFC	-15644.56	50	20	70	7.74
ICFCHSHQQGEECRRWWGCC	-15626.56	35	20	50	7.52
FCFCENHHADERNERRRFC	-15620.96	50	20	60	6.96
ICNCHEHHEGKDNKKWWRFC	-15612.41	50	20	60	7.74
FCFCTKNGRDSRPEQDWRLC	-15608.96	35	20	55	7.71
FCDCCHNHHPEKPERFFRFC	-15607.99	45	20	50	6.99
FCFHHTHEHCERCERWERFC	-15603.62	55	20	60	6.47
FCFWCHESGPNRPERWSRAC	-15602.98	30	20	45	7.72
DCFCHSHHNAEHFRDWRFC	-15600.31	50	20	60	7.11
FCCGHHDDHIEKAERFWRNC	-15598.62	50	20	55	7.11
ICTCHEHHQWEDNRRYTRFC	-15588.21	45	20	70	7.05
ACLCPHHPQPEETHTFFRFC	-15586.65	30	20	45	6.32
FGCCMHRHHPDRNERAWDFC	-15585.66	45	20	50	7.05
ICFCHEHHEGKDNIKWERCC	-15582.39	50	20	55	6.32
NCFCHERHEWEHCRRQIRFC	-15582.25	50	20	60	7.62
QCFTCRHHRAEQNEKWWQFC	-15565.53	35	20	60	7.72
ICRHHEHHEGENNRRWWRFC	-15558.05	55	20	65	7.91
DCFCHHEHQWEAPRAWRFC	-15557.41	45	20	50	7.05
EDFCHHKHHVEKNERQWRFC	-15546.87	60	20	70	7.18
FRFCCHNHHPEKNEEWSRFC	-15546.47	45	20	60	7.05

ตารางที่ 3 แสดงตัวยีนชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 1 region 1 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
FQFCHEHHEWEKCDRWQRC	-15545.15	50	20	60	6.32
ECFNHHHHPPERSERLWRF	-15543.63	50	20	60	7.18
ICSCHEHHEWEKMRRGQRF	-15541.24	50	20	60	7.74
FAFHCHWHGCNNPKGRCFC	-15533.52	30	20	40	7.62
FCACAQRLPEKNERWWRC	-15532.97	35	20	45	7.71
FCFCEQRCETCHGKQWRLC	-15528.67	35	20	50	7.52
FCDCHHCHHDERTERWWRF	-15516.95	55	20	60	6.47
PAFRDHHNHIEKPERWWRC	-15506.89	50	20	55	7.91
CCGCHTHHWDPERWWRYC	-15497.67	45	20	55	7.64
FCNSCQHHRAERTEDWWRF	-15490.72	40	20	60	6.95
AVFCHRRHNAENPKDEYKLC	-15481.97	45	20	60	7.90
QCFCFHGGRSENERWWREC	-15472.80	45	20	60	7.74
ICFCHDHHEHERNRRWFRNC	-15472.35	55	20	65	7.75
DCFCHHGHIEKADRWERFC	-15470.70	55	20	55	6.48
FCFCLQRHENHHPRDGARNC	-15470.51	40	15	55	7.74
FCAHQHHPHLEKNEKWWRPC	-15468.79	45	20	55	7.91
FCFEHHQHCEKNERVERSC	-15468.70	55	15	70	6.48
ECFCHEHHQIRDNRRFQRYC	-15463.28	50	20	70	7.74
ICACHEHHQWKDNRHWTRFC	-15454.65	45	20	60	7.75
LCFCTKNSHQDSPKQDWRLC	-15441.49	30	20	60	7.72
FTFWCHQARTENPKRGWCEC	-15421.12	30	20	50	7.72
FVSSCSHGRPEKTERDWRFC	-15419.96	40	20	60	7.90
SCFCHEHKHIDPERWWRHC	-15419.25	55	20	60	7.75
IVFKCCHVREKPERESTTC	-15409.85	40	20	55	7.72

ตารางที่ 3 แสดงตัวยีนชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 1 region 1 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
ICFGTAKTEPESARRLYHCC	-15387.97	30	20	50	7.72
FCEGCHHARAEGNERWWRFC	-15385.50	40	20	45	6.95
ICFHLKRTEFSPEDSQRHC	-15378.82	45	20	65	7.12
IAFHCHVHKQEKQSGERCFC	-15378.64	40	20	55	7.74
FTFSGEETRCDEMNRNRKCC	-15377.61	45	15	65	7.71
QCFCHDEHHLETQRAWSRFC	-15364.42	40	20	60	6.32
LCFCTKNDHQDSPKQDWRLC	-15354.98	35	20	60	6.79
ECFCPHHFEREKNERWWRSC	-15349.72	50	20	60	6.96
FCFHEHDDHREKTRDWRNC	-15335.49	60	20	70	7.18
FCLCRHRTGYPKDCETCLC	-15331.78	35	20	50	6.70
FCDDCRHHRAERNEQWWQFC	-15330.12	45	20	60	6.05
ICSCHHEQHVERNERWKKFC	-15323.65	50	20	65	7.74
CCFHPQRTQTEVDDPKTPIKFQ	-15318.89	30	20	55	7.90
DSFCHHCHHSEKNERWWMAC	-15313.12	45	20	60	6.48
FCNNHHHHWEKNERFFRNC	-15310.55	45	20	70	7.91
FCAWGHRRHASPETHDQWKLC	-15305.90	35	20	50	7.11
GCFWCQRCRDESSRKMYECC	-15301.09	35	20	55	7.49
FCDHCHHHCDRTRRVYRCC	-15286.22	55	15	65	7.55
FDIRCGRYESQSGERCCLC	-15283.78	30	20	60	6.07
LVFNCQHHREGATDNGAKLC	-15268.75	30	20	50	7.02
NCFRKHDHHLQKNEQDWRF	-15263.55	50	20	70	7.91
CCGCLGHCHIKEMREDRKFC	-15262.65	45	20	45	7.52
LQFRCQHHTHDTDNGAKLC	-15260.56	35	20	60	7.11
LMFHAQEDRASQGNQGHRLC	-15254.58	30	20	55	7.12

ตารางที่ 3 แสดงตัวยีนชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 1 region 1 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
SCFCEHDEWKDKRRNWGFC	-15250.79	50	20	60	6.95
ADFWQDRVTCESNRRCRQFC	-15249.15	35	20	60	7.71
FCQCQARHDPDHEEKLWRLC	-15248.54	45	20	55	6.05
IAFHCHPHKQEKQSGERCFC	-15239.42	40	15	55	7.74
FCNTMQRHDPDHGKKDWRLC	-15235.23	45	20	60	7.90
GCECHPHHWPWPWWRF	-15220.44	50	20	50	7.11
CCFDTAKTEPKSARRCSSCC	-15216.78	30	5	55	7.88
QCTCHEHHSRDKRHIWRF	-15213.50	55	15	70	7.75
LCFCTKHVHEDQGNQGHRLC	-15209.40	35	20	55	7.05
ICERFHHSEPDETHRFRRFC	-15205.55	55	20	65	7.12
ICPCHEHKEIRDNGRAWRF	-15165.96	45	20	50	7.72
ISFTPEKFENKNPDGRWKDC	-15163.74	40	20	60	6.27
QCSWCRHYRCDDPERPIKLC	-15156.35	40	20	55	7.59
LCFHLPRSHEDSPKRDQCQLC	-15154.44	40	20	55	6.95
LCGVLSREETQNPEKSTRVC	-15151.40	30	20	60	6.15
ICDCHEHHECEDTRRWRFH	-15129.94	60	20	65	6.05
ICGWCQRPNDDEEQRTALCC	-15116.76	30	15	50	6.02
LDIGSGRTESAQPEYARKLC	-15099.14	30	20	55	6.23
VHFYAHNKDSDDDGRNHERFC	-15095.18	50	20	70	6.34
PDICLDREDFDKGRKQKFC	-15084.63	55	20	60	7.88
TCIWEHRTEPSEKKIETKLC	-15073.44	45	20	65	6.86
IAFHCHKTGPNQKHYYTFQDC	-15040.14	30	20	60	7.91
LKGCCWHHEGDHGNDFTRFC	-15031.58	40	20	50	6.32
EVFRAQHGRVTNPDNDCKLC	-15021.91	35	20	55	6.86

ตารางที่ 3 แสดงตัวยีนชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 1 region 1 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
FCFPWARHDHEQAKMETKPC	-14992.70	40	20	50	7.02
PCFNHEHHPWKCDRWMRDN	-14990.74	50	20	60	6.33
ECKCDHHGEWERRRWWRFC	-14984.22	60	20	60	6.96
ICEHAEKNDWKDYRRTHRLC	-14983.51	55	20	70	7.90

ตารางที่ 4 แสดงตัวเรียงชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 1 region 2

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
NGIIRREFFERLRHSEDEQREKEPERWR	-15699.25	60.71	21.43	71.43	7.14
VVCIRREAFEGRRRCIEQEQREKDPERWR	-15607.08	53.57	21.43	60.71	7.88
RVIERREWFERLRFCHREQESDRPEEAR	-15601.18	60.71	21.43	67.86	6.98
VHIIAREHFERERTCEQEQKDKKPEEPR	-15583.06	60.71	14.29	71.43	6.09
NALIHWRWFQSHFCENEGKEERPPKPQ	-15578.65	39.29	21.43	60.71	6.07
SCCIYRRWFEEWHFKEEGDKRDKPEKEK	-15554.30	60.71	21.43	67.86	6.87
QPQIRREWFERLRWIEREQEKDRPEESR	-15496.21	57.14	21.43	71.43	6.60
EVEARREWFERLRFQEQEQEKRRPEEWR	-15461.28	64.29	21.43	71.43	6.62
EIISRREWFEQDHFCEEHDKRDKPHKWK	-15456.61	64.29	21.43	71.43	6.35
VIIKAKEAFEKLRFNEDSQRRGPEQGH	-15450.50	46.43	21.43	60.71	7.14
IYIGHKWFEQAPFCESRSQRGHDQPSQ	-15424.51	28.57	21.43	60.71	7.12
KIDARYKKFEQWHFNEREGEKDRPEEWR	-15408.41	60.71	21.43	71.43	7.14
TVILPKEWKERLHFCECRNERDNHPNHG	-15407.66	46.43	21.43	60.71	7.12
SWLCHHEHFHELPRIQNEQARDKGEEWR	-15405.93	50.00	21.43	64.29	6.06
SWAVHYKHFKDWNHAEDKGDKDRHERWN	-15393.96	57.14	21.43	71.43	7.44
SWCVHYRDFLQHKLCDRKGDKNKERSE	-15366.51	53.57	21.43	71.43	7.90
VGIIRREWDSRLRSIEREQEKERPEPHP	-15336.35	53.57	21.43	64.29	7.14
VWQHKEKGEKLRFIENRQPEEKREDPN	-15333.77	53.57	21.43	67.86	7.14
RVICQKRFFYEDLNLPECKQNPNGNERDN	-15259.46	35.71	21.43	64.29	6.23
SVLVREPHPRKQQWIEERQSERSQEOWN	-15244.93	39.29	21.43	71.43	7.14
TGGSRKWEFERLRHIEREGEERLREWE	-15242.15	60.71	21.43	67.86	7.14
SCLVHARWFKLQCPEVKQNTQEQEHCN	-15235.50	32.14	21.43	60.71	7.72
VPQVPKEPFKDRFIECRQRREKPEDWT	-15234.49	50.00	21.43	60.71	6.39
NVILPREWF CRTDFAETNGERDHHSTRD	-15230.31	46.43	21.43	64.29	6.07

ตารางที่ 4 แสดงตัวเรียงชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 1 region 2 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
SDCVHWRWFKTWNICERKGGKDQEEKNN	-15195.69	42.86	21.43	64.29	7.88
CWLEHFQHFTRLPFPPEAQGEERHQHRNE	-15192.16	46.43	21.43	60.71	6.51
TVTVEKRWDEQLRFKEEKQRPEKETWN	-15180.76	50.00	21.43	71.43	6.57

ตารางที่ 5 แสดงตัวยีนยงชนิดเปปไทด์ที่ผ่านเกณฑ์ของจีโรทัยปี 2 region 1

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
SCFHFFHFAAREDEASRFC	-33431.76	45	20	55	7.12
LDFEPEKPEAKQPDRLRKYC	-32991.14	50	20	60	6.32
LDFCQRHMKTNEDQREGKYC	-32966.76	45	20	70	6.86
FCFEPRTPRWKEDDRADRFC	-32927.13	50	20	55	6.21
PCFTLKQDKGRENEREIKFC	-32883.37	45	20	60	7.88
LCFELKRQEGKQPKHEHYC	-32880.52	50	20	65	7.88
CCDFRRWTTNETQRESKLC	-32871.43	35	20	65	7.71
LCFNHEKPQTQDLERERKYC	-32778.52	45	20	70	6.86
HCFDVKKEEGRESEKMIKCC	-32753.53	55	20	60	6.80
FIFEHPRGRDEQPERLSKNC	-32752.35	45	20	60	6.96
LDFEPEKPCAQPDLRLQKYC	-32740.11	40	20	55	6.18
LDFMQRHMKSAESQRDGKCC	-32714.33	40	20	60	7.90
LDFEHHKCEAKKPATLTKYC	-32712.70	45	20	60	7.90
LCFNHEKPQNQQDEWRKYC	-32682.45	35	20	70	6.86
GDFSCRKTEAKQPERDIRYC	-32672.67	45	15	65	7.88
LDFESRHTEAKKPATCMKYC	-32670.75	40	20	60	7.88
IDFAFRDWDAKDPNHKCRGC	-32632.46	45	20	50	6.86
LDCEAKKIEAKSGERLAQYC	-32612.94	40	20	55	6.19
DHFWRDCTCKENERLDKLC	-32608.48	55	20	65	6.05
LDFEHTKTEAKQPTTWAKYC	-32595.43	35	20	65	6.96
LDFERPKCECKPGSQLTKYC	-32579.95	35	20	55	7.71
EDFCRRIEAKQPERWAPFC	-32559.19	40	20	50	6.19
DVDFKKCKPRESERFANQC	-32547.14	45	20	60	7.88
LSFEWHRSNHKQPHHECRYC	-32522.16	45	20	70	7.91

ตารางที่ 5 แสดงตัวยีนชนิดเปปไทด์ที่ผ่านเกณฑ์ของซีโรทัยปี 2 region 1 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
TCFHSKEHQIENSDRVCRW	-32517.17	40	20	65	7.02
LCFEPADVHGHPDKRSIKCG	-32515.99	40	20	45	7.02
LDFEPTKTAAKQADTAVRYC	-32513.42	30	20	55	6.20
LDFELREVKAKQPAANQKCC	-32511.49	35	20	50	7.88
RCFSFEHTEVKQPETHGKFC	-32500.99	40	20	60	7.02
LDFESCHGESKQPRQLPKYC	-32483.34	35	20	60	6.86
LDENLDHAHCKQPATYCKCC	-32480.97	30	20	50	6.90
LDFERPKCEAKPNATLTKYC	-32475.56	35	20	55	7.88
LDFGPEQMKDNQAGRQTKYC	-32471.88	30	20	60	6.20
LSFELPKQEDKHGSKDGKYC	-32471.80	45	20	65	6.96

ตารางที่ 6 แสดงตัวเรียงชนิดเปปไทด์ที่ผ่านเกณฑ์ของซีโรทัยปี 2 region 2

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
MKFVRRYFERFHSQCCKERNREDEQPH	-33366.16	53.57	21.43	71.43	7.90
AWFIYRRNFEERRWCSQRDEDKRGEEDR	-33331.19	57.14	21.43	71.43	6.42
VVFTRRCTDEEFRWCSDGRNRNDRWR	-32929.23	46.43	21.43	67.86	7.88
AWFMHRQHQQHMRWCYPNEEDSDRPRCR	-32778.06	42.86	21.43	64.29	7.91
AWLMRHDREFEDLHLSQPNDRSRNHDPNR	-32774.59	46.43	21.43	67.86	7.39
VWIVNGRWFEDGPHPNPNHNDKNRGNQQ	-32737.17	28.57	21.43	57.14	7.28
HVFVTRQHVANSHFCSNPQRQEQDTNCK	-32658.15	28.57	17.86	67.86	7.91
ACNAYEKQFEWRYPASHGHDHRCDDWR	-32529.98	42.86	17.86	60.71	6.33

ตารางที่ 7 แสดงตัวยีนชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 3 region 1

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
FCFRHEHHRPEKAERDFRFC	-33968.32	60	20	60	7.91
FCFRHEHHRCEKPEREFCFC	-33838.57	55	20	55	6.99
FCFRHENHRPEKEERHFRFC	-33769.20	60	20	65	7.91
FCFRAEHHRPEKTESTFRFC	-33768.71	45	20	60	7.90
FCFRHEHHRPEKTERIERFC	-33732.69	60	20	65	7.91
FCFHHPHNPAPERPERGFRFC	-33725.37	45	20	50	7.91
FCFRDEQHRPERPEREWRFC	-33721.20	55	20	60	6.87
FCFHQHHEPEKNERKFRFC	-33704.02	55	20	65	7.91
FCFHSHHNHPEKSERTFRFC	-33697.30	45	20	65	7.91
ICFHHEHHRCEETKRAFKFC	-33685.59	55	20	60	7.75
FCFRHENHHCEKPERFGRFC	-33665.56	50	20	55	7.74
FVFRHEHHRPEQQERQFRCC	-33655.07	50	20	65	7.91
FCFRGEHHRPEQPERLFRCC	-33628.86	45	20	50	7.74
FCFRHHGHCDNEQREFRFC	-33615.77	50	20	60	7.11
FCCRHHHHPEKEERFFRFC	-33613.77	60	20	60	7.11
FCFHAHHCHPERSERAFRFC	-33602.77	45	20	50	7.75
FCFDAEHHRPEKDERGFRFC	-33596.91	55	20	55	6.06
FCFRPEHHRPEKPERFERFC	-33586.28	55	20	55	7.90
FTFRHHHHPEKTEREFKFC	-33584.06	60	20	70	7.28
FTFHHDRHKPERTEKCFQFC	-33575.45	50	20	65	7.91
FCFHHAHHPPEKSERGFRFC	-33573.96	45	20	50	7.91
FCFHQHQHHPERPERAFRFC	-33558.37	45	20	55	7.91
FTFHSHEHHPERQERFFRCC	-33529.95	50	20	65	7.18
FCFAHEHHRPEKAEREFRFT	-33528.22	55	20	60	7.21

ตารางที่ 7 แสดงตัวยีนชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 3 region 1 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
FCIRHEHHRPEESERISRFC	-33528.20	55	20	65	7.12
FCFRHEHHCCCKPERIPRFC	-33528.16	50	20	50	7.62
FCFDHEHHRPEKPERIERFC	-33515.35	60	20	60	6.33
ICFHHEHHKCEETKRAFEFC	-33512.08	55	20	60	6.48
FCFDHNSHHPEKDERCFRFC	-33503.85	50	20	60	6.32
FCFCPREHRRPERTDRAFNLC	-33503.54	40	20	50	7.72
FCFCHNNHHPEKTEREFPPFC	-33491.57	45	20	55	6.48
FIFCHNNHHPEKEERNFRCC	-33473.91	50	20	60	7.11
FCFHCHHHQHPPERPERLCRFC	-33473.04	45	20	50	7.64
FCFRHEHHRPEEPERNDRFC	-33471.07	60	15	65	6.33
FCFRHEHCRPDETERLGRFC	-33455.65	50	20	55	6.96
FCFRHEHQHPEKEDRAFRFC	-33449.90	55	20	60	7.12
FCFHPHHRHQEKEERAFRFC	-33445.69	55	20	60	7.91
FCFRHEHHRPEKAERFFRDC	-33434.78	60	20	60	7.91
FCFCHHGHHPDREERFTCFG	-33422.98	45	20	50	6.48
FCFHHPPHEHPPERPERAFRFC	-33383.49	50	20	50	7.18
FCFRHHTHHPEKDERFFCQC	-33373.36	50	20	60	7.11
FTFHHDRHKPERAETFFKCC	-33353.98	50	20	60	7.91
FCSRHEHHRPEKDERIFRFC	-33351.96	60	20	65	7.91
FGCHHERHHSEKEERIFCFC	-33342.63	55	20	60	6.48
FCFRTEHHRPEKDKDLPRMC	-33338.37	55	20	60	7.90
FCFTHHPHHIEKTERGFRGC	-33328.44	45	20	55	7.91
FCFHHEQHHCEKAKRFRFC	-33299.75	55	20	60	7.75
FAFRHHDDHHPEKTERFFCQC	-33295.75	50	20	60	7.18

ตารางที่ 7 แสดงตัวยีนชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 3 region 1 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
FCFTHHAHPEKPERQLEFC	-33292.42	45	20	55	6.49
CCFRMEHHRCEKPRNQFCFC	-33291.76	40	20	50	7.88
FCCRHEHHRPEKPERLDRFC	-33290.28	60	15	60	7.74
FCFHHEPHHCEKAKRFERFC	-33284.51	55	20	55	7.75
FQFRHEHHEPERPERFFRQC	-33275.37	55	20	65	7.21
FCFRHEHERVEKPERGTCFC	-33273.46	50	20	55	6.96
YFCFCCREHRPENARGEFRFC	-33269.96	40	20	50	7.59
TCFCHHDDHHIEKTERAFCFC	-33267.37	45	20	55	6.47
FNFCIREHRPEREERGERFC	-33255.04	55	20	60	6.87
FCFHHKHEHPETEERPFRC	-33251.31	55	20	60	6.49
FCQRHEHHRPEKEERLFRFC	-33247.48	60	20	65	7.91
FCFRHEHARPEKPERFFREC	-33240.58	55	20	55	7.90
DCFCHHSHHPEKPEKFFRFC	-33233.15	50	20	55	7.11
FCFCNRDHRPERTNRFDQFC	-33221.52	40	20	60	7.72
TCFCHEHHRCEETKRFFKFC	-33218.12	50	20	60	7.62
FCFCERHHEPEKPKTFFRAC	-33215.18	45	20	50	7.72
FCFHHEHHRCETGRRFFRDC	-33215.16	55	20	60	7.75
FCFRDDQHHSDRTRTGWRIC	-33208.67	45	20	65	7.90
ECFCHHRHHPERSERFFRFC	-33208.24	55	20	60	7.75
FIFCHNNHGERADRQFRAC	-33199.39	45	20	55	7.91
FIFCHEQHHPKEERDHKFC	-33197.04	60	20	65	6.05
GCFRHEHSHCDETKRFFRFC	-33189.22	50	20	60	7.74
SCFRHEEHHPKPKDFFRFC	-33174.14	55	20	60	7.12
FCFCHHDDHHPERTERFFCCK	-33164.98	50	20	55	7.06

