



UNIVERSIDADE D
COIMBRA

Henrique de Vasconcellos Ávila

**DOCKING BASED STRATEGIES FOR THE
VALIDATION OF DEEP LEARNING
GENERATIVE MODELS
AND THE SEARCH FOR NOVEL LIGANDS FOR
A2AR, JAK2, KOR AND USP7**

**Dissertation in the context of the Master in Biochemistry, advised
by Prof. Maryam Abassi and Prof. Paula Veríssimo, presented to
the Faculty of Sciences and Technology - Department of Life
Sciences.**

September 2022

Faculty of Sciences and Technology

Department of Life Sciences

Docking based strategies for the validation of deep learning generative models

And the Search for Novel Ligands for A2aR, JAK2, KOR and USP7.

Henrique de Vasconcellos Avila

Dissertation in the context of the Master in Biochemistry, advised by Prof. Maryam Abassi and Prof. Paula Veríssimo, presented to the Faculty of Sciences and Technology - Department of Life Sciences.

May 2022



UNIVERSIDADE D
COIMBRA

Esta cópia da tese é fornecida na condição de que quem a consulta reconhece que os direitos de autor são pertença do autor da tese e que nenhuma citação ou informação obtida a partir dela pode ser publicada sem a referência apropriada.

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognize that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without proper acknowledgement.

Acknowledgments

Na produção desse trabalho, contei com o apoio de diversas pessoas, sem as quais a concretização do mesmo não seria possível. Gostaria de começar por agradecer aos Professores Joel P. Arrais e Maryam Abbasi, pela tutela e compreensão nessa fase de transição e aprendizado; sua orientação propiciou não apenas direção e conhecimento, mas também grande inspiração para desenvolver esse projeto.

Sou muito grato à Professora Paula C. Veríssimo, pelo apoio por todo meu percurso acadêmico na Universidade de Coimbra. Sua visão singular foi fundamental para o foco de aplicação dessa pesquisa.

Agradeço também a Carlos J. V. Simões, pela solicitude ao ensinar as bases das técnicas de docking e pelo auxílio com os diversos obstáculos encontrados.

Tenho também grande apreciação pelos colegas do LARN, que propiciaram um ambiente coleguismo e cordialidade, com interessantes reuniões e discussões. Em especial, agradeço a Nelson Monteiro e a Tiago Pereira, pelas colaborações, contribuições e sugestões.

Por fim, agradeço aos meus amigos e minha família, particularmente à minha parceira, Gabriela, e à minha mãe. Dos momentos austeros às celebrações, seu apoio foi imprescindível.

...

Acknowledgments

Financing

This research has been funded by the Portuguese Research Agency FCT, through D4 - Deep Drug Discovery and Deployment (CENTRO-01-0145-FEDER029266). This work is funded by national funds through the FCT - Foundation for Science and Technology, I.P., within the scope of the project CISUC - UID/CEC/00326/2020 and by European Social Fund, through the Regional Operational Program Centro 2020.

“A few lines of reasoning can change the way we see the world..”

STEVEN E. LANDSBURG [1, p. 193]

Resumo

A descoberta de medicamentos é um processo altamente demorado, complexo e caro, com baixas taxas de sucesso. Estimativas recentes apontam que para lançar um novo medicamento são necessários, em média, US\$ 1,8 mil milhões e doze anos de trabalho. Essa situação se deve em parte à alta dimensionalidade do espaço químico, que estima-se incluir entre 10^{33} and 10^{60} moléculas sinteticamente acessíveis [2]. Avaliar todo o espaço químico é proibitivamente caro, sendo portanto de extrema importância encontrar maneiras de restringir o âmbito de busca. Para tal, a inteligência artificial foi recentemente de várias maneiras incorporada na descoberta de medicamentos; dentre elas, modelos generativos profundos têm mostrado grande potencial para produzir possíveis candidatos a fármacos. Apesar disso, esta tecnologia ainda está em sua infância e possui algumas falhas fundamentais; por exemplo, tais modelos dificilmente levam em conta informações sobre configurações tridimensionais moleculares ou são validados por métodos realistas.

Este trabalho visa ajudar tal processo com proposto o sistema de validação baseado em docking molecular para modelos generativos de aprendizado profundo, conectando seu vasto potencial à pesquisa de medicamentos.

A metodologia de triagem foi testada por meio de análises de estudos de caso de quatro alvos farmacológicos de alto interesse (A2aR, JAK2, KOR e USP7) e consistiu em três etapas: avaliação de estruturas cristalográficas e ferramentas de docking, e teste e aplicação de triagem de moléculas e validação por meio de simulação de dinâmica molecular. Nesses experimentos, Autodock Fr e Vina demonstraram o mais alto desempenho tanto na previsão precisa da interação molecular quanto no cross-docking. Na segunda etapa, a técnica de consenso de pontuação exponencial foi avaliada, comparada a outros padrões preditivos, e apresentou alta acuidade, posicionando corretamente mais de 85% dos controles positivos dentro de margens muito restritas (5%); a metodologia foi posteriormente aplicada a conjuntos de moléculas produzidas por modelos generativos profundos para design de medicamentos. Na validação, foi demonstrada uma clara diferença na estabilidade ligante-receptor entre as moléculas de melhor e pior pontuação da etapa anterior, indicando a confiabilidade da metodologia proposta.

Observou-se também que, embora a configuração padrão possa ser um pouco demorada, os testes de modelos generativos subsequentes podem ser feitos em um intervalo de tempo muito reduzido, pois os dados resultantes de várias etapas podem

ser simplesmente reutilizados. Além disso, este método não é redundante quando comparado a outras métricas tradicionais, incluindo logP e drug-likeness, e pode ser usado em conjunto com estas para posterior avaliação. Os dados gerados por este método podem também ser usados como feedback para modelos generativos, auxiliando potencialmente em seu treinamento e aumentando a qualidade das moléculas geradas.

Palavras-Chave

Deep Learning, Avaliação de modelos, Redes Neurais Recorrentes, Desenho de Fármacos, Criação de Fármacos, Docking Molecular, Simulação de Dinâmica Molecular, GROMACS, Modelos Generativos Profundos, SMILES.

Abstract

Drug discovery is a highly time-consuming, complex, and expensive process with low success rates. Recent estimates point out that an average of \$1.8 billion and twelve years of work are required to launch a new drug. This state of affairs is partly due to how high-dimensional the chemical space is, as it has been estimated to include between 10^{33} and 10^{60} synthetically accessible molecules [2]. Evaluating the entire chemical space is prohibitively expensive, so it is of the utmost importance to find ways of narrowing down the search space. To this goal, artificial intelligence has recently been incorporated into drug discovery in many forms; among them, deep generative models have shown great potential for producing putative drug candidates. Even so, this technology is still in its infancy and possesses some fundamental flaws; for instance, these models hardly ever account for tridimensional molecular information or are validated through life-like methods.

This work aims to help this process with the provided molecular docking-based validation system for deep learning generative models, bridging their bountiful potential with drug discovery.

The screening methodology was tested through analyses of case studies of four high-interest pharmacologic targets (A2aR, JAK2, KOR, and USP7). It consisted of three stages: crystal structures and docking tools assessment, molecule screening testing and application, and validation through molecular dynamics simulation. In these experiments, Autodock Fr and VINA demonstrated the highest performance on both accurately predicting molecular interaction and cross-docking. In the second stage, the exponential consensus scoring technique was evaluated, compared to other predictive standards, and displayed high acuity, correctly placing over 85% of the positive controls within very strict margins (5%); it was subsequently, applied to sets of molecules produced by deep machine learning for computer-aided drug design. In validation, a clear difference in ligand-receptor stability between the best and worst scoring molecules of the previous stage was demonstrated, indicating the reliability of the proposed methodology.

It was also observed that, although the standard setup can be somewhat lengthy, the tests of subsequent generative models can be done in a far reduced time span, as the resulting data of multiple steps can simply be reused. Moreover, this method is non-redundant when compared to other traditional metrics, including logP and drug-likeness, and can be used in conjunction with these for further evaluation. Also,

the data generated by this method can be used as feedback for generative models, potentially aiding in their training and increasing the quality of the molecules generated.

Keywords

Deep Learning, Benchmarking, Recurrent Neural Networks, Drug Design, Drug Generation, Molecular Docking, Molecular Dynamics Simulation, GROMACS, Deep Generative models, SMILES.

Contents

List of Tables	xix
List of Figures	xxi
Abbreviations	xxv
1 Introduction	1
1.1 Context and Motivation	1
1.2 Objectives	2
1.3 Scientific Outcomes	4
1.4 Document Structure	5
2 Modern Drug Design and Discovery	7
2.1 Artificial intelligence	8
2.2 Deep Generative Models	10
2.3 State of the Art	12
3 Molecular Docking and Dynamics Simulation	17
3.1 Molecular Docking	17
3.1.1 Re-Docking and Cross-Docking	19
3.1.2 Sampling Algorithms	20

3.1.3	Scoring Functions	22
3.1.4	Exponential Consensus Scoring and decoys	23
3.2	Molecular Dynamics Simulation	24
3.3	State of the Art	26
4	Proteins Case Studies	29
4.1	G protein-coupled receptors	29
4.1.1	Adenosine 2a Receptor	31
4.1.2	Kappa Opioid Receptor	32
4.2	Enzymes	33
4.2.1	Janus Kinase 2	33
4.2.2	Ubiquitin-Specific Protease 7	34
5	Materials and Methods	39
5.1	The Proposed Workflow	39
5.2	Datasets	41
5.2.1	Protein structures and preparation	41
5.2.2	Ligands with known binding affinity	41
5.2.3	Deep learning generated molecules	42
5.3	Bioinformatic tools	42
5.3.1	Docking software and structure preparation	42
5.3.2	Fragment analysis, consensus scoring, and drug-likeness prediction	43
5.3.3	Molecular dynamics simulation	43
6	Results and Discussion	45
6.1	Docking Tools Assessment	45
6.2	DGM Molecules Screening	48

6.3	Validation Through Molecular Dynamics Simulations	50
6.4	Additional Chemometric Analyses	50
7	Conclusion	55
7.1	Conclusions	55
7.2	Future Work	56
A	Appendix	59
A.1	Appendix A	59
A.2	Appendix B	64
	Bibliography	71

List of Tables

2.1	GuacaMol and MOSES benchmark metrics.	13
3.1	Overview of the tested docking software with scoring functions and sampling algorithms.	18
4.1	Examples of substrates of USP7 related to carcinomas.	36
5.1	List of the PDB structures used and cognate ligands.	41
6.1	Morgan fingerprint fragments. The table show the correlation weights as related to over- or underprediction of the Morgan fragments displaying higher relevance.	49
6.2	MDS results for (a) A2aR, (b) JAK2 , (c) KOR (d) USP7 ligands. Carbon H bond (CHb), Conventional H bond (CoHb), Distance (Dist.), Unfavorable Donor-Donor (Unf. D-D)	51
A.1	List of A2aR ligands	59
A.2	Listo of JAK2 ligands	60
A.3	List of KOR ligands	62
A.4	List of USP7 ligands	63

List of Figures

2.1	AI's layered structure. Displayed here are the subdivision of AI discussed in this work, directed toward deep learning. The generative models' division, which possesses instances among all the represented groups, is indicated in purple.	9
2.2	NNs' architecture. The figure shows a general representation of how NNs are designed, with the neurons for input, hidden, and input layers represented by red, blue and green circles, respectively.	10
2.3	Example of DGM design. This is a DGM system from Pereira 2021 [68] containing two Generators sharing the same architecture, with an interconnected Predictor by Reinforcement Learning.	11
3.1	Re- and Cross-docking. Re-docking is displayed as a schematic in (a), and the cells with blue background in (b), a schematic of cross-docking. The procedure is often performed with multiple repetitions (c) to eliminate variability. After MD procedures are finished, the RMSD of the resulting poses is measured against the original ligands, and the group result for each crystal structure is aggregated as a mean (d). These scores can then be compared with those achieved by the other MD tools (e)	19
3.2	Visual examples for molecule RMSD. Various poses for the A2aR ligands Adenosine, ZMA and UKA are displayed in poses that resulted in good (under 2Å), acceptable (between 2 and 3Å) and poor (over 3Å) solutions.	20

-
- 4.1 GPCR structure and activation. The activation of GPCRS (b) by coupling with their agonist ligands (a) results in conformational changes that activate their coupled G-protein (c), which is responsible for modulatory action and signal transduction (d) [176]. 30
- 4.2 A2aR structure and binding site. (a) shows A2aR's binding site in dark blue,(b) the tridimensional representation of adenosine bound to A2aR's binding site, and (c) its 2D representation showing residue interactions. A2aR-caffeine interactions are displayed in 3D in (d) and in 2D with residue interactions in (e). 31
- 4.3 KOR structure and binding site. (a) shows KOR's structure (green) with it's binding site (red volume). (b) is a representation of the binding site on a frontal view (from outside of the cell) with the protein (green), binding site residues (cyan) and its ligand (CVV, in orange). In (c), a 2d representation of the primary interactions between CVV and the binding site is provided. 32
- 4.4 OR structure and binding site. (a) shows KOR's binding site in red,(b) N-[(5alpha,6beta)-17-(cyclopropylmethyl)-3-hydroxy-7,8-didehydro-4,5-epoxymorphinan-6-yl]-3-iodobenzamide bound to KOR's binding site, and (c) is the 2D representation showing residue interactions. . . 34
- 4.5 JAK2 structure and binding pockets. The homology domain JH1 is presented in teal, and JH2 in blue, with the activation loop, highlighted in orange. In (a), a detailed view of the enzyme structure is presented, with the binding pocket used in this work displayed in grey. In (b), the protein surface is shown, and the ATP, Helix C, and Activation-loop pocket is indicated. 35
- 4.6 USP7 structure. This image was generated by aligning pdb structures 2f1z, 4z97 and 5fwi. Ubl 1 to 5 are represented in hues of green and blue, TRAF is shown in yellow, the catalytic domain (containing the catalytic cleft, palm, thumb and fingers) in orange and an ubiquitin molecule in docked position is appears in red. 37

5.1	General workflow. The diagram shows an overview of the proposed methodology, with block I standing for the MD tools selection process, II demonstrates the screening of deep learning generated molecules and III shows the validation of the selected molecules through MDS. .	40
6.1	Binding pockets. DoGSiteScorer generated binding pockets for A2aR(a) JAK2(b), KOR(c) and USP7(d)	45
6.2	RMSD heatmaps. Heatmaps expressing the RMSD scores for the cross-docked ligands of the four proteins. (a) represents A2aR, (b) JAK2's, (c) KOR and (d) USP7 structures.	46
6.3	RMSD results presented in the form of box plots.	47
6.4	Binding energy regression plots. The plots display linear regression between the docking score of the three best MD tools and experimentally determined ligand pIC50 or pKi. (a) A2aR, (b) USP7, (c) JAK2, (d) KOR.	48
6.5	A2aR and USP7 ROC curves. (a) represent the ROC curve For A2aR MD results, and (b) that of USP7. The graphs display the performance of ADFR (in red), VINA (in green), exponential consensus scoring (in black) and QED (in purple)	50
6.6	Consensus scoring plots. This distribution shows the MD scoring ranking with ADFR (x-axis) and VINA (y-axis); in this methodology, the closer to zero a molecule is placed, the higher its score. The blue lines indicate the cut-off mark, set as 5% for a more strict evaluation. Molecules have been separated by type as decoy, target or known binder.	52
6.7	A2aR MDS results. The MDS results for the best and worst five molecules are displayed in a and b, respectively. In c the 2d representation of the final pose and interaction of the best consensus scoring molecule with the receptor, d represent the same, but for the worst. .	53
6.8	USP7 MDS results. The MDS results for the best and worst five molecules are displayed in a and b, respectively. In c the 2d representation of the final pose and interaction of the best consensus scoring molecule with the receptor, d represent the same, but for the worst. .	53

6.9	Chemometric analyses. (a) Shows a graph used to estimate compartment-specific absorption, with TPSA at the x-axis and LogP in y, with A2aR data. The consensus scoring is represented in a hue going from blue to yellow, and the zone where blood-brain barrier transposing molecules are expected to fall is delineated by the large green circle. (b) shows the A2aR exponential consensus graph with QED values shown as the dot's hue, with good values (0 and over) in green, intermediate values (between -0.5 and 0) in yellow, and poor results (under -0.5) in red	54
A.1	A2aR regression plots. Results for binding affinity and docking scores. (a) global, (b) AD4, (c) ADFR, (e) PLANTS, (f) rDock, (g) VINA	65
A.2	JAK2 regression plots. Results for binding affinity and docking scores. (a) global, (b) AD4, (c) ADFR, (e) PLANTS, (f) rDock, (g) VINA	66
A.3	KOR regression plots. Results for binding affinity and docking scores. (a) global, (b) AD4, (c) ADFR, (e) PLANTS, (f) rDock, (g) VINA	67
A.4	USP7 regression plots. Results for binding affinity and docking scores. (a) global, (b) AD4, (c) ADFR, (e) PLANTS, (f) rDock, (g) VINA	68
A.5	Ligand stability in MDS. Representation of the positional stability of the best and worst exponential consensus ranking molecules in the MSD experiments. Graphs of the RMSD (y axis) over time (x axis) for the top five, 6th to 10th, and worst five molecules (in relation to target molecule consensus scoring) for (a) A2aR, (b) JAK2, (c) KOR, (d) KOR. Lower values denote ligand positional stability and, therefore, stronger coupling with the receptor.	69

Abbreviations

A2aR Adenosine 2a receptor

AD4 Autodock 4

ADFR Autodock FR

ADMET Absorption, Distribution, Metabolism, Excretion and Toxicity

AI Artificial Intelligence

BBB Blood-Brain Barrier

CNS Central Nervous System

DL Deep Learning

DGM Deep generative model

DTI Drug-Target Interaction

FDA Food and Drug Administration

GPCR G protein-coupled receptor

JAK2 Janus kinase 2

KNN K-Nearest Neighbors

KOR κ - opioid receptor

KOR Kappa opioid receptor

logP octanol-water partition coefficient

LSTM Long Short-Term Memory

MD Molecular Docking

MDS molecular dynamic simulations

ML Machine Learning

MS Molecular Screening

MSE Mean Squared Error

QED quantitative estimate of drug-likeness

QSAR quantitative structure-activity relationship

RF Random Forest

RL Reinforcement Learning

RMSD Root-mean-square deviation

ROC Receiver operating characteristic

SAR Structure-activity relationship

SMILES Simplified Molecular Input Line Entry System

STATs signal transducer and activator of transcription proteins

TPSA Total polar surface area

USP7 Ubiquitin specific peptidase 7

List of Figures

VINA Autodock VINA

Introduction

1.1 Context and Motivation

Molecular docking (MD) is a tool that enables the simulation of binding between two molecular structures into stable states based on interatomic affinities [3]. It is frequently used in structure-based drug design due to its ability to predict the binding-conformation of small molecule ligands to the appropriate target binding site [4,5]. This is particularly useful since many pharmacologically active compounds act by coupling with a target protein and either promoting its activation or inhibiting it [6].

Docking enables the identification of novel compounds of therapeutic interest, predicting ligand-target interactions at a molecular level, or delineating structure-activity relationships (SAR), without a priori knowledge of the chemical structure of other target modulators. It is also a relatively fast tool, which makes it an interesting tool for screening large libraries [7].

MD is now said to be in a mature state of development, with most docking software being to indicate results with great accuracy 80% of the time, with many far surpassing that margin [8,9]. Despite that, the technique is not without its faults; for instance, the estimation of binding energy is still a challenge and, because of that, is held as the most strict test of an MD tool's capacity [10]. Other than that, two other factors are known to commonly interfere with docking results:

- Ligand representations: since the recognition of interaction between ligand and protein depends on polarization, 3D conformation, protonation state, and electrostatic complementarity, the correct generation and treatment of the structure of ligands are imperative [11,12].
- Receptor representations: the primary source of protein structural files are

X-ray crystallography and NMR, very accurate techniques, but that may present problems under certain circumstances and, for instance, omit atoms. Moreover, sometimes portions of these structures are artificially filled through threading, homology modeling, and *de novo* methods, Which are not ideal and may negatively impact model quality, reducing docking acuity [13–16].

Because of this, basic knowledge of the tools and applying strategies to minimize such interferences are paramount. Usually, it is recommended to at least perform cross-docking experiments and follow rigorous structure preparation practices in any MD campaign [17, 18].

As the cost of drug research is consistently increasing, the need for modernization of tools and techniques for drug design is ever more present [19]. Recently, advancements in computer hardware technology and software architecture made *in silico* drug design approach tangible.

Among these, deep generative models (DGM) have promising potential. Unfortunately, most DGMs are trained and benchmarked using extremely simplified parameters and hold little semblance to actual molecular characteristics or ligand-target interaction. With the most commonly used being quantitative structure-activity relationship (QSAR), octanol-water partition coefficient (logP), and quantitative estimate of drug-likeness (QED) [20, 21].

Moreover, the two most well-known evaluation frameworks, GuacaMol and MOSES [22, 23], primarily focus on distribution tasks [24, 25], and present a set of tools unable to assess the molecules’ chemical feasibility, stability, or target affinity [26, 26–28].

MD, being a salient tool in structural molecular biology and computer-assisted drug design, could assist in bringing DGMs to conditions that more closely resemble experimental reality.

1.2 Objectives

The main goal of this thesis is to research, analyze and evaluate docking based methods for the validation of the compounds that generated by the machine learning and deep generative models. The objectives can be summarized as follows:

O1 - Overview MD techniques, molecular screening, DGM benchmarking, and *de novo* drug design.

A thorough review of these topics is to be performed, focusing on expounding their capabilities, flaws, characteristics, and commonly employed techniques.

O2 - Assess a method for benchmarking MD tools conducive to the project end goal.

Several evaluation methodologies exist for MD tools, each gauging these software on different capabilities and characteristics. If possible, methods flexible enough to not depend on fixed pre-established benchmarking sets should be selected to make it possible to be applied with newer crystal structures. This is especially important since ML *de novo* drug design focuses mainly on novel targets with few known active pharmaceutical molecules.

O3 - Select protein targets currently relevant to pharmacologic research and ML model development

Targeting proteins linked to a larger body of specific ML work and overall scientific production would be advantageous. It would make it possible later to apply the benchmark methodology to multiple models and also for it to be of interest to researchers interested in developing new ones.

O4 - Determine the best-suited MD tools and crystal structures for the experiments with each target protein.

Choosing MD tools and crystal structures are vital steps in molecular screening. As such, a thorough selection of these software and files must be performed to maximize the probability of success for this project.

O5 - Assess a molecular screening technique able to select molecules with high affinity for their receptors and validate its capabilities.

Multiple molecular screening protocols for drug design have been devised since the establishment of this technique; these vary in the tools used and the information extracted from the molecules being evaluated. Ideally, techniques involving binding affinity and with high predictive value must be found.

O6 - Create an easily reproducible MD-based protocol to evaluate and compare *de novo* drug design ML models.

This work aims to produce a straightforward and flexible workflow focused on open-source and free tools, capable of selecting the most accurate MD tools and ultimately

assessing the capability of each DGM to produce molecules with good affinity for their target protein.

1.3 Scientific Outcomes

The contributions were submitted to international journals. They are listed, in chronological order, for each type of venue, together with reference to the contribution.

P1 - Nelson RC Monteiro, Carlos JV Simões, Henrique V. Ávila, Maryam Abbasi, José L. Oliveira, and Joel P. Arrais. “Explainable deep drug–target representations for binding affinity prediction”, *BMC bioinformatics* 23, no. 1 (2022): 1-24. Published 17 June 2022.
(**IF :3.169** (2020))
DOI: 10.1186/s12859-022-04767-y

P2 - Henrique V. Ávila, Maryam Abbasi, Joel P Arrais, Paula C. Veríssimo, Carlos J. V. Simões, “Comparison and Benchmark of Molecular Docking Tools for Drug Screening”, VIII EJBCE (Encontro de Jovens Investigadores de Biologia Computacional Estrutural), 20th of December 2021. (Poster presentation)

P3 - Henrique V. Ávila, Tiago Pereira, Joel P Arrais and Maryam Abbasi, “Docking based strategy for the evaluation of deep models in de novo drug design”, *Computational Biology and Chemistry*, 28th of July 2022. (Submitted)

During this work, there was also the opportunity of presenting and discussing the main topics of this dissertation:

- *Molecular Docking Tools and Screening*, Department of Informatics Engineering, University of Coimbra, March 2021.
- *Molecular Dynamics Simulation*, Department of Informatics Engineering, University of Coimbra, April 2021.
- *Application of Docking-based Molecular Screening into de novo drug design*, Department of Informatics Engineering, University of Coimbra, May, 2021.
- *Validation and Evaluation of Deep Generative Models*, Department of Informatics Engineering, University of Coimbra, March 2022.

1.4 Document Structure

The following reading is divided into six parts: state of the art, overview, methods, results, conclusion, and appendices. As we developed this work, we observed it pertains to a niche bridging two vastly different fields, biochemistry, and machine learning. As the professionals of these areas tend to differ profoundly in their focus of knowledge, we found it substantial to make a deeper delineation of the basis of the techniques and concepts involved.

In modern drug design, an overview of the proceedings of drug research is given, with a focus on the contribution of artificial intelligence to this field. There, an explanation is also given on the characteristics and basis of artificial intelligence, deep generative models, the state of the art, and the benchmark for these systems.

Molecular docking and dynamics simulation expound on the intricacies of these two tools, which are the core of the work produced in the thesis. The different component classes and their characteristics are explained, focusing on their strengths, weaknesses, and applications. The last subsection outlines the state of the art in docking-based molecular screening and how the two tools are applied.

In protein case studies, the characteristics of the selected protein targets and their importance for current drug research are displayed. This section is divided into two parts: Enzymes (containing the Janus Kinase 2 and Ubiquitin-Specific Protein 7) and G protein-coupled receptors (with the Adenosine 2a Receptor and Kappa Opioid Receptor).

The overview chapter provides explicit technical and specific knowledge upon which this work is based in the form of a literature review. It was constructed to contain the background information related to drug discovery, the tools used (molecular docking, chemometric analyses, and molecular dynamics simulations), and the datasets employed (protein targets, target-specific ligands, machine learning generated molecules, and decoys).

The remaining chapters present the experimental setup and results obtained in this work, as well as the observations and conclusions we took from the analysis of the resulting data.

We thank you for your time and consideration, and hope that the reading of this document will be pleasant.

Modern Drug Design and Discovery

Drug discovery is the process through which molecules are tested to uncover their therapeutical capabilities and application possibilities. This is a lengthy and costly undertaking, with mean drug costs at 1.3 billion USD and taking 10 to 15 years to clear all phases of clinical trials [29,30]. On top of that, 90% of drug discovery campaigns fail [31], and overall drug discovery costs seem to be steadily increasing over the years [30].

One proposed means for reverting this trend is a change in the classical understanding of drug design, the process of developing new medicaments based on biochemical information of a target or of molecules with known pharmacological activity. Traditionally, drug design and discovery depended on compound libraries and wet-lab experiments, which, despite innovation trends, are still responsible for a vast portion of research costs. But now, the technological advancements of later years have opened an array of possibilities in the camp of computer-aided drug design. Digital screening of far more extensive compound libraries can now be performed in mere minutes, artificial intelligence-aided molecule generation has become a possibility, and experiments that would be cost-intensive and time-consuming can often be quickly performed *in silico* [32–34].

Consequently, bioinformatics has been steadily integrated into pharmaceutical research, and not only for small molecules. For instance, drug delivery systems, biopharmaceuticals, and antibodies can now be designed to have greater specificity, affinity, and stability [35]. Structural analysis can provide insights on function and properties, proteins and small molecules can be altered via specific software, and tests for interaction and overall performance can be done through simulations that

can be highly customized [36, 37].

One of the latest technological improvements to pharmaceuticals has been incorporating artificial intelligence. Deep learning models have shown inordinate promise in the field, exemplified by AlphaFold’s impressive capability to predict the final tridimensional structures based only on their amino acid sequences [38]. And also by the disclosed interest of market giants such as Roche, Johnson & Johnson, Pfizer, and others in their application into drug discovery [39, 40]. Finally, there is the case of Paxlovid, Pfizer’s COVID-19 oral drug, an example of a successful example for machine learning generated pharmaceuticals. However, artificial intelligence models haven’t yet developed a drug from scratch [40, 41].

2.1 Artificial intelligence

Artificial intelligence (AI) is usually defined as the development of systems capable of simulating behaviors commonly associated with humans, such as decision making and judgement [42]. AI is a field with multiple subcategories (Figure 2.1); it is also one in fast expansion and displays extensive permeation in the contemporary world, being found throughout several aspects of daily life. For instance, facial recognition, many spell check and text prediction tools, google search, voice recognition in smartphones and appliances, banking systems, social media, and plagiarism checkers are a few examples of AI being incorporated into our surroundings [43–45].

Machine learning (ML) is a division of artificial intelligence focused on algorithms capable of improving themselves as it works through raw data, be it quantitative or categorical, developing and enhancing their competence to perceive patterns and make decisions. Mostly, these systems conduct classification, clustering, regression, or pattern recognition tasks on large datasets [46–48]. ML is being used in the pharmaceutical sector to create new drug molecules, investigate biological activities and interactions, predict molecular characteristics, and many other applications [49–52]. Recently, there have been even discussions about using these systems in the development of individualized medicine, where the molecular composition and dosage of drugs would be tailored, taking into account the specific biochemical pathways of each person [53–55].

ML algorithms can be classified according to the specificities of their learning approaches into three categories [56, 57]:

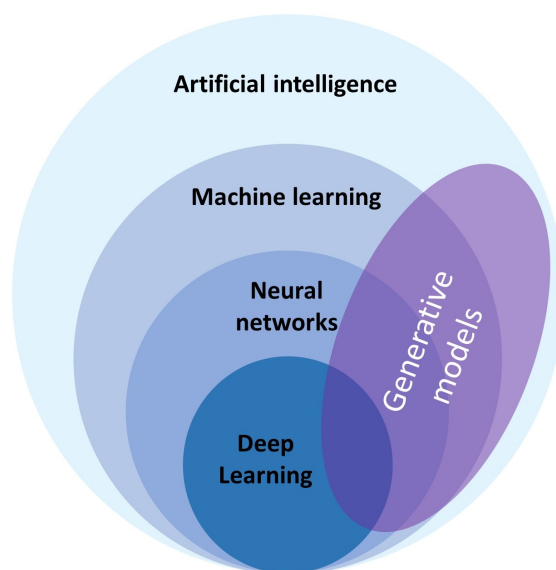


Figure 2.1: AI's layered structure. Displayed here are the subdivision of AI discussed in this work, directed toward deep learning. The generative models' division, which possesses instances among all the represented groups, is indicated in purple.

- **Supervised Learning**, a method that employs labeled data, involving an environment in which inputs correspond to corresponding outputs. A model with this aspect will iteratively adjust its network parameters to match a given ground truth. After successful training, they become capable of providing solutions for external data.
- **Semi-supervised Learning**, which makes use of both labeled and unlabeled data for training.
- **Unsupervised Learning**, represented by systems that are fed only unlabeled information and which are able to uncover relationships and correlations from said data, often availing of clustering, dimensionality reduction, and generative techniques.

