

Machine Learning-Guided Design of Novel Janus Kinase Inhibitors for Rheumatoid Arthritis Treatment

Angela Wu^{1,*}

¹ The Hill School, 860 Beech St, Pottstown, PA 19464, USA

* Corresponding Author Email: angelawu3090@gmail.com

Abstract. Attenuation of autoimmune responses by inhibiting immune activation pathways has been a common strategy for treating autoimmune diseases. The inhibition of JAK2 protein has proven effective in a clinical setting for treating RA. However, most of the JAK inhibitors have limited isoform selectivity, resulting in broad immune suppression that elevates the risk of infection, malignancy, and cardiovascular events. These concerns promote the development of novel JAK inhibitors with improved potency and selectivity. Focusing on the JAK2 protein, we identified *in-silico* novel small molecules with higher binding affinity than the existing drugs through a computational screening process that combines the generation of a drug compound library, development of predictive machine learning model, and molecular dynamics-based modeling. From a generated library of 34,992 candidates, four compounds were identified with enhanced binding affinity relative to existing inhibitors. Notably, the top candidate, M-84, exhibited a binding free energy of - 47.1 kcal/mol, which is 13.4 kcal/mol lower than that of the strongest reference drug. An analysis of the interaction profiles revealed that the additional hydrogen bonding and salt-bridge interactions within the JAK2 binding pocket enabled by the additional sulfonyl and alkylammonium groups contributed to the enhanced binding affinity, suggesting that tailoring functional groups to the chemical environment of the binding pocket can guide the rational design of new inhibitors. Moreover, the computational pipeline presented here is readily transferable to the discovery of potent small-molecule inhibitors targeting other proteins.

Keywords: Molecular design, drug design, artificial intelligence, machine learning, deep learning, rheumatoid arthritis.

1. Introduction

Autoimmune diseases are characterized by dysregulation of immune cell activation against normal tissues and organs [1]. They can affect almost all organ systems of all individuals and present a three to five times higher frequency in women [1], [2], [3]. Most autoimmune diseases show a chronic course that is driven by irreversible tissue damage resulting from sustained autoimmune attacks [3]. This chronicity induces substantial burdens on individuals and healthcare systems and therefore underscores the need for cost-effective strategies for long-term disease management [4]. Rheumatoid arthritis (RA) is one of the prevalent autoimmune diseases and affects approximately 1% of the population globally [5]. RA is marked by autoantibodies of immunoglobulin G (IgG), namely rheumatoid factor (RF), and citrullinated proteins, specifically anti-citrullinated protein antibodies (ACPAs), which primarily targets synovial joints and drives immune-mediated synovial inflammation, leading to progressive cartilage destruction, joint deformity, and ultimately functional disability [6]. Beyond these conditions, RA is also associated with elevated risks of cardiovascular and metabolic conditions such as hypertension (18.6%), diabetes mellitus (6.0%), hyperlipidemia (9.9%), and obesity (4.4%) [7]. The overall average total cost of RA treatment, including direct costs (such as medications), indirect costs (such as unemployment or loss of productivity), and intangible costs (such as quality of life), was estimated to be €14,905 per patient annually, with the cost for society to be €41.6 billion in the United States [5]. These clinical and socioeconomic consequences highlight the need for improved and effective therapeutic strategies.

Janus kinases (JAK) play a key role in the immune activation through the JAK-STAT pathway and thus mediate a broad range of inflammatory and immune responses [8]. JAK proteins couple to multiple cytokine receptors, such as interleukin-6 receptor (IL-6R), interleukin-12 receptor β (IL-

12R β), and interferon- γ receptor 2 (IFN- γ R2) to transduce immune signals [2]. Among them, an elevated expression of JAK2 has been found in the synovial tissues of patients with RA [2] and has been targeted for RA treatment [9]. For instance, Ding *et al.* suggested that the inhibition of JAK2 can mitigate RA conditions by preventing cytokine generation and T- and B-cell activation [2]. However, despite the efficacy of JAK2 inhibitors, they are associated with a range of safety concerns that limit their long-term use [10], [11]. Specifically, treatment with JAK inhibitors has been associated with increased risks of thromboembolic events and serious infections, such as upper and lower respiratory tract and urinary tract infections [5], [7], [9]. For example, the risk of herpes zoster virus infections is increased 1.5 to 2 times in the general RA populations than typically observed [12]. These safety concerns highlight the need for development of novel JAK2 inhibitors with improved selectivity and potency to enable safer and more effective long-term treatment of RA [13].

The drug discovery process is time-consuming, costly, and labor-intensive, often requiring more than a decade and billions of U.S. dollars per approved drug, while yielding success rates of less than 10% [14]. As a result, computational design and screening of drug molecules have been widely applied in pre-clinical stages for prioritization of putative molecules for future experimental tests [15]. In recent years, machine learning (ML) and artificial intelligence (AI) algorithms have emerged as potent tools for drug discovery, with an aim to not only reduce experimental costs but increase the efficiency of the drug discovery process [16], [17], [18]. They have been used in different stages of drug discovery [19], including the generation of compounds [20], hit identification [21], lead optimization [22], as well as the subsequent absorption distribution, metabolism, excretion, and toxicity (ADMET) modeling [23]. In the early stage, when integrated with physics-based methods such as molecular dynamics (MD) simulations and binding free energy calculations [24], [25], these computational techniques provide additional mechanistic insight into ligand-target interactions and allow more informed prioritization of candidates prior to experimental validation.

