Proceedings of the 8th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/2026.38159

Research on GPCR Biased Signaling Pathway Identification Models Based on Multimodal Pharmacodynamic Data Fusion

Yue Wang

Faculty of Pharmacy and Pharmaceutical Sciences, Monash University, Melbourne, Australia
whoami_wy@163.com

Abstract. In the current research, we developed a new method of pharmacodynamic data fusion for predicting biased signaling pathways related to G protein-coupled receptors (GPCRs). The β2-adrenergic receptor in humans (ADRB2) was chosen as the targeted GPCR, while the stable cell line of HEK293T cells with ADRB2 overexpression was used as an assay system. Four pharmacodynamics parameters were considered for data collection, namely cAMP generation, β-arrestin2 recruitment, ERK1/2 activation, and cell viability. The set of 40 molecules was screened as ADRB2 ligands at eight concentrations in triplicate. The process of data processing included four steps sequentially: curve fitting via GraphPad Prism 10, preprocessing in Python 3.11, neural network training in PyTorch 2.2, and SHAP. According to the model outputs, this novel multimodal strategy could classify the ligands as belonging to one of the three classes of biased signaling: Gs-biased, β-arrestin-biased, and balanced. In comparison to single-modality cAMP-based and β-arrestin2-based models, the multimodal model outperformed by its ability to distinguish ligands by their biased signaling profiles, reaching accuracy of 88.7%, macro-F1 of 0.872, and ROC-AUC of 0.931. Selection of key features showed four important features: cAMP pEC50, β-arrestin2 Emax, ERK1/2 max, and pathway ratios. According to SHAP results, cAMP pEC50 and β-arrestin2 Emax were two most important features, having contribution of 21.8% and 19.6%, correspondingly.

Keywords: GPCR, biased signaling, β2-adrenergic receptor, multimodal data fusion, pathway identification

1. Introduction

The coupling of G protein-coupled receptors (GPCRs) with external ligands results in the transmission of extracellular ligands to intracellular activities for effective cellular reactions. Several mechanisms of GPCR signaling include G-protein activation of secondary messengers, β-arrestin recruitment, receptor internalization, kinase phosphorylation, and subsequent changes in transcription. Recent advances in knowledge about GPCRs have made it clear that GPCRs act as allosteric-sensitivities ligands that can adopt different shapes according to their ligands [1]. The

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.38159

ability of GPCRs to exhibit various shapes is a characteristic of biased signaling, which means that the same agonist activates only one of several possible pathways mediated by the GPCR [2].

The biased signals such as β-2adrenergic receptor that is known to utilize Gs and β-arrestins as well as ERK1/2phosphorylation to signal through the cAMP pathway. At that time, β-arrestin was considered responsible for the desensitization and internalisation of receptors, but it appears now that it's also participating in signaling and regulating the effects of drugs, see reference [3]. Thus, does this ligand show Gs-bias or β-arrestin bias, or both?

In the old fashioned way of biased signaling analysis it mostly depends on separate curves of concentration response and pathway ratio which was done manually. But pharmacodynamic response is many-modeled, timed-based, and has assay noise as well as cell harm and pathway interference. A study building a data fusion model of many ways of β2-ARbiased pathway finding. Four kinds of data modalities are generated from the cAMP accumulation, β-arrestin2 recruitment, ERK1/2 phosphorylation, and cell viability test. Fuses curve parameters and time-response features into neural network classifier. Objective: To enhance the pathway identification accuracy and offer understandable feature contribution outcomes for GPCR biased signaling investigation.

2. Literature review

2.1. GPCR biased signaling

The term GPCR biased signaling refers to the selective stimulation of certain pathways in response to particular ligands bound to the same receptor. In case of the β2-adrenergic receptors, the possible signaling outcomes include cAMP formation through the Gs protein, binding of β-arrestin2, and phosphorylation of ERK1/2. Recent studies show that the bias in the selected pathways is controlled by the ligand structure, receptor configuration, involvement of GRK, engagement of β-arrestins, and cell type [4, 5]. It is clear that bias analysis requires comparison of the responses of several signaling cascades rather than the evaluation of one output pathway of the receptor activation.

