



UNIVERSITY OF MESSINA
Department of Chemical, Biological, Pharmaceutical and
Environmental Sciences
Doctorate School in Chemical Sciences (XXXV cycle)
S.S.D. CHIM/08

**Design, synthesis, biological evaluation, and molecular
modeling of inhibitors of enzymes involved in tumor
and viral pathologies**

PhD student:
Giulia Culetta

Supervisor:
Prof. Maria Zappalà

Co-Supervisor:
Prof. Marco Tutone

PhD Coordinator:
Prof. Concetta De Stefano

2021/22 ACADEMIC YEAR

INDEX

1. AIM OF THE WORK	4
2. MOLECULAR MODELING IN DRUG DISCOVERY	6
2.1. INTRODUCTION	6
2.2 LIGAND-BASED DRUG DESIGN (LBDD)	6
2.2.1 QSAR AND PHARMACOPHORE MODELS.....	7
2.3 STRUCTURE-BASED DRUG DESIGN (SBDD)	8
2.3.1 STRUCTURE-BASED PHARMACOPHORE MODELING	8
2.3.2 MOLECULAR DOCKING	8
2.3.3 MOLECULAR DYNAMICS	9
3. DESIGN, SYNTHESIS AND BIOLOGICAL EVALUATION OF NOVEL NON-COVALENT PROTEASOME AND IMMUNOPROTEASOME INHIBITORS	11
3.1 STRUCTURE AND FUNCTIONS OF THE CONSTITUTIVE PROTEASOME	11
3.2 STRUCTURE AND FUNCTIONS OF THE IMMUNOPROTEASOME (IP)	12
3.2.1 INFLAMMATION, VIRAL INFECTIONS, AND AUTOIMMUNE DISEASES.....	13
3.2.2 HEMATOLOGICAL NEOPLASMS	14
3.2.3 NEURODEGENERATIVE DISEASES AND AGING	14
3.3 INHIBITORS OF CONSTITUTIVE PROTEASOME AND IMMUNOPROTEASOME	14
3.3.1 NON-SELECTIVE INHIBITORS	14
3.3.2 SELECTIVE INHIBITORS	18
3.3.3 NON-COVALENT PEPTIDE INHIBITORS	20
3.4 IMMUNOPROTEASOME AND NON-COVALENT INHIBITION: EXPLORATION BY ADVANCED MOLECULAR DYNAMICS AND DOCKING METHODS	21
3.4.2 MATERIALS AND METHODS	22
3.4.2.1 SYSTEM AND LIGANDS PREPARATION	22
3.4.2.2 MD-BINDING SIMULATIONS	22
3.4.2.3 PLAIN MD SIMULATIONS	23
3.4.2.4 BINDING POSE META DYNAMICS (BPMD).....	23
3.4.2.5 INDUCED-FIT DOCKING	24
3.4.2.6 MM-GBSA-BINDING FREE ENERGY CALCULATIONS	24
3.4.3 RESULTS AND DISCUSSION	24
3.4.3.1 MD-BINDING (MDB) ANALYSIS	24

3.4.3.2	INDUCED FIT DOCKING (IFD)	28
3.4.3.3	BINDING POSE METADYNAMICS ANALYSIS	31
3.4.3.4	MM-GBSA-BINDING FREE ENERGY CALCULATIONS	34
3.4.4	CONCLUSION	34

3.5 COMBINED COMPUTATIONAL APPROACHES FOR VIRTUAL SCREENING OF NON-COVALENT β 1I AND β 5I INHIBITORS OF THE IMMUNOPROTEASOME.....35

3.5.1	INTRODUCTION	35
3.5.2	MATERIALS AND METHODS	35
3.5.2.1	PROTEIN PREPARATION, LIGANDS PREPARATION	35
3.5.2.2	PHARMACOPHORE MODELLINGS GENERATION	35
3.5.2.3	MOLECULAR DOCKING	36
3.5.2.4	INDUCED FIT DOCKING (IFD)	36
3.5.2.5	<i>IN VITRO</i> 20S IMMUNOPROTEASOME INHIBITION ASSAYS	36
3.5.3	RESULTS AND DISCUSSION	37
3.5.3.1	β 1I PHARMACOPHORE MODELING	37
3.5.3.2	β 5I DOCKING AND PHARMACOPHORE MODELING	40
3.5.3.3	β 1I AND β 5I VIRTUAL SCREENING	43
3.5.4	CONCLUSION	50

3.6 SYNTHESIS OF NOVEL PROTEASOME INHIBITORS.....51

3.6.1	INTRODUCTION	51
3.6.2	MATERIALS AND METHODS	51
3.6.2.1	CHEMISTRY	51
3.6.2.2	BIOLOGICAL ACTIVITY	56
3.6.3	RESULTS AND DISCUSSION	58
3.6.3.1	SYNTHESIS OF AMIDES 1A-H	58
3.6.3.2	BIOLOGICAL ACTIVITY	58

4. *IN SILICO* DESIGN, SYNTHESIS, AND BIOLOGICAL EVALUATION OF ANTICANCER ARYLSULFONAMIDE ENDOWED WITH ANTI-TELOMERASE ACTIVITY60

4.1.	INTRODUCTION	60
4.2	MATERIALS AND METHODS	60
4.2.1	LIBRARY AND PROTEIN PREPARATION	61
4.2.2	DOCKING STUDIES	61
4.2.3	MYSHAPE APPROACH	61
4.2.4	MM-GBSA FREE ENERGY CALCULATIONS AND MOLECULAR DYNAMICS SIMULATION	61
4.2.5	ADME PREDICTION	62
4.2.6	CHEMISTRY	63
4.2.7	CELL CULTURING AND MTT ASSAY	65
4.3	RESULTS AND DISCUSSION	66
4.3.1	SYNTHESIS	69
4.3.2	CYTOTOXIC ACTIVITY AND SELECTIVITY INDEX	71
4.3.3	DOCKING AND MD SIMULATION OF COMPOUND 2C	72
4.3.4	ADME CALCULATION FOR COMPOUND 2C	75
4.4	CONCLUSION	76

5. COMPARING MOLECULAR DYNAMICS DERIVED PHARMACOPHORE MODELS WITH DOCKING: A STUDY ON CDK-2 INHIBITORS.....77

5.1	INTRODUCTION	77
5.2	MATERIALS AND METHODS	77
5.2.1	PREPARATION OF MOLECULAR STRUCTURES	77
5.2.2	CONVERSION OF MD TRAJECTORIES	77
5.2.3	COMMON HITS APPROACH (CHA) PHARMACOPHORE MODELS.....	78
5.2.4	MOLECULAR DYNAMICS SHARED PHARMACOPHORE (MYSHAPE) MODELS.....	78
5.2.5	DOCKING	78
5.2.6	VIRTUAL SCREENING.....	78
5.3	RESULTS AND DISCUSSION.....	79
5.4	CONCLUSION.....	83

6. EVALUATION OF THE IKKB BINDING OF INDICAXANTHIN BY INDUCED FIT DOCKING, BINDING POSE METADYNAMICS, AND MOLECULAR DYNAMICS.....85

6.1	INTRODUCTION	85
6.2	MATERIALS AND METHODS	86
6.2.1	PROTEIN AND LIGAND PREPARATION	86
6.2.3	BINDING POSE METADYNAMICS.....	86
6.2.4	MD SIMULATIONS	87
6.2.5	MM-GBSA BINDING FREE ENERGY CALCULATION	87
6.3	RESULTS AND DISCUSSION	88
6.3.1	BINDING POSE METADYNAMICS.....	91
6.3.2	MOLECULAR DYNAMICS SIMULATIONS.....	92
6.3.4	MM-GBSA AND BINDING FREE ENERGY ANALYSIS.....	95
6.4	CONCLUSION.....	96

7. EXPLORATION THE SARS-COV-2 PROTEOME IN THE SEARCH OF POTENTIAL INHIBITORS VIA STRUCTURE-BASED PHARMACOPHORE MODELING/DOCKING APPROACH.....97

7.1	INTRODUCTION	97
7.2	MATERIALS AND METHODS	98
7.2.1	LIBRARY PREPARATION	98
7.2.2	HOMOLOGY MODELING AND PROTEIN PREPARATION.....	99
7.2.3	PHARMACOPHORE MODELING.....	99
7.2.4	DOCKING	100
7.2.3	INDUCED-FIT DOCKING AND MM-GBSA.....	100
7.3	RESULTS AND DISCUSSION	101
7.4	CONCLUSION.....	111

8. REFERENCES.....112

1. AIM OF THE WORK

The aim of my Ph.D. project was the design, synthesis, biological evaluation, and molecular modeling of enzyme inhibitors involved in tumor and viral pathologies.

Specifically, this thesis manuscript is divided into three main parts, presenting some of the papers published during my doctoral work. These studies were carried out in cooperation with the University of Palermo under the supervision of Prof. Marco Tutone - Department of Biological, Chemical and Pharmaceutical Sciences and Technologies (STEBICEF).

The first part concerns the study of proteases. Proteases represent one of the most relevant groups of enzymes which catalyze the hydrolysis of peptide bonds. The research group with whom I worked during my Ph.D. has actively been involved in the development of novel proteasome/immunoproteasome inhibitors. In this scenario, during my Ph.D., molecular modeling studies such as MD Binding (MDB), Binding Pose MetaDynamics (BPMD), and Induced Fit Docking (IFD) were carried out on the previously identified non-covalent compound 1 that was shown to inhibit the $\beta 1i$ subunit of the immunoproteasome. The outcomes provided a dynamic point of view for the definition of the pharmacophore features. This dynamic pharmacophore modeling approach was used for the scaffold-hopping of new non-covalent inhibitors.

Additionally, these molecular modeling studies were used to find new inhibitors of the $\beta 5i$ subunit. Thirty-six molecules from three commercial databases were selected to carry out the enzymatic assay for $\beta 1i/\beta 5i$ subunits. Most of the molecules showed activity in the micromolar range, among them, RIM312 inhibits both subunits. The virtual screening campaign was carried out in collaboration with the Fondazione Ri. Med - Molecular Informatic group- under the supervision of Dr. Ugo Perricone.

Moreover, novel proteasome inhibitors were synthesized. The conjugation of these compounds with nano systems based on graphene quantum dots (GQDs) will improve their dispersibility in water and cellular uptake.

The second part concerns molecular modeling studies applied to different anticancer targets.

The first study concerns the design, synthesis, and biological evaluation of arylsulfonamides as telomerase inhibitors. This study was carried out in collaboration with the research group of prof. Tutone. A structure-based approach was carried out to design potential inhibitors of the telomerase active site. The MYSHAPE (Molecular dYnamics SHared PharmacophorE) approach and docking were used to screen an *in-house* library of 126 arylsulfonamide derivatives. Promising hit compounds were synthesized using classical and green methods. Compound 2C was the most active ($IC_{50} = 33 \pm 4 \mu M$) against the K-562 cell line compared with the known telomerase inhibitor BIBR1532 ($IC_{50} = 208 \pm 11 \mu M$). In this study, the biological assays were carried out in collaboration with the University of Palermo under Prof. Mario Allegra - Department of Biological, Chemical, and Pharmaceutical Sciences and Technologies (STEBICEF).

The second study concerns the comparison of MD-derived pharmacophore models with docking on CDK-2 inhibitors. In this study, the performance of MD pharmacophore modeling approaches, the Common Hit Approach (CHA), and the Molecular dYnamics SHARED PharmacophorE (MYSHAPE) approach, were compared with semi-flexible constrained/unconstrained docking. This work aimed to enrich the hit list of a virtual screening on CDK-2 known inhibitors as a case study. The results highlighted that the MYSHAPE approach performs better when multiple target-ligand complexes are available ($ROC_{5\%} = 0.99$).

Moreover, using short MD simulations improves the screening performance ($ROC_{5\%} = 0.98\text{--}0.99$) with respect to docking ($ROC_{5\%} = 0.89\text{--}0.94$).

The third study concerns the evaluation of the IKK β Binding of Indicaxanthin against the active and inactive form, and the allosteric binding site of hIKK β by Induced-Fit Docking, Binding Pose Metadynamics, and MD. The outcomes of this study showed that Indicaxanthin inhibits prevalently the active form of the hIKK β .

The last study concerns the exploration of the SARS-CoV-2 proteome in the search for potential inhibitors via a structure-based pharmacophore modeling/docking approach. Due to the fact the Ph.D. period took place during the pandemic, in the first months of this one, a computational drug repositioning campaign on the DrugBank database was developed. The final selection of the potential inhibitors was made considering the best binding energy for each compound obtained utilizing MM-GBSA calculation. Molecular recognition analysis showed that these compounds interact with the residues found as crucial for each target of the SARS-CoV2 virus. In conclusion, during this doctoral project, it was demonstrated how the use of *in silico* tools could be effective in the drug discovery process. The computational approaches allowed the identification of promising compounds, and the information obtained could be exploited to optimize the identified inhibitors.

2. MOLECULAR MODELING IN DRUG DISCOVERY

2.1. INTRODUCTION

Drug discovery is an extremely laborious and expensive process, requiring an average of 13 years of research and an investment of 1.8 billion US\$ to bring a single drug from the bench to a patient's bedside [1]. Historically, serendipity and trial-and-error have played a major role in the discovery of drugs. With the advent of medicinal chemistry, a more systematic search for new pharmaceutically active compounds began and these were often evaluated using animal experiments. The chemical modification of lead compounds, on a trial-and-error basis, typically led to compounds with improved potency, selectivity, and bioavailability and reduced toxicity. However, this approach is costly and labor- and time-intensive. The advent of Computer Aided Drug Design (CADD) as a powerful tool for the search for potential therapeutic compounds represents advancement compared to high-throughput screening (HTS) as it permits faster and more economical drug development. The goal of drug discovery is to identify a compound that can modulate the effect of a molecular target that regulates a cascade of biological processes related to a disease. The activation or inhibition of a biomolecule function, such as a protein or nucleic acid, results in a therapeutic benefit to the patient. In addition to organic small molecules, new classes of drugs become every day increasingly important, for example, biopharmaceuticals and especially therapeutic antibodies. Different techniques for improving the affinity, selectivity, and stability of these protein-based therapeutics have also been developed [2].

The three major purposes of CADD techniques in the drug discovery campaign are [3]:

- (1) Filter large compound libraries into smaller sets of predicted active compounds. These ones can be tested experimentally leveraging chemical and biological information about ligands and/or targets to identify and optimize new drugs.
- (2) Guide the optimization of the lead compound, whether to increase its affinity and optimize drug metabolism and pharmacokinetics properties such as absorption, distribution, metabolism, excretion, and the potential for toxicity (ADMET).
- (3) Help the rational design of novel compounds by modifying starting molecules or tying fragments into novel chemotypes together.

Based on the availability of the 3D structure of the target protein, CADD techniques are categorized into either of two types, Structure-Based and Ligand-Based Drug Design (SBDD and LBDD) [4].

2.2 LIGAND-BASED DRUG DESIGN (LBDD)

LBDD takes advantage of information from known bioactive compounds (ligands) and is an essential tool when structural information of a biological target is missing or when the molecular design is not directed toward a target-centric approach without precise knowledge of

the mechanism of action. This strategy is also called indirect drug design because, starting from known active compounds on a specific protein, it tries to find the essential chemical features useful for interacting with that target. When the structural information has been collected it is possible to search for chemical similarities between known and new molecules.

2.2.1 QSAR and Pharmacophore Models

LBDD can be exploited to search for a quantitative structure-activity relationship (QSAR). In this approach, ligand structural information is converted into molecular descriptors, and, through mathematical and statistical analysis, one can predict the biological effect of chemicals. Once found the most robust model, the information can be exploited for the prediction of new analogs with better activity, improved understanding, and optimization of the lead compound to congeners with decreased toxicity [5].

The pharmacophore modeling approach is one of the most applied approaches and is based on the indirect building of a pseudo receptor derived from a model that defines the minimum necessary features a molecule must possess to bind to the target [6-7].

A pharmacophore is an abstract description of the structural features necessary for the molecular recognition of ligands by biological macromolecules. The IUPAC defined a pharmacophore as “an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response” [8].

These features may be located on the ligand and include hydrophobic centroids, positive or negative ionizable sites, hydrogen bond acceptors, or donors and aromatic rings (Figure1)

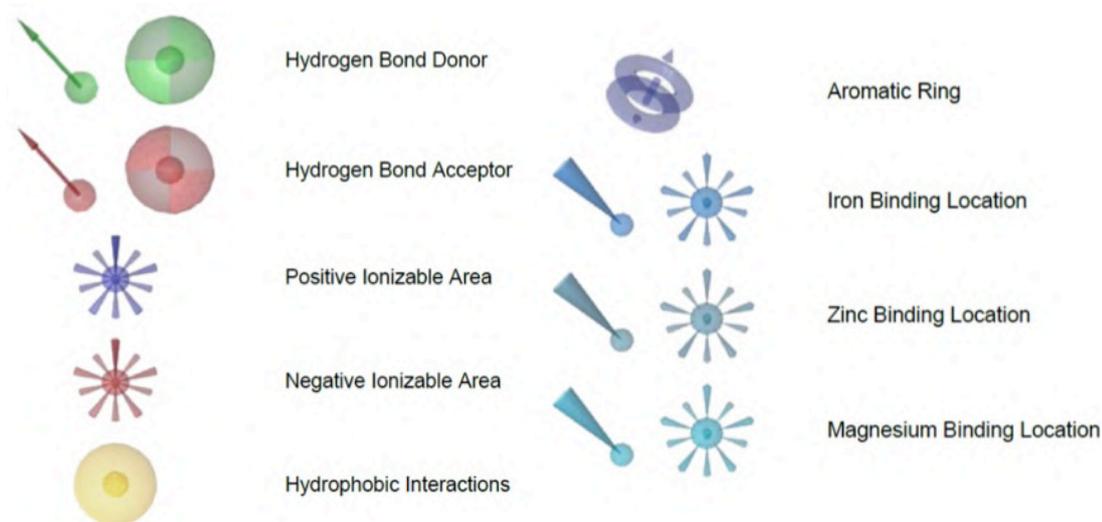


Figure 1: Pharmacophore features as recognized by LigandScout.

The generation of a Ligand-Based pharmacophore model involves several steps:

1. Conformational analysis of the active molecules belonging to the training set.
2. Definition of features for each of the conformations generated.
3. Alignment of the different conformers, to determine the best overlap between the different

groups of pharmacophoric features.

A natural extension to the classical QSAR is the 3D-QSAR, RI-4D-QSAR and MD-QSAR. These methods are a broad term encompassing all those QSAR methods which correlate macroscopic target properties with computed atom-based descriptors derived from the spatial (three-dimensional) representation of the molecular structures [9]. 3D-QSAR correlates activity with non-covalent interaction fields surrounding the molecules such as the Lennard-Jones potential, rather than using experimental constants to define the interatomic interactions [10]. The 4D-QSAR includes conformational flexibility and the freedom of alignment by ensemble averaging in the conventional three-dimensional descriptors found in traditional 3D-QSAR methods [9]. Multidimensional quantitative structure-activity relationship (MD QSAR) is represented by the combination of MD, and the relative computed descriptors, with the generation of QSAR models. This approach provides enhanced predictive power [9].

2.3 STRUCTURE-BASED DRUG DESIGN (SBDD)

SBDD exploits the three-dimensional (3D) structure of the biological target (protein) obtained through methods such as X-ray crystallography, NMR spectroscopy or cryo-EM, to identify putative modulators of the protein activity [11]. When the crystal structure is unavailable, the lack can be overtaken using a homology model. Homology modeling is useful when the query protein is related to other proteins with both a known sequence and structure, named template. The quality of the predicted structure depends on the degree of similarity between the query and the template sequences [12].

In the case of low homology levels, it is possible to assess folding prediction by protein threading. In this technique, also known as fold recognition, each amino acid is assigned a position in a template structure and an evaluation of the best-fit template is made. After the best fit template is selected, the structural model of the sequence is built according the alignment with the best fit template [13].

With the knowledge of the biological target structure, candidate drugs can be optimally designed by medicinal chemists, predicting their binding affinity and selectivity.

The main structure-based techniques are structure-based pharmacophore modeling [14-15], molecular docking [16], and MD [17].

2.3.1 Structure-based Pharmacophore Modeling

Structure-based pharmacophore models aim to be complementary to docking procedures and include the same level of information. However, they are less demanding concerning computational requirements and, therefore, much more efficient [18]. For this reason, it is very suitable for large libraries' virtual screening. Structure-based pharmacophores models can be built both with a macromolecule-ligand complex (holo) structure or a free (apo) structure. The SBDD methods that derive pharmacophore from protein-ligand complexes use the interactions observed between ligand and protein, whereas the SBDD method that derives pharmacophore from apo structure, uses only protein active site information [19].

2.3.2 Molecular Docking

The molecular docking approach characterizes at the atomic level the behavior of small

molecules in the binding site of the target to elucidate fundamental biochemical processes [20]. The docking protocol involves two basic steps: prediction of the ligand conformation, its position and orientation within these sites, as known as pose, and assessment of the binding affinity. These two steps rely on what is defined as a searching algorithm (for pose research) and a simple scoring function for binding affinity calculations.

However, some features, necessary for better accuracy are omitted or underestimated. Molecular docking shows these weaknesses:

- Protein flexibility is not treated extensively.
- Water molecules are treated only implicitly.
- Only a static picture of the binding process is provided.

The earliest docking methods, reported by Kuntz and co-workers in 1982 [21], were based on the lock-and-key theory proposed by Fischer [22], in which the ligand fits into the receptor like lock-and-key. According to this theory, both the ligand and the receptor were treated as rigid bodies.

In the last years, different types of molecular docking have been developed according to ligand and/or receptor flexibility, respectively. In the former, ligand conformations may be generated before docking or within the receptor binding cavity [23]. The proper energetically conformations of ligands are selected using knowledge-based [24] or force field-based methods [25].

In the second case, the “induced-fit” theory [26] created by Koshland takes the Fischer theory a step further, stating that the active site of the protein is continually reshaped during the interaction by ligand. Consequently, it describes the binding events more accurately than the rigid treatment. Induced Fit Docking (IFD) is a protocol that accurately predicts ligand binding modes and the structural changes in the active site of the target.

2.3.3 Molecular Dynamics

The interaction mechanism at the atomic level between the ligand and the active site can be explored using MD. MD in drug design have been demonstrated to give a huge impact on the improvement of drug design strategies by addressing protein folding issues or protein-ligand complex stability through energy profile analysis over time. The first developed MD simulations came up for theoretical physics simulations in the 1950s. In the 1960s, Lennard-Jones potentials were introduced for lipid simulations [28]. Later Karplus and McCammon showed the crucial role of MD simulations in studying biological systems. They used the simulations to obtain different conformations of proteins and nucleic acids.

With the advent of faster, mainly Graphics Processing Unit (GPU)-based clusters, hardware architectures along with the improvement of computation algorithms, classical MD simulations are more important in drug design and development programs [29].

The MD simulation is based on the idea to explore conformational space of the microscopic system on the exam in a time-dependent way.

It is based on the solving the second-order differential equations of Newton’s second law. (Eq. 1)

$$F_i(t) = m_i a_i(t) = - \frac{\partial V(x(t))}{\partial x_i(t)} \quad (1)$$

Where $F_i(t)$ is the force on the i th atom of the system at a given point in the time t ; $a_i(t)$ is the acceleration and m_i is the mass.

The configuration of the system is represented by the vector $x(t)$, which describes the position of the N interacting atoms in the space $x = \{x_1, y_1, z_1 \dots x_N, y_N, z_N\}$.

To improve the exploration of MD, during the last few years, several different enhanced sampling methods have been introduced.

These approaches apply biasing forces or alter the potential energy function to accelerate MD simulations increasing the probability of observing the event of interest, such as ligand-protein (un)binding, treating undruggable sites and unravelling targetable pockets using small molecules, and protein-protein interactions [29]. For example, SUPervised Molecular Dynamics (SuMD) is the most attractive approach to understanding the role of allosteric sites [30], Replica-exchange MD simulation, to accelerate conformational sampling of complex molecular systems [31]. Also, Accelerated MD improves conformational space sampling by reducing energy barriers separating different states of a system [32]. Metadynamics is a method used to estimate the free energy of a system [33].

In the Virtual Screening protocols, the use of MD showed an interesting increase in the screening capability both in terms of the sensitivity of the model and specificity when compared with the classical methods. In the last years, many studies used the information retrieved from the MD trajectory as a starting point to improve structure-based models. Recently, Wieder et al. have proposed the "Common Hits Approach (CHA)". This approach uses the multiple coordinate sets saved during the MD simulations and generates for each frame a pharmacophore model. Pharmacophore models with the same pharmacophore features are pooled to be reduced to only a few hundred representative pharmacophore models. Virtual screening runs are performed with every representative pharmacophore model. The screening results are combined and rescored to generate a single hit list [34].

In the same year, the Molecular dYnamics SHARED PharmacophorE (MYSHAPE) approach by Perricone and co. aims the generation of pharmacophore models using the features common during the MD simulation but not present in the static PDB pharmacophore models [35]. For each structure, MD simulations were carried out, and ligand-protein interactions were analyzed and collected with their appearance frequency. A pharmacophore model was then created using only the common feature patterns that all ligands exhibited during MD simulations. The application of these approaches showed an interesting increase in the screening capability both in terms of the sensitivity of the model and specificity when compared with the classical methods.

3. DESIGN, SYNTHESIS AND BIOLOGICAL EVALUATION OF NOVEL NON-COVALENT PROTEASOME AND IMMUNOPROTEASOME INHIBITORS

3.1 STRUCTURE AND FUNCTIONS OF THE CONSTITUTIVE PROTEASOME

The proteasome, a multi-catalytic enzyme complex, is an essential component of the ubiquitin-proteasome pathway (UPP). It is a key to the selective degradation of 80-90% of intracellular proteins involved in the cell cycle and differentiation, apoptosis, and regulation of transcription factors [36].

The catalytic degradation pathway of the proteasome target proteins consists of:

- Conjugation of a poly-ubiquitination chain to the protein to be degraded, through three enzymes, respectively E1, E2, and E3.
- Degradation of the labeled protein through the 26S protein complex (Figure 2).

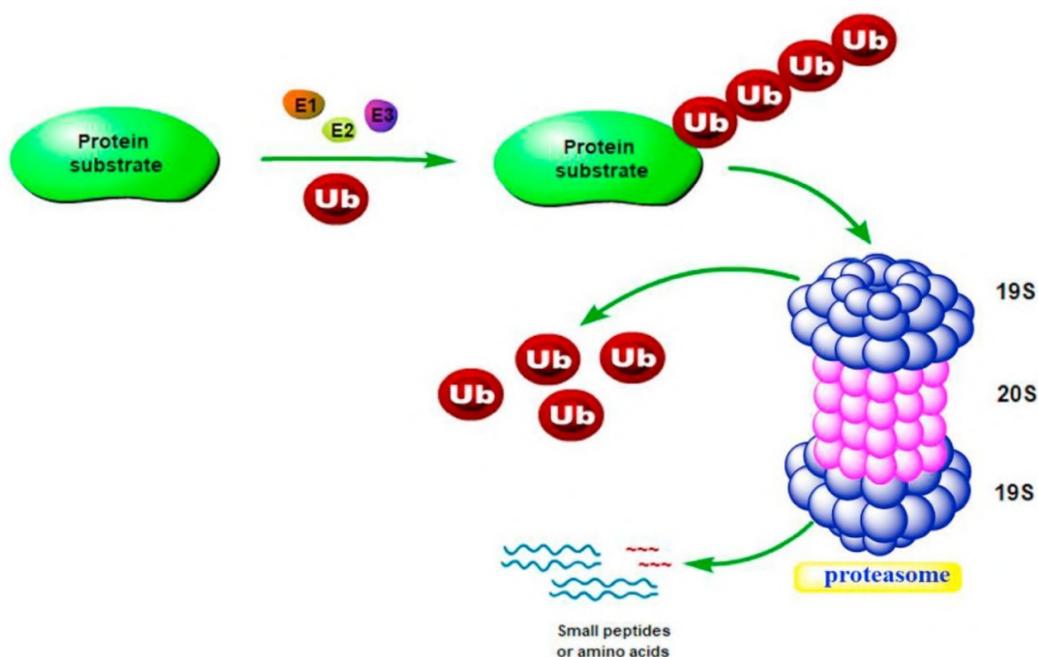


Figure 2: The ubiquitin-proteasome protein degradation pathway.

The 26S proteasome consists of a central catalytic core, the 20S complex, associated with two 19S (or PA700) regulatory structures. The 20S subunit has a cylindrical structure, composed of four superimposed rings, two external α and two internal β complementary. In eukaryotic cells, the α and β rings are each composed of seven distinct subunits, capable of forming four homoheptamers.

The proteolytic activities are closely related to the subunits $\beta 1$, $\beta 2$, and $\beta 5$, respectively with caspases, trypsin, and chymotrypsin activity [37].

The 19S subunits consist of a portion that acts as a base and one that structurally acts as a "lid" of the proteasome. The base is composed of six subunits with ATPase activity and two with structural function, and directly binds the catalytic core 20S; the "lid" portion has instead eight subunits, all devoid of catalytic activity, except one.

The 19S complex performs various functions, including recognizing conformational epitopes of the polyubiquitin chains bound to the substrate.

3.2 STRUCTURE AND FUNCTIONS OF THE IMMUNOPROTEASOME.

It is shown that vertebrates possess a specialized form of the proteasome, the so called immunoproteasome (i20S), expressed mainly in hematopoietic cells, such as lymphocytes and monocytes. It has been noted that in most cells, oxidative stress and pro-inflammatory cytokines are stimuli that lead to over-expression of the immunological variant of the proteasome. The cells that are responsible for antigen presentation (APC) express a baseline level higher than i20S.

One of the main functions of the immunoproteasome is to generate peptides with a hydrophobic C-terminal end, so that they can be quickly processed and can be adapted to the binding pocket of the molecules of the major histocompatibility complex of class I [38]. Structurally, as shown in Figure 3, it has been shown that under the stimulation of pro-inflammatory cytokines such as interferon ($\text{IFN-}\gamma$) and tumor necrosis factor ($\text{TNF-}\alpha$), the constitutive subunits $\beta 1c$, $\beta 2c$, $\beta 5c$, undergo a structural modification, and are replaced by the respective "immunological" subunits, $\beta 1i$ (LMP2, low molecular mass polypeptide 2), $\beta 2i$ (MECL1, multicatalytic endopeptidase similar to 1) and $\beta 5i$ (LMP7, low molecular mass polypeptide 7).

In particular, $\beta 2i$ and $\beta 5i$ maintain the same substrate specificity and at the same time therefore also maintain the same enzymatic activity of the analogous constitutive subunits; on the contrary, $\beta 1i$ carries out a ChT-L activity, which leads mainly to the hydrolysis of carboxy-terminal polypeptides composed of hydrophobic amino acids (in particular Tyr, Trp, Phe, and Leu) [36].

Studies have shown that dendritic cells from mice lacking expression of $\beta 1i$, $\beta 2i$ and $\beta 5i$ had defects in MHC class I antigen presentation.

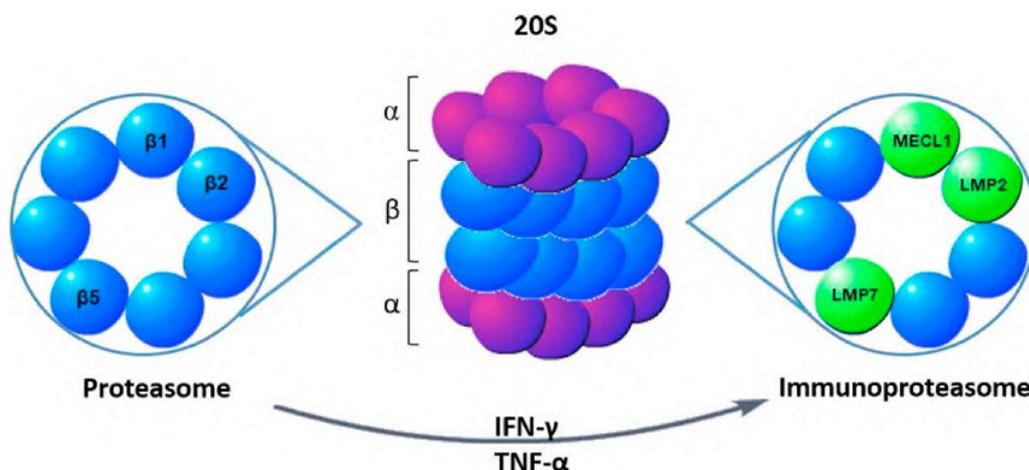


Figure 3: Schematic structure of the 20S proteasome and i20S immunoproteasome.

Regarding the role of IP, it has long been assumed that most tissues of non-hematopoietic origin were normally devoid of IP, unless these tissues faced immunological challenges, and therefore were subjected to stimuli via autocrine or paracrine production of $\text{IFN-}\alpha/\beta$ or $\text{IFN-}\gamma$. Recent studies have confirmed the expression of the immunoproteasome in a wide range of non-immunological cells and organs. In this way, many important functions have emerged such as the regulation of protein homeostasis, cell proliferation, and cell signalization.

The immunoproteasome plays a vital role in antigen presentation due to its hydrolysis capabilities and participates in a variety of immune processes. Furthermore, the immunoproteasome also has an important effect on the regulation of protein homeostasis, cell proliferation, and cytokine gene expression [39]. It is also expressed in non-immune cell types during inflammation or neoplastic transformation [40].

3.2.1 Inflammation, viral infections, and autoimmune diseases

The immunoproteasome is also important for activating the immune system against viral infections. Recent studies have shown, for example, that the reduction or absence of the immunoproteasome in mice reduces the activation of CD8⁺ T cells in hepatitis B virus (HBV) infection and lymphocytic choriomeningitis virus (LCMV) infection. Instead, the decrease of LMP2 (β 2i) reduces inflammatory cytokines (IL-1 β , IL-6, and TNF- α) produced during influenza viral infection [36].

Viral hepatitis C (HCV) infection induces the expression of IFN type I (IFN- β) and immunoproteasome subunits in hepatocytes, while suppression of IFN- β inhibits the expression of the immunoproteasome and IFN type I (IFN- α) induces the expression of the immunoproteasome complex in hepatocytes. Therefore, viral infection, IFN production, and immunoproteasome expression are most often closely related events. In addition to the immunological role of the immunoproteasome, the latter is involved in the pathogenesis of numerous inflammatory diseases through the stimulation of the polarization of T lymphocytes and the activation pathway of the nuclear factor NF- κ B, essential for the transcription of many genes that code for inflammatory cytokines in macrophages [41].

For example, from the literature, it is evident that LMP7 deficiency suppresses the differentiation of naïve CD4⁺ T lymphocytes into Th1 and Th17 cells and instead promotes their differentiation into regulatory T cells (Figure 4) [37]

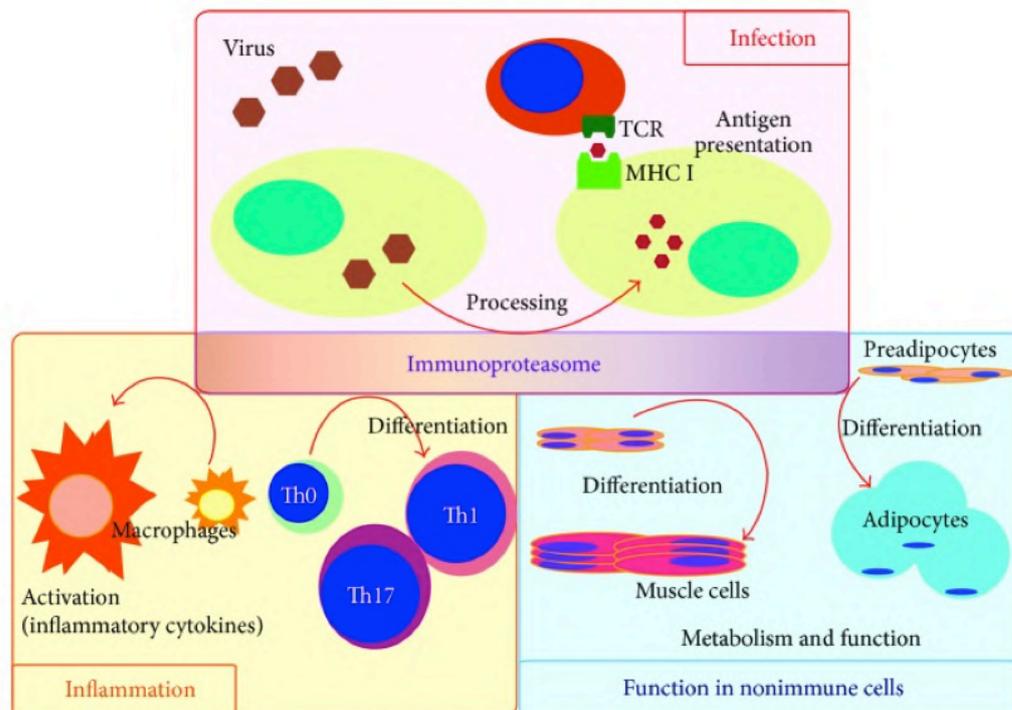


Figure 4: Immunoproteasome as a potential therapeutic target.

Overexpression of the immunoproteasome complex has also been found in diseases such as Crohn's disease, ulcerative colitis, hepatitis, multiple sclerosis, experimental rheumatoid arthritis, and systemic lupus erythematosus [36].

3.2.2 Hematological neoplasms

Cancer cells likely have an altered protein cell cycle that is accompanied by an increased replicative speed, and an increased accumulation of damaged proteins thus conditioning a greater dependence on proteasomal degradation processes. Multiple mechanisms have been proposed to explain the antitumor activity of immunoproteasome inhibitors, they mostly involve inhibiting the degradation of pro-apoptotic and/or regulatory proteins. Cancer cells usually have higher levels of immunoproteasome than normal cells and are more resistant to pro-apoptotic effects, making the immunoproteasome a therapeutic target in the treatment of Multiple Myeloma. Multiple studies show that proteasome inhibitors are more cytotoxic to proliferating neoplastic cells than to normal quiescent cells. [42]

3.2.3 Neurodegenerative diseases and aging

Most central nervous system diseases are related to an increase in the number of central polymeric proteins. Several animal model studies have verified that the immunoproteasome $\beta 1i$ and $\beta 5i$ subunits were overexpressed in central nervous system diseases such as Alzheimer's disease, the most common neurodegenerative disorder [43], Huntington's disease, and sclerosis lateral amyotrophic [44-45].

Furthermore, regarding cellular senescence, it has been found that a high expression of the immunoproteasome in the human brain is increased with aging, this may be related to high concentrations of proteins damaged by oxidative stress and defective ribosomal products (DRP). Therefore, it been stated in numerous studies that immunoproteasome inhibitors have found great use in therapy against neurodegenerative diseases and cellular aging [46].

3.3 INHIBITORS OF CONSTITUTIVE PROTEASOME AND IMMUNOPROTEASOME

The inhibitors of Proteasome and Immunoproteasome can be divided into two main categories, thanks to their different chemical structure which consequently leads to different interactions: the inhibitors capable of forming a covalent bond and the non-covalent inhibitors. [38]

Among the covalent inhibitors, there are boronates, peptide aldehydes, epoxyketones, and β -lactones. To date, they have been approved by the FDA for the treatment of multiple myeloma and several inhibitors are in clinical trials. However, most of them are non-selective inhibitors, therefore they inhibit both the constitutive proteasome and the immunoproteasome, except the inhibitor KZR-616, and the series of analogs associated with it, which have shown a strong selectivity towards the IP [47-48].

3.3.1 Non-selective inhibitors

Bortezomib is the first non-selective immunoproteasome inhibitor approved by the Food and Drug Administration in 2003 for the treatment of second-line multiple myeloma and in 2006 as a treatment for mantle cell lymphoma.

Cell death induced by Bortezomib is related to the stress induction of the smooth endoplasmic reticulum and UPR, Unfolded Protein Response. It can inhibit the NF- κ B inflammatory pathway, initiate apoptosis with the activation of caspase-8 and causes an increase in the production of ROS (Reactive Oxygen Species). [36, 43].

Bortezomib is a dipeptide of boric acid, that binds the β 5 proteasome subunit and blocks the chymotrypsin activity. It can bind the β 1 subunit but with lesser affinity (responsible for a lighter block of caspase-like training). Bortezomib binds the immunoproteasome to a lesser extent.

The mechanism of action of Bortezomib involves the formation of a reversible tetrahedral bond following the reaction of the boric acid dipeptide with the Thr1 residue of the β 5 subunit (Figure 5)[49].

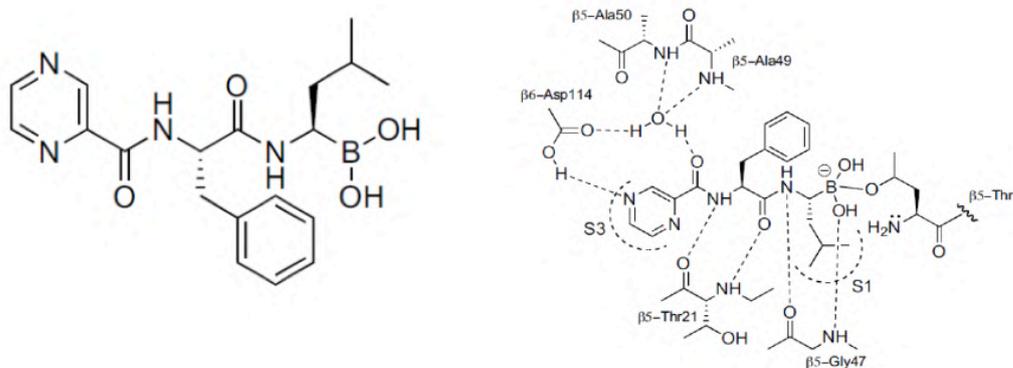


Figure 5: Structure of Bortezomib and its binding mode.

Recent studies have shown that the β 1 and β 2 subunits can also be inhibited by Bortezomib especially when high drug concentrations are reached [39]. However, its therapeutic use is associated with a serious undesirable effect, such as the onset of peripheral neuropathy.

Second-generation proteasome inhibitors have been developed due to the problems described above related to Bortezomib. Among these, there is Carfilzomib, a second-generation inhibitor approved by the FDA in 2012 for the treatment of refractory multiple myeloma. Carfilzomib (Figure 6A) is a tetrapeptide α' , β' -epoxyketone with an *N*-acyl-morpholine portion, obtained from the optimization of epoxomicin, a powerful and natural irreversible inhibitor of the immunoproteasome.

The mechanism of action of Carfilzomib and derivatives involves both the hydroxyl and the amino groups of Thr1, through a double sequential nucleophilic attack on the α' , β' -epoxy ketone portion (Figure 6B).

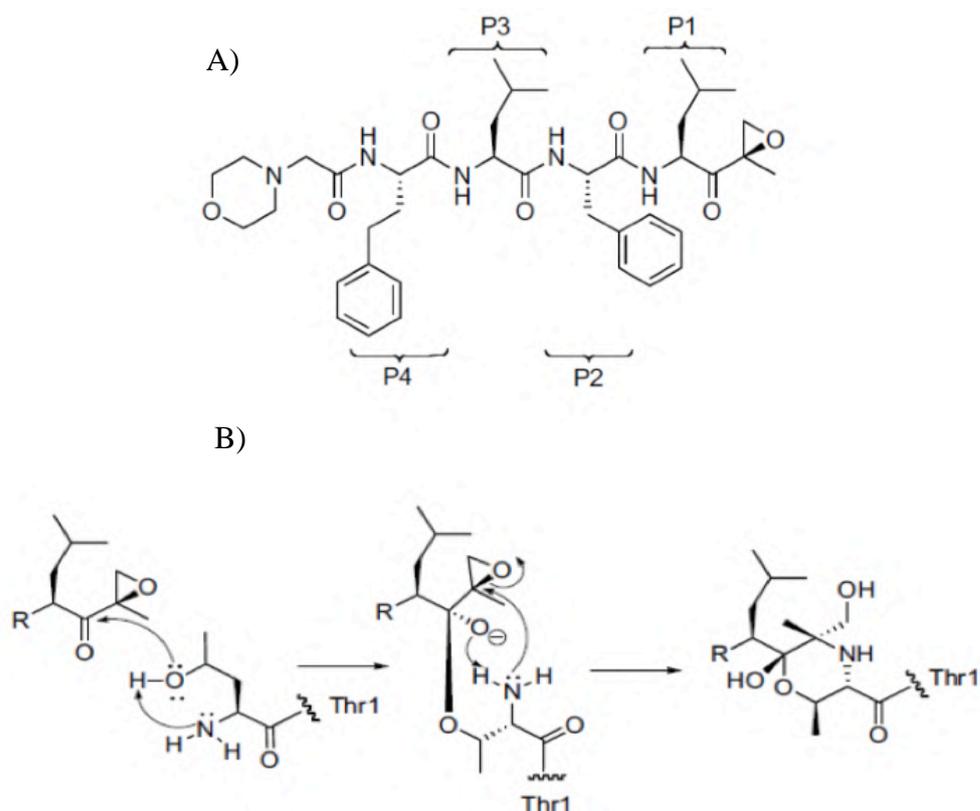


Figure 6: Structure of Carfilzomib (A); Mechanism of action of α,β' -epoxyketones (B).

Ixazomib is an inhibitor approved by the FDA in November 2015, combined with lenalidomide and dexamethasone, for treating patients with multiple myeloma resistant to the previously mentioned therapies. Ixazomib is a prodrug taken orally, as opposed to the previous ones which involved parenteral, intravenous, or subcutaneous administration. *In vivo*, it is metabolized by hydrolysis into the boronic acid derivative MLN2238 (Figure 7) [50].

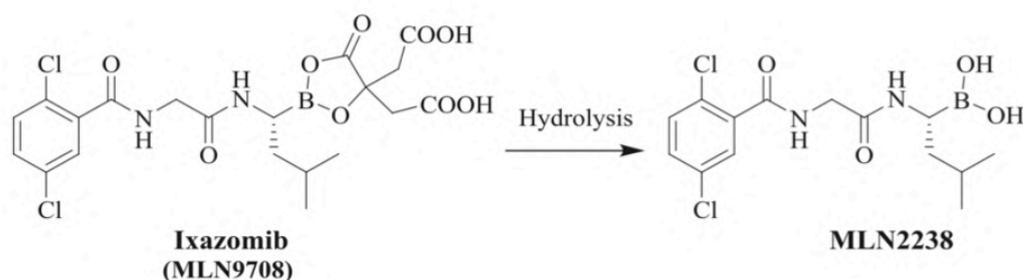


Figure 7: *In vivo* activation of Ixazomib (MLN9708) in MLN2238.

The preferential subunit on which Ixazomib acts is β_5 , however low activity on the β_1 and β_2 subunits have also been observed. It also has a better pharmacokinetic and pharmacodynamic profile and a higher antitumor activity than Bortezomib and Carfilzomib [51-52].

Oprozomib is a tripeptide epoxyketone, structurally like Carfilzomib. Its goal is to bring myeloma cells to apoptosis, together with the inhibition of the angiogenesis process and the migration of myeloma cells. The great advantage of this drug over the others is the oral route of administration, thus significantly improving patient compliance.

From a structural point of view, Oprozomib (Figure 8) has undergone modifications compared to Carfilzomib, the Leu residue in the P1 site of Carfilzomib has been replaced by a Phe residue,

while two hydrophilic residues of O-methyl-serine have been inserted, with stereochemistry S, both in position P2 and P3. The inhibitory activity of Oprozomib is aimed at the $\beta 5i$ and $\beta 5c$ subunits [51,53-55].

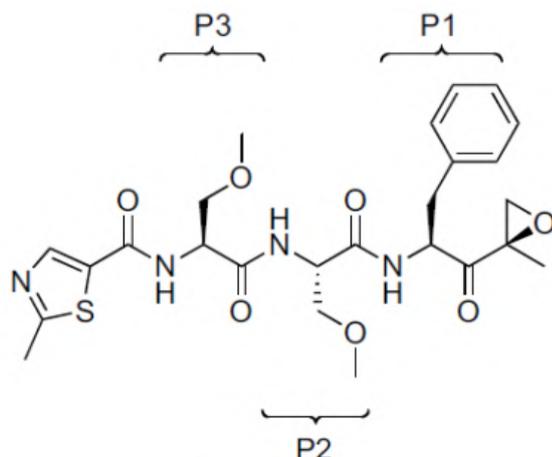


Figure 8: Structure of Oprozomib

Delanzomib and Marizomib (Figure 9A,B) are also non-selective inhibitors of the immunoproteasome undergoing clinical trials. Delanzomib is closely related to Bortezomib, while Marizomib, which is part of the category of β -lactones, can irreversibly bind to the proteasome, inhibiting all three catalytic subunits and thus blocking the $\beta 5$, $\beta 1$ and $\beta 2$. Additionally, it binds to the immunoproteasome causing a better response than Bortezomib. Delanzomib is a boronic acid derivative with a reversible inhibitory action at the level of the $\beta 5$ subunit of proteasome, although less than that of Bortezomib and Carfilzomib [52, 56].

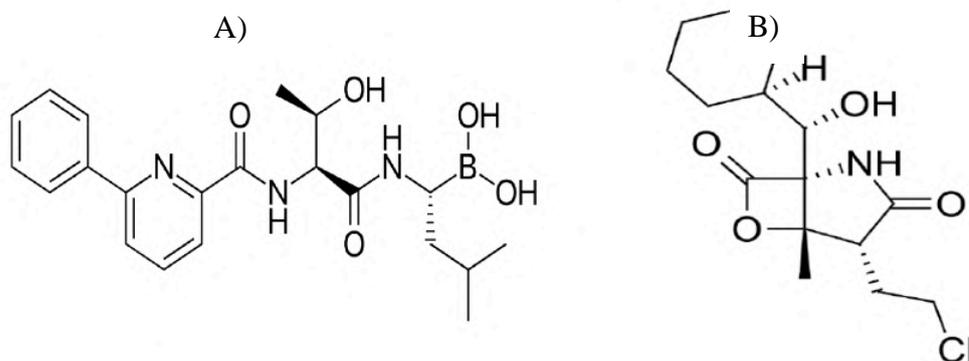


Figure 9: Structure of Delanzomib (A); Marizomib (B)

MG-132 (Figure 10) is a reversible inhibitor of both proteasome and immunoproteasome with IC_{50} s of 100 nM

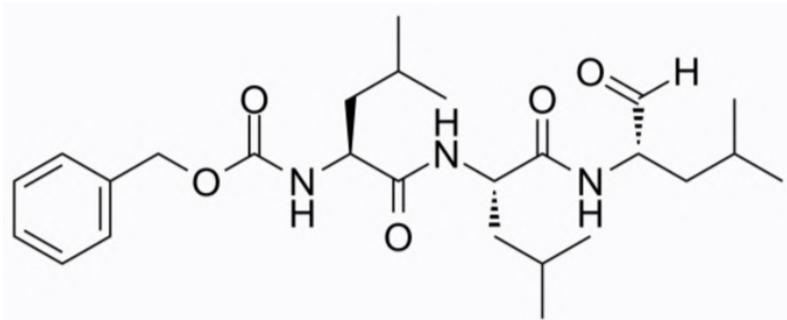


Figure 10: Structure of MG-132

3.3.2 Selective Inhibitors

In the treatment of hematological malignancies, nervous system, and immune system diseases, the development of immunoproteasome-selective inhibitors with less toxicity is of prime importance. The most common selective immunoproteasome inhibitors are the epoxyketone peptides.

ONX-0914 (Figure 11) is the first immunoproteasome-selective epoxyketone peptide inhibitor, which showed better activity towards $\beta 5i$ than the $\beta 5c$ subunit and significantly reduced levels of circulating autoantibodies in the murine model [57].

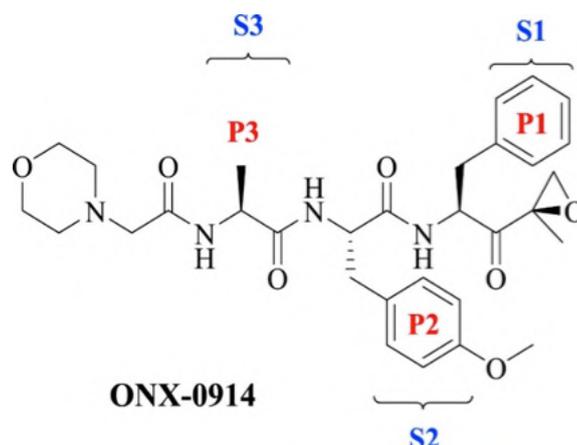


Figure 11: Structure of ONX-0914

Analysis of the structure of ONX-0914 in complex with the constitutive proteasome and the immunoproteasome indicated that the Phe residue at the P1 site induced a critical conformational change in the S1 pocket of $\beta 5c$, instead of only small conformational changes in $\beta 5i$. The N-terminal residues are mostly morpholine, which does not bind to the target, but are essential for improving the solubility and physico-chemical properties of these analogues [57]. Furthermore, preclinical studies have revealed that ONX-0914 could block the production of $\text{INF-}\gamma$ and IL-2 by T cells. Its activity has also proved effective in the treatment of systemic lupus erythematosus and the resolution of an experimentally induced inflammatory phenomenon on mice of about 8-10 weeks. This selective inhibitor for the $\beta 5i$ subunit (LMP7), was able to inhibit the production of cytokines and attenuate the progression of colitis inductively [42]. It also represents a good therapeutic prospect for rheumatoid arthritis, multiple sclerosis, and other degenerative diseases [49, 58].