In this regard, we developed a high-throughput computational virtual screening (HTVS) pipeline that integrates compound library generation, surrogate ML model development, molecular docking, and molecular dynamics-based free energy calculations to identify novel small-molecule JAK2 inhibitors for RA treatment (Figure 1). An existing dataset of validated JAK2 inhibitors and their inhibitory activities was first curated [26], used to train a surrogate model capturing structure-activity relationships, and seed a generative REINVENT4 model for the construction of a new compound library with 34,992 potential JAK2 inhibitor candidates that inherit the scaffold structures from existing molecules [27]. The trained surrogate model was applied to the generated library to predict the inhibitory potency of candidates, from which the top 100 compounds were selected for a further physics-based evaluation. Along with the top five existing JAK2 inhibitors used as reference compounds, the top candidates were evaluated through molecular docking followed by binding free energy calculations using MD simulations. Based on free energy ranking, the four candidates M-84, M-90, M-34, and M-70, were identified to possess greater binding affinity to the JAK2 binding pocket than the reference structures. Notably, the top-ranked compound, M-84, exhibited a binding free energy estimate of 13.4 kcal/mol lower than that of the strongest reference inhibitor. Structural analysis revealed that these candidates retain a binding mode similar to the reference compounds in the JAK2 binding pocket. The functional groups in the top candidates, such as the alkylamine and sulfonyl moieties, enable increased hydrogen bonding and salt-bridge formation and therefore contributed to their improved binding affinity. These site-specific interactions reveal key structure-interaction relationships and provide design insights for the development of new JAK2 inhibitors. Overall, this computational pipeline demonstrates a scalable and transferable strategy for screening and prioritizing potent and mechanism-based drug candidates for RA treatment.

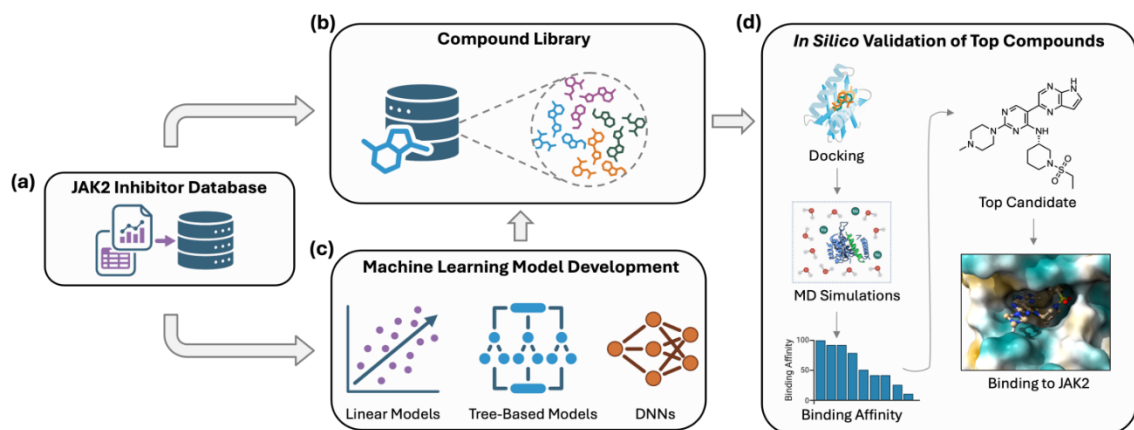


Figure 1. Visualization of the computational virtual screening pipeline for the design of novel JAK2 inhibitors for RA treatment. (a) An existing database comprising of small molecule structures and reported pIC50 values was curated for developing a predictive ML model. (b) The compound library generated by REINVENT4 using the existing inhibitors as seeding compounds. (c) Surrogate model development by training a series of models including linear models, ensemble tree-based models, and neural networks over the database in (a). The model parameters were optimized by cross validation. (d) The *in-silico* validation of model-predicted top candidates using free energy calculations from MD simulations. The interaction profiles of the highest-ranking subset was analyzed to reveal molecular forces stabilizing the ligand-protein binding and design insights for future molecular design. The icons were collected from BioRender.com and the rendering of the complex was created by ChimeraX [28].

2. Methods

2.1. Data Curation of the JAK2 Inhibitors

The initial JAK2 inhibitor dataset was sourced from Popova *et al.* and contains 1911 SMILES-encoded small molecules with experimentally measured pIC50 values [26]. The dataset was used as published without further curation or filtering. Because molecular featurization is crucial for the development of ML models and is influenced by the data size and complexity, we employed two widely used approaches: the extended-connectivity Morgan fingerprints [29] and a pretrained Molformer embedding [30]. In our implementation, Morgan fingerprint encodes each molecule by hashing circular atom-centered neighborhoods (radius = 3) into a fixed 2048-bit vector that captures local substructures and functional patterns. Molformer embeddings provide transformer-based learned representations directly generated from SMILES strings, capturing both the local chemical environment and global molecular context. By utilizing both featurization methods, model performances can be compared across distinct descriptor representations to obtain the optimal model for subsequent predictions.

2.2. Predictive Model Development

With the featurized dataset, our goal was to develop predictive ML models able to estimate the pIC50 values of unseen JAK2 inhibitors. We considered three major classes of ML algorithms that cover a wide range of complexities and biases: (i) linear models, (ii) tree-based ensemble models, and (iii) deep learning models (Figure 1c). Linear models included multiple linear regression (MLR) and regularized LASSO regression. Tree-based ensemble methods included decision trees (DT), random forests (RF), and gradient-boosted ensembles (e.g., CatBoost) [31], which are suitable for capturing nonlinear relationships and interactions without heavy preprocessing. Deep neural networks (DNNs) were also used to assess the capacity of flexible, high-dimensional models to learn more complex structure-activity relationships. For all models, both featurization strategies, Morgan fingerprints and Molformer embeddings, were independently evaluated to identify the optimal pairing of the model

and the representation. In addition to these classical approaches, we also used Chemprop, an advanced message-passing neural network (MPNN) architecture specifically designed for molecular property prediction, to compare its performance with other methods [32].

To develop the optimal model for subsequent prediction, parameter selection and model hyperparameter optimization were conducted using five-fold cross validation, where the dataset was divided into five evenly sized portions. In each evaluation iteration, four folds were used for training, while the remaining fold was used for testing. For each model class, we performed a random grid-search over relevant hyperparameters summarized in Table S1 (e.g., regularization strength for LASSO, tree depth and number of estimators for ensemble models, and network width, depth, dropout, and learning rate for DNNs and Chemprop).

Model performance was quantified on held-out validation folds to avoid overfitting and to ensure that the selected hyperparameters reflected generalizable predictive capability across the whole dataset. Performance of the models was evaluated using three commonly used regression metrics: mean absolute error (MAE), root-mean-square error (RMSE), and the coefficient of determination (R^2). MAE and RMSE respectively measure the average and root-mean-square deviation between predicted and experimental pIC50 values, with lower values indicating better predictive accuracy. R^2 measures the amount of variance in the experimental data explained by the model. A R^2 value close to 1 indicates excellent agreement between predicted and experimental data. The equations for these metrics are shown as the follows:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (1)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (3)$$

where y_i is the experimental value, $\hat{y}_i$ is the model prediction, $\bar{y}$ is the mean of experimental values, and N is the number of samples in the validation set. The final performance of each model was averaged across the five folds, and the mean MAE, RMSE, and R^2 values were reported as the overall predictive accuracy. This averaging process reduces bias from a single train-test split and provides a reliable estimation of the model’s generalizability.

