Using ML to repurpose FDA drugs for the treatment of Diabetic Cardiomyopathy

Vishnu Aravind, Alex Kumar

Abstract:

PARP-1 (Poly ADP Ribose polymerase) functions to repair damage to DNA and is implicated in a variety of diseases including Diabetic Cardiomyopathy (DCM). Unfortunately, there are few treatments for this disease, and the expenses associated with these drugs present barriers to many. With this project, we developed a neural network that was able to distinguish between inhibitors and non-inhibitors of PARP-1 in order to uncover more accessible treatments of DCM. We collected confirmed inhibitors of PARP-1 from PubChem, clustered these compounds, and performed attribute selection. This data was used to develop the neural network which was able to predict inhibitors of PARP-1 with an accuracy of 97% and an AUROC of 0.98. The model was then run on all FDA drugs, and the top 37 predictions were taken. In protein ligand docking simulations, the predicted inhibitors had a significantly better binding affinity for PARP-1 than the control group.

1 | INTRODUCTION

Type 2 Diabetes and Diabetic Cardiomyopathy

Diabetes Mellitus is a chronic disorder characterized by decreased serum insulin, decreased insulin sensitivity, hyperglycemia, and a reduction in pancreatic beta cells (1), and is widespread having affected 422 million individuals in 2014 (2). According to the CDC, in 2019, 28.7 million people in the US (8.7% of the population) had diagnosed diabetes, with about 95 percent of these cases being type 2 diabetes. Among those with Type 2 Diabetes, Heart Failure is one the most common resulting cardiovascular complications, as patients with T2D have up to a 74% higher chance of developing Heart failure (3). Additionally, mortality is quadrupled in T2D patients with heart failure vs patients without (3). This ventricular dysfunction is referred to by Diabetic Cardiomyopathy (DCM), which is defined by abnormal myocardial structure and diminished cardiac performance in individuals with diabetes mellitus that lack other risk factors, such as hypertension and coronary artery disease (4).

PARP-1

Poly ADP-ribose polymerase 1 (PARP-1) is an NAD+ dependent ADP-ribosylating enzyme and its primary function lies in DNA repair. It can also control the accessibility of DNA for RNA polymerase by regulating chromatin structure, and can function as a transcription factor by binding on to motifs in promoter regions (5). This protein has been implicated in a variety of malignant and inflammatory diseases, including ovarian cancer, breast cancer, cardiovascular disease, asthma, arthritis, and diabetes (6). In malignant diseases, PARP-1 enables cells lacking methods of homologous recombination, such as those with BRCA1/2 mutations, to have a method of DNA repair, and thus to avoid cytotoxicity (7). In inflammatory processes, PARP-1 is known to activate NF-kB in response to lipopolysaccharides or TNF-a (8).

DCM Pathogenesis

A significant portion of DCM's pathogenesis is rooted in the over-production of proinflammatory cytokines such as tumor necrosis factor-α (TNF-α) and interleukin-1β (IL-1β), which induce apoptosis in cardiomyocytes, leading to a reduction in contractility and cardiac dysfunction (9, 10). Additionally, the hyperglycemia found in individuals with DCM can trigger the formation of reactive oxygen and nitrogen species, which cause DNA strand breaks and abnormal cell signaling, ultimately resulting in apoptosis.

The oxidative stress and subsequent DNA damage also results in the overactivation of PARP-1, which utilizes NAD+ in order to mend single strand breaks. This excessive activation leads to the rapid depletion of intracellular NAD+ and ATP, resulting in a cellular energy shortage and subsequent apoptosis (11). In a study conducted to investigate the role of PARP-1 in type-2 diabetes induced cardiac complications, diabetic rats showed increased PARP-1 activity compared to control animals, and rats that were treated with 4-aminobenzamide (4-AB), a PARP-1 inhibitor, displayed reduced cardiac and vascular inflammation, reduced hypertension, and alleviated cardiac ischemia (12).

Inhibition of PARP-1

Inhibitors of PARP-1 generally bind the catalytic pocket or to the PARP-chromatin complex and trap the enzyme in an ineffective state at the chromatin. Both ways interfere with ADP-ribosylation, with the latter being much more potent (13). Cancerous cells often are deficient in one of the six major DNA repair pathways (base excision repair, nucleotide excision repair, single strand break repair, homologous recombination, non-homologous end joining, and mismatch repair), and thus the inhibition of another pathway leads to a synthetic lethality. PARP-1 plays a role in the single strand break repair and base excision repair pathways, and therefore its inhibition is key in treating malignancies. In inflammatory and metabolic diseases, however, cells are still proficient in homologous recombination, and can thus mend the double strand breaks which single strand breaks turn into (14). In these types of cells, the inhibition of PARP-1 prevents the rapid depletion of NAD+ and ATP, avoiding a cellular energy crisis and preventing cytotoxicity. There are currently no PARP-1 inhibitors that have been approved by the FDA for use in patients with DCM.

Drug Repurposing and ML

The benefits associated with drug repurposing as opposed to de novo drug development are numerous. Development times are decreased significantly, with de novo drug discovery and development generally taking 10 - 17 years, while repurposed drugs are approved within 3 - 12 years (15, 16). Development costs are brought down as well. The cost of bringing a newly developed drug to market is estimated to be 2-3 billion dollars, compared to the estimated 300 million for a repurposed drug (17). This drop in development costs result in improved prices for consumers and patients, who according, to the American Diabetes Association, are burdened with medical expenditures on average of around $16,000, of which 30% can be attributed to the purchasing of prescription medicine (18). Consumers are also presented with a lower risk of adverse events when using repurposed drugs, as the safety and pharmacokinetics of the

drug are already known (19). Drug repurposing has become even more practical with the introduction of machine learning (ML) techniques to the field of drug discovery. ML has been used in a variety of stages of drug discovery, including target validation, identification of prognostic biomarkers, and the analysis of pathology data in clinical trials (20), and allows for the narrowing down of compounds for wet lab testing for a relatively little cost. Thus, the goal of our study was to utilize ML, specifically a deep neural network, in order to determine which FDA drugs would be able to serve as PARP-1 inhibitors.

2 | METHODS

2.1 | Data Collection

The activity values and canonical SMILES (standardized line notation that represents the molecular structure of a chemical compound as a unique string of characters) of 4,786 inhibitors of PARP-1 were taken from the PubChem Database (22). To ensure the reliability of our analysis, we selected only the top compounds with activity values less than 80 nM. Activity values are a measure of a substance's efficacy or potency, and indicate the concentration required to produce a specific biological response. By focusing on the most active compounds, we aimed to identify potential drug candidates with strong biological activity. Additionally, the SMILES strings of FDA drugs were taken from the Drugbank Database, and the SMILES of random compounds were taken from PubChem to form a control group.

