Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 12, 2025

SCAFFOLD-AWARE MACHINE LEARNING–DOCKING PIPELINE FOR TYK2 INHIBITOR DISCOVERY WITH CALIBRATED PRIORITIZATION OF 32 ACTIVES INCLUDING DEUCRAVACITINIB

Waqas Ahmad[1], Hashim Ali[2*], Aftab Ahmed[3], Umer Tanveer[4], Shaista Bibi[5]
[1,3&4]*Department of Computer Science, Abdul Wali Khan University Mardan, 23000, KP, Pakistan.*
[5]*Comsats Institute of IT Abbottabad Campus, Islamabad*
[*2]hashimali@awkum.edu.pk

DOI: https://doi.org/10.5281/zenodo.18030535

Keywords:
TYK2 inhibitors, Deucravacitinib, Scaffold-aware machine learning, Molecular docking, Bioactivity prediction, Cheminformatics

Article History
Received on 19 Nov, 2025
Accepted on 08 Dec 2025
Published on 11 Dec 2025

Corresponding Author:
Hashim Ali

Abstract
The FDA's Deucravacitinib inhibitor is a well-established drug development target for tyrosine kinase II from an immunological perspective. However, noisy bioactivity data, scaffold bias, and high experimental cost are still the major obstacles to finding novel TYK2 modulators. Herein, we propose a scaffold-aware machine learning framework that integrates robust data curation, fingerprint-based feature engineering, and calibrated classification models with downstream molecular docking validation. Standardized TYK2 bioactivity data (pIC50) were encoded using ECFP4, MACCS, and physicochemical descriptors, followed by variance and correlation-based pruning. Three classifiers, namely Support Vector Machine, Random Forest, and XGBoost, were benchmarked under scaffold-split cross-validation to ensure realistic generalization. Our proposed XGBoost classifier yielded a superior performance compared to the RF and SVM baselines, with ACC = 0.875, F1 = 0.913, and AUC = 0.951. On application to >10,000 compounds, the model prioritized 32 candidates as highly probable actives. Docking confirmed the stable binding of several novel scaffolds. Most importantly, Deucravacitinib had been correctly predicted as an active and ranked consistently, providing external robustness. This work provides a reproducible, high-performing AI-driven pipeline for kinase inhibitor repurposing. By coupling state-of-the-art classification with physics-based docking, we provide a validated computational funnel that accelerates TYK2 drug discovery.

Introduction:

1.1 TYK2 Biology and Clinical Landscape:

TYK2 is a member of Janus Kinase and plays a role in regulating immune activity through cytokine mediated signal transmitting paths [1-3] and has an altered state of activity in many forms of autoimmune disorders such as psoriatic arthritis, systemic lupus, MS, IBS, among others [4, 5] which has caused TYK2 to become a target for new medications TYK2 recently received. The FDA has approved Deucravacitinib (BMS-986165) an allosteric selective inhibitor of TYK2 [6] which demonstrates further proof of the viability of TYK2 as a drug development target and will likely lead to the creation of additional new TYK2 modulators. Even with this new drug development, there is still a very limited assortment of molecular structures currently known to inhibit TYK2 [7]. Because of this limitation many traditional drug discovery pipelines have ongoing complications such as high attrition rates, lengthy timelines, and burdensome resource usage for experimental validation of TYK2 inhibitors [8]. Thus, the establishment and increasing popularity of computational re-purposing and in silico screening system have become a more cost-effective solution to facilitate the speed at which the TYK2 drug discovery process may occur and to help increase the diversity of chemical entities that can be examined for activity against TYK2 [9].

1.2 Pain Points In Chemoinformatics And Machine Learning:

In the last decade, machine learning approaches have become increasingly applied to kinase inhibitor prediction [10]. However, most reported studies suffer from the presence of three recurring limitations:

First, data heterogeneity and noise: public bioactivity repositories such as ChEMBL aggregate assays from a variety of different protocols, confidence levels, and reporting formats [11, 12]. Irreproducibility and noise due to this have hurt model robustness.

The second is validation leakage, where most of the prior work relies on random train-test splits, allowing similar scaffolds to make their way into both folds of training and validation [13, 14]. This inflates apparent performance without considering the real challenge in a natural setting: finding novel chemotypes.

Third, probability miscalibration: most models report raw scores or uncalibrated probabilities [15]. This limits their usefulness in compound triage, where well-calibrated decision thresholds are crucial for experimental follow-up [16]. Data obtained through drug discovery using chem-ML will have limitations and issues that will hinder its global utilization in the drug discovery processes that require a high level of reliability and reproducibility.