2.3. Compound Library Generation

With the optimized predictive model, we next generated a virtual compound library composed of molecules that maintain the core scaffold structures found in the initial dataset. This design choice considers the limitation that ML models often struggle with extrapolation beyond the chemical space represented in their training data [33]. Therefore, by generating molecules with a reasonable amount of structural similarity to the confirmed JAK2 inhibitors, the predictive model is more likely to operate in an interpolative space, where its learned structure-activity relationships remain applicable and predictive performance is more reliable.

For this purpose, we used the AI molecular generation model REINVENT4 [27] to perform scaffold-constrained generation by fixing a predefined core structure and iteratively decorating it with functional groups sampled from its learned chemical vocabulary. In our implementation, the 1911 molecules from the initial dataset were provided to REINVENT4 as scaffold inputs. For each molecule, the model generated 20 highly similar and 20 moderately similar derivatives. After removing duplicate structures, a chemical library of 34,992 molecules was obtained. The distribution of the generated molecules remained similar to that of the initial dataset, as demonstrated by a UMAP [34] projection of the Jaccard similarity matrix of the Morgan fingerprints.

After the generation of the compound library, the optimal predictive model was employed to predict the pIC50 values for each molecule. The predictions were used to rank the molecules and to prioritize candidates for further *in-silico* validation. A total of 100 compounds, along with the top

five compounds from the initial library as baseline reference, were selected by the highest pIC₅₀ values for subsequent molecule modeling. Molecular docking and MD simulations were used to provide a more reliable performance assessment of the putative molecules and to identify top candidates for JAK2 inhibition.

2.4. Molecular Docking

The initial structure of JAK2 (PDB ID: 3FUP) was obtained from the RCSB protein data bank (<https://www.rcsb.org/>). The missing residues in the crystal structure were fixed by the Modeller module in ChimeraX [28]. AutoDock Vina was used to obtain the binding poses of the compounds in JAK2 [35]. We set the exhaustiveness to 32 and used the Vinardo scoring function. Five residues with IDs of 855, 983, 863, 932, 980 were set to be flexible during the docking run to account for the induced-fit effect [36]. For each receptor-ligand pair, only the best pose was kept for subsequent studies. The docked complex was constructed by combining the receptor and the best pose of the ligand.

2.5. MD Simulation and Binding Free Energy Calculation

MD simulations were performed on all protein-ligand complexes from docking using the GROMACS 2024.4 [37]. The Amberff99sb force field was used for the protein structure [38]. The topology of the small molecules was prepared using Acpype with the GAFF2 force field [39], [40]. The prepared complex was placed at the center of a cubic simulation box with a length of 10.8 nm and solvated by the TIP3P water model [41]. The net charge of the system was neutralized by adding counterions, with the inclusion of 150 mM NaCl to mimic physiological ionic strength.

In all simulations, energy minimization was carried out using the steepest descent algorithm to eliminate forces larger than 1000 kJ/mol·nm². For all simulations, the temperature was set to 300 K by a velocity-rescale thermostat with a time constant of 0.1 ps [42]. For NPT simulations, the pressure was maintained at 1 bar by a c-rescale barostat with a time constant of 12 ps and a compressibility of 3.0×10^{-4} bar⁻¹ [43]. The coupling scheme was isotropic for all systems. The time step was set to 2 fs, and Newton's equations of motion were integrated by the leapfrog algorithm [44]. Lennard-Jones interactions were calculated with a cutoff of 1.0 nm, and long-range electrostatics were treated using the Particle Mesh Ewald (PME) method with a real-space cutoff of 1.0 nm, a Fourier grid spacing of 0.16 nm, and a fourth-order interpolation scheme.

All simulations underwent 200 ps NVT equilibration followed by a 10 ns simulation for sampling. Once we obtained the simulation trajectory, binding free energies were estimated using the MM/PBSA approach with gmx_MMPBSA [45]. A multiple-frame scheme was used to obtain a more reliable estimate of the free energy values. The calculation was performed for the final 5 ns simulation comprising 50 frames of trajectories.

3. Results

3.1. Surrogate Model Development for Prediction of JAK2 Inhibition

The model was trained on a dataset of 1,911 JAK2 inhibitors and employed both fingerprint-based and learned molecular representations. We considered a series of models including multiple linear regression (MLR), LASSO, decision tree (DT), random forest (RF), CatBoost, and a deep neural network (DNN), and used two common molecular featurization strategies, including Morgan fingerprints and Molformer embeddings, to represent the molecular structures. Full details of the dataset and molecular featurization are provided in the Methods section.

