



DEPARTMENT OF BIOTECHNOLOGY, CHEMISTRY AND PHARMACY

Ph.D. IN CHEMICAL AND PHARMACEUTICAL SCIENCES
CYCLE XXXV

EXPLOITING MOLECULAR TARGETING IN ANTICANCER
CHEMOTHERAPY: *IN SILICO* STUDIES TO IDENTIFY NOVEL MOLECULES
TARGETING KINASES IN CONVERGENT METABOLIC PATHWAYS IN
CANCER THERAPY

Supervisor

Prof. Fabrizio Manetti

Author

Valentina Marcellini

ACADEMIC YEAR 2019/2022

To my family and Domenico

Summary

In recent years, a confluence of scientific progress has allowed to identify specific targets, as well as multiple signaling routes for cancer cells, leading to more selective, more effective, and less toxic treatments. The validation of compounds was originally based on the validation of the target, despite several of the more effective drugs sometimes have effects unrelated to their hypothetical mechanism. Protein kinases have become interesting molecular targets and a large part of the research has targeted drugs capable of inhibiting the pathogenic kinases. To date, clinical studies have confirmed the important role of kinase inhibitors in cancer therapy.

The first part of this thesis essentially focused on the research of new selective inhibitors of serum and glucocorticoid protein kinase 1 (SGK1). Its role in human tumors has been widely investigated, identifying SGK1 as a key target in cancer progression, having regard to its ability to regulate processes such as cell cycle, invasion, migration, cellular apoptosis, autophagy and others. To date, there are no commercially available drugs against SGK1, and the available inhibitors still require further studies. In this project, starting from the crystal structures of SGK1, through a combined structure- and ligand-based approaches, a pharmacophore model was generated, using Phase software (Schrodinger suite). The hypotheses generated, based on docking poses of known inhibitors MMG (4-(5-phenyl-1*H*-pyrrolo[2,3-*b*]pyridin-3-yl)benzoic acid) and GMG ([4-(5-naphthalen-2-yl-1*H*-pyrrolo[2,3-*b*]pyridin-3-yl)phenyl]acetic acid), have been pruned according to the presence of some essential features for the interaction with the receptor and then were classified by survival score. Then, the selected hypothesis (thereafter reported as ADNRR_1, where A: hydrogen bond acceptor, D: hydrogen bond donor, N: negatively ionizable group, and R: aromatic ring) has been refined by adding excluded volumes. Since the validation process showed that the pharmacophore model was reliable, it was used to perform a virtual screening study. Starting from commercial databases, a multi-conformer optimized library was

generated using Phase tools, which was subsequently screened on the validated pharmacophore model. The filtered molecules were subjected to molecular properties analysis. Only compounds that met the required parameters were analyzed through docking studies and selected based on an appropriate binding mode and extra-precision (XP) Glide-score value. Four compounds were available for purchase and were sent for biological tests. The results provided two hit compounds with the activity of 0.93 μM and 11.9 μM toward SGK1.

Starting from the compound with activity of 0.93 μM , a similarity search was performed on several commercial databases, resulting in a library of 705 analogues. Molecular docking studies were performed, using the Glide SP mode (Schrodinger suite) which led to the prioritization of 5 compounds. Finally, they were evaluated by biological assays on SGK1, finding lower activity in comparison to the parent compound. However, the results obtained have allowed us to make useful observations to direct a future perspective of the study, towards the search for better compounds.

In the second section, the focus is on tyrosine kinase Src. In recent years, a large number of Src family kinase (SFK) targeted compounds have been designed and tested in several preclinical models, confirming the ability of these inhibitors to block cancer progression. Simulations were performed as support to the research study conducted by the group of Professor Sabrina Dallavalle, integrating computational information into a work based on the research of Src inhibitors. Starting from a small internal library of structurally different compounds, the best structures examined had a common indolinone core, which was used as scaffold for the next investigation. Several 3-(hetero)arylideneindolin-2-one substitutes have been designed and synthesized to identify the characteristics that determine activity. Molecular docking studies have helped to suggest a putative binding mode between compounds and the Src binding site and to direct future optimization studies.

Table of Contents

LIST OF TABLE	VII
LIST OF FIGURES	VIII
ABBREVIATIONS USED	IX
PART I	1
INTRODUCTION	2
1.1 <i>SGK1: a Serine/Threonine Protein Kinase</i>	2
1.2 <i>Currently available SGK1 inhibitors</i>	5
AIM OF THE PROJECT	10
MATERIALS AND METHODS	11
3.1 DATABASES AND ONLINE RESOURCES	11
3.1.1 <i>Protein Data Bank</i>	11
3.1.2 <i>DUD-E decoys and validation parameters</i>	12
3.2 PHARMACOPHORE MODELING.....	13
3.3 DOCKING STUDIES	14
3.4 MOLECULAR DYNAMICS SIMULATIONS.....	16
3.5 RMSD ANALYSIS	19
3.6 MM/GBSA CALCULATION	19
3.7 HYDROGEN BONDS ANALYSIS	20
RESULTS AND DISCUSSION	21
4.1 STRUCTURE-BASED AND LIGAND-BASED PHARMACOPHORE MODELING.....	21
4.2 PHARMACOPHORE MODEL VALIDATION AND SCREENING.....	24
4.3 MOLECULAR DOCKING AND VISUAL INSPECTION	27
4.4 BIOLOGICAL RESULTS	32
4.5 MOLECULAR DYNAMICS STUDY OF HIT COMPOUND 2	34
4.6 FOCUSED LIBRARY.....	36
4.7 POST DOCKING ANALYSIS OF SELECTED MOLECULES	40
4.8 BIOLOGICAL RESULTS.....	44
CONCLUSIONS	46
PART II	47
INTRODUCTION	48
6.1 TYROSINE KINASES.....	48
6.2 SRC NON-RECEPTOR TYROSINE KINASE	48
6.3 SRC SIGNALING IN CANCER BIOLOGY	50
6.4 SMALL-MOLECULE TYROSINE KINASE INHIBITORS.....	52
6.5 SRC INHIBITORS.....	53
AIM OF THE PROJECT	56
MATERIAL AND METHODS	57
8.1 LIGAND PREPARATION.....	57
8.2 ADME PREDICTION	57
8.3 DOCKING STUDIES	57
RESULTS AND DISCUSSION	59
9.1 BIOLOGICAL DATASET	59
9.2 PRELIMINARY VALIDATION PROTOCOL.....	65
9.3 DOCKING STUDIES	66
9.4 RESULTS OF ADME ANALYSIS	69

CONCLUSIONS.....71
LIST OF PUBLICATIONS.....73
REFERENCES74

List of Table

TABLE 1. KNOWN INHIBITORS OF SGK1. ADAPTED BY JANG H. ET AL. ⁶	6
TABLE 2. SELF-DOCKING RESULTS OF MMG AND GMG WITH THEIR RESPECTIVE PROTEIN STRUCTURE.	21
TABLE 3. SUMMARY OF PHASE SCORES FOR THE 20 GENERATED HYPOTHESES.	23
TABLE 4. VIRTUAL SCREENING METHOD VALIDATION PARAMETERS.	25
TABLE 5. DOCKING XP GSCORE VALUES OF FOUR HIT COMPOUNDS ON 3HDM.	28
TABLE 6. EVALUATION OF BIOLOGICAL RESULTS COMPARED TO XP GSCORE VALUES.	33
TABLE 7. MMGBSA VALUES OF REPLICAS OF 2.	35
TABLE 8. SELECTED COMPOUNDS FROM DOCKING STUDIES.	37
TABLE 9. EVALUATION OF BIOLOGICAL RESULTS.	44
TABLE 10. EVALUATION OF THE INHIBITORY ACTIVITY OF THE NOVEL INDOLINONE DERIVATIVES. ADAPTED BY PRINCIOTTO S. ET AL. ⁶⁵	61
TABLE 11. CYTOTOXICITY EVALUATION OF SELECTED COMPOUNDS ON HUMAN MCF-7 BREAST. ADAPTED BY PRINCIOTTO ET AL. ⁶⁵	65
TABLE 12. ADME PROPERTIES PREDICTED BY QIKPROP OF BEST FOUR COMPOUNDS AND NATIVE LIGAND AP23464.	70

List of Figures

FIGURE 1. STRUCTURE OF SGK1. (A) THE PROTEIN DOMAIN STRUCTURE. ADAPTED BY XIAO X. ET AL. ⁸	4
FIGURE 2. GLIDE HIERARCHICAL FILTERS.	15
FIGURE 3. PERIODIC BOUNDARY CONDITIONS (PBC).....	17
FIGURE 4. ALIGNMENT OF DOCKING POSES OF MMG (MAGENTA) AND GMG (ORANGE).....	22
FIGURE 5. A) FINAL HYPOTHESIS WAS REFINED WITH THE ADDITION OF THE EXCLUDED VOLUMES.....	24
FIGURE 6. A) HYPOTHESIS VALIDATION AND ENRICHMENT ANALYSIS. B) RECEIVER OPERATING CURVE (ROC). ...	26
FIGURE 7. MOLECULAR INTERACTIONS BETWEEN PROTEIN AND PUTATIVE HITS. THE PROTEIN-LIGAND COMPLEXES FROM THE XP DOCKING SHOWING NATIVE LIGAND MMG (MAGENTA) COMPARED TO(A) COMPOUND 1 (PURPLE); (B) COMPOUND 2 (LIGHT BLUE); (C) COMPOUND 3, (D) COMPOUND 4 (YELLOW).....	31
FIGURE 8. RMSD PLOT OF SGK1 (BLACK) AND LIGAND 2 (RED).....	36
FIGURE 9. ALIGNMENT OF DOCKING POSES OF COMPOUND 2 (PURPLE) AND COMPOUND 5 (ORANGE).....	40
FIGURE 10. ALIGNMENT OF DOCKING POSES OF COMPOUND 2 (PURPLE) AND COMPOUND 6 (GREY).....	41
FIGURE 11. ALIGNMENT OF DOCKING POSES OF COMPOUND 2 (PURPLE) AND COMPOUND 7 (YELLOW).....	42
FIGURE 12. ALIGNMENT OF DOCKING POSES OF COMPOUND 2 (PURPLE) AND COMPOUND 8 (GREEN).....	42
FIGURE 13. ALIGNMENT OF DOCKING POSES OF COMPOUND 2 (PURPLE) AND COMPOUND 9 (BLUE).....	43
FIGURE 14. STRUCTURE OF SRC KINASE IN COMPLEX WITH AMP-PNP (PDB ID: 2SRC).	49
FIGURE 15. SRC SIGNALING PATHWAYS IN CANCER. ADAPTED FROM HSU P. C. ET AL. ⁴⁹	51
FIGURE 16. STRUCTURE OF SRC INHIBITORS ALREADY APPROVED BY THE FDA OR STILL IN CLINICAL PHASE.....	54
FIGURE 17. STRUCTURE OF THE REFERENCE COMPOUND 10.	59
FIGURE 18. GENERAL SYNTHESIS BY KNOEVENAGEL CONDENSATION STARTING FROM 6-CLORO OXINDOLE AND DIFFERENT AROMATIC AND HETEROAROMATIC COMPOUNDS. ADAPTED BY PRINCIOTTO S. ET AL. ⁶⁵	60
FIGURE 19. RE-DOCKING OF NATIVE LIGAND AP23464 IN COMPLEX WITH SRC.	65
FIGURE 20. COMPARISON OF AP23464 (GREEN) AND 13 (LIGHT PINK) BINDING MODES.....	67
FIGURE 21. COMPARISON OF AP23464 (GREEN) AND 30 (PURPLE) BINDING MODES.	68
FIGURE 22. COMPARISON OF AP23464 (GREEN) AND COMPOUND 14 (ORANGE) BINDING MODES.	68

Abbreviations used

AGC	cAMP-dependent Protein Kinase
ALL Ph+	Philadelphia chromosome-positive acute lymphoblastic leukemia
AMP-PNP	Adenylyl-imidodiphosphate
AUAC	Area under the accumulation curve
AUC	Area under the curve
BEDROC	Boltzmann-enhanced discrimination ROC
CAMK	Calcium/Calmodulin regulated Kinases
CK1	Cell Kinase I
CMGC	kinase family that is named after the initials of its subfamily members including: Cyclin-dependent kinase (CDK), Mitogen-activated protein kinase (MAPK), Glycogen synthase kinase (GSK) and CDC-like kinase (CLK)
CSK	C-terminal Src tyrosine kinase
DFG	Asp-Phe-Gly motif
DPI	Diffraction-component precision index
DSKs	Double specificity kinases
DUD-E	Database of Useful (Docking) Decoys – Enhanced
EF	Enrichment factor
EGFR	Epidermal growth factor receptor
ENaC	Epithelial sodium channel
EPH	Erythropoietin-producing hepatocellular carcinoma receptor
FAK	Focal adhesion kinase

FDA	Food and Drug Administration
FGFR	Fibroblast growth factor receptor
GAFF	Generalized Amber Force Field
GBM	Glioblastoma multiforme
GMG	[4-(5-naphthalen-2-yl-1H-pyrrolo[2,3-b]pyridin-3-yl)phenyl]acetic acid
HBA	Hydrogen Bond Acceptor
HBD	Hydrogen Bond Donor
HER2	Human epidermal growth factor receptor 2
HGFR	c-Met/hepatocyte growth factor receptor
IGF-1R	Insulin-like growth factor receptor-1
JAK	Janus kinases
KD	Kinase domain
Kv1.3	K ⁺ voltage-dependent channel
CML Ph ⁺	Philadelphia chromosome positive chronic myeloid leukemia
MCF-7 cell	Michigan Cancer Foundation-7 cell are primarily used as an <i>in vitro</i> model to study breast cancer biology.
MD	Molecular Dynamics
MM/GBSA	Molecular Mechanics Generalized Born Surface Area
MMG	4-(5-phenyl-1H-pyrrolo[2,3-b]pyridin-3-yl)benzoic acid
mTOR	Mammalian target of rapamycin
NB	Neuroblastoma
nRTKs	Non-receptor tyrosine kinases
PBC	Periodic Boundary Conditions
PDB	Protein Data Bank

PDGFR	Platelet-derived growth factor receptors
PI3K	Phosphatidylinositol 3-kinases
PKA	cAMP-dependent protein kinase A
PKB	Protein kinase B
PKC	Protein kinase C
PKs	Protein kinases
PMEMD	Particle Mesh Ewald Molecular Dynamics
RAS/MAPK	Mitogen-activated protein kinases
Rho A	Ras homolog family member A
RIE	Robust initial enhancement
RMSD	Root Mean Square Deviation
ROC	Receiver Operating Characteristics
ROMK1	Renal outer medullary potassium channel
RTKs	Receptor tyrosine kinases
SANDER	Simulated Annealing with NMR-Derived Energy Restraints
SFKs	Src family kinases
SGK1	Serum and glucocorticoid protein kinase 1
SGK2	Serum and glucocorticoid protein kinase 2
SGK3	Serum and glucocorticoid protein kinase 3
SGLT1	Sodium/glucose cotransporter
SH2	Src Homology 2 domain
SH3	Src Homology 3 domain
SH4	Src Homology 4 domain

Src	Non-receptor tyrosine kinases, as it is short for sarcoma.
Srm	Src-related kinase lacking C-terminal regulatory tyrosine and N-terminal myristylation sites
STAT3	Signal transducer and activator of transcription 3
STE	Sterile Ser/Thr Kinases
STKs	Ser/Thr kinases
TKL-kinases	Tyrosine kinase-like kinases
TKR	Tyrosine kinase Receptor
TKs	Tyrosine kinases
UD	Unique domain
VEGFR	Vascular endothelial growth factor receptor
VMD	Visual Molecular Dynamics software
v-Src	Avian sarcoma viral oncogene or viral-src

Part I

***In silico* studies for the identification of new SGK1
enzyme inhibitors.**

Chapter 1

Introduction

Among several post-translational modifications, phosphorylation is one of the most widely investigated process, affecting several important cellular functions such as growth, differentiation, apoptosis, and cell signaling. Protein kinases (PKs) are phosphorylation enzymes that promote the shift of phosphate group from ATP to protein substrates. Therefore, PKs have a pivotal role in regulating activation or inhibition of the activity of specific proteins in a signaling pathway. Depending on the nature of the -OH group that is phosphorylated, it is possible to divide kinases into: tyrosine kinases (TKs), serine/threonine kinases (STKs) and double specificity kinases (DSKs), or group them according to their catalytic domain, into seven large families: AGC, CAMK, CK1, CMGC, STE, TKs and TKL-kinases.

Recent progress in understanding the molecular mechanisms underlying cancer cell signaling has clearly defined the importance of kinases in the carcinogenesis and metastasis of several form of cancer. Due to kinase involvement in cell survival, proliferation and migration, when active or constitutively overexpressed, they are connected with process of oncogenesis. As a result, therapeutic strategies are now directed towards the research and design of new kinase inhibitors to treat human tumors.¹

1.1 SGK1: a Serine/Threonine Protein Kinase

Serum and glucocorticoid kinase 1 (SGK1) is a member of the AGC protein kinases family, to which also belong cAMP-dependent protein kinase A (PKA), protein kinase B (PKB or AKT) and protein kinase C (PKC). Comparative primary sequence studies performed

between SGK1 and other kinases of the AGC family, showed an evident degree of homology of the catalytic domain of the kinases, with a range between 45 and 60%. It occurs in three isoforms in mammals: SGK1, SGK2 and SGK3. While SGK1 and SGK3 isoforms While SGK1 and SGK3 isoforms have ubiquitary expression, SGK2 is confined to the brain, liver, kidneys and pancreas, as reported by Tessier et al.^{2,3} At the transcriptional level, a large number of hormonal and non-hormonal stimuli finely regulates SGK1. Among them there are serum, glucocorticoids and mineralocorticoids, however, its activity can also be modulated by other kinases. It has been observed that SGK1 takes part in the regulation of multiple ion channels, such as the epithelial sodium independent voltage channel ENaC, the K⁺ voltage-dependent channel (Kv1.3), and the renal outer medullary K⁺ channel (ROMK1), as well as the sodium/glucose cotransporter SGLT1 and the glutamine transporter SN1.

SGK1 is now a well-known therapeutic target because it is involved in different pathological disorders such as ischemia, diabetes, or degenerative brain diseases, such as Alzheimer's or Parkinson's disease. Like all kinases, SGK1 has a classical bilobal structure including an N-terminal domain and a C-terminal domain, which contribute to the control of the catalysis and regulation process.⁴ The two structures are connected through a hinge region, representing the central portion of the catalytic domain. The N lobe consists mainly of β -antiparallel strands, three of which contain a highly conserved GXGXXGXV glycine sequence, which contributes in many ways to the protein kinase function. Instead, lobe C is mainly made up of α -helices and contains an important activation segment or better known as activation loop, represented by the Asp-Phe-Gly (DFG) motif, which supports the binding of magnesium ions to ATP by chelating them, and regulating the conversion between the due conformations of the enzyme.^{5,6}

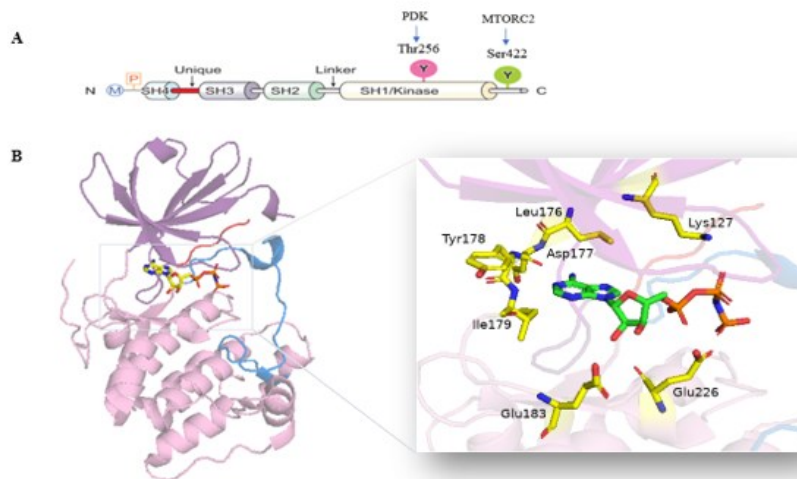


Figure 1. Structure of SGK1. (A) The protein domain structure. Adapted by Xiao X. et al.⁸
 (B) Three-dimensional structure of the kinase domain (left) and the details of ATP-binding site (right),
 in complex with AMP-PNP ligand.