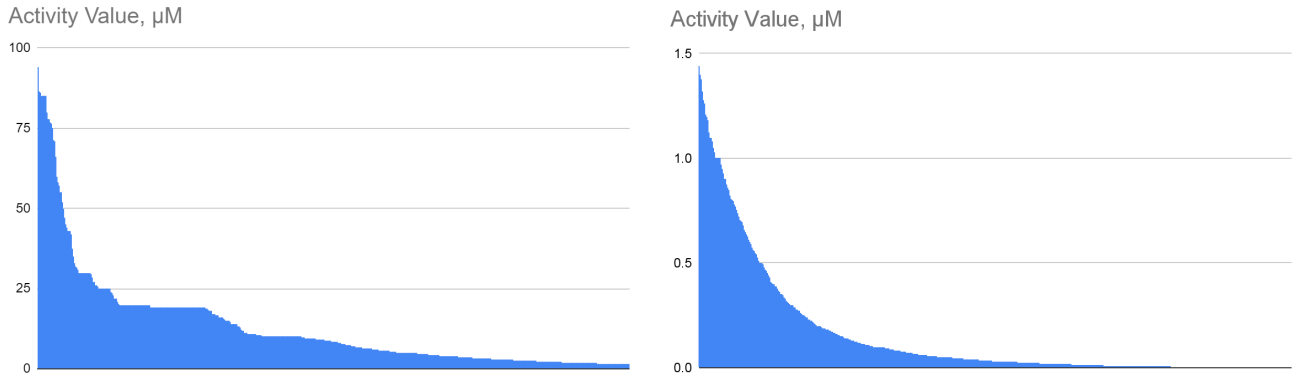


Figure 1 - These two graphs depict the distribution of the different activity values of the compounds. 4062 compounds had activity values less than 1.5 µm, while only 725 compounds had activity values greater than 1.5 µm.

2.2 | Fingerprint Clustering

Fingerprint clustering is a common technique in cheminformatics for grouping molecules based on their structural similarity. By using clustering, we can reduce the complexity of the dataset and create more robust models that are less sensitive to outliers and noise in the data. To cluster the data retrieved from PubChem, we utilized the Butina method from RDKit, a python library for cheminformatics (23). It leverages a user-defined similarity as the only input to the clustering program. This similarity measure is based on the Tanimoto Index, which can be calculated using the below specific formula. In this study, we chose a Tanimoto Index score of 0.47, which ensured that clusters with high similarity were obtained. 234

clusters were formed, with the largest cluster identified containing 318 compounds, and we focused our modeling efforts on this cluster.

$$T(a,b) = \frac{N_c}{N_a + N_b - N_c}$$

Figure 2 - Tanimoto Coefficient used for finding the Tanimoto Index

2.3 | Descriptors

PaDEL-Descriptor is a software tool that generates a wide range of chemical descriptors for a set of compounds (24). It uses the Chemistry Development Kit and some additional descriptor categories. PaDEL was utilized to calculate the molecular descriptors of the clustered compounds. The descriptors are calculated based on the molecular structure encoded in the SMILES notation and can provide 1,875 features that describe the molecular properties of a compound. Some of the features include molecular weight, lipophilicity, hydrogen bonding, and topological parameters such as connectivity and branching. PADEL descriptors were employed to calculate the descriptors for canonical smiles of our largest cluster of known inhibitors.

2.4 | Feature Selection

Infogain is a popular feature selection algorithm that ranks features based on their ability to discriminate between classes in a classification problem (25). The algorithm computes the mutual information between each feature and the class variable, which measures how much information about the class can be inferred from the feature. Features that have a high mutual information score are deemed to be more informative and are ranked higher. After the InfoGain attribute evaluator was used on the training set of inhibitors, FDA drugs, and the control compounds, the number of attributes was decreased from 1,875 to 300.

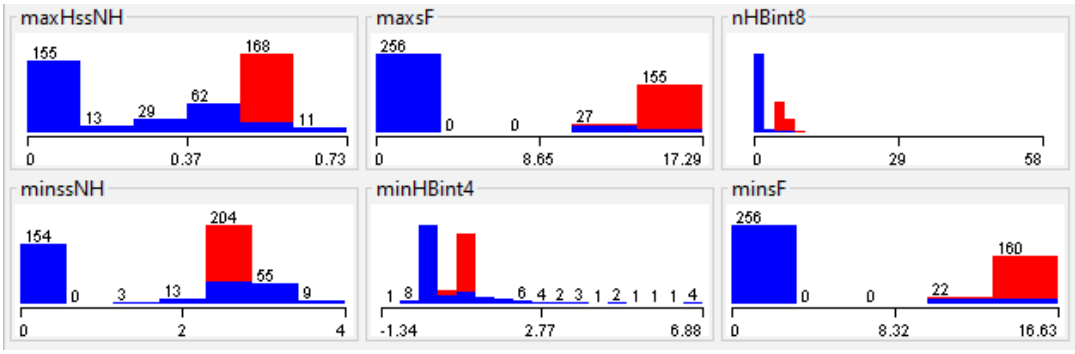


Figure 3 - Highest ranked descriptors by the infogain algorithm from WEKA

2.5 | Neural Network Construction

The Neural Network was created in python and works by splitting the data with an 80-20 train validation split. The 80-20 split is a commonly used method in machine learning to split the dataset into training and validation sets. The training set, which contains 80% of the data, was used to train the model, while the validation set, which contains 20% of the data, was used to evaluate the performance of the model. In order to perform machine learning predictions, we used a multilayer perceptron (aka neural network) made with Keras, a python library for the development of neural networks (26). The model's hyperparameters, namely the number of layers, the number of nodes in each layer, dropout, learning rate, and batch size, were turned with the Keras-Tuner. The process was conducted with Keras-Tuner's Hyperband method, which uses a technique known as successive halving. The Hyperband search method starts by randomly sampling sets of hyperparameter configurations, training and testing them, and then discarding ones that perform poorly. Then, the remaining configurations are trained for a larger number of epochs, again discarding ones that perform poorly. This process is repeated until the best performing configuration can be selected. In order to prevent overtraining in the model, we employed an epoch stop early callback, in which if loss did not decrease for 5 epochs, training would end prematurely.

