

Computational Modelling, Synthesis and Biological Evaluation of Guanidinium-like Derivatives as Protein Kinase Inhibitors for the Treatment of Cancer



**A thesis presented to the University of Dublin for the degree of
Doctor of Philosophy**

by

Viola Previtali, M.Sc.

Under the supervision of Prof. Isabel Rozas

School of Chemistry
Trinity College Dublin

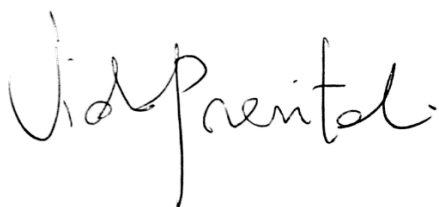
February 2018

Declaration

This work comprises a doctoral thesis submitted for the consideration of Trinity College Dublin.

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work, with due acknowledgement and reference given to the work of others, where appropriate.

I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.



.....

Viola Previtali

February 2018

Trinity College Dublin

Acknowledgements

First, I would like to thank with immense gratitude my supervisor Prof. Isabel Rozas for believing in me since the first day my PhD started. Her support and trust during these four years made me enormously grow both scientifically and personally.

I am also grateful to all the wonderful people in Rozas' group I had the chance to work with. Previous members, Paddy in particular, helped me answering chemistry questions and solving problems and, more importantly, became a dear friend. I would also like to thank the current member of the group (Aaron, Jagdeep and Filippo), who together make a great team. I will definitely miss to work with you. I am immensely grateful to Dr. Cristina Trujillo who patiently introduced me to the computational modelling world. She is not simply an exceptional computational chemist, but a true friend. Thank you also to Dr. Goar Sánchez, whose advices and help helped my quantum world to grow bigger. I will not forget either of you. A word of deep gratitude goes also to Prof. Eric Henon (University of Reims, France), whose expertise in molecular dynamic helped me to explore new computational techniques. Thank you also to Prof. Daniela Zisterer, Prof. James Murray (School of Biochemistry and Immunology, TCD) and Prof. Maria Santos-Martinez (School of Pharmacy, TCD), whose collaboration allowed me to step into the biological world. Thank you also to all the 'biochemistry girls' of the 6th floor, who have always been helping me with routinely cells-problems.

I am thankful to all the administrative and technical staff at Trinity College, in particular, Dr. John O' Brien and Dr. Manuel Ruether for the NMRs and Dr. Martin Feeney and Dr. Gary Hessman for mass spectrometry, who helped me throughout my project.

Last but not least a thankful thought goes to my family, my mum Claudia, my dad Giacomo and my sister Marta. They have always supported me and my dreams in every moment of my life and, especially in these last four years away from home, I have felt them closer than ever.

I am lucky enough to have friends that are always willing to spend some of their time with me during my fast and rare visit home. I have left Italy with dear friends in my heart but I have also found others along the way who have become my second family when so far from home. To all of you, thank you, I will never forget you. A special thank you goes to the best friend of all, Luca. Your unconditional support and love have helped me throughout hard times and have filled of joy the happy ones.

Dedicated to my parents, Claudia and Giacomo and to my sister, Marta.

Abstract

Kinases have emerged as one of the most intensively pursued targets in current medicinal chemistry research, and to date, the FDA has approved 31 small molecule kinase inhibitors (SMKI), with several therapeutic applications, particularly cancer. Previous studies in Rozas' group identified a guanidinium-based lead compound capable of inhibiting B-Raf (kinase frequently altered in the Ras/Raf/MEK/ERK pathway in cancer) through a hypothetical type-III allosteric mechanism. Previous and on-going molecular-modelling studies point towards the idea that this compound forms hydrogen-bond/electrostatic interactions with one of the ATP-phosphates through the mono-substituted guanidinium moiety, thus, positioning the lipophilic (4-Cl-3-CF₃)-Ph group in a hydrophobic pocket of the enzyme. In this research, we have computationally studied the interactions established by this lead compound with (i) isolated ATP (through Density Functional Theory calculations) and (ii) an active phosphorylated ATP-containing B-Raf protein (through Molecular Dynamics simulations and docking), which represents the first model of this type of kinase containing ATP and hence suitable for structure-based design of type-III inhibitors of an active kinase. Moreover, 25 new aromatic derivatives have been synthesized to clearly identify the structural motifs that determine the allosteric inhibition of the Ras/Raf/MEK/ERK pathway and derive the corresponding structure-activity relationships. Several different biochemical assays were conducted to investigate their potential as anticancer agents and protein kinases' inhibitors. Firstly, their cytotoxic effect was assessed in promyelocytic leukaemia cells HL-60, and, for some chosen derivatives, also in breast cancer MCF-7, cervical HeLa, colorectal carcinoma HCT116 and HKH2 and multiple myeloma, H929 and U266B1 cell lines, to clearly establish their therapeutic profile. Secondly, flow-cytometry and western blotting studies were carried out to investigate the apoptosis-inducing activity of these derivatives. Their safety profile was established with assessment of their toxicity against non-tumorigenic epithelial MCF10A and human platelets and their 'druggability' estimated with computer-based prediction of their pharmacokinetics characteristics. The 3-amino-4'-guanidino phenoxy pyridine **66** has distinguished itself among the other compounds for its exclusive high potency against a wide range of cancer cell lines, its ability to induce apoptosis and its low toxicity. Therefore, it now represents a new hit molecule towards the development of type-III allosteric inhibitors of the Ras/Raf/MEK/ERK pathway.

Abbreviations

(COCl)₂: Oxalyl chloride

ADME: Absorption, Distribution, Metabolism and Excretion

ADP: Adenosine diphosphate

AGC: PKA, PKC, PKG

AIM: Atoms in molecules

AnV: Annexin V

Akt: Protein kinase B

ATP: Adenosine triphosphate

BAP: Binding allosteric pocket

BBB: Blood brain barrier

Bcl-2: B-cell lymphoma 2

BCP: Bond critical point

BSSE: Basis set superposition error

BINAP: 2,2'-bis(diphenylphosphino)-1,1'-binaphthyl

CADD: Computer-aided drug design

CAMK: Calcium/calmodulin-dependent protein kinase

CCl₄: Carbon tetrachloride

CH₂Cl₂: Dichloromethane

CHCl₃: Chloroform

CK1: Casein kinase 1

CMGC: CDK, MAPK, GSK3, Cdc2

CNS: Central nervous system

CRC: Colorectal cancer

Cs₂CO₃: caesium carbonate

CuCl₂: Copper (II) chloride

DFT: Density functional theory

DMF: dimethylformamide

DMSO: Dimethyl sulfoxide

EDCI: 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide

EGFR: Epidermal growth factor receptor

ELISA: Enzyme-linked immunosorbent assay

ERK: Extracellular signal-regulated kinase

EtOAc: Ethylacetate

FDA: Food and Drug Administration

GAPDH: Glyceraldehyde 3-phosphate dehydrogenase

GSK3: Glycogen synthase kinase 3

H₂O₂: Hydrogen peroxide

HCl: Hydrochloric acid

HCT116: Human colon cancer cell line

HeLa: Henrietta Lacks

HgCl₂: Mercury (II) chloride

HIA: Human intestinal absorption

HL-60: Human promyelocytic leukemia cells

HPLC: High-performance liquid chromatography

HSQC: Heteronuclear single quantum coherence spectroscopy

IC₅₀: Concentration of an inhibitor where the response (or binding) is reduced by half

IGM: Independent gradient model

IMHB: Intra molecular hydrogen bond

JAK: Janus kinase

K₂CO₃: Potassium carbonate

K₃PO₄: potassium phosphate

LBDD: Ligand-based drug design

LTA: Light transmission aggreometry

MAPK: Mitogen-activated protein kinase

MCF-7: Michigan cancer foundation-7

Mcl-1: Induced myeloid leukemia cell differentiation protein

MEK: Mitogen-activated protein kinase

MeOH: Methanol

MM: Molecular mechanic

MM: Multiple myeloma

mTOR: Mechanistic target of rapamycin

NaH: Sodium hydride

NaOtBu: Sodium tert-butoxide

NBO: Natural bond orbital

NCI: Non covalent interactions

NEt₃: Triethylamine

NMR: Nuclear magnetic resonance

OTf: Triflate or Trifluoromethanesulfonate

PCM: Polarizable continuum model

Pd₂(dba)₃: Tris(dibenzylideneacetone)dipalladium(0)

PDGF: Platelet-derived growth factor

PI: Propidium iodide

PI3K: Phosphoinositide 3-kinase

PKA: protein kinase A

PKI: Protein kinase inhibitors

PPh₃: Triphenylphosphine

PPTS: Pyridinium p-toluenesulfonate

PRP: Platelet rich plasma

PS: Phosphatidylserine

QM: Quantum mechanic

Raf: Rapidly accelerated fibrosarcoma

Rof: Rule of five

RTK: Receptor tyrosine kinase

SAR: Structure activity relationship

SBDD: Structure-based drug design

SMKI: Small molecules kinase inhibitors

SNAR: Nucleophilic aromatic substitution

$\text{SnCl}_2 \cdot 2\text{H}_2\text{O}$: Tin (II) chloride dihydrate

SnCl_4 : Tin (IV) chloride

SOCl_2 : Thionyl chloride

STAT: Signal transducer and activator of transcription

STE: Homologs of yeast Sterile

TCIPA: Tumour cell-induced platelet aggregation

TFA: Trifluoroacetic acid

TFAA: Trifluoroacetic anhydride

$\text{TGF}\alpha/\beta$: Tumour growth factor α/β

THF: Tetrahydrofuran

TKL: Tyrosine kinase-like

TLC: Thin-layer chromatography

TPSA: Topological polar surface area

TR-FRET: Time resolved fluorescence resonance energy transfer

VEGF: Vascular endothelial growth factor

WP: Washed platelets

Xantphos: 4,5-Bis(diphenylphosphino)-9,9-dimethylxanthene

α_2 -AR: alpha-2 adrenergic receptor

Table of Contents

Chapter 1 – Introduction	1
1.1. The burden of cancer: general overview and statistics	1
1.2. Cancer Biology: hints	4
1.2.1. Evolution in cancer and its hallmarks	4
1.2.2. Key pathways in cancer cell signalling	5
1.3. Protein Kinases	8
1.3.1. Protein Kinases and disease	12
1.3.2. Functioning and features of protein kinases	13
1.3.3. Ras/Raf/MEK/ERK kinase pathway and its interconnections	18
1.4. Protein kinase inhibitors	21
1.4.1. Small molecule kinase inhibitors	22
1.4.2. Current challenges in PKI discovery and future directions	27
1.4.3. Protein Kinase inhibition assays	29
1.5. Previous Work Leading Up to the Present Research	33
1.6. References	35
 Chapter 2 – Objectives	 40
2.1. Computational Objectives	41
2.2. Synthetic Objectives	42
2.3. Biological Objectives	44
 Chapter 3 – Computational Modelling	 45
3.1. Computational chemistry in drug discovery	45
3.2. Computational methods	47
3.3. Theoretical studies proposed	50
3.4. Theoretical model of an active form of B-Raf protein kinase	51
3.4.1. Generation of B-Raf systems	53
3.4.2. Analysis of the trajectories obtained	56
3.4.2.1. DFG motif and Mg^{2+} binding loop in the optimised systems	56

3.4.2.2. Analysis of the Activation Loop and Phosphorylation Site	58
3.4.2.3. Analysis of the α C-helix region	61
3.4.2.4. Structure and Activating Interactions of the 2pBRAF_ATP	64
model	
3.5. Molecular Docking Study	69
3.6. Theoretical model of the interaction between phosphates in the ATP molecule	72
and cationic systems	
3.6.1. Arylguanidinium – ATP interactions	74
3.6.1.1. Optimized geometries and interaction energies	74
3.6.1.2. Electron density analysis	75
3.6.2. Arylisouronium – ATP interactions	79
3.6.2.1. Optimized geometries and interaction energies	79
3.6.2.2. Electron density analysis	81
3.7. Conclusions	84
3.8. References	85
 Chapter 4 – Chemical Synthesis	 89
4.1. Introduction	89
4.1.1. Guanidines and their therapeutic applications	89
4.1.2. Synthetic methods for the introduction of the guanidine	90
functionality	
4.1.2.1. Synthesis of acyclic mono-substituted guanidines	91
4.1.2.2. Synthesis of acyclic <i>N,N'</i> -disubstituted guanidines	95
4.2. Synthesis of 3,4'-bis-guanidino phenyloxybenzene derivatives	98
exploring the hydrophobic moiety	
4.2.1. General aspects	98
4.2.2. Synthesis	101
4.3. Synthesis of 3-amino- or 3-amido-4'-guanidino phenyloxybenzene	111
derivatives: exploring the link	
4.3.1. General aspects	111
4.3.2. Synthesis	113
4.3.2.1. Synthesis of arylamino diaromatic mono-guanidines	113
4.3.2.2. Synthesis of arylamido diaromatic mono-guanidines	119

4.4. Synthesis of derivatives with modified cationic moieties: exploring the polar moiety	122
4.4.1. General aspects	122
4.4.2. Synthesis	124
4.4.2.1. Synthesis of 3,4'-bis-guanidino and 3-amino-4'-guanidino phenyloxypyridines	124
4.4.2.2. Synthesis of isouronium derivatives	129
4.4.2.3. Synthesis of the sulfamide derivative	133
4.5. Conclusions	135
4.6. References	137

Chapter 5 – Biochemical evaluation and Physicochemical Properties **143**

5.1. Introduction	143
5.2. Cytotoxicity Evaluation of the Novel Compounds	143
5.2.1. Evaluation of the Cytotoxicity of Compounds in the HL-60 Cell Line	145
5.2.2. Evaluation of the Cytotoxicity of Selected Compounds in the MCF-7 Cell Line	151
5.2.3. Evaluation of the Cytotoxicity of Selected Compounds in the HeLa Cell Line	155
5.2.4. Evaluation of the Cytotoxicity of Selected Compounds in the HKH-2 and HCT116 Cell Lines.	157
5.2.5. Evaluation of the Cytotoxicity of Compounds 1 and 66 in the NCI-H929 and U266B1 Cell Line	161
5.2.6. Flow cytometry	165
5.2.7. Effect of compound 66 on the expression levels of the anti- apoptotic MCL-1 protein in NCI-H929 and U266 cancer cell lines.	168
5.2.8. Thermal Shift analysis in B-Raf ^{V600E} protein	169
5.2.9. Investigation of the Cytotoxicity of compounds 1 and 66 in Healthy Cells and Human Platelets.	171
5.2.9.1. Evaluation of the Cytotoxicity of compound 66 in MCF10A Cell Line	171