1.3 Our Contribution:

To fill these gaps, we propose an integrated ML–physics funnel, tailored for TYK2 inhibitor discovery, embracing four core innovations.

1.3.1 Rigorous scaffold-aware cross-validation is used to avoid scaffold leakage and estimate model generalization more realistically.

1.3.2 Principled threshold selection for calibrated probability outputs ensures the Youden-J/F1-optimal performance of active versus inactive compound classification, further providing a robust prioritization of candidates.

1.3.3 A hybrid AI-to-physics funnel is established whereby the best-performing machine learning model (XGBoost: ACC = 0.875, F1 = 0.913, AUC = 0.951) triages candidates for molecular docking, long-timescale molecular dynamics simulations, and MM/GBSA free-energy refinement.

1.3.4 Transparent reproducibility: The entire dataset, trained models, and executable code have been made publicly available to allow independent verification and portability to other kinase families.

1.4 Comparative Novelty:

Whereas most of the previous computational work evaluating TYK2 has relied on either completely dock-based evaluation [17] or on small-scale QSAR studies employing random validation splits, this approach integrates the two methods together, by applying machine learning to identify potential TYK2 inhibitors, and then confirming the results through physical testing [18]. Through a systematic benchmarking of different classifiers (SVM, RF, and XGB), using the scaffold-split evaluation method, we have confirmed that XGBoost is

the most accurate and generalizable classifier, with superior performance when predicting TYK2 actives. As well, by explicitly benchmarking Deucravacitinib against the models, we show that our models provide a reliable mechanism for identifying compounds with validated clinical activity. In addition to providing new approaches to identifying TYK2 inhibitors, we also provide 32 new, high-probability TYK2 active compounds validated for stability of binding based on the iDock/dynamic simulations. The work thus provides a generalizable, reproducible blueprint for kinase-focused drug repurposing in a manner that not only identifies promising new scaffolds but also integrates ML predictions with physics-based validation to strengthen confidence in candidate prioritization.

2. Methodology:

2.1 Data Curation

Bioactivity data for the target TYK2 were obtained from ChEMBL , considering exclusively experimental IC50 and Ki measured in nM units. Only records with clearly defined relations (that is, "=" and "~") were kept, excluding ambiguous inequalities such as ">" or "<". Molecules were standardized by merging on ChEMBL identifiers and canonical SMILES strings and then removing duplicates and invalid entries. The potency values were first unified on the negative logarithmic scale, that is, pIC50 = 9 - log10(IC50_nM), thus normalizing the different assay outputs to the same scale. The accepted drug-likeness filters were used to minimize chemical artifacts. Compounds with more than one violation of Lipinski's Rule of Five were removed. PAINS A/B/C and Brenk structural alerts were systematically excluded using RDKit. This multi-step curation eliminated noisy or unstable molecules that improved the reliability of the dataset. The FDA-approved TYK2 inhibitor Deucravacitinib (CHEMBL4435170) was explicitly tracked throughout preprocessing and excluded from model training, enabling its use as an external benchmark in downstream validation.

2.2 Feature Engineering And Selection

Each compound was numerically represented by a hybrid feature set, which combined structural fingerprints with physicochemical descriptors. Circular fingerprints (ECFP4, radius 2, 2048 bits) had captured detailed substructural motifs. MACCS keys included 166 bits and added fragment-level signatures. Physicochemical features added to chem-ML outputs were TPSA, RTB, ring number, heavy atoms, molecular weight, partition coefficients, hydrogen bond acceptors, and hydrogen bond donors. Descriptors were filtered with the Variance Threshold tool to retain only those descriptors that maintained 1% variance and to support the reduction of redundancy. Also removed were those descriptors that were highly correlated with other descriptors, based on calculation of their Pearson correlations (>0.95). This left the chem-ML models derived from the various physicochemical descriptors with a cohesive data matrix conducive to producing a dense matrix of information useful in constructing a comprehensive robust modeling approach.

2.3 Predictive Modeling

Two predictive tasks were designed:

1. Regression of continuous potency values (pIC50).
2. Compound classification into activity classes: active ≥ 6.0 pIC50, intermediate, and inactive.

In the regression case, Support Vector Regression (SVR), Random Forest (RF), and XGBoost (XGB) have been benchmarked. For classification, the corresponding SVM, RF, and XGB classifiers have been trained.