Additionally, some ML systems make use of a methodology called Reinforcement Learning. Usually discussed under the purview of semi-supervised and unsupervised learning, this is an approach employed for unknown environments [58], with the most famous example being Google Deep Mind [59, 60].

Some ML models are characterized by their inspiration from biological systems, focusing on emulating the brain's structures. These are given the name of Neural Networks (Figure 2.2), and are usually represented as a set of multiple layers con-

taining multiple sub-units called neurons, responsible for cross-processing input it receives and producing an output [61, 62]. These are classified as input, hidden, and output layers, and a system containing more than one hidden layer is classified as a deep network, and the models employing it as being part of Deep Learning (DL) [62]. This increase in architectural complexity in relation to other ML is justified by, amongst others, demonstrating improved performance in systems containing a higher magnitude of information and better results for unlabeled data [?, 63, 64]. For the purposes of de novo drug design, scientific work seems to focus on the intersection of DL and generative models: Deep Generative Models (DGM).

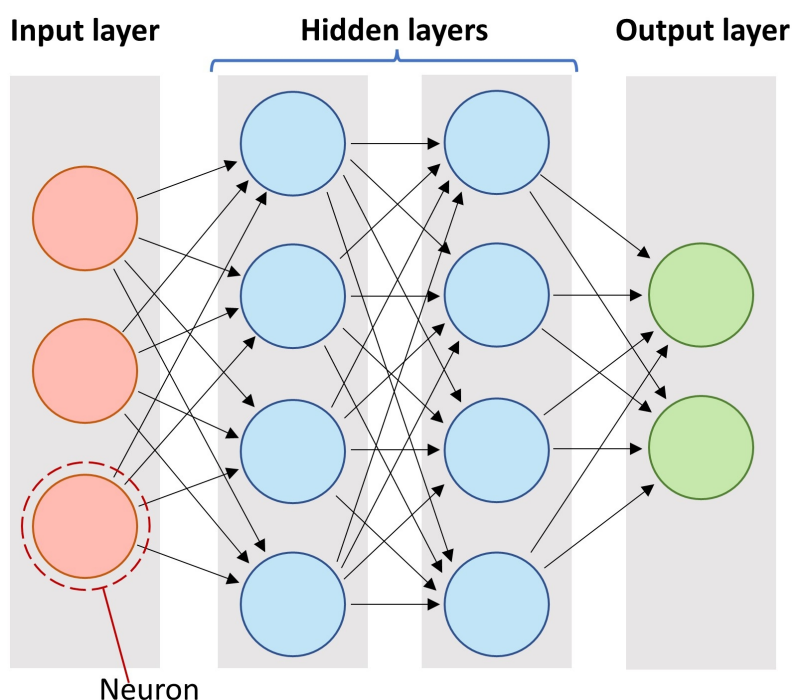


Figure 2.2: NNs' architecture. The figure shows a general representation of how NNs are designed, with the neurons for input, hidden, and input layers represented by red, blue and green circles, respectively.

2.2 Deep Generative Models

Generative models are a subset of machine learning algorithms developed to create new data instances. They are able to detect characteristics present in elements of a dataset and make predictions on what should be the next data point or produce a new one based on similarity and probability. A straightforward example of these is text completion tools, which can predict the next word a person will use in a sentence

while typing. In chemistry, these are usually employed to design new molecules and composites to integrate innovative materials and drugs [65–67]. An example of DGM architecture can be seen in Figure 2.3

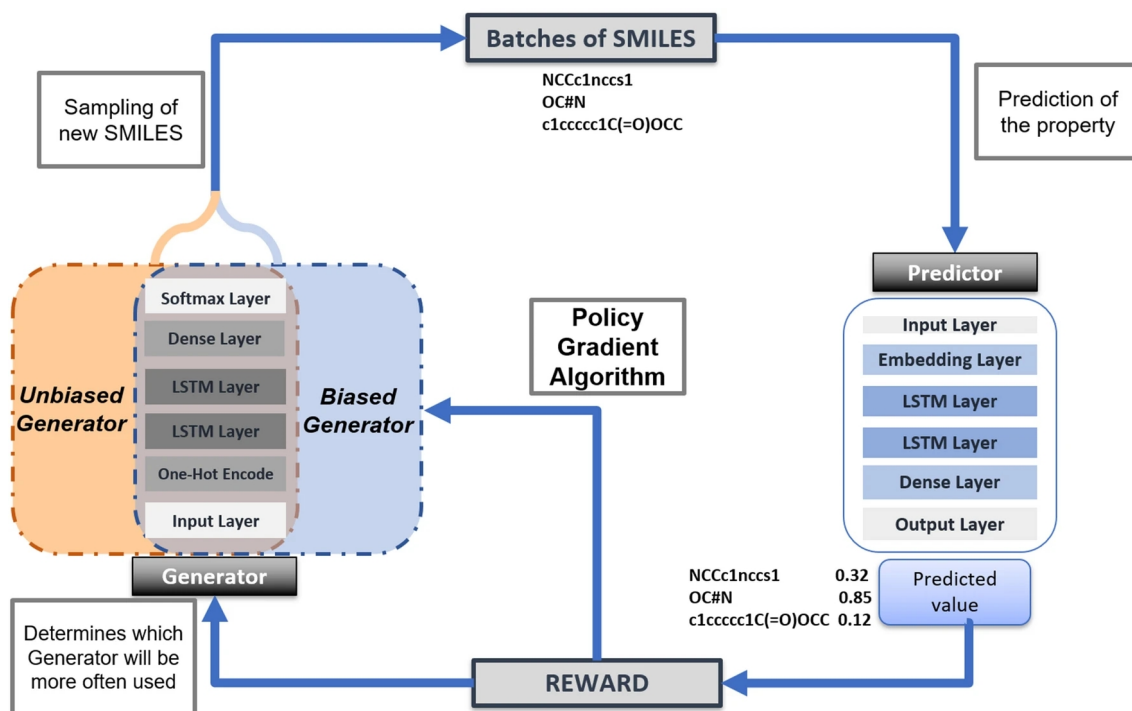


Figure 2.3: Example of DGM design. This is a DGM system from Pereira 2021 [68] containing two Generators sharing the same architecture, with an interconnected Predictor by Reinforcement Learning.

Although these strategies have to be specifically designed for each application, the following steps are common in creating and employing these models: data collection, mathematical descriptors' generation, search for the best subset of variables, model training, and validation [69]. In drug discovery, the data collection step customarily is acquiring a dataset of molecules with distinctive characteristics, such as affinity for a specific target, low toxicity, ease of absorption and laboratory production, and enough solubility to be distributed through the body [70]. These are usually small molecules procured from public repositories, namely ZINC [71], ChEMBL [72], PubChem [73] and others. Also, these models typically use simplified 1D formats such as SMILES and FASTA to reduce the computational workload [74, 75].

The last step, model validation, is usually one of the most critical for these DGMs. This final part of the procedure consists of acuity tests based on the characteristics of the molecules generated in these models. If statistically significant, this determines

that the tool produced is an accurate predictive model. Unfortunately, most of these tools are ultimately validated through tasks such as creating molecules with ideal octanol-water partition coefficient [76] or the maximization of quantitative estimate of drug-likeness [21]; Despite being compelling tests to assess the model’s capability to produce molecules based on a set of tasks, these fail to account for the complexity of realistic tridimensional models and environments [76–79].

2.3 State of the Art

Recently, DGMs have increasingly displayed potential for generating new molecules for pharmaceutical and materials science applications. As the interest in these systems rises, it is also essential to develop accurate benchmarking models that are reproducible and reliably capable of selecting outputs that represent desirable molecules. Regrettably, most of the models in use at the moment present flaws that hinder DGM’s advancement.

In only a decade, the new generation of deep learning has redrawn the basis of how we perceive the production of multiple data genres, with numerous examples of systems capable of independently producing original images, music, and text [80]. Around 2012, deep neural networks began to displace other architectures in the top positions of contests, such as the ImageNet image classification challenge. In that year, a system by Ciresan *et al.* was able to perform handwritten character recognition almost as well as a person [81]; another, by Dahl *et al.* was the best placed in the "Merck Molecular Activity Challenge" [82].

These achievements led to the development of deep systems intended for text, image, and audio generation, in addition to the more traditional classification and regression tasks. And in 2013 and 2014, respectively, the variational autoencoder and generative adversarial network were established [83,84]. Then, in 2016, Gómez-Bombarelli *et al.* published the first molecular autoencoder, opening the gates for *in silico* research for drug molecules [85].

The popularity and success of modern machine learning models have arguably been furthered by the availability of large datasets and standardized benchmarking models, providing a paradigm for model development [23, 86–88]. Regarding organic molecules, the introduction of MoleculeNet [89] was an essential milestone for benchmarking regression and classification tasks. After that, the evaluation framework GuacaMol [25] was proposed with the concept that generative models should be

evaluated based on goal-oriented and distribution learning tasks; this development consequently led to the development of MOSES [24] platform.

GuacaMol and MOSES are two of the most prominent standardized benchmark efforts available. Their vast set of metrics (Table 2.1) can be observed to gain insight into the state of the current DGM assessment tools [22,23]. For starters, both tools rely heavily on distribution tasks [24,25]. Undeniably, these are fundamental for the development of DGMs, but offer far too little on the side of weeding out chemically unrealistic molecules and offer nothing on target affinity [26,27]. Other than that, a high score in these models doesn’t translate to stable or synthetically accessible molecules, especially for goal-directed tasks [26,28].

Table 2.1: GuacaMol and MOSES benchmark metrics.

Metric	Delineation	Type	Model
Filters	Fraction of generated molecules able to pass data construction filters.	Distribution	MOSES, GuacaMol
Unique molecules	Analysis of the fraction of unique molecules	Distribution	MOSES, GuacaMol
Valid molecules	Analysis of the fraction of valid molecules	Distribution	MOSES, GuacaMol
Fragment similarity	BRICS fragment comparison assessing over or under-representation	Distribution	MOSES
Fréchet Distance	Analysis of Fréchet ChemNet chemical space coverage	Distribution	MOSES, GuacaMol
Internal diversity	Detection of model collapse through chemical diversity assessment.	Distribution	MOSES
KL divergence	Kullback–Leibler divergence, a metric related to diversity between training set and generated molecules.	Distribution	GuacaMol
Novelty	Ratio of creation of molecules not found in the training set	Distribution	MOSES, GuacaMol
Properties distribution	Analysis of chemical properties of generated molecules.	Distribution	MOSES
Scaffold similarity	Comparison of Bemis-Murcko scaffolds contained in training and generated sets.	Distribution	MOSES
Nearest neighbour similarity	Tanimoto similarity between molecules and their nearest neighbor molecules.	Distribution	MOSES
Compound quality	Analysis of molecular properties through Walters’ rd_filters implementation.	Goal Directed	GuacaMol
Isomers	Assessment of the model’s capability of producing molecules corresponding to a specific molecular formula.	Goal Directed	GuacaMol
Median Molecules	Conflict task involving maximizing the similarity to multiple molecules simultaneously.	Goal Directed	GuacaMol
Rediscovery	Assessment of the model’s capability of reproducing a specific molecule.	Goal Directed	GuacaMol
Similarity	Evaluation of the model’s competence in generating molecules similar to a target.	Goal Directed	GuacaMol

DGMs are frequently tested on specific goal-oriented tasks, with a majority focusing on maximizing QED and logP [20,90]. These carry little to no correlation to chemical stability or ligand-target interaction and are increasingly seen as meaningless errands [76,91]. It is also important to note that a substantial decrease in scientific rigour in recent DGM works has been reported [92].

Benchmarking sometimes incorporates estimations of target affinity, such as IC50 predictions. This method is a more directly correlated metric, with better applicability to pharmacological molecule research [93,94]. Although, this approach is not without its hurdles, having reported low replicability difficulty of assessment [94–96]. Another pitfall of relying on affinity prediction methodologies is that, although they can in certain circumstances be reasonably accurate, their principles usually translate poorly to molecule construction [97, 98]. For instance, hydrogen bonds are fundamental for protein-ligand interaction and usually heavily influence binding affinities in pharmacologically active molecules and ligands in general [97, 99]. Hence, adding hydrogen donors and acceptors may be perceived as a logical step for affinity increase, albeit this assumption is frequently proven wrong in synthesized molecules [100, 101].

Another interesting element is that tridimensional molecular information is scarcely incorporated in either DGMs or benchmarking [102, 103]. It is well established that the position of chemical groups in a binding site is paramount for agonist and antagonist action. And although it is somewhat more permissive for the latter, alterations as simple as introducing a rigid bond can significantly impair the molecule’s activity in both cases [104–106]. As a more practical example of the disparity between 3D models and simpler systems, the same addition of a hydrogen donor that would increase the molecule’s affinity, polar surface, and logP scoring can just as easily create intramolecular interactions that would change its geometric conformation. Furthermore, this kind of bond could also compete with amino acid residues in the binding site; and that would be difficult for uni- or bi-dimensional models to detect. [107–111].

In conclusion, we observe that DGMs are a budding field with vast promise and the potential to reshape the molecular design and many other industries, but many flaws have yet to be polished before that comes to pass. We are among those that are especially concerned with the simplistic approach that leads programmers to assume that molecules can be treated as text or collections of fragmented images throughout the whole generative process, as this creates unrealistic models and unfeasible

molecules that distance DGMs from the pragmatic pharmacologic application.

Molecular Docking and Dynamics Simulation

This chapter explains the fundamentals of the main tools used in this thesis, MD and MDS. It is divided into three parts: MD, MDS, and a brief review of the state of the art.

3.1 Molecular Docking

Although laboratory means, such as IC50 essays and x-ray crystallography, can be helpful in molecular interaction evaluation, these approaches demand multiple specialized equipments, know-how, and a vast amount of time and resources. Other than that, the success of these techniques is by no means simple to achieve or guaranteed. If the synthesis of any molecule (proteins or small molecules) fails, or if the crystallization process is flawed, the whole process may be wasted. As such, wet experimental models are unsuited for large screening jobs and can heavily impact the cost and duration of drug research efforts. For this reason, computational techniques are often the tools of choice when large screening campaigns are involved, with molecular docking (MD) having a steadfast growth in popularity in later years. It can also be beneficial when crystallographic means fail to determine the conformation or structure of bound ligands, as the results of MD may help formulate a hypothesis on the position of the obscured parts of the small molecule.

MD is dependent on 3D structure files of proteins and ligands. Ligand files can be easily generated through multiple computational means. On the other hand, protein structural files are mainly created through crystallography, but once these are solved and available, MD opens an ample array of possibilities for biochemical research. Most docking software generates multiple poses for the ligand in a predetermined

3. Molecular Docking and Dynamics Simulation

Table 3.1: Overview of the tested docking software with scoring functions and sampling algorithms.

Software	Sampling Algorithm	Scoring Function
Autodock 4	Lamarckian Genetic algorithm	Empirical Physics-based
Autodock FR	Random perturbation, local optimization	Empirical Knowledge-based
Autodock Vina	Random perturbation (Metropolis-Hastings), local optimization (BFGS)	Empirical Knowledge-based
LeDock	simulated annealing evolutionary optimization	Empirical Physics-based
PLANTS	Ant colony optimization	Empirical
rDock	Genetic Algorithm, simulated annealing, simplex minimization	Empirical

search space and evaluates their binding affinity (scoring) to the protein. This capability is conferred by their structure, containing two essential parts: a sample algorithm and a scoring function. Table 3.1 shows the multiple scoring functions and sample algorithms of the docking tools used in this work, and both concepts will be more broadly explained in their respective subsections.

Of course, MD is not a flawless technique and relies on simplifications, and, in this sense, MD is inferior to molecular dynamic simulations (MDS, discussed in a further section). Designed to account for all atomic interactions in a given system, MDS produces results that are arguably very close to real systems but may take more than a day to calculate protein-ligand interactions, even in modern computers with an advanced GPU. Because of this, MDS is rendered impractical for molecular screenings that can search multiple thousands of ligands. Most docking programs currently in use work through semi-flexible docking, that is, an approach in which ligands are treated as flexible structures while the protein conformation is maintained as rigid. Also, water molecules and ions are commonly excluded. These abridgments promote a massive decrease in docking experiments' complexity and computational cost. But, although MD is still far more reliable than the most popularly used methods for protein-ligand and protein-protein interaction evaluation, these two shortcuts are believed to be the most significant culprits in most cases in which docking fails to accurately predict binding.

Fortunately, the field of MD keeps developing at an ever-increasing pace, and many studies report advancements on an array of methods to correct these deficiencies; including flexibility trees, ensemble docking, and elastic potential grids, with multiple reviews on means to manage protein flexibility also being published. These are progressively being incorporated into MD software and screening pipelines. Also

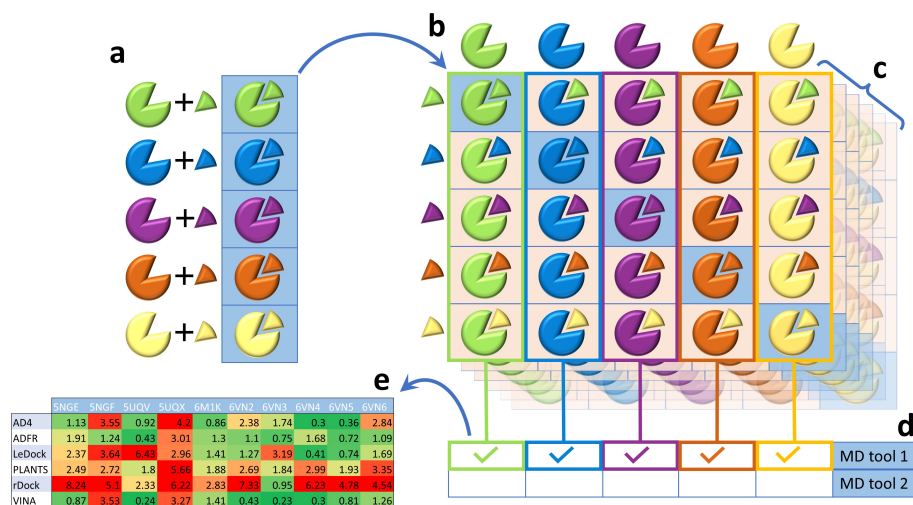


Figure 3.1: Re- and Cross-docking. Re-docking is displayed as a schematic in (a), and the cells with blue background in (b), a schematic of cross-docking. The procedure is often performed with multiple repetitions (c) to eliminate variability. After MD procedures are finished, the RMSD of the resulting poses is measured against the original ligands, and the group result for each crystal structure is aggregated as a mean (d). These scores can then be compared with those achieved by the other MD tools (e)

and multiple strategies have been developed to remedy such flaws. For instance, re- and cross-docking, exponential consensus screening, ensemble docking, and decoy enrichment are strategies that improve docking results.

3.1.1 Re-Docking and Cross-Docking

The processes of re- and cross-docking are commonly used to evaluate MD tools' accuracy and offer insight into whether a specific crystal structure is suitable for docking experiments. Re-docking consists of separating a protein and its cognate ligand from a coordinate file and performing docking of the two structures. Cross-docking works similarly, albeit involving the docking of multiple pre-aligned protein structures to all of their ligands (Figure3.1).

The acuity measurement of these proceedings is given by determining the root-mean-square deviation (RMSD) between the original pose of the co-crystallized ligand and the resulting docked pose, calculated by

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N \delta_i^2}$$

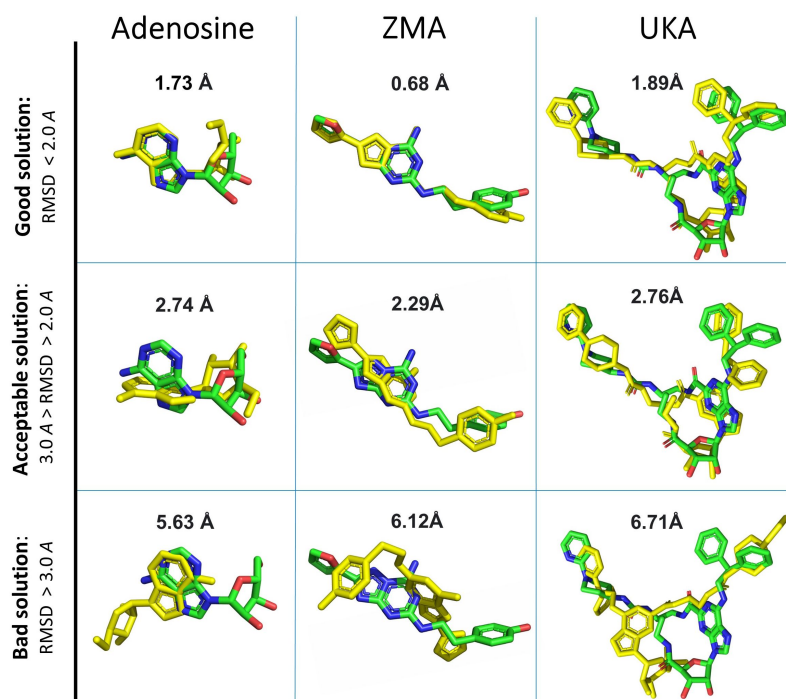


Figure 3.2: Visual examples for molecule RMSD. Various poses for the A2aR ligands Adenosine, ZMA and UKA are displayed in poses that resulted in good (under 2Å), acceptable (between 2 and 3Å) and poor (over 3Å) solutions.

in which δ_i represents the distance between atoms (in Å) of a given i pair, averaged over the N pairs contained in both molecules [112]. A lower RMSD indicates a pose closer to the reference, with values under 2Å being taken as good scores and values between 2Å and 3Å as acceptable, as shown in Figure 3.2. There are many software available that can calculate RMSD, but most cannot take into account ring flip on molecules presenting that display such characteristic [113,114].

3.1.2 Sampling Algorithms

The search through the conformational space is a complex task, as molecules not only present six degrees of freedom (three translational and three rotational) but also many of them display torsional potential through bond rotation [115]. In fact, performing an exhaustive exploration of any binding site, including all possible conformations of a single molecule, without a massive investment of time and computational resources. For this reason, efficient methodologies for conformational space sampling are of utmost importance in MD [116].

The first attempts to overcome this problem treated both ligands and proteins as rigid bodies, reducing computational cost by excluding bond rotation. This is a

simplistic model that represented both ligand and binding site as multiple pharmacophore spheres, with the search algorithm then trying multiple orientations and performing a least-squares fitting of atoms to the center of the spheres [117,118]. Although this approach surely reduces search time, tools employing this strategy also significantly reduce their capacity to find the optimal bind solution, which often causes a decrease in acuity [116].

Since the optimal binding conformation involves the interaction between ligand and protein, taking into account displacements in both structures, this flaw is even more pronounced in predictive docking. As isolated molecules often display conformations different from their bound form, this limitation can result in effective molecules suffering heavy scoring penalties [119].

Hence, most modern docking software work on a compromise: treating ligands as flexible and considering receptor flexibility only for the binding site or regarding the whole protein as rigid. The reasoning for this selection is that since ligands are smaller molecules, not only they are more likely to suffer changes in their conformation, but also, performing calculations for structures of this scale is far more viable. The conformations displayed by the ligand can be assessed by either stochastic or systematic methods [120, 121].

Stochastic methods rely on generating ligand conformations and orientations based on one or multiple values randomly generated every step. The algorithm then discards or adopts the change based on predefined criteria [122,123]. These techniques usually have a higher computational cost than systematic methods. Still, they have the advantage of generating a vast array of molecular conformation groupings, increasing the probability of locating the pose representing the global energy minim [122, 124]. Ant colony optimization, genetic algorithm, Monte Carlo, and tabu search algorithms are examples of this approach [125].

On the other hand, systematic methods are deterministic and execute sampling at predetermined intervals. They are subdivided into three subcategories: fragmentation, conformational ensemble, and exhaustive methods, the principal difference between them being how these handle ligand flexibility. Fragmentation methods divide the molecules into parts, one of these is attached to the protein, and the conformation grows incrementally from that point. Conformational ensemble methods pre-generate multiple ligand poses and perform rigid docking of each of them, and exhaustive methods rotate mobile bonds progressively at a given interval [126–128].

3.1.3 Scoring Functions

Scoring functions are responsible for MD tools' capability of gauging the physical interaction (binding affinity) between ligand and protein. Classical scoring functions are usually classified into three types: force field-based, knowledge-based, and empirical, with the latter being the most commonly used. Other than those, a few docking software also employ machine learning-based functions.

Empirical scoring functions are based on the idea that standard free energy change (ΔG^0) can be split by molecular interaction type and using the information generated by those groups to estimate binding affinity [129]. This can be better exemplified by a simplification of GOLD's Chemscore scoring function [130]:

$$\Delta G_{binding} = \Delta G_0 + \Delta G_{H-bond} + \Delta G_{metal} + \Delta G_{lipo} + \Delta G_{rot}H_{rot}$$

in which the binding free energy $\Delta G_{binding}$ is estimated through the contributions from non-specific molecular interactions ΔG_0 , hydrogen bonds ΔG_{H-bond} metal interactions ΔG_{metal} lipophilic interactions ΔG_{lipo} and rotatable bonds $\Delta G_{rot}H_{rot}$, with ΔG terms generated through multiple linear regression analysis over dozens of protein-ligand complexes. As such, this method is heavily reliant on the quality of the experiments that generate this training set.

Force field (also known as physics-based) methods focus on calculating Van der Waals and electrostatic energies through the molecular mechanics force field to assess binding energy. This method determines molecular interaction information through experimental means and quantum mechanics calculations. Also, every atom is handled as an indivisible particle. It is a known issue that this method presents two primary biases: that polar interactions may be treated with too much weight on the calculations and that they tend to fail to apply enough entropic penalties to large non-polar structures. The following equation represents these functions:

$$E_{bind} = \sum_i \sum_j \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{\varepsilon(r_{ij}) r_{ij}} \right)$$

where A_{ij} and B_{ij} represent van der Waals parameters for a given atom pair ij , d_{ij} the distance between these atoms, partial charges shown as q_i and q_j , and the distance-dependent dielectric function as $\varepsilon(r_{ij})$ [131].

Knowledge-based methods are fundamented on the large and continually growing databases of protein-ligands ligand complexes. In this model, atom pairings are evaluated through statistical analyses based on bond type frequency, employing the Boltzmann law to transform pairwise preferences into potentials translatable to atomic distancing. On docking, the binding score is given by adding the values of each atom pairing between protein and ligand, as illustrated by the following equation:

$$w_{ij}(r) = k_B T \ln \left(\frac{\rho_{ij}(r)}{\rho^*(r)} \right)$$

Where k_B represents the Boltzmann constant, T stands for the system's absolute temperature, $\rho_{ij}(r)$ is the number density of the group comprising a given protein-ligand atom pair at an r distance in the training set, and $\rho^*(r)$ represents the pair density in a stable condition where interatomic forces are at a zero-sum state. The major downside of this method is that interactions that are more rarely found may have poorer characterization [132,133]. These scoring functions offer a balanced approach concerning speed and accuracy when compared with force field and empirical scoring methods. The reason for this is that knowledge-based functions don't have to perform a range of computationally intensive analyses (ab initio calculations for force-field, binding affinities' reproduction for empirical methods) [134].

Machine-learning-based solutions comprise the most recent approach for scoring functions. Unlike the others, which are based on a classical functional form, these apply machine-learning algorithms, for instance, random forest, support vector, neural networks, and deep learning. Even though they usually outperform classical scoring functions, they are rarely incorporated into docking systems, the reason being that they are very dependent on their training datasets [135–137]

3.1.4 Exponential Consensus Scoring and decoys

Consensus scoring is a filtering technique based on combining two or more scoring functions on screening. The concept behind this is that focusing only on molecules with high binding scores can exclude false positives caused by particularities found in either. This approach has been employed in multiple MD screening campaigns and frequently promotes overall performance increases. [138, 139].

Standard consensus scoring is performed by arranging molecules by ranking the

docking scores generated by each MD tool and averaging the arrays generated and is often performed with ligand enrichment of the dataset [139,140].

Ligand enrichment is considered a key metric for the assessment of MD. The technique is based on including, along with the ligand test set, multiple decoy molecules and performing docking with all. These decoys are either known non-binders or structures similar to the ligands but with minor alterations, making them hard for MD tools to discriminate while having topological differences designed to decrease binding chance sharply. Usually, enrichment is at least ten times the number of the test molecules, with a 10% general cut-off based on docking performance [4,16,141,142].

Furthermore, exponential distribution has been combined with traditional consensus systems to great effect. This takes advantage of probability density function characteristics to maximize the difference between good and poor scores [143–145]. This method is better exemplified by the following:

$$P(i) = \sum_j p(r_i^j) = \frac{1}{\sigma_{x\%}^j} \sum_j \exp\left(-\frac{r_i^j}{\sigma_{x\%}^j}\right)$$

Where the rank of a given i molecule is determined by the sum of scores produced after processing the rank r_i^j given by each j MD software. In this system, $\sigma_{x\%}^j$ is an adjustable parameter that can be set independently for each docking software to filter $x\%$ of the molecules [143,146].

3.2 Molecular Dynamics Simulation

Simply put, MDS's foundation is that given an atom system with known positions and types, it is possible to calculate the force exerted on each of them by all other atoms. The atomic interactions are then used to determine the trajectories of all molecules, and the procedure is repeated through multiple steps, each on the order of femtoseconds. This process generates a three-dimensional simulation of the system, detailed to the atomic level and spanning through all the simulated time.

This generates a large amount of data that can be easily extracted and is helpful for research in multiple ways. In the realm of biochemistry, two of these come with heightened interest. First, capturing the coordinates and velocities of every atom at every point in time would be extremely difficult with any experimental technique,

if not outright impossible. And second, all simulated parameters can be perfectly controlled and all conditions known, for instance, the original protein conformation, its exact sequence, if any mutation is present, if there are any ligands bound, what molecules can be found in the environment, protonation states, pressure, temperature, and so forth. This specificity and this richness of information permits performing various similar assays under slightly different conditions, which can help perceive the influence of a vast array of selected variables that would otherwise be unfeasible.

For the calculation of forces present in MDSs, molecular mechanics force fields are used. These are tuned explicitly for quantum mechanics calculations and experimental quantification, with a typical force field being exemplified by the following expression:

$$U = \sum_{intra} + \sum_{RVw}$$

where the first term is composed by:

$$\sum_{intra} = \sum_{bonds} \frac{1}{2} k_b (r - r_0)^2 + \sum_{angles} \frac{1}{2} k_a (\Theta - \Theta_0) + \sum_{torsions} \frac{V_n}{2} [1 + \cos(n\Phi - \delta)] + \sum_{improper} V_{imp}$$

representing intramolecular and local contributions of bond stretching, angle bending, and dihedral and improper torsion; and the second term is formed by:

$$\sum_{RVw} = \sum_{LJ} 4\epsilon_{ij} \left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^6}{r_{ij}^6} \right) + \sum_{elec} \frac{q_i q_j}{r_{ij}}$$

Where terms refer to repulsive and Van der Waals forces (represented by 12-6 Lennard-Jones potential) and Coulombic interactions [147, 148]. It's also noteworthy that comparisons between multiple simulations using different force fields show that these have significantly improved in the last decades [149, 150], although some imperfections persist. Such is the case of covalent bond development, as MDS still cannot reproduce their formation or breakage, with covalent bonds observed at the beginning of the simulation remaining unaltered until the end [151, 152]. Therefore, such imperfections must be considered when designing and interpreting an exper-

iment. For instance, when the need to observe covalent bond changes is present in an experiment, quantum mechanics/molecular mechanics simulations must be considered [153].

