Proceedings of ICMMGH 2026 Symposium: Biomedical Imaging and AI Applications in Neurorehabilitation
DOI: 10.54254/2753-8818/2026.AU29729

Virtual Screening of Potential NF-κB Inducing Kinase Inhibitors Based on Machine Learning and Molecular Docking

Zhiyuan Yang

School of Life Science and Technology, China Pharmaceutical University, Nanjing, China
yangzy@stu.cpu.edu.cn

Abstract. NF-κB inducing kinase (NIK) orchestrates the noncanonical signal pathway and is implicated in chronic inflammation and cancer, making it a compelling yet underexploited therapeutic target. Currently, no NIK-targeted drugs have been approved and traditional discovery routes remain slow and costly, underscoring the need for efficient virtual screening strategies. This study develops an integrated workflow that couples physicochemical prefiltering with machine-learning activity prediction and structure-based docking to rapidly triage large compound library. The two core screening mechanisms in the final pipeline are respectively, pIC_{50} prediction with the best-performing gradient boosting regression model, and molecular docking with key amino acid residues analysis from PLIT and FTmap. Ultimately, 18 promising molecules with potential inhibitory activity against the NIK were selected from a candidate compound library containing 832,335 molecules, all of which were validated to possess novel chemical structures. The pipeline furnishes a rapid, data-driven route to structurally novel NIK chemotypes. Beyond NIK inhibitors, this framework provides a practical foundation for multi-objective filtration toward selective, developable molecules, compressing discovery timelines while preserving scaffold novelty, offering a reusable blueprint for modern hit identification.

Keywords: virtual screening, machine learning, molecular docking, NF-κB inducing kinase, inhibitor

1. Introduction

In eukaryotic cells, nuclear factor κB (NF-κB) regulates gene transcription as a family of transcription factors. In resting cells, inhibitory IκB proteins associate with NF-κB dimers, thereby making them sequestered in the cytoplasm and inactive. Signals from receptors such as CD40, BAFF, and RANK activate NF-κB via the noncanonical pathway: NF-κB inducing kinase (NIK) cooperates with IKKα to drive the processing of p100 (NF-κB2 precursor) into p52, permitting p52/RelB heterodimers to be transported from the cytoplasm to the nucleus [1]. NF-κB governs the expression of numerous immune- and inflammation-related genes. If the dysregulation persists, its signaling pathway is tightly associated with diverse disease processes. Excessive NF-κB activity precipitates chronic inflammation and tissue damage and can disrupt immune tolerance, thereby

© 2025 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: 10.54254/2753-8818/2026.AU29729

triggering autoimmunity [2]. In cancer, the initiation and progression of many solid and hematologic malignancies are promoted by abnormal NF-κB activation through up-regulating anti-apoptotic genes and fostering angiogenesis, proliferation, and migration; it also contributes to resistance to radiotherapy and chemotherapy [3].

Given the pathogenic roles of NF-κB, developing NF-κB inhibitors has become a major line of inquiry. Classical approaches include small molecules that target the IKK complex (e.g., BAY 11-7082) to block IκB phosphorylation and thus terminate NF-κB activation, as well as proteasome inhibitors (e.g., bortezomib) that indirectly suppress NF-κB by preventing IκB degradation—a mechanism already leveraged to treat multiple myeloma and related disorders [4, 5]. To improve target selectivity, recent efforts have shifted to key nodes in the noncanonical pathway. The hub kinase NIK is aberrantly active in multiple inflammatory diseases and cancers and is regarded as a principal driver of pathology [6]. Brightbill et al. developed a highly selective NIK inhibitor and validated it in NZB/W F1 mice, a classic model of systemic lupus erythematosus: the compound inhibited NIK catalytic activity in vitro and, in vivo, ameliorated renal inflammation, lowered cytokine and chemokine expression, reduced autoantibody titers, attenuated kidney injury, and extended survival, outperforming blockade of the BAFF pathway alone [7]. Through structure optimization, Li et al. obtained the novel inhibitor XT2, which shows nanomolar potency against NIK in vitro and confers marked protection in a carbon tetrachloride-induced acute liver injury model, decreasing serum ALT, limiting inflammatory cell infiltration, and suppressing the up-regulation of inflammation-related genes, thereby interrupting hepatic inflammatory pathology [8]. Continued medicinal chemistry refinement is advancing the clinical feasibility of NIK inhibition: using metabolite identification and structure-guided design, Crawford et al. improved systemic clearance and oral bioavailability while maintaining high potency, yielding compounds with more favorable pharmacokinetics [9]. Via a scaffold-hopping strategy, N. Zhang et al. re-engineered the known inhibitor B022 to obtain candidates with superior selectivity and stability that exhibited robust inhibitory activity in cellular and animal models [10]. To date, however, no NIK inhibitor—or any agent directly targeting NF-κB—has received regulatory approval, in part because broad NF-κB blockade risks immunodeficiency.