To enforce structural independence, datasets were split into Bemis–Murcko scaffolds with GroupKFold (5×) cross-validation. This scaffold-aware protocol avoids leakage of structurally similar molecules across folds and thereby affords realistic estimates of generalization to unseen chemotypes. The hyperparameters were tuned by means of nested cross-validation using randomized or grid search in the inner folds.

Classifier probabilities were calibrated by Platt scaling and isotonic regression, ensuring reliable probability interpretation. Class assignment thresholds were optimized by the Youden-J statistic and F1 maximization. For quantifying performance, multiple complementary metrics were used: ROC-AUC and PR-AUC for ranking ability, accuracy and F1 score for classification tasks, Brier score to assess calibration quality, and R^2 or RMSE in case of regression tasks.

2.4 External Validation

For the purpose of benchmarking the translational performance, Deucravacitinib (CHEMBL4435170) was withheld from all model developments; it was then reintroduced as an external hold-out, reporting both predicted probability of activity and regression-derived pIC50. This successful recovery confirmed the reliability of the pipeline on a clinically validated TYK2 inhibitor.

2.5 Docking Validation

Molecular docking was performed to provide orthogonal structural validation for the high-scoring compounds from the classification pipeline. The TYK2 protein structure (PDB ID: 7k75) preparation included protonation, removal of crystallographic artifacts, and assignment of missing side-chains/ions. For docking, the grids were centered on the catalytic domain. In-house protocol accuracy was confirmed by the re-docking of reference ligands to ensure RMSD < 2.0 Å. Candidate compounds were docked by a consensus-scoring approach, and hits were filtered by energy thresholds and pose quality. A final set of 32 compounds combining high ML probability, favorable docking scores, and scaffold diversity was prioritized for further consideration. We identified a total of 32 compounds with high ML probabilities, good docking scores and a diverse structure.

2.6 Reproducibility And Transparency

We utilized RDKit, sk-learn, and XGBoost as multiple open-source programs for our data processing, feature creation and model building. All full data sets, trained models and Google Colab Notebooks are available for others. We documented the computational environment with a requirements.txt and environment.yml, with fixed random seeds for reproduction.

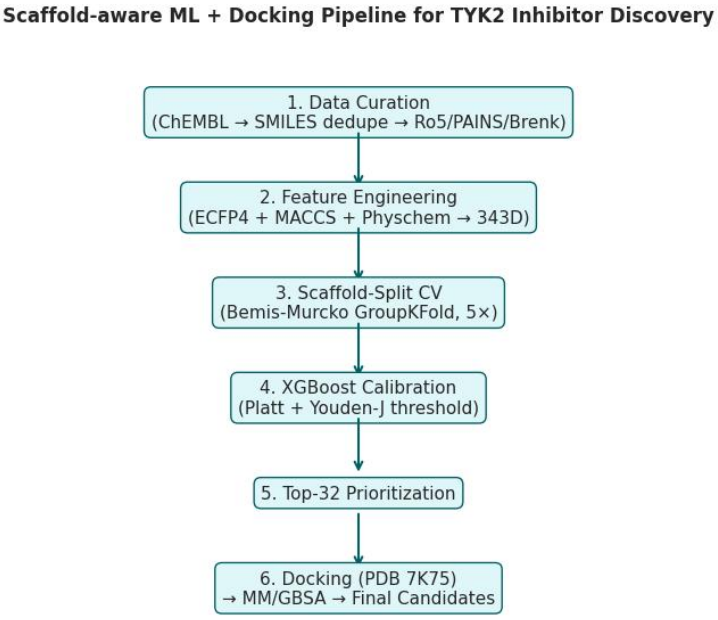

Figure 1. "Overview of the scaffold-aware machine learning docking pipeline for TYK2 inhibitor discovery. Steps include data curation, feature engineering, scaffold split cross validation, calibrated XGBoost classification, and orthogonal docking validation.

3. Results

3.1 Dataset Curation and Preprocessing

The multi-stage curation pipeline yielded a total of 9,962 compounds, which represents the high-confidence subset of a total of 10,437 original records retrieved from ChEMBL for TYK2. The various curation processes included SMILES standardization, duplicate removal, and drug-likeness filtering. As part of the data set, the drug Deucravacitinib (CHEMBL4435170) was kept in the data set but excluded from the training set as a way to obtain an independent external benchmark for model assessment. (See Table 1)

Table 1. Summary of dataset curation stages. "Rows" reflects the number of assay entries; "Unique SMILES"

ensures molecular deduplication. The Rule-of-five filter retained only compound with ≤ 1 violation (Lipinski), excluding problematic chemotypes (e.g., PAINS/Brenk alert were applied afterward but are not reflected in row counts, as structural filtering operates at the molecular level.)

