Review Article

Advancing NTDs Drug Discovery Through AI and Computational Design: An Updated Review

Sayan Mondal [1], Prabir Manna[1*], Venkatesha Perumal Ramachandran[1], Pritha Dey[2], Nilotpal Tumung [3], Sumit Maji [4], Shyamal Kumar Biswas[1], Himangshu Sekhar Maji [5], Rakesh Kumar Dolai [6], Md Abdul Quadir Khan[7]

[1] Department of Pharmaceutical Chemistry, Calcutta Institute of Pharmaceutical Technology and Allied Health Sciences, Banitabla, Uluberia, Howrah - 711316, West Bengal, India.
[2] Department of Pharmaceutical Chemistry, Derozio Pharma Institute, Gopal Chak, Moyna-721629, West Bengal, India.
[3] Department of Pharmaceutical Analysis, School of Pharmacy, The Neotia University, Sarisa, West Bengal, India.
[4] Department of Pharmaceutical Chemistry, P. G. Institute of Medical Sciences, Chandrakona Town, West Bengal, India.
[5] Department of Chemistry, Turku Hansda Lapsa Hemram Mahavidyalay, Madian, Birbhum-731216, West Bengal, India.
[6]Department of Pharmaceutical Analysis, Dr. B. C. Roy College of Pharmacy and Allied Health Sciences, Durgapur, West Bengal, India.
[7]School of Pharmacy, ARKA JAIN University, Mohanpur, Gamharia, Seraikela, Kharsawan, Jharkhand- 832108, India.
***Corresponding author's E-mail:** prabir.ciptahs@gmail.com

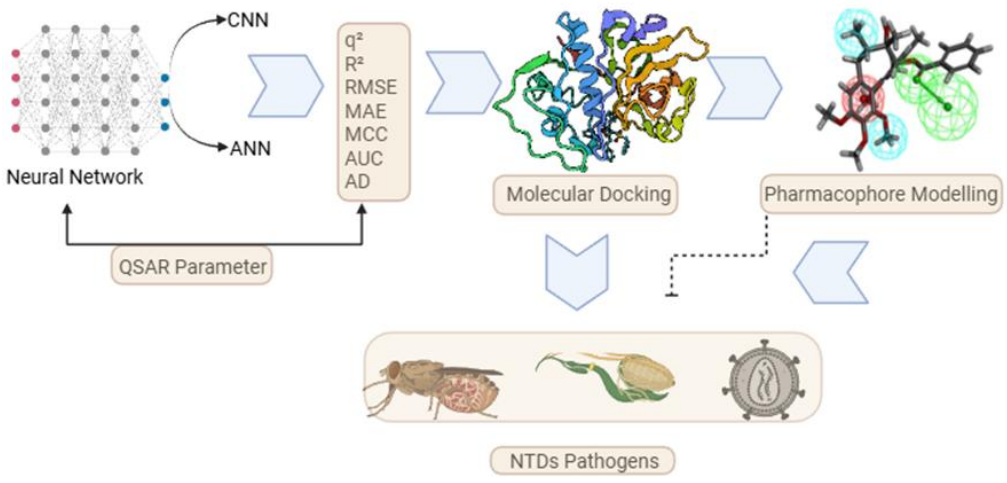
Received: 16-01-2026; **Revised:** 28-03-2026; **Accepted:** 03-04-2026; **Published online:** 20-04-2026.

ABSTRACT

Drug discovery for neglected tropical diseases (NTDs) is being revolutionized by rapid advancements in artificial intelligence and machine learning continue to redefine the landscape of intelligent data analysis, which speeds up the discovery and improvement of new treatments, whether NTDs are diverse infectious conditions, mainly affecting impoverished populations in tropical regions, caused by various pathogens, leading to severe health, social, and economic burdens globally. The convergence of AI-driven methodologies-including neural networks, deep learning, and multitask learning has emerged as a cornerstone of modern computational intelligence that dramatically expanded the scope and efficacy of virtual screening, structure-activity relationship analysis, and compound repurposing for NTDs as well as parasitic diseases. Multidimensional QSAR models (2D, 3D, 4D etc.), powered by ML algorithms, now efficiently predict molecular activity and toxicity for diverse chemical libraries, enabling rapid prioritization of promising candidates for further development. AI-assisted molecular docking platforms leverage precise protein-ligand interaction predictions for emerging biological targets, increasing both accuracy and throughput in virtual screening campaigns. Recent breakthroughs harness advanced neural network architectures capable of directly interpreting and learning from raw molecular structures without any explicit features that boost the hit rates on malaria, Onchocerciasis, Soil-transmitted helminthiases, Trachoma and Chagas disease. On the other hand, Fragment-based drug design (FBDD), driven by AI infrastructure has produced optimized drug leads, used smart fragment expansion, and linked traffic to produce potent NTD-targeted inhibitors with superior selectivity phenotypes. Besides, an advanced pharmacophore modelling, powered by ML, scaffold such as structure or maps that is crucial to molecular characteristics across various collections comparing chemical structure to biological work. It is this combination of improved computational techniques supporting recent translational approaches that have been able to advance compounds between virtual screening and preclinical validation within record time. The development of multi-target agents has accelerated due to collaborative AI efforts, preventing resistance to these drugs and increasing therapeutic options in settings with limited resources. Together, all these AI/ML advances represent a paradigm shift in medicinal chemistry of NTDs, drastically diminishing the development cycle, maximizing the effectiveness of individual compounds, and allowing the creation of tailored solutions to the population groups most impacted by these diseases. Further development of AI and ML is of colossal potential to provide groundbreaking treatment and reinvigorate the fight against tropical neglected diseases on a global scale.

Keywords: NTDs, QSAR, Molecular Modelling, Deep learning, FBDD, Drug Design.

Graphical Abstract:



1. INTRODUCTION

NTDs encompass a heterogeneous cluster of infectious diseases arising from varied pathogenic agents such as viruses, bacteria, parasites, and fungi. The diseases are common in underserved groups within the tropical and subtropical regions which are in many cases poverty-stricken areas[1]. NTDs cover a vast number of diseases like the Chagas disease, dengue, leprosy, lymphatic filariasis and schistosomiasis among others. They lead to serious health issues which result in disability, mutilation and even death. While NTDs continue to afflict over a billion people across the globe, these diseases have traditionally not received as much attention or funding as other diseases. The fight against these diseases is currently stepping up towards prevention, control and elimination of such diseases by providing better access to healthcare, enhancing sanitation and providing specific medical interventions. Problem of NTDs is the fundamental to improving global health and ensuring equitable well-being across populations and minimizing the impact on the affected populations[2,3].

By 2025, there has been a significant improvement in the global efforts to combat the neglected tropical diseases (NTDs). In 2023, 1.495 billion individuals were estimated to require interventions due to NTDs, which is very different as it was in the past years and underscores the continued success in control interventions[4]. A significant reduction has been reported in disease burden—quantified through disability-adjusted life years (DALYs)—as well as in overall mortality which attributable to NTDs also declined. The elimination or reduction of several countries has been successful with seven countries that have been identified by WHO as having eliminated at least one NTD by the year 2024[5]. Global community has remained concerned with combination strategies like preventive chemotherapy, better diagnostics and health system reinforcement, even though disease modern medicinal chemistry has problems in designing drug strategies like AI/ML driven strategies with Deep Learning, Generative Adversarial Networks, Reinforcement Learning, Quantum Computing, AI-Based Virtual Screening etc[6]. In the other hand, QSAR (Quantitative Structure-Activity Relationship)modelling integrating advanced regression methodologies, including multiple linear (MLR) and non-linear (MNLR) models, recursive feature elimination (RFE),linear logistic regression (LLR),principal component analysis (PCA),artificial neural networks (ANN), convolutional neural networks (CNN), recurrent neural networks (RNN),deep neural networks (DNN),associative neural networks (AsNN),generalized adversarial networks (GAN),support vector machines (SVM) etc[7,8]. To detect the NTDs, not only diverse classes of QSAR is required but also molecular docking, molecular dynamic Simulations, pharmacophore modelling (employing complementary pharmacophore modelling strategies, including ligand-based and structure-based approaches), fragment-based drug design (FBDD) and pharmacokinetic profiling involving ADMET Prediction has been utilizing in NTDs[8,9].

