Original article:

TOWARDS COMBATING ANTIBIOTIC RESISTANCE BY EXPLORING THE QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIP OF NDM-1 INHIBITORS

Tianshi Yu[1,2] ID, Aijaz Ahmad Malik[3] ID, Nuttapat Anuwongcharoen[1] ID,
Warawan Eiamphungporn[2] ID, Chanin Nantasenamat[4*] ID, Theeraphon Piacham[2*] ID

[1] Center of Data Mining and Biomedical Informatics, Faculty of Medical Technology, Mahidol University, Bangkok 10700, Thailand
[2] Department of Clinical Microbiology and Applied Technology, Faculty of Medical Technology, Mahidol University, Bangkok 10700, Thailand
[3] Center of Excellence in Computational Molecular Biology, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand
[4] Streamlit Open Source, Snowflake Inc., USA

* **Corresponding authors:** Chanin Nantasenamat, Streamlit Open Source, Snowflake Inc., USA. E-Mail: hellodataprofessor@gmail.com
  Theeraphon Piacham, Department of Clinical Microbiology and Applied Technology, Faculty of Medical Technology, Mahidol University, Bangkok 10700, Thailand.
  E-mail: theeraphon.pia@mahidol.ac.th

ABSTRACT

The emergence of New Delhi metallo-beta-lactamase-1 (NDM-1) has conferred enteric bacteria resistance to almost all beta-lactam antibiotics. Its capability of horizontal transfer through plasmids, amongst humans, animal reservoirs and the environment, has added up to the totality of antimicrobial resistance control, animal husbandry and food safety. Thus far, there have been no effective drugs for neutralizing NDM-1. This study explores the structure-activity relationship of NDM-1 inhibitors. IC_{50} values of NDM-1 inhibitors were compiled from both the ChEMBL database and literature. After curation, a final set of 686 inhibitors were used for machine learning model building using the random forest algorithm against 12 sets of molecular fingerprints. Benchmark results indicated that the KlekotaRothCount fingerprint provided the best overall performance with an accuracy of 0.978 and 0.778 for the training and testing set, respectively. Model interpretation revealed that nitrogen-containing features (KRFPC 4080, KRFPC 3882, KRFPC 677, KRFPC 3608, KRFPC 3750, KRFPC 4287 and KRFPC 3943), sulfur-containing substructures (KRFPC 2855 and KRFPC 4843), aromatic features (KRFPC 1566, KRFPC 1564, KRFPC 1642, KRFPC 3608, KRFPC 4287 and KRFPC 3943), carbonyl features (KRFPC 1193 and KRFPC 3025), aliphatic features (KRFPC 2975, KRFPC 297, KRFPC 3224 and KRFPC 669) are features contributing to NDM-1 inhibitory activity. It is anticipated that findings from this study would help facilitate the drug discovery of NDM-1 inhibitors by providing guidelines for further lead optimization.

Keywords: Antibiotic resistance, beta-lactamase, NDM-1, QSAR, drug discovery, data science

INTRODUCTION

Antibiotic resistance is defined as bacteria having resistance to the effects of antibiotic agents. It has already been listed in the top 10 global public health burdens by the WHO. Every year in the US, there are more than 2.8 million cases of infection by antibiotic-resistant bacteria, and mortality is more than 35000. From amongst all antibiotics, almost 70% of pathogenic bacteria have developed resistance to at least 1 antibiotic (Bush and Bradford, 2016).

Beta-lactamases are enzymes secreted by bacteria to hydrolyze the most widely used antibiotics: beta-lactam antibiotics, rendering them useless and posing a major threat to antibiotic resistance control. There are two categories of beta-lactamases: serine-active-site and zinc-ion-active-site beta-lactamases, according to the Ambler classification system. Beta-lactamases with serine active sites belong to classes A, C and D and those with zinc-ion active sites belong to class B (Bahr et al., 2021). Because the active sites of class B beta-lactamases are coordinated by metal ions, they are also called metallo-beta-lactamases (MBL). Amongst MBLs, New Delhi Metallo-β-lactamase-1 (NDM-1) was first discovered in *K. pneumoniae* and *E. coli* of a Swedish traveller back home from India in the year 2008. *K. pneumoniae* and *E. coli* harboring NDM-1 have spread widely in the Indian subcontinent, China, Southeast Asian countries, and the Middle East. Until now there are a total of 41 genetic variants of NDM-1, ranging from NDM-1 to NDM-41. In comparison with other significant metallo-beta-lactamases such as Imipenemase (IMP-1) and Verona integron-encoded (VIM-2), NDM-1 is mainly harbored by Enterobacteriaceae species (especially *K. pneumoniae* and *E. coli*), while IMP-1, VIM-2 are harbored by *P. aeruginosa* (Naas et al., 2017; Eiamphungporn et al., 2018). Anchored to the outer membrane of gram-negative bacteria as a lipoprotein, NDM-1 possesses higher tolerance of zinc ion depletion than other MBL members as water-soluble periplasmic enzymes. In addition, NDM-1 has highly efficient hydrolytic capabilities against a wide spectrum of beta-lactams, including carbapenems (González et al., 2016).

Clinical bacterial strains expressing NDM-1 are currently susceptible only to certain antibiotics of last resort: colistin, tigecycline, and fosfomycin (Tooke et al., 2019). All those antibiotics of last resort demonstrate significant side effects, which limit their large-scale application. In addition, NDM-1 and its genetic variants can be transmitted by plasmids, which can be facilitated by accelerated globalization and international travellers. Last but not least, plasmids harboring NDM-1 genes are reported to be identified amongst food animals and the environment (Wang et al., 2012; Szmolka and Nagy, 2013; Islam et al., 2017; Parvez and Khan, 2018). The isolation of blaNDM-1 genes in food animals indicates the zoonotic transmission capability and implicates significant food safety issues. While the environmental spread of blaNDM-1, especially through water, poses more pressure on the environmental containment of antimicrobial resistance. All the above aspects of NDM-1 make infection control rather time and resource-consuming.

Until now, although there is noticeable progress on drug discovery for NDM-1 inhibiting molecules, there are no effective inhibitors against NDM-1 that can be applied in clinical practice. Some conventional beta-lactamase inhibitors, such as clavulanic acid, sulbactam, and tazobactam cannot inhibit NDM-1. There are a number of inhibitors *in vitro*, for example, L-captopril, Aspergillomarasmine A, cyclic boronates, etc (Li et al., 2014; Hecker et al., 2020, Tehrani et al., 2020b). Drug discovery for NDM-1 inhibitors faces a series of difficulties, due to the unique structural characteristics of NDM-1, including flexibility of the binding pocket, adaptability of the active site loop, the open shallow cavity of the active site, water reorganization upon ligand binding, relative lacking of knowledge of whole catalysis process and potential off-target effects to physiological metalloenzymes, etc. All the above make the

drug discovery of NDM-1 inhibitors highly challenging and uncertain (Behzadi et al., 2020).

Quantitative structure-activity relationship (QSAR) is a ligand-based drug discovery approach that harnesses mathematical models for correlating physicochemical property information of molecules with their biological activities. The essence of QSAR is based on two principles: (i) structure dictates activity and (ii) molecules with similar structures demonstrate similar bioactivities (Tropsha, 2010). As a methodology in chemistry and drug discovery, QSAR/QSPR has gone through a remarkable transformation since its dawn 60 years ago. As early as the time of Corwin Hansch, QSAR/QSPR modeling processes were performed on a small number of molecules with few molecular descriptors employing multilinear regression. For now, thanks to the development of information technology and artificial intelligence, QSAR/QSPR has evolved to the application of a large dataset, equipped with sophisticated molecular descriptors, advanced machine learning algorithms, and various validation techniques. QSAR/QSPR modeling techniques are now widely used in chemistry, drug discovery, material science, and environmental protection. And they are not for bioactivities only, but also chemical properties prediction, for example, melting point, biodegradation rate, ecotoxicity, blood-brain-barrier penetration, etc (Nantasenamat and Prachayasittikul, 2015; Nantasenamat et al., 2015; Nantasenamat, 2020). Previously, there are some studies on the structure-activity relationship of specific series of NDM-1 inhibitors, for example, Aspergillomarasmine A derivatives, rhodanines and derived enethiol inhibitors, azetidinimines, dipicolinic acid derivatives (Chen et al., 2017; Zhang et al., 2017, 2018; Romero et al., 2021). However, until now, there is no structure-activity relationship study on the comprehensive library of NDM-1 inhibitors.

The OECD has established principles for QSAR modeling consisting of five rules: defined endpoint, unambiguous algorithms, defined applicability domain, model validation, and mechanistic interpretation. The rules involve various steps of QSAR modeling: data compilation, data splitting, machine learning process, evaluation of the robustness and predictability of the model, and mechanistic interpretation of feature importance (Fjodorova et al., 2008; Tropsha, 2010; Piir et al., 2018). In this study, a QSAR classification model was built according to the OECD criteria using the random forest algorithm for investigating the quantitative structure-activity relationship for NDM-1 inhibitors. After hyperparameter optimization, the model was interpreted by selecting important features for exploring their contributions and mechanisms in NDM-1 inhibiting activities.