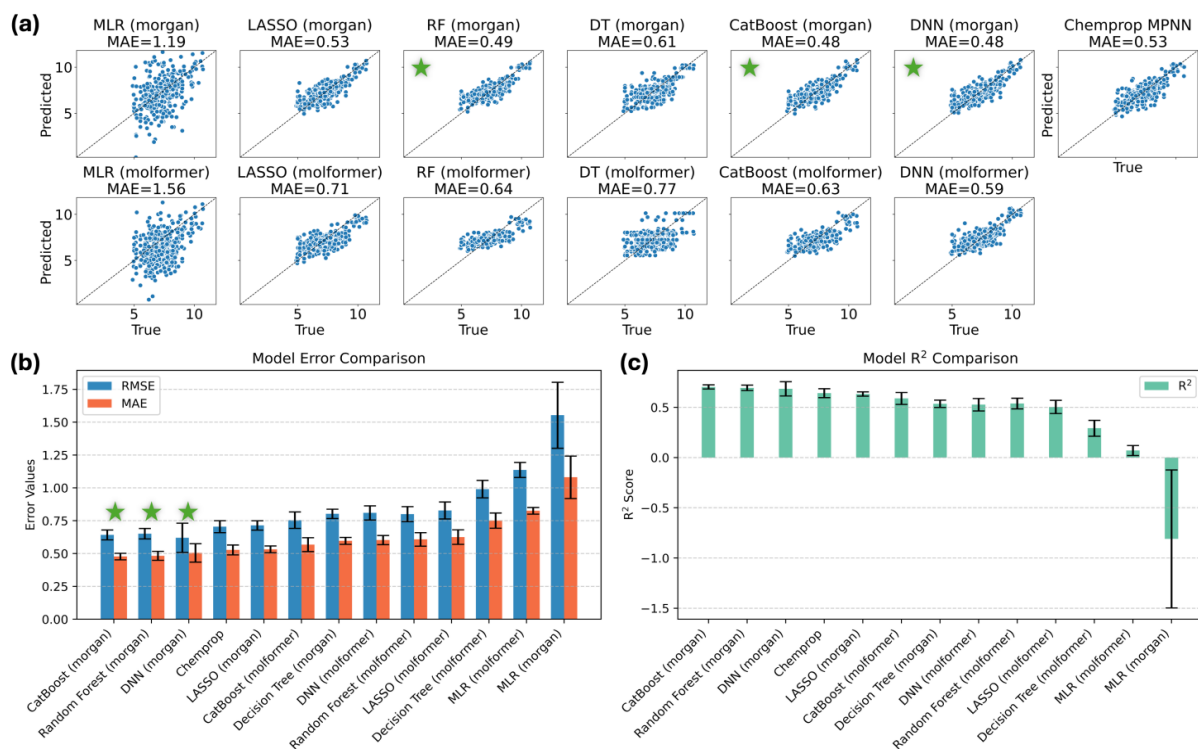


Figure 2. Development of the predicative model for JAK2 inhibition. (a) Representative scatter plots showing the performance of multivariate linear regression mode, regularized linear model, random forest, decision tree, CatBoost, deep neural network, and Chemprop model for the prediction of pIC50 values. The molecules were featurized by Morgan fingerprints, Molformer embeddings and Chemprop. The green stars denote the best learners from the test models. (b) Ranking of the performance of the model by RMSE and MAE values (b) as well as the R² values (c). The mean and error bars were obtained from five-fold cross-validation.

In Figure 2a, we show the representative scatter plots of the predicted and true values for different predictors. We found that the top models including CatBoost, RF, and DNN achieved similar predictive powers with comparable MAE values of 0.48 and 0.49 using the Morgan fingerprints as molecular representations. In contrast, the DT, LASSO, and MLR models showed limited prediction capacity, displaying broad scattering and deviation. Notably, the Chemprop model performed worse than the top conventional models, which is likely due to the relatively small dataset size, because such models typically require large amounts of training data to fully exploit their representational power.

These trends are further reflected in the mean error metrics (Figure 2b) and R² values (Figure 2c) from five-fold cross validations. The CatBoost, RF, DNN models using fingerprints representations have the lowest MAE/RMSE values and the highest R² values. Overall, the results highlight that well-regularized “traditional” ML models remain highly competitive under a limited-data regime. We also note that models using Morgan fingerprints consistently outperformed those using Molformer embeddings (Figure 2a), which is likely due to that Morgan fingerprints provide a task-aligned representation of local substructures that explains structure-property relationship, whereas Molformer offers a broad representation that may require larger and domain-specific datasets to achieve its advantages. Considering the results, we chose CatBoost with Morgan fingerprints as the surrogate predictor for subsequent screening as it provided the best balance of prediction accuracy and efficiency.

3.2. Screen of New Small Molecules Targeting JAK2 Inhibition

We then sought to build a new chemical library and utilize the established model to identify novel JAK2 inhibitors. Briefly, we queried the REINVENT4 model to generate molecules with medium and high Jaccard similarity to the existing molecules in the database by scaffold-constrained

molecular generation, which led to a compound library of 34,992 new molecules labeled by M-1 to M-34992. Subsequently, the optimized ML model was used to predict the inhibitory effect of the putative molecules. The top candidates were then validated by free energy calculation by MD simulations. Details of compound generation, molecular representation, and *in silico* validation protocols are described in the Methods section. The resulting free energy distribution for the top 100 predicted molecules is shown in Figure 3a, where most candidates populate a relatively higher (less favorable) free energy area, while a small subset shows decreased binding free energies compared to the top performing reference compound Ref-5. From the free energy ranking (Figure 3b), we identify four novel molecules, M-84, M-90, M-34, and M-70, that exhibit significantly lower binding free energies than Ref-5, which suggests improved binding affinity to JAK2 protein relative to all reference structures considered. We note that the five reference structures were previously selected as the top molecules from the initial dataset and that a lower binding free energy than any one of these are likely to lead to a new molecule with improved binding affinity.

Focusing on best reference structure Ref-5, Figure 3c shows the key molecular interactions that contribute to its stable binding to the protein. Specifically, we observe hydrogen-bonding interactions between the purine moiety of Ref-5 and residues Leu932 and Glu930. This purine core is additionally buried within the hydrophobic pocket, providing stabilizing nonpolar interactions that further enhance binding affinity. Based on their comparable and more favorable free energy profiles, we hypothesize that the top candidates may adopt binding poses that preserve these key interactions observed for Ref-5, while including additional functional groups that introduce new stabilizing interactions and further improve binding.