According to Zhao et al⁷, the inactive state of the SGK1 kinase domain (KD) has a high degree of structural homology with the PKA. However, SGK1 is characterized by some significant differences compared to kinases of the AGC family. Indeed, in kinases such as PKA, the α C helix is a structural element that plays a very important role, allowing the interaction of the protein with ATP β phosphate.⁸ However, the inactive form of SGK1 lacks the α C helix and assumes an unusual rearrangement compared to other kinases. In addition, there is another significant difference in the conformation of the SGK1 activation loop, characterized by an extended β -strand conformation, which replace α C helix of the other kinases. Despite this, several molecular dynamics studies have shown that, instead, the active form of SGK1 is very similar to other kinases, also presenting the α C helix. However,

the residues that make up the α C helix are less than those reported in the structure of PKA. The hinge region connecting the two lobes of a protein kinase is a highly preserved structural element. As shown in the **Figure 1**, the 6-amino and 1-imido groups of the co-crystallized ATP analog 5-adenyl-imidodiphosphate ligand (AMP-PNP) (PDB ID: 2R5T) form direct hydrogen bonds with main chain of Asp177 and Ile179, by respective carbonyl and amide groups, similar to what can be observed between adenosine of ATP analogs and the hinge region of other kinases.^{5,7}

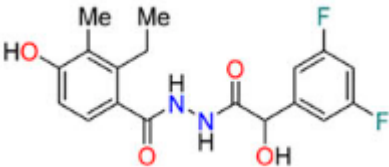
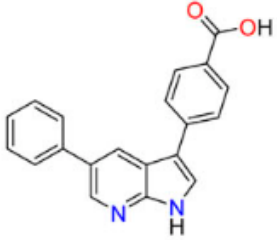
The Lys127 side chain connects with the β -phosphate of AMP-PNP, unlike what occurs in other kinases such as PKA, where the catalytic lysine corresponding to the Lys72, usually interacts with the α C helix Glu91, which in inactive SGK1 is differently reorganized. In addition, there is also a hydrogen interaction between the Glu183 side chain and one of the hydroxyl groups of ribose. Furthermore, there is an interaction between the oxygen atom between α and β the phosphates of AMP-PNP and the side chain of Glu226. The latter residue also interacts through Mg²⁺ ions and Asn227 side chain, with the nitrogen between β and γ phosphates, as reported again by Zhao et Al⁷. During the interaction with AMP-PNP, the glycine-rich cycle adopts a closed conformation. The most important difference with other kinases is in the first part of the activation loop and the DFG motif. From that point on, the peptide chain is closer to the β strand than other kinases, and instead of having the conformation of α C helix, it forms an antiparallel β sheet.⁷

1.2 Currently available SGK1 inhibitors

In recent decades, SGK1 has been identified and described as a major oncogene. Its dysregulation seems to play an essential role in different kind of malignancies, such as brain tumors like glioblastoma and neuroblastoma, or others like breast or prostate cancer. Growing scientific studies has confirmed the anticancer potential of SGK1, given its involvement in the regulation of autophagy, apoptosis and cell cycle, essential functions in many aspect of cancer development and progression.⁹ Although so far the signaling

pathways associated with phosphatidylinositol 3-kinases (PI3K) and mammalian target of rapamycin (mTOR) have been the most widely used as cancer treatment targets, recent studies showed SGK1 as an alternative and independent effector downstream of the PI3K signaling route. So far, only a small number of selective SGK1 inhibitors have been discovered, which show their activity competing with ATP for the active site (**Table 1**).¹⁰ Among these, the N'-homobenzoyl benzohydrazide analog, also known as EMD638683, has been shown to be a selective inhibitor of SGK1 in an assay involving over 69 kinases. It was characterized for the first time for its antihypertensive activity by Ackermann et al (2011).¹¹

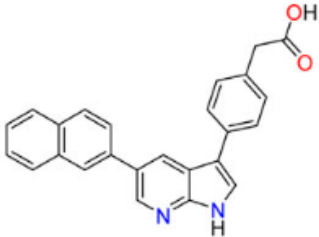
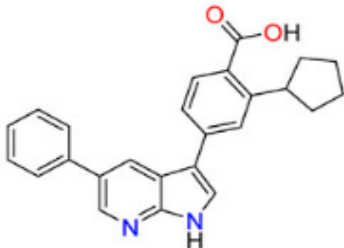

Table 1. Known inhibitors of SGK1. Adapted by Jang H. et al.⁶

Compound code	Structure	SGK1 inhibition	Target disease tested
EMD636883		15% residual SGK1 activity at 1 μ M	Inflammation/ Hypertension
MMG		IC ₅₀ = 40 nM (Fluorescent polarization assay)	N/A

(Continued on following page)

1-Introduction

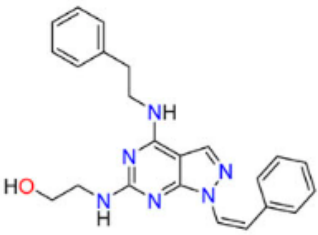
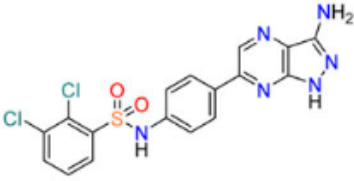
Table 1. (Continued) Known inhibitors of SGK1. Adapted by Jang H. et al.⁶

Compound code	Structure	SGK1 inhibition	Target disease tested
GMG		IC ₅₀ = 63 nM (Fluorescent polarization assay)	N/A
GSK650394		IC ₅₀ = 62 nM (<i>in vitro</i> activity-based scintillation proximity test)	Prostate cancer
QGY-5-114-A		N/A	Colorectal cancer

(Continued on following page)

1-Introduction

Table 1. (Continued) Known inhibitors of SGK1. Adapted by Jang H. et al.⁶

Compound code	Structure	SGK1 inhibition	Target disease tested
SI113		0 residual activity at 12,5 μM	Cancer, including glioblastoma multiforme
14g		$\text{IC}_{50} = 3 \text{ nM}$ (10 μM ATP) $\text{IC}_{50} = 442 \text{ nM}$ (500 μM ATP)	Alpelisib- resistant breast cancer

As reported by the authors, *in vitro* tests showed EMD638683 was found to inhibit SGK1 with an IC_{50} of 3 μM .¹¹ Instead, the azaindole compounds MMG (4-(5-phenyl-1*H*-pyrrolo[2,3-*b*]pyridin-3-yl)benzoic acid) (PDB: 3HDM) and GMG ([4-(5-naphthalen-2-yl-1*H*-pyrrolo[2,3-*b*]pyridin-3-yl)phenyl]acetic acid) (PDB: 3HDN) were co-crystallized with SGK1, providing significant information regarding interaction with the active site. Using azaindole nucleus, they establish two essential hydrogen bond interactions with Asp177 and Ile179 of the hinge region, mimicking ATP adenosine. A further important residue for the interaction with ATP β -phosphate is Lys127, with which both inhibitors interact.¹²

Another SGK1 inhibitor reported by Merck, GSK650394, was predicted to interact with the hinge region residues through a para phenol group and with catalytic lysine through a carbonyl group, respectively. GSK650394 inhibits the enzymatic activity of SGK1 with IC_{50}

values of 62, while on SGK2 it has IC₅₀ value of 103 nm, obtained with an activity-based *in vitro* scintillation proximity tests. From these tests, it was possible to observe greater selectivity for SGK1, than other AGC kinases, such as AKT.³ A series of subsequent studies on GSK650394 analogs, carried out by Liang et al. in 2017, led to the discovery of the QGY-5-114-A, which revealed a slightly lower IC₅₀ value (122.9 μM), compared to that found for GSK650394 (135.5 μM) in the HCT116 tumor cell line. Although GSK650394 and its analog share the same azaindole nucleus, the 2-chlorophenyl group of QGY-5-114-A covers 7-azaindol-free NH, preventing the formation of a crucial hydrogen bond interaction with the hinge region, as reported for GSK650394.¹³ Therefore, it can be assumed that the two compounds have different binding modes. There is no data for SGK1 inhibition for QGY-5-114-A. Screening of an internal library of compounds based on pyrazole [3,4-d]pyrimidine scaffold, already characterized for Src/Abl activity, resulted in the discovery of the selective inhibitor of SGK1, SI113. Based on the known over-expression of SGK1 in glioblastoma multiforme (GBM), a follow-up study was performed on this type of cells in order to assess the effect of SI113. The results on three GMB cell lines (LI, ADF, and A172) showed a significant reduction in cell viability with IC₅₀ values ranging from 9.1 to 11.2 μM, as reported by Talarico et al.¹⁴

In 2015, Halland et al indicated a new SGK1 inhibitor.¹⁵ Starting from 1*H*-pyrazole[3,4-*b*]pyrazine derivatives, only one compound, known as 14g, was selected for the favorable pharmacokinetic profile and IC₅₀ value of 0.003 μM. To date, only a small number of SGK1 inhibitors are available, and information on their selectivity is not entirely clear. Therefore, there is a strong demand for new SGK1 inhibitory compounds, with a strong selectivity profile.¹⁵

Chapter 2

Aim Of The Project

In recent years, a growing interest emerged around SGK1 and its role in multiple pathways involved in the pathogenesis of cancer. Although several SGK1 inhibitors have allowed new therapeutic approaches to be investigated, the prospects for use in oncology are still questionable.³

The involvement of SGK1 in tumor development and metastasis, or in drug metabolism and resistance, allows its use as a potential indicator in prognosis⁹

This project aims to develop new chemicals, focusing on the apical component of signal transduction pathways such as SGK1. The specific goal of the project is the development of various classes of SGK1 small molecule inhibitors exploring the possibility of rational design of new chemical entities with scaffolds different from those studied until now, as well as the virtual screening of commercial libraries through the design of selective pharmacophore models, capable of leading to the selection of compounds with potential inhibitory activity on SGK1.

Chapter 3

Materials And Methods

3.1 Databases and online resources

3.1.1 Protein Data Bank

Founded in 1971 by a group of crystallographers at Brookhaven National Laboratory, the Protein Data Bank (PDB) is today one of the leading global resources for experimental data critical to scientific discovery. Initially conceived as an archive of information, limited to the crystallographic community, as the number of deposited entries increased, it became evident that this enormous amount of data would have a much wider field of use than initially imagined.¹⁶ The database includes 203,863 entries with their atomic coordinates and related experimental data produced with techniques such as crystallography, 3D electron microscopy and nuclear magnetic resonance spectroscopy. All these methods are related to protein structures, nucleic acids and biological macromolecules, as also reported by Burley et al.¹⁷ Each PDB model is identified by a unique 4-digit alphanumeric code and contains a coordinate file, which identifies the atoms present in the structure and their 3D position in space, as well as information on structure, sequence, and experiment. These files are also available in PDBx/mmCIF and XML formats.

As reported in the PDB site coordinates section, (<https://www.wwpdb.org/documentation/file-format-content/format33/sect9.html>)¹⁸ the macromolecule is represented by a text file, organized into data lines, called records. The records describe macromolecular specifications such as x, y, and z coordinates for atoms in protein residues or nucleic acids, identified by the acronym ATOM. Instead, the inhibitor, ions, cofactors, and solvent coordinates are identified as HETATM. The final parts of the polypeptide chains are identified by TER and located at the end of the text file. Furthermore, secondary structures, such as α helices and β sheets, can be recognized as HELIX and SHEET, respectively.¹⁹ The crystallographic experiments are represented by electron density

maps, which define the exact position of each atom.²⁰ The accuracy of the model can be evaluated considering several parameters, among which: the R factor and the resolution. The first also referred to as the reliability factor, indicates the degree of divergence between the crystallographic model and the X-ray diffraction data. In other words, it is a measure of the ability to reproduce observed data, as reported by Miyaguchi et al.²¹ The resolution of the model, instead, indicates the level of precision of the crystallization procedure. One of the parameters used in the selection of crystal models is that developed by Cruickshank, known as diffraction-component precision index (DPI).²² It considers the quality and completeness of the data reported in the model at a specific resolution. In this project of thesis, the 3D crystallographic models used for the study of SGK1 are the PDB complex 3HDM (with the co-crystallized ligand MMG at the resolution of 2.60 Å) and 3HDN (with the co-crystallized ligand GMG at the resolution of 3.10 Å).

3.1.2 DUD-E decoys and validation parameters

The DUD-Enhanced data set (DUD-E) is an online tool specifically designed for evaluating the virtual screening abilities of various docking procedures and their scoring functions. It provides 102 known targets, a series of known active compounds, each of which generates 50 inactive compounds, or decoys, generated to avoid bias in the benchmarking.²³ Different parameters for the validation of the pharmacophore model (ROC, AUC, and EF) have been calculated. The Receiver Operating Characteristics (ROC) curve shows how the model is able to distinguish between active and inactive compounds. The ROC curve provides the positive rate plotted about the false positive.²⁴ The AUC value has a range between 0, if the inactive compounds are found first, and 1 if the opposite occurs. Therefore, to assess model performance, starting from a set of 9 known inhibitors, were produced 438 decoys, through DUD-E implementation. Both groups of compounds have analogue physical-chemical properties, but diverse two-dimensional topological descriptors.

In addition to these parameters, the enrichment factor (EF) statistic parameter was determined to examine model performance. The formula is written below:

$$EF = \frac{Ha(A + D)}{HtA}$$

Where, D = total number of structures, A = number of actives, Ht = total number of hits and Ha = number of active hits. Therefore, when the model successfully separates all active (A) from inactive (D), the ROC curve will have a steep slope and high values of AUC and EF.

3.2 Pharmacophore modeling

Pharmacophore modeling has been widely employed in drug discovery, due to its ability to search for active molecules and optimise them. The pharmacophore model is based on the well-established concept that molecules are similar to each other due to the presence of chemical groups, such as donors or acceptors of hydrogen bonds, or aromatic rings. As a consequence, similar compounds, may interact with a receptor in the same way, resulting in similar biological activity.²⁵ It is possible to generate a pharmacophore model from one or more actives ligands. This model represents the main molecular characteristics needed for the activity, including lipophilic groups, aromatics, hydrogen bonds and charged groups. Once generated, the pharmacophore model can be used to examine databases, separating structures with the same characteristics from others. The Phase module has been widely used for the creation of pharmacophore hypotheses.²⁶ The chemical features of the molecules are coded into pharmacophore features, such as hydrogen bond acceptor (A) or donor (D), negatively (N) or positively (P) charged group, aromatic ring (R) and hydrophobic group (H), provided by the software.²⁷

The first step of this study involves the preparation of ligands, which are optimized and protonated by LigPrep.²⁸ Using Phase, a combined structure and ligand-based approach was used to generate a pharmacophore model. The pre-aligned docking poses of the two

azaindoles inhibitors, MMG and GMG, were imported into the “Develop Pharmacophore hypothesis” tool. Hypotheses generation was performed by setting a minimum of four and a maximum of five features and scoring was done using the default parameters.²⁹

The hypotheses generated have been filtered according to the presence of some essential features for the interaction with the receptor and then, were classified by survival score. The best pharmacophore model includes a hydrogen bond acceptor (A), a hydrogen bond donor (D), a negatively charged group (N), and two aromatic groups (R). Then, the resulting ADNRR_1 hypothesis has been refined by adding excluded volumes, to eliminate the compounds that in screening would clash with the binding pocket.

3.3 Docking studies

The purpose of applying molecular docking methods is to predict the affinity of a ligand for the target macromolecule (protein or nucleic acid) and to assess the energetic terms of the generated complex. Two fundamental aspects are related to the formation of the complex: the generation of poses and their evaluation by means of scoring functions. Docking algorithms do not consider ligands and proteins as flexible objects, so different degrees of freedom have been implemented both translational and rotational.³⁰ Docking software generally samples possible conformations of the ligands within the binding site and then attributes to the poses a score that expresses the affinity of the ligand for the target.

Sampling methods to identify ligand binding geometries are based on the application of a systematic or stochastic approach. Instead, among the strategies regarding the flexibility used in docking simulations, there are three different alternatives: flexible, rigid, and semi-rigid.³¹ In the systematic method, the variation of the structural parameters of the molecule is performed in a gradual manner, modifying the various degrees of freedom (torsional, rotational and translational) in order to obtain the minimum relative and/or absolute energy.

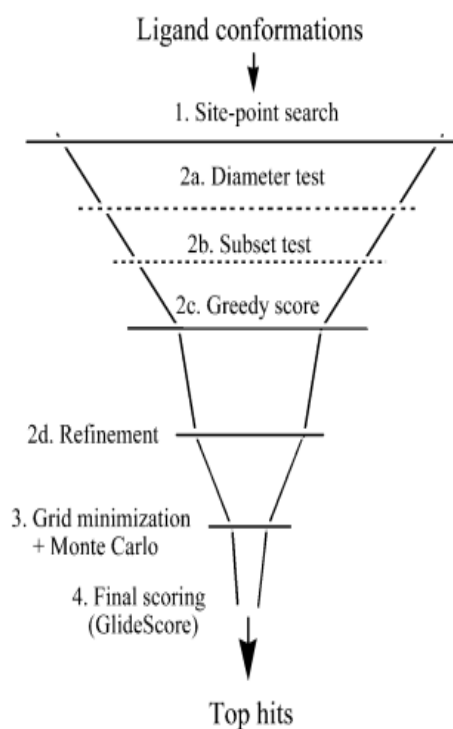


Figure 2. Glide hierarchical filters.

The stochastic method consists, instead, in the random variation of the relative structural parameters to the initial molecule, with consequent increment of the probabilities to find the absolute minimum of energy. Most software deals with the receptor structure as a rigid entity, allowing only the ligand to adapt to the receptor, to limit and manage the computational complexity related to a major number of degrees of receptor freedom.³⁰

In this thesis, the software Glide, an acronym for Grid-based Ligand Docking with Energetics, of the Schrödinger suite, was used. Each identified pose was submitted to a series of hierarchical filters, increasingly restrictive, to identify a limited number of candidates (**Figure 2**).

Only a small number of poses survived to the filters and access to an energy minimization process and a conformational analysis by the Monte Carlo stochastic method. Energy values were calculated employing the OPLS force field. Finally, generated poses are assigned a Gscore (Glide Score) which represents the final evaluation of the pose under consideration.

Glide allows us to operate with three different degrees of accuracy: HTVS (High Throughput Virtual Screening), SP or XP (Standard or Extra-Precision).^{32,33} HTVS allows to evaluate quickly very large libraries of compounds but has the lowest level of precision for conformational sampling. The SP algorithm increases accuracy at the expense of execution speed. Finally, the XP algorithm represents the highest level of accuracy, used exclusively to evaluate the best poses resulting from previous algorithms. Both the calculations were performed in this work. A grid box was generated employing the Receptor Grid Generation tool, and the co-crystallized inhibitor was selected as the center of the grid, leaving untouched the setting of box. One pose per ligand was generated for each compound and all other settings were kept as default.

3.4 Molecular dynamics simulations

A molecular dynamics (MD) simulation is a computational procedure that, by integrating motion equations, enables us to investigate the evolution of a physical and chemical system, at the atomic and molecular level. In order to simulate the motion of atoms, the MD applies the Newton's law of motion. The primary role played by MD is to calculate the forces resulting from the mutual interaction that act on each atom, and then derive the module and direction of the speed vector, using the mass of the atom. Hence, it is possible to know at the relative position of each atom during the simulation, obtaining a series of frames showing what changes take place in geometric arrangement of the structure, providing its trajectory. It is also very important to define the values relative to the time of the simulations, to cover the entire time frame in which the phenomenon occurs, which in this case oscillates between femtoseconds and nanoseconds.³⁴

Molecular dynamics simulations were performed using the AmberTools18 program suite. In a MD simulation, the topology and the coordinate files related to the protein-ligand complex must first be generated. This pivotal step is performed using the LEaP text-based interface, tLeap. The ligand is processed by the Antechamber program to generate topological files, as well as parameters including the total charge and the atomic type, through the use of the Generalized Amber Force Field (GAFF). Very often it is necessary to

parametrize some combinations of atom types involved in bonds, angles or dihedral within molecules. Thus, you must then specify any missing parameter before generating the topology and coordinate files. The Parmchk2 utility allows to check the availability of the required parameters. Instead, through the Pdb4amber tool, the protein file is modified, removing water and hydrogens, so that it can be used in Tleap. The files obtained from these procedures will be used as input in Tleap, loaded together with the relative force fields, to generate the files necessary for the simulation of molecular dynamics.