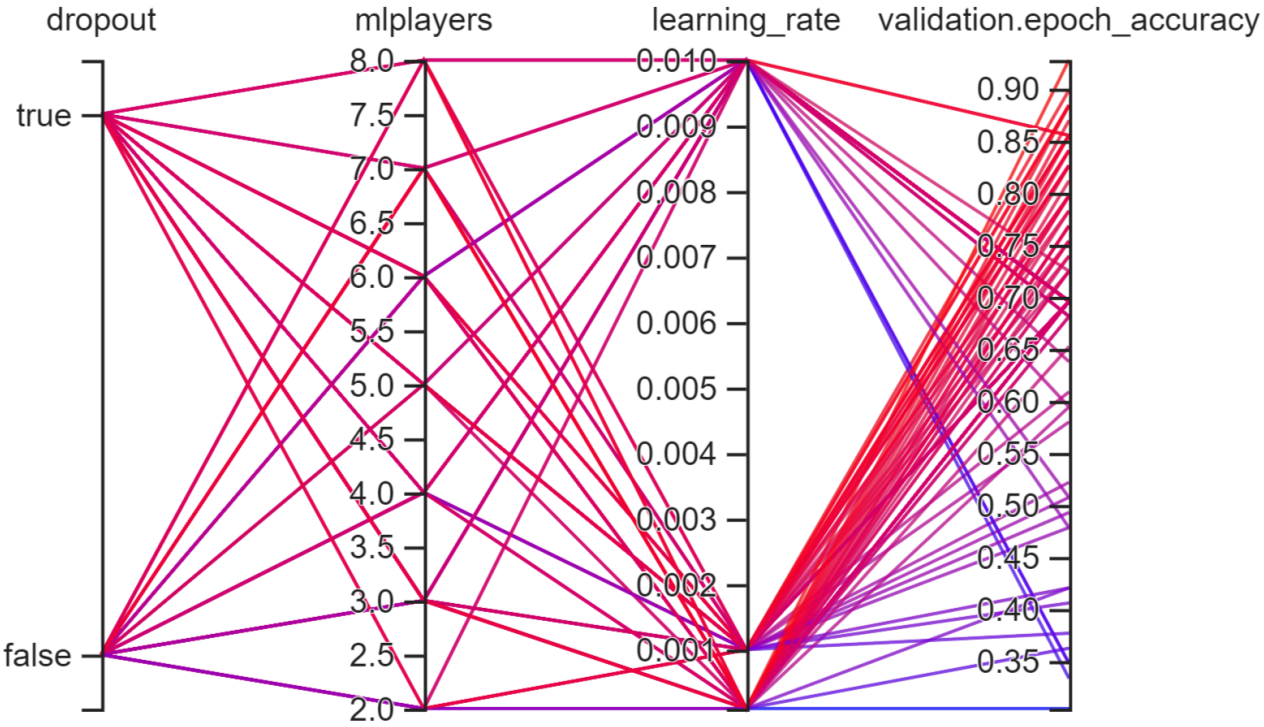


Figure 4 - Visual representation of the hyperparameters used in the tuning process. This representation is produced by TensorBoard (27).

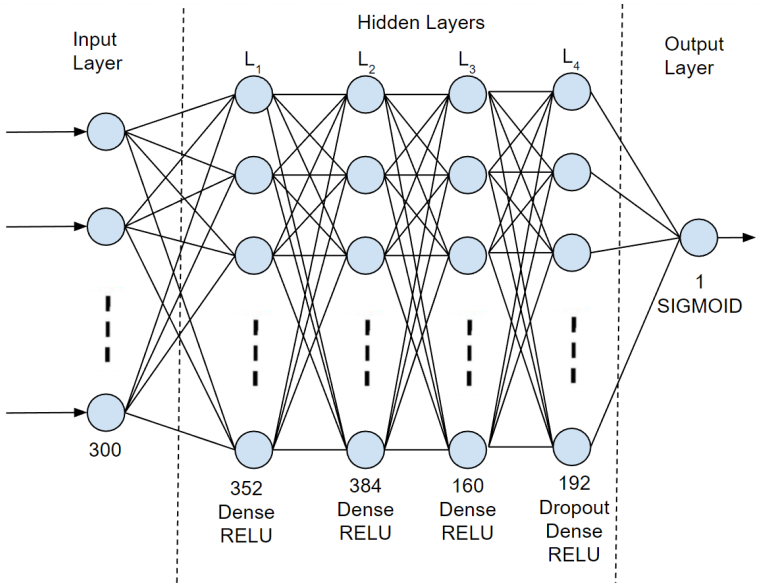


Figure 5 - A schematic of our Neural Network, with an input layer, eight hidden layers, a dropout layer, and then an output layer with a single node.

2.6 | Protein-Ligand Preparation

Before docking could be performed, structural files of PARP-1 and the ligands to be docked would need to be obtained and prepped. For PARP-1, a PDB (Protein Data Bank) file of its binding site was downloaded from the Protein Data Bank Database (28). It was then imported into AutoDock Tools, a software program used for preparing and setting up molecules for docking simulations in the AutoDock suite of programs (29). The software was then used to delete the ligand already present in the binding site (it was olaparib), delete water molecules, add polar hydrogens, add Kollman Charges, and finally convert the file to the PDBQT format. For the ligands, the first step was obtaining the canonical SMILES for each of the predicted inhibitors, confirmed inhibitors, and control compounds from the data we had collected in the first step. Then, using the aforementioned RDKit module, these SMILES were used to get the PDB files of all the ligands. These ligands were then imported into Autodock Tools, where aromatic carbons and rotatable bonds were detected, Gasteiger Charges were added, and non-polar hydrogens were merged. Again, these were all converted to the PDBQT format.

2.7 | Protein-Ligand Docking

AutoDock Vina is a widely-used software program for molecular docking, which is the computational process of predicting how two molecules will interact with each other in three-dimensional space (30, 31). AutoDock Vina works by using a combination of empirical scoring functions and advanced algorithms to generate a large number of possible binding poses (i.e. the different ways that two molecules could interact with each other). It then evaluates each pose based on a variety of factors, such as the stability of the complex, and selects the most promising candidates for further analysis. We employed AutoDock Vina to perform molecular docking of Ligands to the target protein, PARP-1. This was achieved by inputting the grid box dimensions and the prepared ligand into the target protein, allowing for the exploration of possible binding conformations between the Ligands and the active site of the PARP-1 protein. The default dimensions predicted by AutoDock Tools were used to obtain the grid box dimensions.

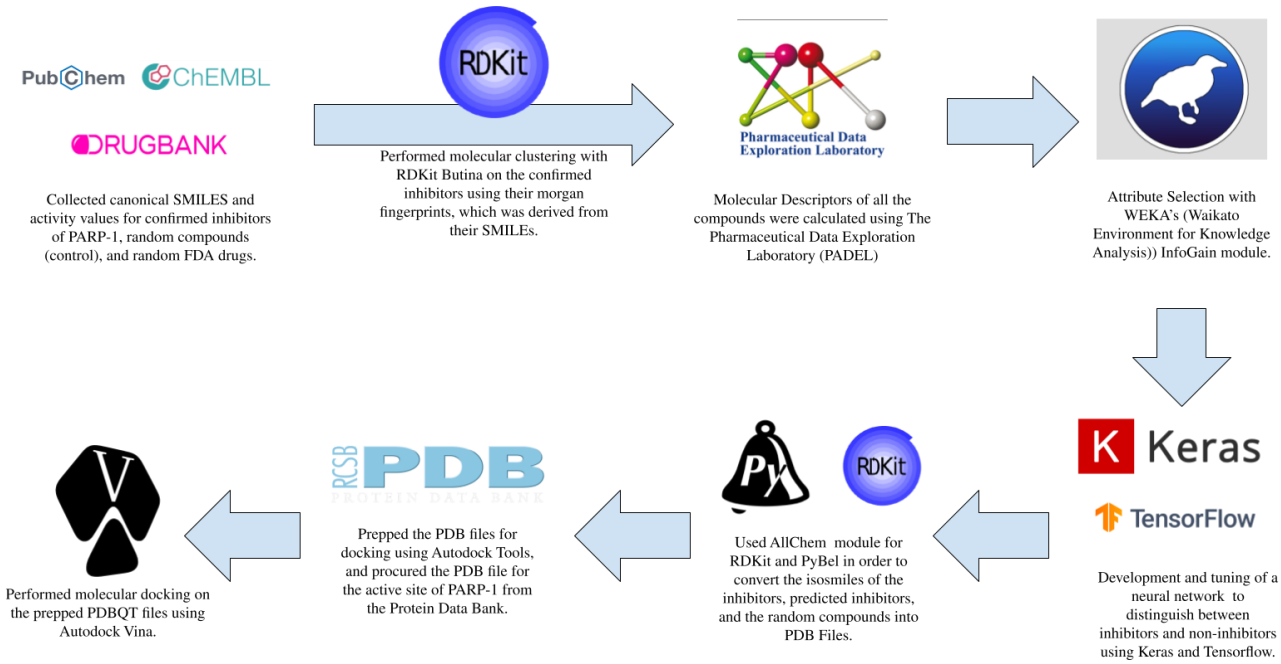


Figure 6 - Overview of the methods utilized in this study

3 | RESULTS

3.1 | Multilayer Perceptron Model

Once the model was tuned, it was trained for 5 times independently on the training data that had been split earlier. The Accuracy vs Epoch graph for all trials can be seen Figure 7. The graph starts at around 0.5 (random guessing) and shows a steady increase in accuracy, indicating that the model is rapidly learning to distinguish between the inhibitor class and the non-inhibitor class.