PR-924 is a $\beta 5i$ -selective epoxyketone tripeptide that blocks the multiple growths of myeloma cells both *in vitro* and *in vivo* (Figure 12A) [59].

KZR-616 is a selective inhibitor obtained following optimization of ONX-0914 and PR-924, it

was the only one that was approved for clinical study in the treatment of systemic lupus erythematosus. It is an epoxyketone tripeptide, that showed inhibitory activity towards the $\beta 5i$, $\beta 2i$ and $\beta 1i$ subunits, less active on the catalytic subunits of the constitutive proteasome [60] (Figure 12B).

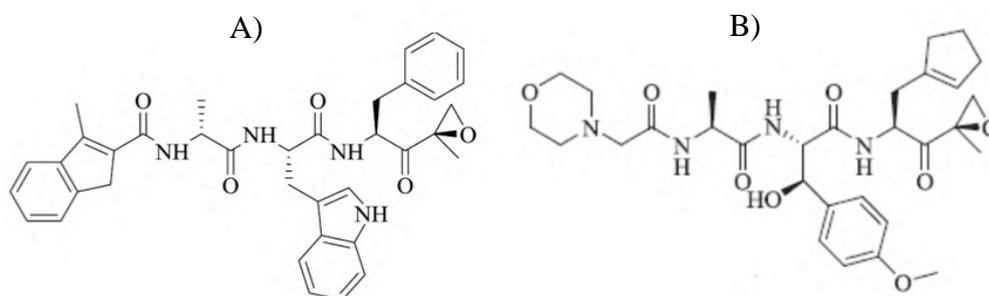


Figure 12: Structure of PR-924 (A); KZR-616 (B)

UK-101 is a potent and selective peptide inhibitor of the $\beta 1i$ subunit [61]. It was obtained through SAR studies on epoxomicin and eponemycin (Figure 13). UK-101 contains a hydrophobic heptanoic tail at the N-terminal, an alanine residue at the P2 position and the hydroxyl group at the P1' site protected with a tertbutyldimethylsilyl (TBDMS) group at the C-terminal moiety, whereas the P1 leucine plays a key role for the inhibition of the ChT-L activity. In *in vitro* assays, UK-101 selectively inhibited $\beta 1i$ subunit with IC_{50} values in the range 104 nM and displays 144- and 10-fold selectivity over $\beta 1c$ ($IC_{50}=15 \mu M$) and $\beta 5$ subunit ($IC_{50}=1 \mu M$), respectively [62]. UK-101 induces cell apoptosis and can be used for the study of prostate cancer

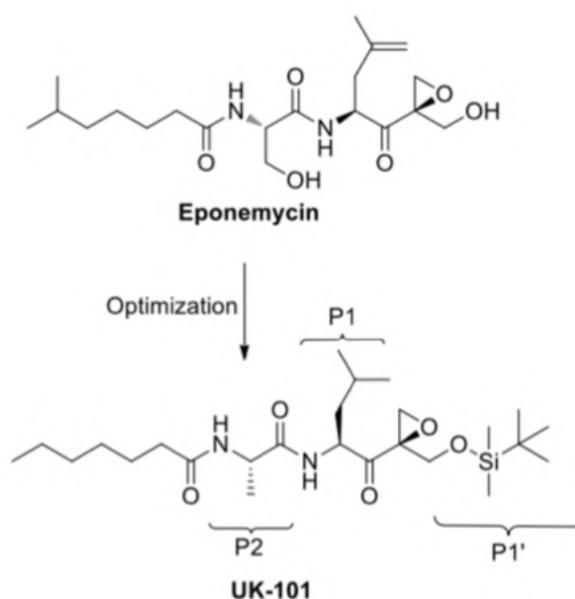


Figure 13: Development of UK-101 from Eponemycin.

The dipeptidyl aldehyde IPSI-001 (Figure 14) showed good selectivity towards immunoproteasome core particles $\beta 1i$ and $\beta 5i$. The mechanism of action proceeds through the nucleophilic attack of the hydroxyl group of Thr1 on the carbonyl carbon, thus leading to the formation of a tetrahedral hemiacetal. It was demonstrated that increasing the hydrophobic group of the P1 amino acid determines a great selectivity towards the immunoproteasome, while

the activity towards the constitutive proteasome decreased [63].

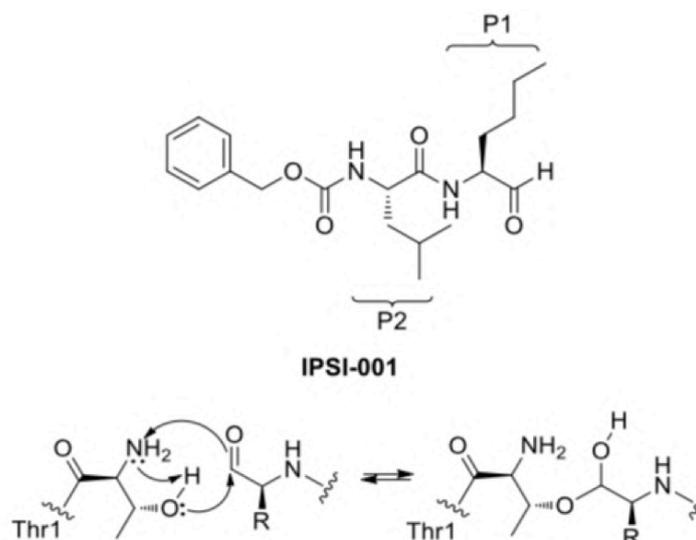


Figure 14: Structure and mechanism of action of IPSI-001

3.3.3 Non-covalent peptide inhibitors

Non-covalent peptide compounds are characterized by the absence of the electrophilic substituent at the C-terminus, as opposed to the covalent ones. The reason why these compounds cannot form the covalent bond with the Thr1 residue present in the active site. However, some may still maintain good proteasome inhibitory activities.

A series of N, C-capped non-covalent dipeptides were designed and synthesized, and each compound of the analog series was suitably modified to obtain the best inhibitory activity against the various types of immunoproteasomes, for example, the insertion of a β -amino acid in the non-covalent peptide skeleton, drastically decreases the inhibitory activities against $\beta 5c$, while keeping the inhibitory activity against $\beta 5i$ high. In the case of PKS2251 (Figure 15), it was obtained through the introduction of β -Ala in position P2 and Asn (tBu) in position P3 [64, 49].

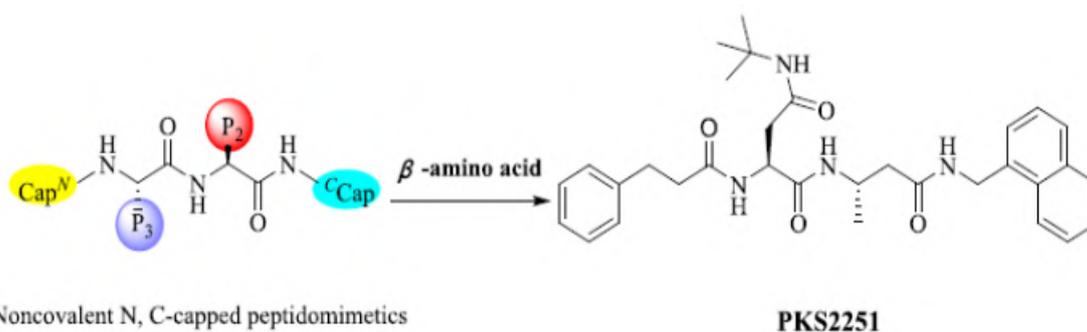


Figure 15: Structure of N, C-capped peptidomimetic, PKS2251

3.4 IMMUNOPROTEASOME AND NON-COVALENT INHIBITION: EXPLORATION BY ADVANCED MOLECULAR DYNAMICS AND DOCKING METHODS

3.4.1 INTRODUCTION

The discovery of selective immunoproteasome inhibitors is pivotal to bringing new chances for the treatment of hematological diseases. Non-covalent inhibition is strongly desirable because it is not responsible for potential toxicity due to off-target binding. Therefore, non-covalent inhibitors may have an improved selectivity and less reactivity and instability [65, 66]. Starting from a panel of selective non-covalent inhibitors of the β 1i and/or β 5i subunits characterized by a 2(1*H*)-pyridone scaffold linked to an amide function [67], an investigation of the binding mechanism of N-Benzyl-2-(2-oxopyridin-1(2*H*)-yl)propanamide (1) was carried out. It proved to be the most potent and selective inhibitor, with a $K_i = 21$ nM against the β 1i subunit. Advanced MD methods, such as MD binding (MDB) [68] and Binding Pose Meta Dynamics (BPMD) [69] were employed. To have a comparative point of view, induced-fit docking (IFD) [70] and BPMD studies was also carried out for the other four compounds, N-benzyl-2-(2-oxopyridin-1(2*H*)-yl)acetamide (2), N-cyclohexyl-3-(2-oxopyridin-1(2*H*)-yl)propanamide (3), N-butyl-3-(2-oxopyridin-1(2*H*)-yl)propanamide (4), and (S)-2-(2-oxopyridin-1(2*H*)-yl)-N,4-diphenylbutanamide (5) that showed high inhibitory activities towards the β 1i subunit (Figure 16).

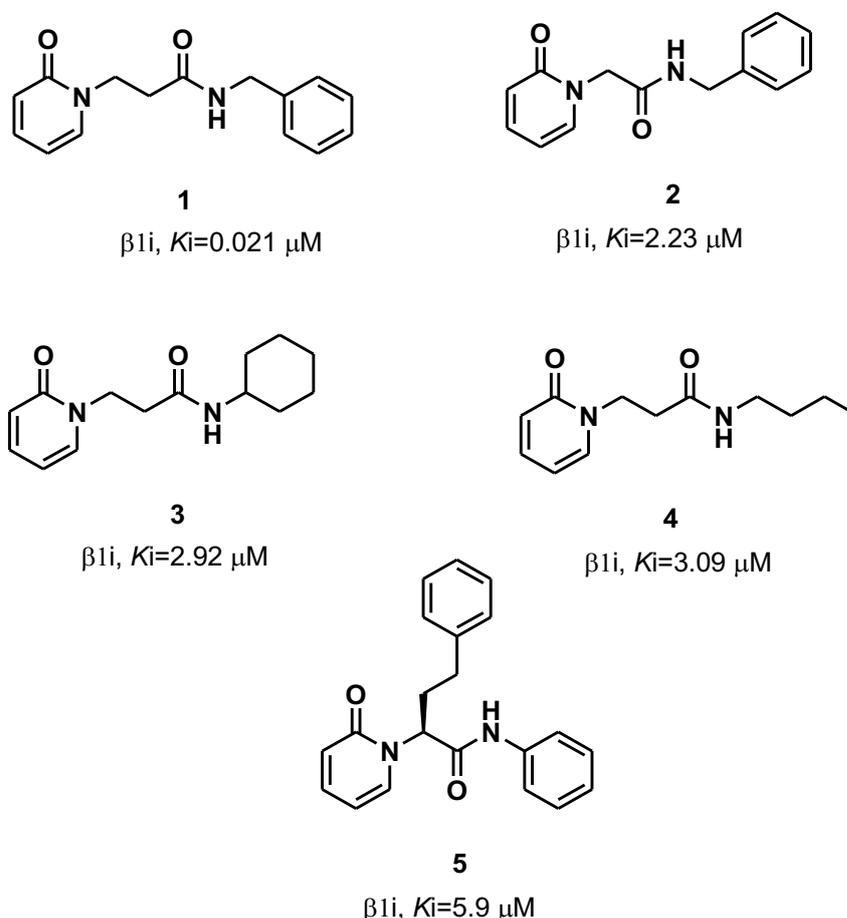


Figure 16: Structures and K_i values of the selective β 1i inhibitors (1-5)

3.4.2 MATERIALS AND METHODS

3.4.2.1 System and Ligands preparation

The catalytic subunit $\beta 1i$ (LMP2 and PSMB9) was extracted from the murine i20S in complex with the inhibitor ONX-0914 (PR-957) bound to the $\beta 5i$ subunit (PDB ID: 3UNF) [71]. Both 20S subunits, murine and human, share a sequence identity of more than 90%, and the few non-identical residues are external to the active sites. As reported in the literature, in the case of covalent co-crystallized inhibitors [72], the reactive residue at the catalytic site was rebuilt after removing the covalent inhibitor by breaking the covalent bond and filling in the open valence. In this case, the involved residue was Thr1. The protein was prepared with the Protein Preparation Wizard [73] included in the Maestro suite (Maestro, Schrödinger, LLC, 2021, New York, NY, USA): adding bond orders and hydrogen atoms to the crystal structure using the OPLS2005 force field. Next, Prime [74] was used to fix missing residues or atoms in the protein and to remove co-crystallized water molecules. The protonation states at pH 7.2 ± 0.2 of the protein and the ligand were evaluated using Epik 3.1 [75]. The hydrogen bonds were optimized through the reorientation of hydroxyl bonds, thiol groups, and amide groups. In the end, the systems were minimized with the value of convergence of the RMSD of 0.3 \AA . Amide 1 was parameterized using the BiKi suite [76] at the AM1-BCC [77] level of theory. Partial charges were derived using the RESP method [78] in Antechamber [70]. Compounds 2–5 were prepared using Schrödinger LigPrep v. 2021-1 (LigPrep, Schrödinger, LLC, 2021, New York, NY, USA). The force field adopted was OPLS2005, and Epik was selected as the ionization tool at pH 7.0 ± 2.0 . Tautomers generation was flagged, and the maximum number of conformers generated was set at 32.

3.4.2.2 MD-Binding Simulations

The MD-binding method [68] within the BiKi suite [76] (BiKi Technologies s.r.l., Genova, Italy) exploits an additive external force that is summed as the regular potential energy of the system to enhance the probability of observing the binding event. The bias is represented by external electrostatic-like forces acting between a subset of the residues of the binding site and the ligand. The intensity of the bias is controlled by the adaptivity rules and gradually switches off as the ligand moves toward the subset of residues; after the conjectured passing of the transition state has occurred, it slowly recovers the behavior of classical unbiased MD [79]. The protocol for MD-binding consists of crucial steps: characterization of the binding pocket using Nano Shaper [78]. Nano Shaper calculations provide a characterization of the binding pocket, which identifies the atoms facing the pocket entrance in the protein structure. This information was then used by BiKi software for the initial ligand positioning outside the binding cavity. Subsequently, an additive external force was made to enhance the sampling of the binding event. Once the ligand was positioned through the “Residue Placement” tool in BiKi, the system was solvated in an orthorhombic box using the TIP3P water model [80]. A suitable number of counterions were added to neutralize the overall system. Then, the whole system underwent energy minimization by using the Amber99SB-ildn force field [81]. According to the standard protocol [68], four different consecutive equilibration steps were performed: (1) 100 ps in the NVT ensemble at 100 K with both the protein backbone and ligand re-strained (1000 kJ/mol nm^2), (2) 100 ps in the NVT ensemble at 200 K with both the protein backbone and the ligand restrained, (3) 100 ps in the NVT ensemble at 300 K with the free protein and the ligand restrained, and (4) 1000 ps in the NPT ensemble at 300 K with the free protein and the ligand restrained. Electrostatic interactions were treated with the cut off method for short-range interactions and with the particle mesh Ewald method for long-range interactions ($rlist = 1.1$

nm, cutoff distance = 1.1 nm, vdW distance = 1.1 nm, and PME order = 4). The constant temperature conditions were provided using the velocity rescale thermostat [82], which is a modification of Berendsen's coupling algorithm [83]. The coordinate output from the last simulation was then used as the input to produce the biased MD. Finally, 20 replica production runs, 20-ns-long in the NVT ensemble at 300 K, were performed for each complex using $C = 0.1$ (the fraction of the felt force, here 10%), a smoothing window size of 1000 samples, and a maximal $K(t)$ of 0.001 (maximal steering constant).

3.4.2.3 Plain MD Simulations

The plain MD simulations were carried out using Desmond 6.5 [84] using the OPLS2005 force field [85] (Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, USA). The complexes were solvated in orthorhombic boxes using the TIP3P water model. Ions were added to neutralize the charges. The systems were minimized and equilibrated at a temperature of 303.15 K and a pressure of 1.013 bar. The system was simulated as an NPT ensemble; a Nose–Hover thermostat and Martyna–Tobia–Klein barostat were used. The integration time step was chosen to be 2 fs. To keep the hydrogen–heavy atom bonds rigid, the SHAKE algorithm was used. A 9 Å cutoff radius was set for the short-range Coulomb interactions, and smooth particle mesh Ewald was used for the long-range interactions. For each replica, 10-ns MD simulations were carried out for a total of 200 ns, with 1.2-ps detection ranges for energy and 4.8 ps for the trajectory frames. The stability of the systems was evaluated using the root mean square deviation (RMSD) of the aligned protein and ligand coordinate set calculated against the initial frame. Visualization and analysis of the MD trajectories were performed using the Desmond simulation analysis tools in Maestro. The trajectory frames were clustered according to the hierarchical cluster linkage method. The 1000 frames recorded in each simulation were clustered considering the binding site conformations into 10 clusters based on the atomic RMSDs.

3.4.2.4 Binding Pose Meta Dynamics (BPMD)

Binding Pose Meta Dynamics (BPMD) is an automated, enhanced sampling, metadynamics-based protocol in which the ligand is forced to move around its binding pose. This method showed the ability to reliably discriminate between the correct ligand binding pose and plausible alternatives generated with docking or plain MD studies [69].

According to the protocol, 10 independent metadynamics simulations of 10 ns were performed using as a collective variable (CV) the measure of the root mean square deviation (RMSD) of the ligand heavy atoms with respect to their starting positions. The alignment before the RMSD calculations was done by selecting protein residues within 3 Å of the ligand. The $C\alpha$ s of these binding site residues were then aligned to those of the first frame of the metadynamics trajectory before calculating the heavy atom RMSD to the ligand conformation in the first frame. The hill height and width were set to 0.05 kcal/mol (about 1/10 of the characteristic thermal energy of the system, kBT) and 0.02 Å, respectively. Before the actual metadynamics run, the system was solvated in a box of SPC water molecules [86], followed by several minimizations and restrained MD steps that allow the system to slowly reach the desired temperature of 300 K, as well as release any bad contacts and/or strain from the initial starting structure. The final snapshot of the short unbiased MD simulation of 0.5 ns was then used as the reference for the following metadynamics production phase. After the simulation, the stability of the ligand during the course was represented by three scores: PoseScore, Persistence Score (PersScore), and Composite Score (CompScore). The PoseScore is indicative of the average RMSD from the starting pose. A steep increase of this value is a symptom that the ligand is not in a well-defined energy minimum and, probably, it might not have been accurately modelled. PersScore

is a measure of the hydrogen bond persistence calculated in the last 2 ns of the simulation that have the same number of hydrogen bonds as the input structure, averaged over all the 10 repeated simulations. It covers a range between 0 and 1, where 0 indicates that either the starting ligand pose did not have any interactions with the target or that the interactions were lost during the simulations, while 1 indicates that the interactions between the starting ligand pose and the last 2 ns of the simulations were retained. CompScore is the linear combination of the PoseScore and PersScore; lower values equate to more stable complexes. Each complex, previously obtained, was run on a single node with a 1 GPU card NVIDIA GeForce RTX2070.

3.4.2.5 Induced-Fit Docking

The induced-fit protocol (IFD)—developed by Schrödinger [70]—is a method for modelling the conformational changes induced by ligand binding. The protocol starts with an initial docking of each ligand using a softened potential (van der Waals radii scaling). Then, a side-chain prediction within a given distance of any ligand pose (5 Å) is performed. Subsequently, a minimization of the same set of residues and the ligand for each protein/ligand complex pose is performed. After this stage, any receptor structure in each pose reflects an induced fit to the ligand structure and conformation. Finally, the ligand is rigorously docked, using Glide XP (Glide, Schrödinger, LLC, 2021, New York, NY, USA), into the induced-fit receptor structure. IFD was performed using a standard protocol, and the OPLS2005 force field was chosen. The receptor box was centered on the active site of β 1i, according to the Nano Shaper calculations. During the initial docking procedure, the van der Waals scaling factor was set at 0.5 for both the receptor and ligand. The Prime refinement step was set on the sidechains of residues within 5 Å of the ligand. For each ligand docked, a maximum of 20 poses was retained to then be redocked in XP mode [87].

3.4.2.6 MM-GBSA-Binding Free Energy Calculations

Prime/MM-GBSA was used for the estimation of ΔG binding. The MM-GBSA approach employs molecular mechanics, the generalized Born model, and the solvent accessibility method to elicit free energies from structural information, circumventing the computational complexity of free energy simulations, wherein the net free energy is treated as a sum of a comprehensive set of individual energy components, each with a physical basis [88]. In our study, the VSGB solvation model was chosen using the OPLS2005 forcefield with a minimized sampling method [74].

3.4.3 RESULTS AND DISCUSSION

3.4.3.1 MD-Binding (MDB) Analysis

With the aim of gaining more insights into non-covalent inhibitors of the immunoproteasome, the enhanced sampling methods were exploited to collecting mechanistic insight into the binding process.

The MDB protocol implemented in BiKi software was performed to simulate the events that elapsed among the compound 1 unbound and the compound 1 entrance in the binding pocket. The advanced proposed MDB protocol requires the identification of the binding pocket with NanoShaper software, characterized by Thr1, Val20, Ser21, Phe31, Lys33, Leu45, Ser46, Gly47, Ser48, Ala49, Ala52, Ser129, and Ser168 residues (Figure 17a).

Compound 1 is positioned with a random orientation at a predetermined distance, measured in terms of the solvation shell around the ligand. Twenty replicas of 20 ns for each entrance

starting from the apo structure were carried out, thus collecting 400 ns of MDB simulation (Figure 17b).

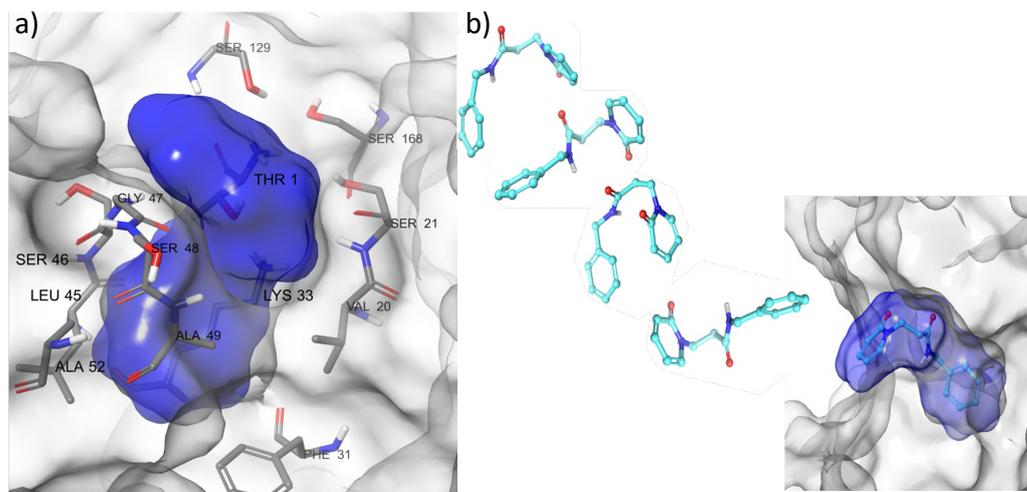


Figure 17: a) Identification of the binding cavity of the $\beta 1i$ subunit (solid blue) by NanoShaper software with the residues involved in the binding pocket; b) Time sequence of compound 1 approaching the active site of the $\beta 1i$ subunit studied by MD bind. The figure is representative of the 20 replicas

The simulations overcome the energetic barrier in an average time of 2 ns, reaching the binding site. For the unavailability of crystallographic structure for non-covalent ligands on $\beta 1$, the RMSD of the protein backbone was used as a reference. All replicas showed a protein backbone RMSD average $<2 \text{ \AA}$, decreasing when the ligand arrived at the binding site. In most replicas, it is observed the ligand entered the active site in the first 8 ns, and in the last 10 ns its refinement at the binding site was observed. Successively, the electrostatic bias was removed, and 10 ns of plain MD simulations of each replica were performed to extend the sampling of the bound states for a total of 600 ns of the simulation.

Twelve replicas showed high stability, with an average value of RMSD 1.5 \AA of the complex, while in eight replicas, the ligand rapidly drifted away.

The trajectories of the 12 stable replicas were clustered. Each trajectory was recorded in 1000 frames and these frames were clustered considering the RMSD of the binding site backbone (12,000 frames total). Each replica returned three representative clusters for a total of 36 MD representative poses. The clustering of the survived complexes trajectories allowed identifying three representative poses (pose 1, pose 2, and pose 3) observed during the simulations. In pose 1, two H-bonds were formed between the oxygen of the amide group and Ser21 and between the hydrogen of the amide group and Gly47. The binding of the ligand was strengthened by several van der Waals contacts between the benzyl group and the residues Val20, Phe31, Lys33, Gly47, Ala49, and Ala52. Val20, Ser21, Ser46, Gly47, Ala49, and Ala52 interacted with the linker between the two rings. The identified pose 2 showed a series of interactions that have not been previously identified: the benzyl group interacts by a pi-stacking interaction with Phe31. This pose was stabilized by several van der Waals contacts. The 2-pyridone moiety showed a series of contacts different from pose 1 (Lys33, Leu45, Ser46, Gly47, Ser48, Ala49, Ala52, Ser129, and Ser168). Ser21, Phe31, Ser46, and Gly47 interacted with the ethylene linker. It is worthy to note the absence of H-bonds in this pose. Pose 3 was characterized by the same H-bonds network observed in pose 1, with Ser21 and Gly47 residues. The 2-pyridone moiety formed one cation-pi-stacking interaction with the epsilon amino group of Lys33. As observed for the other poses, van der Waals contacts strengthened the ligand binding in pose 3. The benzyl group interacted with Ser21, Ala22, Leu45, and Ser46. The 2-pyridone moiety showed contacts with several residues: Phe31, Lys33, Gly47, Ser48, Ala49, and Ala52. The linker

showed interactions with Thr1, Ser46, Gly47, and Ser168. The major differences observed for these poses concerned the orientations of pose 1 and pose 3 related to their interactions with the residues of the binding site. Besides the same H-bond pattern, a flipped orientation of the 2-pyridone and the benzyl moieties was observed. This evidence could reveal that the entrance mode of the ligand occurred in different ways without affecting the binding capability during the MD runs. The folded conformation assumed by the ligand in pose 2 seemed to represent an intermediate conformation. (Figure 18).

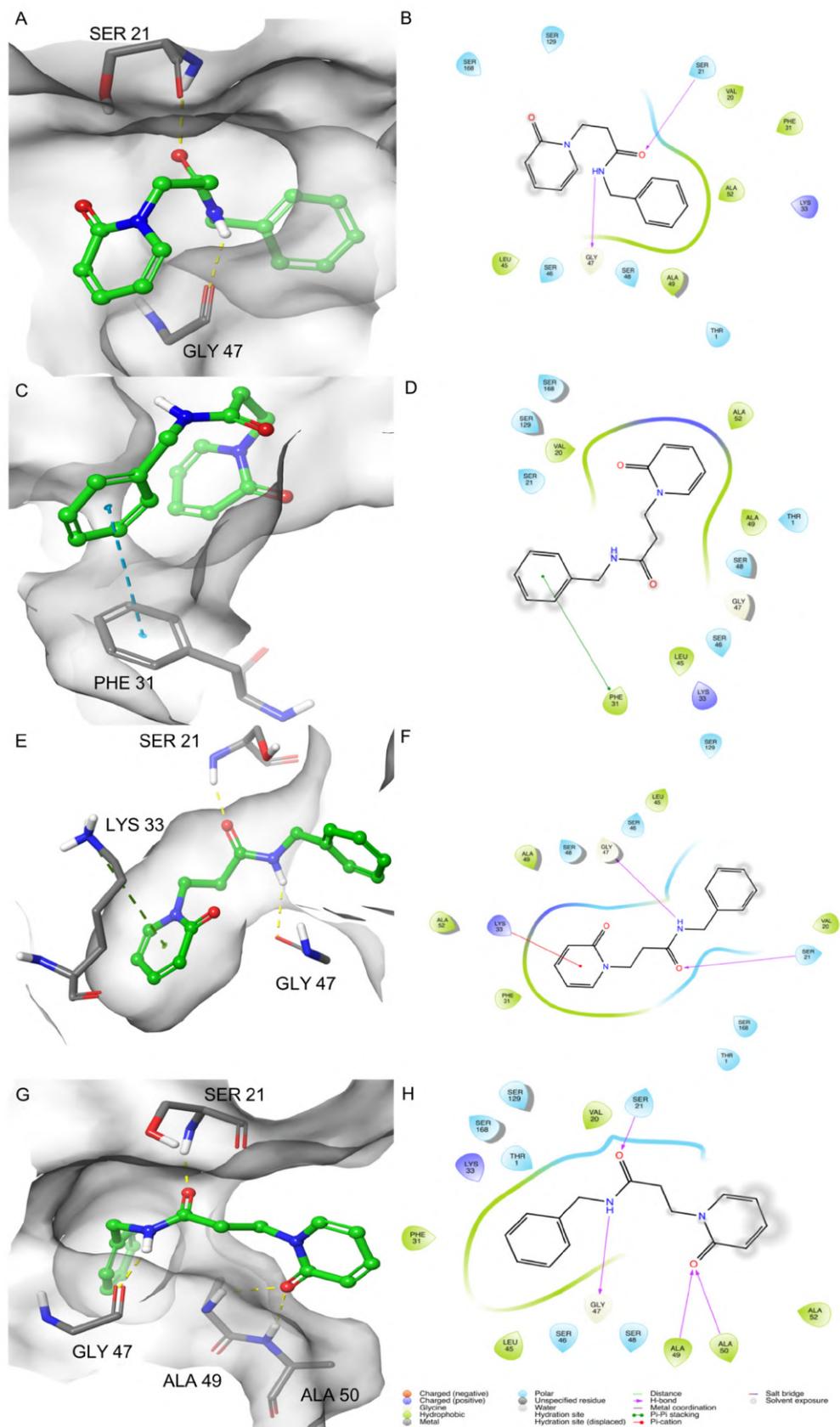


Figure 18: 3D and 2D-binding modes of compound 1: pose 1 (A,B), pose 2 (C,D), and pose 3 (E,F) after the MDB simulations and after IFD (G,H) into the β 1i active site of murine immunoproteasome (PDB ID: 3UNF). In the 3D figures, the H-bonds are represented in yellow dashes, the cation-pi-stacking interactions in green dashes, and the pi-pi stacking in blue dashes

3.4.3.2 Induced Fit Docking (IFD)

The protocol induced-fit docking (IFD) [24] was used to add more information concerning the previous docking studies and compare the results obtained by the MDB. IFD was carried out for compound 1 and for the other encouraging four β 1i inhibitors 2–5. Classical docking was performed for compound 2, while computational studies have not been performed yet for compounds 3–5.

The best IFD pose of 1 reported the same of the MD simulations: the residues Gly47 and Ser21 of the protein establish H-bonds with the NH amide and the carbonyl group of the molecule (Figure 18 G,H). Besides, other hydrogen bonds were found. The carbonyl of 2-pyridone moiety formed other H-bond with Ala49 and Ala50. Concerning the other MDB poses, this was a peculiar difference of the IFD pose that was not observed in the docking study. These residues, together with Ser48, stabilized the ligand binding by van der Waals contacts, such as observed in the previous docking study. Other van der Waals contacts were formed between the benzyl group and Val20, Phe31, Lys33, Leu45, Gly47, and Ala52. Finally, the ethylene linker between the rings interacted with Thr1, Val20, Ser21, Gly47, and Ala49. It is interesting to note that the pi-stacking interactions observed in pose 2 between the benzyl group and Phe31 and the cation-pi-stacking interaction in pose 3 between the 2-pyridone and Lys33 were not evidenced in the IFD pose but only as van der Waals contacts (Figure 18).

Structurally, Compound 2 showed a methylene linker between the 2-pyridone scaffold and the amide function. The best IFD pose of compound 2 showed three H-bonds: Ser21 with the carbonyl of amide, Gly47 with the NH amide and the carbonyl of 2-pyridone. The benzyl moiety of the molecule formed a cation-pi-stacking interaction with Lys33, as also evidenced for 1 in pose 3 (Figure 19 A,B). The cyclohexyl derivative 3 formed four H-bonds: Thr1 establish two H-bonds with the carbonyl of amide and the carbonyl of 2-pyridone; Gly47 formed two H-bonds with NH amide and carbonyl of 2-pyridone (Figure 19 C,D). The best IFD pose of n-butyl derivative 4 was characterized by two H-bonds between the carbonyl and NH of the amide of the molecule with Ser21 and Gly47, respectively. The 2-pyridone moiety formed pi-pi stacking with the Phe31 (Figure 19 E,F). Compound 5 showed two H-bonds, one between Ser21 and carbonyl of amide and the other between Ala49 and carbonyl of 2-pyridone (Figure 19 G,H). Additionally, for these molecules, the recurrent interactions were between the residues Ser21, Gly47, and the amide group, but the pi-stacking interactions with Phe31 and Lys33 could constitute clear evidence of the key role of these residues in the inhibition pattern.

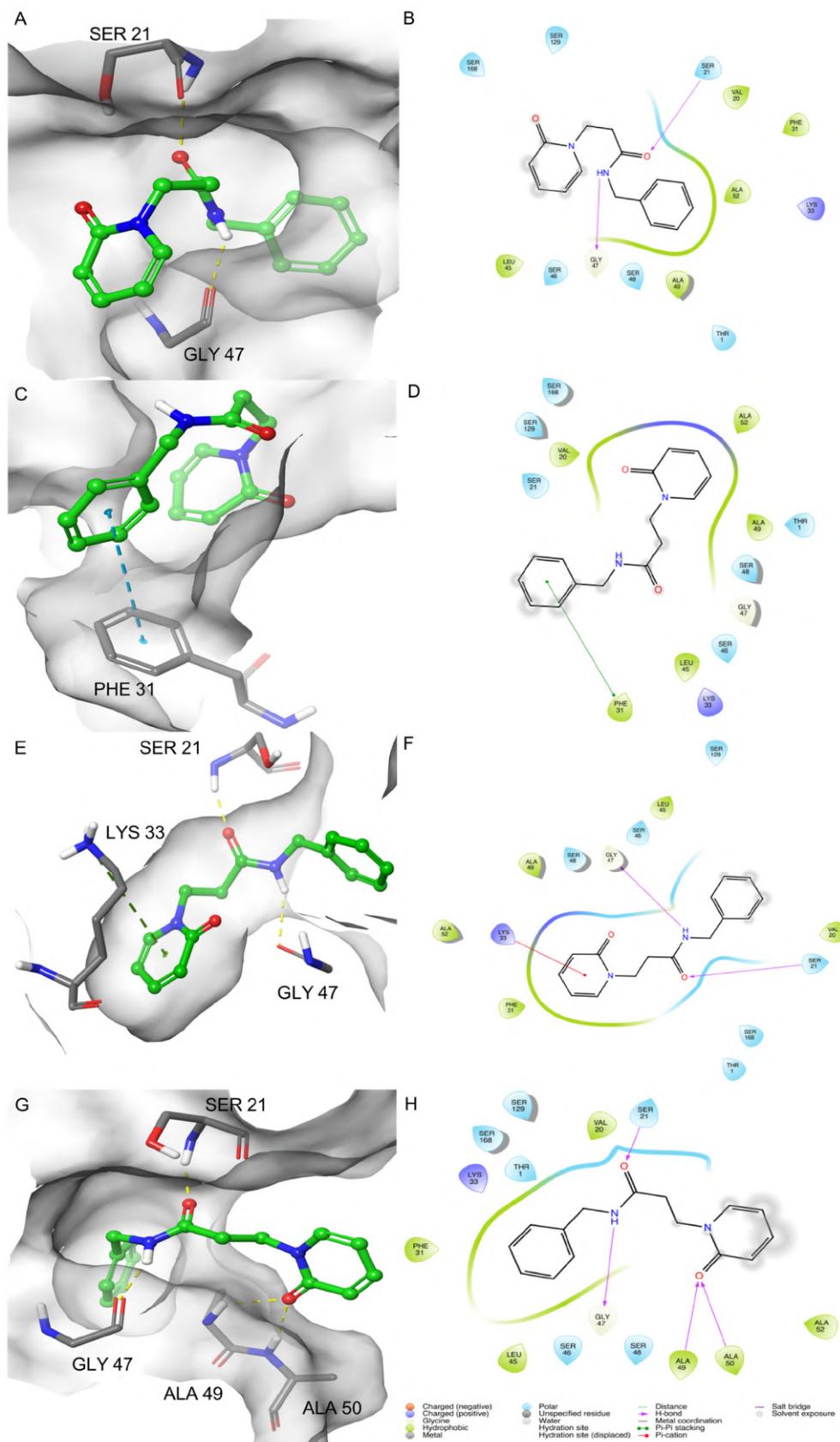


Figure 19: 3D and 2D-binding modes of compound 1: pose 1 (A,B), pose 2 (C,D), and pose 3 (E,F) after the MDB simulations and after IFD (G,H) into the β 1i active site of murine immunoproteasome (PDB ID: 3UNF). In the 3D figures, the H-bonds are represented in yellow dashes, the cation-pi-stacking interactions in green dashes, and the pi-pi stacking in blue dashes

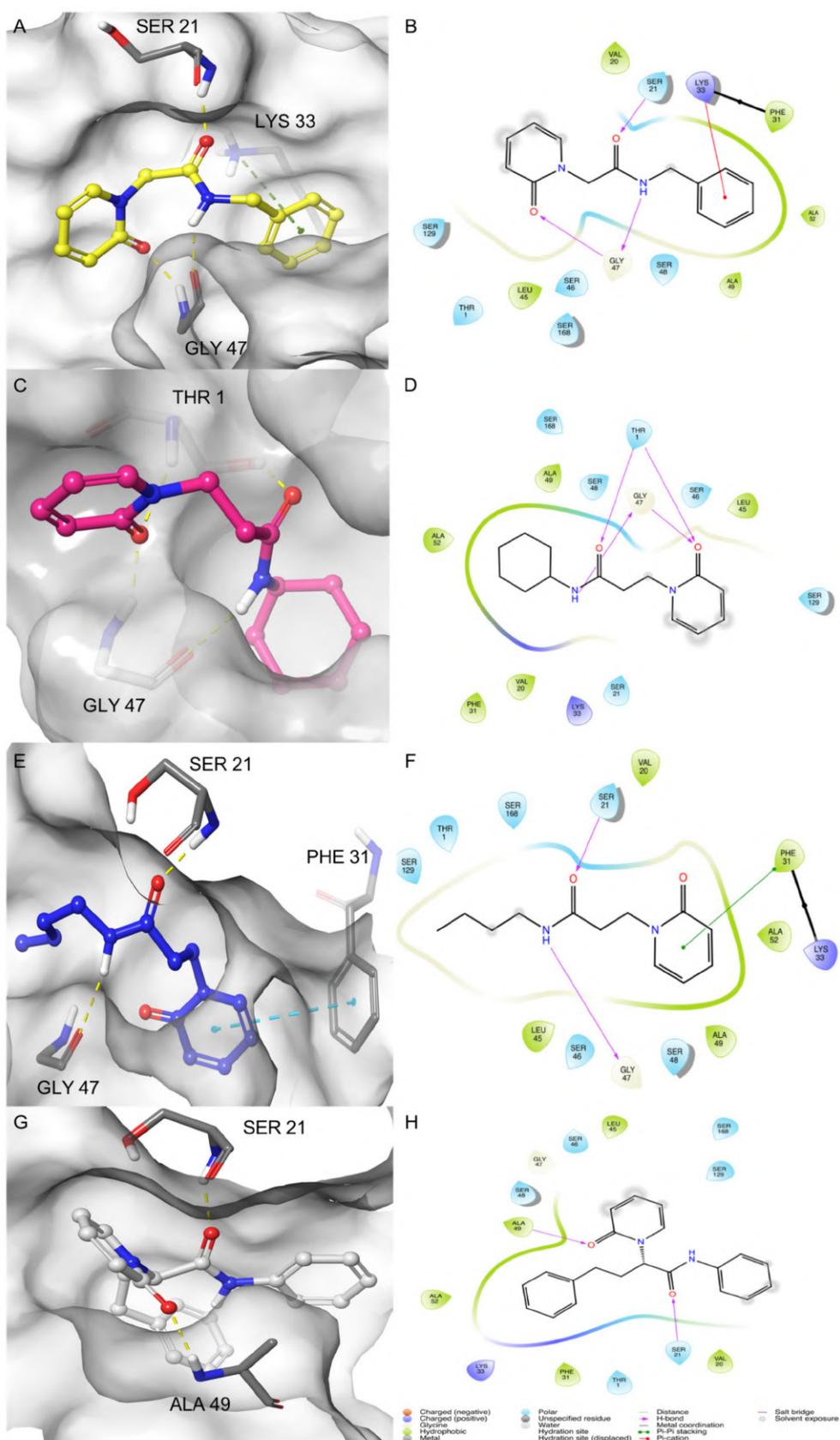


Figure 20: 3D and 2D-binding modes of compound 2 (A,B), compound 3 (C,D), compound 4 (E,F), and compound 5 (G,H) into the $\beta 1i$ active site of murine immunoproteasome (PDB ID: 3UNF) after the IFD study. In the 3D figures, the H-bonds are represented in yellow dashes, the cation-pi-stacking interactions in green dashes, and the pi-pi stacking in blue dashes

3.4.3.3 Binding Pose MetaDynamics Analysis

Binding Pose MetaDynamics (BPMD) is an automated, enhanced sampling, metadynamics-based protocol, in which the ligand is forced to move around its binding pose.

During the metadynamics simulations, pose 1 reached a steady PoseScore = 1.397, considered stable. The PersScore showed that the hydrogen bonds, identified when the metadynamics run started, were kept for 60% of the averaged time (Figure 21A). The H-bond between the NH amide group of the ligand and Gly47 was kept for 88% of the simulation time, while the H-bond between the carbonyl of the ligand and Ser21 for 36% (Figure 21B). The CompScore value was -1.694. Due to the absence of H-bonds recorded, pose 2 showed the same value for the PoseScore and CompScore, 3.129 (Figure 21C). The scores for pose 3 were PoseScore = 3.349, PersScore = 0.223, and CompScore = 2.235, respectively (Figure 21E). As for pose 1, pose 3 kept the H-bond between NH amide and Gly47 as 26% and 18% between carbonyl and Ser21 (Figure 21F).

The PoseScore for the pose of amide 1 obtained by the IFD was 4.576, and the PersScore showed that the hydrogen bonds were kept for 13% of the averaged time. The value of the CompScore was 3.917 (Figure 21G). It is interesting to point out that, of the four H-bonds detected by IFD, the two interactions between the amide group and Ser21 and Gly47 were maintained—in particular, the interaction between NH amide and Gly47 for 43% and 9% between carbonyl and Ser21 (Figure 21H).

The RMSD values and the percentage of the H-bonds retrieved from BPMD studies for the amide 1 in the three MDB poses and in the IFD pose showed that pose 1 could be considered more stable. Pose 1, pose 3, and the IFD pose adopted the same plain conformation and H-bonds between Ser21, Gly47, and the amide group. The differences were in the additional interactions between Ala49, Ala50, and the carbonyl of 2-pyridone, which led to a rotation of 2-pyridone, causing the ring to be specular in the IFD pose and showing a high value of RMSD (4.02 Å). The BPMD analysis was also carried out for compounds 2–5 to evaluate their binding stability with respect to the most active compound of the series, 1. The results of the BPMD calculations are reported in Figure 22. As can be evidenced from the plots, all showed PoseScore values higher than the averaged PoseScore for 1. The hydrogen bonds identified at the start of the MetaDynamics run were maintained for 10-30% of the averaged time (Figure 22B, D, F, H). The CompScore values for compounds 2–5 were 4.750, 4.276, 5.979, and 1.728, respectively

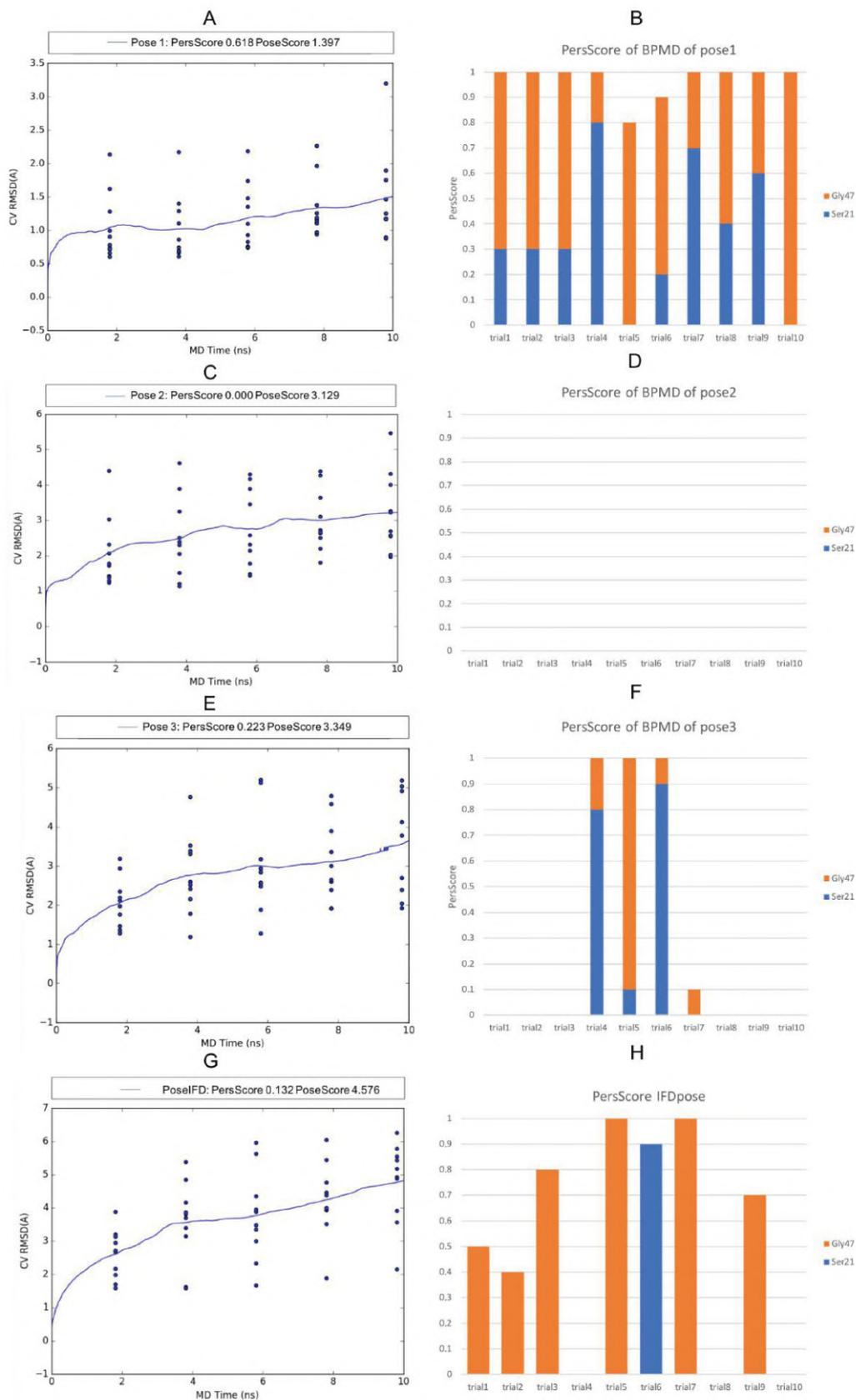


Figure 21: Plots of the RMSD estimate averaged over all 10 trials vs. the simulation time for the Binding Pose MetaDynamics runs: pose 1 (A), pose 2 (C), pose 3 (E), and IFD pose (G). Persistence Score Plots: pose 1 (B) pose 2 (D), pose 3 (F), and IFD pose (H). The blue dots represent the values of the CV RMSD at different times (2 ns, 4 ns, 6 ns, 8 ns, and 10 ns) for each simulation trial. The blue lines represent the mean CV RMSD values along the 10×10 ns of the simulation trials. The orange and blue bars represent the fraction of H-bonds maintained during the simulation for each trial

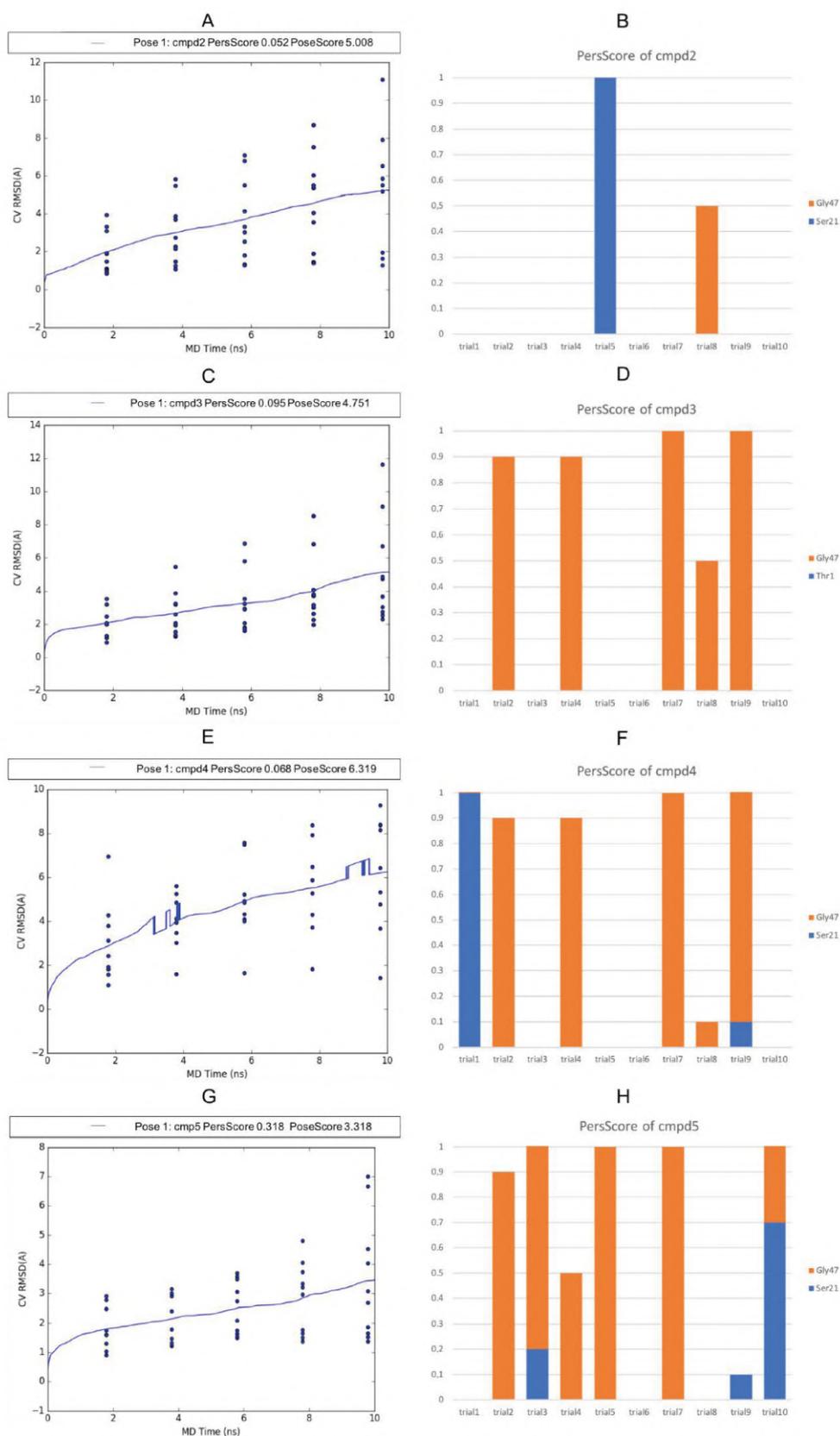


Figure 22: Plots of the RMSD estimate averaged over all 10 trials vs. the simulation time for the Binding Pose MetaDynamics runs: compound 2 (A), compound 3 (C), compound 4 (E), and compound 5 (G). Persistence Score plots of compounds 2 (B), compound 3 (D), compound 4 (F), and compound 5 (H). The blue dots represent the values of the CV RMSD at different times (2 ns, 4 ns, 6 ns, 8 ns, and 10 ns) for each simulation trial. The blue lines represent the mean CV RMSD values along the 10 × 10 ns of the simulation trials. The orange and blue bars represent the fraction of H-bonds maintained during the simulation for each trial

3.4.3.4 MM-GBSA-Binding Free Energy Calculations

Moreover, MM-GBSA-binding free energy calculations for all the complexes were performed. The plot of the calculated ΔG binding vs. the K_i values is reported in Figure 23, and it shows an $R^2 = 0.8071$ (compound 1 $\Delta G = -52.912$ Kcal/mol, compound 2 $\Delta G = -41.684$ Kcal/mol, compound 3 $\Delta G = -41.355$ Kcal/mol, compound 4 $\Delta G = -36.701$ Kcal/mol, and compound 5 $\Delta G = -35.340$ Kcal/mol).