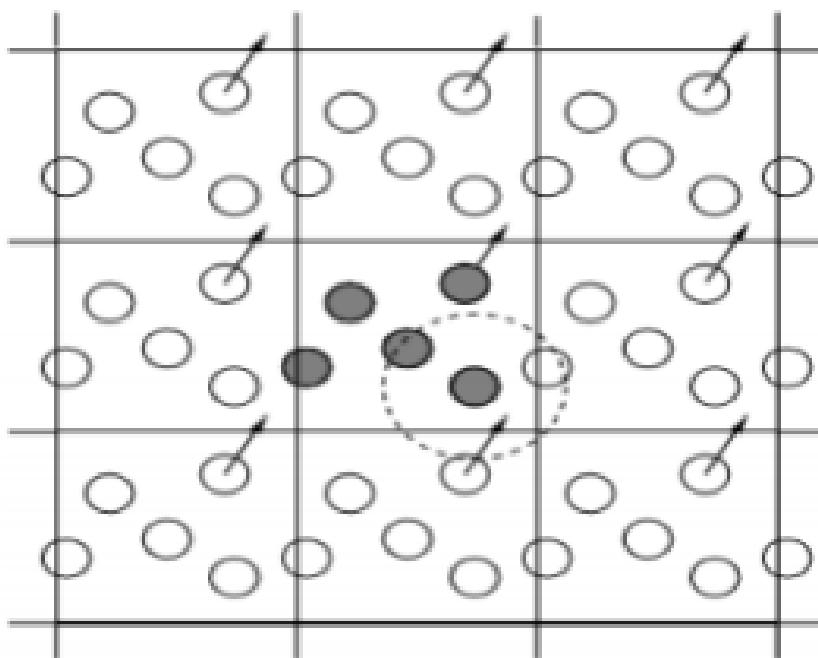


Figure 3. Periodic Boundary Conditions (PBC).

In the next step, the total charge of the protein will be defined, based on the previously loaded file, to bring to zero by adding a specific number of Na^+ or Cl^- ions.

Since the simulation environment must be as close as possible to the real conditions of the macromolecules, the whole complex is placed inside a cubic box filled with water molecules,

adopting specific experimental conditions of solvation known as Periodic Boundary Conditions (PBC), in which the system seems to border other identical systems. In this way, the molecules near the edge of the box are treated as if they were inside a cell and, as shown in **Figure 3**, when one of them exits the box, it will be replaced by an identical molecule coming from an adjacent cell. The procedure ends with the generation of a parameter file/topology (.prmtop) that groups the properties of the atoms, the bonds involved and their connections, and a coordinate file (.inpcrd) related to the system. Then, a minimization of whole system is made, through the use of the software Sander, acronym of "Simulated Annealing with NMR-Derived Energy Restraints". In fact, from a geometric point of view, the files obtained by tLeap may not actually represent the minimum effective corresponding to the force field that we are using.³⁵

The minimization of the system, in this initial phase of the simulation, allows to remove possible steric clash between atoms or similar, that could lead to instability of the system. However, this phase must take place in the presence of restraints, necessary to maintain the coordinates of the atoms in the complex and allowing the minimization of water molecules. Then we move to the simulation phase with the introduction of the program Pmemd, acronym of "Particle Mesh Ewald Molecular Dynamics", an improved version, compared to Sander, in terms of process speed, which implements the use of the GPU in addition to normal processors. As already seen in the previous step, it is possible to insert restrictions in order to keep the coordinates of some atoms fixed within the system. Depending on the temperature, pressure and volume conditions being worked, specific parameters with compatible information must be entered. Normally, in the early stages of simulation, the temperature is zero. It is therefore necessary to administer heat to the system, inevitably altering the calculation of pressure. The aim is to achieve constant temperature and pressure. Therefore, it is advisable to carry out a first phase at constant volume. Then, when the temperature is stable around 300° K, you can move on to the production phase, setting constant pressure and temperature. At the end of MD, the analysis of the results can be performed by applying multiple evaluation criteria. In particular, the Root Mean Square Deviation (RMSD) analysis was performed to assess whether the complex undergoes

significant positional changes during the simulation and to evaluate for examples the exit of the ligand from the active site. Finally, the energy values of the Molecular Mechanics Generalized Born Surface Area (MM/GBSA) were analysed. It provides a useful value for understanding the stability of the complex in terms of free energy. Finally, the analysis of hydrogen bonds allowed us to observe the presence of interactions between ligand and receptor and their duration.³⁶

3.5 RMSD analysis

Through this type of analysis, it is possible to observe the degree of similarity between two poses of the same compound. The calculation, known as RMSD, is generated based on the mean quadratic distance between pose atoms, measured in Angstrom (Å). The RMSD value is expressed by the following equation:

$$\text{RMSd} = \sqrt{\sum id_i^2/n}$$

The lower the RMSD value, the more two poses are similar. In fact, an RMSD value equal to 0 is equivalent to two perfectly identical poses, while if the values are around 2 Å still indicate a good degree of similarity. The procedure is performed using the Cpptraj program, part of the Amber18 suite. The software elaborates the results of the dynamics extrapolating trajectories in Netcdf format.³⁷

3.6 MM/GBSA calculation

MM/GBSA analysis is an extensively used method, available among Amber18 tools. The analysis allows to obtain an estimate of the energy difference (ΔG_{bind} in kcal/mol) between the free and bound states of the ligand and protein, after the minimization phase.

The calculation of the energy associated with the two states is expressed as follows:

$$\Delta G_{bind} = \Delta G_{complex} - \Delta G_{protein} - \Delta G_{ligand}$$

Where ΔG_{bind} is the free binding energy, while $\Delta G_{complex}$, $\Delta G_{protein}$ and ΔG_{ligand} are the energies associated with the complex, protein, and ligand, respectively. The energies are calculated by the equation $\Delta E_{MM} + \Delta G_{GB} + \Delta G_{non\ polar} - T\Delta S$.

The first member is the energy of gas phase interaction between ligand and bound protein (ΔG_{bind}), which considers Van der Waals and electrostatic energies. Instead, ΔG_{GB} and $\Delta G_{non\ polar}$ refer to the polar and non-polar components of desolvation free energy, respectively. $T\Delta S$ is given by the variation of conformational entropy on ligand bound, which has not been included here because they are unnecessary.³⁸

3.7 Hydrogen bonds analysis

Within the overall information content that can be extracted from MD simulations, the analysis of hydrogen bonds allows to understand what are the interactions present and their frequency with respect to the entire simulation time. For this purpose, the Visual Molecular Dynamics (VMD) software was applied to visualize and explore MD simulations, as well as to perform the analysis of hydrogen bonds that occur. Using a graphical interface, a file is generated txt that describes the existing interactions and the residues involved, besides the persistence of the interaction on the total time of the simulation, in percentage value.

Chapter 4

Results And Discussion

4.1 Structure-based and ligand-based pharmacophore modeling

The first step was to create a combined structure-based and ligand-based pharmacophore model based on three-dimensional crystal structures of SGK1 kinase, available in PDB. Based on the presence of a selective co-crystallized SGK1 inhibitor, only the 3HDM and 3HDN crystal structures were selected and imported, excluding the 2R5T structure.

All complexes were in the DFG-out conformation, and lack structural information for the seven residues located at the outer part of the α C-helix. Each of the co-crystallized ligands was extracted from the respective complex and was re-docked in its receptor, using the Glide XP module, to check for the reliability of the docking procedure by evaluation of the Gscore and RMSD values. The results show excellent Gscore values for the two compounds and RMSD values in an appreciable range, thus validating the docking protocol used (**Table 2**).

Table 2. Self-docking results of MMG and GMG with their respective protein structure.

PDB ID	Native ligand ID	Glide XP Gscore (kcal/mol)	RMSD (Å)
3HDM	MMG	-12.511	1.19
3HDN	GMG	-11.615	0.55

Docking poses of MMG and GMG were used to generate a pharmacophore model, taking into account their interactions with the active site (**Figure 4**).

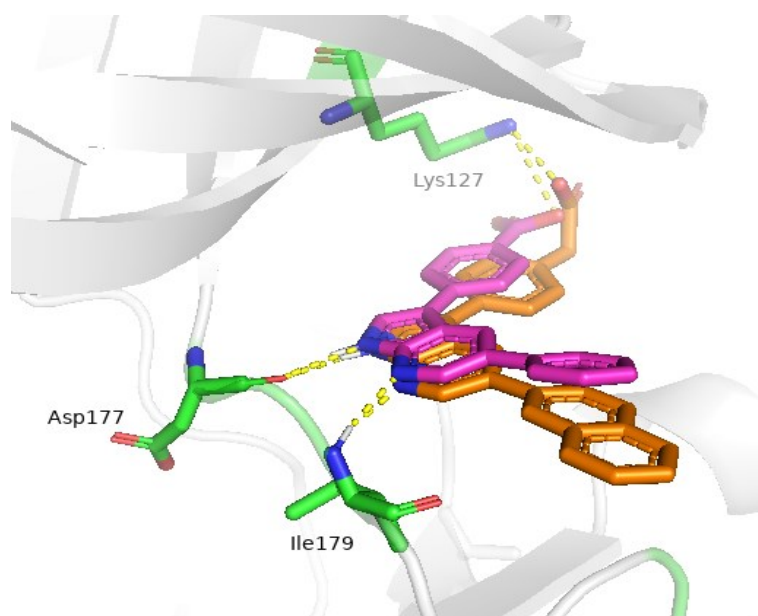


Figure 4. Alignment of docking poses of MMG (magenta) and GMG (orange).

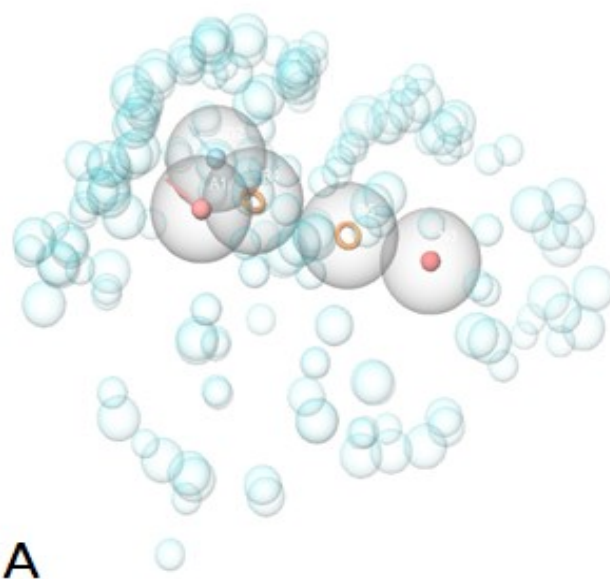
To obtain an excellent combination of features shared by active ligands, the minimum number of site points was set to four and the maximum to five in the Develop Pharmacophore hypothesis tool of the Schrodinger suite. Among 20 hypotheses generated, those lacking the HBA and HBD features, which appeared essential for the interactions with Asp177 and Ile179 of the hinge region, were discarded. The 6 surviving hypotheses were ranked based on the survival score. Then, hypothesis ADNRR_1, with the maximum survival score (4.338), was chosen as the top-ranked hypothesis (**Table 3**). In **Figure 5**, it's possible to observe the 3D spatial arrangement of all the features related to the selected hypothesis, as well as the inter-feature distance constraints. Finally, the selected hypothesis was refined adding the excluded volumes in positions that are sterically claimed by the protein environment.

This process ensures that the compounds derived from virtual screening, match the steric

requirements of the active site, while drastically increasing selectivity.

Table 3. Summary of Phase scores for the 20 generated hypotheses.

Hypo ID	Survival score	Site score	Vector score	Volume score
ADNRR_1	4.338	0.434	0.991	0.599
ADNRR_2	4.309	0.434	0.991	0.599
ADNRR_3	4.239	0.369	0.938	0.599
ADNR_1	3.843	0.540	0.990	0.599
ADNR_2	3.828	0.540	0.990	0.599
ADNR_3	3.820	0.392	0.993	0.599



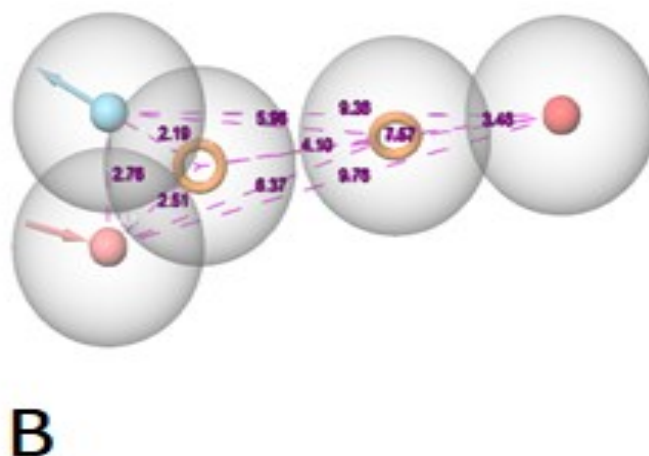


Figure 5. A) Final hypothesis was refined with the addition of the excluded volumes.

B) ADNRR_1 hypothesis and its inter-feature distances. Features are as follows:

H-bond acceptor, dark pink vector; H-bond donor, blue vector; negatively charged group, red sphere; aromatic ring, orange circles.

4.2 Pharmacophore model validation and screening

A preliminary *in silico* external validation, based on screening of a pre-set database, was performed to assess the ability of the pharmacophore model to differentiate between known active compounds and a set of decoys, generated from the DUD-E database. Then, nine active compounds, along with experimental activity provided by Pubchem were used. DUD-E server provided 438 inactive decoys from a subset of the ZINC database, then differentiated according to Lipinski's rules for drug-likeness.

There is a close similarity between the physico-chemical properties of the reference ligands and the generated decoys, however they differ in terms of 2D structure. Potential duplicates were eliminated using a filter in Maestro. In addition, in the early stage of the screening, a multi-conformational database was generated, through the use of the ConfGen algorithm,

which performs a conformational sampling of all databases. Duplicate poses with RMSD value less than 1.0 Å have been removed.⁴⁵ The validation of the combined pharmacophore model included the screening of the prepared database

Table 4. Virtual screening method validation parameters.

Parameters Used for Screening Validation*	Values
Number of actives (1%, 2%, 5%, 10%, 20%)	4;8;9;9;9
% of Actives (1%; 2%; 5%; 10%; 20%)	44.4;88.9;100;100;100
EF (1%; 2%; 5%; 10%; 20%)	50;50;20;9.9;5
ROC	0.99
RIE	16.46
AUAC	0.99
BEDROC (alpha=20.0; 160.9)	1; 1

*EF, enrichment factor; % of actives, percentage of actives; ROC, receiver operating characteristic; BEDROC, Boltzmann-enhanced discrimination ROC; AUAC, area under the accumulation curve; RIE, robust initial enhancement.

Therefore, it was necessary to calculate different enrichment parameters, as indicated in **Table 4**. EF qualitatively expresses the ability of the pharmacophore to differentiate the actives from the decoys. EF (1%) measures the enrichment score for the top 1% of the decoys screened.

The EF score of 1% obtained by the model was 50, which indicates that the screening protocol used has been useful in finding the active compounds among all those filtered. The ROC value represents the position of the actives about the orderly-classified compounds within a determined internal library. The range of values in which ROC is included, is from 0 to 1, where ≥ 0.7 is an optimal performance measurement value. In this study, the ROC value obtained is 0.99, therefore excellent for detecting active molecules.

Then, the ROC value (0.99), the AUAC (0.99), and RIE values (16.46), and a high EF of 50 show that our model can distinguish between truly active substances and decoy compounds.

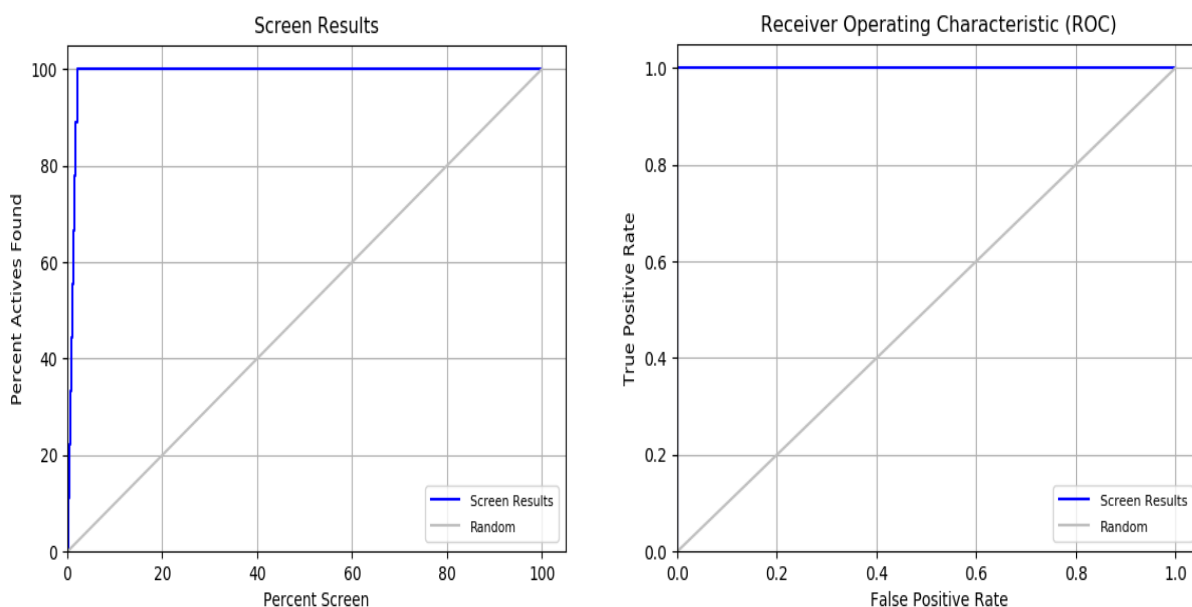


Figure 6. A) Hypothesis validation and enrichment analysis. B) Receiver operating curve (ROC).

The **Figure 6** shows the ROC curve and % screen plot ROC plot. The ROC curve was applied to validate the structure-based pharmacophore model, representing sensitivity versus specificity. Screening the database with the pharmacophore hypothesis, selected using Phase, generated a ROC curve that is far above the diagonal of random hits, indicating that the model can distinguish actives from inactive compounds. The BEDROC value at $\alpha = 20.0$ and $\alpha = 160.9$ confirms the ability to detect the actives in the database at an early stage.

Based on the validation performed, it is possible to confirm that the selected pharmacophore model is suitable for the research of new SGK1 inhibitors. Then, hypothesis ADNRR_1 was used as a 3D query to obtain potential molecules from the Molport chemical database. Database molecules were prepared and filtered according to the Lipinski rule of 5 and Qikprop analysis of the relevant molecular properties was performed. Furthermore, during the screening process all compounds with reactive functional groups, resulting in the presence of false positives, were eliminated.

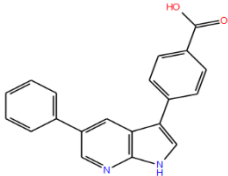
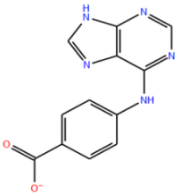
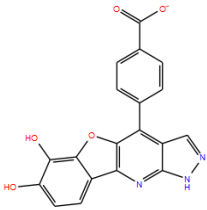
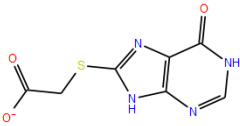
Conformational sampling was performed using the ConfGen search algorithm and the duplicate poses were eliminated. Only those compounds that match at least four characteristics of the hypothesis generated were considered for the following studies. The results were listed according to the degree of alignment to the hypothesis, by calculating the RMSD and other parameters such as site matching, volume terms and vector alignments. Finally, 223 hit compounds with a fitness value between 0.5 and 1.5 were obtained.

4.3 Molecular docking and visual inspection

A molecular docking simulation was performed on 223 screened molecules, using the Glide XP module. (**Table 5**).

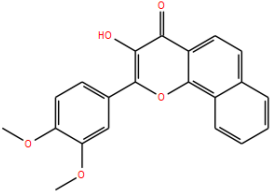
4-Results and Discussion

Table 5. Docking XP Gscore values of four hit compounds on 3HDM.