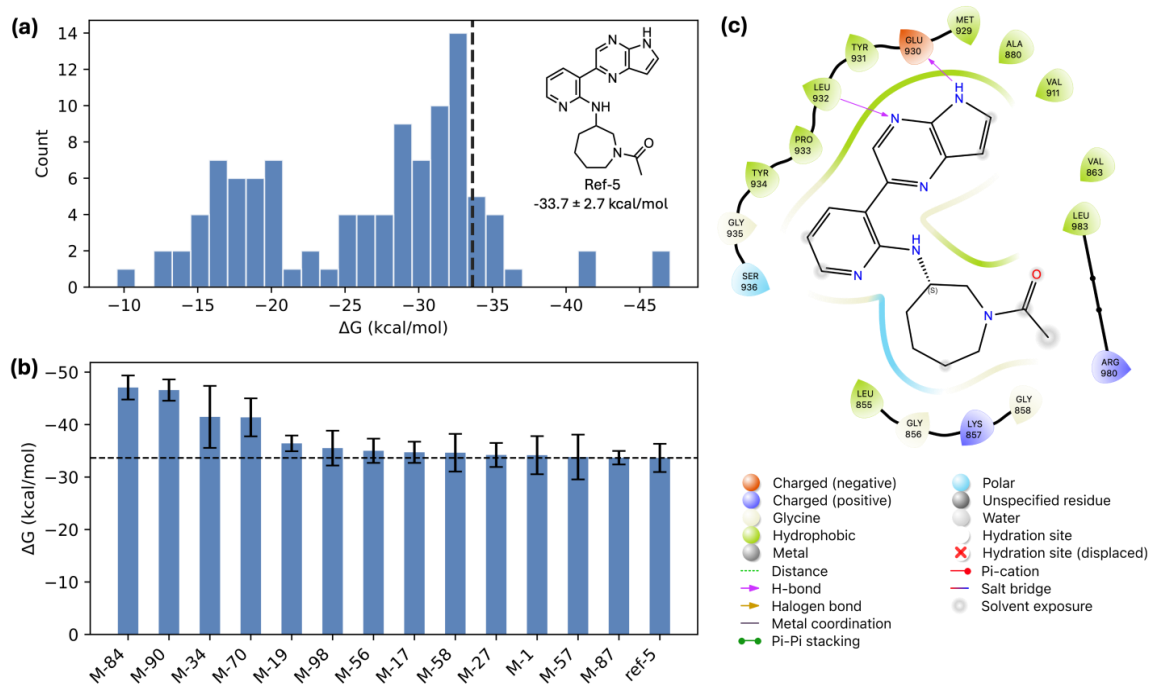


Figure 3. Identification of the new small molecules inhibiting JAK2. (a) A histogram plot showing the distribution of the free energies of the top 100 molecules predicted by the surrogate model. (b) The free energy values of the small molecules with free energies lower than the Ref-5 Structure. (c) A 2D diagram showing the interaction patterns of Ref-5 with the JAK2 protein pocket. The interaction diagram was created by Maestro.

3.3. Identification of New Small Molecules for JAK2 Inhibition

We examined the predicted binding modes of the top candidates and compared the molecular interaction profiles to understand the molecular forces that led to their enhanced binding affinity. As shown in Figure 4a, the four top candidates maintained the same core scaffold as Ref-5 while demonstrating improved binding potencies with a decrease in free energy of 7.7 to 13.4 kcal/mol

(Figure 3a). This structural similarity is expected for ML predictions in which ML models tend to identify top candidates that are chemically similar to the best reference compounds. The preservation of the core scaffold further suggests that the trained model reasonably captures the structure-activity relationship for JAK2 inhibition. Consistent with this structural similarity, the binding pose of M-84 is similar to that of Ref-5 where the purine moiety is inserted into the hydrophobic core of the binding pocket (Figure 4b). We hypothesize that decorating functional groups of these top candidates provide additional stabilizing molecular forces to enhance the binding affinity.

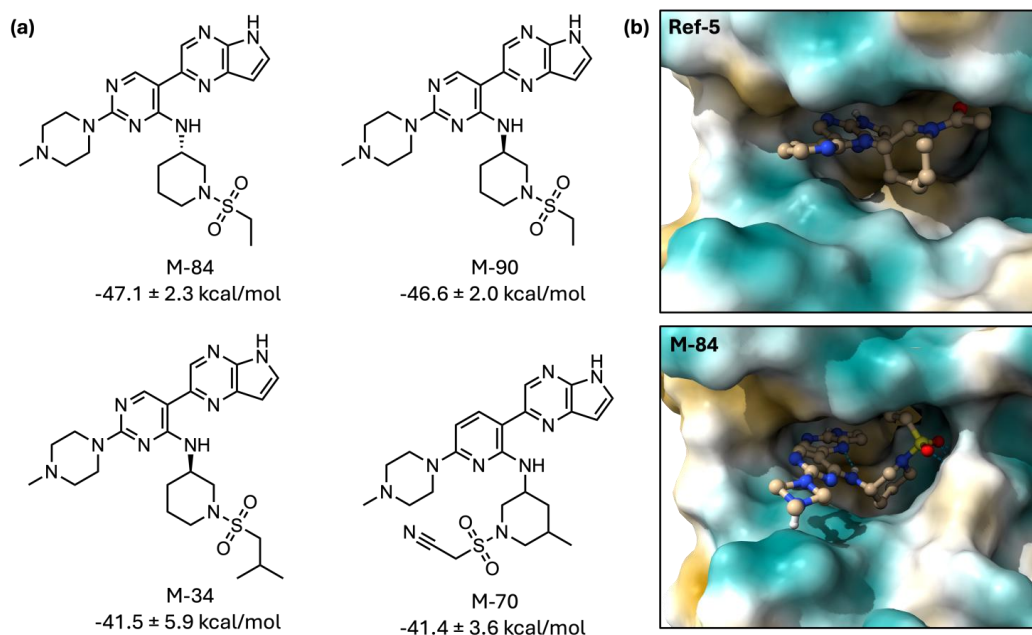


Figure 4. Binding modes of the top candidates identified by free energy calculations. (a) The chemical structures of the top candidates and their associated free energies. (b) The binding pose of Ref-5 and M-84 in the JAK2 protein pocket. The cyan and yellow surface represent respectively the hydrophilic and hydrophobic regions of the protein. The carbon, nitrogen, sulfur, hydrogen, and oxygen atoms are presented in tan, blue, yellow, white, and red colors, respectively. Non-polar hydrogens were omitted for clarity.

We next analyzed the interaction diagrams of the four top-ranked candidates in complex with the protein binding pocket to understand the underlying molecular interactions their enhanced binding affinities and to provide potential design insights for further optimization (Figure 5). All four molecules adopted highly similar predicted binding positions as expected due to their shared core scaffold. The purine moiety in each compound participates in hydrogen bonding with Leu932 and Glu930, while being buried within the hydrophobic core of the pocket to stabilize binding. In addition, these candidates form hydrogen bonds between the sulfonyl group and surrounding residues in the binding pocket. Notably, we hypothesize that the formation of a salt bridge between Asp939 and a protonated amine group in the ligands contributes significantly to binding stabilization through strong electrostatic interactions [46]. Together, these observations highlight that rational optimization of peripheral functional groups to complement the local chemical environment of the binding pocket can further enhance binding affinity and guide the design of improved inhibitors.