The duration of a time step is also pivotal for numeric stability in an experiment and usually is around a few femtoseconds each. Since the biochemical interactions usually observed develop in nanoseconds, microseconds, or longer, this kind of simulation usually contains millions or billions of steps. And considering the millions of interatomic interactions involved, this causes these MDSs analyses to be very computationally intensive. As such, the technological development observed in computer hardware, as well as overall software and specific MDS algorithms, has been pivotal for the accessibility and increase in the practicality of MDS simulations [154]. Among these, advancements in GPU technology have been especially remarkable for the field, making it possible for a personal computer to perform MDSs faster than what was previously achieved by supercomputers, causing an expressive increase in MDS experiments' utility and viability [155].

3.3 State of the Art

Docking-Based Molecular Screening (MS) is a well-defined and prominent technique based on the generation of affinity scores from structural information [6, 156]. With multiple successful screening campaigns found throughout the literature and the promise of a low-cost process with quick and high reward, the strategy has become very popular in the pharmaceutical industry, fostering much improvement and knowledge in the field [157–160].

The first virtual screening based on tridimensional chemical structures took place in 1990; it targeted the dopamine D2 receptor and was able to identify an agonist with a pKi of 6.8. After that, multiple drugs, such as Gefitinib [161], Amprenavir [162], Epalrestat [163], and others, have been discovered through structure-Based MS [164]. And with the introduction of large public compound databases and the evolution of computer hardware, MS has become ever so broadly available, and MD tools are developed and improved with increasing speed and frequency [165, 166].

The technique is not without its complexities and flaws; for instance, frequently using a single MD tool is not enough to warrant enough specificity, and the analysis of binding energies and docking poses may provide conflicting information in some cases [167, 168]. Still, solutions and discussions on these topics abound, and MS has

a long and successful history, making it a prime choice for small molecule analysis [169–171].

MS has evolved into a multivariate approach, encompassing chemometrics analysis, and is grounded on the use of search algorithms and scoring functions (further discussed in the overview chapter). Recently, artificial intelligence has been applied to allow massive searches of the chemical space, encompassing tens of billions of molecules [164, 169, 172].

Proteins Case Studies

This study focused on four human proteins with known pharmacological potential and which currently have substantial research importance described in the following sections. Our test set was divided into two G protein-coupled receptors (GPCRs) Adenosine 2a Receptor (A2aR) and Kappa Opioid Receptor (KOR) and two enzymes Janus Kinase 2 (JAK2) and Ubiquitin-Specific Protease 7(USP7).

4.1 G protein-coupled receptors

GPCRs are a large family of proteins comprising hundreds of different representatives, present in every known animal species. The vast majority of signaling pathways pertaining to physiological functions, and many pathologies, present mediation by GPCRs. For this reason, roughly a third of all clinically approved drugs target this class of receptors.

The proteins of this family present a highly preserved structure common to all of its constituents, with seven transmembrane α -helices linked in line starting from an extracellular N-terminus, forming a channel with three linking loops in the extracellular medium and the remaining three, alongside with the C-terminus- in the intracellular. As its name suggests, these receptors are also found coupled by its C-terminus to a G protein, which fractures in half when an agonist ligand promotes their activation. This split G-protein is then responsible for the next step of signal conduction, as shown in Figure4.1 [173–175].

The A2aR is part of the P1 class of purinergic receptors and the GPCRs superfamily, possessing seven transmembrane alpha-helices, an extracellular N terminus, and an intracellular C-terminus, and capable of transmembrane signaling. The adenosine receptors group has four subdivisions, A1, A2A, A2B, and A3, all of which have adenosine as their primary ligand and mediate their activation. They have

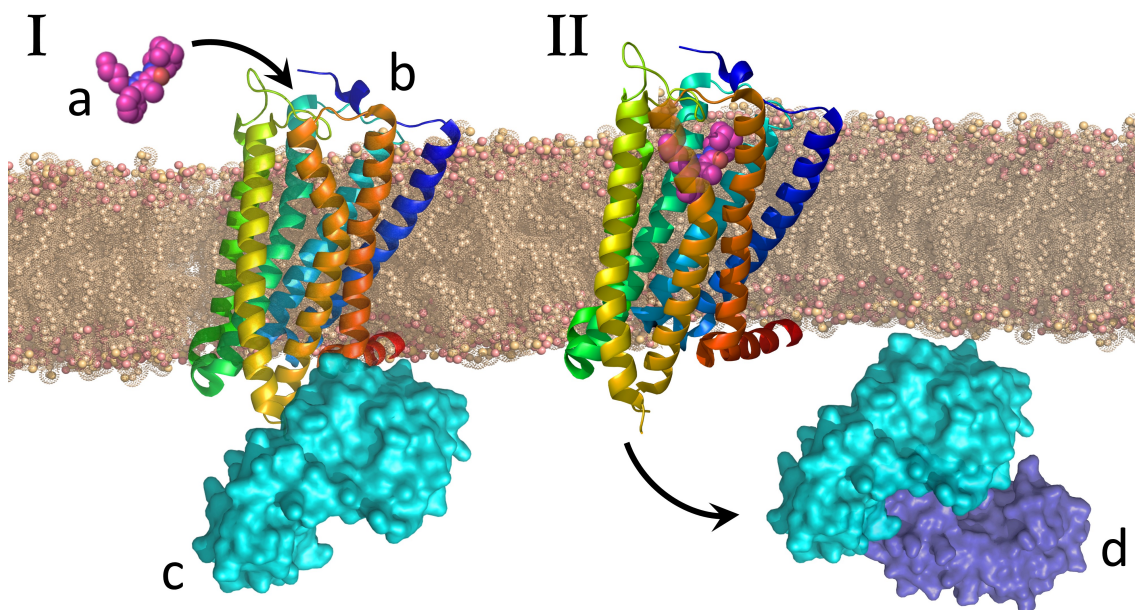


Figure 4.1: GPCR structure and activation. The activation of GPCRS (b) by coupling with their agonist ligands (a) results in conformational changes that activate their coupled G-protein (c), which is responsible for modulatory action and signal transduction (d) [176].

a crucial role in mediating multiple fundamental physiological processes such as immunomodulation, cardiac regulation, angiogenesis, vasocompression, and sleep and mood adjustment. This makes adenosine receptors interesting targets for the regulation of multiple ailments, namely arrhythmias, cardiac hypertrophy, type 2 diabetes, depression and other neurological disorders, asthma, cancer, inflammatory disorders, Parkinson's disease, pain, renal failure, and glaucoma.

The production of accurate models of GPCRs is a well-known problem since their poor stability and variable conformation, but advances in technologies and strategies such as fusion proteins and conformational thermostabilization have generated good results in analyzing such receptors. In the last decade, the A2aR has been placed amongst the most well-characterized GPCRs, as molecular studies of this receptor resulted in the publication of over 30 crystal structures of the human version of the receptor. Also, the recent solving of A1R structure has been very important for structural comparisons for docking sites and understanding ligand specificity between the multiple adenosine receptors.

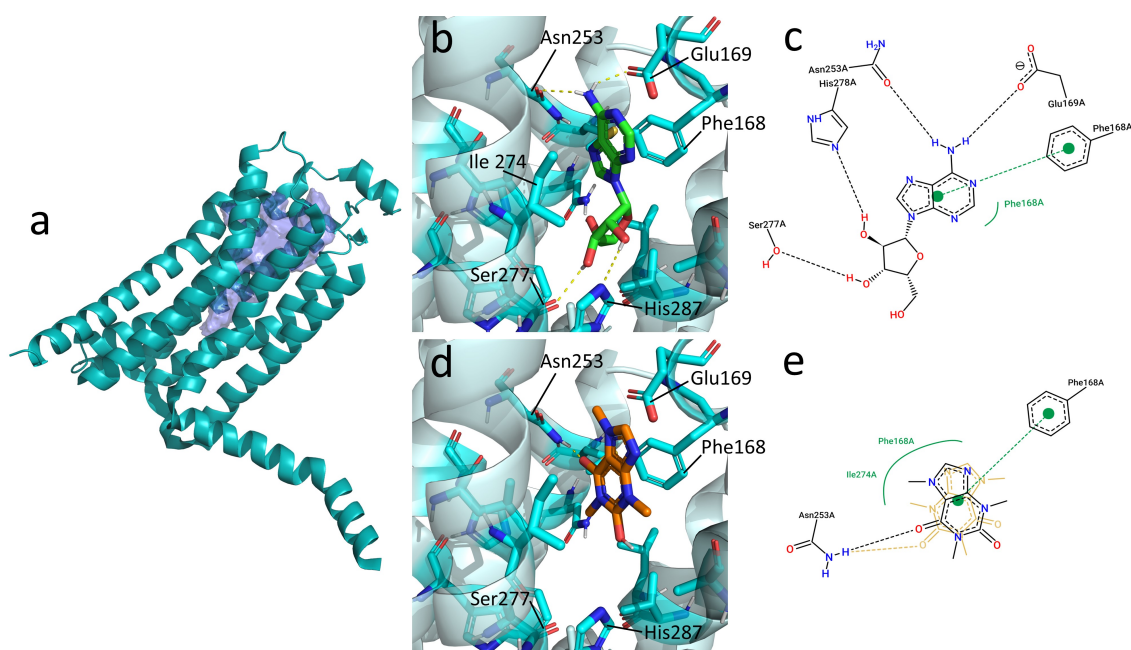


Figure 4.2: A2aR structure and binding site. (a) shows A2aR's binding site in dark blue, (b) the tridimensional representation of adenosine bound to A2aR's binding site, and (c) its 2D representation showing residue interactions. A2aR-caffeine interactions are displayed in 3D in (d) and in 2D with residue interactions in (e).

4.1.1 Adenosine 2a Receptor

The A2aR primary binding site, as shown in Figure 4.2, is determined by the residues associated with interaction with adenosine, its primary agonist, and caffeine, a naturally occurring antagonist. These bind themselves by creating van der Waals interactions with M177, M270 and I274, and π -stacking with F168, precisely by its aromatic ring. Caffeine forms hydrogen bonds with O11 or O13, while adenosine's hydrogen bonds are held with N253. Adenosine also produces Van der Waals interactions with V84, L85, T88, W246, and L249, which interact similarly with various antagonists.

As previously stated, GPCRs can display various conformational states, divided by energy boundaries, with active and inactive states being highly similar among representatives of this receptor family; the interaction between an agonist and its binding site being the fundamental event to surpass said barrier and activate the receptor. Furthermore, it's observed that sodium ions promote negative modulation for multiple GPCRs, and this can be observed in the A2aR by interactions mediated by water molecules with the residues D52, S91, T88, W246, N280, and S281, and

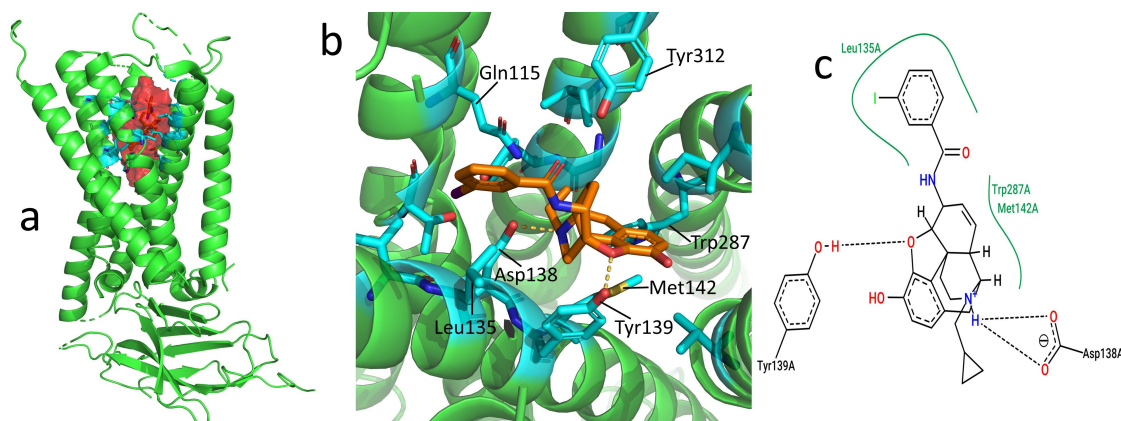


Figure 4.3: KOR structure and binding site. (a) shows KOR's structure (green) with its binding site (red volume). (b) is a representation of the binding site on a frontal view (from outside of the cell) with the protein (green), binding site residues (cyan) and its ligand (CVV, in orange). In (c), a 2d representation of the primary interactions between CVV and the binding site is provided.

directly with D52 and S91.

4.1.2 Kappa Opioid Receptor

The κ OR is one of five known opioid receptors, namely, κ , δ , μ , ζ and nociceptin opioid receptors (κ OR, δ OR, μ OR, ζ OR and NOR), their primary agonists being endogenous peptides (endomorphins, endorphins, enkephalins and dynorphins) [177, 178]. Although the structure of this receptor isn't yet perfectly understood, *kappa*OR's primary binding site is defined by its interactions with dynorphin, its primary agonist, and other strong agonists such as MP1104 (Figure 4.3 [179–181]).

All known opioid receptors promote pain modulation, and opioids, in general, are capable of promoting analgesia without loss of touch, proprioception, or consciousness as side effects. [177, 182] For this reason, these molecules have been highly sought after throughout history and are still extremely important in medical applications [183]. With opioids being widely prescribed worldwide to treat acute and chronic pain [184, 185].

Controversially, many opioids are highly addictive, and from the half-million worldwide deaths attributed to drug abuse, more than 70% are related to opioids [186, 187]. In part, illegal drugs such as fentanyl, heroin and desomorphine, but overprescription also contributes to the problem [188–190]. The issue arises as μ OR and molecules that primarily target this receptor are far more studied [191], but the targeting of

this receptor is linked to not only itch, respiratory depression, and constipation but, more importantly, opioid tolerance and addiction [191, 192].

In this regard, κ OR presents advantageous qualities: as just as μ OR, it is a ubiquitous nociceptive transmission modulator, but its signaling is also connected to reducing pruritis, stress, depression, anxiety, and substance use disorder [191–195]. Compared to other opioid receptors, signaling by *kappa*OR's main agonist has a stronger correlation with negative affect and the stress response generated by drug use [191, 196], making the receptor an attractive target for the research of novel pain modulators with limited side effects.

Also, these signaling pathways can attract attention for the opposite reasons, as κ OR novel antagonists have been linked to the reduction of substance abuse and attenuation of withdrawal effects of alcohol and cocaine [197, 198].

4.2 Enzymes

4.2.1 Janus Kinase 2

JAK2 is a member of the Janus Kinase (JAK) family of non-receptor tyrosine kinases. These proteins associate with cytokines receptors and are fundamental for signal transduction between the extracellular medium to the nucleus. This communication chain ends with the production and activation of the signal transducer and activator of transcription proteins (STATs), being dubbed the JAK/STAT pathway (Figure 4.4), and is related to cell growth and proliferation. This signaling chain is particularly important in controlling blood cell production, and disruptions of it are related to Crohn's disease, essential thrombocythemia, polycythemia vera, primary myelofibrosis, and myeloid disorders and malignancies.

Structurally, all members of the JAK family are composed of seven distinct domains: JAK homology domains 1 to 7 (JH1-7). For the purpose of studying JAK2 inhibitors, the groups which carry the most significance are JH1, and JH2 (tyrosine kinase and pseudokinase domains, respectively) [199–201].

JH1 is the main promoter of the phosphoryl transfer action of JAK family's enzymes and may even be the sole direct contributor. Notably, it's responsible for transphosphorylation of JAK2's Tyr1007 and Tyr1008, present in the kinase activation loop (Figure 4.5), stabilizing the enzyme's active state. Once it's active, JH1 also phosphorylates specific tyrosine residues present in its correlated cytokine receptor,

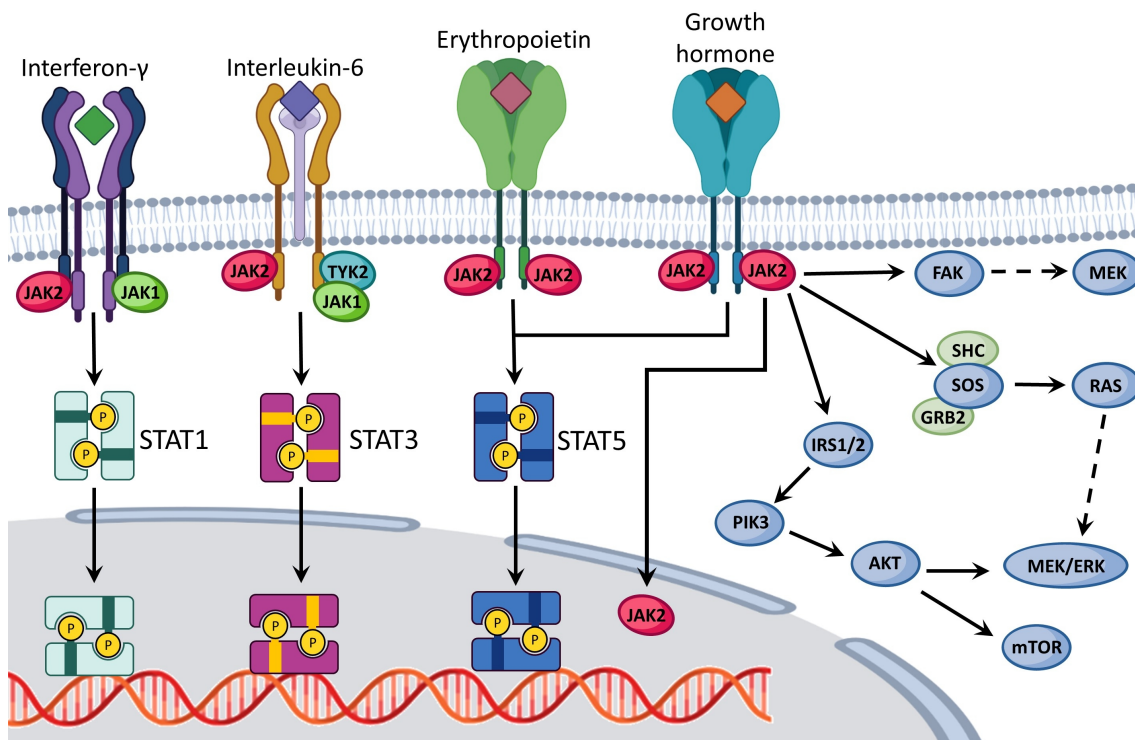


Figure 4.4: OR structure and binding site. (a) shows KOR's binding site in red, (b) N-[(5 α ,6 β)-17-(cyclopropylmethyl)-3-hydroxy-7,8-didehydro-4,5-epoxymorphinan-6-yl]-3-iodobenzamide bound to KOR's binding site, and (c) is the 2D representation showing residue interactions.

STATs, and in JAK molecules [199, 202–204]. JH2 is a negative regulator for JAK enzymes, with JH2 deletion resulting in STATs activation in the absence of ligands for the correlated interferon receptors. This modulation is promoted by the interaction between JH2's N lobe and JH1's N lobe and hinge region [199, 201–203, 205].

The surface of the cleft between these two domains is not only the targets of most JAK2 inhibitors; as it contains the ATP, activation loop, and helix C pockets, as shown in Figure 4.5; but also, the interface between both of them is the site for most known activating mutations of this enzyme (e.g., C618R, H606Q, H608Y, R683G, and V617F) [206–208].

4.2.2 Ubiquitin-Specific Protease 7

USP7 (also known as HAUSP) is a protease that possesses a fundamental role in the stability of proteins involved in DNA regulation and repair, epigenetics, tumor suppression, immune response, viral replication, and multiple signaling pathways.

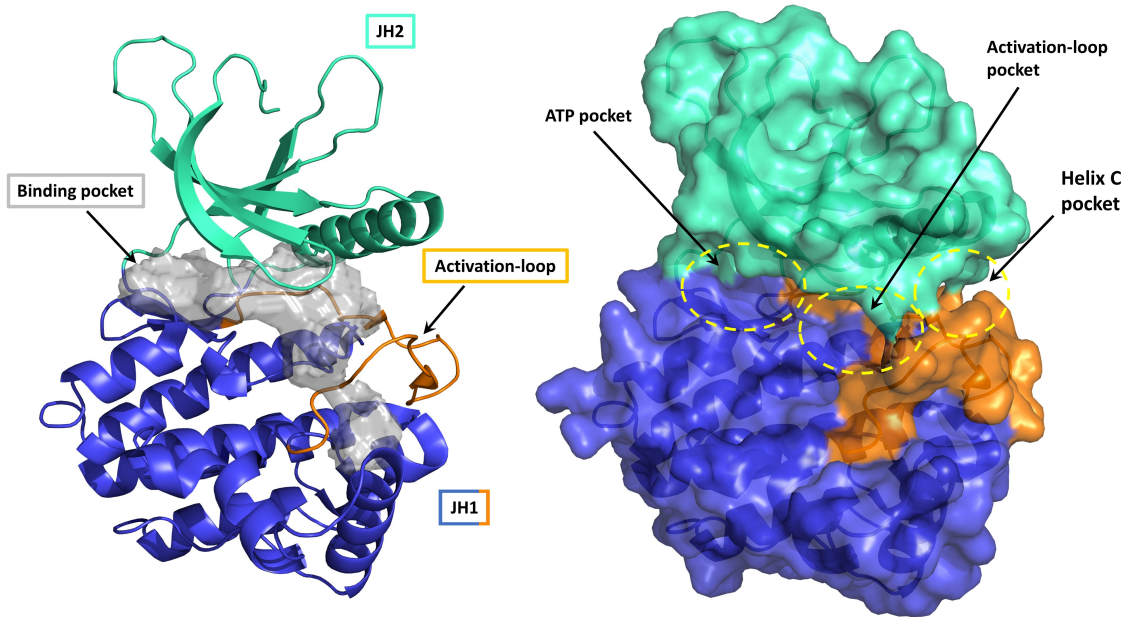


Figure 4.5: JAK2 structure and binding pockets. The homology domain JH1 is presented in teal, and JH2 in blue, with the activation loop, highlighted in orange. In (a), a detailed view of the enzyme structure is presented, with the binding pocket used in this work displayed in grey. In (b), the protein surface is shown, and the ATP, Helix C, and Activation-loop pocket is indicated.

This extensive role gives this enzyme great importance over many cellular processes, from normal metabolism to pathology progression. Table 4.1 shows a list of putative binding partners and substrates linked to USP7 for more straightforward observation.

Its function is that of promoting protein post-translational modifications, more specifically, the deubiquitination of its targets. As ubiquitin acts as a biomarker for, amongst others, protein activation, mobilization to different cellular locations, processing, and degradation, USP7 action can display a multitude of outcomes fundamental to cellular regulation [232–234].

Recently, USP7 was pushed into the spotlight as it's been linked to a seemingly paradoxical role in p53, MDM2, and MDMX regulation. P53 acts as a tumor suppressor, halting cell multiplication by binding to regulatory sites of the DNA or even through initiating apoptosis, whereas MDM2 and MDMX regulate p53 cellular concentration. As cancer cells frequently overcome p53 action by overexpressing MDM2, USP7 inhibition has shown promising results as a cancer treatment [235–237].

Table 4.1: Examples of substrates of USP7 related to carcinomas.

USP7 substrates	Related cancer	reference
Beclin-1	Multiplemyeloma	[209]
CCDC6	Lungneuroendocrinecancer	[210]
CCDC6	Hormone-sensitiveprostatecancer	[211]
CCDC6	Non-smallcelllungcancer	[212]
CHFR	Multipleolidtumors	[213]
CHK1	Acutemyeloidleukemia	[214]
CHK1	Breastcancer	[215]
ER- α	Breastcancer	[216]
Foxp3	Non-smallcelllungcancer	[217]
HDM2	Multiplemyeloma	[218]
HDM2	Coloncancer	[219]
HDM2	Neuroblastoma	[220]
hnRNPA1	Gastriccancer	[221]
MDC1	Cervicalcancer	[222]
MDM2	Gliomas	[223]
MDMX	Multiplemyeloma	[218]
NEK2	Multiplemyeloma	[224]
PHF8	Breastcancer	[225]
PIM2	Multipletypes	[226]
PLK1	Lungcancer	[227]
PTEN	HER2-positivebreastcarcinomas	[228]
RAD18	Chroniclymphocyticleukemia	[229]
TRIM27	Cervicalcancer	[230]
β -catenin	Colorectalcancer	[231]

This enzyme presents a multi-domain architecture, likely to contribute to the complex role it exerts (Figure4.6). It possesses an N-terminal tumor necrosis factor receptor-associated factor related to the binding of its substrate peptides and five ubiquitin-like C-terminal domains, with compounds with USP7 inhibition capabilities typically binding to its active site and impeding ubiquitin coupling [238, 239].

Although many molecules can inhibit USP7, most of these either are unspecific, promote toxicity or bind covalently to their targets, making them poor choices for pharmaceutical applications [235, 238, 240].

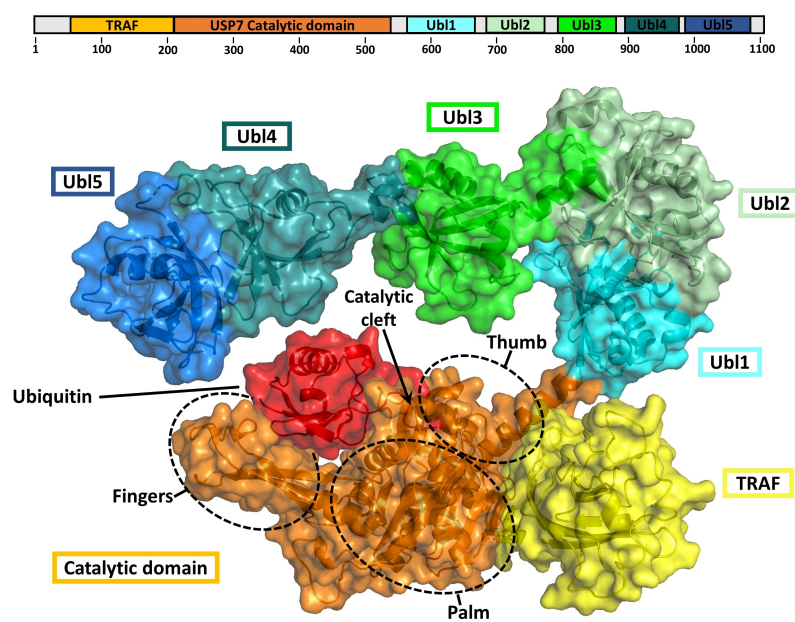


Figure 4.6: USP7 structure. This image was generated by aligning pdb structures 2f1z, 4z97 and 5fwi. Ubl 1 to 5 are represented in hues of green and blue, TRAF is shown in yellow, the catalytic domain (containing the catalytic cleft, palm, thumb and fingers) in orange and an ubiquitin molecule in docked position is appears in red.

Materials and Methods

This chapter is separated into three main parts. In the first section, the general proposed workflow has been illustrated. It is divided into three main portions: I-Docking tools and crystal structures assessment, II-Consensus screening, and III-Screening validation, delineating step by step the processes observed in this study. The second is focused on the datasets used in the production of this work; it contains information on the selected protein crystal structures, selected ligands, and molecules produced by DGM. The third part provides a quick overview of the tools employed and the parameters used for the experiments. The final portion of this section expounds on the tools used for molecular screening and validation.

5.1 The Proposed Workflow

For the MD tool assessment process (Figure 5.1:I), two input datasets are used: the protein crystallographic structures from the PDB database and ligands with a known measure of affinity (pIC₅₀, in our case) for the selected proteins that are then taken from ChEMBL catalogs(Figure 5.1 1a), both sets are prepared as previously discussed.

For the docking experiments, we selected software known for performing well according to the literature and that are also free and accessible to the public; those were: AutoDock 4, AutoDock FR, AutoDock Vina, LeDock, PLANTS, and rDock (Figure 5.1:1b). Cross-docking experiments were performed for each of the four sets of protein structures (Figure 5.1:1c) to determine the crystal structures best suited for the docking of various ligands; these results also produced information on how capable of reproducing experimental data these software are.

The proteins selected with the best overall ligand RMSE scores for each set are used in docking experiments with the ligands with determined experimental intermolec-

5. Materials and Methods

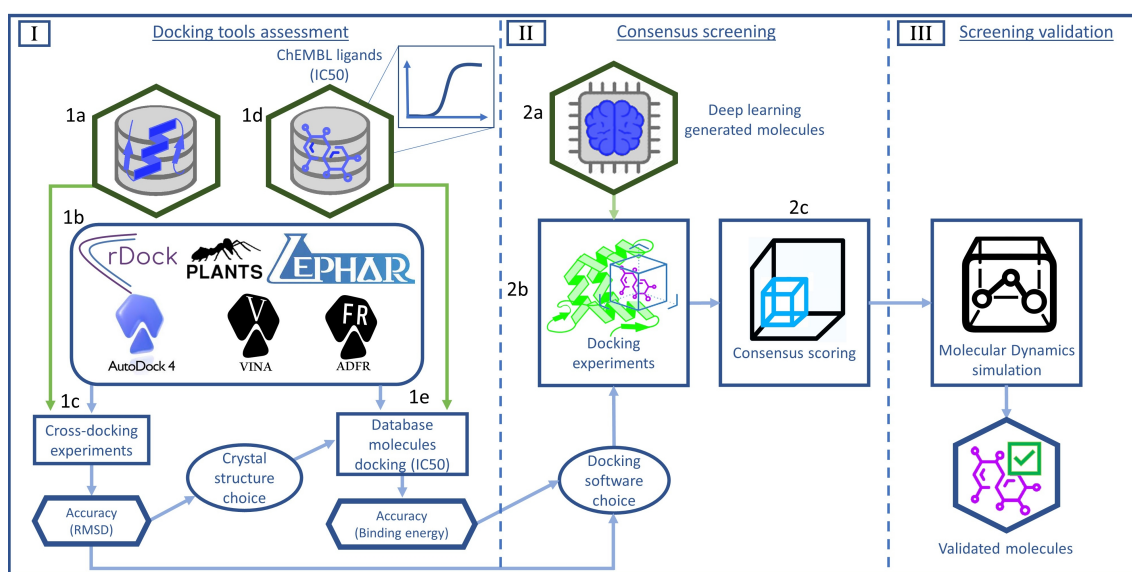


Figure 5.1: General workflow. The diagram shows an overview of the proposed methodology, with block I standing for the MD tools selection process, II demonstrates the screening of deep learning generated molecules and III shows the validation of the selected molecules through MDS.

ular energy. These results are plotted on linear regression models using Spearman's and Pearson's correlations between docking scores and experimental affinities to determine the fidelity of the results for each software (Figure 5.1:1d). The molecules found outside the confidence interval for these regressions are classified as under or over-scored. Molecular fragments based on these molecules are generated by the Monte Carlo technique using RDKit and tabled alongside the original molecule and its classification. This data is then used to train a Random Forest model to determine a correlation between docking bias and chemical and structural properties of the molecules. This information is then used alongside cross-docking data to determine the best choices of docking software to proceed.

After this step, molecule screening by MD (Figure 5.1:II) starts. The set of ligands generated by deep learning techniques (Figure 5.1:2a) is then used in docking experiments to assess their affinity for their according proteins (Figure 5.1:2b), and the docking scores are processed in a consensus score protocol (Figure 5.1:2c).