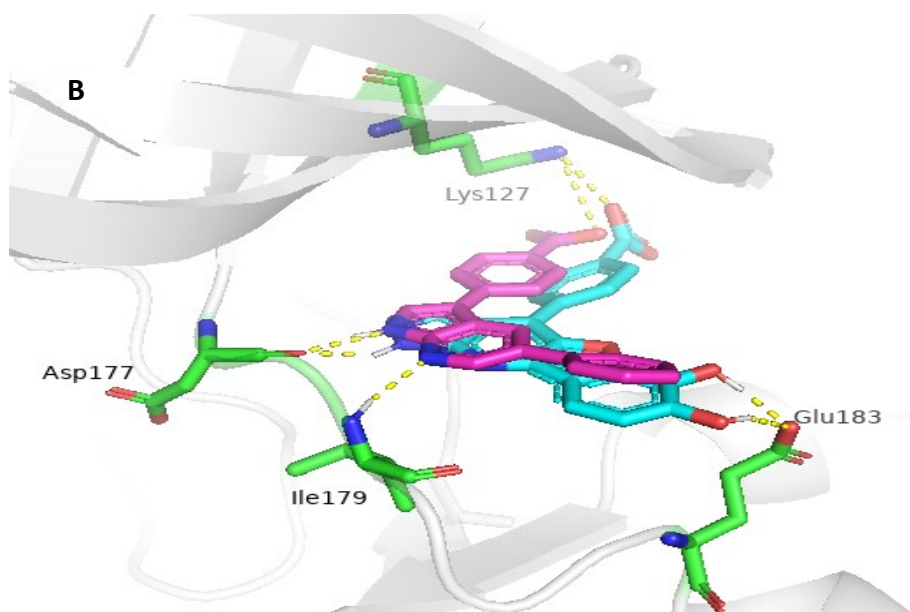
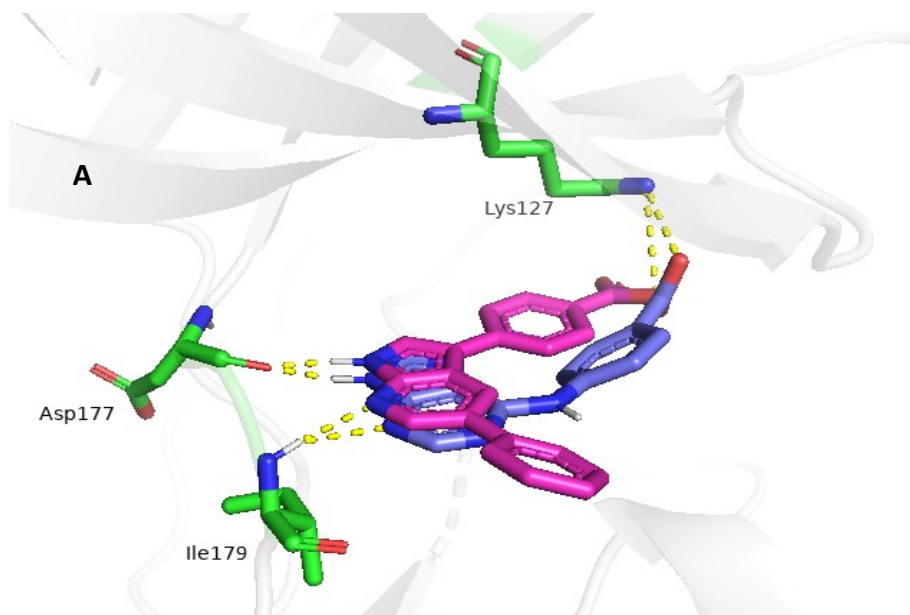
Compound code	Structure	3HDM XP Gscore (kcal/mol)
MMG		-12.511
1		-9.856
2		-9.821
3		-9.201

(Continued on following page)

Table 5. (Continued) Docking XP Gscore values of four hit compounds on 3HDM.

Compound code	Structure	3HDM XP Gscore (kcal/mol)
4		-7.364

To simplify the analysis of the results, the docking simulation was performed only on the 3HDM complex, which in the cross-docking procedure with the 3HDN complex, gave better results in terms of XP Gscore. Molecules selection was performed by Gscore analysis and visual inspection of binding poses to eliminate all compounds that did not have the required binding mode. Interestingly, 1 and the native ligand share a similar pattern of hydrogen bond interactions. Indeed, compound 1 (**Figure 7-A**) forms a double HBA-HBD interaction with Asp177 and Ile179 of the hinge region, mediated by imidazo[4,5-*d*]pyrimidine scaffold, which is superimposed on the azaindole scaffold of MMG.



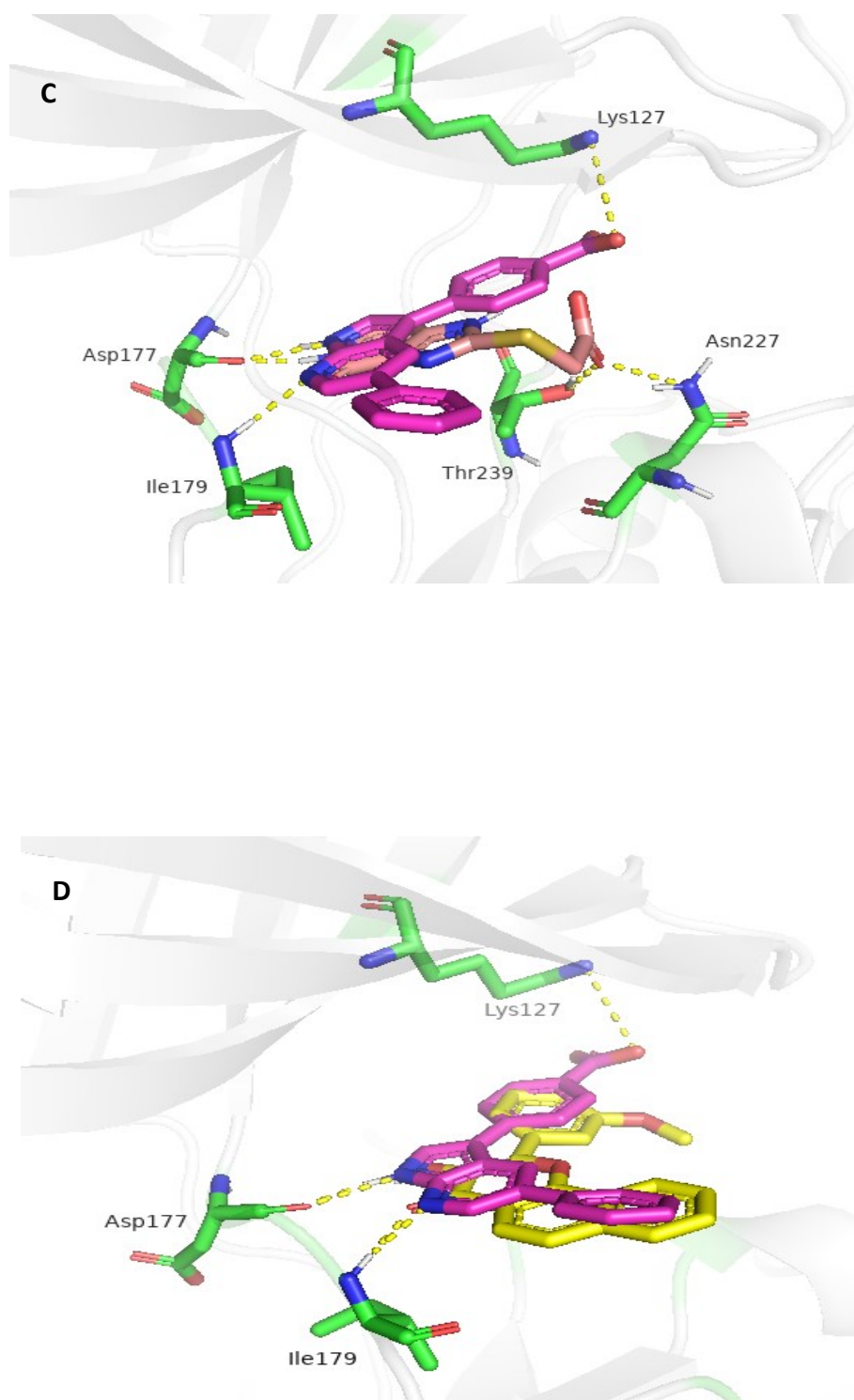


Figure 7. Molecular interactions between protein and putative hits. The protein-ligand complexes from the XP docking showing native ligand MMG (magenta) compared to (A) Compound 1 (purple); (B) Compound 2 (light blue); (C) Compound 3, (D) Compound 4 (yellow).

Moreover, the carboxyl group binds the Lys127 residue, but with a different orientation of the phenyl ring in comparison to that of the native ligand. Similar to the native ligand, compound 2 binds with a single hydrogen bond to Asp177, mediated by pyrazole NH. Moreover, the OH-mediated interaction of the benzoic portion with the Lys127 residue is maintained. An additional hydrogen bond was formed with Glu183, mediated by the hydroxyl groups in C7 on benzofuran moiety that protrudes in the area exposed to the solvent (**Figure 7-B**). The docking pose of compound 3 shows only a hydrogen bond with Asp177, compared to MMG, but maintains interaction with the Lys 127, mediated by the carboxyl group (**Figure 7-C**). Similarly, to MMG, while having a lower Gscore value than the other compounds (-7.364 kcal/mol), 4 shows a double hydrogen bond interaction with the residues of the hinge region, mediated through 3-hydroxy-4*H*-benzocromen-4-one moiety. An additional hydrogen bond is formed between the *p*-methoxy group and Lys127, as in MMG (**Figure 7-D**). Then, the four prioritized compounds were purchased and submitted to biological evaluation.

4.4 Biological results

The activity of selected hit compounds toward SGK1 was then evaluated in kinase assays. The percentage of SGK1 inhibition at 100 μ M and 10 μ M ligand, as well as IC₅₀ of the selected hits and their K_i were evaluated and compared with the Gscore obtained from molecular docking (**Table 6**). Compound 1 (-9.856 kcal/mol) and compound 3 (-9.201 kcal/mol) showed a very weak activity against the target, contrary to the expected suggestions by docking simulation. In the first instance, this may be due to the position of the aromatic core near the negatively charged group. Contrary to the reference compound MMG, which has a similar structure, the conformation assumed by the phenyl ring is not favourable to the stability of the complex. As for the second compound, it can be observed that the absence of feature R next to the carboxyl group, leads to a conformation where the main scaffold is disposed perpendicular to that of the reference molecules. Therefore, the presence of the feature R is essential for the correct interaction of the near carboxyl with the Lys127.

4-Results and Discussion

Compound 4 showed modest activity at 100 μM and was inactive at 10 μM , with very low values of IC_{50} ($21.51 \pm 3.52 \mu\text{M}$) and K_i ($11.88 \mu\text{M}$), probably due to a greater distance from the fundamental residues of the hinge region.

Table 6. Evaluation of biological results compared to XP Gscore values.

Compound	3HDM XP Gscore (kcal/mol)	% of inhibition		IC_{50}^* (μM)	K_i^{**} (μM)
		100 μM	10 μM		
1	-9.856	63.65 ± 2.67	14.09 ± 0.09	-	-
2	-9.821	101.65 ± 4.04	84.70 ± 4.13	1.68 ± 0.39	0.93
3	-9.201	17.09 ± 0.39	5.72 ± 2.32	-	-

Continued on following page)

Table 6. (Continued) Evaluation of biological results compared to XP Gscore values.

Compound	3HDM	% of inhibition		IC ₅₀ *	K _i **
	XP Gscore (kcal/mol)	100 μ M	10 μ M	(μ M)	(μ M)
4	-7.364	88.75 \pm 3.77	11.68 \pm 11.24	21.51 \pm 3.52	11.88

*IC₅₀ value represents concentration of inhibitor needed to reach 50% inhibition. **K_i is the affinity of the inhibitor to the enzyme. Compounds tested were assumed to act as fully ATP-competitive inhibitors.

Finally, the best active compound 2 showed significant inhibition both at low and high concentrations. Moreover, it is the only derivative with micromolar IC₅₀ (1.68 μ M) and sub-micromolar K_i (0.93 μ M). The binding pose of the compound shows that, despite the single interaction with the residue of Asp177 and the hydrogen bond with Lys127, the conformation is stabilized by an additional hydrogen bond with Glu183 placed in the region exposed to the solvent.

4.5 Molecular dynamics study of hit compound 2

The hit compound 2, obtained from the pharmacophore-based study on SGK1, is able of inhibiting the enzyme with a K_i of 0.93 μ M. The complex was subjected to two MD simulations of 50 ns, to confirm that the protocol used replicated the results obtained in the molecular docking study. The post-processing of the trajectory was performed in 0.150 M saline solution, which reflects the physiological conditions. MM/GBSA analysis was performed on an original trajectory obtained by the cpptraj module (AmberTools18),

processed writing one frame every 10. The RMSD plot was carried out with the RMSD Trajectory tool of the VMD software,³⁹ and displayed with the xmgrace interface. There is a small difference in MM/GBSA values between the two replicas (**Table 7**), as well as in terms of maximum values of occupancy. As a consequence, to simplify the following analysis, replica 2 was excluded.

Table 7. MMGBSA values of replicas of 2.

Replica	MM/GBSA (kcal/mol)	Max value of occupancy (%)
1	-47.59	48.19
2	-46.12	51.70

The visual analysis of the trajectory of replica 1, allowed to observe that the molecule maintains the same interactions and binding mode of the starting ligand, without particular positional changes. Specifically, the pyrazole NH-mediated hydrogen bond with the Asp177 is maintained. Similarly, the hydrogen bond between the OH group on the benzoic substituent with Lys127, and that mediated by -OH group on the benzofuranic group with Glu183, repeat as in docking.

However, the paramount interaction with the Ile179 of the hinge region, typical of the classical inhibitors of SGK1, not found in the docking pose, is present in dynamic simulation but assumes a negligible value compared to others. The RMSD plot analysis of SGK1 (black) and ligand (red), stabilized between 0.5 and 2 Å, suggests that hit compound 2 and the

receptor do not undergo significant changes in their position, as also confirmed by the value of MM/GBSA found (**Figure 8**).

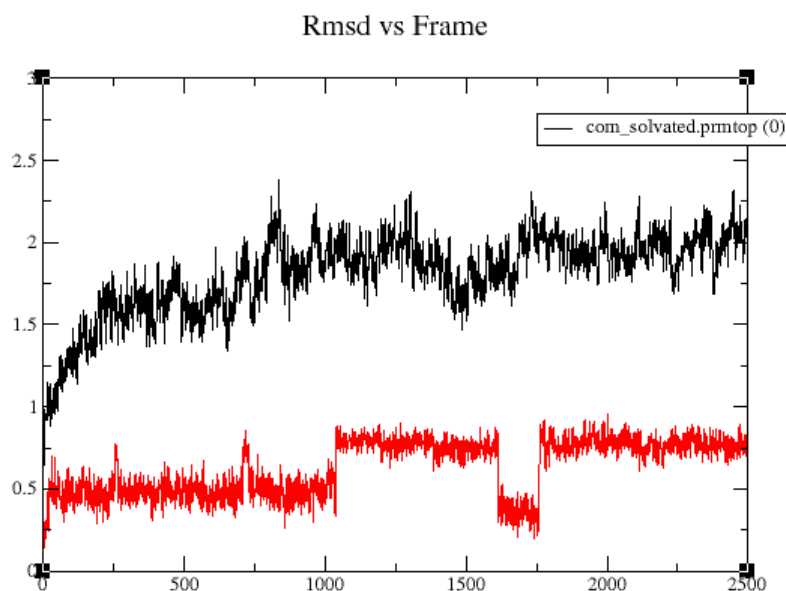


Figure 8. RMSD plot of SGK1 (black) and ligand 2 (red).

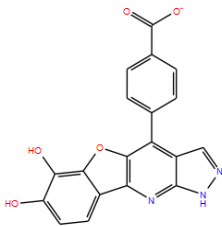
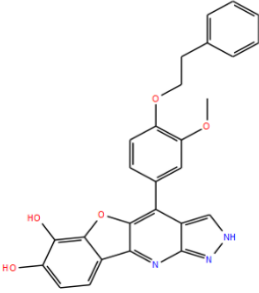
Overall, the results of the trajectory analysis of the complex studied are consistent with those obtained in the previous molecular docking study. They were then used as a starting point for subsequent studies focused on the optimization of the activity profile of the hit compound 2 in order to search for compounds with better K_i .

4.6 Focused library

The analysis of the binding mode of the reference molecule 2 by MD simulations has been exploited for the selection of a series of commercial analogues, to get new hit compounds

with improved K_i . In addition, this study allows further evaluation of whether the interactions observed so far are directly related to the stability of the protein-ligand complex and to the activity of 2. The research for analogues was carried out by similarity on the basis of a 0.8 cut-off compared to 2 and resulted in 705 compounds. As for the choice of receptor to be used for subsequent studies, this has fallen on clusters representing 60% of replication dynamics 1 of the compound 2. With regard to the choice of the receptor to be used for the following studies, clusters representing 60% of the dynamics have been used.

Table 8. Selected compounds from docking studies.

Compound	Structure
Reference	
5	

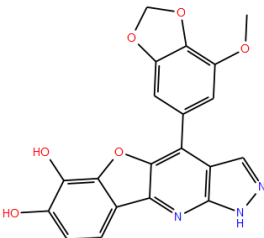
(Continued on following page)

Table 8. (Continued) Selected compounds from docking studies.

Compound	Structure
6	
7	
8	

(Continued on following page)

Table 8. (Continued) Selected compounds from docking studies

Compound	Structure
9	

Based on the RMSD values (1.176 Å: frame 0 and 1; 1.093 Å: frame 0 and 2; 1.396 Å: frame 1 and 2) obtained by superposition of the three receptors two by two, and the population percentages of the clusters (27%, 19%, and 14%, respectively), it was not possible to select a single frame, given the close proximity of the values found. Therefore, it was decided to use all three frames for docking studies. Reduction of the number of compounds was performed gradually, first using the Glide SP mode. The highest score values on the three frames are -10.132, -9.377 and -9.881 kcal/mol, respectively, higher than that of 2 (-9.103 kcal/mol). However, the visual inspection of the results showed frequent cases in which the compounds adopt a binding mode very different from that of the reference ligand. This allowed to decrease the number of filtered molecules to 19. The compounds were then docked via Glide XP mode to improve the accuracy of the procedure and subsequent selection. The high values of XP Gscore on the three frames were -11.998, -9.233, and -11.464

kcal/mol, respectively, also in this case higher or comparable to that of 2 (-9.821 kcal/mol). Following a full inspection procedure of interactions with SGK1, 5 compounds were finally selected (**Table 8**).

4.7 Post docking analysis of selected molecules

The compounds chosen by the docking study share the same main scaffold 4-phenyl-1*H*-benzofuro[3,2-*b*]pyrazolo[4,3-*e*]pyridine-6,7-diol except for compound 8, which has a methyl group instead of OH in C6. All compounds are characterized by a C4 phenyl on the central scaffold, variously decorated.

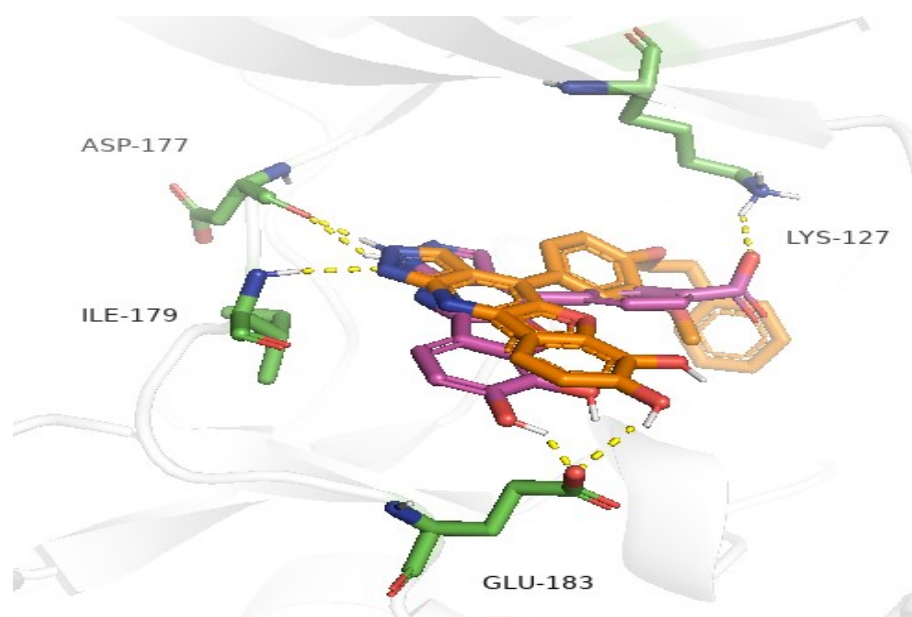


Figure 9. Alignment of docking poses of compound 2 (purple) and compound 5 (orange).

To simplify the visualization of docking poses, frame 0 was selected, on which the best results in terms of Gscore XP were generally obtained. The best compound in the series on all three frames is 5, bearing a 3-methoxy-4-(2-phenylethoxy)phenyl in C4 (**Figure 9**). The pyrazole portion forms the two key hydrogen bonds with Asp177 and Ile179 of the hinge region, while reference compound 2 interacts only with Asp177. Both compounds bind

Glu183 through -OH group present on the dihydroxy-benzofuranic portion. However, interaction with 5 occurs at a greater distance, probably due to the arrangement of the substituent in C4. The conformation assumed by the 2-phenylethoxy substituent also does not allow the compound to form the fundamental interaction with Lys127 in any of the frames analyzed. Compound 6, however, is the second best in terms of Gscore XP on the three frames. The compound, which differs from 8 by the loss of a -CH₂- group on the substituent present in C4, is arranged in the binding pocket so as to form the two hydrogen interactions with Asp177 and Ile179, and the additional interaction with Glu183.