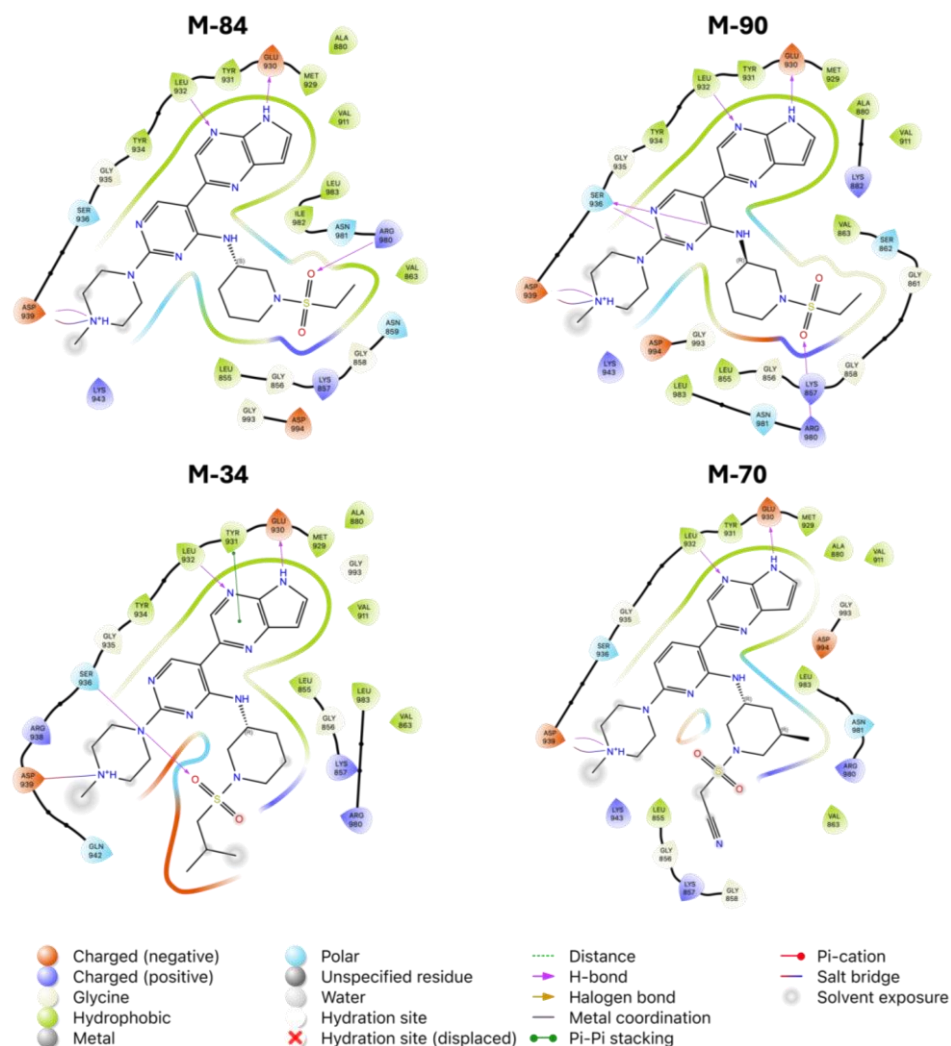


Figure 5. 2D protein-ligand interaction maps for representative scaffold-constrained molecules M-84, M-90, M-34, and M-70 bound in the target binding pocket. Key intermolecular interactions, including hydrogen bonds, π - π stacking, salt bridges, and hydrophobic contacts, are highlighted according to the legend. Residues are colored by physicochemical type, and solvent-exposed regions are indicated.

4. Conclusions and Discussions

Computational design and screening of small molecules has been an effective approach to exploring large chemical spaces and prioritizing novel candidates for experimental testing in drug discovery. In this study, we developed a virtual screening process that combines library generation, predictive surrogate model development, and physics-based modeling to identify novel JAK2 inhibitors with enhanced binding affinity for RA treatment. An initial dataset of known JAK2 inhibitors was curated and used both to seed compound library generation and to train a ML model that captures structure-activity relationships. The trained model was subsequently applied to predict and rank the potency of the generated candidates. For the top 100 molecules, molecular dynamics-based binding free energy calculations were performed to further refine the rankings and performances against existing drugs. From this screening process, four new small molecules (M-84, M-90, M-34, and M-70) were identified with evidently enhanced binding affinity to JAK2, with the top candidate M-84 exhibiting a binding free energy 13.4 kcal/mol lower than that of the best reference inhibitor. Interaction analysis revealed that retention of the core scaffold preserves the canonical binding manner, while the introduction of functional groups such as sulfonyl and alkylammonium moieties allows additional hydrogen bonding and salt-bridge interactions within the

binding pocket, providing design insights for further inhibitor optimization. Moreover, the modular and data-driven nature of this computational pipeline makes it transferable to the discovery of potent small-molecule inhibitors targeting other protein systems.

Several directions for future research are suggested. First, the identified small molecules should be validated in experiments to assess their inhibitory activity against JAK2 using binding and functional assays, along with by *in vivo* evaluation to establish their therapeutic potential. Second, the accuracy of binding free energy ranking could be improved by using more rigorous free energy methods, such as umbrella sampling or metadynamics [47]. Third, building on the identified top candidates, an additional optimization framework could be implemented to iteratively refine molecular structures and further enhance potency [48]. Finally, extending the evaluation of these compounds to other JAK family members would allow a systematic optimization of isoform selectivity toward JAK2 [10]. We anticipate that these efforts would accelerate the translation of the proposed computational discoveries into experimentally validated and clinically relevant drugs for RA treatment.