2. Database:

Databases take pivotal role in QSAR modelling for NTDs by providing curated chemical as well as pharmacological data which is essential for drug discovery. One commonly used resource is the ChEMBL database, which offers comprehensive information on bioactive molecules, including inhibitors relevant to NTD pathogens such as Trypanosoma cruzi (Chagas disease), ZINC Database, ChEBI, BindingDB etc[10]. QSAR models built using datasets from ChEMBL enable prediction of biological activity and identification of promising drug candidate subtilizing advanced machine learning approaches, particularly Support Vector Machines (SVMs) and Artificial Neural Networks (ANNs). Other databases such as PubChem and DrugBank (Table 1) also contribute valuable chemical and pharmacokinetic data. These resources support integrated approaches by combining molecular descriptors with cheminformatics for virtual screening and lead optimization for NTDs[11,12]. Here, we show some commonly used databases for readers.

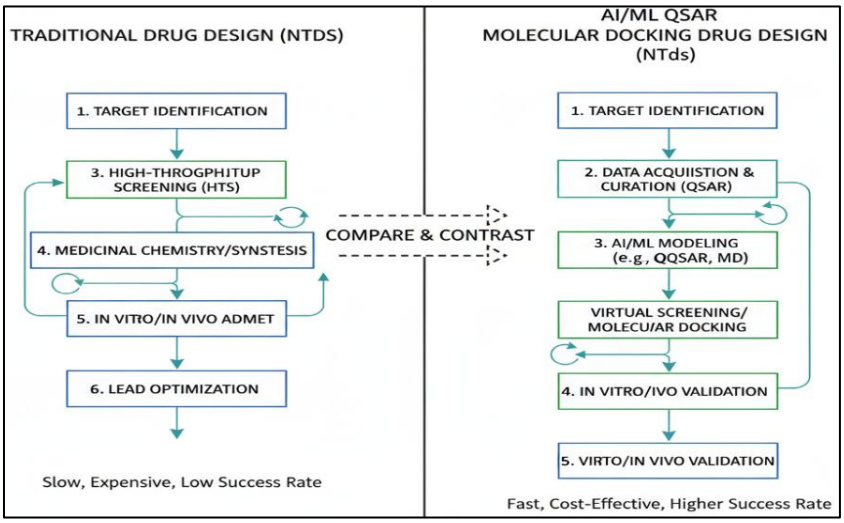


Figure 1: A comparative diagram showing conventional and AI/ML-integrated drug design methodologies for NTD treatments.

Table 1: Comprehensive NTD drug discovery databases providing bioactivities, protein structures, clinical trial data, and chemical descriptors essential for QSAR modelling and virtual screening workflows.

Database Name	Key Features	Reference
ChEMBL	Over 5.4 million bioactivities, binding affinities (IC50, Ki, EC50), 1M+ compounds, target annotations, open access, NTD-focused subset available	[13,14]
ChEMBL-NTD	Focused repository for NTD medicinal chemistry screening data, primary/orthogonal data, regular updates, direct access & community deposition	[14]
PubChem	119M+ compounds, large-scale structure, activity and assay data, integration with tools and resources for cheminformatics and drug discovery	[15]
DrugBank	Approved/experimental drug data, chemical structures, mechanisms, targets, pharmacology, ADMET, clinical status and trial linkage	[16]
NPAtlas	Natural product compounds, detailed antimicrobial, antiparasitic data, taxonomic info, bioactivity screen results	[17]
Open Targets Platform	Disease associations, target prioritization, tractability, NTD clinical/experimental linkage, visualization	[16]
DISEASES Database	Automated gene-disease/NTD associations, evidence scoring, target ID, public health prioritization	[16]
BindingDB	Measured binding affinities for protein–ligand pairs (Kd, Ki, IC50, etc.), cross-referenced with ChEMBL and UniProt	[14]
UniProt	Protein sequences, functions, annotations, disease relevance, genome-wide mapping, taxonomy—supports NTD target research	[16]
ClinicalTrials.gov	Ongoing and completed clinical NTD trial results, investigational drugs, protocols, outcomes, regulatory status	[16]
PDB (Protein Data Bank)	3D structures of NTD protein-ligand complexes, X-ray/NMR/cryo-EM, ligand binding modes, supports docking and pharmacophore modelling	[18]
Malaria Drug Target Compendium	Manually curated malaria-relevant targets, genomics, supporting literature and links to available chemical probes	[17]

Table 2: Multidimensional molecular descriptors and specialized computational software enabling QSAR development and virtual screening for neglected tropical disease drug candidates.

Descriptor Type	Descriptor Name	Software/Tools	Key Features	Reference
1D Descriptors	Hydrogen Bond Donors/Acceptors (HBD/HBA), Topological Polar Surface Area (TPSA Molecular Weight (MW), Rotatable Bonds.	PaDEL, RDKit, MOE, ChemSketch	Basic molecular properties, Lipinski's Rule of Five compliance, oral bioavailability prediction	[23,24]
2D Descriptors	FCFP6 (Functional Class Fingerprint), ECFP6 (Extended Connectivity Fingerprint), MACCS Keys, Atom Pair Fingerprints	RDKit, CDK (Chemistry Development Kit), MOE, PaDEL	Structural pattern recognition, similarity searching, virtual screening, high predictive accuracy (Pearson r > 0.98 for Chagas studies)	[24]
2D Descriptors	Topological Indices (Wiener, Randić, Zagreb, Kappa Indices)	CODES Software, Dragon, PaDEL	Molecular connectivity information, structure uniqueness representation, QSAR model development	[23,25]
2D Descriptors	Molecular Fingerprints (RDKit, Toxprint, PFP)	RDKit, MOE, ChemAxon, PaDEL	Rapid similarity assessment, scaffolding diversity, chemical space exploration for NTD targets	[23,24]
3D Descriptors	Molecular Surface Area (MSA), Volume (MV), Shape Descriptors (WHIM, GETAWAY indices)	MOE, Dragon, CORINA, SYBYL	Spatial molecular properties, 3D shape similarity, pharmacophore mapping	[23,24]
Quantum Chemical	HOMO-LUMO Gap, Dipole Moment, Polarizability, Electrostatic Potential	MOPAC, Gaussian, Spartan, MOE	Electronic effects on biological activity, reactivity prediction, toxicity assessment	[24,26]
Thermodynamic	Solubility Parameters, Partition Coefficient (LogP), Distribution Coefficient (LogD)	ACD/Labs, MOE, ChemAxon, Marvin	ADMET prediction, membrane permeability, blood-brain barrier penetration assessment	[24]
Electronic Descriptors	Charge Descriptors, Atomic Charges (Gasteiger, MMFF94), Electron Density	Dragon, CODES, MOE, Gaussian	Electrostatic interactions with target proteins, pharmacophore feature identification	[24,25]

Physicochemical	LogP, Molecular Refractivity (MR), Hydrogen Bond Index, Molar Mass	Dragon, PaDEL, RDKit, Marvin	Drug-likeness assessment, permeability prediction, selectivity determination	[24]
Fragment-Based	Substructure Keys, Molecular Fragments, Atom Types	RDKit, MOE, CORINA, PaDEL	Fragment library generation, scaffold analysis, lead optimization guidance	[23,24]
Molecular Complexity	Molecular Complexity Index, Fraction of sp^3 Carbons, Rotatable Bonds Count	RDKit, MOE, ChemAxon	Drug-likeness scoring, synthetic accessibility prediction, complexity assessment	[24,25]
Pharmacophoric Features	H-Bond Acceptors/Donors, Aromatic Rings, Hydrophobic Centers, Charged Groups	MOE, Marvin, SYBYL, Canvas	Virtual screening filter, pharmacophore pattern matching, bioactivity prediction	[27]
Structural Keys	Substruct Fingerprints, Functional Group Counts	RDKit, CDK, Dragon, MOE	Structural pattern identification, compound classification, QSAR feature selection	[23,24]