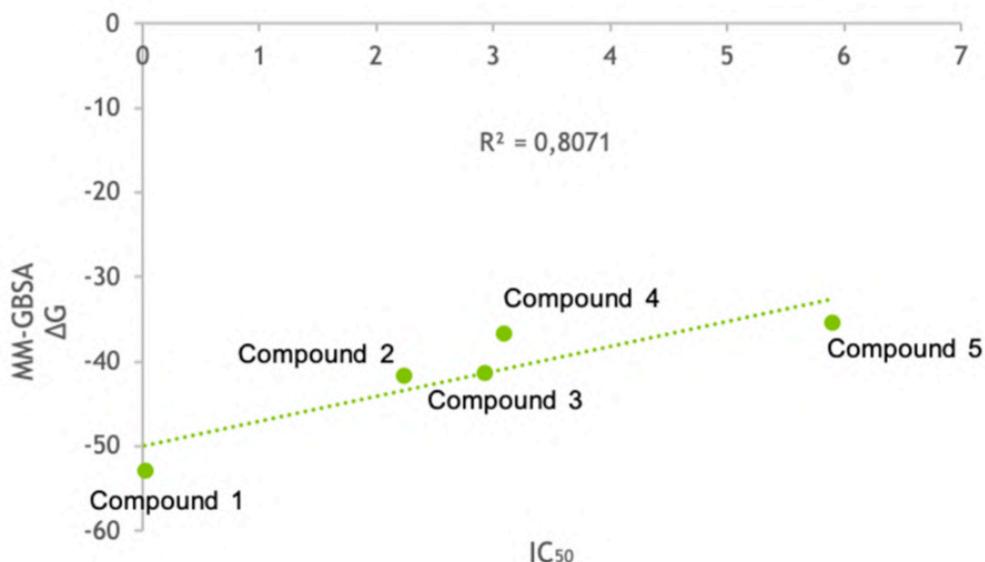


Figure 23: The plot of the calculated ΔG binding vs. the IC_{50} values of compound 1-5

3.4.4 CONCLUSION

In this work, the mechanism of non-covalent inhibition on the $\beta 1$ subunit of the immunoproteasome of the encouraging compound 1 was investigated. Advanced MD methods such as MD binding (MDB) and Binding Pose MetaDynamics (BPMD) and advanced docking methods such as induced-fit docking (IFD) were employed. MD binding allowed analyzing the binding mechanisms and gaining insights into the ligand entrance pathway. Then, plain MD was performed to study the stability and conformational space in the immunoproteasome–ligand complex, thus allowing elucidation of the compound dynamic behavior within the binding cavity. These results were compared with the IFD poses of the other four inhibitors, revealing new key residues in the binding pattern, and confirmed the different binding stability of 1 with respect to the other compounds, 2–5. The collected information and outcome could be to provide a dynamical point of view for the definition of the pharmacophoric features that could be exploited through dynamic pharmacophore modeling approaches for the scaffold-hopping of new non-covalent inhibitors through a virtual screening campaign.

3.5 COMBINED COMPUTATIONAL APPROACHES FOR VIRTUAL SCREENING OF NON-COVALENT β 1I AND β 5I INHIBITORS OF THE IMMUNOPROTEASOME

3.5.1 INTRODUCTION

In the previous chapter, the mechanism of the non-covalent inhibition of compound 1, and derivatives on the β 1 immunoproteasome subunit was investigated using advanced methods of MD and docking.

In this chapter the results of a virtual screening campaign on commercial libraries to identify novel non-covalent inhibitors of the β 1i/ β 5i immunoproteasome subunits are reported.

For the β 1i subunit, the outcomes of the previous dynamic pharmacophore model have been employed.

For β 5i, docking and pharmacophore models have been developed starting from the β 5i subunit co-crystallized with the PR-957 inhibitor (PDB ID: 3UNF).

The interactions of the hit compounds were refined by using Induced Fit Docking (IFD). Finally, hit compounds moved to biological evaluations.

3.5.2 MATERIALS AND METHODS

3.5.2.1 Protein Preparation, Ligands Preparation

The aminoacid chain K of the β 5i subunit with the co-crystallized ligand PR-957 was obtained from the crystallographic structure of 20S mouse immunoproteasome resolution 2.90 Å (PDB ID: 3UNF). The structure was refined using the protein preparation wizard as reported in the previous chapter. The co-crystallized covalent ligand, PR-957, it was rebuilt after removing the covalent inhibitor by breaking the covalent bond and filling in the open valence [89]. PR-957 was prepared and optimized as reported in the previous chapter.

A validation dataset containing active and decoys compounds of β 1 and β 5 for virtual screening were prepared. Active compounds were retrieved from the literature for both subunits, while decoys were obtained from the DUD-E [90] database and filtered to remove duplicates. For each active compound in the dataset, 50 decoys were generated. The commercial databases from Asinex, BioAscent and the database Ri.Med, containing 327.192 compounds, were downloaded in SDF format. They were prefiltered to remove groups classified as Tox-Alerts, Pan Assays INterference compoundS (PAINS) and Rapid Elimination Of Swills (REOS) that can give false positives due to non-specific reactions during the test [91, 92].

The validation databases and VS database were prepared and optimized as reported in the previous chapter.

3.5.2.2 Pharmacophore modellings generation

LigandScout [93] and PHASE [94] were used to generate dynamic pharmacophore models for β 1i and structure-based pharmacophore model based on the experimental structure of the β 5i subunit.

The models were generated using the following merged pharmacophore creation settings: feature tolerance scale factor: 1.0; Maximum number of result pharmacophores: 10; Number of omitted features for merged pharmacophore: 4.

Structure-based pharmacophore model for the β 5i subunit was generated, starting from the PDB coordinates set of β 5i subunit-PR-957.

The resulting models were validated for their performance in distinguishing the active and decoy molecules using specific databases for β 1i and β 5i.

A consensus score was calculated to evaluate the best pharmacophore performance. It combined the normalized scores obtained from the Pharmacophore-Fit Score from LigandScout and the Phase-Fit Score from Phase, as shown in the Eq. 2.

$$\text{Consensus Scoring} = \left\{ \left[\left(\frac{\text{LigandScout Score}}{100} \right) \times 100 \right] + \left[\left(\frac{\text{Phase Score}}{3} \right) \times 100 \right] \right\} \quad (2)$$

3.5.2.3 Molecular docking

The docking study was performed using the Glide docking tool v. 7.9. [95]

The grid box was set on the PR-957 ligand coordinates. The Van der Waals radius was 1 Å, and the partial charge cut-off was 0.15 Å with flexible ligand sampling. Bias sampling torsion penalization for amides with non-planar conformation and Epik state penalties were added to the docking score. To validate the docking protocol and evaluate the ability to reproduce the experimental pose of the co-crystallized ligand, the re-docking test was performed with a SP level precision. The validation of the docking is carried out through the calculation of the RMSD of the cognate ligand that provides indications of the deviation between the experimental pose of the ligand and that reproduced by the docking algorithm. The reference value for RMSD is <2.0 for the scoring function.

The Virtual Screening was performed in three different steps of docking: High Throughput Virtual Screening (HTVS), Standard Precision (SP), and Extra Precision (XP). Each step of docking was run without constraints.

3.5.2.4 Induced Fit Docking (IFD)

Induced Fit Docking (IFD) was performed using the standard protocol, reported in the previous chapter. As in the docking study, the receptor box for β5i was defined on the PR-957 applying no constraints.

The receptor box for β1i was defined on the residues Ser21, Phe31, Ser33 and Gly47. The receptor and ligand Van der Waals scaling were set at 0.5. For Prime refinement, side chains of residues within 5 Å of ligand poses were refined. The maximum of 20 poses per ligand was retained and re-docked in Extra Precision mode (XP).

3.5.2.5 *In vitro* 20S immunoproteasome inhibition assays

Human 20S immunoproteasome, obtained from the human spleen, was purchased from Enzo Life Science. The hydrolysis of the appropriate peptidyl 7-amino-4-methyl coumarin (AMC) substrate was monitored to measure the proteolytic activities of the immunoproteasome. The substrates Suc-Leu-Leu-Val-Tyr-AMC (Bachem) for β5i and Ac-Pro-Ala-Leu-AMC (Biomol GmbH, Hamburg, Germany) for β1i subunits were used at 50 μM. Fluorescence of the hydrolysis product AMC was measured at 30°C with a 380 nm excitation filter and a 460 nm emission filter using an Infinite 200 PRO microplate reader (Tecan, Männedorf, Switzerland). A preliminary screening at 100 μM inhibitor concentrations was carried out on the two proteolytic activities of the immunoproteasome.

An equivalent amount of DMSO as a negative control and compound 1 and MG-132 (a reversible immunoproteasome inhibitor) as a positive control were used for β1i and β5i, respectively. Compounds showing at least 50% inhibition at the screening concentration were then progressed into detailed assays. Continuous assays were performed at seven different concentrations ranging from minimally inhibited to fully inhibiting each proteolytic activity by calculating the dissociation constants K_i of the enzyme–inhibitor complex using the Cheng–Prusoff equation, $K_i = IC_{50} / (1 + [S]K_M^{-1})$. Inhibitor solutions were prepared from stocks in DMSO. Each independent assay was duplicated in 96-well plates with a total volume of 200

mL. For the assay on β 1i and β 5i subunits, a human 20S immunoproteasome was incubated at 30°C at a final concentration of 4 μ g/ml with the inhibitor at seven different concentrations. The reaction buffer comprised: Tris·HCl (50 mM, pH 7.4), KCl (25 mM), NaCl (10 mM), MgCl₂ (1 mM), and 0.03 % SDS. AMC released from substrate hydrolysis was monitored in a kinetic cycle for 10 minutes.

3.5.3 RESULTS AND DISCUSSION

3.5.3.1 β 1i pharmacophore modeling

The previous MD study allowed identifying three representative poses (pose 1, pose 2, and pose 3) that elucidated the dynamic behavior and stability of the compound 1 within the binding cavity. The inhibition interactions pattern comprises two H-bonds between the amide group and Ser21 and Gly47, the benzyl group interacting pi-pi stacking with Phe31; the 2-pyridone moiety made a cation-pi-stacking interaction with the epsilon amino group of Lys33. Starting from these outcomes, a merged dynamic pharmacophore model was constructed and validated for the β 1i subunit.

For each pose, pharmacophore models using LigandScout and PHASE were generated. In LigandScout, pose1 showed three features: one hydrogen bond acceptor (HBA), one hydrogen bond donor (HBD) and hydrophobic interaction (Figure 24A). Pose2, one hydrogen bond acceptor (HBA) and hydrophobic interaction (Figure 24B). Pose3 was characterized by one hydrogen bond acceptor (HBA), one hydrogen bond donor (HBD) and two hydrophobic interactions (Figure 24C). A merged pharmacophore model encompassing all features observed for each pose was constructed and improved by removing the redundant characteristics. The final model was characterized by five features: two hydrogen bond acceptors (HBA), one hydrogen bond donor (HBD), and two hydrophobic interactions (Figure 24D). The model was retrospectively validated using the active and decoys dataset showing values of $AUC_{1\%} = 1$ and $EF_{1\%} = 15.3$.

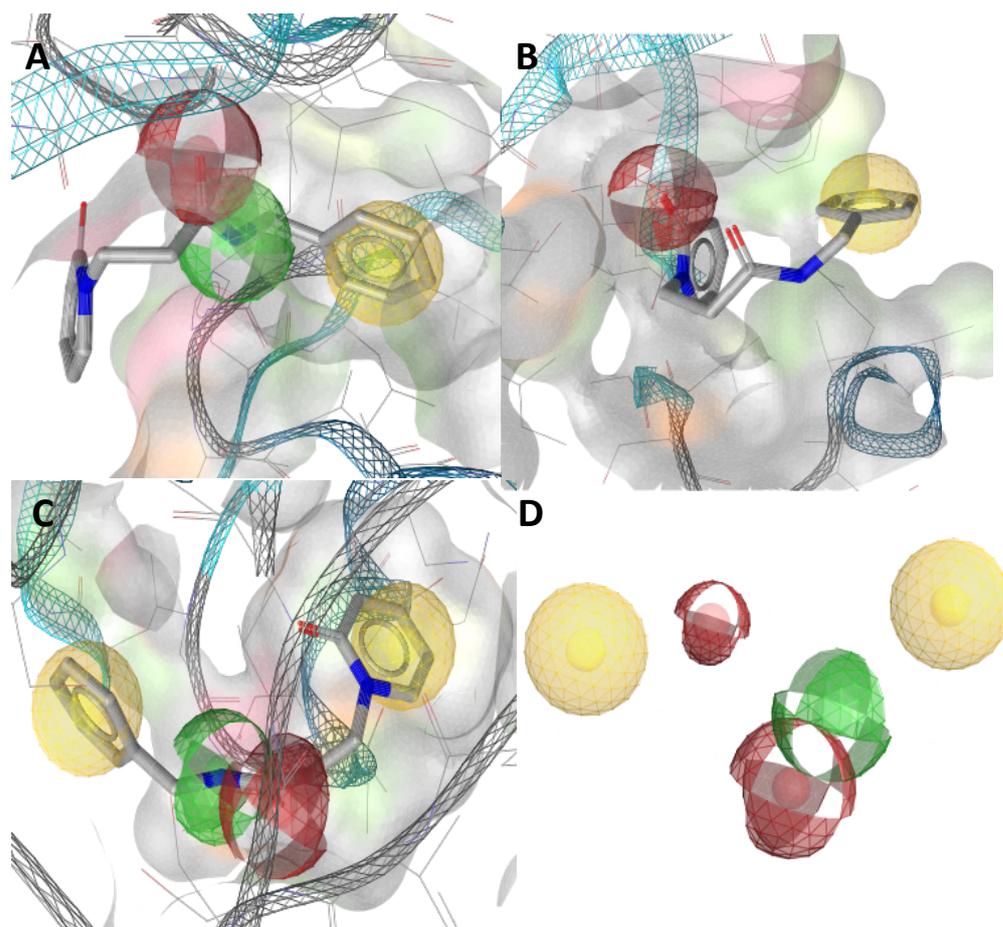


Figure 24: LigandScout pharmacophore model of A) Pose1, B) Pose2, C) Pose3 and D) Merged

The same procedure was carried out using PHASE. Pose1 showed three features: two hydrogen bond acceptors (A1 and A2), one hydrogen bond donor (D3) and two aromatic rings (R4 and R5) (Figure 25A). In Pose2 was retrieved an aromatic ring (R5) (Figure 25B). Pose3 was characterized by one hydrogen bond acceptor (A2), one hydrogen bond donor (D3) and aromatic ring (R4) (Figure 25C). Then, a merged model was obtained by showing the same features observed in the LigandScout model.

As in the LigandScout merged pharmacophore model, the PHASE merged pharmacophore model encloses all features observed in the three poses. Two hydrogen bond acceptors (A1 and A2), one hydrogen bond donor (D3) and two aromatic rings (R4 and R5) (Figure 25D).

As the previous model, this model was retrospectively validated and showed values of $AUC_{1\%} = 0.60$ and $EF_{1\%} = 4.9$.

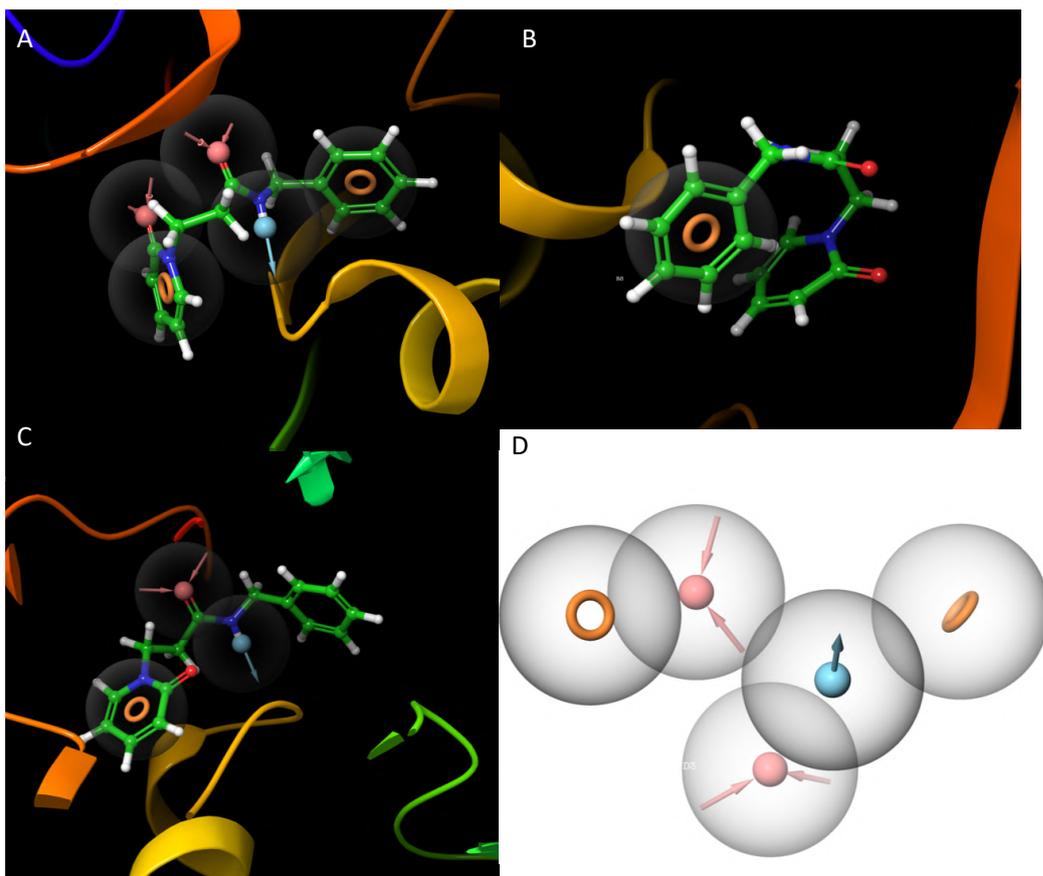


Figure 25: Phase pharmacophore model of A) Pose1, B) Pose2, C) Pose3 and D) Merged

3.5.3.2 β 5i docking and pharmacophore modeling

For the β 5i subunit, structure-based studies based on the experimental structure PDB ID: 3UNF were carried out. Firstly, the β 5i subunit-PR-957 complex was extracted from the crystal immunoproteasome structure and refined. With the aim to design non-covalent inhibitors, the PR-957 inhibitor was structurally modified and the portion of the epoxy ring responsible for the covalent attack on Thr1 was opened (Figure 26 A, B)

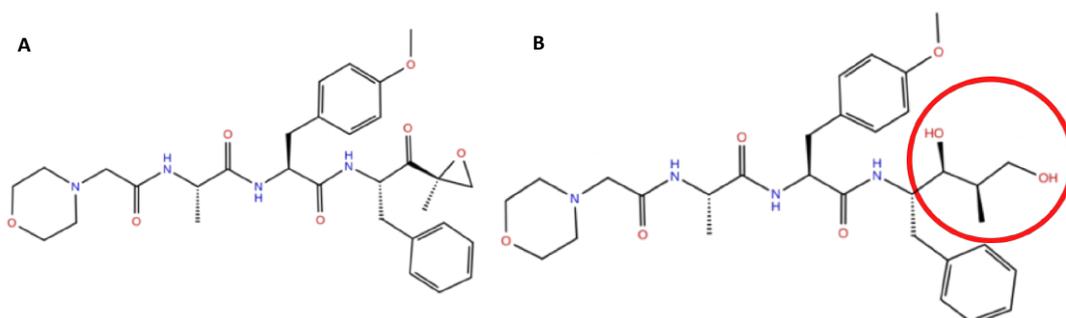


Figure 26: A) PR-957 covalent epoxyketone, B) Modified PR-957, free of the epoxy ring portion (circled in red)

A re-docking with an XP precision level was performed to validate the ability to reproduce the experimental pose of the co-crystallized ligand. The experimental and docking laying just differ in the epoxy portion, which is responsible for the covalent interactions. The root-mean-square deviation (RMSD) value between the re-docked and the crystal conformation of PR-957 was 1.77 Å. The RMSD value less than 2 Å indicates that Glide XP is a reliable method for this study. The docking protocol was further validated using a β 5i validation database of active and decoys. The model showed values of $AUC_{1\%} = 0.90$ and $EF_{1\%} = 27$. The Glide XP energetic terms were calculated using structural and energy information between the PR-957 and the active site were mapped onto pharmacophore sites to develop an e-pharmacophore hypothesis. The pharmacophore features with an energetic value <0.5 kcal/mol were retained and used to compose the pharmacophore hypothesis. The e-pharmacophore hypothesis comprises one hydrogen bond acceptor (A7), two hydrogen bond donors (D8 and D11), and two aromatic rings (R17 and R18). The energy values for the favorable features in the hypothesis are: A3 -0.63 kcal/mol, D8 -1.60 kcal/mol, D11 -0.70 kcal/mol, R17 -0.52 kcal/mol and R18 -1.60 kcal/mol (Figure 27A). The higher absolute value of the feature energy indicates that ligand atom mapping exhibits more potent interaction energy with amino acids. In Figure 27 are also reported the ROC curves for the e-pharmacophore hypothesis with 0 and 1 features omitted

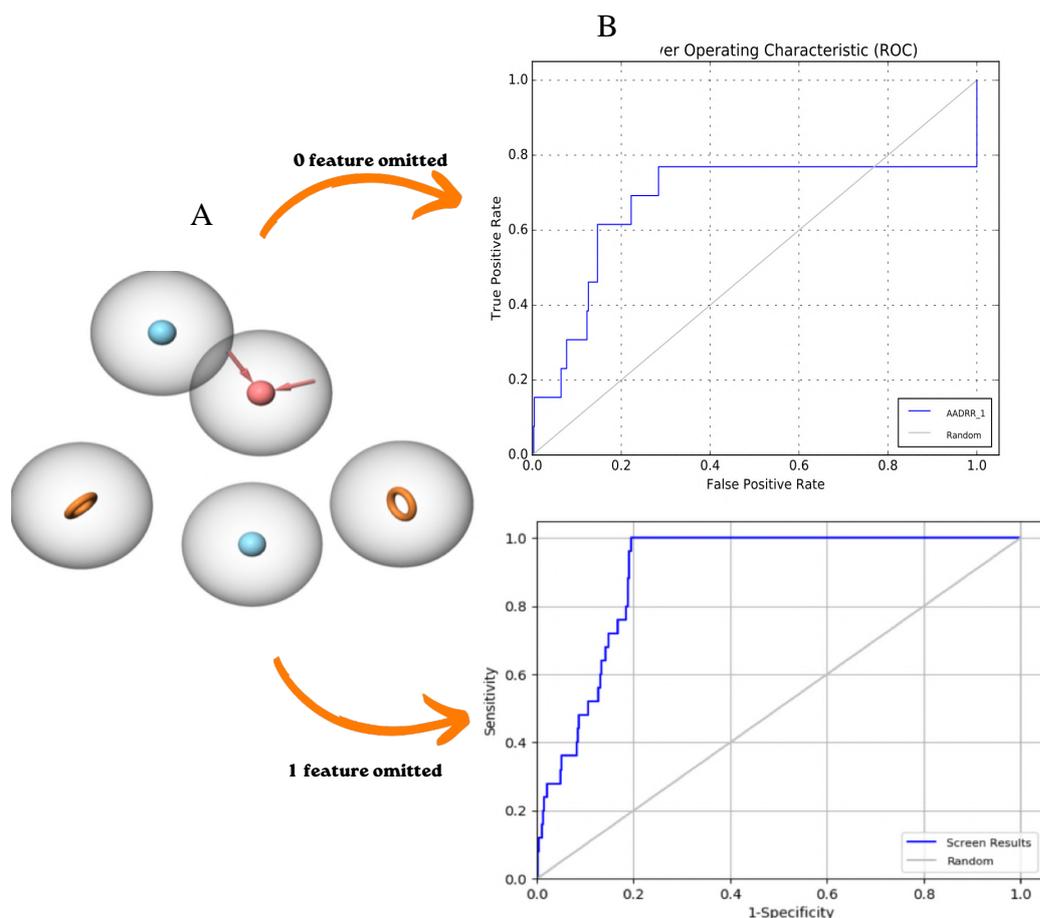


Figure 27: A) E-pharmacophore hypothesis for $\beta 5i$, five features: two Aromatic Rings (Orange), two Hydrogen Bond Donator (Blue) and an Hydrogen Bond Acceptor (Red). B) ROC curves, zero feature omitted and one features omitted

The validation results of the e-pharmacophore hypothesis are presented in Table 1.

Table 1: Matching parameters of the e-pharmacophore

Features omitted	AUC _{1%}	E.F _{1%}
0	0.40	20
1	0.84	8

LigandScout was also used to build a static pharmacophore starting from the the PR-597 docked pose. A 6-characteristic model has been obtained: three hydrogen bond acceptors (HBA), two hydrogen bond donors (HBD) and one aromatic interaction (Figure 28).

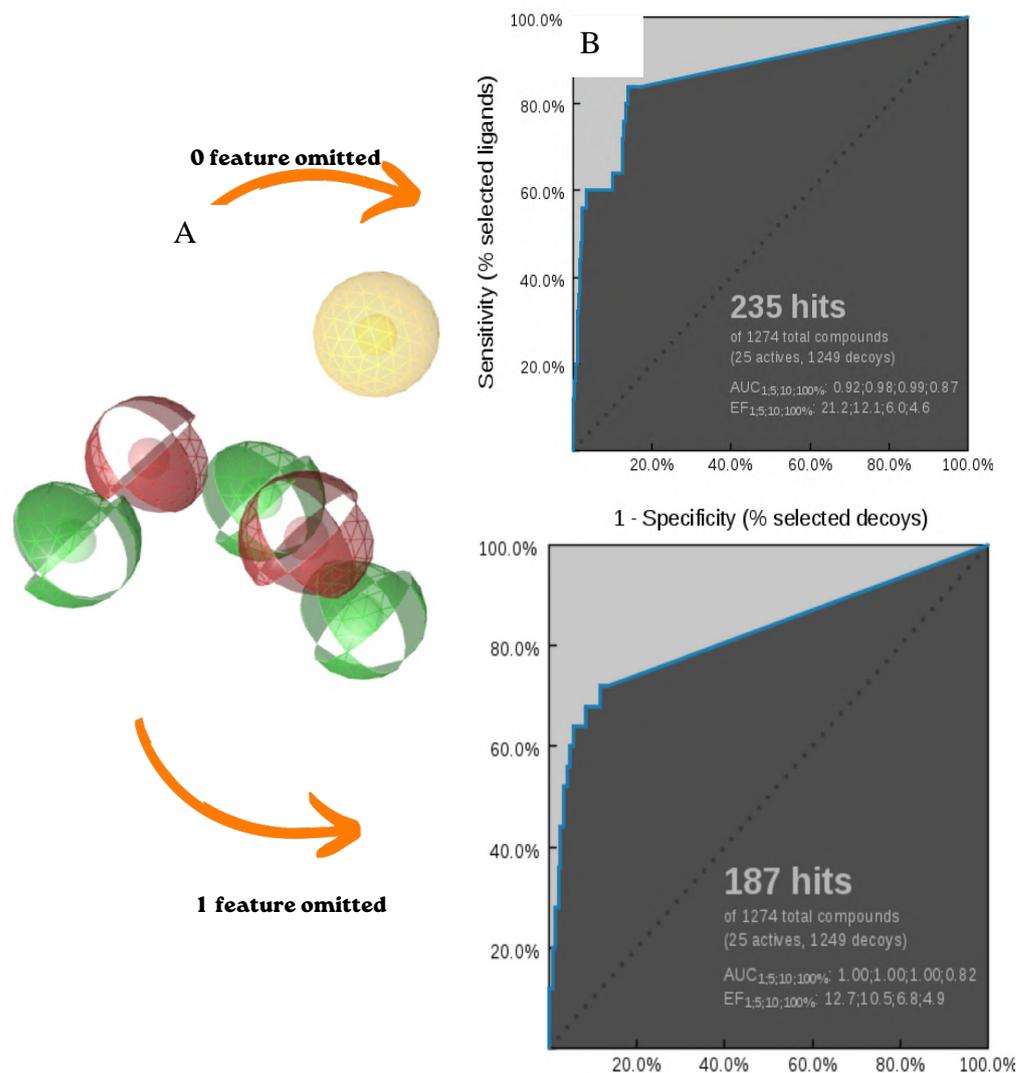


Figure 28: LigandScout model for $\beta 5i$, six features: an Aromatic Interaction (Yellow), three Hydrogen Bond Donator (Green) and two Hydrogen Bond Acceptor (Red). B) ROC curves with 0 feature omitted and one feature omitted

In Figure 28 are also reported the ROC curves for the LigandScout model with 0 and 1 features omitted. The model showed values of $AUC_{1\%} = 0.92$ and $EF_{1\%} = 21.2$ when 0 feature is omitted, and values of $AUC_{1\%} = 1.00$ and $EF_{1\%} = 12$ when 1 feature is omitted (Table 2).

Table 2: Matching parameters of the LigandScout model

Features omitted	$AUC_{1\%}$	$E.F_{1\%}$
0	0.92	21.2
1	1	12

3.5.3.3 β 1i and β 5i virtual screening

The pharmacophore models obtained from the β 1i dynamic studies were used to carry out the virtual screening on the merged commercial library (VS β 1i).

The β 5i models built with the docking scoring function and pharmacophore tools were also used to carry out the virtual screening campaign on the same merged commercial library (VS β 5i). The VS β 5i started using the HTVS, SP and XP docking screening. The 10% of the prioritized ligands were retained from screening for the steps of docking HTVS and SP. Taking into consideration that the previously validated models showed excellent AUC and EF values up to 1% of the screened library, 1% of hits obtained from XP mode were then submitted to pharmacophore matching.

The hits obtained from VS β 1i and VS β 5i were ranked according to the consensus score described in the Materials and Methods section. The consensus score improved the selection and retrieval of the molecules in common between the pharmacophore models for β 1 and β 5. Finally, the hit molecules retrieved after consensus scoring were further evaluated with Induced Fit Docking (IFD) on β 1i and β 5i subunits. IFD confers more flexibility to the protein side chains, allowing the ligands to adjust and optimize binding interactions within the active sites. A visual inspection of IFD outcomes was helpful to enrich the final selection. Thirty-six molecules (20 for β 1i, 11 for β 5i and 5 for both subunits) were selected and purchased for *in vitro* screening against β 1i and β 5i subunit using an enzymatic assay (Table 3).

Table 3: Molecules selected for the enzymatic assay on β 1i and β 5i

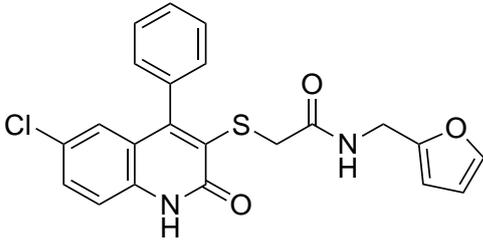
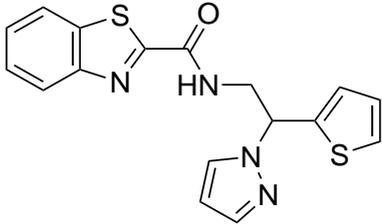
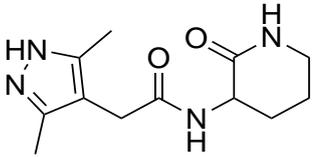
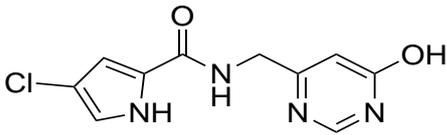
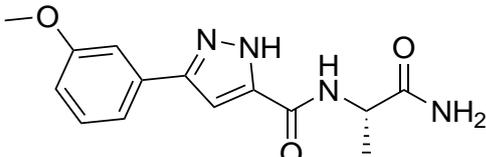
β 1i	β 5i	β 1i/ β 5i
RIM457	RIM436	RIM312
RIM1578	RIM484	RIM518
RIM530	RIM493	RIM671
RIM552	RIM496	RIM701
RIM1361	RIM505	RIM1470
RIM1436	RIM1438	
RIM1500	RIM1485	
RIM1517	RIM1584	
RIM1579	RIM1588	
RIM1580	RIM1590	
RIM1581	RIM1592	
RIM1583		
RIM1586		
RIM1587		
RIM1589		
RIM1582		
RIM1593		
RIM1358		
RIM1585		
RIM1573		

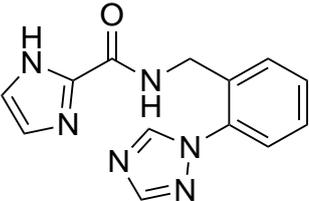
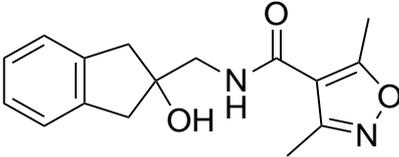
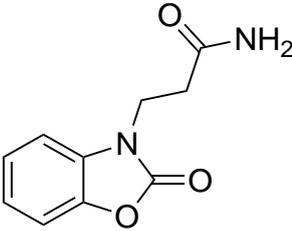
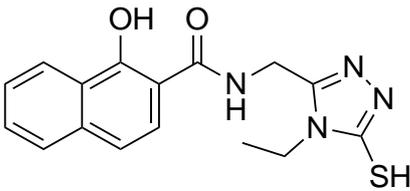
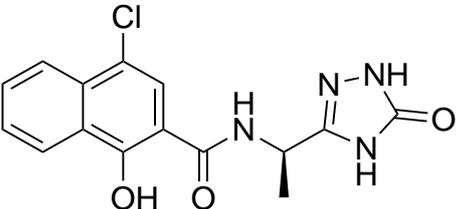
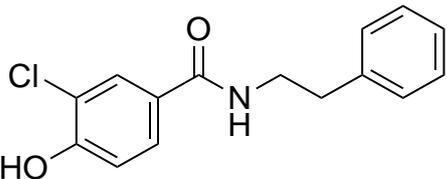
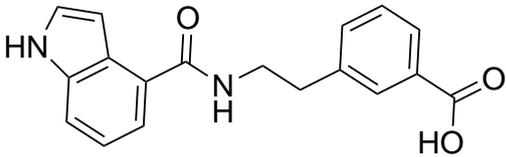
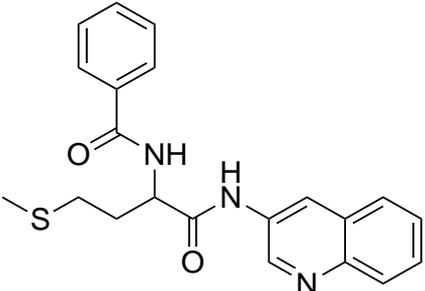
These hits were evaluated for the capability to inhibit both subunits of i20S by measuring the rate of hydrolysis of the appropriate fluorogenic substrate Ac-Pro-Ala-Leu-AMC for β 1i and (Suc-Leu-Leu-Val-Tyr-AMC for β 5i. Compound1 was used as a positive control for β 1i and MG-132 for β 5i.

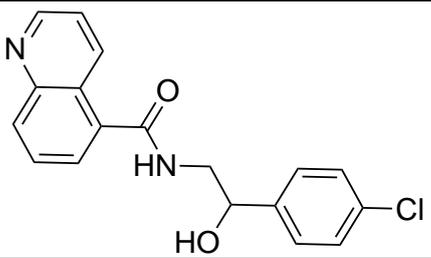
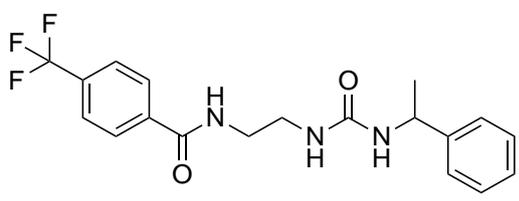
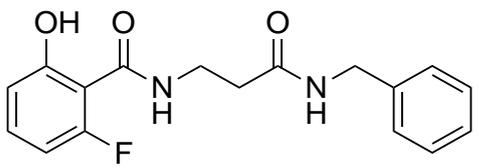
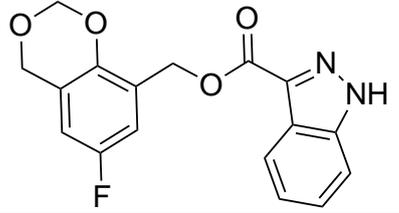
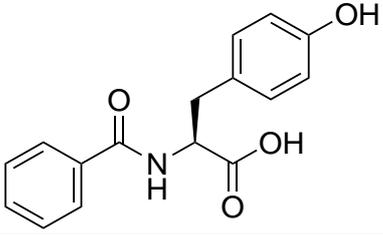
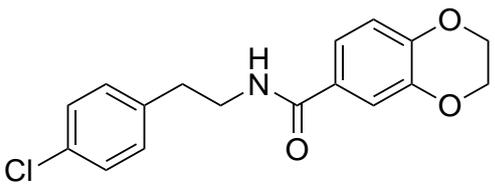
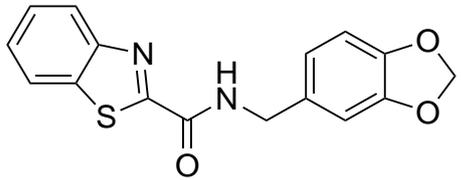
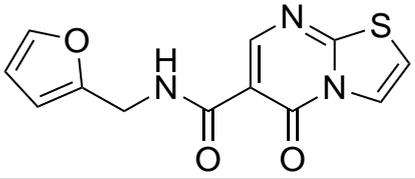
First, compounds underwent preliminary screening on each proteolytic subunit at 100 μM . Compounds able to inhibit the enzymatic activity by more than 50% were characterized in details. Continuous assays were thus performed (progress curve method, at seven different concentrations, ranging from those that minimally inhibited to those that fully inhibited the immunoproteasome subunit) to determine the K_i values (Table 4).

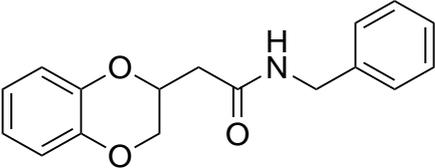
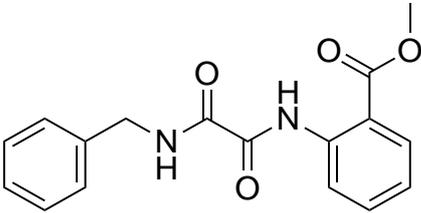
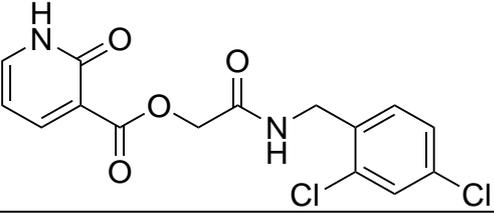
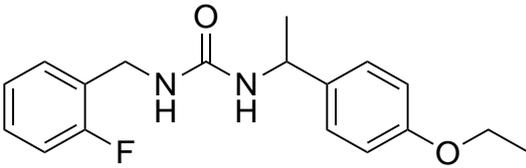
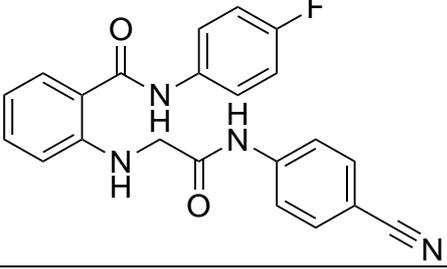
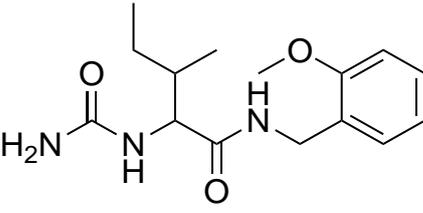
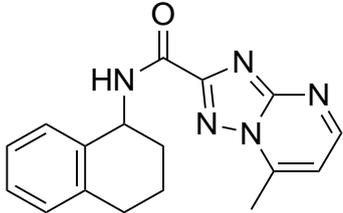
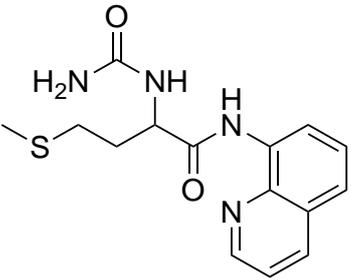
As shown in Table 4, some of the tested compounds were characterized in details (RIM312, RIM436, RIM1436, RIM1438, RIM1500, RIM1589 and RIM1593). Interestingly, as observed in the *in silico* studies, compound RIM312 inhibits both subunits ($K_i = 12.53 \pm 0.18 \mu\text{M}$ for $\beta 1i$ and $31.95 \pm 0.81 \mu\text{M}$ for $\beta 5i$). RIM436 and RIM1589 inhibit $\beta 1i$ subunits, while RIM1436, RIM1438, RIM1500 and RIM1593 inhibit $\beta 5i$.

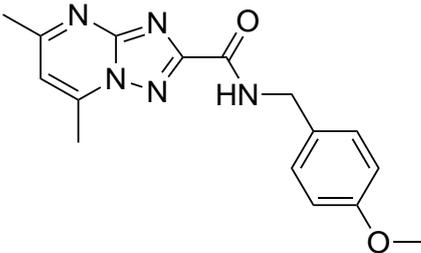
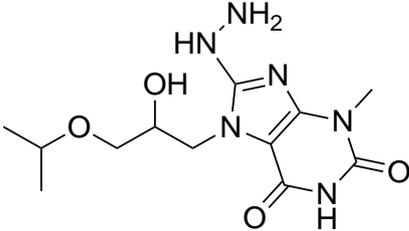
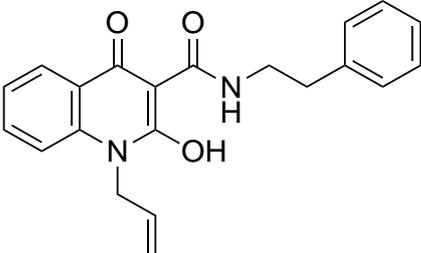
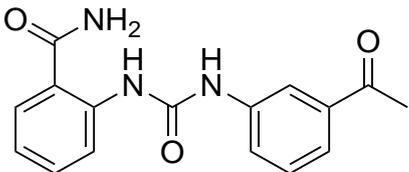
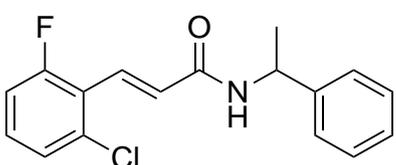
Table 4: Activity of selected compounds on $\beta 1i$ and $\beta 5i$ immunoproteasome. Activity is expressed in μM

Compound	Structure	$\beta 1i$	$\beta 5i$
RIM312		12.53 ± 0.18	31.95 ± 0.81
RIM436		11.84 ± 1.63	9%
RIM457		n.i.	23%
RIM484		n.i.	4%
RIM493		n.i.	17%
RIM496		n.i.	18%

RIM505		n.i.	18%
RIM530		24%	10%
RIM552		n.i.	28%
RIM671		u.s.	u.s.
RIM701		u.s.	u.s.
RIM1358		18%	10%
RIM1361		2%	23%
RIM1436		40%	43.68±4.02

RIM1438		47%	48.30±4.97
RIM1470		24%	23%
RIM1485		46%	22%
RIM1500		33%	12.50±0.77
RIM1517		15%	1%
RIM1573		3%	5%
RIM1578		25%	36%
RIM1579		19%	23%

RIM1580		31%	32%
RIM1581		20%	34%
RIM1582		21%	34%
RIM1583		16%	55%
RIM1584		Fluorescent	Fluorescent
RIM1585		27%	23%
RIM1587		23%	13%
RIM1588		45%	47%

RIM1589		19.53±2.83	23%
RIM1590		30%	37%
RIM1591		32%	34%
RIM1592		17%	42
RIM1593		14%	48.13±4.34
	u.s. under screening, n.i. non inhibition		

RIM312, on $\beta 1i$, showed a pi-pi stacking interaction between Phe31 and the furan aromatic ring, as observed in the previous study. A new interaction, not previously recorded, involves the Thr1 residue with the CO of the dihydroquinoline ring by a hydrogen bond (Figure 29A-B). As in $\beta 1i$, in $\beta 5i$, the furan ring showed the same pi-pi stacking interaction, but with another residue, Tyr169. Thr1 formed a hydrogen bond interaction with the NH of the amide. Finally, two hydrogen bonds were registered between the Thr21 residue with the NH group of the dihydroquinoline and the CO of the amide (Figure 29C-D).

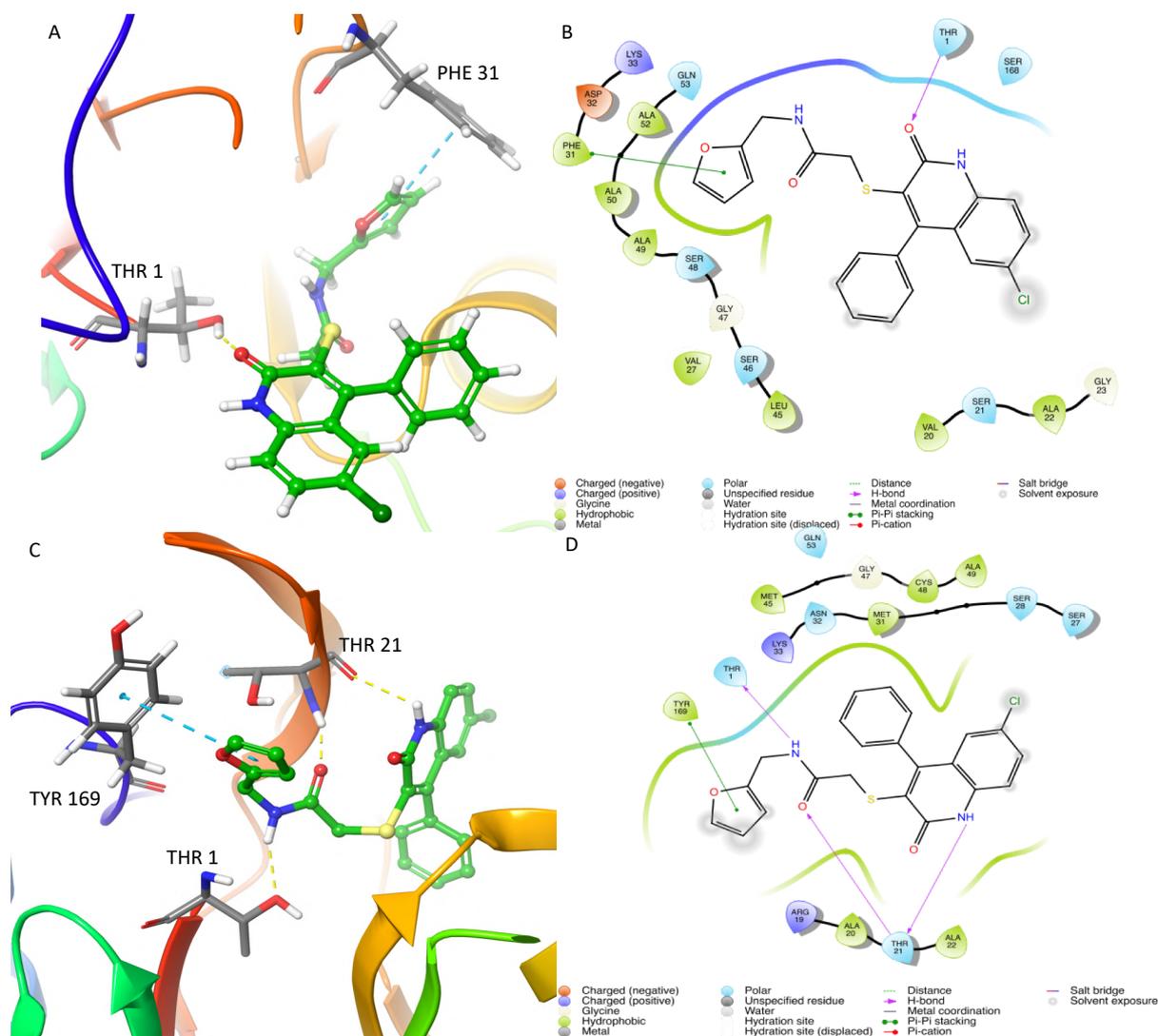


Figure 29: Poses of RIM312 on A-B) $\beta 1i$ and C-D) $\beta 5i$ subunits

3.5.4 CONCLUSION

With the aim to design novel non-covalent inhibitors for β 1i and β 5i subunits, A scaffold-hopping Virtual Screening campaign was employed.

The compounds obtained by VS were docked using Induced Fit Docking (IFD) to evaluate the interaction with the active site. Finally, 36 compounds: 20 for β 1i, 11 for β 5i and 5 in common were chosen and evaluated by biological assay.

Some of them showed activity in the micromolar range on β 1i and β 5i, in particular, as observed in the *in silico* study, compound RIM312 inhibits both subunits ($K_i = 12.53 \pm 0.18 \mu\text{M}$ for β 1i and $31.95 \pm 0.81 \mu\text{M}$ for β 5i).

The cellular inhibition assay on MM.1S and HepG2 cells overexpressing the immunoproteasome are ongoing as far as the optimization of the hits obtained to increase the activity on β 1i and β 5i immunoproteasome subunits.

3.6 SYNTHESIS OF NOVEL PROTEASOME INHIBITORS

3.6.1 INTRODUCTION

The ubiquitin-proteasome system (UPS) is a key pathway involved in the intracellular protein turnover of eukaryotic cells. It is also responsible for the regular cell progression, immune surveillance and homeostasis control. Problems related to the UPS can lead to an uncontrolled cell proliferation and to tumor development. Proteasome is also involved in the degradation of many proto-oncoproteins, which, if not removed from cells, can generate malignancies.

20S proteasome core is characterized by a barrel-like structure with four stacked rings: two inner β -rings, containing the catalytic subunits, and two outer α -rings, whose function is to maintain a gate through which proteins enter the barrel-like structure. The catalytic subunits are $\beta 1$, $\beta 2$, $\beta 5$, which are responsible for the caspase-like, trypsin-like and chymotrypsin-like activities, respectively. It is currently well demonstrated that targeting the $\beta 5$ catalytic subunit is a promising strategy to develop novel anticancer agents both for solid and hematologic tumors [96].

In order to design novel proteasome inhibitors, a series of amide derivatives **1a-h** were synthesized (Figure 30). The designed inhibitors, lacking the electrophilic portion, were designed to act as non-covalent inhibitors, which could prove to be a promising therapeutic strategy, being devoid of the undesirable effects related to irreversible inhibition.

The structure of novel amides bears a pyridone scaffold at the P3 site, as bioisostere of a leucine residue; its function is also to block the peptide in the bioactive conformation [96].

The amino group in position 5 of the pyridone scaffold has been introduced with the aim to increase the polarity of the molecules and with the idea to promote the conjugation of the most active molecules with graphene quantum dots (GQD). GQD are the next generation of carbon-based nanomaterials, endowed of size-dependent photoluminescence and, due to the many reactive oxygen functional groups, endowed of water dispersibility, thus being able to improve the solubility of poor soluble drugs, and at the same time to enhance its cell uptake.

Because of the lack of specificity at the P2 site, a glycine residue was introduced in this position, whereas the amide group at the P1 site was functionalized with hydrophobic aliphatic or aromatic substituents (a-h) for the accommodation into the S1 pocket of the $\beta 5c$ subunit that is large and hydrophobic [97].

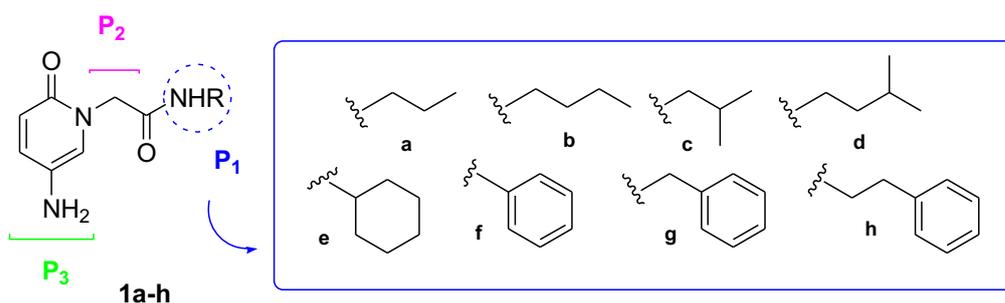


Figure 30: Design of novel proteasome inhibitors

3.6.2 MATERIALS AND METHODS

3.6.2.1 Chemistry

All reagents and solvents were purchased from commercial suppliers and used without any further purification. The reactions conducted with the aid of microwave radiation were performed with a CEM Discover apparatus. Elemental analyses were performed on a C. Erba

Model 1106 (Elemental Analyzer for C, H and N) and the results are within $\pm 0.4\%$ of the theoretical values. Merck silica gel 60 F254 plates were used for analytical TLC; column chromatography was carried out on Merck silica gel (70–230 mesh). $^1\text{H-NMR}$ and $^{13}\text{C-NMR}$ spectra were recorded on a Varian 500 MHz NMR spectrometer operating at frequencies of 499.74 and 125.73 MHz for $^1\text{H-NMR}$, and $^{13}\text{C-NMR}$ spectra, respectively. The residual signal of the deuterated solvent was used as an internal standard. Chemical shifts are given in δ (ppm) and coupling constants (J) in Hz. Splitting patterns are described as singlet (s), broad singlet (bs), doublet (d), triplet (t), quartet (q), multiplet (m), doublet of doublet (dd), or triplet of doublets (td).