References

- [1] D. S. Pisetsky, “Pathogenesis of autoimmune disease,” *Nat. Rev. Nephrol.*, vol. 19, no. 8, pp. 509–524, Aug. 2023, doi: 10.1038/s41581-023-00720-1.
- [2] Q. Ding et al., “Signaling pathways in rheumatoid arthritis: implications for targeted therapy,” *Signal Transduct. Target. Ther.*, vol. 8, no. 1, p. 68, Feb. 2023, doi: 10.1038/s41392-023-01331-9.
- [3] A. D’Orazio et al., “Pathogenesis of rheumatoid arthritis: one year in review 2024,” *Clin. Exp. Rheumatol.*, Sep. 2024, doi: 10.55563/clinexprheumatol/0307ed.
- [4] A. Husivargova Theofanidis, P. Mikula, V. Timkova, Z. Macejova, R. Sanderman, and I. Nagyova, “The impact of socioeconomic disparity in autoimmune diseases,” *Eur. J. Public Health*, vol. 35, no. Supplement_4, p. ckaf161.1084, Oct. 2025, doi: 10.1093/eurpub/ckaf161.1084.
- [5] J. S. Smolen et al., “Rheumatoid arthritis,” *Nat. Rev. Dis. Primer*, vol. 4, no. 1, p. 18001, Feb. 2018, doi: 10.1038/nrdp.2018.1.
- [6] M. Jahid, K. U. Khan, Rehan-Ul-Haq, and R. S. Ahmed, “Overview of Rheumatoid Arthritis and Scientific Understanding of the Disease,” *Mediterr. J. Rheumatol.*, vol. 34, no. 3, p. 284, Sep. 2023, doi: 10.31138/mjr.20230801.oo.
- [7] M. Ausserwinkler et al., “Rheumatoid arthritis is independently associated with metabolic Dysfunction-Associated steatotic liver disease: evidence from the paracelsus 10,000 Population-Based cohort study,” *Rheumatol. Int.*, vol. 45, no. 10, p. 243, Oct. 2025, doi: 10.1007/s00296-025-06002-2.
- [8] X. Hu, J. Li, M. Fu, X. Zhao, and W. Wang, “The JAK/STAT signaling pathway: from bench to clinic,” *Signal Transduct. Target. Ther.*, vol. 6, no. 1, p. 402, Nov. 2021, doi: 10.1038/s41392-021-00791-1.
- [9] Y. Tanaka, Y. Luo, J. J. O’Shea, and S. Nakayamada, “Janus kinase-targeting therapies in rheumatology: a mechanisms-based approach,” *Nat. Rev. Rheumatol.*, vol. 18, no. 3, pp. 133–145, Mar. 2022, doi: 10.1038/s41584-021-00726-8.
- [10] M. Bonelli et al., “Selectivity, efficacy and safety of JAKinibs: new evidence for a still evolving story,” *Ann. Rheum. Dis.*, vol. 83, no. 2, pp. 139–160, Feb. 2024, doi: 10.1136/ard-2023-223850.
- [11] B. Mikaeili, Z. A. Alqahtani, A. L. Hincapie, and J. J. Guo, “Safety of Janus kinase inhibitors in rheumatoid arthritis: a disproportionality analysis using FAERS database from 2012 to 2022,” *Clin. Rheumatol.*, vol. 44, no. 4, pp. 1467–1474, Apr. 2025, doi: 10.1007/s10067-025-07360-9.
- [12] D. A. Pappas et al., “Herpes Zoster Reactivation in Patients with Rheumatoid Arthritis: Analysis of Disease Characteristics and Disease-Modifying Antirheumatic Drugs,” *Arthritis Care Res.*, vol. 67, no. 12, pp. 1671–1678, Dec. 2015, doi: 10.1002/acr.22628.
- [13] A. A. Patel and O. Odenike, “The Next Generation of JAK Inhibitors: An Update on Fedratinib, Momelotonib, and Pacritinib,” *Curr. Hematol. Malig. Rep.*, vol. 15, no. 6, pp. 409–418, Dec. 2020, doi: 10.1007/s11899-020-00596-z.

