**DOCTORAL THESIS**


UCAM
UNIVERSIDAD CATÓLICA
DE MURCIA

# INTERNATIONAL DOCTORAL SCHOOL

*Doctoral Programme in Health Sciences*

## Enhancing Virtual Screening in Drug Discovery Through Consensus-Based Methods

*Author:*
Jochem Nelen

*Supervisor:*
Dr. Horacio Pérez Sánchez

*Murcia, July 2025*

**DOCTORAL THESIS**


UCAM
UNIVERSIDAD CATÓLICA
DE MURCIA

# INTERNATIONAL DOCTORAL SCHOOL

*Doctoral Programme in Health Sciences*

## Enhancing Virtual Screening in Drug Discovery Through Consensus-Based Methods

*Author:*
Jochem Nelen

*Supervisor:*
Dr. Horacio Pérez Sánchez

*Murcia, July 2025*

**THESIS SUPERVISORS' AUTHORISATION**
**FOR THESIS SUBMISSION**

Prof. Dr. Horacio Pérez Sánchez, as Supervisor of the Doctoral Thesis '*Enhancing Virtual Screening in Drug Discovery Through Consensus-Based Methods*' by Mr. Jochem Nelen in the Doctorate Programme in Health Sciences, **authorises its submission**, given that it meets the required conditions for its defence.

Which I hereby sign in compliance with Spanish Royal Decree 99/2011, of 28 January, in Murcia, on July 4, 2025.

Prof. Dr. Horacio Pérez Sánchez

## ABSTRACT

The development of novel drugs is a complex and costly process: accurately identifying active molecules from vast chemical space is an incredibly challenging endeavor. Computational tools can help to speed up the drug discovery process by enabling the exploration of large compound libraries with reduced experimental costs. There are many different computational strategies available, but the efficacy of many methods can vary considerably depending on the scenario and target. One strategy that has been shown to improve consistency and reduce false positives within virtual screening campaigns is the use of consensus strategies, which combine predictions from multiple methods. Further investigating this concept, this thesis presents two novel consensus methods, one structure-based and one ligand-based, developed to improve hit identification through practical and reproducible workflows.

The first, ESSENCE-Dock, is a structure-based consensus docking pipeline that is able to integrate results from several molecular docking methods. By incorporating docking scores, pose similarity, and ligand flexibility, it has been shown to reduce false positives and improve enrichment in the DUD-E benchmark. Specifically, ESSENCE-Dock outperformed both individual docking tools and basic consensus approaches across diverse targets from various protein families. The method has been designed to run efficiently in high-performance computing (HPC) environments and can be run through a simple command-line interface.

The second method, ConFiLiS (*Consensus Fingerprints for Ligand-based Screening*), is a ligand-based virtual screening toolkit that reranks compounds based on their combined similarity scores derived from different molecular fingerprints. Its utility has been demonstrated in the identification of dimethoxycurcumin, a methylated derivative of curcumin, as an inhibitor of ABCC3. This protein target is a transporter associated with chemotherapy resistance in pancreatic cancer. Experimental assays confirmed its ability to reduce cell viability and colony formation in pancreatic cancer cell lines, confirming its potential to act as a chemosensitizing agent in this context.

Finally, this thesis also examines the reliability of bioactivity data in public databases, which is critical for data-driven drug discovery and the development of machine learning models. By analyzing assay variability in ChEMBL, it was shown that comparing potency differences between analogs within the same assay yields more reliable insights than comparing absolute activity values across assays for the same target.

Taken together, these contributions demonstrate how consensus approaches and attention to data quality can improve the reliability and practical impact of virtual screening methods in computational drug discovery.


## KEY WORDS

Drug Discovery, Computational Chemistry, Cheminformatics, Molecular Docking, Molecular Fingerprints, Virtual Screening

## RESUMEN

El desarrollo de nuevos fármacos es un proceso complejo y costoso donde resulta muy difícil identificar moléculas activas con un alto grado de confianza a partir de un vasto espacio químico. Las herramientas computacionales pueden ayudar a acelerar el proceso al permitir la exploración de grandes bibliotecas de compuestos a costes experimentales reducidos. Existen muchas estrategias computacionales diferentes, pero la eficacia de los métodos empleados puede variar considerablemente en distintos escenarios y objetivos. Una forma de mejorar esta coherencia es mediante estrategias de consenso, donde se combinan los resultados los diversos métodos. Esta tesis presenta dos nuevos métodos de consenso, uno basado en estructura y otro en ligandos, desarrollados para mejorar la identificación de aciertos mediante flujos de trabajo prácticos y reproducibles.

En primer lugar, ESSENCE-Dock es un proceso de acoplamiento consensuado basado en estructura que combina los resultados de varios métodos de acoplamiento molecular. Mediante el consenso de las puntuaciones de acoplamiento, la similitud de pose y la flexibilidad del ligando, se ha demostrado que reduce los falsos positivos y mejora el enriquecimiento en la prueba de referencia DUD-E. En concreto, ESSENCE-Dock supera tanto a las herramientas de acoplamiento individuales como a los enfoques básicos de consenso en diversas dianas de varias familias de proteínas. El flujo de trabajo ha sido diseñado para funcionar eficientemente en entornos de computación de alto rendimiento y puede ejecutarse a través de una sencilla interfaz de línea de comandos.

En segundo lugar, ConFiLiS (*Consensus Fingerprints for Ligand-based Screening*) es una herramienta de cribado virtual basada en ligandos que puede jerarquizar compuestos basándose en la similitud de múltiples huellas moleculares. Un ejemplo práctico de su eficacia es la identificación de la dimetoxicurcumina, un derivado metilado de la curcumina, como inhibidor de ABCC3. Se trata de una diana relevante en el cáncer de páncreas, ya que puede facilitar la resistencia a los fármacos contra la quimioterapia. Los ensayos celulares confirmaron su capacidad para reducir la viabilidad celular y la formación de colonias en líneas celulares de cáncer de páncreas, lo que confirma su potencial para actuar como agente quimiosensibilizador en este contexto.

Por último, con el fin de proporcionar información sobre la fiabilidad de las bases de datos de código abierto, se estudió la variabilidad de los ensayos en la base de datos ChEMBL para facilitar el descubrimiento de fármacos y la quimioinformática basados en IA. Mediante la comparación de pares moleculares emparejados a través de ensayos para el mismo objetivo, se demostró que la comparación entre las diferencias de potencia entre los análogos del mismo ensayo son más robustas al ruido que la comparación directa de los valores absolutos de actividad entre estos mismos ensayos.

En conjunto, estos avances demuestran cómo el desarrollo de herramientas computacionales, métodos de consenso y el tratamiento adecuado de la calidad de los datos, permiten mejorar notablemente la fiabilidad y el impacto de los flujos de trabajo de cribado virtual en el descubrimiento de fármacos.


## PALABRAS CLAVE

Descubrimiento de Fármacos, Química Computacional, Quimioinformática, Acoplamiento Molecular, Huellas Moleculares, Cribado Virtual

## ACKNOWLEDGEMENTS

First and foremost, I would like to express my deep gratitude to Prof. Horacio Pérez-Sánchez, my thesis supervisor. His guidance, support, and constant encouragement were instrumental throughout this journey. His readiness to help with any problem, while pushing me to grow as an independent researcher, made this thesis possible.

I am equally thankful to Dr. José Manuel Villalgordo-Soto, CEO of Eurofins-Villapharma, whose scientific collaboration and financial support through the Cátedra Villapharma made this thesis possible. His involvement brought a valuable industrial perspective to my research, and I greatly appreciated the productive and enjoyable collaboration.

Additionally, I would like to thank Prof. Silvia Montoro-García, my tutor, for her help and support with the many practical and administrative matters throughout my PhD.

I would also like to thank my BIO-HPC colleagues Miguel, Alejandro, Carlos, Elena, and Jorge. Their input, camaraderie, and good humor created a very pleasant and stimulating working environment.

I am also grateful to all the researchers I had the opportunity to work with during my PhD. Their feedback, collaboration, and expertise not only helped shape this work, but also supported my growth as a scientist. Among these experiences, I especially value the research stays I was able to undertake at other laboratories. I am grateful to Hans De Winter and Dries Van Rompaey, as well as the entire UAMC group—in particular Joep, Roy, Stijn, and Olivier—for their warm welcome and inspiring scientific atmosphere. My stay at EMBL Grenoble with the McCarthy Team was equally enriching, and I would like to specifically thank Matthew Bowler and Jill von Velsen for their kindness and collaboration.

Along the way, I also met wonderful people who became dear friends. I am especially thankful to Maria Bzówka, whose time in our lab grew into a lasting friendship, and to Yuxin and Xing, whose presence made both my work and daily life in Murcia all the more enjoyable.

Outside of academia, I am fortunate to have many friends who supported me throughout this period. Yrjo, one of my closest friends since high school, has always been there for me. I also want to thank Marijke for her friendship and the many good times, as well as Jasper, Indy, Dries, and Thijs, whose continued companionship since university I continue to cherish.

Furthermore, I want to thank my family for their unwavering love and support. To my grandmother, thank you for always thinking of me, especially when I was far away. Your warmth during every visit home is something I will always cherish. I also want to thank my wonderful nieces, nephews, uncles, and aunts. Your kindness, humor, and interest in my work have always brought joy and balance to my life.

Finally, to my parents, Marleen and Leon—thank you for your endless support, your wisdom, and for always believing in me. You have been my foundation, without which I could not have come to where I am today. To my twin brother Jorne, thank you for your constant encouragement, support, and sense of humor—being on this journey together in different ways means a lot. And to my sister Jente, whose energy, warmth, and support never fail to lift my spirits—thank you. I am endlessly grateful to have you all by my side. *Bedankt voor alles.*

To Leon and Marleen, my parents
To Jorne and Jente, my brother and sister

## GENERAL TABLE OF CONTENTS

## ACRONYMS AND ABBREVIATIONS

**ABFE:** Absolute Binding Free Energy

**ADMET**: Absorption, Distribution, Metabolism, Excretion, and Toxicity

**AI:** Artificial Intelligence

**CADD:** Computer-Aided Drug Design

**CNN**: Convolutional Neural Network

**ConFiLiS**: Consensus Fingerprints for Ligand-based Screening

**DDR1**: Discoidin Domain Receptor 1

**DL**: Deep Learning

**DUck**: Dynamic Undocking

**EMA**: European Medicines Agency

**ESSENCE-Dock**: Effective Structural Screening ENrichment ConsEnsus Dock

**FDA**: Food and Drug Administration

**FEP**: Free Energy Perturbation

**GAN:** Generative Adversarial Networks

**GNN**: Graph Neural Network

**HPC**: High-Performance Computing

**HTS**: High-Throughput Screening

**LBVS**: Ligand-Based Virtual Screening

**MD**: Molecular Dynamics

**ML**: Machine Learning

**MM-GBSA**: Molecular Mechanics with Generalized Born Surface Area

**MMP**: Matched Molecular Pair

**MM-PBSA**: Molecular Mechanics with Poisson–Boltzmann Surface Area

**NET**: Neutrophil Extracellular Trap

**NME**: New Molecular Entity

**PAD4**: Protein-Arginine Deiminase 4

**QSAR**: Quantitative Structure-Activity Relationship

**RBFE**: Relative Binding Free Energy

**RMSD**: Root Mean Square Deviation

**SAR**: Structure-Activity Relationship

**SBVS**: Structure-Based Virtual Screening

**SLURM**: Simple Linux Utility for Resource Management

**TI**: Thermodynamic Integration

**VS**: Virtual Screening

**Wee1**: Wee1-like protein kinase

**XAI**: Explainable Artificial Intelligence

# LIST OF FIGURES AND APPENDIXES

## LIST OF FIGURES

# APPENDICES

# I – INTRODUCTION

# I - INTRODUCTION

## 1.1. THE DRUG DISCOVERY PIPELINE: A COMPLEX AND COSTLY ENDEAVOR

The discovery and development of new therapeutic agents is a difficult, time-consuming, and extremely expensive process. Research indicates that bringing a single drug to market requires a financial investment ranging from hundreds of millions to more than 4 billion USD (1), and can take anywhere from 10 to 15 years (2), depending on the therapeutic area, regulatory path, and nature of the compound. The overall probability of a small molecule or new molecular entity (NME) successfully reaching the market remains low: the average drug candidate has been estimated to only have around a 9.6% likelihood of approval throughout all stages of clinical trials and eventual regulatory approval (3). These metrics highlight the scientific and technical difficulties inherent in drug discovery. They also reflect the complexity of progressing through a highly regulated environment where demonstrating safety and efficacy at every stage is mandatory. A general schematic overview of the drug development process is presented in Figure I.1.

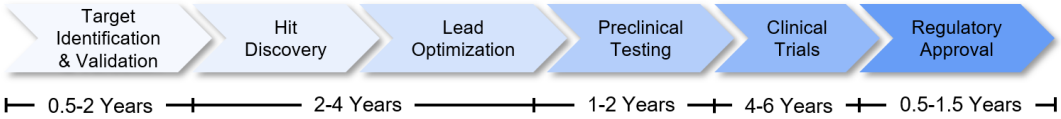



*Figure I.1. Overview of the drug discovery and development pipeline.*

*The drug discovery process spans multiple stages, including target identification, hit discovery, lead optimization, preclinical testing, and clinical trials, each with distinct timelines and success rates. The time estimates for each stage are adapted from Reddy et al. (2).*

Typically, drug discovery first involves target identification and validation (4). The goal is to find a molecular target (e.g. a protein, enzyme, receptor, or gene) that is causally responsible for a specific disease. This early stage is highly interdisciplinary and often involves genomics, proteomics, bioinformatics, and disease biology. Once a target has been identified, it needs to be experimentally confirmed as both biologically relevant and druggable: it should be possible to modify the target's activity by small molecules or biologics in a way that can be therapeutically beneficial.

After target identification, the hit discovery process is started. During this stage, researchers attempt to identify chemical compounds that bind to key regions of the target and cause a change in activity. This procedure is often carried out through time-intensive and costly high-throughput screening (HTS) of commercial compound libraries or fragment-based screening approaches (5). Often, these experimental screenings are preceded by computational methods such as virtual screening (VS), which can focus efforts by pre-selecting promising compounds through *in silico* testing (6). The outcome of this screening phase is usually a list of "hit" compounds: small molecules that are measurable active towards the target in biochemical or biophysical assays. However, these hits will generally not possess desirable druglike properties such as solubility, selectivity or permeability. Nevertheless, these identified hits usually are good starting points but still need to be optimized in subsequent stages.

This refinement process is also known as the lead optimization phase, which very often is chemistry driven and iterative in nature (7). The early hits previously discovered are systematically modified with the goal of improving their potency, selectivity, metabolic stability, solubility, and toxicity profile. Medicinal chemists, pharmacologists, and computational scientists work together to design and synthesize derivatives and analogs, guided by structure–activity relationships (SAR) and absorption, distribution, metabolism, excretion, and toxicity (ADMET) data. This can be resource-intensive in both time and money, and often leads to hundreds of analogs for each hit. The final goal of this stage is to deliver one or more "lead" compounds optimized for activity, selectivity and ADMET properties.

Once the leads are deemed sufficiently optimized, they enter preclinical testing, which involves extensive *in vitro* and *in vivo* studies in order to evaluate pharmacokinetics, toxicology, and safety (8). Preclinical models are meant to resemble human physiology and disease as much as possible, using either cellular systems or animal models. These experiments help determine the compound's dosing parameters, organ distribution, metabolism, and potential off-target effects. It also includes stability studies and formulation development to ensure the compound can be effectively delivered during future clinical trials in humans.

If a compound passes preclinical evaluation, it proceeds to clinical trials, which are conducted in humans and typically involve three stages (9). Phase I of

clinical trials tests safety, tolerability, and pharmacokinetics in a small group of healthy volunteers. Phase II expands to patients with the disease under investigation, assessing efficacy, dosing, and short-term side effects. Phase III involves large-scale studies across diverse patient populations and different clinical settings, comparing the new drug with existing therapies or placebos. Phase III generates much of the data used in regulatory submissions and is usually the most expensive and time-consuming component of clinical development.

Upon successful completion of Phase III, a regulatory filing is prepared and submitted to agencies such as the U.S. Food and Drug Administration (FDA) or the European Medicines Agency (EMA). After approval, the drug can be released on the market, though it continues to be monitored through post-marketing surveillance or Phase IV studies. These studies obtain long-term safety data and may uncover rare adverse effects not detected in earlier trials.

Every phase of the drug discovery and development process is marked by high rates of attrition, long timelines, and increasing costs. Reasons for failures usually are due to lack of efficacy, unexpected toxicity, or poor pharmacokinetics, which often are only discovered in later stages after significant investments. Therefore, there is much interest in optimizing the efficiency and predictive accuracy of the earlier stages, particularly target identification, hit discovery, and lead optimization, which is where computational tools and virtual screening can be the most effective.

## 1.2. *IN SILICO* DRUG DISCOVERY: ACCELERATING EXPLORATION OF CHEMICAL SPACE

Using computational drug discovery methods can help reduce the high costs and extended timelines that are traditionally associated with classical drug development campaigns (10). By applying fast, computational screening of virtual compound libraries during hit and lead identification stages, *in silico* strategies can enable scientists to select the most promising candidates early on in the drug discovery process (11). With techniques such as molecular docking, quantitative structure–activity relationships (QSAR), machine learning (ML) algorithms, and molecular dynamics (MD) simulations, *in silico* methods allow researchers to efficiently screen and rank compounds based on drug-like properties and high predicted binding affinities (12).

One of the main advantages of *in silico* methods is that they have the potential to substantially reduce the quantity of compounds that need to be synthesized and experimentally tested. In this way, computational techniques not only accelerate both hit discovery and lead optimization, but also help reduce resource demands, which is a significant consideration in the high-risk, high-cost world of early drug discovery. At the same time, computational approaches enable exploration of chemical space far beyond what is tractable through classical, experimental HTS alone. Moreover, through application of approaches such as *de novo* drug design and generative artificial intelligence (AI), entirely novel chemotypes can be computationally designed for specific targets, enhancing novelty and diversity of candidate molecules.

Additionally, computer-aided drug design (CADD) can also assist in decision-making during lead optimization by providing detailed molecular insights regarding the factors affecting biological activity, selectivity, and ADME properties. For example, free-energy calculations such as free energy perturbation (FEP) can be used to investigate which chemical modifications are most likely to improve activity and selectivity (13). Similarly, advanced ML models are able to predict biological activity and simultaneously offer explanations for these predictions through explainable artificial intelligence (XAI), which can be used to guide follow-up optimization efforts (14).

Overall, integrating computational methods into drug discovery can accelerate compound design cycles while simultaneously reducing costs. Additionally, these approaches enable researchers to explore larger and more diverse compound libraries that would be difficult or even impossible to access using experimental approaches alone. In this way, computational methods can help researchers to navigate the complexity of chemical space with greater efficiency and precision.

### 1.3. VIRTUAL SCREENING APPROACHES

Virtual screening (VS) is one of the most used computational methods in drug discovery, as it is frequently used to quickly identify bioactive molecules from large chemical libraries (15). VS typically attempts to rank compounds based on their (predicted) affinity for a specific biological target or by comparing their structures

to known active compounds. Thus, VS enables prioritization of compounds during the early hit discovery phase and can guide experimental efforts toward the most promising candidates. This is particularly valuable given the immense, yet sparse chemical space relevant to drug discovery (16).

VS methods are usually categorized into three classes: ligand-based, structure-based, and machine-learning methods (11), as also illustrated in Figure I.2. These three classes are founded on different forms of prior knowledge, such as previously identified active molecules, structural protein data (e.g. 3D crystal structures), or predictive models derived from large datasets. Despite differences in methodology, all VS approaches share the same objective: the effective selection of smaller sub-sets of molecules from large chemical libraries which are enriched in biologically active compounds.

The following sections elaborate on these three virtual screening approaches in further detail, briefly describing their fundamental principles, frequent applications, and recent advances.

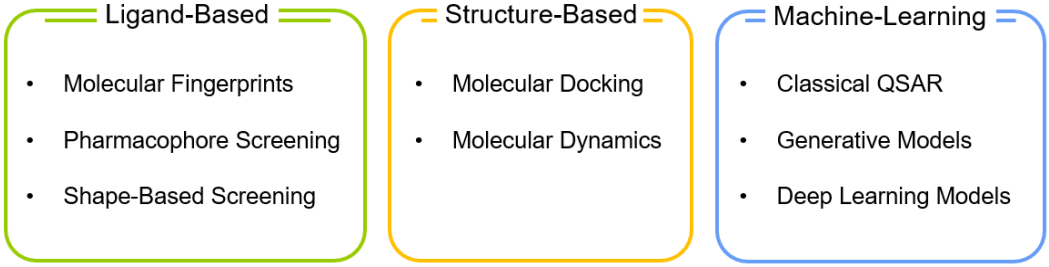



*Figure I.2. Schematic overview of techniques commonly used for virtual screening.*

*General overview of widely used virtual screening approaches in drug discovery, grouped into ligand-based, structure-based, and machine-learning methods.*

### 1.3.1. Ligand-Based Virtual Screening (LBVS)

Ligand-based virtual screening (LBVS) is based on the assumption that molecules with similar chemical or structural features typically exhibit similar biological activities (17). This idea is commonly called the similarity-property principle. Based on this, LBVS uses known bioactive compounds to try to identify new molecules sharing essential features with established query ligands.

A widely used LBVS approach is molecular fingerprint similarity. Classical molecular fingerprints encode chemical substructures into binary or hashed representations, such as Morgan or MACCS fingerprints (18). Figure I.3 illustrates the idea behind substructure fingerprints such as MACCS, where each predefined bit indicates the presence or absence of a specific substructure. The fingerprints of different molecules can be compared using similarity metrics like the Tanimoto coefficient to screen for compounds with similar structural characteristics (19). Once a virtual library has been converted into fingerprints, such similarity calculations can be performed rapidly, making this method particularly suitable and efficient for screening large, pre-processed chemical libraries.

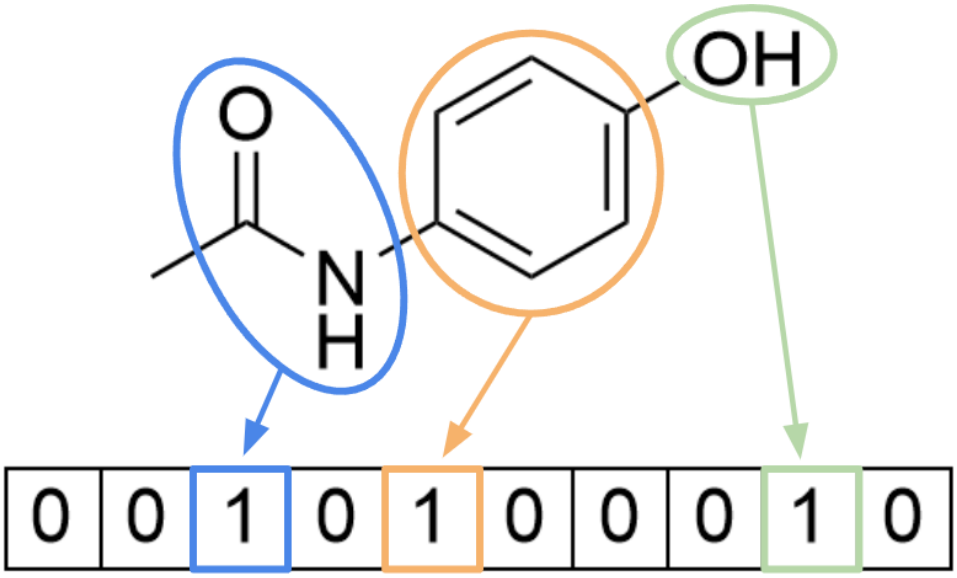



*Figure I.3. Visualization of substructure-based fingerprint encoding.*
*Schematic representation of a substructure fingerprint for paracetamol, where each bit corresponds to the presence (1) or absence (0) of a predefined chemical feature. The fingerprint shown is simplified and illustrative, highlighting the detection of features such as an amide group, an aromatic ring, and an alcohol.*

Shape-based screening provides an alternative approach by comparing the 3D shapes of molecules (20). It evaluates similarity based on the volume overlap or alignment of molecular surfaces, emphasizing spatial rather than purely structural features. Tools like VAMS (21) or ROCS (22) evaluate how well a candidate molecule conforms to the overall shape of a known active, often using Gaussian or electrostatic overlays to quantify similarity.

Extending this concept further, pharmacophore modeling includes specific chemical properties as well as their spatial arrangement (23). These pharmacophore models encode essential features required for biological activity, such as hydrogen bond donors/acceptors, hydrophobic centers, and charged groups. These features are stored as a spatial arrangement, which can be used to screen compound libraries for matching feature patterns. An example pharmacophore model is shown in Figure I.4. Some notable examples of pharmacophore modeling software include LigandScout (24), Pharao (25), and Pharmit (26).

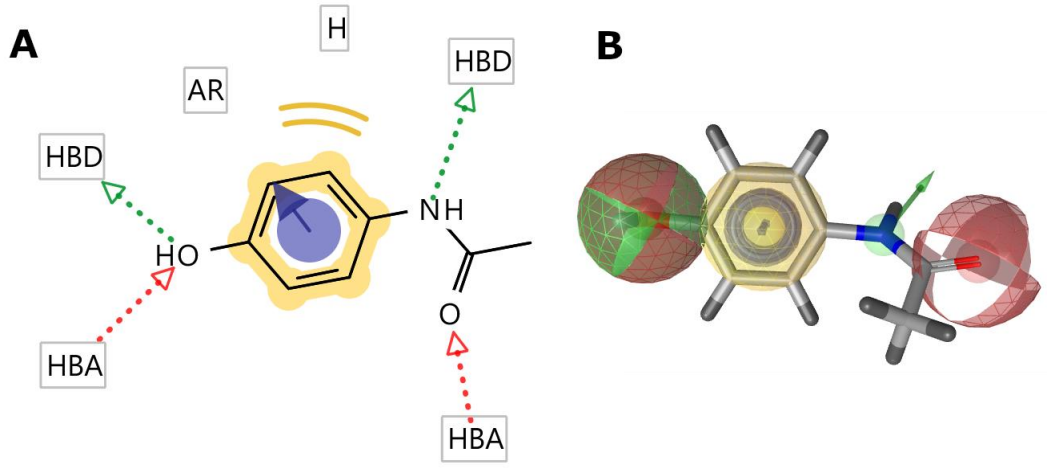



*Figure I.4. Example pharmacophore model of paracetamol in 2D and 3D.*

*Illustrative pharmacophore model of paracetamol as generated by LigandScout (24). A) Two-dimensional view showing key interaction features: hydrogen bond donors (HBD, green), hydrogen bond acceptors (HBA, red), an aromatic ring (AR, blue), and a hydrophobic region (H, yellow). B) Corresponding three-dimensional representation highlighting the spatial orientation of the same pharmacophoric features in 3D space.*

LBVS methods are especially useful when the 3D structure of the biological target is unavailable, poorly resolved, or highly flexible. However, the effectiveness of these approaches hinges on the quality, diversity, and number of known actives. If only a few query ligands are available, or if they are too structurally homogeneous, the resulting models may fail to generalize, potentially overlooking novel chemotypes with distinct but functionally relevant properties. Consequently, while LBVS offers speed and simplicity, it is inherently limited by its dependence on historical data and may struggle to explore unexplored regions of chemical space.

### 1.3.2. Structure-Based Virtual Screening (SBVS)

Structure-based virtual screening (SBVS) uses the 3D structure of a target protein to predict how small molecules might bind to its active or allosteric sites (27). Usually, the protein's structure is determined through methods like X-ray crystallography, cryo-electron microscopy, or computational modeling. One of the most used techniques in SBVS is molecular docking, which computationally simulates the binding of candidate ligands to the target site (28). Molecular docking algorithms aim to generate plausible and optimized binding poses (conformations and orientations) for each ligand and estimate their binding affinities using a scoring function. Scoring functions for protein-ligand interactions are quite diverse, and usually can be classified as either physics-based, empirical, knowledge-based, or machine learning-based; each has advantages and disadvantages, which are thoroughly discussed in a review by Li et al. (29). Some examples of classical and widely used docking methods include AutoDock Vina (30), LeadFinder (31), and Glide (32).

Molecular docking offers a fast and scalable means of screening virtual libraries, and is particularly powerful for scaffold hopping, enabling the identification of novel chemotypes that exploit the same binding site in structurally distinct ways (33). Its effectiveness, however, depends heavily on the accuracy of the scoring function and the treatment of both ligand and receptor flexibility. Most docking tools simplify protein flexibility by considering the receptor as rigid or using limited side-chain adjustments. This can reduce the realism of predicted interactions and can especially hinder performance for targets with high conformational plasticity (34).

To address some of these limitations, SBVS can be complemented with molecular dynamics (MD) simulations, which simulate how protein-ligand interactions change over time (35). MD can refine poses generated by molecular docking by accounting for full atomic flexibility in both the protein and the ligand, and by including explicit water molecules. This allows the system to relax into a more realistic bound state (36). Moreover, MD enables the exploration of binding site flexibility, induced fit effects, and water-mediated interactions, which are often missed in rigid docking protocols.

For better prediction of binding affinities, several advanced scoring methods can be applied after docking. These include MM-GBSA (Molecular Mechanics with Generalized Born Surface Area) and MM-PBSA (Poisson–Boltzmann Surface Area), which estimate the binding free energy by combining molecular mechanics energies with solvation terms from implicit solvent models (37). Although more computationally demanding than standard scoring functions used with molecular docking, MM-GBSA and MM-PBSA generally show better correlation with experimental data and are commonly used to re-rank top candidates (38).

When greater precision is required, alchemical methods like Free Energy Perturbation (FEP) and Thermodynamic Integration (TI) can be used (39). These physics-based approaches simulate the gradual transformation of one ligand into another within the protein binding site, providing accurate estimates of relative binding free energies (RBFE) (13). Absolute binding free energies (ABFE) can also be accurately estimated using similar techniques, but these calculations are typically even more computationally demanding and sensitive to system setup (40). Although they are computationally expensive, FEP and TI have shown excellent accuracy in lead optimization and are increasingly applied in later-stage virtual screening workflows (41).

Despite these advancements, SBVS still faces various challenges. Scoring functions can produce false positives or fail to distinguish between high- and low-affinity binders. Docking pose prediction remains a stochastic process, with high variability across tools and targets (42). Additionally, ligands with very flexible scaffolds or targets where water molecules play key roles can prove especially challenging (43). Furthermore, reliance on rigid crystal structures can neglect relevant conformational states, especially in proteins with dynamic or cryptic binding sites. Nevertheless, when accurate structural information is available, SBVS remains a cornerstone of rational drug design, providing direct insight into ligand-target interactions and supporting the discovery of potent, novel molecules.

### 1.3.3. Machine Learning Approaches

Artificial intelligence (AI), machine learning (ML), and related data-driven approaches have proven effective in supporting modern virtual screening workflows (44). One of the earliest data-driven methods in drug discovery is quantitative structure–activity relationship (QSAR) modeling (45), which often is considered to be a predecessor of modern ML approaches in this field. Traditionally, QSAR has been used to build models that predict not only biological activity, but also physicochemical properties such as solubility or toxicity. These models were often made using either molecular descriptors—features and properties calculated from a molecule's structure, such as molecular weight, number of rotatable bonds, or polar surface area—or molecular fingerprints (46). Classical QSAR models typically used rather simple and interpretable algorithms, such as linear regression or decision trees (47). These approaches laid the groundwork for more advanced, data-driven methods.