This study proposes an integrated screening framework that couples machine learning with molecular docking to overcome the long timelines and high costs of conventional laboratory workflows. Using known active compounds as training data, the author built and optimized six predictive machine-learning models to maximize their discriminative power for compound activity. The top-performing model was then used for stepwise triage of candidate molecules, followed by docking-based evaluation of their binding poses and binding energies to NIK, thereby progressively narrowing the chemical search space in silico. Through this multi-stage filtering pipeline, the author rapidly prioritized putative high-activity candidates from the large REAL compound library.

2. Materials

2.1. Compound libraries

Compounds with reported IC_{50} values and SMILES strings were collected from BindingDB (www.bindingdb.org/bind/index.jsp) and ChEMBL (www.ebi.ac.uk/chembl/). Records were processed in KNIME to remove duplicates and entries with missing fields; IC_{50} units were standardized from nmol/L to mol/L and then converted to pIC_{50}, yielding 1,410 unique molecules that comprised the NIK inhibitor activity dataset. The REAL (readily available for synthesis) database from Enamine (https://www.enamine-genez.com/compound_library/real-database/) is a

DOI: 10.54254/2753-8818/2026.AU29729

large virtual compound collection. For this study, the author selected three drug-discovery oriented sublibraries—Bioactive Libraries, Diversity Libraries, and Targeted Libraries—and, after duplication removal, obtained 832,335 compounds as the candidate library. The NIK activity dataset was then randomly partitioned 4:1 into a training set which contains 1,128 molecules and a testing set which contains 282 molecules.

2.2. NIK protein

Structures of MAP3K14 (NIK) were retrieved from the Protein Data Bank (PDB). The author included entries reported within the last 12 years that were solved by X-ray diffraction at a resolution better than 2.8 Å, yielding ten structures (PDB IDs: 6WPP, 6Z1Q, 6Z1T, 6MYN, 6G4Y, 5T8O, 5T8P, 5T8Q, 4IDT, and 4IDV). Among these, the following are co-crystal complexes with reported, high-potency small molecule inhibitors: 6WPP, 6MYN, 6G4Y, 5T8P, and 4IDT. In 5T8P and 6G4Y, 89% of residues show no outliers in the geometry assessment, and only 5% exhibit a single outlier, indicating superior model quality. Electron density validation identifies outliers for 4% of residues in 5T8P, versus 2% in 6G4Y. Accordingly, the NIK structure from the 6G4Y complex was chosen as the preferred receptor for subsequent analyses. Outlier refers to the abnormal items in the verification results of amino acid residues.

3. Methods

3.1. Virtual screening

Table 1. Results of structure verification and charge verification in the PDB X-ray validation reports [11]

PDB ID	Geometric Quality			Not Modelled	Electron Density Defect
	0 Outlier	1 Outlier	2 Outliers		
4IDT	67%	22%	< 5%	6%	2%
5T8P	89%	5%	0%	6%	4%
6G4Y	89%	5%	0%	6%	2%
6MYN	69%	22%	< 5%	5%	3%
6WPP	82%	9%	0%	9%	6%

Physicochemical properties closely tied to in vivo performance—including rotatable bonds, the octanol–water partition coefficient (logP), hydrogen-bond donors (HBD) and acceptors (HBA), topological polar surface area (TPSA), molecular weight (MW), and heavy-atom count—were first used to prefilter candidates. Confining the library to ranges resembling those of known NIK inhibitors increases hit rates, enriches for novel actives, and focuses attention on drug-like molecules with better membrane permeability and development potential.

Compounds with inhibitor-like physicochemical profiles then entered the machine-learning activity prediction workflow. The author built pIC_{50} regression models using six algorithms—decision tree (DT), Bayesian ridge (BR), gradient boosting regression (GBR), k-nearest neighbors (KNN), random forest (RF), and support vector regression (SVR)—implemented via scikit-learn. Model selection and hyperparameter tuning used five-fold cross-validation (k = 5); randomized

DOI: 10.54254/2753-8818/2026.AU29729

search sampled plausible parameter ranges. In preprocessing, a near-zero-variance filter (threshold 1×10^{-8}) removed quasi-constant, non-informative features. Each regressor was trained and tested on the curated NIK inhibitors dataset, the best model was chosen to score candidates, and an activity threshold was applied to further contract the library. Performance and generalization were assessed by the coefficient of determination (R^2) and mean squared error (MSE)(See Table I).