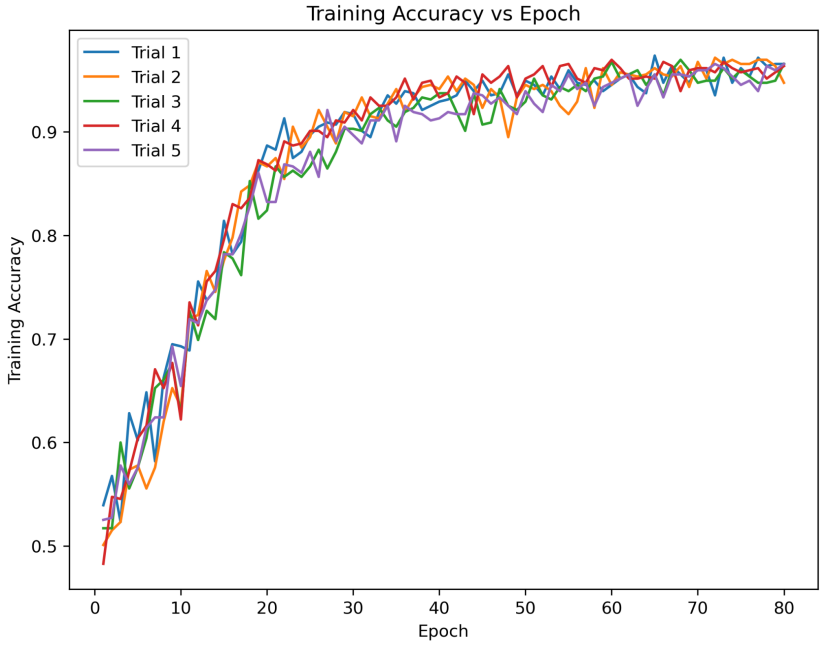


Figure 7 - The Training Accuracy vs Epoch Graph for all 5 trials

After the training period, the model was tested using the validation data. Testing on validation data rather than training data ensures that the model is not benefitting from overfitting. Table 4 depicts the testing metrics for the model from each trial, namely the Validation Accuracy, the Matthews Correlation Coefficient, and the Area Under the ROC (Receiver Operating Characteristic) curve. Accuracy measures the percentage of correct predictions out of the total predictions. The average accuracy over the 5 trials was 0.976, indicating a high percentage of correct predictions. Matthews Correlation Coefficient is a slightly better metric for binary classification methods, as it takes into account false positives, false negatives, true positives, and true negatives, thus being effective when the classes are imbalanced. The average MCC was 0.952, indicating a strong correlation between the predicted and observed classifications. Lastly, the AUC is the area under a plot of the true positive rate against the false positive rate at various threshold settings. The average AUC over the 5 trials was 0.989, indicating that 98.9% percent of the time, the model was able to correctly rank the probabilities of the inhibitor samples higher than the non-inhibitor samples.

Trial	Test Accuracy	MCC	AUROC
Trial 1	0.9596773982	0.9204329801	0.9981747066
Trial 2	0.9677419066	0.9357219295	0.9825293351
Trial 3	0.9919354916	0.9839741502	0.9920469361
Trial 4	0.9838709831	0.9681233591	0.9827900913
Trial 5	0.9758064747	0.9515859379	0.9911342894
Average	0.9758064508	0.9519676714	0.9893350717

Table 1 - Testing Metrics for each of the 5 trials

After the training and testing, the model was used to output prediction scores for a dataset of 2037 FDA-approved drugs with the same 300 descriptors that had been taken previously from InfoGain. Out of these 2037 compounds, the model had predicted 82 compounds to be inhibitors of PARP-1, with 38 receiving a prediction score greater than 0.9.

3.2 | Protein-Ligand Docking

The random FDA control group had an average docking score of -7.824 kcal/mol, while the predicted inhibitors had an average docking score of -10.704 kcal/mol. The t-Test revealed a high magnitude and negative T stat (-17.68), indicating a substantial difference between the two sets of data, with the predicted inhibitors outperforming the control group. The calculated p-value (2.96e-32) is very small, indicating evidence against the null hypothesis and supporting the alternative hypothesis that predicted inhibitors performed significantly better than the control group. Because of the higher binding affinity of the predicted inhibitors, they may have greater efficacy as a treatment option. However, it is important to acknowledge the limitations of our study, as the docking simulations we performed may not perfectly reflect the actual binding affinity of the protein-ligand complex in a biological system. Future studies could aim to address this limitation by performing further experimental validation of the predicted inhibitors to confirm their efficacy in vivo or in clinical trials.