As a last validation step, the molecules selected in consensus scoring as having the highest affinity in each protein set were assessed by MDS using the GROMACS software (Figure 5.1:III).

5.2 Datasets

5.2.1 Protein structures and preparation

The present work focuses on four proteins found recently to be linked to important conditions and that are currently highly sought after in pharmacologic research: A2aR, JAK2, KOR, and USP7.

Binding pocket information was generated by using DoGSiteScorer [241].

All the crystal structures used in this work and their cognate ligands can be found in Table 5.1. All files were retrieved from the protein data bank website [242].

Table 5.1: List of the PDB structures used and cognate ligands.

Protein	PDB code	Ligand	Reference
A2aR	2YDO	Adenosine	[243]
A2aR	3EML	ZM241385	[244]
A2aR	3PWH	ZM241385	[245]
A2aR	4UG2	CGS21680	[246]
A2aR	4UHR	CGS21680	[246]
A2aR	5G53	NEC	[247]
A2aR	5IU4	ZM241385	[248]
A2aR	5IU8	ZM241385	[248]
A2aR	5UIG	8D1	[249]
A2aR	6AQF	ZM241385	[243]
JAK2	6VGL	Ruxolitinib	[250]
JAK2	6VN8	Barictinib	[250]
JAK2	6VNE	2TA	[250]
JAK2	6VNJ	PN4-014	[250]
JAK2	6VNK	Ruxolitinib	[250]
JAK2	6VNL	5W2	[250]
JAK2	6VS3	BL2-057	[250]
JAK2	7Q7I	9I8	[251]
JAK2	7Q7K	9I5	[251]
JAK2	7Q7L	9I2	[251]
KOR	4DJH	JD _{Tic}	[252]
KOR	6B73	CVV	[179]
KOR	6VI4	JDC	[253]
USP7	5NGE	FT671	[238]
USP7	5NGF	FT827	[238]
USP7	5UQV	GNE6640	[226]
USP7	5UQX	GNE6776	[226]
USP7	6M1K	EZF	[254]
USP7	6VN2	R44	[255]
USP7	6VN3	R3Y	[255]
USP7	6VN4	R4D	[255]
USP7	6VN5	R4I	[255]
USP7	6VN6	R4J	[255]

5.2.2 Ligands with known binding affinity

To test the docking software’s scoring acuity, four sets of 50 molecules with known pIC₅₀ or pK_I were downloaded from the ChEMBL database [256], one for each of

the proteins used in this work; all of which are shown in Appendix A.1.

5.2.3 Deep learning generated molecules

The generation of compounds using deep learning methodologies was performed following the procedures implemented in the work of Pereira et al. [257]. In practice, it was trained a model that generates molecules with bespoke properties based on neural networks containing recurrent architectures, namely Long Short-Term Memory layers. This model uses SMILES notation to represent molecules so that the chemical information is perceptible by the computational algorithms. In the first training phase of the Generator, we used a general dataset taken from ChEMBL (chembl identifier 22) so that the model could learn the basic rules of SMILES notation to encode valid and synthesizable molecules. Afterwards, the Generator was re-trained through Reinforcement Learning (RL) to generate compounds with desired biological properties. In this case, it was intended to generate compounds that would bind to Adenosine A2A and USP7 targets. To implement this dynamic, it was necessary to build and train regressor models that assess the ability of molecules to interact with the respective targets through supervised learning. Thus, the RL process was guided by these evaluators, forcing the Generator to explore the regions of the chemical space that favoured the desired interactions. After applying this deep molecular generation framework for the two tasks, the biased generators were used to sample the sets of new molecules with a high probability of inhibiting the Adenosine A2A and USP7 targets.

5.3 Bioinformatic tools

5.3.1 Docking software and structure preparation

Autodock 4 (AD4) [258], Autodock FR (ADFR) [259], Autodock Vina (VINA) [260], LeDock [261], PLANTS [262]

and rDock [263] were used in these experiments.

The search space was defined as a 20x20x20 Å ($v = 8,000 \text{ \AA}^3$) box to encompass all the binding pocket residues. Re-docking experiments were performed in sets of 10 as this sample size has previously produced accurate models. Using Discovery Studio Visualizer, polar hydrogens were added, and all water molecules, ions, and ligands were removed. Cognate ligands were preserved in a separate file for cross-docking.

For AD4, ADFR, Vina, and LeDock, the protein, and ligand files were prepared using AutoDockTools 1.5.6., converted from PDB to PDBQT format, and Gasteiger charges and hydrogens were added. For AD4, grid points were set to 60 and spacing to 0,375 Å. The parameters for the Lamarckian genetic algorithm and the scoring function were left as default. For ADFR, Vina and LeDock, exhaustiveness was set to 8, and additional parameters were left as default.

For plants, proteins and ligands were prepared using spores, and input files were used without alteration. Exhaustiveness was set to 8, and additional parameters were left as default. And for rDock, Maestro was used for processing the used structures. The geometric center of the original ligand was set as the spheres for the two spheres method (radius set as 3.0 Å and 12.0 Å, respectively), defining the cavity.

5.3.2 Fragment analysis, consensus scoring, and drug-likeness prediction

Fragment analysis was performed using Python 3.10's Scikit-learn library to build a Random Forest classifier model and trained on the Morgan fingerprint vectors generated by RDkit 2021.03.1 for each ligand in the data, according to their respective receptors.

Consensus scoring was performed according to the Exponential Consensus Ranking [143] for the results of the deep learning molecules' screening of all the MD tools used. In addition, Screen Explorer [264] was used to generate a ROC curve to evaluate the strategy's acuity.

The online version of the Molsoft software [265] was used to assess the Drug-likeness score of the putative ligand molecules.

5.3.3 Molecular dynamics simulation

GROMACS v2021.4 [266] was used to run MDS for ligands selected after cross-docking at a timeframe of 50 nanoseconds to assess the stability of the docked ligand-protein complexes. These were solvated in a cubical box, respecting a minimum distance of 10 Å from the edges to the ligands. The GROMOS96 53a6 force field was used to prepare these ligand-protein complexes, and Na⁺/Cl⁻ ions were added for charge balancing and neutralizing the system. The system's initial energy reduction was carried out using the steepest descent algorithm (5000 steps). The whole system

was equilibrated for five nanoseconds at 300 K and 1 bar pressure using NVT and NPT ensembles after the first minimization was finished. A Berendsen thermostat was selected, and a temperature coupling of 300K was selected. Parrinello-Rahman was used to set a 1.0 bar pressure and Lennard-Jones and Coulomb interactions for periodic boundary conditions, and PME was used to handle long-range interactions. The chosen time step was 2 femtoseconds, storing coordinates every 10fs, and the final MDS ran for 50 ns. Furthermore, we used Visual Molecular Dynamics 1.9.3 for the calculation of protein-ligand distance and of RMSD throughout the simulation [267].

Results and Discussion

This chapter contains the experimental analysis and results produced throughout this work. It comprises three parts: docking tools assessment, model-generated molecules screening, and validation through molecular dynamics simulations.

6.1 Docking Tools Assessment

Firstly, the six docking tools used were evaluated according to their performance to reproduce the native poses of the ligands found bound to protein's crystal structures used in this work. For this purpose, we used PyMol to align the protein structures and DoGSiteScorer to determine their binding sites, as presented in figure 6.1.

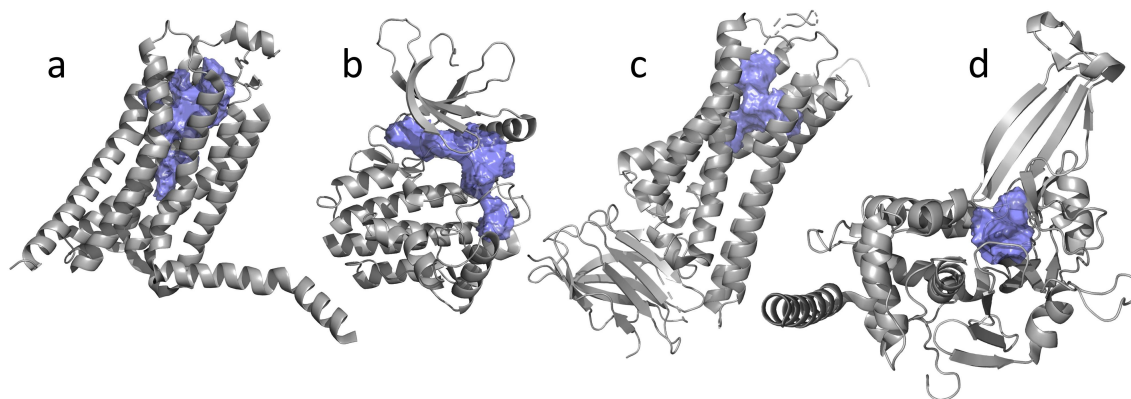


Figure 6.1: Binding pockets. DoGSiteScorer generated binding pockets for A2aR(a) JAK2(b), KOR(c) and USP7(d)

Each software's docking boxes were positioned according to the geometric centre of the binding pocket, and each ligand was docked to every protein structure. At the moment of these experiments, there are only three KOR crystal structures available; so both of the co-crystallized proteins in each file were used. For all other proteins, ten of each kind were selected. The results were gauged by the docked ligand's

RMSD (\AA) against the original pose, with the cutoff mark set as 2\AA . The MD tools expressed varying degrees of success in this task, and the mean RMSD values for the poses generated by each software are displayed in the heatmaps of Figure 6.2, boxplots of these values are shown in Figure 6.3 for ease of visualization.

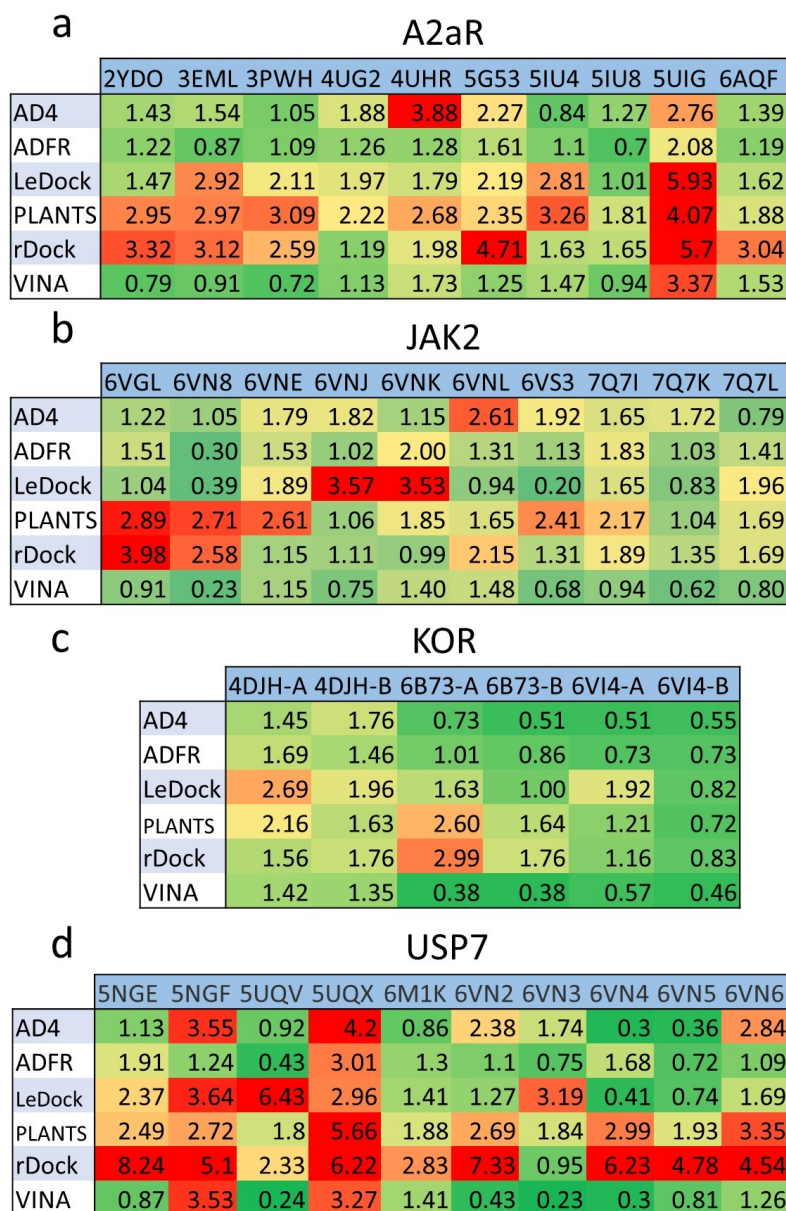


Figure 6.2: RMSD heatmaps. Heatmaps expressing the RMSD scores for the cross-docked ligands of the four proteins. (a) represents A2aR, (b) JAK2's, (c) KOR and (d) USP7 structures.

The tools from the Autodock family consistently presented good results, with RMSDs under 2\AA . ADFR was able to find good solutions 91.7% of the time and mean RMSD of 1.26\AA , VINA 86.1% and 1.05\AA , and AD4 75% and 1.36\AA . LeDock produced inter-

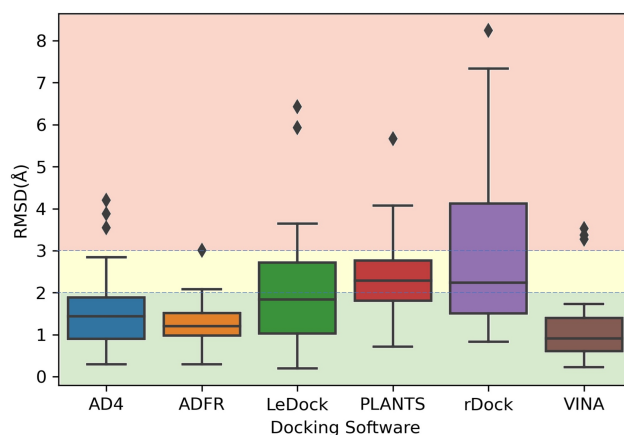


Figure 6.3: RMSD results presented in the form of box plots.

mediate results (58,3% under 2Å and mean RMSD of 2.06; with the worst performers being PLANTS (38.9% and 2.25Å) and rDock(27.8% and 3.06Å), already a sign that these two tools are likely, not adequate for screening campaigns.

The primary analysis results also provided information about the best structures crystal to be used in this experiment. The following selection was made: 5IU8 for A2aR (mean RMSD of 1.23Å), 4DJH-A for KOR (0.65Å), 6VNJ for (0.72Å), and 6VN3 for USP7 (1.45Å).

In sequence, we performed docking of four sets of 50 ligands with known binding affinity, one for each protein, with every MD tool. This is a more nuanced evaluation of their competencies, and the specific capability of providing accurate binding affinities can be of great use in assessing putative ligands. We plotted the docking results against the specific affinity values and performed linear regressions to verify the accuracy of each tool. The three best results and global outcome for each protein can be seen in Figure 6.4 (complete analysis in Appendix Figures A.1 to A.4).

The top two performers for the first experiment, VINA, and ADFR, displayed higher accuracy, with correlations statistically significant. Interestingly, despite being able to position ligands on re-docking adequately, AD4 was not as precise for evaluating intermolecular forces, presenting results close to random chance. An outcome like this is not wholly unexpected, since the accurate estimation of binding energy is a far more stringent test for an MD tool than ligand posing. Still, the software may be focused on accurately posing ligands or present specificities that negatively impact accurate energy estimations.

VINA and ADFR were the best performers in both tests, so these two software were selected for the ensuing screenings.

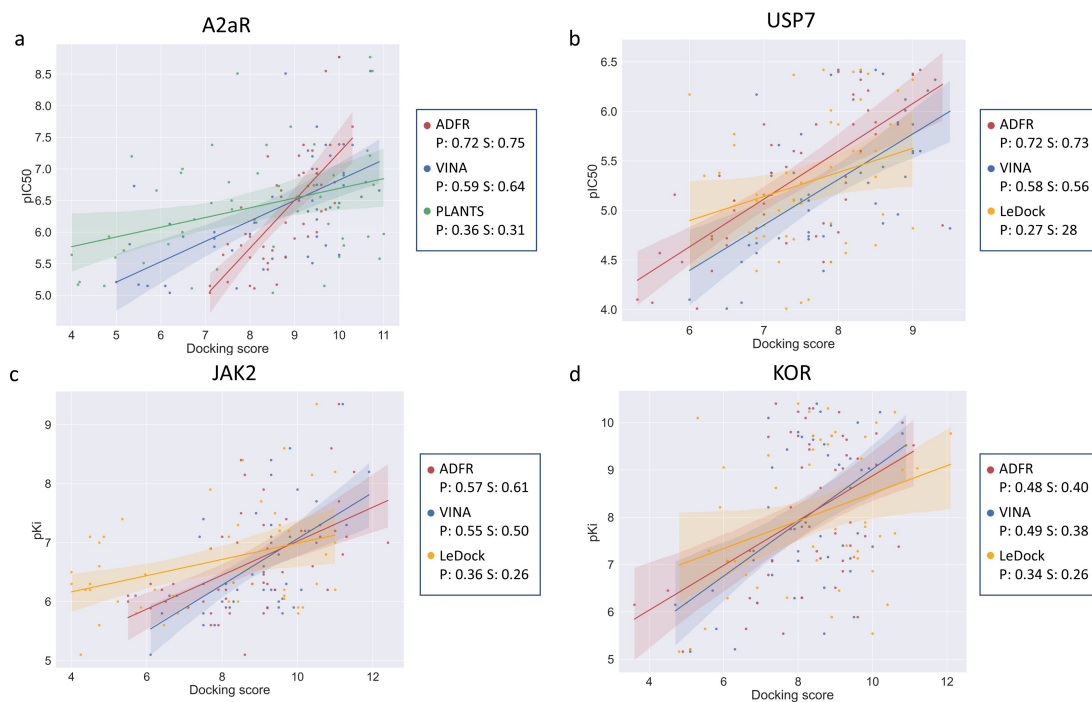


Figure 6.4: Binding energy regression plots. The plots display linear regression between the docking score of the three best MD tools and experimentally determined ligand pIC50 or pKi. (a) A2aR, (b) USP7, (c) JAK2, (d) KOR.

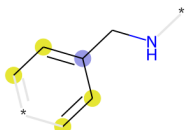
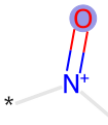
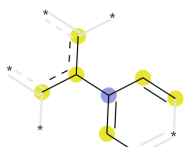
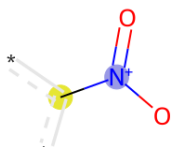
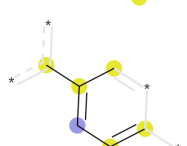
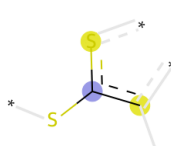
Furthermore, Morgan fingerprint was used to observe tool bias. To this end, molecules that fell outside the confidence interval were classified as under or over-scored and a Random Forest model was produced and trained using python 3's scikit-learn package to determine the correlation between specific molecular fragments and scoring, the results of which can be seen in table6.1

These results indicate that both Vina and ADFR present a positive bias toward large non-polar structures and against small polar ones, as it is known for some docking tools. Fortunately, The molecules found in these groups were different for either of the two MD tools, which suggests that consensus scoring strategies would be enough to eliminate false positives [166,268].

6.2 DGM Molecules Screening

To assess the molecules produced by the DGM, we chose to use the exponential consensus method for MD. As a primary probe into the technique's acuity, we tested the process for the four proteins used in this work using known ligands. Four test groups were created, each comprising one hundred known ligands and three thousand

Table 6.1: Morgan fingerprint fragments. The table show the correlation weights as related to over- or underprediction of the Morgan fragments displaying higher relevance.

Group	Bit	Weight	Group	Bit	Weight	
Overpredicted		0.0352	Underpredicted		0.0230	
						0.0185
						
		0.0227				
		0.0146				

specific DUD-E decoys. The resulting scores were used to create ROC curves for each category. We also took the opportunity to compare one of the most commonly used benchmarking and goals for DGM, and have found out that a very poor performance when applied into discerning drugs from decoys. Figure 6.5 shows the results for A2aR and USP7.

As all the sets performed well, we moved to the docking of all the generated putative ligands against their respective target proteins. The resulting scores of both MD tools were ranked and processed according to the exponential scoring methodology; accordingly, four thousand decoys were also docked and ranked alongside the test molecules. Sample drugs with known affinity for these proteins were also included in each group to provide an extra layer of scrutiny.

To assess structures produced by the DGM, we processed these putative ligands according to exponential consensus methodology. Hence, for each target protein, a group of two thousand specific decoys, the test molecules, and a sample of drugs with known affinity were docked, scored, and processed. The pharmaceutical molecules were added as an extra layer of scrutiny to confirm if the methodology could detect valuable molecules. The results of these screenings can be seen in Figure 6.6

Almost all the drug molecules in this test were placed among the top-scoring ones, indicating that the system can capture molecules of interest. We also observed that a reasonable number of test molecules ranked among the fifth percentile for the

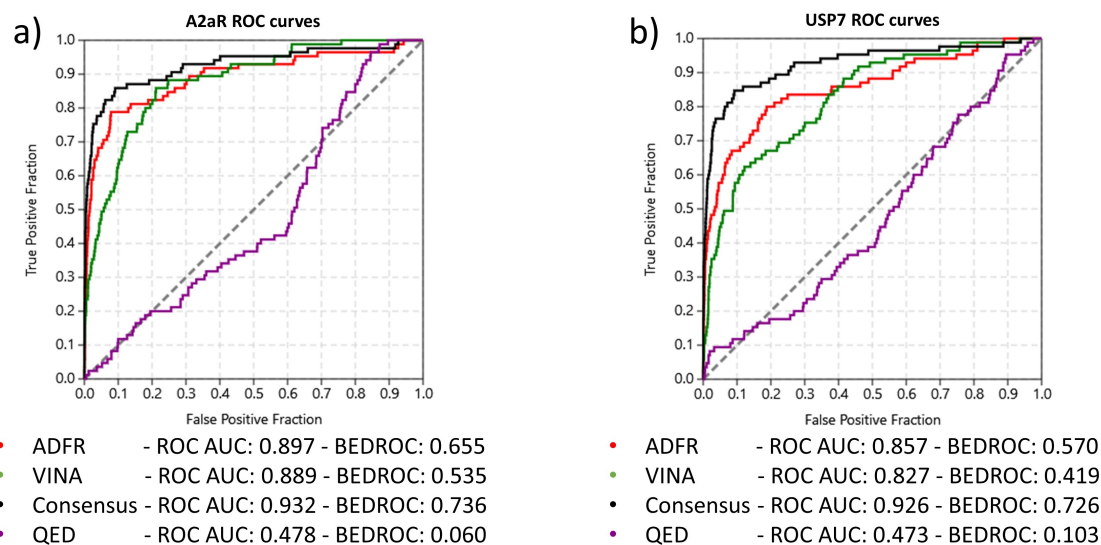


Figure 6.5: A2aR and USP7 ROC curves. (a) represent the ROC curve For A2aR MD results, and (b) that of USP7. The graphs display the performance of ADFR (in red), VINA (in green), exponential consensus scoring (in black) and QED (in purple)

A2aR and USP7 groups. JAK2 and KOR groups presented only 8 and 4 structures with passing scores.

6.3 Validation Through Molecular Dynamics Simulations

The best ten molecules for each of these two groups, as well as the worst five, were selected for validation through MDS in 50ns protein-ligand simulations in GRO-MACS, the results of which can be seen in Table 6.2. Figures 6.7 and 6.8 display stability results for A2aR and USP7, respectively. The complete Stability results can be seen in Appendix Figure A.5.

Even in the case of the KOR set, in which not many DGM molecules were found among the upper five percent screened molecules, MDS results for the best ten structures were clearly better than the worst five. Also, almost all molecules above this cut-out mark presented ligand-protein RMSD under 2\AA , which in this technique indicates relative stability.

6.4 Additional Chemometric Analyses

With the intent of a simple comparison between the above-presented screening methods and some of the most used metrics for DGMs, all test molecules had their

Table 6.2: MDS results for (a) A2aR, (b) JAK2, (c) KOR (d) USP7 ligands. Carbon H bond (CHb), Conventional H bond (CoHb), Distance (Dist.), Unfavorable Donor-Donor (Unf. D-D)**(a) A2aR**

Rank	Label	RMSD	Closest amino acid	Dist. Å
1	71	1.4	Tyr 271 - CoHb	2.56
2	29	1.1	Glu 168 - CoHb	2.25
			Glu 169 - CoHb	2.16
3	41	1.5	Glu 169 - CoHb	2.43
4	68	1.8	Tyr 271 - CoHb	2.14
			Glu 169 - CoHb	2.16
5	70	1.7	Tyr 271 - CoHb	2.11
6	46	1.4	Phe 168 - CoHb	2.18
7	16	2.3	Val 84 - Pi-Sigma	3.54
			Ile 264 - Pi-Sigma	3.62
8	42	1.3	Phe 168 - CoHb	2.39
9	1	1.1	Phe 168 - CoHb	1.62
			Glu 169 - CoHb	2.11
10	21	1.5	Tyr 271 - CoHb	1.64
96	44	3.6	Ala 63 - CoHb	3.8
			Val 84 - CoHb	3.46
97	60	8.4	Phe 168 - Pi-Sigma	2.98
			Glu 169 - Pi-Sigma	1.87
98	48	6.9	Lys 153 - CoHb	2.30
			Tyr 271 - CoHb	2.13
99	70	5.4	Glu 169 - CoHb	1.70
			Tyr 271 - CoHb	1.96
100	12	4.7	His 278 - CoHb	2.07

(b) JAK2

Rank	Label	RMSD	Closest amino acid	Dist. Å
1	12	1.6	Lys 882 - CoHb	2.75
			Asp 994 - CoHb	2.27
			Asn 981 Unf. D-D	1.41
2	11	1.7	Asp 976 - CoHb	2.51
			Tyr 465 - CoHb	2.95
3	24	1.3	Leu 855 - CoHb	2.33
			Ser 936 - CoHb	2.08
4	53	3.1	Leu 855 - Alkyl	5.06
			Val 863 - Alkyl	4.10
			Leu 983 - Alkyl	4.70
5	28	1.9	Lys 587 - CoHb	1.98
6	5	1.2	Lys 882 - CoHb	2.83
			Gly 858 - CoHb	2.6
7	23	1.6	Asn 859 - CoHb	2.29
			Asn 981 - CoHb	2.71
8	9	2.3	Lys 857 - CoHb	2.04
9	6	1.7	Asp 939 - CoHb	2.70
10	3	1.2	Gly 858 - CoHb	2.24
67	27	2.9	Asp 994 - CoHb	3.23
68	6	5.2	Phe 860 - CoHb	2.70
69	39	9.2	Asp 994 - CoHb	2.89
70	4	6.3	Arg 980 - CoHb	2.01
			Asn 981 - CoHb	2.36
71	2	4.9	Phe 860 - CoHb	1.83
			Asp 976 - CoHb	2.52

(c) KOR

Rank	Label	RMSD	Closest amino acid	Dist. Å
1	53	1.2	Asp 138 - CoHb	2.37
			Tyr 139 - CoHb	1.95
			Gln 115 - CoHb	3.01
2	11	1.1	Asp 138 - CoHb	2.23
			Tyr 139 - CoHb	2.39
3	50	1.9	Gln 115 - CoHb	2.38
4	27	1.5	Tyr 312 - CoHb	2.16
			Tyr 320 - CoHb	1.99
			Asp 138 - CoHb	1.97
5	71	1.8	Tyr 139 - CoHb	2.17
			Tyr 312 - CoHb	1.96
6	68	3.4	Tyr 312 - CoHb	2.6
7	11	1.8	Asp 138 - CoHb	2.16
			His 291 - CoHb	2.01
8	10	1.8	Val 134 - CoHb	2.81
			Asp 138 - CoHb	1.90
9	60	4.2	Tyr 312 - CoHb	2.29
			Tyr 312 - CoHb	2.32
10	4	1.2	Tyr 312 - CoHb	2.89
75	2	4.5	Trp 287 Pi-Sigma	3.99
76	60	9.8	Tyr 320 - Pi-Sigma	3.62
77	78	5.7	Gln 115 - CoHb	2.62
			Tyr 320 - Unf. D-D	2.86
78	33	4.3	Gln 115 - CoHb	2.63
			Tyr 320 - Unf. D-D	2.86
79	35	3.6	Gln 115 - CoHb	2.87
			Val 134 - CoHb	2.36

(d) USP7

Rank	Label	RMSD	Closest amino acid	Dist. Å
1	53	1.2	Val 296 - CoHb	2.98
			Tyr 465 - CoHb	2.72
2	72	1.3	Asp 295 - CoHb	2.51
			Tyr 465 - CoHb	2.95
3	18	1.4	Asp 295 - CoHb	2.53
			Tyr 465 - CoHb	2.72
4	28	1.4	Val 296 - CoHb	2.91
5	34	1.8	Arg 408 - CoHb	2.35
			Phe 409 - CoHb	2.86
6	4	1.2	Asp 295 - CoHb	2.08
			Tyr 465 - CoHb	2.98
7	11	1.3	His 294 - CoHb	2.16
			Ile 264 - CoHb	2.43
8	9	1.7	Gln 297 - CoHb	2.89
9	60	1.3	Tyr 465 - CoHb	1.62
10	3	1.2	Leu 406 - CoHb	1.74
96	2	5.2	His 456 - CoHb	2.78
			Tyr 465 - CoHb	2.30
97	13	4.2	Asn 512 - CoHb	2.39
98	41	5.1	Val 296 - CoHb	2.30
99	59	6.3	Val 296 - CoHb	2.06
100	2	7.1	Asn 418 - CoHb	2.16
			Tyr 465 - CoHb	2.40

6. Results and Discussion

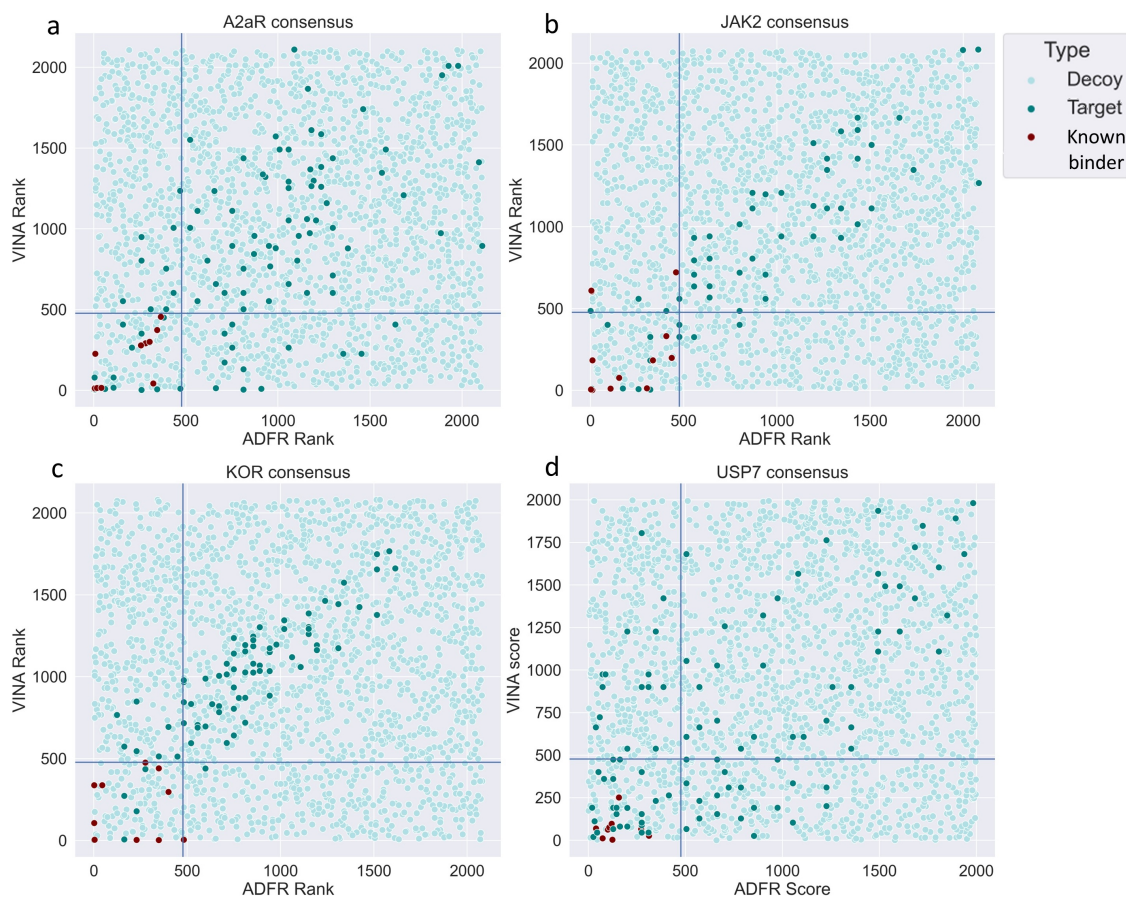


Figure 6.6: Consensus scoring plots. This distribution shows the MD scoring ranking with ADFR (x-axis) and VINA (y-axis); in this methodology, the closer to zero a molecule is placed, the higher its score. The blue lines indicate the cut-off mark, set as 5% for a more strict evaluation. Molecules have been separated by type as decoy, target or known binder.

drug-likeness score, total polar surface area (TPSA), and partition coefficient ($\log P$) calculated, with the results for A2aR displayed in Figure 6.9.