For the BR model, optimization focused on the initial magnitudes of the precision parameters controlling noise and weight regularization, ranging from 10^{-6} to 10^{-3} on a logarithmic scale. For the SVR model, tuning involved the regularization coefficient C (roughly 0.3~100), the kernel width γ (approximately 10^{-3}~0.3), and the ε-insensitive margin (0.01~0.08). The DT model was optimized primarily for structural complexity and pruning strength by adjusting the maximum tree depth (None or integers 3~20), minimum samples per leaf (1~10), minimum samples required for internal node splitting (2~20), and the feature subsampling strategy (among "sqrt," "log2," full, or fractional ratios such as 0.5, 0.7, and 0.9), alongside a cost-complexity pruning coefficient (0~0.01). The GBR model with Huber loss jointly optimized the number of boosting stages (600~1000), learning rate (0.02~0.12), base learner depth (2~4), subsampling rate (0.70~0.95), minimum samples per leaf (2~8), and the robustness quantile parameter α (0.90~0.98). For the KNN model, optimization centered on the number of neighbors k (5~45), the weighting scheme (distance-based or uniform), the Minkowski distance order p (Manhattan or Euclidean), and the leaf size (20~50), balancing computational efficiency and accuracy. Lastly, the RF model was tuned under bootstrap sampling by varying the number of trees (600~1000), maximum depth (None or 6~15), minimum samples per leaf (1~8), minimum samples for node splitting (2~12), and the feature subsampling ratio ("sqrt," "log2," or fractions 0.3, 0.5, 0.7).

Docking was then used to prioritize structures with favorable receptor affinity. Building on the ML-reduced set, the Virtual Screening Workflow in Schrödinger Maestro identified compounds with lower docking scores, which were retained and further filtered by interactions with key residues. Prior to this step, key binding-site residues were designated and both protein and ligands were prepared. Protein ligand interaction fingerprints (PLIF) in MOE were used to map the presence of specific contacts between pocket residues and ligands, and FTMap (https://ftmap.bu.edu) delineated 'hot-spot' regions with strong probe binding, enabling identification of hydrogen-bond forming residues within the pocket. Protein and ligand preparation are conducted in Maestro with OPLS4 force-field via the Protein Preparation Workflow and LigPrep modules, respectively.

3.2. Molecular evaluation

The author denotes the molecules obtained from virtual screening as the result set and the known active NIK inhibitors as the reference set. For each molecule, RDKit was used to compute ECFP4 fingerprints (radius = 2; 2,048 bits), and pairwise structural similarity was quantified by the Tanimoto coefficient. To assess diversity, the author enumerated all pairwise Tanimoto values within the result set and, in parallel, generated Bemis–Murcko scaffolds and reported the number of unique scaffolds and the scaffold ratio. To gauge novelty, for every molecule in the result set, the author calculated its similarity to all compounds in the reference set and defined its maximum similarity as the highest Tanimoto score among those comparisons; the distribution of maximum similarity values was then summarized across the result set.

To compare chemical space coverage, the author performed principal components analysis (PCA) on eight features: MW, logP, TPSA, HBD, HBA, number of rotatable bonds, fraction of sp^3 carbons (Fsp^3), and QED score. Properties of the result and reference sets were z-scored separately (zero mean, unit variance). A two-dimensional PCA model was fitted on the standardized result set, and

DOI: 10.54254/2753-8818/2026.AU29729

the reference set was projected into the same principal-component space. A global bounding box was then defined from the combined 2D coordinates, the PC1×PC2 plane was partitioned into a uniform grid (bins = 60), and the grid cells occupied by each set were recorded; the final overlap was computed from these occupancy maps.

4. Results and discussion

4.1. Physicochemical properties screening

Table 2. Performance of the six regression models on the training set and testing set

Models	R^2_{Train}	MSE_{Train}	R^2_{Test}	MSE_{Test}
BR	0.6085	0.4975	0.5507	0.5640
DT	0.6996	0.3818	0.5645	0.5473
GBR	0.9163	0.1064	0.6630	0.4229
KNN	0.9826	0.0221	0.6117	0.4887
RF	0.9276	0.0920	0.6548	0.4353
SVR	0.8056	0.2471	0.6448	0.4466

From the NIK inhibitor activity dataset, the author selected entries with IC_{50} < 100 nmol/L (n = 796) and summarized the distributions of logP, HBD, HBA, TPSA, MW, number of rotatable bonds, and heavy-atom count; the results are shown in Fig. 1. Most compounds fall within a typical drug-like chemical space, characterized by moderate MW, balanced lipophilicity, and limited numbers of H-bonddonors/acceptors and rotatable bonds. These features align with Lipinski's Rule of Five and Veber's oral bioavailability icriteria [12, 13]. Although a few molecules slightly exceed the classical cutoffs for logP or MW, the set overall remains within ranges typical of small molecule drugs. These physicochemical profiles suggest that potent NIK inhibitors may couple strong target engagement with favorable pharmacokinetics and developability. Applying the parameter ranges in Fig. 1 to the candidate library retained 627,333 compounds.