Nowadays, advances in deep learning (DL) have enabled more sophisticated and flexible approaches to virtual screening. These methods move beyond predefined descriptors and simple models, as they often learn task-specific representations directly from the input data. One example is a graph neural network (GNN) model which was trained to predict antibacterial activity directly from molecular structure (48). This GNN model led to the discovery of halicin, a structurally novel antibiotic with broad-spectrum activity in both *in vitro* and *in vivo* settings, demonstrating the potential of deep learning to identify compounds with mechanisms distinct from known classes. Other efforts have focused on incorporating both ligand and protein information. For instance, iScore is a ML-based scoring function that predicts binding affinity using a combination of ligand descriptors and protein pocket features, without relying on detailed 3D binding poses (49). By circumventing the conformational sampling step typically required in molecular docking, iScore enables fast and efficient screening of large chemical libraries.

Rather than replacing docking altogether, some deep learning approaches aim to improve its core components. Gnina is a fork of AutoDock Vina that complements Vina's robust pose generation with a deep learning–based scoring function (50). The gnina scoring function uses a convolutional neural network

(CNN) that evaluates 3D protein–ligand grids to predict both binding affinity and pose quality (51). The CNN network learns spatial interaction patterns from atomic densities, leading to improved scoring accuracy. Another recent development is DiffDock, a deep learning–based docking method that formulates ligand binding as a generative process using diffusion models (52). Unlike traditional docking, which requires a predefined binding site, DiffDock performs blind docking by simultaneously predicting both the binding pose and pocket location from the protein structure alone. The model is trained to learn to reverse a noise-driven diffusion process to generate accurate ligand poses, showing strong performance on benchmark datasets and highlighting the potential of generative models in structure-based drug discovery.

Beyond pose prediction, generative models are also being developed to design novel molecules directly. DiffLinker, for example, employs a graph-based generative framework to design linkers that connect two fragments within a binding pocket, conditioned on both geometry and 3D structure of the protein–ligand complex (53). This approach enables fragment-based design tailored to specific spatial constraints. DeepTarget, on the other hand, generates novel molecules from protein sequences alone, without requiring structural information or known ligands (54). It combines transformer-based protein encoding with a generative adversarial network (GAN) architecture to infer ligand-like molecules that exhibit favorable docking and affinity scores against target proteins.

Despite their promise, the effectiveness of ML approaches remains closely tied to the quality, diversity, and coverage of training data. Models often perform well within the chemical and biological domains they were trained on, but tend to generalize poorly to novel scaffolds or targets outside the training distribution (55). Limited data, experimental noise, and biases in available datasets remain persistent challenges (56). Still, when combined with domain knowledge and experimental feedback, ML and data-driven methods provide powerful tools for accelerating virtual screening, enriching hit lists, and improving early decision-making in the drug discovery process.

## 1.4. SCALING VIRTUAL SCREENING: THE ROLE OF HIGH-PERFORMANCE COMPUTING

The ability to virtually screen millions of compounds requires not only efficient algorithms but also robust and scalable computational infrastructure. High-performance computing (HPC) environments are crucial in this context, enabling the parallel execution of virtual screening workflows (57).

In many academic and industrial clusters, SLURM (Simple Linux Utility for Resource Management) serves as the job scheduler that distributes tasks across hundreds or thousands of compute nodes (58). Researchers also rely on container technologies such as Singularity to package their software environments, ensuring that docking or prediction pipelines run the same way on any system and improving both portability and reproducibility (59).

As described in section 1.3.1, Ligand-based screening methods compare new compounds to known actives using techniques like fingerprint similarity or three-dimensional shape matching. Each comparison can be run independently, making the calculations embarrassingly parallel (60). This means that these calculations can be distributed efficiently across many cores or accelerated on GPUs with almost no additional overhead. With enough computing power, researchers can screen vast chemical libraries in hours rather than weeks.

Similarly, structure-based methods such as molecular docking are well-suited for parallel execution. Each docking calculation can run in isolation, allowing near-linear scaling with additional computational resources. In practice, the target protein is prepared once, and a virtual compound library is generated with defined protonation states, explicit hydrogens, and valid initial 3D conformers. Software such as MetaScreener can distribute and launch docking jobs across a SLURM-managed cluster, enabling efficient and scalable screening of large libraries (61).

Furthermore, ML–based virtual screening methods can take advantage of GPU acceleration, which is increasingly available in modern HPC clusters. In particular, inference with deep learning models can see substantial performance gains when run on GPUs, enabling fast and efficient screening once models are trained.

Ultimately, the success of virtual screening at scale depends not only on methodological advances but also on the seamless integration of infrastructure, automation, and computational resources. As chemical libraries continue to grow in size and complexity, access to scalable and distributed computing environments will remain a critical enabler for efficient, high-throughput screening in both academic and industrial drug discovery settings.

## 1.5. LIMITATIONS AND SHORTCOMINGS

Despite their widespread adoption, virtual screening methods are often applied in isolation, resulting in fragmented workflows and, in many cases, suboptimal or inconsistent outcomes. Each method has its own strengths but also inherent limitations. Ligand-based approaches depend heavily on known actives and chemical similarity, making them prone to overlooking novel scaffolds that diverge from established structure–activity patterns. Structure-based methods, such as molecular docking can produce inaccurate results due to limitations in scoring functions and errors in pose predictions, leading to false positives or missed opportunities. ML models, although increasingly sophisticated, are highly sensitive to the quality, diversity, and bias of training data, and often struggle to generalize across new targets or unexplored chemical space.

As a result, the reliability and reproducibility of virtual screening outcomes can vary substantially depending on the chosen strategy. These challenges highlight the potential of consensus-based VS strategies, which seek to combine the strengths of different methods while mitigating their individual weaknesses. For instance, integrating compound rankings from multiple docking algorithms has been shown to improve the enrichment of active compounds (62). This principle of consensus, whether applied to docking, ligand similarity, or ML predictions, offers a promising direction for improving VS performance. By integrating results across methods, researchers can try to reduce false positives, improve hit prioritization, and gain greater confidence in their findings.

# II – OBJECTIVES

# II - OBJECTIVES

This thesis was guided by several objectives aimed at improving virtual screening methodologies in drug discovery. The studies focused on both ligand-based and structure-based approaches, with the goal of improving reliability, early enrichment, and chemical diversity in hit identification. This was mainly investigated through the use of consensus approaches, which combine the outputs of different docking algorithms or molecular fingerprint distances with the goal of reducing individual biases and improving predictive accuracy. Additionally, in an effort to support machine learning applications, the data quality and assay noise present in open-source bioactivity databases like ChEMBL was also investigated as one of the objectives. Besides method development, the application of the tools in real-world drug discovery projects were also considered an integral part of the research. The specific objectives are formulated below:

**Objective 1: Development of consensus-based docking method to improve structure-based virtual screening**

The aim of this objective is to improve the reliability and early enrichment of structure-based virtual screening methods by integrating results from multiple docking algorithms. Incorporating metrics such as pose similarity, docking scores, and ligand flexibility will be thoroughly investigated. The final method should be flexible, scalable, and suitable for deployment on high-performance computing (HPC) systems through an accessible command-line interface.

**Objective 2: Investigate consensus fingerprint similarity for flexible ligand-based virtual screening**

This objective targets the improvement of fingerprint-based virtual screening by combining different types of molecular fingerprints in a consensus ranking strategy. The goal is to enhance predictive performance while maintaining chemical diversity and minimizing methodological bias. The resulting method should be lightweight, user-friendly, and capable of efficiently screening large and diverse compound libraries.

**Objective 3: Investigate data noise in public bioactivity data sets to improve QSAR and machine learning applications**

This objective aims to explore assay variability, and the impact on the reliability of public bioactivity data. More specifically, the goal is to quantify systematic noise patterns and investigate strategies which can help to reduce it. The study should inform best practices for data curation and help facilitate the development of improved QSAR and machine learning models in the field of cheminformatics.

**Objective 4: Applying the developed virtual screening methods to proprietary, public, and commercial compound libraries**

This objective focuses on evaluating the performance and generalizability of the developed virtual screening methods in practical applications across a range of chemical libraries. On the one hand, the tools will be applied to the proprietary Eurofins–Villapharma compound library to support in-house drug discovery efforts and guide the optimization of early leads. On the other hand, the methods will also be tested on external libraries, including large commercial collections such as Enamine and MolPort to explore broad chemical space, as well as more focused datasets like DrugBank that are particularly relevant for drug repurposing.

# III – RESEARCH ARTICLES COMPRISING THIS THESIS

# III - RESEARCH ARTICLES COMPRISING THIS THESIS

## 3.1. THEMATIC GROUNDS FOR THE COMPENDIUM OF PUBLICATIONS

The core of this doctoral thesis is structured as a collection of three scientific articles, all of which have been published in high-quality, peer-reviewed journals. The articles are presented in the most suitable sequence to illustrate the methodological development and thematic consistency of the thesis. While every article discusses a particular topic of computational drug discovery, they all have a common goal: the enhancement of virtual screening methods for improved identification of bioactive compounds.

This thesis investigates consensus-based approaches for ligand- and structure-based virtual screening with the aim of enhancing performance. The research also considers the computational challenges of large-scale screening, aiming for a flexible and scalable workflow design, compatible with high-performance computing environments.

The first article details ESSENCE-Dock, which focuses on improving structure-based virtual screening through consensus docking approaches. The second manuscript features the application of ConFiLiS, a ligand-based screening method that uses consensus ranking from different molecular fingerprints. The third considers the impact of experimental noise in publicly available bioactivity data and discusses potential implications for the broader cheminformatics and machine learning communities.

Together, these developments form a coherent body of work that places virtual screening as both scientific and practical drug discovery methodology. Thematic unity of the compendium lies in the overall focus throughout on improving reliability, scalability, and applicability of computational screening protocols at multiple stages of the drug discovery pipeline.

## 3.2. ESSENCE-DOCK: A CONSENSUS-BASED APPROACH TO ENHANCE VIRTUAL SCREENING ENRICHMENT IN DRUG DISCOVERY

| | |
|---|---|
| **Title** | ESSENCE-Dock: A Consensus-Based Approach to Enhance Virtual Screening Enrichment in Drug Discovery |
| **Authors** | Jochem Nelen; Miguel Carmena-Bargueño; Carlos Martínez-Cortés; Alejandro Rodríguez-Martínez; José Manuel Villalgordo-Soto; Horacio Pérez-Sánchez |
| **Journal** | Journal of Chemical Information and Modeling |
| **Year** | 2024 |
| **Volume** | 64 (5) |
| **Pages** | 1605-1614 |
| **DOI** | https://doi.org/10.1021/acs.jcim.3c01982 |
| **IF (JCR 2024)** | 5.3 |
| **Category** | Chemistry, Medicinal; 15/72; Q1<br>Chemistry, Multidisciplinary; 65/239; Q2<br>Computer Science, Information Systems; 46/258; Q1<br>Computer Science, Interdisciplinary Applications; 37/175; Q1 |

**Contribution of PhD Candidate**

The PhD candidate, Jochem Nelen, confirms that he is the first author of the article titled "*ESSENCE-Dock: A Consensus-Based Approach to Enhance Virtual Screening Enrichment in Drug Discovery*". He was primarily responsible for the conception of the study, development and implementation of the methodology, data analysis, manuscript writing, and overall coordination of the research work.

JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

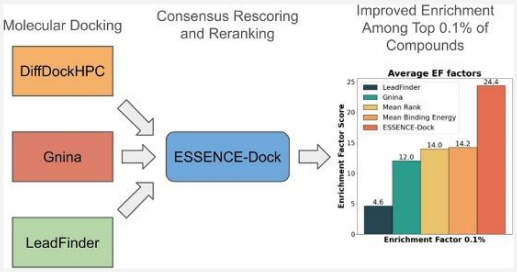
# ESSENCE-Dock: A Consensus-Based Approach to Enhance Virtual Screening Enrichment in Drug Discovery

Jochem Nelen, Miguel Carmena-Bargueño, Carlos Martínez-Cortés, Alejandro Rodríguez-Martínez, José Manuel Villalgordo-Soto, and Horacio Pérez-Sánchez*

Cite This: J. Chem. Inf. Model. 2024, 64, 1605−1614 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

**ABSTRACT:** Drug development is a complex, costly, and time-consuming endeavor. While high-throughput screening (HTS) plays a critical role in the discovery stage, it is one of many factors contributing to these challenges. In certain contexts, virtual screening can complement the HTS, potentially offering a more streamlined approach in the initial stages of drug discovery. Molecular docking is an example of a popular virtual screening technique that is often used for this purpose; however, its effectiveness can vary greatly. This has led to the use of consensus docking approaches that combine results from different docking methods to improve the identification of active compounds and reduce the occurrence of false positives. However, many of these methods do not fully leverage the latest advancements in molecular docking. In response, we present ESSENCE-Dock (Effective Structural Screening ENrichment ConsEnsus Dock), a new consensus docking workflow aimed at decreasing false positives and increasing the discovery of active compounds. By utilizing a combination of novel docking algorithms, we improve the selection process for potential active compounds. ESSENCE-Dock has been made to be user-friendly, requiring only a few simple commands to perform a complete screening while also being designed for use in high-performance computing (HPC) environments.






## INTRODUCTION

The development of novel drugs is a very long and expensive process: depending on the context, the costs of bringing new drugs to market can span from several hundred million to over 4 billion US dollars.[1] Many factors contribute to these high costs, including costly clinical trials and toxicity studies, among others. Another part of the high costs can be attributed to hit discovery through expensive, time-consuming high-throughput screening (HTS).[2] Screening small libraries in this manner is done routinely, but chemical space is vast, and navigating it using HTS exclusively can be very costly, in terms of both time and money.

Virtual screening techniques are capable of computationally identifying potential hit compounds, thereby decreasing the number of compounds to test experimentally. This can be done much faster and cheaper than HTS and thus applied to many more compounds, allowing the exploration of a larger chemical space.[3] In short, virtual screening techniques are valuable tools for lowering costs and speeding up novel medicine development. Many different types of virtual screening methods exist, but one of the most used techniques is molecular docking.[4] For this structure-based screening method, a 3D structure of both the target protein and the ligand needs to be available. The goal of molecular docking is to find the optimal ligand pose, which typically has an associated docking score. Subsequently, this docking score can be used to compare different ligands and rank them to obtain a list of the most promising candidates.

The performance of docking algorithms can vary significantly across different targets.[5] Examples of popular docking methods include AutoDock Vina,[6] and its subsequent forks Smina[7] and Gnina,[8] but also commercial software such as LeadFinder[9] and Glide[10] are routinely used in both industry and academia. These methods can be effective on their own, but consensus docking approaches have been shown to improve reliability.[11] More specifically, by implementing a consensus approach, one can reduce false positives and improve the overall accuracy of results. However, many of these existing consensus methods rely on older docking techniques with a lower accuracy. Furthermore, they often

Received: December 12, 2023
Revised: February 9, 2024
Accepted: February 15, 2024
Published: February 28, 2024





https://doi.org/10.1021/acs.jcim.3c01982
J. Chem. Inf. Model. 2024, 64, 1605−1614

Journal of Chemical Information and Modeling pubs.acs.org/jcim Article

**Table 1. Comparative Performance of ESSENCE-Dock and Reference Docking Methods across 21 DUD-E Targets at Various Enrichment Factors (EF%)[a]**

| | EF5% | | | | | EF1% | | | | | EF0.5% | | | | | EF0.1% | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | LF | GN | MR | MBE | Cons. | LF | GN | MR | MBE | Cons. | LF | GN | MR | MBE | Cons. | LF | GN | MR | MBE | Cons. |
| ADA | 3.9 | 3.2 | 4.3 | 4.1 | 5.6 | 1.1 | 0.0 | 9.8 | 5.4 | 11.9 | 0.0 | 0.0 | 8.5 | 4.3 | 21.3 | 0.0 | 0.0 | 9.9 | 0.0 | 39.7 |
| AKT1 | 1.7 | 3.6 | 2.7 | 3.3 | 1.4 | 1.7 | 8.6 | 4.4 | 6.5 | 2.7 | 2.7 | 12.2 | 6.8 | 8.8 | 2.7 | 0.0 | 26.9 | 6.7 | 16.8 | 3.4 |
| AMPC | 0.8 | 0.8 | 0.8 | 0.0 | 2.9 | 0.0 | 0.0 | 0.0 | 0.0 | 2.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| CASP3 | 2.9 | 4.2 | 5.5 | 5.0 | 5.0 | 9.5 | 4.5 | 10.6 | 11.1 | 9.5 | 11.2 | 4.1 | 13.2 | 15.2 | 17.2 | 5.0 | 0.0 | 14.9 | 24.9 | 39.8 |
| CP3A4 | 2.8 | 1.5 | 2.4 | 2.9 | 2.0 | 4.7 | 1.8 | 2.3 | 4.1 | 1.8 | 8.2 | 2.3 | 4.7 | 7.0 | 0.0 | 11.7 | 5.9 | 5.9 | 5.9 | 0.0 |
| CXCR4 | 2.0 | 0.5 | 1.0 | 0.5 | 0.5 | 0.0 | 0.0 | 2.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| FA7 | 5.6 | 8.6 | 9.0 | 9.3 | 7.5 | 8.7 | 14.8 | 20.1 | 20.9 | 17.4 | 7.0 | 17.4 | 29.7 | 24.4 | 22.7 | 0.0 | 0.0 | 55.8 | 37.2 | 55.8 |
| FABP4 | 0.4 | 8.1 | 3.4 | 5.5 | 6.4 | 0.0 | 21.3 | 2.1 | 8.5 | 10.6 | 0.0 | 25.5 | 0.0 | 8.5 | 8.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| FPPS | 0.0 | 5.6 | 0.0 | 0.0 | 9.4 | 0.0 | 2.4 | 0.0 | 0.0 | 20.1 | 0.0 | 0.0 | 0.0 | 0.0 | 28.0 | 0.0 | 0.0 | 0.0 | 0.0 | 11.7 |
| GCR | 0.7 | 5.3 | 2.9 | 3.9 | 4.7 | 0.0 | 15.8 | 6.2 | 13.5 | 13.9 | 0.0 | 21.8 | 7.0 | 21.0 | 21.0 | 0.0 | 47.3 | 7.9 | 23.7 | 27.6 |
| GLCM | 5.5 | 1.1 | 3.7 | 4.1 | 5.5 | 11.0 | 1.8 | 3.7 | 11.0 | 5.5 | 11.3 | 0.0 | 3.8 | 15.0 | 7.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| HIVPR | 5.1 | 4.5 | 7.1 | 7.0 | 6.5 | 5.4 | 6.2 | 14.2 | 11.2 | 9.7 | 4.5 | 6.0 | 17.6 | 14.2 | 12.0 | 3.8 | 9.4 | 26.3 | 24.4 | 26.3 |
| HIVRT | 2.5 | 3.3 | 3.8 | 3.8 | 4.2 | 2.7 | 3.6 | 5.6 | 5.6 | 8.9 | 1.2 | 4.7 | 7.1 | 5.9 | 13.0 | 0.0 | 9.0 | 9.0 | 6.0 | 18.0 |
| HS90A | 1.8 | 0.0 | 0.2 | 0.5 | 3.9 | 2.3 | 0.0 | 0.0 | 0.0 | 8.0 | 2.2 | 0.0 | 0.0 | 0.0 | 13.5 | 0.0 | 0.0 | 0.0 | 0.0 | 11.2 |
| ITAL | 3.3 | 3.2 | 2.9 | 3.3 | 3.0 | 8.0 | 11.6 | 11.6 | 13.1 | 9.5 | 13.1 | 8.7 | 21.8 | 20.4 | 17.5 | 7.0 | 7.0 | 55.6 | 48.7 | 48.7 |
| KIF11 | 1.2 | 8.3 | 6.9 | 7.8 | 7.4 | 0.0 | 21.4 | 8.6 | 16.3 | 24.0 | 0.0 | 30.9 | 6.9 | 17.2 | 34.3 | 0.0 | 25.7 | 0.0 | 0.0 | 25.7 |
| MK14 | 2.5 | 3.3 | 3.3 | 3.9 | 2.4 | 5.2 | 9.2 | 6.6 | 8.3 | 5.4 | 5.9 | 13.5 | 9.7 | 12.8 | 9.3 | 1.8 | 31.5 | 21.0 | 29.8 | 29.8 |
| NRAM | 0.6 | 1.2 | 3.1 | 2.0 | 7.5 | 0.0 | 0.0 | 2.0 | 0.0 | 17.3 | 0.0 | 0.0 | 0.0 | 0.0 | 29.0 | 0.0 | 0.0 | 0.0 | 0.0 | 53.6 |
| PA2GA | 2.4 | 5.3 | 1.6 | 1.8 | 8.9 | 0.0 | 6.1 | 1.0 | 1.0 | 25.5 | 0.0 | 6.1 | 2.0 | 2.0 | 30.6 | 0.0 | 10.6 | 10.6 | 10.6 | 31.8 |
| TRY1 | 3.0 | 5.9 | 6.0 | 6.1 | 4.0 | 3.6 | 6.5 | 6.7 | 8.2 | 8.2 | 3.1 | 7.6 | 7.6 | 9.4 | 12.0 | 6.8 | 18.1 | 9.1 | 9.1 | 27.2 |
| WEE1 | 14.3 | 17.6 | 15.3 | 16.8 | 17.6 | 41.8 | 60.3 | 61.3 | 61.3 | 53.5 | 47.5 | 61.3 | 61.3 | 61.3 | 59.3 | 61.3 | 61.3 | 61.3 | 61.3 | 61.3 |
| Average | 3.0 | 4.5 | 4.1 | 4.4 | 5.5 | 5.0 | 9.3 | 8.5 | 9.8 | 12.6 | 5.6 | 10.6 | 9.9 | 11.8 | 17.1 | 4.6 | 12.0 | 14.0 | 14.2 | 24.4 |
| Median | 2.5 | 3.6 | 3.3 | 3.9 | 5.0 | 2.3 | 6.1 | 5.6 | 8.2 | 9.5 | 2.2 | 6.0 | 6.9 | 8.8 | 13.5 | 0.0 | 5.9 | 7.9 | 6.0 | 26.3 |

[a]Compares the performance of our ESSENCE-Dock consensus approach (Cons.) with reference methods (LeadFinder—LF, Gnina—GN, mean rank—MR, and mean binding energy—MBE) across 21 DUD-E targets at various enrichment factors (EF%). The table presents EF values for the top 5%, top 1%, top 0.5%, and top 0.1%. Shaded cells indicate method rankings, with green denoting higher comparative EF and red indicating lower comparative EF.

overlook the importance of the predicted pose similarity, despite its demonstrated effectiveness.[12]

To provide a brief overview of existing consensus approaches, we begin with the early yet effective work by Kukol.[13] In this work, different docking rankings from different algorithms were combined in order to determine which combination of docking methods is the most effective. Specifically, the Directory of Useful Decoys (DUD)[14] was used for this purpose. It was reported that AutoDock 4[15] combined with AutoDock Vina yielded the best results overall. Houston and Walkinshaw[16] built on these findings by incorporating a simple form of binding pose similarity for additional consensus, which proved to be quite effective at lowering false positives. Their method started by performing molecular docking using both AutoDock Vina and AutoDock 4. Subsequently, they quantified the binding pose similarity by calculating the root mean square deviation (RMSD) between the binding poses of each ligand. In case the RMSD was larger than 2 Å, the molecule was rejected and filtered out from further calculations due to the disagreement between the two docking methods. If the RMSD was smaller than or equal to 2 Å, the compounds were kept for further analysis and ranking. Although simple, this filter step based on RMSD pose similarity managed to identify many of the decoy compounds, thus demonstrating its value in the whole protocol.

After these developments, consensus docking has been continuously further explored by many researchers. A brief overview of this has been provided in a review of molecular docking written by Torres et al.[11] Some of the most recent consensus docking publications include DockECR[17] and MILCDock.[18] Palacio-Rodríguez et al. described an exponential consensus ranking (ECR) method that uses an exponential distribution for each individual rank of every docking method.[19] Their method is implemented in such a way that compounds which score very well in a single method can still be ranked highly even though they score worse in the other docking methods. The DockECR publication is an in-depth exploration of the ECR method and was collaboratively written with one of the original authors of the ECR paper. They developed an open-source program that has implemented the ECR method and supports several docking methods, such as LeDock,[20] rDock,[21] Smina,[7] and AutoDock Vina.[6] Additionally, they added a binding pose similarity metric, which was shown to improve the results in most cases. Their protocol also supports the use of parallelization, enabling simultaneous docking runs across multiple CPUs.



https://doi.org/10.1021/acs.jcim.3c01982
J. Chem. Inf. Model. 2024, 64, 1605−1614

Finally, one of the latest works regarding consensus docking is MILCDock. Here, the authors describe a consensus docking method that is enhanced by machine learning. To gather training data, the authors performed molecular docking on data sets such as DUD-E[22] using various open-source docking algorithms including AutoDock Vina, AutoDock 4, and rDock. Subsequently, they used the resulting docking data and binding pose similarities to train machine learning models that can be used as scoring function. Their results show substantial improvements when compared to the individual docking runs.

Inspired by this prior research, we developed ESSENCE-Dock: a docking workflow integrating data generated by different docking methods, effectively reducing false positives and enhancing the enrichment of active compounds. This is achieved by prioritizing consensus from different docking techniques, in both docking scores and binding poses, effectively leveraging the strengths of each docking method while mitigating a lot of the individual limitations. High consensus scores from our protocol correlate with a higher likelihood of compound activity, as demonstrated by our reported high enrichment factors across various DUD-E targets compared to the base methods (Table 1). These observed enrichment factors highlight its promising efficacy and applicability in the hit identification process. Our workflow is flexible, user-friendly, and able to be performed using high-performance computing (HPC) clusters. The entire screening process can be streamlined to just four commands: one for each of the individual docking methods—DiffDock,[23] Gnina,[8] and LeadFinder—[9]and a fourth to initiate the actual consensus calculations, which performs the final rescoring. For this, ESSENCE-Dock uses the information from the individual docking runs such as docking scores, binding pose similarity, and ligand flexibility in order to compute a comprehensive consensus score, which is used to finally rerank all of the screened compounds. Additionally, all of the top results can be saved to an interactive PyMOL Session (PSE) file[24] for visual inspection and user convenience. Furthermore, predicted protein−ligand interactions for these top results are calculated using the Protein−Ligand Interaction Profiler (PLIP)[25] and added to the PSE file as well. These protein−ligand interaction results are also saved in tabular and JSON formats for easy integration with potential further postdocking analysis.

In summary, we introduce ESSENCE-Dock (Effective Structural Screening ENrichment ConsEnsus Dock), a novel consensus docking workflow that addresses some of the identified gaps in current consensus docking methods. ESSENCE-Dock leverages state-of-the-art docking techniques and takes into account binding pose similarity and ligand flexibility to calculate the final consensus score. Our workflow is easy to use, only requiring a few commands for the whole protocol. It is also able to run on HPC clusters using Singularity, and has the potential to streamline drug discovery efforts by focusing resources on acquiring and testing compounds with high consensus scores.

## ■ MATERIALS AND METHODS

As established in Introduction, consensus analysis of docking techniques can substantially improve the outcome of a virtual screening campaign.[4] With this in mind, we set out to develop our own consensus docking procedure with a focus on lowering false positives in order to obtain a high enrichment in the top fraction of high-scoring compounds.

As a first step, we attempted to select different types of docking techniques. We opted for methods that make use of very different sampling methods for pose generation and use different types of scoring functions. This makes it more significant when consensus, in both high docking scores and binding poses, is found because the docking algorithms came to the same conclusion by employing completely different methods. We ended up selecting LeadFinder,[9] Gnina,[8] and DiffDock.[23] These docking methods, along with their specific sampling and scoring algorithms, are discussed in more detail below. A schematic overview of the general workflow is provided in Figure 1.

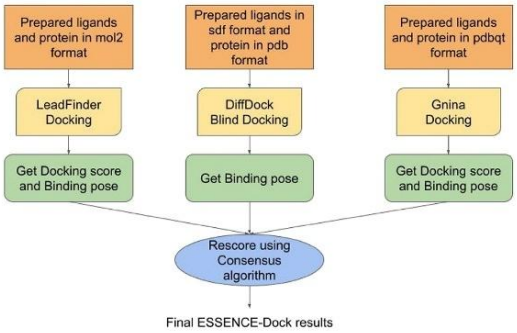



**Figure 1.** Overview of the ESSENCE-Dock Workflow. Each ligand is prepared for and docked using LeadFinder, DiffDock, and Gnina. The resulting docking scores and binding poses are subsequently combined into a final ESSENCE-Dock score, which is used for ranking the compounds.

**Ligand and Protein Preparation.** Proper preparation of protein and ligands is a crucial step when performing molecular docking and helps to ensure optimal results. Each of the docking techniques used in our work, LeadFinder (LF), DiffDock (DD), and Gnina (GN), requires distinct formats for both ligands and proteins. LeadFinder requires the use of the mol2 format for both ligands and proteins, DiffDock performs best with ligands in SDF format and protein structures in PDB format, while Gnina provides optimal results with structures in pdbqt format for both proteins and ligands. This section will detail the procedures used to prepare the ligands and proteins for all of the docking calculations, starting with ligand preparation.

*Ligands.* The selected DUD-E target sets were downloaded from the DUD-E database, with the files "decoys_final.ism" and "actives_final.ism" serving as the starting point for conversion (see Figure 2).

Using Python and RDKit,[26] the smiles were processed, hydrogens were added, and a 3D conformation was generated for each compound using RDKit's ETKDG method. The molecules were saved as separate SDF files. Each file was named according to the ID provided in the original.ism files and was tagged as "active" or "inactive" for more efficient postprocessing after docking.

These SDF files were used as ligands for the DiffDock method but were also converted to mol2 files using ChemAxon's Molconvert tool.[27] This generated 3D coordinates of a low-energy conformer using the MMFF94 force



https://doi.org/10.1021/acs.jcim.3c01982
J. Chem. Inf. Model. 2024, 64, 1605−1614

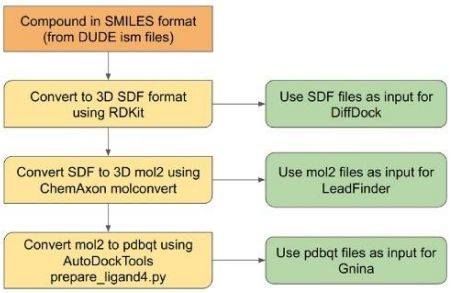



**Figure 2.** Ligand preparation workflow. Input compounds in SMILES format were processed and converted to 3D SDF, mol2, and pdbqt files for docking with DiffDock, LeadFinder, and Gnina, respectively.

field,[28] making the molecules ready to be docked directly using LeadFinder.

Subsequently, these mol2 files were used as input for the conversion to pdbqt format using the "prepare_ligand4.py" script from AutoDockTools (ADT),[15] which were then used for docking with Gnina.

*Protein Structures.* For the protein preparation of the DUD-E targets, the PDB code of the provided receptor.pdb file was used to obtain the final receptor files. A schematic overview of the protocol is given in Figure 3.