MATERIALS AND METHODS

This is a computational study utilizing QSAR modeling for investigating the quantitative structure-activity relationship for NDM-1 inhibitors. QSAR modeling aims to establish a relationship between the intrinsic information of molecules and their endpoint bioactivity values/classes. The design of this study is summarized in schematic diagram in Figure 1.
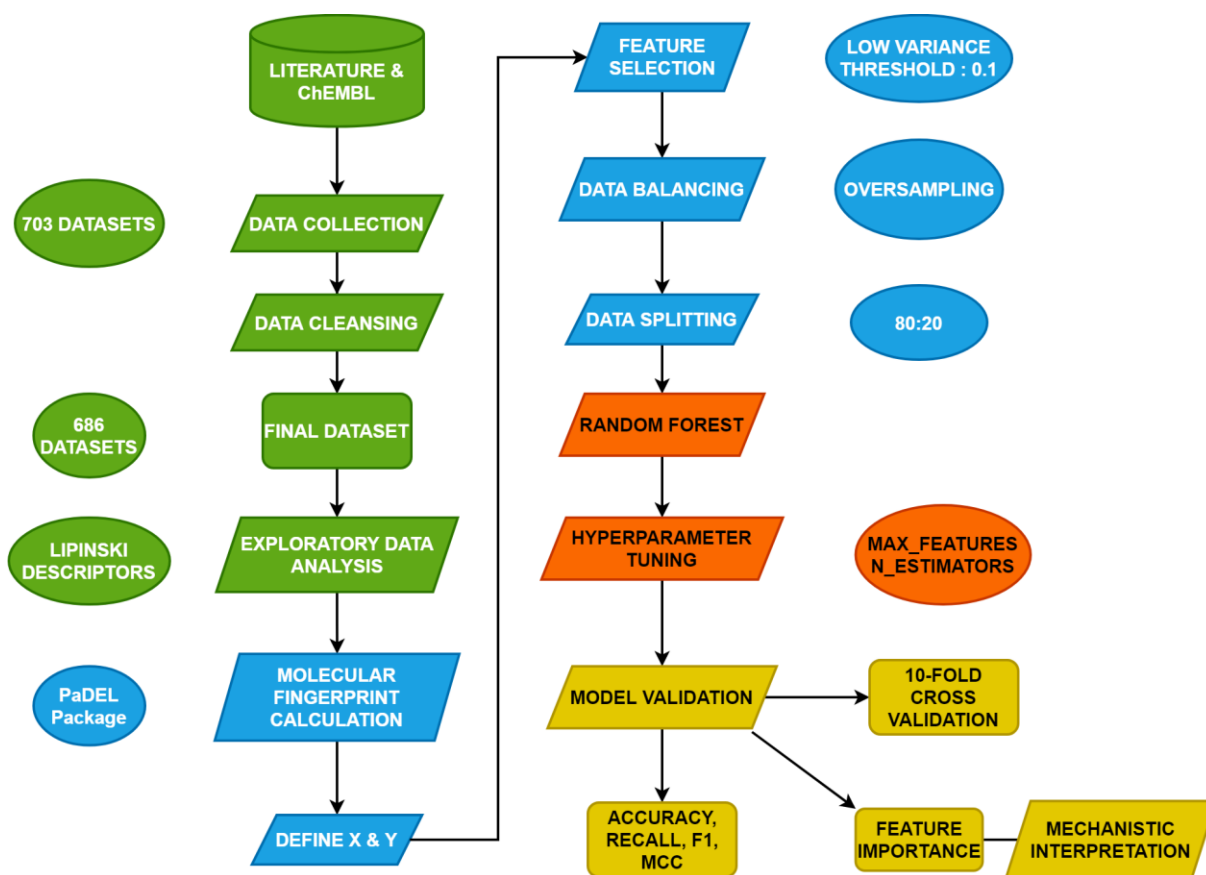
Figure 1: Schematic diagram of the conceptual framework of this study. Different colors symbolize various steps of the study. Green means data collection and data cleansing, blue means descriptor calculation, red means QSAR modeling and orange means validation and mechanistic interpretation.

Data compilation

All biological activity data sets are isolated from the ChEMBL database and a compilation of primary literature. The complete list of origins of data sets are shown in Supplementary Table 1. IC_{50} values are selected exclusively in this study for further investigation. As a result, 703 NDM-1 inhibitor activities have been obtained. After data cleansing which involves removal of redundant data, removal of unqualified data, imputation of missing or unqualified data, there are 686 non-redundant molecules left with available IC_{50} values as the final data sets for modeling.

And above all, for the sake of more straightforward visibility of the bioactivities and interpretability of the values, needs to transform IC_{50} to pIC_{50} values. pIC_{50} is the negative logarithmic value of IC_{50}, and the transformation of value can make the data logarithmic, demonstrating the distribution, diversity, and tendency more visible. Particularly, $IC_{50} \leq 1\mu M$ ($pIC_{50} \leq 6$) were considered as active, while $IC_{50} > 10 \ \mu M$ ($pIC_{50} < 5$) inactive. Those IC_{50} between them are considered intermediate. Shown in Table 1 is the summary of counts of all bioactivity classes.

Table 1: Summary of counts of NDM-1 inhibitors along with bioactivity information. After data balancing, all three classes have the same size of 263.

Class	pIC$_{50}$ values	Count
Inactive	<5.0	263
Intermediate	5.0-6.0	218
Active	≥6.0	205

Molecular fingerprint descriptor calculation

Molecular fingerprints are the representations of a complex form of molecular descriptors. In this study, the PaDEL package was used for calculating 12 sets of molecular fingerprints (Yap, 2011). Details are shown in Table 2.

Klekota-Roth fingerprints is a set of 4860 molecular substructures proposed by Justin Klekota and Fredrick Roth. These substructures originated from chemical libraries and are explored for their privileges in bioactivities. PubChem fingerprints encode molecular fragments information with 881 binary digits

and can be accessed from PubChem. Substructure fingerprint set consists of 307 chemical functional groups by SMARTS patterns. The molecular fingerprint means the presence or absence of any particular fingerprint, and the word count means the number of these particular fingerprints (Yap, 2011). Considering the interpretability for biologists, only KlekotaRoth, KlekotaRothCount, PubChem, Substructure and SubstructureCount fingerprint sets are used for further modeling.

As each set of fingerprints contains tens even hundreds of structural datasets, they can add up to total complexity and give the risk of bias in the model. Therefore, feature selection is necessitated by removing low variance features. By default, all zero variance features should be removed, as zero-variance features have constant value and all instances share the same constant on this feature. The threshold of variance is set to 0.1, which means that features with variance lower than 0.1 are to be removed, without affecting the overall performance of the model, as well as reducing the time and computational resources (Nantasenamat, 2020).

Table 2: All 12 sets of molecular fingerprints in PaDEL software

Fingerprint	Amount	Description
2D AtomsPair	780	Presence of atom pairs at various topological distances
2D AtomsPair-Count	780	Count of atom pairs at various topological distances
Estate	79	Electrotopological state
CDK Extended	1024	Extends fingerprint with additional bits describing ring features
CDK	194	Fingerprint of length 1024 and search depth of 8
CDK Graph-only	1024	Connectivity only
KlekotaRoth	4860	Presence of certain chemical substructures that enrich for biological activities
KlekotaRoth count	4860	Count of certain chemical substructures that enrich for biological activities
MACCS	166	Binary representation of chemical features defined by MACCS keys
PubChem	881	Binary representation of substructures defined by PubChem
Substructure	307	Presence of SMARTS patterns for functional groups
Substructure count	307	Count of SMARTS patterns for functional groups

Data balancing and data splitting

Data balancing is a critical step in classification modeling. Data imbalance occurs when the classes of the dataset are distributed unequally. For classification models, data imbalance is too common to be avoided. Generally, the influence of a mild degree of data imbalance can be ignored, while the influence of significant data imbalance can lead to the unreliability of accuracy. To handle data imbalance, in this study, random oversampling technique is used. Random oversampling by essence is the random duplication of the minority class amongst datasets. Shown in Figure 2 is the comparison between imbalanced data and balanced data. Data splitting is a validation procedure based on the division of the input data set into a training set and a test set (Nantasenamat et al., 2015). In this study, the training set and the testing set will be in the ratio of 80:20. Afterwards, 10-fold cross-validation was performed, as the internal validation procedure, so as to make the most use of the data for ensuring the robustness and reliability of the model.

QSAR modeling and hyperparameter tuning

As a multivariate analysis utilizing machine learning automatic modeling to correlate the independent variables (molecular fingerprint information) with dependent variables (bioactivities), in this study, the QSAR modeling is done by employing a random forest algorithm, through the Jupyter notebook, based on Python scripts. In order to ensure the reproducibility of the model, all the random states during the programming are set to 42 by default.

Hyperparameters of the random forest algorithm can be tuned in order to optimize the model performance. Unlike parameters of a model that can be learned during the training process, hyperparameters are set up before the training begins. In this study, two significant hyperparameters from the random forest algorithm, i.e., the number of decision trees in the forest (n_estimators) and the number of features considered by each tree when splitting a node (max_features) are selected for tuning. The reason as to why they are selected are described as follows. The greater the number of decision trees in the random forest, the better performances they can bring about. However, the increasing number of decision trees can consume more computational resources and slow down the modeling process, without optimizing the model performance significantly. Meanwhile, the number of maximal features by each tree when splitting can impact the model performance to various degrees depending on the size and characters of the dataset. In the next step, GridSearchCV from the scikit-learn Python library is used to evaluate the model performance in a grid-wise manner by exhaustively combining all hyperparameters in the selected range. In this study, the n_estimators hyperparameter is set to 0-700, using 20 as step size. As for the max_features hyperparameter, the employed value range is set to 1-5. The combination of hyperparameters that provide the best performance is selected for modeling.
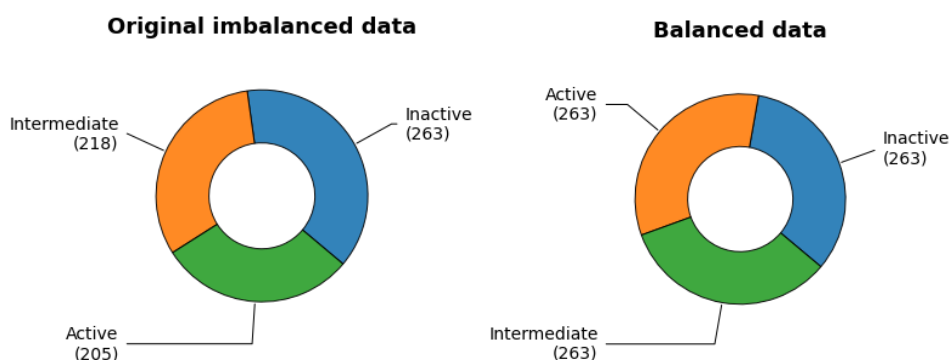

Figure 2: Schematic comparison of original and balanced dataset

Model validation

In order to evaluate the robustness and generalization of the model, validation is performed as an indispensable step of modeling. The measure of validation consists of internal validation and external validation. The internal validation is the validation within the training set, mostly by cross validation. Meanwhile, the external validation is the validation of the holistic model (Shoombuatong et al., 2018).

For classification models, there are a series of parameters to validate the model. In this study, accuracy (ac), recall (re), f1 score (f1) and MCC (Matthew's correlation coefficient) are used for validation. For each set of descriptor/fingerprint, the model will be validated for their performances both in the training set and testing set. The outstanding model will be extracted and further validated through cross validation.

- $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$

- $Recall = \frac{TP}{TP+FN}$

- $Precision = \frac{TP}{TP+FP}$

- $f1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$

- $MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$

Mechanistic interpretation of feature importance

Feature importance refers to assigning a score to the input variables based on how important they are at contributing to the target variable. Random forest algorithm has built-in feature importance of Gini importance (or mean decrease impurity), which is computed from the random forest structure. As random forest algorithm consists of a multitude of decision trees, and each tree can be seen as a set of internal nodes and leaves, the internal nodes determine how to make decisions by dividing datasets into 2 separate subsets. Features for the internal nodes are selected by Gini impurity or information gain in classification model, or variance reduction in regression model. The random forest built-in feature importance is computed during the process of modeling, so that it can save lots of time and computing resources.