4.2. Model evaluation

Table II reports the performance of the six regression models on the training and testing sets. Overall, the algorithms achieved moderate predictive accuracy for NIK inhibitor activity: all models yielded R^2_{Test} values of ~0.55–0.66, a range consistent with typical QSAR benchmarks, for which external validation $R^2 > 0.6$ is often deemed acceptable [14]. GBR, KNN, and RF fit the training data well ($R^2_{Train} > 0.9$), with GBR showing slightly better generalization than KNN and RF. GBR achieved the highest R^2_{Test} (0.6630) and the lowest MSE_{Test} (0.4229). Although KNN produced the strongest training performance (R^2_{Train} = 0.9826), its test error (MSE_{Test} = 0.4887) exceeded that of GBR and RF. BR and DT underperformed (both with $R^2_{Test} < 0.6$), whereas SVR delivered intermediate results.

Fig. 2 plots experimental versus predicted pIC_{50} for the six models. The predictions from GBR, RF and SVR cluster symmetrically around the identity line, with GBR overall lying closest to it. BR and DT show poor alignment with the diagonal, reflecting limited ability to discriminate across the range of experimental values. The stratified, "stepwise" pattern in DT arises because its outputs are restricted to a finite set of leaf means rather than continuous values—an inherent limitation for

DOI: 10.54254/2753-8818/2026.AU29729

complex, continuously varying bioactivity. Consequently, GBR was selected as the optimal predictive model.

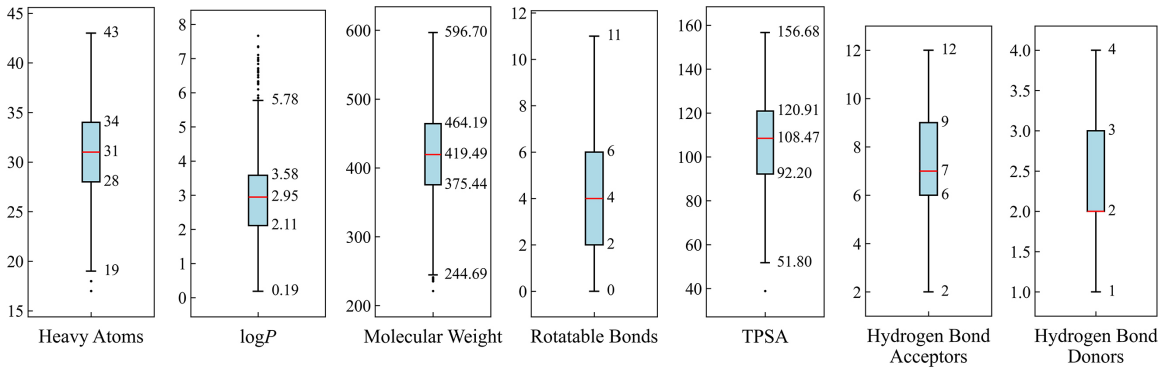


Figure 1. Statistical results of physicochemical properties of molecules in the NIK inhibitors library

4.3. Machine learning model screening

Building on the 627,333 candidates retained after physicochemical filtering, the author predicted activity with the optimal GBR model. As shown in Fig. 3, most predicted pIC_{50} values lie between 4 and 8, corresponding to IC_{50} values of 10–10,000 nmol/L. Using 10 nmol/L (pIC_{50} = 8) as the potency threshold, 397 compounds met the criterion.

4.4. Molecular docking screening

Fig. 5 summarizes PLIF and FTMap analyses of key residues across the five NIK protein–ligand co-crystal structures. As shown in Fig. 5B, ligand contacts with Glu472 and Leu474 are present in all five structures, suggesting these residues mark conserved binding sites. Fig. 5A indicates that Lys431, Leu474, and Asp536 exhibit higher vertical coordinates, reflecting a pronounced tendency to participate in hydrogen bonding and, thus, a greater likelihood of mediating H-bond formation upon ligand engagement. Among them, Leu474 displays the tallest bar and is inferred to directly form hydrogen bonds, whereas Glu472 shows a lower bar, implying a less direct contribution to H-bonding and a potential auxiliary role that facilitates Leu474–ligand interactions. Integrating the PLIF and FTMap results, the author therefore designates Leu474 as the key residue for hydrogen-bonding interactions.