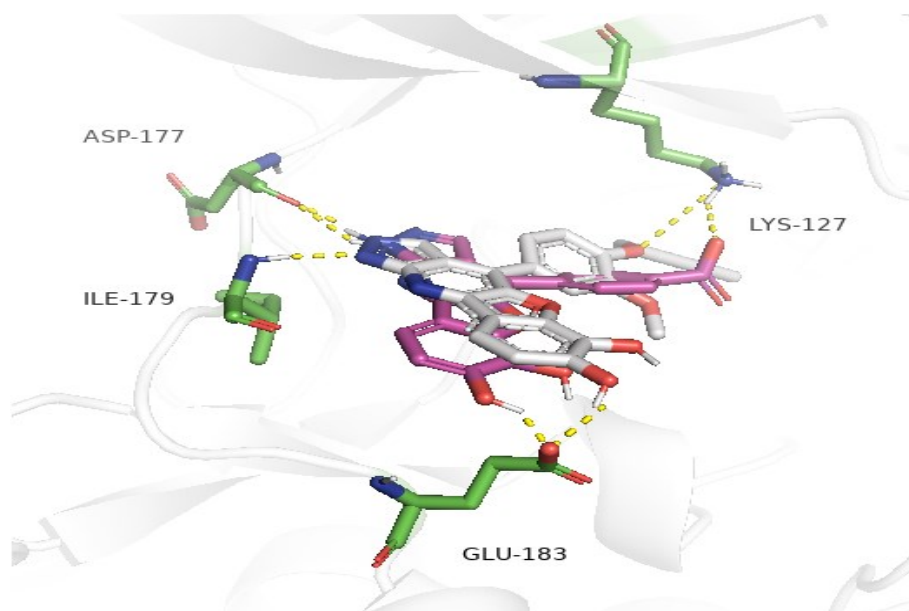


Figure 10. Alignment of docking poses of compound 2 (purple) and compound 6 (grey).

Moreover, frame 0 shows the interaction with the Lys127, which is absent in the others. The presence of the bulky side chain, also in this case, affects the arrangement of the molecule and the binding mode is not exactly similar to the precursor, but still very similar (**Figure 10**).

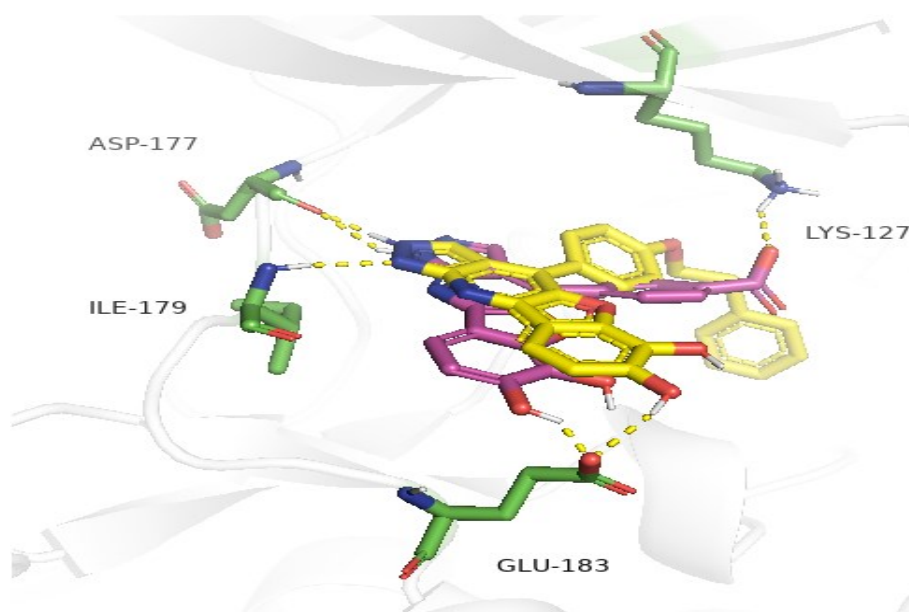


Figure 11. Alignment of docking poses of compound 2 (purple) and compound 7 (yellow).

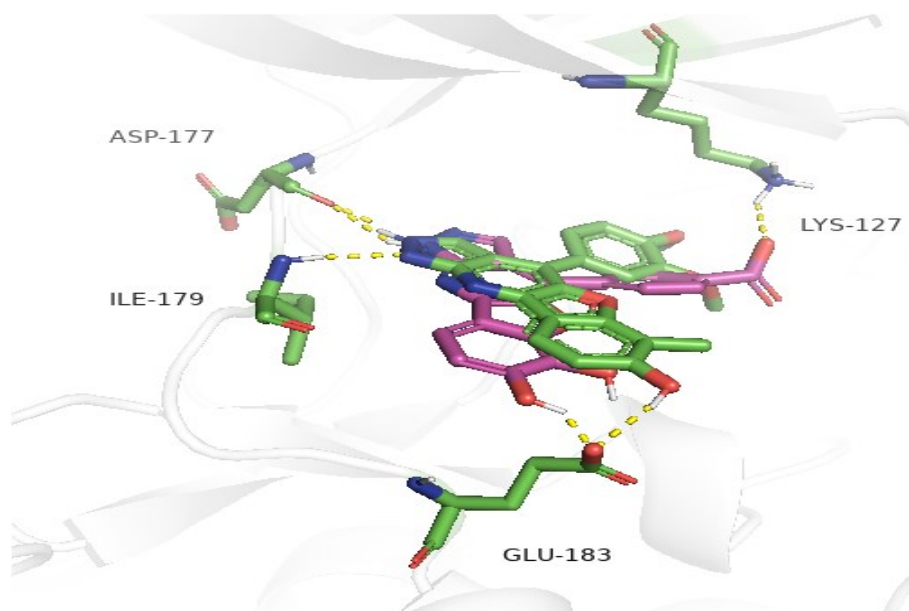


Figure 12. Alignment of docking poses of compound 2 (purple) and compound 8 (green).

Compounds 7, an analogous of 6 without the *m*-methoxy substituent, and the dimethoxy derivative 8, were also prioritized, **Figure 11** and **Figure 12**, respectively. They form the interactions with the fundamental residues of the hinge region and Glu183, as seen for the other compounds, while interaction with the Lys127 is present only in one of the three frames. Finally, compound 9, as shown in **Figure 13**, has a 7-methoxy-1,3-benzodiossol-5-yl cycle in C4, forms a single hydrogen bond with Asp177, as for compound 2, and is the sole ligand to form a double interaction with Glu183.

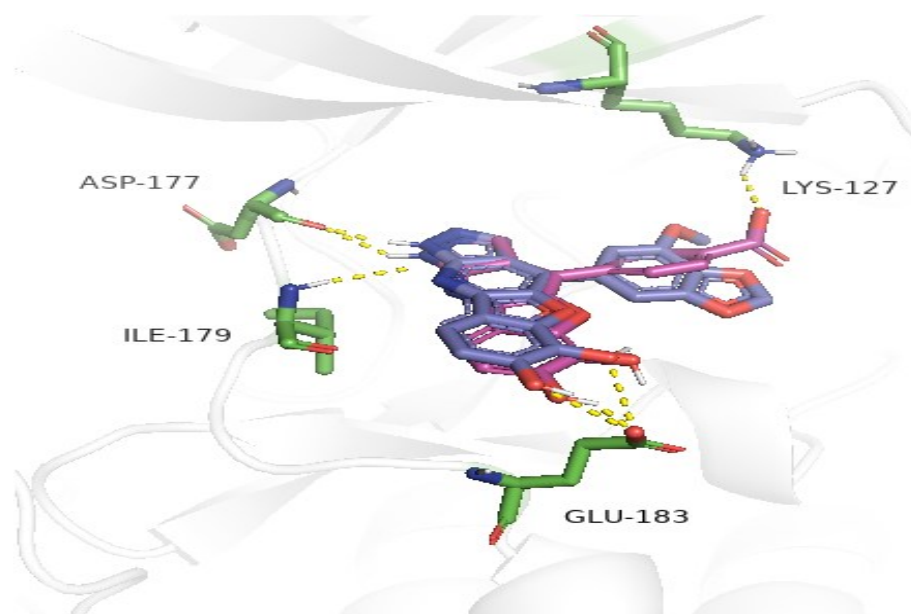


Figure 13. Alignment of docking poses of compound 2 (purple) and compound 9 (blue).

In addition, it forms an additional interaction with Lys127 on a single frame. Overall, the molecules analyzed have an additional interaction with Ile179 of the hinge region compared to the precursor. However, hydrogen bond interaction with Lys127 occurs less frequently. Finally, the biological tests will allow to understand how much and how the interactions observed and the presence of the various substituents affect the K_i values of for the compounds under study.

4.8 Biological results.

The activity of the compounds selected from the docking study has been evaluated in the biological assay towards SGK1 (**Table 9**).

Table 9. Evaluation of biological results.

Compound	% inhibition		IC ₅₀ * (μ M)	K _i ** (μ M)
	100 μ M	10 μ M		
5	84.32 \pm 19.43	37.56 \pm 12.19	-	-
6	96.53 \pm 4.00	86.10 \pm 1.39	5.01 \pm 0.86	2.24
7	96.39 \pm 3.81	69.55 \pm 5.28	11.0 \pm 3.09	4.92
8	n.a	n.a	-	-
9	86.67 \pm 21.20	72.23 \pm 4.42	5.65 \pm 1.54	2.53

*IC₅₀ value represents concentration of inhibitor needed to reach 50% inhibition. **K_i is the affinity of the inhibitor to the enzyme. Compounds tested were assumed to act as fully ATP-competitive inhibitors.

The percentage of enzymatic inhibition of each compound, together with the IC_{50} of the selected hits and their K_i , were evaluated and compared with the values found for the hit compound 2 and then, also, compared with the Gscore obtained from molecular docking study. Among the compounds evaluated, 5, 6 and 9 are found to have activity, albeit lower than that of the precursor. 6 has a good inhibition capacity at both 100 and 10 μM , resulting the best compound in terms of K_i (2.24 μM). Compound 5, an analogue of 8, appears to have a better inhibition capacity at both concentrations, although it shows a weak K_i . Finally, 9 has good inhibitory properties at both concentrations, with a K_i value comparable to that of 6. Finally, despite the best Gscore value of the series, 8 shows a significant inhibition at 100 μM that decreases at 10 μM , but no activity.

Although the results obtained are worse than the those of the starting compound, it is possible to make several considerations for subsequent studies. In particular, when choosing the C4 side chain, it is necessary to correctly evaluate the right distance and the bonding angle with the Lys127, key element for the activity of the compound. Indeed, in order to guarantee the activity, only the interactions seen with Asp177, Ile179 and also the Glu183, are not sufficient. Finally, it is necessary to evaluate a wider range of substituents to be included in C4, perhaps through the design of a new focused library, which allows to further deepen the behavior of ligands within the enzymatic pocket.

Chapter 5

Conclusions

To search a new class of SGK1 inhibitors, a combined structure-based and ligand-based pharmacophore model was created. Based on the aligned docking poses of MMG and GMG inhibitors, the Phase software generated a series of hypotheses, which were filtered and classified according to the survival score. The selected pharmacophore model includes one H-bond acceptor, one H-bond donor, two aromatic rings, and one negatively charged group. Then, the ADNRR_1 hypothesis was refined by including excluded volumes. To assess the ability to differentiate between active and non-active compounds, the hypothesis was validated by using a set of known active and decoys generated by the DUD-E database. Then, ADNRR_1 was used as a query for a virtual screening process, to identify new active compounds. A molecular docking study was performed on the 223 screened molecules, which were analysed based on their interactions with the protein and the value of XP Gscore. The resulting 4 compounds were prioritized and then purchased for biological evaluation. Among them, two compounds (2 and 4) showed micromolar activity towards SGK1 ($K_i = 0.93$ and $11.88 \mu\text{M}$, respectively).

The hit compound 2 was next used as a starting point for the search for new generation compounds with a better activity profile. The compounds found by similarity search were subjected to docking studies and 5 of them were submitted to biological assays. Inhibitory activity toward SGK1 was lower than that of the parent hit. However, a focused library will be built around the parent compound for finding more active analogues.

Part II

***In silico* studies of 3-(hetero)arylideneindolin-2-one derivatives with activity toward Src kinase.**

Chapter 6

Introduction

6.1 Tyrosine kinases

Tyrosine kinases (TKs) are a broad family of protein kinases, which mediate the signalling cascade, assuming a strategic role within multiple biological processes such as programmed cell death (apoptosis), cellular growth and differentiation. Although their activity is strictly regulated in healthy cells, they may undergo mutations or overexpression in various types of tumors, also resulting in metastasis.⁴⁰ Dysregulation of tyrosine kinases can be inhibited by selective inhibitors, an approach considered very promising for anticancer therapies. TKs can be ranked as:

- receptor tyrosine kinases (RTKs)
- non-receptor tyrosine kinases (nRTKs).

Rtk's are a class of transmembrane receptors, to which belong the platelet-derived growth factor receptor (PDGFR) and vascular endothelial growth factor receptor (VEGFR). nRTKs are cytoplasmic proteins consisting of nine families, to which they belong Abl, Ack, Csk, Fak, Fes/Fer, Jak, Src, Syk/Zap70 and Tec. In addition there are also Brl/Sik, Rak/Frk, Rlk/Txk, and Srm, which not included in the above mentioned families.^{40,41}

6.2 Src Non-Receptor Tyrosine Kinase

Src is a nRTK widely studied because of its involvement in multiple cellular processes, such as transcription, proliferation, motility and others. Since the discovery of Rous's sarcoma virus (v-src), which causes cancer in chickens, the oncogenic activity of Src has been related to the pathogenesis of cancer in humans.⁴²

Among the nRTKs, Src family kinases (SFKs) are certainly the most numerous family, interacting directly with different effectors, as well as G-protein-coupled receptors, steroid

receptors, thus mediating a wide range of biological functions such as cell survival and metastasis. The 11 members of this kinase's family have a highly conserved kinase domain organization and are classified into three groups: ubiquitously expressed, with preference for hematopoietic cells and expressed in derived epithelial tissues.⁴³

SFKs share the typical structure represented by Src kinase. Starting from the N terminus, SFKs show a Src-homology 4 (SH4) domain, a unique domain (UD), a Src-homology 3 (SH3) domain and a Src-homology 2 (SH2) domain, a poly-proline linker, a kinase domain (KD), and a C-terminal regulatory region (**Figure 14**).⁴⁴

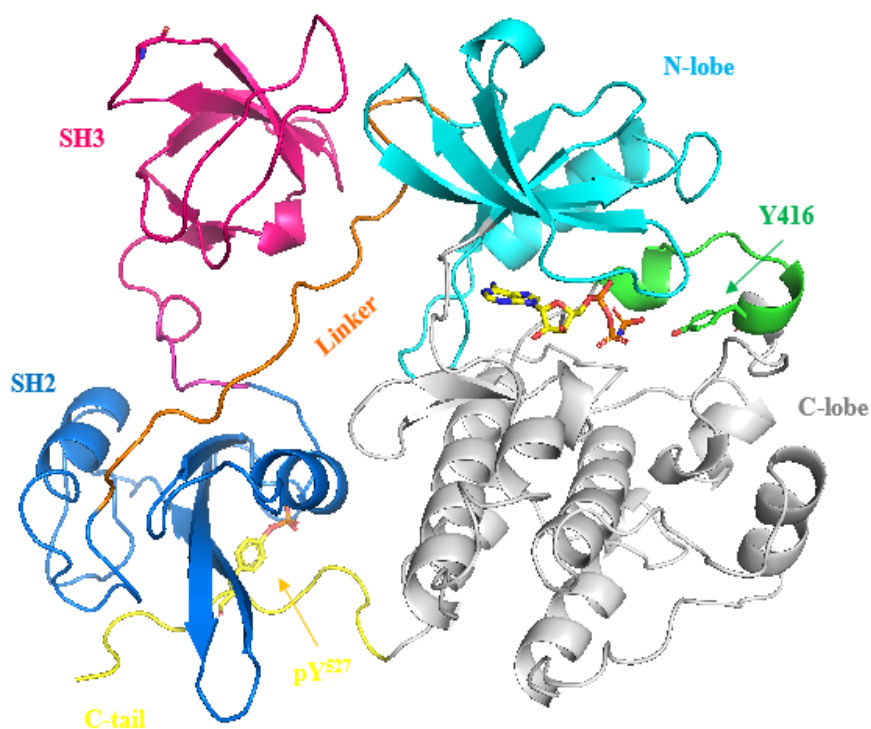


Figure 14. Structure of Src kinase in complex with AMP-PNP (PDB ID: 2SRC).

As in all tyrosine kinases, the SH1 catalytic domain is composed of N-terminal and C-terminal lobes, divided by a catalytic cleft, occupied by ATP. Src has two phosphorylation sites: a Tyr419 that is activated by autophosphorylation, and Tyr530 phosphorylated by C-terminal Src tyrosine kinase (Csk).⁴⁵ In the constitutively inactive Src, SH2 and SH3 assume

a specific arrangement that shields the phosphorylated pTyr530, hindering possible interactions within the catalytic site.

However, under specific conditions, the dephosphorylation of pTyr530, by various tyrosine phosphatase proteins, induces a change in the structure allowing Tyr419 autophosphorylation, resulting in activation of the enzyme and free access of the substrate to the kinase domain. Usually, the active states of a protein kinase are in equilibrium with the inactive ones.⁴⁶ As described by Möbitz et al (2015), the transition between the two states involves the displacement of the two preserved structural elements: the α C helix and the aspartate-phenylalanine-glycine (DFG) motif. When the DFG motif has a "DFG-in" conformation, aspartate is located in front of the ATP binding site, where it coordinates metal ions. Type I kinase inhibitors bind to this conformation (see below for a classification of tyrosine kinase inhibitors).⁴⁷ Instead, in the inactive "DFG-out" conformation, aspartate and phenylalanine are directed in opposite directions, exposing the DFG pocket. In this case, it will be occupied by the type II kinase inhibitors).

Furthermore, for the preserved α C helix, both states can exist. In the " α C-in" state, the glutamic acid within the α C helix forms a salt bridge with β 3-lysine. Conversely, when glutamate is removed from the active site conformation, it is in the " α C-out" conformation.⁴⁸

6.3 Src signaling in Cancer biology

It is well known that SFKs have pleiotropic functions in different cellular activities, and that aberrant expression of Src contributes to different aspects of the development phase of different kinds of cancer, such as some solid colon tumors, lung, prostate and breast, as well as the tumors of nervous system, such as glioblastoma multiforme (GBM) and neuroblastoma (NB).⁴² As reported in **Figure 15** by Hsu et al,⁴⁹ several studies on Src have allowed to define the main partners with which it interacts, such as the epidermal growth factor receptor (EGFR), the human epidermal growth factor receptor 2 (HER2 or ErbB2), the insulin-like growth factor receptor-1 (IGF-1R), the c-Met/hepatocyte growth factor receptor (HGFR) and many others. Through these interactions, Src transduces the survival signal

directly to the effector.⁵⁰

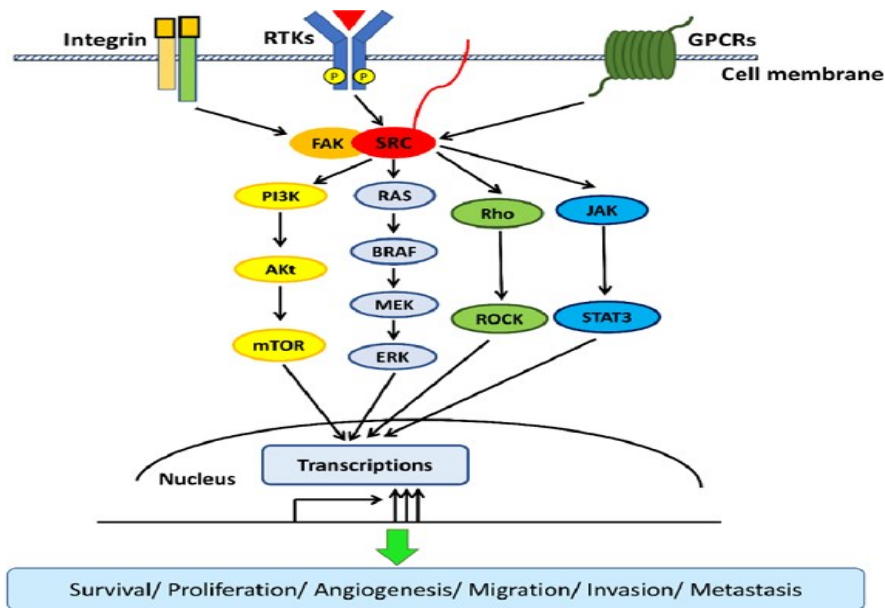


Figure 15. Src signaling pathways in cancer. Adapted from Hsu P. C. et al.⁴⁹

Activation of Src leads to downstream signaling through several pathways, including that of mitogen-activated protein kinase (RAS/MAPK). Src, cooperating with Janus kinases or JAK, can also stimulate activation of the signal transducer and activator of transcription 3 (STAT3) constitutively, which increases the potential for the STAT signaling pathway to control the expression of tumor-related genes. Furthermore, Src is a powerful PI3K/Akt pathway activator, which preserves against pro-apoptotic stimuli by phosphorylation and inactivation of death accelerators like caspase-9, Bad and Bax, as also reported by Seif et al.⁵¹ Src is also known for its key role during tumor metastases mainly for involvement in the control of the cytoskeleton, cell adhesion, migration and invasion. With phosphorylation of focal adhesion kinase (FAK), Src stabilizes other adhesion complexes, such as Ras homolog family member A (Rho A), and improves cellular adhesion to the extracellular matrix. The wide range of cellular activities in which Src participates and its role in cancer development have led to an intensive study of kinases, justifying Src targeting as a

promising anticancer therapy.^{51,52}

6.4 Small-molecule Tyrosine Kinase Inhibitors

Following the discovery of Imatinib (Gleevec®) in 2001 - the first tyrosine kinase inhibitor (TKI) approved by the US Food and Drug Administration (FDA), introduced into the clinic to treat chronic myeloid leukemia (CML), research on TKI has made great strides forward.⁵³ Over 20 years later, the classification of TKI has undergone a major review. They can be classified into 7 groups, according to the mechanisms of action and the modalities of binding to the target:

- **Type I inhibitors:** bind to the active conformation of a kinase, whereby the conserved DFG motif is oriented towards the inside of the kinase (DFG-in). Drugs such as crizotinib, dasatinib, erlotinib, vemurafenib and lapatinib belong to this group.
- **Type II inhibitors:** reversibly bind the inactive conformation of a target kinase, whereby the DFG pattern is directed outward from the ATP binding site (DFG-out). Due to the rotational movement of the DFG pattern outwards, new regions of space nearby to the ATP binding site are delineated that would otherwise not be accessible, making this kind of inhibition highly selective. Imatinib and sorafenib belong to this type of TKI.
- **Type III inhibitors:** bind at a certain distance from the catalytic site of ATP, so as to modulate allosterically the kinase activity. To this class of inhibitors belong Trametinib and Cobimetinib, approved by the FDA.⁵⁴
- **Type IV inhibitors:** are allosteric type inhibitors, which bind to external sites of the ATP binding site, in different way from type III inhibitors. Everolimus, sirolimus and temsirolimus belong to this class.
- **Type V inhibitors:** are considered bivalent inhibitors, which irreversibly bind the active sites of kinase through a ligand tethered to a second molecule that targets a region outside the active site. This class of compounds has not yet received approval for clinical use.