Reproducible research

Reproducibility is defined as the 'closeness of the agreement between the results of measurements of the same measure and carried out under changed conditions of measurement'. Reproducibility of the experiment, whether *in vitro* or *in silico*, is a major concern in science and technology as it is closely related to extensibility of knowledge and reproducibility of outputs (Schaduangrat et al., 2020). As this is a computational study, to maintain reproducibility of the model, all the data sets, source codes are uploaded to GitHub repository, and all random seeds are set at 42. All the above information can be accessed at https://github.com/georgey-uricadd/ic50-dataset.

RESULTS AND DISCUSSIONS

The emergence of NDM-1 has conferred pathogenic bacteria almost full spectrum resistance to beta-lactam antibiotics. However, until now there are no available NDM-1 inhibiting drugs that are applied in practice. Due to the gravity of the global antimicrobial resistance burden, there is an urgent need to develop new drugs to inhibit NDM-1. This study is a QSAR study built according to the OECD criteria using the random forest algorithm for investigating the quantitative structure-activity relationship for NDM-1 inhibitors. After using 12 sets of molecular fingerprints from the PaDEL package for classification model, the KlekotaRothCount fingerprint set stands out with the best model performance. After modeling, hyperparameter tuning and feature importance ranking, top 20 ranked fingerprints, including nitrogen-containing features (KRFPC 4080, KRFPC 3882, KRFPC677,

KRFPC3608, KRFPC 3750, KRFPC4287 and KRFPC3943), sulfur-containing substructures (KRFPC 2855 and KRFPC 4843), aromatic features (KRFPC 1566, KRFPC 1564, KRFPC 1642, KRFPC 3608, KRFPC 4287 and KRFPC 3943), carbonyl features (KRFPC 1193 and KRFPC 3025), aliphatic features (KRFPC 2975, KRFPC 297, KRFPC 3224 and KRFPC 669) are used for mechanistic interpretation of feature importance.

As a reproducible QSAR model to explore the structure-activity relationship of NDM-1 inhibitors, this study utilizes a comprehensive dataset source, interpretable learning algorithm and measurable evaluation metrics for building the model. In addition, after mechanistic interpretation of feature importance, the discoveries from the important features largely provide suggestions and directions for further NDM-1 drug discovery. From a larger point of view, the results and discoveries of the study facilitate combat against antimicrobial resistance.

Exploratory data analysis and visualization

An exploratory data analysis is done to visualize the distribution, diversity, and patterns of the NDM-1 inhibitors. Amongst the 686 inhibitors, there are 263 inactive molecules, 205 active molecules, and 218 intermediate molecules. The molecules are evaluated and compared with Lipinski descriptors as shown in Figure 3.

Lipinski's rule-of-five is a statistical principle to evaluate the drug-likeness properties of orally available drugs by Christopher Lipinski (Lipinski et al., 2001). It consists of molecular weight no more than 500 kDa; octanol-water partition coefficient no more than five; hydrogen bond acceptors no more than ten; hydrogen bond donors no more than five. As a statistical principle, the set of rules is not an absolute principle to define or exclude a molecule as a drug. A molecule that meets all criteria of Lipinski's rule doesn't make it a drug, and vice versa, a molecule that violates 1 or 2 rules of Lipinski's rule doesn't exclude its possibility to be a drug. For example, some prominent drugs such as atorvastatin and montelukast, both violate more than one of the Lipinski rules (Bickerton et al., 2012). According to medicinal chemistry, all drugs play pharmaceutical effects via 3 aspects: molecular shape and size; electronic effects, and solubility profiles. Hereby, molecular weight largely determines the molecular size and shape; nHA and nHD are the electronic effects, and LogP corresponds to solubility profiles. In exploratory data analysis, Lipinski's rule of five properties is evaluated and visualized with regard to the bioactivities of the molecules. From Figure 3, it is clear that the distribution of most molecules in the Lipinski descriptor chemical space satisfies Lipinski criteria of drug likeness. All the descriptors are validated to abide by nonparametric distributions.

When performing Mann-Whitney U test between the two classes, molecular weight gets p-value=6.99e-05, LogP gets p-value=5.90e-10 and nHD gets p-value=0.058, while the p-value for nHA is 0.283. The p-values calculated from Mann-Whitney U test indicate the statistical significance of differences between classes. As seen from Table 3, the overall molecular weight and LogP of molecules in the active class are higher than those in the inactive class. While, the number of hydrogen bond acceptors and donors amongst the two classes are not significantly different.

Classification model and hyperparameter tuning

Classification modeling aims to establish the relationship between molecular fingerprints and bioactivity classes. All 12 sets of fingerprints are used to build the model. Random forest algorithm is set by default to include estimators=500, max_features=3 with random state 42. In Table 4, the performance metrics with default hyperparameters is listed. From the table, KlekotaRothCount fingerprint provides the best overall accuracy, recall, F1 score and MCC values. And due to its interpretability for biologists and pharmacists, it is used for the classification model and feature importance.
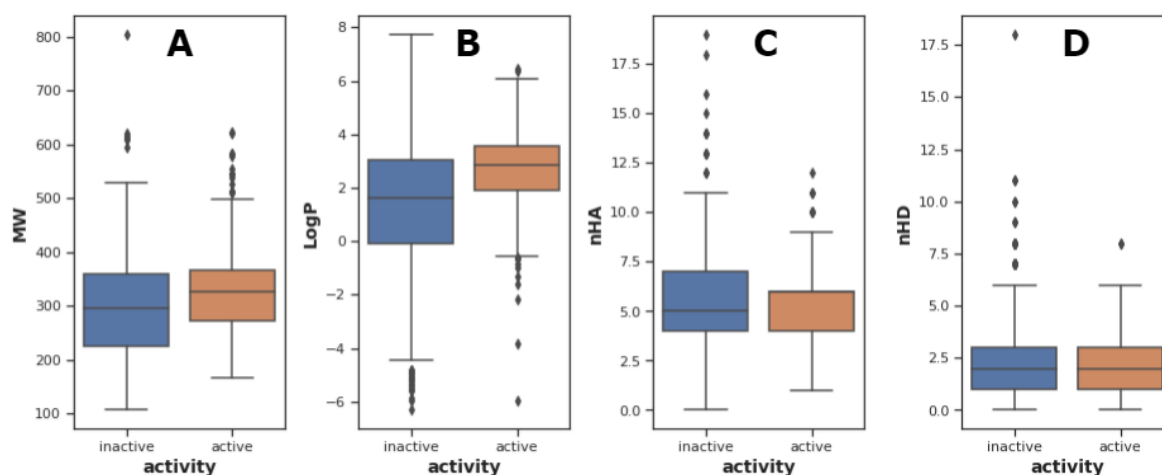
Figure 3: Distribution of Lipinski's descriptors by bioactivities in which all the Lipinski descriptors including **A)** molecular weight (MW), **B)** lipophilicity (LogP), **C)** number of hydrogen bond acceptor atom (nHA) and **D)** number of hydrogen bond donor atom (nHD) are calculated and compared between inactive and active classes.

Table 3: Lipinski descriptor statistics of inactive and active classes

	MW		LogP		nHA		nHD	
	Inactive	Active	Inactive	Active	Inactive	Active	Inactive	Active
Min	106.05	167.02	-6.32	-5.95	0	1	0	0
Max	804.38	622.35	7.74	6.43	19	12	18	8
Median	296.07	325.10	1.60	2.86	5	6	2	2
Mean	303.50	335.92	1.08	2.58	5.75	5.48	2.70	1.99
Skew	0.94	1.01	-0.74	-1.20	1.29	0.33	1.89	0.99

Table 4: Summary of model performance built using 12 sets of molecular fingerprints using default model hyperparameters (n_estimators = 500, max_features = 3, random state=42)

Fingerprint	Training set						External set			
	Ac	Rec	F1	MCC			Ac	Rec	F1	MCC
AtomsPair2D	0.935	0.935	0.935	0.903			0.715	0.715	0.716	0.571
AtomsPair2D Count	0.986	0.986	0.986	0.979			0.766	0.766	0.764	0.651
CDK	0.986	0.986	0.986	0.979			0.766	0.766	0.764	0.651
CDK extended	0.987	0.987	0.987	0.981			0.741	0.741	0.738	0.613
CDK graphonly	0.973	0.973	0.973	0.960			0.791	0.791	0.789	0.69
EState	0.832	0.832	0.832	0.748			0.665	0.665	0.661	0.5
MACCS	0.97	0.97	0.97	0.95			0.722	0.722	0.719	0.586
KlekotaRoth	0.948	0.948	0.948	0.922			0.722	0.722	0.719	0.583
KlekotaRoth Count	0.978	0.978	0.978	0.967			0.766	0.766	0.765	0.650
PubChem	0.956	0.956	0.956	0.934			0.703	0.703	0.699	0.558
Substructure	0.84	0.84	0.84	0.762			0.627	0.627	0.622	0.447
Substructure Count	0.962	0.962	0.962	0.943			0.734	0.734	0.732	0.601

After selecting KlekotaRothCount fingerprint as the one for further modeling, hyperparameters are tuned to optimize the model performance. The results for hyperparameter tuning are demonstrated in Figure 4. The optimal hyperparameter is n_estimators = 150, and max_features = 5. After that, the confusion matrix of its performance is shown in Figure 5.

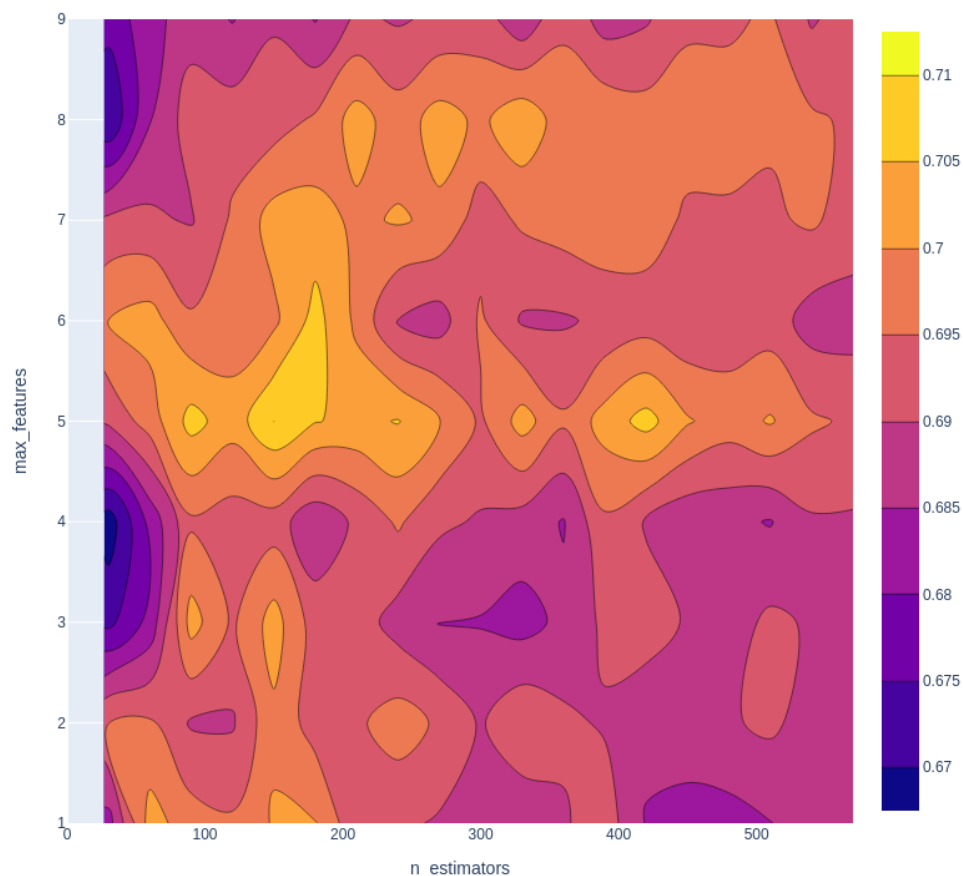Then the optimized parameters are applied to the KlekotaRothCount fingerprint, and the model performance gets improved as shown in Table 5.