DOI: 10.54254/2753-8818/2026.AU29729

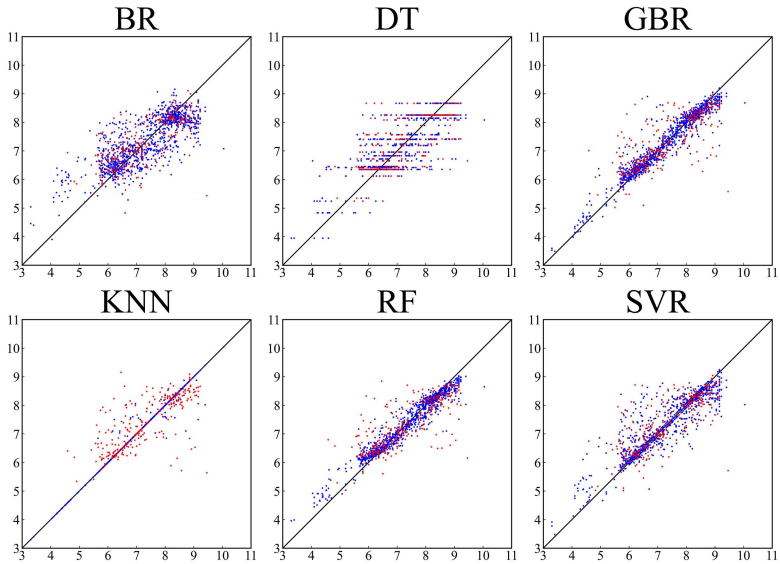


Figure 2. Experimental pIC_{50} (x-axis) versus predicted pIC_{50} (y-axis) for each molecule across the six models. Blue dots are molecules from the training set, and red dots are molecules from the testing set

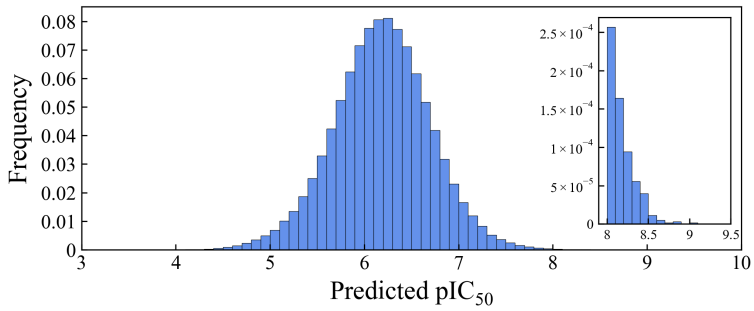


Figure 3. GBR model pIC_{50} prediction results for molecules in candidate library

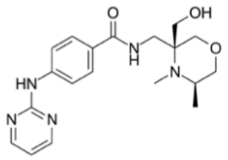
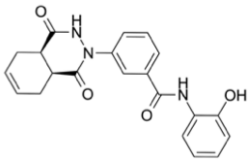
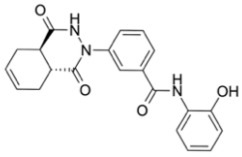
After machine-learning filtration, the 397 molecules were prepared as ligands and docked into the optimal NIK crystal structure (PDB 6G4Y). Applying an additional filter—requiring a hydrogen bond to the key residue Leu474—further narrowed the set to 18. The top three compounds (Table III) all achieved docking scores lower than –9 kcal/mol and had predicted pIC_{50} values > 8. Fig. 5 illustrates their binding modes in the active pocket. Among them, M-2 and M-3 display the most extensive interaction networks: beyond the key residue contact, they form multiple hydrogen bonds with Arg410, Glu475 and Gln481, and engage in a cation-π interaction with Arg410.

4.5. Molecular evaluation

To assess the soundness of our workflow, the author conducted additional analyses on the 18 docked compounds. These molecules were designated the result set, and the 796 known NIK inhibitors with $IC_{50} \leq 100$ nmol/L were designated the reference set. The author then quantified pairwise structural diversity within the result set and compared the two sets in terms of structural novelty and chemical space distribution.