- **Type VI inhibitors:** covalently bind to the kinase target through the interaction of electrophilic groups with nucleophilic cysteines. Examples of this type of inhibitors, are Afatinib, Dacomitinib and neratinib.
- **Type VII inhibitors** are defined as non-classical allosteric inhibitors that target the extra-cellular domain of a TKR, without directly blocking the binding at the binding site. Among of these type of inhibitors SSR128129 affects the extracellular domain of the fibroblast growth factor receptor family (FGFR).^{54,55}

To date there are over 68 TKI approved by the FDA. Most of them are approved to be administered orally, some are for intravenous use (temsirolimus and trilaciclib), while only netarsudil should be given as an ophthalmic solution (eye drops).⁵⁶ TKI represent a new and improved step towards targeted anticancer therapy, which appear to act more specifically on malignant cells than conventional cytotoxic chemotherapy, as reported by Wing Tung Ho, V. et al (2018).⁵⁷ Therefore, they allow to obtain a wider therapeutic window with a lower toxicity profile.⁵⁸ TKI represent second or third-line therapies and are used in combination with traditional cytotoxic chemotherapy. Therefore, it is essential to continue research studies to remedy their limitations, such as resistance, and improve their therapeutic profile.⁵⁷

6.5 Src inhibitors

In recent decades, intensive research has been carried out for the development of Src inhibitors, allowing us to have today five targeted molecules approved by the FDA for the market (**Figure 16**). The first selective Src inhibitor was obtained in 2009, when the X-ray structure of the complex of active c-Src kinase with dasatinib was determined by Getlik and coworkers, now used for the treatment of acute lymphoblastic leukemia (ALL) and CML, in adult and pediatric patients.⁵⁹ Developed by Pfizer and approved by the FDA in 2017 as an inhibitor of TK Abl and c-Src, Bosutinib (SKI-606) was employed for the treatment of CML and known on the market as Bosulif®. It showed 1.2 nM IC₅₀ against c-Src in a cell-free assay.⁶⁰

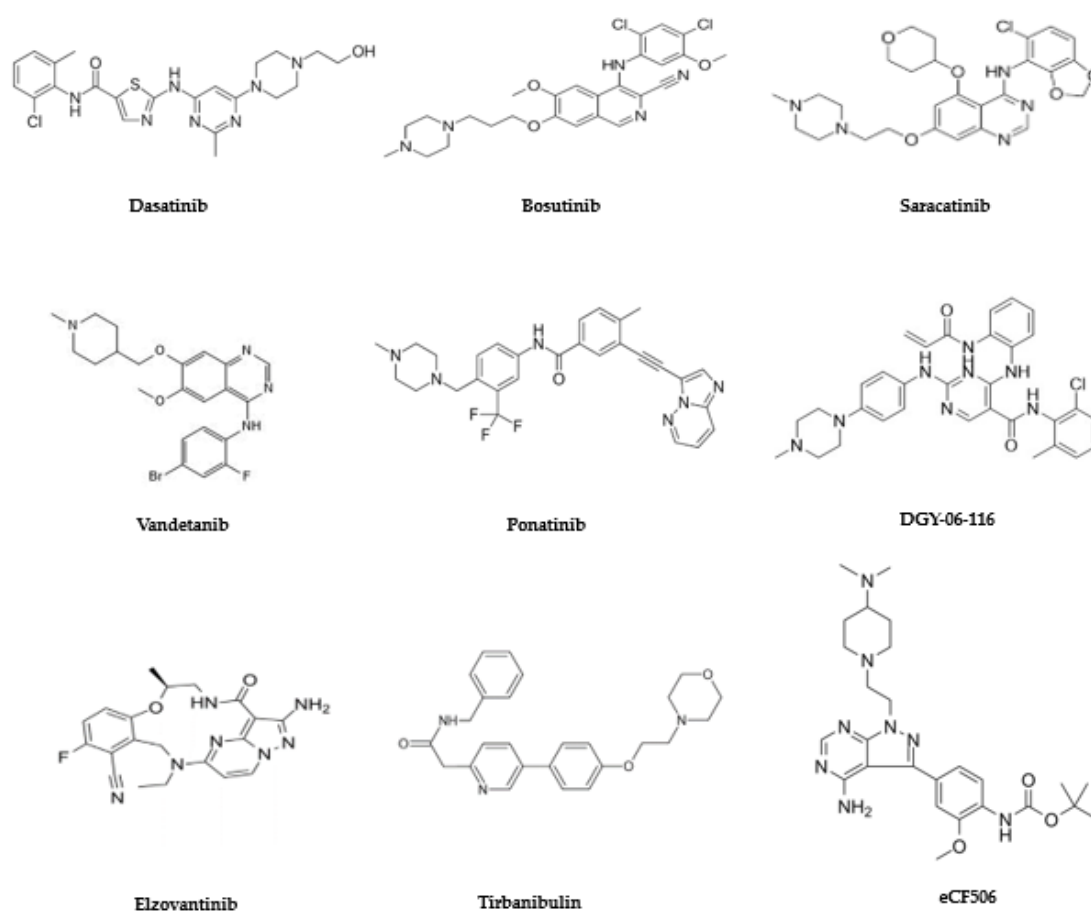


Figure 16. Structure of Src inhibitors already approved by the FDA or still in clinical phase.

Multi-TKI inhibitors have also been designed, in addition to the already known mono-target inhibitors. Multiple TKIs affect kinases or multiple pathways at a time and were designed for multiple diseases, including cancer.

For example, a multi-TKI Iclusig™ (ponatinib) was designed from inhibitor AP24364. Iclusig targets several kinases as Src, Abl, PDGFR and others, and is used as a second-line treatment option in CML Ph⁺ and ALL Ph⁺. In addition, it is the only effective inhibitor for targeting Abl-T135I mutation as well as to treat of lung cancer and GMB.⁶¹

Vandetanib is an oral TKI active against the EGFR family, VEGF, RET, of the ephrin receptor kinase family (EPH), and members of the SFK (Astrazeneca, 2011).⁶² Vandetanib has been approved to treat of CML, ALL and thyroid carcinoma, in adults with positive Philadelphia

chromosome. Finally, Saracatinib, in clinical trial for using in several solid tumors, is a more selective inhibitor of SFK. Finally, DGY-06-116, eCF506 and Tirbanibulin are promising phase I or II c-Src inhibitors.⁶³

Chapter 7

Aim Of The Project

In recent decades, a large amount of compounds targeting SFKs have been designed and tested in pre-clinical models. Despite confirmation that SFKs inhibitors can inhibit tumor progression, especially in solid tumors, their application in clinical trials has proven more complicated than expected. To date research is mainly focused on the inhibition of c-Src, with five competitive c-Src multikinase inhibitors, approved by the FDA for their employment in several kinds of cancer or still in clinical stages.⁶⁴

The present *in silico* study comes from the collaboration with the research group of Professor Dallavalle Sabrina, in support to a work for discovering new c-Src inhibitors. Starting from a small *in-house* library of structurally different molecules, the indolinone core, which characterized the most active compounds, was used as a promising scaffold for the subsequent investigation. Several substituted 3-(hetero)arylideneindolin-2-one have been developed and synthesized to identify the features important for the activity. Molecular docking studies of the resulting active compounds were carried out to provide a putative binding mode within the c-Src binding site and to direct following hit optimization studies.⁶⁵

Chapter 8

Material And Methods

8.1 Ligand Preparation

The 3D structures of a small library of indolinones derivatives and co-crystallized ligand AP23464 were designed in Maestro sketcher panel (Schrodinger suite) and then optimized with Ligprep. Partial atomic charges were attributed, and all ionization states at a pH of 7.4 \pm 0.5 were found. Low energy conformers were generated for each ligand using OPLS3 force field. Energy minimization was performed until it reached a RMSD threshold of 0.01 Å.²⁸

8.2 ADME prediction

QikProp (v 19.2) tool has been applied to predict the ADME profile of the five best results. QikProp predicts physically descriptors and essential pharmaceutical properties of organic molecules, providing optimal ranges valid for 95% of known drugs to refer to for the selection of molecules. Among the calculated descriptors are the logarithm of octanol/water partition coefficient, QPLogP_{o/w} (range is -2.0 to 6.5); the logarithm of aqueous solubility, QPlogS (range is -6.0 to 0.5) the predicted apparent Caco-2 cell membrane permeability, QPPCaco, (<25 poor and >500 great) and the Lipinski's rule of 5 values.

8.3 Docking studies

Molecular docking simulations have been conducted on the three-dimensional crystal structure of c-Src (PDB code: 2BDJ, 2.5 Å resolution), in a complex with inhibitor AP23464. Protein Preparation Wizard panel (Schrodinger suite, version 19.2) was used to perform optimization of the crystal structure, correcting ligand's partial charges, bond orders, adding hydrogens and removing unuseful water molecules. Prime software fixed

conformation of missing side chains.

The pre-processed proteins were optimized first with the help of the PROPKA software, and minimized with an OPLS3 force field. All molecules were docked in the binding site using Glide in SP mode (v 2019.2). Using Receptor Grid Generation tool, a grid box was generated, selecting the inhibitor as centroid, and others parameters to default.⁶⁶

Chapter 9

Results And Discussion

9.1 Biological dataset

Starting from the screening of a small in-house library of compounds with various scaffold, tested on c-Src at a concentration of 100 μM , the best compound 10 (K_i 3.8 μM , **Figure 17**) was used as a reference for the synthesis of new derivatives. Then, the benzylidene moiety at C3 of the indolinone scaffold was variously decorated or replaced with heterocycles.

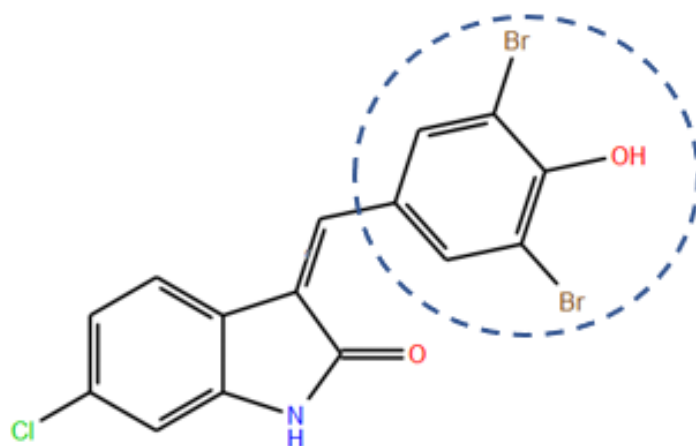


Figure 17. Structure of the reference compound 10.

In particular, a significant number of substituents were inserted in the para position of the pendant phenyl ring of 10, to check for the relationships between different stereoelectronic properties and the activity related to them (**Figure 18**). These include the cyano (12), nitro (13), dimethylamino (14), phenolic (15) and carboxyl (16) groups. Furthermore, in

compounds 23 to 27, the total length of the molecule was increased to eventually fill some empty regions of the active site, using aromatic or aliphatic portions for this purpose, as well as with a terminal acid or external portion. Finally, compounds 28 to 30, characterized by heterocyclic substituents, have been synthesized to study their potential role as hydrogen bond acceptor/donor. The structure and activity values of the new compounds are summarized in **Table 10**. Characterization of all compounds was performed by NMR technique, then the mixtures were tested without separation of the detected isomers (isomers 9:1 to 7:3 E/Z).

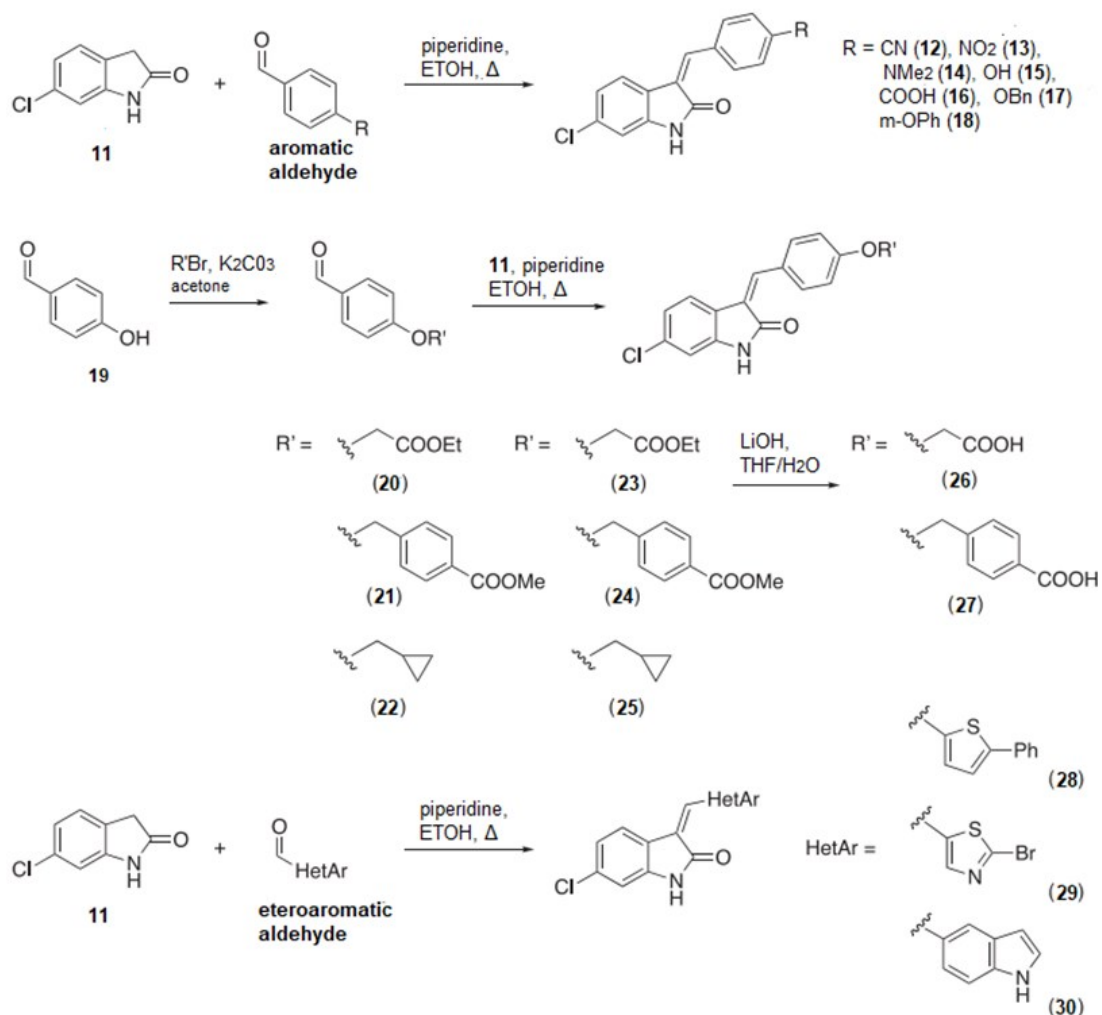
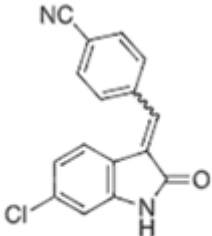
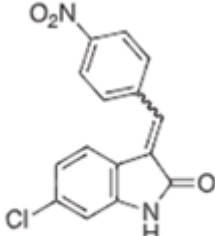
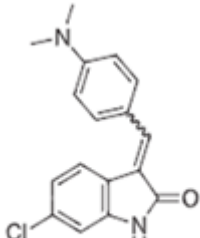
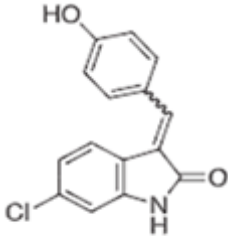


Figure 18. General synthesis by Knoevenagel condensation starting from 6-chloro oxindole and different aromatic and heteroaromatic compounds. Adapted by Princiotta S. et al.⁶⁵

References

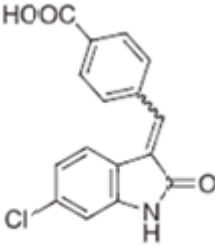
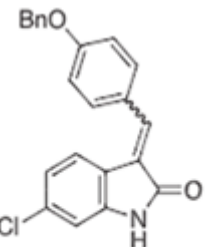
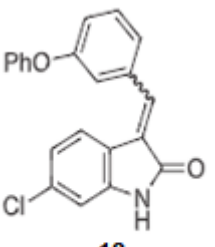
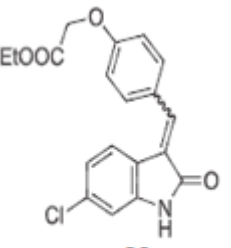
Table 10. Evaluation of the inhibitory activity of the novel indolinone derivatives. Adapted by Princiotta S. et al.⁶⁵

Compound	% Inhibition		IC ₅₀ *(μ M)	K _i **(μ M)
	100 μ M	10 μ M		
Dasatinib	100	100	1.60 \pm 0.22 nM	0.80 \pm 0.11 nM
 12	88.96 \pm 1.35	35.74 \pm 9.66	>10	-
 13	69.94 \pm 0.19	53.03 \pm 18.08	2.43 \pm 1.03	1.64 \pm 0.69
 14	75.72 \pm 7.63	85.35 \pm 6.98	0.71 \pm 0.43	0.48 \pm 0.29
 15	86.86 \pm 2.50	37.90 \pm 0.35	>10	-

(Continued on following page)

References

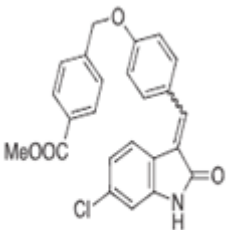
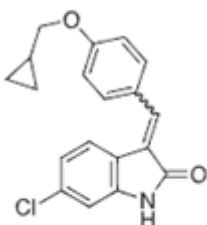
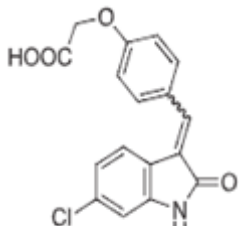
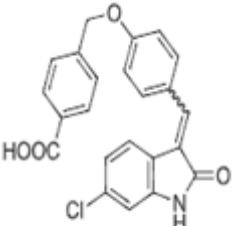
Table 10. (Continued) Evaluation of the inhibitory activity of the novel indolinone derivatives. Adapted by Princiotta S. et al.⁶⁵

Compound	% Inhibition		IC ₅₀ *(μ M)	K _i **(μ M)
	100 μ M	10 μ M		
 <p>16</p>	47.19 \pm 9.33	4.15 \pm 1.28	-	-
 <p>17</p>	84.27 \pm 0.82	16.86 \pm 13.95	-	-
 <p>18</p>	53.62 \pm 12.91	31.48 \pm 3.16	-	-
 <p>23</p>	80.84 \pm 6.47	71.67 \pm 2.27	3.24 \pm 3.43	2.19 \pm 1.03

(Continued on following page)

References

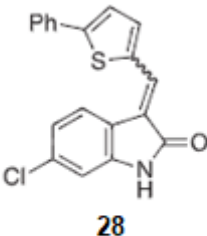
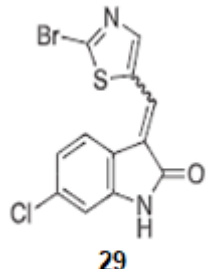
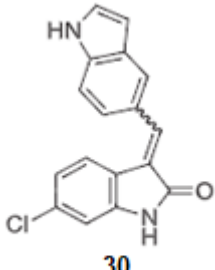
Table 10. (Continued) Evaluation of the inhibitory activity of the novel indolinone derivatives.
Adapted by Princiotta S. et al.⁶⁵

Compound	% Inhibition		IC ₅₀ * (μM)	K _i ** (μM)
	100 μM	10 μM		
 <p style="text-align: center;">24</p>	92.13 ± 3.84	41.05 ± 10.30	-	-
 <p style="text-align: center;">25</p>	34.40 ± 13.26	14.44 ± 2.09	-	-
 <p style="text-align: center;">26</p>	70.56 ± 4.80	0	-	-
 <p style="text-align: center;">27</p>	93.88 ± 0.58	63.39 ± 1.06	5.31 ± 0.47	3.58

(Continued on following page)

References

Table 10. (Continued) Evaluation of the inhibitory activity of the novel indolinone derivatives. Adapted by Princiotta S. et al.⁶⁵

Compound	% Inhibition		IC ₅₀ * (μM)	K _i ** (μM)
	100 μM	10 μM		
 <p style="text-align: center;">28</p>	87.71 ± 3.04	12.83 ± 1.32	-	-
 <p style="text-align: center;">29</p>	76.82 ± 9.22	53.85 ± 26.57	>10	-
 <p style="text-align: center;">30</p>	96.08 ± 1.42	41.20 ± 1.12	9.14 ± 1.74	6.17

*IC₅₀ value represents concentration of inhibitor needed to reach 50% inhibition. **K_i is the affinity of the inhibitor to the enzyme. Compounds tested were assumed to act as fully ATP-competitive inhibitors.