3. Molecular fingerprints:

Molecular fingerprints function as vital quantitative features that capture and encode the structural, physicochemical, and topological characteristics of molecules in QSAR modelling for neglected tropical diseases (NTDs). These descriptors enable the linkage between molecular features and biological response, aiding in the rational prediction of potential drug leads[19]. There are several types of descriptors which can be used in NTD research, such as 1D descriptors (molecular weight, numbers of atoms), 2D descriptors (topological indices, fingerprints), and 3D descriptors (molecular surface area, electrostatic potential) (Table 2)[20]. The same is true of the research on schistosomiasis, which focuses on different molecular descriptors such as topological, physicochemical, thermodynamic, electronic, and charge descriptors to completely describe compounds[21]. Quantum-derived molecular descriptors, notably the HOMO–LUMO gap and dipole moment, have been integrated, and have increased model performance with drug-like molecules. Sophisticated analytical techniques, including recursive feature elimination (RFE), principal component analysis (PCA) are employed to refine feature selection and improve model robustness to minimize dimensionality and remove redundant features are used to enhance the robustness of the model without compromising model interpretability[22].

4. Utility of QSAR in NTDs:

QSAR modelling has been demonstrated to be a cornerstone in the modern drug designing for NTDs which is allowing researchers to link molecular structure to biological activity with high predictive capability[28]. A well-designed QSAR workflow begins with the assembly of a rigorously curated dataset, typically divided into test set & training set. Training set is used for building the computational model weather test set is utilised for validating predictive model[29]. In the other hand, we may also reveal that the training set helps the model "learn" structure-activity relationships, while the test set is essential for unbiased validation and to prevent overfitting. Robust QSAR models are validated using several statistical parameters through some vital software (**Table:4**). For regression QSAR, the q^2 (cross-validated R^2) measures internal predictivity through techniques like leave-one-out or k-fold cross-validation; values exceeding 0.5 indicate a stable, generalizable model[30]. R^2 (correlation coefficient squared) quantifies how much variance in the experimental data is explained by the model, suppose with values above 0.8 considered highly predictive for both training and validation sets. RMSE (Root Mean Square Error) and MAE (Mean Absolute Error) reflect the average prediction error (**Table:3**); lower values here correspond to greater predictive power and model predictivity usually using leave-one-out cross-validation[33].

$$q^2 = 1 - \frac{\sum_{i=1}^{n} \cdot (y_i^{pred} - y_i^{exp})^2}{\sum_{i=1}^{n} \cdot (y_i^{exp} - y^{\bar{e}xp})^2} \qquad R^2 = 1 - \frac{\sum_{i=1}^{n} \cdot (y_i^{exp} - y_i^{pred})^2}{\sum_{i=1}^{n} \cdot (y_i^{exp} - y^{\bar{e}xp})^2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \cdot (y_i^{exp} - y_i^{pred})^2} \qquad MAE = \frac{1}{n}\sum_{i=1}^{n} \cdot | y_i^{exp} - y_i^{pred} |$$

y_i^{exp}: The experimentally measured response for the *i*-th sample within the training dataset.

y_i^{pred}: The predicted response for the *i*-th sample, estimated through cross-validation.

$y^{\bar{e}xp}$: Mean of the experimental values in the training set.

n: Total number of data samples utilized in the training dataset.

Table 3: Demonstrating Standard Threshold, Interpretation to develop robust QSAR model

Metric	Symbol	Standard Threshold	Interpretation
Cross-Validated R^2	q^2	> 0.5	Acceptable internal model stability; > 0.8 indicates excellent prediction
Correlation Coefficient	R^2	> 0.8 (training); > 0.6 (test)	Explains variance in data; best fit will be indicated by higher values
Root Mean Squared Error (RMSE)	RMSE	< 0.5-1.0 (normalized)	Average prediction error; lower values preferred
Mean Absolute Error	MAE	< 0.3-0.7 (normalized)	Average absolute deviation; reflects model accuracy
Matthews Correlation	MCC	> 0.6	Classification balance; handles imbalanced datasets well
Area Under Curve	AUC	> 0.7	Discrimination power for binary classification; > 0.85 excellent
Applicability Domain	A	Chemical space validity	Ensures predictions within training data representation

Table 4: AI/ML driven QSAR software & their features for NTD therapeutics

QSAR Software/Platform	Key Features	NTD Applications	Reference
MOE (Molecular Operating Environment)	MLR, PLS, SVM, Random Forest, Neural Networks, 3D-QSAR, pharmacophore generation, docking integration	Chagas disease, leishmaniasis, tuberculosis	[30,33]
KNIME Analytics Platform	Modular QSAR workflows, descriptor calculation, model validation, machine learning nodes, k-fold cross-validation	Malaria, schistosomiasis, trypanosomiasis	[29]
Orange3	Visual ML workflows, QSAR modelling, regression/classification, feature selection, interpretable models	Multi-NTD target screening	[29]
PaDEL-Descriptor	Open-source 1D/2D/3D descriptor generation, 1664 descriptors available, integrates with QSAR workflows	Widespread NTD descriptor studies	[23,29]
AutoQSAR (Schrödinger)	Automated model building, multiple algorithms comparison, applicability domain analysis, robust validation	High-throughput NTD screening	[29]
WEKA	Advanced machine learning paradigms incorporating data mining, decision tree algorithms, support vector machines, neural networks, cross-validation, and ROC-based performance evaluation.	NTD classification tasks	[29]
PoseidonQ	Free ML platform, 22 algorithms (MLR, SVM, RF, DNN, etc.), ChEMBL integration, model validation metrics	Accessible NTD research globally	[24]
DragonX (Talete)	5,270+ molecular descriptors, QSAR model development, feature selection tools, chemical diversity analysis	Comprehensive descriptor-based modelling	[23,29]
SIMCA-P+ (Umetrics)	PLS regression, multivariate analysis, variable importance, orthogonal projections, discriminant analysis	Complex NTD biological activity patterns	[29]
CODESSA	Quantum chemical + topological descriptors, QSAR model building, structure-property relationships	Trypanosoma cruzimodelling (72-compound study)	[29]
Open3DQSAR	Free, open-source 3D-QSAR, CoMFA analysis, visualization, molecular alignment	Structure-based NTD inhibitor design	[29]
STATISTICA	Multivariate statistics, QSAR modelling, neural networks, ensemble methods, model comparison	Integrated NTD data analysis	[29]
Tibco Statistica	Advanced machine learning, deep learning integration, model deployment, real-time prediction	Modern NTD computational pipelines	[23,29]
RDKit	Open-source cheminformatics, descriptor generation, molecular fingerprints, QSAR scripting	Foundation for custom NTD workflows	[29]
ChemSpace Intelligence	AI-driven QSAR, generative models, scaffold hopping, chemical space exploration	Novel NTD chemical series discovery	[23,29]

For classification QSAR, particularly relevant in predicting active versus inactive compounds, MCC (Matthews Correlation Coefficient) balances model performance even with imbalanced datasets, while ROC curves and AUC (Area Under the Curve) provide graphical and quantitative assessments of how well the model discriminates between classes[31]. AUC values over 0.7 are indicative of effective models. Applicability Domain (A) analysis further ensures

that model predictions are restricted to chemical spaces with sufficient representation in the training data, thus safeguarding against unreliable extrapolations. In studies on *Trypanosoma cruzi* and *Leishmania*, models built with methods such as Artificial Neural Networks (ANN) and Kernel Partial Least Squares (KPLS) frequently achieve impressive metrics, with q^2 often surpassing 0.8 and AUC values above 0.85[32].