Figure 4: Hyperparameter tuning for random forest classifier. The lighter the color, the better the combination of the hyperparameters, from the figure, hyperparameter n_estimators =150, max_features=5 generates the best performance.
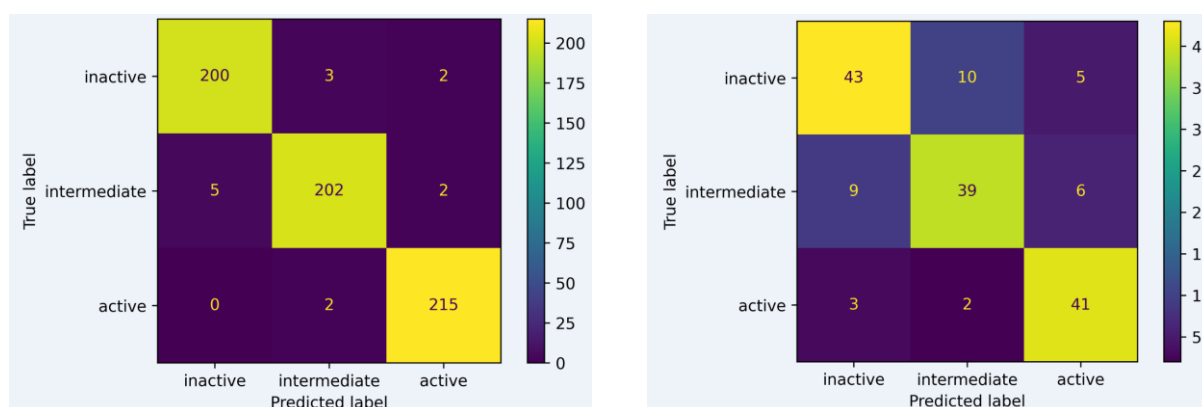

Figure 5: Confusion matrix of experimental versus predicted bioactivity classes of model built using KlekotaRothCount fingerprints as evaluated on training (**A**) and test (**B**) sets after hyperparameter tuning

Table 5: Summary of performance for random forest model built with tuned hyperparameters (n_estimators =150, max_features=5, random state=42).

Fingerprint	Training set				10-fold CV set				External set			
	Ac	*Rec*	*F1*	*MCC*	*Ac*	*Rec*	*F1*	*MCC*	*Ac*	*Rec*	*F1*	*MCC*
KlekotaRoth-Count	0.978	0.978	0.978	0.967	0.753	0.753	0.751	0.633	0.778	0.778	0.777	0.669

Feature importance

Shown in Figure 6 is the intrinsic feature importance ranking of the random forest algorithm. Top 20 ranked features are selected for mechanistic interpretation.

As can be seen from the above feature importance plot, the 20 top-ranked features are listed and described in Table 6. Based on the nature of these KlekotaRothCount fingerprints, which can be categorized as nitrogen-containing features, sulfur-containing features, aromatic features, carbonyl group features, and aliphatic features.

Mechanistic interpretation of feature importance

Mechanistic interpretation of selected features

NDM-1, as a metallo-beta-lactamase, targets amide bonds via nucleophile attack on the carboxyl carbons to hydrolize the substrate antibiotics. NDM-1 has the typical 4-layer sandwich conformation, with 2 β-sheets in the middle, surrounded by 4 α-helixes, and two zinc ions in the active sites (Linciano et al., 2019). Almost all of the zinc ion coordination residues amongst metallo-beta-lactamases are highly conserved. In NDM-1, zinc1 is coordinated by H116, H118, H196, forming a tetrahedral sphere along with a water molecule, and the zinc2 is coordinated by D120, C221, H263, forming trigonal pyramidal sphere along with two water molecules (Linciano et al., 2019). Due to the key roles of zinc ions and the highly conserved zinc-coordinating residues, they have become the primary targets of concurrent NDM-1 inhibitors (Wang et al., 2021).

Inhibitors targeting the active site of NDM-1 work by binding to the zinc ions to form ternary complexes to competitively inhibit the hydrolysis of substrate antibiotics, or by stripping zinc ions from the enzyme so as to prevent the enzyme from launching nucleophile attack on the substrate antibiotics. In this study, NGL viewer, Poseview and LigPlot, are used to visualize and analyze the mechanism of action of NDM-1 inhibitors and confirm the role of selected important features in NDM-1 inhibition (Stierand et al., 2006; Laskowski and Swindells, 2011; Fährrolfes et al., 2017; Rose et al., 2018).

Nitrogen-containing features

From amongst the top 20 ranked features, nitrogen-containing features account for 7 (35 %). They are KRFPC 4080, 3882, 677, 3608, 3750, 4287 and 3943. Amongst the features, KRFPC 677, 3750, 4287 and 3943 belong to amines. Amine is defined as any class of basic organic compounds derived from ammonia by the replacement of hydrogen with alkyl or aryl groups. Amines are nucleophiles that can bond to a variety of electrophiles. Due to the lone pair of electrons, amines are basic in nature, and they can form hydrogen bonds. Shown in Figure 7 is the ligand-enzyme interaction profile for nitrogen-containing feature.

The nitrogen has a lone pair of electrons, with 3 substituents, it may bind to a fourth substituent, leaving a positive charge on the nitrogen atom, which can serve as intermediates for important reactions. Amongst NDM-1 inhibitors, there are a number of significant molecules that contain amines, for example, Aspergillomarasmine A (AMA) and derivatives, methisazone, and derivatives, cefaclor, aryl 2-aminoimidazole derivatives (Linciano et al., 2019). Another feature KRFPC 3608, which is the 2-methylpyridine, is widely seen

in novel categories of anti-NDM 1 agents. For example, N-acylhydrazones and diaryl-substituted thiosemicarbazones contain 2-methylpyridine groups (Gao et al., 2021, Li et al., 2021b). Pyridine is a heterocycle with nitrogen. Due to the nitrogen in the ring, which results in relatively lower electron density of the carbon atoms of the ring, pyridine-containing molecules undergo nucleophilic substitution reactions more easily than corresponding benzene derivatives. The feature is present in a few zinc chelators, such as tris (2-pyridylmethyl)amine (TPA) and N,N,N',N'-

tetrakis (2-pyridylmethyl)ethylenediamine (TPEN) (Schnaars et al., 2018; He et al., 2020). The tris(2-pyridylmethyl) amine (TPA) is a zinc chelator and its derivatives are synthesized as NDM-1 inhibitors by stripping the zinc ions in the active site. This is a tripodal ligand scaffold that is widely used in the coordination and chelation of zinc ions (Huang et al., 2013). Within the trigonal ligand scaffold, there are three identical 2-methylpyridine groups as the metal coordinator. The 2-methylpyridine groups are important in forming the coordination complex.
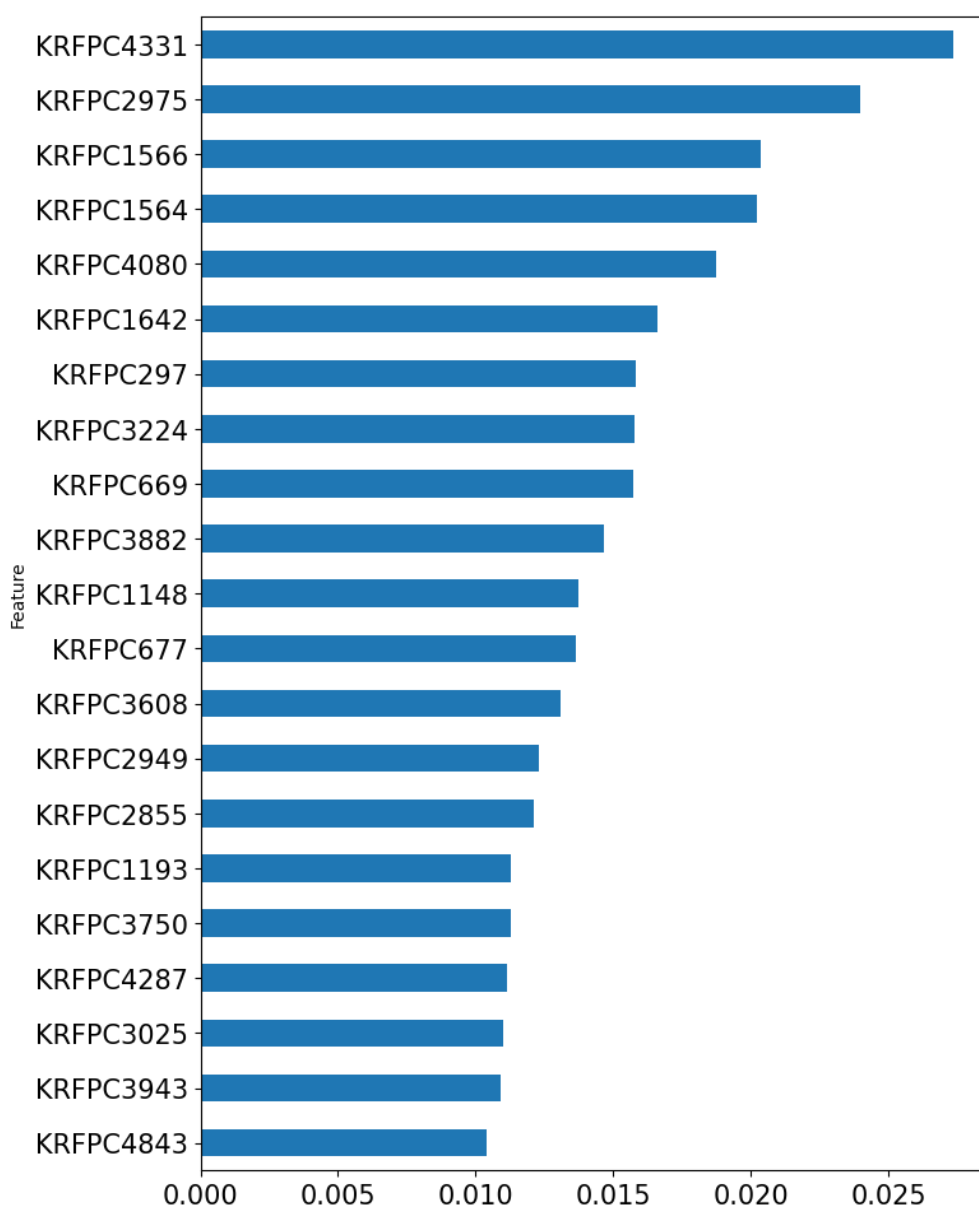

Figure 6: Bar plot of feature importance of the classification model. The top 20 KlekotaRothCount fingerprints are demonstrated in the horizontal bar chart.