Ethyl 2-(5-nitro-2-oxopyridin-1(2H)-yl)acetate (3)

A suspension of NaH at 57% (3.6 g, 0.086 mmol) in dry DMF at 0 °C under N_2 was treated with a solution of 5-nitro-2-pyridone **2** (5 g, 0.036 mmol) in dry DMF. After 1 h, ethyl bromoacetate (5.9 mL, 0.053 mol) was added to the mixture and the reaction was stirred overnight at room temperature. The reaction was quenched with saturated NH_4Cl and the organic phase was extracted with EtOAc. The combined organic phases were washed with saturated NaCl, dried over Na_2SO_4 , filtered and concentrated in vacuo. The crude was purified by column chromatography (light petroleum /EtOAc 5:5) to give compound **3** as a white oil (3.95 g, 61%); $R_f = 0.63$ (light petroleum/EtOAc 5:5); $^1\text{H NMR}$ (500 MHz, CDCl_3) $\delta = 1.26$ (t, 3H), 4.22 (q, $J = 7.1$ Hz, 2H), 4.66 (s, 2H), 6.54 (d, $J = 10.1$ Hz, 1H), 8.07 (dd, $J = 10.1$ Hz and 3.0 Hz, 1H), 8.51 (d, $J = 3.0$ Hz, 1H) ppm; $^{13}\text{C NMR}$ (75 MHz, CDCl_3) $\delta = 29.68, 50.97, 62.60, 119.61, 130.94, 133.78, 139.95, 161.05, 166.39$ ppm.

2-(5-Nitro-2-oxo-2H-pyridin-1-yl)-acetic acid (4)

A solution of ester **3** (3.95 g, 0.017 mol) in MeOH/ H_2O /dioxane (1:1:1) at 0 °C was treated with LiOH (1.25 g, 0.052 mol) and stirred for 12h at room temperature. Then, solvents were evaporated at reduced pressure. The residual was treated with 10% of KHSO_4 . The aqueous layer was extracted with EtOAc and the organic phase was dried over Na_2SO_4 , filtered and concentrated in vacuo to give compound **4** as a brown-yellow powder, used without any further purification (1.76 g, 51%); $^1\text{H NMR}$ (500 MHz, CDCl_3) $\delta = 3.60$ (s, 2H), 5.37 (d, $J = 10.1$ Hz, 1H), 7.00 (dd, $J = 10.1$ Hz and 3.0 Hz, 1H), 8.06 (d, $J = 3.0$ Hz, 1H) ppm.

Procedures for the coupling of amides 6a-h

N-Propyl-2-(5-nitro-2-oxopyridin-1(2H)-yl) acetamide (6a)

To a solution of carboxylic acid **4** (200 mg, 1.01 mmol) in dry DCM and DMF (1.5 mL) cooling to 0°C were added HOBt (204.6 mg, 1.51 mmol) and EDCI (290.3 mg, 1.51 mmol). After 10 minutes, DIPEA (215.2 μl , 1.21 mmol) and propylamine **5a** (99.6 μl , 1.21 mmol) dissolved in DMF were added, and the mixture was stirred at room temperature overnight. The organic layer was washed with a saturated solution of NaCl, dried over Na_2SO_4 , filtered and concentrated under reduced pressure. The crude residue was purified by column chromatography using an eluent mixture EtOAc/light petroleum 7:3 to obtain the coupling product **6a**, white powder, 135.7 mg, yield 56; $R_f = 0.40$ (EtOAc/light petroleum 7:3). $^1\text{H NMR}$ (500 MHz, DMSO) $\delta = 0.90$ (t, $J = 7.1$ Hz, 3H), 1.55 – 1.63 (m, 2H), 3.18 – 3.25 (m, 2H), 4.62 (s, 2H), 6.52 (d, $J = 10.1$ Hz, 1H), 8.23 (d, $J = 10.1$ Hz, 1H), 8.35 (s, 1H), 9.24 (d, $J = 2.6$ Hz, 1H) ppm; $^{13}\text{C NMR}$ (75 MHz, DMSO) $\delta = 12.40, 22.45, 43.20, 51.34, 116.25, 130.82, 132.65, 144.75, 162.34, 165.82$ ppm.

N-Butyl-2-(5-nitro-2-oxopyridin-1(2H)-yl)acetamide (6b)

Following the same procedure used for compound **6a**, a solution of 2-(5-nitro-2-oxo-2H-pyridin-1-yl)-acetic acid **4** (200 mg, 1.01 mmol) was reacted with butylamine **5b** (119.6 μ L, 1.21 mmol). The crude obtained was purified by column chromatography using an eluent mixture of EtOAc/light petroleum 5:5 to obtain the coupling product **6b**, yellow powder, 197.8 mg, yield 75% Rf = 0.45 (EtOAc/light petroleum 6:4). ^1H NMR (500 MHz, DMSO) δ = 0.86 (t, J = 7.3 Hz, 3H), 1.24 – 1.31 (m, 2H), 1.35 – 1.42 (m, 2H), 3.07 (d, J = 5.5 Hz, 2H), 4.67 (s, 2H), 6.47 (dd, J = 10.0 Hz and 2.9 Hz, 1H), 8.15 (dd, J = 10.0 Hz and 2.9 Hz, 1H), 8.24 (s, 1H), 9.13 (d, J = 2.9 Hz, 1H) ppm; ^{13}C NMR (75 MHz, DMSO) δ = 13.62, 19.44, 31.08, 38.42, 51.52, 117.91, 129.54, 133.70, 143.01, 160.83, 165.48 ppm.

N-isobutyl-2-(5-nitro-2-oxopyridin-1(2H)-yl)acetamide (6c)

Following the same procedure used for compound **6a**, a solution of 2-(5-nitro-2-oxo-2H-pyridin-1-yl)-acetic acid **4** (200 mg, 1.01 mmol) was reacted with isobutylamine **5c** (119.9 μ L, 1.21 mmol). The crude obtained was purified by column chromatography using an eluent mixture of EtOAc/light petroleum 8:2 to obtain the coupling product **6c**, white powder, 173.9 mg, yield 68%; Rf = 0.30 (light petroleum/EtOAc 5:5); ^1H NMR (500 MHz, DMSO) δ = 0.83 (d, J = 6.7 Hz, 6H), 1.62 – 1.73 (m, 1H), 2.87 – 2.93 (m, 2H), 4.69 (s, 2H), 6.47 (dd, J = 10.1 Hz and 2.8 Hz, 1H), 8.15 (dd, J = 10.1 Hz and 2.8 Hz, 1H), 8.25 (s, 1H), 9.13 (d, J = 2.8 Hz, 1H) ppm; ^{13}C NMR (75 MHz, DMSO) δ = 20.47, 28.53, 46.68, 52.05, 118.34, 129.77, 134.16, 143.49, 161.29, 166.09 ppm.

N-isopentyl-2-(5-nitro-2-oxopyridin-1(2H)-yl)acetamide (6d)

Following the same procedure used for compound **6a**, a solution of 2-(5-nitro-2-oxo-2H-pyridin-1-yl)-acetic acid **4** (200 mg, 1.01 mmol) was reacted with isopentylamine **5d** (140.6 μ L, 1.21 mmol). The crude obtained was purified by column chromatography using an eluent mixture of EtOAc/light petroleum 7:3 to obtain the coupling product **6d**, white powder, 199.8 mg, yield 74%; Rf = 0.47 (EtOAc/light petroleum 6:4); ^1H NMR (500 MHz, DMSO) δ = 1.01 (d, J = 6.6 Hz, 6H), 1.42 – 1.47 (m, 2H), 1.68 – 1.77 (m, 1H), 3.20 – 3.28 (m, 2H), 4.82 (s, 2H), 6.62 (dd, J = 10.1 Hz and 2.4 Hz, 1H), 8.29 (d, J = 10.1 Hz, 1H), 8.37 (s, 1H), 9.26 (d, J = 2.4 Hz, 1H) ppm; ^{13}C NMR (75 MHz, DMSO) δ = 22.64, 25.32, 37.32, 38.29, 51.85, 118.24, 129.86, 134.04, 143.35, 161.16, 165.77 ppm.

N-cyclohexyl-2-(5-nitro-2-oxopyridin-1(2H)-yl)acetamide (6e)

Following the same procedure used for compound **6a**, a solution of 2-(5-nitro-2-oxo-2H-pyridin-1-yl)-acetic acid **4** (200 mg, 1.01 mmol) was reacted with cyclohexylamine **5e** (138.7 μ L, 1.21 mmol). The crude obtained was purified by column chromatography using an eluent mixture of EtOAc/light petroleum 7:3 to obtain the coupling product **6e**, white powder, 203.1 mg, yield 72%; Rf = 0.35 (EtOAc/light petroleum 5:5); ^1H NMR (500 MHz, DMSO) δ = 0.58 – 0.76 (m, 6H), 0.96 – 1.04 (m, 1H), 1.09 – 1.23 (m, 4H), 4.13 (s, 2H), 5.94 (d, J = 10.1 Hz, 1H), 7.61 (dd, J = 10.0 Hz and 3.1 Hz, 1H), 7.64 (s, 1H), 8.58 (d, J = 3.1 Hz, 1H) ppm; ^{13}C NMR (75 MHz, DMSO) δ = 23.56, 24.86, 32.44, 51.74, 53.89, 114.25, 130.62, 133.94, 146.51, 164.44, 170.89 ppm.

N-phenyl-2-(5-nitro-2-oxopyridin-1(2H)-yl)acetamide (6f)

Following the same procedure used for compound **6a**, a solution of 2-(5-nitro-2-oxo-2H-pyridin-1-yl)-acetic acid **4** (200 mg, 1.01 mmol) was reacted with aniline **5f** (117.7 μ L, 1.21 mmol). The crude obtained was purified by column chromatography using an eluent mixture of EtOAc/light petroleum 6:4 to obtain the coupling product **6f**, white powder, 184.9 mg, yield 67%; Rf = 0.5 (EtOAc/light petroleum 6:4); ^1H NMR (500 MHz, DMSO) δ = 4.76 (s, 2H), 6.66 (d, J = 10.1 Hz, 1H), 7.15 (t, J = 7.2 Hz, 1H), 7.33 (t, J = 7.3 Hz, 2H), 7.50 (d, J = 7.8 Hz, 2H), 8.18 (d, J = 10.1 Hz, 1H), 8.22 (s, 1H), 8.76 (s, 1H) ppm; ^{13}C NMR (75 MHz, DMSO) δ = 52.38, 115.84, 120.38, 124.56, 128.58, 131.95, 133.68, 136.05, 145.33, 164.16, 167.89 ppm.

N-benzyl-2-(5-nitro-2-oxopyridin-1(2H)-yl)acetamide (6g)

Following the same procedure used for compound **6a**, a solution of 2-(5-nitro-2-oxo-2H-pyridin-1-yl)-acetic acid **4** (200 mg, 1.01 mmol) was reacted with benzylamine **5g** (132.2 μ L, 1.21 mmol). The crude obtained was purified by column chromatography using an eluent mixture of EtOAc/light petroleum 6:4 to obtain the coupling product **6g**, white powder, 116 mg, yield 40%; Rf = 0.50 (EtOAc/light petroleum 6:4); ^1H NMR (500 MHz, DMSO) δ = 4.33 (d, J = 5.8 Hz, 2H), 4.78 (s, 2H), 6.52 (d, J = 10.0 Hz, 1H), 7.22 – 7.37 (m, 5H), 8.17 (dd, J = 3.0 Hz, 1H), 8.80 (s, 1H), 9.15 (d, J = 3.0 Hz, 1H) ppm; ^{13}C NMR (75 MHz, DMSO) δ = 45.24, 52.68, 114.25, 125.71, 126.99, 128.14, 130.56, 132.88, 136.74, 147.26, 164.32, 166.90 ppm.

N-phenylethyl-2-(5-nitro-2-oxopyridin-1(2H)-yl)acetamide (6h)

Following the same procedure used for compound **6a**, a solution of 2-(5-nitro-2-oxo-2H-pyridin-1-yl)-acetic acid **4** (200 mg, 1.01 mmol) was reacted with phenethylamine **5h** (152.3 μ L, 1.21 mmol). The crude obtained was purified by column chromatography using an eluent mixture of EtOAc/light petroleum 6:4 to obtain the coupling product **6h**, white powder, 120.4 mg, yield 44%; Rf = 0.45 (EtOAc/light petroleum 6:4). ^1H NMR (500 MHz, DMSO) δ = 2.64 (t, J = 5.4 Hz, 2H), 4.17 (d, J = 5.9 Hz, 2H), 4.69 (s, 2H), 6.87 (d, J = 10.4 Hz, 1H), 7.14 – 7.28 (m, 5H), 8.12 (dd, J = 2.9 Hz, 1H), 8.79 (s, 1H), 8.96 (d, J = 2.6 Hz, 1H) ppm; ^{13}C NMR (75 MHz, DMSO) δ = 35.62, 40.15, 50.28, 117.11, 124.98, 126.22, 128.04, 130.62, 132.44, 139.27, 146.48, 166.52, 169.18 ppm.

2-(5-Amino-2-oxopyridin-1(2H)-yl)-N-propyl-acetamide (1a)

A solution of N-propyl-2-(5-nitro-2-oxopyridin-1(2H)-yl) acetamide **6a** (135.7 mg, 0.57 mmol) in ethanol (5mL) hydrazine (276.5 μ L, 5.7 mmol) and the Pd / C catalyst in catalytic quantity was added in a 10mL microwave vial equipped with a magnetic stirrer and subjected to microwave irradiation at 140 ° C, 50 W for thirty minutes. After this time, an addition of hydrazine (138.2 μ L, 2.8 mmol) and the catalyst was made and irradiated at 140 ° C, 50 W for fifteen minutes. After the reaction time, a filtration on celite with ethanol was carried out. The crude was purified by column chromatography using an eluent mixture of CHCl₃ / MeOH 95: 5 to obtain the final product **1a**, white powder, 39 mg, yield 33% ; Rf = 0.4 (CHCl₃/MeOH 95:5); ^1H NMR (500 MHz, CDCl₃) δ = 0.81 (t, J = 7.1 Hz, 3H), 1.44 (dd, J = 14.6 Hz and 7.3 Hz, 2H), 3.11 (dd, J = 12.9 Hz and 6.2 Hz, 2H), 4.47 (s, 2H), 6.20 (t, J = 6.7 Hz, 1H), 6.56 (d, J = 9.1 Hz, 1H), 6.90 (s, 1H), 7.35 (d, J = 6.7 Hz, 1H) ppm; ^{13}C NMR (75 MHz, CDCl₃) δ = 12.64, 20.45, 40.15, 52.34, 118.64, 121.28, 132.75, 164.25, 167.99 ppm. Elementary Analysis: calculated for C₁₀H₁₅N₃O₂: C 57.40, H 7.23, N 20.08; found C 57.24, H 7.63, N 19.96.

2-(5-Amino-2-oxopyridin-1(2H)-yl)-N-butyl-acetamide (1b)

Compound **1b** was synthesized starting from N-butyl-2-(5-nitro-2-oxopyridin-1(2H)-yl)acetamide **6b** (191.8 mg, 0.76 mmol), following the same procedure used for compound **1a**. The crude was purified by column chromatography using an eluent mixture of CHCl₃ / MeOH 95: 5 to obtain the final product **1b**, white powder, 54 mg, yield 32%, ; R_f = 0.36 (CHCl₃/MeOH 95:5); ¹H NMR (500 MHz, CDCl₃) δ = 0.89 (t, J = 7.3 Hz, 3H), 1.25 – 1.35 (m, 2H), 1.60 – 1.67 (m, 2H), 3.22 (dd, J = 13.0 Hz and 7.1 Hz, 2H), 4.53 (s, 2H), 6.27 (t, J = 6.7 Hz, 1H), 6.63 (d, J = 9.9 Hz, 1H), 6.94 (s, 1H), 7.42 (d, J = 6.8 Hz, 1H) ppm; ¹³C NMR (75 MHz, CDCl₃) δ = 13.76, 20.16, 31.68, 39.94, 54.23, 107.01, 127.10, 138.04, 140.53, 166.59, 167.06 ppm. Elementary Analysis: calculated for C₁₁H₁₇N₃O₂: C 59.17, H 7.67, N 18.82; found C 58.94, H 8.01, N 18.67.

2-(5-Amino-2-oxopyridin-1(2H)-yl)-N-isobutyl-acetamide (1c)

Compound **1c** was synthesized starting from N-isobutyl-2-(5-nitro-2-oxopyridin-1(2H)-yl)acetamide **6c** (173.9 mg, 0.69 mmol), following the same procedure used for compound **1a**. The crude was purified by column chromatography using an eluent mixture of CHCl₃ / MeOH 95: 5 to obtain the final product **1c**, yellow powder, 56 mg, yield 37%; R_f = 0.40 (CHCl₃/MeOH 95:5); ¹H NMR (500 MHz, CDCl₃) δ = 0.87 (d, J = 6.7 Hz, 6H), 1.70 – 1.81 (m, 1H), 3.04 (t, J = 6.3 Hz, 2H), 4.56 (s, 2H), 6.27 (t, J = 6.7 Hz, 1H), 6.64 (d, J = 8.9 Hz, 1H), 7.07 (s, 1H), 7.43 (d, J = 6.7 Hz, 1H) ppm; ¹³C NMR (75 MHz, CDCl₃) δ = 20.21, 28.56, 47.15, 54.56, 107.30, 121.03, 138.30, 140.82, 163.19, 167.47 ppm. Elementary Analysis: calculated for C₁₁H₁₇N₃O₂: C 59.17, H 7.67, N 18.82; found C 59.31, H 7.52, N 18.68.

2-(5-Amino-2-oxopyridin-1(2H)-yl)-N-isopentyl-acetamide (1d)

Compound **1d** was synthesized starting from N-isopentyl-2-(5-nitro-2-oxopyridin-1(2H)-yl)acetamide **6d** (199.8 mg, 0.75 mmol), following the same procedure used for compound **1a**. The crude was purified by column chromatography using an eluent mixture of CHCl₃/MeOH 95: 5 to obtain the final product **1d**, yellow powder, 53 mg, yield 30%; R_f = 0.34 (CHCl₃/MeOH 95:5); ¹H NMR (500 MHz, CDCl₃) δ = 0.81 (d, J = 6.6 Hz, 6H), 1.30 (dd, J = 14.8 Hz and 7.1 Hz, 2H), 1.46 – 1.52 (m, 1H), 3.16 (dd, J = 6.0 Hz, 2H), 4.46 (s, 2H), 6.19 (t, J = 6.7, 1.2 Hz, 1H), 6.55 (d, J = 9.9 Hz, 1H), 6.84 (s, 1H), 7.34 (d, J = 6.7 Hz, 1H) ppm; ¹³C NMR (75 MHz, CDCl₃) δ = 22.62, 26.07, 29.94, 38.25, 54.40, 107.29, 121.02, 138.35, 140.80, 163.16, 167.27 ppm. Elementary Analysis: calculated for C₁₂H₁₉N₃O₂: C 60.74, H 8.07, N 17.71; found C 60.53, H 7.88, N 17.36.

2-(5-Amino-2-oxopyridin-1(2H)-yl)-N-cicloexhyl-acetamide (1e)

Compound **1e** was synthesized starting from N-cyclohexyl-2-(5-nitro-2-oxopyridin-1(2H)-yl)acetamide **6e** (203.1 mg, 0.73 mmol) following the same procedure used for compound **1a**. The crude was purified by column chromatography using an eluent mixture of CHCl₃ / MeOH 95: 5 to obtain the final product **1e**, yellow powder, 60 mg, yield 33%; ; R_f = 0.49 (CHCl₃/MeOH 95:5); ¹H NMR (500 MHz, CDCl₃) δ = 0.74 – 0.84 (m, 8H), 1.46 – 1.62 (2H), 3.57 – 3.67 (m, 1H), 4.44 (s, 2H), 6.19 (t, J = 6.7 Hz, 1H), 6.55 (d, J = 9.0 Hz, 1H), 6.76 (s, 1H), 7.34 (d, J = 6.7 Hz, 1H) ppm; ¹³C NMR (75 MHz, CDCl₃) δ = 22.47, 24.96, 32.76, 52.16, 54.98, 115.60, 120.06, 132.52, 137.86, 166.48, 169.35 ppm. Elementary Analysis: calculated for C₁₃H₁₉N₃O₂: C 62.63, H 7.68, N 16.85; found C 62.91, H 7.83, N 16.53.

2-(5-Amino-2-oxopyridin-1(2H)-yl)-N-phenyl-acetamide (1f)

Compound **1f** was synthesized starting from 2-(5-nitro-2-oxopyridin-1(2H)-yl)-N-phenyl-acetamide **6f** (184.9 mg, 0.68 mmol), following the same procedure used for compound **1a**. The crude was purified by column chromatography using an eluent mixture of CHCl₃ / MeOH 95:5 to obtain the final product **1f**, white powder, 49 mg, yield 30% ; $R_f = 0.42$ (CHCl₃/MeOH 95:5); ¹H NMR (500 MHz, CDCl₃) $\delta = 4.64$ (s, 2H), 6.27 (s, 1H), 6.64 (d, $J = 8.8$ Hz, 1H), 7.00 – 7.07 (m, 2H), 7.37 – 7.52 (m, 5H) ppm; ¹³C NMR (75 MHz, CDCl₃) $\delta = 50.96, 109.45, 121.39, 121.98, 123.70, 124.85, 128.05, 133.23, 136.76, 165.05, 166.84$ ppm. Elementary Analysis: calculated for C₁₃H₁₃N₃O₂: C 64.19, H 5.39, N 17.27; found C 63.98, H 5.28, N 17.41.

2-(5-Amino-2-oxopyridin-1(2H)-yl)-N-benzyl-acetamide (**1g**)

Compound **1g** was synthesized starting from 2-(5-nitro-2-oxopyridin-1(2H)-yl)-N-benzyl-acetamide **6g** (116 mg, 0.40 mmol), following the same procedure used for compound **1a**. The crude was purified by column chromatography using an eluent mixture of CHCl₃ / MeOH 95:5 to obtain the final product **1g**, yellow powder, 35 mg, yield 34%; $R_f = 0.43$ (CHCl₃/MeOH 95:5); ¹H NMR (500 MHz, CDCl₃) $\delta = 4.35$ (d, $J = 5.8$ Hz, 2H), 4.52 (s, 2H), 6.20 (t, $J = 6.7$ Hz, 1H), 6.55 (d, $J = 9.0$ Hz, 1H), 7.15 (d, $J = 7.3$ Hz, 2H), 7.21 – 7.25 (m, 2H), 7.31 – 7.37 (m, 2H) ppm; ¹³C NMR (75 MHz, CDCl₃) $\delta = 42.28, 50.64, 110.06, 120.64, 121.99, 125.52, 128.02, 128.88, 133.78, 138.61, 164.08, 166.91$ ppm. Elementary Analysis: calculated for C₁₄H₁₅N₃O₂: C 65.35, H 5.88, N 16.33; found C 65.51, H 5.61, N 16.64.

2-(5-Amino-2-oxopyridin-1(2H)-yl)-N-phenylethyl-acetamide (**1h**)

Compound **1h** was synthesized starting from 2-(5-nitro-2-oxopyridin-1(2H) -yl)-N-phenylethyl-acetamide **6h** (133.9 mg, 0.44 mmol), following the same procedure used for compound **1a**. The crude was purified by column chromatography using an eluent mixture of CHCl₃ / MeOH 95:5 to obtain the final product **1h**, yellow powder, 42 mg, yield 35%; $R_f = 0.45$ (CHCl₃/MeOH 95:5); ¹H NMR (500 MHz, CDCl₃) $\delta = 2.58$ (t, $J = 5.4$ Hz, 2H), 3.60 (t, $J = 5.4$ Hz, 2H), 4.36 (s, 2H), 6.31 (t, $J = 6.5$ Hz, 1H), 6.45 (d, $J = 6.5$ Hz, 1H), 7.15 (s, 1H), 7.32 – 7.54 (m, 6H) ppm; ¹³C NMR (75 MHz, CDCl₃) $\delta = 36.05, 40.25, 50.06, 108.76, 121.38, 122.99, 124.19, 128.84, 129.03, 132.47, 137.39, 164.22, 167.08$ ppm. Elementary Analysis: calculated for C₁₅H₁₇N₃O₂: C 66.40, H 6.32, N 15.49; found C 66.13, H 6.45, N 15.19.

3.6.2.2 Biological activity

In vitro proteasome inhibition assay

Human 20S proteasome, isolated from human erythrocytes was purchased from Enzo Life Science.

A preliminary screening at 100 μ M inhibitor concentrations was carried out on the three proteolytic activities of proteasome; an equivalent amount of DMSO as a negative control and MG-132, a reversible inhibitor of the proteasome as a positive control was employed. Compounds showing at least 60% inhibition at the screening concentration were then progressed into detailed assays.

The substrate Suc-Leu-Leu-Val-Tyr-AMC (Bachem) for the β 5c subunit was diluted at 50 μ M of reaction buffer (Tris buffer, pH 7.5, 0.03% SDS).

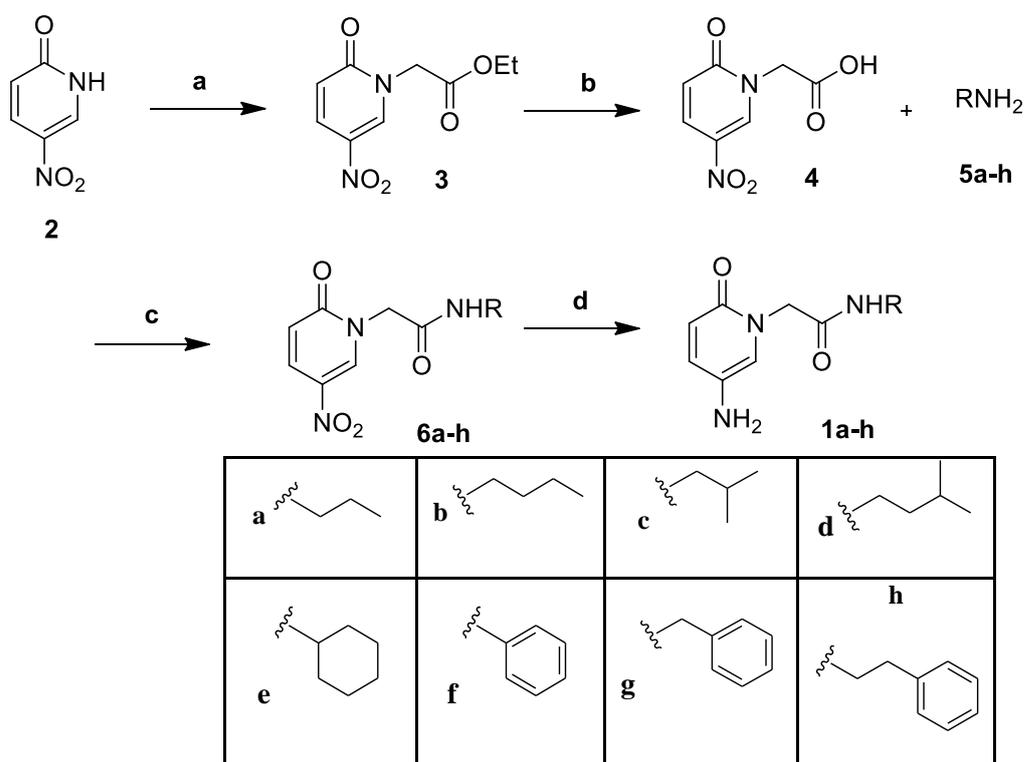
Fluorescence of the product AMC of the substrate hydrolysis was measured at 30°C for 10 minutes with a 380 nm excitation filter and a 460 nm emission filter, using an Infinite 200 PRO microplate reader (Tecan, Männedorf, Switzerland). Compounds showing at least 60% inhibition at the screening concentration were then progressed into detailed assays. Continuous assays were performed at seven different concentrations ranging from minimally inhibited to fully inhibiting each proteolytic activity to calculate the dissociation constants K_i of the

enzyme-inhibitor complex using the Cheng-Prusoff equation $K_i = IC_{50}/(1 + [S] K_m^{-1})$. Inhibitor solutions were prepared from stocks in DMSO. Each independent assay was performed in duplicate in 96-well-plates with a total volume of 200 μ L.

3.6.3 RESULTS AND DISCUSSION

3.6.3.1 Synthesis of amides 1a-h

The synthesis of amides **1a-h** (Scheme 1) started from 5-nitro-2-pyridone **2** that was N-alkylated with ethyl bromoacetate to give ester **3**. Subsequently, acid **4** was obtained by alkaline hydrolysis with LiOH of the ester function. Coupling reactions between the carboxylic acids and the suitable amines **5a-h**, in the presence of EDC-HCl, HOBt, as coupling reagents, and DIPEA as a base, gave the desired amides **6a-h**. The final compounds **1a-h** were obtained by a reduction of the nitro group under microwave irradiation in the presence of hydrazine and palladium catalyst. (Scheme 1).



Scheme 1: Reagents and conditions: a) NaH, DMF, N₂, 0 ° C 1h, then ethyl bromoacetate t.a. o.n. ; b) LiOH, MeOH / H₂O / dioxane (1: 1: 1), 0 ° C, 10 min, then rt o.n. ; c) HOBt, EDCI, DIPEA, CH₂Cl₂ / DMF, 0 ° C, 10 min, then rt o.n. ; d) NH₂NH₂, Pd / C

3.6.3.2 Biological activity

The new amide derivatives have been screened to test their ability to inhibit the human erythrocyte 20S proteasome by measuring the rate of hydrolysis of the appropriate fluorogenic substrate Suc-Leu-Leu-Val-Tyr-AMC, specific for the chymotrypsin-like (ChT-L) activity of the proteasome.

MG-132 (Z-Leu-Leu-Leu-al), a reversible inhibitor of both proteasome and immunoproteasome, was used as a positive control.

First, compounds underwent a preliminary screening on each proteolytic subunit at 100 μM, using an equivalent volume of DMSO, as a negative control, and MG-132, as a positive control.

Furthermore, all the synthesized derivatives are also undergoing biological evaluation both on cell lines of haematological tumors, such as cell lines of multiple myeloma (MM1S, MM1R), and on cell lines of solid tumors, such as hepatocarcinoma (HepG2), breast cancer (AMJ13) and lung cancer (A549).

Table 5: screening at 100 μ M of compounds 1a-f against the β 5 subunit of the constitutive human proteasome

Compound	R	% inhibition at 100 μ M
1a	propyl	46%
1b	butyl	11%
1c	isobutyl	28%
1d	isopentyl	62%
1e	cyclohexyl	53%
1f	phenyl	29%
1g	benzyl	34%
1h	phenylethyl	u.s.
u.s. = under screening		

Table 5 reports the percentages of inhibition of the preliminary screening of compounds **1a-h**. From these preliminary results, it seems that as the length of the carbon chain and branches increases, or as the amide substituent's steric hindrance increases, the inhibition percentage increases. Among all the tested compounds, **1d-1e**, bearing at the P1 site an isopentyl and a cyclohexyl, turned to be active on the CP at 100 μ M.

On these compounds Tian dilution assays at seven different concentrations to determine the dissociation constant K_i for each inhibitor, which is a measure of enzymatic affinity, are in progress.

4. IN SILICO DESIGN, SYNTHESIS, AND BIOLOGICAL EVALUATION OF ANTICANCER ARYLSULFONAMIDE ENDOWED WITH ANTI-TELOMERASE ACTIVITY

4.1. INTRODUCTION

Telomere, a nuclear protein complex, defends the terminal ends of chromosomes from degradation end-to-end fusion and recombination [98–100]. After each cell division cycle, the telomere gradually shortens until the chromosomal DNA is exposed, inducing a DNA damage response [101, 102]. This event helps to maintain the stability of genetic information and protects the genome in a “time-bomb” manner [103]. When the length of telomeres reaches a critical point, cells reach the cycle of termination, ageing, and death [102, 103]. Normal cells cannot survive this progressive shortening. Sometimes, cells can extend telomeres by reactivating telomerase activity or through a telomere replacement elongation mechanism (ALT) to help cells survive the crisis [104]. The reactivation of telomerase is observed in 85–90% of human tumor cells [105]. Telomerase is a reverse transcriptase that contains an RNA template (TER) with its binding domain (TRBD) and reverse transcriptase unit (TERT). Telomerase is a challenging target, because can be inhibited at different stages and with several mechanism. [106–108]. A first approach is based on inhibiting telomerase access to DNA by stabilizing G-quadruplexes formed by a 3' DNA overhang, which has the task to block telomere elongation [108–112]. A second approach includes the use of inhibitors of the enzymatic active site of telomerase including antisense compounds to TER [113,114]. The most promising approach is possibly using compounds that can directly block the enzyme's active center in the catalytic subunit. BIBR1532, which selectively inhibits telomerase activity, was first proposed in 2001 and has been patented as a potential anticancer drug [115,116]. As reported in a recent review, structure-based drug design strategies can be used to design potential inhibitors of the telomerase active site [117]. The first structure showing a co-crystallized inhibitor was published in 2015, revealing the *Tribolium castaneum* full-length catalytic subunit of telomerase in complex with the compound BIBR1532 (PDB ID: 5CQG), which showed that the studied ligand interacts outside of the active center [118]. This experimental crystal structure changed the understanding of the expected mechanism of activity of this compound and possibly other active site-directed molecules. BIBR1532 binds to a highly conserved hydrophobic pocket (FVYL) motif on the outer surface of the thumb domain of telomerase. Several other compounds designed using *in silico* approaches have been tested as potential inhibitors of the catalytic subunit of telomerase [119 -131]. Thus, considering the direct inhibition of TERT as the most promising approach for blocking the action of telomerase and the lack of an approved drug (Imetelstat is the only one to reach clinical trials [117]) makes the search for new and effective inhibitors a hot topic of the drug discovery.

To date, there is just one piece of evidence of aryl sulfonamide derivatives (SEW05920) as an inhibitor of the TERT [132]. In our study, some aryl sulfonamides have been identified as hits, also synthesized using green chemistry approaches, and tested against three cancer cell lines K-562, HCT-116, and MCF-7. The results determined compound 2C as more active than the reference compound BIBR1532. The identification of this novel scaffold, which will undergo optimization, could help to identify new potent telomerase inhibitors

4.2 MATERIALS AND METHODS

4.2.1 Library and Protein Preparation

The *in-house* library used for this work comprised a series of 126 sulfonamide derivatives. The sulfonamide and BIBR1532 structures were built using the builder panel in Maestro, and ligand preparation was carried out by LigPrep (Schrödinger, LLC, New York, NY, 2021). The force field adopted was OPLS4 and Epik 5.5 (Schrödinger, 2021-1) was selected as an ionization tool at pH 7.2 ± 0.2 . The 2.30 Å resolution crystal structure of *Tribolium castaneum* telomerase in complex with the highly specific inhibitor BIBR1532 (PDB ID: 5CQG) [118] was used. The structure was optimized using the Protein Preparation Wizard in Maestro (Schrödinger, 2021-1) adding bond orders and hydrogen atoms to the crystal structure using the OPLS4 force field. The other settings are reported in [133].

4.2.2 Docking Studies

The library was submitted to a docking study using Glide v9.0 [95] in standard precision (SP) with the OPLS4 force field. The grid box was built considering the BIBR1532 as the centroid of the grid. The study was performed using no constraints. Van der Waals radii were set at 0.8, and the partial cutoff was 0.15 with flexible ligand sampling. Bias sampling torsion penalization for amides with non-planar conformation and Epik state penalties were added to the docking score.

4.2.3 MYSHAPE Approach

The evaluation of the ligand-protein interaction patterns was performed by using the MYSHAPE approach[35].

To construct the MYSHAPE, the PDB ligand-protein structure has been imported into LigandScout 4.2.1 [93], and a structure-based pharmacophore model was generated. A shared pharmacophore model was then generated using the structure-based pharmacophore models. This step allows the generation of a model containing only the features that are common and aligned in the PDB models. The PDB ligand was aligned on the previously generated pharmacophore, and the features that were common during the MD simulation but that were not present in the original shared pharmacophore model were added to the previous model. For the newly added pharmacophore features the tolerance radius was increased by 0.15 Å to compensate for small deviations in the 3D coordinates.

4.2.4 MM-GBSA Free Energy Calculations and Molecular Dynamics simulation

The output of docking was used to calculate the ΔG values of the BIBR1532 and molecules using MM-GBSA (molecular mechanics generalized-born surface area) [133,134] (Prime, Schrödinger, LLC, New York, NY, USA, 2021). The VSGB solvation model was chosen using OPLS4 force field with the minimized sampling method.

A 100 ns MD simulation was carried out using a Desmond 6.5 (Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2021) using the OPLS4 force field for the complex TERT/2C. The complex was solvated in a cubic box using the TIP3P. Ions were added to neutralize charges. The systems were minimized and equilibrated at a temperature of 303.15 K and a pressure of 1.013 bar. The system was simulated as an NPT ensemble; a Nose-Hover thermostat and Martyna-Tobia-Klein barostat were used. The integration time step was chosen to be 2 fs. To keep the hydrogen-heavy atom bonds rigid, the SHAKE algorithm was used. A 9 Å cutoff radius was set for the short-range Coulomb interactions, and smooth particle mesh

Ewald was used for the long-range interactions. The detection ranges for energy were 1.2 ps, and 5.0 ps for the trajectory.

4.2.5 ADME Prediction

Drug-likeness, physicochemical properties, lipophilicity, solubility, and pharmaco-kinetics properties were analyzed by the SwissADME web tool (<http://www.swissadme.ch/index.php>, accessed on 30 November 2021) [135].

4.2.6 Chemistry

All melting points were taken on a Büchi melting point M-560 apparatus and were uncorrected. IR spectra were recorded in Bromoform with a JASCO FT-IR spectrophotometer. ¹H and ¹³C NMR spectra were measured at 200 and 50.0 MHz, respectively, in DMSO-d₆ solution and TMS as an internal standard, using a Bruker Avance II Series 200 MHz spectrometer or at 300 and 75 MHz (APT) with a Bruker AC-E spectrometer. Column chromatography was performed with Merck silica gel (230–400 mesh ASTM). Elemental analyses (C, H, and N) were within ±0.4% of theoretical values. The substituted benzylamine derivatives were commercially available and were used without further purification. For all the compounds already cited in the literature, the IR and NMR spectra results were identical to those reported (1A–E). The ¹H and ¹³C NMR spectra of newly synthesized compounds (1F,G and 2A–C) are reported in the original manuscript.

General Methods for the Preparation of N-(R'-benzyl)-4-R-benzenesulfonamides of Type 1

Method X: To a stirred solution of amino derivative (1.5 mmol) in anhydrous THF (20 mL), an equimolar amount of triethylamine (0.21 mL) and the suitable sulfonylchloride were added. The reaction mixture was stirred for the appropriate time-lapse until the disappearance of the starting amine (TLC monitoring). The ammonium salt was collected and taken up with 20 mL H₂O. The aqueous solution was extracted using DCM. The organic layers were combined with the organic mother liquor, dried over Na₂SO₄, and concentrated under reduced pressure to yield the desired sulfonamide of type 1.

Method Y: The amino derivative (5 mmol) was dissolved in dry pyridine, and the suitable sulfonylchloride (5 mmol) was added. The reaction mixture was stirred for the appropriate time-lapse until the disappearance of starting amine (TLC monitoring). The mixture was poured onto ice/water and the solid precipitate was collected by filtration, air-dried, and recrystallized from ethanol.

N-(4-Chlorobenzyl)-4-methylbenzenesulfonamide (1A)

This compound was prepared according to Method X. The reaction mixture was stirred for 24 h at room temperature. Yield 70%. Mp 98 °C ([136] 107–108 °C; [137] 106.1 °C). ¹H NMR ppm: 2.43 (s, 3H, Me), 4.06 (d, J = 6.4 Hz, 2H, CH₂), 5.18 (t, J = 6.0 Hz, 1H, NH), 7.11 (d, J = 8.4 Hz, 2H, Ar-H), 7.20 (d, J = 8.4 Hz, 2H, Ar-H), 7.27 (d, J = 8.0 Hz, 2H, Ar-H), 7.70 (d, J = 8.4 Hz, 2H, Ar-H). ¹³C NMR ppm: 21.6, 46.6, 127.2, 128.8, 129.3, 129.9, 133.7, 135.1, 136.9, and 143.8.

N-(4-Chlorobenzyl)-4-acetylamino benzenesulfonamide (1B)

This compound was prepared according to Method X. The reaction mixture was stirred for 24 h at room temperature and under reflux for an additional 4 h. Yield 64%. Mp 152 °C [138] 172–174 °C). ¹H NMR ppm: 2.10 (s, 3H, Me), 3.08 (bs, 1H, NH), 3.96 (s, 2H, CH₂), 7.30–8.11 (m, 8H, Ar-H), 10.42 (bs, 1H, NH). ¹³C NMR ppm: 24.1, 45.5, 118.6, 127.6, 128.5, 129.4, 131.6, 134.1, 136.9, 142.5, and 169.0.

N-(4-Chlorobenzyl)-4-nitrobenzenesulfonamide (1C)

This compound was prepared according to Method X. The reaction mixture was stirred under reflux for 4.5 h. Yield 78%. Mp 174–176 °C ([139] 168–170 °C). ¹H NMR ppm: 3.80 (d, J = 6.0 Hz, 2H, CH₂), 6.61 (d, J = 8.0 Hz, 2H, Ar-H), 6.85 (d, J = 8.0 Hz, 2H, Ar-H), 7.15 (d, J =

8.0 Hz, 2H, Ar-H), 7.45 (d, J = 8.0 Hz, 2H, Ar-H), 7.56 (t, J = 6.0 Hz, 1H, NH). ¹³C NMR ppm: 45.5, 112.6, 113.5, 125.5, 128.4, 128.9, 129.8, 152.4, and 158.3.

N-(4-Methoxybenzyl)-4-methylbenzenesulfonamide (1D)

This compound was prepared according to Method A X. The reaction mixture was stirred for 3 h at room temperature. Yield 71%. Mp 122–123 °C ([140] 114–117 °C; [141] 122–123 °C from EtOAc). ¹H NMR (300 MHz, DMSO) ppm: 2.38 (s, 3H, Me), 3.72 (s, 3H, Me), 3.89 (s, 2H, CH₂), 6.85 (d, J = 8.4 Hz, 2H, Ar-H), 7.17 (d, J = 8.4 Hz, 2H, Ar-H), 7.38 (d, J = 8.0 Hz, 2H, Ar-H), 7.71 (d, J = 8.4 Hz, 2H, Ar-H), 8.03 (bs, 1H, NH). ¹³C NMR ppm: 20.9, 45.8, 55.0, 113.6, 126.5, 128.9, 129.4, 129.5, 137.8, 142.5, and 158.4.

N-(4-Methoxybenzyl)-4-acetylamino benzenesulfonamide (1E)

This compound was prepared according to Method Y. The reaction mixture was stirred for 1 h at room temperature and under reflux for an additional 2 h. Yield 49%. Mp 192–193 °C ([140] 164–167 °C). ¹H NMR ppm: 1.84 (s, 2H, CH₂), 2.12 (s, 3H, Me), 4.46 (sa, 1H, NH), 7.74 (d, J = 8.5 Hz, 2H), 7.79 (d, J = 8.5 Hz, 2H), 9.41 (sa, 1H, NH). ¹³C NMR ppm: 24.2, 45.6, 55.2, 113.6, 118.6, 127.6, 129.1, 129.5, 134.0, 142.9, 158.5, and 169.4

N-(2,5-Dimethoxybenzyl)-4-methylbenzenesulfonamide (1F)

This compound was prepared according to Method Y. The reaction mixture was stirred for 10 h at room temperature. Yield 80%. Mp 110–111 °C. IR v: 3268 (NH) cm⁻¹. ¹H-NMR ppm: 2.37 (3H, s, Me), 3.65 (6H, s, Me), 3.91 (2H, t, J = 3.9 Hz, CH₂), 6.77–6.83 (3H, m, ArH), 7.10 (2H, d, J = 8.5 Hz), 7.36 (2H, d, J = 7.6 Hz), 7.68 (2H, d, J = 7.6 Hz), 7.89 (1H, t, J = 3.9 Hz, NH). ¹³C NMR ppm: 21.4, 41.3, 55.8, 56.1, 111.8, 113.0, 115.0, 126.7, 127.0, 130.0, 138.2, 143.0, 150.9, and 153.4.

N-(2,5-Dimethoxybenzyl)-4-nitrobenzenesulfonamide (1G)

This compound was prepared according to Method X. The reaction mixture was stirred for 24 h at room temperature. Yield 97%. Mp 129–132 °C. IR v: 3268 (NH), 1523 and 1348 (NO₂) cm⁻¹. ¹H NMR ppm: 3.35 (s, 3H, Me), 3.63 (s, 3H, Me), 4.03 (s, 2H, CH₂), 6.65–6.86 (m, 3H, Ar), 7.96 (2H, d, J = 8.8 Hz), 8.34 (2H, d, J = 8.8 Hz). ¹³C NMR ppm: 46.0, 55.7, 56.1, 111.8, 113.2, 115.5, 124.7, 125.9, 128.4, 147.0, 149.8, 151.0, and 153.2.

Preparation of 1-(4-aminophenyl)-3,5-dimethylpyrazole

A solution of acetylacetone (10 mmol) and 4-nitrophenylhydrazine (10 mmol) in acetic acid was heated under reflux for 9 h. The reaction mixture was cooled to room temperature and poured onto ice/water to give a solid precipitate. Purification by column chromatography (eluant DCM/EtOAc 4:1) gave the nitro compound. Yield 80%, mp 100 °C ([89] 102 °C from EtOH). 1-(4-Nitrophenyl)-3,5-dimethylpyrazole (0.56 mg) dissolved in EtOH (20 mL) Pd/C was added, and the mixture was reduced overnight at room temperature in an H₂ atmosphere (50 psi) in a Parr apparatus. The catalyst was filtered off, and the solution was evaporated under reduced pressure. The amino derivatives were isolated as white crystals. Yield 100%. Mp 84 °C ([90] 82–84 °C from benzene).

General Method for the Preparation of N-[4-(3,5-dimethyl-1H-pyrazol-1-yl)phenyl]-4-R-benzenesulfonamide of Type 2

To a stirred solution of 1-(4-aminophenyl)-3,5-dimethylpyrazole (1.5 mmol) in anhydrous

THF (20 mL), an equimolar amount of triethylamine (0.21 mL) and the suitable sulfonylchloride were added. The reaction mixture was stirred for the appropriate time-lapse until the disappearance of the starting amine (TLC monitoring). The ammonium salt was collected and taken up with 20 mL H₂O. The aqueous solution was extracted using DCM. The organic layers were combined with the organic mother liquor, dried over Na₂SO₄, and concentrated under reduced pressure to yield the desired sul- fonamide.

N-[4-(3,5-dimethylpyrazol-1-yl)phenyl]-4-methylbenzenesulfonamide (2A)

The reaction mixture was stirred for 30 h at room temperature. Yield 96%. Mp 167 °C. ir v: 3392 (NH) cm⁻¹. ¹H NMR ppm: 2.09 (s, 3H, Me), 2.13 (s, 3H, Me), 2.28 (s, 3H, Me), 6.00 (s, 1H, pyrazole-CH), 7.20 (dd, 4H, ArH), 7.27 (d, 2H, ArH), 7.64 (d, 2H, ArH), 10.41 (s, 1H, NH). ¹³C NMR ppm: 12.3, 13.4, 21.3, 107.5, 120.7, 125.5, 125.8, 127.2, 128.8, 130.3, 136.7, 134.0, 144.2, and 148.6.

N-[4-(3,5-dimethylpyrazol-1-yl)phenyl]-4-nitrobenzenesulfonamide (2B)

The reaction mixture was stirred for 3 h at room temperature. The solid residue was recrystallized from ethanol. Yield 57%. Mp 219 °C. ir v: 3392 (NH), 1532, and 1348 (NO₂) cm⁻¹. ¹H NMR ppm: 2.12 (s, 3H, Me), 2.21 (s, 3H, Me), 6.02 (s, 1H, pyrazole-CH), 7.27 (dd, 4H, ArH), 8.19 (dd, 4H, ArH), 10.81 (s, 1H, NH). ¹³C NMR ppm: 12.0, 13.2, 107.0, 120.9, 124.7, 125.0, 128.3, 132.2, 135.5, 136.2, 139.0, 147.7, and 149.9.

Preparation of N-[4-(3,5-dimethylpyrazol-1-yl)phenyl]-4-aminobenzenesulfonamide (2C)

To the nitro derivative, 2B dissolved was in EtOH (20 mL), Pd/C was added, and the mixture was reduced overnight at room temperature in an H₂ atmosphere (50 psi) in a Parr apparatus. The catalyst was filtered off, and the solution was evaporated under reduced pressure. The amino derivative was isolated as white crystals. Yield 100%. Mp 106–110 °C. ir v: 3482 and 3380 (NH₂), 3234 (NH) cm⁻¹. ¹H NMR ppm: 1.92 (s, 3H, Me), 2.33 (s, 3H, Me), 6.00 (s, 1H, pyrazole-CH), 7.30–7.35 (m, 4H, Ar-H), 7.42–7.46 (m, 4H, Ar-H), ¹³C NMR ppm: 12.0, 13.2, 112.6, 119.5, 122.6, 124.8, 125.0, 128.3, 130.3, and 152.9.

4.2.7 Cell Culturing and MTT Assay

Unless stated otherwise, all reagents were from Merck (Milan, Italy) and of the highest purity grade commercially available. All synthesized compounds were dissolved in dimethyl sulfoxide (DMSO) and then diluted in a culture medium so that the effective DMSO concentration did not exceed 0.1% (v/v). HCT-116, MCF-7, K-562, and BALB/3-T3 cell lines were purchased from the American Type Culture Collection, Rockville, MD, USA. Except for BALB/3-T3 cells, which were grown in DMEM, all other cells were cultured in the RPMI-1640 medium. Both DMEM and RPMI-1640 were supplemented with L-glutamine (2 mM), 10% fetal bovine serum (FBS), penicillin (100 U/mL), streptomycin (100 µg/mL), and gentamicin (5 µg/mL). Cells were maintained in the log phase by seeding them twice a week at a density of 3 × 10⁵ cells/mL, in humidified 5% CO₂ atmosphere at 37 °C.

The cytotoxic activity of the synthesized compounds against all the cell lines employed was determined by the MTT colorimetric assay as previously reported [91]. The assay is based on the reduction of 3-(4,5-dimethyl-2-thiazolyl)bromide-2,5-diphenyl-2-H-tetrazolium to purple formazan by the mitochondrial dehydrogenases of living cells. Briefly, cells at the passage that did not exceed the number 20, were seeded into 96-well plates (Corning, New York, NY, USA) at a density of 2.0 × 10⁴ cells/cm², incubated overnight, and then treated with either the compounds or the vehicle (control) for 24 h. Afterwards, the medium was carefully removed,

and 200 μL of 5 mg/mL MTT was added. The supernatant was discarded after a 2 h of incubation at 37 $^{\circ}\text{C}$, and the formazan blue formed dissolved in DMSO. The absorbance at 565 nm of the formazan product was measured using a microplate reader (LTeK, INNO, Seongnam, South Korea), and the value of control cells was taken as 100% of viability. Each experiment was repeated three times in triplicate to obtain the mean values. No differences were found between cells treated with DMSO 0.1% and untreated cells in terms of cell number and viability.

The growth inhibition activity of the tested compounds was defined as the IC_{50} value that represents the concentration of the compound that inhibits 50% of cell viability. IC_{50} values were calculated using the dose-response inhibition model in Prism 8 (GraphPad Software, San Diego, CA, USA). The SI, a measure of the therapeutic potential of the tested compound, was calculated by dividing the IC_{50} for normal BALB/3-T3 cells by the IC_{50} for HCT-116, MCF-7, and K-562 cancer cells.

Two hundred milliliters (6.0×10^4 cells/cm²) of a K-562 cell suspension were plated in each well of a 96-well plate and treated with different concentrations of compounds. An equal volume of DMSO was added to the control well, and the cells were cultured for an additional 24 h; then, 20 μL MTT (5 mg/mL) in the growth medium was added per well, and the plates were incubated at 37 $^{\circ}\text{C}$ for 4 h as reported in [142] with some modifications. Plates were then centrifuged at $400\times g$ for 10 min. Supernatants were removed from the wells, and the reduced MTT dye in each well was solubilized in 200 μL DMSO. Absorbance was measured in a microplate reader (LTeK, INNO, Seongnam, Korea), and the value of the control cells was taken as 100% of viability.