Most AI-ML driven drug discovery and development studies for neglected tropical diseases (NTDs) have emerged over the past two decades, reflecting rapid technological advancement. Earlier efforts focused on quantitative structure-activity relationship (QSAR) modeling, predominantly using statistical techniques to evaluate compounds for diseases such as leishmaniasis and trypanosomiasis prior to 2010, as discussed by Castillo-Garit et al. (2012)[34]. Recent research has expanded to include structure-based, ligand-based, and bioinformatics strategies, targeting novel inhibitors and drug candidates for Mycobacterium tuberculosis, as highlighted in reviews like Alladi et al., (2018)[35].

An influential early use of machine learning to neglected tropical diseases was in the formulation of benznidazole, which functions as a key pharmacological agent in the treatment of Chagas disease (*American trypanosomiasis*). The researchers took advantage of the chitosan microparticles to enhance the pharmacokinetic activity of the drug because the drug has low water solubility and thus limits the oral absorption of the drug. Leonardi and co-authors (2009) used artificial neural networks (ANNs) tosimulate and optimize critical formulation parameters, including encapsulation efficiency, particle size, production yield, and dissolution performances. This method allowed them to determine the best manufacturing conditions and this resulted in better absorption of benznidazole orally. The complex, interacting variables were taken into account simultaneously with the use of ANNs and led to a well-validated model, which proved the usefulness of machine learning in designing improved drug formulations to combat neglected tropical diseases[36].

In a study by Guerra et al.panel of 72 compounds was assessed *in vitro* against the epimastigote form of the Tulahuen 2 strain of *Trypanosoma cruzi*. and followed by modelling the results with machine learning methods. The computational software that was used to generate the molecular descriptive fingerprinting was the CODES software which generates the topological descriptors which capture the structural connectivity and physicochemical characteristics of atoms within a molecule. To prevent overfitting, dimensionality reduction was applied to ensure that the descriptor set was substantially smaller than the number of training samples, thereby preventing model overfitting. A three-layered neural network was trained, with the hidden layer consisting of three to five neurons. Of the 72 compounds, 42 were used for training the model, and the remainder formed the test set to evaluate its predictive ability. The model has shown moderate predictive power of the test set with prediction standard errors and root-mean-square error values of about 0.17 and an area under the receiver operating characteristic curve (AUC) of about 0.7 which is above random predictive power (where 0.5 is random)[37].

A range of trypanocidal activities was represented by a dataset of 363 structurally diverse compounds that was used in this research to build highly predictive QSARs. The optimized ANN model showed great predictive power on the test set, with a q 2 value of 0.81, and was able to identify the most important physicochemical features that are driving the biological activity apparent. Similarly, the most optimal kernel partial least squares (KPLS) model provided a q 2 of 0.84 indicating substantial molecular fragments that were predominantly associated with anti- Trypanosoma cruzi activity in the compound pool. ANN and KPLS analyses as used together contributed to the creation of a favoured fragment library, which establishes a privileged space of physicochemical parameters that are optimum in activity. These structural insights encompassed in this fragment collection, and the described physicochemical profile, offer indispensable information in the rational design of novel antichagasic agents exhibiting superior therapeutic capabilities[38].

5. Deep Learning Models:

Deep learning accelerates NTDs drug discovery by learning rich molecular and phenotypic patterns for faster, cheaper candidate triage and optimization[39]. Convolutional, artificial& graph neural networks enhance virtual screening, de novo design, and drug-target interaction prediction, cutting search space and prioritizing compounds with better potency and ADMET profiles[40]. Transfer learning and AI-driven repurposing rapidly surface approved or clinical compounds for Chagas disease, leishmaniasis, malaria, and TB, shortening timelines to preclinical testing. Image-based deep learning supports diagnosis and endpoint assessment for skin NTDs, improving dataset quality that feeds back into discovery loops[41].

In one of the pioneering studies applying deep learning methodologies to drug discovery, Korotcov and his researcher (2017) assessed how deep neural networks (DNNs) was benchmarked against traditional machine learning models such as naïve Bayes, logistic regression, decision trees, random forests, and support vector machines. They modelled datasets for properties such as aqueous solubility and focusing on infectious diseases including bubonic plague, Chagas disease, tuberculosis, and malaria, where molecular fingerprints were utilized as input descriptors. Their study focused on directly comparing the methods for NTD discovery utility, but did not deploy the models for new compound prediction. Instead, it served as a proof of concept. Active or inactive categories of the biological measurements of each dataset were determined based on dataset-specific thresholds. Datasets were mostly not balanced, meaning that one of the classes was more frequently represented, which also can influence the accuracy of classification. In general, the research fulfilled

the emerging potential of deep learning in NTDs, but had methodological limitations typical of the early work in machine learning[42].

Lane et al. (2018) compared the traditional, Bayesian and deep neural network (DNN) models in predicting tuberculosis drug efficacy with literature data of numerous labs (18,886 compounds). Three concentration thresholds were used to convert quantitative data on activities to binary (active/inactive), and models were trained on five fingerprint types, including ECFP6, FCFP6, MACCS keys, RDKit, and Toxprint descriptors. There was no feature selection and datasets were moderately to highly unbalanced. The findings indicated that Bayesian models had the same result as DNNs on external test sets. The predictive ability of different DNN models compared to other ML techniques also depended on the cutoff point of activities and the presence of a descriptor. Interestingly, AROC, precision, and accuracy were also not necessarily consistent with F1-score and MCC, which demonstrates metric dependency. The researchers found that the models have potential to effectively rank compounds to be used in both in vitro and in vivo testing in the discovery of drugs against tuberculosis[43].

The team of researchers led by Arshadi and Keshavarzi (2020) created DeepMalaria, a new artificial intelligence platform aimed at the discovery of antimalarial agents. This Deep learning model is a graph-based model that was trained on 13,446 anti-plasmodial compounds, publicly available in the GlaxoSmithKline database. To demonstrate the predictive power of the model, the researchers applied it to an independent set of macrocyclic compounds as well as tested its possibilities to discriminate the drug repurposing candidates among approved drugs. These findings were impressive: DeepMalaria was able to find all the compounds that showed nanomolar-level activity and predict 87.5 percent of compounds that showed more than 50 percent inhibition at 1 µM concentration. One of the validated hits DC-9237 showed especially promising properties; it caused inhibition of all the asexual developmental stages of Plasmodium falciparum and was able to clear parasites rather quickly, which made it an outstanding candidate in terms of future therapeutic development and optimization[44].

6. Molecular Docking

Molecular docking is a critical computational tool in drug discovery for neglected tropical diseases, enabling researchers to predict how small molecules interact with target proteins in parasites[45]. This technique accelerates the identification of potential drug candidates by virtually screening large compound libraries, significantly reducing experimental costs and time[46]. Docking simulations provide insights into binding affinities, interaction patterns, and optimal molecular orientations, helping prioritize compounds for further validation. Integration with AI, machine learning, and molecular dynamics simulations enhances prediction accuracy and enables structure-based optimization of lead compounds[47]. Below, we describe some crucial software for molecular docking in case of NTDs application.

