Prediction of Activators for Pathogen Sensing Receptors using Machine Learning

Pyaree Mohan Dash[1,2], Pratiti Bhadra[1], Volkhard Helms[1] , Bernd Bufe[2]

[1] Center for Bioinformatics, Saarland Informatics Campus, Saarland University, D-66041 Saarbrücken, Germany.

[2] Department of Informatics and Microsystems Technology, University of Applied Sciences Kaiserslautern, Germany.

bernd.bufe@hs-kl.de

Abstract. Formyl peptide receptors (FPRs) are G protein-coupled receptors (GPCRs) that are predominantly expressed in the immune system, where they play a critical role in detecting bacterial invasion and inflammatory responses [1] through detection of pathogen-derived formylated peptides [2]. Recent studies highlighted an involvement of FPRs in various diseases [1, 3], such as bacterial and viral infections, Alzheimer's and prion diseases, immunodeficiency, diabetes, and cancer. Given the sheer importance of FPRs, there is an immediate need for a better understanding of the mode of action of these receptors. A current challenge in FPR research is their well-documented capability to intact with an extremely vast number of structurally diverse ligands such as bacterial and virus-derived peptides, various small non-peptide molecules, and even some lipid-derivatives, that lack any obvious common structural motifs [1]. Because of the high potential of FPRs as a therapeutic target, we developed a computational method to predict FPR ligands using machine learning. Moreover, we can provide experimental evidence that our computation models are promising data mining tools that are useful tools to identify FPR activators from a vast amount of bacterial amino-acid sequence information that is contained in public databases.

The human genome encodes the three FPR genes FPR1, FPR2, and FPR3. In this study, we focused on FPR1 and FPR2. The proposed agonist prediction classifiers utilize amino-acid composition and physicochemical properties as features. Our optimized prediction models showed high test accuracy (FPR1: 82% and FPR2: 90%), Matthew's correlation coefficient (MCC) of 0.5 (FPR1) and 0.6 (FPR2), and area under the receiver operating characteristic curve (AUC-ROC) score of 0.76 (FPR1) and 0.90 (FPR2). To demonstrate the performance of the proposed prediction models in the real world, we screened the Escherichia coli K12 proteome and selected 30 novel peptides (20 predicted as activators and 10 as non-activators) for experimental validation. Human embryonic kidney (HEK-293T) cells were used to perform a cell-based calcium flux assay using Molecular Devices' Flex station. The experimental validation showed a true negative rate of 90% (9/10 non-activators) and a true positive rate of 80% (18/20 activators). Furthermore, our study also sheds light on the physio-chemical properties of FPR agonists and antagonists. A feature descriptor analysis revealed that FPR1 is activated by peptides with higher aromaticity, low hydrophobicity, low volume, and high density when compared to the peptide activators of FPR2. Moreover, the gene set annotation analysis of the predicted FPR agonists indicated that FPR1 and FPR2 activators are involved in different metabolic processes and transport systems related to bacterial stress responses. This indicates that our models can be used to mine novel information on the biological function of FPRs, which is potentially helpful for the rational design of therapeutic approaches.

Keywords: Formyl peptide receptors, pathogen sensing, machine learning, gene ontology

References

1. The sensing of bacteria: emerging principles for the detection of signal sequences by formyl peptide receptors. Bufe B, Zufall F. Biomol Concepts. 2016 Jun 1;7(3):205-14. doi: 10.1515/bmc-2016-0013.
2. Recognition of bacterial signal peptides by mammalian formyl peptide receptors: a new mechanism for sensing pathogens. Bufe B, Schumann T, Kappl R, Bogeski I, Kummerow C, Podgórska M, Smola S, Hoth M, Zufall F. J Biol Chem. 2015 Mar 20;290(12):7369-87. doi: 10.1074/jbc.M114.626747.
3. Bacterial MgrB peptide activates chemoreceptor Fpr3 in mouse accessory olfactory system and drives avoidance behavior. Bufe B, Teuchert Y, Schmid A, Pyrski M, Pérez-Gómez A, Eisenbeis J, Timm T, Ishii T, Lochnit G, Bischoff M, Mombaerts P, Leinders-Zufall T, Zufall F. Nat Commun. 2019 Oct 25;10(1):4889. doi: 10.1038/s41467-019-12842-x.