5.2.9.2. Investigation of the Effect of compounds 1 and 66 on Platelet Function	172
5.3. Physicochemical Properties	177
5.4. Cancer Cell Cytotoxicity and Physicochemical Properties of compound 66	182
5.5. Conclusions	185
5.6. References	187
Chapter 6 – Conclusions and Future Work	191
6.1. Conclusions	191
6.1.1. Computational Modelling	191
6.1.2. Synthesis	192
6.1.3. Biochemistry and physicochemical properties	193
6.2. Future Work	195
6.2.1. Future Computational Modelling	195
6.2.2. Future Protein Kinase Inhibitors	195
6.2.3. Future Biological Studies of our Compounds	197
6.3. References	198
Chapter 7 – Experimental	199
7.1. Molecular Modelling	199
7.1.1. Modelling of the target and ligand-target interaction study	199
7.1.1.1. Starting B-Raf structures preparation	199
7.1.1.2. Molecular Dynamics simulation	199
7.1.1.3. 1D Free Energy Landscapes (1D-FEL)	200
7.1.1.4. NCI calculations	201
7.1.1.5. Molecular Docking	202
7.1.2. Quantum Chemistry studies	202
7.1.2.1. Optimized geometries and binding energies	202
7.1.2.2. Bonding characteristics and electron density	203
7.2. Synthetic Chemistry	203
7.2.1. Materials and Methods	203
7.2.2. Purity Assessment of Hydrochloride Salts	205

7.2.3. General Procedures	205
7.2.4. Synthesis and Characterisation	209
7.3. Biochemical Procedures	296
7.3.1. Materials	296
7.3.2. General Procedures	296
7.3.2.1. Cell Storage/Cryopreservation	296
7.3.2.2. Use of Haemocytometer	297
7.3.2.3. Preparation of the Drugs' Solutions	297
7.3.3. Growth and Maintenance of the Cell Lines	297
7.3.4. AlamarBlue® Cell Viability Assay	299
7.3.5. Effect on Human Platelets	299
7.3.6. Flow Cytometric Analysis	301
7.3.7. Western Blotting	302
7.3.8. Thermal Melting B-Raf ^{V600E}	303
7.4. References	304

Appendix

Chapter 1

Chapter 1 - Introduction

1.1. The burden of cancer: general overview and statistics

Cancer is a burden upon the society in which we live today. This is even more relevant in a scientific environment where researchers routinely discuss cancer with a pragmatic point of view; however, cancer has a deeper human and societal cost. Behind every diagnosis, genome sequence, therapy and statistic lies a human being and their family, together with their fears and hopes. Today the battle against cancer is not only about successful therapies, instead it starts with prevention strategies, early detection and affordable personalized therapies which have all been developed through the expertise and collaboration of numerous interdisciplinary and translational fields, starting from early stage discovery through to clinical application.

The collection and analysis of cancer statistics is a fundamental tool utilized to both understand and control cancer. This data not only depicts the affliction of the disease but can also indicate cause-effect relationships and the impact of socioeconomic differences which can inform the assignment of priorities for cancer control programs. Epidemiological data on incidence of cancer and related mortality shows wide variation across worldwide regions. The data presented in this thesis are taken from the latest edition of World Cancer Report (WHO, 2012)¹ and GLOBOCAN (2012)², the two most current and complete reports published on cancer to date.

Cancer is a major cause of morbidity and mortality, with approximately 14 million new cases and 8 million cancer-related deaths in 2012, affecting populations in all countries and all regions, but with marked differences worldwide. Among men (Figure 1.1.1), the five most common sites of cancer diagnosed in 2012 were lung (16.7% of the total), prostate (15.0%), colorectum (10.0%), stomach (8.5%), and liver (7.5%). Among women (Figure 1.1.1), the five most common incident sites of cancer were breast (25.2% of the total), colorectum (9.2%), lung (8.7%), cervix (7.9%), and stomach (4.8%).

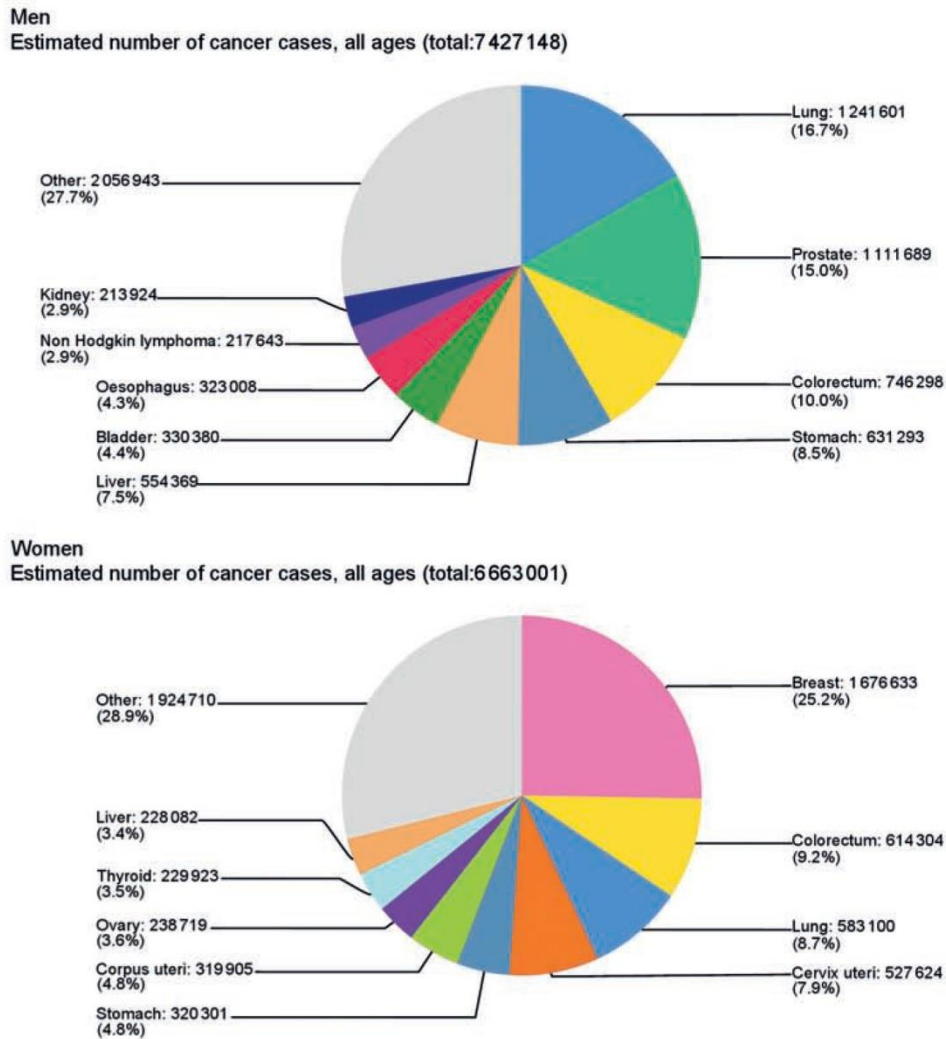


Figure 1.1.1. Estimated world cancer incidence proportions by major sites, in men, and in women, 2012. Adapted from reference 1.¹

Among men (Figure 1.1.2), lung cancer has the highest incidence and mortality rates (23.6%), the mortality rate for prostate cancer is considerably lower than that of lung cancer (6.6%), despite its high rate of incidence (15%). Stomach, liver, colorectum and oesophageal cancers are three of the other major cancers in men that show relatively poor survival and hence high mortality rates.

Among women (Figure 1.1.2), breast cancer has the highest mortality rate (14.7%). Together with its high incidence rate, breast cancer is the most common and aggressive cancer in women, despite the improvement of early screening tests and biomarkers. The second highest mortality is observed for lung cancer (13.8%), and then colorectal (9%), cervical (7.5%) and stomach cancer (7.2%).

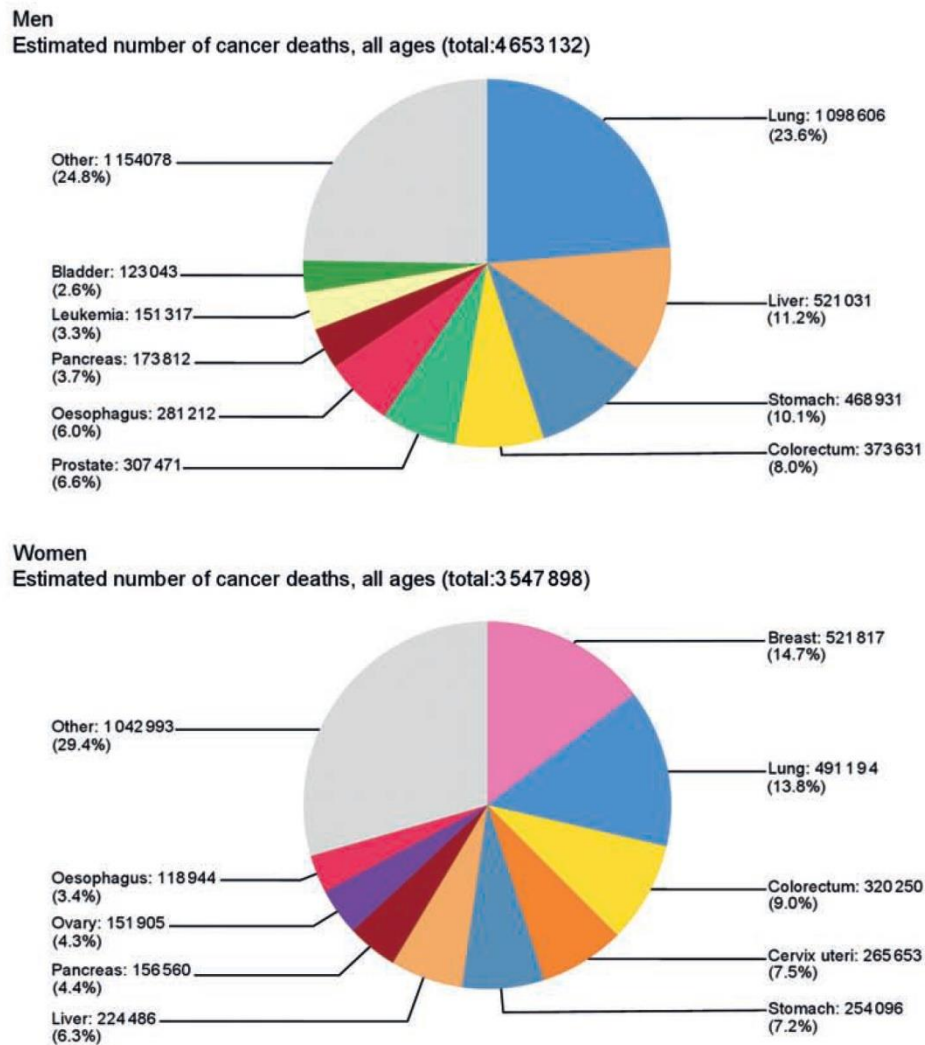


Figure 1.1.2. Estimated world cancer mortality proportions by major sites, in men, and in women, 2012. Adapted from reference 1.¹

The ‘Other’ category (Figure 1.1.1. and Figure 1.1.2.) comprises a wide variety of cancers among which leukemia, multiple myeloma and melanoma of the skin must be highlighted. According to a cancer report published by Siegel et al. (2016) concerning the United States of America in the year 2016, leukemia was responsible for 4% of deaths (men and women) and melanoma of the skin had an incidence of 6% in men and 3% in women.³

Further details of cancer incidence and mortality may be considered fundamental, but a full description is outside the scope of this thesis. As a conclusion, it is estimated that

by 2035 the annual number of new cancers across all ages will grow by 70% compared with 2012 estimates.⁴ These figures highlight the massive prevalence and importance of cancer, how it affects all humankind and brings the urgent need for successful and alternative therapeutic strategies into sharp focus.

1.2. Cancer Biology: hints

Cancer statistics help to record and examine health changes over time and provide clues as to the environmental, lifestyle and behavioural risks responsible for cancer etiology.

Chronic infections play a major role in common cancers in parts of Africa and Asia but are less common in Europe and North America. In highly developed and high-income countries, the diet exerts an influence on cancer advancement due to energy-rich food intake and associated obesity linked to low levels of physical activity. Exposure to tobacco, hazardous workplaces, alcohol intake and excessive sun are also reflected in the incidence of particular cancer types.

1.2.1. Evolution in cancer and its hallmarks

Behind each cancer lie cellular genomic mutations which drive the healthy cell through a tumorigenesis process that eventually leads to cancer-related deaths. The types and subtypes of cancer and related cancer-causing mutations differ considerably across the world populations as a result of genetic diversity within our communities. These mutations can vary in number from a few (10–20) to, more commonly, hundreds or thousands. The great majority of these mutational are relatively harmless ‘passengers’, but few of them are more relevant and deleterious ‘drivers’. In 1976, Peter Nowell defined cancer as “an evolutionary process that is driven by stepwise, somatic-cell mutations with sequential, subclonal selection” (Figure 1.2.1).⁵

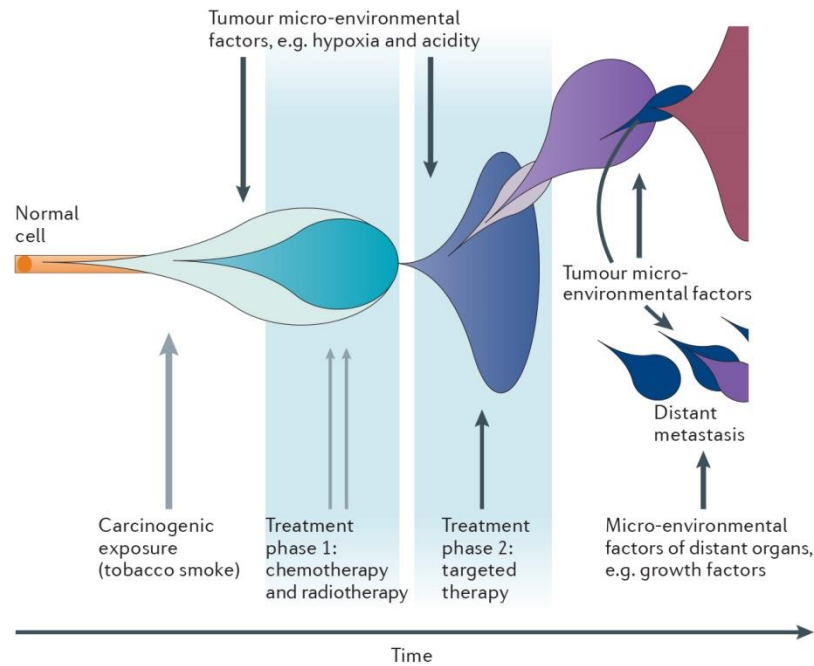


Figure 1.2.1. The evolution of clonal populations and the role of the surrounding environment in their evolutionary adaptation. Adapted from Yates *et al.*, 2012.⁶

It is now well established that cancer evolves by repetitive processes of clonal expansion, genetic diversification and clonal selection. External forces can act on the cancer genome to generate heterogeneity and influence the subclonal structure. Therapeutic intervention may temporarily destroy cancer clones, but it can also provide a potent selective pressure for the expansion of resistant variants or subclones (Figure 1.2.1). Moreover, there is now growing evidence that cancer cells behave as communities and increasing attention is directed toward the cooperative behaviour of subclones that can influence disease progression.⁷ The Darwinian character of cancer is the primary reason for therapeutic failure, but it may also hold the key to more effective control.⁸

1.2.2. Key pathways in cancer cell signalling

Hanahan and Weinberg in 2000 suggested, in an outstanding publication, that cancer cell genotypes result from six essential alterations in cell physiology that must be acquired, known as the *hallmarks of cancer*.⁹ The first three acquired capabilities are:

(i) self-sufficiency in growth signals, (ii) insensitivity to growth-inhibitory (antigrowth) signals and (iii) evasion of programmed cell death (apoptosis).