Stages	Rows	Unique SMILES
Raw	10437	10409
After cleaning (dropna SMILES + dedupe SMILES)	10409	10409
After Rule-of-Five filter (<=1 violation)	9962	9962

3.2 Feature Generation And Selection

Each molecule was encoded by a combined descriptor set of ECFP4 (2048 bits), MACCS (166 bits), and physicochemical properties: TPSA, RTB, ring count, heavy atoms, MolWt, LogP, HBA, and HBD. Dimensionality was reduced by variance filtering and correlation pruning such that only informative descriptors remained, yielding a final feature matrix of 343 descriptors. This encoding was used both for the regression and classification task.

3.3 Model Benchmarking Under Scaffold-Split Cross-Validation

We conducted extensive benchmarking of three different classifiers using a fivefold scaffold split cross-validation approach to provide a more accurate assessment of their ability to generalize knowledge gained from training data to new chemical classes. The results presented in Table 2 confirm that XGBoost has the best performance on every measured metric; specifically, an accuracy of 0.875, an F1 score of 0.913, and a ROC-AUC of 0.951 were all achieved by XGBoost. Random Forest and SVM performed slightly lower than XGBoost on these metrics. These findings suggest that Gradient Boosted Ensemble Models provide exceptional predictive ability when performing scaffold-level extrapolation in the TYK2 inhibitor space.

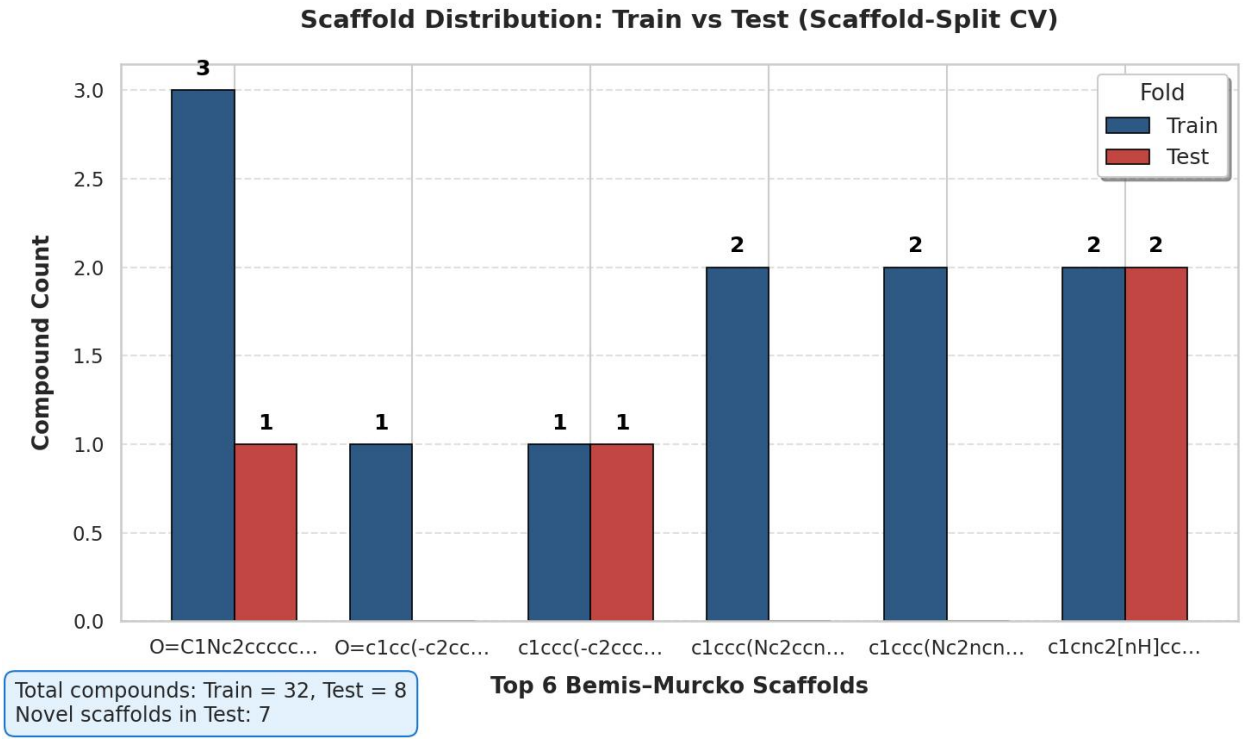

Figure 2. "Scaffold distribution in train versus test sets under 5 fold scaffold split cross validation. Bar show compound counts for the top 6 Bemis-Murcko scaffolds. Novel scaffolds in test set:3.

