Abstract

Molecular selection in drug discovery bears resemblance to problems in the stock portfolio. Therefore, the stock portfolio method can be applied in this molecular selection problem. This study used the IC50 bio-activity value to inhibit the growth of the PTP1B protein in diabetes mellitus by using a portfolio optimization method. The SPEA2 algorithm is used to determine the pareto-optimal set for multi-objective optimization problems. There are several parameters used such as population size, archive size, maximum generation, crossover probability and mutasi probability. The solution obtained from the SPEA2 algorithm will then form an Efficient Frontier in the form of a two-dimensional graphic where each axis describes the diversity value and the probability of success of the portfolio. The data used in this study were obtained from www.ebi.ac.uk/chembl as many as 1452 pieces of data. The test was carried out 3 times with each test using a different number of molecules, namely 5 molecules, 10 molecules and 20 molecules. Each test also uses a different number of generations, such as 100, 200 and 300. The results obtained indicate that the greater the number of generations does not affect the results of the graph. Because in several tests, a certain number of generations has obtained optimal results because there has been no significant change. The number of molecules also affects the value of diversity and the value of expected return, where the greater the number of molecules used, the greater the value of expected return and resulting diversity. The range of diversity values and expected return values are also influenced by the number of molecules.

Keywords: **portfolio, drug discovery, SPEA2, efficient frontier**