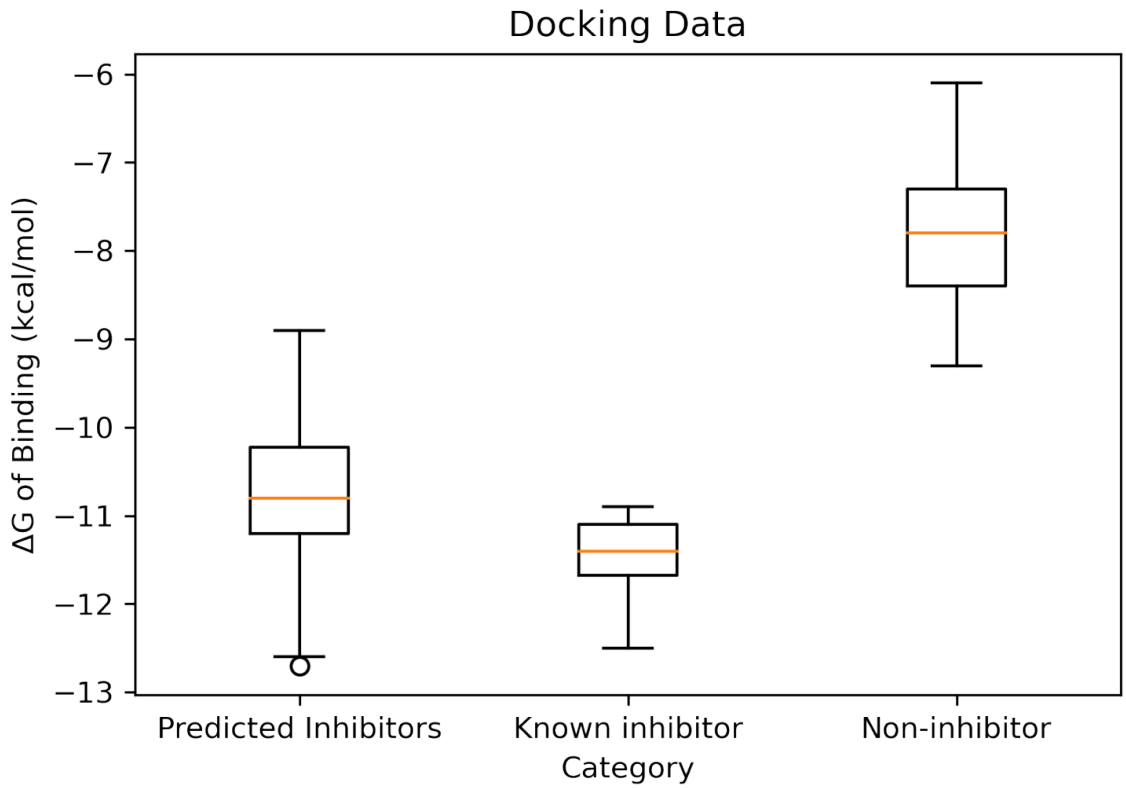
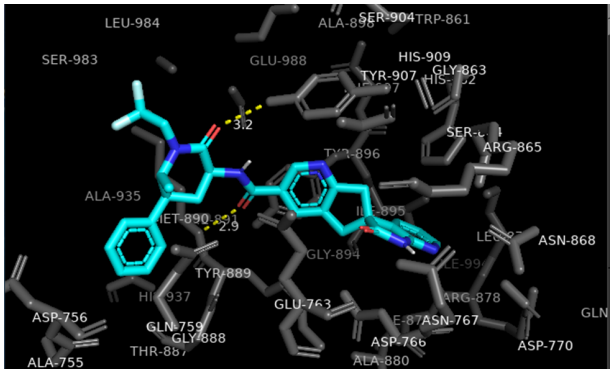


Figure 8 - Docking scores (kcal/mol) for top 35 Predicted Inhibitors, 147 known inhibitors, 147 random FDA-approved drugs

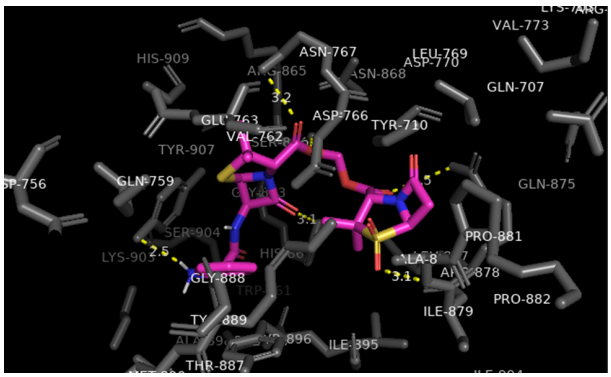
Statistic	Predicted Inhibitor	Control (Random FDA drugs)
Mean	-10.704	-7.824
Variance	0.719584	0.5802
T stat	-17.68	
p-value	2.96E-32	