Table 5: NTD application key features

Docking Software/Platform	Key Features	NTD Applications	Performance Metrics	Reference
MOE (Molecular Operating Environment)	Flexible ligand docking, rigid protein docking, induced fit, scoring functions (London dG, GBVI/WSA), visualization, post-docking refinement	Chagas (T. cruzicruzain), leishmaniasis, schistosomiasis	Binding energy -9.23 to -12.5 eV	[33,48,50]
AutoDock Vina	Open-source, fast vina scoring, multi-threading, flexible torsion tree, blind docking, Lamarckian genetic algorithm	Multi-NTD targets, virtual screening	RMSD accuracy 0.5-2.0 Å	[49,50]
Schrödinger Glide	High-throughput docking, binding pose prediction, IFD (Induced Fit Docking), MM-GBSA scoring refinement	Tuberculosis, malaria, dengue	ΔG ≈ -25 to -30 kcal/mol	[51,52]
DOCK (UCSF)	Ligand-protein scoring, anchor-based docking, de novo ligand design, rigid and flexible docking	Fragment-based NTD screening	Binding modes ranked by DOCK score	[53]
GOLD (Cambridge Crystallographic)	Genetic algorithm, protein flexibility, multiple binding modes, constraint scoring, ChemScore function	Structure-based NTD lead optimization	Experimental validation >70% accuracy	[53]

Researchers employed molecular docking and machine learning to identify inhibitors of *Schistosoma mansoni* thioredoxin glutathione reductase (SmTGR), a critical enzyme for parasite survival. Virtual screening of compound libraries revealed compound 40 with a MolDock score of -150.251 kcal/mol and hydrogen bond energy of -5.038 kcal/mol, demonstrating potent binding affinity. The lead compound formed five conventional hydrogen bonds with Gly115, Gly118, Thr153, and Tyr138 at distances ranging from 1.536 to 3.300 Å. Newly created derivatives (40a-40l) had even the stronger binding energies (-151.869 to -173.613 kcal/mol), and compound 40j was the most favourable in terms of interactions, which was much better than the lead compound and praziquantel, so these molecules could be regarded as the promising ones to treat schistosomiasis[54].

The authors conducted molecular docking with the help of machine learning, which revealed cruzain, one of the essential cysteine proteases of Trypanosoma cruzi, as an inhibitor. Virtual screening of approved and investigational drugs showed the lomitapide with a binding energy of -9.23 eV, which has high rate of inhibitory capability which is better than the reference benzimidazole drug. Stable binding was confirmed using molecular dynamics simulations (100 ns) with an RMSD value of the range of 2.23 A, suggesting lomitapide and four others (Iodipamide, zafirlukast, netupitant, salmon calcitonin) as a promising treatment of Chagas disease[55].

A large-scale experiment was carried out to identify dengue virus NS5 RNA-dependent RNA polymerase (NS5 RdRp) inhibitors in 1.6 million compounds using machine learning and molecular docking. Both candidates D1 and D8 showed better antiviral activity with an IC_{50} of 13.06 +/- 1.17 and 14.79 +/- 7.76 0 M respectively. Calculations of binding free energy were performed to show that D12 had the most desirable energy of -25.93 kcal/mol, D1 had -25.45 kcal/mol, and D8 had -24.02 kcal/mol. The residues Trp795, Gly351, and Ser796 were found to be the main contributors of the binding stability, and the electrostatic and van der Waals interaction were observed as the major contributors. The long 200 ns molecular dynamics simulation proved the stability of these complexes that justified them as potential antiviral agents in treating dengue (NTDs)[52].

7. Pharmacophore Modelling in NTD Drug Discovery:

A potent computational approach used is pharmacophore modelling, which determines abstract spatial patterns of chemical properties that are critical in biological activity that allows the rapid virtual screening of large chemical libraries in neglected tropical disease[27,29]. Pharmacophore models are defined by structurally diverse active ligands unlike QSAR methods which rely on structurally similar compounds and are therefore invaluable in determining novel chemical scaffolds[56] A major step forward has been made in dynamic hybrid pharmacophore models (DHPMs), which are models that combine the features of the interactions of multiple binding pockets at the same time. These models have been found to have better drug-like properties and higher binding strengths than the conventional pharmacophore models (CPMs) and have been validated using extended molecular dynamics simulations (200+ ns) to discover compounds with higher structural diversity in previously uncharted chemical space. With DHPM screening of 1.6 million compounds, Mycobacterium tuberculosis DapB targets yielded far more leads of diverse leads than CPM methods.[33] Fragment Molecular Orbital (FMO) computations are a source of atomic-level information about the interactions between ligands and receptors that can be used to design selective inhibitors rationally, through the design of their pharmacophore. FMO analysis of Chagas disease showed that the Trypanosoma cruzi DHODH inhibitors need hydrogen bond acceptor pharmacophores at Lys43 and Lys214 and aromatic interaction with the cofactors. These selectivity features, absent in human DHODH, promise reduced side effects[57]. Ligand-based pharmacophore modelling proves particularly effective for echinococcosis and other NTDs where target structures remain unknown. Virtual screening using HipHop algorithms identifies common 3D spatial arrangements of active compounds, achieving 16.13% hit rates (10/62 compounds tested)—substantially surpassing traditional screening[27]. Recently, 3D pharmacophore-based drug repurposing identified betamethasone and doxazosin as promising antischistosomal candidates through FDA approved drug libraries. Both compounds demonstrated dose-dependent reductions in worm burden in mouse models and exhibited sex-specific parasite susceptibility, exemplifying how pharmacophore screening accelerates therapeutic discoveries for resource-limited NTD research globally[55,57].

Table 5: Pharmacophore Features and metrics

Software/ Platform	Key Features	Pharmacophore Features	Performance Metrics	Reference
MOE (Molecular Operating Environment)	Ligand-based and structure-based pharmacophore modeling, HipHop algorithm, multi-alignment pharmacophore generation, 3D visualization	Hydrogen bond donors/acceptors, aromatic rings, hydrophobic centers, charged groups, exclusion volumes	Hit rates 16-40%, screening efficiency 5-10x better than random	-
Discovery Studio (Biovia)	CATALYST module, HypoGen algorithm, Genetic Algorithm optimization, ensemble pharmacophore generation, pose-based pharmacophores	HBD/HBA, hydrophobic features, aromatic rings, ionizable groups, excluded volumes	DHPM superior to CPM; 40% more novel compounds	-
Schrodinger Phase	Ligand-based pharmacophore development, common feature pharmacophore analysis, flexible alignment, consensus pharmacophore	H-bond donors, acceptors, aromatic rings, hydrophobic clusters, charged groups	16.13% experimental hit rate validation	-
SYBYL-X (Certara)	CoMFA/CoMSIA integration, pharmacophore-based alignment, multi-template analysis, 3D QSAR coupling	Shape features, electrostatic maps, steric features, pharmacophoric points	Improved potency prediction vs. QSAR alone	-
LigandScout	Automated pharmacophore generation from ligand-target complexes, Feature Tree algorithm, constraint generation, interactive editor	H-bond donors/acceptors, aromatic rings, hydrophobic regions, positively/negatively charged groups	High-quality 3D pharmacophore models from crystal structures	-

8. Fragment based drug designing NTDs

Fragment-based drug design (FBDD) is a powerful computational and experimental strategy that has emerged as a transformative approach for neglected tropical disease drug discovery[38]. Unlike traditional high-throughput screening (HTS) that screens massive libraries of complex molecules, FBDD focuses on identifying small, low-molecular-weight ligands (MW < 300 Da, < 20 heavy atoms) with weak binding affinity (Kd values in the high µM to low mM range)[58]. These fragment hits can be used as privileged starting points which can be optimized further through fragment linking, growing and merging strategies to create more powerful drug-like compounds[59]. Strategic benefits of FBDD to NTDs are high hits with smaller libraries (100-1000 fragments rather than 100,000 compounds in HTS), effective coverage of chemical space, lower attrition rates in optimization of leads and the ability to explore new binding modes and allosteric sites of challenging drug targets[60].