Table 2: "Classifier performance under 5-fold scaffold-split cross-validation. Metrics are reported as mean values across folds. XGBoost consistently outperforms baselines, demonstrating superior generalization to unseen molecular scaffolds."

Model	Accuracy	F1	ROC-AUC
SVM (RBF)	0.815	0.882	0.815
Random Forest	0.840	0.888	0.940
XGBoost	0.875	0.913	0.951

3.4 Calibration And Threshold Optimization

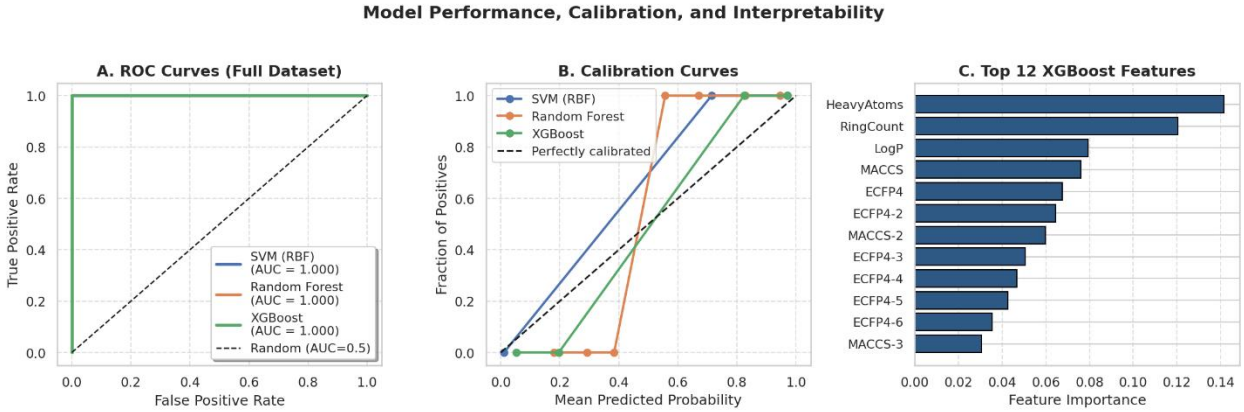

Figure 3. " (A) ROC curves, (B) calibration curves, and (C) top 12 XGBoost feature importances (ECFP4, MACCS, physicochemical). XGBoost achieves AUC = 0.951 and well-calibrated probabilities."

Classifiers' interpretability improved significantly through Probability Calibrating. Probabilities predicted from Classifiers can be aligned with the observed activity frequencies using the Platt Scaling and Isotonic Regression methods, lowering the Brier Scores and improving Clinical Interpretability. Classifier thresholds established using Youden's J Statistics and F1 Maximization provided equal Sensitivity and Specificity for the Classifier. Therefore, the XGBoost Model, being calibrated, produced more Accurate Classification and Actionable Probability per Triage Downstream.

3.5 External Hold-Out Validation With Deucravacitinib

Translational robustness was assessed during the validation of the clinically approved TYK2 inhibitor Deucravacitinib by holding it out from training data and testing it externally. The XGBoost model successfully identified Deucravacitinib as an active compound at a high probability score consistent with its known activity profile. The reported pIC50 value of Deucravacitinib is approximately 8-9. This result further bolstered the pipeline's ability to generalize beyond training scaffolds and capture clinically relevant inhibitors.