- [14] A. Sertkaya, T. Beleche, A. Jessup, and B. D. Sommers, “Costs of Drug Development and Research and Development Intensity in the US, 2000-2018,” *JAMA Netw. Open*, vol. 7, no. 6, p. e2415445, Jun. 2024, doi: 10.1001/jamanetworkopen.2024.15445.
- [15] A. V. Sadybekov and V. Katritch, “Computational approaches streamlining drug discovery,” *Nature*, vol. 616, no. 7958, pp. 673–685, Apr. 2023, doi: 10.1038/s41586-023-05905-z.
- [16] A. R. Bhat and S. Ahmed, “Artificial intelligence (AI) in drug design and discovery: A comprehensive review,” *Silico Res. Biomed.*, vol. 1, p. 100049, 2025, doi: 10.1016/j.insci.2025.100049.
- [17] F. J. N. Ferreira and A. S. Carneiro, “AI-Driven Drug Discovery: A Comprehensive Review,” *ACS Omega*, vol. 10, no. 23, pp. 23889–23903, Jun. 2025, doi: 10.1021/acsomega.5c00549.
- [18] V. Patel and M. Shah, “Artificial intelligence and machine learning in drug discovery and development,” *Intell. Med.*, vol. 2, no. 3, pp. 134–140, Aug. 2022, doi: 10.1016/j.imed.2021.10.001.
- [19] D. B. Catacutan, J. Alexander, A. Arnold, and J. M. Stokes, “Machine learning in preclinical drug discovery,” *Nat. Chem. Biol.*, vol. 20, no. 8, pp. 960–973, Aug. 2024, doi: 10.1038/s41589-024-01679-1.
- [20] K. Atz et al., “Prospective de novo drug design with deep interactome learning,” *Nat. Commun.*, vol. 15, no. 1, p. 3408, Apr. 2024, doi: 10.1038/s41467-024-47613-w.
- [21] D. Boldini, L. Friedrich, D. Kuhn, and S. A. Sieber, “Machine Learning Assisted Hit Prioritization for High Throughput Screening in Drug Discovery,” *ACS Cent. Sci.*, p. acscentsci.3c01517, Mar. 2024, doi: 10.1021/acscentsci.3c01517.
- [22] S. Chen et al., “Deep lead optimization enveloped in protein pocket and its application in designing potent and selective ligands targeting LTK protein,” *Nat. Mach. Intell.*, vol. 7, no. 3, pp. 448–458, Feb. 2025, doi: 10.1038/s42256-025-00997-w.
- [23] K. Swanson et al., “ADMET-AI: a machine learning ADMET platform for evaluation of large-scale chemical libraries,” *Bioinformatics*, vol. 40, no. 7, p. btae416, Jul. 2024, doi: 10.1093/bioinformatics/btae416.
- [24] M. Badaoui et al., “Combined Free-Energy Calculation and Machine Learning Methods for Understanding Ligand Unbinding Kinetics,” *J. Chem. Theory Comput.*, vol. 18, no. 4, pp. 2543–2555, Apr. 2022, doi: 10.1021/acs.jctc.1c00924.
- [25] A. Ghidini, E. Serra, and A. Cavalli, “On Free Energy Calculations in Drug Discovery,” *Acc. Chem. Res.*, vol. 58, no. 20, pp. 3137–3145, Oct. 2025, doi: 10.1021/acs.accounts.5c00465.
- [26] M. Popova, O. Isayev, and A. Tropsha, “Deep reinforcement learning for de novo drug design,” *Sci. Adv.*, vol. 4, no. 7, p. eaap7885, Jul. 2018, doi: 10.1126/sciadv.aap7885.
- [27] H. H. Loeffler et al., “Reinvent 4: Modern AI-driven generative molecule design,” *J. Cheminformatics*, vol. 16, no. 1, p. 20, Feb. 2024, doi: 10.1186/s13321-024-00812-5.
- [28] T. D. Goddard et al., “UCSF ChimeraX: Meeting modern challenges in visualization and analysis,” *Protein Sci.*, vol. 27, no. 1, pp. 14–25, Jan. 2018, doi: 10.1002/pro.3235.
- [29] S. Zhong and X. Guan, “Count-Based Morgan Fingerprint: A More Efficient and Interpretable Molecular Representation in Developing Machine Learning-Based Predictive Regression Models for Water Contaminants’ Activities and Properties,” *Environ. Sci. Technol.*, vol. 57, no. 46, pp. 18193–18202, Nov. 2023, doi: 10.1021/acs.est.3c02198.
- [30] J. Pan, “Large language model for molecular chemistry,” *Nat. Comput. Sci.*, vol. 3, no. 1, pp. 5–5, Jan. 2023, doi: 10.1038/s43588-023-00399-1.
- [31] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, “CatBoost: unbiased boosting with categorical features,” 2017, arXiv. doi: 10.48550/ARXIV.1706.09516.
- [32] E. Heid et al., “Chemprop: A Machine Learning Package for Chemical Property Prediction,” *J. Chem. Inf. Model.*, vol. 64, no. 1, pp. 9–17, Jan. 2024, doi: 10.1021/acs.jcim.3c01250.
- [33] Z.-W. Zhao, M. Del Cueto, and A. Troisi, “Limitations of machine learning models when predicting compounds with completely new chemistries: possible improvements applied to the discovery of new non-fullerene acceptors,” *Digit. Discov.*, vol. 1, no. 3, pp. 266–276, 2022, doi: 10.1039/D2DD00004K.
- [34] L. McInnes, J. Healy, and J. Melville, “UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction,” 2018, arXiv. doi: 10.48550/ARXIV.1802.03426.

- [35] J. Eberhardt, D. Santos-Martins, A. F. Tillack, and S. Forli, "AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings," *J. Chem. Inf. Model.*, vol. 61, no. 8, pp. 3891–3898, Aug. 2021, doi: 10.1021/acs.jcim.1c00203.
- [36] D. Lu, D. Luo, Y. Zhang, and B. Wang, "A Robust Induced Fit Docking Approach with the Combination of the Hybrid All-Atom/United-Atom/Coarse-Grained Model and Simulated Annealing," *J. Chem. Theory Comput.*, vol. 20, no. 14, pp. 6414–6423, Jul. 2024, doi: 10.1021/acs.jctc.4c00653.
- [37] M. J. Abraham et al., "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers," *SoftwareX*, vol. 1–2, pp. 19–25, Sep. 2015, doi: 10.1016/j.softx.2015.06.001.
- [38] C. Tian et al., "ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution," *J. Chem. Theory Comput.*, vol. 16, no. 1, pp. 528–552, Jan. 2020, doi: 10.1021/acs.jctc.9b00591.
- [39] A. W. Sousa Da Silva and W. F. Vranken, "ACPYPE - AnteChamber PYthon Parser interfacE," *BMC Res. Notes*, vol. 5, no. 1, p. 367, Dec. 2012, doi: 10.1186/1756-0500-5-367.
- [40] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, "Development and testing of a general amber force field," *J. Comput. Chem.*, vol. 25, no. 9, pp. 1157–1174, Jul. 2004, doi: 10.1002/jcc.20035.
- [41] P. Mark and L. Nilsson, "Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K," *J. Phys. Chem. A*, vol. 105, no. 43, pp. 9954–9960, Nov. 2001, doi: 10.1021/jp003020w.
- [42] G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling," *J. Chem. Phys.*, vol. 126, no. 1, p. 014101, Jan. 2007, doi: 10.1063/1.2408420.
- [43] M. Bernetti and G. Bussi, "Pressure control using stochastic cell rescaling," *J. Chem. Phys.*, vol. 153, no. 11, p. 114107, Sep. 2020, doi: 10.1063/5.0020514.
- [44] W. F. Van Gunsteren and H. J. C. Berendsen, "A Leap-frog Algorithm for Stochastic Dynamics," *Mol. Simul.*, vol. 1, no. 3, pp. 173–185, Mar. 1988, doi: 10.1080/08927028808080941.
- [45] M. S. Valdés-Tresanco, M. E. Valdés-Tresanco, P. A. Valiente, and E. Moreno, "gmx_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS," *J. Chem. Theory Comput.*, vol. 17, no. 10, pp. 6281–6291, Oct. 2021, doi: 10.1021/acs.jctc.1c00645.
- [46] D. S. Spassov, M. Atanasova, and I. Doytchinova, "A role of salt bridges in mediating drug potency: A lesson from the N-myristoyltransferase inhibitors," *Front. Mol. Biosci.*, vol. 9, p. 1066029, Jan. 2023, doi: 10.3389/fmolb.2022.1066029.
- [47] Q. Liao, "Enhanced sampling and free energy calculations for protein simulations," in *Progress in Molecular Biology and Translational Science*, vol. 170, Elsevier, 2020, pp. 177–213. doi: 10.1016/bs.pmbts.2020.01.006.
- [48] P. I. Frazier, "A Tutorial on Bayesian Optimization," 2018, arXiv. doi: 10.48550/ARXIV.1807.02811.