4.3 RESULTS AND DISCUSSION

For the study, the *Tribolium castaneum* full-length catalytic subunit of telomerase in complex with the compound BIBR1532 (PDB 5CQG) was selected. The crystal structure was optimized by completing and refining the entire structure and optimizing amide groups of asparagine (Asn) and glutamine (Gln) as well as the imidazole ring in histidine (His); then, the protonation states of histidine, (His), aspartic acid (Asp), glutamic acid (Glu), and the tautomeric states of histidine were calculated. Starting from the PDB coordinates set of BIBR1532, a static pharmacophore model was created by using LigandScout containing six features (Figure 31A):

- Four hydrophobic features
- One H-bond acceptor
- One anionic feature

The pharmacophore features were decreased to five, the two hydrophobic features on the naphthyl ring of BIBR1532 were fused into one, and the tolerance radius for the new pharmacophore feature was increased by 0.15 \AA to compensate for small deviations (Figure 31B). To improve the performance of the virtual screening (VS) process, the recent MYSHAPE (Molecular dYnamics SHared PharmacophorE) approach was used [51,52]. In this approach, the exploration of the protein conformations by MD coupled with the pharmacophore modeling improved the result of the VS concerning the corresponding model generated from the PDB coordinates set. To build the MYSHAPE model, 20 ns of MD simulation of the BIBR1532–protein complex was run. A new interaction was retrieved such as a hydrogen bond interaction of the carbonyl group of BIBR1532 with the Met482 by a water bridge. The new MYSHAPE pharmacophore feature was added to the original pharmacophore model. In addition, in this case, the tolerance radius for the added pharmacophore feature was in- creased by 0.15 \AA to

compensate for small deviations in the 3D coordinates of the different conformations. The MYSHAPE pharmacophore model had six features (Figure 31C):

- Three hydrophobic features;
- Two H-bond acceptors;
- One anionic feature.

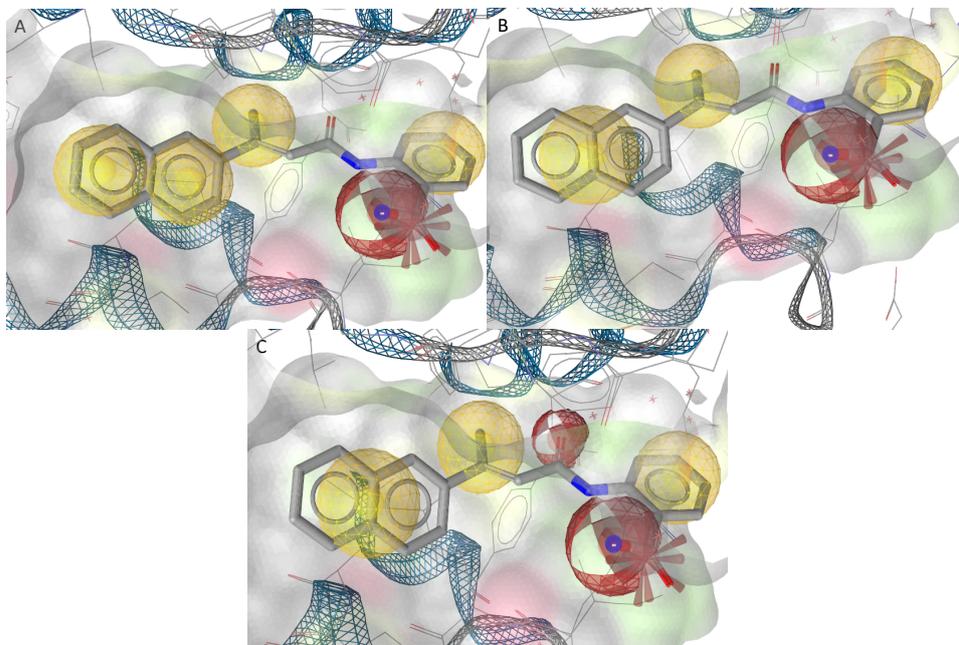


Figure 31: Pharmacophore models derived from the X-ray pose of BIBR1532 (PDB ID: 5CQG): (A) static pharmacophore model from the PDB with six features—4 hydrophobic features (yellow spheres), 1 H-bond acceptor (red sphere), and one anionic feature (red star); (B) static pharmacophore model after fusing with 5 features—3 hydrophobic features (yellow spheres), 1 H-bond acceptor (red sphere), and 1 anionic feature (red star); (C) MYSHAPE pharmacophore model—3 hydrophobic features (yellow spheres), 2 H-bond acceptors (red spheres), and 1 anionic feature (red star)

This model was used to screen the *in-house* library of arylsulfonamide derivatives.

At the same time as the pharmacophore modeling, standard precision (SP) molecular docking studies were performed using Glide [53], considering the conserved hydrophobic pocket (FVYL motif) where BIBR1532 binds. The docking studies were performed centering the docking boxes on the 3D coordinates of BIBR1532. The RMSD of BIBR1532 was calculated showing a value of 0.2 Å. The residues that characterized the binding site were Phe478, Met482, Met483, Arg486, Phe494, Gly495, Ile497, Trp498, Ile550, Tyr551, Gly553, Lys552, Leu554, and Arg557. BIBR1532 established an H-bond interaction and a salt bridge between the oxygen of the amide group and Arg486 and aromatic H-bond interactions between the naphthyl ring and Phe494 and Ile550 (Figure 32).

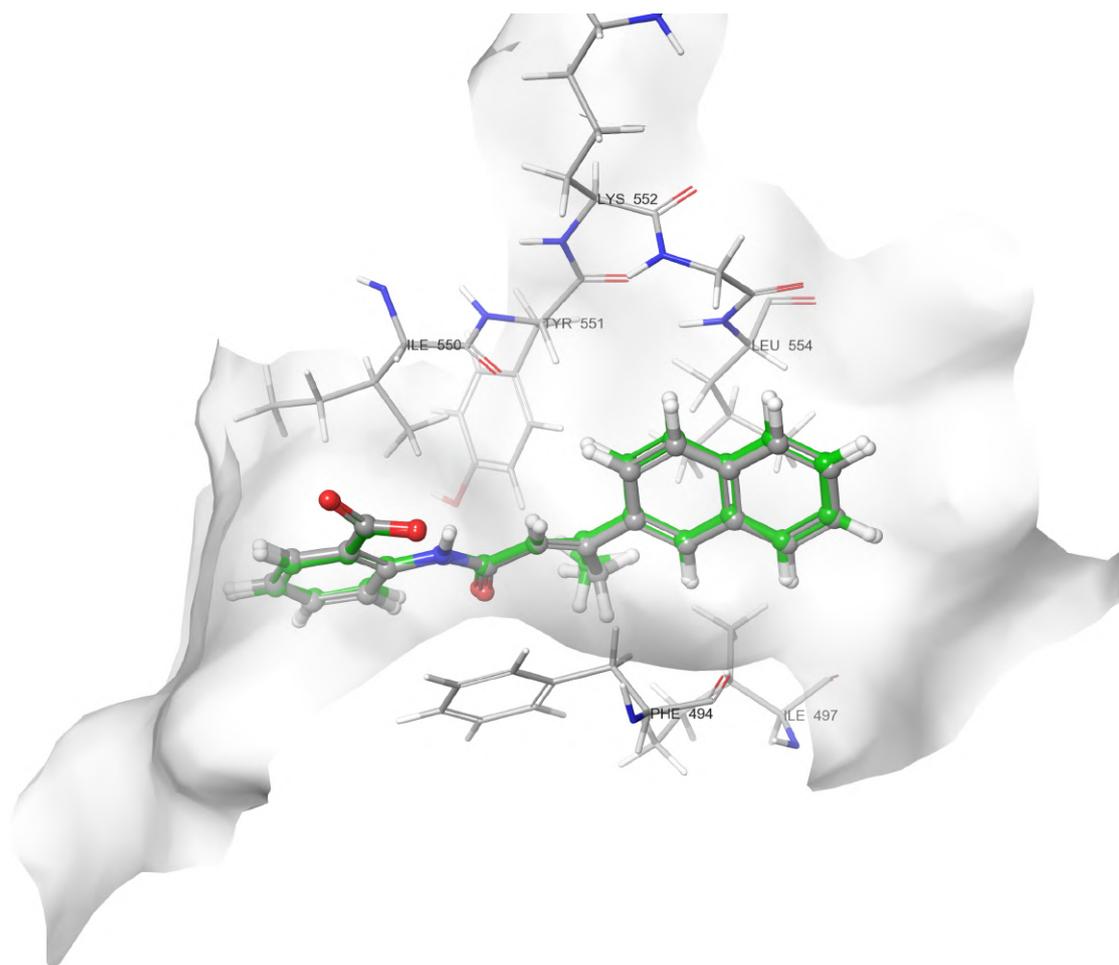


Figure 32: Binding mode after re-docking of BIBR1532 (grey) superimpose on BIBR1532 cocrystallized (green): yellow, dashed lines represent the H-bond interaction and, the purple, dashed lines represent the salt bridge between the O of the carboxyl group and Arg486, and the light blue dashed lines represent the aromatic H-bond interactions between the naphthyl ring and Phe494 and Ile550

The *in-house* library was used to perform the VS by docking. MM-GBSA-binding free energy calculations were performed for the molecules obtained from the VS and compared with BIBR1532. Hit compounds that exhibited ΔG binding values (1A, 1B, 1C, 1E, 1G, 2B, and 2C, $-70.75/-62.97$ kcal/mol) lower than that obtained for BIBR1532 (-62.76 kcal/mol) were selected for synthesis and *in vitro* tests together with hit compounds retrieved using the MYSHAPE approach (1D, 1F, and 2A) (Figure 32).

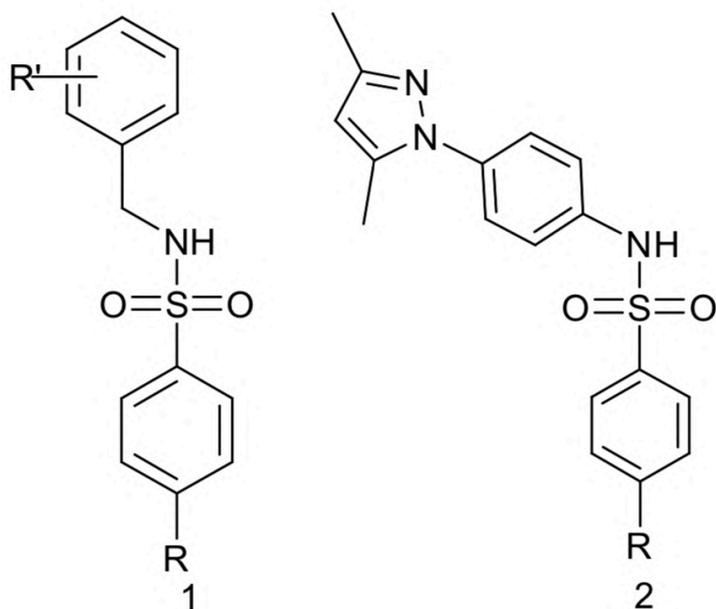
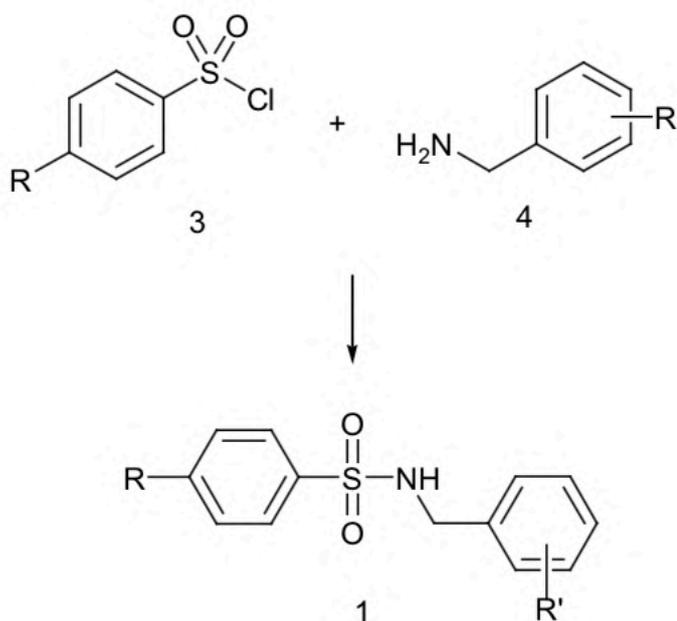


Figure 33: Structures of sulfonamides 1 and 2 prepared in this work (R = Me, NHAc; NO₂; NH₂), (R' = 4 Cl; 4-OMe; 2,5-diOMe)

4.3.1 Synthesis

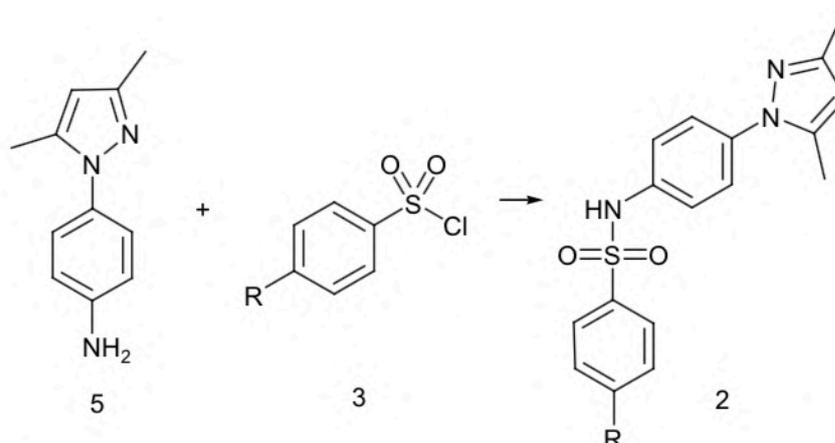
For the preparation of sulfonamide compounds of type 1 and 2 (Figure 33), several procedures are present in the literature [143, 144, 138], and a few of them were also proposed recently in light of using more environmentally friendly green chemistry approaches [145, 146].

Our approach to targeting derivatives of type 1 and 2 involved reactions of the sulfonyl chloride 3 and suitable benzylamines 4 or amine 5, according to Scheme 2 and Scheme 3. The benzylamines were commercially available, whereas the 1-(4-aminophenyl)-3,5-dimethylpyrazole (5) was prepared in two steps from acetylacetone and 4-nitrophenylhydrazine and the subsequent reduction of the nitro group with H₂ and Pd/C, using standard literature methodology.



1	R	R'
A	Me	4-Cl
B	NHAc	4-Cl
C	NO ₂	4-Cl
D	Me	4-OMe
E	NHAc	4-OMe
F	Me	2,5-di-OMe
G	NO ₂	2,5-di-OMe

Scheme 2: The general method for the synthesis of benzyl-sulfonamides of type 1



Scheme 3: General method for the synthesis of sulfonamide 2, R = Me (2A); NO₂ (2B); NH₂ (2C)

In all cases, initially, the reactions were carried out under classical literature conditions, according to Methods X or Y (see Section 4.2.6) affording derivatives 1 in good to high isolated yields (64–97%) and slightly lower (49%) only in the case of 1E, probably because of the solubility problem of the starting material.

The aminophenyl-3,5-dimethylpyrazole 5 also reacted with sulfonylchloride 3 in anhydrous THF in the presence of an equimolar amount of triethylamine, affording sulfonamides 2A, B (Scheme 3).

The possibility of avoiding the use of dangerous/dry solvents was also considered. The pilot reaction of benzylamine and para-toluensulfonyl chloride, in Schotten-Baumann conditions under pH control with Na₂CO₃, as discussed in [147], was explored, but the result was the incomplete conversion to the desired sulfonamide. Recently, a new environmentally safe methodology was reported for the preparation of a wide range of aliphatic- and aryl-sulfonamides [148], and although the methodology was not extended to investigate the reactivity of benzylamines, it was decided to explore whether this kind of amine resulted in a suitable substrate to prepare sulfonamides via this route. The green method afforded yields from good to quantitative, giving in some cases a similar efficient conversion of the reactants into products, if compared to Method X (Table 6). The major concern of this methodology is the competitive reactions leading to bis-sulfonylation which, however, in the case benzylamines were not observed, suggesting that this procedure has preferable results, especially when electron-withdrawing groups are present in both the reactants (as in the case of entry 1C).

Table 6: Comparison of the yields obtained using classical and green chemistry methods.

1	R	R'	Method X	No Solvent*
C	NO ₂	4-Cl	78%	100%
D	Me	4-OMe	100%	60%
E	NHAc	4-OMe	80%	70%

*According to [148], the amine (2 mmol) and anhydrous NaHCO₃ (approximately 1 g) were ground together into a fine powder, and arylsulfonyl chloride (2 mmol) was added under vigorous stirring at room temperature. The progress of the reaction was monitored by TLC until the conversion of the amine was completed.

Compound 2C was prepared by reduction of the corresponding nitro compound with H₂/Pd in ethanol a quantitative yield.

4.3.2 Cytotoxic Activity and Selectivity Index

To select the most promising cytotoxic agents, the newly synthesized compounds were initially evaluated on human colorectal carcinoma (HCT-116) [69,70], human breast adenocarcinoma (MCF-7) [149, 150], and human chronic myeloid leukemia (K-562) [71,72] cell lines that express the active telomerase.

As shown in Table 7, except for compound 2C, the antiproliferative efficacy of the compounds was modest and only evident at the highest micro-submillimolar concentrations. Moreover, compounds 1B, 1D, 1E, 1F, 1G, and 2B exhibited solubility issues that prevented the evaluation of their cytotoxicity.

Table 7: Cytotoxicity of the synthesized compounds on human tumor cell lines HCT-116, MCF-7, and K-562. Cells were treated for 24 h in the absence or the presence of the compound tested at the indicated concentration, and cell viability was assessed by an MTT test. Values are expressed as the mean \pm SD of three separate experiments conducted in triplicate. * Sol: solubility issue.

Compound	% Inhibition Cell Proliferation														
	HCT-116					MCF-7					K-562				
	Compound Concentration (μ m)														
	10	75	150	300	400	10	75	150	300	400	10	75	150	300	400
1A	0	0	19 \pm 2	29 \pm 3	40 \pm 2	0	0	0	12 \pm 1	13 \pm 1	0	0	0	33 \pm 2	38 \pm 2
1B	0	5 \pm 0.2	23 \pm 2	* Sol	* Sol	0	15 \pm 1	21 \pm 2	* Sol	* Sol	0	5 \pm 0.3	21 \pm 1	* Sol	* Sol
1C	0	0	2 \pm 0.1	8 \pm 0.2	15 \pm 2	0	5 \pm 0.1	15 \pm 1	19 \pm 1	30 \pm 2	0	12 \pm 1	13 \pm 1	16 \pm 1	25 \pm 1
1D	0	0	0	* Sol	* Sol	0	0	0	* Sol	* Sol	0	0	0	* Sol	* Sol
1E	0	5 \pm 0.1	* Sol	* Sol	* Sol	0	20 \pm 2	* Sol	* Sol	* Sol	0	16 \pm 1	* Sol	* Sol	* Sol
1F	0	0	6 \pm 0.3	* Sol	* Sol	0	0	0	* Sol	* Sol	0	0	0	* Sol	* Sol
1G	0	5 \pm 0.2	28 \pm 2	* Sol	* Sol	0	12 \pm 1	35 \pm 2	* Sol	* Sol	0	2 \pm 0.1	37 \pm 3	* Sol	* Sol
2A	0	5 \pm 0.1	30 \pm 1	44 \pm 2	55 \pm 3	0	0	0	0	0	0	13 \pm 1	42 \pm 2	55 \pm 3	60 \pm 3
2B	0	0	37 \pm 2	* Sol	* Sol	0	17 \pm 1	34 \pm 2	* Sol	* Sol	0	19 \pm 1	23 \pm 2	* Sol	* Sol
2C	15 \pm 2	31 \pm 2	56 \pm 2	75 \pm 2	91 \pm 4	10 \pm 1	19 \pm 3	59 \pm 2	76 \pm 2	85 \pm 2	35 \pm 2	70 \pm 3	76 \pm 1	80 \pm 3	85 \pm 2
BIBR-1532	25 \pm 2	47 \pm 3	56 \pm 3	89 \pm 2	95 \pm 5	12 \pm 2	23 \pm 2	57 \pm 3	74 \pm 2	91 \pm 3	10 \pm 2	19 \pm 3	32 \pm 1	62 \pm 3	79 \pm 2

On the other hand, compound 2C revealed an interesting activity, and it was possible to calculate its IC₅₀ (Table). Interestingly, this value on K-562 cells appeared 3.6-fold lower than that on HCT-116 and 4.2-fold lower than that on MCF-7 cells. Moreover, in the K-562 cell line tested, compound 2C showed an IC₅₀ 6.8-fold lower than the reference compound BIBR1532, and an IC₅₀ comparable to BIBR1532 in the MCF-7 cell line.

Table 8: IC₅₀ of 2C against HCT-116, MCF-7, K-562, and BALB/3-T3 cell lines. Values were calculated by plotting the percentage viability versus concentration on a logarithmic graph and are expressed as the mean \pm SD of three separate experiments conducted in triplicate.

Compound	IC ₅₀ (μ m)			
	HCT-116	MCF-7	K-562	BALB/3-T3
BIBR-1532	70 \pm 4	120 \pm 8	208 \pm 11	N/A
2C	109 \pm 8	127 \pm 9	33 \pm 4	324 \pm 15

Selective cytotoxicity is a pivotal requirement for anticancer drugs. To determine the selectivity of compound 2C, its cytotoxicity against the cancer cell lines employed (i.e., HCT-116, MCF-7, and K-562) was compared with that against the non-cancerous, murine, embryonic, fibroblast cell line BALB/3-T3. As shown in Table 8, the IC₅₀ of compound 2C on BALB/3-T3 cells was remarkably higher than those on the tumoral cell lines tested. Moreover, the calculated selectivity indexes (SIs) of compound 2C for HCT-116, MCF-7, and K-562 cells were 2.9, 2.5, and 9.8, respectively. Relevantly, these values are above the accepted threshold for antitumor drugs (SI = 2.0) [151, 152].

Although selectivity for cancer cells cannot be easily derived from the comparison of toxicity parameters in different cell cultures, these data indicate that compound 2C shows preferential toxicity towards cancer cells. Along these lines, the potential use of compound 2C as a novel lead molecule for the development of more potent and selective antiproliferative agents can, therefore, be envisaged.

4.3.3 Docking and MD simulation of compound 2C

The analysis of the best docking pose of compound 2C showed a pi–pi stacking interaction between Phe494 and the central aromatic ring as observed in BIBR1532, where the same residue is involved in two H-aromatic bonds with the naphthyl ring (Figure 32). The same

residues Phe494, Asp493, and Gly495 establish positive van der Waals (vdW) contacts with the aniline group of 2C. Other positive vdW contacts are formed by the sulfonamide moiety of 2C and Ile550, Gly553, and Leu554. A major part of positive vdW contacts are established by the dimethylpyrazolyl moiety with Met482, Met483, Arg486, Phe494, and Ile550 (Figure 34).

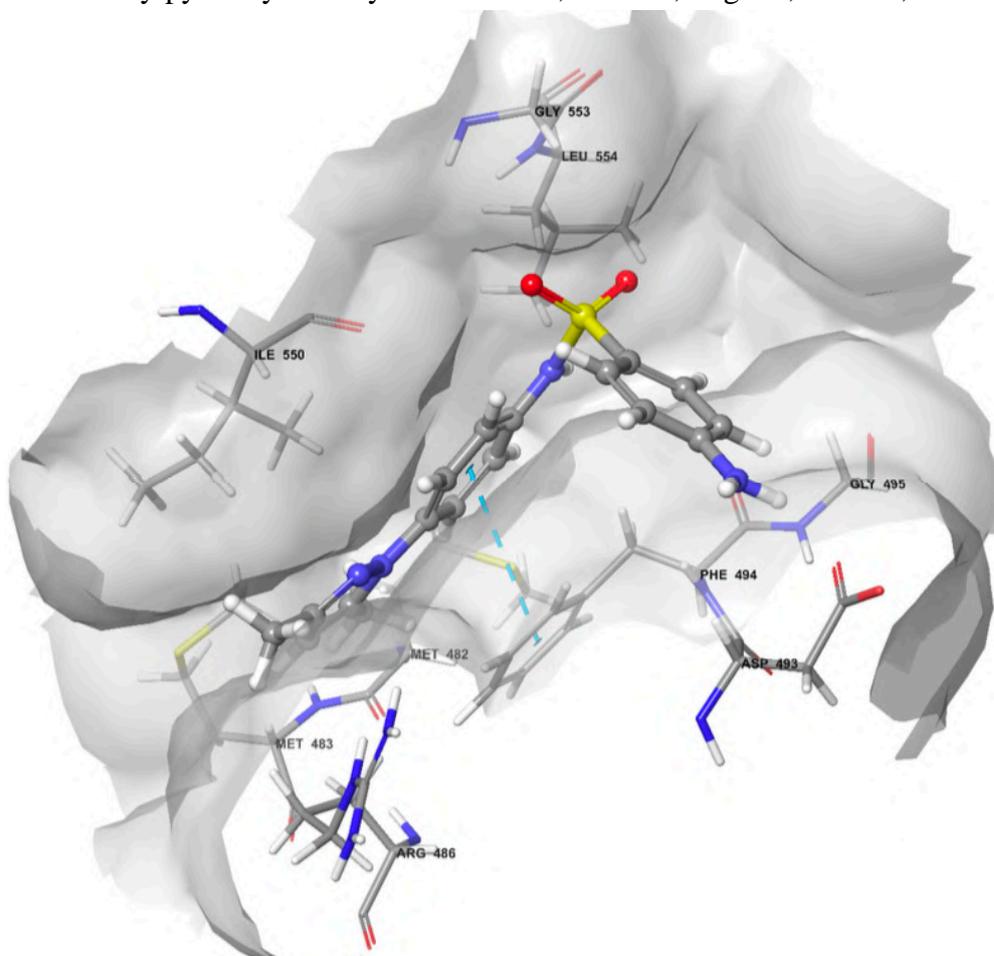


Figure 34: Binding mode of the best docking pose of 2C after the docking study. The light blue, dashed lines represent the pi-pi stacking. The displayed residues establish positive vdW contacts with the compound

Subsequently, 100 ns MD simulation of compound 2C at the telomerase active site showed good stability and evidenced Phe494 as the key residue also in BIBR1532. But the more interesting evidence of the simulation regarded that each moiety of compound 2C was involved in key interactions with some residues of the active site: the pyrazole moiety in a cation-pi-stacking interaction with Arg557, the two phenyl rings in the pi-stacking interaction with Phe494, and the sulfonamide moiety in the H-bond interaction with Ile550 and Gly553 (Figure 35).

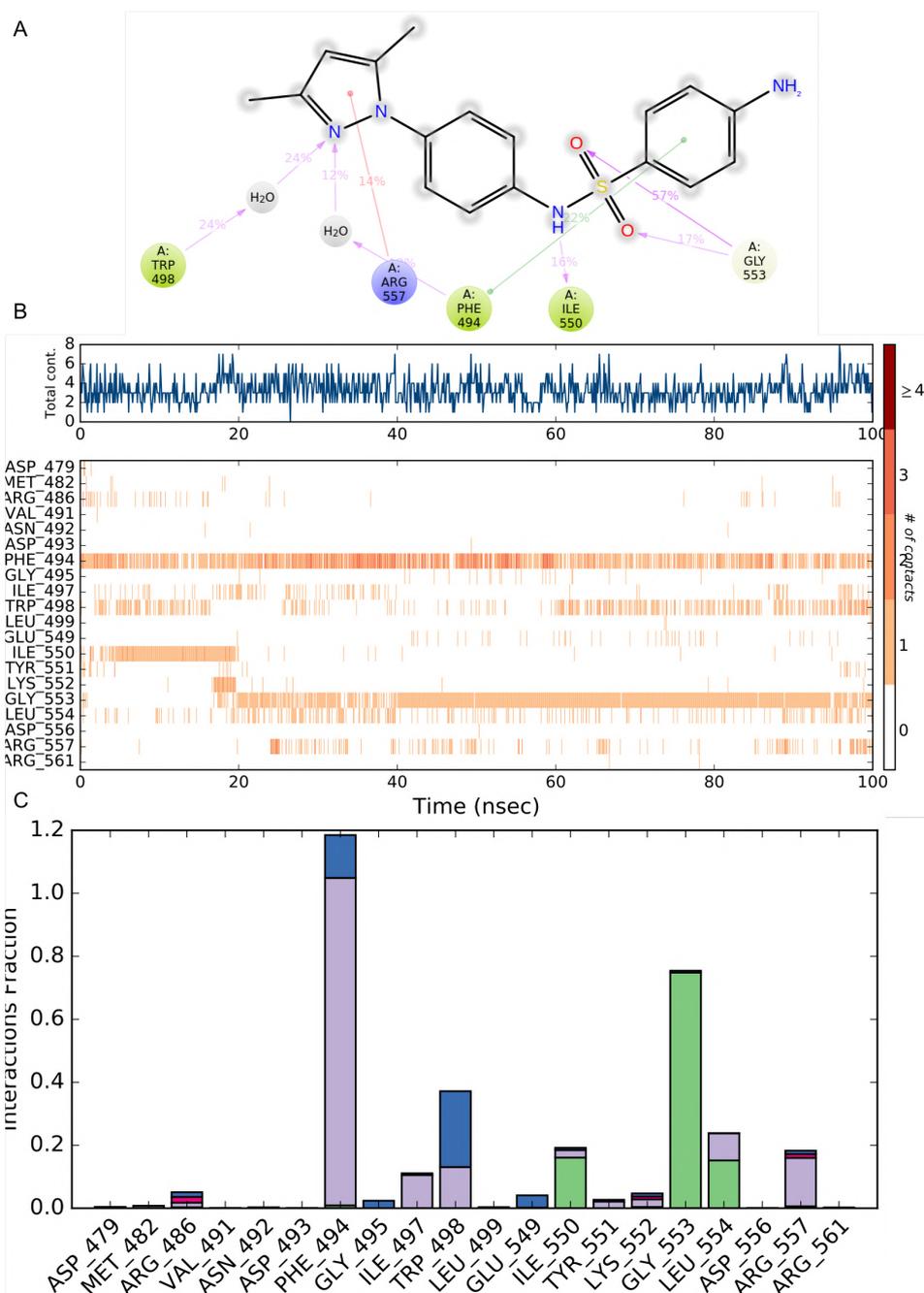


Figure 35: Protein–compound 2C contacts: (A) the interactions that occur for more than 10% of the simulation time are reported; (B) a timeline representation of the interactions and contacts (H-bonds, hydrophobic, ionic, and water bridges) during the simulation; (C) histogram of protein–compound 2C interactions fractions: hydrogen bonds (green), hydrophobic (purple), ionic (red), and water bridges (blue)

The system reached equilibrium quickly and fluctuated around the average RMSD value of <3 Å. The average value of ligand vs. protein RMSD of ~ 6.4 Å indicated strong stability of 2C in the binding pocket compared to the ligand vs. ligand RMSD of ~ 1.81 Å (Figure 36). The analysis of the trajectory evidenced that the pi-stacking interaction between the aniline ring of 2C and Phe494 occurred at 57%. This output evidence a shifting in the pi-stacking interaction with respect to the docking where the central phenyl ring was involved, but it clearly defines Phe494 as a key residue in the inhibition pattern. As also reported in Figure 34B, C, Phe494 established at least one contact for the entire duration of the simulation and, for a major part of the time, 2–3 contacts with the phenyl ring of 2C. Gly553 showed for 57% and 17% of the time two H-bonds interactions with the oxygen of the sulfonamide group emerging after ~ 18 ns of

simulation. Interestingly, after 22 ns of the simulation a new π -cation interaction appeared for 14% of the time between Arg557 and the pyrazole ring, even though it was in a spotted fashion. During the first 20 ns (16% of the total time), the protein–ligand interaction stabilized with an H-bond between Ile550 and the NH of the sulfonamide group. It is interesting evidence of the role of a water molecule as a water bridge between Trp498 and the pyrazole ring. Other residues involved in the protein–ligand interaction but with a minor role were Arg486, Ile497, Trp498, Tyr551, Lys552, and Leu554.

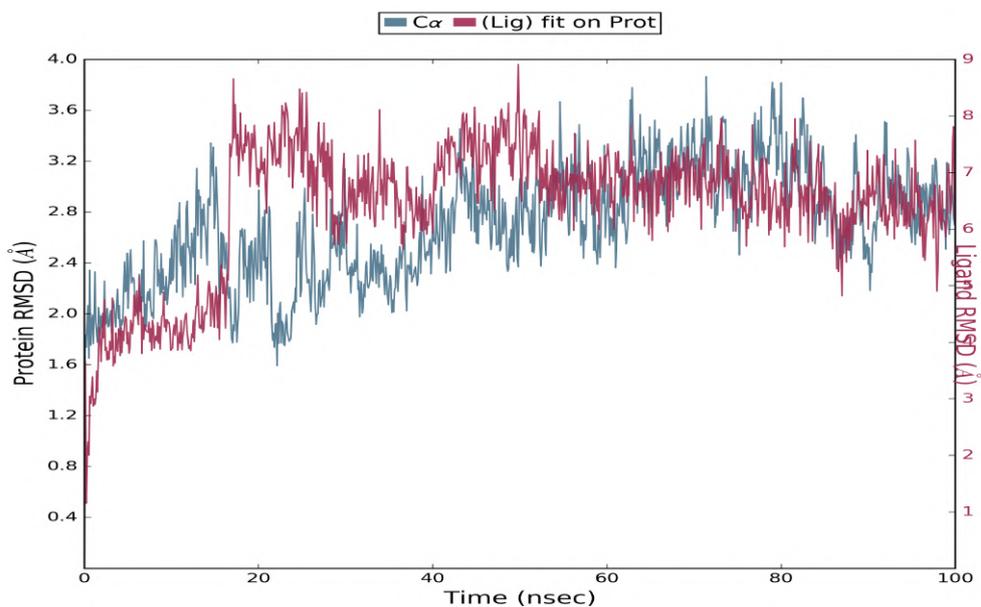


Figure 36: Compound 2C and protein RMSD during MD simulation (100 ns)

4.3.4 ADME Calculation for Compound 2C

ADME calculation for compound 2C was performed using the SwissADME web tool [135]. In the hexagon drug-likeness graph (Figure 37), each vertex represents a parameter that defines a bioavailable drug. The pink regions represent the optimum range of the following six properties: lipophilicity = 2.53 (XLOGP3 between -0.7 and $+5.0$), size = 342.42 g/mol (MW between 150 and 500 g/mol), polarity = 98.39 Å² (TPSA between 20 and 130 Å²), solubility (log S not higher than 6), saturation = 0.12 (fraction of carbons in the sp³ hybridization not less than 0.25), and flexibility (no more than nine rotatable bonds). It was found that compound 2C was slightly outside the pink area on one side due to the inconformity of its insaturation. The pharmacokinetics analysis showed that compound 2C is probably not a P-glycoprotein (P-gp) substrate and not a BBB permeant. It has potentially good gastrointestinal (GI) absorption and ABS, up to nearly 70%. The %ABS, a very functional physiochemical variable, defines a drug's transport properties. It was calculated according to the equation $\%ABS = 109 - (0.345 \times TPSA)$ [153, 154]. TPSA values below 98.39 Å² characterize a significant permeability in the cellular plasma membrane. It has a bioavailability score of 0.55, which means good pharmacokinetic properties according to the Rule of 5 by Lipinski [155].

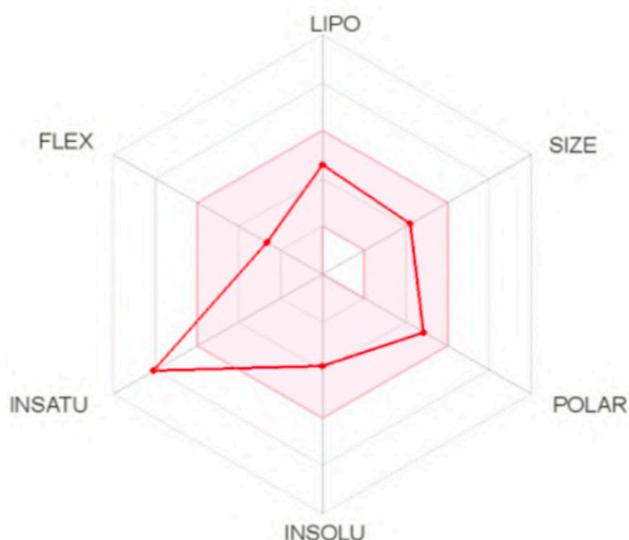


Figure 37: The bioavailability radar (the pink area exhibits the optimal range of a particular property) for compound 2C (LIPO = lipophilicity as in XLOGP3; SIZE = size as in molecular weight; POLAR = polarity as TPSA (topological polar surface area); INSOLU = insolubility in water by log S scale; INSATU = insaturation per fraction of carbons in the sp^3 hybridization; FLEX = flexibility per rotatable bonds)

4.4 CONCLUSION

In the past decade, there has been tangible progress in the definition of the role of telomerase in tumor progression and its promising target for developing cancer therapeutics.

With the aim of identifying the most promising cytotoxic agents, a combined structure-based strategy was developed to screen an *in-house* library of 126 aryl sulfonamide derivatives. Firstly, the MYSHAPE approach was used to identify the hit compounds. Successively, docking and calculation of the binding free energy using the MM-GBSA method identified other hits. One of the selected hits (2C) revealed the IC_{50} results were comparable with BIBR1532 for the HCT-116 cell line and the MCF7 cell line, but several times lower in the K-562 cell lines.

These findings were confirmed by the evidence of the selective cytotoxicity against the non-cancerous, mu-rine, embryonic, fibroblast cell line BALB/3-T3. Also, compound 2C showed a good calculated pharmacokinetic profile. Further, the MD simulation of compound 2C at the telomerase active site showed good stability and evidenced Phe494 as the key residue also in BIBR1532. In conclusion, compound 2C is an interesting hit compound that deserves, for the reasons stated above, further investigation in terms of the optimization of its structure to obtain more active telomerase inhibitors.

5. COMPARING MOLECULAR DYNAMICS DERIVED PHARMACOPHORE MODELS WITH DOCKING: A STUDY ON CDK-2 INHIBITORS

5.1 INTRODUCTION

In this study, a comparative study between two known MD-derived pharmacophore modeling protocols, the Common Hit Approach (CHA) and the Molecular dYnamics SHARed PharmacophorE (MYSHAPE) approach [34, 35] with classical semi-flexible constrained/unconstrained docking using CDK2 inhibitors as a case study was performed. Ligand-based and structure-based pharmacophore models represent a powerful description of the target-ligand interaction patterns [156]. Consequently, pharmacophore models can be used to perform virtual screening of molecular libraries [157]. Pharmacophore models together with docking analysis present a unique opportunity to harvest the enormous available chemical space for drug-like molecules [158–161]. It is now widely known that proteins and small molecules are dynamic entities, which can perform a wide range of motions, reflecting the vibrations of individual bonds and collective large structural movements [162,163].

For this reason, the use of a single frame of a dynamic system provides scarce information about the conformational flexibility of the ligand and about the motions of the residues near the binding pocket [164]. The incorporation of dynamic features in pharmacophore modeling represents the new frontier, and in the recent past, different approaches were explored [165–170]. This study was performed using CDK2/inhibitors complexes because cyclin-dependent kinases are one of the most attractive targets for oncology [171–174]. Moreover, the presence of the Protein Data Bank web page (www.PDB.org) [175] of over 140 complexes bound with an inhibitor in the ATP binding pocket aided the selection of the complexes allowing this study to be comprehensive.

The CHA and MYSHAPE approaches perform consensus pharmacophore-based virtual screening on the conformational ensemble of the protein-ligand complex obtained by means of MD simulations. Pharmacophore models were created using the common feature pattern of the ligands during the MD simulations following the CHA workflow. The best performing four PDB target-ligand complexes in CHA were selected: 3QTU, 3RK7, 4FKT, 4FKV. Thereafter, further pharmacophore models were created according to the MYSHAPE workflow. In the end, classical docking studies were performed. The models were used to carry out a virtual screening on a dataset of actives and decoys, to define if the use of dynamic features allows to enrich and increase the screening capability compared to the results obtained from CHA and MYSHAPE approaches. The validation process points out which modelling technique could be more suitable and able to predict a CDK-2 inhibitory activity and give the guidelines for other prospective studies.

5.2 MATERIALS AND METHODS

5.2.1 Preparation of molecular structures

Selection of the 149 CDK2/inhibitor complexes, preparation of the proteins, MD simulations settings are reported in our previous paper [176]

5.2.2 Conversion of MD trajectories

The output files of MD trajectories were processed using VMD software (Visual Molecular Dynamics) [177]. MD complexes were desolvated, and ions eliminated in order to allow a better analysis of the only interactions between ligand and protein. Then, the .cms (Composite Model System) files output format of the trajectories were converted to .dcd files, representing the trajectories of the complexes. The settings are reported in our previous paper [176].

5.2.3 Common Hits Approach (CHA) pharmacophore models

Pharmacophore Models Generation by means of the Common Hits Approach (CHA) are reported in previous papers [176, 178]

5.2.4 Molecular dYnamics SHARED PharmacophorE (MYSHAPE) models

Pharmacophore Models Generation by means of the MYSHAPE is reported in a previous paper [35]

5.2.5 Docking

Using the pharmacophore interaction pattern obtained from the MD simulations, constraints were set on the docking grid. Specifically, positional constraints were imposed considering aromatic interactions, hydrogen bonds, and hydrophobic interactions with the ligand according to the Glide grid constraints panel workflow. For each ligand–protein system analyzed, a docking grid with and without constraints was generated. Standard precision (SP) molecular docking with and without constraints was performed using Glide 4.7 [95]. Ligands were considered according to the setting of previous chapters.

5.2.6 Virtual screening

To validate the virtual screening performance of the pharmacophore models and the docking calculation, a screening library of active and decoy compounds were retrieved from the DUD-E site [90] containing molecules (676 Active and 28,121 Decoys) generated specifically for the CDK-2 and filtered to remove duplicates. The Dataset obtained was prepared as a Virtual-Screening Database, in .lbd format, through Idbgen (extension present in the LigandScout 4.2 package) which allowed obtaining the best conformation of the ligand (at low energy) between the 200 the application can calculate. The tautomers were considered as separate molecules and those molecules that were duplicated or whose conformation calculation had failed were eliminated. The same dataset was used for docking calculation. Receiver Operating Characteristics (ROC) graphs [179–180] were generated, and the Area Under the Curve (AUC) was calculated to validate the virtual screening performance of the pharmacophore models. ROC graphs and values were also used to evaluate the virtual screening capability of the docking runs. In contrast to pharmacophore models, the AUC was not calculated for different fractions of the screening database, but a numeric representation of the ROC area underneath the curve was obtained. This ROC value can be interpreted as the probability that an active compound will appear before an inactive compound, and is calculated as follows in Eq. (2):

$$ROC = \left(\frac{AUAC}{R_i} - \frac{R_a}{2R_i} \right)$$

in which AUAC is the area under the accumulation curve, R_i is the ratio of inactive molecules to the total number of compounds in the screening library, and R_a is the ratio of active compounds to the total number of entries in the screening library. The ROC values were generated using the “enrichment calculator” python script provided by Schrodinger.

5.3 RESULTS AND DISCUSSION

The snapshots saved during the MD simulations plus the coordinates of the PDB file were processed according to CHA and MYSHAPE. According to CHA, the pharmacophore models were generated starting from every single snapshot, and subsequently, a feature vector (represented as a bit string) was generated for each pharmacophore model. This procedure allows aggregating the distinct vectors by counting how many times that particular vector, which is a specific combination of pharmacophore features, was identified during MD simulations [180]. Then, this number will be shown as the number of appearances (appearance count). Therefore, instead of using 1000 individual feature vectors, a smaller number of distinct feature vectors was obtained, which were observed one or more times during MD simulations. Distinct feature vectors that were observed at least 2 times are named Representative Pharmacophore Models (RPMs). The nearly 30,100 RPMs obtained from 149 MD ligand-protein complexes were screened against a validation dataset obtained from the DUD-E site containing molecules (676 Active and 28,121 Decoys) generated specifically for the CDK-2. For each RPM, a hit-list was obtained, so that for each ligand-protein complex several hit-lists were obtained. The multiple RPM hit-lists were combined into a single list named RPM-HIT-LIST, consisting only of unique compounds. The molecules in this list have been ranked according to how many times they are present in the hit-lists. For example, if a molecule is present in many hit-lists it is classified with a higher score than one that appears only in a few hit-lists [176]. To assess the performance of the pharmacophore model in distinguishing between Active and Decoys of DUD-E screening dataset, ROC curves were calculated and analyzed plotting the number of True Positives (TPR) on False Positives (FPR). The performance of RPM models was evaluated using ROC curve at 5% and 100%, considering for each complex the first 50 hits.

The best performance crystallographic complexes (4FKV, 3QTU, 4FKT, 3RK7) were selected to perform the MYSHAPE approach and to compare MYSHAPE versus CHA performance. As described, the protein/ligand interactions were investigated searching for a common pattern (Figure 39 and Figure 39).

Figure 39 shows the four pharmacophore models obtained from the best PDB complexes. The model of 4FKV shows three HBDs, four HBAs, and two hydrophobic interactions. The 3QTU pharmacophore model shows three HBDs, three HBAs, and two hydrophobic interactions. The 4FKV model has an HBD, three HBAs, and two hydrophobic interactions. The 3RK7 model shows three HBDs, two HBAs, and two hydrophobic interactions. The common interaction pattern was used to generate a shared model of PDBs. The shared model has one HBD, one HBA, and one hydrophobic interaction. After MD simulations, new features were retrieved for each ligand-protein complex. Figure 39 shows the four MD's pharmacophore models of 4FKV, 3QTU, 4FKT, and 3RK7. In the orange circles, the newly identified features are depicted. During the MD simulation, for 4FKV two more features were found with respect to the PDB: one HBA for the nitrogen (imidazole) with Lys33; one HBD for the NH (imidazole) with Val163. In the case of 3QTU, two new features are evidenced: HBA for sulfone oxygen with Thr14; one HBD for primary amine with Leu83. For 4FKT, one new feature is found: HBD for the amine with Leu83. No new features resulted in 3RK7.

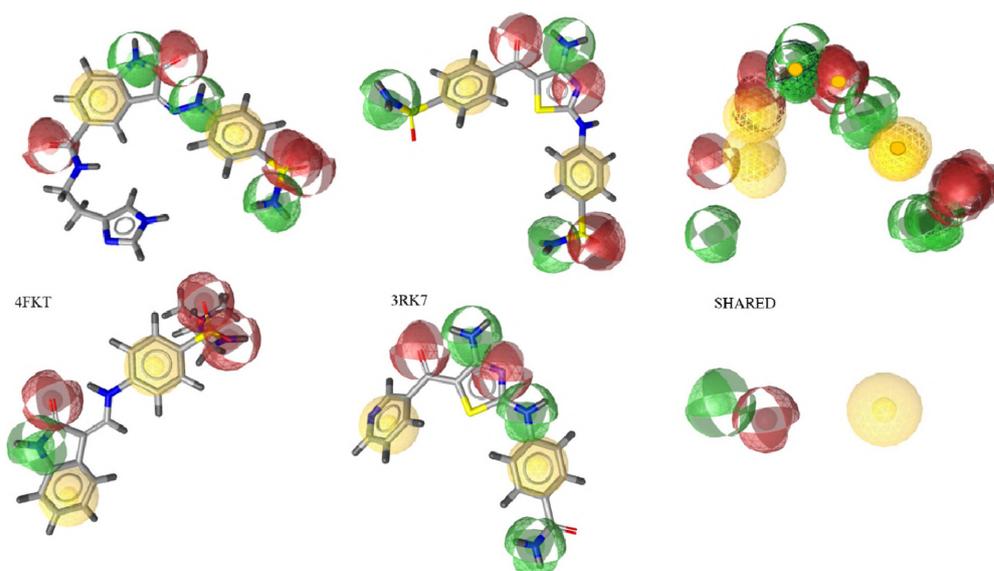


Figure 38: PDB obtained pharmacophore models. Green spheres for HBD (hydrogen bond donor), red spheres for HBA (hydrogen bond acceptor), and yellow spheres for hydrophobic interaction. The common features retrieved from the PDB models are depicted by a full orange circle

The MYSHAPE model, obtained after MD simulations, has one more feature with respect to the shared model of PDBs. (Figure 39)

The virtual screening performance of five different pharmacophore models (PDB models and MYSHAPE model) was carried out for every system by using the same screening library used before for the CHA Table 8 shows the four pharmacophore models created by using the crystal structure, and the generated MYSHAPE model, together with their virtual screening performance. Virtual screening results were evaluated using the ROC-AUC values for 5% and 100 % of the screening database. Analyzing the results, the ROC values for the five pharmacophore models (4FKV, 3QTU, 4FKT, 3RK7, and MYSHAPE), MYSHAPE shows high performance for 5%, 100%. The default PDB pharmacophore models retrieve a lower number hit- list (9, 11, 36, and 50 hits, respectively) than the MYSHAPE hit list (3582 hits). The calculated ROC values from the MYSHAPE model have values much higher than the default pharmacophore models obtained from the PDB structures.

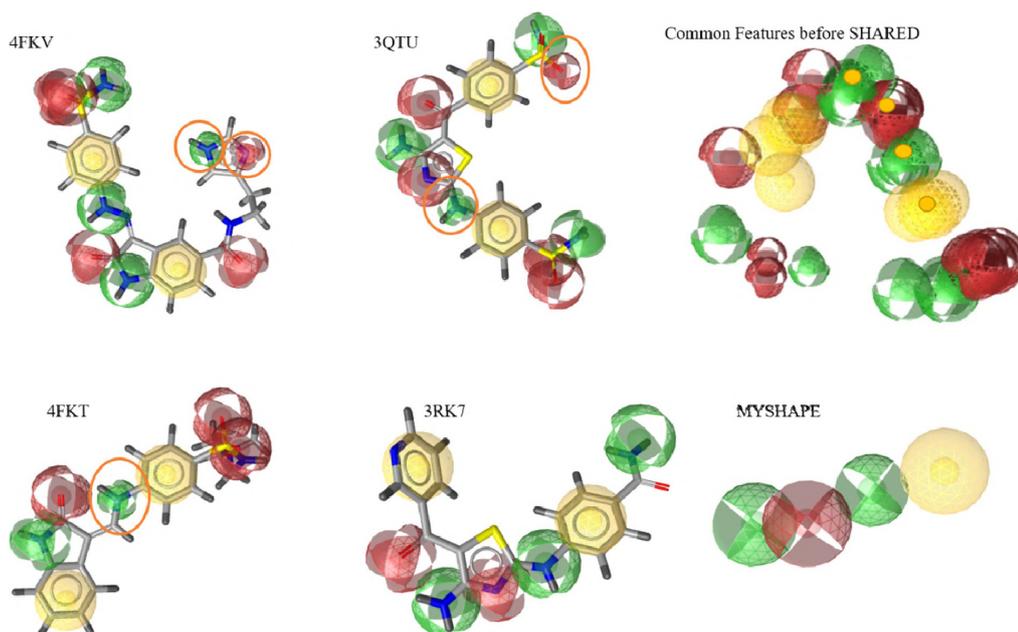


Figure 39: Pharmacophore models obtained after MD. Green spheres for HBD, red spheres for HBA and yellow spheres hydrophobic interaction. The new features retrieved are depicted in orange circle. The shared features are depicted by a full orange circle

Docking grids were constructed for the four different investigated–ligand–protein complexes with CDK-2. These were built with constraints derived from the MD simulations or without constraints. The SP docking was used to analyze the change in the virtual screening capability of using grids with constraints. ROC values were used as metrics to compare the performance of the approaches. Table 9 explains the values of SP docking with and without constraints, for each complex.

The SP docking outcomes are satisfactory for all investigated systems. Looking at the ROC graphs for all four complexes, the SP docking favours active molecules over inactive compounds. The ROC graph for 3QTU shows the best virtual screening performance of the systems (0.92 and 0.97, without and with constraints respectively). The use of grid constraints improved the virtual screening capability, for two systems over four (3QTU and 4FKT), especially for the 4FKT model which showed the maximum value of identified actives.

As mentioned, CHA showed an early enrichment ($AUC_{5\%}$) over 0.97 in the four best-performing systems (PDB: 4FKV, 4FKT, 3QTU, 3RK7) which were used to carry out another pharmacophore modelling approach such as MYSHAPE. The application of the MYSHAPE model showed an interesting increase in the screening early recognition capability ($ROC_{5\%}$) both in terms of sensitivity of the model and specificity when compared to the four PDB models. Last but not least, MYSHAPE features have been used as constraints to be used in docking calculation. The use of these interactions pattern to create the docking grid showed an improvement in the early recognition of actives compounds, in particular for 4FKT which showed a better improvement in the ROC value. Even if it should be considered that each system showed higher values when the constraints have been applied to docking calculation. Comparing the two pharmacophore modeling approaches together, and the docking adding MD constraints versus classical approach such as pharmacophore screening starting by the PDB structures and docking, the application of information derived from MD strongly improve the ROC values, in particular related to CHA and MYSHAPE (Table 10). Docking with constraints outcomes slightly improves with respect to docking without constraints.