2.2. Multimodal pharmacodynamic data

Multimodal pharmacodynamic data include different biological effects obtained from the same ligand-receptor interaction. During the study of GPCRs, the cAMP measurement test assesses the involvement of the Gs signaling pathway; the β-arrestin2 recruitment test assesses the involvement of the arrestin signaling pathway; the phosphorylation test measures ERK1/2 activation; and the viability test assesses any toxicity caused by increased concentration and, hence, affecting the responses observed. Recent studies conducted on the β2-adrenergic receptor show that Gαs, β-arrestin, and GRKs have an intricate relationship; hence, using one type of test to determine signaling bias is not enough [6].

2.3. Data fusion for pathway identification

Data fusion integrates pathway-specific pharmacodynamic features before pathway assignment. In this study, fused features include pEC50, Emax, response slope, area under the curve, time-point response, pathway ratio, and viability-adjusted response. Machine learning has already been applied to ligand-bias prediction and molecular property modeling, showing that fused representations can capture signal patterns that are difficult to detect with single descriptors [7, 8]. Multimodal deep learning further supports intermediate fusion, where modality-specific branches learn separate

DOI: 10.54254/2755-2721/2026.38159

representations before a shared classifier performs final prediction [9, 10]. This logic is suitable for GPCR pathway identification because each assay modality reflects a different but related signaling dimension.

3. Experimental methods

3.1. Receptor system and ligand dataset

The subject of investigation is the human beta 2-adrenergic receptor expressed via ADRB2 gene. Cells HEK293T stably transfected with ADRB2 were cultured with DMEM, 10% fetal bovine serum, 1% penicillin-streptomycin under 37°C, 5% CO_2 conditions. Cell-seeding for assay was performed at a seeding density of 20,000 cells per well 24 hours prior to ligand addition in 96-well plates. The described seeding density is used to keep stable levels of receptors and avoid overconfluency, obtaining similar values of luminescent or fluorescent signals in pathway assays.

Ligand set consists of 40 β2-adrenergic receptor ligands with isoproterenol acting as reference agonist. Ligands were tested in 8 concentrations, from 0.1 nM to 100 μM. Reference agonists' wells, vehicle wells, blank wells, and test ligand wells were included into the experiment design to avoid the normalization effect between plates. Identifiers of ligands were kept identical for all assays, so cAMP, β-arrestin2, ERK1/2, and viability data could be collected for each ligand at its corresponding concentrations. Wells that failed quality control (pipetting errors, abnormal background or absent signal) were marked before feature extraction.

3.2. Pharmacodynamic assays and data recording

The level of cAMP activation was monitored through the GloSensor cAMP Assay. Cells were preincubated with the GloSensor reagent before the ligands treatment, and then the luminescence was measured after 30 minutes post stimulation by a BioTek Synergy H1 microplate reader. This approach corresponds to Gs-dependent pathway of the signal transduction. The recruitment of β-arrestin2 was measured by means of NanoBiT complementation assay, when the levels of luminescence were detected after 5, 10, 20, and 30 minutes of the ligand stimulation. Phosphorylation of ERK1/2 was analyzed using phospho-ERK1/2 HTRF Assay; cells were stimulated with the compounds for 5 minutes, and afterwards lysed and read in plates. Cell viabilities were determined using CellTiter-Glo after 24 hours of stimulation in order to evaluate the potential toxic effect of the ligands at high concentrations.

The raw data were exported from the Gen5 software in .csv format for each ligand and modality. Blank wells were corrected and all responses were normalized to the maximum isoproterenol-induced activity on the corresponding plate. Concentration-response curves are fitted in GraphPad Prism 10 using a four-parameter logistic model as shown in Equation (1):

$$R_m(c) = Bottom_m + \frac{Top_m - Bottom_m}{1+10^{(LogEC50_m - log_{10}c)\times Hill_m}} \quad (1)$$

where $R_m(c)$ is the fitted response of modality m at concentration c, $Bottom_m$ and Top_m are the lower and upper asymptotes, $LogEC50_m$ is the logarithmic half-maximal effective concentration, and $Hill_m$ is the slope factor. This formula can be applied since the pharmacological response of GPCRs usually displays a sigmoid shape in relation to concentration. The values obtained from the curve include pEC50, Emax, slope, area under the curve, maximum response, and fitted response.