Table 6: Feature importance from RF model built using KlekotaRothCount fingerprint

Feature name	SMARTS pattern	Description
KRFPC 4331	O	Oxygen
KRFPC 2975	C	Carbon
KRFPC 1566	[!#1]c1[cH][cH][cH][cH]c1[!#1]	1,2-substituted benzene
KRFPC 1564	[!#1]c1[cH][cH][cH][cH]1	Phenyl group
KRFPC 4080	N	Nitrogen
KRFPC 1642	[!#1]c1[cH][cH]c([!#1])[cH][cH]1	1,4-substituted benzene
KRFPC 297	[!#1][CH2][!#1]	Aliphatic carbon
KRFPC 3224	CC	Carbon-carbon bond
KRFPC 669	[!#1][CH3]	Methyl group
KRFPC 3882	CN	Carbon-nitrogen bond
KRFPC 1148	[!#1][OH]	Hydroxyl group
KRFPC 677	[!#1][NH][!#1]	Secondary amine
KRFPC 3608	Cc1ccccn1	2-methylpyridine
KRFPC 2949	[OH]	Hydroxyl group
KRFPC 2855	[!#1]S[!#1]	Sulfide
KRFPC 1193	[!#1]C(=O)[!#1]	Ketone
KRFPC 3750	CCN	Ethylamine
KRFPC 4287	NCc1ccccc1	Benzylamine
KRFPC 3025	C=O	Carbonyl group
KRFPC 3943	CNc1ccccc1	N-methylaniline
KRFPC 4843	S	Sulfur

Sulfur-containing features

There are 2 sulfur-containing features: KRFPC 2855 and KRFPC 4843. KRFPC 2855 is the sulfur with two R groups, while KRFPC 4843 is the indication of sulfur in the molecule. The presence of sulfur in an organic molecule is indicative of an organosulfur compound. Amongst NDM-1 inhibitors, there are many of them containing the sulfur, such as ʟ-captopril, ᴅ-captopril, thiol derivatives, mercapto acid derivatives, thioester derivatives, etc (Li et al., 2014; Zhang et al., 2019; Ma et al., 2021). Thiol group is capable of interacting directly with the zinc ion to cripple the enzyme. While for sulfur with two R groups, sulfide group, it can form hydrogen bonds with residues to stabilize the complex as shown in Figure 8. For example, molecular docking has indicated that sulfur on the amino acid thioester derivative 1 can form hydrogen bond with serine 223, which facilitates the complex and affinity (Zhang et al., 2019).

Aromatic features

There are 6 features that contain aromatic rings. They are KRFPC 1566, 1564, 1642, 3608, 4287 and 3943. KRFPC 1566, 1564 and 1642 are phenyl groups of various substitutions. The presence of phenyl groups amongst NDM-1 inhibitors is very common. For example, ebselen, some thiol-containing compounds, cyclic boric acid derivatives, and pyridinedicarboxylic acid derivatives (Li et al., 2014). The presence of benzene rings is very important for forming hydrophobic interactions and electrostatic interactions with NDM-1 active site residues. For example, in rhodanine derivatives, the benzene rings can form hydrophobic interactions with Gln60, Trp87, Gly160, these interactions facilitate the binding to the active sites to make it more potently anchored (Xiang et al., 2018). On the other hand, benzene rings can affect the water solubility of compounds. The presence of benzene rings can increase the hydrophobicity of molecules. The latter three aromatic features all contain nitrogens, and have been discussed in nitrogen-containing features.
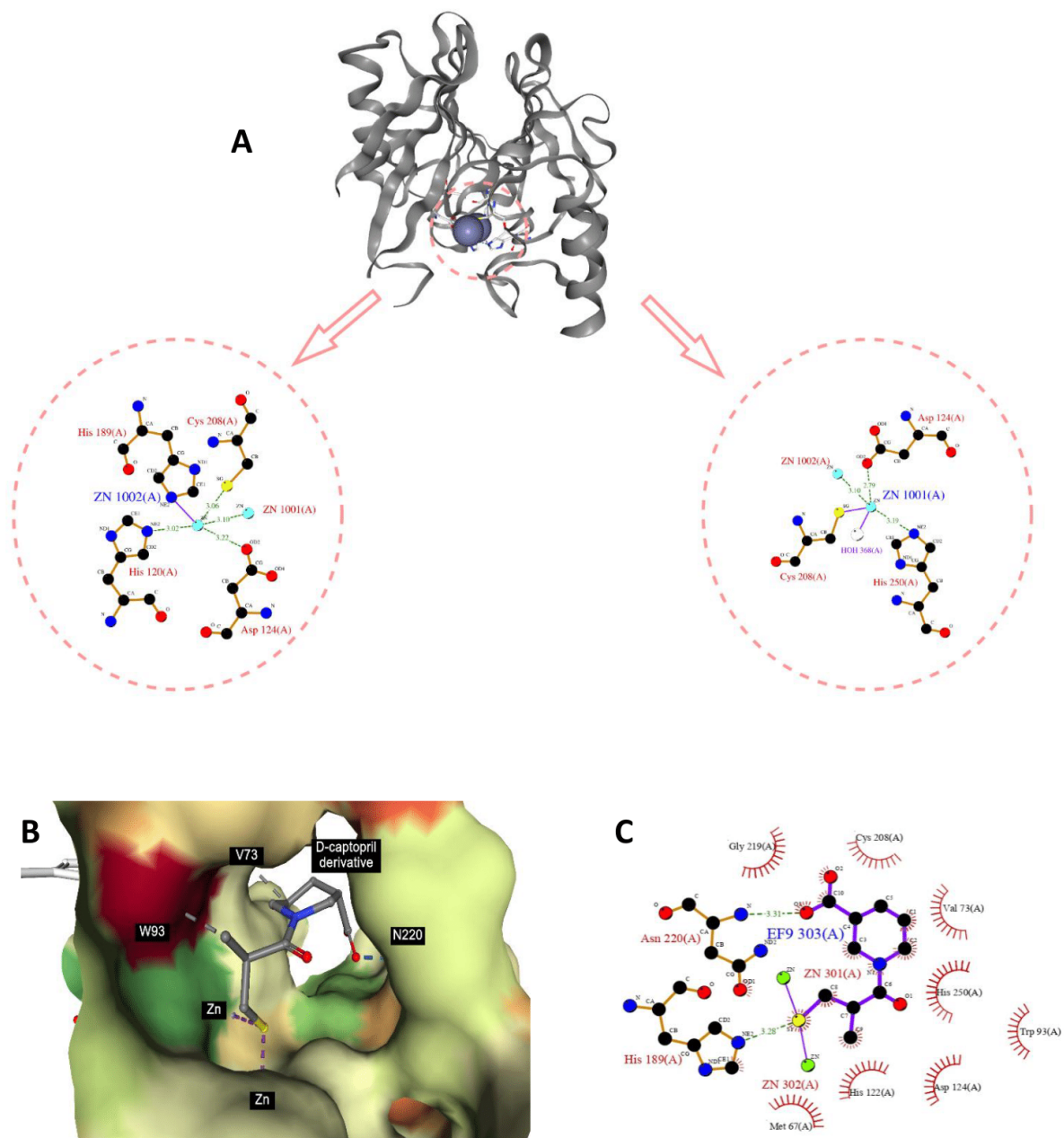
Figure 7: NDM-1 crystal structure and active site. From PDB 3S0Z (**A**). The binding modality of ᴅ-captopril derivative against the active site of NDM-1 (PDB ID: 6LJ4) (**B**) and 2D interaction diagram generated by Ligplot+ (**C**), which highlight the role of nitrogen atom and the adjacent aliphatic carbons in stabilization of the protein-ligand complex.
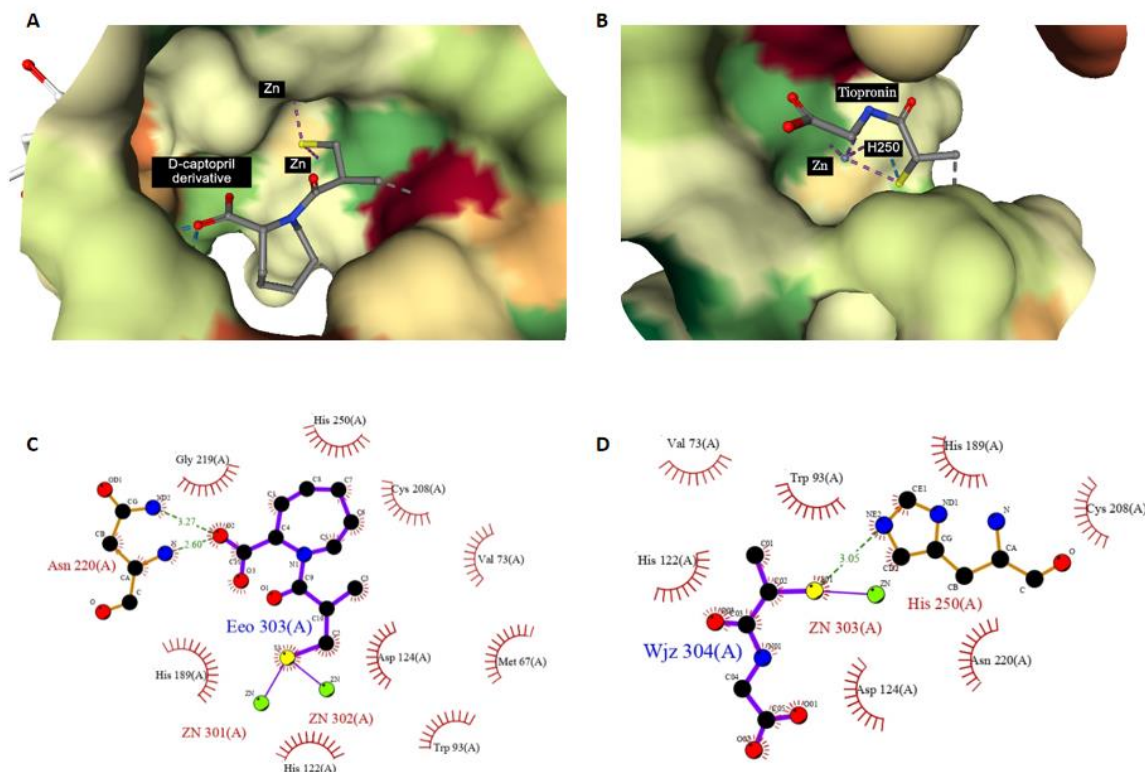
Figure 8: Ligplot illustration of thiol group in ligand/NDM-1 interaction. **A** and **C** are the D-captopril derivative from PDB 6LIZ, **B** and **D** are the tiopronin from PDB 5A5Z.