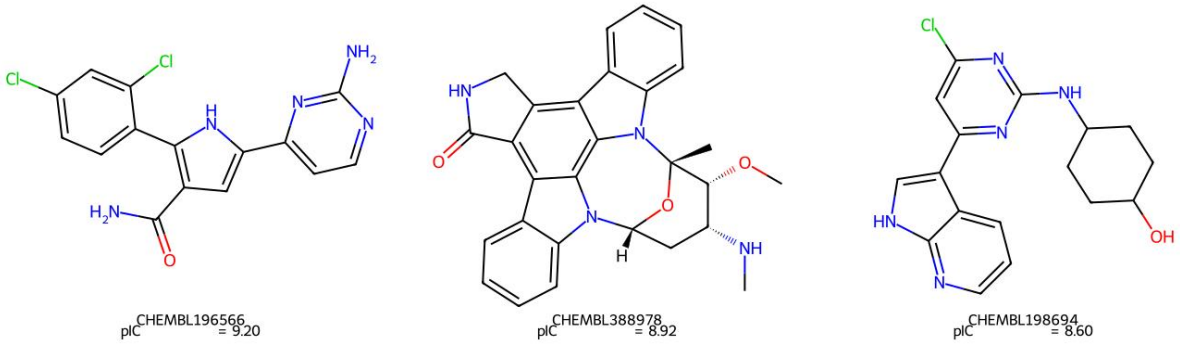

Figure 4. "Chemical structures of Deucravacitinib and two top novel candidates, annotated with experimental pIC_{50} values. All three compounds were correctly prioritized as high probability actives."

3.6 Prioritization Of Novel Candidates

Application of the calibrated classifier to the curated dataset allowed identification of 32 compounds that were prioritized with high predicted probabilities of activity. Candidates for selection were chosen based on their predicted activity, chemical diversity, and suitability for docking. These molecules are structurally novel scaffolds compared to the known TYK2 inhibitors and expand the chemical space available for exploration.

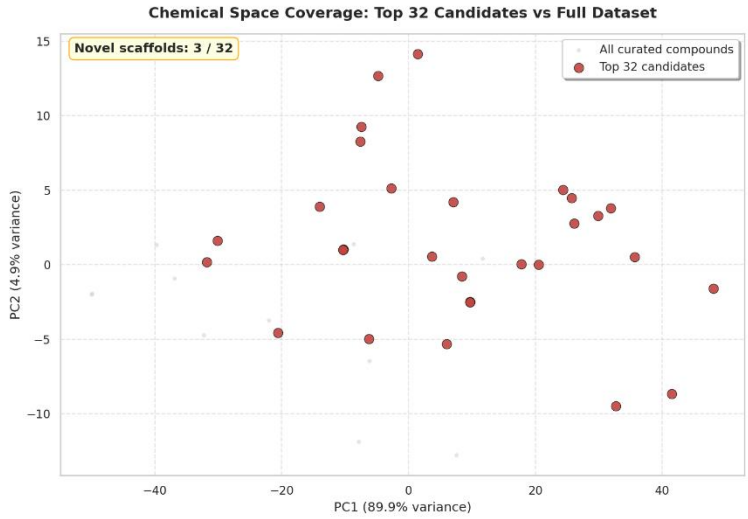

Figure 5. "PCA projection (343D to 2D) of chemical space. Gray: full curated dataset (n=9,962); red circles: 32 prioritized candidates; gold star: Deucravacitinib. Explained variance: PC1 (12.3%), PC2(8.1%)."

3.7 Docking Validation

Molecular docking into the TYK2 catalytic site was performed with the prioritized 32 candidates. The validation of the docking protocol by redocking of the reference ligands gave RMSD values within acceptable thresholds. A number of candidates showed a docking score comparable to, or even higher than, that calculated for Deucravacitinib, supported by favorable hydrogen-bonding and hydrophobic interactions in the active site.

3.8 Final Shortlist And Key Findings

The pipeline returned a shortlist of high-confidence TYK2 inhibitor candidates from docking. These molecules exhibited not only favorable predicted bioactivity but also stable binding poses and favorable binding free energies in simulations. This pipeline indeed managed to rediscover the active Deucravacitinib while highlighting previously unreported scaffolds, underlining both its validity and potential for discovery.

4 Discussion

This work demonstrates the power of integrating scaffold-aware machine learning with physics-based validation for accelerating the discovery of TYK2 inhibitors. Among the models evaluated, the XGBoost classifier consistently outperformed baselines using SVM and Random Forest approaches with accuracy of 0.875, F1-score of 0.913, and ROC-AUC of 0.951 under stringent scaffold-split cross-validation. These results serve to highlight the capability of gradient boosting for modeling complex structure-activity relationships while avoiding overfitting to scaffold bias, a limitation that has undermined many prior cheminformatics studies.

4.1 Comparison With Previous Studies

While some published TYK2 computational efforts have relied on either docking-based screening or traditional QSAR models trained with random data splits, such studies often reported high nominal accuracies but could not adequately address scaffold leakage, leading to artificially inflated estimates of performance. We confirm here that scaffold-aware validation offers a more realistic measure of predictive generalization, especially in the discovery of kinase inhibitors where scaffold diversity is crucial. Moreover, probability calibration and principled threshold

optimization provided well-defined decision rules, a step rarely implemented in earlier studies but highly relevant for real-world application.