In addition, cytotoxicity of chosen 5 molecules was also evaluated on the breast cancer MCF-7 cell line, using the Dasatinib as the reference compound (**Table**). As reported by Princiotta S. et al,⁶⁵ thiazole 29 was able to exert an activity comparable to Dasatinib, 23 and 30 derivatives showed significant, but not optimal, cytotoxic activity. However, the cell-free assay had identified 14 as the best inhibitor, despite it had no cytotoxic activity⁶⁵

Table 11. Cytotoxicity evaluation of selected compounds on human MCF-7 breast. Adapted by Princiotta et al.⁶⁵

Compound	MCF-7 (IC ₅₀ , μM)
Dasatinib	27±1
13	>100
14	>100
23	85±2
29	33±2
30	68±2

9.2 Preliminary validation protocol

To propose a reasonable binding mode of 3-(hetero)arylideneindolin-2-one derivatives, obtained from 10, all molecules were docked into c-Src binding site and evaluated using the Standard Precision (SP) mode of Glide.

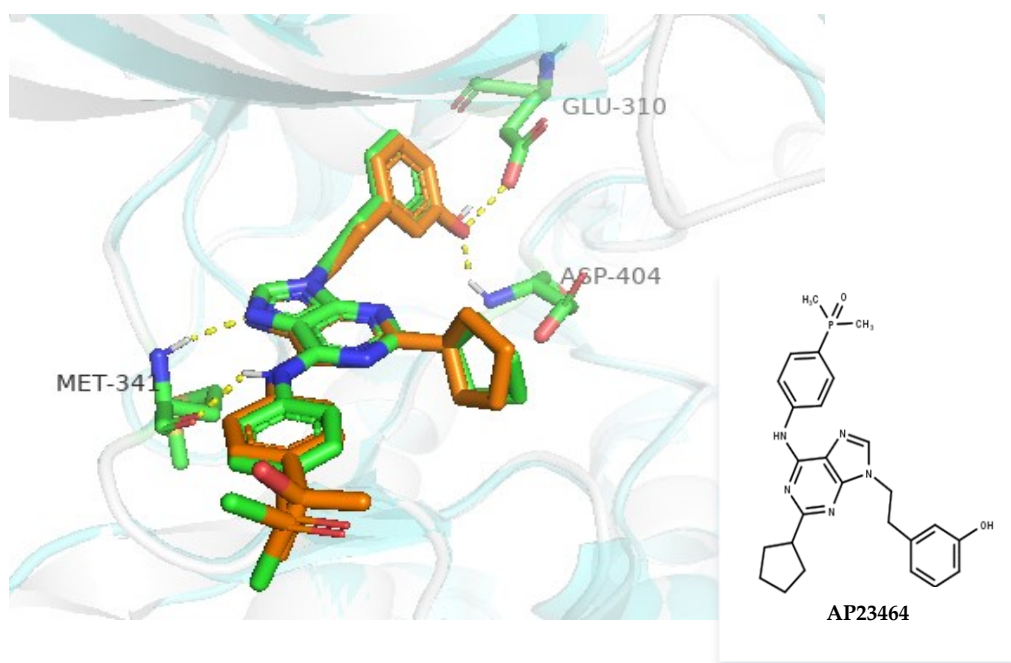


Figure 19. Re-docking of native ligand AP23464 in complex with Src.

The three-dimensional structure of c-Src in complex with 3-[2-(2-cyclopentyl)-6-[[4-

(dimethyl-phosphoryl)phenyl]amino}-9H-purin-9-yl)etil]phenol (AP23464), has been imported from PDB (ID: 2BDJ, resolution = 2.50 Å, IC₅₀ = 0,45 nM). In order to appreciate the reliability of the docking procedure, the native ligand AP23464 was re-docked within the binding pocket (Gscore = -12.110 kcal/mol). The obtained RMSD value = 0.903 Å, validated the docking procedure. The resulting docking pose replicates the interactions observed in literature, showing the two hydrogen bonds between aniline NH and N7 of AP23464 with carbonyl of the Met341 backbone and the amide nitrogen, respectively. The 3-hydroxyphenylethyl in N9 on the purine scaffold, through the hydroxyl group forms a hydrogen bond with the NH of the backbone of Asp404 and the carboxyl of the side chain Glu310 (**Figure 19**).

9.3 Docking studies

Since the new indolinones have been evaluated as *E/Z* isomers in bioassays, isomerization has also been considered during calculations. For each structure, the docking scores highlighted that the isomer E has been favored to the isomer Z. In general, the derivatives exhibited a similar interaction pattern, where the acceptor-donor hydrogen bonding motif, corresponding to the N9 group and aniline NH of AP23464, was emulated by the lactam moiety of the indolinone derivatives.

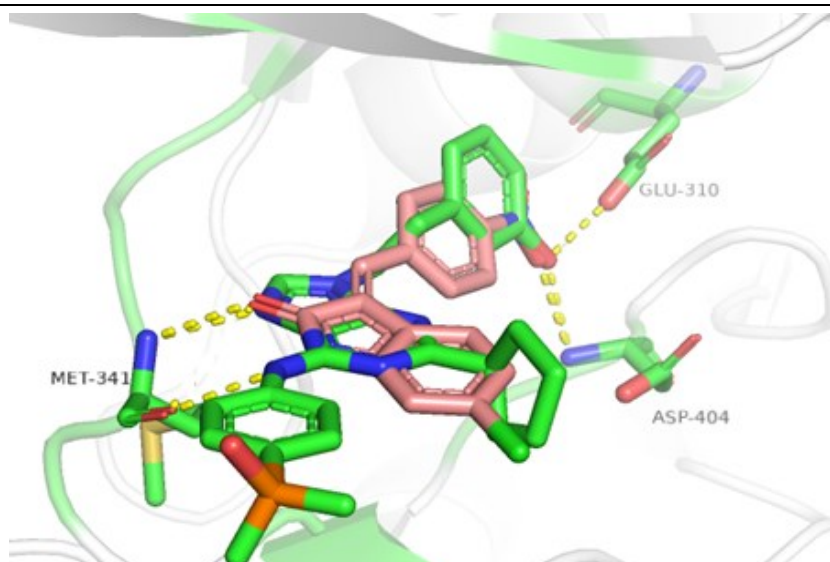


Figure 20. Comparison of AP23464 (green) and 13 (light pink) binding modes.

Through this system, the derivatives bind the fundamental residue Met341. In addition, the chlorophenyl portion and the cyclopentyl ring of AP23464 were superimposed. Lastly, the aryl substituent attached to C3 could imitate the phenolic part of the native ligand. An example of this behavior is given by 13, (Gscore = -7.980 kcal/mol), which maintained the two classical interactions with hinge region essential residue Met341 (**Figure 19**). In addition, the compound formed a hydrogen bond interaction with the Asp404 mediated by nitro group, mimicking phenolic moiety of AP23464. As seen in biological results, 13 shows a single-digit micromolar IC_{50} and K_i , highlighting the role of the nitro substituent to improve the stability of the complex. As shown in **Figure 20**, compound 30 also assumed the same binding mode, maintaining essential interactions with Met341. However, through the NH of peripheral indole ring, 30 forms a hydrogen bond interaction with the side chain of Glu310, confirming its potential role as a hydrogen-bond donor. On the other hand, the biological activity detected is lower than expected from the computational study, 30 being the best compound in terms of score (Gscore value = -8.761 kcal/mol). Probably, the steric encumbrance and rigidity of the indole substituent within the side pocket, influenced the conformational freedom of the compound, leading to lower IC_{50} and K_i values.

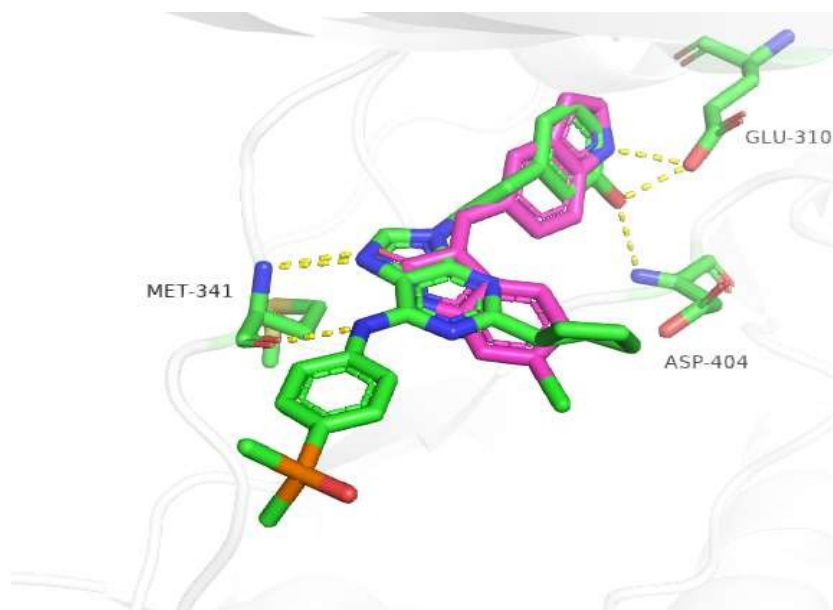


Figure 21. Comparison of AP23464 (green) and 30 (purple) binding modes.

Instead, the computational study does not allow a clear interpretation of the binding affinity for compound 14 (Gscore = -7.638 kcal/mol). Therefore, except for the canonical contacts with Met341, no other convenient interactions have been observed that explain the biological results (**Figure 22**). In fact, the only interactions observed are hydrophobic and mediated by the peripheral of dimethylanilino portion with the Ala403 (**Figure 22**).

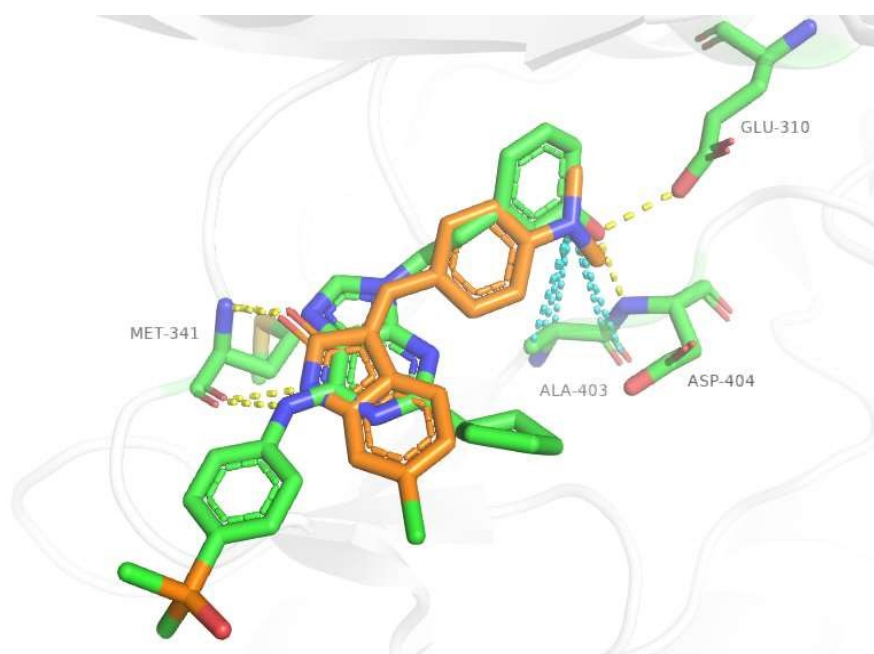


Figure 22. Comparison of AP23464 (green) and compound 14 (orange) binding modes.

Growing the volume and the length of the phenyl ring, pending from the central core, prevented this molecular moiety from accessing the hydrophobic side pocket, unlike the phenyl ring of AP23464. For example, 27 established two hydrogen bonds with Cys277 and Phe278, mediated by the carboxyl function. At the same time, it maintained the classic interactions with Met341. Conversely, the corresponding methyl ester 24 could not establish the anchoring points necessary for interaction with the active site, due to the steric clashes generated by the methyl substituent, that led to a different binding mode.

It is therefore possible to affirm that:

- the biological activity observed in this series of compounds is probably due to the isomer E;
- the conformation adopted by the compounds always appears to have as a common element the HBA-HBD interaction of the indolinone moiety and the Met341 residue of the hinge region;
- insertion of the nitro group in para to the phenyl ring leads to improved activity compared to the parent compound 10;
- the addition of heterocyclic substituents, such as indole, did not lead to an improvement in activity;
- the addition of the substituent dimethylamino significantly improves activity, but computational studies have not been able to determine the mechanisms that justify the values found.

9.4 Results of ADME analysis

Physicochemical and pharmacokinetic parameters, like partition coefficient (QPLogP_{o/w}), solubility (QPlogS) and cell permeation (QPPCaco), were predicted with QikProp module of Schrödinger and are shown in **Table** . Although lower than those of AP23464, the calculated properties revealed that QPlogP_{o/w} values are in the range 2.5-3.6. In addition,

the QPLogS value for inhibitors is around the mean value of -4.3, unlike that of the reference compound.

Table 12. ADME properties predicted by QikProp of best four compounds and native ligand AP23464.

Compound	QPLogPo/w	QPLogS	QPPCaco	Lipinski RO5
AP23464	4.193	-7.142	562.138	0
13	2.465	-4.036	177.260	0
14	3.586	-4.719	1410.192	0
29	2.812	-4.256	843.217	0
30	3.390	-4.642	801.139	0

In terms of permeability, predicted values are mostly higher than recommended (>500 great). Therefore, the computational prediction provided for 14 does not agree with cytotoxic tests, where the compound is ineffective ($IC_{50} > 100 \mu\text{m}$). Moreover, although within an acceptable permeability range, 13 has the lowest value in the series. This result, due to the nitro group, allows to find a probable justification of its ineffectiveness in cytotoxic tests ($IC_{50} > 100 \mu\text{m}$). Finally, none of the inhibitors violates Lipinski's rule of 5.

Chapter 10

Conclusions

A small internal library of compounds has been examined in order to find new c-Src inhibitors. Biological tests confirmed that the best compounds shared the indolinone core, and compound 10 (K_i 3.8 μM) was used as a reference to further optimization studies. The addition of several substituents on the benzylidene moiety at C3 of the indolinone scaffold of 10, led to the synthesis of 3-(hetero)arylideneindolin-2-ones. All compounds have been subjected to NMR analysis, showing the presence of isomeric mixtures (9:1 to 7:3 E/Z isomers). Most of them are able to exert an inhibitory activity against Src >70% at 100 μM , however, only some of them also showed relevant activity at 10 μM . Later, the best compounds were analyzed towards the cell line of human breast cancer MCF-7, exhibiting limited antiproliferative activity. Among these, 29 showed cytotoxicity comparable to Dasatinib, even if no activity toward Src was found, suggesting possible off-target effects. In contrast, 13, proved ineffective in the cellular assay (IC_{50} >100 μM), despite having the best inhibition profile in the enzymatic assay. This result contrasts with the value found by the Qikprop prediction, which showed high permeability for the compound. This suggests there are other intracellular mechanisms to affect the activity of the compound.

Computational study has also allowed to hypothesize that the biological activity is probably due to the isomer E, as well as configuring a common binding mode for the analyzed compounds. In fact, post docking analysis has shown the interactions model HBA-HBD with the fundamental residue of the hinge region, Met341, mediated by the lactam moiety of the indolinone core, constantly repeated. In addition, the aryl portion attached to C3 may mimic the phenolic part of the native ligand by fitting into the side pocket. The computational result is in accordance with the biological data of compound 13, showing the importance of the nitro substituent in determining activity in the sub micromolar range. In

addition, ADME analysis allows to understand how the same group affects membrane permeability and, consequently, the result in cell.

On the contrary, computational studies do not allow us to justify the behavior of 14, characterized by a dimethylanilino substituent, which does not establish further useful interactions, although it is the best in enzymatic assays. Furthermore, compound 25, through the NH of the peripheral indole portion, forms an additional hydrogen bond with Glu310, providing discrete IC_{50} and K_i values. However, the value found does not lead to an improvement in activity, allowing bulky substituents, such as indole, to be excluded for further research. Finally, the increase in length brought about in some substituents was not found to be favorable for the interactions with the lateral pocket of the enzyme. Therefore, further studies will be necessary to improve the pharmacological profile of these derivatives. The results of molecular docking studies, which were performed to propose a binding mode within the c-Src binding site, will guide and integrate future work.

List Of Publications

Princiotto, S., Musso, L., Manetti, F., Marcellini, V., Maga, G., Crespan, E., Perini, C., Zaffaroni, N., Beretta, G.L. and Dallavalle, S., 2022. Synthesis and biological activity evaluation of 3-(hetero) arylideneindolin-2-ones as potential c-Src inhibitors. *Journal of enzyme inhibition and medicinal chemistry*, 37(1), pp. 2382-2394.