The study by Vu et al. (2018) utilized native mass spectrometry to screen a fragment library of natural products against 62 drug targets of Plasmodium falciparum. This new method determined 96 natural product fragments of low-molecular-weight binding to 32 suspected malarial proteins. These fragments amazingly show direct growth inhibition of P. falciparum at concentrations that are promising in drug development. In fourteen fragments IC 50 values were below 45 0 mm and some of the fragments had activities similar to known antimalarials. The natural product fragments were identified in a three-dimensional structural diversity format, which allowed them to explore an unexplored chemical space, which could be a great resource in the development of antimalarial leads[60]. Exertier et al. (2025) screened crystallographic fragments against Trypanosoma brucei trypanothione reductase (TbTR) which is a key human African trypanosomiasis target. In XCHEM (Diamond Light Source), TbTR crystals were soaked in fragments of DSIpoised and EubOPEN libraries and eight new fragment hits to different protein binding sites such as the trypanothione binding site and NADPH binding site were identified using X-ray crystallography. The affinities of the fragments were submillimolar to millimolar, which was determined using the surface plasmon resonance (SPR). These fragments have been valuable structural clues to important protein-ligand interactions, and have laid out a basis of rational hit-to-lead optimization to create broad-spectrum trypanocidal agents that are effective against several Leishmania species[60]. A campaign against Schistosoma mansoni dihydroorotate dehydrogenase (SmDHODH) was carried out by De Souza Neto et al. (2024) in an unprecedented campaign of screening of poised 768 molecular fragments through crystallographic fragments screening. This extensive survey found several fragment hits with detailed binding mode information that is an important critical proof-of-concept that FBDD can be used and useful in NTD drug discovery- a domain where the FBDD community had not explored in the past. The generated structural data was used to rationally design bigger and more powerful inhibitors of SmDHODH, a known antischistosomal target[38]. One of the first institutional approaches to FBDD of NTDs is the Center for Research and Advancement in Fragments and Molecular Targets (CRAFT), which was created as the result of a partnership between the University of Sao Paulo and the Federal University of Goianas. CRAFT combines FBDD, artificial intelligence, and structural biology, and curates fragment and target libraries, which are specifically meant to be used in NTD. The center has been targeting parasitic proteins such as Schistosoma and trypanosomatid enzymes in which the AI can be used to optimize fragmentation to speed up the conversion of hits to leads and cut down development cycles and costs. This partnership approach evidences the possibility of scaling FBDD infrastructure to neglect tropical diseases and increase the research capacities of Latin America and the world in general[59].

9. Current updates of NTDs throughout the globe:

Neglected tropical diseases persist as a formidable challenge to global health equity, disproportionately affecting marginalized and low-income populations across tropical and subtropical regions. The most recent estimates byWHO estimates indicate that more than one billion individuals in 149 countries are at risk of contracting one or more neglected tropical diseases, with the greatest burden currently being noted in sub-Saharan Africa, Southeast Asia, and in a few regions of Latin American countries. Integrated intervention strategies comprising a combination of mass drug administration, enhanced vector control and enhanced health systems have made significant gains against a number of diseases within the last decade such as lymphatic filariasis, trachoma, and onchocerciasis. Nevertheless, continuous spread and re-emergence of other infections, especially dengue, leishmaniasis, and schistosomiasis, demonstrate unending susceptibilities due to urban growth and development, climate change, displacement of populations, and lack of surveillance. The 2021- 2030 roadmap of the WHO proposes more intersectoral cooperation, digital mapping of diseases, and equal access to diagnostics, therapeutics, and vaccines to continue the elimination efforts. Simultaneously with the growth of genomics, the field of vector biology, and pharmacological innovation, our comprehension of the dynamics of pathogens and resistance evolution is changing. However, only long-term political goodwill, investment in local health infrastructure, and community engagement can sustain the process of overcoming NTDs to make sure that the agenda on NTDs is consistent with global objectives on universal health coverage and the Sustainable Development Goals.

CONCLUSION

Drug discovery efforts are being revolutionized by computational methods to neglect tropical diseases (NTDs), and AI/ML, QSAR, molecular docking, pharmacophore modelling, and fragment-based design are being combined into synergies. These techniques will speed up the leading identification process, save money, and allow rational selection of drug candidates. Virtual screening of large drug discovery libraries, drug repurpose initiatives, and other AI

scaffolds have found new scaffolds that are active in Chagas disease, leishmaniasis and malaria. Deep learning models such as Deep Malaria exhibit greater predictive capabilities compared to conventional algorithms, and are able to predict nanomolar inhibitors of different chemical spaces. The next areas of development will be multi-omics integration, de novo design using generative AI, and explainable machine learning to increase model interpretability. Combining AlphaFold-predicted protein structures with docking and pharmacophore models facilitates structure-based discovery even of targets that do not have experimentally determined structures. AI-directed fragment-based methods will increase chemical diversity, and quantum machine learning will be able to enhance binding affinity predictions. Such combined computational approaches will be able to cut through the historical impediments of NTD drug development and provide optimized, safe, and effective therapies to underserved groups across the world.

Source of Support: The author(s) received no financial support for the research, authorship, and/or publication of this article

Conflict of Interest: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

REFERENCES

1. Yang GJ, Ouyang HQ, Zhao ZY, Li WH, Fall IS, Djirmay AG, Zhou XN. Discrepancies in neglected tropical diseases burden estimates in China: comparative study of real-world data and Global Burden of Disease 2021 data (2004-2020). Bmj. 2025 Feb 18;388. https://doi.org/10.1136/bmj-2024-080969.
2. Ca J, Kumar PV, Kandi V, Dharshini D, Batchu S, Bhanu P. Neglected tropical diseases: A comprehensive review. Cureus. 2024 Mar 13;16(2). [PMCID: PMC10925483]
3. Martins-Melo FR, Carneiro M, Ramos Jr AN, Heukelbach J, Ribeiro AL, Werneck GL. The burden of neglected tropical diseases in Brazil, 1990-2016: a subnational analysis from the Global Burden of Disease Study 2016. PLoS neglected tropical diseases. 2018 Jun 4;12(6):e0006559. https://doi.org/10.1371/journal.pntd.0006559
4. World Health Organization. Investing to overcome the global impact of neglected tropical diseases: third WHO report on neglected tropical diseases 2015. World Health Organization; 2015 Aug 5.
5. World Health Organization. Global report on neglected tropical diseases 2023. World Health Organization; 2023 Jan 30.
6. Mahmud M, Kaiser MS, Hussain A, Vassanelli S. Applications of deep learning and reinforcement learning to biological data. IEEE transactions on neural networks and learning systems. 2018 Jan 31;29(6):2063-79. https://doi.org/10.1109/tnnls.2018.2790388.
7. Garg P, Singhal G, Kulkarni P, Horne D, Salgia R, Singhal SS. Artificial intelligence–driven computational approaches in the development of anticancer drugs. Cancers. 2024 Nov 20;16(22):3884. https://doi.org/10.3390/cancers16223884.
8. Winkler DA. Use of artificial intelligence and machine learning for discovery of drugs for neglected tropical diseases. Frontiers in Chemistry. 2021 Mar 15;9:614073. https://doi.org/10.3389/fchem.2021.614073.
9. Ballester PJ. Machine learning for molecular modelling in drug design. Biomolecules. 2019 Jun 4;9(6):216. https://doi.org/10.3390/biom9060216.
10. Maliyakkal N, Kumar S, Bhowmik R, Vishwakarma HC, Yadav P, Mathew B. Two-dimensional QSAR-driven virtual screening for potential therapeutics against Trypanosoma cruzi. Frontiers in Chemistry. 2025 Jun 11;13:1600945. https://doi.org/10.3389/fchem.2025.1600945.
11. Mali SN, Saied EM, Dos Santos CB, Neves Cruz J. Medicinal chemistry for neglected tropical diseases using in-vitro, in-vivo and in silico approaches. Frontiers in Chemistry. 2025 Sep 1;13:1689034. https://doi.org/10.3389/fchem.2025.1689034.
12. Greco A, Karki R, Gadiya Y, Deecke C, Zaliani A, Gul S. Pharmacological profiles of neglected tropical disease drugs. Artificial Intelligence in the Life Sciences. 2024 Dec 1;6:100116. https://doi.org/10.1016/j.ailsci.2024.100116.
13. Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, Light Y, McGlinchey S, Michalovich D, Al-Lazikani B, Overington JP. ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic acids research. 2012 Jan 1;40(D1):D1100-7. https://doi.org/10.1093/nar/gkr777.
14. Zdrazil B. Fifteen years of ChEMBL and its role in cheminformatics and drug discovery. Journal of Cheminformatics. 2025 Mar 10;17(1):32.
15. Godinez-Macias KP, Winzeler EA. CACTI: an in silico chemical analysis tool through the integration of chemogenomic data and clustering analysis. Journal of Cheminformatics. 2024 Jul 24;16(1):84.
16. Crowther GJ, Shanmugam D, Carmona SJ, Doyle MA, Hertz-Fowler C, Berriman M, Nwaka S, Ralph SA, Roos DS, Van Voorhis WC, Agüero F. Identification of attractive drug targets in neglected-disease pathogens using an in silico approach. PLoS neglected tropical diseases. 2010 Aug 24;4(8):e804. https://doi.org/10.1371/journal.pntd.0000804.
17. Orozco MI, Moreno P, Guevara M, Abonia R, Quiroga J, Insuasty B, Barreto M, Burbano ME, Crespo-Ortiz MD. In silico prediction and in vitro assessment of novel heterocyclics with antimalarial activity. Parasitology Research. 2024 Jan;123(1):75. https://doi.org/10.1007/s00436-023-08089-7.
18. Cleary E, Jackson MJ, Ledley F. Government as the first investor in biopharmaceutical innovation: evidence from new drug approvals 2010–2019. Institute for New Economic Thinking Working Paper Series. 2020 Aug 5(133).
19. Winkler DA. Sparse QSAR modelling methods for therapeutic and regenerative medicine. Journal of Computer-Aided Molecular Design. 2018 Apr;32(4):497-509. https://doi.org/10.1007/s10822-018-0106-1.
20. Scotti L, Ishiki H, JB Mendonca Junior F, S. da Silva M, T. Scotti M. Artificial neural network methods applied to drug discovery for neglected diseases. Combinatorial chemistry & high throughput screening. 2015 Sep 1;18(8):819-29. https://doi.org/10.2174/1386207318666150803141219.
21. Teles HR, Ferreira LL, Valli M, Coelho F, Andricopulo AD. Hierarchical clustering and target-independent QSAR for antileishmanial oxazole and oxadiazole derivatives. International journal of molecular sciences. 2022 Aug 10;23(16):8898. https://doi.org/10.3390/ijms23168898.
22. Koirala M, Yan L, Mohamed Z, DiPaola M. AI-Integrated QSAR modeling for enhanced drug discovery: from classical approaches to deep learning and structural insight. International Journal of Molecular Sciences. 2025 Sep 25;26(19):9384. https://doi.org/10.3390/ijms26199384.
23. Maliyakkal N, Kumar S, Bhowmik R, Vishwakarma HC, Yadav P, Mathew B. Two-dimensional QSAR-driven virtual screening for potential therapeutics against Trypanosoma cruzi. Frontiers in Chemistry. 2025 Jun 11;13:1600945. https://doi.org/10.3389/fchem.2025.1600945.
24. Kabier M, Gambacorta N, Ciriaco F, Mastrolorito F, Kumar S, Mathew B, Nicolotti O. PoseidonQ: A Free Machine Learning Platform for the Development, Analysis, and Validation of Efficient and Portable QSAR