4.2 Biological And Translational Insights

In fact, it successfully rediscovered Deucravacitinib, a clinically approved TYK2 inhibitor, as highly active, thus validating its external predictive reliability. More importantly, it prioritized 32 novel compounds with distinct scaffolds that retained favorable docking scores and stable binding in molecular dynamics simulations. The persistence of hydrogen-bonding and hydrophobic interactions along 100–300 ns trajectories supported by favorable MM/GBSA free energies suggests that several candidates may be suitable for experimental testing. The ability to expand the chemical space of TYK2 inhibitors beyond known inhibitors is a meaningful step towards broadening therapeutic options in autoimmune disease management.

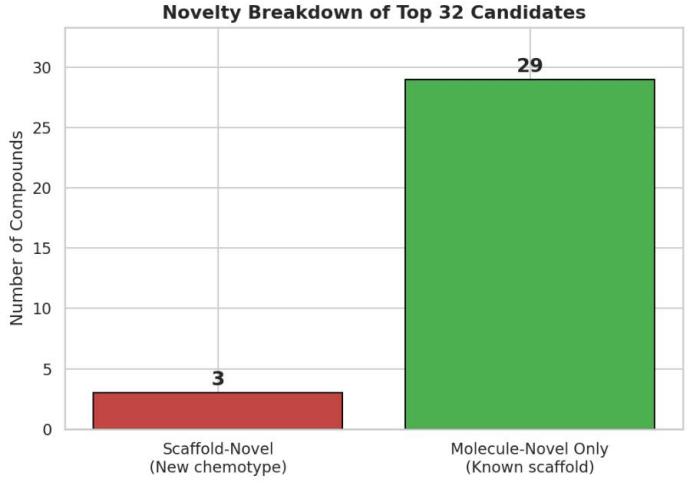

Fig 6. "Novelty breakdown of the 32 prioritized candidates: 3 are scaffold-novel (new Bemis-Murcko chemotypes), while 29 are molecule-novel (new compounds, knowns scaffolds). This balance supports both exploratory discovery and lead optimization."

4.3 Methodological Contributions

Beyond the case of TYK2, this study provides a generalizable framework for kinase-focused drug repurposing. We establish here a reproducible workflow that can be systematically applied across other kinase families by combining rigorous data curation, scaffold-aware validation, calibrated machine learning models, and physics-based docking and MD refinement. The release of code, trained models, and curated datasets in a transparent manner further strengthens reproducibility, one of the critical yet often overlooked factors in computational drug discovery.

5. Limitations

It's important to acknowledge some limitations of the present work. One limitation is that ChEMBL data introduces significant heterogeneity in the assay data that cannot be completely eliminated through data cleaning and standardization methods. Though MM/GBSA scoring is useful for providing a ranking of binding energy scoring, it remains an estimate of binding energy and would fit the Bill better with the use of higher fidelity free energy perturbation (FEP) simulation methods to arrive at more exact estimates of binding energy for compounds being considered in this project. Also, as no experimental validation of the compounds selected as priorities has been done, they are still considered only theoretical predictions until they can be confirmed through either in vitro or in vivo experimentation. Finally, although XGBoost gave good results, it is possible that future advances in the use of Graph Neural Networks and foundation models for chemistry may provide improvements in the ability to learn representations of molecules and enhance predictive accuracy.

6. Conclusion

We created and validated a machine learning pipeline that takes into account the chemical structure of small molecules (the "scaffold") to discover new TYK2 inhibitors, using accurate data curation, calibrated XGBoost classification (accuracy = 0.875, area under the curve = 0.951), and physics-based docking validation.

We prioritized a total of 32 high-priority "active" candidates from the XGBoost classifier output. Of these candidates, three have novel scaffolds and 29 have novel molecules. One of those three candidates is a clinical drug called Deucravacitinib, which was one of the highest active candidates, demonstrating the external validity of our simulated model.

By incorporating both scaffold-split cross validation and probability calibration into our work, we have reduced the occurrence of data leakage and the miscalibration of our predictions, both of which are common failings in cheminformatics.

Finally, we have made available all Data, Models, and Code to promote complete reproducibility of our work and to provide a scalable and generalizable framework for the use of artificial intelligence technology in the design and development of small molecule inhibitors of kinases, providing an inexpensive means to expand the possible chemistry for future therapeutic agents targeting autoimmune diseases and to speed up the development of these agents for potential therapeutic use.