DOI: 10.54254/2753-8818/2026.AU29729

Table 3. Screening results

ID	Structure	Docking Score (kcal/mol)	Predicted pIC_{50}
M-1		–10.181	8.228
M-2		–9.452	8.137
M-3		–9.030	8.137

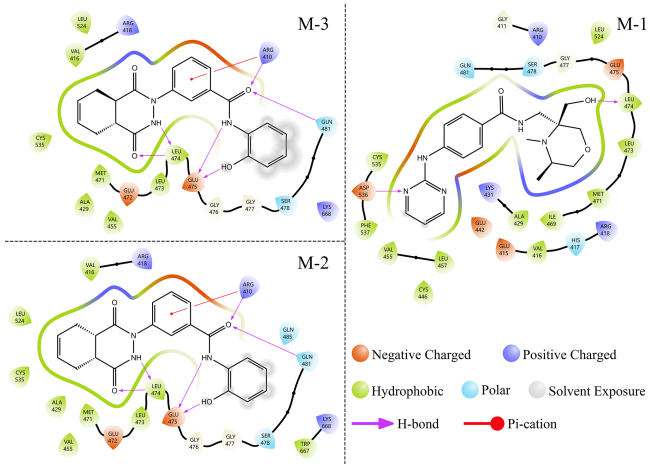


Figure 4. Interaction patterns between ligands and NIK active site

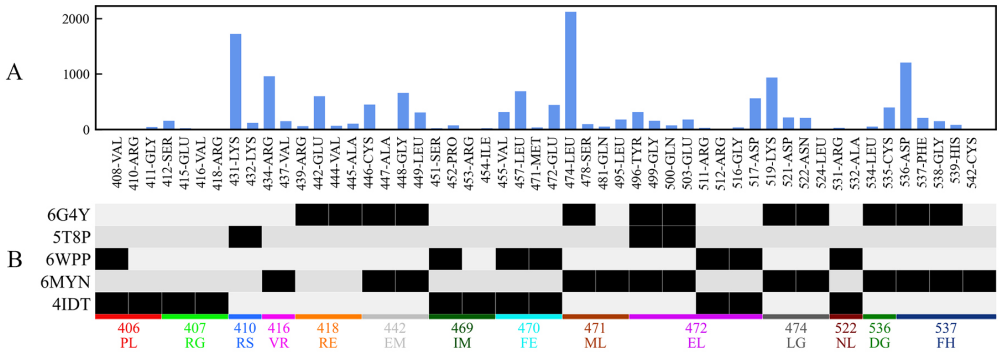


Figure 5. A: FTMap analysis result. B: PLIF analysis result

DOI: 10.54254/2753-8818/2026.AU29729

4.5.1. Structural diversity

At the scaffold level, the result set exhibits substantial structural heterogeneity. It contains 14 unique scaffolds; with 18 active compounds in total, the corresponding scaffold ratio is 0.778. This indicates that most molecules do not share a common core, and scaffold types are broadly distributed. Fingerprint-based similarity corroborates this conclusion. Using ECFP4 (radius = 2; 2,048 bits) with the Tanimoto coefficient for pairwise comparisons, the mean pairwise similarity within the result set is 0.199, consistent with generally low local-environment similarity and high overall diversity.

Fig. 6A shows the distribution of Tanimoto coefficients in the result set. Most values cluster in the 0.12–0.20 range, indicating predominantly low pairwise similarity; small secondary peaks at intermediate similarity (0.25–0.40) suggest a handful of closer neighbor pairs; and sparse bars near 1.0 reflect a few highly similar pairs arising from stereoisomers—molecules sharing the same scaffold but differing in the configuration of chiral centers.

4.5.2. Structural novelty

Fig. 6B depicts the distribution of each molecule's maximum similarity to the reference set. The values span ~0.20–0.32, with a mean of ~0.253 (min 0.205, max 0.324), forming a unimodal, left-shifted profile without a right tail. This pattern indicates that no compound in the result set has a close structural "neighbor" in the reference set; even the most similar pairs reach only ~0.32. The histogram further shows that the result set lies globally displaced from the reference set in fingerprint space, with no clustering near known chemical neighborhoods. Taken together with the diversity analyses, these findings indicate that the REAL-derived result set is scaffold-rich and exhibits low scaffold similarity to the reference set—demonstrating substantial structural novelty.

4.5.3. Chemical space

The gridded Jaccard index is 0.0014, a near-zero value indicating that the two sets occupy almost entirely distinct grid cells in the same coordinate frame, with only a negligible intersection. Given the disparate sample sizes (18 vs. 1,410), this extremely low Jaccard score implies that the result set's descriptor distribution lies largely in sparse or uncovered regions of the reference set—i.e., it adopts a markedly "outlying" spatial layout. While the Grid Jaccard depends on grid resolution, the conclusion is robust at the present setting (bins = 60): modest changes in resolution would cause only minor numerical fluctuation and would not alter the judgment.