The first step after importing the PDB structure in Schrödinger's Maestro (version 2020-4)[29] was to remove all solvents, ions, and ligands. Subsequently, the Protein Preparation Wizard in Maestro was used to properly (re)assign bond orders, add missing hydrogens, create disulfide bonds where necessary, and fill in missing side chains. Next, Maestro's system builder was used to charge the protein using the OPLS3e force field.[30] Concretely, this was accomplished by minimizing the box volume, defining the solvent model as none, the box shape as orthorhombic, and a 10 Å buffer along every dimension.

After these steps, the processed protein was saved in both mol2 and pdb format. The mol2 file was directly used as the receptor file for the LeadFinder docking calculations, while the pdb file was used as an input for DiffDock. The pdb file also underwent further processing to create a pdbqt file using AutoDockTools (ADT) from MGLTools (version 1.5.7).[15] Concretely, hydrogens were added with the settings "All hydrogens" and "noBondOrder", followed by assigning the AD4 type to all atoms. As a last step, the Gasteiger charge was computed, and the resulting file was saved as a.pdbqt file for use in the Gnina docking process.

**Docking Techniques.** In the ESSENCE-Dock consensus approach, we deliberately utilized docking methods with different sampling and scoring functions, enhancing the significance of consensus findings in both docking scores and binding poses. Here, we will briefly discuss the individual docking techniques used for the current implementation of our ESSENCE-Dock workflow, detailing their methodology for pose sampling and scoring functions as well as the docking parameters we used to process the input ligands.

*LeadFinder.* LeadFinder (LF), developed by BioMolTech and distributed by Cresset, uses a speed- and accuracy-enhanced genetic algorithm for ligand conformation generation.[9] This algorithm creates a diversified pool of conformations that are subsequently ranked using the LF scoring function. The output includes the best compounds with their docking score, predicted binding energy, and corresponding docked pose in mol2 format. LF's scoring function is semiempirical and explicitly accounts for various types of molecular interactions.[9] Overall, LF is a performant docking algorithm that has been shown to perform well across various protein targets.[31]

We ran the LF calculations using MetaScreener v1.1,[32] an open-source and publicly available tool that allows a user to perform molecular docking on high-performance computing (HPC) clusters. This meant that LF could run in parallel, reducing the overall run time substantially and making it feasible to screen larger libraries. The LF docking was executed using LeadFinder version 2104, along with standard variables and a grid size, where *X*, *Y*, and *Z* were all set to 30. As user input, only the docking coordinates, as well as the protein structure and the directory containing the input ligands, were included as input variables.

*Gnina.* Another docking method we used in our consensus docking approach is Gnina.[8] Gnina is a fork of Smina,[7] which in turn is a fork of AutoDock Vina (Vina).[6] In contrast with LeadFinder, which utilizes a genetic algorithm to explore a ligand's conformational space, Gnina employs Monte Carlo sampling to generate a diverse ensemble of ligand conformations. It also implemented an ensemble of convolutional neural networks (CNN) as a scoring function, which outperforms the scoring functions of its predecessors Vina and Smina in most cases.[8] Although Vina's scoring function is still

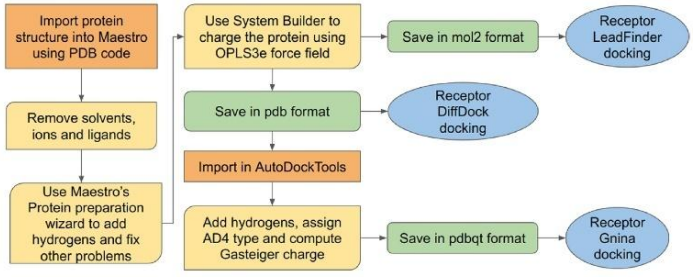



**Figure 3.** Protein structure preparation workflow. Protein structures are imported using their PDB code in Maestro. After processing, the structures are saved as both PDB and mol2 files, which are used for DiffDock and LeadFinder, respectively. Additionally, the PDB file is converted to pdbqt format using AutoDockTools for Gnina docking.



https://doi.org/10.1021/acs.jcim.3c01982
J. Chem. Inf. Model. 2024, 64, 1605−1614

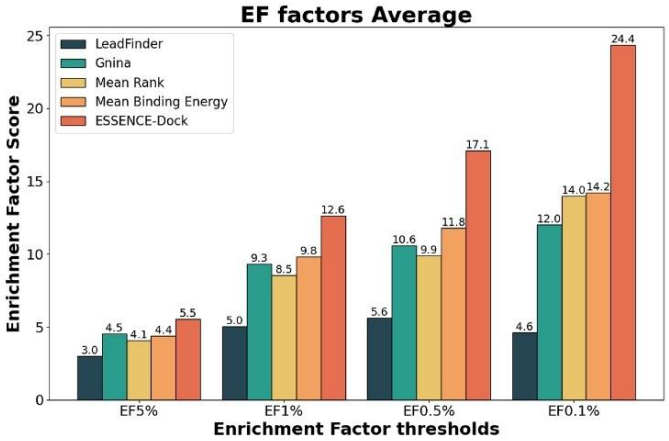



**Figure 4.** Bar chart of average enrichment factors (EF) across 21 various DUD-E targets at various EF thresholds. This figure provides a comparative analysis of the average EFs across 21 distinct DUD-E targets. The *Y*-axis represents the average EF, while the *X*-axis indicates the corresponding EF threshold values. The color-coded bars correspond to the different ranking methods: LeadFinder, Gnina, Mean Rank, Mean Binding Energy, and ESSENCE-Dock.

used during pose generation and refinement for computational efficiency, Gnina's CNN pose score ensemble model rescores and sorts the final poses. These settings, found to be the most efficient by the authors, were also used for our Gnina docking calculations. As a metric for the final predicted binding energy, we used the Vina score of the most optimal pose, as determined by the CNN pose score for further processing.

For our calculations, the Gnina docking algorithm was implemented in MetaScreener v1.1, again providing support to run the calculations massively in parallel on HPC clusters. Because of the CNN scoring function, Gnina is able to run a lot faster on a GPU compared to a CPU. Despite this, the parallel execution of calculations on an HPC cluster made processing the selected DUD-E data sets still feasible. Gnina Version 1.0 was used with standard settings, a box size of 30 in all dimensions, along with the user-provided protein, directory of ligands, and docking coordinates.

*DiffDock.* The final docking technique we integrated into our consensus docking workflow is DiffDock.[23] DiffDock is a novel, state-of-the-art blind docking method based on geometric deep learning. DiffDock refines randomized ligand conformations at various protein locations via a trained reverse diffusion process, simultaneously determining the best binding pocket and conformation. This eliminates the need for defining a predesignated search box, as required by LeadFinder and Gnina. Because this is a blind docking method, the whole protein and all of its potential binding pockets are explored.

Unlike the other methods, DiffDock does not provide a scoring function. It does offer a confidence score indicating prediction certainty. Although this is useful, it does not allow a user to directly estimate how potent a ligand binds, making it difficult to rank different ligands accurately among each other. Nonetheless, DiffDock is capable of predicting a ligand's binding pose with state-of-the-art accuracy, so it is still very useful to compare it to the binding poses generated by LeadFinder and Gnina.

To facilitate the use of DiffDock in virtual screening, we developed a fork of DiffDock named DiffDockHPC. Like MetaScreener, it can split up the ligand data set and launch it as separate jobs, meaning the calculations can run in parallel. This significantly speeds up calculations and is especially useful when processing more extensive libraries. DiffDockHPC is fully open-source and freely available at https://github.com/Jnelen/DiffDockHPC. Like Gnina, DiffDock can be greatly accelerated by GPUs, but running the calculations massively in parallel on CPUs still makes it feasible for it to be a part of the consensus docking workflow.

**Processing of Docking Results.** In this section, we describe the processing of docking results, the details of our scoring formula, and the final output files that our ESSENCE-Dock workflow generates.

*Scoring Formula.* In order to effectively rank compounds, we developed a scoring formula that considers not only individual docking scores but also incorporates the binding pose similarity and the ligand's flexibility. The complete ESSENCE-Dock rescoring formula is presented in eq 1:

$$\text{ESSENCE-Dock Score} = \sqrt{\frac{2}{\text{RMSD average}}} \times \text{MBE} \times \sqrt{\ln(\text{RB} + 1) + 1} \quad (1)$$

Here, RMSD average denotes the average RMSD (as shown in eq 2) and MBE represents the mean binding energy (computed as per eq 3). Both $\text{BindingEnergy}_{\text{GN}}$ and $\text{BindingEnergy}_{\text{LF}}$ are denoted in kcal/mol. Gnina's final docking score is based on Vina's scoring function, applied to the most optimal pose as determined by Gnina's CNN pose score. RB corresponds to the number of rotatable bonds. The RMSD and the number of rotatable bonds are calculated using OpenBabel's obrms and obrotamer functionalities, respectively.[33] During the RMSD calculations, obrms takes



https://doi.org/10.1021/acs.jcim.3c01982
J. Chem. Inf. Model. 2024, 64, 1605−1614

potential internal molecular symmetry into account, providing more robust and accurate results.

$$\text{RMSD average} = \frac{\text{RMSD}_{\text{DD\_GN}} + \text{RMSD}_{\text{DD\_LF}} + \text{RMSD}_{\text{GN\_LF}}}{3} \quad (2)$$

$$\text{MBE} = \frac{\text{BindingEnergy}_{\text{GN}} + \text{BindingEnergy}_{\text{LF}}}{2} \quad (3)$$

The core principle underpinning the ESSENCE-Dock formula is the concept of consensus, which includes both predicted binding energies and binding poses. Our starting point was a consensus of the docking scores: the mean binding energy, which showed the best performance when comparing the individual docking methods along with their mean and median scores (Figure 4). With this as the starting point, we attempted to improve this baseline by including a form of binding pose consensus. In the literature, an RMSD difference smaller than or equal to 2 is often chosen as a threshold for similar poses.[11] Hence, we incorporated the concept of binding pose consensus by dividing 2 by the average RMSD. If the average RMSD is less than 2, the overall score improves; otherwise, the final score will become worse. To prevent a disproportionate impact of really low or high consensus values, the square root is taken to balance out the impact of extremes. Finally, to mitigate a potential bias toward rigid structures with few rotational bonds, we also introduced a flexibility term based on the number of rotational bonds.

To conclude, we point out that only the average RMSD, docking scores, and number of rotatable bonds are required as input metrics. In this sense, our approach is flexible and can be used with basically any other docking program that produces a docking pose and score. For user convenience, the whole ESSENCE-Dock approach has been implemented in MetaScreener v1.1, meaning that the consensus rescoring and postprocessing can be executed using just a single command.

*Enrichment Factors.* To validate the protocol, we selected a diverse set of DUD-E targets and computed enrichment factors (EFs) at different thresholds. We compared these EFs with those obtained from individual docking methods (LeadFinder and Gnina) as well as a more basic consensus docking approach using the mean binding energy and mean rank as scoring metrics. The EF[34] for the top 5, 1, 0.5, and 0.1% was calculated using eq 4:

$$\text{EF} = \frac{\text{actives EF\%/total EF\%}}{\text{all actives/all compounds}} \quad (4)$$

The EF compares the fraction of active compounds within a specific top percentage of ranked compounds (actives EF%) to the overall distribution of actives in the data set. This essentially compares the distribution of actives at the top of the ranking compared to random chance. A high EF signifies a higher concentration of true positives (TPs) with few false positives (FPs) within that top fraction, while a low EF suggests that many FPs ranked at the top. This makes it a valuable metric to assess a method's effectiveness in early-stage hit identification.

**Output Files.** After scoring and ranking all the compounds using eq 1, ESSENCE-Dock is able to generate an interactive PyMOL[24] session (PSE) file. This PSE file contains the protein structure and the top 50 results (user-configurable) based on the ESSENCE-Dock consensus scores. Each entry in the PSE file displays the three docked structures from LeadFinder, DiffDock, and Gnina, enabling an easy visual assessment of binding pose similarity. Additionally, the file includes predictions of protein−ligand interactions as predicted by PLIP,[25] offering insights into the potential binding interactions. Furthermore, these PLIP results are also summarized in both JSON and spreadsheet-style CSV files for convenient further analysis. Finally, all of the information that ESSENCE-Dock uses (the individual docking scores, RMSD values, etc.) for all compounds is saved in an output CSV file as well, summarizing all of the key data for every docked compound.

## RESULTS AND DISCUSSION

To demonstrate the efficacy of ESSENCE-Dock, we applied our workflow to a set of 21 distinct DUD-E targets, representing a diverse range of protein families and biological contexts. After docking and ranking of the compounds, the EFs were calculated at different fractions. EFs are a measure of the method's ability to identify active compounds among the top-ranked candidates, with higher EF values indicating better performance (see Materials and Methods and eq 4 for the exact EF formula). We evaluated EFs for the individual DUD-E targets at various fractions: the top 5%, top 1%, top 0.5%, and top 0.1%, in order to analyze ESSENCE-Dock's efficacy in different scenarios. To compare and benchmark ESSENCE-Dock's performance, we also computed the EFs at the same threshold levels using the rankings from the individual docking methods: LeadFinder (LF) and Gnina (GN). We also calculated the EFs for DiffDock by ranking the compounds based on the calculated Confidence metric. However, because this was not intended to be used as a virtual screening metric, we refer to Table S1 of the supplementary data. Additionally, we calculated two basic consensus metrics using the individual GN and LF docking calculations: the mean binding energy (MBE), which provides insights into the collective binding affinities, and the mean rank (MR), which offers a measure of ranking consistency across different docking methods. All of these results are summarized in Table 1. The average results are also compared by using a bar chart in Figure 4. The same information for the median and all of the results for the individual DUD-E targets can be found in the Supplementary Data (Figures S1−S23).

As illustrated in Table 1, our ESSENCE-Dock method often outperformed the other methods, especially at the smaller EF fractions of EF0.5% and EF0.1%. In the case of EF0.1%, ESSENCE-Dock was surpassed by only another method in five cases: AKT1, CP3A4, GCR, ITAL, and MK14. Even when our approach performed worse, it often showed notable enrichment. Only in the cases of AKT1 and CP3A4 did ESSENCE-Dock substantially underperform other methods. For the other three DUD-E targets, ESSENCE-Dock's performance remained competitive with high EFs of 27.6, 48.7, and 29.8 for GCR, ITAL, and MK14, respectively. In some other cases such as AMPC, CXCR4, FABP4, and GLCM, none of the methods manage to identify actives in the top 0.5 and 0.1% of ranked compounds, resulting in EFs of 0. However, an important note is that these specific data sets are relatively small (ranging from only ∼2900 to ∼3850 compounds). This means that at these top enrichments, only a few compounds are considered, which can partly explain why all of the methods score substantially worse here.



https://doi.org/10.1021/acs.jcim.3c01982
J. Chem. Inf. Model. 2024, 64, 1605−1614

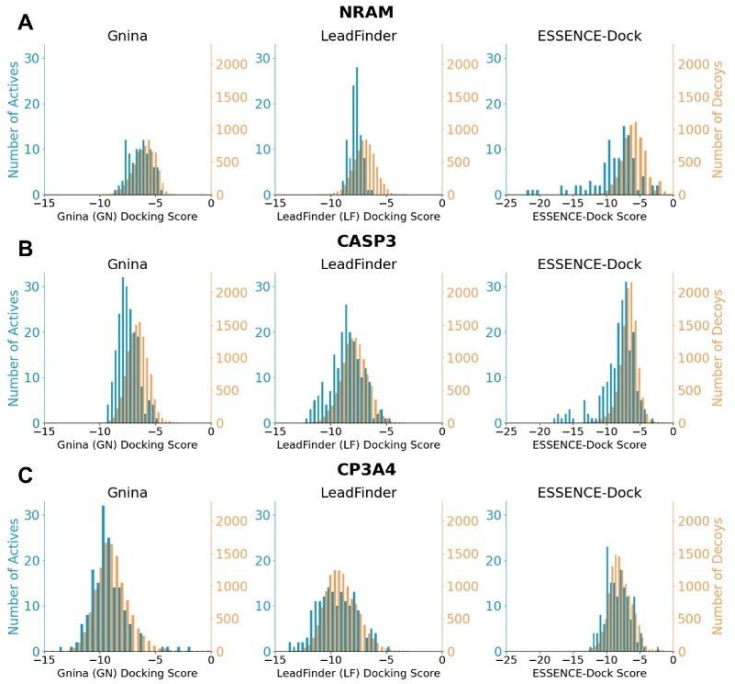



**Figure 5.** Comparative docking performance of Gnina, LeadFinder, and ESSENCE-Dock for NRAM, CASP3, and CP3A4 from the DUD-E data set. This figure presents the distributions of actives (depicted in blue) and decoys (represented in orange) based on their respective docking scores for the NRAM (A), CASP3 (B), and CP3A4 (C) data sets included in the DUD-E database. Note that the figure employs differently scaled axes to accommodate the varying quantity of actives (left axis) compared with decoys (right axis). In all cases, a lower score (more negative) indicates a better-ranked sample.

The individual EFs were averaged for each method and EF threshold, and are summarized in Figure 4. This figure clearly demonstrates that our approach achieved the highest average enrichment scores across all of the evaluated thresholds. While the improvement at EF5% is relatively modest, the improvement becomes more evident as the EF threshold decreases. This is especially the case with EF0.1%, where ESSENCE-Dock has an EF that is five times higher than that of LeadFinder and two times higher than that of Gnina. Furthermore, even in comparison to the mean rank and mean binding energy consensus methods, ESSENCE-Dock is still nearly twice as effective.

Figure 5 shows histograms comparing the ranking performance of the individual docking methods (GN and LF) and the ESSENCE-Dock approach for three distinct scenarios of varying performance: NRAM (A), CASP3 (B), and CP3A4 (C). In the NRAM scenario, ESSENCE-Dock substantially outperforms both individual docking methods. For CASP3, the individual methods already exhibit strong independent performance, and combining the results using the ESSENCE-Dock consensus approach further enhances the overall ranking. Conversely, CP3A4 demonstrates a scenario with limited consensus between methods, resulting in a lower performance by ESSENCE-Dock. These specific examples are explored in greater detail below. Other targets could have been chosen to be featured here, and similar figures for all of the tested DUD-E data sets are provided in the Supplementary Data (Figures S24–S44). Each figure within this comparison features two separate histograms: one representing the distribution of active compounds (in blue) and the other depicting the distribution of decoy compounds (in orange). Of note are varying quantities of active and decoy compounds, which lead to separate *Y*-axes for each group. This distinction is caused by the significantly larger number of decoys compared with active compounds. For all docking methods, lower (more negative) scores on the *X*-axis indicate better result.

In the example of NRAM (Figure 5A), the individual docking methods GN and LF do not excel at assigning better docking scores to the active compounds compared to the decoys. This is reflected in the overlapping distributions between the actives and decoys in these cases. In contrast, the assigned ESSENCE-Dock scores produce a distinct cluster of active compounds distributed within the range of −25 to around −15, clearly separated from the decoy compounds. These trends align with the quantitative data presented in Table 1, where our ESSENCE-Dock method at EF0.5% and EF0.1% for NRAM demonstrates substantial enrichment while the other methods yield EFs of zero.

In the case of CASP3 (Figure 5B), the individual docking methods GN and LF demonstrate better performance compared to the NRAM case, successfully ranking a higher number of actives near the top based on their respective docking scores. Still, ESSENCE-Dock can outperform the



https://doi.org/10.1021/acs.jcim.3c01982
J. Chem. Inf. Model. 2024, 64, 1605−1614

other methods using its consensus approach. Again, this is reflected in Table 1, where at EF0.5% and especially EF0.1%, ESSENCE-Dock shows a higher EF than the individual docking methods, as well as the more simplistic MR and MBE consensus results. Interestingly, again, there is a distinct cluster of active compounds present between the −20 and −15 scores.

Finally, Figure 5C shows an example of CP3A4 from the DUD-E database. This is a case where ESSENCE-Dock performs worse than the other methods, as can also be seen in the EF0.1% results for CP3A4 in Table 1. Even the individual docking methods, GN and LF provide better enrichment than ESSENCE-Dock in this specific example. While these results for ESSENCE-Dock on the CP3A4 data set seem disappointing, they are also informative. The observation that ESSENCE-Dock assigns lower scores to nearly all compounds in this data set indicates its effectiveness in detecting a lack of consensus, which, in turn, suggests that the results may be unreliable.

In this context, the ESSENCE-Dock score can also serve as a confidence metric for the consensus results. When the score is relatively bad (e.g., >−15), it typically indicates that ESSENCE-Dock detects low consensus. This suggests that there is substantial disagreement among the individual docking methods and the results should be interpreted with caution. This is corroborated by the findings for most other data sets where ESSENCE-Dock performs poorly (AMPC, CP3A4, CXCR4, FABP4, and GLCM), which typically showed low ESSENCE-Dock scores (>−15), even for the top-ranked compounds, indicating low confidence in these results. We believe that the ability of ESSENCE-Dock scores to act as a quality measure for the docking results is a valuable feature. Specifically, during virtual screening campaigns using ESSENCE-Dock, a certain score cutoff (e.g., −15 or −20) could be applied. Compounds scoring below the threshold would be rejected, allowing researchers to effectively filter and prioritize compounds, reducing false positives and focusing on candidates with a more robust consensus.

Finally, we briefly discuss the general runtimes of both the individual docking runs and the ESSENCE-Dock calculations themselves. More detailed information is provided for three DUD-E examples—MK14, TRY1, and WEE1—in the supplementary data: Figure S45 for the individual docking times and Figure S46 for the consensus time. These data sets vary in size and can thus give a good indication for the approximate runtimes across various scenarios.

In general, the Gnina calculations took the most amount of time. This is also clear from Figure S45 and can probably be attributed to the fact that these calculations were run using only CPUs, while Gnina's CNN scoring algorithm has been optimized to be run on a GPU. The Gnina docking calculations were assigned four CPUs per job, which was able to accelerate the initial generation and optimization of binding poses compared to running it with only one CPU. However, the final step of CNN scoring took up the most amount of time and seemed to be unchanged regardless of the number of assigned CPUs. Performing the Gnina calculations using GPUs would accelerate the CNN scoring and reduce the overall runtime substantially.

Additionally, DiffDock and LeadFinder took about the same amount of wall-clock time to complete in most cases. However, even though their runtime was similar, DiffDock and LeadFinder do not have the same computational cost: DiffDock jobs were assigned four CPU cores, while LeadFinder jobs ran using only one. Note that DiffDock was also designed to be run on a GPU, but when it is not, it can still take advantage of multiple CPU cores if they are available. Finally, LeadFinder was developed to be run on CPUs and, thus, was the most optimized. This translated into the lowest runtimes most of the time. A lot of these general trends can be clearly seen in Figure S45.

Specific details about the ESSENCE-Dock consensus runtimes are presented in Figure S46. During this process, the individual docking calculations for every compound are processed and integrated by computing their final ESSENCE-Dock score. The time required to generate a PyMOL Session file with PLIP interactions is not included here. The consensus calculations are performed using a Python script, which also makes use of subprocess calls to obrms and obrotamers to calculate the RMSD and determine the number of rotatable bonds. The time spent in the subprocess calls is not included in the user time. Instead, it is listed as system time, and therefore, we include all three metrics as returned from the Unix "time" command. ESSENCE-Dock consensus calculations also scale linearly, as is evident from Figure S46 as well. This is corroborated by $R^2$ value of 0.99 for system time and 1.00 for user and real times. The total runtime can get more significant if the data sets are really large, but processing small- to medium-sized libraries is still very manageable, only taking up about 15 min for a small data set (like the WEE1 example; ∼6250 compounds) and about 70 min for larger ones (like the MK14 example; ∼36,500 compounds).

From these results, it is clear that our novel ESSENCE-Dock approach is a powerful consensus docking technique that is able to provide substantial enrichment across various families of proteins. On top of that, it also provides helpful output files, including a PSE file of the top results along with their predicted PLIP interactions. The predicted protein−ligand interactions are also summarized in useful JSON and CSV files.

## ■ CONCLUSIONS AND OUTLOOK

We have presented ESSENCE-Dock, an easy-to-use consensus docking workflow that focuses on providing a high enrichment of active compounds. In its current implementation, it combines three different docking methods (DiffDock, Gnina, and LeadFinder) and considers the predicted binding pose similarities, docking scores, and ligand flexibility. Additionally, the top results can be visualized in an interactive PyMOL session with PLIP-predicted protein−ligand interactions. The predicted protein−ligand interactions are also provided in useful output formats for postprocessing. ESSENCE-Dock can be run on HPC clusters by using Singularity and Slurm, making it feasible to apply the workflow on larger virtual libraries.

Future work could explore the use of new docking methods, as ESSENCE-Dock is a flexible framework that can easily be made to work with other docking methods. When new, better-performing docking methods are developed, they can be quickly and easily integrated into the workflow. Additionally, it could be interesting to train a machine learning model with all of the input data that can be generated by ESSENCE-Dock. This would allow us to move to an ML-based scoring function, which could be more robust and outperform the current method, leading to better overall results.



https://doi.org/10.1021/acs.jcim.3c01982
J. Chem. Inf. Model. 2024, 64, 1605−1614

## ■ ASSOCIATED CONTENT

### Data Availability Statement

DiffDockHPC (software for the DiffDock calculations): https://github.com/Jnelen/DiffDockHPC. Metascreener (Software for the Gnina and LeadFinder docking calculations as well as the ESSENCE-Dock consensus workflow): https://github.com/bio-hpc/metascreener. The DUD-E docking data and ESSENCE-Dock consensus results: https://zenodo.org/doi/10.5281/zenodo.10025839.

### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.3c01982.

DiffDock enrichment factors table; Complementary figures for individual DUD-E targets; Docking and ESSENCE-Dock consensus runtimes (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Author

**Horacio Pérez-Sánchez** − *Structural Bioinformatics and High Performance Computing Research Group (BIO-HPC), HiTech Innovation Hub, UCAM Universidad Católica de Murcia, Murcia 30107, Spain;* orcid.org/0000-0003-4468-7898; Email: hperez@ucam.edu

### Authors

**Jochem Nelen** − *Structural Bioinformatics and High Performance Computing Research Group (BIO-HPC), HiTech Innovation Hub, UCAM Universidad Católica de Murcia, Murcia 30107, Spain; Health Sciences PhD Program, Universidad Católica de Murcia UCAM, Murcia 30107, Spain;* orcid.org/0000-0002-9970-4950

**Miguel Carmena-Bargueño** − *Structural Bioinformatics and High Performance Computing Research Group (BIO-HPC), HiTech Innovation Hub, UCAM Universidad Católica de Murcia, Murcia 30107, Spain; Health Sciences PhD Program, Universidad Católica de Murcia UCAM, Murcia 30107, Spain;* orcid.org/0000-0002-2735-4859

**Carlos Martínez-Cortés** − *Structural Bioinformatics and High Performance Computing Research Group (BIO-HPC), HiTech Innovation Hub, UCAM Universidad Católica de Murcia, Murcia 30107, Spain;* orcid.org/0000-0002-5029-3089

**Alejandro Rodríguez-Martínez** − *Structural Bioinformatics and High Performance Computing Research Group (BIO-HPC), HiTech Innovation Hub, UCAM Universidad Católica de Murcia, Murcia 30107, Spain; Health Sciences PhD Program, Universidad Católica de Murcia UCAM, Murcia 30107, Spain;* orcid.org/0000-0002-1251-8072

**José Manuel Villalgordo-Soto** − *Eurofins-Villapharma, Parque Tecnológico de Fuente Álamo, Murcia E-30820, Spain*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.3c01982



### Funding

J.N. was funded by Cátedra Villapharma-UCAM. M.C.-B is a predoctoral student funded by Plan Propio de Investigación, UCAM. Supercomputing resources in this work have been supported by the Plataforma Andaluza de Bioinformática of the University of Málaga (https://www.scbi.uma.es/site/), by the supercomputing infrastructure of the NLHPC (ECM-02, Powered@NLHPC), and by the Extremadura Research Centre for Advanced Technologies (CETA−CIEMAT), funded by the European Regional Development Fund (ERDF). CETA−CIEMAT belongs to CIEMAT and the Government of Spain.

### Notes

The authors declare no competing financial interest.


## ■ ABBREVIATIONS

ADT: AutoDock Tools
CNN: convolutional neural network
Cons.: consensus
CPU: central processing unit
CSV: comma-separated values (file)
DD: DiffDock
DUD(-E): Directory of Useful Decoys(-Enhanced)
ECR: exponential consensus ranking
EF: enrichment factor
GN: Gnina
GPU: graphics processing unit
HTS: high-throughput screening
HPC: high-performance computing
LF: LeadFinder
MBE: mean binding energy
ML: machine learning
MR: mean rank
PDB: Protein Data Bank
PLIP: protein−ligand interaction profiler
PSE: PyMOL session (file)
RAM: random-access memory
RB: rotatable bonds
RMSD: root mean square deviation
VS: virtual screening

## ■ REFERENCES


(1) Schlander, M.; Hernandez-Villafuerte, K.; Cheng, C. Y.; Mestre-Ferrandiz, J.; Baumann, M. How Much Does It Cost to Research and Develop a New Drug? *A Systematic Review and Assessment. PharmacoEconomics.* **2021**, *39* (11), 1243−69.

(2) Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe, E. W. Computational Methods in Drug Discovery. *Pharmacol Rev.* **2014**, *66* (1), 334−95.

(3) Gloriam, D. E. Bigger is better in virtual drug screens. *Nature.* **2019**, *566* (7743), 193−4.

(4) Pinzi, L.; Rastelli, G. Molecular Docking: Shifting Paradigms in Drug Discovery. *Int. J. Mol. Sci.* **2019**, *20* (18), 4331.

(5) Blanes-Mira, C.; Fernández-Aguado, P.; de Andrés-López, J.; Fernández-Carvajal, A.; Ferrer-Montiel, A.; Fernández-Ballester, G. Comprehensive Survey of Consensus Docking for High-Throughput Virtual Screening. *Molecules.* **2023**, *28* (1), 175.

(6) Trott, O.; Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455−61.

(7) Koes, D. R.; Baumgartner, M. P.; Camacho, C. J. Lessons Learned in Empirical Scoring with smina from the CSAR 2011 Benchmarking Exercise. *J. Chem. Inf Model.* **2013**, *53* (8), 1893−904.

(8) McNutt, A. T.; Francoeur, P.; Aggarwal, R.; Masuda, T.; Meli, R.; Ragoza, M.; et al. Gnina 1.0: molecular docking with deep learning. *J. Cheminf.* **2021**, *13* (1), 43.

(9) Stroganov, O. V.; Novikov, F. N.; Stroylov, V. S.; Kulkov, V.; Chilov, G. G. Lead finder: an approach to improve accuracy of protein-ligand docking, binding energy estimation, and virtual screening. *J. Chem. Inf Model.* **2008**, *48* (12), 2371−85.

