Dimensionality reduction in peptide Quantitative structure–retention relationships using global optimization algorithms

Petar Zuvela, 유 준†, Katarzyna Macur[1],
Tomasz Bączek[1]
부경대학교; [1]Medical University of Gdańsk
(jayliu@pknu.ac.kr†)

Nowadays, thousands of variables can be calculated directly from the molecular structure. This makes dimensionality reduction a crucial aspect of In Quantitative Structure–Retention Relation–ships (QSRR) model development. In this work, five global optimization algorithms: Genetic Algorithms (GA), Artificial Bee Colony (ABC), Firefly Algorithm (FA), and Flower Pollination Algorithm (FPA), were used for this purpose. QSRR models were developed for 83 peptides originating from eight model proteins, with Partial Least Squares (PLS) as regression of choice.

Three performance criteria: prediction error, number of selected variables, and computational cost were used for evaluation of the algorithms. GA was found to be superior with a low error of 5.534% for 9 selected variables and the lowest computational cost. Developed GA–PLS model was extensively validated. Y–randomization procedure was performed, and an independent validation set was used with 102 peptides originating from two Bacillus subtilis proteomes.