Table 2 - t-Test & statistical analysis of docking results

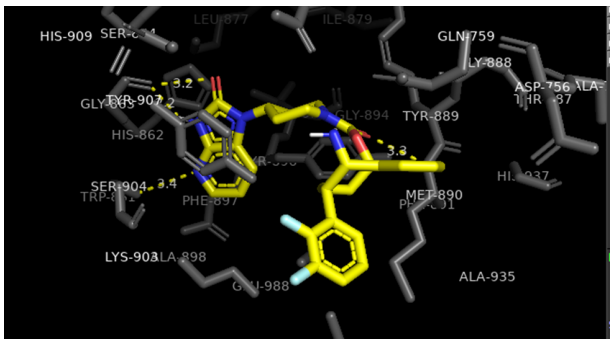
I



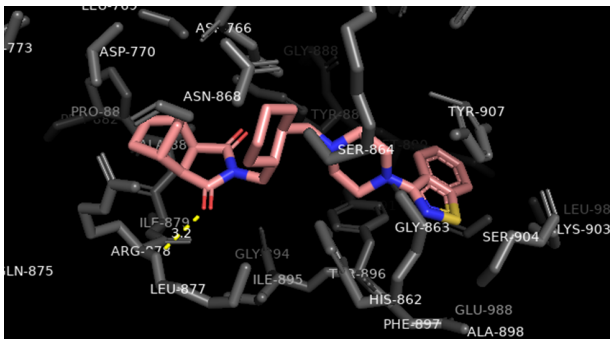
II



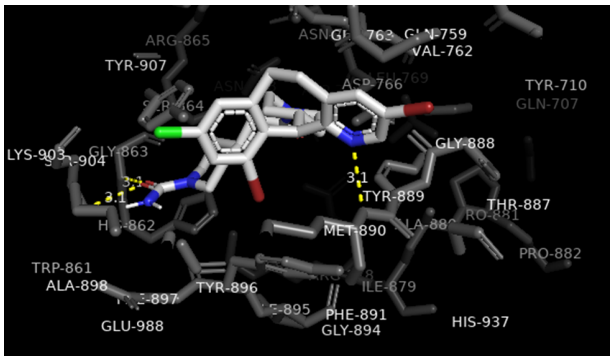
III



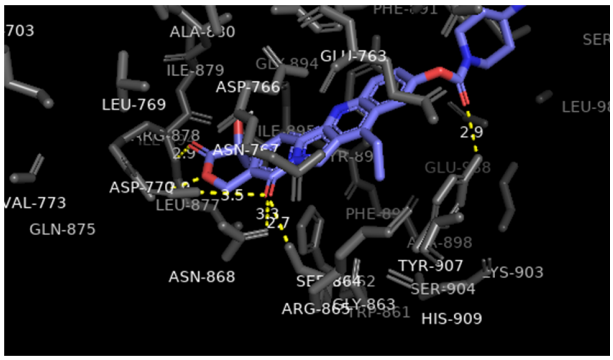
IV



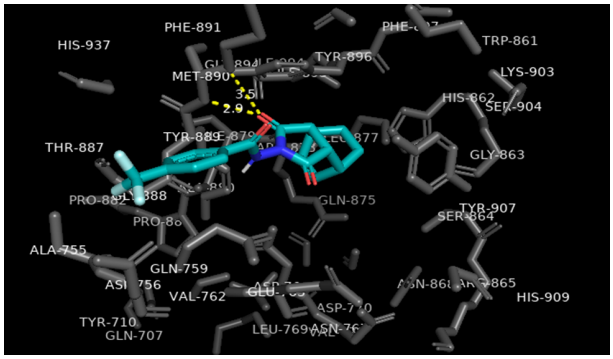
V



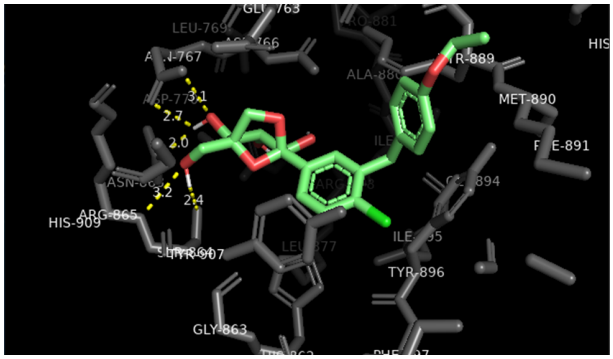
VI



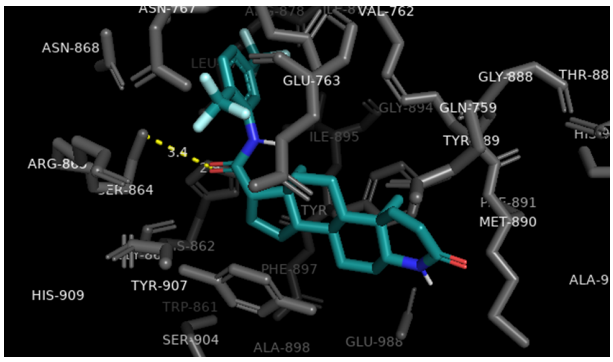
VII



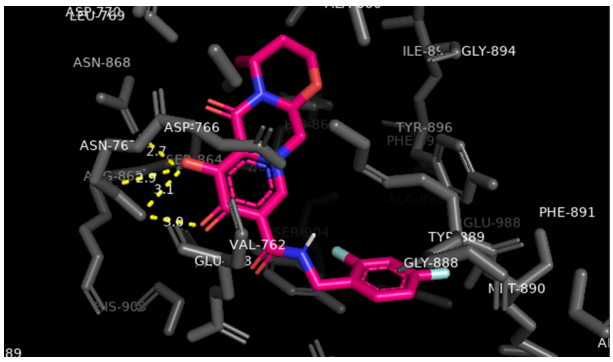
VIII



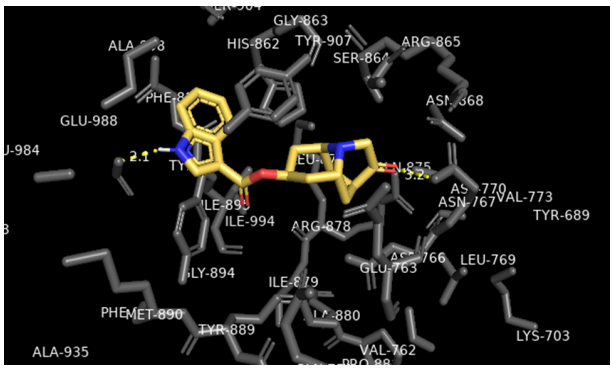
IX



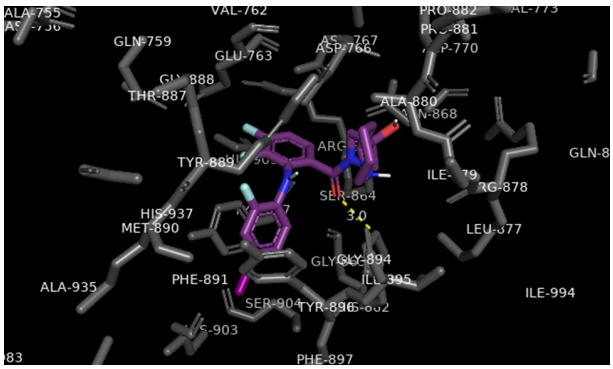
X



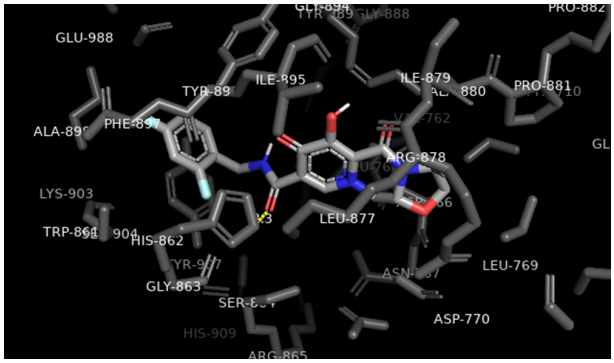
XI



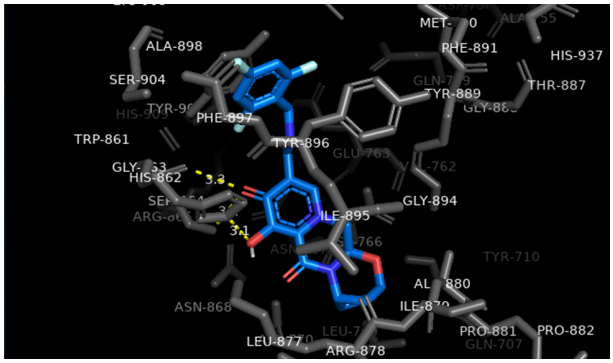
XII



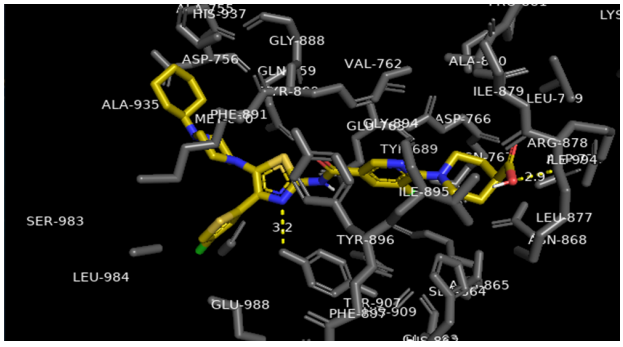
XIII



XIV



XV



XVI

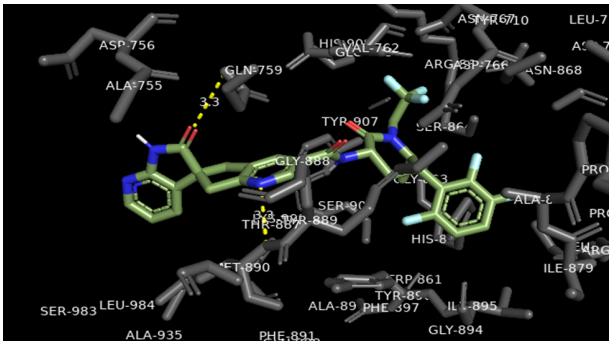


Figure 7 - Docking of the highest ranked drugs by the model with the PARP-1 catalytic pocket: (I)Ubrogepant; (II)Sultamicillin; (III)Rimegepant; (IV)Lurasidone; (V)Lonafarnib; (VI)Irinotecan; (VII)Tecovirimat; (VIII)Ertugliflozin; (IX)Dutasteride; (X)Dolutegravir; (XI)Dolasetron; (XII)Cobimetinib; (XIII)Cabotegravir; (XIV)Bictegravir; (XV)Avatrombopag; (XVI)Atogepant

Residue	Number of Interactions (%)	Average Distance
ASN-767	0.67	3.0625
MET-890	0.50	3.13
ARG-878	0.42	3.2
SER-864	0.416	2.9
GLY-863	0.34	2.95
TYR-907	0.33	3.10
HIS-862	0.083	3.08

Table 3 -Residues most frequently contacting inhibitors.