Carbonyl features

There are two carbonyl features, KRFPC 1193 and KRFPC 3025. Due to the high polarity of the carbonyl bond, they are polar and hydrogen bond acceptors. Carbonyl groups exist in ketones, aldehydes or carboxylic acids. The carbon-oxygen double bond forms hydrogen bond with adjacent NDM-1 active site residues as shown in Figure 9.

Aliphatic features

KRFPC 2975, 297, 3224, 669 belong to aliphatic features. KRFPC 669 is the methyl group. As a non-polar functional group, it is important in forming non-polar covalent bonds, or hydrophobic interactions with adjacent residues as shown in Figure 10.

Limitation of the study

First and foremost, as this is a pure computational study on drug discovery, there's a lack of *in vitro* experiment support. Secondly, although the study utilizes data from reliable literature and accredited chemical databases, the scale and diversity of the dataset are not large and broad enough to cover more portions of chemical space. Last but not least, currently, there are some novel categories of compounds, such as peptidomimetics, oligomer nucleic acid drugs, nanoparticles (Sully et al., 2017; Kazi et al., 2020), etc that may play roles as NDM-1 inhibitors, this study doesn't involve these novel compounds. This is also due to the lack of related data in this field.
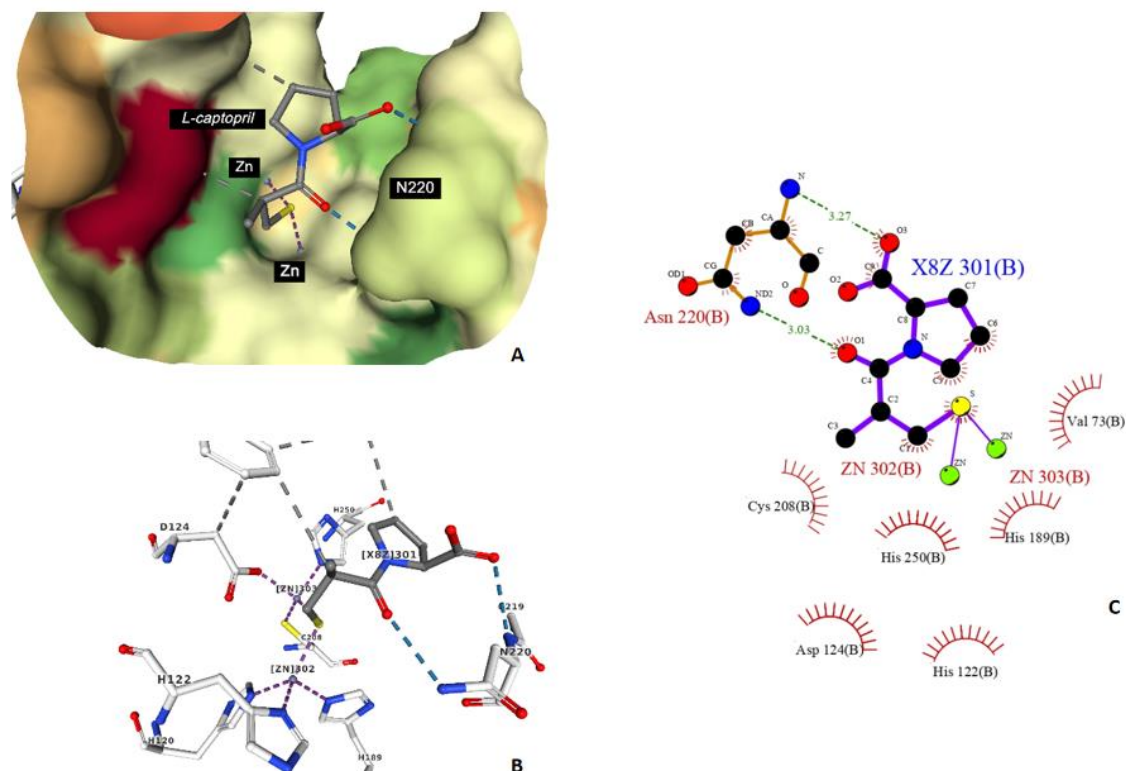
Figure 9: The binding modality of L-captopril against the active site of NDM-1 (PDB ID: 4EXS) (**A** and **B**) and 2D interaction diagram generated by Ligplot+ (**C**), the carbonyl groups act as hydrogen bond acceptor with Asn220 to facilitate the pocket formation.
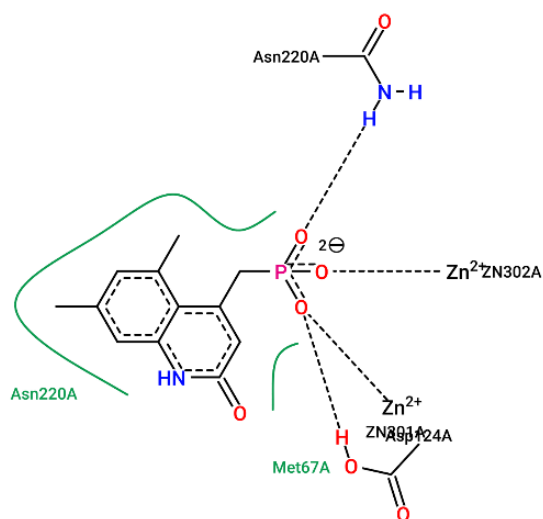

Figure 10: Ligplot illustration of methyl group in ligand/NDM-1 interaction. This is heteroaryl phosphonate derivative with NDM-1, from PDB 6D1J. The two methyl groups form hydrophobicity with Asn220A, marked by the green spline section.

CONCLUSION

Antimicrobial resistance is a significant global challenge. NDM-1 as a metallo-beta-lactamase is one of the most perplexing factors in public health and infection control by causing antimicrobial resistance. However, until now, there are no effective drugs for NDM-1 and the structure-activity relationship remains largely unknown. In this study, a QSAR model using currently available NDM-1 inhibiting compounds has been made, via a random forest algorithm and 12 sets of molecular fingerprints from the PaDEL package. And metrics for model performance evaluation have indicated that KlekotaRothCount fingerprint set provides the best performance and model robustness. Feature importance ranking after hyperparameter tuning has demonstrated that amine group, 2-methylpyridine group, sulfur, aromatic features, carbonyl groups and certain aliphatic hydrocarbons are the features that contribute to NDM-1 inhibiting activities. The findings from this

study can facilitate drug discovery of NDM-1 inhibitors and can be a guideline for further optimization.

Conflict of interest

The authors declare no conflict of interest.

REFERENCES

Bahr G, González LJ, Vila AJ. Metallo-β-lactamases in the age of multidrug resistance: from structure and mechanism to evolution, dissemination, and inhibitor design. Chem Rev. 2021;121:7957-8094.

Behzadi P, García-Perdomo HA, Karpiński TM, Issakhanian L. Metallo-ß-lactamases: a review. Mol Biol Rep. 2020;47:6281-94.

Bickerton GR, Paolini GV, Besnard J, Muresan S, Hopkins AL. Quantifying the chemical beauty of drugs. Nat Chem. 2012;4:90-8.

Brem J, van Berkel SS, Zollman D, Lee SY, Gileadi O, McHugh PJ, et al. Structural basis of metallo-β-lactamase inhibition by captopril stereoisomers. Antimicrob Agents Chemother. 2016a;60:142-50.

Brem J, Cain R, Cahill S, McDonough MA, Clifton IJ, Jiménez-Castellanos J-C, et al. Structural basis of metallo-β-lactamase, serine-β-lactamase and penicillin-binding protein inhibition by cyclic boronates. Nat Commun. 2016b;7:12406.

Büttner D, Kramer JS, Klingler F-M, Wittmann SK, Hartmann MR, Kurz CG, et al. Challenges in the development of a thiol-based broad-spectrum inhibitor for metallo-β-lactamases. ACS Infect Dis. 2018;4:360-72.

Bush K, Bradford PA. β-lactams and β-lactamase inhibitors: An overview. Cold Spring Harb Perspect Med. 2016,6(8):a025247.

Cain R, Brem J, Zollman D, McDonough MA, Johnson RM, Spencer J, et al. *In silico* fragment-based design identifies subfamily B1 metallo-β-lactamase inhibitors. J Med Chem. 2018;61:1255-60.

Chen AY, Thomas PW, Stewart AC, Bergstrom A, Cheng Z, Miller C, et al. Dipicolinic acid derivatives as inhibitors of New Delhi Metallo-β-lactamase-1. J Med Chem. 2017;60:7267-83.

Chen AY, Thomas CA, Thomas PW, Yang K, Cheng Z, Fast W, et al. Iminodiacetic Acid as a novel metal-binding pharmacophore for new delhi metallo-β-lactamase inhibitor development. ChemMedChem. 2020; 15:1272-82.

Chen C, Liu Y, Zhang Y-J, Ge Y, Lei J-E, Yang K-W. The assemblage of covalent and metal binding dual functional scaffold for cross-class metallo-β-lactamases inhibition. Future Med Chem. 2019;11:2381-94.

Chen F, Bai M, Liu W, Kong H, Zhang T, Yao H, et al. H2dpa derivatives containing pentadentate ligands: An acyclic adjuvant potentiates meropenem activity *in vitro* and in vivo against metallo-β-lactamase-producing Enterobacterales. Eur J Med Chem. 2021;224: 113702.

Chigan J-Z, Hu Z, Liu L, Xu Y-S, Ding H-H, Yang K-W. Quinolinyl sulfonamides and sulphonyl esters exhibit inhibitory efficacy against New Delhi metallo-β-lactamase-1 (NDM-1). Bioorg Chem. 2022a;120: 105654.

Chigan J-Z, Li J-Q, Ding H-H, Xu Y-S, Liu L, Chen C, et al. Hydroxamates as a potent skeleton for the development of metallo-β-lactamase inhibitors. Chem Biol Drug Des. 2022b;99:362-72.