ตารางที่ 7 แสดงตัวเรียงชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 3 region 1 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
YCFRHEHHRPEKPERMFTGC	-33152.26	50	20	60	7.91
FCFSHEHHRPDHPERIGRPC	-33141.22	50	15	55	7.18
FCFCHHKHHVETTDNRWRQC	-33138.99	45	20	65	7.75
FCFEVREHRPERSERRFRDC	-33138.00	60	20	65	7.90
FSFHHDEHHPERQERFKRFC	-33137.65	60	20	70	7.28
FCFRHHGHHIERMERSRDC	-33130.57	55	20	65	7.91
FTFEFRHNRNDSSDRACRFC	-33129.58	40	20	65	7.90
FCFRHDQHHSDQERQCFRLC	-33124.99	45	20	65	7.05
FTFCFHHQRTDRTDRHDFRC	-33104.85	50	20	70	7.91
ACFCMREHDPERARREFRFC	-33093.41	50	20	50	7.72
FCTRMEHHRPEKPERFHSFC	-33084.67	50	20	60	7.91
TCFCDRHHRPEKTERNFYFC	-33078.92	45	20	65	7.72
FCFRHECHRPKEKPERLFRFC	-33076.39	55	20	55	7.72
FCFCSQEHRPERGDRAFRLC	-33073.96	45	20	55	6.86
FCFDMAEHRPERTDRANRFC	-33068.86	45	20	55	6.86
FVFRHHEHHPKEKTERFDRGC	-33062.79	60	20	65	7.28
FCDRHEHHRPERPERFFRFC	-33060.47	60	20	60	7.91
FCPRHEHCHCEKAERFFRFC	-33059.16	50	20	50	7.62
FCFRHEEHRPEPPEQIFRQC	-33054.72	45	20	55	6.06
FTFRHHAHHCERMETGFRQC	-33047.56	45	20	60	7.91
FCFCFEHMHADRTERTHRDC	-33026.25	50	20	60	6.32
SCFHHECHHCEKPERFFRFC	-33021.27	50	20	55	7.06
TCFCVERHRRPERSERFDRFC	-33019.64	50	20	60	7.72
FTGHHDRHKPERTETFFHFC	-33006.42	50	20	65	7.28

ตารางที่ 7 แสดงตัวบ่งชี้ชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 3 region 1 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
FCFCFREDRAERPERFNRTC	-33005.54	45	20	55	7.71
FTFHPDRHKPERDETFHHGC	-32994.75	50	20	60	6.34
NCFRTREHEPEKTERFFRFC	-32987.01	50	20	65	7.90
DCFRHTDHHCDNARRFFRFC	-32980.37	50	20	60	7.74
FCFGIRDHRPERTDRSFREC	-32968.02	50	20	60	7.90
ISFHHEHHDCDQPRRFWKQC	-32956.82	50	20	65	7.18
FCFRDSTHRNDGERQEWRLC	-32951.46	45	20	65	6.86
FCFCHSHHVEKTGRYDREC	-32946.57	50	20	65	7.11
YCFCHDDHHEKGETYWRNC	-32943.82	45	20	65	6.48
LCFHLEKTEAERERTCCRMC	-32941.15	45	20	55	6.74
FAFRGERPEPERTETCKKFC	-32928.18	45	15	55	7.88
FTFHHDHFHGERTEKCTQFC	-32918.22	45	20	65	7.12
FCFESQEHRPERTRRAFVGC	-32915.52	40	20	55	7.90
LCFRHDHCHPDGSRQMWRTC	-32910.76	40	20	55	7.74
FTFHHDFHKPERTEKCFQPC	-32903.46	45	20	60	7.12
FIFTSNEHRQERPKNDFRCC	-32903.18	40	20	65	7.90
FCFRHRHLHPDREERDFRDC	-32889.97	65	20	65	7.12
FCDCHHPHHTEKTERIWRFC	-32878.23	50	20	60	7.11
TCFCEERYRPEKAERTFKFC	-32869.09	45	20	60	7.71
FCERHEHHRPEKDETFFRFC	-32867.14	60	20	65	6.33
FEFRNEHHPPEREERFFRGC	-32848.76	55	20	60	6.07
FCFCHHAHHEKTEDFCRRC	-32841.39	50	15	55	7.06
FCDRHEHHSPEKPERMFRFC	-32830.30	55	20	60	7.12
FTFNHDRHKPERTATFDCFC	-32815.07	40	20	60	7.02

ตารางที่ 7 แสดงตัวยีนชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 3 region 1 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
FCNCHHEHHIDRTERFTRFC	-32804.08	50	20	65	7.11
FCFCIREFRPECEERREPRHS	-32801.47	50	20	55	7.72
FCARHDHNHCDRMEQAWRFC	-32800.38	45	20	55	7.05
FTFNFRHERCDRTETDQRFC	-32792.98	45	20	70	6.86
NCFRHDHRSRDREKQLWKFC	-32767.61	55	20	70	7.90
FCDDHHEHHPERPERIFRFC	-32760.53	60	20	60	6.05
FCQHNNHHPDERTERFFRFC	-32760.31	50	20	65	7.18
FCGRVEHHRPECNERSFQFT	-32759.62	40	20	60	7.02
FCFTHHSHHDEKERFKWRTC	-32757.66	55	20	70	7.91
FCNHHTHHRPEKPERFCTFC	-32756.27	45	15	60	7.75
ECFHHIRHHPEKQERAFRFC	-32752.87	55	20	60	7.91
FCFRQERCECEKPENMIRRC	-32741.90	45	20	55	7.58
FCFLKHEHRGERTRREECLC	-32734.46	55	20	60	7.72
NCFCIEHHRPEKTERFFNAC	-32728.49	40	20	55	6.95
FCGRHEHHQIECEERHFRFC	-32719.34	55	20	60	6.48
ICIRHDEHRSDDGERGYQKLC	-32718.43	50	20	65	7.02
FVFCERTHKPERTDRPFHGC	-32697.66	45	20	55	7.90
RCFCHHAHVHVEKEERDFCFC	-32692.99	55	20	55	6.47
ICFCHEKERCEDEGRRFFRDC	-32670.26	55	20	55	6.74
LDIEGRNHNPRECRREFKLC	-32668.98	50	20	60	7.90
FDIHCCQHHSQQANRQWKLC	-32662.76	35	20	65	7.11
FCSCIREFRPERERRDFCPC	-32658.67	45	20	50	7.58
FQFCQERYRPERSEREFRCQ	-32637.86	45	20	70	7.88
HCFRVEHHRPDRSEREFYRC	-32635.16	60	20	70	7.91

ตารางที่ 7 แสดงตัวยีนชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 3 region 1 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
FCFRHNHSDTERTGWRDC	-32629.71	45	15	70	7.12
FCFRHEHHAPERPERVFREH	-32627.79	60	20	60	7.28
DVFRHEHHRPERERRFDFFC	-32624.45	65	20	65	7.21
FDFHEDRTEACKERRCCRFM	-32624.34	55	20	60	6.80
QCFRCEHHRLEKPEHMFRCQ	-32611.08	50	20	60	7.74
LAIECQHHECQQANRQWKLC	-32601.79	30	20	55	6.95
FCDCHHQHHEKEERWFCFC	-32601.25	50	20	55	6.05
FCFHDNKTAEHESKRQVALC	-32591.33	40	20	60	7.02
LCQRHDNHHDRERQEWRFC	-32589.11	55	15	70	7.12
LDCHHEKHSCDPSKRAWKLC	-32568.09	50	15	60	7.74
LCFRSHNMCDHTQTLRRCC	-32558.30	35	20	60	7.61
FCDHHAHHWPEKPERFSRFC	-32538.05	50	20	55	7.18
FTEHHDHRHKPERTETFFHFC	-32536.57	55	20	70	6.51
LDIECRHHESNQARQWKLC	-32534.20	40	20	65	7.02
FCQHHTHEHPERPERLFRFC	-32531.75	50	20	60	7.18
FCGCFDSQHCEPERFSRFC	-32531.57	35	20	50	6.74
CCTCVGDYRPERTERFFRCC	-32523.39	35	20	50	7.49
SCFCHDMHHVEREERIARAC	-32510.97	50	20	55	6.32
LDFDHDEPRAGHMRHTRRLC	-32494.72	55	20	60	7.21
LCFHHECTEGGPPKNCFKLC	-32494.37	30	20	40	6.90
SCFRHEHHRPEKPHRMFCEC	-32491.64	55	15	60	7.75
NCFAFHHMHIEREERKTRPC	-32480.85	50	20	60	7.91
RCFHEHHIHPEKEKEREFRFC	-32479.57	65	20	65	7.91
LDITCQHHDTRQIPTHHKLC	-32463.80	40	20	65	7.18

ตารางที่ 7 แสดงตัวยีนชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 3 region 1 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
LDIEGDNHNPRECRRCFKLC	-32462.07	45	20	55	6.80
HCFCdreYNPEKTERFFRNC	-32457.19	45	20	65	6.80
PCTHHKHMHPEREERFLRFC	-32449.72	55	20	60	7.91
IIFHHESHRCDEGKQVTEKC	-32449.61	50	20	65	6.33
CCGCVRErERPEKAERFWRFC	-32435.09	45	20	45	7.58
DCCRKEHHRPEKPERFFYFC	-32417.72	55	20	60	7.72
LDFRPHQNEGGKQEQGVKLC	-32413.22	35	20	55	6.96
LDITCQHHETrQIPThGKLC	-32407.56	35	20	60	7.11
FDFQTRKYHPGECERGTLc	-32406.61	35	20	60	6.86
SCPCHNNHHMEKGERIWRFC	-32403.80	45	20	55	7.75
KDFHHDHRHKPERCEKCFQFC	-32402.13	60	15	65	7.74
VDFPHQNHGPGKCERITKLC	-32399.12	35	20	50	7.90
FISNHHGHAEPTDRSFRFD	-32395.42	45	20	65	6.51
LDIAGTHSHPEPDKTTcRLC	-32389.67	35	15	55	6.05
EHFCHATHCERSDRFFRFC	-32383.08	50	20	60	7.11
LEFHHEKHEIKQKEDYKGC	-32377.24	60	20	70	6.34
LCFEWKKNEAEPRSDDRKLC	-32369.19	50	20	60	6.23
LDFHHQKHEAKIEKRQEHLC	-32342.00	60	20	70	7.28
LDITKDNHRPDECRRPFKLC	-32336.56	50	20	60	7.90
ISIRLPNCKDENAEQARKLC	-32334.86	35	20	55	7.88
LDIECQHHSQHQAQWKLc	-32325.90	40	20	65	6.33
FCNDAREYRAEKTRNMAHFC	-32319.91	40	20	60	7.90
FCkCHHRQHcERADRFcFC	-32305.62	45	20	50	7.90
LCFGCERQEPEQRATHKKLC	-32300.17	40	15	55	7.72

ตารางที่ 7 แสดงตัวเรียงชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 3 region 1 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
MDIRPHRTHCNQQEQKLTLC	-32299.20	35	20	65	7.90
RVFCHDDHHPERSGRFECFC	-32291.28	50	20	55	7.11
FCFPHHSDHVEKPERFPNPC	-32288.40	40	20	50	6.33
HCFCEEQPRFCKERRERFRC	-32286.00	50	20	55	7.59
LDIDCQAGKGTHQQRWKLC	-32277.35	30	20	55	7.90
LDIPERCHRPDECRRPFKLC	-32270.82	50	20	50	7.72
FDFKQQPCECRESGKAALLC	-32269.49	30	20	45	6.06
SCFHHEHHRCCQTPRVFRWD	-32255.46	45	20	60	7.75
REFCEREFPEREERIGRFC	-32251.42	60	20	60	6.29
LDVHCQHNCSEQTPTLRKLC	-32245.65	30	20	60	6.95
FCECPREHRPDRSERFPRFC	-32244.97	50	15	55	7.72
LPIELRETAANRNDKHFANFC	-32242.31	35	20	55	6.96
LDIEPRCHRAKNHERGYELC	-32241.17	50	20	60	7.02
TVFHLDTCKANRNDHHANFC	-32240.52	35	20	60	7.11
LDITCQHHESNQARQHKLC	-32239.41	35	15	70	7.11
STFHGDRHKPCTTETFYQFC	-32238.39	30	20	65	7.02
IDIGWSECRCSDRQQRKLC	-32234.72	40	20	60	6.11
FPFINDKTEAKQQERGTFQFG	-32234.61	30	20	60	6.38
LDIRHDEHHLDDGHHQGRRC	-32228.33	55	15	70	6.51
LDITLDKGKSEQPTCWKNC	-32222.56	30	20	60	6.11
KDFCVERNQPEKSECHWRFC	-32221.70	45	20	60	6.80
YFTTTHRHGGDDATLDKLC	-32213.65	40	20	65	6.33
FCERHEHSTCDETKRVFRFH	-32212.08	55	20	70	7.12
LDIEHQKHSCDPGATLRKLC	-32210.60	40	20	55	7.02

ตารางที่ 7 แสดงตัวยีนชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 3 region 1 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
HVFRTECRSDGERTGKRLC	-32207.26	50	15	65	7.90
LDVHCQHNSQQENRQRFIC	-32206.23	30	20	65	6.95
LICHHEKDENTTGKRAWKLC	-32195.70	45	20	60	7.90
LDITCQHGDAQHDHRC SKLC	-32184.89	40	15	60	6.32
FPFPKDNHRPDHVERGTQFG	-32176.92	40	20	55	7.28
FDIHCQHHESSQQANRQLKEC	-32169.54	40	15	70	6.33
FAFIPDNRRSGKCENDTQFG	-32163.59	30	20	55	6.20
FCTAVTKHETDNTDRDKFC	-32163.29	40	15	65	7.90
ECFCFRHRRCCEECERFKRFC	-32147.50	55	20	55	7.88
DTFHHEETPEAGRGTRIWKLC	-32146.82	40	20	55	7.12
LNIDCQAHKGTHDETHLKLK	-32145.99	40	20	60	6.33
LCIEEQKTEATRTEITLRKLK	-32125.32	40	20	65	6.21
LDIELEKRRPQNSEGCPKLC	-32109.76	40	20	55	6.19
LIPHHEKHSCDPLKTA AKLC	-32108.45	40	20	50	7.91
LDIALKNHDAQQDP THGKLC	-32101.71	35	20	55	6.05
FCHCHRPQHFD DSKRRVEFC	-32100.22	50	20	60	7.74
IMITQQHCNSDHAHNCRKLC	-32089.94	30	20	60	7.74
QVFHIRETEAQRNDHHRNFC	-32086.90	45	20	70	7.21
LDGQMRNHNHCERREWKLK	-32084.14	45	20	60	7.72
LCCRCCQHNADKDKTLRTLC	-32082.89	35	15	55	7.88
VDFPMQNHPGKCERGT KLC	-32082.80	30	20	45	7.90
LDIEHDKCELKGNKQGRKLK	-32081.12	50	20	60	7.88
AIFCQPTHHIERSDRFERPC	-32047.09	40	20	55	7.02
ATFEWHDDCCATKGSRALKLC	-32042.08	30	20	45	7.72

ตารางที่ 7 แสดงตัวเรียงชนิดเปปไทด์ที่ได้ผ่านเกณฑ์ของศิริราชปี 3 region 1 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
FCECVERSRDEKEERFFRKC	-32039.62	60	20	65	6.21
FDFDHGQATKENSATKFKFN	-32037.13	35	20	65	7.12
LLFPKDNHRPDQCERGTFQH	-32029.38	40	20	60	7.12
WCIRLQKTEAAPKDSHTCLC	-32020.90	30	20	50	7.72

ตารางที่ 8 แสดงตัวยีนชนิดเปปไทด์ที่ผ่านเกณฑ์ของจีโรทัยปี 3 region 2

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
TYIFRQEDFRRCRMICGERKDDHHPRHE	-32955.19	57.14	21.43	67.86	7.91
RRFCERRDFEQWRLPWTNNHSHSDHHPWE	-32852.12	50.00	21.43	67.86	7.28
AYMIRRHQHFEDLRAAEQRDGRDTPPWN	-32851.45	39.29	21.43	60.71	7.14
SCLIRRWFFERLHLCEQEAGRHHHTPPQN	-32724.36	35.71	21.43	57.14	7.91
MYLNRREAFQRLHNAEQRDEEHHTPPWN	-32693.64	42.86	21.43	67.86	6.35
CYEIRREDFERLRLARQGDGDEKHHHTPPWN	-32638.25	50.00	21.43	64.29	7.12
ANMKRREWFERIEMPEQAGEHHHTPRWN	-32554.07	46.43	21.43	60.71	6.09
AYGIRREAFERLRQVYQEDRDHHCPTN	-32488.29	42.86	21.43	64.29	7.12
AYFAHRQHFDTHLMDDSGKKQETPTWN	-32450.29	39.29	21.43	67.86	6.51
RYFLRKDQDDNPRCVVEEREKERHRPIWE	-32424.15	60.71	21.43	71.43	6.98
ASCVRRDWWFFNLKACSAEGNTRHPWE	-32379.89	32.14	21.43	50.00	6.86
MTILQRRHIDEWKCNDDGSEPNNKPGWH	-32378.33	35.71	21.43	60.71	7.12
ACFIRRKRSDDLPLPEERNNEETPTWE	-32350.53	42.86	17.86	64.29	6.37
IALVRRQANENARMGYRQDQRDHEQWE	-32309.16	39.29	21.43	67.86	7.14
ARLIRRQWFEDLRANHDCSANDETPKWN	-32296.93	39.29	21.43	60.71	6.96
MTTRQRQQFEDWRLPPTNNHKSDDHHPWE	-32295.31	42.86	17.86	71.43	7.44
AYMRPREEFFHHPGPWNDDKNEKHGHE	-32291.93	50.00	21.43	60.71	6.06
QTHIQRQQFEDLNLPWARTHHSNDHPWE	-32256.23	35.71	21.43	67.86	6.05
CYIRREWAERLRLSTQEDEEHHTPPWN	-32239.20	46.43	21.43	67.86	6.07
AYQVKRQWFSDWHLPEERNNEETPTTK	-32226.49	39.29	21.43	71.43	7.12
AWFGRLHHFNDWPCNDQGSEERNKPKCN	-32197.78	35.71	17.86	57.14	7.02
ACFAHNHHFFDPELMERSRENTHDPRWN	-32180.78	42.86	21.43	60.71	6.05
AYPIRRDYACHLKLKLCYNRDEENNHPSTE	-32160.57	39.29	21.43	67.86	6.06
AFFCHNHHFNDLPLLADSGKEPPNPETHN	-32156.92	28.57	21.43	50.00	6.05

ตารางที่ 8 แสดงตัวยีนชนิดเปปไทด์ที่ผ่านเกณฑ์ของซีโรทัยปี 3 region 2 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
SIFMKREGFFDWKCNDFGSEKQSGKPAE	-32153.11	35.71	21.43	53.57	6.32
MTERKRQQFEDWRLPWTNNHPPSDHHPWE	-32137.15	42.86	21.43	64.29	6.35
ASMVRRQTFEQARHAEYEAGEHKMPWPN	-32117.96	35.71	21.43	57.14	7.28
APLIRRQGFHVGMPEPESRETKTTPWPN	-32104.97	32.14	21.43	50.00	7.14
AGTLHRQHFDRFMGCHTSDNREDDPQWK	-32092.84	42.86	17.86	64.29	6.34
AYMIRRENRDDQHSIRPESIREQPPTWQ	-32083.50	39.29	21.43	67.86	7.14

ตารางที่ 9 แสดงตัวบ่งชี้ชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 4 region 1

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
ICFCEHPRWERNREDTRFC	-16420.96	50	20	60	6.96
ICFCCHHPRCEKNKEPIRFC	-16318.78	40	20	45	7.88
GCFCTHHPRWERQREDIRFC	-16255.94	45	20	55	7.74
FCFCHHEHRTKETKELIRHC	-16169.89	55	20	65	7.75
ACFHAHHPRWRENCECLRFC	-16169.07	40	20	45	7.62
ICFHAHHPRTRENGETLRF	-16166.52	40	20	55	7.91
FCFCHDDHHWQENREMDRPC	-16091.85	50	20	60	6.05
THFCHDDHRNDKPREMVRF	-16090.48	55	20	65	7.91
ICFCAHHPRWDENREMPRHC	-16079.79	45	20	50	7.05
FCFCHHEHRENKKNKEMIRNC	-16074.47	50	20	65	7.74
CQFCHHNHREEKNRELLRFC	-16063.90	50	20	65	7.74
FCFCPHQPRCDRQRELNCFC	-16044.98	30	20	45	7.50
ECFCHNHPRWEKNKEMLRTC	-16044.20	45	20	60	7.72
IVFCHQHPRDEKQREKQTFC	-16035.41	45	20	65	7.90
FCFSAHHPAWERQKEALRNC	-16031.51	35	20	50	7.90
FCFHVHHPRWEKNRETNREC	-16021.74	50	20	65	7.91
FCECHEHPRWERQREGLRFC	-16020.00	50	20	55	6.96
NCFGDDHPRWERPSEGLRFC	-16018.72	40	20	50	7.02
PCFHCHCHPRWDKNREELRYC	-16016.85	50	20	60	7.74
ICFHCEHPRWEKNREMPRNC	-16006.32	45	20	55	7.72
ICFCDHHPRNEKQKELIKSC	-16005.19	45	20	60	7.72
ICFCHDHPRQKEDREMLASC	-16003.11	45	20	55	6.05
SVFCHAHHGWEKNREATRFC	-15984.67	40	20	55	7.91
ICFCEHPRSKENGEMIRHC	-15982.94	45	20	55	7.05

ตารางที่ 9 แสดงตัวยีนชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 4 region 1 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
FCECHHCPRWDKNKEMGRFC	-15963.68	45	20	50	7.61
HCFWHHEHRTKNDREPIRFC	-15960.17	55	20	65	7.91
ECFHRHHPRWREDREMIRTC	-15958.84	60	20	65	7.91
LCFHTHRDRNKESCSLCRCC	-15955.33	40	15	60	7.88
QTFCHSHHWHRENKEMIRTC	-15951.41	45	20	70	7.91
ICDCHNHPRWHENREMRRFC	-15949.86	50	20	60	7.74
LCFMTHRDRNKESCSLCRCC	-15949.27	35	20	55	7.88
NCACTHDPKWEKQKELLRFC	-15943.52	40	20	55	7.72
PCFCHTHPREKENREMPVFC	-15935.38	40	20	50	6.95
DCFCFHHRWEKNRELERCC	-15925.96	55	20	60	6.99
ICDCHHTPRWKENREMGRFC	-15922.24	45	20	55	7.72
DCFCSHHPRCAEQKEMLRFC	-15901.95	40	20	50	6.90
FCNHHEHHRWERPRELERFC	-15901.74	60	20	65	7.18
ICPCHHTHRWDKNREMDRFC	-15901.66	50	20	60	7.74
ECFCDDHPRWERQREMLRKC	-15894.33	55	20	60	7.72
ICPHHDHPRWEKNREMLRGC	-15886.41	50	20	55	7.91
ICFHHAHPRWRENREMSRDD	-15884.07	55	20	65	7.21
ITFCCHDHPRAERNREMI SRC	-15878.77	45	20	60	7.90
AHHCCHHPRWNNPDRLIQFC	-15868.21	35	20	50	7.75
SCACHHCHRWKENDEWLRFCC	-15856.41	45	20	55	6.99
ICQCCHHPRWDRNKEMIRDC	-15839.15	45	20	55	7.61
GCEHGHHPRWKEDREMLRFFC	-15834.94	55	20	55	7.12
DCFCCKHHPRWKENREMI RDC	-15808.55	55	20	60	7.72
FCCYHAHNHGKENREMI RGC	-15801.16	40	20	55	7.74