(10) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; et al. Glide: A New Approach for Rapid,

Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47* (7), 1739−49.
(11) Torres, P. H. M.; Sodero, A. C. R.; Jofily, P.; Silva, F. P., Jr Key Topics in Molecular Docking for Drug Design. *Int. J. Mol. Sci.* **2019**, *20* (18), 4574.
(12) Tuccinardi, T.; Poli, G.; Romboli, V.; Giordano, A.; Martinelli, A. Extensive Consensus Docking Evaluation for Ligand Pose Prediction and Virtual Screening Studies. *J. Chem. Inf Model.* **2014**, *54* (10), 2980−6.
(13) Kukol, A. Consensus virtual screening approaches to predict protein ligands. *Eur. J. Med. Chem.* **2011**, *46* (9), 4661−4.
(14) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking Sets for Molecular Docking. *J. Med. Chem.* **2006**, *49* (23), 6789−801.
(15) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30* (16), 2785−91.
(16) Houston, D. R.; Walkinshaw, M. D. Consensus Docking: Improving the Reliability of Docking in a Virtual Screening Context. *J. Chem. Inf Model.* **2013**, *53* (2), 384−90.
(17) Ochoa, R.; Palacio-Rodriguez, K.; Clemente, C. M.; Adler, N. S. DockECR: Open consensus docking and ranking protocol for virtual screening of small molecules. *J. Mol. Graph Model.* **2021**, *109*, No. 108023.
(18) Morris, C. J.; Stern, J. A.; Stark, B.; Christopherson, M.; Della Corte, D. MILCDock: Machine Learning Enhanced Consensus Docking for Virtual Screening in Drug Discovery. *J. Chem. Inf. Model.* **2022**, acs.jcim.2c00705.
(19) Palacio-Rodríguez, K.; Lans, I.; Cavasotto, C. N.; Cossio, P. Exponential consensus ranking improves the outcome in docking and receptor ensemble docking. *Sci. Rep.* **2019**, *9* (1), 5142.
(20) Zhang, N.; Zhao, H. Enriching screening libraries with bioactive fragment space. *Bioorg. Med. Chem. Lett.* **2016**, *26* (15), 3594−7.
(21) Ruiz-Carmona, S.; Alvarez-Garcia, D.; Foloppe, N.; Garmendia-Doval, A. B.; Juhos, S.; Schmidtke, P.; et al. rDock: A Fast, Versatile and Open Source Program for Docking Ligands to Proteins and Nucleic Acids. *PLOS Comput. Biol.* **2014**, *10* (4), No. e1003571.
(22) Mysinger, M. M.; Carchia, M.; Irwin, J. J.; Shoichet, B. K. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *J. Med. Chem.* **2012**, *55* (14), 6582−94.
(23) Corso, G.; Stärk, H.; Jing, B.; Barzilay, R.; Jaakkola, T. DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking. *arXiv* **2022**, arXiv-2210.
(24) Schrödinger, L. L. C. *The PyMOL Molecular Graphics System*. Version 1.8. 2015.
(25) Salentin, S.; Schreiber, S.; Haupt, V. J.; Adasme, M. F.; Schroeder, M. PLIP: fully automated protein−ligand interaction profiler. *Nucleic Acids Res.* **2015**, *43* (W1), W443.
(26) *RDKit: Open-source cheminformatics*. RDKit; **2023** https://github.com/RDKit/RDKit.
(27) *Molconvert* | Chemaxon Docs. https://docs.chemaxon.com/display/docs/Molconvert.md.
(28) Halgren, T. A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **1996**, *17* (5−6), 490−519.
(29) Schrödinger Release 2020−4; *Maestro*, Schrödinger, LLC: New York, NY, 2020.
(30) Roos, K.; Wu, C.; Damm, W.; Reboul, M.; Stevenson, J. M.; Lu, C.; et al. OPLS3e: Extending Force Field Coverage for Drug-Like Small Molecules. *J. Chem. Theory Comput.* **2019**, *15* (3), 1863−74.
(31) Novikov, F. N.; Stroylov, V. S.; Zeifman, A. A.; Stroganov, O. V.; Kulkov, V.; Chilov, G. G. Lead Finder docking and virtual screening evaluation with Astex and DUD test sets. *J. Comput. Aided Mol. Des.* **2012**, *26* (6), 725−35.
(32) *bio-hpc/metascreener: Collection of scripts that integrates docking, virtual screening, similarity and molecular modeling programs*. https://github.com/bio-hpc/metascreener/tree/main.
(33) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An open chemical toolbox. *J. Cheminf.* **2011**, *3* (1), 33.
(34) Krüger, D. M.; Evers, A. Comparison of Structure- and Ligand-Based Virtual Screening Protocols Considering Hit List Complementarity and Enrichment Factors. *ChemMedChem.* **2010**, *5* (1), 148−58.



https://doi.org/10.1021/acs.jcim.3c01982
*J. Chem. Inf. Model.* 2024, 64, 1605−1614

## 3.3. TARGETING DRUG RESISTANCE IN CANCER: DIMETHOXYCURCUMIN AS A FUNCTIONAL ANTIOXIDANT TARGETING ABCC3

| | |
|---|---|
| **Title** | Targeting Drug Resistance in Cancer: Dimethoxycurcumin as a Functional Antioxidant Targeting ABCC3 |
| **Authors** | Jochem Nelen; Valeria Naponelli; José Manuel Villalgordo-Soto; Marco Falasca; Horacio Pérez-Sánchez |
| **Journal** | Antioxidants |
| **Year** | 2025 |
| **Volume** | 14 (5) |
| **Pages** | 18 |
| **DOI** | https://doi.org/10.3390/antiox14050599 |
| **IF (JCR 2024)** | 6.6 |
| **Category** | Biochemistry & Molecular Biology; 45/319; Q1<br>Chemistry, Medicinal; 8/72; Q1<br>Food Science & Technology; 22/181; Q1 |

**Contribution of PhD Candidate**

The PhD candidate, Jochem Nelen, confirms that he is the first author of the article titled "*Targeting Drug Resistance in Cancer: Dimethoxycurcumin as a Functional Antioxidant Targeting ABCC3*". He was primarily responsible for the conception of the study, development and implementation of the methodology, data analysis, manuscript writing, and overall coordination of the research work.

*antioxidants*

MDPI

*Article*

# Targeting Drug Resistance in Cancer: Dimethoxycurcumin as a Functional Antioxidant Targeting ABCC3

**Jochem Nelen [1,2], Valeria Naponelli [3], José Manuel Villalgordo-Soto [4], Marco Falasca [3,*] and Horacio Pérez-Sánchez [1,*]**

[1] Structural Bioinformatics and High Performance Computing Research Group (BIO-HPC), HiTech Innovation Hub, UCAM Universidad Católica de Murcia, 30107 Murcia, Spain; jnelen@ucam.edu
[2] Health Sciences PhD Program, Universidad Católica de Murcia UCAM, Campus de los Jerónimos n°135, Guadalupe, 30107 Murcia, Spain
[3] University of Parma, Department of Medicine and Surgery, Via Volturno 39, 43125 Parma, Italy; valeria.naponelli@unipr.it
[4] Eurofins-Villapharma Research, 30320 Murcia, Spain; josemanuel.villalgordo@discovery.eurofinseu.com
* Correspondence: marco.falasca@unipr.it (M.F.); hperez@ucam.edu (H.P.-S.)

**Abstract:** The development of new anticancer therapies remains challenging due to tumor heterogeneity and the frequent emergence of multidrug resistance (MDR). Natural products have garnered increasing attention as alternative or complementary therapeutic agents due to their bioactivity and reduced toxicity. Polyphenols, particularly curcumin and its derivatives, have shown promise in modulating signaling pathways, enhancing chemosensitivity, and overcoming drug resistance. The anticancer potential of dimethoxycurcumin, a chemically modified curcumin derivative identified through consensus fingerprint similarity screening, was investigated for its potential to inhibit ABCC3 (MRP3)—a member of the ATP-binding cassette (ABC) transporter family implicated in drug efflux, tumor cell survival, and resistance. In vitro experiments demonstrated that dimethoxycurcumin significantly reduced cancer cell viability and colony formation, indicating a strong inhibitory effect on ABCC3 function. These results suggest that dimethoxycurcumin may sensitize cancer cells to chemotherapy by targeting resistance pathways. The data presented contribute to the growing body of evidence suggesting that bioactive plant-derived compounds, including chemically modified derivatives, may hold therapeutic potential in oncology by modulating multidrug resistance pathways. Targeting ABC transporters with natural compound derivatives could offer a promising strategy for developing more effective and less toxic anticancer therapies.

**Keywords:** cancer drug resistance; virtual screening; curcumin derivatives; molecular fingerprints; pancreatic cancer; ABC transporters; antioxidants; natural products



Academic Editors: María Garrido, Javier Espino and Jonathan Delgado-Adamez

Received: 28 March 2025
Revised: 12 May 2025
Accepted: 12 May 2025
Published: 16 May 2025

**Citation:** Nelen, J.; Naponelli, V.; Villalgordo-Soto, J.M.; Falasca, M.; Pérez-Sánchez, H. Targeting Drug Resistance in Cancer: Dimethoxycurcumin as a Functional Antioxidant Targeting ABCC3. *Antioxidants* **2025**, *14*, 599. https://doi.org/10.3390/antiox14050599



## 1. Introduction

Advancing anticancer therapies is a significant challenge due to the intricate nature of tumor biology and the persistent issue of drug resistance [1]. While conventional small-molecule drugs have been highly effective in cancer treatment, their long-term use can be limited by toxicity, off-target effects, and the emergence of multidrug resistance (MDR). MDR arises through various mechanisms, including enhanced drug efflux, metabolic adaptation, and alterations in drug targets, ultimately reducing therapeutic efficacy [2]. This issue is especially critical in pancreatic ductal adenocarcinoma (PDAC), an aggressive malignancy with one of the highest cancer-related mortality rates. Its poor prognosis stems

from multiple factors, including the rapid emergence of drug resistance, which significantly limits treatment effectiveness [3].

ABC transporters play a crucial role in cancer resistance by actively exporting drugs from cells, reducing intracellular concentrations, and promoting chemoresistance [4]. These membrane transporters are involved in various mechanisms of drug resistance, affecting the efficacy of many chemotherapeutic agents, including cisplatin and taxanes [5]. One notable member of this family is ATP-binding cassette subfamily C member 3 (ABCC3), also known as multidrug resistance-associated protein 3 (MRP3). ABCC3 plays a key role in drug disposition and has been implicated in cancer cell survival and proliferation, particularly in pancreatic cancer [6]. In addition to its role in drug resistance, it has been found that ABC transporters actively contribute to cancer progression by extruding bioactive molecules that can influence the tumor microenvironment [7–10]. Their role in modulating drug response and promoting cancer progression makes them a relevant target for potential therapeutic interventions. Strategies to counteract ABC transporter-mediated resistance include the development of small-molecule inhibitors, combination therapies to enhance drug retention, and natural compound-based approaches that modulate transporter activity [11]. Targeting these mechanisms could improve chemotherapy efficacy and help overcome treatment resistance in cancers such as PDAC, where drug efflux plays a significant role in therapeutic failure. Notably, ABCC3 inhibitors have the potential to not only enhance drug retention but also impede tumor progression by blocking the extrusion of bioactive molecules that contribute to the tumor microenvironment and cancer cell survival [10].

The challenges of drug resistance in cancer have led to growing interest in alternative strategies, such as natural product-based compounds, which may complement existing treatments by targeting resistance pathways while offering potentially improved safety profiles [12]. Plant-derived bioactive compounds, particularly polyphenols such as flavonoids [13], hold promise in oncology due to their antioxidative effects, regulation of signaling pathways, and potential to enhance chemosensitivity [14]. Antioxidants play a crucial role in mitigating oxidative stress, a factor implicated in cancer progression and resistance to therapy [15]. Their ability to modulate the tumor microenvironment, influence immune responses, and synergize with conventional treatments could make them valuable in cancer management. Several plant-derived compounds have already demonstrated clinical success, underscoring the importance of further exploration in this area. For instance, paclitaxel, a taxane originally extracted from *Taxus brevifolia*, is widely used to treat various malignancies, including breast, ovarian, and lung cancers [16]. Similarly, vinblastine and vincristine, alkaloids derived from *Catharanthus roseus*, are essential components of chemotherapy regimens for hematologic malignancies and solid tumors [17]. Additionally, camptothecin, isolated from *Camptotheca acuminata*, has led to the development of topoisomerase inhibitors such as irinotecan and topotecan, which are used in colorectal, ovarian, and small-cell lung cancers [18]. These examples highlight the immense potential of plant-derived compounds in oncology and reinforce the need to investigate novel bioactive molecules with improved efficacy and safety profiles. Beyond these already well-established plant-derived chemotherapeutics, other natural compounds continue to attract attention for their potential anticancer properties.

Taken together, these findings highlight the urgent need for novel therapeutic strategies capable of overcoming multidrug resistance in PDAC by targeting key molecular mechanisms such as ABC transporter activity. In particular, ABCC3 (MRP3) represents a compelling target due to its established role in drug efflux and tumor progression. To explore potential inhibitors, a ligand-based virtual screening campaign was carried out using a structurally diverse compound library, including both synthetic molecules and natural product derivatives. This approach aimed to identify small-molecule modulators

*Antioxidants* **2025**, *14*, 599 

of ABCC3 that may enhance chemosensitivity and contribute to improved therapeutic outcomes in pancreatic cancer.

## 2. Materials and Methods

### *2.1. Virtual Screening Using Consensus Fingerprint Similarity*

The diverse and proprietary in-house Eurofins-VillaPharma library was processed to generate three types of molecular fingerprints: Avalon [19], Circular (Extended Connectivity Fingerprint with a diameter of 6, ECFP6) [20], and PubChem Substructure fingerprints [21], using the PyFingerprint package (version 3.0) [22]. This package leverages the RDKit [23] for Avalon fingerprint generation, and the Chemistry Development Kit (CDK) [24] for PubChem and Circular (ECFP6) fingerprints. As part of the fingerprint conversion process, the input molecules were sanitized and pre-processed, including steps such as charge normalization, salt removal, and tautomer standardization, to ensure consistent and reliable representations. The resulting fingerprints were encoded as binary vectors, which enabled efficient and rapid similarity calculations for downstream analyses.

Initially, a custom script was developed to compare the query structure against the library. This script converted the input query on-the-fly into Avalon, PubChem, and Circular (ECFP6) fingerprints and subsequently compared them to the corresponding fingerprints in the Eurofins-VillaPharma library using the Euclidean distance metric. To ensure comparability across fingerprints, the distances were normalized using z-score normalization, where each value was scaled based on the mean and standard deviation of its distribution, resulting in a standardized score for each fingerprint type.

The final consensus score for each compound was obtained by averaging these normalized scores across all three fingerprint types. Compounds were then ranked based on their consensus scores, with higher-ranking compounds selected for further evaluation. This consensus approach attempts to reduce fingerprint-specific biases and provide a more robust prediction of potential hits, while also increasing the chemical diversity among the top-ranked compounds.

To facilitate reproducibility and improve usability, this initial screening script was further developed into ConFiLiS v1.0 (Consensus Fingerprints for Ligand-based Screening), an open-source tool that refines the original methodology with improved preprocessing and performance, while maintaining the core principles of the original approach. The tool, along with detailed documentation and usage instructions, is publicly available on GitHub: https://github.com/Jnelen/ConFiLiS (accessed on 27 March 2025).

### *2.2. Hit Comparison Against BindingDB*

To assess the structural novelty of the identified hits, all known ABCC3-active compounds were retrieved from BindingDB [25] and converted into Morgan fingerprints (radius 2, 2048 bits) using RDKit. The Morgan fingerprinting method, which is based on extended-connectivity circular fingerprints (ECFP), was chosen for its ability to effectively capture molecular substructures and functional groups that contribute to biological activity, making it particularly suitable for structural similarity analysis in drug discovery.

To quantify structural similarity, Tanimoto similarity coefficients were computed between each hit and all known ABCC3-active compounds present in the BindingDB dataset. The Tanimoto coefficient, a widely used metric in cheminformatics, measures the degree of overlap between two molecular fingerprint representations, yielding a numerical value between 0 (no similarity) and 1 (identical structures). This analysis enabled the identification of the closest known ABCC3-associated compound in BindingDB for each hit based on structural similarity. This step was critical in determining whether the identified hits represented novel chemical scaffolds or bore a close resemblance to previously reported in-

hibitors. Following the computational similarity analysis, a manual review was performed to evaluate the similarity of each hit to the closest known ABCC3-associated compounds. This curation step was essential for prioritizing candidates with novel chemical scaffolds, thereby minimizing the risk of rediscovering previously known inhibitors while maximizing the potential for identifying structurally distinct ABCC3-targeting compounds.

In addition to the structural similarity assessment, in silico ADME property predictions were carried out to evaluate the drug-likeness and pharmacokinetic profiles of the selected compounds. SMILES representations of the hits were converted to SDF format using RDKit and imported into Schrödinger's Maestro (version 2024-3) [26]. Explicit hydrogen atoms were added, and 3D structures were generated using the Maestro interface. The QikProp module [27] was subsequently used to calculate key physicochemical and pharmacokinetic parameters, including predicted logP, aqueous solubility, and human oral absorption.

Finally, molecular docking was performed to investigate the potential binding modes of the prioritized hit compounds. Specifically, blind docking was conducted using AutoDock Vina [28] through MetaScreener [29], targeting all α-carbon atoms across the protein structure to identify potential binding sites without prior knowledge of specific active regions [30]. The resulting docking poses were clustered based on spatial proximity, and the most populated clusters were examined to identify key binding hotspots. From these, the top-scoring pose from the cluster within the main channel was selected for protein–ligand interaction analysis. To characterize the predicted binding interactions, the selected protein–ligand complex was processed using Protein–Ligand Interaction Profiler (PLIP) [31]. Molecular docking poses and PLIP interaction diagrams were rendered using PyMOL (version 2.6.2) [32], to produce high-resolution figures for visualization.

### *2.3. Viability Assays*

HPAF-II, BxPC-3, and CFPAC-1 cells were seeded in 96-well plates at a density of 5000 cells per well and allowed to adhere for 24 h under standard culture conditions (37 °C, 5% $CO_2$). Cells were then treated with the indicated compounds at varying concentrations for an additional 72 h to assess dose-dependent effects on viability.

Following the treatment period, cell viability was evaluated by measuring the metabolic activity of live cells. The cells were first fixed with 3% paraformaldehyde for 10 min at room temperature to preserve morphology and prevent detachment. They were then stained with 0.5% crystal violet for 15 min to label adherent cells. Excess dye was removed by gently washing the wells with deionized water, and the plates were left to air-dry completely.

To quantify cell viability, the bound crystal violet dye was solubilized using 0.1 M sodium citrate, and absorbance was measured at 590 nm using a microplate reader (EnSpire, PerkinElmer, Waltham, MA, USA). The optical density (OD) values obtained reflected the relative number of viable, adherent cells, providing a quantitative measure of treatment efficacy. All experiments were performed in triplicate to ensure reproducibility, and background absorbance was subtracted using wells containing only the staining solution. Graphics and $IC_{50}$ calculations were carried out using GraphPad Prism v10.0.

### *2.4. Clonogenic Assays*

BxPC-3 cells were seeded in 6-well plates at a density of 500 cells per well and incubated in a complete medium at 37 °C in a humidified atmosphere containing 5% $CO_2$ for 12 days to allow colony formation. The medium was refreshed every three days to maintain optimal growth conditions. For treatment experiments, the medium was supplemented with the indicated concentrations of DMSO or test compounds, ensuring continuous drug exposure throughout the incubation period. After 12 days, colony formation was assessed

by first fixing the cells with 4% paraformaldehyde for 20 min at room temperature to preserve the cellular structure and prevent detachment. The colonies were then stained with 0.5% crystal violet solution for 30 min to enhance visualization. Excess dye was removed by gently rinsing the wells multiple times with deionized water, and the plates were air-dried.

Colony formation was visualized using a bright-field microscope, and images were captured for documentation. Colonies were manually counted to quantify clonogenic potential. Alternatively, to achieve higher sensitivity and automated quantification, fixed colonies were stained with high-content screening (HCS) CellMask™ Deep Red (cat. n. H32721, Thermo Fisher Scientific, Eugene, OR, USA) and 4′,6-diamidino-2-phenylindole (DAPI, cat. n. D1306, Thermo Fisher Scientific, Eugene, OR, USA). Fluorescent images were acquired and analyzed using the IN-Cell Analyzer 2200 (GE Healthcare Life Sciences, Marlborough, MA, USA) to provide a more precise and reproducible assessment of colony formation.

All experiments were conducted in triplicate to ensure statistical reliability, and appropriate controls were included to account for background staining and autofluorescence.

### *2.5. Statistical Analysis*

Results are presented as mean ± standard error of the mean (SEM). One-way ANOVA was performed to evaluate the significance of differences between treatment groups and the control group, followed by Dunnett's post hoc test. Statistical analyses were conducted using GraphPad Prism version 10.0, and a $p$-value < 0.05 was considered statistically significant.

## 3. Results

### *3.1. Consensus Fingerprint-Based Virtual Screening for ABCC3 Inhibitor Discovery*

Potential ABCC3 inhibitors were identified using a consensus fingerprint-based virtual screening approach using known ABCC3-active compounds from BindingDB as a reference. Compound BDBM50302828, with a reported $IC_{50}$ of 1.1 μM, served as the starting query structure. At the time of selection, it was the most potent small-molecule drug for ABCC3 listed in BindingDB, making it the most suitable candidate.

To maximize chemical diversity and ensure broad structural coverage, three types of molecular fingerprints—Avalon, Circular (ECFP6), and PubChem substructure fingerprints—were generated for both the query compound and the screening library. These fingerprints were subsequently used to calculate Euclidean distances between the reference compound and each library entry.

A consensus score was derived by averaging the normalized distances across all three fingerprint types, ensuring that hits were not disproportionately influenced by any single fingerprinting method. This approach prioritizes compounds that exhibit a consistent degree of similarity across multiple fingerprinting techniques, thus improving the likelihood of identifying true ABCC3 inhibitors rather than false positives arising from fingerprint-specific artifacts.

For the final selection, a consensus score cutoff of −2.5 was applied, meaning that only compounds with the highest overall structural similarity to the reference compound across all fingerprint types were considered for further evaluation. Applying this threshold resulted in five top-ranked compounds, which were selected for experimental validation. This threshold was chosen to balance hit diversity with structural relevance to focus on candidates with high potential for ABCC3 inhibition while maintaining chemical diversity. Selected hits were subsequently subjected to further computational and experimental validation to assess their potential as novel ABCC3 inhibitors.

Table 1 summarizes the Euclidean distances for each fingerprint type, along with their normalized counterparts (Dist and Norm, respectively). The consensus score was computed as the average of these normalized distances and served as the basis for the final ranking of compounds. Only compounds with a consensus score below −2.5 were considered for further testing.

**Table 1.** Top-ranked compounds from the Eurofins-VillaPharma library identified using ConFiLiS, based on consensus fingerprint similarity. The table includes Euclidean distances (Dist) and their corresponding z-score normalized values (Norm) for each fingerprint type: Avalon, Circular (ECFP6), and PubChem. The final consensus score was computed as the average of the three normalized distances. Only compounds with a consensus score below −2.5 were selected for further biological evaluation and are shown here.

| CompoundID | Avalon Dist | Avalon Norm | Circular Dist | Circular Norm | PubChem Dist | PubChem Norm | Consensus Avg |
|---|---|---|---|---|---|---|---|
| **1** | 9.747 | −3.345 | 9.899 | −2.963 | 9.747 | −2.733 | −3.013 |
| **2** | 10.149 | −3.005 | 9.950 | −2.854 | 10.488 | −2.002 | −2.620 |
| **3** | 12.369 | −1.128 | 9.327 | −4.196 | 10.050 | −2.434 | −2.586 |
| **4** | 12.961 | −0.628 | 9.592 | −3.626 | 9.055 | −3.415 | −2.557 |
| **5** | 12.806 | −0.759 | 9.849 | −3.072 | 8.775 | −3.692 | −2.508 |

The compound structures, along with the BDBM50302828 reference compound, are shown in Figure 1. It is noteworthy that these five compounds represent three distinct scaffolds. Compounds **1** and **2** share structural features, including a sulfonamide moiety, which is also present in the reference compound. Additionally, Compounds **4** and **5** feature a 5,7-dimethoxy-2,2-dimethylchromane-based scaffold linked to a benzamide moiety.

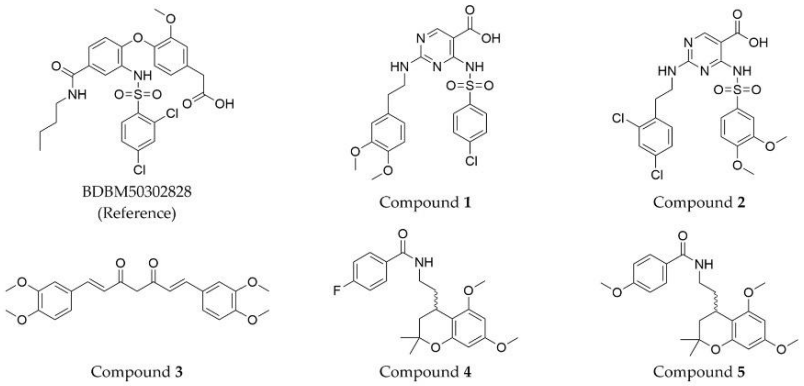



**Figure 1.** Chemical structures of predicted ABCC3 inhibitors. The reference compound, BDBM50302828, is presented alongside the five top-ranking compounds selected for testing. Compound **3** was identified as dimethoxycurcumin, a methylated natural derivative of curcumin, while Compounds **1**, **2**, **4**, and **5** are structurally diverse synthetic compounds without natural product origins. Compounds **4** and **5** were synthesized as racemic mixtures, as indicated by the wavy bond notation.

This structural diversity suggests that different scaffolds capture distinct features from the query structure, potentially interacting with ABCC3 through varied binding modes. Notably, Compound **3**, identified as dimethoxycurcumin, exhibits a high degree of structural similarity to curcumin, differing only by the methylation of both hydroxyl groups into methoxy groups.

In summary, the screening approach yielded a set of structurally diverse hits, underscoring the potential of consensus fingerprint similarity techniques in virtual screening workflows. The observed chemical diversity among the top-ranked hits highlights the utility of integrating multiple fingerprint types to mitigate model-specific biases and expand the structural search space.

### *3.2. BindingDB Similarity Analysis*

To gain a deeper understanding of the structural characteristics of the identified hits, a comparison was made against known ABCC3-active compounds from BindingDB. This analysis aimed to uncover potential structural relationships between the predicted hits and previously reported ABCC3 inhibitors. In total, 1849 ABCC3-related entries were retrieved and used as a reference set. Each hit was compared to this reference set by calculating Tanimoto similarity scores using Morgan (ECFP4) fingerprints, and the closest known compound was identified for each hit.

The results of this similarity analysis are compiled into Table 2, which lists each predicted hit, its closest known ABCC3 compound, and the corresponding Tanimoto similarity score calculated using the Morgan (ECFP4) fingerprint. A subsequent manual assessment was performed to evaluate the distinctiveness of each hit, ensuring that structurally novel candidates were prioritized for further testing while minimizing the likelihood of rediscovering known inhibitors.

**Table 2.** Structural similarity to known ABCC3 inhibitors from BindingDB. Each row shows the predicted compound, its most structurally similar known compound, and the corresponding Tanimoto similarity score using the Morgan (ECFP4) fingerprint. The synthesis of compounds **4** and **5** yielded racemic mixtures, as indicated by the wavy bond, used to represent stereochemical ambiguity.

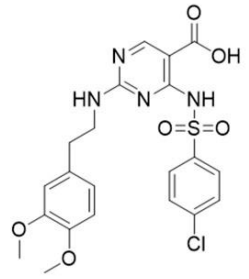
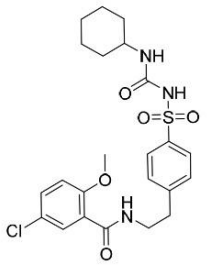
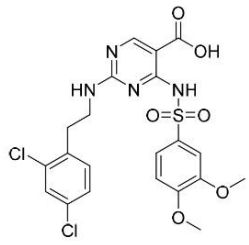
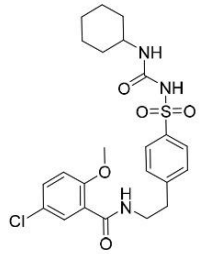
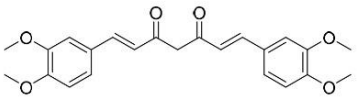
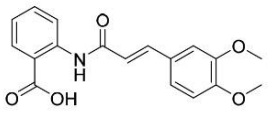
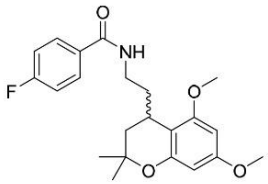
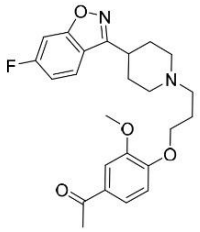
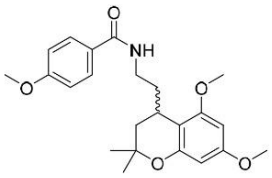
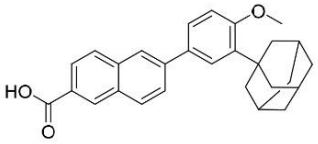
| Predicted Compound | Closest BindingDB Compound | ECFP4 Tanimoto Similarity |
|---|---|---|
|  Compound **1** |  BDBM50012957 | 0.3810 |
|  Compound **2** |  BDBM50012957 | 0.3222 |

**Table 2.** *Cont.*

| Predicted Compound | Closest BindingDB Compound | ECFP4 Tanimoto Similarity |
|---|---|---|
|  Compound **3** |  BDBM21613 | 0.5556 |
|  Compound **4** |  BDBM50034043 | 0.2637 |
|  Compound **5** |  BDBM50048280 | 0.2703 |

None of the predicted hits exactly matched any known ABCC3 inhibitors deposited in BindingDB. Compounds **1** and **2** were both structurally closest to the same compound, BDBM50012957, with Tanimoto similarity scores of 0.3810 and 0.3222, respectively. This structural similarity is likely due to the presence of a common sulfonamide moiety. Compound **3**, also known as dimethoxycurcumin, exhibited the highest similarity to any of the known ABCC3-active compounds, with a Tanimoto similarity of 0.5556. Although this similarity is relatively high, the structure was deemed sufficiently distinct to warrant testing, given the presence of amide and benzoic acid moieties in the BindingDB compound, which are not found in the predicted hit compound. Furthermore, the natural product-like characteristics of this compound contributed to its selection for further investigation. Finally, Compounds **4** and **5** were found to be the most distinct from any of the known compounds, with Tanimoto similarities of 0.2637 and 0.2703, respectively. Interestingly, despite their high structural similarity—differing only by the presence of either a fluoro or methoxy group on the para position of the terminal benzene ring—their closest known compounds were different.

To further evaluate the drug-likeness and pharmacokinetic profiles of the selected hits, in silico ADME predictions were conducted using Schrödinger's QikProp module [27]. Key physicochemical and pharmacokinetic parameters, such as predicted logP, aqueous solubility, and human oral absorption, are presented in Table 3. These results provide insights into the predicted drug-likeness and pharmacokinetic behavior of the selected compounds.