The mechanism of *acquired self-sufficiency in growth signals* derives from alterations of components in the downstream signalling circuits of the cell which receive signals generated by trans-membrane receptors at the cellular surface through activation by growth factors. Cancer cells can achieve such physiological states through different mechanisms; for example, they can acquire the ability to synthesize growth factors like tumour growth factor α (TGF α) to which they are responsive *via* a positive feedback signalling loop. Moreover, transmembrane receptor overexpression (e.g., RTK receptor tyrosine kinase) or extracellular matrix receptors (integrins) may enable the cell to become hyper-responsive to their activation (Figure 1.2.2).⁹ The Ras/Raf/MEK/ERK (also known as MAPK/ERK) cascade plays a central role here, receiving, processing and enhancing the signals emitted by growth factors receptors and integrins. This cascade is the central hub of the integrated signalling circuit of the cell, from where other effector pathways can radiate and enable extracellular signals to elicit changes in gene expression through cross-talking connections (Figure 1.2.2).⁹ Normal cells control their proliferation through different mechanisms of antigrowth signals, but cancer cells must evade these cascades to reach another acquired capability, the *insensitivity to antigrowth signals*. Most antiproliferative signals are controlled not only by anti-growth factors (e.g. TGF β), but also by alterations to their receptors which, in a hypophosphorylated state, sequester and alter the function of E2F transcription factor which controls cell proliferation and cell cycle (Figure 1.2.2).⁹

The destiny of cancer cells is not only controlled by an unbalanced equilibrium between antigrowth signals and cell proliferation, but also through an acquired *resistance toward apoptosis*, another hallmark of cancer cell. In a healthy cell, intra- and extracellular sensors can activate this death pathway in response to abnormalities, like DNA damage. However, in cancer cells resistance to apoptosis is most commonly acquired by mutation and subsequent loss of a proapoptotic regulator of the *p53* tumour suppressor gene, a key sensor of DNA damage, which can induce the apoptotic effector cascade (Figure 1.2.2). Many of the signals converge on the mitochondria, which release cytochrome C in response to proapoptotic signals and this works in concert with caspase 9 to control cell death. Members of the Bcl-2 family of proteins, whose members have either proapoptotic (Bax, Bak, Bid, Bim and BAD) or antiapoptotic

(Bcl-2, Bcl-XL, or Mcl-1) function, are fundamental to the control of apoptosis and its attrition in cancer.⁹

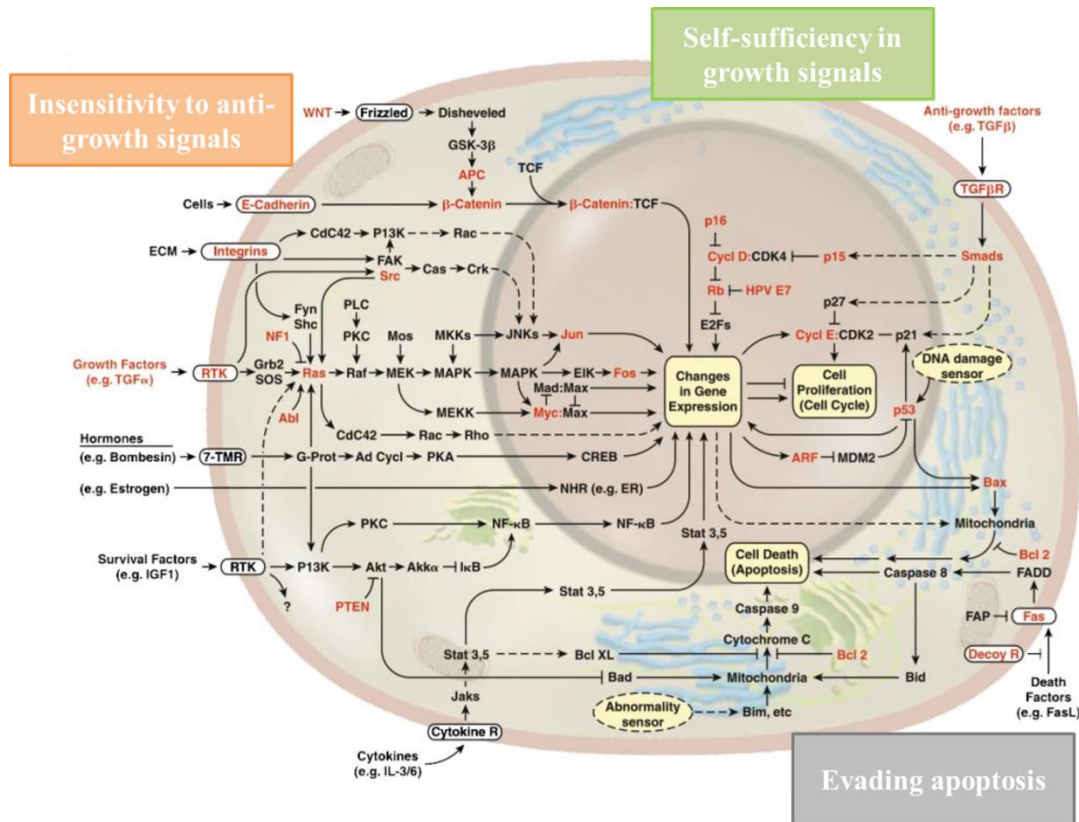


Figure 1.2.2. Key signalling pathways for cell biology. Adapted from Hanahan and Weinberg, 2000.⁹

Hence, cancer cell proliferation depends on an equilibrium established by the three aforementioned acquired capabilities: self-sufficiency in growth signals, insensitivity to antigrowth signals and evasion of apoptosis. Another of the key requirements for tumour growth is sustained angiogenesis and associated supply of oxygen and nutrients. Several receptors for growth factors like VEGF and PDGF are important mediators of vascular recruitment. Intracellular kinase pathways, such as Akt and mTOR, serve to integrate signalling from both growth factors and nutrients with resultant promotion of tumour neovascularization. Moreover, VEGFR and the epithelial-mesenchymal transition receptor (EMT) are of key importance to the invasive, metastatic phenotype of cancer cells and promoting cell migration to facilitate tumour invasion of better oxygenated tissues.¹⁰

Two other emerging hallmarks have been identified in the pathogenesis of some and perhaps all cancers. One involves the capability to modify, or reprogram, cellular metabolism in order to most effectively support neoplastic proliferation.¹¹ Oncogenic activation of the PI3K pathway (in particular Akt) promotes glucose uptake and utilization and may be the driver of the aerobic glycolysis phenotype.¹⁰ The second emerging hallmark allows cancer cells to evade immunological destruction, by T- and B-lymphocytes, macrophages, and natural killer cells in particular.¹¹

1.3. Protein Kinases

Adenosine triphosphate (ATP) (Figure 1.3.1) is present in cells at a concentration of 1 to 5 mM and its hydrolysis to adenine diphosphate (ADP) and inorganic phosphate is central to almost all biological processes.¹⁰ Protein kinases are enzymes that transfer the inorganic phosphate from ATP to their targets. Protein phosphorylation by protein kinases is a universal switching mechanism that regulates almost all cellular processes and, over the past years, intensive research has been conducted to understand their activity (Figure 1.3.1). For every protein kinase, a corresponding phosphatases exists which is responsible for removal of the phosphate from their common target; however, less is known concerning the mechanisms which regulate the activity of these phosphatases (Figure 1.3.1).¹⁰

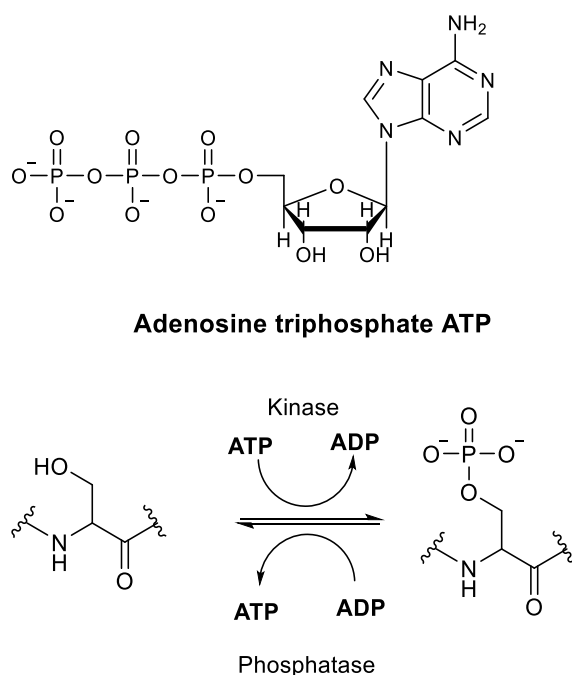


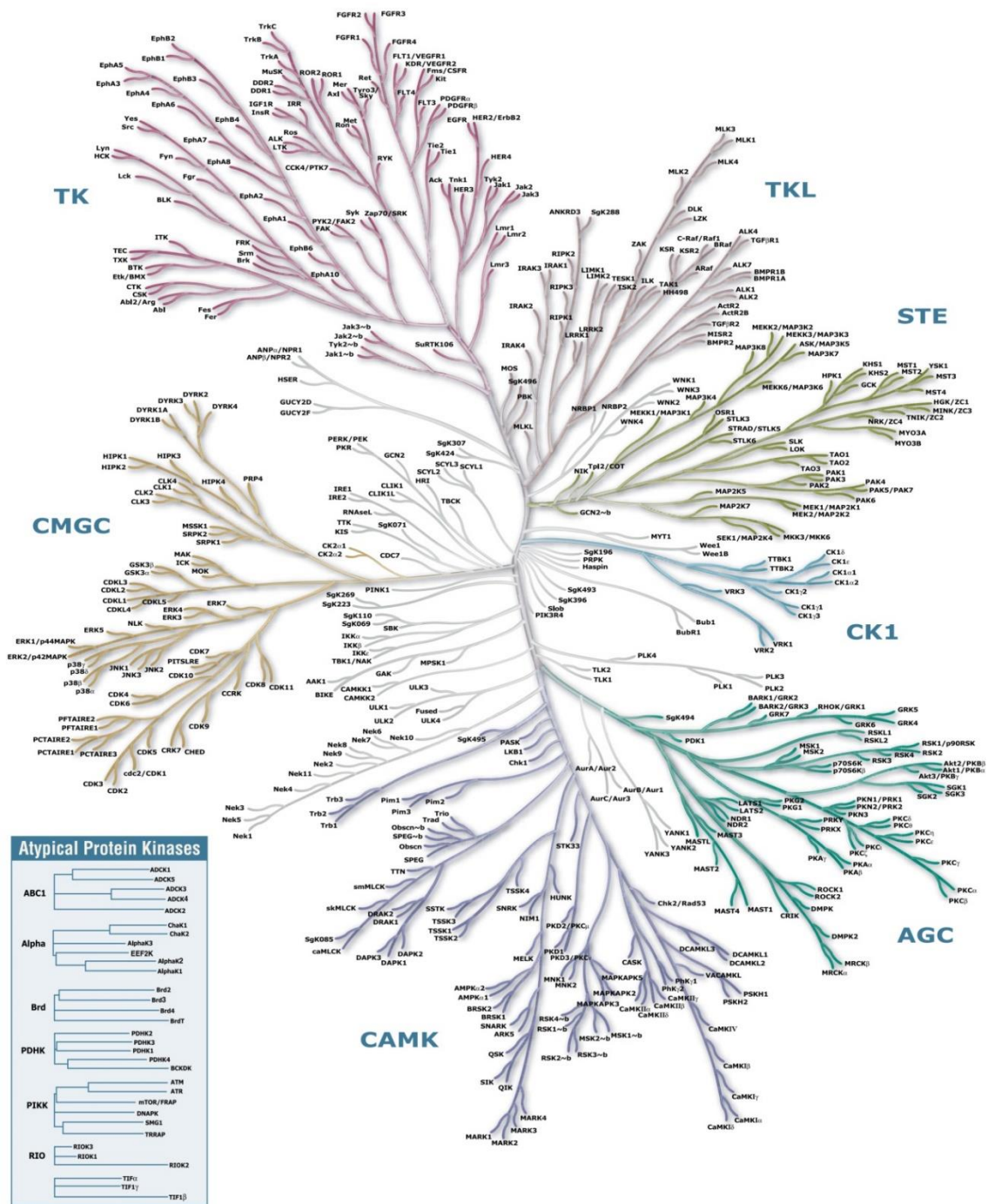
Figure 1.3.1. Protein kinase/phosphatase “phospho” transfer/elimination reaction.