ตารางที่ 9 แสดงตัวยีนชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 4 region 1 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
FCDHHLHPRWEQQRECMRCC	-15796.16	45	20	55	6.99
DCFYHPHPRWERQKESIRPC	-15777.95	45	20	60	7.90
FCNPCHQPKWEEAKENLRFC	-15768.15	35	20	50	6.80
FCECHHSHRHRENREMLRFC	-15753.51	55	20	65	7.75
LDFCHTKHTAPDETKALCFC	-15743.52	35	20	50	6.04
LDFCTHGRCKESKSTTMC	-15726.40	35	15	65	7.72
FCDCHAHPRWDKDRMHRFC	-15700.36	55	20	55	7.05
LCFHHGRHPQERPEDWIHNC	-15697.69	45	20	55	6.49
DCFHNCHQRIKENRESMRFC	-15681.42	45	20	65	7.72
FRFIHRHRCKPAPEEDQYC	-15676.22	50	20	60	7.91
HCTMTHTRDKNDREMITFC	-15656.60	45	20	70	7.90
FDCCHHVPRWKEDKEMNRRC	-15655.77	55	20	60	7.72
PCFCHHAWGEEKPRESIRFS	-15640.07	40	20	50	7.02
IFFGRQDHRGEHHD TACTCM	-15632.36	40	20	55	6.33
IVECHHGHWHWEKNRENLRDC	-15626.11	55	20	65	6.49
FCVNNHHAPDWRENREMPRDA	-15609.79	45	20	55	6.06
CFFCGKEKTTAKEVKQLGECC	-15589.63	30	20	45	6.04
FCEHQHHPRWCEDREVL RNC	-15561.65	50	20	60	6.32
FCEHCNHPRWEKNREFIREC	-15549.59	50	20	60	6.96
LDFVCDTTHCTSQKHGHYC	-15547.41	30	20	65	6.32
WCRCHNRDSWDKEREMTRFC	-15530.40	50	20	65	7.72
LDFEHDAKGRPSRTCRTVC	-15513.71	40	15	55	7.72
SVFQNHHSRWKPSREEINDC	-15503.38	40	20	70	7.12
LCFEHKKKTCPDQNGALLRTT	-15480.78	30	20	55	7.90

ตารางที่ 9 แสดงตัวเรียงชนิดเปปไทด์ที่ผ่านเกณฑ์ของที่ร้อยละ 4 region 1 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
ACFDHKESNLKESRHLHYCQ	-15470.29	45	20	70	7.11
HVFCWEKTRNHDKEKGNQYC	-15470.14	45	20	70	7.90
LDVPGKRAEWNNHNCDCRITYC	-15451.28	35	20	55	6.79
WCFDTKESNSKEVTPHARIS	-15450.60	35	20	65	6.96
WCRCTRECRNHESRMVACLC	-15446.52	35	20	55	7.88
FCFDTKESNSKEVYPGSKAC	-15445.88	30	20	60	6.13
IDFHNNHNRGRQEAEMYAPC	-15442.30	35	20	60	6.05
WCFELSKTNAKEVHKTQQCS	-15429.91	30	20	65	7.88

ตารางที่ 10 แสดงตัวบ่งชี้ชนิดเปปไทด์ที่ได้ที่ผ่านเกณฑ์ของซีโรทัยป์ 4 region 2 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
TVFVRRREEFERWRACDSSHDDNRRSR	-16443.10	50.00	21.43	71.43	7.02
TVFCRREEFERWRFCTSDSRHQDENRWR	-16423.27	50.00	21.43	71.43	7.90
TVFIRREDFERWRFCECDSHHRDHRRER	-16406.29	64.29	21.43	71.43	7.91
TQFIRREEFERWRFCTCDSHHEDNRRWR	-16392.06	53.57	21.43	71.43	7.90
TVFVRRREEAERWRFCCSDSRREDENRWR	-16385.24	53.57	21.43	67.86	7.88
DTFIRREEFERWRFQCDSNHQDNRWR	-16328.79	50.00	21.43	71.43	7.90
TVGIRREEFERWRFCDCESRHRDENRWR	-16304.31	57.14	21.43	67.86	7.90
IVFIRREEFERDRGCYCDSKREDNRQCR	-16293.87	53.57	21.43	67.86	7.71
GCFCRREEFERWRFCHCDWRRENNHRWC	-16292.01	50.00	21.43	57.14	7.88
TVFIRREECERWRFYCDSREDNRDR	-16275.21	53.57	21.43	67.86	7.71
TVFIRREEFERDRFCDCSHHRDTRRWR	-16274.22	60.71	21.43	71.43	7.90
TVFIRREEFERPRFCACDSRHHDENRWR	-16267.23	53.57	21.43	64.29	7.90
TVFDRREEFERARFCYCDSRHEDNNRWR	-16264.76	53.57	21.43	71.43	6.87
TVFIRREEDERWRFCDSSHKKDNQRWR	-16252.01	53.57	21.43	71.43	7.90
TVFIRREEFERHRFCSCSRREDNGQWR	-16250.81	50.00	21.43	67.86	7.90
TVFDRREEFERWNFCYCDGREHDHRRPR	-16246.09	57.14	21.43	71.43	7.02
TVFTIRREECERWPFCMSRREDNNRWR	-16229.30	46.43	21.43	64.29	7.71
TVFDRREEQERWAFYCDSRHHDENRWR	-16226.93	53.57	21.43	71.43	6.07
TVFVRPEEFERWRFCCSDSRHRDDNRTR	-16223.07	50.00	21.43	67.86	7.90
TDFIRREEFERWRCYCDSRREDNRQWQ	-16212.60	50.00	21.43	71.43	6.29
TVFIKREEAERCRCFCYSDSAHHDNRHWR	-16211.90	50.00	21.43	67.86	7.91
TVCIAREEFERWRFYCESRREENRRDR	-16206.48	53.57	21.43	67.86	7.71
TVFVRRREEAERWHFCYCDSRHQDNRREK	-16202.64	53.57	21.43	71.43	7.90
TWFBVHREHFDRWRLCCSDGRREDNPPN	-16189.73	46.43	21.43	60.71	6.06

ตารางที่ 10 แสดงตัวบ่งชี้ชนิดเปปไทด์ที่ได้ที่ผ่านเกณฑ์ของวิธีโรทีย์ปี 4 region 2 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
TVFIRIEEAERWRFQCDSHHQDHRRSR	-16184.23	50.00	21.43	67.86	7.91
TVFIRREEGERNRFYCDSQHHDNRQWR	-16160.49	46.43	21.43	71.43	7.90
TCFIRDEEFERWRFYCESRRKENRRDA	-16133.19	53.57	21.43	67.86	7.71
TVFGRREESERWPFYCSDGRREDNRQWQ	-16131.51	42.86	21.43	67.86	6.35
TVFCRREEIERTDFCPCDSRHHQDNRWR	-16124.64	50.00	17.86	67.86	7.72
THFIRREEPEKWRFCYCDSHDHDNTKWR	-16086.96	53.57	21.43	71.43	7.12
CVFDRREEPERWRFYCDSRHQDNRRT	-16066.79	50.00	21.43	67.86	7.72
IPFIRREEFERGRFCQCDSDRENERRWR	-16056.02	53.57	21.43	64.29	7.88
TDCIRREEFERWRFCDSCRREDENRWR	-16033.72	57.14	17.86	71.43	6.27
TVSCRDEEFERMRFCYCDSRHRDERRWR	-16026.00	57.14	21.43	71.43	7.72
TVAIRGEEFERWRFCTDSHHDDNRRRLR	-16019.55	50.00	21.43	64.29	7.12
TVFSMREEAERWRFYCESRREDNNRDR	-16011.34	50.00	21.43	71.43	6.30
TNSWRREEFERWRFYCDSHHGDNARWR	-15992.69	46.43	21.43	67.86	7.90
TFFVHKETPERWVRCCEDSRREDNNRWR	-15966.34	50.00	21.43	67.86	7.90
TFGVRREEPERWRFYCDSRHRDENRWD	-15958.22	53.57	21.43	67.86	6.87
TVFIRREEPEEQPFYCETRRQDNRWS	-15941.68	42.86	21.43	67.86	6.28
TFAQRREEIERWRFYCDSRREDENRWC	-15940.30	50.00	21.43	71.43	6.30
THGIHREWFERGRFCCPDVRREDNNSWR	-15933.90	46.43	21.43	60.71	7.90
AVNIRCEEFEDWRFYCSDKERKDNRRCR	-15911.57	53.57	21.43	67.86	7.71
TWIIQRRTFEKAELACCHGHDREHKDWE	-15905.47	50.00	21.43	60.71	6.33
TWFBVHREAFLKWRQCCSDGNREDQNSGN	-15894.35	32.14	21.43	60.71	6.86
GWAVHREHGDRWRFCDPRREDENRW	-15882.08	50.00	21.43	57.14	7.02
TYFFHRRQAFERLESADPKGRDPGNHHWN	-15854.36	39.29	21.43	60.71	7.39
TWFAHREEFRRQDLCCRDGKREDNCNNWN	-15844.90	46.43	17.86	64.29	6.80

ตารางที่ 10 แสดงตัวบ่งชี้ชนิดเปปไทด์ที่ได้ที่ผ่านเกณฑ์ของซีโรทัยปี 4 region 2 (ตารางต่อเนื่อง)

sequence	score	% hydrophilic	%hydrophobic	% forming H bond	Iso.point
TVFAKGRWFTDLRHCDGHPHQNEERWN	-15827.50	39.29	21.43	60.71	7.21
TWVHVHREWFRGARFCQSDGRRDDNNGEN	-15809.44	39.29	21.43	60.71	6.96
TWVHVHREWCRRYEFCCPDGRRNDNNPEN	-15804.11	39.29	21.43	60.71	6.80
TVWDRHEEDERWRTCYCDSPHRHHMRCR	-15788.95	57.14	17.86	71.43	7.12
TVGVRREESEKLDFIYCQKQHSEHSHGH	-15772.29	46.43	21.43	71.43	6.51
AWCCQKWFERLSLPWTKGADHHEHRHN	-15734.33	39.29	21.43	57.14	7.91
SVFVNKQFFSQLECCRNARGHHETHQN	-15727.85	28.57	21.43	64.29	7.91
TWLCRRQWFDDLLHLCHTAGRNKHDTEEN	-15709.88	42.86	21.43	64.29	6.33
VIFVKEQAFDQQRQLCHPGSPDRHDDHHN	-15702.26	42.86	21.43	60.71	6.05
TWLCHREAMAKWGYCCSTPRRHDCDGWN	-15701.21	32.14	21.43	50.00	7.61
AVIQERRWFKECEPIWDHPHNRPNRERHN	-15692.08	46.43	21.43	60.71	7.21
TWLCDREAMAKWGYCCDRAQRGDNNNGWN	-15629.55	28.57	21.43	50.00	6.09
TPFGSREWFERDRFTYCEKPNNEHNRAH	-15624.01	42.86	17.86	67.86	7.12
PPFVHREWFERTIECSDSRREDNNRWR	-15599.61	46.43	21.43	67.86	6.96
TVWCQRQWMDLLRHCDQRPNGPEKREG	-15597.54	42.86	17.86	60.71	7.02
TWLCHREAMAKWGPCWDRAQRGDNNNGWN	-15592.80	28.57	21.43	46.43	7.90
TWLPHREAMAKWGYCCSTPRSHDCDGWN	-15587.02	28.57	21.43	50.00	6.95
TVHIRDTEFERQHFCPCESKRNDRCARWR	-15582.74	46.43	17.86	64.29	7.72
TQLLHKQWFERLPLPRTDAEQEGNNHAR	-15579.45	35.71	21.43	60.71	7.28
KCIVHRRDFDEQPLPYTAKAETNHCRFQ	-15543.98	39.29	21.43	60.71	7.90
NQCVHRHWFSKLPLPDDAQPSDHENHWN	-15534.32	35.71	21.43	60.71	6.05
RWLCHREAMWRWNICRSDGCDEDDNNKTS	-15526.71	39.29	21.43	60.71	6.80

ส่วนที่ 2 สกรีนหายาหรือสารต้านไวรัสเดงกี

จากการสกรีนหายาหรือสารต้านไวรัสเดงกี ด้วยการทำ molecular docking โดยใช้ AutoDock version 4.0 ผลที่ได้สามารถสกรีนได้ยาหรือสารต้านไวรัสเดงกีที่สามารถจับกับโครงสร้างไวรัสเดงกี 4 ซีโรทัยป์ ด้วยค่า K_i ที่ต่ำที่สุด ได้จำนวนทั้งสิ้น 1,175 ลิแกนด์ จากลิแกนด์ทั้งหมด 966,297 ลิแกนด์ ซึ่งเป็นลิแกนด์ได้ทำการคัดกรองที่มีขนาดสอดคล้องกับขนาดของ active site ของโปรตีน type 2 ของทั้งสี่โปรตีนดังกล่าวข้างต้น (ลิแกนด์ มีขนาดมากกว่า $500 \text{ \AA}^3$ และมีขนาดไม่เกิน $5000 \text{ \AA}^3$) ผลของลิแกนด์ที่ได้ ดังตารางที่ 11 สามารถแบ่งเป็นกลุ่มยาที่ได้จากฐานข้อมูล ChemBank (Set1) จำนวน 59 ลิแกนด์, ChemPDB (Set2) จำนวน 60 ลิแกนด์, KEGG Ligand (Set3) จำนวน 91 ลิแกนด์, Anti-HIV NCI (Set4) จำนวน 162 ลิแกนด์, Drug-likeness NCI (Set5) จำนวน 229 ลิแกนด์, Not annotated NCI (Set6) จำนวน 49 ลิแกนด์, AKos GmbH (Set7) จำนวน 258 ลิแกนด์, Asinex Ltd (Set8) จำนวน 236 ลิแกนด์ และฐานข้อมูล Tim Tec subset (Set9) จำนวน 31 ลิแกนด์ สำหรับโปรตีน 2FOM ไม่มีการรันในกลุ่มยาชุดที่ 7

ลิแกนด์ที่ได้จากการสกรีนยาด้วยการรัน docking ของโปรตีน 1OKE, 1R6A, 2BMF และ 2FOM ในแต่ละซีโรทัยป์ แสดงดังตารางที่ 12, 13, 14 และ 15 ตามลำดับ ซึ่งแต่ละตาราง ประกอบด้วยชุดของกลุ่มยาหมายเลขของยาในแต่ละกลุ่ม, สูตรเคมีและค่า K_i ที่คำนวณได้จากโปรแกรม AutoDock

ตารางที่ 12 ผลของกลุ่มยาที่ได้จากการสกรีนด้วยโปรแกรม AutoDock ของกลุ่มโปรตีนเนอีลอป (IOKE) type 1, 2, 3 และ 4 (ตารางต่อเนื่อง)

No.	เนอีลอป ซีโรทัยป์ 1				เนอีลอป ซีโรทัยป์ 2				เนอีลอป ซีโรทัยป์ 3				เนอีลอป ซีโรทัยป์ 4			
	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula
25	set1	1177	1.01E-04	C15H14O3	set7	152142	8.43E-10	C48H38N4O2	set9	6505	1.50E-05	C16H13BN4O2S	set5	169710	1.91E-09	C30H39FN2O2
26	set1	625	1.02E-04	C11H8O5	set5	178859	8.91E-10	C33H50N4O6	set4	11875	1.53E-05	C21H22N2	set7	141763	1.94E-09	C32H29Cl2N3OS
27	set6	14680	1.03E-04	C15H11N3O	set4	5044	9.33E-10	C38H54O4	set5	170352	1.56E-05	C22H15N3S	set3	8452	2.00E-09	C30H26O10
28	set6	15220	1.04E-04	C16H12N2O4	set4	4820	9.38E-10	C36H52O15	set2	2615	1.60E-05	C20H20N2O5S	set7	183794	2.05E-09	C30H30N2O5
29	set1	2313	1.06E-04	C15H20O4	set7	241593	9.50E-10	C29H23BN2O2	set4	7704	1.62E-05	C22H26N2O	set3	7037	2.07E-09	C30H46O7
30	set1	2139	1.08E-04	C16H13NO4	set7	159806	9.55E-10	C35H30N2O6	set4	7471	1.70E-05	C23H25NO	set9	317	2.20E-09	C23H26N4O4S3
31	set2	3617	1.08E-04	C6HB4N3	set7	290161	9.55E-10	C37H27N3O2	set9	1494	1.70E-05	C20H15FN2O2S2	set7	182985	2.27E-09	C32H29NO2
32	set2	1131	1.10E-04	C15H24	set5	156273	9.67E-10	C34H56O2	set9	4901	1.72E-05	C19H12N4O4S	set5	166656	2.31E-09	C33H45NO3S
33	set1	774	1.12E-04	C14H8O6	set3	7154	9.83E-10	C34H28O9	set9	6352	1.73E-05	C20H16ClNO2S2	set7	183785	2.37E-09	C34H38N2O6
34	set2	3206	1.13E-04	C12H17NO7S	set7	290155	1.05E-09	C41H31N3O2	set7	169421	1.82E-05	C24H26N2S	set1	1372	2.42E-09	C28H44O
35	set2	2290	1.15E-04	C18H15NO2	set7	185828	1.07E-09	C34H26N2O2	set9	988	1.85E-05	C18H18N2O3S2	set7	183364	2.50E-09	C28H25BN2O4
36	set1	898	1.17E-04	C14H8O4	set7	158182	1.16E-09	C38H36ClN3O2	set4	10392	1.88E-05	C16H12N4OS	set7	183825	2.71E-09	C35H32N2O5
37	set5	184016	1.18E-04	C21H16O3	set7	142609	1.21E-09	C35H35Cl2N3OS	set6	14280	1.90E-05	C19H12Cl3N3O2S	set3	7067	2.77E-09	C28H38O6
38	set9	3453	1.18E-04	C16H16BrN3	set7	142346	1.22E-09	C34H23Cl2N3OS	set5	181653	1.91E-05	C23H22	set7	183736	2.77E-09	C26H30N2O5
39	set1	2311	1.19E-04	C12H7Cl3O2	set5	169083	1.27E-09	C31H46O6S4	set5	164598	1.93E-05	C20H12O3S	set5	161932	2.78E-09	C29H39N3O2S
40	set1	234	1.21E-04	C15H18N2S	set7	186704	1.29E-09	C36H26N2O2	set7	182999	2.03E-05	C24H20N2O2	set6	14810	2.87E-09	C26H33NO2
41	set2	2553	1.21E-04	C14H10N2O4	set7	183753	1.30E-09	C30H29N3O6	set9	656	2.03E-05	C19H12N2O4	set5	175033	2.88E-09	C28H49NO
42	set1	1859	1.23E-04	C16H21NO2	set7	186168	1.30E-09	C40H22N2O2	set2	2090	2.04E-05	C20H19IN2	set4	5782	2.91E-09	C28H14O4S2
43	set9	5411	1.24E-04	C17H11ClO4	set7	159805	1.32E-09	C33H26N2O6	set4	8744	2.05E-05	C17H14N2O2S	set3	7038	3.04E-09	C30H48O7
44	set6	14986	1.27E-04	C15H12N2O	set5	156487	1.33E-09	C32H55NOS	set9	4767	2.05E-05	C19H14ClNO	set5	189902	3.13E-09	C29H46N2O4
45	set6	14613	1.29E-04	C10H9N5	set5	179331	1.34E-09	C37H34N2O	set9	803	2.07E-05	C17H19N5S	set6	15179	3.14E-09	C34H31NO3
46	set1	750	1.30E-04	C9H5I2NO	set7	188382	1.35E-09	C32H22N2O2	set6	14278	2.08E-05	C17H13Cl3N2O2S	set8	25451	3.18E-09	C28H26ClN3O3S
47	set2	3627	1.30E-04	C12H7Cl3O2	set7	290128	1.36E-09	C35H23N3O2	set8	29510	2.08E-05	C21H26N4OS	set3	7407	3.20E-09	C32H32O14
48	set6	13714	1.31E-04	C10H4Br2O4	set2	1353	1.38E-09	C46H50NO2	set9	500	2.09E-05	C16H13ClNO4O2S	set3	4030	3.22E-09	C28H44O
49	set1	2099	1.33E-04	C15H14O3	set7	142827	1.38E-09	C37H34BN3O2	set9	5924	2.11E-05	C16H12Br2N6O2	set5	161936	3.24E-09	C29H37N3O2S
50	set1	1226	1.36E-04	C15H14O6	set4	12866	1.40E-09	C39H24N6O5	set5	176769	2.19E-05	C22H34O2	set3	5021	3.32E-09	C30H42O

ตารางที่ 12 ผลของกลุ่มยาที่ได้จากการสกรีนด้วยโปรแกรม AutoDock ของกลุ่มโปรตีนเอนโดลอป (IOKE) type 1, 2, 3 และ 4 (ตารางต่อเนื่อง)