*Antioxidants* **2025**, *14*, 599



**Table 3.** Predicted physicochemical and pharmacokinetic properties of selected ABCC3 inhibitor candidates using Schrödinger's QikProp. For Compounds **4** and **5**, properties for both the R and S configurations were predicted. Properties include molecular weight (mol_MW), dipole moment, solvent-accessible surface areas (SASA for total surface area, FOSA for hydrophobic component, and FISA for hydrophilic component), polar surface area (PSA), number of rotatable bonds (#rotor), hydrogen bond donors and acceptors (donorHB, accptHB), lipophilicity (QPlogPo/w), aqueous solubility (QPlogS), Caco-2 cell permeability (QPPCaco), brain–blood barrier partitioning (QPlogBB), predicted human oral absorption (%Human Oral Absorption), and the number of violations of Lipinski's Rule of Five and Jorgensen's Rule of Three.

| Properties | Compound 1 | Compound 2 | Compound 3 | Compound 4 (R) | Compound 4 (S) | Compound 5 (R) | Compound 5 (S) |
|---|---|---|---|---|---|---|---|
| mol_MW | 492.933 | 527.378 | 396.439 | 387.45 | 387.45 | 399.486 | 399.486 |
| dipole | 3.204 | 7.76 | 1.557 | 6.978 | 5.167 | 8.733 | 4.167 |
| SASA | 710.817 | 675.128 | 759.736 | 707.6 | 709.866 | 734.525 | 724.248 |
| FOSA | 235.466 | 169.765 | 438.151 | 390.289 | 393.435 | 483.11 | 479.647 |
| FISA | 210.469 | 218.471 | 82.343 | 46.848 | 46.83 | 46.498 | 38.271 |
| PSA | 145.07 | 139.776 | 81.826 | 58.054 | 58.196 | 66.341 | 66.298 |
| #rotor [1] | 10 | 10 | 12 | 6 | 6 | 7 | 7 |
| donorHB | 2 | 2 | 0 | 1 | 1 | 1 | 1 |
| accptHB | 9.5 | 9.5 | 7 | 4.75 | 4.75 | 5.5 | 5.5 |
| QPlogPo/w | 3.164 | 3.262 | 4.215 | 5.281 | 5.285 | 5.128 | 5.157 |
| QPlogS | −4.487 | −4.104 | −4.995 | −6.636 | −6.677 | −6.464 | −6.28 |
| QPlogBB | −1.928 | −1.698 | −1.094 | −0.177 | −0.179 | −0.356 | −0.259 |
| QPPCaco | 25 | 21 | 1640 | 3561 | 3562 | 3588 | 4295 |
| %Human Oral Absorption | 71 | 57 | 100 | 100 | 100 | 100 | 100 |
| RuleOfFive Violations | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| RuleOfThree Violations | 0 | 1 | 0 | 1 | 1 | 1 | 1 |

[1] #: number of rotatable bonds (rotors).

To further characterize the selected hit compounds beyond structural similarity and ADME properties, molecular docking was performed to predict their potential binding modes within the ABCC3 transporter. Blind docking using AutoDock Vina enabled an unbiased search for potential binding sites across the protein surface. Among the docking results, the highest-scoring pose located within the main channel was selected as the most probable binding mode. To gain structural insight, interaction profiles were generated using the Protein–Ligand Interaction Profiler (PLIP), which can identify and visualize hydrogen bonds, hydrophobic contacts, and other relevant non-covalent interactions. The resulting binding poses and interaction maps are presented in Figure 2, providing a visual summary of the predicted docking poses and interactions for each compound. Table 4 provides a summary of the interacting residues along with the corresponding Vina scores for all docking poses.

*Antioxidants* **2025**, *14*, 599



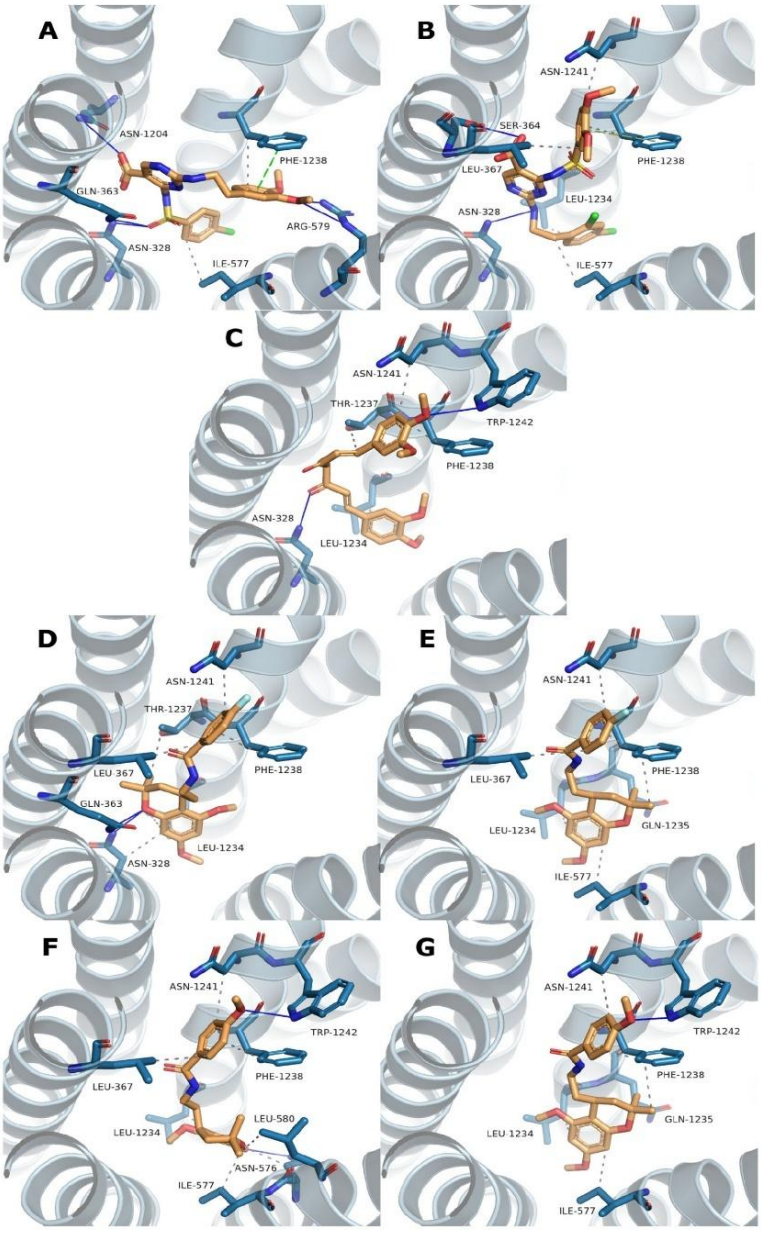



**Figure 2.** Predicted binding poses and interaction profiles of selected ABCC3 hit compounds. AutoDock Vina blind docking identified top-ranked poses in the ABCC3 transporter's main channel. Key non-covalent interactions—hydrogen bonds (solid lines), hydrophobic contacts (gray dashed), and π–π stacking (green dashed)—were identified using PLIP. Panels (**A**–**C**) show Compounds **1**–**3**; (**D**,**E**) and (**F**,**G**) depict the R and S enantiomers of Compounds **4** and **5**, respectively.

*Antioxidants* **2025**, *14*, 599



**Table 4.** Predicted protein–ligand interactions of selected ABCC3 hit compounds based on top blind docking poses. Docking scores (in kcal/mol), shown at the top of the table, reflect predicted binding affinities, with more negative values indicating stronger interactions. Each row summarizes interactions between a compound and specific ABCC3 residues, using the top-ranked pose from AutoDock Vina blind docking in the main translocation channel. Residue contacts were identified using PLIP, which detects key non-covalent interactions. A "Y" indicates at least one interaction with a residue; "N" indicates none. Stereochemical configurations (R/S) are specified for Compounds **4** and **5**.

| | **Compound 1** | **Compound 2** | **Compound 3** | **Compound 4 (R)** | **Compound 4 (S)** | **Compound 5 (R)** | **Compound 5 (S)** |
|---|---|---|---|---|---|---|---|
| Vina Score | −9.54 | −9.44 | −8.07 | −8.32 | −8.29 | −8.00 | −8.21 |
| ASN328 | Y | Y | Y | Y | N | N | N |
| GLN363 | Y | N | N | Y | N | N | N |
| SER364 | N | Y | N | N | N | N | N |
| LEU367 | N | Y | N | Y | Y | Y | N |
| ASN576 | N | N | N | N | N | Y | N |
| ILE577 | Y | Y | N | N | Y | Y | Y |
| ARG579 | Y | N | N | N | N | N | N |
| LEU580 | N | N | N | N | N | Y | N |
| ASN1204 | Y | N | N | N | N | N | N |
| LEU1234 | N | Y | Y | Y | Y | Y | Y |
| GLN1235 | N | N | N | N | Y | N | Y |
| THR1237 | N | N | Y | Y | N | N | N |
| PHE1238 | Y | Y | Y | Y | Y | Y | Y |
| ASN1241 | N | Y | Y | Y | Y | Y | Y |
| TRP1242 | N | N | Y | N | N | Y | Y |

### *3.3. Determine Activity Using Cellular Assays*

The biological activity of the five selected compounds was evaluated using a panel of human pancreatic cancer cell lines, including HPAF-II, BxPC-3, and CFPAC-1. To assess their impact on cell growth, an initial screening was performed by treating the cells with a fixed concentration of 10 μM for HPAF-II and BxPC-3, as shown in Figure 3. Among the tested compounds, Compound **3**, a methylated derivative of curcumin, exhibited the highest activity.

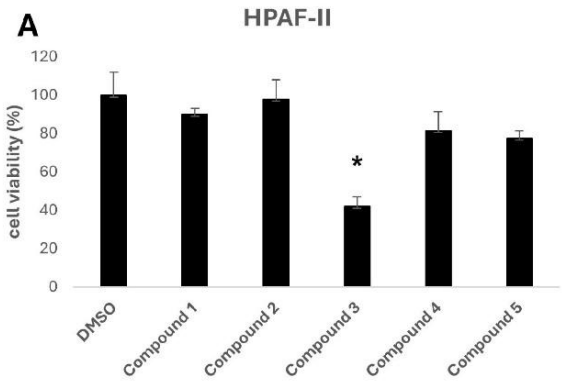



**Figure 3.** *Cont.*

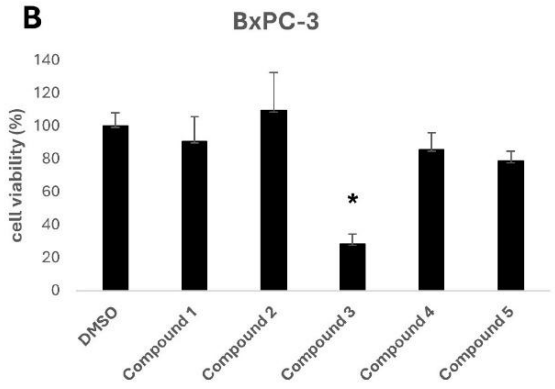



**Figure 3.** Cell viability of HPAF-II (**A**) and BxPC-3 (**B**) cells after 72 h treatment with 10 μM of Compounds **1**–**5** or DMSO (vehicle control). Viability was assessed by crystal violet staining and is expressed as a percentage. Data are presented as mean ± SEM from three independent experiments. Statistical analysis was performed using repeated measures of one-way ANOVA, followed by Dunnett's post hoc test to compare treatments to the control. * $p < 0.05$ was considered statistically significant.

To further investigate the potential anti-tumorigenic properties of the selected compounds, a two-dimensional colony formation assay was conducted using the BxPC-3 cell line (Figure 4). At a concentration of 10 μM, Compound **3** again stood out as the most effective candidate, completely inhibiting colony formation, indicating its potential for disrupting cancer cell proliferation.

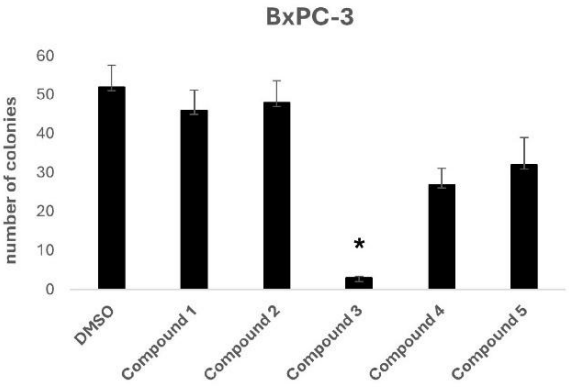



**Figure 4.** Effects of tested compounds (10 μM) on BxPC-3 cell colony formation. Cells were treated with 10 μM of each compound (Compounds **1**–**5**) or DMSO (vehicle control) and incubated for 12 days. Data are presented as mean ± SEM of three independent experiments. Statistical analysis was performed using repeated measures of one-way ANOVA, followed by Dunnett's post hoc test to compare treatments to the control. * $p < 0.05$ was considered statistically significant.

To further characterize its potency, Compound **3** was tested across a range of concentrations (1 nM to 50 μM) to determine the $IC_{50}$ value using a crystal violet viability assay for three different cancer cell lines: HPAF-II, BxPC-3, and CFPAC-1. Compound **3** demonstrated substantial growth inhibition, yielding $IC_{50}$ values of 11.03 μM, 12.90 μM, and 2.91 μM for HPAF-II, BxPC-3, and CFPAC-1, respectively (Figure 5).

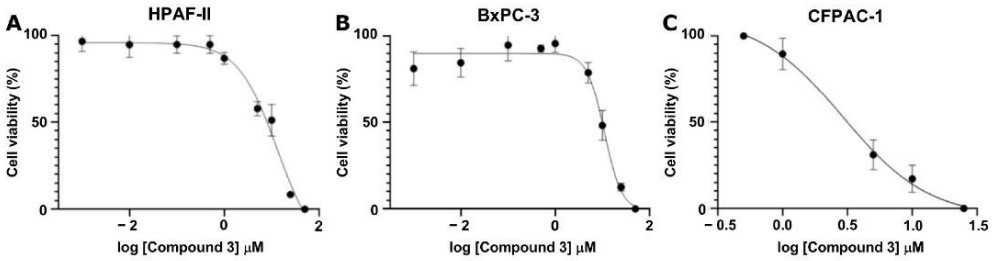



**Figure 5.** Dose-Response curve of Compound **3** for $IC_{50}$ determination in (**A**) HPAF-II (0.001–50 μM), (**B**) BxPC-3 (0.001–50 μM), and (**C**) CFPAC-1 (0.1–25 μM) cell lines. Data are presented as mean ± SEM of the three independent experiments.

## 4. Discussion

Multidrug resistance (MDR) remains a central obstacle in effective chemotherapy, particularly in aggressive solid tumors such as pancreatic ductal adenocarcinoma (PDAC). Central to this resistance mechanism are ATP-binding cassette (ABC) transporters, with ABCC3 (MRP3) emerging as a critical mediator of chemotherapeutic drug efflux and tumor progression. Identifying novel and potent modulators of ABCC3 is, therefore, crucial in advancing therapeutic strategies to overcome resistance and improve patient outcomes.

ConFiLiS, a virtual screening tool employing consensus fingerprint similarity scoring, was used to identify structurally diverse inhibitors of ABCC3 from the proprietary Eurofins-VillaPharma compound library. By integrating multiple molecular fingerprint types—Avalon, ECFP6, and PubChem substructure—this method mitigated fingerprint-specific biases and facilitated the discovery of structurally varied candidates with potential pharmacological activity. The resulting compounds fall into three distinct chemotypes: sulfonamide-containing molecules (Compounds **1** and **2**), dimethoxy-2,2-dimethylchromane-based scaffolds linked to a benzamide moiety (Compounds **4** and **5**), and a methylated curcumin derivative, dimethoxycurcumin (Compound **3**).

Structural similarity analysis against known ABCC3 ligands in BindingDB reinforced the novelty of the predicted compounds. Dimethoxycurcumin demonstrated a relatively high structural similarity to a known ABCC3-active compound (Tanimoto coefficient 0.5556), yet retained distinct pharmacophoric differences, particularly the absence of amide and benzoic acid functionalities, which supported its further investigation. The observed structural differences, despite sharing common molecular features with known inhibitors, underscore the chemical novelty achievable through unbiased ligand-based screening, emphasizing subtle yet potentially impactful modifications that could affect biological activity or potency.

To further guide early evaluation, key physicochemical and pharmacokinetic properties were predicted using QikProp (Table 3). All selected candidates exhibited general favorable drug-like characteristics, including high predicted human oral absorption, acceptable solubility profiles, and minimal violations of Lipinski's and Jorgensen's rules. These in silico predictions served as a screening filter to identify any major outliers that might warrant caution before biological testing. However, all compounds fell within acceptable ranges and were, therefore, included in downstream assays.

Among the tested ligands, Compound **3** (dimethoxycurcumin) displayed the most favorable overall profile, combining high solubility, strong permeability, and zero rule-of-five and rule-of-three violations. In contrast, compounds **1** and **2** showed notably low Caco-2 cell permeability and reduced predicted oral absorption (71% and 57%, respectively), which may indicate potential absorption limitations in vivo. Compound **5** and especially Com-

pound **4** exhibited relatively low aqueous solubility, with values predicted to fall below the typical range measured for most drugs, which could suggest potential formulation challenges. However, predicted permeability and oral absorption were favorable, warranting inclusion in the experimental studies.

To further evaluate the binding mode of the selected compounds, molecular docking was performed using a blind docking approach with AutoDock Vina. Although in some cases the highest-scoring pose corresponded to an allosteric site, we deliberately selected poses within the central translocation channel of the ABCC3 transporter due to their greater biological relevance. Importantly, the difference in docking scores between the allosteric and channel-bound poses was minimal (<0.3 kcal/mol), supporting the validity of this selection. Within the channel, all compounds were placed in a similar binding region despite differences in chemical structure (Figure 2), suggesting a common site of interaction. Still, the predicted binding poses themselves were quite diverse in orientation and contact patterns, indicating flexibility in how different ligands may engage the site. Interestingly, the R and S configurations of Compounds **4** and **5** were predicted to adopt binding poses with a highly similar orientation of the benzamide moiety, while the 5,7-dimethoxy-2,2-dimethylchromane scaffold at the stereocenter displayed nearly flipped positions. Interaction profiling using PLIP (Table 4) identified 15 predicted binding residues across all compounds. PHE1238 formed interactions with every pose, while LEU1234 and ASN1241 were involved in all except that of Compound **1**. The PLIP interaction profiles closely mirrored the differences and similarities observed in the docking poses, complementing the visual representations in Figure 2.

In vitro evaluation confirmed the biological relevance of dimethoxycurcumin, which demonstrated potent activity across multiple PDAC cell lines. At a concentration of 10 μM, it substantially reduced cell viability and completely inhibited colony formation, key indicators of anti-proliferative efficacy and the ability to impair long-term clonogenic potential. The compound's $IC_{50}$ values ranged from 2.91 to 12.90 μM, depending on the cell line, highlighting consistent anti-tumor activity in a physiologically relevant concentration range. Although these $IC_{50}$ values are higher in absolute terms (~1 order of magnitude) than that of the original query compound, BDBM50302828—reported to inhibit ABCC3 with an $IC_{50}$ of 1.1 μM in a membrane vesicle-based assay—it is important to consider the distinct nature of the experimental models. BDBM50302828 was evaluated in a cell-free biochemical system designed to isolate transporter function [33], whereas dimethoxycurcumin's activity was assessed using whole-cell viability assays. These assays inherently integrate additional layers of biological complexity, including compound uptake, intracellular metabolism, off-target interactions, and potential synergistic mechanisms. As such, a direct numerical comparison of $IC_{50}$ values between these systems may be misleading; instead, the results should be interpreted qualitatively, with an emphasis on the translational relevance of whole-cell phenotypic responses.

Beyond its role as a transporter inhibitor, dimethoxycurcumin's antioxidant properties may act synergistically to enhance its anticancer activity. Oxidative stress is a well-established driver of cancer progression and chemoresistance, with high levels of reactive oxygen species (ROS) influencing tumor cell survival, metastasis, and drug sensitivity [34]. While dimethoxycurcumin shows slightly lower antioxidant efficacy than curcumin, particularly in scavenging peroxyl radicals, it remains equally effective against superoxide radicals and exhibits similar activity in antioxidant pathways such as heme oxygenase-1 induction [35]. By simultaneously inhibiting ABCC3-mediated drug resistance and mitigating oxidative stress, dimethoxycurcumin may exert a dual mechanism of action, increasing its therapeutic potential.

Furthermore, the identification of dimethoxycurcumin extends the promising therapeutic potential of curcumin derivatives. Although the broad anticancer and antioxidant effects of curcumin have been extensively studied [36], specific targeting of ABCC3 has remained largely unexplored. Previous work has reported curcumin's effects on ABCC1 (MRP1) [37] and ABCC5 (MRP5) [38], attributing transporter inhibition largely to its polyphenolic core structure. The present findings suggest ABCC3 as an additional, therapeutically relevant target for optimized curcumin analogs, highlighting the potential improvement of structural modifications, such as methylation, in enhancing transporter specificity and anticancer efficacy.

An acknowledged limitation in the clinical development of curcumin is its poor pharmacokinetic profile, including low solubility, limited bioavailability, and rapid systemic clearance [39]. Dimethoxycurcumin has been reported to exhibit improved metabolic stability [40], which may enhance its therapeutic performance compared to curcumin. However, further improvements in solubility and bioavailability may still be needed to optimize its clinical potential. Advanced delivery strategies such as cyclodextrin (CD) encapsulation, liposomes, and nanoparticle-based systems have shown promise in overcoming these limitations [41,42]. Cyclodextrins, cyclic oligosaccharides derived from starch degradation, form inclusion complexes with hydrophobic molecules, thereby improving their solubility and metabolic stability [43]. In particular, β-cyclodextrin and its derivatives, such as hydroxypropyl-β-cyclodextrin (HP-β-CD) and randomly methylated-β-cyclodextrin (RM-β-CD), have been shown to enhance the aqueous solubility and bioavailability of curcumin [44]. Applying similar formulation approaches to dimethoxycurcumin may further improve its pharmacokinetic properties and support its development as a clinically viable anticancer agent.

Moving forward, several promising avenues warrant further exploration. Firstly, elucidating the precise molecular interactions between dimethoxycurcumin and ABCC3 via experimental biochemical studies could provide invaluable insights into its mechanism of transporter inhibition. Secondly, assessing the compound's specificity toward ABCC3 relative to other ABC transporters would better define its potential therapeutic window and minimize off-target effects. Finally, in vivo efficacy studies, particularly in relevant animal models of PDAC, combined with existing chemotherapeutics, could provide critical preclinical validation. Optimizing pharmacokinetics through strategic formulation approaches—such as those previously explored for curcumin analogs—could further enhance dimethoxycurcumin's in vivo stability, improving its therapeutic potential.

## 5. Conclusions

Multidrug resistance (MDR) in pancreatic ductal adenocarcinoma (PDAC) remains a critical barrier to effective chemotherapy. A consensus fingerprint-based virtual screening approach was used to identify structurally diverse candidate inhibitors of the ABC transporter ABCC3. Following the initial hit selection, predicted pharmacokinetic properties were evaluated, and molecular docking was performed to explore potential binding modes. Docking results indicated that the compounds occupied a similar region within the ABCC3 translocation channel, though their binding poses varied in orientation and interaction patterns, reflecting scaffold diversity.

In vitro evaluation of the top-ranked compounds revealed variable levels of growth inhibition across pancreatic cancer cell lines. In contrast to the moderate or limited activity observed in most candidates, dimethoxycurcumin consistently demonstrated the strongest effects. In viability assays performed on BxPC-3 and HPAF-II cells, treatment with 10 μM dimethoxycurcumin led to a significant reduction in cell viability. In colony formation assays conducted using BxPC-3 cells, only dimethoxycurcumin resulted in complete sup-

pression of colony growth. Based on this pronounced activity, a dose–response analysis was subsequently performed for dimethoxycurcumin across three cell lines—BxPC-3, HPAF-II, and CFPAC-1—confirming low micromolar $IC_{50}$ values in all models.

The convergence of computational predictions with experimental outcomes supports the potential of dimethoxycurcumin to modulate ABCC3 activity and impair drug resistance mechanisms. In addition to transporter inhibition, the antioxidant properties associated with this compound may contribute to broader anti-tumor effects, including the modulation of oxidative stress and the enhancement of chemosensitivity.

These findings reinforce the value of integrated computational and experimental strategies in anticancer drug discovery and highlight dimethoxycurcumin as a promising candidate for further development. Future work should focus on validating ABCC3-specific inhibition, characterizing selectivity across ABC transporters, and assessing pharmacokinetic behavior in vivo, with attention to formulation techniques that may enhance bioavailability.

**Author Contributions:** Conceptualization, J.N. and H.P.-S.; methodology, J.N.; software, J.N.; validation, J.N. and V.N.; formal analysis, J.N. and V.N.; investigation, J.N. and V.N.; resources, J.M.V.-S., M.F., and H.P.-S.; data curation, J.N.; writing—original draft preparation, J.N.; writing—review and editing, V.N., J.M.V.-S., M.F. and H.P.-S.; visualization, J.N. and V.N.; supervision, J.M.V.-S., M.F. and H.P.-S. All authors have read and agreed to the published version of the manuscript.

**Funding:** J.N. is funded by Cátedra Villapharma-UCAM. Powered@NLHPC: This research was partially supported by the supercomputing infrastructure of the NLHPC (CCSS210001). Supercomputing resources in this work have been supported by the Plataforma Andaluza de Bioinformática of the University of Málaga.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All data supporting the findings of this study are available from the corresponding author upon reasonable request. However, data derived from or related to the Eurofins-VillaPharma compound library are proprietary and cannot be shared due to confidentiality agreements.

**Conflicts of Interest:** M.F. is a member of LIPOVEXA S.r.l., a spin-off company focused on developing innovative treatments for diabetes, obesity and liver health J.M.V.-S. is affiliated with Eurofins-VillaPharma, a company focused on synthesis of novel compounds for drug discovery. The other authors declare no conflicts of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| ABCC3 | ATP-binding cassette subfamily C member 3 |
| (HP/RM-β-)CD | (Hydroxypropyl/randomly methylated-β)-Cyclodextrin |
| CDK | Chemistry Development kit |
| ConFiLiS | Consensus Fingerprints for Ligand-based Screening |
| DAPI | 4′,6-diamidino-2-phenylindole |
| Dist | Euclidean distance |
| ECFP(4/6) | Extended Connectivity Fingerprint (with a diameter of 4/6) |
| HCS | High-content screening |
| MDR | Multidrug resistance |
| MRP3 | Multidrug resistance-associated protein 3 |
| Norm | Normalized Euclidean distance |
| OD | Optical density |
| PDAC | Pancreatic ductal adenocarcinoma |
| PLIP | Protein–Ligand Interaction Profiler |
| ROS | Reactive oxygen species |

## References


1. Hussain, S.; Singh, A.; Nazir, S.U.; Tulsyan, S.; Khan, A.; Kumar, R.; Bashir, N.; Tanwar, P.; Mehrotra, R. Cancer Drug Resistance: A Fleet to Conquer. *J. Cell. Biochem.* **2019**, *120*, 14213–14225. [CrossRef] [PubMed]
2. Bukowski, K.; Kciuk, M.; Kontek, R. Mechanisms of Multidrug Resistance in Cancer Chemotherapy. *Int. J. Mol. Sci.* **2020**, *21*, 3233. [CrossRef] [PubMed]
3. Adamska, A.; Elaskalani, O.; Emmanouilidi, A.; Kim, M.; Abdol Razak, N.B.; Metharom, P.; Falasca, M. Molecular and Cellular Mechanisms of Chemoresistance in Pancreatic Cancer. *Adv. Biol. Regul.* **2018**, *68*, 77–87. [CrossRef]
4. Robey, R.W.; Pluchino, K.M.; Hall, M.D.; Fojo, A.T.; Bates, S.E.; Gottesman, M.M. Revisiting the Role of ABC Transporters in Multidrug-Resistant Cancer. *Nat. Rev. Cancer* **2018**, *18*, 452–464. [CrossRef]
5. Leonard, G.D.; Fojo, T.; Bates, S.E. The Role of ABC Transporters in Clinical Practice. *Oncologist* **2003**, *8*, 411–424. [CrossRef]
6. Adamska, A.; Ferro, R.; Lattanzio, R.; Capone, E.; Domenichini, A.; Damiani, V.; Chiorino, G.; Akkaya, B.G.; Linton, K.J.; De Laurenzi, V.; et al. ABCC3 Is a Novel Target for the Treatment of Pancreatic Cancer. *Adv. Biol. Regul.* **2019**, *73*, 100634. [CrossRef]
7. Fletcher, J.I.; Williams, R.T.; Henderson, M.J.; Norris, M.D.; Haber, M. ABC Transporters as Mediators of Drug Resistance and Contributors to Cancer Cell Biology. *Drug Resist. Updat. Rev. Comment. Antimicrob. Anticancer Chemother.* **2016**, *26*, 1–9. [CrossRef]
8. Falasca, M.; Linton, K.J. Investigational ABC Transporter Inhibitors. *Expert Opin. Investig. Drugs* **2012**, *21*, 657–666. [CrossRef]
9. Emmanouilidi, A.; Casari, I.; Gokcen Akkaya, B.; Maffucci, T.; Furic, L.; Guffanti, F.; Broggini, M.; Chen, X.; Maxuitenko, Y.Y.; Keeton, A.B.; et al. Inhibition of the Lysophosphatidylinositol Transporter ABCC1 Reduces Prostate Cancer Cell Growth and Sensitizes to Chemotherapy. *Cancers* **2020**, *12*, 2022. [CrossRef]
10. Adamska, A.; Domenichini, A.; Capone, E.; Damiani, V.; Akkaya, B.G.; Linton, K.J.; Di Sebastiano, P.; Chen, X.; Keeton, A.B.; Ramirez-Alcantara, V.; et al. Pharmacological Inhibition of ABCC3 Slows Tumour Progression in Animal Models of Pancreatic Cancer. *J. Exp. Clin. Cancer Res.* **2019**, *38*, 312. [CrossRef]
11. Juan-Carlos, P.-D.M.; Perla-Lidia, P.-P.; Stephanie-Talia, M.-M.; Mónica-Griselda, A.-M.; Luz-María, T.-E. ABC Transporter Superfamily. An Updated Overview, Relevance in Cancer Multidrug Resistance and Perspectives with Personalized Medicine. *Mol. Biol. Rep.* **2021**, *48*, 1883–1901. [CrossRef] [PubMed]
12. Pathak, K.; Pathak, M.P.; Saikia, R.; Gogoi, U.; Sahariah, J.J.; Zothantluanga, J.H.; Samanta, A.; Das, A. Cancer Chemotherapy via Natural Bioactive Compounds. *Curr. Drug Discov. Technol.* **2022**, *19*, e310322202888. [CrossRef] [PubMed]
13. Quideau, S.; Deffieux, D.; Douat-Casassus, C.; Pouységu, L. Plant Polyphenols: Chemical Properties, Biological Activities, and Synthesis. *Angew. Chem. Int. Ed.* **2011**, *50*, 586–621. [CrossRef]
14. Hazafa, A.; Rehman, K.-U.-; Jahan, N.; Jabeen, Z. The Role of Polyphenol (Flavonoids) Compounds in the Treatment of Cancer Cells. *Nutr. Cancer* **2020**, *72*, 386–397. [CrossRef]
15. Slika, H.; Mansour, H.; Wehbe, N.; Nasser, S.A.; Iratni, R.; Nasrallah, G.; Shaito, A.; Ghaddar, T.; Kobeissy, F.; Eid, A.H. Therapeutic Potential of Flavonoids in Cancer: ROS-Mediated Mechanisms. *Biomed. Pharmacother. Biomed. Pharmacother.* **2022**, *146*, 112442. [CrossRef]
16. Weaver, B.A. How Taxol/Paclitaxel Kills Cancer Cells. *Mol. Biol. Cell* **2014**, *25*, 2677–2681. [CrossRef]
17. Domínguez, A.R.; Márquez, A.; Gumá, J.; Llanos, M.; Herrero, J.; Nieves, M.A.d.l.; Miramón, J.; Alba, E. Treatment of Stage I and II Hodgkin's Lymphoma with ABVD Chemotherapy: Results after 7 Years of a Prospective Study. *Ann. Oncol.* **2004**, *15*, 1798–1804. [CrossRef]
18. Khaiwa, N.; Maarouf, N.R.; Darwish, M.H.; Alhamad, D.W.M.; Sebastian, A.; Hamad, M.; Omar, H.A.; Orive, G.; Al-Tel, T.H. Camptothecin's Journey from Discovery to WHO Essential Medicine: Fifty Years of Promise. *Eur. J. Med. Chem.* **2021**, *223*, 113639. [CrossRef]
19. Gedeck, P.; Rohde, B.; Bartels, C. QSAR—How Good Is It in Practice? Comparison of Descriptor Sets on an Unbiased Cross Section of Corporate Data Sets. *J. Chem. Inf. Model.* **2006**, *46*, 1924–1936. [CrossRef]
20. Rogers, D.; Hahn, M. Extended-Connectivity Fingerprints. *J. Chem. Inf. Model.* **2010**, *50*, 742–754. [CrossRef]
21. CDK Development Team. PubChem Substructure Fingerprint. Available online: https://github.com/cdk/orchem/blob/master/doc/pubchem_fingerprints.txt (accessed on 12 March 2025).
22. Ji, H.; Deng, H.; Lu, H.; Zhang, Z. Predicting a Molecular Fingerprint from an Electron Ionization Mass Spectrum with Deep Neural Networks. *Anal. Chem.* **2020**, *92*, 8649–8653. [CrossRef] [PubMed]
23. RDKit: Open-Source Cheminformatics. Available online: https://www.rdkit.org (accessed on 12 March 2025).
24. Willighagen, E.L.; Mayfield, J.W.; Alvarsson, J.; Berg, A.; Carlsson, L.; Jeliazkova, N.; Kuhn, S.; Pluskal, T.; Rojas-Chertó, M.; Spjuth, O.; et al. The Chemistry Development Kit (CDK) v2.0: Atom Typing, Depiction, Molecular Formulas, and Substructure Searching. *J. Cheminformatics* **2017**, *9*, 33. [CrossRef]
25. Liu, T.; Hwang, L.; Burley, S.K.; Nitsche, C.I.; Southan, C.; Walters, W.P.; Gilson, M.K. BindingDB in 2024: A FAIR Knowledgebase of Protein-Small Molecule Binding Data. *Nucleic Acids Res.* **2025**, *53*, D1633–D1644. [CrossRef]
26. *Schrödinger Release 2024-3: Maestro*; Schrödinger, LLC: New York, NY, USA, 2024.
27. *Schrödinger Release 2024-3: QikProp*; Schrödinger, LLC: New York, NY, USA, 2024.