Protein kinases are considered one of the largest protein families in humans.^{12,13} The kinase enzymes in this family play important roles in cell growth, metabolism, differentiation, and apoptosis by catalyzing phosphorylation of serine, threonine or tyrosine residues. Due to the large number of different types of kinases, almost every biochemical transduction pathway is influenced and regulated by phosphorylation events governed by protein kinases. Targeted inhibition of protein kinases has therefore become an attractive therapeutic strategy in the treatment of relevant diseases.¹⁴

Based on sequence similarity, Hanks and Hunter discovered conserved sequence motifs within the kinase core that they classified into 12 subdomains.¹² Only after sequencing of the human genome, Manning *et al.* catalogued the protein kinase complement of the human genome, known as the kinome (Figure 1.3.2), showing the phylogenetic relationships of all the known protein kinases within the human genome. There are over 500 known kinases, encoded by ~1.7% of the human genome and they constitute the second largest set of drug targets after G-protein-coupled receptors.¹⁵ Families and sub-families are clustered based on their sequence similarity and biological function. The kinome is divided into eight subgroups: AGC (containing PKA, PKG, PKC families), CAMK (calcium/calmodulin-dependent protein kinase), CK1 (casein kinase 1), CMGC

(containing cyclin-dependent kinases-CDKs, mitogen activated protein kinases-MAPKs, glycogen synthase kinase 3-GSK3 and Cdc2-like kinases-CLK families), STE (homologs of yeast Sterile 7, Sterile 11, Sterile 20 kinases), TK (tyrosine kinase), TKL (tyrosine kinase-like). Their division is also based on the protein residue that they phosphorylate. For instance, the TKs form a distinct group, whose members phosphorylate proteins at tyrosine residues, whereas enzymes in all other groups phosphorylate primarily serine and threonine residues. Kinases that belong to different groups can interact with and phosphorylate each other, for example MEK, a STE7 kinase, phosphorylates ERK, one of the best-studied mammalian signalling cascades (CMGC).

Atypical kinases (Figure 1.3.2) are kinases that do not exhibit sequence similarity with conventional protein kinases, but are reported to have protein kinase activity. These include ATM and ATR kinases, mTOR, DNA-dependent protein kinase (DNA-PK) and eEF2 kinase (eEF-2 K). Finally, another class of kinases are lipid kinases which act on lipid molecules rather than proteins as their substrates. PI3-kinase (PI3K) is a well-characterized member of this group; other members include diacylglycerol kinase and sphingosine (SphK) kinase.

Figure 1.3.2. The human kinome.¹⁵

1.3.1. Protein Kinases and disease

It is easy to understand that if protein phosphorylation regulates most aspects of cellular biology, abnormal phosphorylation can logically be the cause or consequence of disease. Thus, it has been found that deregulation of protein kinases is implicated in a number of diseases including cancer, diabetes, and inflammation.¹⁶

Many kinases have been found to be involved in cancer development and studies have uncovered a large number of kinases bearing mutations that can lead to tumour development. Altered key signalling pathways in cancer (Figure 1.2.2.) largely involve protein kinases of which deregulation can alter cell biology. Some of the most commonly studied kinases are those that are closely related to or involved in the acquisition of the cancer hallmarks, also reported in Figure 1.3.3.¹⁷

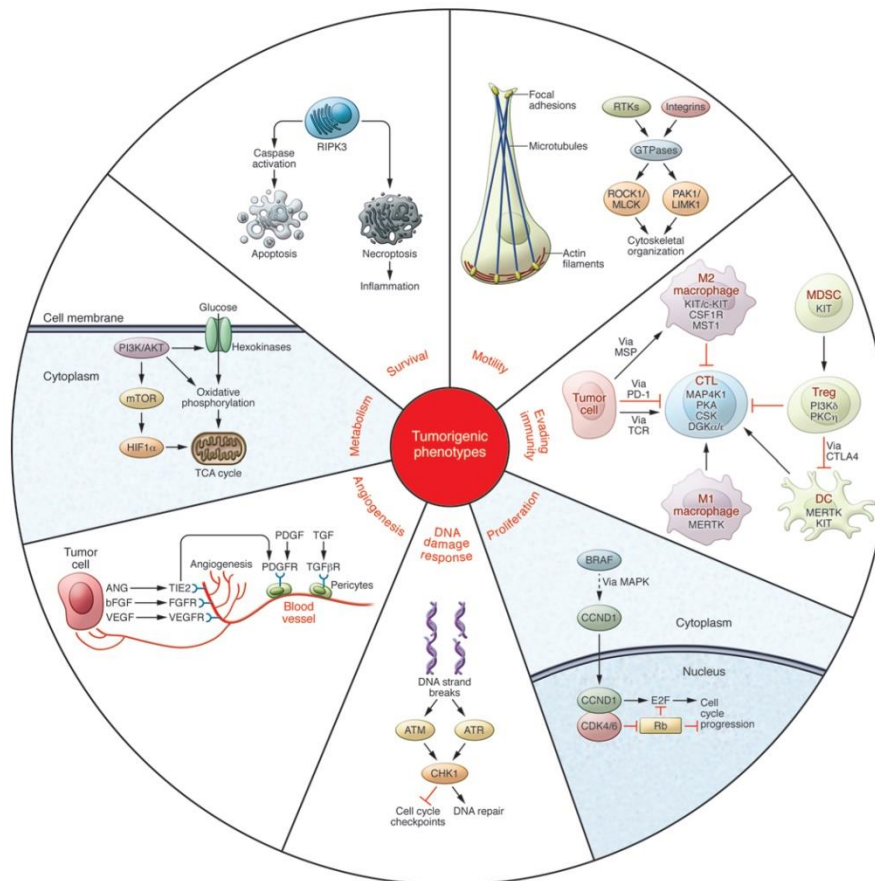


Figure 1.3.3. Altered key signalling pathways by oncogenic alteration of kinases. Adapted from Gross *et al.*, 2015.¹⁸

Although cancer still remains the major focus of drug discovery nowadays, the number of kinase inhibitors undergoing clinical trials for the treatment of other diseases is also increasing.¹⁹ Deregulation of kinase function has been implicated in different disorders, including immunological, metabolic and infectious diseases.²⁰ Diverse types of mutations implicate certain kinases and alter the normal function of the immune system. For example, almost 150 mutations have been reported in Bruton tyrosine kinase (BTK), which encodes a key regulator of B-cell development.²¹ Endocrine and metabolic disorders can derive from a mutation in AKT2, which encodes a ubiquitous STK (serine/threonine kinase) that has a key downstream function in activating the insulin receptor resulting in several pathologies.²¹ Development of kinase-targeted therapies for central nervous system (CNS) diseases remains a challenge, since the problem is not only the design of inhibitors that exhibit sufficient affinity and selectivity for a molecular target (or set of targets), but also insuring that these inhibitors have the appropriate molecular profile for CNS penetrance.²² One example of CNS protein kinase targets is glycogen synthase kinase 3 (GSK3) which is a serine-threonine protein kinase and is involved in numerous cellular processes. Not surprisingly upregulation of GSK3 activity has been linked to numerous human pathological conditions, including diabetes, muscle hypertrophy, cancer, Alzheimer's disease, stroke, sleep disorders, and neuropsychiatric and mood disorders. Currently, diverse small-molecule GSK3 inhibitors have shown preclinical efficacy in animal models for Alzheimer's or Parkinson's disease, amyotrophic lateral sclerosis (ALS), depression and ischemic stroke.²²

1.3.2. Functioning and features of protein kinases

Internal architecture of eukaryotic protein kinases is controlled by a highly dynamic and finely regulated set of conserved residues that act as switches for the control of most biological events in eukaryotic cells. Sequence analysis and surface comparison of different serine-threonine and tyrosine kinases reveals residues whose spatial positions are highly conserved among eukaryotic protein kinases.²³ Although sequence was enough to determine the conserved regions, the first crystallization of a protein kinase, PKA, revealed not only the overall architecture of the kinase core including where all

the conserved residues were located, but also the way in which a kinase can recognize a peptide/protein substrate.^{24,25}

For the sake of simplicity in describing the kinase' domains and the motifs, PKA will be taken as an example and the numbering of its individual residues will be used in the discussion, as the relevant residues are conserved in every kinase.²⁶ The most important domains and motifs are represented in Figure 1.3.4.

Kinases consist of an upper N-lobe, rich in β sheets, and a lower C-lobe, rich in α helices. The ATP/Mg²⁺ is buried at the base of the cleft between the two lobes, known as the 'hinge' region. Among the conserved key functional regions for kinase activation, we find the α C-helix, the DFG, HRD and APE motifs, and the activation loop. The activation segment is defined as the region between and including two conserved motifs (DFG at the N terminus and APE at the C terminus). Different studies have been carried out in order to gain an understanding of how the conformation of this activation segment, fundamental for the control of kinase activity, controls a unique set of signals to turn on or off the kinase activity.^{27,28}

The conserved DFG motif marks the beginning of the activation loop²⁸ and consists of: (i) a conserved aspartate (Asp184), which is known to chelate a Mg²⁺ which, in turn, interacts with the phosphates of the ATP molecule for phosphotransfer;²⁹ (ii) a conserved phenylalanine (Phe185), which makes a significant contribution to the DFG flip state from active to inactive kinase, disrupting fundamental interconnections with two hydrophobic residues in α C-helix;²⁸ and (iii) a conserved glycine (Gly186), the role of which remains unclear even though it is an integral part of all protein kinases.³⁰ The DFG motif can assume one of two conformations based on the orientation of the central phenylalanine side chain: DFG-in and DFG-out. The conformation of the DFG motif affects the ATP binding and the overall catalytic activity of the kinase. For example, distortion of DFG motif is a consequence of the V600E mutated form of B-Raf, where the salt bridge formed between Glu600 and Lys507 keeps the activation loop in the 'DFG-in' conformation and renders the mutant protein constitutively active.³¹

The electrostatic interaction between a particular phospho-residue (Thr197 in PKA) and a basic pocket that is conserved among kinases (known as RD pocket) is critical for driving the conformational change in the activation loop.²⁸ Assembly of the activation

segment into an active conformation is typically mediated by phosphorylation either by *cis* or *trans* autophosphorylation or by another activating kinase.³²

The HRD motif forms the catalytic site and is crucial to kinase function as it plays a key role in phospho-transfer to the peptide substrate. It is composed of three different amino-acids: a histidine, an arginine and an aspartate (H, R and D). The His residue is not only a conserved amino-acid of the HRD motif, but is also the last residue of the regulatory spine (R-spine) which aids the alignment of the kinase in an active state.³³ In PKA and some of the other AGC kinases, this His is replaced with a Tyr (Tyr164). In most eukaryotic protein kinases the second residue in the HRD motif is arginine Arg165.³³ Finally, the last residue is Asp166 which is required for the orientation of the hydroxyl group of the substrate peptide and the transfer of the phosphoryl group.³⁴

The peptide substrate is known to bind in the C-lobe, in such a way that it can reach the Asp166 in the HRD region and the proximal Lys168 responsible for catalysis.³⁵ The APE motif is important in this sense as it constitutes the end of the activation segment and forms part of the substrate-binding pocket.²¹ In PKA, Glu208 is part of the conserved APE motif (together with Ala206 and Pro207) and exerts a conserved interaction with Arg280, which appears to be key to maintaining correct interactions and conformational changes associated with the α F-helix.³⁶

A new model of kinase regulation based on the assembly of hydrophobic spines has recently been proposed. The kinase core is built around a hydrophobic centre that consists of a hydrophobic helix (α F-helix) and two hydrophobic spines (C-spine and R-spine) that are formed by non-contiguous residues from different sequences.³⁷

The C-spine (or catalytic spine), consists of a series of hydrophobic residues (Val57, Ala70, Met128, Leu172, Leu173, Ile174, Leu227, Met231) and is completed after ATP binds.³⁷

The R-spine consists of four residues, RS1 to RS4, two from the C-lobe and two from the N-lobe: RS1 is the tyrosine residue from the HRD motif in the catalytic loop (Tyr164); RS2 is the phenylalanine from the DFG motif in the activation segment (Phe185); RS3 is a conserved aliphatic residue from the α C-helix (Leu95); and RS4 is a residue that belongs to β 5 sheet (Leu106). Recent studies indicate that mutations which

lock the R-spine in the active conformation result in a kinase with constitutive activity, independent of Ras, dimerization, or activation loop phosphorylation for its activity.^{38,39}

Assembling of two hydrophobic spines, supported by the α F-helix, connects the N- and C- lobes, forming a hinge region where the ATP molecule can bind for catalysis to occur with the help of a glycine-rich loop.

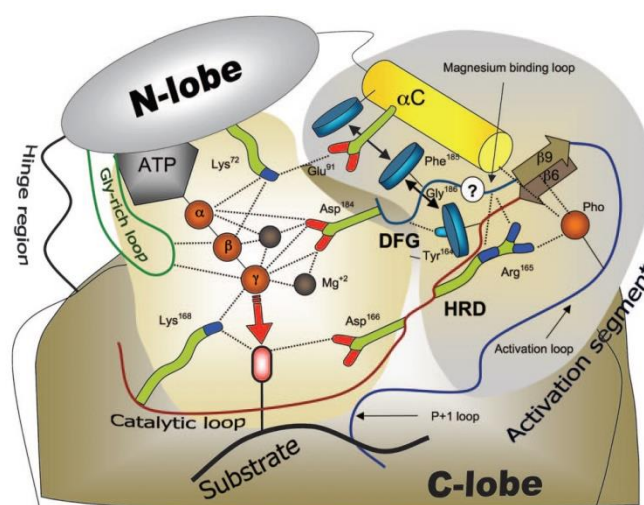
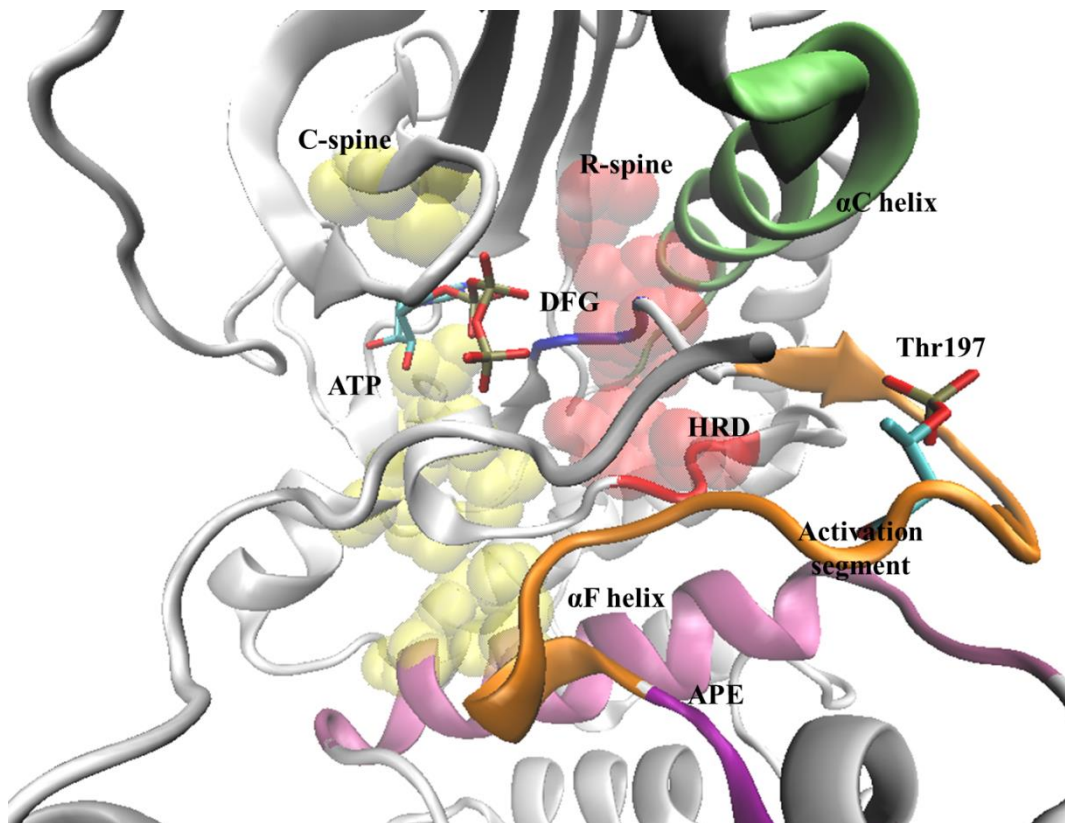