No.	เอนโดลอป ซีโรทัยป์ 1				เอนโดลอป ซีโรทัยป์ 2				เอนโดลอป ซีโรทัยป์ 3				เอนโดลอป ซีโรทัยป์ 4			
	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula
51	set1	1106	1.38E-04	C16H18N2O	set7	183376	1.40E-09	C29H25N3O4	set9	1256	2.24E-05	C17H16N2O3S	set4	4862	3.35E-09	C32H46O9
52	set1	920	1.39E-04	C15H10O2	set7	142911	1.41E-09	C34H32O3N3OS	set5	188484	2.26E-05	C23H20N2O2	set7	125805	3.36E-09	C27H28O3N3O2S
53	set1	668	1.39E-04	C14H10O3	set7	141890	1.42E-09	C35H36O3N3OS	set5	181608	2.27E-05	C25H17N3O2	set7	233515	3.40E-09	C25H18O2N2O4
54	set1	1074	1.40E-04	C14H12O3S	set1	2216	1.45E-09	C38H38N4O4	set9	4377	2.28E-05	C24H17NO4	set3	7358	3.44E-09	C26H36O11
55	set2	751	1.40E-04	C15H13BrFNO2	set3	7044	1.45E-09	C29H40O8	set9	3379	2.29E-05	C23H16BrClFNO3	set3	1747	3.50E-09	C30H50O
56	set2	362	1.40E-04	C14H24O2	set7	152250	1.48E-09	C36H30N4O4	set5	169824	2.30E-05	C20H18N2O	set7	156907	3.70E-09	C29H20O3NO5S
57	set1	1557	1.41E-04	C15H18	set7	242535	1.49E-09	C24H25BrClNO	set5	183260	2.35E-05	C24H24	set7	142403	3.71E-09	C31H27BrClN3OS
58	set1	1488	1.42E-04	C13H14N2	set7	183871	1.50E-09	C34H29N3O6	set9	1604	2.35E-05	C23H20N4O2S2	set7	141081	3.73E-09	C31H29O4N4OS
59	set2	3658	1.42E-04	C13H14N2	set7	141538	1.51E-09	C38H37N3O2	set5	157464	2.41E-05	C20H12O3S	set7	183788	3.79E-09	C31H32N2O4
60	set2	2801	1.42E-04	C14H14O	set3	9042	1.52E-09	C45H73NO16	set5	157995	2.42E-05	C19H14N2O	set7	182936	3.79E-09	C34H27NO2
61	set6	13774	1.42E-04	C11H7BrO4	set7	158134	1.52E-09	C38H37N3OS	set7	241771	2.47E-05	C24H27NO2	set5	187263	3.83E-09	C29H22O3NO5O2S
62	set4	5803	1.44E-04	C19H18N2O2	set7	284608	1.52E-09	C28H16O2N2O2S2	set5	166774	2.52E-05	C19H13N5	set4	12307	3.94E-09	C31H42O2N2O4
63	set4	12499	1.44E-04	C16H21N3OS	set7	240525	1.53E-09	C28H19O3N2O3	set9	7474	2.52E-05	C18H14Br2N4O2	set7	142922	3.95E-09	C31H26BrCl2N9OS
64	set2	355	1.47E-04	C14H11N	set7	158133	1.54E-09	C38H37N3O3	set7	182988	2.54E-05	C24H19NO2	set7	119269	3.96E-09	C28H28N2O4S
65	set6	15217	1.47E-04	C13H11NO5	set7	242497	1.55E-09	C31H27BrN2O2	set9	5827	2.55E-05	C21H22N6O	set3	7032	4.00E-09	C30H46O7
66	set6	13756	1.48E-04	C12H10ClNO5	set7	183828	1.61E-09	C34H30N2O5	set4	4676	2.59E-05	C19H23N3O2	set7	233517	4.14E-09	C25H18O3NO2
67	set1	121	1.52E-04	C17H14N2	set7	247805	1.61E-09	C29H25N3O2S	set2	2088	2.62E-05	C16H20N2O	set7	119258	4.16E-09	C25H23O3N2O2S
68	set6	14345	1.52E-04	C14H13NO3S	set7	183785	1.67E-09	C34H38N2O6	set6	14222	2.63E-05	C14H11ClN4OS	set8	27788	4.22E-09	C28H24N4O3S
69	set2	839	1.53E-04	C18H16O3	set7	242500	1.67E-09	C31H27O3N2O2	set5	184407	2.65E-05	C21H14O2	set7	114272	4.23E-09	C32H25N3O2S
70	set2	838	1.57E-04	C18H16O3	set7	147889	1.69E-09	C34H31F4N3OS	set6	14697	2.65E-05	C19H8N2O	set7	183374	4.27E-09	C25H25N3O4
71	set9	3594	1.59E-04	C19H15NO	set7	183754	1.69E-09	C34H32N2O4	set4	4294	2.67E-05	C15H9BrCl2N2S	set7	141302	4.32E-09	C33H32O3N3OS
72	set4	12308	1.60E-04	C19H31NO2	set3	1571	1.70E-09	C37H67NO13	set5	171282	2.68E-05	C19H13ClO3	set7	118796	4.40E-09	C25H19N3O4S
73	set6	2542	1.61E-04	C10H16	set7	159796	1.71E-09	C35H30N2O6	set8	29323	2.70E-05	C21H26N4OS	set8	1653	4.41E-09	C30H28N4OS
74	set6	15219	1.63E-04	C13H7NO5	set7	183861	1.74E-09	C35H32N2O4	set8	29301	2.71E-05	C20H24N4OS	set7	183782	4.52E-09	C30H30N2O4
75	set1	1307	1.64E-04	C14H12O4	set7	190141	1.74E-09	C38H32N2O2	set6	14956	2.73E-05	C17H13N5O	set5	154857	4.58E-09	C29H49I
76	set2	122	1.65E-04	C13H15NO5S	set3	8400	1.75E-09	C30H18O10	set4	8551	2.75E-05	C20H24O4	set3	6450	4.71E-09	C28H44O2

ตารางที่ 12 ผลของกลุ่มยาที่ได้จากการสกรีนด้วยโปรแกรม AutoDock ของกลุ่มโปรตีนเอนโดลอป (IOKE) type 1, 2, 3 และ 4 (ตารางต่อเนื่อง)

No.	เอนโดลอป ซีโรทัยป์ 1				เอนโดลอป ซีโรทัยป์ 2				เอนโดลอป ซีโรทัยป์ 3				เอนโดลอป ซีโรทัยป์ 4			
	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula
77	set1	1162	1.66E-04	C13H14N2O	set3	6366	1.78E-09	C27H44O3	set5	159747	2.76E-05	C24H22O3	set5	172561	4.72E-09	C24H28N4O4S2
78	set1	556	1.66E-04	C14H12BrN5	set7	159793	1.78E-09	C35H30N2O8	set4	3443	2.78E-05	C22H30O2S	set1	1090	4.78E-09	C27H36F2O5
79	set1	1505	1.66E-04	C15H14ClNO3	set7	141635	1.79E-09	C37H34FN3O2	set6	14954	2.78E-05	C18H15N5O	set8	23789	4.83E-09	C27H29ClN6O4
80	set6	15083	1.66E-04	C15H18N2O2S	set4	4433	1.81E-09	C45H73NO16	set7	182938	2.79E-05	C24H20N2O	set7	183829	4.85E-09	C34H29N3O6
81	set1	557	1.68E-04	C15H12O5	set3	8464	1.84E-09	C34H28O22	set9	6659	2.79E-05	C21H19NO3	set7	276617	4.87E-09	C27H29N3OS2
82	set6	15221	1.68E-04	C17H14N2O4	set7	292887	1.84E-09	C38H28N2O2	set4	5786	2.80E-05	C22H21N	set3	7035	4.89E-09	C30H44O8
83	set9	521	1.68E-04	C14H18N4S2	set7	241588	1.86E-09	C29H23ClN2O2	set1	67	2.83E-05	C21H21NO	set3	3826	4.93E-09	C29H48O
84	set6	14702	1.69E-04	C14H10N2O2	set7	159807	1.87E-09	C33H26N2O8	set1	2314	2.83E-05	C21H25N3O2	set5	156487	4.97E-09	C32H55NOS
85	set6	14697	1.71E-04	C19H8N2O	set8	26022	1.87E-09	C37H37N7O	set1	2233	2.84E-05	C13H18N2O2	set7	183525	4.97E-09	C32H28N2O4
86	set6	14217	1.73E-04	C14H12N4O2	set7	183826	1.91E-09	C34H29BrN2O4	set7	183010	2.84E-05	C24H21N3O	set7	183741	5.07E-09	C31H32N2O4
87	set1	1151	1.74E-04	C18H22O2	set7	142239	1.94E-09	C33H31ClN4O3S	set4	7639	2.86E-05	C21H16BrNO3	set8	17674	5.14E-09	C27H33N3OS
88	set2	560	1.74E-04	C14H11N	set3	9784	1.95E-09	C28H191N5O5	set9	6315	2.87E-05	C19H16Br2N4O	set5	174803	5.28E-09	C30H42O2
89	set6	14683	1.75E-04	C13H10N6	set7	183868	1.95E-09	C35H32N2O5	set8	26355	2.91E-05	C26H31N3OS	set7	119212	5.28E-09	C28H29FN2O2S
90	set2	1566	1.78E-04	C9H81N2S	set7	116567	1.96E-09	C31H29ClFN3O3S	set4	10311	2.92E-05	C21H22N6S	set7	278828	5.32E-09	C30H34N4O4S
91	set6	15216	1.79E-04	C12H9NO5	set7	183825	1.97E-09	C35H32N2O5	set8	22840	2.93E-05	C27H31N3O3S	set7	240522	5.32E-09	C28H17Cl2N3O5
92	set6	14115	1.79E-04	C13H10N2S	set7	117127	1.98E-09	C36H33N5O7S	set5	187864	2.98E-05	C20H16N2O3	set7	273243	5.45E-09	C28H23N3O2S
93	set9	4462	1.81E-04	C19H14ClNO	set7	183750	1.99E-09	C30H29BrN2O4	set1	611	3.07E-05	C19H18N2O2	set8	11282	5.60E-09	C24H21ClN2O3S
94	set1	1059	1.84E-04	C13H10Cl2O2	set7	183866	1.99E-09	C35H32N2O4	set5	162753	3.08E-05	C27H22N2O	set8	25346	5.62E-09	C28H26FN3O3S
95	set4	12501	1.84E-04	C18H23N3S	set7	107328	2.03E-09	C35H29N3OS	set5	161146	3.08E-05	C24H22OS	set7	264490	5.66E-09	C33H38N2O5S2
96	set1	387	1.86E-04	C12H81	set7	195499	2.03E-09	C34H26Cl2N2O4S2	set3	5783	3.11E-05	C20H12	set8	25223	5.67E-09	C27H31N3O3S
97	set6	14043	1.86E-04	C15H14N2O	set7	142435	2.04E-09	C33H30Cl2FN3OS	set5	185788	3.11E-05	C21H30O3	set7	183816	5.69E-09	C32H32N2O5
98	set6	14610	1.87E-04	C10H9N5O	set7	142350	2.05E-09	C34H33Cl2N3OS	set4	10149	3.14E-05	C20H18N4	set3	7030	5.71E-09	C30H44O7
99	set1	1960	1.88E-04	C16H16N2O2	set7	266147	2.13E-09	C35H29N3OS	set9	4462	3.14E-05	C19H14ClNO	set6	14149	5.74E-09	C29H22Cl2N4O6S
100	set1	1158	1.89E-04	C12H12N2O2S	set4	8334	2.14E-09	C35H26N4	set2	2386	3.15E-05	C19H17NO3	set2	843	5.78E-09	C28H28N4O3
101									set9	5334	3.15E-05	C15H12ClN5OS	set7	128463	5.78E-09	C29H34N2O3S

ตารางที่ 13 ผลของกลุ่มยาที่ได้จากการสกรีนด้วยโปรแกรม AutoDock ของกลุ่มโปรตีนเอ็นเอสหัวแม่ทูลิทรานส์ฟอเรส (1R6A) type 1, 2, 3 และ 4

No.	เอ็นเอสหัวแม่ทูลิทรานส์ฟอเรส ซีโรทัยป์ 1			เอ็นเอสหัวแม่ทูลิทรานส์ฟอเรส ซีโรทัยป์ 2			เอ็นเอสหัวแม่ทูลิทรานส์ฟอเรส ซีโรทัยป์ 3			เอ็นเอสหัวแม่ทูลิทรานส์ฟอเรส ซีโรทัยป์ 4		
	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula
1	set5	31570	7.36E-10	C47H78O14	set4	1471	6.75E-12	C61H72C2O10	set5	67516	8.27E-10	C56H42O13
2	set5	85763	8.00E-10	C34H26N6O12S4	set5	67516	2.52E-11	C56H42O13	set7	233486	1.59E-09	C34H28CIN3O8S
3	set5	67516	1.09E-09	C56H42O13	set5	85763	5.17E-11	C34H28N6O12S4	set7	142239	1.98E-09	C33H31CIN4O3S
4	set7	143077	1.28E-09	C33H29CIN3N4O3S	set4	38534	1.11E-10	C37H36N8O12S6	set7	142852	2.13E-09	C33H30CIN2N4O3S
5	set7	142239	1.72E-09	C33H31CIN4O3S	set4	432	1.69E-10	C56H44O13	set7	142590	2.66E-09	C34H33CIN4O3S
6	set7	142852	1.73E-09	C33H30CIN2N4O3S	set4	41982	2.52E-10	C59H72CIN2N2O10	set7	143077	2.71E-09	C33H29CIN3N4O3S
7	set3	9815	1.84E-09	C33H36N4O6	set2	3360	2.95E-10	C35H25N9O22S6	set7	158241	2.84E-09	C33H30CIN2N4O3S
8	set4	24114	1.97E-09	C49H52N4O6S2	set8	219453	2.99E-10	C35H26CIN3O6S	set7	141901	2.86E-09	C33H31FN4O3S
9	set7	142590	2.06E-09	C34H33CIN4O3S	set5	73862	3.82E-10	C37H36N2O10S3	set7	142238	3.07E-09	C33H31CIN4O3S
10	set4	502	2.13E-09	C48H28O30	set8	130972	4.31E-10	C42H32N2O6	set7	158276	3.44E-09	C33H29CIN3N4O3S
11	set8	287829	2.23E-09	C31H29N3O6S4	set6	3310	4.86E-10	C34H28N6O20S6	set8	219453	3.45E-09	C35H26CIN3O6S
12	set4	41228	2.53E-09	C27H27N6O15S3	set8	139026	7.49E-10	C27H20N4O8	set7	60553	3.61E-09	C41H32N4O2
13	set8	219453	2.76E-09	C35H26CIN3O6S	set2	34	8.04E-10	C16H16N8O19P3	set5	186260	3.79E-09	C30H38N4O8
14	set7	158221	2.81E-09	C32H28CIN2N4O3S	set5	28671	8.31E-10	C38H30N6O12	set7	158242	4.20E-09	C33H30CIN2N4O3S
15	set8	219454	2.83E-09	C35H26CIN3O6S	set8	267925	9.28E-10	C28H325N5O6S	set7	142848	4.27E-09	C33H30CIN2N4O3S
16	set5	124908	2.91E-09	C43H44N4O6S2	set5	100190	9.94E-10	C60H42N3O6P3	set8	110742	5.35E-09	C30H36CIN7O5S2
17	set2	1254	2.98E-09	C33H36N4O6	set5	75079	1.00E-09	C33H26N8O16S4	set7	158209	5.41E-09	C34H33CIN4O3S
18	set7	336430	3.25E-09	C31H24N8O4S2	set5	73856	1.01E-09	C30H20N4O11S3	set7	182945	5.56E-09	C54H44N2O4
19	set7	142238	3.30E-09	C33H31CIN4O3S	set4	4527	1.02E-09	C9H18N6O18S6	set3	4268	5.68E-09	C33H38N4O6
20	set5	18085	3.47E-09	C31H33N3O6S2	set2	1321	1.15E-09	C34H32FeN4O10	set4	21294	5.92E-09	C32H20N2O6S2
21	set7	158241	3.49E-09	C33H30CIN2N4O3S	set2	3544	1.17E-09	C42H42FeN4O16	set4	4527	5.95E-09	C9H18N6O18S6
22	set7	141901	3.53E-09	C33H31FN4O3S	set4	26760	1.21E-09	C26H10N8O12	set8	130183	6.28E-09	C35H22N2O8
23	set4	36393	3.67E-09	C36H29NO4	set5	143021	1.21E-09	C22H16N4O14S4	set7	142575	6.38E-09	C34H33CIN4O3S
24	set7	142848	3.72E-09	C33H30CIN2N4O3S	set7	187434	1.24E-09	C33H25N7O6S2	set4	41228	6.40E-09	C27H27N6O15S3

ตารางที่ 13 ผลของกลุ่มยาที่ได้จากการสกรีนด้วยโปรแกรม AutoDock ของกลุ่มโปรตีนเอ็นเอสพหามะเร็งที่ 1 (R6A) type 1, 2, 3 และ 4 (ตารางต่อเนื่อง)

No.	เอ็นเอสพหามะเร็งที่ 1				เอ็นเอสพหามะเร็งที่ 2				เอ็นเอสพหามะเร็งที่ 3				เอ็นเอสพหามะเร็งที่ 4			
	Ligand set	ID	Ki	Formula	Ligand set	ID	Ki	Formula	Ligand set	ID	Ki	Formula	Ligand set	ID	Ki	Formula
25	set7	158276	3.72E-09	C33H29Cl3N4O3S	set4	811	1.30E-09	C49H60Cl2O6	set2	1972	6.50E-09	C49H56FeN4O6	set7	233509	9.63E-09	C35H30Cl2N2O6S
26	set4	11	3.77E-09	C28H20N4O12S4	set5	32093	1.44E-09	C30H20N4O22S6	set8	219454	6.79E-09	C35H26ClN3O6S	set7	158209	1.09E-08	C34H33ClN4O3S
27	set3	4733	3.94E-09	C44H52N4O17	set5	122957	1.59E-09	C22H12N6O12	set7	141900	7.76E-09	C33H31FN4O3S	set7	266225	1.11E-08	C34H25ClN4O5S
28	set7	233524	3.96E-09	C31H21Cl5N2O4S	set2	1489	1.60E-09	C30H28N5O9P	set8	172286	8.09E-09	C33H24ClN5O2S2	set7	141900	1.12E-08	C33H31FN4O3S
29	set7	366	4.59E-09	C33H34N6O4	set8	194372	1.69E-09	C29H24N4O10	set5	68144	8.18E-09	C41H30O27	set7	142573	1.17E-08	C34H33ClN4O3S
30	set5	99010	4.72E-09	C34H26N6O13S4	set5	85251	1.73E-09	C28H20N4O13S4	set7	141826	8.31E-09	C34H34N4O3S	set7	233524	1.18E-08	C31H21Cl5N2O4S
31	set7	182740	4.77E-09	C46H44N2O4	set8	219454	1.74E-09	C35H26ClN3O6S	set7	182005	8.51E-09	C33H24ClN5O2S2	set8	110742	1.22E-08	C30H36ClN7O5S2
32	set7	158209	5.12E-09	C34H33ClN4O3S	set7	334902	1.88E-09	C30H23N5O5	set4	23149	8.86E-09	C24H17N5O9S	set4	21294	1.24E-08	C32H20N2O6S2
33	set7	158242	5.14E-09	C33H30Cl2N4O3S	set5	180899	1.91E-09	C36H54N4O4	set7	142573	9.20E-09	C34H33ClN4O3S	set5	75647	1.29E-08	C38H22N4O4
34	set4	18824	5.77E-09	C22H14Cl4N4O8	set8	151417	2.04E-09	C28H24N4O8	set7	336430	9.28E-09	C31H24N8O4S2	set4	23149	1.40E-08	C24H17N5O9S
35	set4	4527	5.79E-09	C9H18N6O18S6	set4	498	2.09E-09	C41H32O27	set8	130239	9.32E-09	C26H14Br2N2O8	set8	130268	1.40E-08	C33H22N2O10S2
36	set5	15595	5.88E-09	C33H34N4O6	set4	5873	2.12E-09	C34H32ClFeN4O4	set3	1860	9.38E-09	C42H62O16	set7	107144	1.43E-08	C33H22Cl2N2O24S
37	set2	783	5.97E-09	C33H34N4O6	set7	233524	2.24E-09	C31H21Cl5N2O4S	set8	198680	9.67E-09	C33H23Cl3N6O2S2	set7	142238	1.47E-08	C33H31ClN4O3S
38	set5	136626	6.00E-09	C35H34N4O12	set4	141	2.31E-09	C33H25N5O12S4	set5	33205	9.71E-09	C47H55N3O13	set7	141826	1.49E-08	C34H34N4O3S
39	set3	8444	6.41E-09	C37H30O17	set7	334912	2.34E-09	C29H20ClN5O5	set7	233499	1.02E-08	C33H26ClN3O8S	set8	172286	1.56E-08	C33H24ClN5O2S2
40	set5	49828	6.71E-09	C45H72O3	set8	248564	2.35E-09	C36H23Cl2N3O5	set3	7131	1.03E-08	C45H74O17	set5	186260	1.62E-08	C30H38N4O8
41	set2	3359	6.72E-09	C19H32N8O10S3	set5	182384	2.37E-09	C36H54N4O4	set7	158266	1.03E-08	C33H29ClN4O3S	set4	20634	1.63E-08	C25H17ClN8O8S
42	set5	75647	7.38E-09	C38H22N4O4	set8	130948	2.43E-09	C33H18N2O9	set5	136626	1.04E-08	C35H34N4O12	set5	157247	1.64E-08	C38H30N4
43	set3	458	7.46E-09	C33H34N4O6	set2	1974	2.44E-09	C34H34FeN4O4	set5	557	1.04E-08	C42H62O16	set7	117127	1.65E-08	C36H33N5O7S
44	set7	142573	7.59E-09	C34H33ClN4O3S	set7	233487	2.47E-09	C34H28Cl2N2O6S	set8	203835	1.04E-08	C22H14ClN3O6	set7	145385	1.68E-08	C26H23N5O9S
45	set7	142575	7.68E-09	C34H33ClN4O3S	set8	219914	2.51E-09	C24H18N6O6S2	set7	53701	1.05E-08	C43H36N4O2	set5	158259	1.69E-08	C40H20Br4
46	set5	122957	7.76E-09	C22H12N6O12	set5	113656	2.54E-09	C28H24N8O8	set7	142929	1.07E-08	C31H25Cl3N4O3S	set7	182005	1.71E-08	C33H24ClN5O2S2
47	set4	18823	7.80E-09	C22H12Cl2N4O8	set7	142587	2.61E-09	C34H33ClN4O3S	set7	117127	1.09E-08	C36H33N5O7S	set7	233444	1.71E-08	C28H22ClN3O5S
48	set7	141900	7.91E-09	C33H31FN4O3S	set5	99010	2.70E-09	C34H26N6O13S4	set7	80359	1.10E-08	C25H25N5O3S	set7	233487	1.72E-08	C34H28Cl2N2O6S
49	set7	141826	7.91E-09	C34H34N4O3S	set7	58476	2.73E-09	C33H21N5O2S2	set4	38147	1.16E-08	C45H74O17	set5	186261	1.76E-08	C32H42N4O8
50	set5	6206	7.98E-09	C30H38O8	set5	105127	2.81E-09	C28H18BrN3O12S	set8	144951	1.16E-08	C35H22I2O4	set5	176132	1.82E-08	C44H30O2

ตารางที่ 13 ผลของกลุ่มยาที่ได้จากการสกรีนด้วยโปรแกรม AutoDock ของกลุ่มโปรตีนเอ็นเอสพหามะเร็ง (1R6A) type 1, 2, 3 และ 4 (ตารางต่อเนื่อง)