28. Trott, O.; Olson, A.J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461. [CrossRef]
29. BIO-HPC Research Group. MetaScreener. Available online: https://github.com/bio-hpc/metascreener (accessed on 12 March 2025).
30. Tapia-Abellán, A.; Angosto-Bazarra, D.; Martínez-Banaclocha, H.; De Torre-Minguela, C.; Cerón-Carrasco, J.P.; Pérez-Sánchez, H.; Arostegui, J.I.; Pelegrin, P. MCC950 Closes the Active Conformation of NLRP3 to an Inactive State. *Nat. Chem. Biol.* **2019**, *15*, 560–564. [CrossRef]
31. Salentin, S.; Schreiber, S.; Haupt, V.J.; Adasme, M.F.; Schroeder, M. PLIP: Fully Automated Protein–Ligand Interaction Profiler. *Nucleic Acids Res.* **2015**, *43*, W443–W447. [CrossRef]
32. Schrödinger, LLC. *The PyMOL Molecular Graphics System*, Version 2.6.2; Schrödinger, LLC: New York, NY, USA, 2025.
33. Morgan, R.E.; van Staden, C.J.; Chen, Y.; Kalyanaraman, N.; Kalanzi, J.; Dunn, R.T.I.; Afshari, C.A.; Hamadeh, H.K. A Multifactorial Approach to Hepatobiliary Transporter Assessment Enables Improved Therapeutic Compound Development. *Toxicol. Sci.* **2013**, *136*, 216–241. [CrossRef]
34. Liou, G.-Y.; Storz, P. Reactive Oxygen Species in Cancer. *Free Radic. Res.* **2010**, *44*, 479–496. [CrossRef]
35. Kunwar, A.; Barik, A.; Sandur, S.K.; Indira Priyadarsini, K. Differential Antioxidant/pro-Oxidant Activity of Dimethoxycurcumin, a Synthetic Analogue of Curcumin. *Free Radic. Res.* **2011**, *45*, 959–965. [CrossRef]
36. Giordano, A.; Tommonaro, G. Curcumin and Cancer. *Nutrients* **2019**, *11*, 2376. [CrossRef]
37. Chearwae, W.; Wu, C.-P.; Chu, H.-Y.; Lee, T.R.; Ambudkar, S.V.; Limtrakul, P. Curcuminoids Purified from Turmeric Powder Modulate the Function of Human Multidrug Resistance Protein 1 (ABCC1). *Cancer Chemother. Pharmacol.* **2006**, *57*, 376–388. [CrossRef] [PubMed]
38. Li, Y.; Revalde, J.L.; Reid, G.; Paxton, J.W. Modulatory Effects of Curcumin on Multi-Drug Resistance-Associated Protein 5 in Pancreatic Cancer Cells. *Cancer Chemother. Pharmacol.* **2011**, *68*, 603–610. [CrossRef]
39. Jiang, M.; Gan, Y.; Li, Y.; Qi, Y.; Zhou, Z.; Fang, X.; Jiao, J.; Han, X.; Gao, W.; Zhao, J. Protein-Polysaccharide-Based Delivery Systems for Enhancing the Bioavailability of Curcumin: A Review. *Int. J. Biol. Macromol.* **2023**, *250*, 126153. [CrossRef]
40. Teymouri, M.; Barati, N.; Pirro, M.; Sahebkar, A. Biological and Pharmacological Evaluation of Dimethoxycurcumin: A Metabolically Stable Curcumin Analogue with a Promising Therapeutic Potential. *J. Cell. Physiol.* **2018**, *233*, 124–140. [CrossRef]
41. Pinho, E.; Grootveld, M.; Soares, G.; Henriques, M. Cyclodextrins as Encapsulation Agents for Plant Bioactive Compounds. *Carbohydr. Polym.* **2014**, *101*, 121–135. [CrossRef]
42. Gayathri, K.; Bhaskaran, M.; Selvam, C.; Thilagavathi, R. Nano Formulation Approaches for Curcumin Delivery- a Review. *J. Drug Deliv. Sci. Technol.* **2023**, *82*, 104326. [CrossRef]
43. Dhiman, P.; Bhatia, M. Pharmaceutical Applications of Cyclodextrins and Their Derivatives. *J. Incl. Phenom. Macrocycl. Chem.* **2020**, *98*, 171–186. [CrossRef]
44. Patro, N.M.; Sultana, A.; Terao, K.; Nakata, D.; Jo, A.; Urano, A.; Ishida, Y.; Gorantla, R.N.; Pandit, V.; Devi, K.; et al. Comparison and Correlation of in Vitro, in Vivo and in Silico Evaluations of Alpha, Beta and Gamma Cyclodextrin Complexes of Curcumin. *J. Incl. Phenom. Macrocycl. Chem.* **2014**, *78*, 471–483. [CrossRef]

## 3.4. MATCHED PAIRS DEMONSTRATE ROBUSTNESS AGAINST INTER-ASSAY VARIABILITY

| | |
|---|---|
| **Title** | Matched pairs demonstrate robustness against inter-assay variability |
| **Authors** | Jochem Nelen; Horacio Pérez-Sánchez; Hans De Winter; Dries Van Rompaey |
| **Journal** | Journal of Cheminformatics |
| **Year** | 2025 |
| **Volume** | 17 (1) |
| **Pages** | 8 |
| **DOI** | https://doi.org/10.1186/s13321-025-00956-y |
| **IF (JCR 2024)** | 5.7 |
| **Category** | Chemistry, Multidisciplinary; 58/239; Q1<br>Computer Science, Information Systems; 35/258; Q1<br>Computer Science, Interdisciplinary Applications; 32/175; Q1 |

**Contribution of PhD Candidate**

The PhD candidate, Jochem Nelen, confirms that he is the first author of the article titled "*Matched pairs demonstrate robustness against inter-assay variability*". He was primarily responsible for the conception of the study, development and implementation of the methodology, data analysis, manuscript writing, and overall coordination of the research work.

Nelen *et al. Journal of Cheminformatics* (2025) 17:8
https://doi.org/10.1186/s13321-025-00956-y

Journal of Cheminformatics

BRIEF REPORT Open Access

# Matched pairs demonstrate robustness against inter-assay variability

Jochem Nelen[1,2], Horacio Pérez-Sánchez[1], Hans De Winter[3*] and Dries Van Rompaey[4]

**Abstract**

Machine learning models for chemistry require large datasets, often compiled by combining data from multiple assays. However, combining data without careful curation can introduce significant noise. While absolute values from different assays are rarely comparable, trends or differences between compounds are often assumed to be consistent. This study evaluates that assumption by analyzing potency differences between matched compound pairs across assays and assessing the impact of assay metadata curation on error reduction. We find that potency differences between matched pairs exhibit less variability than individual compound measurements, suggesting systematic assay differences may partially cancel out in paired data. Metadata curation further improves inter-assay agreement, albeit at the cost of dataset size. For minimally curated compound pairs, agreement within 0.3 pChEMBL units was found to be 44–46% for $K_i$ and $IC_{50}$ values respectively, which improved to 66–79% after curation. Similarly, the percentage of pairs with differences exceeding 1 pChEMBL unit dropped from 12 to 15% to 6–8% with extensive curation. These results establish a benchmark for expected noise in matched molecular pair data from the ChEMBL database, offering practical metrics for data quality assessment.

**Keywords** Matched structural pairs, Assay noise, Data curation, ChEMBL, Machine learning

## Introduction

In recent work, Landrum and Riniker demonstrated that combining $IC_{50}$ or $K_i$ measurements from different experiments can introduce substantial noise to datasets [1]. They demonstrated this by comparing assay measurements from the ChEMBL32 database [2] across different assays for the same target. Similarly, previous studies have highlighted the challenges posed by assay variability and prediction errors in biological datasets, emphasizing their impact on model reliability and the interpretation of results [3, 4]. Landrum and Riniker did report that data curation can partially help to mitigate these effects, but that the overall amount of noise remains high. In the current era of machine learning (ML) and artificial intelligence (AI), where the quality of training data is crucial for building accurate and robust models, understanding and quantifying the noise introduced by combining data from multiple sources is essential. Noisy data can lead to reduced model performance and unreliable predictions, underscoring the importance of identifying best practices for data collection and integration. A common assumption is that within-assay comparisons between different compounds are more reliable than direct between-assay comparisons, a rationale which is also used for Matched Molecular Pair Analysis (MMPA) [5]. MMPA can be used to propose new compounds or to calculate the potential effect of a given modification to a compound. The MMPA

*Correspondence:
Hans De Winter
hans.dewinter@uantwerpen.be
[1] Structural Bioinformatics and High Performance Computing Research Group (BIO-HPC), HiTech Innovation Hub, UCAM Universidad Católica de Murcia, 30107 Murcia, Spain
[2] Health Sciences PhD Program, Universidad Católica de Murcia UCAM, 30107 Murcia, Spain
[3] Department of Pharmaceutical Sciences, Faculty of Pharmaceutical, Biomedical and Veterinary Sciences, University of Antwerp, Universiteitsplein 1A, 2610 Wilrijk, Belgium
[4] Drug Discovery Data Sciences, Janssen Pharmaceutica NV, Turnhoutseweg 30, 2340 Beerse, Belgium

paradigm has also given rise to various machine learning approaches which aim to improve upon the performance of standard MMPA [6–8]. However, the impact of assay variability on these methods remains underexplored. In this study, we expand on Landrum and Riniker's previous research by identifying and leveraging structural analogs within the ChEMBL32 database to further investigate the relative variability in bioactivity measurements across different assays. By focusing on matched molecular pairs, we aim to provide a quantified estimate of the expected noise, which is helpful to contextualize the performance of models trained on such matched pair data.

## Method

We analyzed pairwise differences in $K_i$ and $IC_{50}$ values among structural analogs across assays to evaluate the consistency of trends in the ChEMBL database [2]. We built our data curation process upon the established workflow of Landrum and Riniker [1], which investigated assay noise by comparing affinity measurements directly across different assays. For ease of comparison, we also used ChEMBL32 as the primary database, and we performed a similar minimal and maximal curation procedure for all $K_i$ and $IC_{50}$ data in ChEMBL, in order to investigate if more careful curation can improve overall trends and data quality. We note that the extension of our work to new versions of ChEMBL is straightforward.

The minimal curation procedure involved a relatively lenient filtering process, where assays were considered comparable if they were measured for the same protein target and had at least five compounds in common. Additionally, activity curation was applied to filter out compounds with identical activity values. This process also removed compounds with activity values that differed by exactly 3.0 log units to account for a unit error, such as incorrect annotation of units (e.g., μM instead of nM or vice versa). The aim of activity curation was to avoid accidental duplicates in the results, as these could greatly impact the findings. By incorporating these specific criteria and filtering steps, we ensured that the curation process mostly removed true duplicates, as it is improbable for different measurements to coincidentally report the exact same affinity value.

The maximal curation (maxcur) procedure also applied the same basic curation steps as the minimal curation process (mincur), including activity curation and requiring an overlap of five or more compounds between two assays. Furthermore, several additional filtering steps were included before considering two assays to be compatible for the maxcur procedure. In short, only compounds with a high confidence score as assigned by ChEMBL were kept, as well as data that came directly from published papers. Additionally, any assays that were associated with mutant/variant proteins or had multiple assays for the same target were filtered out. Finally, assays

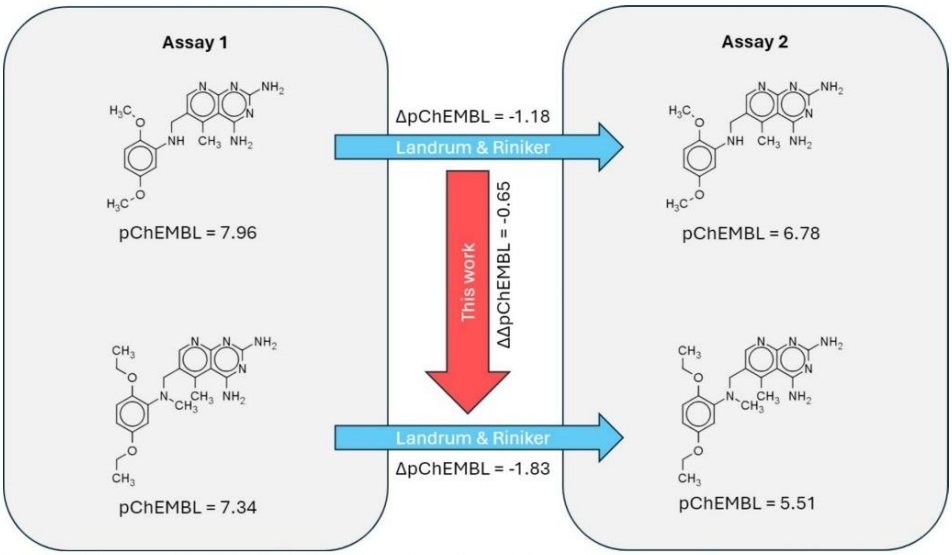



**Fig. 1** Schematic Overview of ΔΔpChEMBL Calculation for Matched Molecular Pairs

**Table 1** Pairwise metrics and dataset characteristics for the matched pairs of structural analogs

| Curation level | MAE | F > 0.3 | F > 1.0 | Assay pairs | Compound pairs |
|---|---|---|---|---|---|
| $IC_{50}$ minimal | 0.33 | 0.54 | 0.12 | 1,338 | 372,683 |
| $IC_{50}$ maximal | 0.18 | 0.34 | 0.06 | 24 | 1,773 |
| $K_i$ minimal unpruned | 0.36 | 0.56 | 0.18 | 584 | 36,189 |
| $K_i$ maximal unpruned | 0.40 | 0.62 | 0.43 | 279 | 2,900 |
| $K_i$ minimal pruned | 0.35 | 0.55 | 0.15 | 219 | 32,363 |
| $K_i$ maximal pruned | 0.03 | 0.21 | 0.08 | 8 | 311 |

Summary of pairwise metrics (Median Absolute Error, fractions > 0.3 and 1.0) and dataset characteristics (assay pairs and compound pairs) for all the tested datasets and curation procedures. The pruned datasets have all assays associated with Carbonic Anhydrase filtered away

were filtered based on their conditions metadata such as the type of assay (e.g. binding assay, functional assay, ...) and the target organism, and only assays that shared identical assay conditions metadata were considered compatible. For a more in-depth overview and reasoning of all of the individual curation steps, we refer to the original manuscript, where the same procedure was used [1]. We also investigated some potential "medium" curation procedures, aiming to balance the number of retained pairs with a low noise level. Starting from the maximal curation settings, we sequentially disabled individual settings which we deemed to be retain scientific validity (e.g. relaxing strict requirements on matching exact assay conditions or including lower-confidence results). However, none of the tested configurations consistently improved both $IC_{50}$ and $K_i$ datasets. Typically, relaxing the curation settings resulted in either minimal gains in new datapoints or a substantial increase in noise, with none of the options examined herein offering a suitable middle ground. Additional details can be found in the Supplementary Information (SI): Figure S1 presents the $IC_{50}$ data, Figure S2 shows the $K_i$ data, and Table S1 summarizes the performance statistics for both.

Following the data curation process, we identified structural analogs between compatible assays for the minimal and maximal curated $IC_{50}$ and $K_i$ data. This identification process involved two stages. Pairs within the same assay were identified using rdRascalMCES, an RDkit [9] implementation of RASCAL [10], which is an algorithm that uses bond matching to efficiently identify the Maximum Common Edge Subgraphs (MCES) between two molecules. The rdRascalMCES method was selected for its ability to directly identify the largest common scaffold and quickly terminate calculations for pairs with initial low Johnson similarity estimates, making it well-suited for analyzing large datasets. For the specific rdRascalMCES settings, we used 0.6 as the Johnson similarity threshold and allowed partial aromatic rings to match. We refer to Figure S3 in the supplementary information for concrete examples. Subsequently, we compared the identified intra-assay pairs with their compatible assays, as determined by the initial curation step. If an exact match between two pairs across assays was found, it was considered a matched pair of structural analogs. All relevant data such as the associated ligand and assay ChEMBL IDs, reported pChEMBL affinity values and protein target ID were retained for subsequent analysis. For each matched pair, we calculated the difference in pChEMBL values within the same assay, representing the relative potency difference between the two molecules. We then compared the intra-assay differences for the matched pair across the two assays, defining this difference between the intra-assay differences as the ΔΔpChEMBL. Figure 1 provides a schematic overview of the calculation of the ΔΔpChEMBL for a matched pair.

The final part of the workflow involved conducting a statistical analysis based on earlier work by Landrum and Riniker [1]. The Median Absolute Error (MAE), as well as the fractions of values exceeding 0.3 (F > 0.3) and 1.0 (F > 1.0) were computed for each dataset. In contrast to the work of Landrum and Riniker, the $R^2$ and Kendall Tau metrics were not included, as these correlation metrics would be sensitive to both the arbitrary assignment of the direction of the transformation (e.g. depending on the direction of the transformation, signs may be positive or negative), as well as the range spanned by the assay measurements. Additionally, we generated a histogram of the ΔΔpChEMBL values to provide a complementary overview of the overall distribution of differences. RDKit version 2024.03.1 was used for pair analytics and cheminformatics. Data processing was performed using pandas 2.2.2, and plots were generated with matplotlib 3.8.4.

For the $K_i$ datasets, we also applied data pruning inspired by the original manuscript's findings [1], which identified that several Carbonic Anhydrase (CA) assays were responsible for a substantial part of the noise observed. However, instead of only pruning away the assays indicated in their work, we decided to remove all data associated with CA assays. This decision was made after manually inspecting the pruned data, which revealed that a cluster of large outliers associated with CA assays persisted even after excluding the assays indicated in the original paper (Figure S4). The Supplementary Information includes comparative figures (Figure S5-6) and details regarding the pruned assays for each procedure. Although both methods produce similar trends and figures, the more extensive pruning yields

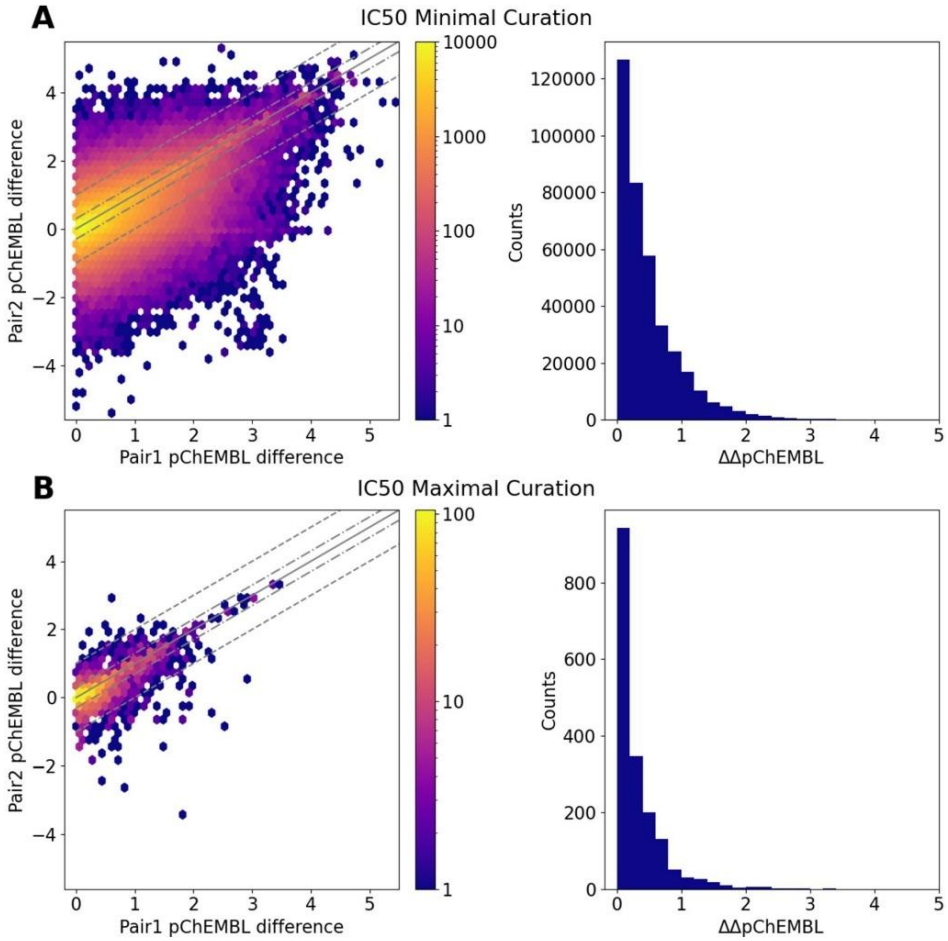



**Fig. 2** Hexbin plots (left) and histograms (right) for the ΔΔpChEMBL $IC_{50}$ data. **A** Minimal curation procedure applied to the $IC_{50}$ data resulting in 372,683 data points. **B** Maximal curation procedure performed for the $IC_{50}$ data resulting in 1,773 data points. The colors indicate the number of datapoints in that area. To highlight the 0.3 and 1.0 log difference, dotted lines are drawn as guidelines

better results, with improved pairwise metrics and fewer outliers.

## Results and discussion

We first sought to estimate the inter-assay differences in potency between the differences of matched compound pairs present in combined data sets. Table 1 summarizes the pairwise metrics for the mincur and maxcur procedures applied to both the $IC_{50}$ and $K_i$ data. Specifically, we evaluated the noise of each dataset using three metrics: the Median Absolute Error (MAE), and the fraction of values exceeding 0.3 and 1.0 log units. These metrics were chosen based on the original manuscript's thresholds, which consider differences smaller than 0.3 log units to be within acceptable experimental noise limits, and differences greater than 1 log unit to be indicative of significant discrepancies. Additionally, the number of assay pairs, and total compound pairs (data points) are provided to give context regarding the dataset sizes.

The $IC_{50}$ data, also visualized in Fig. 2, demonstrates the impact of curation on data quality. The minimal curation

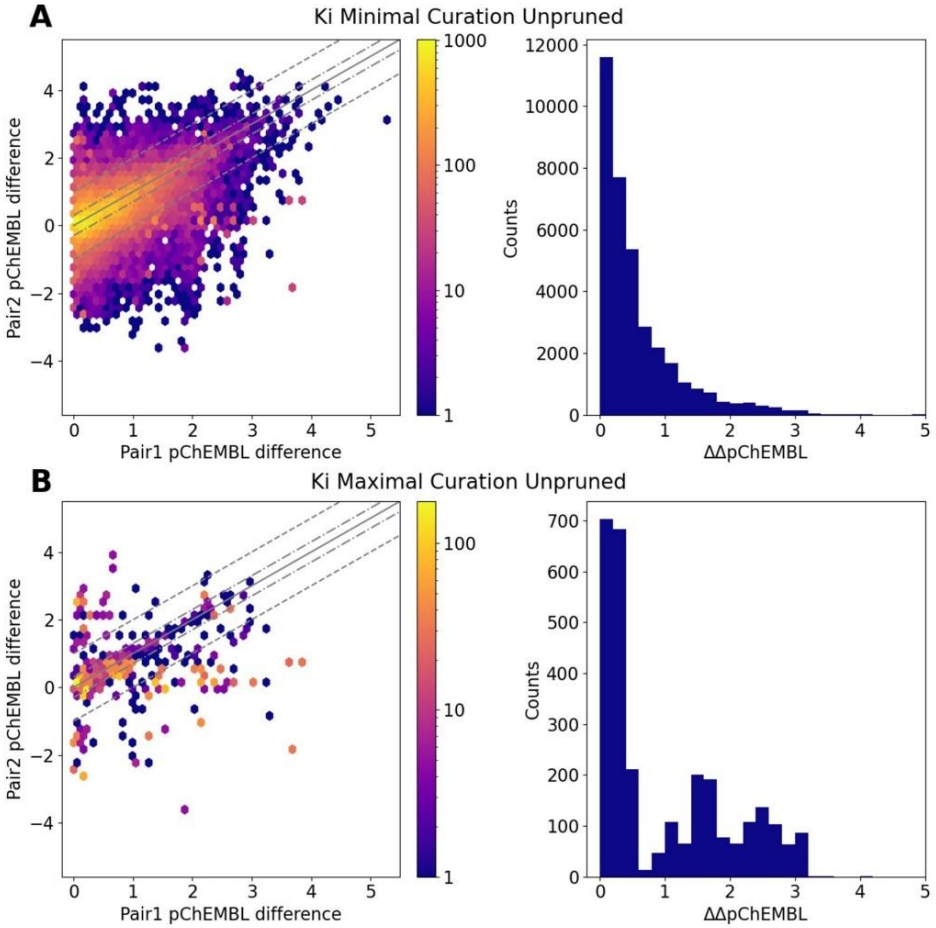



**Fig. 3** Hexbin plots (left) and histograms (right) for the unpruned ΔΔpChEMBL $K_i$ data. **A** Minimal curation procedure applied to the $K_i$ data resulting in 36,189 data points. **B** Maximal curation procedure performed for the $K_i$ data resulting in 2,900 data points. The colors indicate the number of datapoints in that area. To highlight the 0.3 and 1.0 log difference, dotted lines are drawn as guidelines.

dataset (Fig. 2A) contains just under 372,700 data points, which is reduced to around 1,770 with maxcur (Fig. 2B). As shown in Table 1, this reduction is accompanied by substantial improvements in data quality, including a 45% reduction in MAE, a 37% decrease in F > 0.3, and a 50% decrease in F > 1.0.

Interestingly, the unpruned $K_i$ data (Fig. 3) show a different trend. The metrics worsen when going from the minimal to the maximal curation procedure (Table 1). This is especially apparent for F > 1.0, which saw an increase from 18 to 43%. Both datasets are visualized in Figs. 3A and B respectively. The histogram in Fig. 3B also displays an irregular pattern, indicating the presence of a substantial amount of noise in the data. This observation prompted us to perform data pruning as described in the methods section.

After pruning away all data associated with Carbonic Anhydrase assays, substantial improvements are observed (Fig. 4). The pruned minimal curation data (Fig. 4A) show minimal changes in pairwise metrics

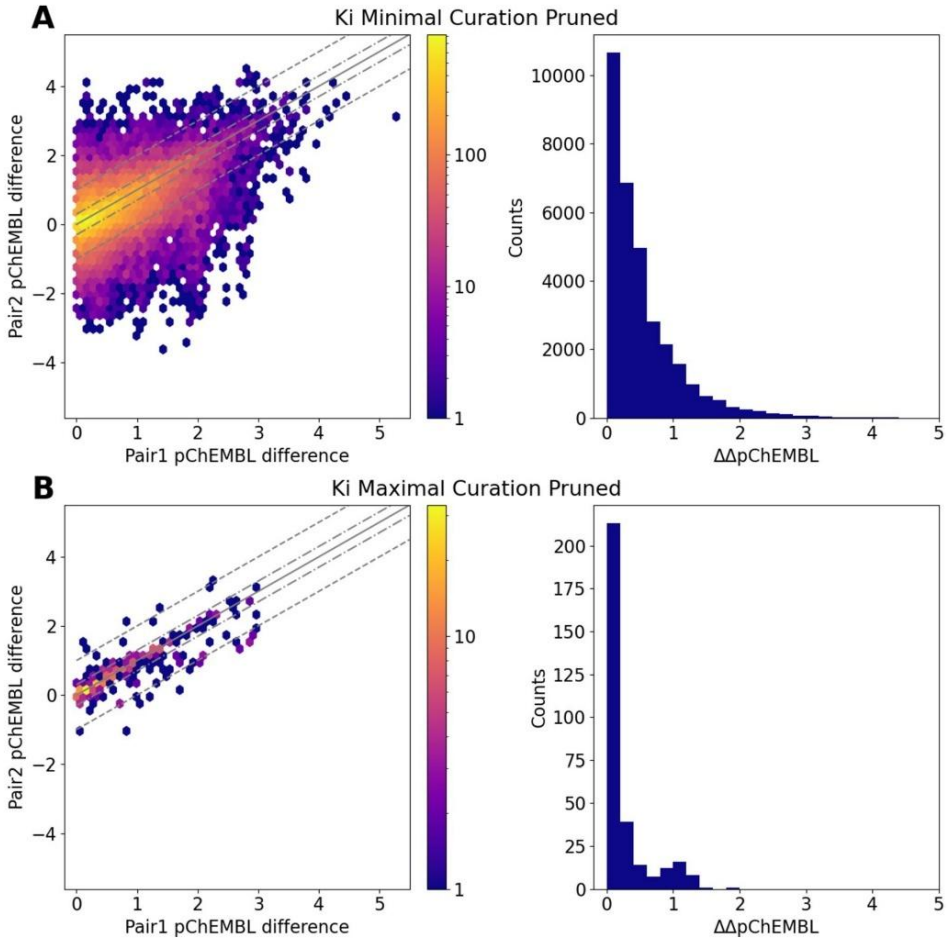



**Fig. 4** Hexbin plots (left) and histograms (right) for the unpruned ΔΔpChEMBL $K_i$ data. **A** Minimal curation procedure performed for the $K_i$ data resulting in 32,363 data points. **B**: Maximal curation procedure applied to the $K_i$ data resulting in 311 data points. The colors indicate the number of datapoints in that area, while lines indicating the 0.3 and 1.0 log difference are drawn as dotted lines

compared to the unpruned minimal curation data. In contrast to the unpruned datasets, the curation of the pruned $K_i$ data (Fig. 4B) does lead to substantial improvements. As quantified in Table 1, the MAE sees more than a 90% reduction, the $F > 0.3$ is reduced by over 60%, and $F > 1.0$ is halved. However, the pruned $K_i$ maxcur data are reduced from 32,363 to only 311 data points.