DOI: 10.54254/2755-2721/2026.38159

3.3. Data fusion and model construction

For this research work, we adopt Python version 3.11 for data cleaning, data fusion, and dataset generation. The assay observations are fused through ligand IDs, concentrations, replicate indexes, and modality IDs via the Pandas library. Missing data due to invalid wells are imputed based on the median of the same modality and the same concentration. We normalize all data points to eliminate the effect of large values in the assay outputs on the learning process. The fused feature matrix contains cAMP, β-arrestin2, ERK1/2, viability, and pathway ratio features. The labels include Gs-biased, β-arrestin-biased, or balanced signaling based on the normalized ratio.

We construct the multimodal neural network using PyTorch version 2.2. Our network includes four different input branches, which contain cAMP, β-arrestin2, ERK1/2, and viability feature sets, respectively. Each branch is composed of two dense layers with ReLU activation. All branch output vectors are combined in a fusion layer, which are further fed into a softmax classifier. The objective function for our model is defined as in Equation (2):

$$\mathscr{L} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{3} y_{ik} log(\widehat{p}ik) + \lambda \sum l = 1^{L} |W_l|_2^2 \tag{2}$$

where N is the number of samples, k represents the three bias classes, y_{ik} is the true one-hot label, $\widehat{p}ik$ is the predicted probability, W_l is the weight matrix of layer l, and λ is the L2 regularization coefficient. This function is correct for three-class pathway classification because the cross-entropy term optimizes softmax class probability, while the regularization term reduces overfitting in a small ligand dataset. SHAP is used after training to identify the most influential features contributing to bias classification.

4. Experimental process

4.1. Cell treatment and signal acquisition

Cells are seeded into 96-well plates 24 hours before ligand stimulation. On the test day, culture medium is replaced with assay buffer, and ligand working solutions are prepared according to the fixed eight-concentration gradient. Each plate contains reference agonist wells, blank wells, vehicle control wells, and test ligand wells. For the cAMP assay, cells are incubated with GloSensor reagent before ligand addition, and luminescence is recorded after 30 minutes. For the β-arrestin2 assay, luminescence is recorded at four time points. For the ERK1/2 assay, cells are stimulated for 5 minutes and then lysed. For the viability assay, CellTiter-Glo reagent is added after 24 hours of ligand exposure.

4.2. Data preprocessing and bias label construction

Raw signals are corrected by subtracting blank-well values. Each pathway response is normalized against the maximal isoproterenol response on the same plate. Concentration-response parameters are exported from GraphPad Prism 10. The processed file includes pEC50, Emax, curve slope, fitted response value, time-point response, AUC, and viability-adjusted response for each ligand and pathway. Bias labels are constructed from pathway response ratios. A ligand is labeled as Gs-biased when the normalized cAMP response exceeds the β-arrestin2 response by the fixed threshold. A ligand is labeled as β-arrestin-biased when β-arrestin2 recruitment exceeds the cAMP response by

DOI: 10.54254/2755-2721/2026.38159

the same threshold. Ligands without clear pathway dominance are labeled as balanced. The final labeled dataset is saved as gpcr_bias_dataset.csv.

4.3. Model training and evaluation

The dataset is split into three subsets with a ratio of 70%, 15%, and 15% for training, validation, and test datasets, respectively. A random seed value of 42 is used for consistency throughout the experiments. The multimodal neural network model is trained for 200 epochs using the Adam optimization algorithm, cross entropy as the loss function, learning rate of 0.001, and a batch size of 8. Performance metrics include accuracy, precision, recall, F1 score, macro averaged area under the receiver operating characteristics curve, and confusion matrices. An additional comparison model, namely random forests, is trained using the same concatenation of features. Analysis through SHAP is performed on neural network results to determine which features have the most impact on classification.

5. Experimental results

5.1. Multimodal pharmacodynamic feature distribution

The dataset contained 40 ligands, 8 concentrations, 3 replicates, and 4 pharmacodynamic modalities, generating 3,840 raw assay records before feature aggregation. After quality control, 3,712 valid records were retained (see Table 1). cAMP responses showed the highest dynamic range, while β-arrestin2 recruitment showed stronger time-dependent variation. ERK1/2 phosphorylation provided intermediate downstream information, and viability adjustment reduced false pathway dominance at 50-100 μM.