With the outcome of these analyses, no meaningful correlation was present between these metrics, making them non-redundant. Furthermore, these seem to be good additional filters, enabling the selection of molecules with better pharmacological properties. For instance, we observed that for A2aR, one-third of the DGM molecules, including some with high consensus scores, are unlikely to be able to transpose the blood-brain barrier, a feature fundamental for drugs targeting this receptor.

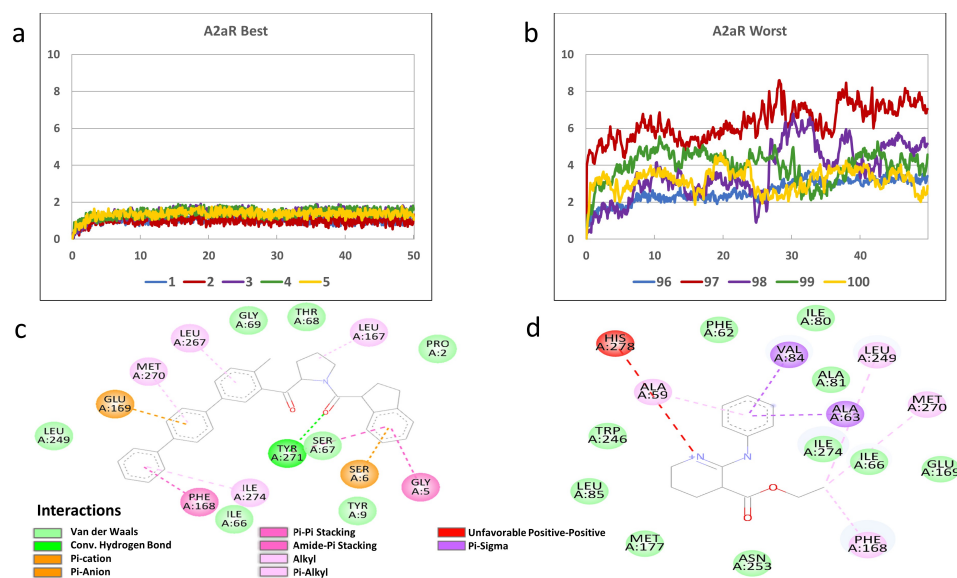


Figure 6.7: A2aR MDS results. The MDS results for the best and worst five molecules are displayed in a and b, respectively. In c the 2d representation of the final pose and interaction of the best consensus scoring molecule with the receptor, d represent the same, but for the worst.

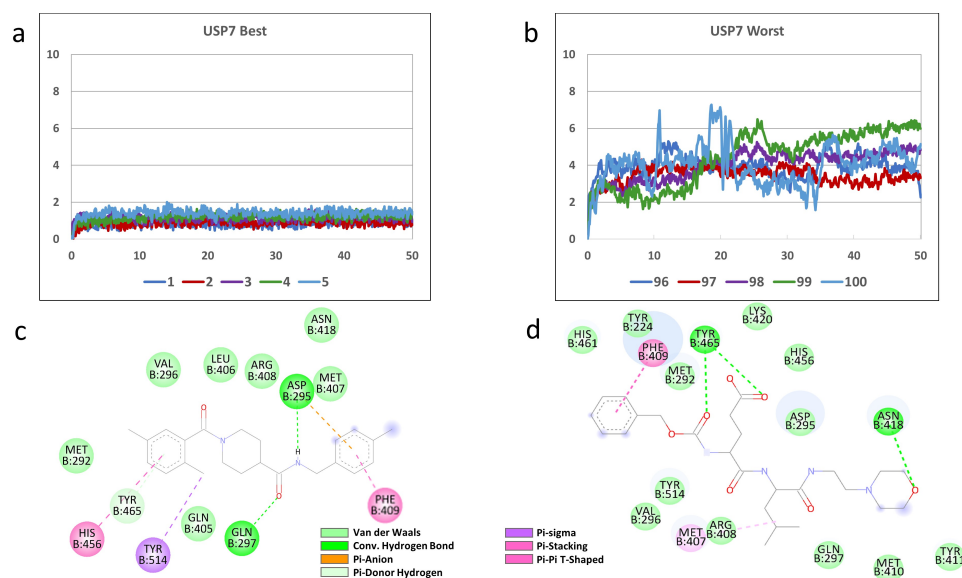


Figure 6.8: USP7 MDS results. The MDS results for the best and worst five molecules are displayed in a and b, respectively. In c the 2d representation of the final pose and interaction of the best consensus scoring molecule with the receptor, d represent the same, but for the worst.

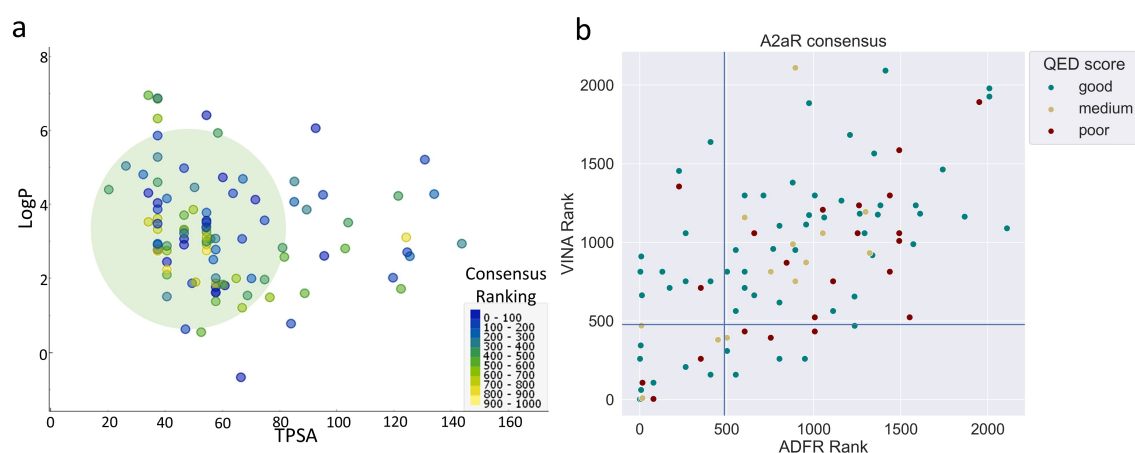


Figure 6.9: Chemometric analyses. (a) Shows a graph used to estimate compartment-specific absorption, with TPSA at the x-axis and LogP in y, with A2aR data. The consensus scoring is represented in a hue going from blue to yellow, and the zone where blood-brain barrier transposing molecules are expected to fall is delineated by the large green circle. (b) shows the A2aR exponential consensus graph with QED values shown as the dot's hue, with good values (0 and over) in green, intermediate values (between -0.5 and 0) in yellow, and poor results (under -0.5) in red

Conclusion

7.1 Conclusions

In this study, comprehensive studies on small molecule screening and ML models for *de novo* drug design were performed, focusing on the intersection between both techniques and model validation. This assessment concluded that, despite ML's displayed potential in *de novo* drug design, a glaring flaw is found in model development and benchmarking: tridimensional molecular information is rarely accounted for. This situation is likely to impact development, as the research community focuses on models lacking adequate validation and that may focus on the production of information not transferable to actual biochemical conditions.

As a solution for this, a validation including MD tools was proposed since this technique has a relatively low computational cost and is able to encompass the use of more realistic molecular calculations. Since the choice of MD software and base crystal structures is a momentous step of molecular screening, a more flexible approach encompassing the assessment of each MD tool was preferred. To this means, multiple methodologies for accuracy testing were studied, with cross-docking analysis and linear regression between docking score and binding energy being selected in the end. The former was incorporated as it is a well-established methodology for software acuity and crystal structure evaluation, and the latter was added since binding energy estimations are a more direct means to assess protein-ligand interaction. Considering the importance of accessibility in benchmarking, we opted for using freeware and open source software, with a selection of six of the best well-known MD tools (AD4, ADFR, LeDock, PLANTS, rDOCK, and VINA) being tested for acuity in the end. ADFR and VINA displayed great results overall, outperforming the other applications by large margins.

Exponential consensus scoring was incorporated to better evaluate the molecules generated by the ML models into the molecular screening. The effectiveness of this technique was gauged through the generation of ROC curves and the use of known binder molecules as positive controls, and DUD-E-generated decoys were used for sample enrichment. In this, a significant increase in predictive performance was observed when applying exponential consensus scoring, and the control molecules were displayed inside the cutoff mark, even when very strict margins (5%) were used.

As a final validation step, MDS of the systems containing the best ten and worst five scoring molecules of each set and their respective target protein were performed, making use of parameters found in well-established protocols. The resulting data pointed to a clear divide between groups, with the overwhelming majority of the best molecules displaying displacement under 2Å and the worst with much higher RMSD, indicating that the model is able to select molecules with high affinity for their specific binding sites.

Additionally, we observe that, although the first implementation of this process may take a while to execute, much of the process can be skipped on subsequent tests for the same binding sites. For instance, it is not necessary to repeat the MD tools evaluation. The information stored from the experiments can also be directly compared with new data if other MD software is tested. Also, the scores of decoy molecules for each MD tool can be re-introduced against new ligands for consensus score ranking.

We expect that this system can bridge the impressive potential of DGMs and the representational metrics provided by MD and MDS. Although, it is essential to observe that this system was developed as an additional implement to, and not a substitute for, other chemometric evaluations and does not account for molecular properties other than structural affinity. Thus, qualities such as chemical stability, toxicity, and biodisponibility, among others, must still be assessed through other means.

7.2 Future Work

To better place the methodology suggested in this work, testing multiple *de novo* drug design DL models and comparing their outcomes would be useful. Other than that, although the fraction of molecules displayed inside the cutoff mark is a useful metric, developing other means for comparison between models would be an inter-

esting proposal, which could perhaps incorporate multiple molecular characteristics and encompass the use of chemometrics.

Using the results' high-ranking molecules selected by the screening process as feedback for DGMs training and comparing their performance with similar systems or previous iterations of the same model would also be of interest. It could perhaps demonstrate the method's capability as not only a benchmarking tool but also a means of directly improving DGMs.

Another important step would be directly comparing the performance of the method conceived in this work with evaluation frameworks such as GuacaMol and MOSES, since those bestride the benchmark of DL-based *de novo* drug design models. Thus, allowing for the analysis of weak points and bestowing insight on areas of interest for improvement.

Furthermore, as the knowledge needed to accurately employ MD and MDS tools represents a significant barrier to performing the techniques described in this work, this project was originally planned to produce a program able to automate most of its processes and with an interface that would simplify its usage.

Lastly, even though MDS is a well-established technique with high predictive performance, chemically synthesizing and *in vitro* testing of high-scoring molecules would impart information paramount for the validation of this model. Perhaps even producing a molecule candidate for drug test trials.

7. Conclusion

A

Appendix

A.1 Appendix A

Lists containing all the ligands used in docking experiments for each target protein.

Table A.1: List of A2aR ligands

N ^o	ChEMBL ID	pIC50	AD4	ADFR	LeDock	PLANTS	rDock	VINA
1	CHEMBL1081334	-8	-7.34	-7.7	-138.77	-30.4	-6.1	6.1
2	CHEMBL1081335	-8	-7.23	-5.5	-99.93	-26.1	-8.8	8.8
3	CHEMBL1076503	-8.8	-6.99	-8.9	-126.14	-7.8	-8.1	8.1
4	CHEMBL1076504	-8.3	-8.06	-5.2	-149.04	-20.5	-8.5	8.5
5	CHEMBL1077749	-8.1	-9.05	-7.3	-114.13	-12.3	-7.7	7.7
6	CHEMBL1077750	-8.4	-7.46	-8.7	-118.84	-11.8	-8.3	8.3
7	CHEMBL1077751	-8.7	-7.1	-7.8	-107.57	-18.1	-10.9	10.9
8	CHEMBL1077752	-9.1	-8.04	-9.3	-140.5	-9.9	-10.1	10.1
9	CHEMBL1077902	-9.2	-7.53	-7.4	-112.13	-21	-10.1	10.1
10	CHEMBL1077915	-9.5	-8.97	-6.5	-126.38	-30.5	-9.5	9.5
11	CHEMBL1077943	-9.1	-8.16	-8.1	-138.05	-14.6	-8.5	8.5
12	CHEMBL2398483	-8.7	-8.15	-5.5	-133.67	-33.1	-10.5	10.5
13	CHEMBL1077985	-7.1	-8.46	-9.8	-97.86	-8.1	-5.7	5.7
14	CHEMBL1078001	-9.9	-9.37	-9.9	-121.02	-28.3	-10.6	10.6
15	CHEMBL1078030	-7.5	-7.79	-7.6	-93.26	-10.6	-5	5
16	CHEMBL1078085	-9.9	-7.9	-6.8	-117.75	-27.5	-10	10
17	CHEMBL1078086	-9.6	-7.84	-7.2	-147.93	-13.2	-9.5	9.5
18	CHEMBL1078087	-7.4	-8.43	-5.1	-91.81	-14.8	-10	10
19	CHEMBL1078347	-7.9	-8.47	-7.7	-149.81	-25.8	-8	8
20	CHEMBL1078348	-9.3	-9.16	-7.1	-130.28	-12.8	-5.4	5.4
21	CHEMBL1078349	-9.5	-7.87	-8	-112.38	-23.5	-7.2	7.2
22	CHEMBL1079043	-9.2	-8.46	-8.3	-139.53	-7	-6.2	6.2
23	CHEMBL1079044	-8.3	-6.92	-5.9	-129.32	-21.4	-7.2	7.2
24	CHEMBL1079059	-8.6	-8.25	-8.4	-92.98	-17.3	-5.5	5.5
25	CHEMBL1079115	-9.5	-8.42	-5.6	-112.39	-27.2	-9.4	9.4
26	CHEMBL1079116	-7.5	-9.16	-6.9	-109.79	-6.2	-5.3	5.3
27	CHEMBL1079117	-9	-7.84	-6.4	-142.19	-24.1	-7.2	7.2
28	CHEMBL1079118	-9.4	-9.68	-9.3	-146.77	-7.2	-9.3	9.3
29	CHEMBL1079132	-7.8	-8.42	-9.7	-141.35	-24.3	-9.8	9.8
30	CHEMBL1079141	-7.8	-8.41	-5.5	-142.39	-13.3	-9.7	9.7
31	CHEMBL1079142	-9.7	-8.83	-5.2	-121.6	-18.2	-9.5	9.5
32	CHEMBL1079232	-8.1	-8	-6.3	-114.62	-14.1	-9.5	9.5
33	CHEMBL1079894	-9.6	-9.47	-8.7	-133.82	-11.1	-9.3	9.3
34	CHEMBL4167557	-9.5	-9.43	-9.9	-122.64	-19.8	-8.8	8.8

A. Appendix

35	CHEMBL1080067	-8.5	-8.17	-9.8	-101.49	-22.2	-7.6	7.6
36	CHEMBL1080068	-9.1	-9.19	-5.5	-147.26	-28.7	-9.8	9.8
37	CHEMBL1081873	-8	-9.87	-6.2	-146.93	-13.7	-6.5	6.5
38	CHEMBL1082063	-8.4	-8.06	-6.2	-105.18	-16.6	-9.4	9.4
39	CHEMBL1080421	-8.8	-7.68	-5	-141.52	-24.6	-9.3	9.3
40	CHEMBL1081160	-9.1	-6.6	-9.8	-121.54	-26.3	-9.4	9.4
41	CHEMBL290106	-9.3	-8.14	-7	-98.76	-13.5	-9.5	9.5
42	CHEMBL113142	-9.2	-8.15	-9.2	-117.91	-29.6	-9.7	9.7
43	CHEMBL4064795	-8.4	-6.73	-9.7	-102.98	-24	-9.4	9.4
44	CHEMBL2381758	-7.1	-6.75	-7.6	-129.2	-5.5	-6.2	6.2
45	CHEMBL1258397	-9.8	-7.94	-8.8	-128.02	-17.2	-10.3	10.3
46	CHEMBL2024116	-9.7	-7.59	-6.1	-147.7	-30.2	-10.7	10.7
47	CHEMBL4741609	-10	-9.06	-9.3	-147.28	-32.1	-10	10
48	CHEMBL2381760	-8.6	-7.73	-9.9	-136.81	-32.5	-10.1	10.1
49	CHEMBL2381761	-10	-7.61	-9.1	-136.92	-19.2	-10.1	10.1
50	CHEMBL21572	-10.3	-8.17	-6.4	-132.52	-12.8	-9.5	9.5

Table A.2: Listo of JAK2 ligands

N ^o	ChEMBL ID	pKi	AD4	ADFR	LeDock	PLANTS	rDock	VINA
1	CHEMBL445332	8.21	-6.4	-7.8	-5.9	-128.1	-50.4	-7.2
2	CHEMBL573214	9.72	-9.9	-8	-8.9	-103.9	-42.1	-8.9
3	CHEMBL19019	9.05	-7.8	-8.1	-6	-104.5	-47.6	-7.3
4	CHEMBL267495	9.64	-7.7	-9.3	-8.6	-106.1	-37.5	-8.4
5	CHEMBL70	7.07	-7.3	-6.1	-6.3	-107.8	-49.5	-7.3
6	CHEMBL281786	9.43	-7	-8.3	-8.4	-128.4	-32.4	-7.7
7	CHEMBL301160	10.1	-7.2	-8.2	-5.3	-123.4	-38.2	-7.2
8	CHEMBL508112	6.3	-8.6	-8.1	-7.7	-103.3	-35.8	-8.1
9	CHEMBL322717	5.54	-7.3	-7.6	-10	-125.5	-49.6	-8.7
10	CHEMBL322717	9.77	-6.4	-10.1	-12.1	-148.5	-38.6	-10.8
11	CHEMBL267495	7.66	-8.7	-9	-10.6	-127.2	-49.1	-9.6
12	CHEMBL1081093	7.28	-7.6	-8.5	-9.6	-96.0	-38.1	-8
13	CHEMBL127463	9.52	-8.2	-11.1	-10.9	-111.2	-31.1	-10.9
14	CHEMBL127969	9.24	-8.6	-9.6	-9.7	-126.9	-49.8	-9.4
15	CHEMBL136421	5.89	-7.1	-9.6	-8.9	-110.5	-30.5	-9.5
16	CHEMBL1642760	9.03	-9.1	-9.2	-11.2	-112.4	-48.2	-9.9
17	CHEMBL1814705	7.29	-6.7	-6.7	-5.6	-123.5	-28.4	-6.8
18	CHEMBL1921841	9	-8	-9	-10.8	-142.9	-49.7	-9.7
19	CHEMBL2022296	7.4	-8.4	-9.4	-9.3	-124.0	-48.3	-9.2
20	CHEMBL2048776	7.6	-6.1	-9.8	-10.2	-101.6	-36.4	-8.6

21	CHEMBL2112345	6.51	-9.2	-9	-6.9	-91.7	-30	-8.7
22	CHEMBL216640	10.23	-4	-8	-8.3	-109.2	-33.9	-8.6
23	CHEMBL219319	7.75	-9.5	-7.8	-9	-148.8	-30.5	-7.9
24	CHEMBL2386901	7.8	-6.8	-8.3	-9.8	-102.5	-35.6	-8.1
25	CHEMBL2387196	7.08	-7.2	-9.2	-6.1	-116.9	-27.5	-8
26	CHEMBL3086747	9.8	-8.8	-7.8	-9	-92.5	-28.3	-8
27	CHEMBL240657	6.86	-7.9	-9.2	-10	-99.1	-30.7	-9.4
28	CHEMBL514662	7.89	-7.5	-9.4	-9.8	-141.7	-42.8	-8.2
29	CHEMBL261024	7.16	-5	-9.5	-9.5	-97.9	-44.5	-8.8
30	CHEMBL266314	7.38	-8.6	-10.7	-8.4	-99.8	-41.1	-10
31	CHEMBL271797	10.4	-7.5	-7.4	-8	-142.7	-26.8	-8.5
32	CHEMBL286411	5.21	-8.1	-5.1	-5.1	-102.8	-36.4	-6.3
33	CHEMBL3084634	6.29	-9.5	-6.7	-6.1	-121.4	-41.4	-7.9
34	CHEMBL3262089	9.31	-7.9	-9.3	-7.2	-129.4	-29.7	-8.8
35	CHEMBL3262367	10	-8.6	-10.8	-10.1	-143.9	-50.7	-9.6
36	CHEMBL3264441	8.62	-8.6	-8.6	-7.4	-131	-40.1	-9.4
37	CHEMBL328411	7.13	-10.7	-8.6	-9.6	-114	-40.9	-9.4
38	CHEMBL593583	7.09	-7.5	-7.3	-8.5	-144	-41.4	-6.8
39	CHEMBL58646	10.22	-7.3	-9.1	-10.6	-150.5	-25.8	-10.2
40	CHEMBL3581754	10.3	-3.1	-8.4	-8.8	-97	-41.4	-8.3
41	CHEMBL3639941	6.15	-8.4	-3.6	-10.4	-113.2	-43.5	-4.7
42	CHEMBL3647958	5.64	-7.2	-6.6	-5.5	-145.8	-42.5	-5.8
43	CHEMBL3698904	6.19	-7	-6.9	-5	-108.8	-48.2	-6.9
44	CHEMBL3665419	6.78	-6.7	-7.2	-6.6	-103.5	-46.5	-7.3
45	CHEMBL3665424	5.16	-6.2	-4.9	-4.8	-128.6	-26.1	-5.1
46	CHEMBL3665425	6.45	-6.7	-4.5	-9.1	-120.9	-39.2	-5.7
47	CHEMBL3678965	9.11	-8	-10.1	-8.3	-139.4	-25.9	-9.1
48	CHEMBL3678975	9.7	-5.3	-7.2	-9.8	-139.5	-43.3	-8.1
49	CHEMBL3678977	9.54	-7.8	-8.3	-8.8	-126	-27.9	-8.7
50	CHEMBL3678981	9.72	-6.2	-9.2	-8.4	-123.4	-40.4	-8.4

Table A.3: List of KOR ligands

N^o	ChEMBL ID	pKi	AD4	ADFR	LeDock	PLANTS	rDock	VINA
1	CHEMBL445332	8.21	-6.4	-7.8	-5.9	-128.1	-50.4	-7.2
2	CHEMBL573214	9.72	-9.9	-8	-8.9	-103.9	-42.1	-8.9
3	CHEMBL19019	9.05	-7.8	-8.1	-6	-104.5	-47.6	-7.3
4	CHEMBL267495	9.64	-7.7	-9.3	-8.6	-106.1	-37.5	-8.4
5	CHEMBL70	7.07	-7.3	-6.1	-6.3	-107.8	-49.5	-7.3
6	CHEMBL281786	9.43	-7	-8.3	-8.4	-128.4	-32.4	-7.7
7	CHEMBL301160	10.1	-7.2	-8.2	-5.3	-123.4	-38.2	-7.2
8	CHEMBL508112	6.3	-8.6	-8.1	-7.7	-103.3	-35.8	-8.1
9	CHEMBL322717	5.54	-7.3	-7.6	-10	-125.5	-49.6	-8.7
10	CHEMBL322717	9.77	-6.4	-10.1	-12.1	-148.5	-38.6	-10.8
11	CHEMBL267495	7.66	-8.7	-9	-10.6	-127.2	-49.1	-9.6
12	CHEMBL1081093	7.28	-7.6	-8.5	-9.6	-96	-38.1	-8
13	CHEMBL127463	9.52	-8.2	-11.1	-10.9	-111.2	-31.1	-10.9
14	CHEMBL127969	9.24	-8.6	-9.6	-9.7	-126.9	-49.8	-9.4
15	CHEMBL136421	5.89	-7.1	-9.6	-8.9	-110.5	-30.5	-9.5
16	CHEMBL1642760	9.03	-9.1	-9.2	-11.2	-112.4	-48.2	-9.9
17	CHEMBL1814705	7.29	-6.7	-6.7	-5.6	-123.5	-28.4	-6.8
18	CHEMBL1921841	9	-8	-9	-10.8	-142.9	-49.7	-9.7
19	CHEMBL2022296	7.4	-8.4	-9.4	-9.3	-124	-48.3	-9.2
20	CHEMBL2048776	7.6	-6.1	-9.8	-10.2	-101.6	-36.4	-8.6
21	CHEMBL2112345	6.51	-9.2	-9	-6.9	-91.7	-30	-8.7
22	CHEMBL216640	10.2	-4	-8	-8.3	-109.2	-33.9	-8.6
23	CHEMBL219319	7.75	-9.5	-7.8	-9	-148.8	-30.5	-7.9
24	CHEMBL2386901	7.8	-6.8	-8.3	-9.8	-102.5	-35.6	-8.1
25	CHEMBL2387196	7.08	-7.2	-9.2	-6.1	-116.9	-27.5	-8
26	CHEMBL3086747	9.8	-8.8	-7.8	-9	-92.5	-28.3	-8
27	CHEMBL240657	6.86	-7.9	-9.2	-10	-99.1	-30.7	-9.4
28	CHEMBL514662	7.89	-7.5	-9.4	-9.8	-141.7	-42.8	-8.2
29	CHEMBL261024	7.16	-5	-9.5	-9.5	-97.9	-44.5	-8.8
30	CHEMBL266314	7.38	-8.6	-10.7	-8.4	-99.8	-41.1	-10

31	CHEMBL271797	10.4	-7.5	-7.4	-8	-142.7	-26.8	-8.5
32	CHEMBL286411	5.21	-8.1	-5.1	-5.1	-102.8	-36.4	-6.3
33	CHEMBL3084634	6.29	-9.5	-6.7	-6.1	-121.4	-41.4	-7.9
34	CHEMBL3262089	9.31	-7.9	-9.3	-7.2	-129.4	-29.7	-8.8
35	CHEMBL3262367	10	-8.6	-10.8	-10.1	-143.9	-50.7	-9.6
36	CHEMBL3264441	8.62	-8.6	-8.6	-7.4	-131	-40.1	-9.4
37	CHEMBL328411	7.13	-10.7	-8.6	-9.6	-114	-40.9	-9.4
38	CHEMBL593583	7.09	-7.5	-7.3	-8.5	-144	-41.4	-6.8
39	CHEMBL58646	10.22	-7.3	-9.1	-10.6	-150.5	-25.8	-10.2
40	CHEMBL3581754	10.3	-3.1	-8.4	-8.8	-97	-41.4	-8.3
41	CHEMBL3639941	6.15	-8.4	-3.6	-10.4	-113.2	-43.5	-4.7
42	CHEMBL3647958	5.64	-7.2	-6.6	-5.5	-145.8	-42.5	-5.8
43	CHEMBL3698904	6.19	-7	-6.9	-5	-108.8	-48.2	-6.9
44	CHEMBL3665419	6.78	-6.7	-7.2	-6.6	-103.5	-46.5	-7.3
45	CHEMBL3665424	5.16	-6.2	-4.9	-4.8	-128.6	-26.1	-5.1
46	CHEMBL3665425	6.45	-6.7	-4.5	-9.1	-120.9	-39.2	-5.7
47	CHEMBL3678965	9.11	-8	-10.1	-8.3	-139.4	-25.9	-9.1
48	CHEMBL3678975	9.7	-5.3	-7.2	-9.8	-139.5	-43.3	-8.1
49	CHEMBL3678977	9.54	-7.8	-8.3	-8.8	-126	-27.9	-8.7
50	CHEMBL3678981	9.72	-6.2	-9.2	-8.4	-123.4	-40.4	-8.4

Table A.4: List of USP7 ligands

N ^o	ChEMBL ID	pIC50	AD4	ADFR	LeDock	PLANTS	rDock	VINA
1	CHEMBL1762621	4.85	-9.35	-9.4	-13.9	-148.99	-35	-7.9
2	CHEMBL2159495	5.38	-5.15	-7.1	-13.19	-145.45	-32.45	-7.4
3	CHEMBL2159497	5.16	-6.4	-7.1	-5.52	-107.07	-4.74	-7.6
4	CHEMBL2159498	5.1	-6.03	-6.6	-11.27	-135.86	-25.52	-7.6
5	CHEMBL2159498	4.71	-6.34	-6.3	-8.44	-121.72	-15.31	-7.6
6	CHEMBL2159498	5.1	-5.79	-6.6	-11.58	-137.37	-26.61	-7.6
7	CHEMBL2159499	5.66	-7	-7.1	-6.53	-112.12	-8.39	-7.2
8	CHEMBL2159501	5.77	-7.57	-8.5	-8.85	-123.74	-16.77	-7.5
9	CHEMBL2159502	5.11	-6.73	-8.1	-10.67	-132.83	-23.33	-7.5
10	CHEMBL2159503	6.42	-3.94	-8.3	-6.73	-113.13	-9.11	-8
11	CHEMBL2159504	6.37	-6.61	-7	-7.33	-116.16	-11.3	-7.9
12	CHEMBL2159505	6.4	-6.47	-8	-9.66	-127.78	-19.69	-8.3
13	CHEMBL2159506	4.72	-6.61	-7.5	-12.38	-141.41	-29.53	-7.2
14	CHEMBL2159507	6.42	-9.04	-9.1	-5.21	-105.56	-3.65	-8.5
15	CHEMBL2159508	6.38	-7.61	-9	-4.51	-102.02	-1.09	-8.6

A. Appendix

16	CHEMBL4061087	4.82	-7.96	-7.9	-4.4	-101.52	-5.73	-9.5
17	CHEMBL4062409	5.16	-5.58	-5.8	-8.24	-120.71	-14.58	-7
18	CHEMBL4063215	5.07	-6.45	-7	-10.26	-130.81	-21.88	-8.5
19	CHEMBL4064074	4.07	-7.14	-8.1	-10.77	-133.33	-23.7	-7.7
20	CHEMBL4065173	5.35	-5.12	-6.4	-10.36	-131.31	-22.24	-7.1
21	CHEMBL4065700	5.28	-6.66	-8.1	-12.38	-141.41	-29.53	-8.4
22	CHEMBL4069093	5.38	-5.89	-7.1	-10.57	-132.32	-22.97	-8.3
23	CHEMBL4069137	4.1	-6.69	-7.8	-12.08	-139.9	-28.44	-8
24	CHEMBL4070513	5	-6.96	-6.6	-9.76	-128.28	-20.05	-7.6
25	CHEMBL4070982	6.32	-8.27	-8.3	-6.53	-112.12	-8.39	-9.3
26	CHEMBL4073748	4.74	-5.5	-7.3	-12.99	-144.44	-31.72	-7.6
27	CHEMBL4074910	5.6	-6.5	-7.4	-10.46	-131.82	-22.6	-9
28	CHEMBL4074910	5.34	-7.15	-8.1	-5.62	-107.58	-5.1	-8.9
29	CHEMBL4076247	6.12	-8	-8.2	-7.13	-115.15	-10.57	-8.9
30	CHEMBL4076247	6.01	-7.59	-8.4	-5.92	-109.09	-6.2	-8.9
31	CHEMBL4076471	4.67	-8.71	-8.3	-8.24	-120.71	-14.58	-8.9
32	CHEMBL4079672	4.74	-7.04	-7.7	-4.61	-102.53	-1.46	-7.8
33	CHEMBL4079850	5.58	-5.52	-7.5	-6.12	-110.1	-6.93	-9
34	CHEMBL4081422	4.87	-6.14	-7.8	-7.74	-118.18	-12.76	-8.6
35	CHEMBL4081614	4.57	-6.01	-7.6	-5.52	-107.07	-4.74	-8.9
36	CHEMBL4081911	4.65	-6.93	-8.5	-9.96	-129.29	-20.78	-8.8
37	CHEMBL4087728	5.44	-6.91	-7.3	-9.56	-127.27	-19.32	-8.6
38	CHEMBL4090158	4.39	-5.41	-7.3	-11.98	-139.39	-28.07	-7.8
39	CHEMBL4090799	6.21	-7.69	-8.4	-4.2	-100.51	-5	-9.2
40	CHEMBL4091535	4.48	-6.99	-7.8	-13.19	-145.45	-32.45	-8.7
41	CHEMBL4092836	6.17	-5.54	-6.2	-6.83	-113.64	-9.48	-6.9
42	CHEMBL4092976	5.87	-6	-8	-5.92	-109.09	-6.2	-8.8
43	CHEMBL4092976	5.89	-6.56	-8.5	-7.84	-118.69	-13.13	-8.8
44	CHEMBL4092976	5.87	-8.3	-8.2	-8.95	-124.24	-17.14	-8.8
45	CHEMBL4094344	4.78	-5.38	-8	-5.72	-108.08	-5.47	-7.6
46	CHEMBL4097329	5.6	-6.75	-8.3	-5.11	-105.05	-3.28	-9.1
47	CHEMBL4097744	4.01	-7.77	-8.3	-13.9	-148.99	-35	-7.8
48	CHEMBL4097876	5.87	-5.47	-7.7	-10.46	-131.82	-22.6	-9
49	CHEMBL4098430	4.48	-7.5	-7.9	-6.32	-111.11	-7.66	-7.9
50	CHEMBL4100447	4.96	-6.28	-8	-11.98	-139.39	-28.07	-8.8

A.2 Appendix B

Graphs of regression plots of all docking experiments against binding affinity values (pKi or pIC50) and MDS RMSD results.