Table 2 presents a summary and comparison of our matched pair data (ΔΔpChEMBL) with the results reported in the original paper [1], which directly compared assay measurements (ΔpChEMBL). The errors observed by Landrum and Riniker are a combination of annotation errors, intra-assay variability (i.e. normal experimental variability) and inter-assay differences (e.g. different types of readout technology, different concentrations of medium or substrate, etc.).

Our data shows improved pairwise metrics in all but one comparison—in which noise from a single target had a drastic impact—indicating that comparing pairs of structural analogs reduces noise. Inter-assay differences, such as different substrate concentrations or lipophilicity-driven interactions with assay medium or

**Table 2** Comparison of matched pair data with the original manuscript results [1]

| Curation level | Type | MAE | F > 0.3 | F > 1.0 | Assay pairs | Compound pairs |
|---|---|---|---|---|---|---|
| $IC_{50}$ minimal | Ref | 0.50 | 0.64 | 0.27 | 1,358 | 38,022 |
| | Pair | 0.33 | 0.54 | 0.12 | 1,338 | 372,683 |
| $IC_{50}$ maximal | Ref | 0.27 | 0.48 | 0.13 | 26 | 340 |
| | Pair | 0.18 | 0.34 | 0.06 | 24 | 1,773 |
| $K_i$ minimal | Ref | 0.52 | 0.67 | 0.30 | 587 | 7,734 |
| | Pair | 0.36 | 0.56 | 0.18 | 584 | 36,189 |
| $K_i$ maximal | Ref | 0.47 | 0.69 | 0.32 | 282 | 2,434 |
| | Pair | 0.40 | 0.62 | 0.43 | 279 | 2,900 |
| $K_i$ maximal pruned | Ref | 0.45 | 0.58 | 0.25 | 9 | 115 |
| | Pair | 0.03 | 0.21 | 0.08 | 8 | 311 |

Note that the compound pairs for *Pair-type* datasets refers to cross-assay matched structural pairs, rather than pairs of individual compounds as in the *Ref-type* datasets. Additionally, reference pruning was less strict, leaving some Carbonic Anhydrase assays. These were removed in our workflow after identifying a cluster of outliers (Figure S4)

assay materials [11] will often have similar effects on the observed potencies of structurally similar compounds, resulting in cancellation of these sources of error when looking at the difference in potency. Consequently, assessing pairwise differences appears more robust than comparing direct assay measurements, though some variability still persists as shown in our analysis.

The single metric on which the matched pair data performed worse was the F > 1.0 for the unpruned $K_i$ maxcur data. This anomaly is due to the large fraction of noisy samples present in that dataset, which is amplified further due to the formation of pairs between them. It is also worth noting that while the pruning methods employed differ slightly between the two studies, the trends remain consistent, and we provide the data using the same pruning procedure in the Supplementary Information (Table S2) for comparison.

## Conclusions

In conclusion, our study finds that analyzing matched pairs is more robust to inter-assay variability than directly comparing assay measurements across assays, likely driven by the cancellation of systematic differences between assays. For minimally curated compound pairs, 44–46% exhibited a difference within 0.3 pChEMBL units, compared to only 33–36% for direct assay measurements. Similarly, the proportion of compound pairs with differences exceeding 1 pChEMBL unit was substantially lower for matched pairs (12–15%) compared to direct assay comparisons (27–30%). Additionally, our results show that careful data curation can help to mitigate noise introduced by combining datasets. Following curation, 66–79% of the pairs agreed within 0.3 pChEMBL units, up from 42–52% for direct assay comparisons. Furthermore, the proportion of compound pairs with differences exceeding 1 pChEMBL unit (F > 1.0) was reduced to 6–8% after curation, compared to 13–25% for direct comparisons.

Datasets with less noise enable the development of more performant machine learning models. However, stricter data curation also results in discarding a substantial amount of data. Practitioners should carefully consider this tradeoff when assembling datasets for MMPA or deep learning MMP methods. To facilitate further exploration, we provide all datasets used in this study, along with a Jupyter notebook that enables researchers to experiment with different curation settings and assess the impact of their choices on noise levels and robustness. Finally, our analysis reaffirms that careful examination of datasets for potential artifacts can improve data quality, as evidenced by the differences in pairwise metrics between the unpruned and pruned datasets.

**Abbreviations**

| | |
|---|---|
| AI | Artificial intelligence |
| CA | Carbonic anhydrase |
| MAE | Median absolute error |
| Maxcur | Maximal curation |
| MCES | Maximum common edge subgraphs |
| Mincur | Minimal curation |
| ML | Machine learning |
| MMPA | Matched molecular pair analysis |

## Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13321-025-00956-y.

Supplementary material 1.

**Acknowledgements**

We would like to thank Dr. Greg Landrum and Dr. Jeremy R. Ash for their thoughtful comments and insightful suggestions which helped to improve our methodology, results, and conclusions.

**Author contributions**
JN conducted the formal analysis, created the figures, and drafted the original manuscript. HDW and DVR conceptualised and supervised the study, and revised the manuscript to enhance its quality. HPS contributed to the refinement and finalization of the manuscript. All authors reviewed and approved the final version of the manuscript.

**Funding**
Jochem Nelen was funded by Cátedra Villapharma-UCAM.

**Availability of data and materials**
The code and datasets used in this analysis are publicly available on GitHub at https://github.com/Jnelen/ChEMBL_MatchedPairsAnalysis.

## Declarations

**Competing interests**
The authors declare no competing interests.

Received: 7 August 2024 Accepted: 11 January 2025
Published online: 20 January 2025

**References**

1. Landrum GA, Riniker S (2024) Combining IC50 or Ki values from different sources is a source of significant noise. J Chem Inf Model 64:1560–1567. https://doi.org/10.1021/acs.jcim.4c00049
2. Mendez D, Gaulton A, Bento AP et al (2019) ChEMBL: towards direct deposition of bioassay data. Nucleic Acids Res 47:D930–D940. https://doi.org/10.1093/nar/gky1075
3. Kramer C, Dahl G, Tyrchan C, Ulander J (2016) A comprehensive company database analysis of biological assay variability. Drug Discov Today 21:1213–1221. https://doi.org/10.1016/j.drudis.2016.03.015
4. Brown SP, Muchmore SW, Hajduk PJ (2009) Healthy skepticism: assessing realistic model performance. Drug Discov Today 14:420–427. https://doi.org/10.1016/j.drudis.2009.01.012
5. Kramer C, Fuchs JE, Whitebread S et al (2014) Matched Molecular Pair Analysis: Significance and the Impact of Experimental Uncertainty. J Med Chem 57:3786–3802. https://doi.org/10.1021/jm500317a
6. Jin W, Yang K, Barzilay R, Jaakkola T (2019) Learning Multimodal Graph-to-Graph Translation for Molecular Optimization
7. Fralish Z, Chen A, Skaluba P, Reker D (2023) DeepDelta: predicting ADMET improvements of molecular derivatives with deep learning. J Chemin 15:101. https://doi.org/10.1186/s13321-023-00769-x
8. Fralish Z, Skaluba P, Reker D (2024) Leveraging bounded datapoints to classify molecular potency improvements. RSC Med Chem. https://doi.org/10.1039/D4MD00325J
9. RDKit. In: RDKit: Open-Source Cheminformatics Software. https://www.rdkit.org/. Accessed 13 Jun 2024
10. Raymond JW, Gardiner EJ, Willett P (2002) RASCAL: calculation of graph similarity using maximum common edge subgraphs. Comput J 45:631–644. https://doi.org/10.1093/comjnl/45.6.631
11. Palmgrén JJ, Mönkkönen J, Korjamo T et al (2006) Drug adsorption to plastic containers and retention of drugs in cultured cells under in vitro conditions. Eur J Pharm Biopharm 64:369–378. https://doi.org/10.1016/j.ejpb.2006.06.005

## Publisher's Note

# IV – SUMMARY AND DISCUSSION OF RESULTS

# IV - SUMMARY AND DISCUSSION OF RESULTS

## 4.1. SUMMARY OF THE RESULTS

### 4.1.1. Article 1: ESSENCE-Dock: A Consensus-Based Approach to Enhance Virtual Screening Enrichment in Drug Discovery

This article presents ESSENCE-Dock (*Effective Structural Screening ENrichment ConsEnsus Dock*), a consensus docking method developed to improve the early identification of active compounds in virtual screening campaigns. The motivation for this work lies in the limitations of individual molecular docking algorithms, which often produce inconsistent results across different protein targets and tend to yield high rates of false positives. While other consensus docking strategies have been described before, the majority of the current approaches utilize older docking methods and usually don't incorporate predicted pose similarity, a feature that lies at the basis of ESSENCE-Dock.

In its published form, ESSENCE-Dock uses three different docking tools: LeadFinder, Gnina, and DiffDock. These three docking methods all feature fundamentally different pose sampling strategies and scoring functions. After docking, a final consensus score is calculated by combining docking scores, predicted similarity of binding pose (quantified as the root-mean-square deviation, RMSD), and ligand flexibility (computed from the number of rotatable bonds). This consensus score can subsequently be used to re-rank the compounds, prioritizing agreement of binding modes as predicted by the different docking methods.

The protocol has been integrated into MetaScreener (https://github.com/bio-hpc/metascreener), enabling efficient execution on high-performance computing clusters with only a few commands required to complete the entire screening workflow. ESSENCE-Dock also automatically produces detailed output that includes an interactive PyMOL session file with the best-ranked compounds together with their predicted protein–ligand interactions, along with convenient spreadsheet files of the results for further analysis.

A comparison across 21 targets from the DUD-E dataset showed that, on average, ESSENCE-Dock achieves higher enrichment factors than stand-alone docking methods and simple consensus baselines, particularly at stringent cutoffs such as the top 0.1% of ranked molecules. These results confirm the effectiveness of the ESSENCE-Dock to help prioritize active hit compounds. In addition, it was observed that the consensus score itself can be used as a proxy for result confidence, with low-scoring consensus frequently corresponding to disagreement between approaches and associating with unsuccessful enrichment, thus allowing more informed filtering of low-confidence predictions.

ESSENCE-Dock thus presents a robust, scalable, and convenient solution to structure-based virtual screening. By systematic combination of complementary docking strategies and use of agreement on binding poses, it effectively enhances the confidence of hit identification early in drug discovery.

### 4.1.2. Article 2: Targeting Drug Resistance in Cancer: Dimethoxycurcumin as a Functional Antioxidant Targeting ABCC3

This paper reports on the identification and *in vitro* validation of dimethoxycurcumin as a potential modulator of ABCC3 (MRP3), an ATP-binding cassette transporter involved in multidrug resistance and tumor development in pancreatic cancer. Dimethoxycurcumin was identified by using ConFiLiS (*Consensus Fingerprints for Ligand-based Screening*; publicly available at https://github.com/Jnelen/ConFiLiS), a ligand-based virtual screening method combining Avalon, ECFP6, and PubChem substructure fingerprints into consensus similarity scores. By averaging the normalized Euclidean distances of all fingerprint types, ConFiLiS aims to improve chemical diversity while reducing fingerprint bias from singular fingerprints in hit selection.

The in-house Eurofins-Villapharma library was screened using ConFiLiS, with the most active small-molecule ABCC3 inhibitor in BindingDB, BDBM50302828, used as the query. After screening, the highest ranked candidates were compared to all known ABCC3 inhibitors documented in BindingDB to ensure their novelty.

Initial screening of the compounds revealed that dimethoxycurcumin, a methylated derivative of curcumin, was the most active. These results were also confirmed in studies using three different pancreatic cancer cell lines. Dimethoxycurcumin inhibited BxPC-3 cell viability and colony formation at 10 μM. Dose–response experiments demonstrated low micromolar $IC_{50}$ values for all three cell lines, with the most potent effect against CFPAC-1 cells. In addition, dimethoxycurcumin possesses strong antioxidant activity that may contribute synergistically to its therapeutic benefit in this context.

In short, this study demonstrates the utility of ConFiLiS as an effective ligand-based virtual screen approach and identifies dimethoxycurcumin as a potential lead compound for further optimization. Its dual role as an ABCC3 modulator and antioxidant may have therapeutic benefits in the treatment of pancreatic cancer, particularly in the context of ABC transporter-mediated drug resistance.

### 4.1.3. Article 3: Matched pairs demonstrate robustness against inter-assay variability

This article investigates the reliability of using matched molecular pairs (MMP) as a way to potentially mitigate assay variability in large bioactivity databases. This is a significant challenge in machine learning for drug discovery, where both more and higher-quality data is needed to improve QSAR and AI models. It has been a known issue that absolute bioactivity values (e.g., $IC_{50}$ or $K_i$) across different assays are not necessarily directly comparable due to differences in assay conditions, annotation errors, and measurement schemes. Such inconsistencies can severely impair training data quality utilized in training machine learning models.

In an attempt to remedy this issue, the article investigates whether using potency differences between matched molecular compound pairs (i.e. structurally similar analogs) can be used as a more consistent comparative measure. Examining Ki and $IC_{50}$ values from the public ChEMBL32 database, the study employs statistical analysis to assess whether inter-assay variability is partially canceled out when considering relative differences between such analogs, as well as quantifying how and if data curation can affect this behavior.

Two levels of data curation were employed: a minimal curation step that filters away outliers and duplicates, and a maximal curation step that applies stricter assay metadata filtering and excludes low-confidence entries. Additionally, assays associated with carbonic anhydrase were excluded from the analysis after it was found that they contained a disproportionate amount of dataset noise.

The results show that differences in potency between matched pairs (ΔΔpChEMBL) are much more consistent than absolute differences (ΔpChEMBL). After maximal curation, 66-79% of matched pairs agreed within 0.3 log units, while for direct measurements this was only 42-52%. Furthermore only 6–8% exceeded a 1 log unit difference when using ΔΔpChEMBL, compared to 13–25% for unpaired measurements. This highlights the benefit of matched pair analysis for obtaining more consistent and less noisy datasets.

Overall, the paper provides insights in the noise present in open-source bioactivity databases, and how MMP data can be used to reduce inter-assay noise and increase the reliability of data used in cheminformatics and machine learning pipelines.

## 4.2. DISCUSSION OF THE RESULTS

### 4.2.1. Discussion

Each article in this thesis either uses or describes a tool or methodology to address different but complementary stages of early drug discovery, ranging from compound screening to lead optimization.

ESSENCE-Dock contributes to the field of structure-based virtual screening as a consensus framework that integrates docking scores, pose similarity, and ligand flexibility. Benchmarking showed consistent early enrichment of actives across diverse protein targets. In practical applications, ESSENCE-Dock has been used to identify high-affinity inhibitors in various contexts. One early example came from a drug repurposing project using a primitive version of what would later evolve into the published ESSENCE-Dock method. In this case, compounds were prioritized based on binding pose agreement between LeadFinder (31) and AutoDock Vina (30). Screening DrugBank (63), a database containing

investigational and FDA-approved drugs, led to the identification of gilteritinib as an inhibitor of protein-arginine deiminase 4 (PAD4). PAD4 is of interest due to its role in the formation of neutrophil extracellular traps (NETs) (64), which have been implicated in inflammation and cancer (65). After further experimental validation and characterization, a patent application was filed with the Spanish Patent Office (application number P202430866).

Encouraged by this result, the method was further developed into the current ESSENCE-Dock framework. Subsequent applications include a drug repurposing project targeting discoidin domain receptor 1 (DDR1), a receptor of interest in the context of cancer (66). To this end, DrugBank was initially screened using ESSENCE-Dock to identify potential repurposing candidates. Subsequent experimental $IC_{50}$ determination resulted in the identification of three nanomolar inhibitors for DDR1, with no previously reported activity for this target (67).

ESSENCE-Dock also performed well when applied to the proprietary Eurofins-Villapharma library. Two concrete examples here are projects involving Wee1-like protein kinase (Wee1) and p38$\alpha$ (MAPK14). Wee1 is a key regulator during the cell cycle, making it a relevant target in various cancers (68). The protein p38$\alpha$ is of interest because of its involvement in various cellular processes, including inflammation and cancer (69). After screening the in-house Villapharma library for these two different protein targets, various low μM to high nM inhibitors were identified in a first round of ESSENCE-Dock screening (70,71). In subsequent optimization rounds, these compounds were further improved to very potent, low nM inhibitors with desirable drug-like properties. In short, ESSENCE-Dock has shown itself as an effective method for virtual screening, providing a high enrichment of active compounds across various scenarios.

Another tool developed during this thesis is ConFiLiS. This ligand-based virtual screening method combines individual similarities from multiple fingerprints into a joint, consensus similarity score. As detailed in Article 2 (Section 3.3), ConFiLiS was used to identify dimethoxycurcumin as a potential ABCC3 inhibitor from the in-house Villapharma library (72). Beyond just virtual screening, ConFiLiS has also been effective at assisting with exploring the chemical space around a scaffold. Specifically, after screening the Villapharma library for p38$\alpha$ with ESSENCE-Dock, the top hit was used as a query for ConFiLiS (71). From here,

the top 100 closest compounds were further clustered using Butina clustering (73), which resulted in a diverse but manageable set of compounds to test in a second round. Furthermore, the integrated fingerprint generation utility that comes with the ConFiLiS toolkit can also be used for the efficient construction of datasets of classical fingerprints for the purpose of training ML models.

Finally, the insights and datasets provided in the third article can be used to train ML models to aid lead optimization efforts. This analysis confirms that potency differences between MMPs are more robust to inter-assay variability than absolute affinity values. This makes MMP data highly suitable for developing transformation-based delta models, which are trained on input pairs and the differences in their properties. Using the curated MMP datasets which were part of the analysis, a ML model could be trained to more accurately predict the properties of close analogs.

In summary, each tool presented in this thesis can be used for specific tasks within the broader pipeline of early drug discovery. ConFiLiS enables efficient ligand-based screening, ESSENCE-Dock supports structure-based consensus docking, and MMP models trained with curated datasets can be used during lead optimization to guide compound refinement. ConFiLiS can also be used to explore chemical space around a hit compound and support SAR analysis by retrieving compounds that are similar but still feature some degree of diversity. While each tool functions independently, they can also be combined effectively. For example, ConFiLiS can be used to filter libraries prior to docking with ESSENCE-Dock, or its fingerprints can complement the training of MMP models. Together, these tools offer a flexible and modular framework for hit identification and optimization.

### 4.2.2. Potential Limitations

While the presented methods are effective across a range of applications, each comes with some limitations that should be considered when selecting them for specific use cases.

ConFiLiS is designed as a general-purpose toolkit for ligand-based screening. It supports many different molecular fingerprint types and offers flexibility to users, but currently relies on established, classical fingerprints. Additionally, the

optimal combination of fingerprint descriptors has not been systematically evaluated at this point in time, which may limit performance in certain contexts. Therefore, ConFiLiS should be viewed as a flexible toolkit rather than a fully optimized screening pipeline.

ESSENCE-Dock follows a similar approach, featuring a general scoring function which only requires docking poses and scores as input. This gives flexibility and allows users to integrate results from most docking software. While the current configuration featuring Gnina (50), LeadFinder (31), and DiffDock (52) has demonstrated strong enrichment performance, it is not guaranteed to be optimal: different combinations of docking methods could potentially provide better results. Furthermore, recent advances in docking and structure prediction, such as AlphaFold3 (74) and Boltz (75), could be integrated to further improve accuracy. Additionally, the current ESSENCE-Dock implementation can sometimes produce a large number of high-scoring hits, which may require a secondary filtering step. In such cases, more elaborate methods like MM-GB/PBSA calculations (76) or steered molecular dynamics such as dynamic undocking (DUck) (77) can be employed to refine hits.

For the curated matched molecular pairs datasets, the main limitation lies in the complexity of model development. Although MMP-based data are more usually more accurate and interpretable, they are also more specific, and building general models requires careful treatment of transformation directionality, context sensitivity, and data sparsity. This adds an additional layer of complexity compared to conventional regression or classification approaches, making it more difficult to work with.

### 4.2.3. Future Perspectives

Several directions can be pursued to extend and improve the tools presented in this work.

For ESSENCE-Dock, a promising avenue could be the development of a ML-based scoring function. At this time, only docking scores, binding pose similarity, and ligand flexibility are used in the consensus calculations. However, many docking methods produce additional outputs that are currently not incorporated

when scoring compounds. Examples include DiffDock's confidence scores and Gnina's predicted binding affinity. Integrating these values into a unified, data-driven scoring model could yield more accurate rankings and improve enrichment performance further. Additionally, such a model could incorporate additional features such as protein pocket characteristics or ligand strain, enabling a more nuanced compound assessment.

For ConFiLiS, a clear opportunity lies in exploring and optimizing fingerprint configurations for specific use cases. While the current implementation supports several classical fingerprints, exploring newer representations, such deep learning-derived descriptors (78), could also enhance performance.

An important direction for matched pair analysis is improved integration with DL techniques. Methods that predict property changes, rather than absolute values, align well with the transformation-based nature of MMP data. Future efforts could explore hybrid models that combine the interpretability of transformation rules with the flexibility of neural networks or ensemble methods. Furthermore, model-agnostic explainability tools such as SHAP (79) or LIME (80) could be applied to provide interpretable insights into what drives predicted changes, helping to bridge data-driven modeling with medicinal chemistry intuition. This would make MMP-based models more accessible and actionable in practical settings.

Together, these developments would further increase the flexibility, precision, and practical usability of the tools developed in this thesis. All developed methods and tools that are discussed in this thesis are open-source, allowing integration into broader discovery workflows and facilitating reproducibility and community-driven improvement. Their continued development can hopefully support better informed decision-making in early drug discovery.

# V – CONCLUSIONS

# V - CONCLUSIONS

This thesis revolves around three complementary articles which all focus on improving virtual screening methods in the context of drug discovery. This includes ESSENCE-Dock, a consensus docking framework, ConFiLiS a consensus fingerprint method for ligand-based screening, and a data quality analysis pipeline based on matched molecular pairs analysis. These tools can be used in different stages of early drug discovery, from initial library screening to lead optimization.

ESSENCE-Dock was developed as a structure-based screening method with a flexible, scalable workflow that integrates docking scores, pose similarity, and ligand flexibility across multiple docking algorithms. ConFiLiS is a lightweight but flexible toolkit to perform consensus molecular fingerprint analysis by incorporating multiple fingerprints to identify similar compounds while improving hit diversity. A matched molecular pair analysis of the ChEMBL32 database revealed that relative potency differences between assays are more consistent compared to direct absolute affinity values across assays, resulting in datasets with less noise which can be used to build more robust machine learning models.

Both ESSENCE-Dock and ConFiLiS were applied to in-house and public compound libraries. Within the proprietary Eurofins–Villapharma library, these tools successfully identified promising initial hits, while also aiding with further compound optimization, resulting in various nanomolar hits across various protein targets. Furthermore, similar virtual screening campaigns were performed with external datasets, including DrugBank and large commercial libraries, demonstrating their effectiveness across diverse chemical spaces and practical relevance for real-world drug discovery workflows.

Together, these contributions show how methodologically diverse but complementary computational tools can support more reliable and efficient virtual screening, ultimately aiding decision-making in early-stage drug discovery.

# VI – BIBLIOGRAPHICAL REFERENCES

# VI -BIBLIOGRAPHICAL REFERENCES

1. Schlander M, Hernandez-Villafuerte K, Cheng CY, Mestre-Ferrandiz J, Baumann M. How Much Does It Cost to Research and Develop a New Drug? A Systematic Review and Assessment. Pharmacoeconomics. 2021;39(11):1243–69.

2. Reddy VDK, Banaganapalli B, Rajitha G. Drug Discovery: Concepts and Approaches. In: Shaik NA, Hakeem KR, Banaganapalli B, Elango R, editors. Essentials of Bioinformatics, Volume I: Understanding Bioinformatics: Genes to Proteins [Internet]. Cham: Springer International Publishing; 2019 [cited 2025 Jun 27]. p. 319–34. Available from: https://doi.org/10.1007/978-3-030-02634-9_14

3. Mullard A. Parsing clinical success rates. Nat Rev Drug Discov. 2016 Jul 1;15(7):447–447.

4. Rasul A, Riaz A, Sarfraz I, Khan SG, Hussain G, Zara R, et al. Target Identification Approaches in Drug Discovery. In: Scotti MT, Bellera CL, editors. Drug Target Selection and Validation [Internet]. Cham: Springer International Publishing; 2022 [cited 2025 Jun 30]. p. 41–59. Available from: https://doi.org/10.1007/978-3-030-95895-4_3

5. Keserű GM, Makara GM. Hit discovery and hit-to-lead approaches. Drug Discov Today. 2006 Aug 1;11(15):741–8.

6. da Silva Rocha SFL, Olanda CG, Fokoue HH, Sant'Anna CMR. Virtual Screening Techniques in Drug Discovery: Review and Recent Applications. Curr Top Med Chem. 2019;19(19):1751–67.

7. Temml V, Kutil Z. Structure-based molecular modeling in SAR analysis and lead optimization. Comput Struct Biotechnol J. 2021 Jan 1;19:1431–44.

8. van Rijt A, Stefanek E, Valente K. Preclinical Testing Techniques: Paving the Way for New Oncology Screening Approaches. Cancers. 2023 Jan;15(18):4466.

9. Piantadosi S. Clinical Trials: A Methodologic Perspective. John Wiley & Sons; 2024. 789 p.

10. Sadybekov AV, Katritch V. Computational approaches streamlining drug discovery. Nature. 2023 Apr;616(7958):673–85.

11. Reddy AS, Pati SP, Kumar PP, Pradeep HN, Sastry GN. Virtual screening in drug discovery -- a computational perspective. Curr Protein Pept Sci. 2007 Aug;8(4):329–51.

12. Paul D, Sanap G, Shenoy S, Kalyane D, Kalia K, Tekade RK. Artificial intelligence in drug discovery and development. Drug Discov Today. 2021 Jan;26(1):80–93.

13. York DM. Modern Alchemical Free Energy Methods for Drug Discovery Explained. ACS Phys Chem Au. 2023 Nov 22;3(6):478–91.

14. Rodríguez-Pérez R, Bajorath J. Interpretation of Compound Activity Predictions from Complex Machine Learning Models Using Local Approximations and Shapley Values. J Med Chem. 2020 Aug 27;63(16):8761–77.

15. Oliveira TA de, Silva MP da, Maia EHB, Silva AM da, Taranto AG. Virtual Screening Algorithms in Drug Discovery: A Review Focused on Machine and Deep Learning Methods. Drugs Drug Candidates. 2023 Jun;2(2):311–34.

16. Korn M, Ehrt C, Ruggiu F, Gastreich M, Rarey M. Navigating large chemical spaces in early-phase drug discovery. Curr Opin Struct Biol. 2023 Jun 1;80:102578.

17. López-Pérez K, Avellaneda-Tamayo JF, Chen L, López-López E, Juárez-Mercado KE, Medina-Franco JL, et al. Molecular similarity: Theory, applications, and perspectives. Artif Intell Chem. 2024 Dec 1;2(2):100077.

18. Muegge I, Mukherjee P. An overview of molecular fingerprint similarity search in virtual screening. Expert Opin Drug Discov. 2016 Feb;11(2):137–48.

19. Bajusz D, Rácz A, Héberger K. Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? J Cheminformatics. 2015 May 20;7(1):20.

20. Kumar A, Zhang KYJ. Advances in the Development of Shape Similarity Methods and Their Application in Drug Discovery. Front Chem. 2018 Jul 25;6:315.

21. Koes DR, Camacho CJ. Shape-Based Virtual Screening with Volumetric Aligned Molecular Shapes. J Comput Chem. 2014 Sep 30;35(25):1824–34.

22. ROCS; OpenEye Scientific Software: Santa Fe, NM. [Internet]. Available from: https://www.eyesopen.com/rocs

23. Giordano D, Biancaniello C, Argenio MA, Facchiano A. Drug Design by Pharmacophore and Virtual Screening Approach. Pharmaceuticals. 2022 May 23;15(5):646.

24. Wolber G, Langer T. LigandScout: 3-D Pharmacophores Derived from Protein-Bound Ligands and Their Use as Virtual Screening Filters. J Chem Inf Model. 2005 Jan 1;45(1):160–9.

25. Taminau J, Thijs G, De Winter H. Pharao: pharmacophore alignment and optimization. J Mol Graph Model. 2008 Sep;27(2):161–9.

26. Sunseri J, Koes DR. Pharmit: interactive exploration of chemical space. Nucleic Acids Res. 2016 Jul 8;44(Web Server issue):W442–8.

27. Li Q, Shah S. Structure-Based Virtual Screening. In: Wu CH, Arighi CN, Ross KE, editors. Protein Bioinformatics: From Protein Modifications and Networks to Proteomics [Internet]. New York, NY: Springer; 2017 [cited 2025 Jul 1]. p. 111–24. Available from: https://doi.org/10.1007/978-1-4939-6783-4_5

28. Paggi JM, Pandit A, Dror RO. The Art and Science of Molecular Docking. Annu Rev Biochem. 2024 Aug 2;93(Volume 93, 2024):389–410.

29. Li J, Fu A, Zhang L. An Overview of Scoring Functions Used for Protein–Ligand Interactions in Molecular Docking. Interdiscip Sci Comput Life Sci. 2019 Jun 1;11(2):320–8.

30. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010 Jan 30;31(2):455–61.

31. Stroganov OV, Novikov FN, Stroylov VS, Kulkov V, Chilov GG. Lead Finder: An Approach To Improve Accuracy of Protein−Ligand Docking, Binding Energy Estimation, and Virtual Screening. J Chem Inf Model. 2008 Dec 22;48(12):2371–85.

32. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. J Med Chem. 2004 Mar 1;47(7):1739–49.

33. Lyu J, Wang S, Balius TE, Singh I, Levit A, Moroz YS, et al. Ultra-large library docking for discovering new chemotypes. Nature. 2019 Feb;566(7743):224–9.

34. Lexa KW, Carlson HA. Protein flexibility in docking and surface mapping. Q Rev Biophys. 2012 Aug;45(3):301–43.

35. De Vivo M, Masetti M, Bottegoni G, Cavalli A. Role of Molecular Dynamics and Related Methods in Drug Discovery. J Med Chem. 2016 May 12;59(9):4035–61.

36. Sakano T, Mahamood MdI, Yamashita T, Fujitani H. Molecular dynamics analysis to evaluate docking pose prediction. Biophys Physicobiology. 2016;13:181–94.

37. Genheden S, Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opin Drug Discov. 2015 May 4;10(5):449–61.

38. Chen F, Liu H, Sun H, Pan P, Li Y, Li D, et al. Assessing the performance of the MM/PBSA and MM/GBSA methods. 6. Capability to predict protein–protein binding free energies and re-rank binding poses generated by protein–protein docking. Phys Chem Chem Phys. 2016 Aug 10;18(32):22129–39.