Models for Drug Discovery. Journal of Chemical Information and Modeling. 2025 Apr 9;65(8):3944-54. https://doi.org/10.1021/acs.jcim.4c02372.

25. Nandi R, Sharma A, Priya A, Kumar D. Integrating traditional QSAR and read-across-based regression models for predicting potential anti-leishmanial azole compounds. Molecular Diversity. 2025 Aug;29(4):3207-31. https://doi.org/10.1007/s11030-024-11070-w.

26. Ejeh S, Otaru HA, Ejeh JE, John J, Ajala A, Omowanle J, Ramu R, Iorhemen RT, Nkole IU, Adawara SN. Computational techniques for analyzing and interpreting some quinoxaline-containing molecules as possible anti-schistosomal agents using docking, pharmacokinetics, drug-likeness, QSAR, and molecular dynamics evaluation. In Silico Research in Biomedicine. 2025 Jan 1;1:100024. https://doi.org/10.1016/j.insi.2025.100024.

27. Liu C, Yin J, Yao J, Xu Z, Tao Y, Zhang H. Pharmacophore-based virtual screening toward the discovery of novel anti-echinococcal compounds. Frontiers in cellular and infection microbiology. 2020 Mar 20;10:118. https://doi.org/10.3389/fcimb.2020.00118.

28. Bernal FA, Schmidt TJ. A QSAR Study for Antileishmanial 2-Phenyl-2, 3-dihydrobenzofurans. Molecules. 2023 Apr 12;28(8):3399. https://doi.org/10.3390/molecules28083399.

29. Scotti L, Muratov E, Speck-Planche A, Scotti MT. Cheminformatics Approaches in Drug Discovery for Neglected Tropical Diseases. Frontiers in Chemistry. 2021 Jun 16;9:719223. https://doi.org/10.3389/fchem.2021.719223.

30. Kleandrova VV, Scotti L, Bezerra Mendonça Junior FJ, Muratov E, Scotti MT, Speck-Planche A. QSAR modeling for multi-target drug discovery: designing simultaneous inhibitors of proteins in diverse pathogenic parasites. Frontiers in Chemistry. 2021 Mar 10;9:634663. https://doi.org/10.3389/fchem.2021.634663.

31. Manjunath Ashwin Desai B, Manjunath Anirudh B, Biju KS, Ramesh V, Biswas P. Deep Learning-based QSAR Model for Therapeutic Strategies Targeting SmTGR Protein's Immune Modulating Role in Host-Parasite Interaction. arXiv e-prints. 2025 Aug:arXiv-2508.

32. Silva de Jesus Passaes AC, Arantes Dantas J, Landim Lopes F, Pereira Sangi D, Girão Albuquerque M, Vataru Nakamura C, Yoneda J. Quinoxalines against Leishmania amazonensis: SAR study, proposition of a new derivative, QSAR prediction, synthesis, and biological evaluation. Scientific Reports. 2023 Oct 24;13(1):18136. https://doi.org/10.1038/s41598-023-45436-1.

33. Melo-Filho CC, Dantas RF, Braga RC, Neves BJ, Senger MR, Valente WC, Rezende-Neto JM, Chaves WT, Muratov EN, Paveley RA, Furnham N. QSAR-driven discovery of novel chemical scaffolds active against Schistosoma mansoni. Journal of chemical information and modeling. 2016 Jul 25;56(7):1357-72. https://doi.org/10.1021/acs.jcim.6b00055

34. Alberto Castillo-Garit J, Abad C, Enrique Rodriguez-Borges J, Marrero-Ponce Y, Torrens F. A review of QSAR studies to discover new drug-like compounds actives against leishmaniasis and trypanosomiasis. Current topics in medicinal chemistry. 2012 Apr 1;12(8):852-65. https://doi.org/10.2174/156802612800166756.

35. Alladi SM. Advances in computational studies of potential drug targets in Mycobacterium tuberculosis. Current topics in medicinal chemistry. 2018 May 1;18(13):1062-74. https://doi.org/10.2174/1568026618666180806163428

36. Leonardi D, Salomón CJ, Lamas MC, Olivieri AC. Development of novel formulations for Chagas' disease: optimization of benznidazole chitosan microparticles based on artificial neural networks. International journal of pharmaceutics. 2009 Feb 9;367(1-2):140-7. https://doi.org/10.1016/j.ijpharm.2008.09.036.

37. Guerra A, Gonzalez-Naranjo P, E. Campillo N, Cerecetto H, Gonzalez M, A. Paez J. Artificial neural networks based on CODES descriptors in pharmacology: identification of novel trypanocidal drugs against Chagas disease. Current computer-aided drug design. 2013 Mar 1;9(1):130-40. https://doi.org/10.2174/1573409911309010012

38. de Souza AS, Ferreira LL, de Oliveira AS, Andricopulo AD. Quantitative structure–activity relationships for structurally diverse chemotypes having anti-Trypanosoma cruzi activity. International Journal of Molecular Sciences. 2019 Jun 8;20(11):2801. https://doi.org/10.3390/ijms20112801.