Figure 1.3.4. *Top:* Structure of PKA and conserved motifs: the structural parts are labelled according to present literature. Yellow: C-spine; red: R-spine; lime: α C-helix; pink: α F-helix; blue: DFG motif; red: HRD motif; orange: activation segment; purple: APE motif.²⁶ *Bottom:* cartoon explanation of PKA catalytic domain. Adapted from Kornev *et al.*, 2006.³⁰

Kornev *et al.* proposed in 2006 a general model for protein kinase activation.³⁰ In a starting inactive conformation (Figure 1.3.5.B), DFG and HRD motifs and the spines are distorted, destabilizing the molecule causing it to adopt different inactive conformations. Activation starts (Figure 1.3.5.A) with phosphorylation at Thr197 (PKA numeration) that re-arranges the Mg^{2+} -binding loop positioning the DFG aspartate in an appropriate orientation for interaction with ATP through Mg^{2+} binding and building up the hydrophobic spine. The α -helix flips inward towards the ATP hinge region to complete the spine formation and secure it through Lys72-Glu91 interaction. Positioning of Asp166 from the HRD motif will facilitate the interaction with the peptide substrate and its subsequent phosphorylation.³⁰

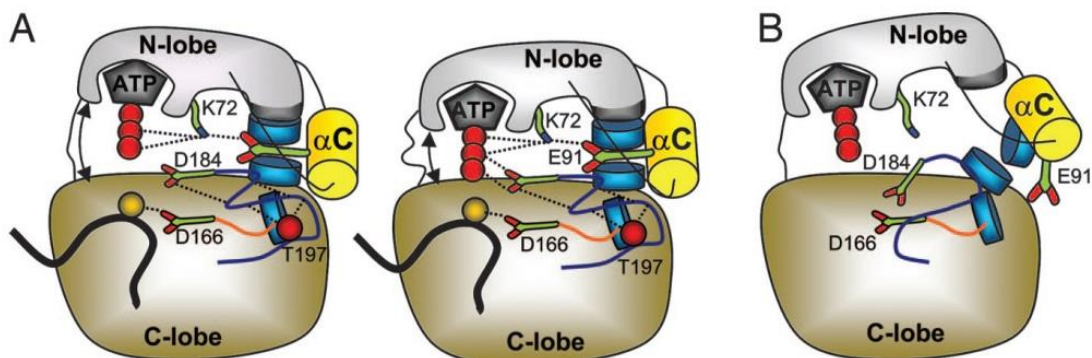


Figure 1.3.5. Kinase activation mechanism. Adapted from Kornev *et al.*, 2006.³⁰

The activation process can also involve other kinase specific events like trans-membrane translocations and dimerization, as is the case for B-Raf.⁴⁰ It has been previously reported that, in B-Raf, Arg506 plays a role in the dimerization process. Recent studies have demonstrated how the positively charged C-terminal end of the α C-helix composed of Arg506 and Lys507 can form several inter-protomer salt bridges

with the phosphorylated NtA motif (SSDD motif at the N-terminus of the kinase), promoting dimerization.⁴¹

1.3.3. Ras/Raf/ ERK MEK/ kinase pathway and its interconnections

As briefly eluded to, the Ras/Raf/MEK/ERK signal transduction pathway is a conserved kinases' cascade that regulates different aspects of cell biology in response to growth factors, cytokines and hormones.⁴²

Classical activation results from growth factors binding to a tyrosine kinase receptor (TKR) such as EGFR, thus triggering receptor dimerization, activation, and transphosphorylation.⁸ Adaptor proteins (e.g. Shc, Grb2, Sos) then associate with the receptor's phosphorylated intracellular domains and recruit guanine nucleotide exchange factors (GEFs) to the cell membrane which activate Ras (H-Ras, N-Ras, K-Ras). GEFs promote formation of Ras-GTP by catalysing the replacement of GDP with GTP, whereas GTPase-activating proteins (GAPs) stimulate GTP hydrolysis to GDP.⁴³ Then, Raf activation is initiated by association with Ras-GTP through the Ras binding domain (RBD) situated at the N-terminal part of the kinase. Raf proteins act as key points in Ras/Raf/MEK/ERK pathway and despite several published attempts to characterize their complex regulation,⁴³ the exact mechanism of Raf activation remains largely unknown. There are three isoforms of Raf (A-Raf, B-Raf and C-Raf) and each of them differs in the details of their activity regulation.⁴⁴ B-Raf is considered to be the most active of the three because its highest affinity and stronger stimulation towards MEK, while A-Raf has instead the lowest activity due to the weakest binding affinity toward Ras.⁴⁵ Raf dimerization (homo- and hetero-) is required for activation of wild-type Raf, and it can also constitute a mechanism of escape used by cancer cells to bypass eventual inhibition with small-molecule inhibitors. This is known as the Raf inhibitor paradox, where non saturating concentrations of inhibitor can cause transactivation of the Raf dimer to the active state or even an adaptive resistant response.⁴⁶

Following recruitment to the cell membrane, conformational changes, phosphorylation and dimerization promote Raf serine/threonine kinase activity, consequently activating MEK and ERK.^{40,42}

Components of this pathway are frequently mutated in cancer with approximately one-third of human tumours expressing a constitutively activated mutant form of Ras and approximately 8% of tumours expressing an activated form of B-Raf.⁴⁶ B-Raf mutation, specifically V600E, is responsible for 50-70% of melanomas; in contrast, non-small cell lung cancer is mainly driven by mutation at receptor EGFR (10-20%) and N-Ras mutation (15-20%), as shown in Figure 1.2.3.⁴⁷

Many of the transcription factors activated by the Ras/Raf/MEK/ERK pathway are involved in cell proliferation and differentiation (e.g. AP1, c-Myc), in addition, many growth factor genes have binding sites for transcription factors, activated by the Ras/Raf/MEK/ERK pathway, located in their promoter regions. Thus, aberrant activation of this pathway may establish an autocrine/paracrine loop, resulting in self-sufficiency in proliferative signals and continuous stimulation of cell growth, with consequences that facilitate acquisition of all the different hallmarks of a cancer cell previously described (Figure 1.3.6.).⁴⁷

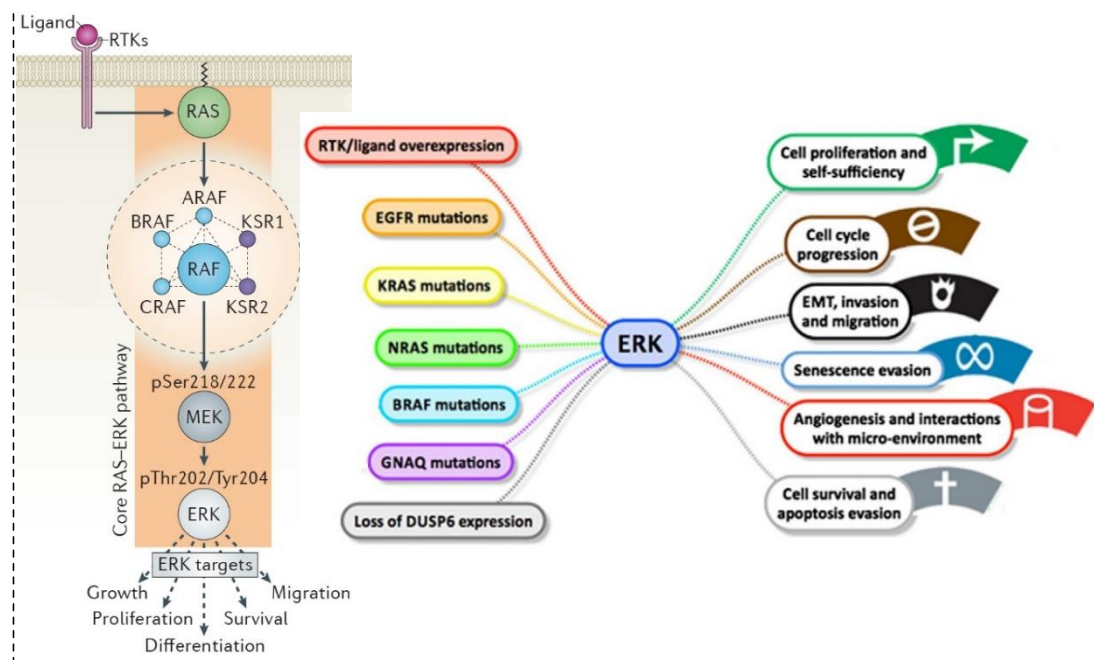


Figure 1.3.6. *Left:* the Ras/Raf/MEK/ERK signal transduction pathway; *Right:* main molecular alterations leading to ERK activation in cancer and biological consequences. Adapted from Lavoie *et al.*, 2015⁴³ and Neuzillet *et al.*, 2014.⁴⁷

The Ras/Raf/MEK/ERK pathway not only plays a critically important role in controlling gene expression and preventing apoptosis, but it also establishes crucial interactions with other key cell signalling pathways and exerts its functions through these interconnections (Figure 1.3.7.).

Activated Ras has the ability to phosphorylate two additional pathways besides the Raf/MEK/ERK cascade; the PI3K/Akt/mTOR and the Ral guanine nucleotide exchange factors (Ral-GEFs) pathways (Figure 1.3.7.).⁴⁸ The Ras/Raf/MEK/ERK and Ras/PI3K/PTEN/Akt pathways interact with each other to regulate growth and tumorigenesis. Moreover, Raf is downregulated by Akt, indicating a cross-talk between the two pathways. On the other hand, RalGEF and/or Ral activation is involved in Ras-mediated activation of various transcription factors that include, for example, phosphorylation and activation of c-Jun through JNK MAPK activation, STAT3 through Src tyrosine kinase activation and NF- κ B gene regulation (Figure 1.3.7.).⁴⁹ Another established interaction of the Ras/Raf/MEK/ERK pathway is its cross-talk with the JAK/STAT pathway. The relationship between these cascades is highly complex with each of these pathways intersecting at multiple levels.⁵⁰

Regulation of apoptosis by the Ras/Raf/MEK/ERK cascade can also be achieved as a result of phosphorylation of apoptotic regulatory molecules such as Mcl-1, Bim and caspase 9.⁵¹ Moreover, a non-MEK/ERK mediated function of C-Raf has also been proposed in which it is targeted to the mitochondrial membrane by Bcl-2, protecting cells from apoptosis and results in phosphorylation of BAD, proapoptotic Bcl-2 homolog (Figure 1.3.7.).⁵²

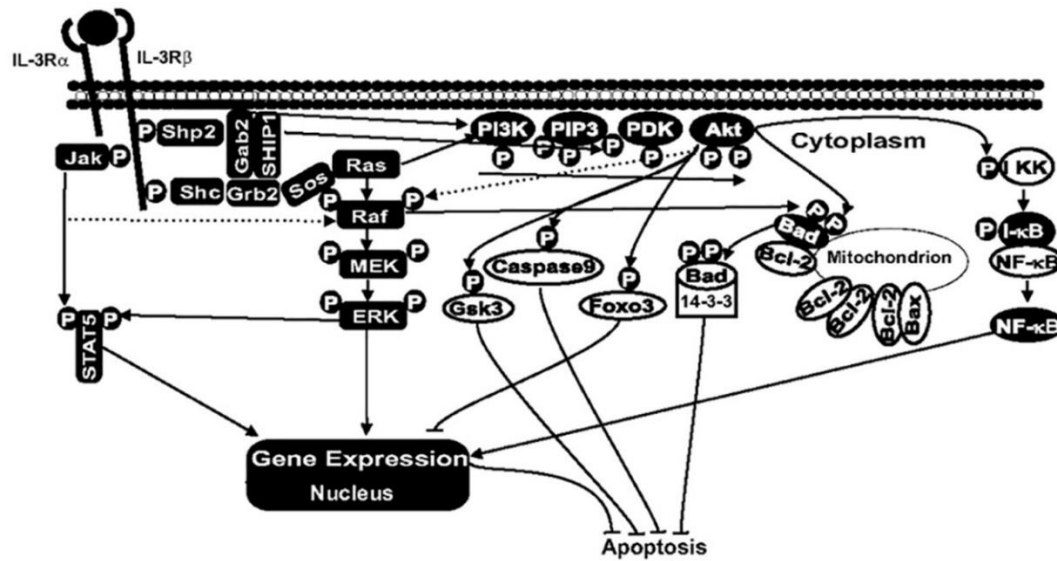


Figure 1.3.7. The Ras/Raf/MEK/ERK kinase pathway and its interconnections. Adapted from Steelman *et al.*, 2004.^{49,53}

1.4. Protein kinases inhibitors

The prevalence and influence of protein kinase deregulation on disease development presents a unique opportunity for therapeutic intervention and, hence, there has been considerable interest and effort invested in developing pharmacological tools targeting kinases. So much so that they have become one of the most intensively pursued classes of drug target.²⁰

Despite the interest in and success of protein kinase inhibitors as therapeutics, drug discovery has focused on a relatively small subset of kinase targets.²⁰ This has led to drug-development programs focusing on similar sets of scaffolds, covering only a small fraction of the kinome, for which evidence of therapeutic utility already exists (Figure 1.4.1).²⁰ As showed in Figure 1.4.1, the most targeted protein kinases are the tyrosine protein kinases, in particular VEGFR and EGFR. Over the past five years, serine/threonine kinases (B-Raf) and also lipid kinases (PI3K) have acquired much interest; however, tyrosine kinase inhibitors, including dual specificity inhibitors, still account for 77% of all approved small molecule kinase inhibitors (SMKIs or PKIs) targeting most of the common tyrosine kinases.²⁰ It is not surprising that more than 100 kinases (~25%) have unknown function and that about 50% are largely

uncharacterized, with still little indication of their biological function or role in disease development.

