

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาวิธีการทางสถิติที่เหมาะสมในการจำแนกกลุ่มสารประกอบที่มีฤทธิ์ในการยับยั้งการทำงานของเอนไซม์ HIV-1 เพื่อประโยชน์ในการคัดกรองสารประกอบที่เอื้อต่อการพัฒนาให้มีความรวดเร็วมากยิ่งขึ้น ซึ่งข้อมูลที่ใช้คือคุณลักษณะของสารประกอบทั้งหมด 69 ชนิด ซึ่งได้มาโดยอาศัยความสัมพันธ์เชิงปริมาณระหว่างโครงสร้างทางโมเลกุลและคุณสมบัติของสารประกอบ โดยเป็นสารประกอบในโครงสร้าง 1pro, 1izh, 1d4y, 1qbt, และ 1w5v เท่ากับ 37, 19, 5, 6 และ 3 ชนิด ตามลำดับ สารประกอบทั้งหมดถูกจำแนกเป็น 2 กลุ่มคือ กลุ่มสารประกอบที่มีฤทธิ์ในการยับยั้งการทำงานของเอนไซม์ HIV-1 สูง และ ต่ำ จำนวน 21 และ 48 ชนิด ตามลำดับ สำหรับการวิเคราะห์เพื่อจำแนกกลุ่มของสารประกอบจากคุณลักษณะของสารประกอบทุกโครงสร้าง และแต่ละโครงสร้าง (เฉพาะ โครงสร้าง 1pro และ 1izh) โดยใช้วิธีซิมการวิเคราะห์การจำแนกกลุ่ม และการวิเคราะห์องค์ประกอบหลัก ร่วมกับการคัดเลือกตัวแปรด้วยวิธีโมเดลลิ่งพาวเวอร์และวิธีพีชเชอร์เวท พบว่าในการวิเคราะห์ทุกโครงสร้าง การวิเคราะห์การจำแนกกลุ่มร่วมกับการคัดเลือกตัวแปรด้วยวิธี Modeling Power มีความสามารถในการจำแนกกลุ่มสารประกอบได้เหมาะสม ส่วนการวิเคราะห์ในโครงสร้าง 1pro พบว่าการวิเคราะห์องค์ประกอบหลักร่วมกับการคัดเลือกตัวแปรด้วยวิธีโมเดลลิ่งพาวเวอร์ มีความสามารถในการจำแนกกลุ่มสารประกอบได้เหมาะสม และสำหรับการวิเคราะห์ในโครงสร้าง 1izh พบว่าการวิเคราะห์การจำแนกกลุ่มหรือการวิเคราะห์องค์ประกอบหลักร่วมกับการคัดเลือกตัวแปรด้วยวิธีพีชเชอร์เวท มีความสามารถในการจำแนกกลุ่มสารประกอบได้อย่างเหมาะสม

The aim of this research was to classifier the inhibitor activity compounds against Human Immunodeficiency Virus Protease (HIV-1) using the suitable statistical methods in order to give a better performance in process of High Throughput Screening for drug discovery. The Quantitative Structure – Activity Relationship (QSAR) properties of 69 compounds were used for analysis. All compounds derived from 5 structures such as lpro, lizh, 3d4y, lqbt and lw5v with 37, 19, 5 ,6 and 3 compounds respectively. 21 compounds were determined in high activity compounds against HIV-1 (High active groups) and 48 compounds in low activity anti – HIV compounds (Low active groups). Classification methods including Soft Independent Modeling Class Analogy (SIMCA), Discriminant Analysis (DA), and Principal Component Analysis (PCA) in combination with the selection variable techniques by Modeling Power or Fisher's weight were proposed to identify the activity of a compounds to its QSAR properties, which is quantified by various explanatory variables. First, all structures of compounds were used to analyze and second, only some structures of compounds (lpro and lizh structures). The result showed that when we analyzed all structures of compounds, DA with Modeling Power method produced a high level of corrected classification. With compounds in lpro structure, the high corrected classification was found by PCA combined with Modeling Power method. With compounds in lizh structure, we found that DA or PCA combined with Fisher's weight were appropriate method to identify the activity of each compounds.