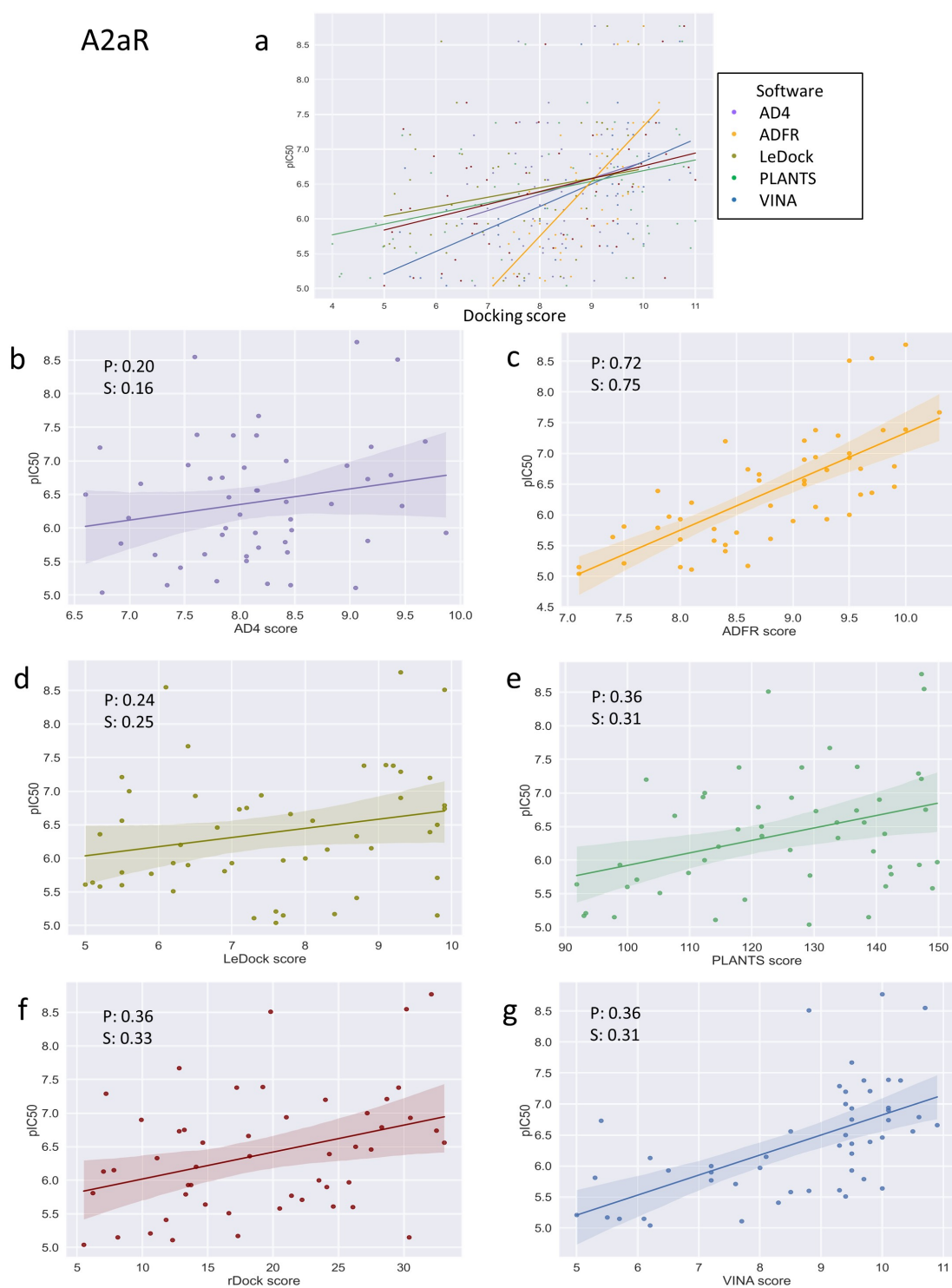


Figure A.1: A2aR regression plots. Results for binding affinity and docking scores. (a) global, (b) AD4, (c) ADFR, (e) PLANTS, (f) rDock, (g) VINA.

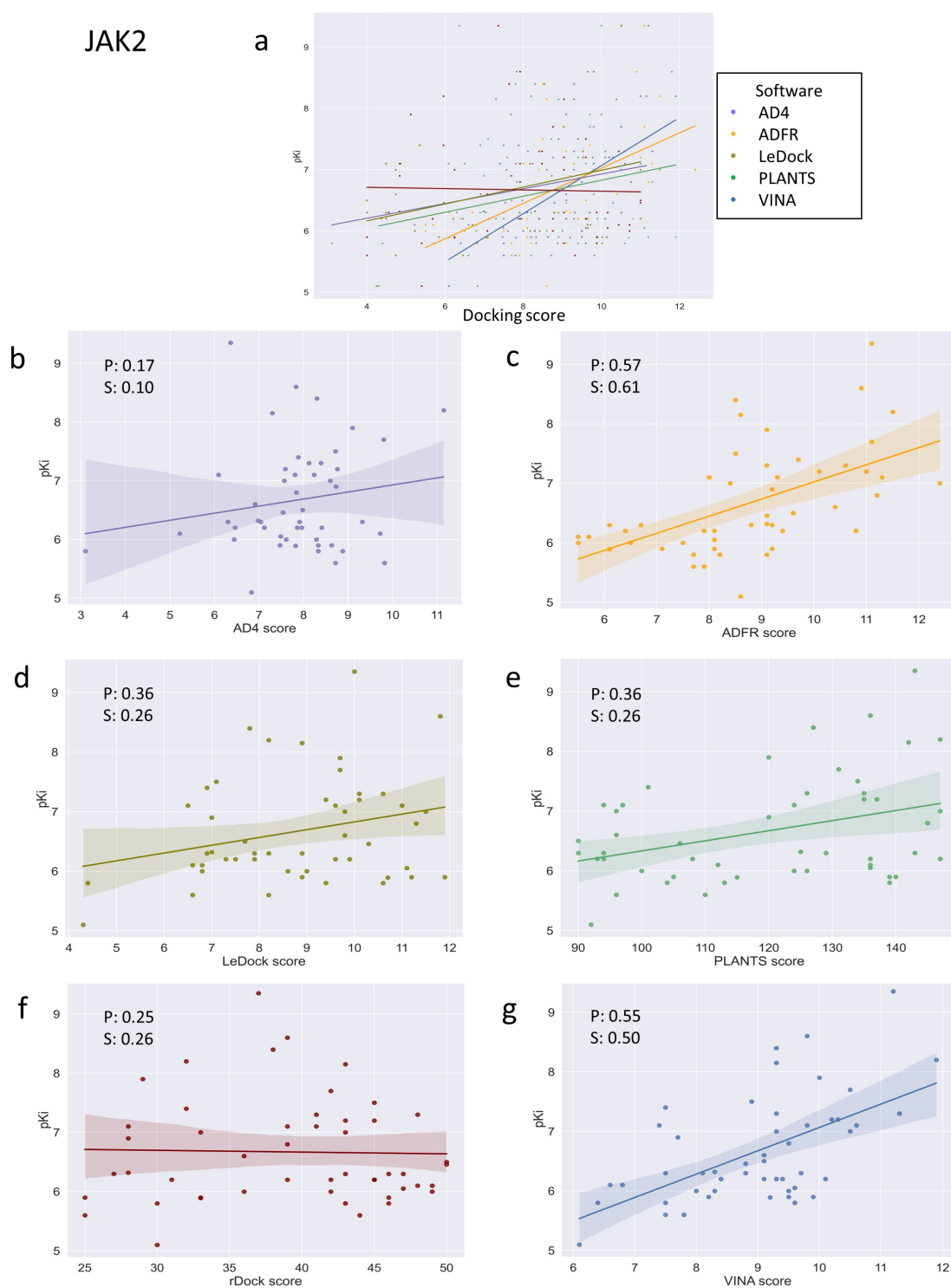


Figure A.2: JAK2 regression plots. Results for binding affinity and docking scores. (a) global, (b) AD4, (c) ADFR, (e) PLANTS, (f) rDock, (g) VINA

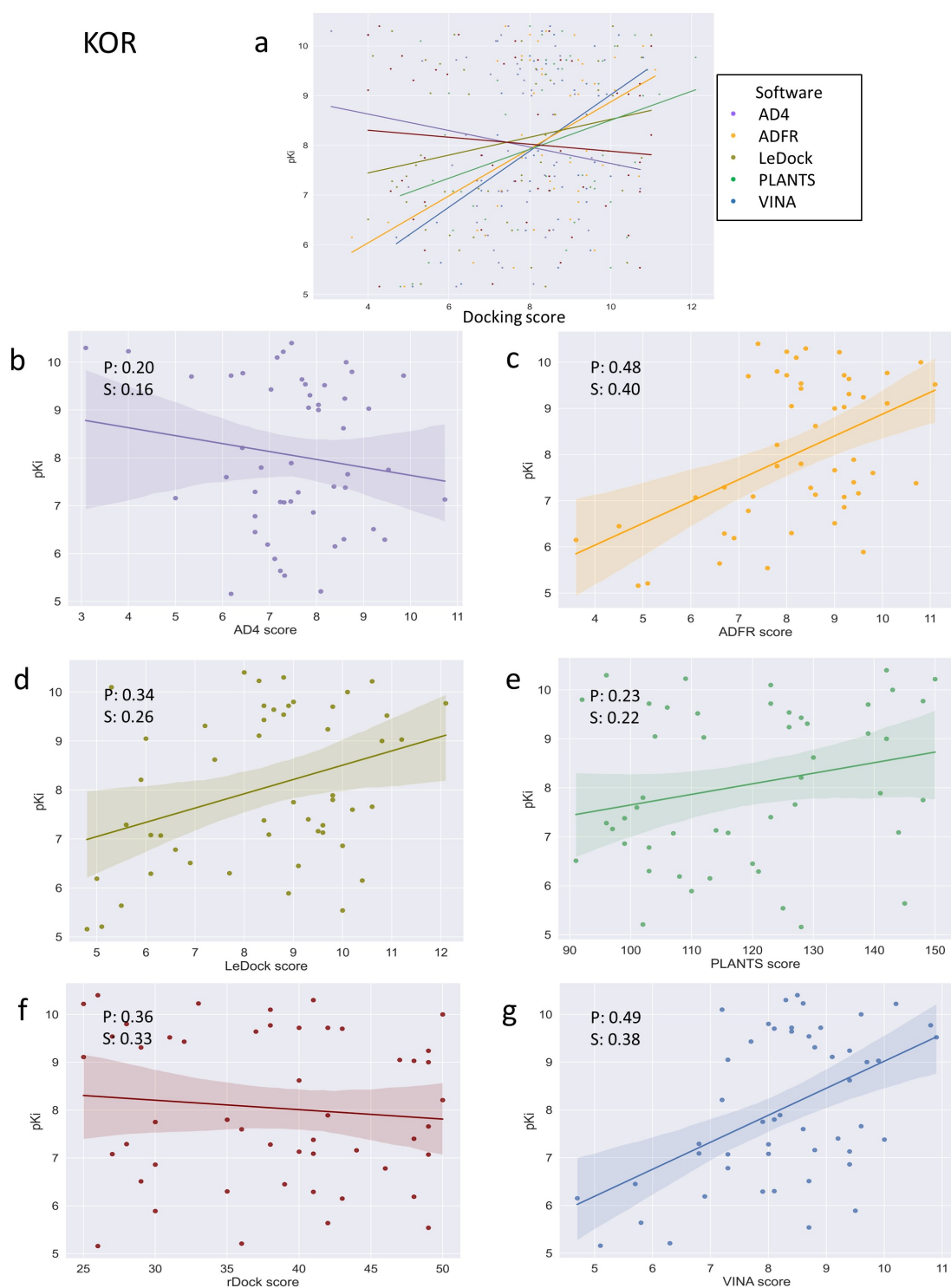


Figure A.3: KOR regression plots. Results for binding affinity and docking scores. (a) global, (b) AD4, (c) ADFR, (e) PLANTS, (f) rDock, (g) VINA

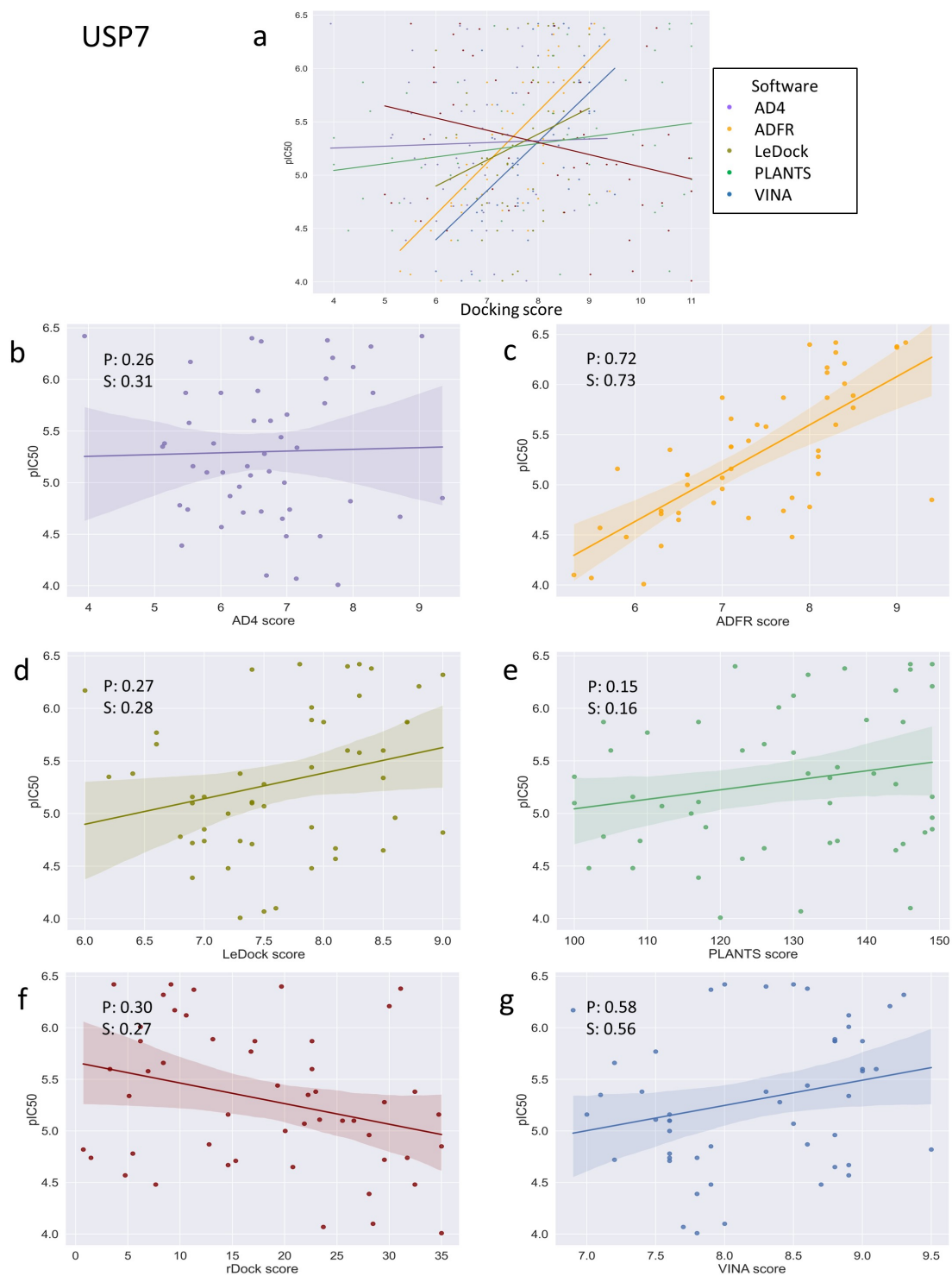


Figure A.4: USP7 regression plots. Results for binding affinity and docking scores. (a) global, (b) AD4, (c) ADFR, (e) PLANTS, (f) rDock, (g) VINA

Appendix C

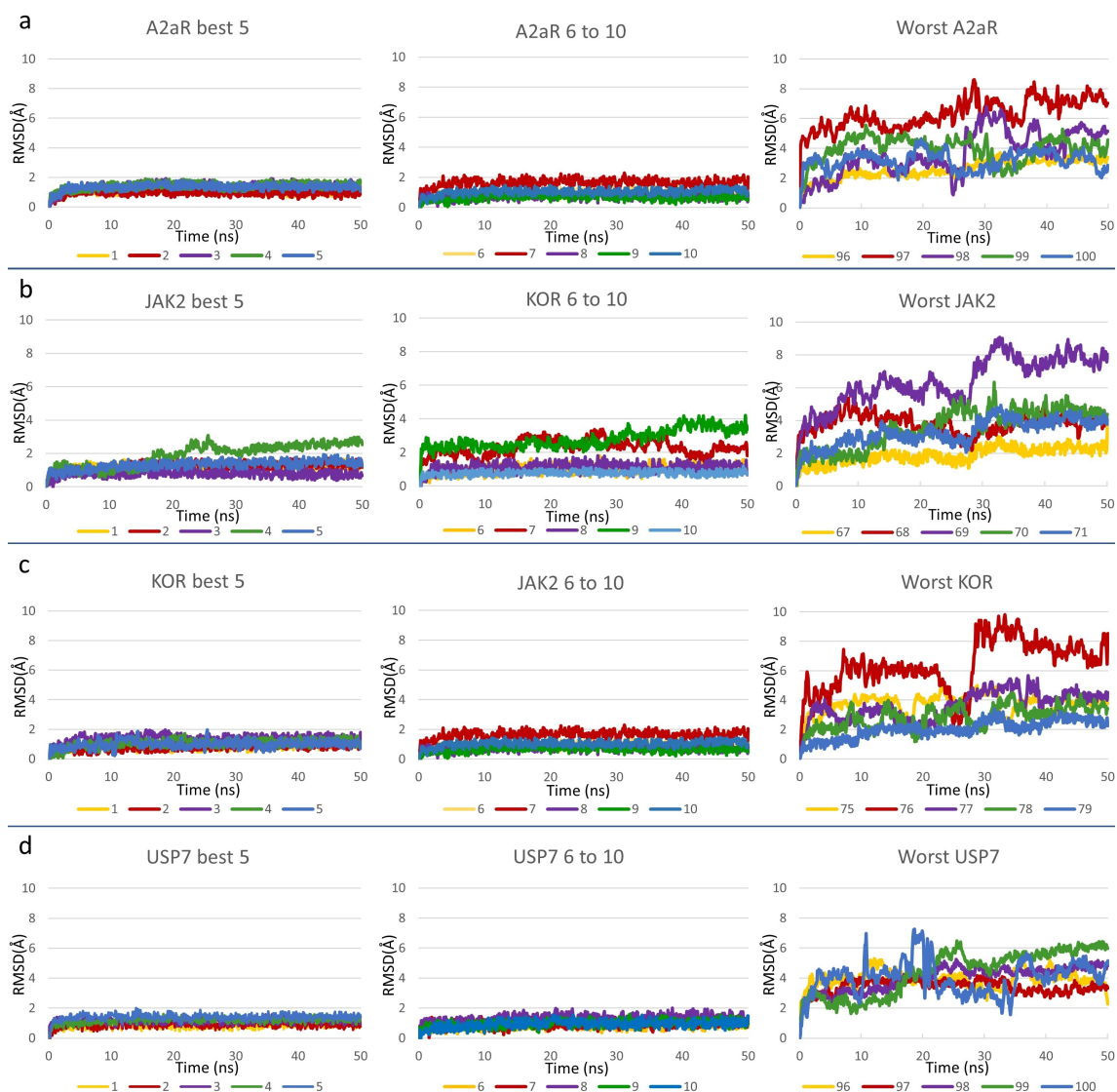


Figure A.5: Ligand stability in MDS. Representation of the positional stability of the best and worst exponential consensus ranking molecules in the MSD experiments. Graphs of the RMSD (y axis) over time (x axis) for the top five, 6th to 10th, and worst five molecules (in relation to target molecule consensus scoring) for (a) A2aR, (b) JAK2, (c) KOR, (d) KOR. Lower values denote ligand positional stability and, therefore, stronger coupling with the receptor.

Bibliography

- [1] E. Landsburg Steven, “The armchair economist: Economics and everyday life,” 2012.
- [2] P. G. Polishchuk, T. I. Madzhidov, and A. Varnek, “Estimation of the size of drug-like chemical space based on gdb-17 data,” *Journal of computer-aided molecular design*, vol. 27, no. 8, pp. 675–679, 2013.
- [3] O. Trott and A. J. Olson, “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, pp. 455–461, 2010.
- [4] B. J. Bender, S. Gahbauer, A. Lutten, J. Lyu, C. M. Webb, R. M. Stein, E. A. Fink, T. E. Balius, J. Carlsson, J. J. Irwin *et al.*, “A practical guide to large-scale docking,” *Nature protocols*, vol. 16, no. 10, pp. 4799–4832, 2021.
- [5] D. B. Kitchen, H. Decornez, J. R. Furr, and J. Bajorath, “Docking and scoring in virtual screening for drug discovery: methods and applications,” *Nature reviews Drug discovery*, vol. 3, no. 11, pp. 935–949, 2004.
- [6] S. Forli, R. Huey, M. E. Pique, M. F. Sanner, D. S. Goodsell, and A. J. Olson, “Computational protein–ligand docking and virtual drug screening with the autodock suite,” *Nature protocols*, vol. 11, no. 5, pp. 905–919, 2016.
- [7] P. Agrawal, H. Singh, H. K. Srivastava, S. Singh, G. Kishore, and G. P. Raghava, “Benchmarking of different molecular docking methods for protein-peptide docking,” *BMC bioinformatics*, vol. 19, no. 13, pp. 105–124, 2019.
- [8] R. Thomsen and M. H. Christensen, “Moldock: a new technique for high-accuracy molecular docking,” *Journal of medicinal chemistry*, vol. 49, no. 11, pp. 3315–3321, 2006.

- [9] E. D. Boittier, Y. Y. Tang, M. E. Buckley, Z. P. Schuurs, D. J. Richard, and N. S. Gandhi, "Assessing molecular docking tools to guide targeted drug discovery of cd38 inhibitors," *International journal of molecular sciences*, vol. 21, no. 15, p. 5183, 2020.
- [10] M. Akhter, "Challenges in docking: Mini review," *JSM chemistry*, 2016.
- [11] K. M. Elokely and R. J. Doerksen, "Docking challenge: protein sampling and molecular docking performance," *Journal of chemical information and modeling*, vol. 53, no. 8, pp. 1934–1945, 2013.
- [12] I. W. Davis, K. Raha, M. S. Head, and D. Baker, "Blind docking of pharmaceutically relevant compounds using rosettaligand," *Protein science*, vol. 18, no. 9, pp. 1998–2002, 2009.
- [13] I. Muegge, "The effect of small changes in protein structure on predicted binding modes of known inhibitors of influenza virus neuraminidase: Pmf-scoring in dock4," *MEDICINAL CHEMISTRY RESEARCH*, vol. 9, no. 7-8, pp. 490–500, 1999.
- [14] M. Schapira, R. Abagyan, and M. Totrov, "Nuclear hormone receptor targeted virtual screening," *Journal of medicinal chemistry*, vol. 46, no. 14, pp. 3045–3059, 2003.
- [15] S. C. Hopkins, R. D. Vale, and I. D. Kuntz, "Inhibitors of kinesin activity from structure-based computer screening," *Biochemistry*, vol. 39, no. 10, pp. 2805–2814, 2000.
- [16] S. L. McGovern and B. K. Shoichet, "Information decay in molecular docking screens against holo, apo, and modeled conformations of enzymes," *Journal of medicinal chemistry*, vol. 46, no. 14, pp. 2895–2907, 2003.
- [17] J. J. Sutherland, R. K. Nandigam, J. A. Erickson, and M. Vieth, "Lessons in molecular recognition. 2. assessing and improving cross-docking accuracy," *Journal of chemical information and modeling*, vol. 47, no. 6, pp. 2293–2302, 2007.
- [18] C. Dequeker, Y. Mohseni Behbahani, L. David, E. Laine, and A. Carbone, "From complete cross-docking to partners identification and binding sites predictions," *PLoS computational biology*, vol. 18, no. 1, p. e1009825, 2022.

-
- [19] M. Dickson and J. P. Gagnon, “Key factors in the rising cost of new drug discovery and development,” *Nature reviews Drug discovery*, vol. 3, no. 5, pp. 417–429, 2004.
- [20] C. Bilodeau, W. Jin, T. Jaakkola, R. Barzilay, and K. F. Jensen, “Generative models for molecular discovery: Recent advances and challenges,” *Wiley Interdisciplinary Reviews: Computational Molecular Science*, p. e1608, 2022.
- [21] G. R. Bickerton, G. V. Paolini, J. Besnard, S. Muresan, and A. L. Hopkins, “Quantifying the chemical beauty of drugs,” *Nature chemistry*, vol. 4, no. 2, pp. 90–98, 2012.
- [22] A. Zhavoronkov, Y. A. Ivanenkov, A. Aliper, M. S. Veselov, V. A. Aladinskiy, A. V. Aladinskaya, V. A. Terentiev, D. A. Polykovskiy, M. D. Kuznetsov, A. Asadulaev *et al.*, “Deep learning enables rapid identification of potent ddr1 kinase inhibitors,” *Nature biotechnology*, vol. 37, no. 9, pp. 1038–1040, 2019.
- [23] D. C. Elton, Z. Boukouvalas, M. D. Fuge, and P. W. Chung, “Deep learning for molecular design—a review of the state of the art,” *Molecular Systems Design & Engineering*, vol. 4, no. 4, pp. 828–849, 2019.
- [24] D. Polykovskiy, A. Zhebrak, B. Sanchez-Lengeling, S. Golovanov, O. Tatanov, S. Belyaev, R. Kurbanov, A. Artamonov, V. Aladinskiy, M. Veselov *et al.*, “Molecular sets (moses): a benchmarking platform for molecular generation models,” *Frontiers in pharmacology*, vol. 11, p. 565644, 2020.
- [25] N. Brown, M. Fiscato, M. H. Segler, and A. C. Vaucher, “Guacamol: benchmarking models for de novo molecular design,” *Journal of chemical information and modeling*, vol. 59, no. 3, pp. 1096–1108, 2019.
- [26] P. Renz, D. Van Rompaey, J. K. Wegner, S. Hochreiter, and G. Klambauer, “On failure modes in molecule generation and optimization,” *Drug Discovery Today: Technologies*, vol. 32, pp. 55–63, 2019.
- [27] L. L. Grant and C. S. Sit, “De novo molecular drug design benchmarking,” *RSC Medicinal Chemistry*, vol. 12, no. 8, pp. 1273–1280, 2021.
- [28] W. Gao and C. W. Coley, “The synthesizability of molecules proposed by generative models,” *Journal of chemical information and modeling*, vol. 60, no. 12, pp. 5714–5723, 2020.