NO.	เอ็นเอสพหามะเร็งชนิดที่ 1				เอ็นเอสพหามะเร็งชนิดที่ 2				เอ็นเอสพหามะเร็งชนิดที่ 3				เอ็นเอสพหามะเร็งชนิดที่ 4			
	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula
51	set5	143021	8.05E-09	C22H16N4O14S4	set8	130239	2.84E-09	C26H14B2N2O8	set7	277498	1.19E-08	C28H35N5O6S	set5	31570	1.84E-08	C47H78O14
52	set7	182945	8.14E-09	C54H44N2O4	set7	107144	2.86E-09	C33H22Cl2N2O4S	set8	110743	1.19E-08	C28H32ClN7O5S2	set4	20848	1.87E-08	C38H20Cl6O8S6
53	set7	142819	8.28E-09	C34H32Cl2N4O4	set8	160487	2.95E-09	C26H18N2O8	set4	27636	1.20E-08	C22H13Cl4N3O4S2	set5	119776	1.93E-08	C36H24N9O18P3
54	set5	175135	8.36E-09	C22H16N2O10S3	set7	113461	2.97E-09	C31H24N2O5	set7	142151	1.20E-08	C35H36N4O3S	set5	165513	1.94E-08	C27H21N3O13
55	set5	97208	8.45E-09	C28H20N2O8S2	set4	3008	2.98E-09	C27H19N3O8S2	set4	28419	1.25E-08	C26H26BrN5O6	set7	182743	1.97E-08	C43H46N2O4
56	set4	35889	8.79E-09	C39H54O5	set4	41708	3.00E-09	C59H72Cl2N2O10	set8	127095	1.27E-08	C45H54O4	set7	233499	1.98E-08	C33H26ClN3O8S
57	set3	1980	8.86E-09	C40H38N4O16	set7	148180	3.01E-09	C24H16N6O9	set7	107144	1.29E-08	C33H22Cl2N2O4S	set7	142916	2.07E-08	C37H33Cl2N3O5
58	set8	144951	9.35E-09	C35H22I2O4	set8	211485	3.01E-09	C33H25N7O6S2	set7	128814	1.32E-08	C29H26N2O2S2	set4	41934	2.12E-08	C47H40O2P2
59	set4	21319	9.52E-09	C28H24N10O8	set7	195839	3.04E-09	C24H16N6O9	set7	182743	1.34E-08	C43H46N2O4	set7	142610	2.12E-08	C35H35Cl2N3O5
60	set8	222627	9.56E-09	C36H24N6O4	set8	202685	3.08E-09	C28H19N3O3	set7	183599	1.35E-08	C30H32N2O4	set4	27636	2.13E-08	C22H13Cl4N3O4S2
61	set7	129183	9.72E-09	C22H12N2O8S2	set8	130976	3.10E-09	C28H14N4O9	set7	143051	1.35E-08	C33H30BrClN4O4	set7	142346	2.16E-08	C34H33Cl2N3O5
62	set7	334902	1.01E-08	C30H23N5O5	set1	241	3.18E-09	C44H68O13	set5	71454	1.36E-08	C20H14N2O10S3	set5	147713	2.20E-08	C27H20N4O4
63	set7	158253	1.03E-08	C31H25Cl3N4O3S	set8	130273	3.20E-09	C26H14B2N2O10S	set5	175135	1.37E-08	C22H16N2O10S3	set5	49828	2.22E-08	C45H72O3
64	set7	80359	1.03E-08	C25H25N5O3S	set8	179440	3.21E-09	C40H31ClN4O4S	set7	350671	1.37E-08	C27H18N4O9S2	set5	115301	2.22E-08	C30H33ClN3O10
65	set5	45256	1.04E-08	C30H26N4O6	set3	8453	3.23E-09	C41H30O27	set7	141822	1.39E-08	C34H34N4O3S	set7	128814	2.24E-08	C29H26N2O2S2
66	set7	142225	1.06E-08	C34H33ClN4O4	set7	233486	3.24E-09	C34H28ClN3O8S	set4	11	1.42E-08	C28H20N4O12S4	set7	142708	2.24E-08	C33H30Cl3N3O5
67	set7	142649	1.08E-08	C33H30Cl2N4O4	set8	211414	3.26E-09	C27H16N4O7S	set7	142819	1.45E-08	C34H32Cl2N4O4	set7	141881	2.26E-08	C35H36ClN3O5
68	set8	130178	1.08E-08	C37H26N2O8	set8	186934	3.29E-09	C36H24N6O4	set5	147713	1.48E-08	C27H20N4O4	set8	110743	2.28E-08	C28H32ClN7O5S2
69	set5	186260	1.09E-08	C30H38N4O8	set5	46117	3.31E-09	C34H36O8	set7	141634	1.48E-08	C31H27ClN4O3S	set7	143054	2.30E-08	C33H30BrClN4O4
70	set5	73862	1.09E-08	C37H36N2O10S3	set1	2166	3.33E-09	C17H12N2O12S3	set7	233524	1.49E-08	C31H21Cl5N2O4S	set7	158266	2.31E-08	C33H29Cl4N3O5
71	set7	142755	1.11E-08	C33H31BrN4O4	set4	41934	3.35E-09	C47H40O2P2	set7	142857	1.52E-08	C31H26BrClN4O4	set7	158134	2.34E-08	C38H37N3O5
72	set3	9454	1.13E-08	C26H43NO8S2	set5	41061	3.36E-09	C32H32N2O11S2	set7	142406	1.52E-08	C31H26Cl2N4O3S	set7	233486	2.35E-08	C34H28ClN3O8S
73	set7	233499	1.14E-08	C33H26ClN3O8S	set7	60789	3.39E-09	C19H11N7O10	set7	142708	1.55E-08	C33H30Cl3N3O5	set8	66528	2.35E-08	C29H21Br2N3O3
74	set7	334899	1.20E-08	C26H20N6O6	set8	167685	3.40E-09	C43H26N6O5	set7	141240	1.55E-08	C34H34FN3O5	set4	28419	2.39E-08	C26H26BrN5O6
75	set7	334912	1.21E-08	C29H20ClN5O5	set5	135968	3.44E-09	C25H30N6O8	set7	142008	1.57E-08	C34H33ClFN3O5	set7	142609	2.39E-08	C35H35Cl2N3O5
76	set8	169755	1.22E-08	C40H22N2O6	set8	217507	3.45E-09	C29H23ClN2O4	set5	49828	1.59E-08	C45H72O3	set7	158221	2.44E-08	C32H28Cl2N4O3S

ตารางที่ 13 ผลของกลุ่มยาที่ได้จากการสกัดด้วยโปรแกรม AutoDock ของกลุ่มโปรตีนเอนไซม์ฟอสเฟอเรส (1R6A) type 1, 2, 3 และ 4 (ตารางต่อเนื่อง)

NO.	เอ็นเอสห้ามกิลทราสส์เฟอเรส ซีโรทัยป์ 1			เอ็นเอสห้ามกิลทราสส์เฟอเรส ซีโรทัยป์ 2			เอ็นเอสห้ามกิลทราสส์เฟอเรส ซีโรทัยป์ 3			เอ็นเอสห้ามกิลทราสส์เฟอเรส ซีโรทัยป์ 4		
	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula
77	set7	71697	1.25E-08	C26H28N4O5	set5	90671	3.56E-09	C25H20N6O12S2	set8	211465	1.59E-08	C33H25N7O6S2
78	set7	117890	1.27E-08	C23H14N4O9S	set8	200446	3.56E-09	C39H28N2	set5	48515	1.64E-08	C42H42N4
79	set5	165513	1.30E-08	C27H21N3O13	set2	1973	3.57E-09	C34H34FeN4O4	set7	141989	1.65E-08	C32H29CIN4O3S
80	set1	2162	1.31E-08	C42H70O11	set8	167614	3.60E-09	C34H20N2O8	set7	142225	1.67E-08	C34H33CIN4O4
81	set7	141906	1.31E-08	C33H31CIN4O4	set5	119776	3.63E-09	C36H24N9O18P3	set7	233444	1.69E-08	C28H22CIN3O5S
82	set7	119117	1.33E-08	C35H36N2O3	set1	2330	3.65E-09	C51H62O17	set7	143066	1.69E-08	C34H32CIN4O4S
83	set8	125389	1.33E-08	C24H14N4O10S3	set5	136626	3.68E-09	C35H34N4O12	set7	142609	1.69E-08	C35H35CIN3O3S
84	set3	4272	1.34E-08	C33H42N4O6	set2	717	3.70E-09	C40H56	set2	1254	1.70E-08	C33H36N4O6
85	set5	71454	1.34E-08	C20H14N2O10S3	set3	6781	3.73E-09	C34H34N2O16	set7	117879	1.71E-08	C23H14N4O9S
86	set4	17507	1.36E-08	C33H38N4O6	set5	21934	3.73E-09	C33H50O5S	set4	18824	1.81E-08	C22H14CIN4N4O8
87	set7	128814	1.36E-08	C29H26N2O2S2	set8	224370	3.81E-09	C24H16CIN3O4S	set8	130178	1.81E-08	C37H26N2O8
88	set7	142759	1.37E-08	C33H31BnN4O4	set8	129158	3.89E-09	C29H16CIN4N4O12	set8	167614	1.82E-08	C34H20N2O8
89	set4	38147	1.38E-08	C45H74O17	set7	111677	3.90E-09	C34H26CIN5O5S	set7	141872	1.83E-08	C35H36CIN3O3S
90	set8	135902	1.41E-08	C26H22N2O8	set8	171211	3.91E-09	C26H15B2N5O5	set5	122957	1.86E-08	C22H12N6O12
91	set7	182743	1.43E-08	C43H46N2O4	set8	260069	3.92E-09	C22H15CIN6O5S	set7	158221	1.86E-08	C32H28CIN2N4O3S
92	set8	187015	1.46E-08	C34H18N4O10	set7	233522	3.97E-09	C40H34CIN3O6S	set6	14315	1.87E-08	C28H40CIN3O2
93	set8	198680	1.46E-08	C33H23CIN6O2S2	set7	278629	3.98E-09	C34H26CIN5O5S	set5	9322	1.88E-08	C42H62O16
94	set5	75642	1.48E-08	C32H18N4O4	set4	23098	3.99E-09	C38H28N10O7	set5	24267	1.88E-08	C39H59N3O9
95	set3	7131	1.49E-08	C45H74O17	set5	165513	4.01E-09	C27H21N3O13	set7	125263	1.91E-08	C28H21CIN4O5S
96	set3	9577	1.51E-08	C27H24N4O6	set7	234999	4.02E-09	C33H26CIN3O8S	set7	107385	1.92E-08	C34H25CIN4O5S
97	set7	142647	1.51E-08	C33H30CIN2N4O4	set4	2145	4.03E-09	C24H17N3O7S3	set7	142411	1.92E-08	C31H26CIN2N4O3S
98	set7	145385	1.52E-08	C26H23N5O9S	set7	235503	4.13E-09	C33H28BnCIN2O6S	set4	36393	1.95E-08	C36H29N4O4
99	set8	130183	1.52E-08	C35H22N2O8	set3	5047	4.15E-09	C34H32CIN4O4	set7	158238	2.00E-08	C34H32CIN2N4O4
100	set5	9322	1.54E-08	C42H62O16	set8	152328	4.15E-09	C29H25N3O3	set7	58229	2.01E-08	C20H11N7O12
101	set5	26395	1.54E-08	C37H42O3	set8	173743	4.15E-09	C34H20N6O4				
102	set7	117879	1.54E-08	C23H14N4O9S								
103	set8	167485	1.54E-08	C34H22CIN2O4								

ตารางที่ 14 ผลของกลุ่มยาที่ได้จากการสกรีนด้วยโปรแกรม AutoDock ของกลุ่มโปรตีนเฮลิคเลส (2BMF) type 1, 2, 3 และ 4

No.1	เฮลิคเลส ซีโรทัยป์ 1				เฮลิคเลส ซีโรทัยป์ 2				เฮลิคเลส ซีโรทัยป์ 3				เฮลิคเลส ซีโรทัยป์ 4			
	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula
1	set7	182945	1.04E-12	C54H44N2O4	set4	9617	1.34E-12	C70H90N12O17	set5	125938	1.21E-12	C65H58N2O3S2	set4	3468	1.12E-12	C55H64N12P4
2	set7	55201	1.27E-12	C44H28Br2N6O2	set5	100190	1.38E-12	C60H42N3O6P3	set5	43551	3.55E-12	C56H82	set5	49346	1.38E-12	C60H80N4O6
3	set5	21936	1.91E-12	C58H82O8	set3	8539	2.75E-12	C60H56O11	set4	11262	5.99E-12	C50H44N2O6	set4	17887	1.40E-12	C60H88O4
4	set3	8539	2.08E-12	C60H56O11	set6	7414	3.15E-12	C51H84O22	set7	182945	1.02E-11	C54H44N2O4	set7	182945	2.58E-12	C54H44N2O4
5	set4	18996	2.23E-12	C50H62O18	set4	33212	3.37E-12	C63H55NO6	set5	180022	1.25E-11	C64H55BrO5	set8	170383	2.74E-12	C47H40N4O5S2
6	set5	165643	3.25E-12	C50H30N4O8	set7	182945	3.38E-12	C54H44N2O4	set4	37808	1.75E-11	C54H82N2O3	set4	37808	2.82E-12	C54H82N2O3
7	set6	11961	4.22E-12	C51H82O20	set4	7617	3.78E-12	C56H79N3O12	set8	170382	1.83E-11	C47H42N4O4S2	set5	32685	3.36E-12	C43H59N4O13
8	set5	43551	4.52E-12	C56H82	set6	11438	3.90E-12	C64H78N10O10	set4	11263	1.85E-11	C54H48N2O6	set4	7424	5.44E-12	C52H68N4O12
9	set7	57147	5.14E-12	C46H32Cl2N6O2	set4	41934	5.82E-12	C47H40O2P2	set4	41934	2.47E-11	C47H40O2P2	set8	170382	5.60E-12	C47H42N4O4S2
10	set7	182979	5.41E-12	C54H44N2O4	set8	170383	6.33E-12	C47H40N4O5S2	set3	8434	2.56E-11	C41H28O26	set4	6997	6.15E-12	C45H53N5O16
11	set8	130927	5.75E-12	C59H32N2O8	set4	25216	7.09E-12	C80H50N8O8	set5	100190	2.67E-11	C60H42N3O6P3	set5	19466	8.52E-12	C45H62O2
12	set8	130922	5.79E-12	C52H24N2O14	set4	24114	8.11E-12	C49H52N4O6S2	set8	170383	3.29E-11	C47H40N4O5S2	set5	100190	9.20E-12	C60H42N3O6P3
13	set4	23884	5.94E-12	C59H58O18	set4	17981	8.19E-12	C56H40O12	set5	33202	3.56E-11	C47H66N2O13	set5	123478	9.30E-12	C42H36Br2N2P2
14	set7	296849	6.77E-12	C53H40N2O2	set4	40665	1.02E-11	C71H54N6O15S3	set8	130922	3.75E-11	C52H24N2O14	set4	17442	9.99E-12	C50H48N6O4
15	set7	296791	6.92E-12	C47H32N2	set3	7137	1.16E-11	C57H96O28	set4	17888	4.04E-11	C75H110O5	set5	49828	1.11E-11	C45H72O3
16	set7	182981	6.97E-12	C51H46N2O4	set5	43551	1.17E-11	C56H82	set7	182981	4.31E-11	C51H46N2O4	set1	2330	1.29E-11	C51H62O17
17	set4	28198	7.88E-12	C43H32O14S	set8	130969	1.40E-11	C54H30N4O12	set3	7138	4.66E-11	C45H74O17	set5	143538	1.34E-11	C45H42FeN3S6
18	set5	155035	8.62E-12	C61H54O8	set7	182979	1.42E-11	C54H44N2O4	set3	9014	5.32E-11	C45H73NO14	set4	20924	1.40E-11	C60H104
19	set6	7414	8.84E-12	C51H84O22	set5	54763	1.49E-11	C41H30O27	set4	41138	5.51E-11	C48H42N4O4S2	set7	182740	1.41E-11	C46H44N2O4
20	set8	170383	1.05E-11	C47H40N4O5S2	set4	7398	1.55E-11	C49H62N8O16	set7	182949	5.60E-11	C51H46N2O4	set5	29197	1.70E-11	C40H50O4
21	set2	259	1.26E-11	C52H76N2O13	set4	22458	1.60E-11	C52H83NO22	set3	9669	5.67E-11	C55H74N3O21S4	set7	182843	1.71E-11	C42H36N2O4
22	set7	56751	1.41E-11	C48H38N6O2	set4	6437	1.67E-11	C53H62N2O13	set4	1342	6.17E-11	C50H66N4O4	set5	32973	1.85E-11	C44H59N3O13
23	set4	38635	1.42E-11	C49H48N6O4	set4	7035	1.70E-11	C50H71N3O12	set8	197035	6.33E-11	C47H40N4O5S2	set5	9431	2.03E-11	C45H73NO15
24	set4	27686	1.43E-11	C59H68Cl2O8	set3	8453	1.83E-11	C41H30O27	set4	24114	6.34E-11	C49H52N4O6S2	set3	7121	2.12E-11	C51H84O23

ตารางที่ 14 ผลของกลุ่มยาที่ได้จากการสกรีนด้วยโปรแกรม AutoDock ของกลุ่มโปรตีนเฮลิคเลส (2BMF) type 1, 2, 3 และ 4 (ตารางต่อเนื่อง)

No.	เฮลิคเลส ซีโรทัยป์ 1			เฮลิคเลส ซีโรทัยป์ 2			เฮลิคเลส ซีโรทัยป์ 3			เฮลิคเลส ซีโรทัยป์ 4		
	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula
25	set8	167448	1.48E-11	C47H32N2O2	set8	170382	2.00E-11	C47H42N4O4S2	set5	9861	6.55E-11	C60H88N2O8S
26	set4	37132	1.51E-11	C59H72Cl2O6	set4	38984	2.05E-11	C72H100N4O8	set4	28517	6.69E-11	C51H36N2O5S
27	set8	161279	1.52E-11	C48H38N4O6	set3	6271	2.13E-11	C53H67N9O10S	set4	6535	7.64E-11	C53H77N3O12
28	set7	296794	1.55E-11	C53H40N2O2	set4	42359	2.70E-11	C58H94O26	set6	7414	7.65E-11	C51H84O22
29	set4	40454	1.57E-11	C54H46N4O4	set4	501	3.01E-11	C48H28O30	set4	4178	7.80E-11	C38H58S
30	set4	7423	1.61E-11	C49H61N7O16	set4	7423	3.17E-11	C49H61N7O16	set8	130930	7.87E-11	C50H24N2O14
31	set4	41933	1.62E-11	C47H40O2P2	set7	296848	3.21E-11	C49H36N2O4	set4	40053	8.25E-11	C55H48N4O4
32	set5	80548	1.63E-11	C62H30N2O6	set4	10778	3.48E-11	C39H63NO11	set3	9045	8.37E-11	C50H83NO21
33	set4	24114	1.65E-11	C49H52N4O6S2	set8	197035	3.58E-11	C47H40N4O5S2	set4	27686	8.54E-11	C59H68Cl2O8
34	set4	41715	1.81E-11	C52H86O22	set4	7402	3.61E-11	C53H77N3O12	set5	12824	9.45E-11	C48H35Cl7O7
35	set4	40665	1.81E-11	C71H54N6O15S3	set8	127095	3.80E-11	C45H54O4	set5	80548	9.80E-11	C62H30N2O6
36	set4	20924	1.87E-11	C60H104	set4	37132	3.86E-11	C59H72Cl2O6	set5	110510	1.02E-10	C48H42N2O8S
37	set3	6271	1.94E-11	C53H67N9O10S	set8	167448	4.26E-11	C47H32N2O2	set1	2330	1.03E-10	C51H62O17
38	set8	197035	1.96E-11	C47H40N4O5S2	set4	41139	4.44E-11	C56H46N4O4	set6	8445	1.12E-10	C48H28Br6O6S4
39	set4	28193	1.97E-11	C61H72Cl2O8	set4	7973	4.49E-11	C51H66N4O13	set6	11438	1.18E-10	C64H78N10O10
40	set4	11262	1.98E-11	C50H44N2O6	set3	7149	4.59E-11	C48H78O18	set7	182740	1.18E-10	C48H44N2O4
41	set3	9038	2.29E-11	C45H73NO15	set4	40454	4.70E-11	C54H46N4O4	set4	17440	1.24E-10	C48H44N6O4
42	set5	158252	2.37E-11	C70H54O4	set4	38635	4.88E-11	C49H48N6O4	set5	143538	1.29E-10	C45H42FeN3S6
43	set4	41934	2.60E-11	C47H40O2P2	set4	7424	5.14E-11	C52H68N4O12	set7	182993	1.29E-10	C54H52N2O4
44	set4	7731	2.62E-11	C64H69N3O19	set4	37093	5.15E-11	C57H84O19	set4	41236	1.33E-10	C45H74O18
45	set3	7121	2.70E-11	C51H84O23	set5	80548	5.19E-11	C62H30N2O6	set7	393	1.33E-10	C43H48N4O4
46	set7	193358	2.85E-11	C49H34Br2N2O4	set7	187948	5.22E-11	C46H34N8O6S2	set4	7035	1.34E-10	C50H71N3O12
47	set5	68144	2.91E-11	C41H30O27	set3	7459	5.34E-11	C55H62N10	set4	28193	1.36E-10	C61H72Cl2O8
48	set7	296793	2.92E-11	C49H36N2O2	set8	167440	5.48E-11	C42H28N6O2	set8	127584	1.36E-10	C48H48N4O2
49	set7	187949	2.94E-11	C43H28Cl2N6O3S2	set5	49346	5.61E-11	C60H80N4O6	set3	7170	1.38E-10	C48H82O18
50	set5	9431	3.02E-11	C45H73NO15	set4	11262	5.62E-11	C50H44N2O6	set5	10839	1.43E-10	C44H42Br2N2O6

ตารางที่ 14 ผลของกลุ่มยาที่ได้จากการสกรีนด้วยโปรแกรม AutoDock ของกลุ่มโปรตีนเฮลิคเลส (2BMF) type 1, 2, 3 และ 4 (ตารางต่อเนื่อง)

No.	เฮลิคเลส ซีโรทัยป์ 1				เฮลิคเลส ซีโรทัยป์ 2				เฮลิคเลส ซีโรทัยป์ 3				เฮลิคเลส ซีโรทัยป์ 4			
	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula
51	set4	6535	3.04E-11	C53H77N3O12	set5	16785	5.79E-11	C41H30O27	set6	11961	1.45E-10	C51H82O20	set3	9014	5.82E-11	C45H73NO14
52	set4	37808	3.07E-11	C54H82N2O3	set5	9861	5.92E-11	C60H88N2O8S	set7	152146	1.46E-10	C41H33N5O2	set5	32718	5.95E-11	C45H62N2O12
53	set3	7143	3.12E-11	C51H84O22	set4	271	6.14E-11	C41H30O27	set4	41933	1.56E-10	C47H40O2P2	set5	61203	6.07E-11	C45H73NO15
54	set4	7398	3.16E-11	C49H62N8O16	set4	40053	6.26E-11	C55H44N4O4	set8	169744	1.62E-10	C48H50O6P2	set4	7946	6.19E-11	C51H66N4O13
55	set4	6437	3.49E-11	C53H62N2O13	set6	8445	6.29E-11	C48H28B6O6S4	set7	296858	1.66E-10	C49H34Br2N2O4	set7	296858	6.40E-11	C49H34Br2N2O4
56	set7	152029	3.72E-11	C46H34N6O2	set4	28198	6.30E-11	C43H32O14S	set5	146619	1.69E-10	C54H52N2P4	set5	176132	6.88E-11	C44H30O2
57	set4	27099	3.90E-11	C62H40N12O4S4	set4	499	6.48E-11	C41H30O27	set5	56348	1.71E-10	C39H36O9	set5	38858	7.06E-11	C45H78O13
58	set4	7402	3.92E-11	C53H77N3O12	set3	7126	6.54E-11	C50H82O23	set3	8193	1.89E-10	C44H34O22	set4	17452	7.22E-11	C52H52N6O8
59	set5	68403	3.92E-11	C48H40N4O11	set3	7131	6.55E-11	C45H74O17	set5	31770	1.84E-10	C51H66N4O12	set4	8075	7.31E-11	C44H60N4O12
60	set4	7000	4.26E-11	C53H61N3O14	set8	130930	6.74E-11	C50H24N2O14	set4	38147	1.88E-10	C45H74O17	set1	1355	7.35E-11	C42H69O11
61	set7	393	4.30E-11	C43H48N4O4	set2	105	7.08E-11	C49H56N6O2	set5	45535	1.91E-10	C53H80O10	set8	241115	7.38E-11	C46H58O8
62	set5	119898	4.45E-11	C40H22S2	set4	42596	7.25E-11	C54H36Cl8N8O6S4	set3	8539	1.96E-10	C60H56O11	set4	28198	7.42E-11	C43H32O14S
63	set7	182949	4.57E-11	C51H46N2O4	set8	162738	7.36E-11	C46H42N4O8	set8	130972	2.00E-10	C42H32N2O6	set7	185161	7.52E-11	C44H50O6
64	set4	7617	4.73E-11	C56H79N3O12	set4	8004	7.70E-11	C53H69N5O13	set5	162551	2.03E-10	C42H26N4O3	set3	9037	7.94E-11	C45H73NO15
65	set3	9021	4.81E-11	C50H83NO20	set8	130955	8.07E-11	C40H16Br8N2O6	set5	49828	2.07E-10	C45H72O3	set8	126912	8.31E-11	C55H58O4
66	set7	296865	4.91E-11	C45H36Cl4N2O4	set4	28193	8.28E-11	C61H72Cl2O8	set8	135868	2.38E-10	C51H56O4	set7	182743	8.36E-11	C43H46N2O4
67	set5	50915	4.98E-11	C51H75N3O3	set5	49349	9.03E-11	C58H70N4O6	set3	7182	2.39E-10	C53H86O21	set5	9576	8.51E-11	C48H78O19
68	set7	305745	5.13E-11	C43H32Br4N2O4	set3	9427	9.04E-11	C60H92N12O10	set4	40054	2.46E-10	C56H50N4O4	set6	10954	8.59E-11	C50H83NO21
69	set8	127584	5.28E-11	C48H48N4O2	set3	8193	9.26E-11	C44H34O22	set4	17452	2.51E-10	C52H52N6O8	set5	146619	8.68E-11	C54H52N2P4
70	set7	296858	5.37E-11	C49H34Br2N2O4	set8	197036	9.31E-11	C51H50N6O7S2	set6	7416	2.59E-10	C51H82O23	set8	124153	8.69E-11	C49H46N2O2
71	set7	57881	5.49E-11	C48H38N6O2	set4	40054	9.32E-11	C56H50N4O4	set7	57881	2.66E-10	C48H38N6O2	set3	9042	8.75E-11	C45H73NO16
72	set6	10954	5.61E-11	C50H83NO21	set3	5213	9.36E-11	C46H64N14O12S2	set8	130984	2.67E-10	C50H40N2O6	set7	193400	8.83E-11	C39H30N4S2
73	set7	305752	5.80E-11	C45H30Cl2N2O4S2	set4	17888	9.49E-11	C75H110O5	set4	499	2.75E-10	C41H30O27	set5	68440	9.79E-11	C49H46N4O6
74	set5	180022	5.84E-11	C64H55BrO5	set5	12824	9.85E-11	C48H35Cl7O7	set5	2955	2.81E-10	C42H52N2O6	set3	7159	1.01E-10	C55H86O25
75	set4	40375	5.96E-11	C52H38N4O6S	set3	8435	1.00E-10	C41H30O27	set7	182952	2.89E-10	C54H52N2O4	set4	10778	1.01E-10	C39H63NO11
76	set7	266293	6.12E-11	C38H27Cl2N3O5S	set7	182993	1.01E-10	C54H52N2O4	set4	40452	2.85E-10	C56H50N4O4	set5	44889	1.01E-10	C41H37Cl2FN4O11S