Posters for International Conferences/Workshops

Marcellini, V., Perini C., Crespan, E. and Manetti, F. Exploiting synergy in molecular targeted anticancer chemotherapy: Virtual screening approaches to identify novel molecules targeting convergent metabolic pathways in cancer. *European School of Medicinal Chemistry ESMEC, 2022, 3-7 July, Urbino.*

References

- (1) Malki, A.; Elruz, R. A.; Gupta, I.; Allouch, A.; Vranic, S.; Al Moustafa, A. E. Molecular Mechanisms of Colon Cancer Progression and Metastasis: Recent Insights and Advancements. *Int. J. Mol. Sci.* **2021**, *22* (1), 1–24. <https://doi.org/10.3390/ijms22010130>.
- (2) Tessier, M.; Woodgett, J. R. Serum and Glucocorticoid-Regulated Protein Kinases: Variations on a Theme. *J. Cell. Biochem.* **2006**, *98* (6), 1391–1407. <https://doi.org/10.1002/jcb.20894>.
- (3) Sang, Y.; Kong, P.; Zhang, S.; Zhang, L.; Cao, Y.; Duan, X.; Sun, T.; Tao, Z.; Liu, W. SGK1 in Human Cancer: Emerging Roles and Mechanisms. *Front. Oncol.* **2021**, *10* (January), 1–17. <https://doi.org/10.3389/fonc.2020.608722>.
- (4) Mohammad, T.; Siddiqui, S.; Shamsi, A.; Alajmi, M. F.; Hussain, A.; Islam, A.; Ahmad, F.; Hassan, M. I. Virtual Screening Approach to Identify High-Affinity Inhibitors of Serum and Glucocorticoid-Regulated Kinase 1 among Bioactive Natural Products: Combined Molecular Docking and Simulation Studies. *Molecules* **2020**, *25* (4). <https://doi.org/10.3390/molecules25040823>.
- (5) Akhoun, B. A.; Gandhi, N. S.; Pandey, R. Computational Insights into the Active Structure of SGK1 and Its Implication for Ligand Design. *Biochimie* **2019**, *165*, 57–66. <https://doi.org/10.1016/j.biochi.2019.07.007>.
- (6) Jang, H.; Park, Y.; Jang, J. Serum and Glucocorticoid-Regulated Kinase 1: Structure, Biological Functions, and Its Inhibitors. *Front. Pharmacol.* **2022**, *13* (November), 1–17. <https://doi.org/10.3389/fphar.2022.1036844>.
- (7) Zhao, B.; Lehr, R.; Smallwood, A. M.; Ho, T. F.; Maley, K.; Randall, T.; Head, M. S.; Koretke, K. K.; Schnackenberg, C. G. Crystal Structure of the Kinase Domain of Serum and Glucocorticoid-Regulated Kinase 1 in Complex with AMP-PNP. *Protein Sci.* **2007**, *16* (12), 2761–2769. <https://doi.org/10.1110/ps.073161707>.
- (8) Xiao, X.; Yang, Y.; Mao, B.; Yan Cheng, C.; Ni, Y. Emerging Role for SRC Family

-
- Kinases in Junction Dynamics during Spermatogenesis. *Reproduction* **2019**, *157* (3), R85–R94. <https://doi.org/10.1530/REP-18-0440>.
- (9) Cicenás, J.; Meskinyte-Kausiliene, E.; Jukna, V.; Rimkus, A.; Simkus, J.; Soderholm, D. SGK1 in Cancer: Biomarker and Drug Target. *Cancers (Basel)*. **2022**, *14* (10), 1–10. <https://doi.org/10.3390/cancers14102385>.
- (10) Beck, J. T.; Ismail, A.; Tolomeo, C. Targeting the Phosphatidylinositol 3-Kinase (PI3K)/AKT/Mammalian Target of Rapamycin (MTOR) Pathway: An Emerging Treatment Strategy for Squamous Cell Lung Carcinoma. *Cancer Treat. Rev.* **2014**, *40* (8), 980–989. <https://doi.org/10.1016/j.ctrv.2014.06.006>.
- (11) Ackermann, T. F.; Boini, K. M.; Beier, N.; Scholz, W.; Fuchß, T.; Lang, F. Cellular Physiology Biochemistry and Biochemistr y EMD638683 , a Novel SGK Inhibitor with Antihyper- Tensive Potency. *Cell. Physiol. Biochem.* **2011**, 137–146.
- (12) Zhang, H.; Shen, C.; Zhang, H. R.; Qi, H. Z.; Hu, M. L.; Luo, Q. Q. Identification of Novel Inhibitors Targeting SGK1 via Ensemble-Based Virtual Screening Method, Biological Evaluation and Molecular Dynamics Simulation. *Int. J. Mol. Sci.* **2022**, *23* (15), 1–21. <https://doi.org/10.3390/ijms23158635>.
- (13) Liang, X.; Lan, C.; Zhou, J.; Fu, W.; Long, X.; An, Y.; Jiao, G.; Wang, K.; Li, Y.; Xu, J.; Huang, Q.; Xu, B.; Xiao, J. Development of a New Analog of SGK1 Inhibitor and Its Evaluation as a Therapeutic Molecule of Colorectal Cancer. *J. Cancer* **2017**, *8* (12), 2256–2262. <https://doi.org/10.7150/jca.19566>.
- (14) Talarico, C.; Dattilo, V.; Lucia D’Antona; Barone, A.; Amodio, N.; Belviso, S.; Musumeci, F.; Abbruzzese, C.; Bianco, C.; Trapasso, F.; Schenone, S.; Alcaro, S.; Ortuso, F.; Florio, T.; Paggi, M. G.; Perrotti, N.; Amato, R. SI113, a SGK1 Inhibitor, Potentiates the Effects of Radiotherapy, Modulates the Response to Oxidative Stress and Induces Cytotoxic Autophagy in Human Glioblastoma Multiforme Cells. *Oncotarget* **2016**, *7* (13), 15868–15884. <https://doi.org/10.18632/oncotarget.7520>.
- (15) Halland, N.; Schmidt, F.; Weiss, T.; Saas, J.; Li, Z.; Czech, J.; Dreyer, M.; Hofmeister, A.; Mertsch, K.; Dietz, U.; Stru, C.; Nazare, M. Sano Fi R&D, Industriepark Ho “Chst Building G838, D-65926 Frankfurt Am Main, Germany Leibniz-Institut Fu “r

-
- Molekulare Pharmakologie (FMP), Robert-Rössle-Straße 10, 13125 Berlin-Buch, Germany. **2015**, 10–15.
- (16) Kurisu, G. Fifty Years of Protein Data Bank in the Journal of Biochemistry. *J. Biochem.* **2022**, *171* (1), 3–11. <https://doi.org/10.1093/jb/mvab133>.
- (17) Burley, S. K.; Berman, H. M.; Kleywegt, G. J.; Markley, J. L. HHS Public Access. **2018**, 627–641. <https://doi.org/10.1007/978-1-4939-7000-1>.
- (18) <https://www ww p d b . o r g / d o c u m e n t a t i o n / f i l e - f o r m a t - c o n t e n t / f o r m a t 3 3 / s e c t 9 . h t m l>.
- (19) Zardecki, C.; Dutta, S.; Goodsell, D. S.; Lowe, R.; Voigt, M.; Burley, S. K. PDB-101: Educational Resources Supporting Molecular Explorations through Biology and Medicine. *Protein Sci.* **2022**, *31* (1), 129–140. <https://doi.org/10.1002/pro.4200>.
- (20) 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* (3), 1528–1539. <https://doi.org/10.1021/acs.jcim.9b00843>.
- (21) Miyaguchi, I.; Sato, M.; Kashima, A.; Nakagawa, H.; Kokabu, Y.; Ma, B.; Matsumoto, S.; Tokuhisa, A.; Ohta, M.; Ikeguchi, M. Machine Learning to Estimate the Local Quality of Protein Crystal Structures. *Sci. Rep.* **2021**, *11* (1), 1–13. <https://doi.org/10.1038/s41598-021-02948-y>.
- (22) Blow, D. M. Rearrangement of Cruickshank's Formulae for the Diffraction-Component Precision Index. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2002**, *58* (5), 792–797. <https://doi.org/10.1107/S0907444902003931>.
- (23) Bafna, D.; Ban, F.; Rennie, P. S.; Singh, K.; Cherkasov, A. Computer-Aided Ligand Discovery for Estrogen Receptor Alpha. *Int. J. Mol. Sci.* **2020**, *21* (12), 1–49. <https://doi.org/10.3390/ijms21124193>.
- (24) Thesnaar, L.; Bezuidenhout, J. J.; Petzer, A.; Petzer, J. P.; Cloete, T. T. Methylene Blue Analogues: In Vitro Antimicrobial Minimum Inhibitory Concentrations and in Silico Pharmacophore Modelling. *Eur. J. Pharm. Sci.* **2021**, *157*, 105603. <https://doi.org/10.1016/j.ejps.2020.105603>.
- (25) 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**, *10* (4), 1–20.
<https://doi.org/10.1002/wcms.1468>.
- (26) Giordano, D.; Biancaniello, C.; Argenio, M. A.; Facchiano, A. Drug Design by Pharmacophore and Virtual Screening Approach. *Pharmaceuticals* **2022**, *15* (5), 1–16.
<https://doi.org/10.3390/ph15050646>.
- (27) Bhuvaneshwari, S.; Sankaranarayanan, K. Identification of Potential CRAC Channel Inhibitors: Pharmacophore Mapping, 3D-QSAR Modelling, and Molecular Docking Approach. *SAR QSAR Environ. Res.* **2019**, *30* (2), 81–108.
<https://doi.org/10.1080/1062936X.2019.1566172>.
- (28) Rilascio S - Schrödinger, LLC, New York, NY, 2019. Ligprep.
- (29) KB, S.; Kumari, A.; Shetty, D.; Fernandes, E.; DV, C.; Jays, J.; Murahari, M. Structure Based Pharmacophore Modelling Approach for the Design of Azaindole Derivatives as DprE1 Inhibitors for Tuberculosis. *J. Mol. Graph. Model.* **2020**, *101*, 107718.
<https://doi.org/10.1016/j.jmgm.2020.107718>.
- (30) Torres, P. H. M.; Sodero, A. C. R.; Jofily, P.; Silva-Jr, F. P. Key Topics in Molecular Docking for Drug Design. *Int. J. Mol. Sci.* **2019**, *20* (18), 1–29.
<https://doi.org/10.3390/ijms20184574>.
- (31) Stanzione, F.; Giangreco, I.; Cole, J. C. *Use of Molecular Docking Computational Tools in Drug Discovery*, 1st ed.; Elsevier B.V., 2021; Vol. 60.
<https://doi.org/10.1016/bs.pmch.2021.01.004>.
- (32) Krupenev, D.; Boyarkin, D.; Iakubovskii, D. Improvement in the Computational Efficiency of a Technique for Assessing the Reliability of Electric Power Systems Based on the Monte Carlo Method. *Reliab. Eng. Syst. Saf.* **2020**, *204* (October 2019), 107171. <https://doi.org/10.1016/j.ress.2020.107171>.
- (33) Sirous, H.; Campiani, G.; Calderone, V.; Brogi, S. Discovery of Novel Hit Compounds as Potential HDAC1 Inhibitors: The Case of Ligand- and Structure-Based Virtual Screening. *Comput. Biol. Med.* **2021**, *137* (May 2021), 104808.
<https://doi.org/10.1016/j.compbimed.2021.104808>.
- (34) Wicaksana, A.; Rachman, T. 濟無No Title No Title No Title. *Angewandte Chemie*

-
- International Edition*, 6(11), 951–952. 2018, pp 10–27.
- (35) D.A. Case, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, T. Kurtzman, T.S. Lee. University of California, San Francisco 2018.
- (36) Kapral, R.; Ciccotti, G. Molecular Dynamics: An Account of Its Evolution. *Theory Appl. Comput. Chem. First Forty Years* **2005**, 425–441. <https://doi.org/10.1016/B978-044451719-7/50059-7>.
- (37) Brüschweiler, R. Efficient RMSD Measures for the Comparison of Two Molecular Ensembles. *Proteins Struct. Funct. Genet.* **2003**, 50 (1), 26–34. <https://doi.org/10.1002/prot.10250>.
- (38) Genheden, S.; Ryde, U. The MM / PBSA and MM / GBSA Methods to Estimate Ligand-Binding Affinities. **2015**, 1–13. <https://doi.org/10.1517/17460441.2015.1032936>.
- (39) Yamada, Y.; Gohda, S.; Abe, K.; Togo, T.; Shimano, N.; Sasaki, T.; Tanaka, H.; Ono, H.; Ohba, T.; Kubo, S.; Ohkubo, T.; Sato, S. Carbon Materials with Controlled Edge Structures. *Carbon N. Y.* **2017**, 122 (October 1995), 694–701. <https://doi.org/10.1016/j.carbon.2017.07.012>.
- (40) Zangouei, A. S.; Barjasteh, A. H.; Rahimi, H. R.; Mojarrad, M.; Moghbeli, M. Role of Tyrosine Kinases in Bladder Cancer Progression: An Overview. *Cell Commun. Signal.* **2020**, 18 (1), 1–14. <https://doi.org/10.1186/s12964-020-00625-7>.
- (41) Sudhesh Dev, S.; Zainal Abidin, S. A.; Farghadani, R.; Othman, I.; Naidu, R. Receptor Tyrosine Kinases and Their Signaling Pathways as Therapeutic Targets of Curcumin in Cancer. *Front. Pharmacol.* **2021**, 12 (November), 1–26. <https://doi.org/10.3389/fphar.2021.772510>.
- (42) Luo, J.; Zou, H.; Guo, Y.; Tong, T.; Ye, L.; Zhu, C.; Deng, L.; Wang, B.; Pan, Y.; Li, P. SRC Kinase-Mediated Signaling Pathways and Targeted Therapies in Breast Cancer. *Breast Cancer Res.* **2022**, 24 (1), 99. <https://doi.org/10.1186/s13058-022-01596-y>.
- (43) Martellucci, S.; Clementi, L.; Sabetta, S.; Mattei, V.; Botta, L.; Angelucci, A. Src Family Kinases as Therapeutic Targets in Advanced Solid Tumors: What We Have

-
- Learned so Far. *Cancers (Basel)*. **2020**, *12* (6). <https://doi.org/10.3390/cancers12061448>.
- (44) Boggon, T. J.; Eck, M. J. Structure and Regulation of Src Family Kinases. *Oncogene* **2004**, *23* (48 REV. ISS. 7), 7918–7927. <https://doi.org/10.1038/sj.onc.1208081>.
- (45) Roskoski, R. Src Protein-Tyrosine Kinase Structure, Mechanism, and Small Molecule Inhibitors This Paper Is Dedicated to the Memory of Prof. Donald F. Steiner (1930-2014) - Advisor, Mentor, and Discoverer of Proinsulin. *Pharmacol. Res.* **2015**, *94*, 9–25. <https://doi.org/10.1016/j.phrs.2015.01.003>.
- (46) Li, Y.; Bao, Y.; Zheng, H.; Qin, Y.; Hua, B. The Nonreceptor Protein Tyrosine Kinase Src Participates in Every Step of Cancer-Induced Bone Pain. *Biomed. Pharmacother.* **2021**, *141* (June), 111822. <https://doi.org/10.1016/j.biopha.2021.111822>.
- (47) Möbitz, H. The ABC of Protein Kinase Conformations. *Biochim. Biophys. Acta - Proteins Proteomics* **2015**, *1854* (10), 1555–1566. <https://doi.org/10.1016/j.bbapap.2015.03.009>.
- (48) Adams, J. A. Activation Loop Phosphorylation and Catalysis in Protein Kinases: Is There Functional Evidence for the Autoinhibitor Model? *Biochemistry* **2003**, *42* (3), 601–607. <https://doi.org/10.1021/bi020617o>.
- (49) Hsu, P. C.; Yang, C. T.; Jablons, D. M.; You, L. The Crosstalk between Src and Hippo/YAP Signaling Pathways in Non-Small Cell Lung Cancer (NSCLC). *Cancers (Basel)*. **2020**, *12* (6), 1–24. <https://doi.org/10.3390/cancers12061361>.
- (50) Belli, S.; Esposito, D.; Servetto, A.; Pesapane, A.; Formisano, L.; Bianco, R. C-Src and EGFR Inhibition in Molecular Cancer Therapy: What Else Can We Improve? *Cancers (Basel)*. **2020**, *12* (6), 1–16. <https://doi.org/10.3390/cancers12061489>.
- (51) Seif, F.; Khoshmirsafa, M.; Aazami, H.; Mohsenzadegan, M.; Sedighi, G.; Bahar, M. The Role of JAK-STAT Signaling Pathway and Its Regulators in the Fate of T Helper Cells. *Cell Commun. Signal.* **2017**, *15* (1), 1–13. <https://doi.org/10.1186/s12964-017-0177-y>.
- (52) Varkaris, A.; Katsiampoura, A. D.; Araujo, J. C.; Gallick, G. E.; Corn, P. G. Src Signaling Pathways in Prostate Cancer. *Cancer Metastasis Rev.* **2014**, *33* (2–3), 595–606. <https://doi.org/10.1007/s10555-013-9481-1>.

-
- (53) Iqbal, N.; Iqbal, N. Imatinib: A Breakthrough of Targeted Therapy in Cancer. *Chemother. Res. Pract.* **2014**, *2014*, 1–9. <https://doi.org/10.1155/2014/357027>.
- (54) Lee, P. Y.; Yeoh, Y.; Low, T. Y. A Recent Update on Small-Molecule Kinase Inhibitors for Targeted Cancer Therapy and Their Therapeutic Insights from Mass Spectrometry-Based Proteomic Analysis. *FEBS J.* **2022**, 1–20. <https://doi.org/10.1111/febs.16442>.
- (55) Roskoski, R. Properties of FDA-Approved Small Molecule Protein Kinase Inhibitors. *Pharmacol. Res.* **2019**, *144*, 19–50. <https://doi.org/10.1016/j.phrs.2019.03.006>.
- (56) Ayala-Aguilera, C. C.; Valero, T.; Lorente-Macías, Á.; Baillache, D. J.; Croke, S.; Unciti-Broceta, A. Small Molecule Kinase Inhibitor Drugs (1995–2021): Medical Indication, Pharmacology, and Synthesis. *J. Med. Chem.* **2022**, *65* (2), 1047–1131. <https://doi.org/10.1021/acs.jmedchem.1c00963>.
- (57) Wing Tung Ho, V.; Yue Tan, H.; Wang, N.; Feng, Y. Cancer Management by Tyrosine Kinase Inhibitors: Efficacy, Limitation, and Future Strategies. *Tyrosine Kinases as Druggable Targets Cancer* **2019**. <https://doi.org/10.5772/intechopen.82513>.
- (58) Arora, A.; Scholar, E. M. Role of Tyrosine Kinase Inhibitors in Cancer Therapy. *J. Pharmacol. Exp. Ther.* **2005**, *315* (3), 971–979. <https://doi.org/10.1124/jpet.105.084145>.
- (59) Jha, V.; Macchia, M.; Tuccinardi, T.; Poli, G. Three-Dimensional Interactions Analysis of the Anticancer Target c-Src Kinase with Its Inhibitors. *Cancers (Basel)*. **2020**, *12* (8), 1–36. <https://doi.org/10.3390/cancers12082327>.
- (60) Rusconi, F.; Piazza, R.; Vagge, E.; Gambacorti-Passerini, C. Bosutinib: A Review of Preclinical and Clinical Studies in Chronic Myelogenous Leukemia. *Expert Opin. Pharmacother.* **2014**, *15* (5), 701–710. <https://doi.org/10.1517/14656566.2014.882898>.
- (61) Tan, F. H.; Putoczki, T. L.; Stylli, S. S.; Luwor, R. B. Ponatinib: A Novel Multi-Tyrosine Kinase Inhibitor against Human Malignancies. *Onco. Targets. Ther.* **2019**, *12*, 635–645. <https://doi.org/10.2147/OTT.S189391>.
- (62) Musumeci, F.; Radi, M.; Brullo, C.; Schenone, S. Vascular Endothelial Growth Factor (VEGF) Receptors: Drugs and New Inhibitors. *J. Med. Chem.* **2012**, *55* (24), 10797–10822. <https://doi.org/10.1021/jm301085w>.

- (63) Williams, E.; Bagarova, J.; Kerr, G.; Xia, D. D.; Place, E. S.; Dey, D.; Shen, Y.; Bocobo, G. A.; Mohedas, A. H.; Huang, X.; Sanderson, P. E.; Lee, A.; Zheng, W.; Economides, A. N.; Smith, J. C.; Yu, P. B.; Bullock, A. N. Saracatinib Is an Efficacious Clinical Candidate for Fibrodysplasia Ossificans Progressiva. *JCI Insight* **2021**, *6* (8).
<https://doi.org/10.1172/jci.insight.95042>.
- (64) Yamaoka, T.; Kusumoto, S.; Ando, K.; Ohba, M.; Ohmori, T. Receptor Tyrosine Kinase-Targeted Cancer Therapy. *Int. J. Mol. Sci.* **2018**, *19* (11), 1–35.
<https://doi.org/10.3390/ijms19113491>.
- (65) Princiotta, S.; Musso, L.; Manetti, F.; Marcellini, V.; Maga, G.; Crespan, E.; Perini, C.; Zaffaroni, N.; Beretta, G. L.; Dallavalle, S. Synthesis and Biological Activity Evaluation of 3-(Hetero) Arylideneindolin-2-Ones as Potential c-Src Inhibitors. *J. Enzyme Inhib. Med. Chem.* **2022**, *37* (1), 2382–2394.
<https://doi.org/10.1080/14756366.2022.2117317>.
- (66) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes. *J. Med. Chem.* **2006**, *49* (21), 6177–6196. <https://doi.org/10.1021/jm051256o>.