In the PCA scatter plot (Fig. 6C), the reference set forms a dense, wedge-like cloud whose principal axis extends along positive PC1, with samples concentrated at lower PC2 values. By contrast, points from the result set are sparse and reside predominantly in the upper half of the plot; several lie beyond the upper boundary of the reference cloud, with only a few falling near its periphery. This spatial relationship is fully consistent with the near-zero grid-overlap inferred from the Jaccard metric.

DOI: 10.54254/2753-8818/2026.AU29729

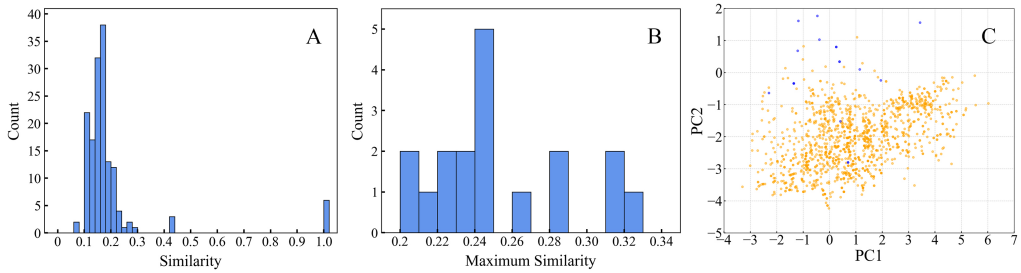


Figure 6. A: Similarity distribution of molecular pairs in the result set. B: Distribution of maximum similarity for molecules in the result set relative to the reference set. C: Chemical spatial distribution of principal component analysis; blue dots are molecules from the result set, and orange dots are molecules from the reference set

Notably, the multi-stage screening framework in this study exhibits advantages over existing virtual screening strategies for NIK inhibitors. For instance, compared with the single molecular docking approach used by Li et al. [8] (which yielded a hit rate of ~0.002% from a 500,000-compound library), our integration of physicochemical filtering, GBR-based activity prediction, and residue-guided docking achieved a higher hit rate (~0.00216%, 18 hits from 832,335 compounds) while ensuring structural novelty. Additionally, relative to the single machine learning model (RF) employed by Crawford et al. [9] for NIK inhibitor screening, our GBR model showed better generalization (R^2_{Test} = 0.663 vs. 0.58 reported in [9]), reducing false positives in the preliminary screening stage. These comparisons indicate that the combined strategy balances screening efficiency and hit quality, addressing the trade-off between speed and accuracy in large-library virtual screening.

5. Conclusion

This study integrated physicochemical filtering, machine learning, and molecular docking to virtually screen the 832,335-compound REAL library, ultimately yielding 18 novel, putative actives. The author first constrained candidates by the physicochemical profiles of known NIK inhibitors—logP, HBD, HBA, TPSA, MW, number of rotatable bonds, and heavy-atom count—retaining 627,333 compounds with similar properties. Next, six regression models were trained to predict pIC_{50}; the gradient boosting regression (GBR) performed best on the test set and was then used to score the candidate pool, keeping compounds with predicted IC_{50} below 100 nmol/L. Finally, docking-based triage—guided by FTMap, PLIF, and the identification of Leu474 as a key hydrogen-bonding residue—produced 18 candidates, three of which achieved docking scores better than –9 kcal/mol, indicative of stable protein engagement. Given NIK's association with human lymphoid malignancies, suitable NIK inhibitors could inspire new therapeutic avenues. The present virtual-screening strategy thus identifies molecules that both mirror the properties of known actives and exhibit favorable receptor binding, providing a practical, effective route—and proof of concept—for discovering structurally novel, potentially active chemotypes from large libraries.

Nonetheless, the study is constrained by synthetic feasibility and a lack of experimental validation. The hits have not been systematically assessed for synthesizability, stereochemical and functional-group compatibility, route length, or chemical stability; nor have in vivo stability and safety been evaluated. Future work should incorporate multi-objective optimization at the terminal screening stage, jointly modeling predicted potency, selectivity, synthetic accessibility (SA score/route complexity), chemical stability, and ADMET liabilities. Such refinements could advance

DOI: 10.54254/2753-8818/2026.AU29729

these virtual hits toward higher-activity, higher-selectivity, and more developable lead compounds, closing the loop from in silico discovery to experimental confirmation and medicinal chemistry optimization.