ตารางที่ 14 ผลของกลุ่มยาที่ได้จากการสกรีนด้วยโปรแกรม AutoDock ของกลุ่มโปรตีนเฮลิคเลส (2BMF) type 1, 2, 3 และ 4 (ตารางต่อเนื่อง)

No.	เฮลิคเลส ซีโรทัยป์ 1				เฮลิคเลส ซีโรทัยป์ 2				เฮลิคเลส ซีโรทัยป์ 3				เฮลิคเลส ซีโรทัยป์ 4			
	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula
77	set7	296847	6.18E-11	C43H26Cl2N2O2S2	set7	55734	1.04E-10	C45H40Cl4N4O2	set8	169755	2.89E-10	C40H22N2O6	set4	41934	1.02E-10	C47H40O2P2
78	set5	9861	6.33E-11	C60H88N2O8S	set6	7524	1.06E-10	C62H49NO15	set3	8435	2.90E-10	C41H30O27	set5	49348	1.04E-10	C58H64N4O6
79	set3	7138	6.41E-11	C45H74O17	set4	9106	1.09E-10	C41H36N4O8S4	set4	41139	2.92E-10	C56H46N4O4	set2	3346	1.05E-10	C43H45ClN3O13
80	set7	58676	6.65E-11	C45H30Br2N4O2	set7	182749	1.10E-10	C46H52N2O4	set4	7402	2.93E-10	C53H77N3O12	set7	296829	1.06E-10	C39H30N4S2
81	set4	41714	6.84E-11	C52H86O22	set3	7138	1.14E-10	C45H74O17	set5	16785	2.93E-10	C41H30O27	set3	7188	1.10E-10	C48H76O18
82	set7	60654	7.08E-11	C46H34N6O2	set5	50159	1.14E-10	C41H26O26	set5	68039	2.94E-10	C46H75NO14	set3	7025	1.12E-10	C48H78O18
83	set5	85764	7.09E-11	C42H70N6O8S4	set5	50186	1.15E-10	C41H26O26	set5	124308	3.02E-10	C48H34O2	set4	4861	1.12E-10	C45H73NO15
84	set7	193576	7.09E-11	C43H26Cl2N2O2S2	set4	17452	1.16E-10	C52H52N6O8	set8	255994	3.02E-10	C34H26ClN5OS	set8	136814	1.12E-10	C47H53NO4
85	set5	4664	7.18E-11	C40H42N4O10	set5	64692	1.17E-10	C48H64N2O17	set7	193358	3.03E-10	C49H34Br2N2O4	set4	7438	1.15E-10	C51H65ClN4O12
86	set7	193676	7.19E-11	C46H30O6	set8	130905	1.19E-10	C52H32N2O8	set6	7524	3.06E-10	C62H49NO15	set7	182952	1.15E-10	C54H52N2O4
87	set8	130946	7.28E-11	C45H20N2O13	set5	166362	1.20E-10	C44H24N4O8	set4	271	3.07E-10	C41H30O27	set7	182979	1.15E-10	C54H44N2O4
88	set4	38147	7.40E-11	C45H74O17	set8	169755	1.21E-10	C40H22N2O6	set4	41716	3.08E-10	C52H86O23	set7	182949	1.15E-10	C51H46N2O4
89	set4	40054	7.43E-11	C56H50N4O4	set7	182740	1.22E-10	C46H44N2O4	set5	39115	3.08E-10	C45H73NO15	set8	151475	1.15E-10	C46H31N3O6
90	set7	305757	7.67E-11	C49H32Br4N2O4	set4	40670	1.25E-10	C59H44N12O9	set5	55118	3.09E-10	C46H76O17	set5	50915	1.18E-10	C51H75N3O3
91	set1	2330	7.91E-11	C51H62O17	set4	6484	1.26E-10	C44H51N5O16	set5	119898	3.10E-10	C40H22S2	set4	36522	1.19E-10	C43H37Br3O10
92	set3	7131	8.06E-11	C45H74O17	set5	165643	1.29E-10	C50H30N4O8	set4	4144	3.17E-10	C42H38Cl2N2O6	set8	196680	1.20E-10	C33H23Cl3N6O2S2
93	set5	2955	8.28E-11	C42H52N2O6	set8	151444	1.32E-10	C41H24N4O10	set7	182979	3.18E-10	C54H44N2O4	set8	130972	1.28E-10	C42H32N2O6
94	set7	353484	8.58E-11	C40H31ClN4O2S	set8	637	1.33E-10	C46H34N4S2	set4	37914	3.20E-10	C50H32N2O	set2	51	1.29E-10	C49H44N8O5
95	set7	60249	8.77E-11	C44H28Br2N4O3	set5	33211	1.35E-10	C48H55N5O13	set4	41234	3.28E-10	C50H82O23	set4	38147	1.29E-10	C45H74O17
96	set8	167688	9.15E-11	C43H26I2N4O	set5	155270	1.35E-10	C44H32O2	set5	158259	3.34E-10	C40H20Br4	set5	24266	1.29E-10	C40H61NO6
97	set7	296869	9.18E-11	C43H34Br2N2O4	set5	25760	1.35E-10	C42H38O20	set3	7145	3.36E-10	C39H62O12	set4	41711	1.30E-10	C45H72O17
98	set4	40670	9.20E-11	C59H44N12O9	set6	11835	1.35E-10	C28Cl2O54	set3	7120	3.52E-10	C45H72O17	set3	7211	1.33E-10	C42H70O14
99	set7	194158	9.23E-11	C55H36N2	set4	1342	1.38E-10	C50H66N4O4	set8	26136	3.54E-10	C35H39N7O2	set3	7186	1.35E-10	C51H82O21
100	set7	59450	9.23E-11	C44H30N6O2	set8	130922	1.38E-10	C52H24N2O14	set4	7398	3.59E-10	C49H62N8O16	set4	37090	1.35E-10	C39H64O13

ตารางที่ 15 ผลของกลุ่มยาที่ได้จากการสกรีนด้วยโปรแกรม AutoDock ของกลุ่มโปรตีนเอ็นเอสสามโปรตีนเอส (2FOM) type 1, 2, 3 และ 4

No.	เอ็นเอสสามโปรตีนเอสที่ร้ายป์ 1				เอ็นเอสสามโปรตีนเอสที่ร้ายป์ 2				เอ็นเอสสามโปรตีนเอสที่ร้ายป์ 3				เอ็นเอสสามโปรตีนเอสที่ร้ายป์ 4			
	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula
1	set1	907	1.91E-10	C37H67NO13	set3	1938	6.12E-11	C44H73NO17	set1	907	5.16E-12	C37H67NO13	set8	182134	3.47E-09	C32H25Br2Cl2NO2
2	set2	1563	2.44E-09	C37H67NO13	set5	67516	8.66E-11	C56H42O13	set3	1571	8.40E-12	C37H67NO13	set8	127095	1.39E-08	C45H54O4
3	set8	130984	1.20E-08	C50H40N2O6	set4	7424	1.29E-10	C52H68N4O12	set8	130972	7.91E-10	C42H32N2O6	set5	86091	3.95E-08	C34H16Cl2N2O6
4	set5	54092	1.92E-08	C46H50N6O2	set8	170383	7.11E-10	C47H40N4O5S2	set5	123478	2.23E-09	C42H36Br2N2P2	set5	176323	4.88E-08	C36H18N4O4S
5	set8	172286	1.99E-08	C33H24ClN5O2S2	set5	43551	7.17E-10	C56H82	set8	162763	3.49E-09	C38H28O2	set8	193355	5.76E-08	C25H28N8O4
6	set8	130972	2.11E-08	C42H32N2O6	set4	7399	1.15E-09	C48H58N6O17	set5	26395	6.45E-09	C37H42O3	set5	176132	5.93E-08	C44H30O2
7	set8	136814	2.15E-08	C47H53NO4	set3	8193	1.26E-09	C44H34O22	set8	55598	1.00E-08	C31H32ClN7O	set8	36501	6.05E-08	C31H31N7O3
8	set5	177945	2.22E-08	C35H25N3O3	set8	128587	2.40E-09	C35H24N8	set8	169755	1.26E-08	C40H22N2O6	set4	11252	6.39E-08	C27H30O16
9	set8	169755	2.28E-08	C40H22N2O6	set5	156284	2.71E-09	C64H52O8	set8	26136	1.30E-08	C35H39N7O2	set8	36503	7.89E-08	C29H33N7O3
10	set8	126360	3.17E-08	C37H28O4	set5	21936	3.38E-09	C58H82O8	set8	63882	1.37E-08	C32H35N7O2	set8	217526	8.58E-08	C33H22Cl2N2O4S
11	set8	314669	3.71E-08	C28H30N6O2S	set2	105	3.49E-09	C49H56N6O2	set3	8320	1.54E-08	C40H36O11	set3	7211	8.61E-08	C42H70O14
12	set8	189031	3.76E-08	C29H19Cl3N2O2	set3	8453	3.53E-09	C41H30O27	set5	83177	1.54E-08	C46H60Br2O6	set5	92231	9.64E-08	C36H33NO3
13	set5	4664	3.88E-08	C40H42N4O10	set5	124233	3.78E-09	C53H48N2O3S	set8	26022	1.54E-08	C37H37N7O	set2	1191	1.03E-07	C33H40N4O6
14	set5	163483	4.01E-08	C40H26O2	set8	222615	3.87E-09	C40H24N4O	set8	209422	1.67E-08	C48H45NO3	set8	176217	1.06E-07	C30H31Br2N7O
15	set5	82803	4.18E-08	C44H34	set1	275	4.40E-09	C51H79NO13	set8	35672	1.71E-08	C30H30ClN7O	set8	173197	1.08E-07	C34H27BrN4O4
16	set8	145477	4.35E-08	C40H26O2	set8	167683	4.79E-09	C43H28N4O	set8	24405	1.75E-08	C35H39N7O2	set5	179331	1.17E-07	C37H34N2O
17	set8	191353	4.40E-08	C33H22Cl2N2O2	set8	198679	4.85E-09	C33H24Cl2N6O2S2	set3	8903	1.88E-08	C37H34N2O6	set3	8058	1.29E-07	C40H36O12
18	set8	45443	4.94E-08	C31H27FN6O2	set5	56543	4.90E-09	C41H28O27	set5	179331	1.97E-08	C37H34N2O	set3	8057	1.36E-07	C40H36O12
19	set5	162384	4.96E-08	C43H42N4	set1	1762	4.98E-09	C51H79NO13	set8	126879	2.00E-08	C35H22Cl2O4	set8	179439	1.40E-07	C38H29BrN4O4S
20	set4	17980	5.02E-08	C42H32O9	set8	144951	5.08E-09	C35H22I2O4	set8	135868	2.13E-08	C42H50O4	set2	1492	1.43E-07	C30H46N4O2
21	set5	183330	5.42E-08	C30H22O4	set8	222622	5.25E-09	C43H28N4O3	set4	9469	2.18E-08	C37H44P2	set2	863	1.45E-07	C32H37N3O3S
22	set3	9038	5.45E-08	C45H73NO15	set4	7973	5.77E-09	C51H66N4O13	set5	31185	2.19E-08	C36H41N3O3	set8	36523	1.46E-07	C31H31N7O3
23	set8	182130	5.49E-08	C32H26BrCl2NO2	set4	10778	6.40E-09	C39H63NO11	set8	172286	2.35E-08	C33H24ClN5O2S2	set8	163389	1.49E-07	C34H29ClN4O3
24	set8	149672	5.57E-08	C31H32N4O2	set4	17887	6.52E-09	C60H88O4	set8	255994	2.60E-08	C34H26ClN5OS	set8	36000	1.53E-07	C31H39N7O

ตารางที่ 15 ผลของกลุ่มยาที่ได้จากการสกรีนด้วยโปรแกรม AutoDock ของกลุ่มโปรตีนเอ็นเอสสามโปรตีนเอส (2FOM) type 1, 2, 3 และ 4 (ตารางต่อเนื่อง)

No.	เอ็นเอสสามโปรตีนเอส ซีโรทัยป์ 1				เอ็นเอสสามโปรตีนเอส ซีโรทัยป์ 2				เอ็นเอสสามโปรตีนเอส ซีโรทัยป์ 3				เอ็นเอสสามโปรตีนเอส ซีโรทัยป์ 4			
	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula
25	set8	149675	5.63E-08	C31H32N4O3	set8	170382	6.63E-09	C47H42N4O4S2	set8	126907	2.63E-08	C35H22Br2O4	set1	2216	1.54E-07	C38H38N4O4
26	set3	7276	6.03E-08	C35H40N4O3	set4	7586	7.37E-09	C48H63N3O12	set8	198680	2.63E-08	C33H23Cl3N6O2S2	set5	119898	1.64E-07	C40H22S2
27	set8	198679	6.06E-08	C33H24Cl2N6O2S2	set5	158259	8.20E-09	C40H20Br4	set8	129130	2.68E-08	C35H22I2O4	set8	35925	1.71E-07	C32H35N7O
28	set8	185312	6.38E-08	C32H27Cl2NO2	set8	167687	8.45E-09	C43H26I2N4O	set5	155061	2.69E-08	C35H52O2	set5	83784	1.74E-07	C36H22N2O6
29	set8	167688	6.45E-08	C43H26I2N4O	set2	1353	8.75E-09	C46H50N6O2	set8	126360	2.72E-08	C37H28O4	set5	46913	1.75E-07	C34H35NO13S
30	set8	237530	6.54E-08	C29H22N2O4S3	set2	3297	8.99E-09	C51H79NO13	set8	31512	2.73E-08	C32H34N6O4	set8	169245	1.76E-07	C31H20O4S2
31	set8	134468	6.93E-08	C34H24N2O2	set5	176132	9.39E-09	C44H30O2	set5	177861	2.76E-08	C20H12N10O3	set2	571	1.78E-07	C34H55NO3
32	set8	172281	7.13E-08	C32H22ClN5O2S2	set8	167689	9.43E-09	C45H32N4O	set4	7911	2.83E-08	C55H69N3O12	set5	177945	1.90E-07	C35H25N3O3
33	set3	8058	7.62E-08	C40H36O12	set5	53211	9.63E-09	C42H42N4O2	set8	206707	2.89E-08	C35H25I2N2	set8	144951	1.95E-07	C35H22I2O4
34	set8	73802	7.80E-08	C28H40N6O2	set8	183394	9.87E-09	C34H34N8O4	set5	53211	3.02E-08	C42H42N4O2	set5	34481	1.96E-07	C34H32N2O
35	set8	22778	7.89E-08	C29H29N3O3S	set8	197035	1.00E-08	C47H40N4O5S2	set8	35781	3.03E-08	C31H39N7O	set4	1425	2.00E-07	C35H40O12
36	set8	90601	8.24E-08	C28H32N10O	set8	130907	1.00E-08	C52H32N2O8	set8	126923	3.08E-08	C39H28O4	set8	36506	2.05E-07	C29H27N7O4
37	set8	166990	8.40E-08	C42H28N4O3	set8	167448	1.07E-08	C47H32N2O2	set8	34869	3.13E-08	C30H36ClN5O3S	set8	167676	2.06E-07	C36H24I2N4
38	set8	167681	8.47E-08	C48H34N6O	set5	28670	1.08E-08	C38H30Cl4N2O4	set8	127508	3.23E-08	C38H33NO	set8	53372	2.08E-07	C36H40FN3O8S
39	set5	74932	8.69E-08	C40H55NO2	set5	56348	1.13E-08	C39H36O9	set8	35385	3.35E-08	C29H41N7O	set1	1223	2.15E-07	C27H30O16
40	set4	5926	8.71E-08	C42H66O17	set5	68144	1.16E-08	C41H30O27	set5	157247	3.36E-08	C38H30N4	set2	1725	2.28E-07	C31H28N6O
41	set8	610	9.05E-08	C33H22ClN3O5S2	set8	219645	1.17E-08	C40H32N2O2	set5	37219	3.83E-08	C35H41N5O5	set5	55325	2.28E-07	C38H40N10
42	set5	56348	9.23E-08	C39H36O9	set8	130922	1.19E-08	C52H24N2O14	set8	70867	3.92E-08	C29H36N6O3	set5	176227	2.30E-07	C34H28N6
43	set3	6846	9.45E-08	C32H39O20	set8	126912	1.20E-08	C55H58O4	set8	198679	4.05E-08	C33H24Cl2N6O2S2	set4	12766	2.37E-07	C28H19N5O5S
44	set5	73178	9.86E-08	C34H14Br2O2	set5	82803	1.21E-08	C44H34	set8	35879	4.22E-08	C30H37N7O	set2	771	2.41E-07	C31H37NO4
45	set8	26022	1.00E-07	C37H37N7O	set5	172656	1.21E-08	C33H18N4O9S2	set3	8399	4.28E-08	C33H40O19	set5	3574	2.52E-07	C31H36N4
46	set8	180889	1.04E-07	C45H34N2O6	set5	86091	1.23E-08	C34H16Cl2N2O6	set8	35746	4.35E-08	C29H41N7O	set8	35781	2.56E-07	C31H39N7O
47	set8	126502	1.06E-07	C30H24N2O5S2	set5	36717	1.24E-08	C54H90O	set5	28669	4.44E-08	C38H34N2O4	set8	152371	2.68E-07	C31H18N2O2
48	set4	14816	1.07E-07	C32H28N2O2S	set3	7127	1.26E-08	C45H72O17	set8	170382	4.55E-08	C47H42N4O4S2	set8	223544	2.80E-07	C39H33N3O
49	set8	24405	1.08E-07	C35H39N7O2	set8	125138	1.27E-08	C34H26ClNOPS	set4	11252	4.58E-08	C27H30O16	set5	46117	2.83E-07	C34H36O8
50	set8	35385	1.09E-07	C29H41N7O	set5	37219	1.30E-08	C35H41N5O5	set8	150930	4.61E-08	C24H28ClN7O	set4	10401	2.87E-07	C36H24ClN5O2S

ตารางที่ 15 ผลของกลุ่มยาที่ได้จากการสกรีนด้วยโปรแกรม AutoDock ของกลุ่มโปรตีนเอ็นเอสสามโปรตีน (2FOM) type 1, 2, 3 และ 4 (ตารางต่อเนื่อง)