- [29] O. J. Wouters, M. McKee, and J. Luyten, “Estimated research and development investment needed to bring a new medicine to market, 2009-2018,” *Jama*, vol. 323, no. 9, pp. 844–853, 2020.
- [30] T. Takebe, R. Imai, and S. Ono, “The current status of drug discovery and development as originated in united states academia: the influence of industrial and academic collaboration on drug discovery and development,” *Clinical and translational science*, vol. 11, no. 6, pp. 597–606, 2018.
- [31] D. Sun, W. Gao, H. Hu, and S. Zhou, “Why 90% of clinical drug development fails and how to improve it?” *Acta Pharmaceutica Sinica B*, 2022.
- [32] W. Yu and A. D. MacKerell, “Computer-aided drug design methods,” in *Antibiotics*. Springer, 2017, pp. 85–106.
- [33] M. Hassan Baig, K. Ahmad, S. Roy, J. Mohammad Ashraf, M. Adil, M. Haris Siddiqui, S. Khan, M. Amjad Kamal, I. Provazník, and I. Choi, “Computer aided drug design: success and limitations,” *Current pharmaceutical design*, vol. 22, no. 5, pp. 572–581, 2016.
- [34] Y. Gilad, K. Nadassy, and H. Senderowitz, “A reliable computational workflow for the selection of optimal screening libraries,” *Journal of cheminformatics*, vol. 7, no. 1, pp. 1–17, 2015.
- [35] K. T. Luu, E. Kraynov, B. Kuang, P. Vicini, and W.-Z. Zhong, “Modeling, simulation, and translation framework for the preclinical development of monoclonal antibodies,” *The AAPS journal*, vol. 15, no. 2, pp. 551–558, 2013.
- [36] S. D. Satyanarayana, M. Krishna, P. P. Kumar, and S. Jeeretty, “In silico structural homology modeling of nif a protein of rhizobial strains in selective legume plants,” *Journal of Genetic Engineering and Biotechnology*, vol. 16, no. 2, pp. 731–737, 2018.
- [37] Z. S. Hashemi, M. Zarei, M. K. Fath, M. Ganji, M. S. Farahani, F. Afsharnouri, N. Pourzardosht, B. Khalesi, A. Jahangiri, M. R. Rahbar *et al.*, “In silico approaches for the design and optimization of interfering peptides against protein–protein interactions,” *Frontiers in Molecular Biosciences*, vol. 8, p. 282, 2021.
- [38] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko *et al.*, “Highly accu-

- rate protein structure prediction with alphafold,” *Nature*, vol. 596, no. 7873, pp. 583–589, 2021.
- [39] C. Réda, E. Kaufmann, and A. Delahaye-Duriez, “Machine learning applications in drug development,” *Computational and structural biotechnology journal*, vol. 18, pp. 241–252, 2020.
- [40] J. Walker, “Machine learning drug discovery applications—pfizer, roche, gsk, and more,” *Available at: emerj.com/ai-sector-overviews/machine-learning-drug-discovery-applications-pfizer-roche-gsk*. Accessed [September 16, 2019], 2019.
- [41] E. Mahase, “Covid-19: Pfizer’s paxlovid is 89% effective in patients at risk of serious illness, company reports,” 2021.
- [42] L. Da Xu, Y. Lu, and L. Li, “Embedding blockchain technology into iot for security: A survey,” *IEEE Internet of Things Journal*, vol. 8, no. 13, pp. 10 452–10 473, 2021.
- [43] A. L. Guzman, “Beyond extraordinary: Theorizing artificial intelligence and the self in daily life,” in *A Networked Self and Human Augmentics, Artificial Intelligence, Sentience*. Routledge, 2018, pp. 83–96.
- [44] M. Koch, “Artificial intelligence is becoming natural,” *Cell*, vol. 173, no. 3, p. 533, 2018.
- [45] R. S. Lee, *Artificial intelligence in daily life*. Springer, 2020.
- [46] J. L. Blanco, A. B. Porto-Pazos, A. Pazos, and C. Fernandez-Lozano, “Prediction of high anti-angiogenic activity peptides in silico using a generalized linear model and feature selection,” *Scientific Reports*, vol. 8, no. 1, pp. 1–11, 2018.
- [47] C. R. Munteanu, E. Fernández-Blanco, J. A. Seoane, P. Izquierdo-Novo, J. Angel Rodriguez-Fernandez, J. Maria Prieto-Gonzalez, J. R. Rabunal, and A. Pazos, “Drug discovery and design for complex diseases through qsar computational methods,” *Current pharmaceutical design*, vol. 16, no. 24, pp. 2640–2655, 2010.
- [48] I. García, C. R. Munteanu, Y. Fall, G. Gómez, E. Uriarte, and H. González-Díaz, “Qsar and complex network study of the chiral hmgr inhibitor structural diversity,” *Bioorganic & medicinal chemistry*, vol. 17, no. 1, pp. 165–175, 2009.

- [49] Y. Liu, S. Tang, C. Fernandez-Lozano, C. R. Munteanu, A. Pazos, Y.-z. Yu, Z. Tan, and H. González-Díaz, “Experimental study and random forest prediction model of microbiome cell surface hydrophobicity,” *Expert Systems with Applications*, vol. 72, pp. 306–316, 2017.
- [50] P. Riera-Fernandez, C. R. Munteanu, J. Dorado, R. Martin-Romalde, A. Duardo-Sanchez, and H. Gonzalez-Diaz, “From chemical graphs in computer-aided drug design to general markov-galvez indices of drug-target, proteome, drug-parasitic disease, technological, and social-legal networks,” *Current computer-aided drug design*, vol. 7, no. 4, pp. 315–337, 2011.
- [51] P. Shirvani and A. Fassihi, “Molecular modelling study on pyrrolo [2, 3-b] pyridine derivatives as c-met kinase inhibitors: a combined approach using molecular docking, 3d-qsar modelling and molecular dynamics simulation,” *Molecular Simulation*, vol. 46, no. 16, pp. 1265–1280, 2020.
- [52] B. Suay-Garcia, J. I. Bueso-Bordils, A. Falcó, M. T. Pérez-Gracia, G. Antón-Fos, and P. Alemán-López, “Quantitative structure–activity relationship methods in the discovery and development of antibacterials,” *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 10, no. 6, p. e1472, 2020.
- [53] F. S. Collins and H. Varmus, “A new initiative on precision medicine,” *New England journal of medicine*, vol. 372, no. 9, pp. 793–795, 2015.
- [54] M. E. Ozer, P. O. Sarica, and K. Y. Arga, “New machine learning applications to accelerate personalized medicine in breast cancer: rise of the support vector machines,” *Omics: a journal of integrative biology*, vol. 24, no. 5, pp. 241–246, 2020.
- [55] O. Khan, J. H. Badhiwala, G. Grasso, and M. G. Fehlings, “Use of machine learning and artificial intelligence to drive personalized medicine approaches for spine care,” *World neurosurgery*, vol. 140, pp. 512–518, 2020.
- [56] S. Uddin, A. Khan, M. E. Hossain, and M. A. Moni, “Comparing different supervised machine learning algorithms for disease prediction,” *BMC medical informatics and decision making*, vol. 19, no. 1, pp. 1–16, 2019.
- [57] V. Nasteski, “An overview of the supervised machine learning methods,” *Horizons. b*, vol. 4, pp. 51–62, 2017.

-
- [58] M. Z. Alom, T. M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M. S. Nasrin, M. Hasan, B. C. Van Essen, A. A. Awwal, and V. K. Asari, “A state-of-the-art survey on deep learning theory and architectures,” *Electronics*, vol. 8, no. 3, p. 292, 2019.
- [59] J. Powles and H. Hodson, “Google deepmind and healthcare in an age of algorithms,” *Health and technology*, vol. 7, no. 4, pp. 351–367, 2017.
- [60] E. Gibney, “Google secretly tested ai bot,” *Nature*, vol. 541, no. 7636, p. 142, 2017.
- [61] C. M. Bishop, “Neural networks and their applications,” *Review of scientific instruments*, vol. 65, no. 6, pp. 1803–1832, 1994.
- [62] W. Samek, G. Montavon, S. Lapuschkin, C. J. Anders, and K.-R. Müller, “Explaining deep neural networks and beyond: A review of methods and applications,” *Proceedings of the IEEE*, vol. 109, no. 3, pp. 247–278, 2021.
- [63] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [64] V. P. Dwivedi, C. K. Joshi, T. Laurent, Y. Bengio, and X. Bresson, “Benchmarking graph neural networks,” *arXiv preprint arXiv:2003.00982*, 2020.
- [65] B. Sanchez-Lengeling and A. Aspuru-Guzik, “Inverse molecular design using machine learning: Generative models for matter engineering,” *Science*, vol. 361, no. 6400, pp. 360–365, 2018.
- [66] G. Harshvardhan, M. K. Gourisaria, M. Pandey, and S. S. Rautaray, “A comprehensive survey and analysis of generative models in machine learning,” *Computer Science Review*, vol. 38, p. 100285, 2020.
- [67] C. G. Turhan and H. S. Bilge, “Recent trends in deep generative models: a review,” in *2018 3rd International Conference on Computer Science and Engineering (UBMK)*. IEEE, 2018, pp. 574–579.
- [68] T. Pereira, M. Abbasi, B. Ribeiro, and J. P. Arrais, “Diversity oriented deep reinforcement learning for targeted molecule generation,” *Journal of cheminformatics*, vol. 13, no. 1, pp. 1–17, 2021.

- [69] C. Fernandez-Lozano, M. Gestal, C. R. Munteanu, J. Dorado, and A. Pazos, “A methodology for the design of experiments in computational intelligence with multiple regression models,” *PeerJ*, vol. 4, p. e2721, 2016.
- [70] M. E. Burton, *Applied pharmacokinetics & pharmacodynamics: principles of therapeutic drug monitoring*. Lippincott Williams & Wilkins, 2006.
- [71] T. Sterling and J. J. Irwin, “Zinc 15–ligand discovery for everyone,” *Journal of chemical information and modeling*, vol. 55, no. 11, pp. 2324–2337, 2015.
- [72] A. Gaulton, L. J. Bellis, A. P. Bento, J. Chambers, M. Davies, A. Hersey, Y. Light, S. McGlinchey, D. Michalovich, B. Al-Lazikani *et al.*, “ChEMBL: a large-scale bioactivity database for drug discovery,” *Nucleic acids research*, vol. 40, no. D1, pp. D1100–D1107, 2012.
- [73] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu *et al.*, “Pubchem 2019 update: improved access to chemical data,” *Nucleic acids research*, vol. 47, no. D1, pp. D1102–D1109, 2019.
- [74] P. Carracedo-Reboredo, J. Liñares-Blanco, N. Rodríguez-Fernández, F. Cedrón, F. J. Novoa, A. Carballal, V. Maojo, A. Pazos, and C. Fernandez-Lozano, “A review on machine learning approaches and trends in drug discovery,” *Computational and Structural Biotechnology Journal*, vol. 19, pp. 4538–4558, 2021.
- [75] P. Gautam, A. Jaiswal, T. Aittokallio, H. Al-Ali, and K. Wennerberg, “Phenotypic screening combined with machine learning for efficient identification of breast cancer-selective therapeutic targets,” *Cell chemical biology*, vol. 26, no. 7, pp. 970–979, 2019.
- [76] J. Meyers, B. Fabian, and N. Brown, “De novo molecular design and generative models,” *Drug Discovery Today*, vol. 26, no. 11, pp. 2707–2715, 2021.
- [77] P. Gramatica and A. Sangion, “A historical excursus on the statistical validation parameters for qsar models: a clarification concerning metrics and terminology,” *Journal of chemical information and modeling*, vol. 56, no. 6, pp. 1127–1131, 2016.

-
- [78] D. Janz, J. van der Westhuizen, B. Paige, M. J. Kusner, and J. M. Hernández-Lobato, “Learning a generative model for validity in complex discrete structures,” *arXiv preprint arXiv:1712.01664*, 2017.
- [79] Z. Chen, M. R. Min, S. Parthasarathy, and X. Ning, “A deep generative model for molecule optimization via one fragment modification,” *Nature Machine Intelligence*, vol. 3, no. 12, pp. 1040–1049, 2021.
- [80] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M. A. Fadhel, M. Al-Amidie, and L. Farhan, “Review of deep learning: Concepts, cnn architectures, challenges, applications, future directions,” *Journal of big Data*, vol. 8, no. 1, pp. 1–74, 2021.
- [81] A. Giusti, D. C. Cireşan, J. Masci, L. M. Gambardella, and J. Schmidhuber, “Fast image scanning with deep max-pooling convolutional neural networks,” in *2013 IEEE International Conference on Image Processing*. IEEE, 2013, pp. 4034–4038.
- [82] G. E. Dahl, N. Jaitly, and R. Salakhutdinov, “Multi-task neural networks for qsar predictions,” *arXiv preprint arXiv:1406.1231*, 2014.
- [83] D. P. Kingma and M. Welling, “Auto-encoding variational bayes,” *arXiv preprint arXiv:1312.6114*, 2013.
- [84] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial networks,” *arXiv preprint arXiv:1406.2661*, 2014.
- [85] R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams, and A. Aspuru-Guzik, “Automatic chemical design using a data-driven continuous representation of molecules,” *ACS central science*, vol. 4, no. 2, pp. 268–276, 2018.
- [86] Y. Bian and X.-Q. Xie, “Generative chemistry: drug discovery with deep learning generative models,” *Journal of Molecular Modeling*, vol. 27, no. 3, pp. 1–18, 2021.
- [87] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009, pp. 248–255.

- [88] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [89] Z. Wu, B. Ramsundar, E. N. Feinberg, J. Gomes, C. Geniesse, A. S. Pappu, K. Leswing, and V. Pande, "Moleculenet: a benchmark for molecular machine learning," *Chemical science*, vol. 9, no. 2, pp. 513–530, 2018.
- [90] A. Jain, M. Heinonen, and S. Kaski, "Multi-target optimization for drug discovery using generative models."
- [91] J. Boitreau, V. Mallet, C. Oliver, and J. Waldispuhl, "Optimol: optimization of binding affinities in chemical space for drug discovery," *Journal of Chemical Information and Modeling*, vol. 60, no. 12, pp. 5658–5666, 2020.
- [92] D. Sculley, J. Snoek, A. Wiltschko, and A. Rahimi, "Winner's curse? on pace, progress, and empirical rigor," 2018.
- [93] F. Ahmadi Moughari and C. Eslahchi, "A computational method for drug sensitivity prediction of cancer cell lines based on various molecular information," *PloS one*, vol. 16, no. 4, p. e0250620, 2021.
- [94] G. Adam, L. Rampásek, Z. Safikhani, P. Smirnov, B. Haibe-Kains, and A. Goldenberg, "Machine learning approaches to drug response prediction: challenges and recent progress," *NPJ precision oncology*, vol. 4, no. 1, pp. 1–10, 2020.
- [95] F. Firoozbakht, B. Yousefi, and B. Schwikowski, "An overview of machine learning methods for monotherapy drug response prediction," *Briefings in Bioinformatics*, vol. 23, no. 1, p. bbab408, 2022.
- [96] M. Joo, A. Park, K. Kim, W.-J. Son, H. S. Lee, G. Lim, J. Lee, D. H. Lee, J. An, J. H. Kim *et al.*, "A deep learning model for cell growth inhibition ic50 prediction and its application for gastric cancer patients," *International journal of molecular sciences*, vol. 20, no. 24, p. 6276, 2019.
- [97] D. Chen, N. Oezguen, P. Urvil, C. Ferguson, S. M. Dann, and T. C. Savidge, "Regulation of protein-ligand binding affinity by hydrogen bond pairing," *Science advances*, vol. 2, no. 3, p. e1501240, 2016.
- [98] M. Gurusaran, P. Sivarajan, K. Dinesh Kumar, P. Radha, K. Thulaa Tharshan, S. Satheesh, K. Jayanthan, R. Ilaiyaraja, J. Mohanapriya, D. Michael

- et al.*, “Hydrogen bonds computing server (hbcs): An online web server to compute hydrogen-bond interactions and their precision,” *Journal of Applied Crystallography*, vol. 49, no. 2, pp. 642–645, 2016.
- [99] S. Zhou, Y. Liu, S. Wang, and L. Wang, “Effective prediction of short hydrogen bonds in proteins via machine learning method,” *Scientific reports*, vol. 12, no. 1, pp. 1–10, 2022.
- [100] J. D. Chodera and D. L. Mobley, “Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design,” *Annual review of biophysics*, vol. 42, p. 121, 2013.
- [101] V. Lafont, A. A. Armstrong, H. Ohtaka, Y. Kiso, L. Mario Amzel, and E. Freire, “Compensating enthalpic and entropic changes hinder binding affinity optimization,” *Chemical biology & drug design*, vol. 69, no. 6, pp. 413–422, 2007.
- [102] F. Imrie, T. E. Hadfield, A. R. Bradley, and C. M. Deane, “Deep generative design with 3d pharmacophoric constraints,” *Chemical science*, vol. 12, no. 43, pp. 14 577–14 589, 2021.
- [103] K. Papadopoulos, K. A. Giblin, J. P. Janet, A. Patronov, and O. Engkvist, “De novo design with deep generative models based on 3d similarity scoring,” *Bioorganic & Medicinal Chemistry*, vol. 44, p. 116308, 2021.
- [104] E. Kangas and B. Tidor, “Optimizing electrostatic affinity in ligand–receptor binding: Theory, computation, and ligand properties,” *The Journal of chemical physics*, vol. 109, no. 17, pp. 7522–7545, 1998.
- [105] N. J. Ataie, Q. Q. Hoang, M. P. Zahniser, Y. Tu, A. Milne, G. A. Petsko, and D. Ringe, “Zinc coordination geometry and ligand binding affinity: the structural and kinetic analysis of the second-shell serine 228 residue and the methionine 180 residue of the aminopeptidase from *Vibrio proteolyticus*,” *Biochemistry*, vol. 47, no. 29, pp. 7673–7683, 2008.
- [106] D. D. Wang, L. Ou-Yang, H. Xie, M. Zhu, and H. Yan, “Predicting the impacts of mutations on protein-ligand binding affinity based on molecular dynamics simulations and machine learning methods,” *Computational and structural biotechnology journal*, vol. 18, pp. 439–454, 2020.

- [107] A. K. Bronowska, “Thermodynamics of ligand-protein interactions: implications for molecular design,” in *Thermodynamics-Interaction Studies-Solids, Liquids and Gases*. IntechOpen, 2011.
- [108] J. Kostal, “Computational chemistry in predictive toxicology: status quo et quo vadis?” in *Advances in molecular toxicology*. Elsevier, 2016, vol. 10, pp. 139–186.
- [109] M. Brand, J. Clayton, M. Moroglu, M. Schiedel, S. Picaud, J. P. Bluck, A. Skwarska, H. Bolland, A. K. Chan, C. M. Laurin *et al.*, “Controlling intramolecular interactions in the design of selective, high-affinity ligands for the crebbp bromodomain,” *Journal of medicinal chemistry*, vol. 64, no. 14, pp. 10 102–10 123, 2021.
- [110] S. Raschka, A. J. Wolf, J. Bemister-Buffington, and L. A. Kuhn, “Protein–ligand interfaces are polarized: discovery of a strong trend for intermolecular hydrogen bonds to favor donors on the protein side with implications for predicting and designing ligand complexes,” *Journal of computer-aided molecular design*, vol. 32, no. 4, pp. 511–528, 2018.
- [111] S. Nickerson, S. T. Joy, P. S. Arora, and D. Bar-Sagi, “An orthosteric inhibitor of the ras–sos interaction,” *The Enzymes*, vol. 34, pp. 25–39, 2013.
- [112] I. Kufareva and R. Abagyan, “Methods of protein structure comparison,” in *Homology modeling*. Springer, 2011, pp. 231–257.
- [113] U. Weininger, K. Modig, and M. Akke, “Ring flips revisited: ^{13}C relaxation dispersion measurements of aromatic side chain dynamics and activation barriers in basic pancreatic trypsin inhibitor,” *Biochemistry*, vol. 53, no. 28, pp. 4519–4525, 2014.
- [114] A. N. Jain, A. E. Cleves, Q. Gao, X. Wang, Y. Liu, E. C. Sherer, and M. Y. Reibarkh, “Complex macrocycle exploration: parallel, heuristic, and constraint-based conformer generation using forcegen,” *Journal of computer-aided molecular design*, vol. 33, no. 6, pp. 531–558, 2019.
- [115] L. G. Ferreira, R. N. Dos Santos, G. Oliva, and A. D. Andricopulo, “Molecular docking and structure-based drug design strategies,” *Molecules*, vol. 20, no. 7, pp. 13 384–13 421, 2015.

- [116] Z. Wang, H. Sun, X. Yao, D. Li, L. Xu, Y. Li, S. Tian, and T. Hou, “Comprehensive evaluation of ten docking programs on a diverse set of protein–ligand complexes: the prediction accuracy of sampling power and scoring power,” *Physical Chemistry Chemical Physics*, vol. 18, no. 18, pp. 12 964–12 975, 2016.
- [117] W. Kabsch, “A discussion of the solution for the best rotation to relate two sets of vectors,” *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography*, vol. 34, no. 5, pp. 827–828, 1978.
- [118] ———, “A solution for the best rotation to relate two sets of vectors,” *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography*, vol. 32, no. 5, pp. 922–923, 1976.
- [119] A. Sethi, K. Joshi, K. Sasikala, and M. Alvala, “Molecular docking in modern drug discovery: Principles and recent applications,” *Drug discovery and development-new advances*, vol. 2, pp. 1–21, 2019.
- [120] S. Saikia and M. Bordoloi, “Molecular docking: challenges, advances and its use in drug discovery perspective,” *Current drug targets*, vol. 20, no. 5, pp. 501–521, 2019.
- [121] L. Banchi, M. Fingerhuth, T. Babej, C. Ing, and J. M. Arrazola, “Molecular docking with gaussian boson sampling,” *Science advances*, vol. 6, no. 23, p. eaax1950, 2020.
- [122] C. Gardiner, *Stochastic methods*. Springer Berlin, 2009, vol. 4.
- [123] X.-Y. Meng, H.-X. Zhang, M. Mezei, and M. Cui, “Molecular docking: a powerful approach for structure-based drug discovery,” *Current computer-aided drug design*, vol. 7, no. 2, pp. 146–157, 2011.
- [124] G. Jones, P. Willett, R. C. Glen, A. R. Leach, and R. Taylor, “Development and validation of a genetic algorithm for flexible docking,” *Journal of molecular biology*, vol. 267, no. 3, pp. 727–748, 1997.
- [125] M. Su, Q. Yang, Y. Du, G. Feng, Z. Liu, Y. Li, and R. Wang, “Comparative assessment of scoring functions: the casf-2016 update,” *Journal of chemical information and modeling*, vol. 59, no. 2, pp. 895–913, 2018.
- [126] S. Sarfaraz, I. Muneer, and H. Liu, “Combining fragment docking with graph theory to improve ligand docking for homology model structures,” *Journal of computer-aided molecular design*, vol. 34, no. 12, pp. 1237–1259, 2020.

- [127] R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry *et al.*, “Glide: a new approach for rapid, accurate docking and scoring. 1. method and assessment of docking accuracy,” *Journal of medicinal chemistry*, vol. 47, no. 7, pp. 1739–1749, 2004.
- [128] T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard, and J. L. Banks, “Glide: a new approach for rapid, accurate docking and scoring. 2. enrichment factors in database screening,” *Journal of medicinal chemistry*, vol. 47, no. 7, pp. 1750–1759, 2004.
- [129] M. L. Verdonk, J. C. Cole, M. J. Hartshorn, C. W. Murray, and R. D. Taylor, “Improved protein–ligand docking using gold,” *Proteins: Structure, Function, and Bioinformatics*, vol. 52, no. 4, pp. 609–623, 2003.
- [130] O. Korb, T. Stutzle, and T. E. Exner, “Empirical scoring functions for advanced protein–ligand docking with plants,” *Journal of chemical information and modeling*, vol. 49, no. 1, pp. 84–96, 2009.
- [131] O. Sperandio, M. A. Miteva, F. Delfaud, and B. O. Villoutreix, “Receptor-based computational screening of compound databases: the main docking-scoring engines,” *Current Protein and Peptide Science*, vol. 7, no. 5, pp. 369–393, 2006.
- [132] J. Li, A. Fu, and L. Zhang, “An overview of scoring functions used for protein–ligand interactions in molecular docking,” *Interdisciplinary Sciences: Computational Life Sciences*, vol. 11, no. 2, pp. 320–328, 2019.
- [133] H. F. Velec, H. Gohlke, and G. Klebe, “Drugscorecsd knowledge-based scoring function derived from small molecule crystal data with superior recognition rate of near-native ligand poses and better affinity prediction,” *Journal of medicinal chemistry*, vol. 48, no. 20, pp. 6296–6303, 2005.
- [134] F. Stanzione, I. Giangreco, and J. C. Cole, “Use of molecular docking computational tools in drug discovery,” *Progress in Medicinal Chemistry*, vol. 60, pp. 273–343, 2021.
- [135] M. A. Khamis, W. Gomaa, and W. F. Ahmed, “Machine learning in computational docking,” *Artificial intelligence in medicine*, vol. 63, no. 3, pp. 135–152, 2015.

-
- [136] M. Wójcikowski, P. J. Ballester, and P. Siedlecki, “Performance of machine-learning scoring functions in structure-based virtual screening,” *Scientific Reports*, vol. 7, no. 1, pp. 1–10, 2017.
- [137] Q. U. Ain, A. Aleksandrova, F. D. Roessler, and P. J. Ballester, “Machine-learning scoring functions to improve structure-based binding affinity prediction and virtual screening,” *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 5, no. 6, pp. 405–424, 2015.
- [138] V. Scardino, M. Bollini, and C. N. Cavasotto, “Combination of pose and rank consensus in docking-based virtual screening: the best of both worlds,” *RSC Advances*, vol. 11, no. 56, pp. 35 383–35 391, 2021.
- [139] M. A. Llanos, M. E. Gantner, S. Rodriguez, L. N. Alberca, C. L. Bellera, A. Talevi, and L. Gavernet, “Strengths and weaknesses of docking simulations in the sars-cov-2 era: the main protease (mpro) case study,” *Journal of Chemical Information and Modeling*, vol. 61, no. 8, pp. 3758–3770, 2021.
- [140] G. Poli, A. Martinelli, and T. Tuccinardi, “Reliability analysis and optimization of the consensus docking approach for the development of virtual screening studies,” *Journal of Enzyme Inhibition and Medicinal Chemistry*, vol. 31, no. sup2, pp. 167–173, 2016.
- [141] M. M. Mysinger, M. Carchia, J. J. Irwin, and B. K. Shoichet, “Directory of useful decoys, enhanced (dud-e): better ligands and decoys for better benchmarking,” *Journal of medicinal chemistry*, vol. 55, no. 14, pp. 6582–6594, 2012.
- [142] R. M. Stein, Y. Yang, T. E. Balius, M. J. O’Meara, J. Lyu, J. Young, K. Tang, B. K. Shoichet, and J. J. Irwin, “Property-unmatched decoys in docking benchmarks,” *Journal of Chemical Information and Modeling*, vol. 61, no. 2, pp. 699–714, 2021.
- [143] K. Palacio-Rodríguez, I. Lans, C. N. Cavasotto, and P. Cossio, “Exponential consensus ranking improves the outcome in docking and receptor ensemble docking,” *Scientific reports*, vol. 9, no. 1, pp. 1–14, 2019.
- [144] I. Lans, E. Anoz-Carbonell, K. Palacio-Rodríguez, J. A. Aínsa, M. Medina, and P. Cossio, “In silico discovery and biological validation of ligands of fad synthase, a promising new antimicrobial target,” *PLoS computational biology*, vol. 16, no. 8, p. e1007898, 2020.

- [145] S. S. Ericksen, H. Wu, H. Zhang, L. A. Michael, M. A. Newton, F. M. Hoffmann, and S. A. Wildman, “Machine learning consensus scoring improves performance across targets in structure-based virtual screening,” *Journal of chemical information and modeling*, vol. 57, no. 7, pp. 1579–1590, 2017.
- [146] R. D. Gupta and D. Kundu, “Generalized exponential distribution: different method of estimations,” *Journal of Statistical Computation and Simulation*, vol. 69, no. 4, pp. 315–337, 2001.
- [147] M. González, “Force fields and molecular dynamics simulations,” *École thématique de la Société Française de la Neutronique*, vol. 12, pp. 169–200, 2011.
- [148] S. Chmiela, H. E. Sauceda, K.-R. Müller, and A. Tkatchenko, “Towards exact molecular dynamics simulations with machine-learned force fields,” *Nature communications*, vol. 9, no. 1, pp. 1–10, 2018.
- [149] F. Martín-García, E. Papaleo, P. Gomez-Puertas, W. Boomsma, and K. Lindorff-Larsen, “Comparing molecular dynamics force fields in the essential subspace,” *PLoS One*, vol. 10, no. 3, p. e0121114, 2015.
- [150] J. A. Harrison, J. D. Schall, S. Maskey, P. T. Mikulski, M. T. Knippenberg, and B. H. Morrow, “Review of force fields and intermolecular potentials used in atomistic computational materials research,” *Applied Physics Reviews*, vol. 5, no. 3, p. 031104, 2018.
- [151] K. Lindorff-Larsen, P. Maragakis, S. Piana, M. P. Eastwood, R. O. Dror, and D. E. Shaw, “Systematic validation of protein force fields against experimental data,” *PloS one*, vol. 7, no. 2, p. e32131, 2012.
- [152] D. A. Luzik, O. N. Rogacheva, S. A. Izmailov, M. I. Indeykina, A. S. Kononikhin, and N. R. Skrynnikov, “Molecular dynamics model of peptide-protein conjugation: Case study of covalent complex between sos1 peptide and n-terminal sh3 domain from grb2,” *Scientific reports*, vol. 9, no. 1, pp. 1–18, 2019.
- [153] S. Ahmadi, L. Barrios Herrera, M. Chehelamirani, J. Hostaš, S. Jalife, and D. R. Salahub, “Multiscale modeling of enzymes: Qm-cluster, qm/mm, and qm/mm/md: a tutorial review,” *International Journal of Quantum Chemistry*, vol. 118, no. 9, p. e25558, 2018.