Christopeit T, Leiros H-KS. Fragment-based discovery of inhibitor scaffolds targeting the metallo-β-lactamases NDM-1 and VIM-2. Bioorg Med Chem Lett. 2016;26:1973-7.

Cui D-Y, Yang Y, Bai M-M, Han J-X, Wang C-C, Kong H-T, et al. Systematic research of H2dedpa derivatives as potent inhibitors of New Delhi metallo-β-lactamase-1. Bioorg Chem. 2020;101:103965.

Eiamphungporn W, Schaduangrat N, Malik AA, Nantasenamat C. Tackling the antibiotic resistance caused by class A *β*-lactamases through the use of *β*-lactamase inhibitory protein. Int J Mol Sci. 2018;19 (8):2222.

Fährrolfes R, Bietz S, Flachsenberg F, Meyder A, Nittinger E, Otto T, et al. ProteinsPlus: a web portal for structure analysis of macromolecules. Nucleic Acids Res. 2017;45:W337-43.

Farley AJM, Ermolovich Y, Calvopiña K, Rabe P, Panduwawala T, Brem J, et al. Correction to "Structural Basis of Metallo-β-lactamase Inhibition by N-Sulfamoylpyrrole-2-carboxylates". ACS Infect Dis. 2021; 7:3124.

Fjodorova N, Novich M, Vrachko M, Smirnov V, Kharchevnikova N, Zholdakova Z, et al. Directions in QSAR modeling for regulatory uses in OECD member countries, EU and in Russia. J Environ Sci Health C Environ Carcinog Ecotoxicol Rev. 2008;26:201-36.

Gao H, Li J-Q, Kang P-W, Chigan J-Z, Wang H, Liu L, et al. N-acylhydrazones confer inhibitory efficacy against New Delhi metallo-β-lactamase-1. Bioorg Chem. 2021;114:105138.

Ge Y, Xu LW, Liu Y, Sun LY, Gao H, Li JQ, Yang K. Dithiocarbamate as a valuable scaffold for the inhibition of metallo-β-lactmases. Biomolecules. 2019;9 (11):699.

Ge Y, Kang P-W, Li J-Q, Gao H, Zhai L, Sun L-Y, et al. Thiosemicarbazones exhibit inhibitory efficacy against New Delhi metallo-β-lactamase-1 (NDM-1). J Antibiot. 2021;74:574-9.

González LJ, Bahr G, Nakashige TG, Nolan EM, Bonomo RA, Vila AJ. Membrane anchoring stabilizes and favors secretion of New Delhi metallo-β-lactamase. Nat Chem Biol. 2016;12:516-22.

Grigorenko VG, Khrenova MG, Andreeva IP, Rubtsova MY, Lev AI, Novikova TS, Detusheva EV, Fursova NK, Dyatlov IA, Egorov AM. Drug repurposing of the unithiol: inhibition of metallo-β-lactamases for the treatment of carbapenem-resistant gram-negative bacterial infections. Int J Mol Sci. 2022;23(3): 1834.

Guo Y, Yang Y, Xu X, Li L, Zhou Y, Jia G, et al. Metallo-β-lactamases inhibitor fisetin attenuates meropenem resistance in NDM-1-producing Escherichia coli. Eur J Med Chem. 2022;231:114108.

He S, Zou Y, Zhan M, Guo Q, Zhang Y, Zhang Z, et al. Zinc Chelator N,N,N',N'-Tetrakis(2-Pyridylmethyl)Ethylenediamine Reduces the Resistance of Mycobacterium abscessus to Imipenem. Infect Drug Resist. 2020;13:2883-90.

He Y, Zhou S, Sun W, Li Q, Wang J, Zhang J. Emerione A, a novel fungal metabolite as an inhibitor of New Delhi metallo-β-lactamase 1, restores carbapenem susceptibility in carbapenem-resistant isolates. J Glob Antimicrob Resist. 2022;28:216-222.

Hecker SJ, Reddy KR, Lomovskaya O, Griffith DC, Rubio-Aparicio D, Nelson K, et al. Discovery of Cyclic Boronic Acid QPX7728, an Ultrabroad-Spectrum Inhibitor of Serine and Metallo-β-lactamases. J Med Chem. 2020;63:7491-507.

Hinchliffe P, Tanner CA, Krismanich AP, Labbé G, Goodfellow VJ, Marrone L, et al. Structural and kinetic studies of the potent inhibition of metallo-β-lactamases by 6-Phosphonomethylpyridine-2-carboxylates. Biochemistry. 2018;57:1880-92.

Hu L, Yang H, Yu T, Chen F, Liu R, Xue S, et al. Stereochemically altered cephalosporins as potent inhibitors of New Delhi metallo-β-lactamases. Eur J Med Chem. 2022;232:114174.

Huang Z, Zhang X, Bosch M, Smith SJ, Lippard SJ. Tris(2-pyridylmethyl)amine (TPA) as a membrane-permeable chelator for interception of biological mobile zinc. Metallomics. 2013;5:648-55.

Islam MA, Islam M, Hasan R, Hossain MI, Nabi A, Rahman M, Goessens WHF, Endtz HP, Boehm AB, Faruque SM. Environmental spread of New Delhi metallo-β-lactamase-1-producing multidrug-resistant bacteria in Dhaka, Bangladesh. Appl Environ Microbiol. 2017;83(15):e00793-17.

Jackson AC, Zaengle-Barone JM, Puccio EA, Franz KJ. A cephalosporin prochelator inhibits New Delhi metallo-β-lactamase 1 without removing zinc. ACS Infect Dis. 2020;6:1264-72.

Jackson AC, Pinter TBJ, Talley DC, Baker-Agha A, Patel D, Smith PJ, et al. Benzimidazole and benzoxazole zinc chelators as inhibitors of metallo-β-lactamase NDM-1. ChemMedChem. 2021;16:654-61.

Jin WB, Xu C, Cheung Q, Gao W, Zeng P, Liu J, et al. Bioisosteric investigation of ebselen: Synthesis and *in vitro* characterization of 1,2-benzisothiazol-3(2H)-one derivatives as potent New Delhi metallo-β-lactamase inhibitors. Bioorg Chem. 2020;100:103873.

Kang P-W, Su J-P, Sun L-Y, Gao H, Yang K-W. 3-Bromopyruvate as a potent covalently reversible inhibitor of New Delhi metallo-β-lactamase-1 (NDM-1). Eur J Pharm Sci. 2020;142:105161.

Kazi MI, Perry BW, Card DC, Schargel RD, Ali HB, Obuekwe VC, et al. Discovery and characterization of New Delhi metallo-β-lactamase-1 inhibitor peptides that potentiate meropenem-dependent killing of carbapenemase-producing Enterobacteriaceae. J Antimicrob Chemother. 2020;75:2843-51.

Koteva K, Sychantha D, Rotondo CM, Hobson C, Britten JF, Wright GD. Three-dimensional structure and optimization of the metallo-β-lactamase inhibitor Aspergillomarasmine A. ACS Omega. 2022;7:4170-84.

Krasavin M, Zhukovsky D, Solovyev I, Barkhatova D, Dar'in D, Frank D, et al. RhII -Catalyzed De-symmetrization of Ethane-1,2-dithiol and Propane-1,3-dithiol yields metallo-β-lactamase inhibitors. ChemMedChem. 2021;16:3410-7.

Laskowski RA, Swindells MB. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. J Chem Inf Model. 2011;51:2778-86.

Legru A, Verdirosa F, Hernandez J-F, Tassone G, Sannio F, Benvenuti M, et al. 1,2,4-Triazole-3-thione compounds with a 4-ethyl alkyl/aryl sulfide substituent are broad-spectrum metallo-β-lactamase inhibitors with re-sensitization activity. Eur J Med Chem. 2021; 226:113873.

Li J-Q, Gao H, Zhai L, Sun L-Y, Chen C, Chigan J-Z, et al. Dipyridyl-substituted thiosemicarbazone as a potent broad-spectrum inhibitor of metallo-β-lactamases. Bioorg Med Chem. 2021a;38:116128.

Li J-Q, Sun L-Y, Jiang Z, Chen C, Gao H, Chigan J-Z, et al. Diaryl-substituted thiosemicarbazone: A potent scaffold for the development of New Delhi metallo-β-lactamase-1 inhibitors. Bioorg Chem. 2021b;107: 104576.

Li N, Xu Y, Xia Q, Bai C, Wang T, Wang L, et al. Simplified captopril analogues as NDM-1 inhibitors. Bioorg Med Chem Lett. 2014;24:386-9.

Linciano P, Cendron L, Gianquinto E, Spyrakis F, Tondi D. Ten Years with New Delhi Metallo-β-lactamase-1 (NDM-1): From Structural insights to inhibitor design. ACS Infect Dis. 2019;5:9-34.

Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev. 2001;46: 3-26.

Liu S, Jing L, Yu Z-J, Wu C, Zheng Y, Zhang E, et al. ((S)-3-Mercapto-2-methylpropanamido)acetic acid derivatives as metallo-β-lactamase inhibitors: Synthesis, kinetic and crystallographic studies. Eur J Med Chem. 2018a;145:649-60.

Liu S, Zhou Y, Niu X, Wang T, Li J, Liu Z, et al. Magnolol restores the activity of meropenem against NDM-1-producing Escherichia coli by inhibiting the activity of metallo-beta-lactamase. Cell Death Discov 2018b;4: 28.

Liu S, Zhang J, Zhou Y, Hu N, Li J, Wang Y, et al. Pterostilbene restores carbapenem susceptibility in New Delhi metallo-β-lactamase-producing isolates by inhibiting the activity of New Delhi metallo-β-lactamases. Br J Pharmacol. 2019a;176:4548-57.

Liu XL, Shi Y, Kang JS, Oelschlaeger P, Yang KW. Amino acid thioester derivatives: a highly promising scaffold for the development of metallo-β-lactamase L1 inhibitors. ACS Med Chem Lett. 2015;6:660-4.

Liu X-L, Xiang Y, Chen C, Yang K-W. Azolylthioacetamides as potential inhibitors of New Delhi metallo-β-lactamase-1 (NDM-1). J Antibiot. 2019b;72: 118-21.

Ma G, Wang S, Wu K, Zhang W, Ahmad A, Hao Q, et al. Structure-guided optimization of D-captopril for discovery of potent NDM-1 inhibitors. Bioorg Med Chem. 2021;29:115902.

Muhammad Z, Skagseth S, Boomgaren M, Akhter S, Fröhlich C, Ismael A, et al. Structural studies of triazole inhibitors with promising inhibitor effects against antibiotic resistance metallo-β-lactamases. Bioorg Med Chem. 2020;28:115598.