Table 1. Multimodal pharmacodynamic feature distribution

Modality / feature group	Valid records	pEC50	Emax	AUC	Signal variability
cAMP accumulation	944	7.36±0.41	0.94±0.13	5.82±0.46	0.118±0.026
β-arrestin2 recruitment	928	6.91±0.52	0.71±0.18	4.63±0.57	0.146±0.031
ERK1/2 phosphorylation	936	6.74±0.47	0.68±0.16	4.21±0.49	0.139±0.028
Cell viability	904	—	0.93±0.09	6.14±0.38	0.074±0.019
Fused pathway-ratio features	40	0.42±0.17	0.31±0.14	0.58±0.21	0.096±0.024*

Note: *$p<0.05$ for reduced variability after multimodal normalization.

5.2. Biased signaling pathway classification performance

The multimodal neural network achieved higher pathway identification performance than single-modality and random forest baselines. The test accuracy reached 0.887±0.034, with macro-F1 of 0.872±0.041 and macro-AUC of 0.931±0.026. cAMP-only classification confused several balanced ligands with Gs-biased ligands, while β-arrestin2-only classification missed ligands with weak but temporally stable recruitment (see Figure 1). Multimodal fusion reduced these errors by combining concentration-response and time-response features.

DOI: 10.54254/2755-2721/2026.38159

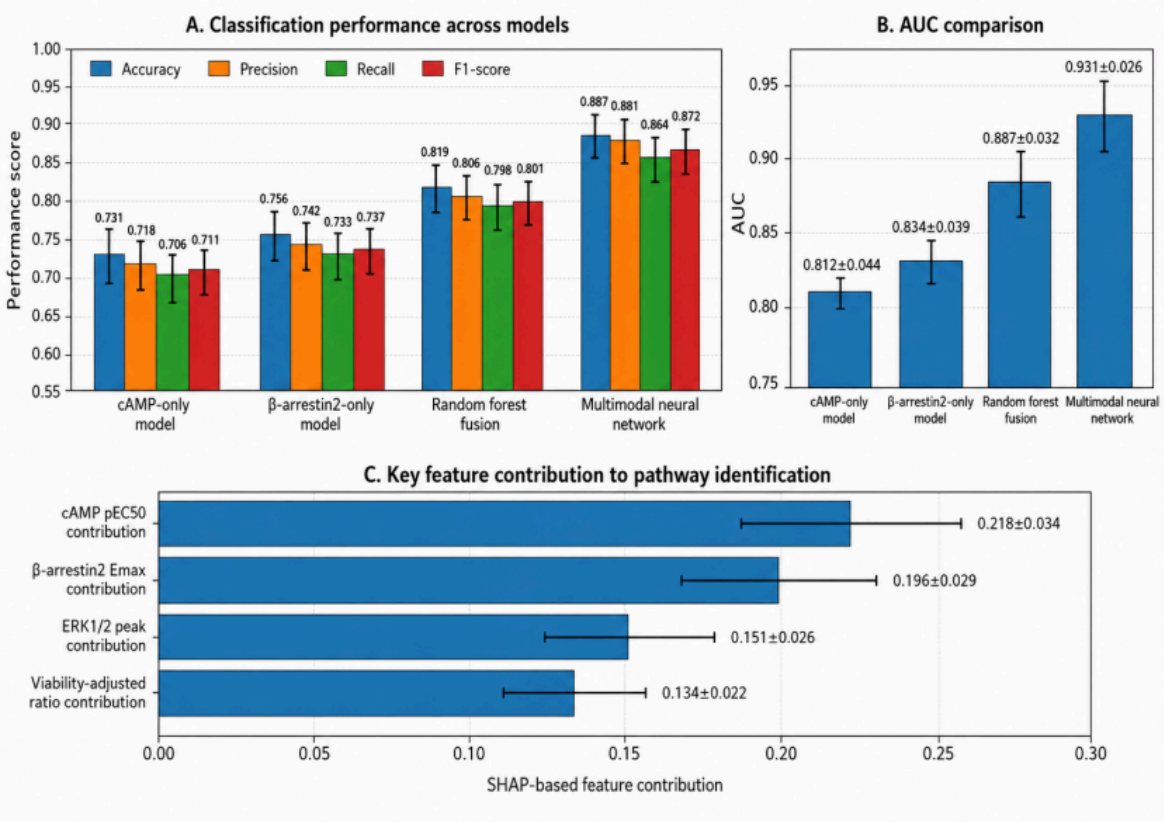

Figure 1. Biased signaling pathway classification and feature contribution results