-
- [154] P. Robustelli, S. Piana, and D. E. Shaw, “Developing a molecular dynamics force field for both folded and disordered protein states,” *Proceedings of the National Academy of Sciences*, vol. 115, no. 21, pp. E4758–E4766, 2018.
- [155] J. Glaser, T. D. Nguyen, J. A. Anderson, P. Lui, F. Spiga, J. A. Millan, D. C. Morse, and S. C. Glotzer, “Strong scaling of general-purpose molecular dynamics simulations on gpus,” *Computer Physics Communications*, vol. 192, pp. 97–107, 2015.
- [156] F. A. Opo, M. M. Rahman, F. Ahammad, I. Ahmed, M. A. Bhuiyan, and A. M. Asiri, “Structure based pharmacophore modeling, virtual screening, molecular docking and admet approaches for identification of natural anti-cancer agents targeting xiap protein,” *Scientific reports*, vol. 11, no. 1, pp. 1–17, 2021.
- [157] C. McInnes, “Virtual screening strategies in drug discovery,” *Current opinion in chemical biology*, vol. 11, no. 5, pp. 494–502, 2007.
- [158] P. Badrinarayan and G. Narahari Sastry, “Virtual high throughput screening in new lead identification,” *Combinatorial chemistry & high throughput screening*, vol. 14, no. 10, pp. 840–860, 2011.
- [159] M. Seifert, J. Kraus, and B. Kramer, “Virtual high-throughput screening of molecular databases.” *Current opinion in drug discovery & development*, vol. 10, no. 3, pp. 298–307, 2007.
- [160] Ö. H. Omar, M. Del Cueto, T. Nemataram, and A. Troisi, “High-throughput virtual screening for organic electronics: a comparative study of alternative strategies,” *Journal of Materials Chemistry C*, vol. 9, no. 39, pp. 13 557–13 583, 2021.
- [161] A. Barker and D. Andrews, “Discovery and development of the anticancer agent gefitinib, an inhibitor of the epidermal growth factor receptor tyrosine kinase,” in *Introduction to Biological and Small Molecule Drug Research and Development*. Elsevier, 2013, pp. 255–281.
- [162] A. Wlodawer and J. Vondrasek, “Inhibitors of hiv-1 protease: a major success of structure-assisted drug design,” *Annual review of biophysics and biomolecular structure*, vol. 27, no. 1, pp. 249–284, 1998.
- [163] S. Dadashpour, T. Tuylu Kucukkilinc, O. Unsal Tan, K. Ozadali, H. Irannejad, and S. Emami, “Design, synthesis and in vitro study of 5, 6-diaryl-1,

- 2, 4-triazine-3-ylthioacetate derivatives as cox-2 and β -amyloid aggregation inhibitors,” *Archiv der Pharmazie*, vol. 348, no. 3, pp. 179–187, 2015.
- [164] M. Batool, B. Ahmad, and S. Choi, “A structure-based drug discovery paradigm,” *International journal of molecular sciences*, vol. 20, no. 11, p. 2783, 2019.
- [165] A. Varela-Rial, M. Majewski, and G. De Fabritiis, “Structure based virtual screening: Fast and slow,” *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 12, no. 2, p. e1544, 2022.
- [166] F. Wang, F.-X. Wu, C.-Z. Li, C.-Y. Jia, S.-W. Su, G.-F. Hao, and G.-F. Yang, “Acid: a free tool for drug repurposing using consensus inverse docking strategy,” *Journal of cheminformatics*, vol. 11, no. 1, pp. 1–11, 2019.
- [167] T. J. Ewing and I. D. Kuntz, “Critical evaluation of search algorithms for automated molecular docking and database screening,” *Journal of computational chemistry*, vol. 18, no. 9, pp. 1175–1189, 1997.
- [168] F. D. Prieto-Martínez, M. Arciniega, and J. L. Medina-Franco, “Molecular docking: current advances and challenges,” *TIP. Revista especializada en ciencias químico-biológicas*, vol. 21, 2018.
- [169] N. A. Murugan, A. Podobas, D. Gadioli, E. Vitali, G. Palermo, and S. Markidis, “A review on parallel virtual screening softwares for high-performance computers,” *Pharmaceuticals*, vol. 15, no. 1, p. 63, 2022.
- [170] R. Teramoto and H. Fukunishi, “Supervised consensus scoring for docking and virtual screening,” *Journal of chemical information and modeling*, vol. 47, no. 2, pp. 526–534, 2007.
- [171] D.-D. Li, X.-F. Meng, Q. Wang, P. Yu, L.-G. Zhao, Z.-P. Zhang, Z.-Z. Wang, and W. Xiao, “Consensus scoring model for the molecular docking study of mtor kinase inhibitor,” *Journal of Molecular Graphics and Modelling*, vol. 79, pp. 81–87, 2018.
- [172] A. Dhakal, C. McKay, J. J. Tanner, and J. Cheng, “Artificial intelligence in the prediction of protein–ligand interactions: recent advances and future directions,” *Briefings in Bioinformatics*, vol. 23, no. 1, p. bbab476, 2022.
- [173] V. V. Gurevich and E. V. Gurevich, “Gpcr signaling regulation: the role of grks and arrestins,” *Frontiers in pharmacology*, vol. 10, p. 125, 2019.

-
- [174] D. M. Rosenbaum, S. G. Rasmussen, and B. K. Kobilka, “The structure and function of g-protein-coupled receptors,” *Nature*, vol. 459, no. 7245, pp. 356–363, 2009.
- [175] N. R. Latorraca, A. Venkatakrisnan, and R. O. Dror, “Gpcr dynamics: structures in motion,” *Chemical reviews*, vol. 117, no. 1, pp. 139–155, 2017.
- [176] B. Li, R. K. Rampal, and Z. Xiao, “Targeted therapies for myeloproliferative neoplasms,” *Biomarker Research*, vol. 7, no. 1, p. 1–8, 2019.
- [177] A. Dhaliwal and M. Gupta, “Physiology, opioid receptor,” 2019.
- [178] C. W. Stevens, “The evolution of vertebrate opioid receptors,” *Frontiers in bioscience: a journal and virtual library*, vol. 14, p. 1247, 2009.
- [179] T. Che, S. Majumdar, S. A. Zaidi, P. Ondachi, J. D. McCorvy, S. Wang, P. D. Mosier, R. Uprety, E. Vardy, B. E. Krumm *et al.*, “Structure of the nanobody-stabilized active state of the kappa opioid receptor,” *Cell*, vol. 172, no. 1-2, pp. 55–67, 2018.
- [180] A. Mafi, S.-K. Kim, and W. A. Goddard, “The atomistic level structure for the activated human κ -opioid receptor bound to the full gi protein and the mp1104 agonist,” *Proceedings of the National Academy of Sciences*, vol. 117, no. 11, pp. 5836–5843, 2020.
- [181] C. O’Connor, K. L. White, N. Doncescu, T. Didenko, B. L. Roth, G. Czaplicki, R. C. Stevens, K. Wüthrich, and A. Milon, “Nmr structure and dynamics of the agonist dynorphin peptide bound to the human kappa opioid receptor,” *Proceedings of the National Academy of Sciences*, vol. 112, no. 38, pp. 11 852–11 857, 2015.
- [182] E. N. Brown, K. J. Pavone, and M. Naranjo, “Multimodal general anesthesia: theory and practice,” *Anesthesia and analgesia*, vol. 127, no. 5, p. 1246, 2018.
- [183] S. Brill, G. M. Gurman, and A. Fisher, “A history of neuraxial administration of local analgesics and opioids,” *European Journal of Anaesthesiology*, vol. 20, no. 9, p. 682–689, 2003.
- [184] J. V. Pergolizzi, J. A. LeQuang, G. K. Berger, and R. B. Raffa, “The basic pharmacology of opioids informs the opioid discourse about misuse and abuse: a review,” *Pain and therapy*, vol. 6, no. 1, pp. 1–16, 2017.

- [185] S. Kerrigan and B. A. Goldberger, “Opioids,” in *Principles of forensic toxicology*. Springer, 2020, pp. 347–369.
- [186] W. H. Organization *et al.*, “International classification of diseases for mortality and morbidity statistics (11th revision),” 2021. [Online]. Available: <https://www.who.int/standards/classifications/classification-of-diseases>
- [187] “World drug report 2021.” [Online]. Available: <https://www.unodc.org/unodc/data-and-analysis/wdr2021.html>
- [188] A. Osborn, “Russia declares “total war” on the country’s drug problem,” 2011.
- [189] T. Seyler, I. Giraudon, A. Noor, J. Mounteney, and P. Griffiths, “Is europe facing an opioid epidemic: What does european monitoring data tell us?” *European Journal of Pain*, vol. 25, no. 5, pp. 1072–1080, 2021.
- [190] M. Tyndall, “An emergency response to the opioid overdose crisis in canada: a regulated opioid distribution program,” *Cmaj*, vol. 190, no. 2, pp. E35–E36, 2018.
- [191] C. Chavkin and G. F. Koob, “Dynorphin, dysphoria, and dependence: the stress of addiction,” *Neuropsychopharmacology*, vol. 41, no. 1, p. 373, 2016.
- [192] R. Al-Hasani and M. R. Bruchas, “Molecular mechanisms of opioid receptor-dependent signaling and behavior,” *The Journal of the American Society of Anesthesiologists*, vol. 115, no. 6, pp. 1363–1381, 2011.
- [193] M. R. Bruchas and B. L. Roth, “New technologies for elucidating opioid receptor function,” *Trends in pharmacological sciences*, vol. 37, no. 4, pp. 279–289, 2016.
- [194] A. P. Kardon, E. Polgár, J. Hachisuka, L. M. Snyder, D. Cameron, S. Savage, X. Cai, S. Karnup, C. R. Fan, G. M. Hemenway *et al.*, “Dynorphin acts as a neuromodulator to inhibit itch in the dorsal horn of the spinal cord,” *Neuron*, vol. 82, no. 3, pp. 573–586, 2014.
- [195] A. Cowan, G. B. Kehner, and S. Inan, “Targeting itch with ligands selective for κ opioid receptors,” *Pharmacology of itch*, pp. 291–314, 2015.
- [196] G. F. Koob, C. L. Buck, A. Cohen, S. Edwards, P. E. Park, J. E. Schlosburg, B. Schmeichel, L. F. Vendruscolo, C. L. Wade, T. W. Whitfield Jr *et al.*,

- “Addiction as a stress surfeit disorder,” *Neuropharmacology*, vol. 76, pp. 370–382, 2014.
- [197] E. R. Butelman, V. Yuferov, and M. J. Kreek, “ κ -opioid receptor/dynorphin system: genetic and pharmacotherapeutic implications for addiction,” *Trends in neurosciences*, vol. 35, no. 10, pp. 587–596, 2012.
- [198] A. Karkhanis, K. M. Holleran, and S. R. Jones, “Dynorphin/kappa opioid receptor signaling in preclinical models of alcohol, drug, and food addiction,” *International Review of Neurobiology*, vol. 136, pp. 53–88, 2017.
- [199] F. P. Santos and S. Verstovsek, “Jak2 inhibitors: what’s the true therapeutic potential?” *Blood reviews*, vol. 25, no. 2, pp. 53–63, 2011.
- [200] F. Giordanetto and R. T. Kroemer, “Prediction of the structure of human janus kinase 2 (jak2) comprising jak homology domains 1 through 7,” *Protein engineering*, vol. 15, no. 9, pp. 727–737, 2002.
- [201] P. Saharinen and O. Silvennoinen, “The pseudokinase domain is required for suppression of basal activity of jak2 and jak3 tyrosine kinases and for cytokine-inducible activation of signal transduction,” *Journal of Biological Chemistry*, vol. 277, no. 49, pp. 47 954–47 963, 2002.
- [202] P. Saharinen, K. Takaluoma, and O. Silvennoinen, “Regulation of the jak2 tyrosine kinase by its pseudokinase domain,” *Molecular and cellular biology*, vol. 20, no. 10, pp. 3387–3395, 2000.
- [203] S. R. Hubbard, “Mechanistic insights into regulation of jak2 tyrosine kinase,” *Frontiers in endocrinology*, vol. 8, p. 361, 2018.
- [204] O. Silvennoinen and S. R. Hubbard, “Molecular insights into regulation of jak2 in myeloproliferative neoplasms,” *Blood, The Journal of the American Society of Hematology*, vol. 125, no. 22, pp. 3388–3392, 2015.
- [205] M. Chen, A. Cheng, F. Candotti, Y.-J. Zhou, A. Hymel, A. Fasth, L. D. Notarangelo, and J. J. O’Shea, “Complex effects of naturally occurring mutations in the jak3 pseudokinase domain: evidence for interactions between the kinase and pseudokinase domains,” *Molecular and Cellular Biology*, vol. 20, no. 3, pp. 947–956, 2000.

- [206] S. S. Jatiani, S. J. Baker, L. R. Silverman, and E. P. Reddy, “Jak/stat pathways in cytokine signaling and myeloproliferative disorders: approaches for targeted therapies,” *Genes & cancer*, vol. 1, no. 10, pp. 979–993, 2010.
- [207] K. Qamar and M. Saboor, “Jak 2 and stat proteins; a mini review,” *Biomedica*, vol. 34, no. 4, p. 232, 2018.
- [208] A. Kowalik, A. Kowalska, J. Kopczynski, A. Walczyk, E. Nowak, E. Wypior-kiewicz, R. Chodurska, L. Pieciak, and S. Gozdz, “Occurrence other than v600e mutation in the braf gene in papillary thyroid carcinoma,” in *Endocrine Abstracts*, vol. 35. Bioscientifica, 2014.
- [209] J. Xia, Y. He, B. Meng, S. Chen, J. Zhang, X. Wu, Y. Zhu, Y. Shen, X. Feng, Y. Guan *et al.*, “Nek2 induces autophagy-mediated bortezomib resistance by stabilizing beclin-1 in multiple myeloma,” *Molecular oncology*, vol. 14, no. 4, pp. 763–778, 2020.
- [210] U. Malapelle, F. Morra, G. Ilardi, R. Visconti, F. Merolla, A. Cerrato, V. Napolitano, R. Monaco, G. Guggino, G. Monaco *et al.*, “Usp7 inhibitors, downregulating ccdc6, sensitize lung neuroendocrine cancer cells to parp-inhibitor drugs,” *Lung Cancer*, vol. 107, pp. 41–49, 2017.
- [211] F. Morra, F. Merolla, V. Napolitano, G. Ilardi, C. Miro, S. Paladino, S. Staibano, A. Cerrato, and A. Celetti, “The combined effect of usp7 inhibitors and parp inhibitors in hormone-sensitive and castration-resistant prostate cancer cells,” *Oncotarget*, vol. 8, no. 19, p. 31815, 2017.
- [212] F. Morra, C. Luise, F. Merolla, I. Poser, R. Visconti, G. Ilardi, S. Paladino, H. Inuzuka, G. Guggino, R. Monaco *et al.*, “Fbxw7 and usp7 regulate ccdc6 turnover during the cell cycle and affect cancer drugs susceptibility in nsclc,” *Oncotarget*, vol. 6, no. 14, p. 12697, 2015.
- [213] S. Giovinazzi, V. M. Morozov, M. K. Summers, W. C. Reinhold, and A. M. Ishov, “Usp7 and daxx regulate mitosis progression and taxane sensitivity by affecting stability of aurora-a kinase,” *Cell Death & Differentiation*, vol. 20, no. 5, pp. 721–731, 2013.
- [214] M. Cartel, P.-L. Mouchel, M. Gotanègre, L. David, S. Bertoli, M.-D. Mas, A. Besson, J.-E. Sarry, S. Manenti, C. Didier *et al.*, “Inhibition of ubiquitin-specific protease 7 sensitizes acute myeloid leukemia to chemotherapy,” *Leukemia*, vol. 35, no. 2, pp. 417–432, 2021.

-
- [215] P. Zhang, Y. Wei, L. Wang, B. G. Debeb, Y. Yuan, J. Zhang, J. Yuan, M. Wang, D. Chen, Y. Sun *et al.*, “Atm-mediated stabilization of zeb1 promotes dna damage response and radioresistance through chk1,” *Nature cell biology*, vol. 16, no. 9, pp. 864–875, 2014.
- [216] W. M. Swetzig, J. Wang, and G. M. Das, “Estrogen receptor alpha ($era/esr1$) mediates the p53-independent overexpression of mdm4/mdmx and mdm2 in human breast cancer,” *Oncotarget*, vol. 7, no. 13, p. 16049, 2016.
- [217] S. Sharma, S.-C. Yang, L. Zhu, K. Reckamp, B. Gardner, F. Baratelli, M. Huang, R. K. Batra, and S. M. Dubinett, “Tumor cyclooxygenase-2/prostaglandin e2-dependent promotion of foxp3 expression and cd4+ cd25+ t regulatory cell activities in lung cancer,” *Cancer research*, vol. 65, no. 12, pp. 5211–5220, 2005.
- [218] D. Chauhan, Z. Tian, B. Nicholson, K. S. Kumar, B. Zhou, R. Carrasco, J. L. McDermott, C. A. Leach, M. Fulciniti, M. P. Kodrasov *et al.*, “A small molecule inhibitor of ubiquitin-specific protease-7 induces apoptosis in multiple myeloma cells and overcomes bortezomib resistance,” *Cancer cell*, vol. 22, no. 3, pp. 345–358, 2012.
- [219] K. Becker, N. Marchenko, G. Palacios, and U. M. Moll, “A role of hausp in tumor suppression in a human colon carcinoma xenograft model,” *Cell Cycle*, vol. 7, no. 9, pp. 1205–1213, 2008.
- [220] Y. Fan, J. Cheng, S. Vasudevan, J. Dou, H. Zhang, R. Patel, I. Ma, Y. Rojas, Y. Zhao, Y. Yu *et al.*, “Usp7 inhibitor p22077 inhibits neuroblastoma growth via inducing p53-mediated apoptosis,” *Cell death & disease*, vol. 4, no. 10, pp. e867–e867, 2013.
- [221] H. Zhang, T. Deng, R. Liu, T. Ning, H. Yang, D. Liu, Q. Zhang, D. Lin, S. Ge, M. Bai *et al.*, “Caf secreted mir-522 suppresses ferroptosis and promotes acquired chemo-resistance in gastric cancer,” *Molecular cancer*, vol. 19, no. 1, pp. 1–17, 2020.
- [222] D. Su, S. Ma, L. Shan, Y. Wang, Y. Wang, C. Cao, B. Liu, C. Yang, L. Wang, S. Tian *et al.*, “Ubiquitin-specific protease 7 sustains dna damage response and promotes cervical carcinogenesis,” *The Journal of clinical investigation*, vol. 128, no. 10, pp. 4280–4296, 2018.

- [223] M. B. Falana and Q. O. Nurudeen, “Evaluation of phytochemical constituents and in vitro antimicrobial activities of leaves extracts of *calotropis procera* against certain human pathogens,” *Notulae Scientia Biologicae*, vol. 12, no. 2, pp. 208–221, 2020.
- [224] R. Franqui-Machin, M. Hao, H. Bai, Z. Gu, X. Zhan, H. Habelhah, Y. Jethava, L. Qiu, I. Frech, G. Tricot *et al.*, “Destabilizing nek2 overcomes resistance to proteasome inhibition in multiple myeloma,” *The Journal of clinical investigation*, vol. 128, no. 7, pp. 2877–2893, 2018.
- [225] Q. Wang, S. Ma, N. Song, X. Li, L. Liu, S. Yang, X. Ding, L. Shan, X. Zhou, D. Su *et al.*, “Stabilization of histone demethylase phf8 by usp7 promotes breast carcinogenesis,” *The Journal of clinical investigation*, vol. 126, no. 6, pp. 2205–2220, 2016.
- [226] L. Kategaya, P. Di Lello, L. Rougé, R. Pastor, K. R. Clark, J. Drummond, T. Kleinheinz, E. Lin, J.-P. Upton, S. Prakash *et al.*, “Usp7 small-molecule inhibitors interfere with ubiquitin binding,” *Nature*, vol. 550, no. 7677, pp. 534–538, 2017.
- [227] S.-B. Shin, C.-H. Kim, H.-R. Jang, and H. Yim, “Combination of inhibitors of usp7 and plk1 has a strong synergism against paclitaxel resistance,” *International Journal of Molecular Sciences*, vol. 21, no. 22, p. 8629, 2020.
- [228] A. Gallardo, E. Lerma, D. Escuin, A. Tibau, J. Munoz, B. Ojeda, A. Barnadas, E. Adrover, L. Sánchez-Tejada, D. Giner *et al.*, “Increased signalling of egfr and igf1r, and deregulation of pten/pi3k/akt pathway are related with trastuzumab resistance in her2 breast carcinomas,” *British journal of cancer*, vol. 106, no. 8, pp. 1367–1373, 2012.
- [229] A. Agathangelou, E. Smith, N. J. Davies, M. Kwok, A. Zlatanou, C. E. Oldreive, J. Mao, D. Da Costa, S. Yadollahi, T. Perry *et al.*, “Usp7 inhibition alters homologous recombination repair and targets cll cells independently of atm/p53 functional status,” *Blood, The Journal of the American Society of Hematology*, vol. 130, no. 2, pp. 156–166, 2017.
- [230] J. Cai, H.-Y. Chen, S.-J. Peng, J.-L. Meng, Y. Wang, Y. Zhou, X.-P. Qian, X.-Y. Sun, X.-W. Pang, Y. Zhang *et al.*, “Usp7-trim27 axis negatively modulates antiviral type i ifn signaling,” *The FASEB Journal*, vol. 32, no. 10, pp. 5238–5249, 2018.

-
- [231] T. An, Y. Gong, X. Li, L. Kong, P. Ma, L. Gong, H. Zhu, C. Yu, J. Liu, H. Zhou *et al.*, “Usp7 inhibitor p5091 inhibits wnt signaling and colorectal tumor growth,” *Biochemical pharmacology*, vol. 131, pp. 29–39, 2017.
- [232] S. M. Heaton, N. A. Borg, and V. M. Dixit, “Ubiquitin in the activation and attenuation of innate antiviral immunity,” *Journal of Experimental Medicine*, vol. 213, no. 1, pp. 1–13, 2016.
- [233] W. Wu, A. Koike, T. Takeshita, and T. Ohta, “The ubiquitin e3 ligase activity of brca1 and its biological functions,” *Cell division*, vol. 3, no. 1, pp. 1–10, 2008.
- [234] D. Komander, “The emerging complexity of protein ubiquitination,” *Biochemical Society Transactions*, vol. 37, no. 5, pp. 937–953, 2009.
- [235] Y. M. Ohol, M. T. Sun, G. Cutler, P. R. Leger, D. X. Hu, B. Biannic, P. Rana, C. Cho, S. Jacobson, S. T. Wong *et al.*, “Novel, selective inhibitors of usp7 uncover multiple mechanisms of antitumor activity in vitro and in vivo,” *Molecular Cancer Therapeutics*, vol. 19, no. 10, pp. 1970–1980, 2020.
- [236] F. Colland, E. Formstecher, X. Jacq, C. Reverdy, C. Planquette, S. Conrath, V. Trouplin, J. Bianchi, V. N. Aushev, J. Camonis *et al.*, “Small-molecule inhibitor of usp7/hausp ubiquitin protease stabilizes and activates p53 in cells,” *Molecular cancer therapeutics*, vol. 8, no. 8, pp. 2286–2295, 2009.
- [237] N. J. Schauer, X. Liu, R. S. Magin, L. M. Doherty, W. C. Chan, S. B. Ficarro, W. Hu, R. M. Roberts, R. E. Iacob, B. Stolte *et al.*, “Selective usp7 inhibition elicits cancer cell killing through a p53-dependent mechanism,” *Scientific reports*, vol. 10, no. 1, pp. 1–15, 2020.
- [238] A. P. Turnbull, S. Ioannidis, W. W. Krajewski, A. Pinto-Fernandez, C. Heride, A. C. Martin, L. M. Tonkin, E. C. Townsend, S. M. Buker, D. R. Lancia *et al.*, “Molecular basis of usp7 inhibition by selective small-molecule inhibitors,” *Nature*, vol. 550, no. 7677, pp. 481–486, 2017.
- [239] A. Pozhidaeva, G. Valles, F. Wang, J. Wu, D. E. Sterner, P. Nguyen, J. Weinstock, K. S. Kumar, J. Kanyo, D. Wright *et al.*, “Usp7-specific inhibitors target and modify the enzyme’s active site via distinct chemical mechanisms,” *Cell Chemical Biology*, vol. 24, no. 12, pp. 1501–1512, 2017.

- [240] I. Lamberto, X. Liu, H.-S. Seo, N. J. Schauer, R. E. Iacob, W. Hu, D. Das, T. Mikhailova, E. L. Weisberg, J. R. Engen *et al.*, “Structure-guided development of a potent and selective non-covalent active-site inhibitor of usp7,” *Cell chemical biology*, vol. 24, no. 12, pp. 1490–1500, 2017.
- [241] A. Volkamer, D. Kuhn, F. Rippmann, and M. Rarey, “Dogsitescorer: A web server for automatic binding site prediction, analysis and druggability assessment,” *Bioinformatics*, vol. 28, no. 15, p. 2074–2075, 2012.
- [242] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, “The protein data bank,” *Nucleic acids research*, vol. 28, no. 1, pp. 235–242, 2000.
- [243] G. Lebon, T. Warne, P. C. Edwards, K. Bennett, C. J. Langmead, A. G. Leslie, and C. G. Tate, “Agonist-bound adenosine a2a receptor structures reveal common features of gpcr activation,” *Nature*, vol. 474, no. 7352, pp. 521–525, 2011.
- [244] V.-P. Jaakola, M. T. Griffith, M. A. Hanson, V. Cherezov, E. Y. Chien, J. R. Lane, A. P. Ijzerman, and R. C. Stevens, “The 2.6 angstrom crystal structure of a human a2a adenosine receptor bound to an antagonist,” *Science*, vol. 322, no. 5905, pp. 1211–1217, 2008.
- [245] A. S. Doré, N. Robertson, J. C. Errey, I. Ng, K. Hollenstein, B. Tehan, E. Hurrell, K. Bennett, M. Congreve, F. Magnani *et al.*, “Structure of the adenosine a2a receptor in complex with zm241385 and the xanthines xac and caffeine,” *Structure*, vol. 19, no. 9, pp. 1283–1293, 2011.
- [246] G. Lebon, P. C. Edwards, A. G. Leslie, and C. G. Tate, “Molecular determinants of cgs21680 binding to the human adenosine a2a receptor,” *Molecular pharmacology*, vol. 87, no. 6, pp. 907–915, 2015.
- [247] B. Carpenter, R. Nehmé, T. Warne, A. G. Leslie, and C. G. Tate, “Structure of the adenosine a2a receptor bound to an engineered g protein,” *Nature*, vol. 536, no. 7614, pp. 104–107, 2016.
- [248] E. Segala and D. Guo, “Controlling the dissociation of ligands from the adenosine a2a receptor through modulation of salt bridge strength,” *Journal of medicinal chemistry*, vol. 59, no. 13, pp. 6470–6479, 2016.

- [249] B. Sun, P. Bachhawat, M. L.-H. Chu, M. Wood, T. Ceska, Z. A. Sands, J. Mercier, F. Lebon, T. S. Kobilka, and B. K. Kobilka, “Crystal structure of the adenosine a2a receptor bound to an antagonist reveals a potential allosteric pocket,” *Proceedings of the National Academy of Sciences*, vol. 114, no. 8, pp. 2066–2071, 2017.
- [250] R. R. Davis, B. Li, S. Y. Yun, A. Chan, P. Nareddy, S. Gunawan, M. Ayaz, H. R. Lawrence, G. W. Reuther, N. J. Lawrence *et al.*, “Structural insights into jak2 inhibition by ruxolitinib, fedratinib, and derivatives thereof,” *Journal of medicinal chemistry*, vol. 64, no. 4, pp. 2228–2241, 2021.
- [251] C. R. Wellaway, I. R. Baldwin, P. Bamborough, D. Barker, M. A. Bartholomew, C.-w. Chung, J. P. Evans, N. J. Fazakerley, P. Homes *et al.*, “Investigation of janus kinase (jak) inhibitors for lung delivery and the importance of aldehyde oxidase metabolism,” *Journal of medicinal chemistry*, 2021.
- [252] H. Wu, D. Wacker, M. Mileni, V. Katritch, G. W. Han, E. Vardy, W. Liu, A. A. Thompson, X.-P. Huang, F. Carroll *et al.*, “Structure of the human κ -opioid receptor in complex with jdtic,” *Nature*, vol. 485, no. 7398, pp. 327–332, 2012.
- [253] T. Che, J. English, B. E. Krumm, K. Kim, E. Pardon, R. H. Olsen, S. Wang, S. Zhang, J. F. Diberto, N. Sciaky *et al.*, “Nanobody-enabled monitoring of kappa opioid receptor states,” *Nature communications*, vol. 11, no. 1, pp. 1–12, 2020.
- [254] M. Li, S. Liu, H. Chen, X. Zhou, J. Zhou, S. Zhou, H. Yuan, Q.-L. Xu, J. Liu, K. Cheng *et al.*, “N-benzylpiperidinol derivatives as novel usp7 inhibitors: Structure–activity relationships and x-ray crystallographic studies,” *European Journal of Medicinal Chemistry*, vol. 199, p. 112279, 2020.
- [255] P. R. Leger, D. X. Hu, B. Biannic, M. Bui, X. Han, E. Karbarz, J. Maung, A. Okano, M. Osipov, G. M. Shibuya *et al.*, “Discovery of potent, selective, and orally bioavailable inhibitors of usp7 with in vivo antitumor activity,” *Journal of medicinal chemistry*, vol. 63, no. 10, pp. 5398–5420, 2020.
- [256] A. Gaulton, A. Hersey, M. Nowotka, A. P. Bento, J. Chambers, D. Mendez, P. Mutowo, F. Atkinson, L. J. Bellis, E. Cibrián-Uhalte *et al.*, “The chembl

- database in 2017,” *Nucleic acids research*, vol. 45, no. D1, pp. D945–D954, 2017.
- [257] T. Pereira, M. Abbasi, B. Ribeiro, and J. P. Arrais, “Diversity oriented deep reinforcement learning for targeted molecule generation,” *Journal of Cheminformatics*, vol. 13, no. 1, p. 1–17, 2021.
- [258] G. M. Morris, H. Ruth, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, and A. J. Olson, “Software news and updates autodock4 and autodocktools4: Automated docking with selective receptor flexibility,” *Journal of Computational Chemistry*, vol. 30, no. 16, p. 2785–2791, 2009.
- [259] P. A. Ravindranath, S. Forli, D. S. Goodsell, A. J. Olson, and M. F. Sanner, “Autodockfr: Advances in protein-ligand docking with explicitly specified binding site flexibility,” *PLoS Computational Biology*, vol. 11, no. 12, p. 1–28, 2015.
- [260] O. Trott and A. J. Olson, “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, pp. NA–NA, 2009.
- [261] H. Zhao and A. Caffisch, “Discovery of zap70 inhibitors by high-throughput docking into a conformation of its kinase domain generated by molecular dynamics,” *Bioorganic & medicinal chemistry letters*, vol. 23, no. 20, pp. 5721–5726, 2013.
- [262] O. Korb, T. Stützel, and T. E. Exner, “Plants: Application of ant colony optimization to structure-based drug design,” in *International workshop on ant colony optimization and swarm intelligence*. Springer, 2006, pp. 247–258.
- [263] S. Ruiz-Carmona, D. Alvarez-Garcia, N. Foloppe, A. B. Garmendia-Doval, S. Juhos, P. Schmidtke, X. Barril, R. E. Hubbard, and S. D. Morley, “rdock: A fast, versatile and open source program for docking ligands to proteins and nucleic acids,” *PLoS Computational Biology*, vol. 10, no. 4, p. 1–8, 2014.
- [264] C. Empereur-Mot, J.-F. Zagury, and M. Montes, “Screening explorer—an interactive tool for the analysis of screening results,” *Journal of chemical information and modeling*, vol. 56, no. 12, pp. 2281–2286, 2016.
- [265] L. Molsoft, “Drug-likeness and molecular property prediction,” 2018.

- [266] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, “Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers,” *SoftwareX*, vol. 1-2, pp. 19–25, 2015.
- [267] W. Humphrey, A. Dalke, and K. Schulten, “Vmd: visual molecular dynamics,” *Journal of molecular graphics*, vol. 14, no. 1, pp. 33–38, 1996.
- [268] D. R. Houston and M. D. Walkinshaw, “Consensus docking: improving the reliability of docking in a virtual screening context,” *Journal of chemical information and modeling*, vol. 53, no. 2, pp. 384–390, 2013.