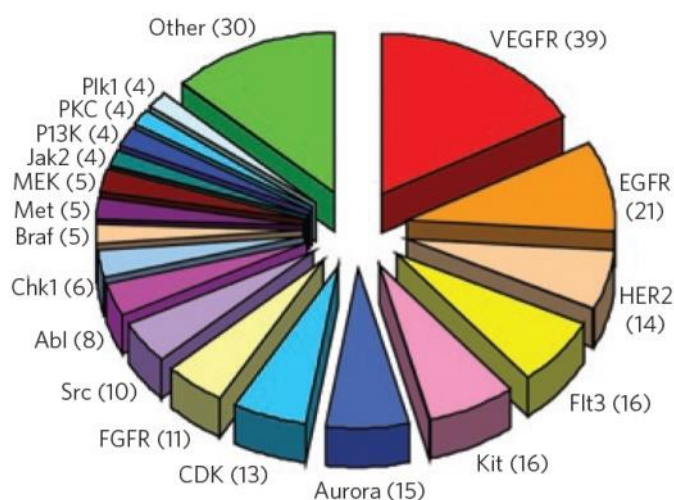
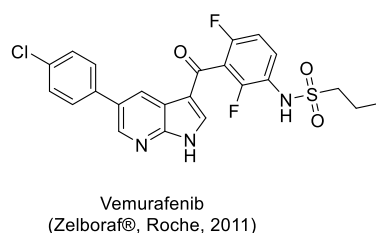
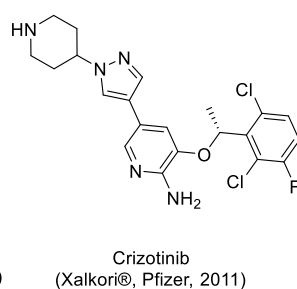
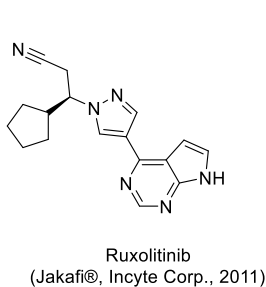
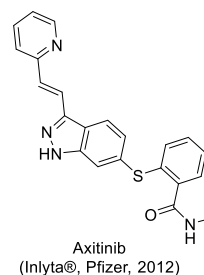
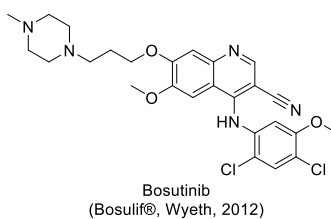
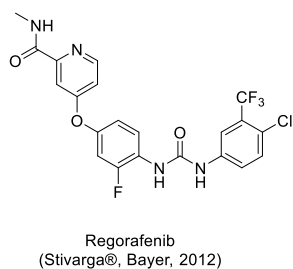
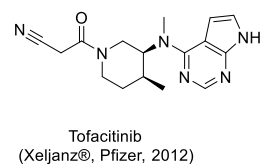
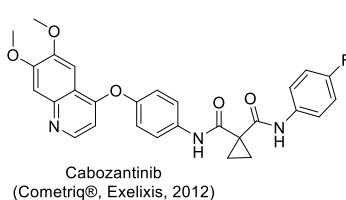
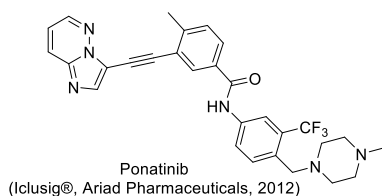
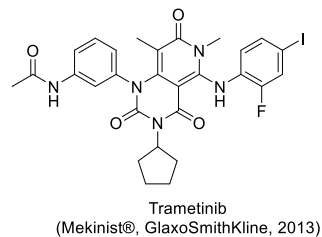
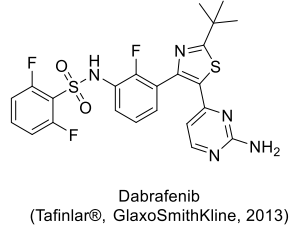
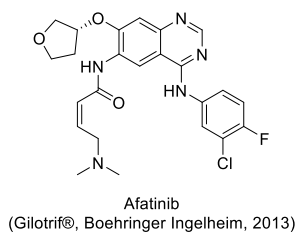
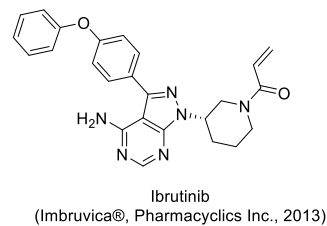
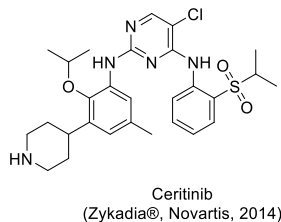
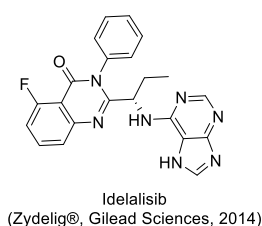
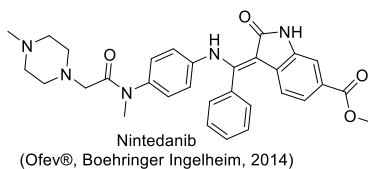
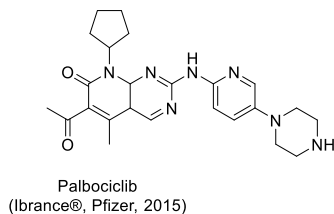
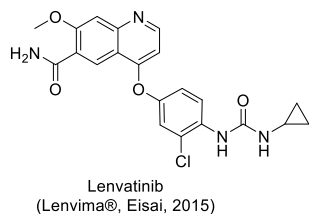
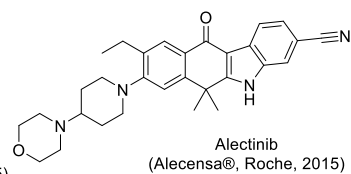
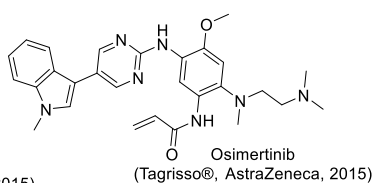
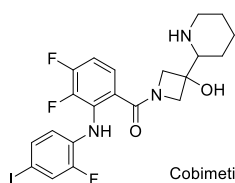


Figure 1.4.1. Estimation of targeted kinases in clinical trials. Adapted from Fedorov *et al.*, 2010.⁵⁴

Nowadays, more modern techniques can be employed to help address this problem and find new kinase targets. For example, mass-spectrometry based techniques and phospho-proteomics are powerful tools capable of revealing cancer signatures. These have been utilized to analyse changes in the proteome and phospho-proteome of primary cells derived from epithelial ovarian cancer (EOC) compared to healthy tissues, revealing a role for the kinase CDK7 in EOC cell proliferation.⁵⁵

1.4.1. Small molecule kinase inhibitors

Understanding the structural basis and mechanism of kinase inhibition is crucial to the goal of developing inhibitors to target every member of the kinome which can be selectively tuned to achieve a different mode of inhibition. To date, the US FDA has approved 31 small-molecule protein kinase inhibitors (SMKIs or PKI). Their structure, together with their therapeutic indication are reported in Figure 1.4.2.⁵⁶⁻⁵⁸



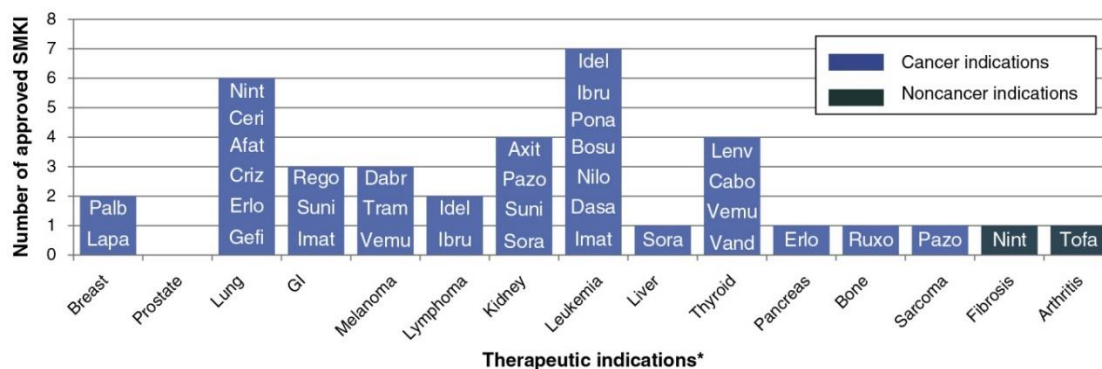
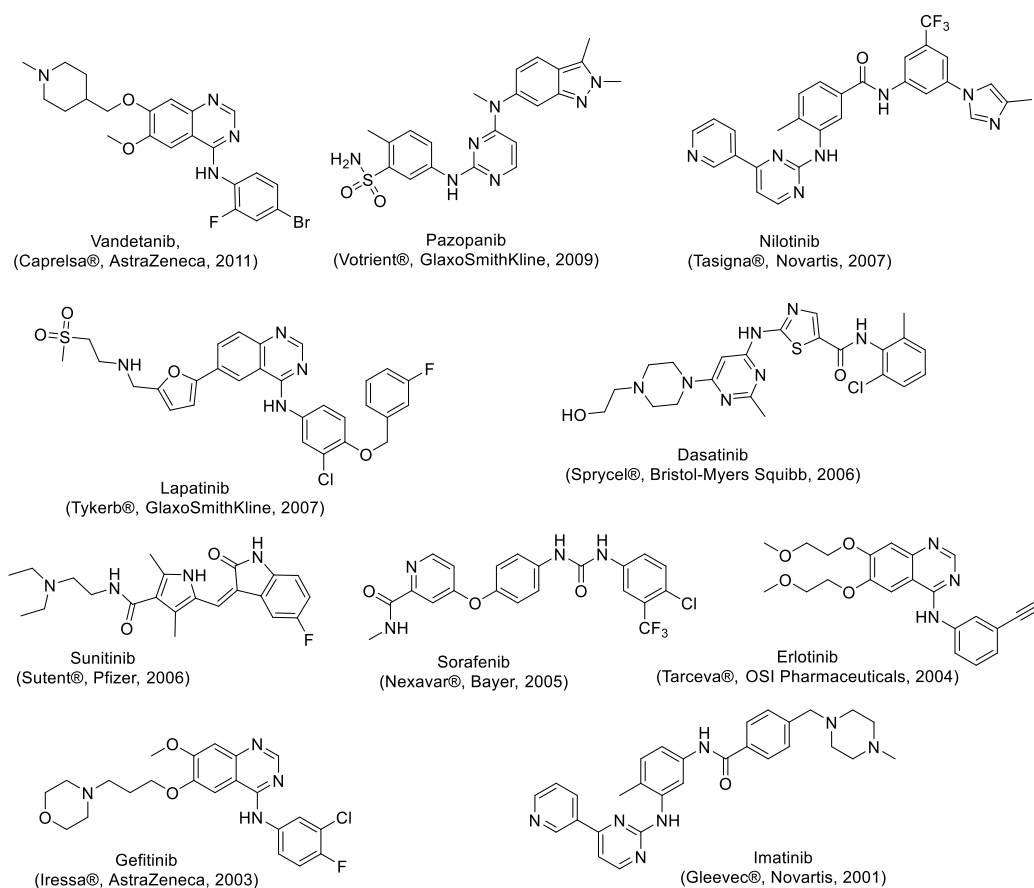


Figure 1.4.2. US FDA has approved 31 SMKIs (in brackets year of approval) and their therapeutic indication. Adapted from Wu *et al.*, 2016.⁵⁶⁻⁵⁸

Among these, five small molecule were specifically designed and approved to inhibit the Ras/Raf/MEK/ERK pathway. Two RAF inhibitors (vemurafenib, zelboraf; Genentech/ Plexxikon and dabrafenib, tafenlar; GlaxoSmithKline) and two MEK inhibitors (trametinib, mekinist; GSK and cobimetinib, cotellic, Genentech) have been approved against melanoma expressing BRAF^{V600E/K} mutation.⁴⁶ More recently, in

2015, osimertinib (tagrisso, AstraZeneca) was approved as EGFR inhibitor for the treatment of non-small cell lung cancer with EGFR^{T790M} mutations.

There are different types of PKIs classified based on their mode of action and binding:

Type I inhibitors (such as vemurafenib, zelboraf): ATP-competitive inhibitors that recognize the so-called active conformation of the kinase (DFG-in) and constitute the majority of known PKIs. Type I inhibitors typically consist of a heterocyclic ring system that occupies the purine binding site and are the most widely pursued class (many kinase inhibitors have been synthesized to mimic ATP) (Figure 1.4.3).²⁰

Type II inhibitors (such as imatinib, gleevec or sorafenib, nexavar): ATP competitive inhibitors which recognize the inactive conformation of the kinase (DFG-out). Movement of the kinase to an inactive state exposes an additional hydrophobic binding site directly adjacent to the ATP binding site which is also occupied by this type of inhibitor (Figure 1.4.3).²⁰

Type III, known also as type IV⁵⁹ or allosteric inhibitors (such as trametinib, mekinist): This third class of compounds bind to a site outside the ATP-binding site thus modulating kinase activity in an allosteric manner. Inhibitors belonging to this category tend to exhibit the highest degree of kinase selectivity because they exploit binding sites and regulatory mechanisms that are unique to a particular kinase (Figure 1.4.3).²⁰

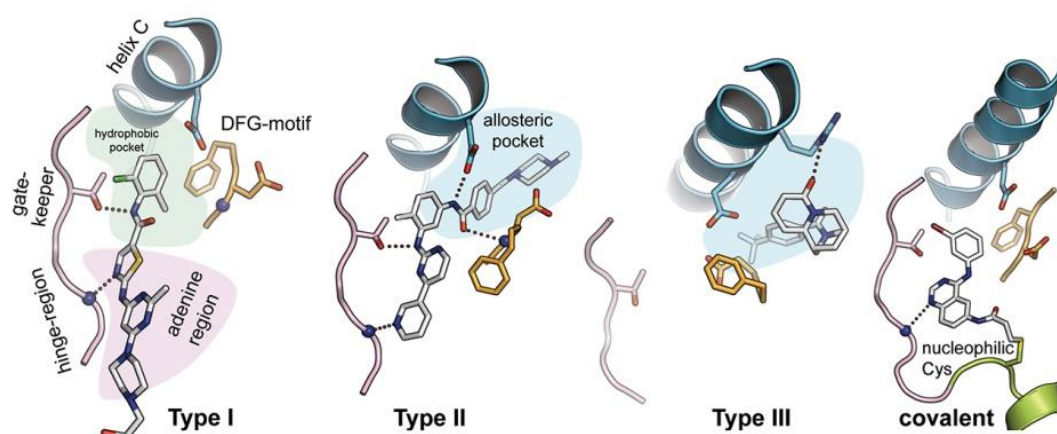


Figure 1.4.3. Representation of the mode of action for type I, type II, type III and covalent inhibitors. Adapted from Fang *et al.*, 2013.⁶⁰