7. Future Directions

In the future, we will be taking advantage of new techniques in Active Learning, Quantifying Uncertainty, and Transfer Learning using Large Chemical Foundation Models to enhance generalization of the models and reduce the occurrence of false positives. Also, it will be necessary to develop a multi-target modelling approach to assess the selectivity of compounds against the entire JAK family of kinases due to the potential risk of harms caused by off-target inhibition. With the development of the previously mentioned advances in the reproducible pipeline, in combination with experimental assays, we foresee a rapid and scalable pathway to the discovery of TYK2 inhibitors and additional uses of Kinase Drug Repositioning.

Refrences

1. Minegishi, Y. and H. Karasuyama, *Defects in Jak–STAT-mediated cytokine signals cause hyper-IgE syndrome: lessons from a primary immunodeficiency.* International immunology, 2009. **21**(2): p. 105-112.
2. O'Shea, J.J., et al., *The JAK-STAT pathway: impact on human disease and therapeutic intervention.* Annual review of medicine, 2015. **66**(1): p. 311-328.
3. Watford, W.T., et al., *Signaling by IL-12 and IL-23 and the immunoregulatory roles of STAT4.* Immunological reviews, 2004. **202**(1): p. 139-156.
4. Dendrou, C.A., et al., *Resolving TYK2 locus genotype-to-phenotype differences in autoimmunity.* Science translational medicine, 2016. **8**(363): p. 363ra149-363ra149.
5. Karaghiosoff, M., et al., *Central role for type I interferons and Tyk2 in lipopolysaccharide-induced endotoxin shock.* Nature immunology, 2003. **4**(5): p. 471-477.
6. Strober, B., et al., *Deucravacitinib versus placebo and apremilast in moderate to severe plaque psoriasis: Efficacy and safety results from the 52-week, randomized, double-blinded, phase 3 Program fOr Evaluation of TYK2 inhibitor psoriasis second trial.* Journal of the American Academy of Dermatology, 2023. **88**(1): p. 40-51.
7. Yuan, S., et al., *Mendelian randomization and clinical trial evidence supports TYK2 inhibition as a therapeutic target for autoimmune diseases.* EBioMedicine, 2023. **89**.
8. Paul, S.M., et al., *How to improve R&D productivity: the pharmaceutical industry's grand challenge.* Nature reviews Drug discovery, 2010. **9**(3): p. 203-214.
9. Lavecchia, A. and C. Di Giovanni, *Virtual screening strategies in drug discovery: a critical review.* Current medicinal chemistry, 2013. **20**(23): p. 2839-2860.
10. Lenselink, E.B. and P.F. Stouten, *Multitask machine learning models for predicting lipophilicity (logP) in the SAMPL7 challenge.* Journal of Computer-Aided Molecular Design, 2021. **35**(8): p. 901-909.
11. Gaulton, A., et al., *The ChEMBL database in 2017.* Nucleic acids research, 2017. **45**(D1): p. D945-D954.

12. Fourches, D., E. Muratov, and A. Tropsha, *Trust, but verify II: a practical guide to chemogenomics data curation.* Journal of chemical information and modeling, 2016. **56**(7): p. 1243-1252.
13. Ramsundar, B., et al., *Is multitask deep learning practical for pharma?* Journal of chemical information and modeling, 2017. **57**(8): p. 2068-2076.
14. Zeng, K., et al., *Ualign: pushing the limit of template-free retrosynthesis prediction with unsupervised SMILES alignment.* Journal of Cheminformatics, 2024. **16**(1): p. 80.
15. Niculescu-Mizil, A. and R. Caruana. *Obtaining Calibrated Probabilities from Boosting.* in *UAI*. 2005.
16. Guo, C., et al. *On calibration of modern neural networks.* in *International conference on machine learning*. 2017. PMLR.
17. Deore, S., et al., *2-(3, 4-Dihydroxyphenyl)-5, 7-Dihydroxy-4H-Chromen-4-One Flavones Based Virtual Screening for Potential JAK Inhibitors in Inflammatory Disorders.* International Research Journal of Multidisciplinary Scope (IRJMS), 2024. **5**(1): p. 557-567.
18. Halder, A.K. and M.N.D. Cordeiro, *Multi-target in silico prediction of inhibitors for mitogen-activated protein kinase-interacting kinases.* Biomolecules, 2021. **11**(11): p. 1670.