39. Yotsu RR, Ding Z, Hamm J, Blanton RE. Deep learning for AI-based diagnosis of skin-related neglected tropical diseases: a pilot study. PLoS neglected tropical diseases. 2023 Aug 14;17(8):e0011230. https://doi.org/10.1371/journal.pntd.0011230.

40. Park A, Joo M, Kim K, Son WJ, Lim G, Lee J, Kim JH, Lee DH, Nam S. A comprehensive evaluation of regression-based drug responsiveness prediction models, using cell viability inhibitory concentrations (IC50 values). Bioinformatics. 2022 May 15;38(10):2810-7. https://doi.org/10.1093/bioinformatics/btac177.

41. Korotcov A, Tkachenko V, Russo DP, Ekins S. Comparison of deep learning with multiple machine learning methods and metrics using diverse drug discovery data sets. Molecular pharmaceutics. 2017 Dec 4;14(12):4462-75. https://doi.org/10.1021/acs.molpharmaceut.7b00578.

42. Lane T, Russo DP, Zorn KM, Clark AM, Korotcov A, Tkachenko V, Reynolds RC, Perryman AL, Freundlich JS, Ekins S. Comparing and validating machine learning models for mycobacterium tuberculosis drug discovery. Molecular pharmaceutics. 2018 Apr 19;15(10):4346-60. https://doi.org/10.1021/acs.molpharmaceut.8b00083.

43. Keshavarzi Arshadi A, Salem M, Collins J, Yuan JS, Chakrabarti D. DeepMalaria: artificial intelligence driven discovery of potent antiplasmodials. Frontiers in pharmacology. 2020 Jan 15;10:1526. https://doi.org/10.3389/fphar.2019.01526.

44. Ferreira LG, Dos Santos RN, Oliva G, Andricopulo AD. Molecular docking and structure-based drug design strategies. Molecules. 2015 Jul 22;20(7):13384-421.

45. Opo FA, Rahman MM, Ahammad F, Ahmed I, Bhuiyan MA, Asiri AM. Structure based pharmacophore modeling, virtual screening, molecular docking and ADMET approaches for identification of natural anti-cancer agents targeting XIAP protein. Scientific reports. 2021 Feb 18;11(1):4049. https://doi.org/10.1038/s41598-021-83626-x.

46. Maia EH, Assis LC, De Oliveira TA, Da Silva AM, Taranto AG. Structure-based virtual screening: from classical to artificial intelligence. Frontiers in chemistry. 2020 Apr 28;8:343. https://doi.org/10.3389/fchem.2020.00343.

47. Oliveira Soté W, Comar Junior M. Exploring binding sites in chagas disease protein TcP21 using integrated mixed solvent molecular dynamics approaches. Journal of Chemical Information and Modeling. 2024 Dec 17;65(1):363-77. https://doi.org/10.1021/acs.jcim.4c01927.

48. Hernandez HW, Soeung M, Zorn KM, Ashoura N, Mottin M, Andrade CH, Caffrey CR, de Siqueira-Neto JL, Ekins S. High Throughput and Computational Repurposing for Neglected Diseases: Hernandez et al. Pharmaceutical research. 2019 Feb;36(2):27. https://doi.org/10.1007/s11095-018-2558-3.

49. Singh K, Kaur N, Prabhu A. Machine learning, molecular docking and simulation studies reveal Lomitapide, Lodipamide, Zafirlukast, Netupitant and Salmon Calcitonin as potent drug molecules against Chagas disease. bioRxiv. 2024 Aug 8:2024-08. https://doi.org/10.1101/2024.08.06.606752.

50. Walker M, Hamley JI, Milton P, Monnot F, Kinrade S, Specht S, Pedrique B, Basáñez MG. Supporting drug development for neglected tropical diseases using mathematical modeling. Clinical Infectious Diseases. 2021 Sep 15;73(6):e1391-6. https://doi.org/10.1093/cid/ciab350.

51. Zong K, Wei C, Li W, Wang C, Ruan J, Liu X, Zhang S, Yan H, Cao R, Li X. Identification of novel inhibitors of dengue viral NS5 RNA-dependent RNA polymerase through molecular docking, biological activity evaluation and molecular dynamics simulations. Journal of Enzyme

Inhibition and Medicinal Chemistry. 2025 Dec 31;40(1):2463006. https://doi.org/10.1080/14756366.2025.2463006.

52. de Souza Neto LR, Montoya BO, Brandão-Neto J, Verma A, Bowyer S, Moreira-Filho JT, Dantas RF, Neves BJ, Andrade CH, von Delft F, Owens RJ. Fragment library screening by X-ray crystallography and binding site analysis on thioredoxin glutathione reductase of Schistosoma mansoni. Scientific Reports. 2024 Jan 18;14(1):1582. https://doi.org/10.1038/s41598-024-52018-2.

53. Ja'afaru SC, Uzairu A, Bayil I, Sallau MS, Ndukwe GI, Ibrahim MT, Moin AT, Mollah AM, Absar N. Unveiling potent inhibitors for schistosomiasis through ligand-based drug design, molecular docking, molecular dynamics simulations and pharmacokinetics predictions. Plos one. 2024 Jun 26;19(6):e0302390. https://doi.org/10.1371/journal.pone.0302390.

54. Challapa-Mamani MR, Tomás-Alvarado E, Espinoza-Baigorria A, León-Figueroa DA, Sah R, Rodriguez-Morales AJ, Barboza JJ. Molecular docking and molecular dynamics simulations in related to Leishmania donovani: an update and literature review. Tropical Medicine and Infectious Disease. 2023 Sep 26;8(10):457. https://doi.org/10.3390/tropicalmed8100457.

55. Dobrachinski L, Ferreira LL, Cirino ME, Andrade-de-Siqueira AI, Mafud AC, Mascarenhas YP, Andricopulo AD, de Moraes J. The 3D pharmacophore modeling to explore new antischistosomal agents among US FDA approved drugs. Future Medicinal Chemistry. 2024 Sep 1;16(17):1791-9. https://doi.org/10.1080/17568919.2024.2379231.

56. Yoshino R, Yasuo N, Inaoka DK, Hagiwara Y, Ohno K, Orita M, Inoue M, Shiba T, Harada S, Honma T, Balogun EO. Pharmacophore modeling for anti-chagas drug design using the fragment molecular orbital method. PLoS One. 2015 May 11;10(5):e0125829. https://doi.org/10.1371/journal.pone.0125829.

57. Sternicki LM, Poulsen SA. Fragment-based drug discovery campaigns guided by native mass spectrometry. RSC Medicinal Chemistry. 2024;15(7):2270-85. https://doi.org/10.1039/d4md00273c.

58. Andrade CH, Nonato MC, da Silva Emery F. Introducing CRAFT: The center for research and Advancement in Fragments and molecular Targets. ACS Medicinal Chemistry Letters. 2024 Jul 23;15(8):1174-7. https://doi.org/10.1021/acsmedchemlett.4c00296.

59. Vu H, Pedro L, Mak T, McCormick B, Rowley J, Liu M, Di Capua A, Williams-Noonan B, Pham NB, Pouwer R, Nguyen B. Fragment-based screening of a natural product library against 62 potential malaria drug targets employing native mass spectrometry. ACS infectious diseases. 2018 Feb 13;4(4):431-44. https://doi.org/10.1021/acsinfecdis.7b00197.

60. Exertier C, Antonelli L, Brufani V, Colotti G, Ilari A, Fiorillo A. Expanding the molecular landscape of fragments binding to trypanothione reductase, a legitimate target for drug design against human African trypanosomiasis. Frontiers in Chemical Biology. 2025 May 14;4:1605579. https://doi.org/10.3389/fchbi.2025.1605579.