No.	เอ็นเอสสามโปรตีนเอส ซีโรทัยป์ 1				เอ็นเอสสามโปรตีนเอส ซีโรทัยป์ 2				เอ็นเอสสามโปรตีนเอส ซีโรทัยป์ 3				เอ็นเอสสามโปรตีนเอส ซีโรทัยป์ 4			
	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula
51	set8	35676	1.12E-07	C31H33N7O	set8	130905	1.31E-08	C52H32N2O8	set5	53914	4.70E-08	C39H56O3	set8	64136	2.88E-07	C32H35N7O
52	set8	41680	1.15E-07	C26H23FN2O5S	set1	274	1.33E-08	C38H35N5O6S2	set8	83336	4.76E-08	C25H29FN6O2	set8	36004	2.88E-07	C30H37N7O
53	set5	94447	1.17E-07	C37H54O2	set8	167488	1.33E-08	C38H29N3O5	set8	163387	4.82E-08	C33H28N4O5S	set8	217433	2.93E-07	C33H28CINO4
54	set8	17773	1.18E-07	C27H24CIN3O3S	set8	172286	1.33E-08	C33H24CIN5O2S2	set5	55345	4.85E-08	C31H28CIN3O	set8	207893	2.93E-07	C35H18N2O2
55	set8	152836	1.20E-07	C33H25NO3	set8	36523	1.34E-08	C31H31N7O3	set5	156007	4.87E-08	C39H26N2O	set8	169419	2.95E-07	C34H28CIN2O
56	set5	44069	1.21E-07	C38H36N2O11S	set2	259	1.41E-08	C52H76N2O13	set8	169756	4.87E-08	C40H20CIN2O6	set5	10569	3.03E-07	C36H26CIN3O6
57	set8	75666	1.21E-07	C29H36N6O2	set5	68836	1.41E-08	C38H50O6	set3	9784	5.03E-08	C28H19IN5O5	set2	1292	3.04E-07	C30H40N4
58	set5	55345	1.25E-07	C31H28CIN3O	set8	167447	1.41E-08	C42H29N3O3	set1	939	5.08E-08	C40H56	set8	35274	3.08E-07	C28H32N8OS
59	set5	70306	1.26E-07	C34H16O2	set5	125938	1.43E-08	C65H58N2O3S2	set5	133601	5.31E-08	C34H26N2O2	set5	103057	3.10E-07	C30H11BIO2
60	set8	35925	1.31E-07	C32H35N7O	set3	8446	1.45E-08	C45H38O16	set5	68930	5.34E-08	C41H62O13	set5	177851	3.20E-07	C20H11CIN5O2
61	set8	206014	1.32E-07	C35H29CIN3O	set4	1239	1.50E-08	C44H26O12	set5	76303	5.42E-08	C34H26O4	set8	169790	3.20E-07	C44H24N2O6
62	set8	132710	1.32E-07	C25H28CIN5	set5	179331	1.50E-08	C37H34N2O	set5	58534	5.42E-08	C38H40N4O	set8	241740	3.26E-07	C30H20BNN3OS
63	set8	25346	1.34E-07	C28H26FN3O3S	set5	124308	1.50E-08	C48H34O2	set5	62932	5.46E-08	C35H44N2O3	set8	54500	3.27E-07	C31H33N3OS2
64	set4	16971	1.35E-07	C30H24N6O6S2	set1	2330	1.51E-08	C51H62O17	set4	17	5.47E-08	C33H40O19	set5	116037	3.30E-07	C40H34N6
65	set8	129130	1.35E-07	C35H22I2O4	set5	187552	1.51E-08	C51H42F6N6O6	set5	5764	5.49E-08	C39H36O10S2	set5	15677	3.32E-07	C30H44N2O
66	set3	8454	1.36E-07	C39H32O15	set3	8445	1.52E-08	C41H30O26	set3	8637	5.57E-08	C26H30O12	set8	145247	3.34E-07	C36H34O6
67	set8	35553	1.37E-07	C28H34N6O	set1	2281	1.55E-08	C78H136O20	set8	128936	5.73E-08	C34H24N6O4	set8	54511	3.41E-07	C31H35N3OS2
68	set8	169245	1.40E-07	C31H20O4S2	set2	1982	1.56E-08	C40H54O2	set8	35254	5.75E-08	C34H36FN5O4S	set8	205866	3.43E-07	C29H30N2O4
69	set5	141453	1.41E-07	C36H24N6O2	set8	179441	1.56E-08	C38H28Br2N4O4S	set8	35674	5.75E-08	C30H37N7O	set5	45018	3.46E-07	C29H28N4
70	set3	8481	1.45E-07	C28H22O6	set8	167413	1.57E-08	C41H32N2O2	set8	95469	5.94E-08	C30H33N5O2S	set3	7475	3.52E-07	C29H28N4
71	set8	130272	1.46E-07	C38H28O4	set8	179439	1.59E-08	C38H29BN4O4S	set8	54511	5.94E-08	C31H35N3OS2	set2	1716	3.53E-07	C26H30N4O3
72	set3	8463	1.48E-07	C28H10O16	set8	126907	1.61E-08	C35H22Br2O4	set3	8442	6.02E-08	C30H26O12	set5	57813	3.61E-07	C34H32N2O
73	set8	235681	1.48E-07	C28H20CIN3O3S2	set1	499	1.63E-08	C41H43N3O7S	set4	10896	6.03E-08	C32H41N2O2	set8	208455	3.71E-07	C36H33CIN4O4
74	set5	46117	1.53E-07	C34H36O8	set5	162551	1.68E-08	C42H26N4O3	set8	90601	6.03E-08	C28H32N10O	set2	862	3.73E-07	C33H38N2O2S
75	set8	26168	1.56E-07	C27H30N6O	set5	55345	1.70E-08	C31H28CIN3O	set8	35273	6.05E-08	C30H28N8OS	set5	176186	3.80E-07	C28H20N6
76	set5	113089	1.61E-07	C28H20O	set5	191784	1.71E-08	C30H22CIN8O4S4	set8	302992	6.11E-08	C34H32NO	set5	26395	3.83E-07	C37H42O3

ตารางที่ 15 ผลของกลุ่มยาที่ได้จากการสกรีนด้วยโปรแกรม AutoDock ของกลุ่มโปรตีนเอ็นเอสสามโปรตีนเอส (2FOM) type 1, 2, 3 และ 4 (ตารางต่อเนื่อง)

No.	เอ็นเอสสามโปรตีนเอส ซีโรทัยป์ 1				เอ็นเอสสามโปรตีนเอส ซีโรทัยป์ 2				เอ็นเอสสามโปรตีนเอส ซีโรทัยป์ 3				เอ็นเอสสามโปรตีนเอส ซีโรทัยป์ 4			
	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula	Ligand set	Ligand ID	Ki	Formula
77	set5	68836	1.61E-07	C38H50O6	set8	126360	1.71E-08	C37H28O4	set5	105583	6.31E-08	C34H18O2	set8	217422	3.87E-07	C34H29ON2O
78	set5	31186	1.62E-07	C31H41N3O2	set5	54092	1.74E-08	C46H50N6O2	set5	48838	6.32E-08	C34H35NO10	set4	9037	3.96E-07	C37H36
79	set8	23255	1.63E-07	C35H38FN3O2	set8	222614	1.75E-08	C42H28N4O	set8	16089	6.32E-08	C26H22ClN3O3S	set3	8442	3.98E-07	C30H26O12
80	set8	41922	1.64E-07	C27H26N4O3S2	set4	17980	1.76E-08	C42H32O9	set8	73802	6.33E-08	C28H40N6O2	set8	36524	4.06E-07	C29H35N7O3
81	set8	222622	1.64E-07	C43H28N4O3	set5	180022	1.76E-08	C64H55BrO5	set3	8459	6.34E-08	C30H26O12	set8	166914	4.08E-07	C36H18N4O2S
82	set8	41852	1.66E-07	C26H24N4O3S2	set4	17440	1.78E-08	C48H44N6O4	set8	102864	6.53E-08	C25H28N6S	set8	126907	4.09E-07	C35H22Br2O4
83	set8	125138	1.66E-07	C34H26ClNOPS	set5	76303	1.80E-08	C34H26O4	set8	211465	6.77E-08	C33H25N7O6S2	set8	45810	4.23E-07	C31H29N7O2
84	set4	20528	1.67E-07	C27H18N2S4	set3	8434	1.82E-08	C41H28O26	set8	35785	6.79E-08	C30H37N7O	set5	28422	4.27E-07	C30H46O9
85	set5	190101	1.67E-07	C26H19NO2S	set3	8539	1.83E-08	C60H56O11	set5	43962	7.15E-08	C40H39N3O10	set5	177143	4.27E-07	C28H22F2N6O
86	set8	125261	1.67E-07	C41H24N4	set6	4739	1.83E-08	C56H64N12P4	set8	206709	7.32E-08	C36H28N2O	set8	35553	4.27E-07	C28H34N6O
87	set9	3716	1.70E-07	C24H28BrNO4	set2	336	1.87E-08	C53H78N2O12	set8	242678	7.50E-08	C33H31BrN2O2S	set8	70220	4.33E-07	C28H27ClN6O
88	set5	87735	1.71E-07	C32H18Cl2N2O6S	set4	7617	1.87E-08	C56H79N3O12	set8	125252	7.54E-08	C35H22Br2O4	set4	1239	4.36E-07	C44H26O12
89	set5	163141	1.72E-07	C30H34O3	set5	82801	1.87E-08	C38H30O	set8	36058	7.55E-08	C28H39N7O2	set5	58183	4.38E-07	C32H22Cl2N4O2
90	set8	202363	1.74E-07	C26H20N4O4S	set5	43962	1.91E-08	C40H39N3O10	set3	5815	7.58E-08	C24H26ClFN4O	set8	26073	4.47E-07	C29H35N7O2
91	set8	126912	1.76E-07	C55H58O4	set8	34605	1.91E-08	C37H40N4O8S2	set8	47076	7.59E-08	C35H25Cl2N3O2	set5	46377	4.48E-07	C31H32N2O4
92	set8	126879	1.81E-07	C35H22Cl2O4	set8	169358	1.94E-08	C25H23FN8O3	set5	18847	7.60E-08	C26H34N2O2	set5	48838	4.49E-07	C34H35NO10
93	set5	157247	1.82E-07	C38H30N4	set2	51	1.97E-08	C49H44N8O5	set8	71448	7.68E-08	C26H29N7O3	set8	129011	4.49E-07	C35H30S2
94	set5	123478	1.83E-07	C42H36Br2N2P2	set5	179429	1.97E-08	C52H34Br2O4	set8	160592	7.79E-08	C33H22O6S2	set5	87968	4.51E-07	C35H20N2O5
95	set8	125796	1.86E-07	C31H20O6	set8	167688	2.01E-08	C43H26I2N4O	set2	717	7.81E-08	C40H56	set8	90601	4.75E-07	C28H32N10O
96	set8	95649	1.89E-07	C24H33N9	set5	155270	2.02E-08	C44H32O2	set8	70866	7.82E-08	C31H34N6O2	set5	43782	4.76E-07	C36H54ClNO2
97	set8	75191	1.89E-07	C28H30N6O2	set5	42595	2.08E-08	C34H33Cl2N3O10	set5	30818	7.83E-08	C38H56N2O2	set8	166983	4.77E-07	C36H22Cl2I2N4
98	set8	211465	1.89E-07	C33H25N7O6S2	set5	160031	2.11E-08	C32H20Br2N4O2S2Zn	set5	141311	7.83E-08	C28H46	set5	80546	4.90E-07	C34H16O4
99	set5	127257	1.90E-07	C28H24N2O2S	set5	58183	2.11E-08	C32H22Cl2N4O2	set8	160593	8.10E-08	C33H20O7S2	set8	202363	4.92E-07	C26H20N4O4S
100	set5	176175	1.90E-07	C32H26N4	set1	2336	2.14E-08	C62H64N2O20	set5	161427	8.21E-08	C42H40N4	set8	35680	4.95E-07	C29H35N7O

สรุปและวิจารณ์ผลการทดลอง

ส่วนที่ 1 การออกแบบ peptide inhibitor

ได้พัฒนาเทคนิคทางคอมพิวเตอร์โดยใช้วิธีเจเนติกอัลกอริทึมร่วมกับการใช้สกอริงฟังก์ชัน 2 ชนิดในการคำนวณเป็นค่าคะแนนฟิตเนส ให้เหมาะสมกับการหาตัวยับยั้งชนิดเปปไทด์ โดยช่วยแก้ปัญหาในด้านระยะเวลาที่ใช้ในการคำนวณ และยังทำให้สามารถออกแบบสายเปปไทด์ได้อย่างมีประสิทธิภาพมากกว่าวิธีการที่เคยมีรายงานมาก่อนหน้านี้ เนื่องจากช่วยให้สามารถหลบ (escape) จากจุดต่ำสุดสัมพัทธ์ได้อย่างมีประสิทธิภาพ จึงส่งผลให้สามารถค้นหาเปปไทด์ที่มีค่าคะแนนฟิตเนสที่ดีได้อย่างรวดเร็วกว่าวิธีการที่เคยมีรายงานมาก่อนหน้านี้ และเมื่อนำเอาเทคนิคนี้มาคำนวณหาตัวยับยั้งชนิดเปปไทด์ของแต่ละซีโรทัยป์ จะได้ประมาณ 300,000 ลำดับต่อหนึ่งตำแหน่งต่อหนึ่งวิธีในแต่ละซีโรทัยป์

ซึ่งพบว่าคะแนนที่ได้จากโครงสร้าง wild type และ mutant มีความแตกต่างกันประมาณ 6 – 10% ซึ่งแสดงให้เห็นว่าตัวยับยั้งชนิดเปปไทด์ที่ได้ออกแบบใหม่นั้นมีความสามารถในการยับยั้งเพิ่มสูงขึ้น นอกจากนี้ยังได้พัฒนาเทคนิคทางคอมพิวเตอร์เพิ่มเติมโดยสามารถหาตัวยับยั้งชนิดเปปไทด์ที่มีลักษณะที่เหมาะสมที่สุดจาก 300,000 ลำดับ ให้เหลือเป็น 20-100 ลำดับ ในแต่ละซีโรทัยป์ เพื่อนำไปทดสอบในห้องปฏิบัติการต่อไป

ส่วนที่ 2 สกรีนยาหรือสารต้านไวรัสเดงกี

ผู้วิจัยได้ลิสต์ของยาหรือสารต้านไวรัสเดงกี ที่เป็น candidate ซึ่งคาดว่าจะส่งผลต่อโปรตีนเอนVELOPE (envelope), เอ็นเอสห้าเมทิลทรานสเฟอเรส (NS5 Methyltransferase), เอ็นเอสสามโปรตีเอส (NS3/NS2 protease), และเฮลิเคส (helicase) ของไวรัสเดงกี type1, 2, 3 และ 4 ด้วยเทคนิค molecular docking โดยได้ลิแกนด์ที่มีค่า K_i ต่ำสุดใน 100 อันดับแรกของแต่ละโปรตีนในแต่ละซีโรทัยป์ ลิแกนด์ที่ได้เป็น approved drugs และ drug-like compounds จากฐานข้อมูล 9 แหล่งที่รวบรวมไว้ในเว็บไซต์ <http://ligand.info> ได้จำนวนทั้งหมด 1,175 ลิแกนด์ นอกจากนี้ ผู้วิจัยยังได้พัฒนาวิธีการสกรีนยาก่อนการดocking โดยพิจารณาการใช้ขนาดของ binding site ของโปรตีนที่สอดคล้องกับขนาดของลิแกนด์ เพื่อลดระยะเวลาในการดocking ลิแกนด์ที่มีจำนวนมากได้ อย่างไรก็ตาม โครงสร้างลิแกนด์ที่ได้เหล่านี้ อาจจะต้องทำการศึกษาในเชิงลึกต่อไปว่า โครงสร้างลิแกนด์มีผลในระดับของการยับยั้งการทำงานของโปรตีนดังกล่าว ได้มากน้อยเพียงใด ซึ่งประโยชน์ที่ได้รับ สามารถนำไปสู่การคัดเลือกโครงสร้างลิแกนด์ที่มีประสิทธิภาพสูง ก่อนนำไปใช้ในการทดลองทางห้องปฏิบัติการเพื่อหาค่า inhibitory activity ของยาในลิสต์ที่เป็น candidate นี้ต่อไป

หนังสืออ้างอิง

1. Freedman DO, Weld LH, Kozarsky PE, Fisk T, Robins R, von Sonnenburg F, *et al.* **Spectrum of disease and relation to place of exposure among ill returned travelers.** *N Engl J Med* 2006,354:119-130.
2. Molyneux EM, Maitland K. **Intravenous fluids--getting the balance right.** *N Engl J Med* 2005,353:941-944.
3. Wills BA, Nguyen MD, Ha TL, Dong TH, Tran TN, Le TT, *et al.* **Comparison of three fluid solutions for resuscitation in dengue shock syndrome.** *N Engl J Med* 2005,353:877-889.
4. Oishi K, Saito M, Mapua CA, Natividad FF. **Dengue illness: clinical features and pathogenesis.** *J Infect Chemother* 2007,13:125-133.
5. Pang T, Cardoso MJ, Guzman MG. **Of cascades and perfect storms: the immunopathogenesis of dengue haemorrhagic fever-dengue shock syndrome (DHF/DSS).** *Immunol Cell Biol* 2007,85:43-45.
6. Green S, Rothman A. **Immunopathological mechanisms in dengue and dengue hemorrhagic fever.** *Curr Opin Infect Dis* 2006,19:429-436.
7. Melino S, Paci M. **Progress for dengue virus diseases. Towards the NS2B-NS3pro inhibition for a therapeutic-based approach.** *Febs J* 2007,274:2986-3002.
8. Shiryayev SA, Kozlov IA, Ratnikov BI, Smith JW, Lebl M, Strongin AY. **Cleavage preference distinguishes the two-component NS2B-NS3 serine proteinases of Dengue and West Nile viruses.** *Biochem J* 2007,401:743-752.
9. Shiryayev SA, Ratnikov BI, Aleshin AE, Kozlov IA, Nelson NA, Lebl M, *et al.* **Switching the substrate specificity of the two-component NS2B-NS3 flavivirus proteinase by structure-based mutagenesis.** *J Virol* 2007,81:4501-4509.
10. Beasley DW. **Recent advances in the molecular biology of west nile virus.** *Curr Mol Med* 2005,5:835-850.
11. Mukhopadhyay S, Kuhn RJ, Rossmann MG. **A structural perspective of the flavivirus life cycle.** *Nat Rev Microbiol* 2005,3:13-22.
12. Leung D, Schroder K, White H, Fang NX, Stoermer MJ, Abbenante G, *et al.* **Activity of recombinant dengue 2 virus NS3 protease in the presence of a truncated NS2B co-factor, small peptide substrates, and inhibitors.** *J Biol Chem* 2001,276:45762-45771.

13. Yin Z, Patel SJ, Wang WL, Wang G, Chan WL, Rao KR, *et al.* **Peptide inhibitors of Dengue virus NS3 protease. Part 1: Warhead.** *Bioorg Med Chem Lett* 2006,16:36-39.
14. Chanprapaph S, Saparpakorn P, Sangma C, Niyomrattanakit P, Hannongbua S, Angsuthanasombat C, Katzenmeier G. **Competitive inhibition of the dengue virus NS3 serine protease by synthetic peptides representing polyprotein cleavage sites.** *Biochem Biophys Res Commun* 2005,330:1237-1246.
15. Yin Z, Patel SJ, Wang WL, Chan WL, Ranga Rao KR, Wang G, *et al.* **Peptide inhibitors of dengue virus NS3 protease. Part 2: SAR study of tetrapeptide aldehyde inhibitors.** *Bioorg Med Chem Lett* 2006,16:40-43.
16. Kiat TS, Phippen R, Yusof R, Ibrahim H, Khalid N, Rahman NA. **Inhibitory activity of cyclohexenyl chalcone derivatives and flavonoids of fingerroot, Boesenbergia rotunda (L.), towards dengue-2 virus NS3 protease.** *Bioorg Med Chem Lett* 2006,16:3337-3340.
17. Koonin EV. **Computer-assisted identification of a putative methyltransferase domain in NS5 protein of flaviviruses and lambda 2 protein of reovirus.** *J Gen Virol* 1993,74 (Pt 4):733-740.
18. Tan BH, Fu J, Sugrue RJ, Yap EH, Chan YC, Tan YH. **Recombinant dengue type 1 virus NS5 protein expressed in Escherichia coli exhibits RNA-dependent RNA polymerase activity.** *Virology* 1996,216:317-325.
19. Crance JM, Scaramozzino N, Jouan A, Garin D. **Interferon, ribavirin, 6-azauridine and glycyrrhizin: antiviral compounds active against pathogenic flaviviruses.** *Antiviral Res* 2003,58:73-79.
20. Modis Y, Ogata S, Clements D, Harrison SC. Structure of the dengue virus envelope protein after membrane fusion. *Nature* 2004,427:313-319.
21. Costin JM, Jenwitheesuk E, Lok S-M, Hunsperger E, Conrads KA, *et al.*, **Structural Optimization and De Novo Design of Dengue Virus Entry Inhibitory Peptides.** *PLoS Negl Trop Dis* 2010, 4(6): e721
22. Hrobowski YM, Garry RF, Michael SF. **Peptide inhibitors of dengue virus and West Nile virus infectivity.** *Virol J* 2005,2:49.
23. Courageot MP, Frenkiel MP, Dos Santos CD, Deubel V, Despres P. **Alpha-glucosidase inhibitors reduce dengue virus production by affecting the initial steps of virion morphogenesis in the endoplasmic reticulum.** *J Virol* 2000,74:564-572.

24. Schul W, Liu W, Xu HY, Flamand M, Vasudevan SG. **A dengue fever viremia model in mice shows reduction in viral replication and suppression of the inflammatory response after treatment with antiviral drugs.** *J Infect Dis* 2007,195:665-674.
25. Whitby K, Pierson TC, Geiss B, Lane K, Engle M, Zhou Y, *et al.* **Castanospermine, a potent inhibitor of dengue virus infection in vitro and in vivo.** *J Virol* 2005,79:8698-8706.
26. Samudrala R, Moult J. **Determinants of side chain conformational preferences in protein structures** *Protein Eng.* 1998, 11(11):991–997.
27. Samudrala R and Moult J, **An all-atom distance-dependent conditional probability discriminatory function for protein structure prediction.** *J Mol Biol* 1998, 275: 895–916.
28. Canutescu AA, Shelenkov AA, Dunbrack RL., **A graph-theory algorithm for rapid protein side-chain prediction.** *Jr Protein Sci.* 2003, 12(9):2001–2014.
29. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten KJ. **Scalable molecular dynamics with NAMD.** *Comput. Chem.* 2005, 26 (16):1781.
30. Morris G.M., *et al.*, **AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility.** *J Comput Chem* 2009, 30:2785-2791.

Output จากโครงการวิจัยที่ได้รับทุนจาก สกว.

ได้ทำการตีพิมพ์ผลงานวิจัย เรื่อง “Pre-docking filter for protein and ligand 3D structures” เมื่อเดือนธันวาคม ปี 2551 ในวารสารต่างประเทศ Bioinformation (*Wilantho A, Tongsima S, Jenwitheesuk E. Pre-docking filter for protein and ligand 3D structure. Bioinformation 2008;3:189-193.*)

รายละเอียดในภาคผนวก ก

Pre-docking filter for protein and ligand 3D structures

Alisa Wilantho¹, Sissades Tongsim¹ and Ekachai Jenwitheesuk^{1,*}

¹National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency, 113 Thailand Science Park, Phahonyothin Road, Klong 1, Klongluang, Pathumtani 12120, Thailand;
Ekachai Jenwitheesuk* - Email : ekachai@biotec.or.th; Phone: 66 2564 6700; Fax: 66 2564 6701; * Corresponding author

received November 15, 2008; accepted December 01, 2008; published December 31, 2008

Abstract:

Virtual drug screening using protein-ligand docking techniques is a time-consuming process, which requires high computational power for binding affinity calculation. There are millions of chemical compounds available for docking. Eliminating compounds that are unlikely to exhibit high binding affinity from the screening set should speed-up the virtual drug screening procedure. We performed docking of 6353 ligands against twenty-one protein X-ray crystal structures. The docked ligands were ranked according to their calculated binding affinities, from which the top five hundred and the bottom five hundred were selected. We found that the volume and number of rotatable bonds of the top five hundred docked ligands are similar to those found in the crystal structures and corresponded with the volume of the binding sites. In contrast, the bottom five hundred set contains ligands that are either too large to enter the binding site, or too small to bind with high specificity and affinity to the binding site. A pre-docking filter that takes into account shapes and volumes of the binding sites as well as ligand volumes and flexibilities can filter out low binding affinity ligands from the screening sets. Thus, the virtual drug screening procedure speed is increased.

Keywords: virtual drug screening; ligand volume; protein binding site; docking filter; binding affinity

Background:

Virtual screening techniques are becoming increasingly more important in drug discovery. A popular method for virtual screening is molecular docking [1, 2], which selects small-molecule structures from databases such as ChemBank [3], ChemPDB [4], KEGG [5], and NCI [6] and docks them into the protein binding site [7]. These processes involve the prediction of binding energies and analysis of molecular binding modes, which are time consuming and computationally expensive. The two-dimensional (2D) fingerprint technique, a virtual screening method which measures the structural similarity of molecules has been developed to address the above problems [8]. The similarity search is based upon the “similar property principle”, which states that molecules that are structurally similar are likely to have similar properties [9]. This technique uses a ligand with known chemical properties, inhibitory activities, or binding modes for a target of interest as a reference for searching similar ligands in the database regardless of the shape and size of the protein binding site. The accuracy of this method depends on which similarity coefficient is used [10-12], and the Tanimoto coefficient is most popularly employed [13]. Based on the “lock-and-key” principle, we propose a novel pre-docking procedure that matches the sizes of the ligand with the protein binding site, and optimizes the grid-box size before docking. This simple procedure dramatically reduces the size of screening ligand sets, significantly reducing time and effort required for virtual drug screening.