Table 8: Default PDB pharmacophore and MYSHAPE models with their screening parameters. Red spheres = HBA, Green spheres = HBD, Yellow spheres = hydrophobic feature, Blue ring = aromatic feature

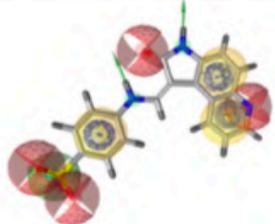
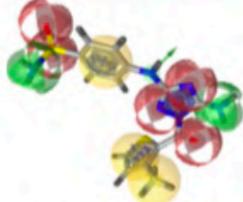
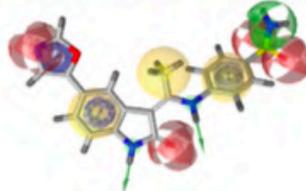
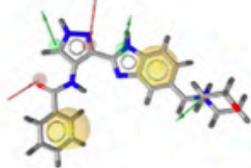
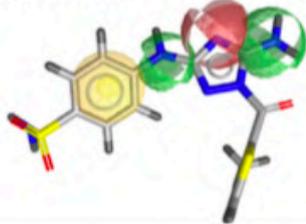
PDB ID	Pharmacophore Model	hits	ROC 5%	ROC 100%
4FKV		9	0.63	0.51
3QTU		11	0.63	0.51
4FKT		36	0.94	0.52
3RK7		50	0.76	0.55
	MYSHAPE	3582	0.99	0.58
				

Table 9: ROC curves obtained with docking without constraints and with constraints

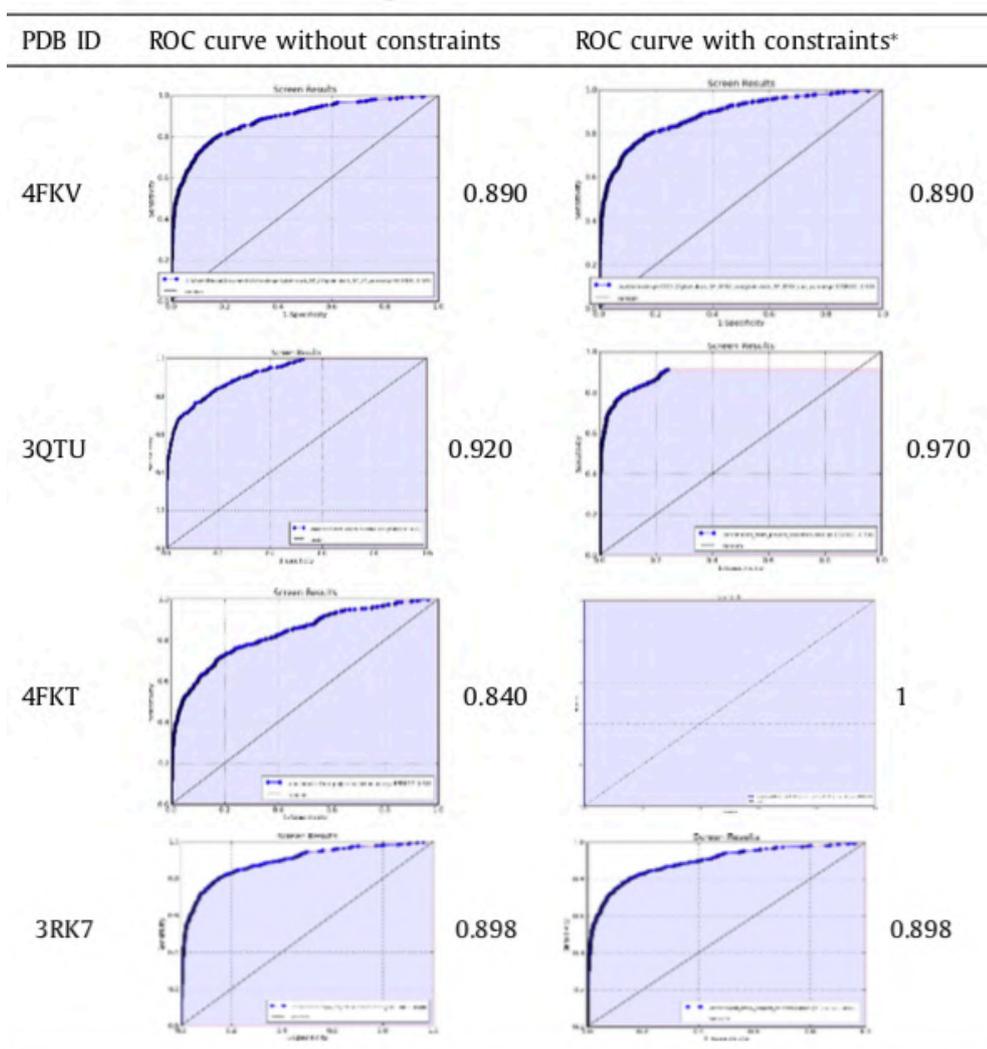


Table 10: Comparison of the five approaches used

	PDB Starting Structure	CHA	MYSHAPE	Docking without constraints	Docking with constraints
PDB ID	ROC _{5%}	ROC _{5%}	ROC _{5%}	ROC	ROC
4FKV	0.63	1	0.99	0.89	0.89
3QTU	0.63	0.97		0.92	0.97
4FKT	0.94	0.97		0.84	1
3RK7	0.76	0.97		0.90	0.90
mean	0.74	0.98		0.89	0.94

5.4 CONCLUSION

The results of these compared studies showed that the use of the MD trajectories snapshot should be mandatory to improve pharmacophore-based virtual screening. The CHA and the MYSHAPE were revealed to increase the performance, but they should be used relatively to

the availability of target-ligand complexes. In fact, in our case study on CDK2 inhibitors, the CHA showed a sharp improvement with respect to each complex. Consequently, this type of approach is useful when only one complex is available. If many target-ligand complexes are available, it would be advisable to use the MYSHAPE which leads to a shared pharmacophore model with all the complexes adding the common features obtained from the MD simulations. Both the MD-derived pharmacophore approaches showed a better performance than docking. The average two-times longer docking calculation with constraints requires a higher computational effort, but the improved screening outcomes legitimate the use of MD simulations as an ameliorative tool. The cost arising from synthesizing and testing false positives is an order of magnitude higher than performing accurate modelling and virtual screening analysis. Therefore, every chance to increase true positive hits even to a slight extent should be exploited.

6. EVALUATION OF THE IKKB BINDING OF INDICAXANTHIN BY INDUCED FIT DOCKING, BINDING POSE METADYNAMICS, AND MOLECULAR DYNAMICS

6.1 INTRODUCTION

Indicaxanthin is a betalain pigment belonging to the betalain class compounds (Figure 40). Indicaxanthin has been demonstrated to exert significant reducing, anti-oxidative, anti-inflammatory, spasmolytic, and neuromodulatory effects both *in vitro* and *in vivo* [181]

In addition to its redox-modulating and anti-inflammatory properties, it has also been shown to exert antiproliferative effects against melanoma cells being also able to significantly impair tumor progression in a mouse model of cutaneous melanoma [182]. Mechanistic evaluations have individuated the inhibition of the nuclear factor- κ B (NF- κ B) signalling, a key event underlying the antitumoral effects exerted by the pigment.

The NF- κ B transcription factor family consists of five different DNA-binding proteins that form a variety of homodimers and heterodimers regulating both the innate and adaptive immune responses [183]. More interestingly, besides its modulating effects on the entire inflammatory response, NF- κ B is responsible for accelerating cancer progression, metastasis, angiogenesis, and drug resistance [184]. Indeed, in several types of cancer, both in malignant cells and in the tumor microenvironment, NF- κ B is constitutively activated, and only rarely is such activation due to NF- κ B-related genetic alterations [185-186].

Given that constitutive or drug-induced upregulation of NF- κ B activity is associated with chemoresistance, NF- κ B is considered one of the most relevant targets for discovering new active compounds [183]. Melanoma is aggressive skin cancer, notoriously resistant to current cancer therapies [187]. Permanent activation of NF- κ B signaling in melanoma has been reported to proceed through the activation of the so-called canonical pathway [188-191]. Through this process, activation of NF- κ B depends on the degradation of its specific inhibitors (I κ B) consisting of I κ B α , I κ B β , and I κ B ϵ . Typically, I κ Bs bind to NF- κ B complexes, inhibiting their DNA binding and keeping them in a predominantly, inactive cytoplasmic form. Tumor microenvironmental stimuli can lead to the activation of a large cytoplasmic protein complex: the I κ B kinase (IKK). The precise nature of this molecular mechanism remains to be elucidated but it contains IKK α , IKK β , and IKK γ as the three seminal components. The phosphorylated and thus activated IKK complex is responsible for the phosphorylation of I κ B, marking it for degradation through the proteasomal degradation machinery. The free NF- κ B dimers (p50-p65) can, then, translocate from the cytoplasm to the nucleus, bind to DNA, and regulate gene transcription. Typical targets within the classical NF- κ B signaling include genes encoding pro-inflammatory cytokines, growth factors, chemokines, matrix metalloproteinases, pro-proliferative proteins, anti-apoptotic proteins, pro-inflammatory enzymes, angiogenic factors, and adhesion molecules [185]. As a primary druggable mediator of canonical NF- κ B signaling, the IKK β enzyme inhibition has been the historical focus of drug development pipelines. Thousands of compounds with activity against IKK β have been characterized, with much demonstrating promising efficacy in pre-clinical models of cancer. However, severe on-target toxicities and other safety concerns associated with systemic IKK β inhibition have so far prevented the clinical approval of any IKK β inhibitors [192].

Indicaxanthin has been demonstrated to inhibit I κ B α degradation in melanoma A375 cells at 100 μ M, a concentration at which it impairs NF- κ B signaling and inhibits 50% cell proliferation (IC₅₀) [182]. In light of this evidence, an *in silico* approach to evaluate Indicaxanthin's inhibitory activity against IKK β was employed. To this end, molecular modeling and simulation

techniques including induced-fit docking (IFD) [70], binding pose metadynamics (BPMD) [69], MD simulations, and MM-GBSA (molecular mechanics-generalized Born surface area) free energy calculation have been performed.

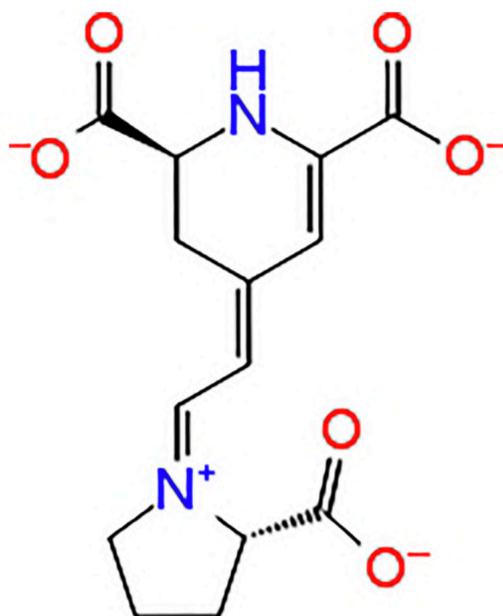


Figure 40: Indicaxanthin structure

6.2 MATERIALS AND METHODS

6.2.1 Protein and Ligand Preparation

A crystal structure of human IKK β (hIKK β) at 2.83 Å resolution, which is partially phosphorylated and bound to the staurosporine analogue K252a (PDB ID: 4KIK) [193] was used. The structure was optimized using the Protein Preparation Wizard in Maestro (Schrödinger, 2017) adding bond orders and hydrogen atoms to the crystal structure using the OPLS3 force field. Prime was used to fix missing residues or atoms in the protein and to remove co-crystallized water molecules. PROPKA was used to check for the protonation state of ionizable protein groups (pH = 7.0). The hydrogen bonds were optimized through the reorientation of hydroxyl bonds, thiol groups, and amide groups. In the end, the system was minimized with the value of convergence of the RMSD of 0.3 Å. Indicaxanthin and staurosporine analog K252a were prepared using LigPrep. The force field adopted was OPLS3 and Epik 3.9 (Schrödinger, 2017-1) was selected as an ionization tool at pH 7.2 \pm 0.2. Tautomers generation was unflagged and the maximum number of conformers generated was set at 32, as reported in the previous studies.

6.2.2 Induced Fit Docking

The grid boxes for the binding sites of Chain A (inactive form) and B (active form) were built considering the co-crystallized ligand staurosporine analogue K252a as a centroid. For the allosteric sites of Chain A, the amino acid residues previously identified by Liu et al. [193] were considered crucial for centring the docking grid.

6.2.3 Binding Pose Metadynamics

See as reported in the section 3.4.2.

6.2.4 MD Simulations

The plain MD simulations were carried out using Desmond 4.9 [84] and OPLS3 force field [80]. The complexes were solvated in orthorhombic boxes using the TIP3P water model. Ions were added to neutralize charges. The systems were minimized and equilibrated at a temperature of 303.15 K and a pressure of 1.013 bar. The system was simulated as an NPT ensemble; a Nose–Hover thermostat and Martyna–Tobia–Klein barostat were used. The integration time step was chosen to be 2 fs. To keep the hydrogen–heavy atom bonds rigid, the SHAKE algorithm was used. A 9 Å cutoff radius was set for the short-range Coulomb interactions, and smooth particle mesh Ewald was used for the long-range interactions. For each system, out 100 ns MD was carried, with 1.2 ps detection ranges for energy, and 4.8 ps for the trajectory frames. Visualization and analysis of the MD trajectories were performed using Desmond simulation analysis tools in Maestro.

6.2.5 MM-GBSA Binding Free Energy Calculation

The MM-GBSA approach [133] was applied to the snapshots extracted from the 100 ns production MD trajectories. Protein-ligand binding free energy using MM-GBSA was calculated as the difference between the energy of the bound complex and the energy of the unbound protein and ligand. In this work MM-GBSA calculations were also achieved in Prime software [134]; the entropy term $-T\Delta S$ was not calculated to reduce computational time. The VSGB solvation model was chosen using OPLS3 FF with minimized sampling method.

6.3 RESULTS AND DISCUSSION

The hIKK β protomer adopts a trimodular structure that closely resembles that of *Xenopus laevis* (α IKK β): an N-terminal kinase domain (KD), a central ubiquitin-like domain (ULD), and a C-terminal scaffold/dimerization domain (SDD). The selected crystal structure has one protomer in the active conformation with phosphorylated Ser177 and Ser181 (Chain B), and the other protomer is in the inactive conformation with the same residues of serine unphosphorylated in the activation loop (Chain A). Recently, Liu and co. identified a druggable allosteric site between KD and ULD.

To evaluate the binding capability of Indicaxanthin into the hIKK β , a series of computational studies with increasing accuracy, induced-fit docking (IFD), binding pose metadynamics (BPMD), unbiased MD followed by MM-GBSA (molecular mechanics-generalized Born surface area continuum solvation) free energy calculation was performed. The computational studies began by selecting the crystal structure of the human IKK β (PDB ID: 4KIK) [193] bound to the staurosporine analogue K252a in the KD (Figure 41A, B).

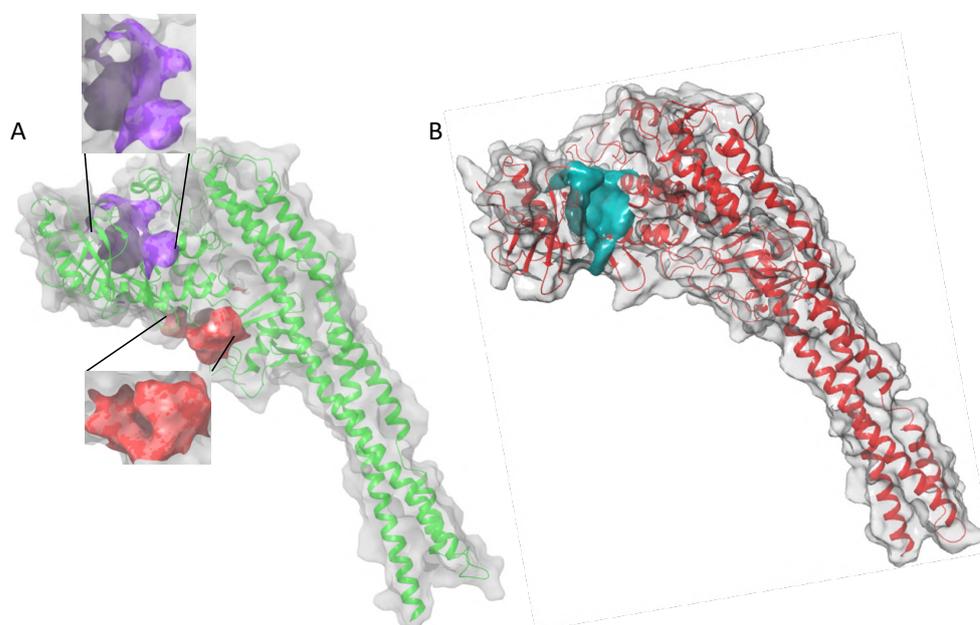


Figure 41: 3D structure of hIKK forms: (A) the inactive form, Chain A, in purple the KD binding pocket, in red the allosteric binding site; (B) the active form, Chain B, in cyan the KD binding pocket of the inactive Chain A with a volume of 551 Å³. This pocket is surrounded by three α helices (Arg118-Ser127, Leu265-Leu273, and Leu303-His313) of KD and a loop (Thr368-Leu386) between two β sheets of ULD. They suggested that small molecules binding in the pocket between KD and ULD probably interfere with the kinase function by disrupting the interaction between these two domains

The crystal structure was optimized by completing and refining the missing loops and residues and optimizing amide groups of asparagine (Asn) and glutamine (Gln), and the imidazole ring in histidine (His); and predicting protonation states of histidine, His, aspartic acid (Asp) and glutamic acid (Glu) and tautomeric states of histidine. Next, docking studies have been performed centering the docking boxes on the 3D coordinates of ligand co-crystallized, both in the inactive Chain A and in the active Chain B. The RMSD of K252a in Chain B and A was calculated showing values 0.8–0.18 Å respectively. Another docking box for the allosteric binding pocket was centered on the previously identified residues by Liu and Co [193].

Docking of K252a in Chain A showed a single H-bond interaction with Cys99 (hinge region), with the carbonyl of the lactam ring. Other H-bonds were found between the amine of the K252a lactam group and Glu97 and the hydroxyl group and Glu149. Moreover, an aromatic H-bond

was found with Asp166 (DLG triad). In chain B, K252a, as in Chain A, established the same interactions with Glu97, Cys99 (hinge region), and Glu149, with the only difference in another H-bond interaction between the carbonyl group of ether group and Thr23, and an aromatic H-bond with Tyr98 (Figure 42A, B). The IFD of K252a and Indicaxanthin in the two sites previously identified, and the IFD of Indicaxanthin in the allosteric binding site were performed.

In the best IFD result considering the Chain A (Docking score -6.166 kcal/mol, Table 12), the carboxyl group of the pyrrolidinium ring interacts with Thr23 and Gly24 forming two H-bonds, while the Cys99 simultaneously interacts with the carboxyl groups and the nitrogen of pyridine moiety. The carboxyl group of pyridine in position 11 interacts with Tyr98 (hinge region) with another H-bond and with Lys106 (solvent-accessible region) through one salt bridge. The carboxyl groups in the pyridine moiety establish two aromatic H-bonds with Tyr98. Finally, another salt bridge involves the pyrrolidinium nitrogen and the Asp103. In Chain B (Docking score -7.293), Indicaxanthin showed similar interactions as in Chain A. Two H-bond interactions involve the carboxyl group of groups of pyridine interacts through H-bond with Tyr98, Lys106. The last interactions were a salt bridge between two carboxyl groups of the pyridine and Lys106 (Figure 42C, D).

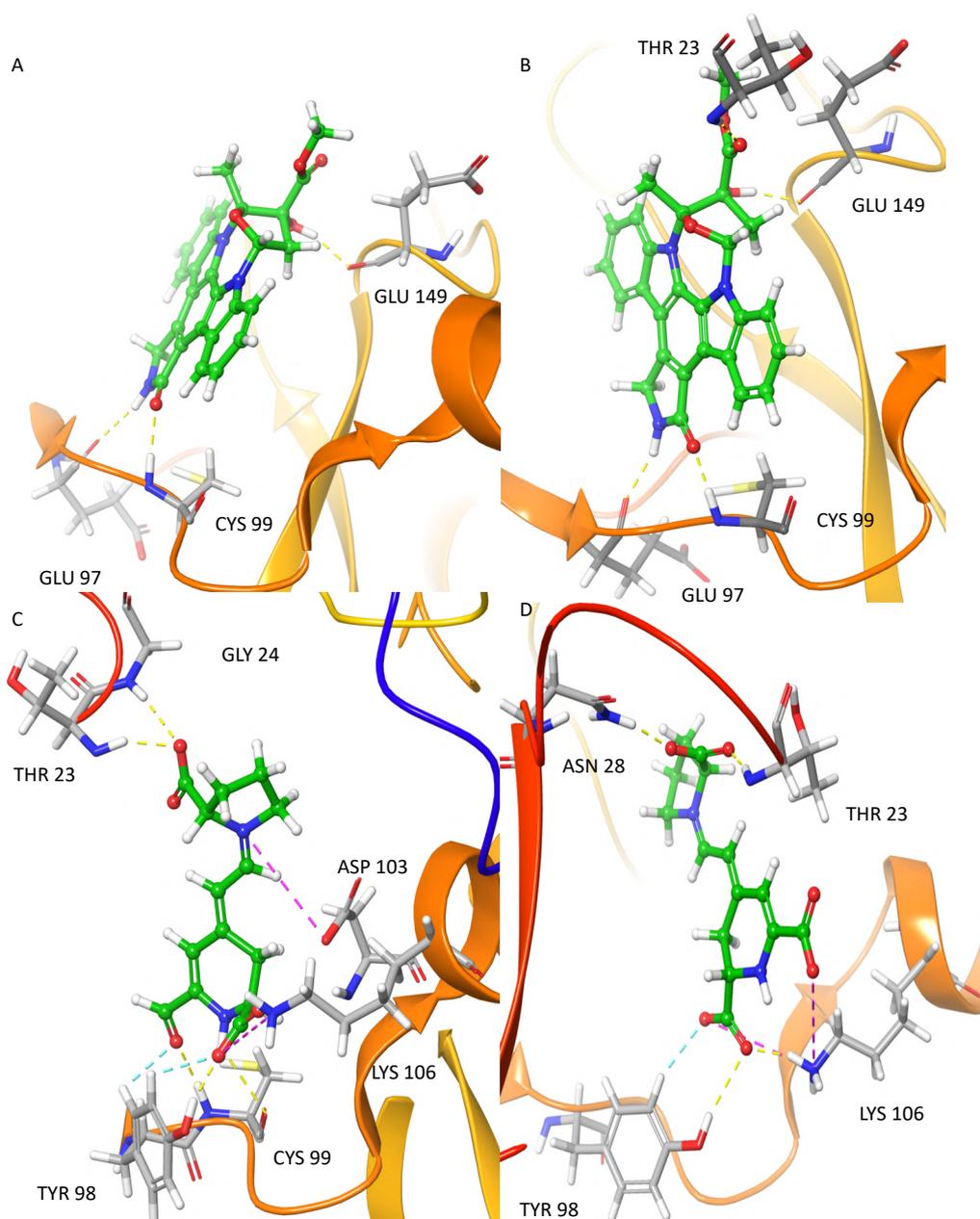


Figure 42: IFD pose of K252a in the inactive Chain A (A); IFD pose of K252a in the active Chain B (B); IFD pose of Indicaxanthin in the inactive Chain A (C); IFD pose of Indicaxanthin in the active Chain B (D). H-bonds interactions are represented in yellow dashes, aromatic H-bonds are represented in light blue dashes, the salt bridges in purple dashes

The IFD study of Indicaxanthin in the allosteric site (Docking score -6.117) showed a single H-bond interaction was found between one of the carboxyl group of pyridine and His380; another interaction was a salt bridge, between the nitrogen of pyrrolidinium moiety and Asp373 (Figure 43).

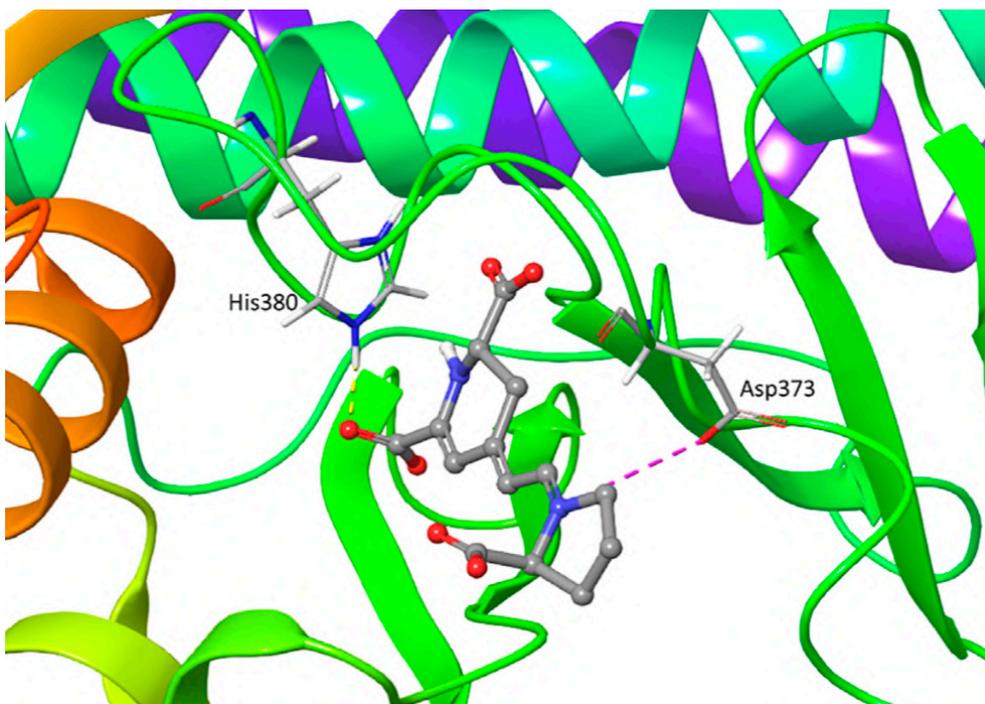


Figure 43: IFD pose of Indicaxanthin in the inactive Chain A allosteric binding site. H-bonds interactions are represented in yellow dashes, the salt bridges in purple dashes

Moreover, docking simulations were performed with three known orthosteric inhibitors: MLN120B ($IC_{50} = 60$ nM) [194]; the imidazo [1,2-*a*]quinaxoline derivative 6a (IC_{50} 324 nM) [195]; LASSBio-1524 ($IC_{50} = 20$ μ M) [196]; and one allosteric inhibitor: BMS345521 ($IC_{50} = 300$ nM) [194]. The docking scores of these known inhibitors are consistent with the experimental IC_{50} , considering that Indicaxanthin showed an $IC_{50} = 100\mu$ M. Additionally, the key residues identified for these known inhibitors are like those identified for Indicaxanthin (Cys99, Asp103, Lys106). The allosteric inhibitor BMS345521 showed key interactions with His380, already identified for indicaxanthin, Asn308, Ser127. The 2D and 3D figures of the docked reference inhibitors together with key interactions and docking scores are reported

6.3.1 Binding pose metadynamics

BPMD was used to evaluate the stability of the K252a best poses (in terms of docking scores) obtained from IFD studies into the binding site in Chain A and Chain B to evaluate the reliability of the IFD poses obtained as validation of the docking scoring functions. The simulation performed on K252a pose in Chain A showed a PoseScore of 1.206. The PersScore showed that the hydrogen bonds were kept for 81% of the simulation time. The interactions by the lactam ring were confirmed to stabilize the molecule, it was observed for 98.2% of the simulation time, interactions between the amine group and Glu97, for 97.3% of the simulation time interactions between the carbonyl oxygen and Cys99. The H-bond interaction between the hydroxyl group and Glu149 was kept for 48.2% of the simulation time. The CompScore of -2.85 confirmed that the starting molecule pose is stable in the active site.

The results of the K252a pose in Chain B showed a PoseScore of 0.909, the PersScore proves that for 63.9% of the simulation time, the H-bonds were maintained. As in Chain A, the interactions by the lactam ring were confirmed 100% of the simulation time, the interaction between the carbonyl oxygen and Cys99, and 98.2% of the simulation time interaction between the amine group and Glu97. The same for H-bond interaction between the hydroxyl group and Glu149 was kept for 57.3% of the time. The CompScore, was -2.28 . The result obtained

confirmed the accuracy of the identified IFD poses (Table11).

Table 11 BPMD and docking scores (kcal/mol) for the Indicaxanthin and K252a complexes.

	Docking Score	Pose Score	Pers Score	Comp Score
Indi/Chain A	-6.166	3.133	0.12	2.53
Indi/Chain B	-7.293	4.631	0.04	4.44
K252a/Chain A	-13.700	1.206	0.81	-2.85
K252a/Chain B	-14.121	0.909	0.64	-2.28
Indi/Chain A allosteric	-6.117	>6	0.0	>6

Successively the BPMD simulations have been performed at the binding site in Chain A, the best pose of Indicaxanthin reached a steady PoseScore of 3.133, considered stable, while PersScore showed that the hydrogen bonds identified at the start of the metadynamics run were kept for 12% of the averaged time. In particular, the H-bond between carboxyl oxygen and Cys99. The CompScore value was 2.53.

At the binding site in chain B, the averaged RMSD of Indicaxanthin reached a steady PoseScore of 4.631, PersScore of 0.04. Three H-bond were recorded during the ten replicas. H-bond interaction between the carboxyl group of pyrrolidinium moiety and Thr23 was kept for 11.8% of the simulation time, the same group showed an ulterior H-bond with Asn28 for 7.3% of the simulation time. The last interaction was between one of the carboxyl groups of pyrimidine and Lys106 for 2.7% of the simulation time. This interaction was supported by salt bridges between the two inferior carboxyl groups and Lys106. The last score, CompScore was 4.44.

6.3.2 Molecular Dynamics Simulations

Starting from the previous IFD poses which BPMD analysis showed to be accurate, unbiased MD simulations was performed to explain the stability of Indicaxanthin as an inhibitor against the two forms of hIKK β , active and inactive, and allosteric inhibitor of the inactive form. Additionally, to understand if the protein target undergoes conformational alteration after interacting with Indicaxanthin. Five systems have been generated and submitted each for 100 ns in MD simulations (Chain A-Indicaxanthin, Chain B-Indicaxanthin, Chain A allosteric site-Indicaxanthin, Chain A-K252a, Chain B-K252a). Calculation of RMSD, RMSF, number, and types of protein-ligand contacts have been carried to have a more detailed analysis of Indicaxanthin-target complexes compared to the co-crystallized ligand.

The Chain A-K252a system reached equilibrium quickly and fluctuated around the average value of 3 Å, the low average value of ligand RMSD 1.8 Å indicated strong stability of K252a in the binding pocket as expected due to the low number of rotatable bonds and the eight fused rings. The Chain B-K252a behaviour is slightly different. The system reached equilibrium after ~10 ns and the fluctuation of the protein is higher than the previous system analyzed (~4.3 Å). The same evidence is reported for the average ligand RMSD that is higher (3.72Å). This first analysis showed that the phosphorylation of the serine residues in the active form (Chain B) could confer more flexibility to the binding pocket in the KD, as confirmed by the higher fluctuations of K252a despite the rigid structure (Figure 44).

The RMSF plot of Chain A is quite comparable for the residues of the active site with Δ RMSF <1 Å. The same could be observed for Chain B in which the Δ RMSF <0.5 Å. Analyzing the trajectories of Chain A and Chain B with the co-crystallized ligand K252a, it is interesting to

note that the residues involved in the interactions are variable both in terms of the type of contacts and in the time of interactions. (Figure 45)

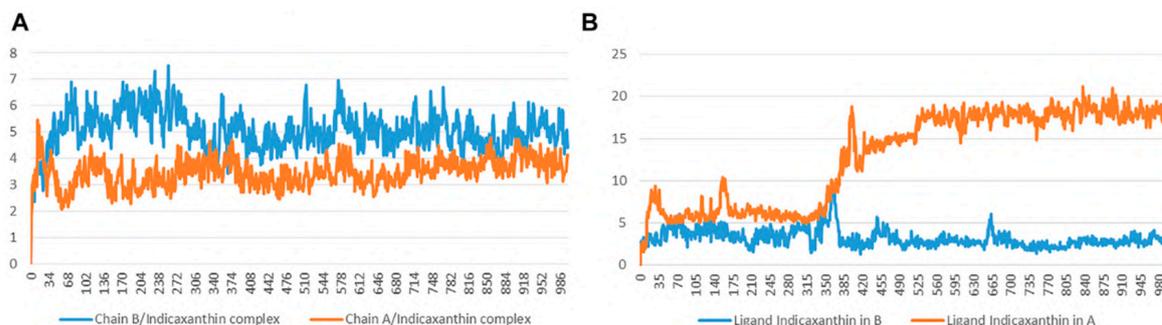


Figure 44: Time evolution of the RMSD (Å) values of backbone atoms (A) and ligand Indicaxanthin (B) in the active form (Chain B) and the inactive form (Chain A)

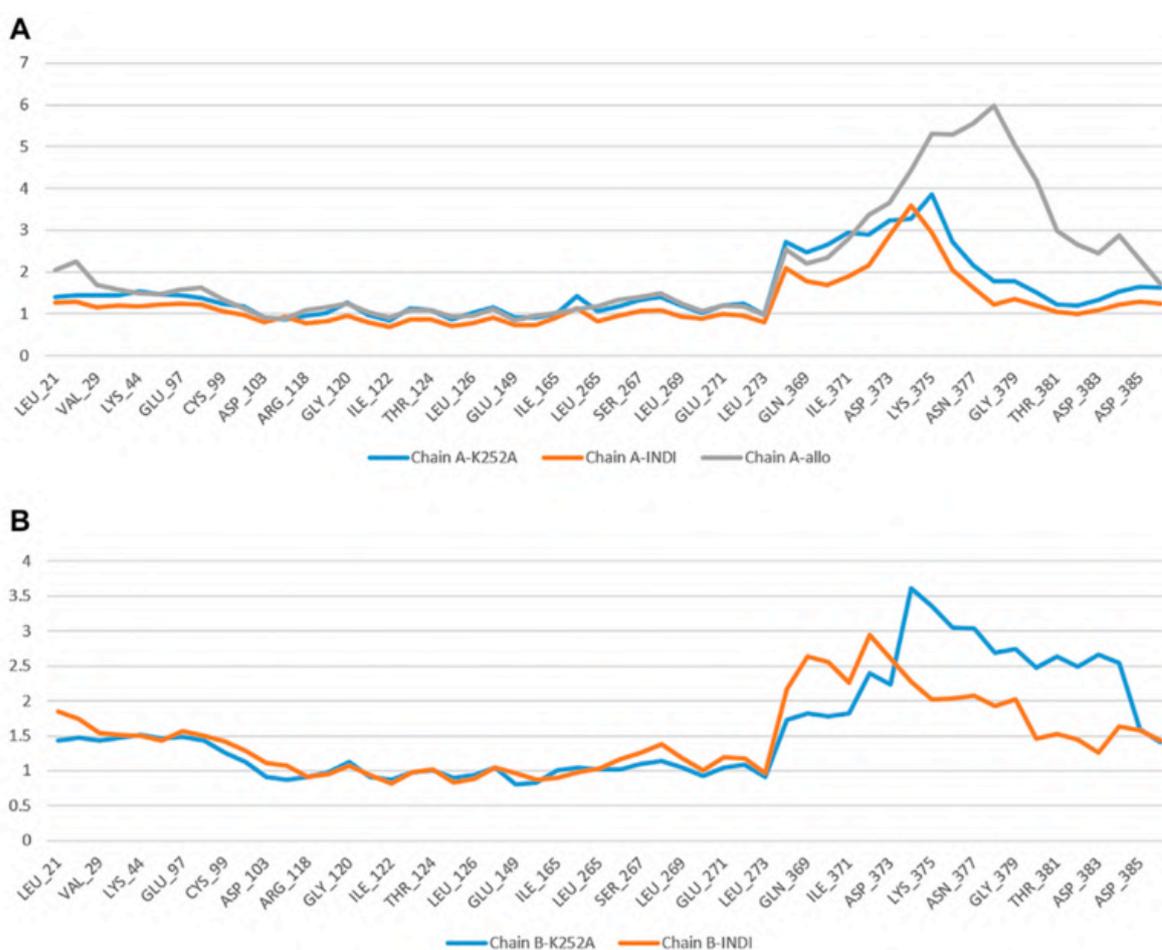


Figure 45: The comparative RMSF values for the complexes of backbone atoms of Chain A (A) and Chain B (B) bound to K252a, Indicaxanthin, and of the allosteric pocket

Analyzing the trajectories of Chain A and Chain B with the co-crystallized ligand K252a, it is interesting to note that the residues involved in the interactions are variable both in terms of the type of contacts and in the time of interactions. As reported in Figure 46, K252a has four key residues involved: Met96, Glu97, Cys99, and Ile165 in both active and inactive forms. The H-bonds with Glu97 and Cys99 were stable all 100ns simulations long. Hydrophobic contacts of Ile165 are maintained for a major part of the simulations, while hydrophobic contacts with

Met96 are involved in both simulations but just for a little fraction of time. The differences in the two complexes regard the less interaction time of the H-bond with Glu149 in Chain B. In Chain A, an H-Bond with Gln175 is observed during the simulation in a discontinuous fashion, while in Chain B, Hydrophobic contacts are observed with Phe26 for about 50% of the simulation time. Other interactions are involved in both complexes such as with Leu21, Thr23, Val29, Ala42, Tyr98, and Val152 (Figure 46).

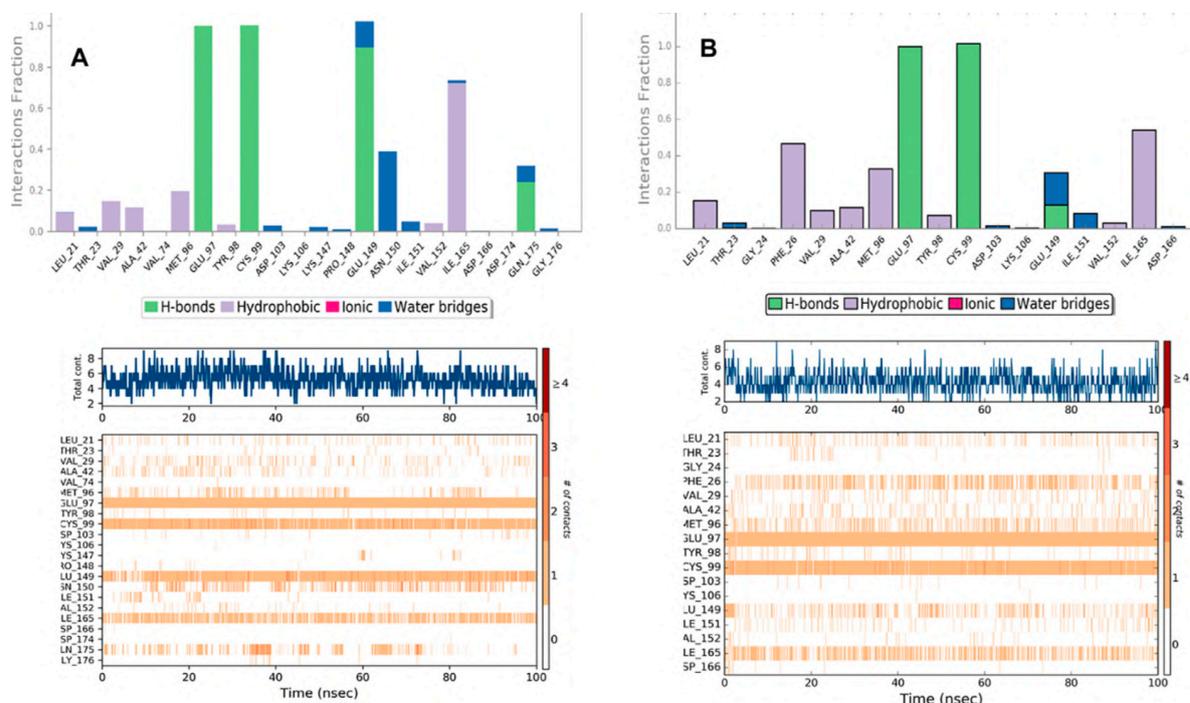


Figure 46: Protein-Ligand contacts: K252a in Chain A (A) K252a in Chain B (B)

The behaviour of Indicaxanthin, when bound to the two forms of IKK, is somewhat different in terms of interactions, but above in terms of residence in the binding pocket. As previously commented, in Chain A after ~38 ns. Indicaxanthin moves from the deep pocket but remains over it by interacting with H-bonds and ionic interaction with Arg427, Arg575, Arg579, and Arg582 acting as a lid. In the first 40ns, the interactions of Indicaxanthin are both ionic and H-bond interactions involving the residues of the pocket such as residues 20–24 and 103–106. Indicaxanthin remains confined into the binding pocket for all the simulation time when Chain B is considered. Even though the residues involved in the protein-ligand interactions differ from those involved in the interaction with K252a. Indicaxanthin interacts for the major of the time with Thr23, Asn28, Arg47, and Lys106 through H-bonds, but a key role in the binding stability is due to the ionic interaction with Asp103 and Lys 106 (Figure 47).

The stability of the Indicaxanthin's binding is also mediated by several water bridges that are hydrogen-bonded protein-ligand interactions mediated by a water molecule.

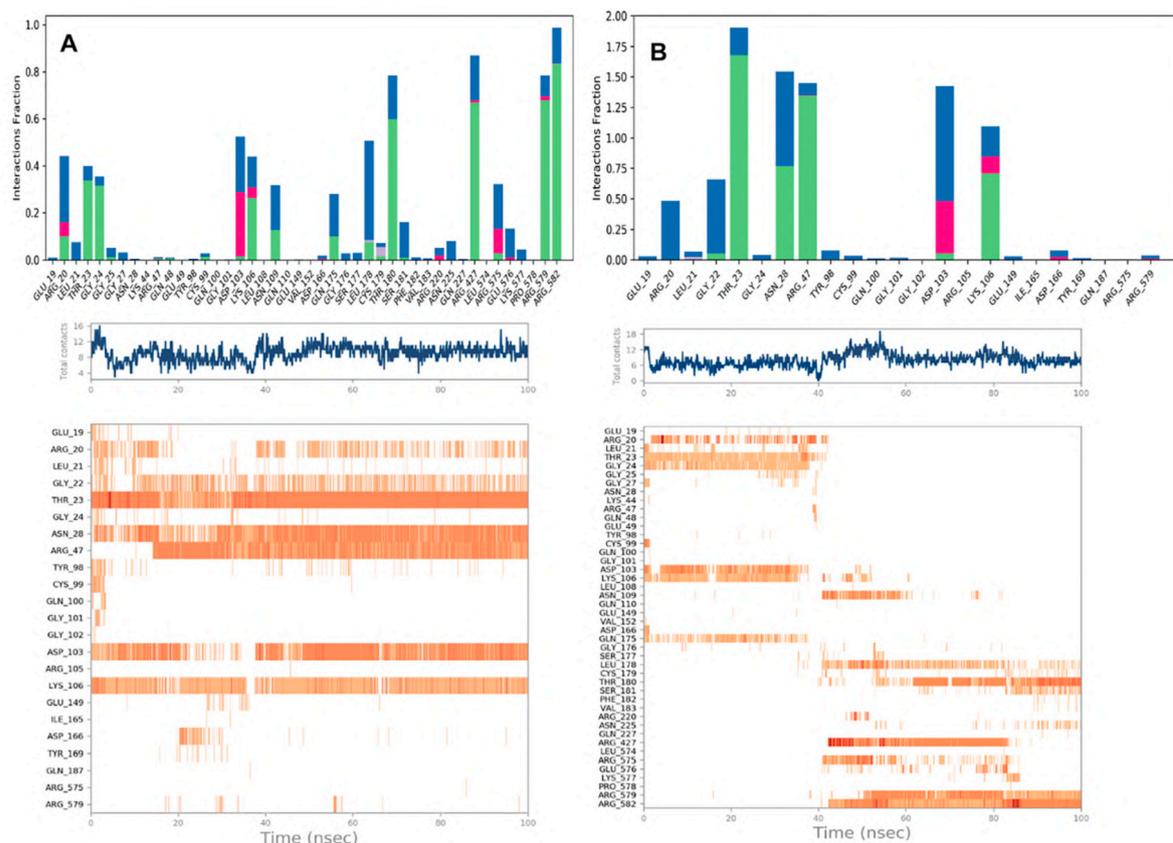


Figure 47: Protein-Ligand contacts: Indicaxanthin in Chain A (A) Indicaxanthin in Chain B (B). Legend is the same as Figure 46

6.3.4 MM-GBSA and Binding Free Energy Analysis

To understand the biophysical basis of molecular recognition of Indicaxanthin with the inactive form of IKK, and active form of IKK, a molecular mechanics-generalized Born surface area MM-GBSA approach was used. For binding-free calculation of Indicaxanthin and K252a/protein systems, 101 frames from 100 ns (every 1 ns) were retrieved to calculate ΔG_{bind} . A summary of binding components in the binding free energy is reported in Table 12.

Table 12: Predicted MM-GBSA free energies (kcal/mol) and individual energy terms of the Indicaxanthin-target complexes and K252a-target complexes

	ΔG_{bind}	ΔG_{Coul}	ΔG_{Hbond}	ΔG_{Lipo}	ΔG_{Solv}	ΔG_{vdW}
Indi/Chain B	-22.2 ± 4.3	67.5 ± 11.5	-3.9 ± 1.0	-4.5 ± 0.8	-59.2 ± 11.2	-24.6 ± 4.3
Indi/Chain A	-20.7 ± 4.7	10.6 ± 30.3	-4.3 ± 1.5	-3.4 ± 1.3	-3.2 ± 0.28	-22.3 ± 5.45
K252a/Chain B	-72.02 ± 4.1	-16.78 ± 3.0	-1.33 ± 0.2	-20.62 ± 1.3	24.2 ± 2.1	-57.6 ± 2.5
K252a/Chain A	-74.45 ± 4.0	-22.42 ± 0.3	-1.68 ± 0.3	-18.67 ± 0.3	26.6 ± 1.9	-57.96 ± 2.9

Furthermore, in Figure 48 it is reported the free energy landscape of Indicaxanthin in Chains A and B. As expected from previous analysis of the MD trajectories, Indicaxanthin showed a higher binding affinity for the Chain B concerning Chain A ($\Delta G_{\text{bind}} = -22.2 \pm 4.3$ kcal/mol, ΔG_{bind} respectively) that could be justified by the shifting of Indicaxanthin towards the mouth of the binding pocket. Comparing the binding free energy values of Indicaxanthin and K252a, they are consistent with the experimental evidence (K252a $\Delta G_{\text{bind}} = -72.02 \pm 4.1$ kcal/mol for Chain B, and $\Delta G_{\text{bind}} = -74.45 \pm 4.0$ kcal/mol for Chain A).

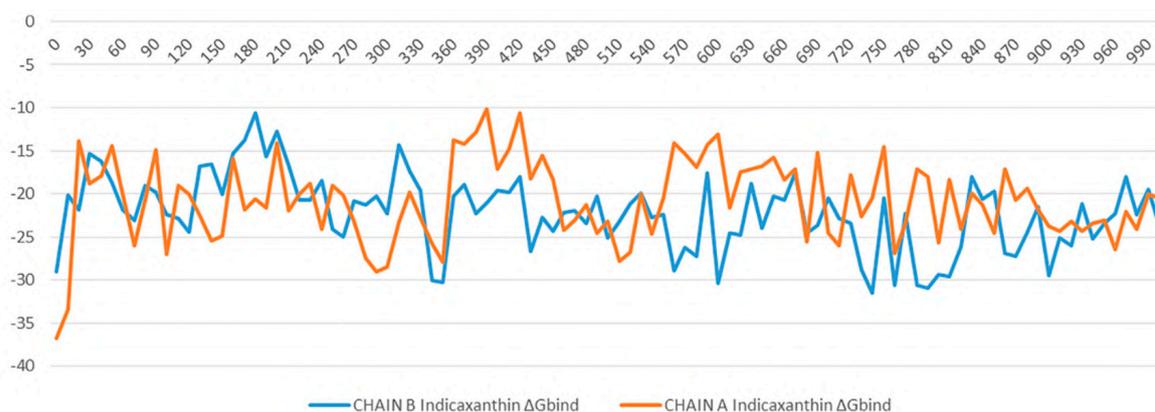


Figure 48: The binding energy (ΔG expressed in kcal/mol) landscape of Indicacaxanthin bound to Chain A and Chain B

6.4 CONCLUSION

The stability of Indicaxanthin-hIKK β complexes compared to K252a, a co-crystallized inhibitor, was assessed by using Induced fit docking, binding pose metadynamics, and MD. Finally, MM-GBSA free energy calculations have been performed to establish what form of IKK Indicaxanthin prefers. Induced fit docking results showed that the binding of Indicaxanthin with the active form, the inactive form, and the allosteric site of hIKK β showed the strongest stability with the active form. MD trajectories analysis (RMSD, RMSF, and protein-ligand contacts number and along the time) also showed that Indicaxanthin enhanced the stability of the active form at the same level as the known inhibitor K252a. The stability of the inactive form complex with Indicaxanthin is quite similar but it did not reach the quality of the active form. Contrariwise, even though for 40 ns over 100 ns Indicaxanthin can bind the allosteric pocket, it should not be considered an allosteric inhibitor of hIKK β .

Along these lines, present results further suggest the molecule as a useful nutraceutical tool in combo therapy i.e., with other therapeutical agents targeting different checkpoints of melanoma development.

The currently demonstrated ability of Indicaxanthin to inhibit the active form of hIKK β may, then, suggest the phytochemical as a new lead compound to synthesize novel and more potent IKK β inhibitors for the treatment of cancer and inflammation-related conditions.

7. EXPLORATION THE SARS-COV-2 PROTEOME IN THE SEARCH OF POTENTIAL INHIBITORS VIA STRUCTURE-BASED PHARMACOPHORE MODELING/DOCKING APPROACH

7.1 INTRODUCTION

Coronaviruses (CoVs) are one of the major pathogens that primarily targets the human respiratory system which caused previous outbreaks such as the severe acute respiratory syndrome (SARS)-CoV and the Middle East respiratory syndrome (MERS)-CoV. At the last of 2019, the novel coronavirus SARS-CoV-2 has become a pandemic threat (COVID-19) to public health. It is a respiratory disease causing fever, fatigue, dry cough, muscle aches, shortness of breath and some instances lead to pneumonia [197]. The SARS-CoV-2 genome comprises 29,903 nucleotides, with 10 Open Reading Frames (ORFs). The 3' terminal regions encode structural viral proteins: whereas the 5' terminal ORF1ab encodes two viral replicase polyproteins pp1a and pp1b. The proteolytic cleavage of pp1a and pp1b produces 16 nonstructural proteins (nsp1 to nsp16). Among these, there are nsp3, the papain-like protease (PLpro) and nsp5, the 3-chymotrypsin-like protease (3CLpro, also known as the main protease (Mpro). The viral polyprotein processing is essential for maturation and infectivity of the virus (Figure 49) [198]. Because of the crucial roles, these two proteases are important targets for antiviral drug design. Moreover, the virus encoded for other proteins that could be potential targets of antiviral drugs. The mature proteins of SARS-CoV-2 are: host translation inhibitor nsp1 (nsp1); nonstructural protein 2 (nsp2); papain-like proteinase (PLpro); nonstructural protein 4 (nsp4); 3C-like proteinase (3CLpro), nonstructural protein 6 (nsp6), nonstructural protein 7 (nsp7), nonstructural protein 8 (nsp8), nonstructural protein 9 (nsp9), nonstructural protein 10 (nsp10), RNA-directed RNA polymerase (Pol/RdRp), helicase (Hel), guanine-N7 methyltransferase (ExoN/nsp14), uridylyate-specific endoribonuclease (NendoU/nsp15), 2'-O-ribose methyltransferase (nsp16), Spike glycoprotein (S glycoprotein), protein 3a, Envelope small membrane protein (E protein), Membrane protein (M protein), nonstructural protein 6 (nsp6), protein 7a, nonstructural protein 7b (nsp7b), nonstructural protein 8 (nsp8), nucleoprotein (NC), ORF10 protein. These proteins can form hetero-oligomeric complexes such as: nsp7/nsp8 hetero-oligomeric complex; nsp7/nsp8/Pol hetero-oligomeric complex; nsp10/nsp14 hetero-oligomeric complex; nsp10/nsp16 hetero-oligomeric complex; Spike glycoprotein/hACE2 hetero-oligomeric complex. Anti-coronavirus therapies can be split into two main approaches: the first approach is to act on the human immune system or human cells level, and the other approach is to focus on coronavirus itself [199]. In exploring novel therapies for COVID-19, researchers are using computational approaches to aid in the discovery of potential candidates [200]. In particular, *in silico* drug repurposing, also named drug repositioning, is a strategy used to identify novel uses for existing approved and investigational drugs. This strategy offers numerous advantages over traditional drug development pipelines that suffer risks failure in preclinical or early stage clinical trials due to safety and/or toxicological issues. On the contrary, the drug repurposing strategy reduces this risk by using drugs that have demonstrated safety records from previous trials. The real advantage of drug repurposing is that preclinical and early stage clinical trials do not need to be repeated. This determines cost reductions compared to traditional drug development [5,201-213]. The number of *in silico* studies on drug repositioning against SARS-CoV2 is growing rapidly in these last months. A major part of these studies is focused on the repurposing of approved and

investigational drugs against the 3CLpro or Mpro by using both ligand-based approaches and structure-based approaches. Structure-based approaches are related to different docking analysis [214–224]. In another work, Battisti and coworkers used two different approaches related to docking and pharmacophore combined with MD to perform virtual screening of a large database of compounds on 10 different SARS-CoV-2 proteins [225]. To our knowledge, Touret and coworkers performed, to date, the only *in vitro* screening of an FDA approved chemical library which revealed potential inhibitors of SARS-CoV-2 replication [216]. Nevertheless, the identification of potential inhibitors is still challenging for all the researchers involved in the field. In this study, a computational analysis of the proteins encoded by the SARS-CoV-2 genes was performed. Such an analysis was used as a starting point for a druggability assessment and a computational drug repurposing work-frame. First, high-quality protein structures were built employing homology modeling or exploiting existing experimental structures. Starting from the models, a computational assessment was done to find out a druggable binding pocket for those proteins of which catalytic site is not known in the literature. The best druggable sites found in the previous analysis, together with the catalytic sites reported in the literature, were then used to build structure-based pharmacophore models. In the end, these models were used to screen the DrugBank library (approved and investigational drugs) [226] as a first screening approach.