Table 4 -Top 30 performing FDA drugs in Autodock Vina

Docking Score (kcal/mol)	Name of predicted Inhibitor	Functions	Physiological Interactions
-12.7	Dutasteride	Used for the treatment of symptomatic benign prostatic hyperplasia	Inhibits type II 5a reductase, preventing 5a-dihydrotestosterone formation
-12.6	Rimegepant	Used to treat migraines in adults	Antagonist of Calcitonin Gene-Related peptide type 1 receptor

-12.1	Lonafarnib	Used to decrease the mortality associated with Hutchinson-Gilford progeria syndrome	Inhibitor of farnesyl transferase
-11.9	Bictegravir	Used to treat HIV infections	Antagonist of HIV-1 reverse transcriptase and integrase
-11.7	Irinotecan	Used to treat metastatic carcinoma of the colon or rectum	Inhibitor of DNA topoisomerase I
-11.6	Trospium	Used to treat the symptoms of overactive bladder	Antagonizes the effect of acetylcholine on muscarinic receptors
-11.4	Flecainide	Used to manage atrial fibrillation and paroxysmal supraventricular tachycardias	Inhibitor of fast sodium channels, delayed potassium channels, and ryanodine receptors
-11.3	Edoxaban	Used for reducing the risk of stroke and systemic embolism	Inhibitor of Coagulation Factor X
-11.3	Sitagliptin	Used for the management of type 2 diabetes mellitus	Inhibitor of Dipeptidyl Peptidase 4
-11.3	Eravacycline	Used to treat complicated intra-abdominal infections	Inhibitor of 30S ribosomal protein S4
-11.2	Adapalene	Used to treat acne vulgaris in adolescents and adults	Agonist of Retinoic Acid Receptor beta, gamma, RXR-beta, RXR-gamma, RXR-alpha, antagonist of Toll-like Receptor 2
-11.2	Trifluoperazine	Used to treat depression, anxiety, and agitation	Antagonist of Dopaminergic D1 and D2 receptors
-11.1	Atogepant	Used for the preventative therapy of episodic migraine headaches	Antagonist of Calcitonin gene-related peptide type 1 receptor

-11.1	Empagliflozin	Used to manage type 2 diabetes mellitus	Inhibitor of Sodium/glucose cotransport 2
-11.1	Flupentixol	Used to treat schizophrenia and depression	Antagonist of Dopamine D1 and D2 receptors, and 5-hydroxytryptamine receptor 2A
-10.9	Zuclopenthixol	Used for the management of schizophrenia	Antagonist of Dopamine D1, D5, and D2 receptors
-10.9	Fluphenazine	Used to treat patients requiring long-term neuroleptic therapy	Antagonist of Dopamine D1 and D2 receptor
-10.9	Periciazine	Used with other medications to treat aggressiveness, impulsiveness, and hostility associated with psychiatric conditions	Antagonist of Dopamine D1 receptor, and Alpha-2A Adrenergic Receptor
-10.8	Ubrogepant	Used in the acute treatment of migraine with or without aura	Antagonist of calcitonin gene-related peptide type 1 receptor
-10.8	Dolutegravir	Used for the treatment of HIV-1 infections	Inhibitor of HIV-1 integrase
-10.8	Cabotegravir	Used for treatment and pre-exposure prophylaxis of HIV-1 infection.	Inhibitor of HIV-1 integrase
-10.7	Lurasidone	Used to treat schizophrenia and depressive episodes	Antagonist of Dopamine D2 Receptor and 5-hydroxytryptamine receptor 2A
-10.7	Tecovirimat	Used to treat smallpox, monkeypox, and cowpox	Inhibitor of Envelope Protein F13
-10.6	Cobimetinib	Used to treat unresectable or metastatic melanoma	Inhibitor of Dual specificity mitogen-activated protein kinase kinase 1

-10.5	Dolasetron	Used in chemotherapy and postoperatively to prevent nausea and vomiting	Antagonist of 5-hydroxytryptamine receptor 3A
-10.5	Bicalutamide	Used to treat Stage D2 metastatic carcinoma of the prostate	Antagonist of Androgen Receptor
-10.4	Ivacaftor	Used to treat cystic fibrosis	Potentiator of Cystic fibrosis transmembrane conductance regulator
-10.3	Avatrombopag	Used to treat thrombocytopenia	Agonist of Thrombopoietin Receptor
-10.3	Vorapaxar	used to reduce thrombotic cardiovascular events	Antagonist of Proteinase-activated receptor 1
-10.3	Dapagliflozin	used in the management of type 2 diabetes mellitus	Inhibitor/Antagonist of Sodium/glucose cotransporter 2

In Table 6, the deep neural network score ranks the top drug candidates. Each compound is scored by the model, outputting a number between 0 and 1, with a higher score indicating a higher similarity to the known inhibitor, and thus a higher probability of being a PARP-1 inhibitor. The model's score serves as a probability of PARP-1 inhibition. The drug candidate with the highest probability of inhibition, according to the model, is Lonafarnib, with a prediction score of (0.999964). However, the most reliable approach to identifying the strongest candidate is to combine the rankings of the protein-ligand docking with the neural network results. Fenoprofen also achieved a high docking score of -12.1 kcal/mol, indicating its potential as a PARP-1 inhibitor. Given the high docking score and strong backing from the machine learning model, Lonafarnib is a promising candidate for PARP-1 inhibition.

Table 5 - Top 20 predicted inhibitors by the Model's Prediction Score

Compound Name	Model Prediction Score	Docking Score (kcal/mol)
Lonafarnib	0.999964	-12.1
Ioflupane	0.999133	-10.1
Bictegravir	0.998813	-11.9
Atogepant	0.997899	-11.1
Cobimetinib	0.997713	-10.6
Sultamicillin	0.996654	-10.2
Lurasidone	0.996568	-10.7
Dolutegravir	0.995013	-10.8
Irinotecan	0.992985	-11.7
Avatrombopag	0.992983	-10.3
Dutasteride	0.992951	-12.7
Rimegepant	0.990847	-12.6
Ubrogepant	0.988249	-10.8
Cabotegravir	0.987173	-10.8
Ertugliflozin	0.986705	-10.2
Dolasetron	0.986392	-10.5

Tecovirimat	0.986151	-10.7
Adapalene	0.984745	-11.2
Empagliflozin	0.983438	-11.1
Bicalutamide	0.975614	-10.5

4 | Discussion

In this study, we revealed the potential ability of 82 FDA-approved to effectively inhibit PARP-1. The model that was developed achieved an average validation accuracy of 0.976, Matthews Correlation Coefficient of 0.952, and AUC of 0.989, displaying its ability to accurately distinguish between inhibitors and non-inhibitors of PARP-1, and to predict compounds that could potentially inhibit the protein. The efficacy of the model was further corroborated by the excellent performance of the predicted inhibitors in protein-ligand docking simulations, with the predicted inhibitors performing significantly better than non-inhibitors.With this, we have efficiently and economically highlighted 82 FDA drugs that could be implemented in further in vitro and in vivo trials, and eventually, clinical trials.

The practicality of repurposing FDA approved drugs is undeniable. Development times and costs plummet, making drugs available quicker and more cheap for patients who suffer from PARP-1 related diseases, such as Diabetic Cardiomyopathy. Additionally, the drugs have already been through the stringent safety tests of clinical tests, thus ensuring a lower risk of adverse effects, and allowing for a quick progression of the drug through the FDA pipeline.