Other type of inhibitors include covalent inhibitors, which are capable of forming an irreversible, covalent bond to the kinase active site, most frequently by reacting with a nucleophilic cysteine residue (Figure 1.4.3). Recently, these covalent inhibitors have received increasing attention because of their high binding affinity and selectivity.⁶¹ The FDA has recently approved three irreversible inhibitors, afatinib and osimertinib as EGFR inhibitors for the treatment of non-small-cell lung cancer (NSCLC) and ibrutinib as BTK inhibitor for B-cell malignancies. Lately, new insights into cysteine reactivity have been defined through a structure-based binding site cysteine data set, which can contribute towards the development of further covalent kinase inhibitors.⁶¹

Type V (such as levantinib, lenvima) inhibitors instead refer to bivalent inhibitors that can target different sites simultaneously.⁶² As previously mentioned, allosteric inhibitors usually display excellent selectivity profiles toward a particular kinase, but they usually lack the potency of the more common type I inhibitors. Chemically linking an allosteric moiety with a type I inhibitor to form a bivalent inhibitor, could improve both selectivity and affinity.⁶²

Type I and II PKIs are so far the most developed type of inhibitors followed by covalent, type III and V inhibitors (Figure 1.4.4 *left*). Type I are known to be the most potent inhibitors, but despite their potency they are known to lack selectivity due to the highly conserved nature of the hinge region across all kinases. However, some of the type I inhibitors (such as vemurafenib) belong to a class of very high selective inhibitors (Figure 1.4.4 *right*).^{31,56} Other type II inhibitors (such as sorafenib) have been developed to bind more selectively to their target, but have instead been classified as promiscuous inhibitors.⁵⁶ Despite limited examples, type III inhibitors, such as trametinib, are likely to be more selective than type I and II inhibitors because of the exclusive utilization of allosteric binding sites outside the ATP-binding pocket.⁵⁶

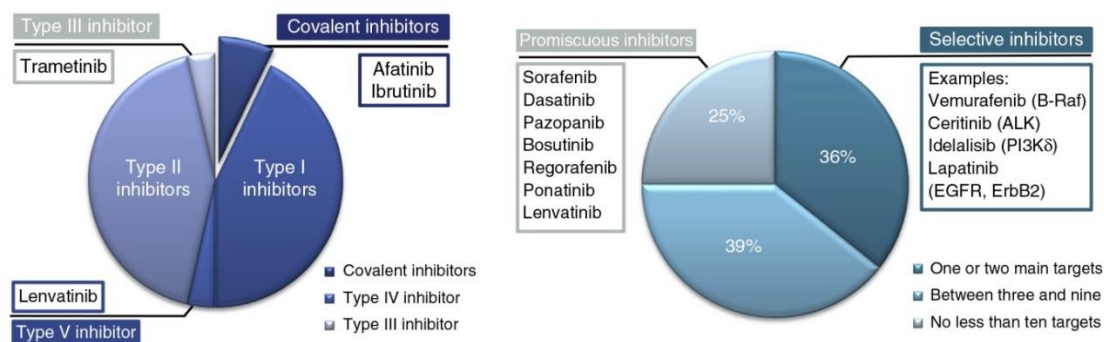


Figure 1.4.4. Classification (*left*) and selectivity (*right*) of the FDA approved SMKIs. Adapted from Wu *et al.*, 2016.⁵⁶

1.4.2. Current challenges in PKI discovery and future directions

Selectivity and resistance are two of the most prominent issues that affect PKI discovery programs. Given the widespread importance of kinases and phosphatases in cell biology (and disease), controlled enzyme pharmacology is essential for the development of successful therapies.

Targeted therapies represent the state-of-the-art in preclinical and clinical oncology research. The use of imatinib (type II, gleevec) in the treatment of chronic myelogenous leukemia (CML) or vemurafenib (type II, zelboraf) for the treatment of melanoma serves as a proof of concept that inhibition of specific (or few) kinases can improve the medical responses to treatment of such disease.^{31,63} Most approved PKIs potentially inhibit multiple targets, but, as exemplified by imatinib, such accidental off-target activities can extend the utility of a drug to treat other tumour types driven by a distinct kinase.⁶⁴

Type III inhibitors have acquired much importance due to their potential specificity against a particular kinase. Development of allosteric inhibitors that target previously undiscovered binding sites is not an easy task given the difficulties in identifying and characterizing them. For example, trametinib is the only allosteric type III MEK inhibitor of the Ras/RafMEK/ERK pathway approved for mutated B-Raf(V600E) melanoma.^{46,65} Chemical biology together with computational modelling comprised an enormous contribution providing experimental characterization of the process by which such binding occurs.⁶⁶

Nevertheless, resistance to targeted therapy is very likely to occur.⁴⁶ Strategies to overcome this issue include the development of dual kinase inhibitors that target different kinases involved in the same pathway. For example, dasatinib (sprycel) is an oral Bcr-Abl tyrosine kinase inhibitor and Src tyrosine kinase inhibitor.⁶⁷ Moreover, the dual Raf/MEK inhibitor RO5126766 (CH5126766) may be a potential therapy for Ras-mutated tumour cells.⁶⁸ Pan-Raf inhibitors in development have shown antitumour efficacy in preclinical models, with the ability to overcome Raf paradoxical activation in Ras mutated driven tumours.⁶⁹

Despite the efficacy of such ‘magic bullets’, which are considered the targeted therapies, cancer is an evolving disease and emergence of secondary resistance to growth inhibitory drugs limits the efficacy of targeted therapies in different patients.⁷⁰ As previously described, tumours only rarely depend on a single regulatory pathway for growth and survival and there are a number of different key signalling pathways upon which cancer cells can rely on to proliferate. Variable response to targeted therapies has therefore led to the emergence of personalized medicine. The objective of this approach is to match these targeted drugs with the molecular and genetic characteristics, both pre-existing and evolving during therapy of each tumour, for each patient.⁷⁰

Personalized therapies can rely on a fine *ad hoc* tuning of targeted therapies based on a specific cancer genome and, accordingly, the hallmarks of cancer elaborated by Hanahan and Weinberg (in 2000 and 2011, respectively) present different points of intervention. Enabling characteristics, such as genomic instability and mutation, have focused attention on epigenetic regulation which is crucial for proper programming and reading of the genome; deregulation of which can lead to cancer.⁷¹ Epigenetic reader domains of the bromodomain family have recently emerged as new targets for cancer therapy. It has also been demonstrated that several clinical kinase inhibitors also inhibit bromodomains with therapeutically relevant potencies and as such are best classified as dual kinase-bromodomain inhibitors.⁷²

Another emerging hallmark concerns avoidance of immune destruction by cancer cells; therefore, a drive to develop personalized medicines which can activate antitumor immunity by either non-specific immune stimulation, immune checkpoint blockade (e.g. anti-PD-L1, ipilimumab, yervoy) and vaccination strategies, is now at its cutting edge.⁷³

In conclusion, there is tremendous potential for synergistic combinations of immunotherapy, epigenetic targets and conventional cancer treatments in a more personalized anticancer therapy guided by specific biomarkers, that gives hope to future cancer drug discovery.⁷⁴

1.4.3. Protein Kinase inhibition assays

The challenges which accompany kinases' inhibitors development are reflected by the number of FDA-approved drugs with such activity, which remains very low in comparison with the hundreds of molecules in drug discovery programs.⁷⁵ To accomplish this goal, reliable assays of kinase inhibition are needed and the most established ones are herein reported.

Biochemical kinase assays can be divided into two classes: (i) activity assays, which directly or indirectly quantify the catalytic product (*i.e.* the phosphorylated substrate) and (ii) binding assays, which quantitatively measure the competitive binding of small molecules to the ATP-binding site.

Classical methods to test kinase activity involve the quantification of the phosphorylation reaction by detection of the phosphorylated product or the change in the ratio of ATP to ADP. The traditional biochemical method used to achieve this goal is the radioisotope filtration binding assay, in which the reactions are performed in the presence of radiolabelled γ -ATP (Figure 1.4.5). The reaction mixtures are then spotted onto filter papers, which bind the radiolabelled catalytic product, while unreacted phosphate is removed by washing the filter paper. The radioisotope filter binding assay is considered the gold standard for kinase profiling.⁷⁶

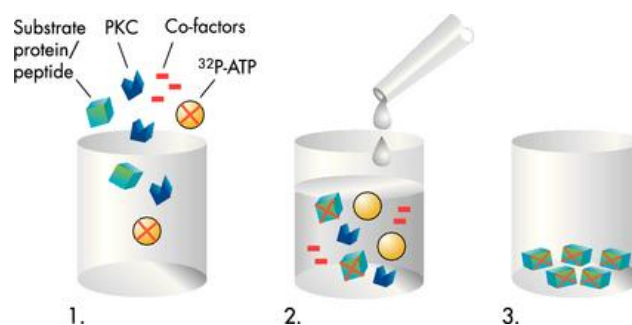
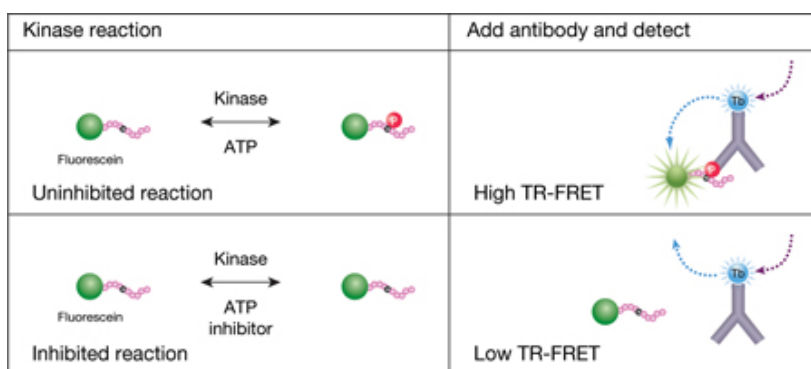


Figure 1.4.5. Illustration of radioisotope filtration binding assay.

This method directly detects the true product without the use of modified substrates, coupling enzymes, or detection antibodies. Despite the sensitivity and resolution of this assay the use of radioisotope materials is a serious handicap for its use in high-throughput screening (HTS).

Fluorescence (or time-resolved fluorescence) resonance energy transfer (FRET and TR-FRET) involves transfer of non-radiative energy from a donor fluorophore to a close-proximity acceptor fluorophore (Figure 1.4.6). One type of activity-based TR-FRET assay employs a peptide substrate labelled with an acceptor fluorophore, and an anti-phosphopeptide detection antibody labelled with a donor fluorophore.⁷⁶

**Figure 1.4.6.** Illustration of fluorescence resonance energy transfer (TR-FRET).

Mobility shift assays use a fluorophore-labelled substrate and employ electrophoresis to separate the more-negatively charged phosphorylated product from the unphosphorylated substrate. The product is subsequently quantified by measuring fluorescence intensity (Figure 1.4.7).⁷⁶

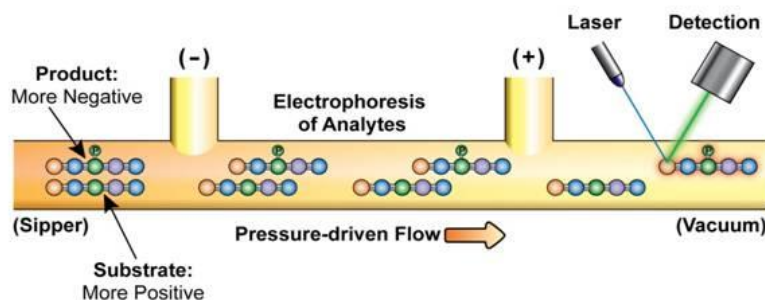


Figure 1.4.7. Illustration of mobility shift assays.

Competition binding assays quantify the binding of small molecules to the kinase active site (ATP site), rather than measuring catalytic product. Typically, a standard active site-binding inhibitor is immobilized on a solid support or conjugated to a tracer molecule. This standard inhibitor competes with the test compound for binding to the protein kinase domain. Competition binding assays are performed in the absence of ATP and peptide substrate.

Thermal shift assays operates on the principle that ligand binding alters the thermal stability of proteins. A ligand bound to a protein (e.g. to its active site) has the propensity to increase its thermal stability and hence its melting/denaturation temperature (T_m) through newly formed ligand-protein interactions. Monitoring protein denaturation upon heating via fluorescence-based detection methods allows detection of ligand binding to proteins such as enzymes.⁷⁷

The value of non-ATP competitive kinase inhibitors (like type III inhibitors) has been widely recognized; however, currently there are no direct screening methods available to test such compounds. In many cases, compounds initially found via high throughput screening (HTS) are identified as non-ATP site binders by enzyme kinetics and biophysical assays, but the hypothesized binding mechanism must be confirmed by co-crystallization and the use of X-ray crystallography.⁷⁸

To investigate the mechanism of action of hypothetical type III inhibitors different methods have been developed, however there is still a need for improvement. The most accessible of these methods are the kinetic experiments which are capable of

identifying an allosteric inhibitor through double-reciprocal plotting of data according to the Lineweaver–Burk equation and plot (Figure 1.4.8).

$$\frac{1}{V} = \frac{K_m}{V_{\max}} \frac{1}{[S]} + \frac{1}{V_{\max}}$$

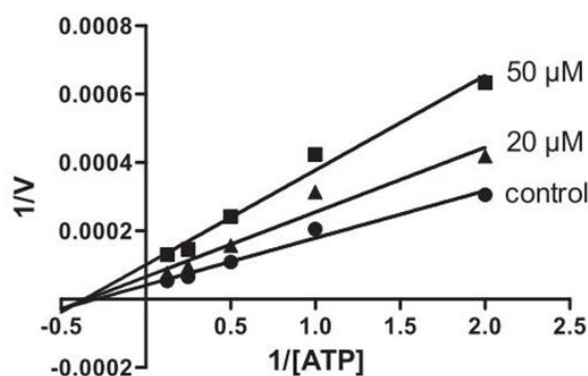


Figure 1.4.8 Lineweaver–Burk equation and an example of Lineweaver–Burk plot which shows non-competitive inhibition.⁷⁹ Here, $V_{\max}$ represents the maximum rate achieved by the system, at saturating substrate concentration; the Michaelis constant K_m is the substrate concentration at which the reaction rate is half of $V_{\max}$ and $[S]$ is the concentration of the substrate.