39. Cournia Z, Chipot C, Roux B, York DM, Sherman W. Free Energy Methods in Drug Discovery—Introduction. In: Free Energy Methods in Drug Discovery: Current State and Future Directions [Internet]. American Chemical Society; 2021 [cited 2025 Jul 1]. p. 1–38. (ACS Symposium Series; vol. 1397). Available from: https://doi.org/10.1021/bk-2021-1397.ch001

40. Clark F, Robb GR, Cole DJ, Michel J. Automated Adaptive Absolute Binding Free Energy Calculations. J Chem Theory Comput. 2024 Sep 24;20(18):7806–28.

41. Li Z, Wu C, Li Y, Liu R, Lu K, Wang R, et al. Free energy perturbation–based large-scale virtual screening for effective drug discovery against COVID-19. Int J High Perform Comput Appl. 2023 Jan;37(1):45–57.

42. Blanes-Mira C, Fernández-Aguado P, de Andrés-López J, Fernández-Carvajal A, Ferrer-Montiel A, Fernández-Ballester G. Comprehensive Survey of Consensus Docking for High-Throughput Virtual Screening. Molecules. 2023 Jan;28(1):175.

43. Huang SY, Zou X. Advances and Challenges in Protein-Ligand Docking. Int J Mol Sci. 2010 Aug 18;11(8):3016–34.

44. Thaingtamtanha T, Ravichandran R, Gentile F. On the application of artificial intelligence in virtual screening. Expert Opin Drug Discov. 2025 Jul 3;20(7):845–57.

45. Neves BJ, Braga RC, Melo-Filho CC, Moreira-Filho JT, Muratov EN, Andrade CH. QSAR-Based Virtual Screening: Advances and Applications in Drug Discovery. Front Pharmacol [Internet]. 2018 Nov 13 [cited 2025 Jul 2];9. Available from: https://www.frontiersin.org/journals/pharmacology/articles/10.3389/fphar.2018.01275/full

46. Grisoni F, Ballabio D, Todeschini R, Consonni V. Molecular Descriptors for Structure–Activity Applications: A Hands-On Approach. In: Nicolotti O, editor. Computational Toxicology: Methods and Protocols [Internet]. New York, NY: Springer; 2018 [cited 2025 Jul 2]. p. 3–53. Available from: https://doi.org/10.1007/978-1-4939-7899-1_1

47. Selassie C, Verma RP. History of Quantitative Structure–Activity Relationships. In: Burger's Medicinal Chemistry and Drug Discovery

[Internet]. 1st ed. Wiley; 2010 [cited 2025 Jul 2]. p. 1–96. Available from: https://onlinelibrary.wiley.com/doi/10.1002/0471266949.bmc001.pub2

48. Stokes JM, Yang K, Swanson K, Jin W, Cubillos-Ruiz A, Donghia NM, et al. A Deep Learning Approach to Antibiotic Discovery. Cell. 2020 Feb 20;180(4):688-702.e13.

49. Mahdizadeh SJ, Eriksson LA. iScore: A ML-Based Scoring Function for De Novo Drug Discovery. J Chem Inf Model. 2025 Mar 24;65(6):2759–72.

50. McNutt AT, Francoeur P, Aggarwal R, Masuda T, Meli R, Ragoza M, et al. GNINA 1.0: molecular docking with deep learning. J Cheminformatics. 2021 Jun 9;13(1):43.

51. Ragoza M, Hochuli J, Idrobo E, Sunseri J, Koes DR. Protein-Ligand Scoring with Convolutional Neural Networks. J Chem Inf Model. 2017 Apr 24;57(4):942–57.

52. Corso G, Stärk H, Jing B, Barzilay R, Jaakkola T. DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking [Internet]. arXiv; 2023 [cited 2025 Jun 26]. Available from: http://arxiv.org/abs/2210.01776

53. Igashov I, Stärk H, Vignac C, Schneuing A, Satorras VG, Frossard P, et al. Equivariant 3D-conditional diffusion model for molecular linker design. Nat Mach Intell. 2024 Apr;6(4):417–27.

54. Chen Y, Wang Z, Wang L, Wang J, Li P, Cao D, et al. Deep generative model for drug design from protein target sequence. J Cheminformatics. 2023 Mar 28;15(1):38.

55. Wallach I, Heifets A. Most Ligand-Based Classification Benchmarks Reward Memorization Rather than Generalization. J Chem Inf Model. 2018 May 29;58(5):916–32.

56. Yang J, Shen C, Huang N. Predicting or Pretending: Artificial Intelligence for Protein-Ligand Interactions Lack of Sufficiently Large and Unbiased Datasets. Front Pharmacol [Internet]. 2020 Feb 25 [cited 2025 Jul 2];11. Available from: https://www.frontiersin.org/journals/pharmacology/articles/10.3389/fphar.2020.00069/full

57. Li Z, Li H, Yu K, Luo HB. Perspective of drug design with high-performance computing. Natl Sci Rev. 2021 Dec 1;8(12):nwab105.

58. Yoo AB, Jette MA, Grondona M. SLURM: Simple Linux Utility for Resource Management. In: Feitelson D, Rudolph L, Schwiegelshohn U, editors. Job Scheduling Strategies for Parallel Processing. Berlin, Heidelberg: Springer; 2003. p. 44–60.

59. Kurtzer GM, Sochat V, Bauer MW. Singularity: Scientific containers for mobility of compute. PLOS ONE. 2017 May 11;12(5):e0177459.

60. Herlihy M, Shavit N. The Art of Multiprocessor Programming, Revised Reprint. Elsevier; 2012. 537 p.

61. bio-hpc. bio-hpc/metascreener [Internet]. 2025 [cited 2025 Jul 1]. Available from: https://github.com/bio-hpc/metascreener

62. Kukol A. Consensus virtual screening approaches to predict protein ligands. Eur J Med Chem. 2011 Sep 1;46(9):4661–4.

63. Knox C, Wilson M, Klinger CM, Franklin M, Oler E, Wilson A, et al. DrugBank 6.0: the DrugBank Knowledgebase for 2024. Nucleic Acids Res. 2024 Jan 5;52(D1):D1265–75.

64. Lewis HD, Liddle J, Coote JE, Atkinson SJ, Barker MD, Bax BD, et al. Inhibition of PAD4 activity is sufficient to disrupt mouse and human NET formation. Nat Chem Biol. 2015 Mar;11(3):189–91.

65. Wang Y, Yang K, Li J, Wang C, Li P, Du L. Neutrophil extracellular traps in cancer: From mechanisms to treatments. Clin Transl Med. 2025 Jun 13;15(6):e70368.

66. Tian Y, Bai F, Zhang D. New target DDR1: A "double-edged sword" in solid tumors. Biochim Biophys Acta Rev Cancer. 2023 Jan;1878(1):188829.

67. Nelen J, Song X, Liu L, Zhao L, Tu Z, Wang Z, et al. Enhancing Drug Repurposing with Consensus Docking: Discovery of Novel DDR1 Inhibitors [Internet]. Rochester, NY: Social Science Research Network; 2025 [cited 2025 Jun 25]. Available from: https://papers.ssrn.com/abstract=5129660

68. Matheson CJ, Backos DS, Reigan P. Targeting WEE1 Kinase in Cancer. Trends Pharmacol Sci. 2016 Oct 1;37(10):872–81.

69. Gupta J, Nebreda AR. Roles of p38α mitogen-activated protein kinase in mouse models of inflammatory diseases and cancer. FEBS J. 2015;282(10):1841–57.

70. Nelen J, Sarkar B, Jimenez-Bernad J, Darragh AC, Hanna AM, Servant NB, et al. Discovery and biological evaluation of WEE1 inhibitors through virtual screening and medicinal chemistry approaches for cancer treatment [Internet]. ChemRxiv; 2025 [cited 2025 Jun 27]. Available from: https://chemrxiv.org/engage/chemrxiv/article-details/685aad773ba0887c33527b15

71. Nelen J, Goettert M, Forster M, Breuils L, Villalgordo-Soto JM, Laufer SA, et al. Computational Discovery and Optimization of a Potent, Selective, and Drug-like Scaffold for p38α Inhibition [Internet]. ChemRxiv; 2025 [cited 2025 Jun 25].

Available from: https://chemrxiv.org/engage/chemrxiv/article-details/6857d9ab3ba0887c33f1e41c

72. Nelen J, Naponelli V, Villalgordo-Soto JM, Falasca M, Pérez-Sánchez H. Targeting Drug Resistance in Cancer: Dimethoxycurcumin as a Functional Antioxidant Targeting ABCC3. Antioxidants. 2025 May;14(5):599.

73. Butina D. Unsupervised Data Base Clustering Based on Daylight's Fingerprint and Tanimoto Similarity: A Fast and Automated Way To Cluster Small and Large Data Sets. J Chem Inf Comput Sci. 1999 Jul 26;39(4):747–50.

74. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature. 2024 Jun;630(8016):493–500.

75. Passaro S, Corso G, Wohlwend J, Reveiz M, Thaler S, Somnath VR, et al. Boltz-2: Towards Accurate and Efficient Binding Affinity Prediction [Internet]. bioRxiv; 2025 [cited 2025 Jun 26]. p. 2025.06.14.659707. Available from: https://www.biorxiv.org/content/10.1101/2025.06.14.659707v1

76. Yang M, Bo Z, Xu T, Xu B, Wang D, Zheng H. Uni-GBSA: an open-source and web-based automatic workflow to perform MM/GB(PB)SA calculations for virtual screening. Brief Bioinform. 2023 Jul 1;24(4):bbad218.

77. Ruiz-Carmona S, Schmidtke P, Luque FJ, Baker L, Matassova N, Davis B, et al. Dynamic undocking and the quasi-bound state as tools for drug discovery. Nat Chem. 2017 Mar;9(3):201–6.

78. Lžičař M, Gamouh H. CHEESE: 3D Shape and Electrostatic Virtual Screening in a Vector Space [Internet]. ChemRxiv; 2024 [cited 2025 Jun 26]. Available from: https://chemrxiv.org/engage/chemrxiv/article-details/67250915f9980725cfcd1f6f

79. Lundberg SM, Lee SI. A Unified Approach to Interpreting Model Predictions. In: Advances in Neural Information Processing Systems [Internet]. Curran Associates, Inc.; 2017 [cited 2025 Jun 26]. Available from: https://papers.nips.cc/paper_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html

80. Ribeiro MT, Singh S, Guestrin C. "Why Should I Trust You?": Explaining the Predictions of Any Classifier [Internet]. arXiv; 2016 [cited 2025 Jun 26]. Available from: http://arxiv.org/abs/1602.04938

# VII – APPENDICES

# VII - APPENDICES

## 7.1. APPENDIX 1. QUALITY OF INCLUDED THE PUBLICATIONS

All publications included in this thesis have been published in high-quality, peer-reviewed journals that are well regarded within their respective scientific fields. This is supported by strong journal metrics, with impact factors ranging from 5.3 to 6.6. Each journal is ranked in the first quartile (Q1) of one or multiple categories.

### 7.1.1. ESSENCE-Dock: A Consensus-Based Approach to Enhance Virtual Screening Enrichment in Drug Discovery

| ISSN | EISSN |
|---|---|
| 1549-9596 | 1549-960X |
| **JCR ABBREVIATION** | **ISO ABBREVIATION** |
| J CHEM INF MODEL | J. Chem Inf. Model. |

**Journal Information**

| EDITION | CATEGORY |
|---|---|
| Science Citation Index Expanded (SCIE) | COMPUTER SCIENCE, INTERDISCIPLINARY APPLICATIONS<br>CHEMISTRY, MULTIDISCIPLINARY<br>COMPUTER SCIENCE, INFORMATION SYSTEMS<br>CHEMISTRY, MEDICINAL |

| LANGUAGES | REGION | 1ST ELECTRONIC JCR YEAR |
|---|---|---|
| English | USA | 2005 |

**Publisher Information**

| PUBLISHER | ADDRESS | PUBLICATION FREQUENCY |
|---|---|---|
| AMER CHEMICAL SOC | 1155 16TH ST, NW, WASHINGTON, DC 20036 | 12 issues/year |

*Figure VII.1. Key publication details of Journal of Chemical Information and Modeling.*

CATEGORY
CHEMISTRY, MEDICINAL
15/72

| JCR YEAR | JIF RANK | QUART ILE | JIF PERCENTILE |
|---|---|---|---|
| 2024 | 15/72 | Q1 | 79.9 |
| 2023 | 10/72 | Q1 | 86.8 |

CATEGORY
CHEMISTRY, MULTIDISCIPLINARY
65/239

| JCR YEAR | JIF RANK | QUART ILE | JIF PERCENTILE |
|---|---|---|---|
| 2024 | 65/239 | Q2 | 73.0 |
| 2023 | 61/231 | Q2 | 73.8 |

CATEGORY
COMPUTER SCIENCE, INFORMATION SYSTEMS
46/258

| JCR YEAR | JIF RANK | QUART ILE | JIF PERCENTILE |
|---|---|---|---|
| 2024 | 46/258 | Q1 | 82.4 |
| 2023 | 34/250 | Q1 | 86.6 |

CATEGORY
COMPUTER SCIENCE, INTERDISCIPLINARY APPLICATIONS
37/175

| JCR YEAR | JIF RANK | QUART ILE | JIF PERCENTILE |
|---|---|---|---|
| 2024 | 37/175 | Q1 | 79.1 |
| 2023 | 28/170 | Q1 | 83.8 |

*Figure VII.2. Journal Impact Factor (JIF) rankings and percentiles for Journal of Chemical Information and Modeling across its indexed subject categories.*

### 7.1.2. Targeting Drug Resistance in Cancer: Dimethoxycurcumin as a Functional Antioxidant Targeting ABCC3

| ISSN | EISSN |
|---|---|
| N/A | 2076-3921 |
| **JCR ABBREVIATION** | **ISO ABBREVIATION** |
| ANTIOXIDANTS-BASEL | Antioxidants |

**Journal Information**

| EDITION | CATEGORY |
|---|---|
| Science Citation Index Expanded (SCIE) | FOOD SCIENCE & TECHNOLOGY<br>CHEMISTRY, MEDICINAL<br>BIOCHEMISTRY & MOLECULAR BIOLOGY |

| LANGUAGES | REGION | 1ST ELECTRONIC JCR YEAR |
|---|---|---|
| English | SWITZERLAND | 2018 |

**Publisher Information**

| PUBLISHER | ADDRESS | PUBLICATION FREQUENCY |
|---|---|---|
| MDPI | ST ALBAN-ANLAGE 66, CH-4052 BASEL, SWITZERLAND | 12 issues/year |

*Figure VII.3. Key publication details of Antioxidants.*

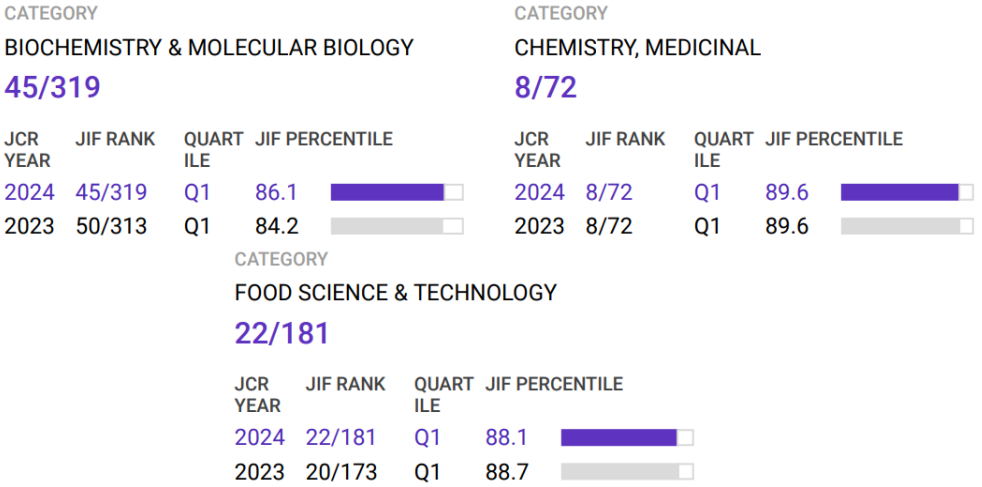



*Figure VII.4. Journal Impact Factor (JIF) rankings and percentiles for Antioxidants across its indexed subject categories.*

### 7.1.3. Matched pairs demonstrate robustness against inter-assay variability

CATEGORY

CHEMISTRY, MULTIDISCIPLINARY

58/239

| JCR YEAR | JIF RANK | QUARTILE | JIF PERCENTILE |
|---|---|---|---|
| 2024 | 58/239 | Q1 | 75.9 |
| 2023 | 43/231 | Q1 | 81.6 |

CATEGORY

COMPUTER SCIENCE, INFORMATION SYSTEMS

35/258

| JCR YEAR | JIF RANK | QUARTILE | JIF PERCENTILE |
|---|---|---|---|
| 2024 | 35/258 | Q1 | 86.6 |
| 2023 | 19/250 | Q1 | 92.6 |

CATEGORY

COMPUTER SCIENCE, INTERDISCIPLINARY APPLICATIONS

32/175

| JCR YEAR | JIF RANK | QUARTILE | JIF PERCENTILE |
|---|---|---|---|
| 2024 | 32/175 | Q1 | 82.0 |
| 2023 | 18/170 | Q1 | 89.7 |

*Figure VII.5 Key publication details of Journal of Cheminformatics.*

ISSN

1758-2946

EISSN

1758-2946

JCR ABBREVIATION

J CHEMINFORMATICS

ISO ABBREVIATION

J. Cheminformatics

## Journal Information

EDITION

Science Citation Index Expanded (SCIE)

CATEGORY

COMPUTER SCIENCE, INTERDISCIPLINARY APPLICATIONS
CHEMISTRY, MULTIDISCIPLINARY
COMPUTER SCIENCE, INFORMATION SYSTEMS

LANGUAGES

English

REGION

ENGLAND

1ST ELECTRONIC JCR YEAR

2011

## Publisher Information

PUBLISHER

BMC

ADDRESS

CAMPUS, 4 CRINAN ST, LONDON N1 9XW, ENGLAND

PUBLICATION FREQUENCY

1 issues/year

*Figure VII.6. Journal Impact Factor (JIF) rankings and percentiles for Journal of Cheminformatics across its indexed subject categories.*

## 7.2. APPENDIX 2. RESEARCH OUTPUT AND CONTRIBUTIONS

### 7.2.1. Publications

**Targeting Drug Resistance in Cancer: Dimethoxycurcumin as a Functional Antioxidant Targeting ABCC3**
Jochem Nelen, Valeria Naponelli, José Manuel Villalgordo Soto, Marco Falasca, Horacio Pérez Sánchez
*Antioxidants*, 16 May 2025
DOI: 10.3390/antiox14050599
Impact factor (JCR 2024): 6.6
Journal ranking: 8 out of 72 in Chemistry, Medicinal (Q1)

**Matched Pairs Demonstrate Robustness Against Inter-Assay Variability**
Jochem Nelen, Horacio Pérez Sánchez, Hans De Winter, Dries Van Rompaey
*Journal of Cheminformatics*, 20 January 2025
DOI: 10.1186/s13321-025-00956-y
Impact factor (JCR 2024): 5.7
Journal ranking: 35 out of 258 in Computer Science, Information Systems (Q1)

**Interaction of an Anticancer Oxygenated Propenylbenzene Derivative with Human Topoisomerase II $\alpha$ and Actin: Molecular Modeling and Isothermal Titration Calorimetry Studies**
Joanna Grzelczyk, Horacio Pérez Sánchez, Jochem Nelen, Miguel Carmena Bargueño, Ilona Gałązka Czarnecka, Grażyna Budryn, Dawid Hernik, Elisabetta Brenna, Filip Boratyński
*Journal of Thermal Analysis and Calorimetry*, 2 September 2024
DOI: 10.1007/s10973-024-13569-8
Impact factor (JCR 2024): 3.1
Journal ranking: 24 out of 79 in Thermodynamics (Q2)

**Enhancing MD Simulations: ASGARD's Automated Analysis for GROMACS**

Alejandro Rodríguez Martínez, Jochem Nelen, Miguel Carmena Bargueño, Carlos Martínez Cortés, Irene Luque, Horacio Pérez Sánchez

*Journal of Biomolecular Structure and Dynamics*, 22 March 2024

DOI: 10.1080/07391102.2024.2349527

Impact factor (JCR 2024): 2.4

Journal ranking: 47 out of 79 in Biophysics (Q3)

**ESSENCE-Dock: A Consensus-Based Approach to Enhance Virtual Screening Enrichment in Drug Discovery**

Jochem Nelen, Miguel Carmena Bargueño, Carlos Martínez Cortés, Alejandro Rodríguez Martínez, José Manuel Villalgordo Soto, Horacio Pérez Sánchez

*Journal of Chemical Information and Modeling*, 28 February 2024

DOI: 10.1021/acs.jcim.3c01982

Impact factor (JCR 2024): 5.3

Journal ranking: 46 out of 258 in Computer Science, Information Systems (Q1)

**Antcin-B, a Phytosterol-Like Compound from Taiwanofungus camphoratus, Inhibits SARS-CoV-2 3-Chymotrypsin-Like Protease (3CLPro) Activity In Silico and In Vitro**

Gyaltsen Dakpa, Senthil Kumar, Jochem Nelen, Horacio Pérez Sánchez, Sheng-Yang Wang

*Scientific Reports*, 10 October 2023

DOI: 10.1038/s41598-023-44476-x

Impact factor (JCR 2024): 3.8

Journal ranking: 25 out of 134 in Multidisciplinary Sciences (Q1)

### 7.2.2. Preprints

#### *7.2.2.1. First Author Preprints*

**Discovery and biological evaluation of WEE1 inhibitors through virtual screening and medicinal chemistry approaches for cancer treatment**

Jochem Nelen, Bidisha Sarkar, Javier Jimenez-Bernad, Antonia C. Darragh, Andrew M. Hanna, Nicole B. Servant, Mirza Jahic, José Manuel Villalgordo-Soto, Gunda I. Georg, Horacio Pérez-Sánchez

*ChemRxiv*, 27 June 2025

DOI: 10.26434/chemrxiv-2025-wccvn

**Computational Discovery and Optimization of a Potent, Selective, and Drug-like Scaffold for p38$\alpha$ Inhibition**

Jochem Nelen, Marcia Goettert, Michael Forster, Laure Breuils, José Manuel Villalgordo-Soto, Stefan A. Laufer, Horacio Pérez-Sánchez

*ChemRxiv*, 25 June 2025

DOI: 10.26434/chemrxiv-2025-l8jlq

**Enhancing Drug Repurposing with Consensus Docking: Discovery of Novel DDR1 Inhibitors**

Jochem Nelen, Xiaojuan Song, Lianchao Liu, Lijie Zhao, Zhengchao Tu, Zhen Wang, Ke Ding, Horacio Pérez Sánchez

*SSRN Preprint Server*, 10 February 2025

DOI: 10.2139/ssrn.5129660

#### *7.2.2.2. Co-author Preprints*

**STELLAR Enables CPU-Based Fragment Docking and Reconstruction of Large Peptides Without AI Assistance**

Alejandro Rodríguez Martínez, Jochem Nelen, Miguel Carmena Bargueño, Carlos Martínez Cortés, Horacio Pérez Sánchez

*ChemRxiv*, 3 June 2025

DOI: 10.26434/chemrxiv-2025-nmn5s

**Freshness flavour perception via quantum-inspired and quantum ligand-based virtual screening**

Antón Rodríguez Otero, Mariamo Mussa Juane, María Paredes Ramos, Jochem Nelen, Alejandro Borrallo Rentero, Horacio Pérez Sánchez, Andrés Gómez Tato, Jose M López Vilariño

*Research Square*, 19 February 2025

DOI: 10.21203/rs.3.rs-6062772/v1

### 7.2.3. Patent

**Tratamiento de enfermedades causadas por netosis, o de enfermedades asociadas a trampas extracelulares de neutrófilos (NET)**

Inventors: Constantino Martínez Gómez, Laura Zapata Martínez, Irene Martínez Martínez, Sonia Águila Martínez, Rocío González Conejero, Horacio Pérez Sánchez, Jochem Nelen

Application No.: P202430866

Country: Spain

Date of Registration: 24/10/2024

### 7.2.4. Conference Presentations

#### *7.2.4.1.* ***Oral Communications***

(Presenter in all cases)

**ESSENCE-Dock: A Robust and Reliable Consensus Docking Method for Identifying Active Compounds**

X Jornadas de Investigación y Doctorado, Murcia, 21/06/2024

Organized by Universidad Católica San Antonio de Murcia

**Visualizing Molecular Impact: Understanding the Role of Substructures in Drug Activity Prediction**

II Congreso Estatal de Estudiantes de Biociencias (CEEBI), Granada, 18–21/07/2023

Organized by Universidad de Granada

**Unpacking the Black Box: Understanding Machine Learning Models to Aid the Drug Discovery Process**
IX Jornadas de Investigación y Doctorado, Murcia, 23/06/2023
Organized by Universidad Católica San Antonio de Murcia

**Visualizing the Impact of Chemical Substructures on Compound Activity for Improving the Drug Discovery Process**
III Symposium on Chemical and Physical Sciences for Young Researchers, Murcia, 15–16/06/2023
Organized by Faculty of Chemistry, University of Murcia
**Award: Best Flash Presentation – Academia de Ciencias de la Región de Murcia**

**Optimizing Consensus Docking by Incorporating Predicted Binding Pose Similarity: A Wee1 Case Study**
VII Jornadas Doctorales, Murcia, 04–06/07/2022
Organized by Universidad de Murcia

**Optimizing Consensus Docking by Incorporating Predicted Binding Pose Similarity: A Wee1 Case Study**
VIII Jornadas de Investigación y Doctorado, Murcia, 24/06/2022
Organized by Universidad Católica San Antonio de Murcia

**Optimizing Consensus Docking by Incorporating Predicted Binding Pose Similarity: A Wee1 Case Study**
Simposio de Estudiantes Hispanohablantes de Bioinformática y Biología Computacional, 02–03/06/2022
Organized by SEH2Bioinfo

#### *7.2.4.2. **Poster Presentations***

(Presenter unless otherwise noted)

**Gilteritinib inhibits PAD4 activity and neutrophil extracellular trap Formation** (Co-author)
ISTH 2025 Congress, Washington, D. C., 21-25/06/2025
Organized by International Society on Thrombosis and Haemostasis

**Leveraging Consensus: A Novel Approach to Molecular Docking for Improved Enrichment of Active Compounds**
1er Congreso SEBiBC, Valencia, 16–18/10/2024
Organized by Sociedad Española de Bioinformática y Biología Computacional

**ESSENCE-DOCK: A Unified Framework for Integrating Docking Methods and Enhancing Hit Identification**
24th European QSAR Symposium, Barcelona, 22–26/09/2024
Organized by QSAR, Chemoinformatics and Modeling Society

**Utilizing Binding Pose Similarity after Consensus Docking to Aid in the Discovery of Novel Wee1 Inhibitors**
II Congreso Internacional de Investigación Biosanitaria para Jóvenes Investigadores, Murcia, 15–16/09/2022
Organized by Universidad de Murcia

**Using Consensus Molecular Docking for the Discovery of Wee1 Inhibitors in the Context of Cancer**
II Symposium on Chemical and Physical Sciences for Young Researchers, Murcia, 24–25/03/2022
Organized by Universidad de Murcia

### 7.2.5. Participation in Funded Projects

**2025 Spanish Supercomputing Network – 1st Call**
*Project title: Discovery of Novel Compounds from Large Chemical Spaces Targeting Fascin for Inhibiting Colorectal Cancer*
*Project ID:* BCV-2025-1-0004
*Resource awarded:* 5,000,000 CPU hours (supercomputing time)

**2025 Fundación Séneca – "Desarrollo De Nuevos Inhibidores de Fascina Como Agentes Anticancerígenos" Call**
*Programme:* Development of scientific and technical research by competitive research groups
*Project ID:* 22616/PI/24
*Funding awarded:* €72,800

**2024 Spanish Supercomputing Network – 2nd Call**

*Project title: Discovery of Novel Compounds from Large Chemical Spaces Targeting Fascin for Inhibiting Colorectal Cancer*
*Project ID:* BCV-2024-2-0013
*Resource awarded:* 5,000,000 CPU hours (supercomputing time)

### 7.2.6. Science Outreach

**Accelerating Drug Development with Computational Methods: A Practical Overview**
Hi-Tech Seminar, 14/03/2025
Organized by Universidad Católica San Antonio de Murcia, Murcia, Spain

**Descodificando la caja negra: Mejorando y guiando el descubrimiento de fármacos usando inteligencia artificial e interpretabilidad de las subestructuras moleculares**
*Hidden Nature*, Revista 19, 05/01/2024
Type of production: Popular science article

**The Discovery of Wee1 Inhibitors in the Context of Cancer: an *in Silico* Approach**
Guest Lecture, 25/05/2023
Organized by Dokuz Eylül Üniversitesi, Turkey

### 7.2.7. Open-source GitHub Contributions

As part of my commitment to open science and reproducibility, I have developed and contributed to several open-source software projects. These tools and libraries are freely available to the scientific community, aiming to accelerate research in computational drug discovery, molecular modeling, and cheminformatics by promoting transparency, collaboration, and innovation.

#### *7.2.7.1. Maintained Repositories*

For the following projects, I am either the original creator or the primary maintainer of key components and functionalities. My responsibilities include development, ongoing maintenance, and ensuring the tools remain accessible and useful to the scientific community.

**MetaScreener**
Implemented Gnina integration and ESSENCE-Dock consensus scoring as part of a broader virtual screening framework.
https://github.com/bio-hpc/metascreener

**ConFiLiS**
Consensus-based fingerprint similarity library for ligand screening.
https://github.com/Jnelen/ConFiLiS

**ChEMBL_MatchedPairs Notebook**
Interactive Jupyter notebook for matched molecular pair analysis using ChEMBL data. Includes datasets and reproducible workflows.
https://github.com/Jnelen/ChEMBL_MatchedPairsAnalysis

**DynamicBindHPC**, **TorsionalDiffusionHPC**, **DiffDockHPC**
Adaptations of popular molecular docking/AI methods for high-performance computing environments.
https://github.com/Jnelen/DynamicBindHPC
https://github.com/Jnelen/DiffDockHPC
https://github.com/Jnelen/TorsionalDiffusionHPC

#### *7.2.7.2. Code Contributions to External Projects*

These are open-source projects developed by external groups. While I am not the original author of these projects, I have actively used these tools in my research and contributed enhancements such as new features, performance improvements, and bug fixes to support their continued development.

**OpenDUck**
Enabled batch mode for the OpenMM backend and introduced a launcher script.
https://github.com/CBDD/openduck

**AEV-PLIG**
Improved multicore processing and runtime efficiency.
https://github.com/isakvals/AEV-PLIG

**Spyrmsd**
Enabled backend selection and added multicore script functionality.
https://github.com/RMeli/spyrmsd

**Spectrophore**
Refactored CLI utility, added continuous integration, and improved usability.
https://github.com/silicos-it/spectrophore

**PyFingerprint**
Extended support for additional molecular fingerprints and resolved existing bugs.
https://github.com/hcji/PyFingerprint

**Uni-GBSA**
Bug fixes and enhancements for usability and documentation.
https://github.com/dptech-corp/Uni-GBSA