In silico research can efficiently and cheaply elucidate promising drug candidates to expedite the creation of new treatments, and as new technologies emerge and accessible drug data continues to grow, machine learning and other computational algorithms will continue to improve and advance the pharmaceutical field.

References

1. Banday, Mujeeb Z et al. “Pathophysiology of diabetes: An overview.” Avicenna journal of medicine vol. 10,4 174-188. 13 Oct. 2020, doi:10.4103/ajm.ajm_53_20
2. “Diabetes.” World Health Organization (WHO), 16 September 2022, https://www.who.int/news-room/fact-sheets/detail/diabetes. Accessed 3 March 2023.
3. Gulsin, Gaurav S et al. “Diabetic cardiomyopathy: prevalence, determinants and potential treatments.” Therapeutic advances in endocrinology and metabolism vol. 10 2042018819834869. 27 Mar. 2019, doi:10.1177/2042018819834869
4. Jia, Guanghong et al. “Diabetic Cardiomyopathy: An Update of Mechanisms Contributing to This Clinical Entity.” Circulation research vol. 122,4 (2018): 624-638. doi:10.1161/CIRCRESAHA.117.311586
5. Ko, Hui Ling, and Ee Chee Ren. “Functional Aspects of PARP1 in DNA Repair and Transcription.” Biomolecules vol. 2,4 524-48. 12 Nov. 2012, doi:10.3390/biom2040524
6. Ke, Yueshuang et al. “The Role of PARPs in Inflammation-and Metabolic-Related Diseases: Molecular Mechanisms and Beyond.” Cells vol. 8,9 1047. 6 Sep. 2019, doi:10.3390/cells8091047
7. Morales, Julio et al. “Review of poly (ADP-ribose) polymerase (PARP) mechanisms of action and rationale for targeting in cancer and other diseases.” Critical reviews in eukaryotic gene expression vol. 24,1 (2014): 15-28. doi:10.1615/critreveukaryotgeneexpr.2013006875
8. Bohio, Ameer Ali et al. “c-Abl-Mediated Tyrosine Phosphorylation of PARP1 Is Crucial for Expression of Proinflammatory Genes.” Journal of immunology (Baltimore, Md. : 1950) vol. 203,6 (2019): 1521-1531. doi:10.4049/jimmunol.1801616
9. Qin, Wei-Dong et al. “Poly(ADP-ribose) polymerase 1 inhibition protects cardiomyocytes from inflammation and apoptosis in diabetic cardiomyopathy.” Oncotarget vol. 7,24 (2016): 35618-35631. doi:10.18632/oncotarget.8343
10. Haudek, Sandra B et al. “TNF provokes cardiomyocyte apoptosis and cardiac remodeling through activation of multiple cell death pathways.” The Journal of clinical investigation vol. 117,9 (2007): 2692-701. doi:10.1172/JCI29134
11. Szabó, C et al. “DNA strand breakage, activation of poly (ADP-ribose) synthetase, and cellular energy depletion are involved in the cytotoxicity of macrophages and smooth muscle cells exposed to peroxynitrite.” Proceedings of the National Academy of Sciences of the United States of America vol. 93,5 (1996): 1753-8. doi:10.1073/pnas.93.5.1753
12. Esraa M. Zakaria, Hany M. El-Bassossy, Nabila N. El-Maraghy, Ahmed F. Ahmed, Abdelmoneim A. Ali, "PARP-1 inhibition alleviates diabetic cardiac complications in experimental animals.",

European Journal of Pharmacology vol. 791 (2016): 444-454. https://doi.org/10.1016/j.ejphar.2016.09.008

13. Jacob O. Spiegel, Bennett Van Houten, Jacob D. Durrant, "PARP1: Structural insights and pharmacological targets for inhibition", DNA Repair vol. 103 (2021): 103125. https://doi.org/10.1016/j.dnarep.2021.103125.
14. Feiyue Zheng, Yi Zhang, Shuang Chen, Xiang Weng, Yuefeng Rao, Hongmei Fang, "Mechanism and current progress of Poly ADP-ribose polymerase (PARP) inhibitors in the treatment of ovarian cancer", Biomedicine & Pharmacotherapy vol. 123 (2020): 109661. https://doi.org/10.1016/j.biopha.2019.109661.
15. Novack, Gary D. “Repurposing medications.” The ocular surface vol. 19 (2021): 336-340. doi:10.1016/j.jtos.2020.11.012
16. Deotarse, P. P., et al. "Drug repositioning: a review." Int. J. Pharma Res. Rev 4 (2015): 51-58.
17. Stransky, Nicolai et al. “Can Any Drug Be Repurposed for Cancer Treatment? A Systematic Assessment of the Scientific Literature.” Cancers vol. 13,24 6236. 13 Dec. 2021, doi:10.3390/cancers13246236
18. American Diabetes Association. “Economic Costs of Diabetes in the U.S. in 2017.” Diabetes care vol. 41,5 (2018): 917-928. doi:10.2337/dci18-0007
19. Pushpakom, Sudeep et al. “Drug repurposing: progress, challenges and recommendations.” Nature reviews. Drug discovery vol. 18,1 (2019): 41-58. doi:10.1038/nrd.2018.168
20. Vamathevan, Jessica et al. “Applications of machine learning in drug discovery and development.” Nature reviews. Drug discovery vol. 18,6 (2019): 463-477. doi:10.1038/s41573-019-0024-5
21. Beatson, Erica L et al. “PARP inhibitors on the move in prostate cancer: spotlight on Niraparib & update on PARP inhibitor combination trials.” American journal of clinical and experimental urology vol. 10,4 252-257. 15 Aug. 2022
22. National Center for Biotechnology Information. "PubChem Protein Summary for Protein P09874, Poly [ADP-ribose] polymerase 1 (human)" PubChem, https://pubchem.ncbi.nlm.nih.gov/protein/P09874. Accessed 4 March, 2023.
23. RDKit: Open-source cheminformatics; http://www.rdkit.org
24. Yap, C. W. (2011) PaDEL‐descriptor: An open source software to calculate molecular descriptors and fingerprints. Journal of Computational Chemistry, 32, 1466–1474. https://doi.org/10.1002/jcc.21707
25. Whitten, I.H.; Frank, E.; Hall, M.A. Appendix B: The WEKA Workbench. Data Mining, 4th ed.; Morgan Kaufmann: Cambridge, MA, USA, 2016; pp. 553–571

26. Chollet, François, et al. "Keras." Keras, 2015, https://keras.io.
27. Abadi, Martín, et al. "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, software available from tensorflow.org.
28. Berman, H.M., Henrick, K., and Nakamura, H. "Announcing the Worldwide Protein Data Bank." Nature Structural Biology, vol. 10, no. 12, 2003, pp. 980.
29. Morris, G.M., et al. "Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility." J. Computational Chemistry, vol. 16, 2009, pp. 2785-91.
30. Eberhardt, J., Santos-Martins, D., Tillack, A.F., and Forli, S. "AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings." Journal of Chemical Information and Modeling, vol. 61, no. 5, 2021, pp. 2048-58.
31. Trott, O., and Olson, A.J. "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading." Journal of Computational Chemistry, vol. 31, no. 2, 2010, pp. 455-461.