5.3. Feature contribution and model interpretation

SHAP results indicated that cAMP pEC50, β-arrestin2 Emax, ERK1/2 peak response, and viability-adjusted pathway ratio were the most important variables for pathway assignment. Gs-biased ligands showed higher normalized cAMP potency and stronger cAMP AUC. β-arrestin-biased ligands showed delayed but sustained β-arrestin2 recruitment. Balanced ligands showed smaller pathway-ratio differences and lower SHAP separation values. These findings suggest that pathway identification depends on integrated pharmacodynamic patterns rather than one endpoint response.

6. Conclusion

This study establishes a multimodal pharmacodynamic data fusion model for identifying β2-adrenergic receptor biased signaling pathways. Four assay modalities were integrated, including cAMP accumulation, β-arrestin2 recruitment, ERK1/2 phosphorylation, and cell viability. The multimodal neural network outperformed single-modality and random forest baselines in pathway classification. Results show that pathway bias is best identified through fused potency, efficacy, temporal response, downstream kinase activation, and viability-adjusted ratio features. SHAP interpretation further confirms that cAMP pEC50 and β-arrestin2 Emax are the most influential indicators, while ERK1/2 and viability features improve classification reliability. The proposed model provides a practical framework for GPCR ligand screening and can be extended to other receptor systems after receptor-specific assay redesign and larger ligand validation.

DOI: 10.54254/2755-2721/2026.38159

References

[1] Zhang, M., Zhang, H., Feng, Z., & Liu, Y. (2024). G protein-coupled receptors (GPCRs): Advances in structures, mechanisms and drug discovery. *Signal Transduction and Targeted Therapy, 9*, 88.
[2] Kenakin, T. (2024). Bias translation: The final frontier? *British Journal of Pharmacology, 181*(13), 2375–2388.
[3] Kee, T. R., Kang, J., & Kim, K. M. (2024). The multifaceted functions of β-arrestins and their therapeutic potential in neurodegenerative diseases. *Experimental & Molecular Medicine, 56*, 147–161.
[4] Jiang, Q., & Dwivedi-Agnihotri, H. (2025). How ligands achieve biased signaling toward arrestins. *Biochemistry, 64*(2), 198–214.
[5] Matthees, E. S. F., Filor, J. C., Jaiswal, N., et al. (2024). GRK specificity and Gβγ dependency determines the potential of a GPCR for arrestin-biased agonism. *Communications Biology, 7*, 802.
[6] Burghi, V., Lambert, N. A., & Gutkind, J. S. (2023). Gαs is dispensable for β-arrestin coupling but dictates GRK selectivity and is predominant for gene expression regulation by β2-adrenergic receptor. *Journal of Biological Chemistry, 299*, 105293.
[7] Sanchez, J. E., KC, G. B., Franco, J., Allen, W. J., Garcia, J. D., & Sirimulla, S. (2021). BiasNet: A model to predict ligand bias toward GPCR signaling. *Journal of Chemical Information and Modeling, 61*(9), 4190–4199.
[8] Lu, X., Xie, L., Xu, L., Mao, R., Chang, S., & Xu, X. (2024). Multimodal fused deep learning for drug property prediction. *Briefings in Bioinformatics, 25*(3), bbae171.
[9] Cao, B., Xia, Y., Shen, Y., & Zhang, C. (2024). Predictive dynamic fusion. *Proceedings of the 41st International Conference on Machine Learning, 235*, 6087–6107.
[10] Rollins, Z. A., Cheng, A., & Warren, G. L. (2024). MolPROP: Molecular property prediction with multimodal language and graph learning. *Journal of Cheminformatics, 16*, 57.