Naas T, Oueslati S, Bonnin RA, Dabos ML, Zavala A, Dortet L, et al. Beta-lactamase database (BLDB) - structure and function. J Enzyme Inhib Med Chem. 2017;32:917-9.

Nantasenamat C. Best practices for constructing reproducible QSAR models. In: Roy K (ed): Ecotoxicological QSARs (pp 55-75). New York, NY: Springer, 2020.

Nantasenamat C, Prachayasittikul V. Maximizing computational tools for successful drug discovery. Expert Opin Drug Discov. 2015;10:321-9.

Nantasenamat C, Worachartcheewan A, Jamsak S, Preeyanon L, Shoombuatong W, Simeon S, et al. AutoWeka: toward an automated data mining software for QSAR and QSPR studies. Methods Mol Biol. 2015; 1260:119-47.

Palica K, Vorácová M, Skagseth S, Andersson Rasmussen A, Allander L, Hubert M, et al. Metallo-β-lactamase inhibitor phosphonamidate monoesters. ACS Omega. 2022;7:4550-62.

Parvez S, Khan AU. Hospital sewage water: a reservoir for variants of New Delhi metallo-β-lactamase (NDM)- and extended-spectrum β-lactamase (ESBL)-producing Enterobacteriaceae. Int J Antimicrob Agents. 2018;51:82-8.

Piir G, Kahn I, García-Sosa AT, Sild S, Ahte P, Maran U. Best practices for QSAR model reporting: physical and chemical properties, ecotoxicity, environmental fate, human health, and toxicokinetics endpoints. Environ Health Perspect. 2018;126:126001.

Rehman MT, AlAjmi MF, Hussain A, Rather GM, Khan MA. High-throughput virtual screening, molecular dynamics simulation, and enzyme kinetics identified ZINC84525623 as a potential inhibitor of NDM-1. Int J Mol Sci. 2019;20(4):819.

Romero E, Oueslati S, Benchekroun M, D'Hollander ACA, Ventre S, Vijayakumar K, et al. Azetidinimines as a novel series of non-covalent broad-spectrum inhibitors of β-lactamases with submicromolar activities against carbapenemases KPC-2 (class A), NDM-1 (class B) and OXA-48 (class D). Eur J Med Chem. 2021;219:113418.

Rose AS, Bradley AR, Valasatava Y, Duarte JM, Prlic A, Rose PW. NGL viewer: web-based molecular graphics for large complexes. Bioinformatics. 2018;34:3755-8.

Santucci M, Spyrakis F, Cross S, Quotadamo A, Farina D, Tondi D, et al. Computational and biological profile of boronic acids for the detection of bacterial serine- and metallo-β-lactamases. Sci Rep. 2017;7:17716.

Schaduangrat N, Lampa S, Simeon S, Gleeson MP, Spjuth O, Nantasenamat C. Towards reproducible computational drug discovery. J Cheminform. 2020; 12:9.

Schnaars C, Kildahl-Andersen G, Prandina A, Popal R, Radix S, Le Borgne M, et al. Synthesis and preclinical evaluation of TPA-based zinc chelators as metallo-β-lactamase inhibitors. ACS Infect Dis. 2018;4:1407-22.

Shen B, Yu Y, Chen H, Cao X, Lao X, Fang Y, et al. Inhibitor discovery of full-length New Delhi metallo-β-lactamase-1 (NDM-1). PLoS One. 2013;8:e62955.

Shi C, Bao J, Sun Y, Kang X, Lao X, Zheng H. Discovery of Baicalin as NDM-1 inhibitor: Virtual screening, biological evaluation and molecular simulation. Bioorg Chem. 2019;88:102953.

Shin WS, Nguyen ME, Bergstrom A, Jennings IR, Crowder MW, Muthyala R, et al. Fragment-based screening and hit-based substructure search: Rapid discovery of 8-hydroxyquinoline-7-carboxylic acid as a low-cytotoxic, nanomolar metallo β-lactamase inhibitor. Chem Biol Drug Des. 2021;98:481-92.

Shoombuatong W, Schaduangrat N, Nantasenamat C. Towards understanding aromatase inhibitory activity via QSAR modeling. EXCLI J. 2018;17:688-708.

Skagseth S, Akhter S, Paulsen MH, Muhammad Z, Lauksund S, Samuelsen Ø, et al. Metallo-β-lactamase inhibitors by bioisosteric replacement: Preparation, activity and binding. Eur J Med Chem. 2017;135:159-73.

Song G-Q, Wang W-M, Li Z-S, Wang Y, Wang J-G. First identification of isatin-β-thiosemicarbazones as novel inhibitors of New Delhi metallo-β-lactamase-1: Chemical synthesis, biological evaluation and molecular simulation. Chin Chem Lett. 2018;29:899-902.

Stierand K, Maass PC, Rarey M. Molecular complexes at a glance: automated generation of two-dimensional complex diagrams. Bioinformatics. 2006;22:1710-6.

Su J, Liu J, Chen C, Zhang Y, Yang K. Ebsulfur as a potent scaffold for inhibition and labelling of New Delhi metallo-β-lactamase-1 *in vitro* and *in vivo*. Bioorg Chem. 2019;84:192-201.

Sully EK, Geller BL, Li L, Moody CM, Bailey SM, Moore AL, et al. Peptide-conjugated phosphorodiamidate morpholino oligomer (PPMO) restores carbapenem susceptibility to NDM-1-positive pathogens *in vitro* and i*n vivo*. J Antimicrob Chemother. 2017;72:782-90.

Szmolka A, Nagy B. Multidrug resistant commensal Escherichia coli in animals and its impact for public health. Front Microbiol. 2013;4:258.

Tehrani KHME, Brüchle NC, Wade N, Mashayekhi V, Pesce D, van Haren MJ, et al. Small molecule carboxylates inhibit metallo-β-lactamases and resensitize carbapenem-resistant bacteria to meropenem. ACS Infect Dis. 2020a;6:1366-71.

Tehrani KHME, Fu H, Brüchle NC, Mashayekhi V, Prats Luján A, van Haren MJ, et al. Aminocarboxylic acids related to aspergillomarasmine A (AMA) and ethylenediamine-N,N'-disuccinic acid (EDDS) are strong zinc-binders and inhibitors of the metallo-beta-lactamase NDM-1. Chem Commun. 2020b;56:3047-9.

Tooke CL, Hinchliffe P, Bragginton EC, Colenso CK, Hirvonen VHA, Takebayashi Y, et al. β-Lactamases and β-Lactamase Inhibitors in the 21st Century. J Mol Biol. 2019;431:3472-500.

Tropsha A. Best practices for QSAR model development, validation, and exploitation. Mol Inform. 2010; 29:476-88.

van Haren MJ, Tehrani KHME, Kotsogianni I, Wade N, Brüchle NC, Mashayekhi V, et al. Cephalosporin prodrug inhibitors overcome metallo-β-lactamase driven antibiotic resistance. Chem Eur J. 2021;27:3806-11.

Wang T, Xu K, Zhao L, Tong R, Xiong L, Shi J. Recent research and development of NDM-1 inhibitors. Eur J Med Chem. 2021;223:113667.

Wang Y, Wu C, Zhang Q, Qi J, Liu H, Wang Y, et al. Identification of New Delhi metallo-β-lactamase 1 in Acinetobacter lwoffii of food animal origin. PLoS One. 2012;7:e37152.

Wang Y, Sun X, Kong F, Xia L, Deng X, Wang D, Wang J. Specific NDM-1 inhibitor of isoliquiritin enhances the activity of meropenem against NDM-1-positive *Enterobacteriaceae in vitro*. Int J Environ Res Public Health. 2020;17(6):2162.

Xiang Y, Chang Y-N, Ge Y, Kang JS, Zhang Y-L, Liu X-L, et al. Azolylthioacetamides as a potent scaffold for the development of metallo-β-lactamase inhibitors. Bioorg Med Chem Lett. 2017;27:5225-9.

Xiang Y, Chen C, Wang W-M, Xu L-W, Yang K-W, Oelschlaeger P, et al. Rhodanine as a potent scaffold for the development of broad-spectrum metallo-β-lactamase inhibitors. ACS Med Chem Lett. 2018;9:359-64.

Yahiaoui S, Voos K, Haupenthal J, Wichelhaus TA, Frank D, Weizel L, et al. N-Aryl mercaptoacetamides as potential multi-target inhibitors of metallo-β-lactamases (MBLs) and the virulence factor LasB from Pseudomonas aeruginosa. RSC Med Chem. 2021;12:1698-708.

Yang Y, Guo Y, Zhou Y, Gao Y, Wang X, Wang J, et al. Discovery of a novel natural allosteric inhibitor that targets NDM-1 against *Escherichia coli*. Front Pharmacol. 2020;11:581001.

Yap CW. PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. J Comput Chem. 2011;32:1466-74.

Yu Z-J, Liu S, Zhou S, Li H, Yang F, Yang L-L, et al. Virtual target screening reveals rosmarinic acid and salvianolic acid A inhibiting metallo- and serine-β-lactamases. Bioorg Med Chem Lett. 2018;28:1037-42.

Zhai L, Zhang Y-L, Kang JS, Oelschlaeger P, Xiao L, Nie S-S, et al. Triazolylthioacetamide: A valid scaffold for the development of New Delhi metallo-β-lactamase-1 (NDM-1) inhibitors. ACS Med Chem Lett. 2016;7:413-7.

Zhang D, Markoulides MS, Stepanovs D, Rydzik AM, El-Hussein A, Bon C, et al. Structure activity relationship studies on rhodanines and derived enethiol inhibitors of metallo-β-lactamases. Bioorg Med Chem. 2018;26:2928-36.

Zhang J, Wang S, Wei Q, Guo Q, Bai Y, Yang S, et al. Synthesis and biological evaluation of Aspergillomarasmine A derivatives as novel NDM-1 inhibitor to overcome antibiotics resistance. Bioorg Med Chem. 2017;25:5133-41.

Zhang Y-J, Liu X-L, Wang W-M, Chen C, Zhao M-H, Yang K-W. Amino acid thioesters exhibit inhibitory activity against B1-B3 subclasses of metallo-β-lactamases. Chem Pharm Bull. 2019;67:135-42.

Zhao B, Zhang X, Yu T, Liu Y, Zhang X, Yao Y, et al. Discovery of thiosemicarbazone derivatives as effective New Delhi metallo-β-lactamase-1 (NDM-1) inhibitors against NDM-1 producing clinical isolates. Acta Pharm Sin B. 2021;11:203-21.

Zhou Y, Guo Y, Sun X, Ding R, Wang Y, Niu X, et al. Application of oleanolic acid and its analogues in combating pathogenic bacteria *in vitro/vivo* by a two-pronged strategy of β-lactamases and hemolysins. ACS Omega. 2020;5:11424-38.