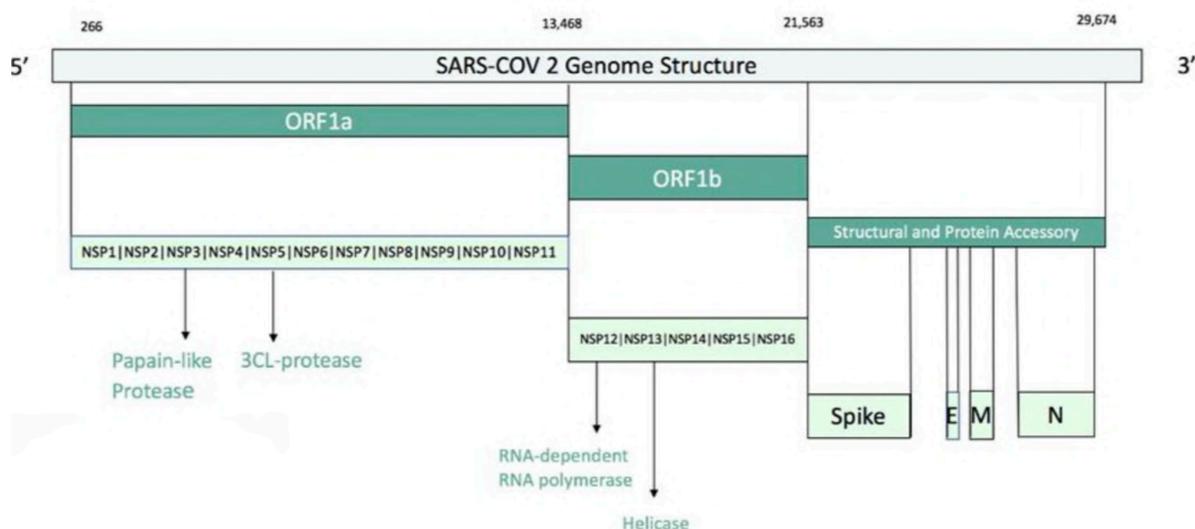


Figure 49: SARS-COV2 genome structure

7.2 MATERIALS AND METHODS

7.2.1 Library Preparation

A total of 8752 experimental, investigational, and approved molecules were downloaded from the DrugBank database (www.drugbank.ca). First, the database molecules were prepared using Schrödinger LigPrep v. 2018-4. The force field adopted was OPLS3e and Epik [75] was selected as an ionization tool at $\text{pH } 7.0 \pm 2.0$. The database obtained was prepared as a Pharmacophore Screening database, in *.lbd format, through Idbgen (extension present in the LigandScout 4.3 [15] package), which allowed obtaining the best conformation of the ligand (at low energy) between the 200 the application can calculate. The tautomers were considered as separate molecules and those molecules that were duplicated or whose conformation calculation had failed were eliminated.

7.2.2 Homology Modeling and Protein Preparation

The full SARS-CoV-2 proteome based on the NCBI reference sequence NC_045512, which is identical to GenBank entry MN908947 and annotations from UniProt, was modeled in the SWISS-MODEL [227] workspace (swissmodel.expasy.org/workspace). Only for 10 proteins, it was possible to obtain high-quality models and experimental structures that were considered for further analysis. Investigated proteins are 3C-like protease (3CLpro), papain-like protease (PLpro), guanine-N7 methyltransferase (nsp14), uridylate-specific endoribonuclease (NendoU/nsp15), nsp4, nsp7/nsp8 supercomplex, nsp9, nsp7/nsp8/nsp12 hetero-oligomeric complex, helicase (Hel), 2'-O-ribose methyltransferase (nsp16). For each structure, templates with the highest identity available at the time of this study (25 March 2020) were selected and respective models were generated.

For 3C-like protease (3CLpro) the crystal structure of the COVID-19 main protease (PDB ID: 6LU7) was available. The structure of papain-like protease (PLpro) of SARS virus (PDB ID: 3E9S) was used as a template of the human coronavirus papain-like model (82.86% sequence identity). This one was the best available experimental structure at the time of the study (25 March 2020). On 27 May 2020, the crystal structure of PLpro of SARS-CoV-2 was released (PDB ID: 6WZU). The overlapping of the model and the experimental structure was performed. The RMSD value of 3.99 Å shows that the two structures are identical unless few residues in the C-terminal. For guanine-N7 methyltransferase (nsp14), the SARS-related coronavirus (PDB ID: 5C8S) was used as template that shows 95.07% of sequence identity. For uridylate-specific endoribonuclease (NendoU/nsp15), the experimental structure was reported in the Protein Data Bank [34] (PDB ID: 6W01). The crystal structure of nsp4 from mouse hepatitis virus A59 (PDB ID: 3VCB) was used as a template of SARS-CoV-2 nsp4 (61.36% sequence identity). The crystal structure of SARS-CoV super complex of nonstructural proteins (PDB ID: 2AHM) was chosen as a template of nsp7/nsp8 supercomplex (97.86% sequence identity). For nsp9 the template of nsp9 from SARS-coronavirus (PDB ID: 1UW7) was used. It shares a sequence identity of 97.35%. The X-ray structure of SARS coronavirus nsp7/8/12 (PDB ID: 6NUR) was selected as a template of nsp7/nsp8/nsp12 hetero-oligomeric complex (96.70% sequence identity). The crystal structure of SARS-coronavirus helicase (PDB ID: 6JYT) was used as template for SARS-CoV-2 helicase (Hel). It shows a high sequence identity (99.83%). On 29 July 2020, the experimental structure of SARS-CoV-2 helicase (PDB ID: 6ZSL) was released. The overlapping of our model and the experimental structure shows a RMSD value of 4.17 Å. This means a quite identical structure unless some loops (See Supplementary information of the original paper). The crystal structure of nsp16/nsp10 SARS coronavirus complex (PDB ID: 2XYQ) was chosen as a template of the model of 2'-O-ribose methyltransferase (nsp16) with 93.45% sequence identity. The models obtained and the PDBs were refined using the protein preparation wizard tool of Maestro Suite Software as reported in the previous chapter.

7.2.3 Pharmacophore Modeling

Pharmacophore model generation was performed using LigandScout 4.3. The structures were imported into LigandScout. 3C-like proteinase, PLpro, nsp14, nsp15, nsp16–nsp10 are protein–ligand complexes, while nsp4, nsp9, nsp10–nsp14, helicase, nsp7–nsp8 supercomplex, nsp12 are targets without ligand-bound. For protein-ligand complexes, a structure-based pharmacophore model was generated. When the model showed more features, to improve the performance of virtual screening, the features for the binding were considered, in other cases the features were omitted until hits were found. The calculate pockets tool has been used to find the binding pockets for the structures without ligand bound. A grid was calculated over the

entire protein structure and grid points were evaluated according to their buriedness and their number of neighboring grid points. Isocontour surfaces were generated. Then, a model was created by selecting the nature and number of six features according to the features shown in the protein–ligand complexes utilizing “Create Apo Site Grids”. Next, the pharmacophore model was generated for each one. The obtained pharmacophore models were used as a query to screen the DrugBank library. For apo protein, such an approach allows evaluating if a putative binding site is suitable for ligand binding.

Pharmacophore screening was preferred to be used prior to docking for two reasons. First, it exploits a rapid screening technique that is crucial in the first stage of the virtual screening cascade. Indeed, this is very common to use it as a first step in a virtual screening campaign on large databases [228]. Second, the structure-based pharmacophore uses a static conformation of protein side chains, while the docking funnel here used was set to have a gradually increasing precision with a final step of IFD that allow user to simulate side-chains-induced fit based on the ligand.

7.2.4 Docking

The hits identified by the virtual screening were submitted to a docking study using Glide [95] in standard precision (SP) with the OPLS3e [27] force field. The crystal structures were optimized using protein preparation wizard in Maestro as reported in the previous chapter. For protein-ligand complexes, the grid boxes were built considering the ligands as centroid. In contrast, for apoproteins, the amino acid residues, previously identified by LigandScout as crucial, were considered for centering the docking grid. The docking study was performed using the Glide docking tool, in extra precision (XP) using no constraints. Van der Waals radii were set at 0.8 and the partial cutoff was 0.15 and flexible ligand sampling. Bias sampling torsion penalization for amides with nonplanar conformation and Epik state penalties were added to the docking score.

7.2.3 Induced-Fit Docking and MM-GBSA

Induced Fit Docking protocol (IFD) -developed by Schrödinger [24]-is a method for modeling the conformational changes induced by ligand binding. This protocol models induced-fit docking of one or more ligands using the following steps as also reported in [87].

IFD was performed using a standard protocol and the OPLS3e force field was chosen. The receptor box was centered on the co-crystallized ligands on the crucial residues identified within the binding site. During the initial docking procedure, the van der Waals scaling factor was set at 0.5 for both receptor and ligand. Prime refinement step was set on side chains of residues within 5 Å of the ligand. For each ligand docked, a maximum of 20 poses was retained to be then redocked at XP mode. IFD calculation was followed by prime/MM-GBSA for the estimation of $\Delta G_{\text{binding}}$. For the prime/MM-GBSA method see as reported in the section 3.4.2.6.

7.3 RESULTS AND DISCUSSION

SARS-CoV-2 caused the outbreak of coronavirus disease 2019 (COVID-19) threatening global health security. To date, no approved antiviral drugs or vaccines are available against COVID-19 although several clinical trials are underway. In this framework, computational methods offer an immediate and scientifically sound basis to potentially design highly specific inhibitors against important viral proteins and guide the antiviral drug discovery process [229].

In this study, SARS-CoV-2 encoded proteins were analyzed from PDB structures and homology models were generated by using the most similar PDB crystal structures as templates. For the homology models created, starting from the high similarity between SARS-CoV-2 proteins and some available crystal structures from SARS-CoV, ligand coordinates of the available most similar crystals were exploited for the structure-based pharmacophore creation. Below the analyzed proteins and the related pharmacophore maps composition are reported.

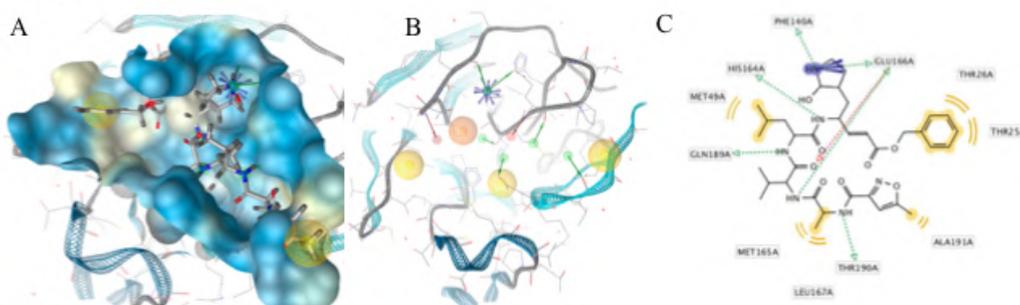
3C-like proteinase (3CLpro), also termed the main protease, cleaves most of the sites in the polyproteins and the products are nonstructural proteins (NSPs), which assemble into the replicase–transcriptase complex (RTC). The binding site of the main protease consists of a conserved catalytic dyad, i.e., Cys145 and His41 with other crucial residues, which are Phe140, Leu141, Asn142, Gly143, Ser144, Cys145, Met165, Glu166, Gln189 and Thr190 [222] (Figure 50A). The pharmacophore model was developed on the co-crystallized ligand (N3) that is present in the PDB ID 6LU7; this ligand was covalently bound to Cys145. N3 was modified, by breaking the covalent bond and filling in open valence. The final pharmacophore showed 12 features: 2 H-bond acceptors (HBAs) interacting one with Glu166 and the other with Gly143; 4 H-bond donors (HBDs) which interact, respectively, with Phe140, His164, Glu166, Gln189 and Thr190; 4 hydrophobic features interacting with Thr25, Thr26, Met49 and Ala191; and a negative ionizable area with Glu166 (Figure 50B,C). Papain-like protease (PLpro) cleaves the nsp1/2, nsp2/3 and nsp3/4 boundaries. It works with 3CLpro to cleave the polyproteins into NSPs [230]. It showed in the active site residues Gly164, Asp165, Arg166, Glu168, Pro248, Pro249, Tyr 265, Gly267, Asn268, Tyr 269, Gln270, Cys271, Gly272, Tyr274 and Thr302 (Figure 50D). The pharmacophore model was developed on the co-crystallized ligand present in the PDB ID 3E9S. The pharmacophore map was composed of 7 features: 1 HBA with Gln270; 2 HBDs, one with Tyr265 and the other with Tyr269; and 4 hydrophobic interactions with Leu163, Met209, Tyr274 and Thr302 (Figure 50E,F).

Guanine-N7 methyltransferase (nsp14) is important for viral replication and transcription. The N-terminal exoribonuclease (ExoN) domain plays a proofreading role in the prevention of lethal mutagenesis and the C-terminal domain functions as a guanine-N7 methyltransferase (N7-MTase) for mRNA capping [230]. The models were developed using as template the PDB ID 5C8S, which shows nsp14 in complex with its activator Nonstructural protein10 (nsp10) and two functional ligands: S-adenosyl-L-homocysteine (SAH) and guanosine-P3-adenosine-5,5' -triphosphate (G3A). One molecule of nsp10 interacts with ExoN of nsp14 to stabilize it and stimulate its activity. SAH and G3A bind the guanine-N7 methyltransferase site. The SAH binding pocket contains residues Trp292, Gly333, Asp352, Phe367 and Tyr368 [218] (Figure 50G). The derived pharmacophore model showed 5 features: 1 HBA with Tyr368, 3 HBDs, two with Asp352 and one with Tyr368 and hydrophobic interaction with Val290 (Figure 50H,I). The binding pocket engaging G3A contains the following residues: Trp292, Arg310, Gly333, Pro335, Lys336, Asn386, Asn388, Tyr420 and Phe426 (Figure 50J). Therefore, the derived pharmacophore model showed 10 features: 4 HBAs which interacted, respectively, with Cys309, Arg310, Trp385, Asn388, 3 HBDs, two with Cys309 and one with Asn422, 3 negative ionizable features at the 3 phosphate groups and an aromatic ring with Phe426 (Figure 50K,L). Nonstructural protein 16 (nsp16) also termed 2'-O-methyltransferase is activated only by the

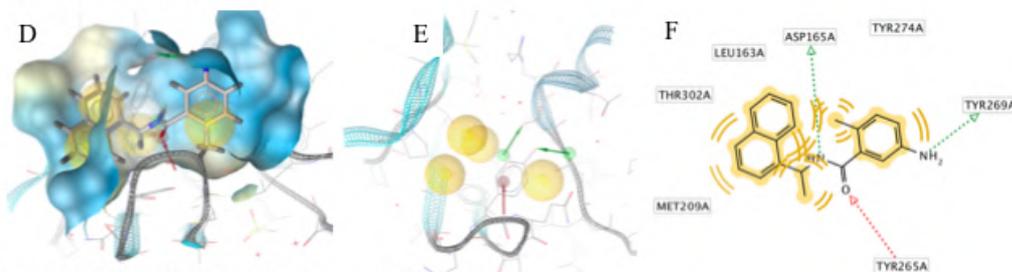
binding of nsp10. The structure of the nsp16–nsp10 complex from SARS-COV-2 with 1.80 Å of resolution (PDB ID: 6W4H) was considered. This complex shows S-adenosylmethionine (SAM) in the binding site. It forms hydrogen bonds with Asp6928, Tyr6930, Asp6897 and Cys6913 (Figure 50M). The derived pharmacophore model on the co-crystallized ligand showed 9 features: 4 HBAs with Gly248 and Thr341, 1 HBD with His250 and 2 negative ionizable areas with Gly248 and Lys290 (Figure 50N,O).

Moreover, the nsp16–nsp10 SARS coronavirus complex (PDB ID: 2XYQ), which shows S-adenosyl-L-homocysteine (SAH) in the binding site was used. SAH forms hydrogen bonds with Lys46, Asp130, Lys170 e Glu203 (Figure 50P). The derived pharmacophore model showed 9 features: 4 HBAs with Asn43, Leu100, Tyr Cys115, 4 HBDs with Gly71 and Asp99, 2 negative ionizable areas with Asp130 (Figure 50Q,R).

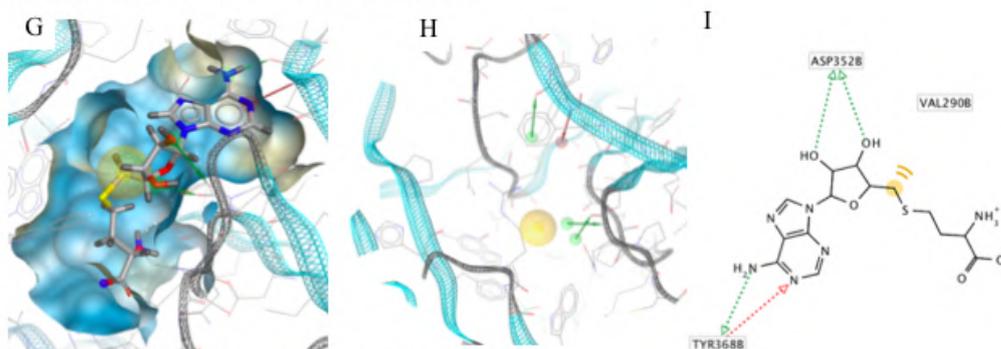
3C-like protease



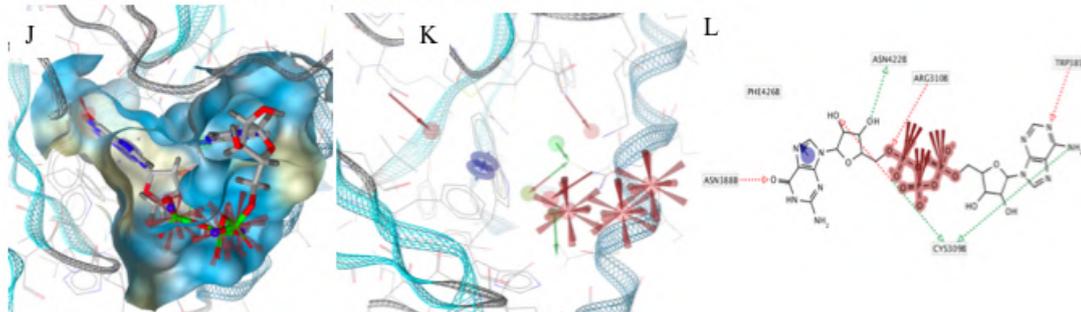
Papain-like Protease



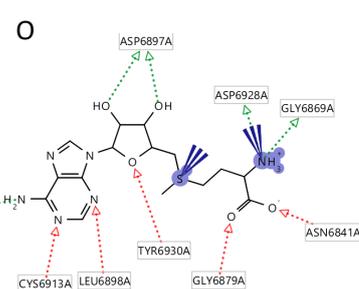
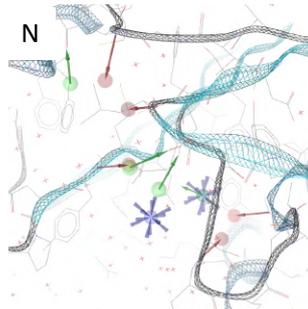
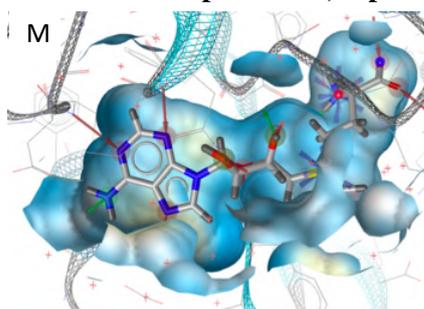
Guanine-N7 methyltransferase, nsp14 (SAH)



Guanine-N7 methyltransferase, nsp14 (G3A)



Non-structural protein 16, nsp16



Non-structural protein 16 and 10, nsp16-10

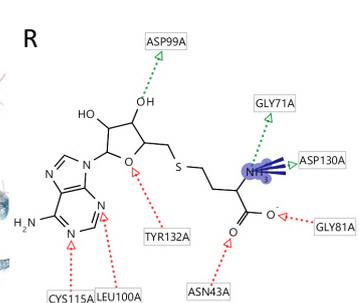
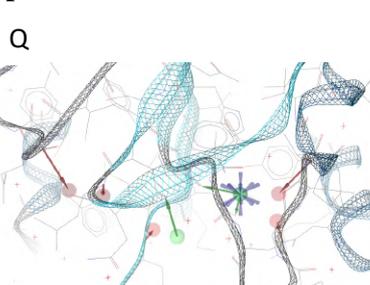
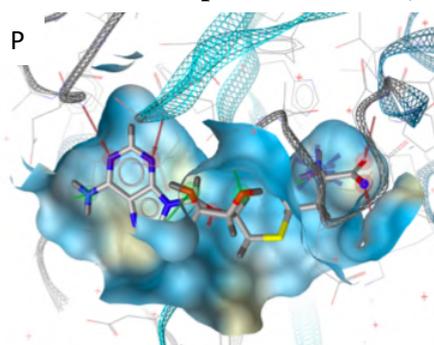


Figure 50: Pharmacophore modeling of ligand–protein complexes. For each structure molecular surface of the active site the co-crystallized ligand (A,D,G,J,M,P), structure-based pharmacophore model (B,E,H,K,N,Q) and ligand interactions (C,F,I,L,O,R) are shown

The other pharmacophore models were developed exploring the apoprotein surfaces as follows: uridylate-specific endoribonuclease (NendoU/nsp15) forms a hexameric endoribonuclease, that preferentially cleaves 3' of uridines. It is one of the RNA-processing enzymes encoded by the coronavirus [231]. Exploring the apoprotein surface, a potential active site was found, and a pharmacophore model was generated (Figure 51A). It contained the following residues: Thr166, Arg198, Asp267 and Ser273. The pharmacophore model showed 3 features: 2 HBDS and one hydrophobic feature.

Nonstructural protein 4 (nsp4) is localized at the endoplasmic reticulum membrane when expressed alone, but this protein can be recruited into the replication complex in infected cells [231]. After scanning the protein surface, a potential binding pocket was identified containing residues Leu417, Thr460 and Arg464. The derived pharmacophore model showed 6 features: 2 HBAs, 2 HBDS and a hydrophobic feature (Figure 51B).

Nonstructural protein 9 (nsp9), encoded by ORF1a, does not present a designated function, but is most likely involved with viral RNA synthesis. The crystal structure suggests that the protein is dimeric, whereas nsp9 binds RNA and interacts with nsp8 [232]. The potential identified binding site contains the following residues: Gly38, Arg39, Ser59 and Thr64. The derived pharmacophore model showed 6 features: 2 HBAs, 2 HBDS and one hydrophobic feature (Figure 51C).

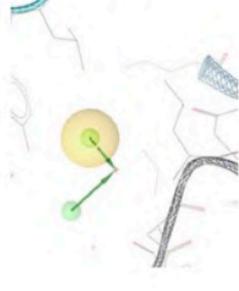
Helicase (hel) catalyzes the unwinding of duplex oligonucleotides into single strands in an NTP-dependent manner. The structure of SARS-CoV-2 nsp13 adopted a triangular pyramid shape comprising five domains. Among these, there are two “RecA-like” domains, 1A (261–441 a.a.) and 2A (442–596 a.a.) and 1B domain (150–260 a.a.) forming the triangular base, while N-terminal zinc-binding domain (ZBD) (1–99 a.a.) and stalk domain (100–149 a.a.), which connects ZBD and 1B domain, are arranged at the apex of the pyramid [222]. Exploring the apoprotein surface, two putative binding sites were found, pocket A and pocket B. Pocket

A contained residues from the stalk domain (Lys139, Lys146), 1B domain (Asn179) and 1A domain (Cys309, Arg339) and domain 1B (Thr228-Thr231) important for helicase activity. Pocket B contained residue from the N-terminal zinc-binding domain, ZBD domain, (Ile20, Arg21, Arg22) and stalk domain (Arg129). The pharmacophore models obtained for each pocket have the same 6 features: 2 HBAs, 2 HBDs and two hydrophobic features (Figure 51D-E).

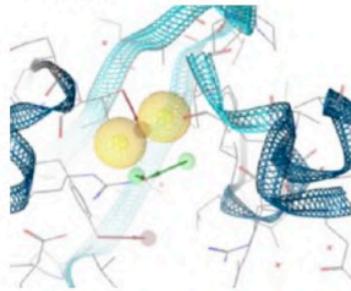
Nonstructural protein 7 and 8 (nsp7–nsp8) supercomplex are essential cofactors for Nsp12 polymerase [227]. Two putative active sites were found: pocket A and pocket B. Pocket A between chains C, G and H, pocket B between chain G–H of nsp8. The pocket A showed as residues: Glu50 of chain C; Thr124 and Arg190 of chain G; Glu5, Arg57, of chain H. The pocket B of chains G–H of nsp8 showed the residues: Arg57 and Asp64 of chain G; Leu122 and Thr123 of chain H. The pharmacophore model showed 6 features each: 2 HBAs, 2 HBDs and two hydrophobic features (Figure 51F,G).

Nonstructural protein 12 bound to nsp7-8 co-factors (nsp7–nsp8–nsp12) hetero-oligomeric complex is an RNA-dependent RNA polymerase. It is bound to its essential co-factors nsp7 and nsp8 greatly stimulates the replication and transcription activities of the polymerase. The nsp12 contains a polymerase domain (a.a. 398–919) that assumes a structure resembling a cupped “right hand”. The polymerase domain consists of a finger domain (a.a. 398–581, 628–687), a palm domain (a.a. 582–627, 688–815) and a thumb domain (a.a. 816–919). CoV nsp12 also contains a nidovirus-unique N-terminal extension (a.a. 1–397) [222]. The putative active sites, pocket A and pocket B were found into conserved motif regions (A–G) possessed of all polymerases [227]. Pocket A contained residues of N-terminal extension Thr246 and Arg249; pocket B contained residues of N-terminal extension Tyr129, His133, Asn138 and motif D (Ala706–Asp711), the pharmacophore model showed 6 features: 2 HBAs, 2 HBDs and two hydrophobic features (Figure 51H,I).

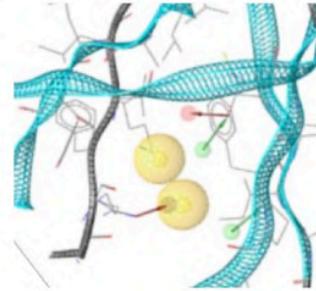
Uridylate-specific endoribonuclease,
nsp15



Non-structural protein 4,
nsp4

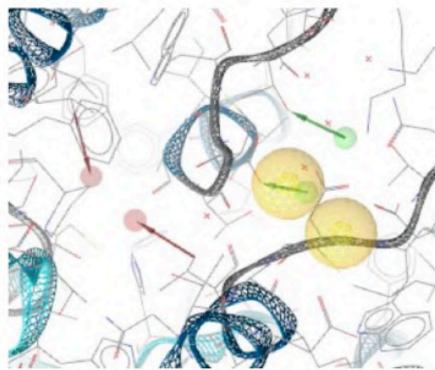


Non-structural protein 9,
nsp9

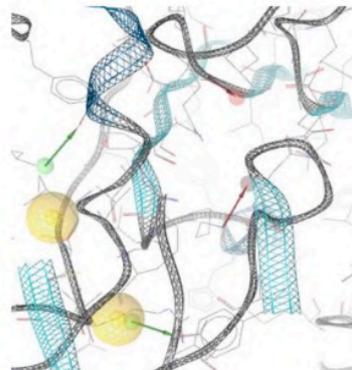


Helicase

Pocket A
D

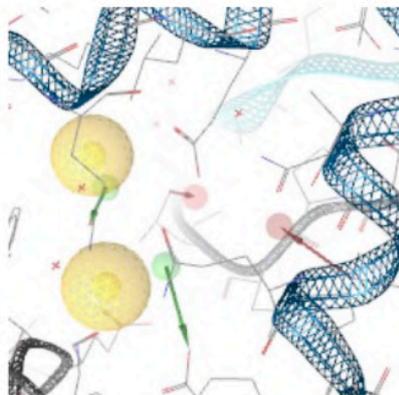


Pocket B
E

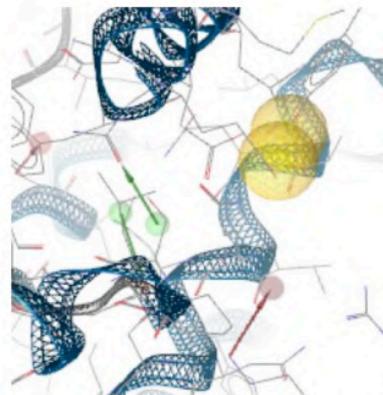


Non-structural protein 7 and 8, nsp7-nsp8

Pocket A
F

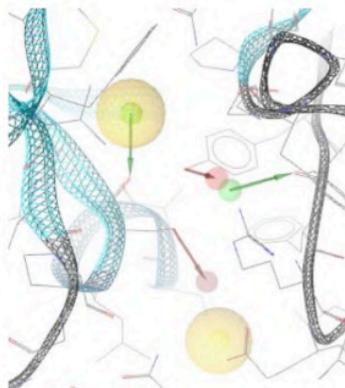


Pocket B
G



Non-structural protein 7-8 and 12, nsp7-nsp8-nsp12

Pocket A
H



Pocket B
I

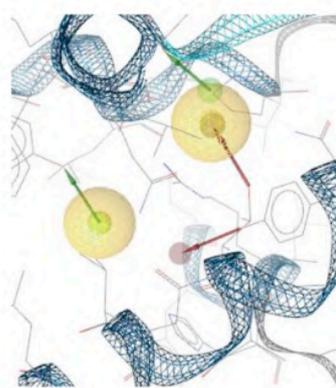


Figure 51: Pharmacophore modeling of the apoproteins. The first step was to find putative active sites, then the pharmacophore models were built in the identified pockets. Four of the structures had two active sites. (A) Uridylate-specific endoribonuclease, nsp15; (B) nsp4; (C) nsp9; (D) Helicase pocket A; (E) Helicase pocket B; (F) nsp7-nsp8 pocket A; (G) nsp7-nsp8 pocket B; (H) nsp7-8-12 pocket A; (I) nsp7-8-12 pocket B

The identified pharmacophore models were used to perform a virtual screening against the DrugBank database of experimental, investigational and approved drugs considering as a first filter.

The hits found were submitted to docking studies to evaluate the poses and interactions at the putative active site. First, XP docking was performed and subsequently, the highest-ranked hits were submitted to induced fit docking analysis and MM-GBSA calculation to further filter. For just one protein (nsp16), no hits were identified in the DrugBank database. At the end of the computational exploration, a total of 34 hits were identified for all the explored targets. Among these compounds, 26 are experimental drugs, 5 investigational drugs and 3 approved drugs. The summary results were reported in the Supporting Information. In the main text, the molecular recognition analysis for the best binder hits for each target will be discussed. The rest of the identified hits, docking scores, $\Delta G_{\text{binding}}$, and protein–ligand interactions is reported in a table in Supplementary Information as well as 2D ligand interaction diagrams of the best binders.

The best docked hit molecule for 3CL-protease is the experimental drug DB082309, a phenyl pyrroline derivative ($\Delta G = -72.56$ kcal/mol). This compound is characterized by an H-bond between the carbonyl oxygen with Asn142, but the principal contribution to the binding is given by the $\Delta G_{\text{vdW}} = -52.56$ kcal/mol and the $\Delta G_{\text{liPo}} = -23.65$ due to the 2 aromatic rings (phenyl and O-difluorophenyl) of the molecules which are located in two hydrophobic pockets (Leu140, Phe141, Leu167, Pro168) and the piperazine moiety interacting with His41 and Met49 (Figure 52A).

The most promising drug candidate for papain-like protease is the experimental drug DB07358 ($\Delta G = -50.662$ kcal/mol), a benzamide derivative. In our study, the experimental drug DB07358 forms three H-bonds with Tyr269, Gln270 and Tyr274. Moreover, the binding is characterized by a strong pi-stacking of the thiazol moiety with the phenyl ring of Tyr269 and phenylamino moiety with the phenyl ring of Tyr274 ($\Delta G_{\text{liPo}} = -19.47$ kcal/mol, $\Delta G_{\text{vdW}} = -38.50$ kcal/mol) (Figure 52B).

Top-ranked guanine-N7-methyltransferase (nsp14) hit is the experimental drug DB02933 as known as 5'-deoxy-5'-(methylthio)-tubercidin ($\Delta G = -65.07$ kcal/mol). This compound was previously identified as an inhibitor of the h-S-methyl-5'-thioadenosine phosphorylase and bacterial methylthioadenosinucleosidase. The compound 5'-deoxy-5'-(methylthio)-tubercidin showed 3 H-bond interactions with Asn386, Asn388 and Glu302, but the most contribution to the binding energy is due to pyrrole pyrimidine moiety, which establishes strong pi-stacking interaction with Tyr420 and Phe426 (Figure 52C).

Considering the NendoU/nsp15 protein, the most promising compound is the experimental drug DB01792 as known Adenylyl-(3'-5')-uridine 3'-monophosphate ($\Delta G = -63.169$ kcal/mol). The compound showed a high number of H-bond interactions with several different residues (Thr166, Ser197, Glu264, Asp272, Tyr278) (Figure 52D).

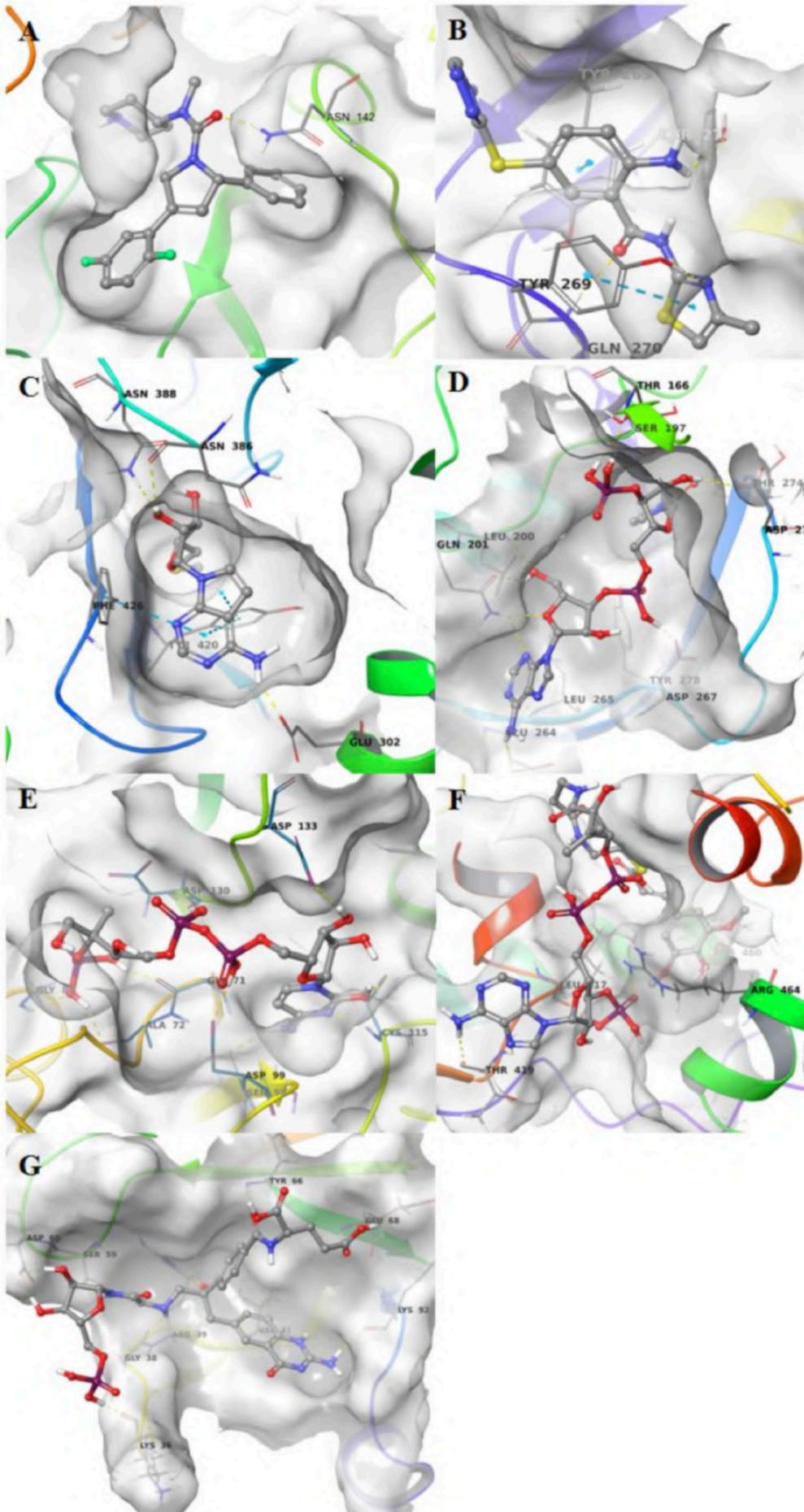
The experimental drug DB01859 resulted in the hit related to the nsp16. The compound is also known as 4-diphosphocytidyl-2-C-methyl-D-erythritol 2-phosphate ($\Delta G = -25.204$ kcal/mol). It showed 9 H-bond interactions with Gly71, Ala72, Gly81, Ser98, Asp99, Asp130 and Asp133. Residue Cys115 showed 2 H-bonds (Figure 52E).

The top-ranked compound for nsp4 is the experimental drug sinapoyl-coA ($DG = -80.73$ kcal/mol). The binding of sinapoyl-CoA in the nsp4 pocket is influenced by a high number of H-bonds with several different residues (Leu417, Thr419, Arg464, Thr460) (Figure 52F).

The experimental drug DB02794 resulted in the best binding hit related to the nsp9. Due to the

presence in the scaffold of many oxygen atoms, DB02794 establishes many H-bond interactions involving Lys36, Gly38, Arg39, Ser59, Asp60, Glu68. Other H-bond interactions involve some nitrogen of the experimental drug and the residues Gly38, Ser59 and Lys92. The strong net of H-bond interactions is reflected by a $\Delta G_{\text{coul}} = -84.35$ kcal/mol, partially compensated by a loss of binding energy due to the solvation contribution $\Delta G = +68.45$ kcal/mol. It is worthy to note that the next top-ranked hits for nsp9 are 3 approved drugs (ioxilan, Pemetrexed, and isoprenaline), which could be of particular interest due to the status “approved”, which would allow to use them in clinical trials (Figure 52G).

For the helicase, the apo binding pocket analysis identified 2 different putative binding sites. The most promising candidate drug-binding pocket A is the experimental drug 4-hydroxybenzoyl-coA ($\Delta G_{\text{bind}} = -91.90$ kcal/mol). The interactions that this compound establishes with the pocket A are characterized by several H-bonds, most of which formed by the three phosphate moieties with Lys139, Arg339, Asn361, Arg390. Other H-bond interactions are among the hydroxyl and carbonyl oxygens and Lys139, Glu142, Lys146, Asp179, His230, Cys309, Arg339, Arg390. Moreover, the purine moiety establishes pi-stacking interactions with the imidazole moiety of His230. Regarding the top-ranked compound in pocket B, this is the experimental drug DB02136, a cephalosporin analog, ($\Delta G_{\text{bind}} = -75.81$ kcal/mol). This compound interacts with the residues Ile20, Arg21, Arg22, Arg129, Glu136 forming H-Bonds with carbonyl and hydroxyl oxygen atoms, but the binding mode is strengthened by an important contribution of $\Delta G_{\text{vdW}} = -71.94$ kcal/mol (Figure 52H,I).



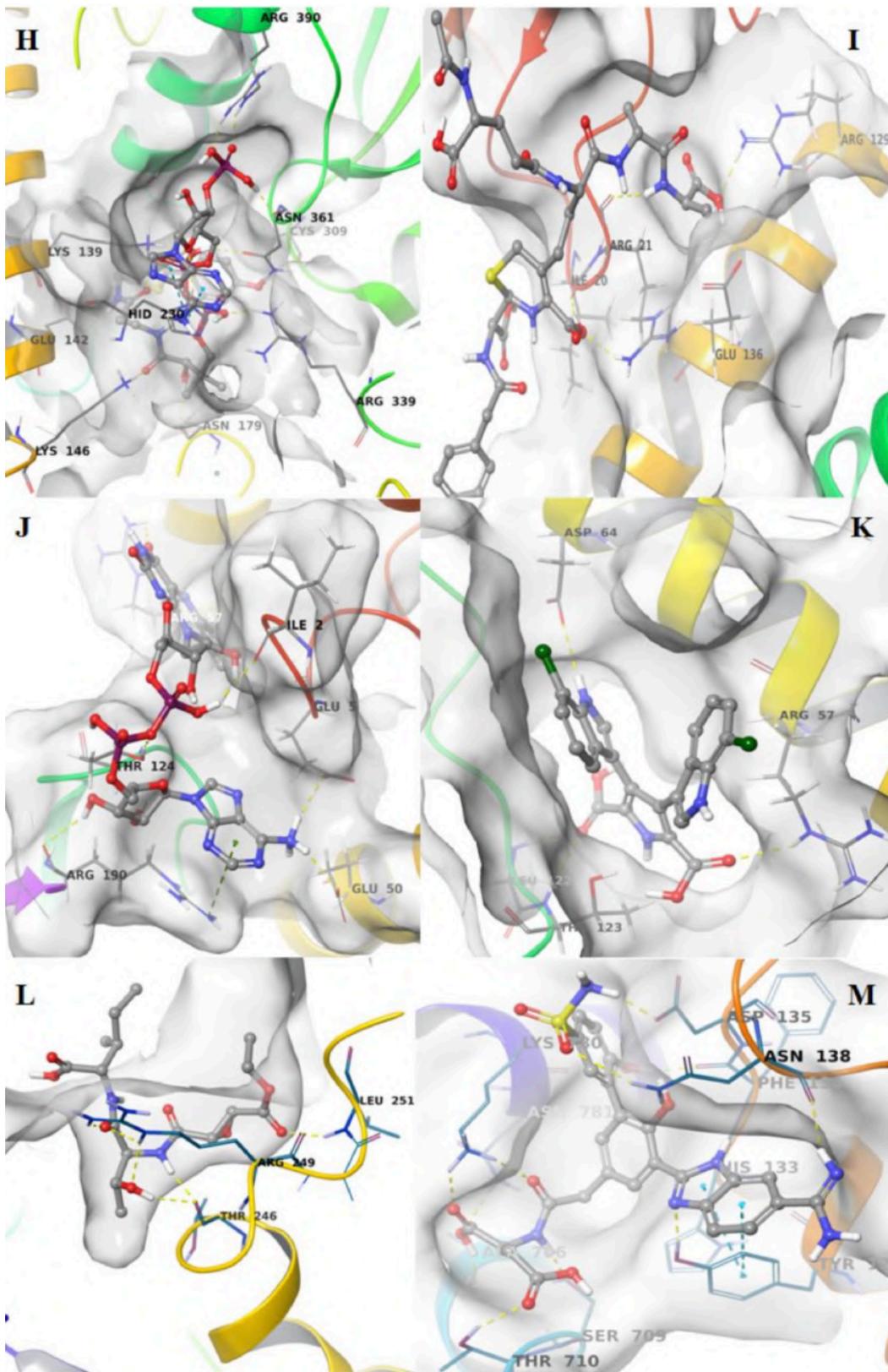


Figure 52: (A) DB08239 binding pose in 3C-like protease; (B) DB07358 binding pose in papain-like protease; (C) DB02933 binding pose in guanine-N7-methyltransferase (nsp14); (D) DB01792 binding pose in NendoU/nsp15; (E) DB01859 binding pose in nsp16; (F) synapoylCoenzyme A binding pose in nsp4; (G) DB02794 binding pose in nsp9.(H) DB04579 binding pose in helicase, pocket A; (I) PCI-27483 binding pose in helicase, pocket B; (J) flavin-N7 protonated-adenine dinucleotide binding pose in nsp7-8, pocket A; (K) DB06955 binding pocket in nsp7-8, pocket B; (L) DB04579 binding pose in nsp7-8-12, pocket A; (M) PCI-27483 in nsp7-8-12, pocket B.

Furthermore, for the supercomplex nsp7–nsp8, two different pockets were found. The most promising candidate for pocket A is the experimental drug flavin-N7 protonated-adenine dinucleotide ($\Delta G = 78.86$ kcal/mol). The flavin moiety interacts with the residue Arg57 forming 2 H-bonds. These latter are also formed among the phosphate and Thr190, the ribose moiety and Arg190 and the purine moiety and Ile2, Ile3, Ile4. Moreover, the binding interaction is strengthened by ionic interactions among the NH_3^+ and the glutamic residues 5 and 50. The residue Arg190 interacts with the purine moiety employing pi-stacking interactions. The top-ranked compound for pocket B is the experimental drug DB06955 ($\Delta G = -58.14$ kcal/mol), a pyrrole-indole derivative, interacting with Arg57, Asp64, Leu122 and Thr123 employing H-bond interactions (Figure 52J,K).

Last, but not least, for the hetero-oligomeric complex nsp7–nsp8–nsp12 two different pockets were identified. In pocket A, the most promising compound is the experimental peptide analog DB04579 ($\Delta G = -57.10$ kcal/mol) interacting with the residues Thr246, Arg249, Leu251, Ser255 through H-bond interactions. The most promising compound for the pocket B is the investigational drug PCI-27483, a phenyl benzimidazole derivative to date used for the treatment of the pancreatic adenocarcinoma. The binding mode is characterized by several H-bond interactions involving His133, Phe134, Asp135, Asn138, Ala708, Ser709, Thr710, Lys780 and Asn781. The indole moiety is further involved in pi-stacking interactions with Tyr129 (Figure 52L,M).

7.4 CONCLUSION

The recently emerged SARS-CoV-2 caused a major outbreak of COVID-19 and instigated a widespread fear and has threatened global health security because there are no approved therapies for treating. In the attempt to try to speed up the search for new inhibitors of the virus replication, in this study, a computational drug repositioning campaign on the DrugBank database of experimental, investigational, and approved drugs was performed. The aim of using such a restricted database had the rationale to identify potential lead compounds to quickly test *in vitro* and *in vivo* as they passed toxicity tests. The proteome of SARS-CoV-2 was analyzed and using homology modeling were identified the high-quality models of proteins. A structure-based pharmacophore modeling study was performed to identify pharmacophore features for each target. Successively, the pharmacophore models were used to perform a virtual screening against the DrugBank library. After a docking study, a total of 34 hits was identified for all the explored targets (3CL-protease, papain-like protease, guanine-N7-methyltransferase nsp14, nsp16, NendoU/nsp15, nsp4, nsp9, helicase, nsp7–nsp8 supercomplex and nsp7–nsp8–nsp12 hetero-oligomeric complex). Among these compounds, 26 are experimental drugs, five investigational drugs and three approved drugs. The final selection of the potential inhibitors was made considering the best binding energy for each compound obtained utilizing MM-GBSA calculation. Molecular recognition analysis showed that these compounds interact with the residues found as crucial for each target. These drugs can be further explored against the successful inhibition of COVID-19. Moreover, a set of hot spot residues and pharmacophore features for each target, which makes substantial contributions to the protein–ligand binding are also identified. This achievement can facilitate us to rationally design novel selective inhibitors targeting SARS-CoV-2, not comprised in the DrugBank. The results of this study offer a double important hint for anti-COVID19 drug discovery campaigns. On one side, it shows putative repurposing drugs to be adopted as a single therapy or in combination with other therapies. On the other side, our deep studies attempted to map out the main binding hot spots for the most important SARS-CoV-2 proteins, opening an important route to the design of new molecules to test.

8. REFERENCES

- [1] Paul, S.M.; Mytelka, D.S.; Dunwiddie, C.T.; Persinger, C.C.; Munos, B.H.; Lindborg, S.R.; Schacht, A.L. How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat Rev Drug Discov.* **2010**, 9, 203-14.
- [2] Shirai, H.; Prades, C.; Vita, R.; Marcatili, P.; Popovic, B.; Xu, J.; Overington, J. P.; Hirayama, K.; Soga, S.; Tsunoyama, K.; Clark, D.; Lefranc, M. P.; Ikeda, K. Antibody informatics for drug discovery. *Biochim Biophys Acta*, **2014**, 1844, 2002–15.
- [3] Van Drie JH. Computer-aided drug design: the next 20 years. *J Comput Aided Mol Des.* **2007**, 21, 591–601.
- [4] Arya, H.; Coumar, M.S. Lead identification and optimization. *The Design & Development of Novel Drugs and Vaccines.* **2021**, 31-63.
- [5] Lauria, A.; Tutone, M.; Almerico, A.M. Virtual lock-and-key approach: The in-silico revival of Fischer model using molecular descriptors. *Eur. J. Med. Chem.* **2011**, 46, 4274-80.
- [6] Wolber, G.; Sippl, W. Pharmacophore Identification and Pseudo-Receptor Modeling. *The Practice of Medicinal Chemistry: Fourth Edition.* **2015**, 489–510.
- [7] Langer, T.; Wolber, G., Pharmacophore definition and 3D searches. *Drug Discov. Today Technol.* **2004**, 1, 203–207.
- [8] Wermuth, C.G.; Ganellin, C.R.; Lindberg, P.; Mitscher, L.A. Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). *Pure Appl. Chem.* **1998**, 70, 1129-43.
- [9] Verma, J.; Khedkar, V.M.; Coutinho, E.C. 3D-QSAR in Drug Design - A Review. *Curr. Top. Med. Chem.* **2010**, 10, 95–115.
- [10] Kubinyi, H. QSAR and 3D QSAR in drug design. *Drug Discov. Today.* **1997**, 2, 457–67.
- [11] Schneider, M.; Fu, X.; Keating, A.E. X-ray vs. NMR structures as templates for computational protein design. *Proteins.* **2009**, 77, 97–110.
- [12] Robinson, S.W.; Afzal, A.M.; Leader, D.P. Bioinformatics: Concepts, Methods, and Data. *Handbook of Pharmacogenomics and Stratified Medicine.* **2014**, 259-87.
- [13] Peng, J.; Xu, J. Low-homology protein threading. *Bioinformatics.* **2010**, 26
- [14] Langer, T. Pharmacophores for medicinal chemists: a personal view. *Fut Med Chem.* **2011**, 3, 8.
- [15] Gerhard, W.; Langer, T. LigandScout: 3-D Pharmacophores Derived from Protein-Bound Ligands and Their Use as Virtual Screening Filters. *J Chem Inf Model.* **2004**, 45, 160-9.
- [16] Morris, G.M.; Lim-Wilby, M. Molecular docking. *Methods Mol. Biol.* **2008**, 443, 365-82.
- [17] Hospital, A.; Goñi, JR; Orozco, M; Gelpi, J. Molecular dynamics simulations: advances and applications. *Adv Appl Bioinform Chem.* **2015**; 8, 37-47.
- [18] Langer, T.; Hoffmann, RD. Pharmacophore modelling: applications in drug discovery. *Expert Opin Drug Discov.* **2006**, 1, 261-7.
- [19] Pirhadi, S.; Shiri, F.; Ghasemi, JB. Methods, and applications of structure-based pharmacophores in drug discovery. *Curr Top Med Chem.* **2013**, 13, 1036-47.
- [20] McConkey, B.J.; Sobolev, V.; Edelman, M. The performance of current methods in ligand-protein docking. *Curr Science.* **2002**, 83, 845-55.
- [21] Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* **1982**, 161, 269-88.
- [22] Fischer, E. Einfluss der configuration auf die wirkung der enzyme. *Ber. Dt. Chem. Ges.* **1894**, 27, 2985-93.
- [23] Kearsley, S.K.; Underwood, D.J.; Sheridan, R.P.; Miller, M.D. Flexibase: A way to enhance the use of molecular docking methods. *J. Comput. Aided. Mol. Des.* **1994**, 8, 565-82.
- [24] Klebe, G.; Mietzner, T. A fast and efficient method to generate biologically relevant conformations. *J. Comput. Aided. Mol. Des.* **1994**, 8, 583-606.