Methodology:

Preparation of ligand and protein structures for docking

Preparation of ligand structures

The ligands used in this study were downloaded from Ligand.Info Meta-Database [14]. Ligands set-1 consisted of 2344 structures from ChemBank and set-2 consisted of 4009 structures from ChemPDB. The downloaded ligands in the SDF format were first converted to the Protein Data Bank (PDB) format using Open Babel [15]. The Gasteiger charges and rotatable bonds were then assigned to the PDB ligands using AutoDockTool [16]. All rotatable bonds were allowed to move freely.

Preparation of protein structures

Twenty-one protein X-ray crystal structures from the Protein Data Bank [17] were downloaded. The proteins and their PDB structure identifiers (PDB ID) are given in Table 1 (supplementary material). Of the twenty-one protein structures, sixteen have co-crystallized ligands (X-ray ligand) in the binding site. The ligand contained in each protein structure was removed from the binding site and saved to a new file. The missing atoms in each protein structure were searched for and fixed using SwissPDB [18]. The Gasteiger charges and the solvation term were then added to the protein structure using the AutoDockTool.

Ligand SMILES string similarity search

The ligands extracted from the X-ray crystal structure obtained from the previous step were converted to the SMILES string format, and used as an input for similarity search against the ligands in the ChemBank and the ChemPDB sets in step of the Preparation of ligand structures using Tanimoto coefficient cutoffs of 0.5, 0.6, and 0.7, respectively.

Calculation of ligand molecular volume

The volumes of the ligands in the screening set were calculated using Mol_Volume version 1.0 [19]. The van der Waals radii value for each atom type was derived from the CHARMM 22 force field. The radius of the spherical probe (R_PROBE) was set to 2.0 Å, and the GRID_STEP was set to 0.5 Å. The volumes of the ligands extracted from the X-ray crystal structures were calculated using the same protocol. The extracted ligand name (X-ray ligand) and its calculated volumes are shown in Table 2 (supplementary material).

Calculation of protein binding site molecular volume

Protein binding site volumes were calculated using the CASTp server (<http://sts-fw.bioengr.uic.edu/castp>) [20]. The solvent probe radius used for volume calculation was 1.4 Å. CASTp identifies all surface pockets with the chosen volume values, and then displays them on the computer screen. Pockets calculated by CASTp that matched the pocket resolved by crystallography were selected, and the volume of that calculated pocket was taken as the volume of the protein binding site.

Protein-ligand docking

Grid-box generation

The grid parameter file of each protein was generated using AutoDockTool. A grid-box was generated that was large enough to cover the entire protein binding site and accommodate all ligands to move freely. The number of grid points in x, y, and z-axes were 60×60×60. The distance between two connecting grid points was 0.375 Å. The center of the ligand in the X-ray crystal structure was used as the center of the grid-box. For protein structures that do not have ligands in the binding site, the center of the binding site was estimated from the structure and taken as the center of the grid-box.

Ligand docking

AutoDock4 and a Lamarckian Genetic Algorithm (LGA) [21] were used for protein-fixed ligand-flexible docking calculations. Ten search attempts (ga_run parameter) were performed for each ligand. The maximum number of energy evaluations before the termination of LGA run was 2500000 and the maximum number of generations of the LGA run before termination was 27000. Other docking parameters were set to the software's default values. After docking, the ligands were ranked according to their protein-ligand affinity (calculated inhibitory constant, K_i).

Discussion:

Size and shape of protein binding site

In this study, we categorized protein binding sites according to their sizes and shapes. Protein binding sites were classified as small (less than 1200 Å³) or large

(greater than or equals to 1200 Å³). Protein binding site shapes were classified as either simple or complex. The protein binding site classifications are shown in Table 1 (supplementary material). Sixteen protein structures had co-crystallized ligands bound in the binding site. The calculated volumes show that the majority of ligands are larger (305–5922 Å³) than the binding sites (1040–2690 Å³) in particular those in the small binding site group. However, the average volumes of the ligands (1684 Å³) and the binding sites (1638 Å³) are very similar. The protein binding site typically accommodates 50–70% of the ligand, with the remainder of the ligand occupying pockets adjacent to the binding site. For example, 50% of the GDP ligand (volume = 1460 Å³) was contained within the small binding pocket (volume = 594 Å³) of the “Filamenting temperature-sensitive mutant Z” protein (PDB ID: 1RQ7), while the rest of GDP ligand occupied pockets close to or floating over the binding site. This suggests that the optimal ligand size can potentially exceed the binding site volume.

X-ray ligand docking and ranking

To verify the docking procedure utilized in this work, we re-docked the original X-ray ligand back to its corresponding protein binding site. The X-ray ligands along with all other ligands in the screening set were ranked according to the calculated K_i . The X-ray ligands were ranked in the top ten percentiles and were also able to move back to the original positions with the root mean square deviations of less than 3 Å.

Docked ligand size and flexibility

The top 500 and the bottom 500 ligands ranked according to the K_i value for each protein were selected for further analysis. A scatter plot of the molecular volumes and the number of active torsion bonds for these ligands is shown in **Figure 1**. The top 500 ligands are clearly coincident with the sixteen X-ray ligands, which occupy volumes of 800–2800 Å³, whereas the bottom 500 ligands occupy volumes outside this range, with 95% much smaller (300–900 Å³). There does not appear to be any correlation between the number of active torsion bonds and calculated K_i ; however, the majority of the top 500 ligands have twenty or fewer active torsion bonds. These data suggest that ligands with high binding affinity are constrained by their size (volume 800–2800 Å³) and flexibility (20 or fewer active torsion bonds). For untested ligands, these parameters could be useful to prioritize docking calculations, so that priority is given to ligands of optimal size and flexibility.

Optimal size of the grid-box

In this study, a very large grid-box (10830 Å³, 22.125 Å on each side) was used because we wanted to ensure that the grid-box could cover the entire binding site, and that all ligands in the screening sets had enough space to enter and move freely in the grid-box. The volume of the grid-box was 10830 Å³ while the volume of the largest protein binding site was only 5921.8 Å³ (PDB ID: 1N8W). We hypothesized that using a very large grid-box would allow the binding of some ligands to extend beyond the actual binding site, with non-specific binding into adjacent pockets. We tested this hypothesis by generating minimal

grid-boxes that perfectly encompassed the entire binding site for each protein. The box sizes and their dimensions

are shown in Table 3 (supplementary material). The top 500 ligands were left on protein at the docked positions.

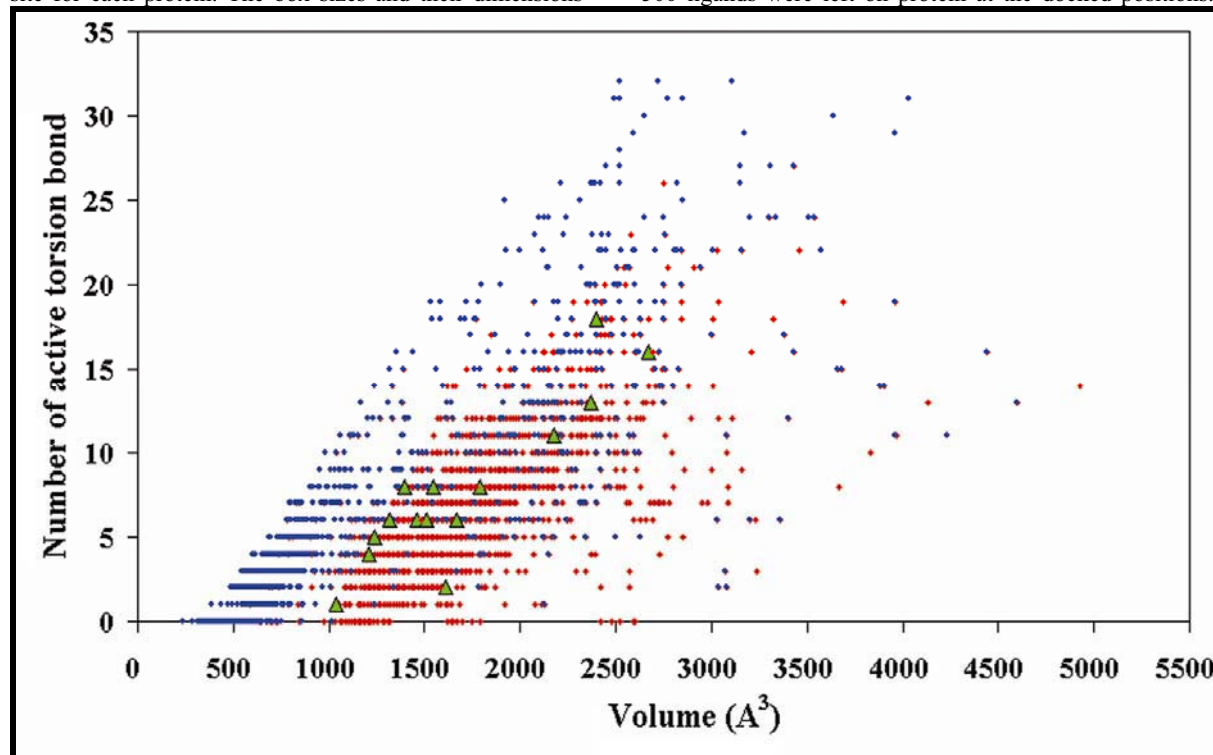


Figure 1: The volume and the number of active torsion bonds of the top 500 (red) and the bottom 500 docked ligands (blue) ranked according to the calculated inhibitory constant (K_i). Of the twenty-one protein X-ray crystal structures used in this study, sixteen structures had ligand bound in the binding site. The top 500 ligands generally had structural profiles in terms of volume and number of active torsion bond similar to those of the X-ray ligands (green) while the bottom 500 ligands were, on average, 800-900 Å³ smaller than the X-ray ligands.

Filtering of the top 500 docked ligands was performed to test how ligands occupy space beyond the protein binding site. Six thresholds of decreasing stringency, allowing progressively more of the ligand atoms to be outside the minimal grid-box were used (Table 4 in supplementary material). It is clear that on average, the top 500 ligands cannot fit entirely within the minimal grid-boxes extending outside them, since 10.8% of the ligands were rejected, even when a very relaxed 30% threshold was employed (Table 4 in supplementary material). Visual inspection of the docked structures revealed that the protein binding sites contain at least one opening space, and parts of the docked ligands were always outside of the minimal box on this side. The rejected ligands might be either too large or too long to fit entirely within the minimal box, or their chemical properties may not match perfectly well with the binding pocket so that parts of their structures bind preferentially with adjacent pockets.

Reducing the grid-box size would significantly reduce CPU time for docking calculation, an important consideration for drug-discovery when potentially millions of compounds are screened. However, it is clear from our data that this would also increase the false negative rate, leading to some high binding-affinity ligands to be missed. These false-negatives would likely include molecules with long linear shapes, or with branches which extend beyond the target binding site and bind to adjacent pockets, in

particular on the opening space side (see above). We propose that the optimal grid-box size allows approximately two-thirds of ligand molecule to occupy the target binding site, with the remaining one third able to bind with adjacent pockets. Grid-boxes of this size provide the optimal balance between the number of screening ligands and the CPU time required for docking.

SMILES strings similarity of the docked ligands and the X-ray ligands

We further explored whether the top 500 and the bottom 500 ligands docked on each protein were chemically similar to the X-ray ligand extracted from the protein-ligand co-crystal structures using the SMILES strings similarity search. The results show that, in general, more of the top 500 ligands matched with the X-ray ligands than of the bottom 500 ligands (Table 5 in supplementary material). On the other hand, even at a Tanimoto coefficient of 0.5, only thirteen ligands in the top500 list matched with the X-ray ligand. This indicates that most of the potential hits were chemically dissimilar to the X-ray ligands, yet similar in size as discussed above. Conversely, at a Tanimoto coefficient of 0.5, six out of the bottom 500 ligands matched with the X-ray ligand, suggesting that these ligands although similar to the X-ray ligand have chemical properties that are unfavorable to interactions with the binding site.

Conclusion:

There are millions of ligand structures currently available in public databases. Virtual screening of these ligands against a protein target using protein-ligand docking methods requires lengthy calculations on a high performance computer. This is a major obstacle that prevents several research groups in academia, especially those with limited computer resources to conduct research in this field. To overcome this problem, several computational techniques have been developed to reduce the calculation time. In this study, we show that ligand size may be used as an initial criterion for prioritizing ligands for docking. The ligands that are greatly different in size to that of the X-ray ligand, or to the binding site volume may be set to have low priority for screening or removed from the screening set, since these usually have low calculated binding affinity (K_i). On the other hand, ligands of optimum size within $\pm 1000 \text{ \AA}^3$ of the X-ray ligand or binding site volume may be assigned higher priority since they tend to bind with higher calculated affinity. In addition to clustering ligands into groups, adjusting the grid-box size would also help limit the number of ligands to be screened. The optimum grid-box allows one-third of the ligand to lie outside of the target binding site. Our new screening procedure, which takes into account the ligand size, the binding site volume, and the grid-box size, is easy to perform and could significantly reduce time and effort required for virtual drug screening.

Acknowledgment:

This investigation received financial support from the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Disease (TDR), National Center for Genetic Engineering and Biotechnology, Thailand (grant number BT-B-02-IG-GI-5101), and Thailand Research Fund (grant number DBG5180013) to Ekachai Jenwitheesuk. The author would like to thank Philip Shaw, Michal Guerquin, Siriphan Manochewa and people in the Biostatistics and Informatics laboratory for their critical reading and valuable comments.

References:

- [01] G. M. Morris and M. Lim-Wilby, *Methods Mol Biol.*, 443: 365 (2008) [PMID: 18446297]
- [02] J. Holyoake *et al.*, *Biophys J*, 91: 84 (2006) [PMID: 16980356]
- [03] K. P. Seiler *et al.*, *Nucleic Acids Res.*, 36: 351 (2008) [PMID: 17947324]
- [04] D. Dimitropoulos *et al.*, *Curr Protoc Bioinformatics*, 14: 14 (2006) [PMID: 18428761]
- [05] S. Goto *et al.*, *Nucleic Acids Res.*, 30: 402 (2002) [PMID: 11752349]
- [06] K. Arnold, *J Natl Cancer Inst.*, 95: 1192 (2003)[PMID: 12928341]
- [07] G. Schneider and H.J. Bohm, *Drug Discov. Today*, 7: 64 (2002) [PMID: 11790605]
- [08] P. Willett, *Drug Discov. Today*, 11: 1046 (2006) [PMID: 17129822]
- [09] M. A. Johnson and G.M. Maggiora, *John Wiley* (1990)
- [10] A. Talevi *et al.*, *J Enzyme Inhib Med Chem.*, 22: 253 (2007) [PMID: 17674806]
- [11] T. Kogej *et al.*, *J. Chem. Inf. Model*, 46: 1201 (2006) [PMID: 16711740]
- [12] G. W. Adamson and J. A. Bush, *J Chem Inf Comput Sci.*, 15: 15 (1975) [PMID: 1127038]
- [13] X. Chen and C. H. Reynolds, *J Chem Inf Comput Sci.*, 42: 1407 (2002)[PMID: 12444738]
- [14] M. von Grotthuss *et al.*, *Comb Chem High Throughput Screen*, 7: 757 (2004)[PMID: 15578937]
- [15] R. Guha *et al.*, *J Chem Inf Model*, 46: 99 (2006) [PMID: 16711717]
- [16] G. M. Morris *et al.*, *J Computational Chemistry*, 19: 1639 (1998)
- [17] H. M. Berman *et al.*, *Nucleic Acids Res.*, 28: 235 (2000)
- [18] N. Guex and M.C. Peitsch *Electrophoresis*, 18: 2714 (1997) [PMID: 9504803]
- [19] L. Kale *et al.*, *J Comp Phys.*, 151: 283 (1999)
- [20] J. Dundas *et al.*, *Nucleic Acids Res.*, 34: 116 (2006) [PMID: 16844972]
- [21] D. S. Goodsell *et al.*, *J Mol Recognit.*, 9: 1 (1996) [PMID: 8723313]

Edited by P. Kanguane

Citation: Wilantho *et al.*, Bioinformation 3(5): 189-193 (2008)

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.

Supplementary material

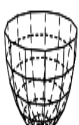
Binding site shape		PDB ID	
		Binding site volume	
		Small	Large
Simple		1OKE	NA
		1CET	NA
		1ZNY	1M4D
		1C3V	1LQU
		2FOM	1P44
		1ZAU	1EYE
		1K44	1F61
			2DEN
Complex	Irregular shape and varied depth	1R6A	1ENY
		1RQ7	
	L-shape with varied depth	1DF7	1MRN
			1N2B
			2C27
			1N8W

Table 1: Shape and size of the protein binding sites is shown. The binding sites were divided into groups according to their shapes (simple and complex) and sizes (small; volume $<1200 \text{ \AA}^3$ and large; volume $\geq 1200 \text{ \AA}^3$).

PDB ID	Protein		Ligand		
	Name	Binding site volume($\text{\AA}^3$)	Name	volume ($\text{\AA}^3$)	No. of active torsion
1C3V	Dihydrodipicolinate reductase	668	NA	NA	NA
2FOM	NS2B-NS3 protease	828	NA	NA	NA
1K44	Nucleoside diphosphate kinase	1012	NA	NA	NA
1F61	Isocitrate lyase	2735	NA	NA	NA
2DEN	Helicase	5071	NA	NA	NA
1R6A	Methyltransferase	305	RVP	1210	4
1RQ7	Filamenting temperature-sensitive mutant Z	594	GDP	1460	6
1ZNY	Guanylate kinase	664	GDP	1460	6
1DF7	Dihydrofolate reductase	752	MTX	1790	8
1ZAU	NAD ⁺ dependent DNA ligase	904	AMP	1240	4
1OKE	Envelope glycoprotein E	1044	BOG	1310	6
1CET	Lactate dehydrogenase	1074	CLQ	1410	8
1MRN	Thymidylate kinase	1268	T5A	2690	16
1M4D	Aminoglycoside 2'-N-acetyltransferase	1422	TOY	1680	6

1LQU	Mycobacterial oxidoreductase	1474	NDP	2400	13
1ENY	Enoyl-acyl carrier protein reductase	1657	NAD	2180	11
1P44	Enoyl reductase	1692	GEQ	1600	2
1N2B	Pantothenate synthetase	2008	APC	1570	8
1EYE	6-Hydroxymethyl-7,8-dihydropteroate synthase	2278	PMM	1040	3
2C27	Mycothioli synthase	3150	MA8	1540	6
1N8W	Malate synthase	5922	COA	2370	18

Table 2: Details of protein and ligand X-ray crystal structures used in this study.

PDB ID	Box center x, y, z axes	coordinate on x, y, z axes	No. of axis points on x, y, z axes
1C3V	134.796	25.327 22.087	40×50×40
2FOM	-4.451	-10.488 17.047	50×40×50
1K44	26.365	42.416 51.116	40×40×40
1F61	37.471	66.151 4.837	40×44×40
2DEN	-0.719	1.110 52.523	40×50×60
1R6A	16.498	-52.304 16.284	40×40×46
1RQ7	-8.274	37.324 7.065	40×40×40
1ZNY	24.782	8.849 30.726	40×40×40
1DF7	1.827	27.794 10.958	40×40×40
1ZAU	-10.553	43.872 68.997	40×40×40
1OKE	-11.491	80.375 45.517	60×40×40
1CET	35.592	12.727 18.358	40×40×40
1MRN	25.094	14.419 2.243	40×50×46
1M4D	20.818	29.061 13.435	40×40×40
1LQU	-4.075	-3.490 5.319	50×40×40
1ENY	-1.139	33.283 13.766	45×45×40
1P44	15.210	14.125 8.359	40×46×40
1N2B	34.878	35.004 40.786	40×40×40
1EYE	32.818	4.458 38.753	40×40×40
2C27	4.671	3.124 18.325	40×40×40
1N8W	15.581	32.585 77.100	40×50×40

Table 3: The coordinates of the grid-box center and the number of grid points on the x, y, and z axes in grid-boxes of minimum size is given. The distance between grid points was 0.375 Å.

PDB ID	Percentage of ligand atoms allowed outside of the box					
	5%	10%	15%	20%	25%	30%
1C3V	64.2	59.8	54.6	51.2	49.2	46.4
2FOM	59.2	40.0	23.4	13.2	8.6	5.6
1K44	73.4	60.6	45.8	30.6	18.8	11.4
1F61	88.4	76.6	65.2	50.4	38.2	27.2
2DEN	6.4	3.0	1.2	0.4	0.4	0.4
1R6A	50.8	36.8	24.2	16.2	11.8	8.4
1RQ7	52.0	39.0	25.6	16.8	9.0	5.8
1ZNY	78.4	63.2	51.8	39.2	27.6	18.6
1DF7	37.8	23.4	13.8	8.4	5.4	3.0
1ZAU	52.0	47.0	39.6	31.4	21.8	14.4

1OKE	12.8	7.6	3.6	1.8	1.0	0.8
1CET	61.4	47.6	34.8	23.2	15.0	8.4
1MRN	42.6	24.2	14.2	5.4	1.8	0.8
1M4D	38.0	29.0	19.8	13.6	9.2	5.2
1LQU	34.6	17.8	10.4	5.6	3.0	1.8
1ENY	40.2	22.6	11.8	5.6	2.2	1.2
1P44	41.4	31.6	22.4	14.0	9.6	6.0
1N2B	64.0	53.8	44.8	33.8	24.2	15.8
1EYE	95.6	89.2	78.8	64.2	49.4	39.6
2C27	42.8	29.0	20.4	14.0	6.2	3.6
1N8W	49.4	27.2	16.2	9.6	4.8	2.8
Average	51.6	39.4	29.6	21.3	15.1	10.8

Table 4: Percent of ligands rejected from the top500 ligand list docked with twenty-one protein binding sites and minimal grid-boxes (see Table 3). Six thresholds for rejection were tested, in which the percentage of ligand atoms outside of the grid-box was varied.

PDB ID	Ligand name	TC = 0.7		TC = 0.6		TC = 0.5	
		Top 500	Bottom 500	Top 500	Bottom 500	Top 500	Bottom 500
1R6A	RVP	1	0	19	0	47	0
1ZNY	GDP	0	0	1	1	2	14
1DF7	MTX	7	0	15	0	24	0
1ZAU	AMP	2	0	3	0	7	6
1OKE	BOG	0	0	0	0	0	0
1CET	CLQ	0	0	0	0	0	0
1MRN	T5A	1	0	11	2	23	16
1M4D	TOY	0	0	0	0	1	0
1LQU	NDP	8	0	13	4	21	4
1ENY	NAD	1	0	4	0	5	3
1P44	GEQ	0	0	1	0	28	0
1N2B	APC	0	0	7	8	13	9
1EYE	PMM	0	0	0	0	0	1
2C27	MA8	0	0	4	0	19	0
1N8W	COA	0	27	0	28	0	30
Total		20	27	78	43	190	83

Table 5: SMILES string similarity search of X-ray ligands in sixteen protein-ligand co-crystal X-ray structures against the top 500 and the bottom 500 docked ligands using Tanimoto coefficients (TC) of 0.5, 0.6, and 0.7, respectively.