References

[1] Guo, Q., Jin, Y., Chen, X., Ye, X., Shen, X., Lin, M., Zeng, C., Zhou, T., & Zhang, J. (2024). "NF-kappaB in biology and targeted therapy: new insights and translational implications". Signal Transduct Target Ther, 9(1), 53.

[2] Mao, H., Zhao, X., & Sun, S. C. (2025). "NF-kappaB in inflammation and cancer". Cell Mol Immunol, 22(8), 811-839.

[3] Wang, Y., & Rui, L. (2024). "The druggable target potential of NF-κB-inducing kinase (NIK) in cancer". International Journal of Translational Medicine, 5(1).

[4] Rauert-Wunderlich, H., Siegmund, D., Maier, E., Giner, T., Bargou, R. C., Wajant, H., & Stuhmer, T. (2013). "The IKK inhibitor Bay 11-7082 induces cell death independent from inhibition of activation of NFkappaB transcription factors". PLoS One, 8(3), e59292.

[5] Boccadoro, M., Morgan, G., & Cavenagh, J. (2005). "Preclinical evaluation of the proteasome inhibitor bortezomib in cancer therapy". Cancer Cell Int, 5(1), 18.

[6] Cheng, J., Feng, X., Li, Z., Zhou, F., Yang, J. M., & Zhao, Y. (2021). "Pharmacological inhibition of NF-kappaB-inducing kinase (NIK) with small molecules for the treatment of human diseases". RSC Med Chem, 12(4), 552-565.

[7] Brightbill, H. D., Suto, E., Blaquiere, N., Ramamoorthi, N., Sujatha-Bhaskar, S., Gogol, E. B., Castanedo, G. M., Jackson, B. T., Kwon, Y. C., Haller, S., Lesch, J., Bents, K., Everett, C., Kohli, P. B., Linge, S., Christian, L., Barrett, K., Jaochico, A., Berezhkovskiy, L. M., …Ghilardi, N. (2018). "NF-kappaB inducing kinase is a therapeutic target for systemic lupus erythematosus". Nat Commun, 9(1), 179.

[8] Li, Z., Li, X., Su, M.-B., Gao, L.-X., Zhou, Y.-B., Yuan, B., Lyu, X., Yan, Z., Hu, C., Zhang, H., Luo, C., Chen, Z., Li, J., & Zhao, Y. (2020). "Discovery of a potent and selective NF-κB-inducing kinase (NIK) inhibitor that has anti-inflammatory effects in vitro and in vivo". Journal of Medicinal Chemistry, 63(8), 4388-4407.

[9] Crawford, J. J., Feng, J., Brightbill, H. D., Johnson, A. R., Wright, M., Kolesnikov, A., Lee, W., Castanedo, G. M., Do, S., Blaquiere, N., Staben, S. T., Chiang, P. C., Fan, P. W., Baumgardner, M., Wong, S., Godemann, R., Grabbe, A., Wiegel, C., Sujatha-Bhaskar, S., …Landry, M. L. (2023). "Filling a nick in NIK: Extending the half-life of a NIK inhibitor through structure-based drug design". Bioorg Med Chem Lett, 89, 129277.

[10] Zhang, N., Shen, S., Yang, M., He, S., Liu, C., Li, H., Lu, T., Liu, H., Hu, Q., Tang, W., & Chen, Y. (2024). "Design, synthesis, and biological evaluation of a novel NIK inhibitor with anti-inflammatory and hepatoprotective effects for sepsis treatment". J Med Chem, 67(7), 5617-5641.

[11] Worldwide Protein Data Bank (wwPDB). (2024) "Full wwPDB X-ray structure validation report". RCSB PDB. https: //files.rcsb.org/validation/view/6g4y_full_validation.pdf; https: //files.rcsb.org/validation/view/5t8p_full_validation.pdf; https: //files.rcsb.org/validation/view/6wpp_full_validation.pdf; https: //files.rcsb.org/validation/view/6myn_full_validation.pdf; https: //files.rcsb.org/validation/view/4idt_full_validation.pdf. (Retrieved 2025-10-07).

[12] Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2001). "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings". Advanced Drug Delivery Reviews, 46(1), 3-26.

[13] Veber, D. F., Johnson, S. R., Cheng, H.-Y., Smith, B. R., Ward, K. W., & Kopple, K. D. (2002). "Molecular properties that influence the oral bioavailability of drug candidates". Journal of Medicinal Chemistry, 45(12), 2615-2623.

[14] Alexander, D. L., Tropsha, A., & Winkler, D. A. (2015). "Beware of R: Simple, unambiguous assessment of the prediction accuracy of QSAR and QSPR models". J Chem Inf Model, 55(7), 1316-1322.