- [25] Wang, Q.; Pang, Y.P. Preference of small molecules for local minimum conformations when binding to proteins. *PLoS One*, **2007**, 2.
- [26] Koshland, D.E. Jr. Correlation of Structure and Function in Enzyme Action. *Science*. **1963**, 142, 1533-41.
- [27] Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R., Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects, *J. Med. Chem.* **2006**, 49, 534.
- [28] Rahman, A. 1964. Correlations in the motion of atoms in liquid argon. *Physical Review*. **1964**, 136.
- [29] De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* **2016**, 59, 4035-61.
- [30] Sabbadin, D.; Moro, S. Supervised Molecular Dynamics (SuMD) as a Helpful Tool To Depict GPCR–Ligand Recognition Pathway in a Nanosecond Time Scale. *J. Chem. Inf. Model.* **2014**, 54, 372-76.
- [31] Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett.* **1999**, 314, 141-51.
- [32] Abrams, C.; Bussi, G. Enhanced Sampling in Molecular Dynamics Using Metadynamics, Replica-Exchange, and Temperature-Acceleration. *Entropy*. **2013**, 16, 163–199.
- [33] Laio, A.; Parrinello, M. Escaping free-energy minima. *PNAS*, **2002**, 99, 12562-66.
- [34] Wieder, M.; Garon, A.; Perricone, U.; Boresch, S.; Seidel, T.; Almerico, A.M.; Langer, T. Common Hits Approach, Combining pharmacophore modeling and molecular dynamics simulations. *J. Chem. Inf. Model.* **2017**, 57, 365-85
- [35] Perricone, U.; Wieder, M.; Seidel, T.; Langer, T.; Padova, A.; Almerico, A. M.; Tutone, M. A Molecular Dynamics-Shared Pharmacophore Approach to Boost Early-Enrichment Virtual Screening: A Case Study on Peroxisome Proliferator-Activated Receptor α . *ChemMedChem*. **2017**, 12, 1399-1407.
- [36] Xi, J.; Zhuang, R.; Kong, L.; He, R.; Zhu, H.; Zhang, J. Immunoproteasome-selective inhibitors: An overview of recent developments as potential drugs for hematologic malignancies and autoimmune diseases. *Eur. J. Med. Chem.* **2019**, 182, 111646.
- [37] Maccari, R.; Ettari, R.; Adornato, A.; Na, A.; Wolber, G.; Bitto, A.; Mannino, F.; Aliquò, F.; Bruno, G.; Nicolò, F.; Previti, S.; Grasso, S.; Zappalà, M.; Ottanà, R. Identification of 2-thioxoimidazolidin-4-one derivatives as novel non-covalent proteasome and immunoproteasome inhibitors. *Bioorganic Med. Chem. Lett.* **2018**, 28, 278-83.
- [38] Ferrington, D.A.; Gregerson, D.S. Immunoproteasomes: Structure, function, and antigen presentation. *Progress in Molecular Biology and Translational Science*. **2012**, 109, 75-112.
- [39] Ebstein, F.; Kloetzel, P.M.; Krüger, E.; Seifert, U. Emerging roles of immunoproteasomes beyond MHC class I antigen processing. *Cell. Mol. Life Sci.* **2012**, 69, 2543–58.
- [40] Ichihara, A. Tanaka, K. Roles of proteasomes in cell growth. *Mol. Biol. Rep*, **1995**, 21, 49-52.
- [41] Kimura, H.; Caturegli, P.; Takahashi, M.; Suzuki, K. New Insights into the Function of the Immunoproteasome in Immune and Nonimmune Cells. *J. Immunol. Res.* **2015**, 1-8.
- [42] Wu, P.Y.; Hanlon, M.; Eddins, M.; Tsui, C.; Rogers, R.S.; Jensen, J.P.; Matunis, M.J.; Weissman, A.M.; Wolberger, C.P.; Pickart C.M. A conserved catalytic residue in the ubiquitin-conjugating enzyme family. *EMBO J.* **2003**, 22, 5241-50.
- [43] Wagner, L.K.; Gillig, K.E.; Schormann, E.; Kloetzel, P.M.; Heppner, F.L.; Krüger, E.; Prokop, S. Immunoproteasome deficiency alters microglial cytokine response and improves cognitive deficits in Alzheimer’s disease-like APPS1 mice. *Acta Neuropathol. Commun.* **2017**, 5, 52.
- [44] Liu, Y.; Qiao, F.; Leiferman, P.C.; Ross, A.; Schlenker, E.H.; Wang, H. FOXOs modulate proteasome activity in human-induced pluripotent stem cells of Huntington’s disease and their derived neural cells. *Hum. Mol. Genet.* **2017**, 26, 4416-28.

- [45] Lin, G.; Mao, D.; Bellen H.J. Amyotrophic Lateral Sclerosis Pathogenesis Converges on Defects in Protein Homeostasis Associated with TDP-43 Mislocalization and Proteasome-Mediated Degradation Overload. *Fly Model. Hum. Dis., Academic Press.* **2017**, 111-71.
- [46] Seifert, U.; Bialy, L.P.; Ebstein, F.; Bech-Otschir, D.; Voigt, A.; Schröter, F.; Prozorovski, T.; Lange, N.; Steffen, J.; Rieger, M.; Kuckelkorn, U.; Aktas, O.; Kloetzel, P.M.; Krüger, E. Immunoproteasomes preserve protein homeostasis upon interferon-induced oxidative stress. *Cell.* **2010**, 142, 613-24.
- [47] Richardson, P.G.; Hideshima, T.; Anderson, K.C. Bortezomib (PS-341): A Novel, First-in-Class Proteasome Inhibitor for the Treatment of Multiple Myeloma and other Cancers. *Cancer Control.* **2003**, 10, 361-69.
- [48] Barr, P.; Fisher, R.; Friedberg, J. The Role of Bortezomib in the Treatment of Lymphoma. *Cancer Invest.* **2007**, 25, 766-75.
- [49] Ettari, R.; Previti, S.; Bitto, A.; Grasso, S.; Zappalà, M. Immunoproteasome-Selective Inhibitors: A Promising Strategy to Treat Hematologic Malignancies, Autoimmune and Inflammatory Diseases. *Curr. Med. Chem.* **2016**, 23, 1217-38.
- [50] Kupperman, E.; Lee, E.C.; Cao, Y.; Bannerman, B.; Fitzgerald, M.; Berger, A.; Yu, J.; Yang, Y.; Hales, P.; Bruzzese, F.; Liu, J.; Blank, J.; Garcia, K.; Tsu, C.; Dick, L.; Fleming, P.; Yu, L.; Manfredi, M.; Rolfe, M.; Bolen, J. Evaluation of the proteasome inhibitor MLN9708 in preclinical models of human cancer. *Cancer Res.* **2010**, 70, 1970-80.
- [51] Kuhn, D.J.; Orłowski, R.Z. The immunoproteasome as a target in hematologic malignancies., *Semin. Hematol.* **2012**, 49, 258-62.
- [52] Chauhan, D.; Hideshima, T.; Anderson, K.C. A novel proteasome inhibitor NPI-0052 as an anticancer therapy. *Br. J. Cancer.* **2006**, 95, 961-65.
- [53] Floden, C.; Combs, A. NIH Public Access. *Bone.* **2012**, 23, 1-7.
- [54] Li, Y.Y.; Jones, S.J. Drug repositioning for personalized medicine. *Genome Med.* **2012**, 4, 1-14.
- [55] Afifi, S.; Michael, A.; Lesokhin, A. Immunotherapy: A New Approach to Treating Multiple Myeloma with Daratumumab and Elotuzumab. *Ann. Pharmacother.* **2016**, 50, 555-68.
- [56] Potts, B. C.; Albitar, M. X.; Anderson, K. C.; Baritaki, S.; Berkers, C.; Bonavida, B.; Chandra, J.; Chauhan, D.; Cusack, J. C.; Jr, Fenical, W.; Ghobrial, I. M.; Groll, M.; Jensen, P. R.; Lam, K. S.; Lloyd, G. K.; McBride, W.; McConkey, D. J.; Miller, C. P.; Neuteboom, S. T.; Oki, Y.; Palladino, M.A. Marizomib, a Proteasome Inhibitor for All Seasons: Preclinical Profile and a Framework for Clinical Trials. *Curr. Cancer Drug Targets.* **2011**, 11, 254-84.
- [57] Muchamuel, T.; Basler, M.; Aujay, M.A.; Suzuki, E.; Kalim, K.W.; Lauer, C.; Sylvain, C.; Ring, E.R.; Shields, J.; Jiang, J.; Shwonek, P.; Parlati, F.; Demo, S. D.; Bennett, M. K.; Kirk, C. J.; Groettrup, M. A selective inhibitor of the immunoproteasome subunit LMP7 blocks cytokine production and attenuates progression of experimental arthritis. *Nat. Med.* **2009**, 15, 781-87.
- [58] Huber, E.M.; Basler, M.; Schwab, R.; Heinemeyer, W.; Kirk, C.J.; Groettrup, M.; Groll, M. Immuno- and constitutive proteasome crystal structures reveal differences in substrate and inhibitor specificity. *Cell.* **2012**, 148, 727-38.
- [59] Singh, A.V.; Bandi, M.; Aujay, M.A.; Kirk, C.J.; Hark, D.E.; Raju, N.; Chauhan, D.; Anderson, K.C. PR- 924, a selective inhibitor of the immunoproteasome subunit LMP-7, blocks multiple myeloma cell growth both *in vitro* and *in vivo*. *Br. J. Haematol.* **2011**, 152, 155-63.
- [60] Johnson, H.W.B.; Lowe, E.; Anderl, J.L.; Fan, A.; Muchamuel, T.; Bowers, S.; Moebius, D.C.; Kirk, C.; McMinn, D.L. Required Immunoproteasome Subunit Inhibition Profile for Anti-Inflammatory Efficacy and Clinical Candidate KZR-616 ((2 S,3 R)- N-((S)-3-(Cyclopent-1-en-1-yl)-1-((R)-2- methyloxiran-2-yl)-1-oxopropan-2-yl)-3-hydroxy-3-(4-methoxyphenyl)-2-((S)-2-(2-morpholin. *J. Med. Chem.* **2018**, 6, 11127-43.

- [61] Ho, Y. K.; Bargagna-Mohan, P.; Wehenkel, M.; Mohan, R.; Kim, K. B. LMP2- specific inhibitors: chemical genetic tools for proteasome biology. *Chem Biol.* **2007**, *14*, 419-30.
- [62] de Bruin, G.; Huber, E. M.; Xin, B. T.; Van Rooden, E. J.; Al-Ayed, K.; Kim, K. B.; Kisselev, A. F.; Driessen, C.; Van der Stelt, M.; Van der Marel, G. A.; Groll, M.; Overkleeft, H. S. Structure-based design of β 1i or β 5i specific inhibitors of human immunoproteasomes. *J Med Chem.* **2014**, *57*, 6197-209.
- [63] Kuhn, D. J.; Hunsucker, S. A.; Chen, Q.; Voorhees, P. M.; Orłowski, M.; Orłowski, R. Z. Targeted inhibition of the immunoproteasome is a potent strategy against models of multiple myeloma that overcomes resistance to conventional drugs and nonspecific proteasome inhibitors. *Blood.* **2009**, *113*, 4667-76.
- [64] Singh, P.K.; Fan, H.; Jiang, X.; Shi, L.; Nathan, C.F.; Lin, G. Immunoproteasome β 5i-Selective Dipeptidomimetic Inhibitors. *ChemMedChem.* **2016**, 2127-31.
- [65] Allardyce, D.J.; Bell, C.M.; Loizidou, E.Z. Argyrin B, a Non-Competitive Inhibitor of the Human Immunoproteasome Exhibiting Preference for β 1i. *Chem. Biol. Drug Des.* **2019**, *94*, 1556-67.
- [66] Zhan, W.; Singh, P.K.; Ban, Y.; Qing, X.; Ah Kioon, M.D.; Fan, H.; Zhao, Q.; Wang, R.; Sukenick, G.; Salmon, J. Structure– Activity Relationships of Noncovalent Immunoproteasome β 5i-Selective Dipeptides. *J. Med. Chem.* **2020**, *63*, 13103-23
- [67] Ettari, R.; Cerchia, C.; Maiorana, S.; Guccione, M.; Novellino, E.; Bitto, A.; Grasso, S.; Lavecchia, A.; Zappalà, M. Development of Novel Amides as Noncovalent Inhibitors of Immunoproteasomes. *ChemMedChem.* **2019**, *14*, 842-52.
- [68] Spitaleri, A.; Decherchi, S.; Cavalli, A.; Rocchia, W. Fast Dynamic Docking Guided by Adaptive Electrostatic Bias: The MD-Binding Approach. *J. Chem. Theory Comput.* **2018**, *14*, 1727-36.
- [69] Fusani, L.; Palmer, D.S.; Somers, D.O.; Wall, I.D. Exploring Ligand Stability in Protein Crystal Structures Using Binding Pose Metadynamics. *J. Chem. Inf. Model.* **2020**, *60*, 1528-39.
- [70] Sherman, W.; Day, T.; Jacobson, M.P.; Friesner, R.A.; Farid, R. Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects. *J. Med. Chem.* **2006**, *49*, 534-53.
- [71] Huber, E.M.; Basler, M.; Schwab, R.; Heinemeyer, W.; Kirk, C.J.; Groettrup, M.; Groll, M. Immuno- and constitutive proteasome crystal structures reveal differences in substrate and inhibitor specificity. *Cell.* **2012**, *148*, 727-38.
- [72] Selvaraj, C.; Panwar, U.; Dinesh, D.C.; Boura, E.; Singh, P.; Dubey, V.K.; Singh, S.K. Microsecond MD Simulation and Multiple-Conformation Virtual Screening to Identify Potential Anti-COVID-19 Inhibitors Against SARS-CoV-2 Main Protease. *Front. Chem.* **2021**, *8*, 595273.
- [73] Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments. *J. Comput Aided Mol Des.* **2013**, *27*, 221-34.
- [74] Prime; Schrödinger, LLC: New York, NY, USA, **2021**.
- [75] Epik; Schrödinger, LLC: New York, NY, USA, **2021**.
- [76] Decherchi, S.; Bottegoni, G.; Spitaleri, A.; Rocchia, W.; Cavalli, A. BiKi Life Sciences: A New Suite for Molecular Dynamics and Related Methods in Drug Discovery. *J. Chem. Inf. Model.* **2018**, *58*, 219-24.
- [77] Lexa, K.W.; Carlson, H.A. Protein flexibility in docking and surface mapping. *Q. Rev. Biophys.* **2012**, *45*, 301-43.
- [78] Decherchi, S.; Rocchia, W. A general and Robust Ray-Casting-Based Algorithm for Triangulating Surfaces at the Nanoscale. *PLoS ONE*, **2013**, *8*, e59744.
- [79] Motta, S.; Callea, L.; Giani Tagliabue, S.; Bonati, L. Exploring the PXR Ligand-Binding Mechanism with Advanced Molecular Dynamics Methods. *Sci. Rep.* **2018**, *8*, 16207.
- [80] Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926-35.

- [81] Lindorff-Larsen, J.K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J.L.; Dror, R.O.; Shaw, D.E. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins Struct. Funct. Genet.* **2010**, *78*, 1950-58.
- [82] Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
- [83] Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.F.; DiNola, A.; Haak, J.R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684-90.
- [84] Bowers, K.J.; Chow, E.; Xu, H.; Dror, R.O.; Eastwood, M.P.; Gregersen, B.A.; Klepeis, J.L.; Kolossvary, I.; Moraes, M.A.; Sacerdoti, F.D. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. In *Proceedings of the 2006 ACM/IEEE SC/06 Conference (SC'06); Association for Computing Machinery: Tampa.* **2006**; ISBN 978-0-7695-2700-0.
- [85] Robertson, M.J.; Tirado-Rives, J.; Jorgensen, W.L. Improved Peptide and Protein Torsional Energetics with the OPLS-AA Force Field. *J. Chem. Theory Comput.* **2015**, *11*, 3499-3509.
- [86] Dyer, K.M.; Perkyns, J.S.; Stell, G.; Pettitt, B.M. Site-renormalised Molecular Fluid Theory: On the Utility of a Two-site Model of Water. *Mol Phys.* **2009**, *107*, 423-31.
- [87] Almerico, A.M.; Tutone, M.; Pantano, L.; Lauria, A. Molecular Dynamics Studies on Mdm2 Complexes: An Analysis of the Inhibitor Influence. *Biochem Biophys Res Commun.* **2012**, *424*, 341-47.
- [88] Li, J.; Abel, R.; Zhu, K.; Cao, Y.; Zhao, S.; Friesner, R. A., "The VSGB 2.0 model: a next generation energy model for high resolution protein structure modeling," *Proteins.* **2011**, *79*, 2794-812.
- [89] Huber, E.M.; Basler, M.; Schwab, R.; Heinemeyer, W.; Kirk, C.J.; Groettrup, M.; Groll, M. Immuno- and Constitutive Proteasome Crystal Structures Reveal Differences in Substrate and Inhibitor Specificity. *Cell.* **2012**, *148*, 727-38.
- [90] 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*, 6582-94.
- [91] Lipinski, C.A. Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov. Today Technol.* **2004**, *1*, 337-41.
- [92] Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des.* **2013**, *27*, 221-34.
- [93] 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**, *45*, 160-69.
- [94] Phase, Schrödinger, LLC, New York, NY, **2021**.
- [95] Glide, Schrödinger, LLC, New York, NY, **2021**.
- [96] Huber E.M.; Groll M. Inhibitors for the immuno- and constitutive proteasome: Current and future trends in drug development. *Angew. Chem Int Ed Engl.* **2012**, *51*, 8708-20.
- [97] Genin, E.; Reboud-Ravaux, M.; Vidal, J. Proteasome inhibitors: recent advances and new perspectives in medicinal chemistry. *Curr Top Med Chem.* **2010**, *10*, 232-56.
- [98] Cong, Y.-S.; Wright, W.E.; Shay, J.W. Human Telomerase and Its Regulation. *Microbiol Mol Biol Rev.* **2002**, *66*, 407-25.
- [99] Wojtyla, A.; Gladych, M.; Rubis, B. Human telomerase activity regulation. *Mol Biol Rep.* **2011**, *38*, 3339-49.
- [100] Wang, F.; Podell, E.R.; Zaug, A.J.; Yang, Y.; Baciou, P.; Cech, T.R.; Lei, M. The POT1-TPP1 telomere complex is a telomerase processivity factor. *Nature.* **2007**, *445*, 506-10.
- [101] Masutomi, K.; Yu, E.Y.; Khurts, S.; Ben-Porath, I.; Currier, J.L.; Metz, G.B.; Brooks, M.W.; Kaneko, S.; Murakami, S.; DeCaprio, J.A. Telomerase maintains telomere structure in normal human cells. *Cell.* **2003**, *114*, 241-53.

- [102] Shay, J.W.; Wright, W.E. Role of telomeres and telomerase in cancer. *Semin Cancer Biol.* **2011**, *21*, 349-53.
- [103] Blackburn, E.H. Telomere states and cell fates. *Nature.* **2000**, *408*, 53–56.
- [104] Ramlee, M.K.; Wang, J.; Toh, W.X.; Li, S. Transcription regulation of the human telomerase reverse transcriptase (hTERT) gene. *Genes.* **2016**, *7*, 50.
- [105] Hanahan, D.; Weinberg, R.A. Hallmarks of cancer: The next generation. *Cell.* **2011**, *144*, 646-74.
- [106] Ruden, M.; Puri, N. Novel anticancer therapeutics targeting telomerase. *Cancer Treat Rev.* **2013**, *39*, 444-56
- [107] Arndt, G.M.; MacKenzie, K. New prospects for targeting telomerase beyond the telomere. *Nat Rev Cancer.* **2016**, *16*, 508-24.
- [108] Jäger, K.; Walter, M. Therapeutic Targeting of Telomerase. *Genes.* **2016**, *7*, 39.
- [109] Neidle, S. Human telomeric G-quadruplex: The current status of telomeric G-quadruplexes as therapeutic targets in human cancer. *FEBS J.* **2009**, *277*, 1118-25.
- [110] Islam, M.K.; Jackson, P.J.; Rahman, K.M.; Thurston, D.E. Recent advances in targeting the telomeric G-quadruplex DNA sequence with small molecules as a strategy for anticancer therapies. *Futur Med Chem.* **2016**, *8*, 1259-90.
- [111] Rahman, K.M.; Tizkova, K.; Reszka, A.P.; Neidle, S.; Thurston, D.E. Identification of novel telomeric G-quadruplex-targeting chemical scaffolds through screening of three NCI libraries. *Bioorganic Med Chem Lett.* **2012**, *22*, 3006-10.
- [112] Lauria, A.; Terenzi, A.; Bartolotta, R.; Bonsignore, R.; Perricone, U.; Tutone, M.; Martorana, A.; Barone, G.; Almerico, A.M. Does ligand symmetry play a role in the stabilization of DNA g-quadruplex host-guest complexes? *Curr Med Chem.* **2014**, *21*, 2665-90.
- [113] Barata, P.; Sood, A.K.; Hong, D.S. RNA-targeted therapeutics in cancer clinical trials: Current status and future directions. *Cancer Treat. Rev.* **2016**, *50*, 35-47.
- [114] Asai, A.; Oshima, Y.; Yamamoto, Y.; Uochi, T.; Kusaka, H.; Akinaga, S.; Yamashita, Y.; Pongracz, K.; Pruzan, R.; Wunder, E.; et al. A novel telomerase template antagonist (GRN163) as a potential anticancer agent. *Cancer Res.* **2003**, *63*, 3931-39.
- [115] Damm, K.; Hemmann, U.; Garin-Chesa, P.; Huel, N.; Kauffmann, I.; Priepke, H.; Niestroj, C.; Daiber, C.; Enenkel, B.; Guilliard, B.; et al. A highly selective telomerase inhibitor limiting human cancer cell proliferation. *EMBO J.* **2001**, *20*, 6958–6968, doi:10.1093/emboj/20.24.6958.
- [116] Man, R.-J.; Chen, L.-W.; Zhu, H.-L. Telomerase inhibitors: a patent review (2010–2015). *Expert Opin. Ther. Patents* **2016**, *26*, 679–688.
- [117] Baginski, M.; Serbakowska, K. *In silico* design of telomerase inhibitors. *Drug Discov. Today* **2020**, *25*, 1213–1222.
- [118] Bryan, C.; Rice, C.; Hoffman, H.; Harkisheimer, M.; Sweeney, M.; Skordalakes, E. Structural Basis of Telomerase Inhibition by the Highly Specific BIBR1532. *Structure.* **2015**, *23*, 1934-42.
- [119] Alam, M.S.; Lee, D.-U. Synthesis, biological evaluation, drug-likeness, and *in silico* screening of novel benzylidene-hydrazone analogues as small molecule anticancer agents. *Arch Pharmacol Res.* **2015**, *39*, 191-201.
- [120] Cai, Z.-Y.; Yang, Y.; Liu, X.-H.; Qi, X.-B. Novel 3-(1-acetyl-5-(substituted-phenyl)-4,5-dihydro-1H-pyrazol-3-yl)-7-fluoro-2H-chromen-2-one Derivatives: Synthesis and Anticancer Activity. *Lett Drug Des Discov.* **2010**, *7*, 640-43.
- [121] Liu, X.-H.; Li, J.; Shi, J.B.; Song, B.-A.; Qi, X.-B. Design and synthesis of novel 5-phenyl-N-piperidine ethanone containing 4,5-dihydropyrazole derivatives as potential antitumor agents. *Eur J Med Chem.* **2012**, *51*, 294-99.
- [122] Liu, X.-H.; Liu, H.-F.; Chen, J.; Yang, Y.; Song, B.-A.; Bai, L.-S.; Liu, J.-X.; Zhu, H.-L.; Qi, X.-B. Synthesis and molecular docking study of novel coumarin derivatives containing

- 4,5-dihydropyrazole moiety as potential antitumor agents. *Bioorganic Med Chem Lett.* **2010**, 20, 5705-08.
- [123] Luo, Y.; Zhang, S.; Qiu, K.-M.; Liu, Z.-J.; Yang, Y.-S.; Fu, J.; Zhong, W.-Q.; Zhu, H.-L. Synthesis, biological evaluation, 3D-QSAR studies of novel aryl-2H-pyrazole derivatives as telomerase inhibitors. *Bioorganic Med Chem Lett.* **2013**, 23, 1091-95.
- [124] Luo, Y.; Zhou, Y.; Fu, J.; Zhu, H.-L. 4,5-Dihydropyrazole derivatives containing oxygen-bearing heterocycles as potential telomerase inhibitors with anticancer activity. *RSC Adv.* **2014**, 4, 2390-13.
- [125] Chen, Y.Y.; Wu, X.Q.; Tang, W.J.; Shi, J.B.; Li, J.; Liu, X.H. Novel dihydropyrazole-chromen: Design and modulates hTERT inhibition proliferation of MGC-803. *Eur J Med Chem.* **2016**, 110, 65-75.
- [126] Xiao, X.; Ni, Y.; Jia, Y.-M.; Zheng, M.; Xu, H.-F.; Xu, J.; Liao, C. Identification of human telomerase inhibitors having the core of N -acyl-4,5-dihydropyrazole with anticancer effects. *Bioorganic Med Chem Lett.* **2016**, 26, 1508-11.
- [127] Kalathiya, U.; Padariya, M.; Baginski, M. Molecular Modeling and Evaluation of Novel Dibenzopyrrole Derivatives as Telomerase Inhibitors and Potential Drug for Cancer Therapy. *IEEE/ACM Trans. Comput Biol Bioinform.* **2014**, 11, 1196-1207.
- [128] Liu, X.-H.; Liu, H.-F.; Shen, X.; Song, B.-A.; Bhadury, P.S.; Zhu, H.-L.; Liu, J.-X.; Qi, X.-B. Synthesis and molecular docking studies of novel 2-chloro-pyridine derivatives containing flavone moieties as potential antitumor agents. *Bioorganic Med Chem Lett.* **2010**, 20, 4163-67.
- [129] Zheng, Q.-Z.; Zhang, X.-M.; Xu, Y.; Cheng, K.; Jiao, Q.-C.; Zhu, H.-L. Synthesis, biological evaluation, and molecular docking studies of 2-chloropyridine derivatives possessing 1,3,4-oxadiazole moiety as potential antitumor agents. *Bioorganic Med Chem.* **2010**, 18, 7836-41.
- [130] Saraswati, A.P.; Relitti, N.; Brindisi, M.; Gemma, S.; Zisterer, D.; Butini, S.; Campiani, G. Raising the bar in anticancer therapy: recent advances in, and perspectives on, telomerase inhibitors. *Drug Discov Today.* **2019**, 24, 1370-88.
- [131] Chen, X.; Tang, W.-J.; Shi, J.B.; Liu, M.M.; Liu, X.-H. Therapeutic strategies for targeting telomerase in cancer. *Med Res Rev.* **2020**, 40, 532-85.
- [132] Wong, L.H.; Unciti-Broceta, A.; Spitzer, M.; White, R.; Tyers, M.; Harrington, L. A Yeast Chemical Genetic Screen Identifies Inhibitors of Human Telomerase. *Chem Biol.* **2013**, 20, 333-40.
- [133] Allegra, M.; Tutone, M.; Tesoriere, L.; Attanzio, A.; Culetta, G.; Almerico, A.M. Evaluation of the IKK β Binding of Indicaxanthin by Induced-Fit Docking, Binding Pose Metadynamics, and Molecular Dynamics. *Front. Pharmacol.* **2021**, 12.
- [134] Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA Methods to Estimate Ligand-Binding Affinities. *Expert Opin Drug Discov.* **2015**, 10, 449-61.
- [135] Daina, A.; Michielin, O.; Zoete, V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep.* **2017**, 7, 42717.
- [136] Jiang, Y.; Wang, Q.Q.; Liang, S.; Hu, L.-M.; Little, R.D.; Zeng, C.-C. Electrochemical Oxidative Amination of Sodium Sulfinates: Synthesis of Sulfonamides Mediated by NH₄I as a Redox Catalyst. *J Org Chem.* **2016**, 81, 4713-19.
- [137] Shi, F.; Kin Tse, M.; Zhou, S.; Pohl, M.-M.; Radnik, J.; Hübner, S.; Jähnisch, K.; Brückner, A.; Beller, M. Green and Efficient Synthesis of Sulfonamides Catalyzed by Nano-Ru/Fe₃O₄. *J Am Chem Soc.* **2009**, 131, 1775-79.
- [138] Ge, S.; Zhong, H.; Ma, X.; Zheng, Y.; Zou, Y.; Wang, F.; Wang, Y.; Hu, Y.; Li, Y.; Liu, W. Discovery of secondary sulphonamides as IDO1 inhibitors with potent antitumour effects *in vivo*. *J Enzyme Inhib Med Chem.* **2020**, 35, 1240-57.

- [139] Lai, Y.; Ge, S.; Zheng, Y.; Zou, Y.; Wang, F.; Xu, Q.; Guo, W.; Wang, Y.; Hu, Y. Benzene sulfonamide IDO1 inhibitor, its preparation method and application. China Patent 106928101A, 7 July 2017.
- [140] Molander, G.A.; Fleury-Brégeot, N.; Hiebel, M.-A. Synthesis and Cross-Coupling of Sulfonamidomethyltrifluoroborates. *Org Lett.* **2011**, 13, 1694-97.
- [141] Martínez-Asencio, A.; Ramón, D.J.; Yus, M. N-Alkylation of poor nucleophilic amines and derivatives with alcohols by a hydrogen autotransfer process catalyzed by copper(II) acetate: Scope and mechanistic considerations. *Tetrahedron.* **2011**, 67, 3140-49.
- [142] Allegra, M.; De Cicco, P.; Ercolano, G.; Attanzio, A.; Busà, R.; Cirino, G.; Tesoriere, L.; Livrea, M.A.; Ianaro, A. Indicaxanthin from *Opuntia Ficus Indica* (L. Mill) impairs melanoma cell proliferation, invasiveness, and tumor progression. *Phytomedicine.* **2018**, 50, 19-24.
- [143] Doobary, S.; Sedikides, A.T.; Caldora, H.P.; Poole, D.L.; Lennox, A.J.J. Electrochemical Vicinal Difluorination of Alkenes: Scalable and Amenable to Electron-Rich Substrates. *Angew Chem Int.* **2019**, 59, 1155-60.
- [144] Yang, Y.; Meng, X.; Zhu, B.; Jia, Y.; Cao, X.; Huang, S. A Micellar Catalysis Strategy for Amidation of Alkynyl Bromides: Synthesis of Ynamides in Water. *Eur J Org Chem.* **2019**, 1166-69.
- [145] Huang, M.; Li, Y.; Liu, J.; Lan, X.B.; Liu, Y.; Zhao, C.; Ke, Z. A bifunctional strategy for N-heterocyclic carbene-stabilized iridium complex-catalyzed: N-alkylation of amines with alcohols in aqueous media. *Green Chem.* **2019**, 21, 219-24.
- [146] Li, Q.Q.; Xiao, Z.F.; Yao, C.Z.; Zheng, H.X.; Kang, Y.B. Direct Alkylation of Amines with Alcohols Catalyzed by Base. *Org Lett.* **2015**, 17, 5328-31.
- [147] Deng, X.; Mani, N.S. A facile, environmentally benign sulfonamide synthesis in water. *Green Chem.* **2006**, 8, 835-38.
- [148] Massah, A.; Kazemi, F.; Azadi, D.; Farzaneh, S.; Aliyan, H.; Naghash, H.; Momeni, A. A Mild and Chemoselective Solvent-Free Method for the Synthesis of N-Aryl and N-Alkylsulfonamides. *Lett Org Chem.* **2006**, 3, 235-41.
- [149] Raymond, E.; Sun, D.; Izbicka, E.; Mangold, G.; Silvas, E.; Windle, B.; Sharma, S.; Soda, H.; Laurence, R.; Davidson, K.; et al. A human breast cancer model for the study of telomerase inhibitors based on a new biotinylated-primer extension assay. *Br. J. Cancer.* **1999**, 80, 1332-41.
- [150] Akiyama, M.; Yamada, O.; Kanda, N.; Akita, S.; Kawano, T.; Ohno, T.; Mizoguchi, H.; Eto, Y.; Anderson, K.C.; Yamada, H. Telomerase overexpression in K562 leukemia cells protects against apoptosis by serum deprivation and double-stranded DNA break inducing agents, but not against DNA synthesis inhibitors. *Cancer Lett.* **2002**, 178, 187-97.
- [151] Badisa, R.B.; Darling-Reed, S.F.; Joseph, P.; Cooperwood, J.S.; Latinwo, L.M.; Goodman, C.B. Selective cytotoxic activities of two novel synthetic drugs on human breast carcinoma MCF-7 cells. *Anticancer Res.* **2009**, 29, 2993-96.
- [152] Le, M.-T.; Mai, T.T.; Huynh, P.N.H.; Tran, T.-D.; Thai, K.-M.; Nguyen, Q.-T. Structure-based discovery of interleukin-33 inhibitors: A pharmacophore modelling, molecular docking, and molecular dynamics simulation approach. *SAR QSAR Environ Res.* **2020**, 31, 883-904.
- [153] Azam, F.; Madi, A.M.; Ali, H.I. Molecular Docking and Prediction of Pharmacokinetic Properties of Dual Mechanism Drugs that Block MAO-B and Adenosine A2A Receptors for the Treatment of Parkinson's Disease. *J Young Pharm.* **2012**, 4, 184-192.
- [154] Zhao, Y.H.; Abraham, M.H.; Le, J.; Hersey, A.; Luscombe, C.N.; Beck, G.; Sherborne, B.; Cooper, I. Rate-limited steps of human oral absorption and QSAR studies. *Pharm Res.* **2002**, 19, 1446-57.
- [155] Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev.* **1997**, 23, 3-25.

- [156] C.J. Krusemark, Drug design: structure- and ligand-based approaches. Edited by Kenneth M. Merz Jr., Dagmar Ringe, and Charles H. Reynolds. Cambridge and New York: Cambridge University Press. 274 p.; ISBN: 978-0-521-88723-6. 2010, *Q Rev Biol.* **2012**, doi:10.1086/665453.
- [157] Acharya, C.; Coop, A.; Polli, J.E.; MacKerell, A.D. Recent advances in ligand-based drug design: relevance and utility of the conformationally sampled pharmacophore approach. *Curr Comput Aided-Drug Des.* **2010**, 7, 10-22.
- [158] Pasha, F.A.; Muddassar, M.; Neaz, M.M.; Cho, S.J. Pharmacophore and docking-based combined in-silico study of KDR inhibitors. *J Mol Graph Model.* **2009**, 1, 54-61.
- [159] Hu, B.; Lill, M.A. PharmDock: a pharmacophore-based docking program, *J Cheminform.* **2014**, 6, 14.
- [160] Almerico, A.M.; Tutone, M.; Ippolito, M.; Lauria, A. Molecular modelling and QSAR in the discovery of HIV-1 integrase inhibitors. *Curr Comput. -Aided Drug Des.* 2007, 3, 214-33.
- [161] Steindl, T.M.; Crump, C.E.; Hayden, F.G.; Langer, T. Pharmacophore modeling, docking, and principal component analysis based clustering: combined computer-assisted approaches to identify new inhibitors of the human rhinovirus coat protein. *Int J Med Chem.* **2005**, 48, 6250-60.
- [162] König, I.; Zarrine-Afsar, A.; Aznauryan, M.; Soranno, A.; Wunderlich, B.; Dingfelder, F.; Stüber, J. C.; Plückthun, A.; Nettels, D.; Schuler, B. Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells. *Nat Methods.* **2015** 12, 773-79.
- [163] Torbeev, V. Y.; Raghuraman, H.; Hamelberg, D.; Tonelli, M.; Westler, W. M.; Perozo, E.; Kent, S. B. Protein conformational dynamics in the mechanism of HIV-1 protease catalysis. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, 108, 20982-87.
- [164] Mirjalili, V.; Feig, M. Protein structure refinement through structure selection and averaging from molecular dynamics ensembles. *J. Chem. Theory Comput.* **2013**, 9, 1294-1303.
- [165] Schaller, D.; Šribar, D.; Noonan, T.; Deng, L.; Nguyen, T.N.; Pach, S.; Machalz, D.; Bermudez, M.; Wolber, G. Next generation 3D pharmacophore modeling, Wiley Interdiscip. *Rev Comput Mol Sci.* **2020**, e1468.
- [166] Bottegoni, G.; Rocchia, W.; Rueda, M.; Abagyan, R.; Cavalli, A. Systematic exploitation of multiple receptor conformations for virtual ligand screening. *PLoS ONE.* **2011**, 6, e18845.
- [167] Spyrikis, F.; Benedetti, P.; Decherchi, S.; Rocchia, W.; Cavalli, A.; Alcaro, S.; Ortuso, F.; Baroni, M.; Cruciani, G. A pipeline to enhance ligand virtual screening: integrating molecular dynamics and fingerprints for ligand and proteins. *J Chem Inf Model.* **2015**, 55, 2256-74.
- [168] Wieder, M.; Perricone, U.; Boresch, S.; Seidel, T.; Langer, T. Evaluating the stability of pharmacophore features using molecular dynamics simulations. *Biochem Biophys Res Commun.* **2016**, 470, 685-89.
- [170] Schuetz, D.A.; Seidel, T.; Garon, A.; Martini, R.; Körbel, M.; Ecker, G.F.; Langer, T. GRAIL: GRids of phArmacophore Interaction fieLds, *J Chem Theory Comput.* **2018**, 14, 4958-70.
- [171] Sánchez-Martínez, C.; Gelbert, L.M.; Lallena, M.J.; De Dios, A. Cyclin dependent kinase (CDK) inhibitors as anticancer drugs. *Bioorganic Med Chem Lett.* **2015**, 25, 3420-35.
- [172] Tutone, M.; Almerico, A.M. Recent advances on CDK inhibitors: an insight by means of *in silico* methods. *Eur J Med Chem.* **2017**, 15, 300-15.
- [173] Bose, P.; Simmons, G.L.; Grant, S. Cyclin-dependent kinase inhibitor therapy for hematologic malignancies. *Expert Opin Investig Drugs.* **2013**, 22, 723-38
- [174] Zou, H.Z.; Xie, S.Y.; Yang, J.J.; Chen, J.X.; Ren, Y.Q.; Wei, I. Towards more accurate pharmacophore modeling: multicomplex-based comprehensive pharmacophore map and most-frequent-feature pharmacophore model of CDK2. *J Mol Graph Model.* **2008**, 27, 430-38.
- [175] Moorhouse, M.; Barry, P. The Protein Databank, in: Bioinforma. Biocomput, *Wiley Perl.* **2005**,

- [176] Tutone, M.; Culetta, G.; Livecchi, L.; Almerico, A.M. A definitive pharmacophore modelling study on CDK2 ATP pocket binders: tracing the path of new virtual high-throughput screenings. *Curr Drug Discov Technol.* **2019**, *17*, 740-47.
- [177] Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J Mol Graph.* **1995**, *14*, 33-38.
- [178] Perricone, U.; Wieder, M.; Seidel, T.; Langer, T.; Padova, A.; Almerico, A.M.; Tutone, M. A molecular dynamics-shared pharmacophore approach to boost early-enrichment virtual screening: a case study on peroxisome proliferator-activated receptor α , *ChemMedChem.* **2017**, *12* 1399–1407.
- [179] Fawcett, T. An introduction to ROC analysis, *Pattern Recognit. Lett.* **2006**, *27*, 861-74.
- [180] Truchon, J.F.; Bayly, C.I. Evaluating virtual screening methods: good and bad metrics for the “early recognition” problem. *J Chem Inf Model.* **2007**, *47*, 488-508.
- [181] Allegra, M.; Tutone, M. Tesoriere, L.; Almerico, A.M.; Culetta, G.; Livrea, M.A.; Attanzio, A. Indicaxanthin, a Multi-Target Natural Compound from *Opuntia Ficus-Indica* Fruit: From its Poly-Pharmacological Effects to Biochemical Mechanisms and Molecular Modelling Studies. *Eur J Med Chem.* **2019**, *179*, 753-64.
- [182] Allegra, M.; De Cicco, P.; Ercolano, G.; Attanzio, A.; Busà, R.; Cirino, G. Indicaxanthin from *Opuntia Ficus Indica* (L. Mill) Impairs Melanoma Cell Proliferation, Invasiveness, and Tumor Progression. *Phytomedicine.* **2018**, *50*, 19-24.
- [183] Taniguchi, K.; Karin, M. NF- κ B, Inflammation, Immunity and Cancer: Coming of Age. *Nat Rev Immunol.* **2018**, *18*, 309-24.
- [184] Zhang, Q.; Lenardo, D.; Baltimore, D. 30 Years of NF- κ B: A Blossoming of Relevance to Human Pathobiology. *Cell.* **2017**, *168*, 37–57.
- [185] Staudt, L. M. Oncogenic Activation of NF-kappaB. *Perspect Biol.* **2010**, *2*, a000109.
- [186] Ben-Neriah, Y.; Karin, M. Inflammation Meets Cancer, with NF- κ B as the Matchmaker. *Nat Immunol.* **2011**, *12*, 715-23.
- [187] Van Herck, Y.; Antoranz, A.; Andhari, M. D.; Milli, G.; Bechter, O.; De Smet, F. Multiplexed Immunohistochemistry and Digital Pathology as the Foundation for Next-Generation Pathology in Melanoma: Methodological Comparison and Future Clinical Applications. *Front Oncol.* **2021**, *11*, 636681.
- [188] Yang, J.; Richmond, A. Constitutive IkappaB Kinase Activity Correlates with Nuclear Factor-kappaB Activation in Human Melanoma Cells. *Cancer Res.* **2001**, *61*, 4901-09.
- [189] Dhawan, P.; Richmond, A. A Novel NF-Kappa B-Inducing Kinase- MAPK Signaling Pathway Up-Regulates NF-Kappa B Activity in Melanoma Cells. *J Biol Chem.* **2002**, *277*, 7920-28.
- [190] Amiri, K. I.; Richmond, A. Role of Nuclear Factor-Kappa B in Melanoma. *Cancer Metastasis Rev.* **2005**, *24*, 301-13.
- [191] Amschler, K.; Schön, M. P.; Pletz, N.; Wallbrecht, K.; Erpenbeck, L.; Schön, M. NF-kappaB Inhibition through Proteasome Inhibition or IKKbeta Blockade Increases the Susceptibility of Melanoma Cells to Cytostatic Treatment through Distinct Pathways. *J Invest Dermatol.* **2010**, *130*, 1073-86.
- [192] Prescott, J.; Cook, S. Targeting IKK β in Cancer: Challenges and Opportunities for the Therapeutic Utilisation of IKK β Inhibitors. *Cells.* **2018**, *7*, 115.
- [193] Liu, H.; Liang, H.; Meng, H.; Deng, X.; Zhang, X.; Lai, L. A Novel Allosteric Inhibitor that Prevents IKK β Activation. *Medchemcomm.* **2018**, *9*, 239–243.
- [194] Prescott, J.; Cook, S. Targeting IKK β in Cancer: Challenges and Opportunities for the Therapeutic Utilisation of IKK β Inhibitors. *Cells.* **2018**, *7*, 115.
- [195] Moarbess, G.; Guichou, J. F.; Paniagua-Gayraud, S.; Chouchou, A.; Marcadet, O.; Leroy, F. New IKK Inhibitors: Synthesis of New Imidazo[1,2-A] quinoxaline Derivatives Using Microwave Assistance and Biological Evaluation as IKK Inhibitors. *Eur J Med Chem.* **2016**, *115*, 268-74.

- [196] Guedes, I. A.; Freitas, R. H.; Cordeiro, N. M.; do Nascimento, T. S.; Valerio, T. S.; Fernandes, P. D. LASSBio-1829 Hydrochloride: Development of a New Orally Active N-Acylhydrazone IKK2 Inhibitor with Anti-inflammatory Properties. *Chem Med Chem*. **2016**, *11*, 234-44.
- [197] Rothan, H.A.; Byrareddy, S.N. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *J Autoimmun*. **2020**, *109*, 102433.
- [198] Yu, R.; Chen, L.; Lan, R.; Shen, R.; Li, P. Computational screening of antagonists against the SARS-CoV-2 (COVID-19) coronavirus by molecular docking. *Int J Antimicrob Agents*. **2020**, *2*, 3–8.
- [199] Wu, C.; Liu, Y.; Yang, Y.; Zhang, P.; Zhong, W.; Wang, Y.; Wang, Q.; Xu, Y.; Li, M.; Li, X. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharm Sin B*. **2020**, *10*, 766-88.
- [200] Ciliberto, G.; Cardone, L. Boosting the arsenal against COVID-19 through computational drug repurposing. *Drug Discov Today*. **2020**, *26*, 946-94.
- [201] Pushpakom, S.; Iorio, F.; Eyers, P.A.; Escott, K.J.; Hopper, S.; Wells, A.; Doig, A.; Guilliams, T.; Latimer, J.; McNamee, C. Drug repurposing: Progress, challenges and recommendations. *Nat Rev Drug Discov*. **2018**, *18*, 41-58.
- [202] Oprea, T.I.; Mestres, J. Drug repurposing: Far beyond new targets for old drugs. *AAPS J*. **2012**, *14*, 759-63.
- [203] Vanhaelen, Q.; Mamoshina, P.; Aliper, A.M.; Artemov, A.; Lezhnina, K.; Ozerov, I.; Labat, I.; Zhavoronkov, A. Design of efficient computational workflows for *in silico* drug repurposing. *Drug Discov Today*. **2017**, *22*, 210-22
- [204] March-Vila, E.; Pinzi, L.; Sturm, N.; Tinivella, A.; Engkvist, O.; Chen, H.; Rastelli, G. On the integration of *in silico* drug design methods for drug repurposing. *Front Pharmacol*. **2017**, *8*, 298.
- [205] Liu, Z.; Fang, H.; Reagan, K.; Xu, X.; Mendrick, D.L.; Slikker, W.; Tong, W. *In silico* drug repositioning-what we need to know. *Drug Discov. Today*. **2013**, *18*, 110–115.
- [206] Oprea, T.I.; Bauman, J.E.; Bologna, C.G.; Buranda, T.; Chigaev, A.; Edwards, B.S.; Jarvik, J.W.; Gresham, H.D.; Haynes, M.K.; Hjelle, B. Drug repurposing from an academic perspective. *Drug Discov. Today Ther. Strateg*. **2011**, *8*, 61–69.
- [207] Lauria, A.; Tutone, M.; Barone, G.; Almerico, A.M. Multivariate analysis in the identification of biological targets for designed molecular structures: The BIOTA protocol. *Eur. J. Med. Chem*. **2014**, *75*, 106–110.
- [208] Corsello, S.M.; Bittker, J.A.; Liu, Z.; Gould, J.; McCarren, P.; Hirschman, J.E.; Johnston, S.E.; Vrcic, A.; Wong, B.; Khan, M. The Drug Repurposing Hub: A next-generation drug library and information resource. *Nat. Med*. **2017**, *23*, 405-8.
- [209] Farha, M.A.; Brown, E.D. Drug repurposing for antimicrobial discovery. *Nat. Microbiol*. **2019**, *4*, 565–577.
- [210] Sleire, L.; Førde-Tislevoll, H.E.; Netland, I.A.; Leiss, L.; Skeie, B.S.; Enger, P.Ø. Drug repurposing in cancer. *Pharmacol Res*. **2017**, *124*, 74–91.
- [211] Cha, Y.; Erez, T.; Reynolds, I.J.; Kumar, D.; Ross, J.; Koytiger, G.; Kusko, R.; Zeskind, B.; Risso, S.; Kagan, E.; et al. Drug repurposing from the perspective of pharmaceutical companies. *Br. J. Pharmacol*. **2018**, *175*, 168–180.
- [212] Tutone, M.; Perricone, U.; Almerico, A.M. Conf-VLKA: A structure-based revisit of the Virtual Lock-and-key Approach. *J. Mol. Graph. Model*. **2017**, *71*, 50–57.
- [213] Tutone, M.; Almerico, A.M. The *In Silico* Fischer Lock-and-Key Model: The Combined Use of Molecular Descriptors and Docking Poses for the Repurposing of Old Drugs. Targeting Enzymes for Pharmaceutical Development. *Methods Mol. Biol*. **2020**, *2089*, 29–39.
- [214] Gao, J.; Zhang, L.; Liu, X.; Li, F.; Ma, R.; Zhu, Z.; Zhang, J.; Wu, J.; Shi, Y.; Pan, Y. Repurposing Low-Molecular-Weight Drugs Against the Main Protease of Severe Acute Respiratory Syndrome Coronavirus 2. *J Phys Chem Lett*. **2020**, *11*, 7267-72.

- [215] Meyer-Almes, F.J. Repurposing approved drugs as potential inhibitors of 3CL-protease of SARS-CoV-2: Virtual screening and structure based drug design. *Comput Biol Chem.* **2020**, 88.
- [216] Touret, F.; Gilles, M.; Barral, K.; Nougairède, A.; van Helden, J.; Decroly, E.; de Lamballerie, X.; Coutard, B. *In vitro* screening of a FDA approved chemical library reveals potential inhibitors of SARS-CoV-2 replication. *Sci. Rep.* **2020**, 10, 13093.
- [217] Shyr, Z.A.; Gorshkov, K.; Chen, C.Z.; Zheng, W. Drug discovery strategies for SARS-CoV-2. *J Pharmacol Exp Ther.* **2020**, 374.
- [218] Cavasotto, C.; Di Filippo, J. *In silico* Drug Repurposing for COVID-19: Targeting SARS-CoV-2 Proteins through Docking and Consensus Ranking. *Mol Inform.* **2020**, 40:e2000115.
- [219] Wang, J. Fast Identification of Possible Drug Treatment of Coronavirus Disease-19 (COVID-19) through Computational Drug Repurposing Study. *J Chem Inf Model.* **2020**, 60, 3277-86.
- [220] Ferraz, W.R.; Gomes, R.A.; S Novaes, A.L.; Goulart Trossini, G.H. Ligand and structure-based virtual screening applied to the SARS-CoV-2 main protease: An *in silico* repurposing study. *Future Med. Chem.* **2020**, 12:1815-1828.
- [221] Zhou, Y.; Hou, Y.; Shen, J.; Huang, Y.; Martin, W.; Cheng, F. Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discov.* **2020**, 6, 14.
- [222] Mirza, M.U.; Froeyen, M. Structural elucidation of SARS-CoV-2 vital proteins: Computational methods reveal potential drug candidates against main protease, Nsp12 polymerase and Nsp13 helicase. *J Pharm Anal.* **2020**, 1:320-328.
- [223] Harrison, C. Coronavirus puts drug repurposing on the fast track. *Nat. Biotechnol.* **2020**, 38, 379–381.
- [224] Gordon, D.E.; Jang, G.M.; Bouhaddou, M.; Xu, J.; Obernier, K.; White, K.M.; O’Meara, M.J.; Rezelj, V.V.; Guo, J.Z.; Swaney, D.L.; et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature.* **2020**, 583, 459-68.
- [225] Battisti, V.; Wieder, O.; Garon, A.; Seidel, T.; Urban, E.; Langer, T. A Computational Approach to Identify Potential Novel Inhibitors against the Coronavirus SARS-CoV-2. *Mol. Inform.* **2020**, 39:e2000090.
- [226] Wishart, D.S.; Knox, C.; Guo, A.C.; Shrivastava, S.; Hassanali, M.; Stothard, P.; Chang, Z.; Woolsey, J. DrugBank: A comprehensive resource for *in silico* drug discovery and exploration. *Nucleic Acids Res.* **2006**, 34, D668–D672.
- [227] Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; de Beer, T.A.P.; Rempfer, C.; Bordoli, L. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, 46, W296–W303.
- [228] Schaller, D.; Šribar, D.; Noonan, T.; Lihua Deng, L.; Nguyen, T.N.; Pach, S.; Machalz, D.; Bermudez, M.; Wolber, G. Next generation 3D pharmacophore modeling. *WIREs Comput Mol Sci.* **2020**, 10, e1468.
- [229] Arya, R.; Das, A.; Prashar, V.; Kumar, M. Potential inhibitors against papain-like protease of novel coronavirus (SARS-CoV-2) from FDA approved drugs. *Chemrxiv. Org.* **2020**.
- [230] Gil, C.; Ginex, T.; Maestro, I.; Nozal, V.; Barrado-Gil, L.; Cuesta-Gejjo, M.A.; Urquiza, J.; Ramírez, D.; Alonso, C.; Campillo, N.E. COVID-19: Drug targets and potential treatments. *J. Med. Chem.* **2020**, 63, 12359-12386.
- [231] Xu, X.; Lou, Z.; Ma, Y.; Chen, X.; Yang, Z.; Tong, X.; Zhao, Q.; Xu, Y.; Deng, H.; Bartlam, M. Crystal Structure of the C-Terminal Cytoplasmic Domain of Non-Structural Protein 4 from Mouse Hepatitis Virus A59. *PLoS ONE*, **2009**, 4, e6217.
- [232] Sutton, G.; Fry, E.; Carter, L.; Sainsbury, S.; Walter, T.; Nettleship, J.; Berrow, N.; Owens, R.; Gilbert, R.; Davidson, A.; et al. The nsp9 Replicase Protein of SARS-Coronavirus, Structure and Functional Insights. *Structure*, **2004**, 12, 341–353.