In order to determine a compound's binding kinetics to a non-ATP site, research teams have developed a surface plasmon resonance (SPR) competition assay, measuring the binding of test compounds in the presence of a high affinity ATP site inhibitor.⁷⁸ Others have performed affinity based screening followed by mass spectroscopy (ASMS).⁷⁸

Screening of kinases is frequently performed *in vitro* using only the catalytic domain of the kinase; however, reproducing the results first obtained in a HTS or biochemical assay both with full length kinases and in a cellular environment is fundamental, since the presence of additional domains or subunits may affect the function of the kinase and its interaction with its related proteins.⁸⁰

To obtain a thorough understanding of the selectivity profile of a particular drug, multiple screens should be conducted in both enzymatic and cellular assays.⁸¹ An important challenge to be addressed is the production of compounds which display

potency both in cells as well as assays utilizing the purified enzyme. Non-kinase targets of protein kinase inhibitors are very common drawbacks.⁸² In this circumstance it is crucial to establish that the cellular effect observed results from inhibiting the desired target and not from action at other kinases that could, directly or indirectly, contribute to the observed biology in cell proliferation phenotypic assays.^{82,83}

1.5. Previous Work Leading Up to the Present Research

Considering the structural similarities between the kinases' inhibitor sorafenib and 4,4'-bis-guanidinium derivatives that had been prepared in Rozas' laboratory,⁸⁴ coupled with their interesting cytotoxicity displayed in different cancer cell lines, five new families were prepared. These consisted of unsubstituted and aryl-substituted 3,4'-bis-guanidinium (**Families A** and **C-F**, Figure 1.5.1), 3,4'-bis-2-aminoimidazolinium (**Family B**, Figure 1.5.1) and 3-acetamide-4'-(4-chloro-3-trifluoromethylphenyl) guanidinium derivatives (**Family G**, Figure 1.5.1). Each of these families were subsequently studied to understand whether this activity was a result of kinase inhibition.⁸⁴

A commercial test on a small panel of kinases indicated that the 4,4'-bis-guanidiniums, which had previously been shown to be cytotoxic, did not inhibit the tested kinases (CK-1 δ , GSK-3, ERK 1/2, p38 MAPK, VEGFR); however, some of the newly prepared 3,4'-substituted derivatives (in particular **Families E** and **F**) showed up to 99% of inhibition at 10 μ M for the C-Raf/MEK-1 pathway and high toxicity in cancer cell lines, for which a potential mode of binding was proposed from *in silico* docking studies.⁸⁵

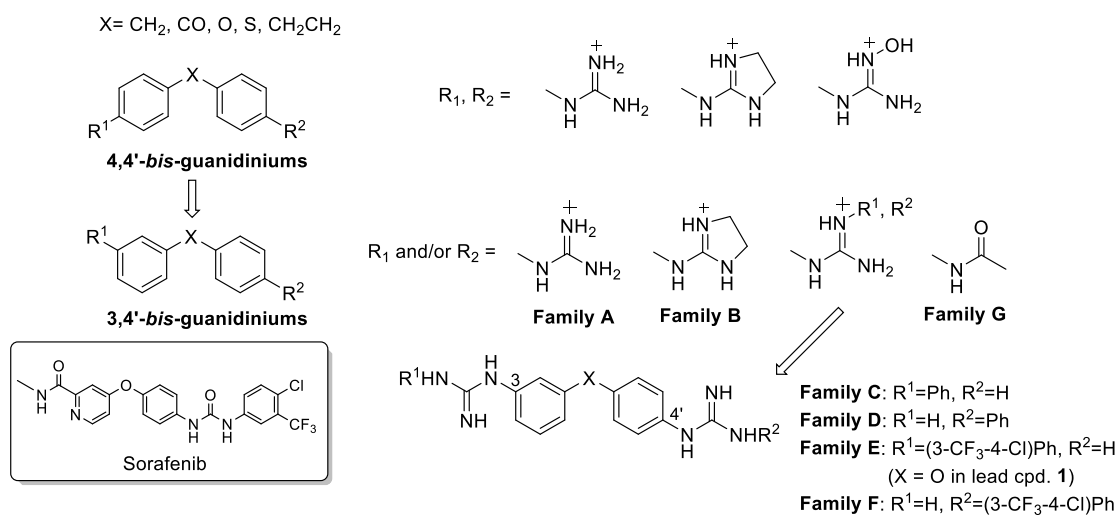


Figure 1.5.1: Different kinase inhibitor families previously prepared in Rozas' group.

These docking studies showed that when ATP was present in the active kinase (3PP1 – MEK1), all resulting complexes show a guanidinium-phosphate electrostatic interaction, with the most favourable complexes orienting the 4-chloro-3-trifluoromethylphenyl moiety into a hydrophobic pocket of the kinase (Figure 1.5.2.).

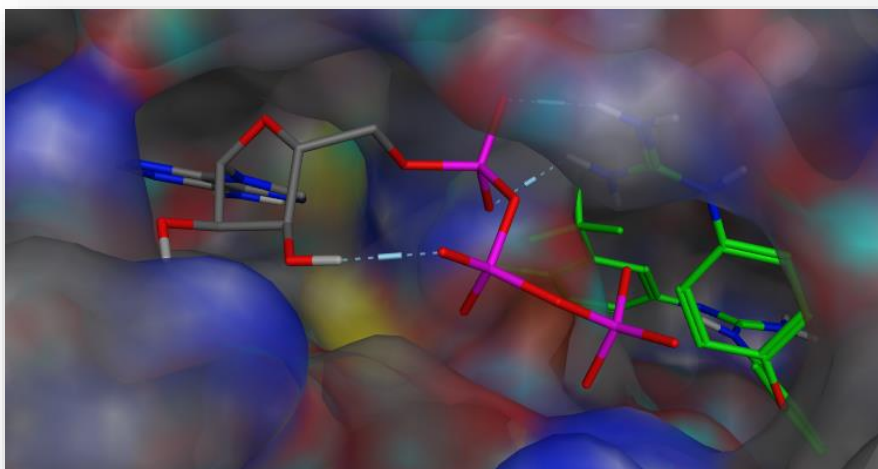


Figure 1.5.2: VDW surface of 3PP1 binding site, showing ATP, interaction with compound 1 and hydrophobic pocket where 4-chloro-3-trifluoromethylphenyl moiety is positioned.

Taking this as the preferred mode of binding, and considering both the results from the commercial panel and the cell cytotoxicity, the activity of the ligands previously prepared by Rozas and coworkers could potentially be explained as follows:

- i) **Families A** and **B** have poor affinity for kinases, as they place a cationic group in the hydrophobic pocket;
- ii) Phenylguanidiniums from **Families C** and **D** have higher affinity for C-Raf/MEK-1 since they can place the phenyl ring in the hydrophobic pocket, and maintain a guanidinium-phosphate electrostatic interaction;
- iii) Improvements in the cytotoxic activity of the compounds were obtained in **Family E** (as in lead compound **1**), when the 4-chloro-3-trifluoromethylphenyl moiety was inserted in the guanidinium moiety in the 3-position instead of the 4'-position, with no significative differences among different X links.
- iv) The largest C-Raf/MEK-1 affinity was shown by **Families F** which positions the hydrophobic 4-chloro-3-trifluoromethylphenyl ring in a hydrophobic pocket and the unsubstituted guanidinium in contact with a phosphate group of ATP.

Based on these results, it was proposed to design novel and improved guanidine-like derivatives of **Families E** and **F**, which would subsequently undergo a more detailed biological evaluation to elucidate their activity as PKIs.

1.6. References

1. <http://publications.iarc.fr/Non-Series-Publications/World-Cancer-Reports/World-Cancer-Report-2014>.
2. <http://globocan.iarc.fr/>.
3. Siegel, R. L.; Miller, K. D.; Jemal, A., *CA Cancer J. Clin.* **2016**, 66, 7.
4. Ferlay, J.; Soerjomataram, I.; Dikshit, R.; Eser, S.; Mathers, C.; Rebelo, M.; Parkin, D. M.; Forman, D.; Bray, F., *Int. J. Cancer* **2015**, 136, E359.
5. Nowell, P. C., *Current Contents/Life Sciences* **1988**, 19.
6. Yates, L. R.; Campbell, P. J., *Nat. Rev. Genet.* **2012**, 13, 795.

7. Tabassum, D. P.; Polyak, K., *Nat. Rev. Cancer* **2015**, *15*, 473.
8. Greaves, M.; Maley, C. C., *Nature* **2012**, *481*, 306.
9. Hanahan, D.; Weinberg, R. A., *Cell* **2000**, *100*, 57.
10. Matthews, D. J.; Gerristen, M. E., *Targeting protein kinases for cancer therapy* John Wiley & Sons, **2010**.
11. Hanahan, D.; Weinberg, R. A., *Cell* **2011**, *144*, 646.
12. Hanks, S. K.; Hunter, T., *FASEB J.* **1995**, *9*, 576.
13. Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S., *Science* **2002**, *298*, 1912.
14. Liao, J. J.-L., *J. Med. Chem.* **2007**, *50*, 409.
15. <https://www.cellsignal.com/>
16. Cohen, P., *Nat. Rev. Drug Discov.* **2002**, *1*, 309.
17. Futreal, P. A.; Coin, L.; Marshall, M.; Down, T.; Hubbard, T.; Wooster, R.; Rahman, N.; Stratton, M. R., *Nat. Rev. Cancer* **2004**, *4*, 177.
18. Gross, S.; Rahal, R.; Stransky, N.; Lengauer, C.; Hoeflich, K. P., *J. Clin. Invest.* **2015**, *125*, 1780.
19. Cohen, P.; Alessi, D. R., *ACS Chem. Biol.* **2013**, *8*, 96.
20. Zhang, J. M.; Yang, P. L.; Gray, N. S., *Nat. Rev. Cancer* **2009**, *9*, 28.
21. Lahiry, P.; Torkamani, A.; Schork, N. J.; Hegele, R. A., *Nat. Rev. Genet.* **2010**, *11*, 60.
22. Chico, L. K.; Van Eldik, L. J.; Watterson, D. M., *Nat. Rev. Drug Discov.* **2009**, *8*, 892.
23. Kornev, A. P.; Taylor, S. S.; Ten Eyck, L. F., *Proc. Natl. Acad. Sci.* **2008**, *105*, 14377.
24. Knighton, D. R.; Zheng, J. H.; Teneyck, L. F.; Xuong, N. H.; Taylor, S. S.; Sowadski, J. M., *Science* **1991**, *253*, 414.
25. Knighton, D. R.; Zheng, J. H.; Teneyck, L. F.; Ashford, V. A.; Xuong, N. H.; Taylor, S. S.; Sowadski, J. M., *Science* **1991**, *253*, 407.
26. Taylor, S. S.; Kornev, A. P., *Trends Biochem. Sci.* **2011**, *36*, 65.
27. Adams, J. A., *Biochemistry* **2003**, *42*, 601.
28. Nolen, B.; Taylor, S.; Ghosh, G., *Molecular cell* **2004**, *15*, 661.

29. Adams, J. A., *Chem. Rev.* **2001**, *101*, 2271.
30. Kornev, A. P.; Haste, N. M.; Taylor, S. S.; Ten Eyck, L. F., *Proc. Natl. Acad. Sci.* **2006**, *103*, 17783.
31. Bollag, G.; Tsai, J.; Zhang, J. Z.; Zhang, C.; Ibrahim, P.; Nolop, K.; Hirth, P., *Nat. Rev. Drug Discov.* **2012**, *11*, 873.
32. McClendon, C. L.; Kornev, A. P.; Gilson, M. K.; Taylor, S. S., *Proc. Natl. Acad. Sci.* **2014**, *111*, 16973.
33. Zhang, L.; Wang, J.-C.; Hou, L.; Cao, P.-R.; Wu, L.; Zhang, Q.-S.; Yang, H.-Y.; Zang, Y.; Ding, J.-P.; Li, J., *Sci. Rep.* **2015**, *5*, 10115.
34. Johnson, L. N.; Noble, M. E.; Owen, D. J., *Cell* **1996**, *85*, 149.
35. Madhusudan; Akamine, P.; Xuong, N.-H.; Taylor, S. S., *Nat. Struct. Mol. Biol.* **2002**, *9*, 273.
36. Kannan, N.; Neuwald, A. F., *J. Mol. Biol.* **2005**, *351*, 956.
37. Taylor, S. S.; Keshwani, M. M.; Steichen, J. M.; Kornev, A. P., *Philos. Trans. R. Soc. London, Ser. B* **2012**, *367*, 2517.
38. Hu, J. C.; Stites, E. C.; Yu, H. Y.; Germino, E. A.; Meharena, H. S.; Stork, P. J. S.; Kornev, A. P.; Taylor, S. S.; Shaw, A. S., *Cell* **2013**, *154*, 1036.
39. Hu, J.; Ahuja, L. G.; Meharena, H. S.; Kannan, N.; Kornev, A. P.; Taylor, S. S.; Shaw, A. S., *Mol. Cell. Biol.* **2015**, *35*, 264.
40. Kolch, W., *Biochem. J.* **2000**, *351*, 289.
41. Jambrina, P. G.; Rauch, N.; Pilkington, R.; Rybakova, K.; Nguyen, L. K.; Kholodenko, B. N.; Buchete, N. V.; Kolch, W.; Rosta, E., *Angew. Chem. Int. Ed.* **2016**, *55*, 983.
42. Wan, P. T.; Garnett, M. J.; Roe, S. M.; Lee, S.; Niculescu-Duvaz, D.; Good, V. M.; Jones, C. M.; Marshall, C. J.; Springer, C. J.; Barford, D.; Marais, R.; Cancer Genome, P., *Cell* **2004**, *116*, 855.
43. Lavoie, H.; Therrien, M., *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 281.
44. Baccarini, M., *FEBS Lett.* **2005**, *579*, 3271.
45. Zhang, B.-H.; Guan, K.-L., *EMBO J.* **2000**, *19*, 5429.
46. Samatar, A. A.; Poulikakos, P. I., *Nat. Rev. Drug Discov.* **2014**, *13*, 928.
47. Neuzillet, C.; Tijeras-Raballand, A.; de Mestier, L.; Cros, J.; Faivre, S.; Raymond, E., *Pharmacol. Ther.* **2014**, *141*, 160.
48. Saini, K. S.; Loi, S.; de Azambuja, E.; Metzger-Filho, O.; Saini, M. L.; Ignatiadis, M.; Dancey, J. E.; Piccart-Gebhart, M. J., *Cancer Treat. Rev.* **2013**, *39*, 935.

49. Neel, N. F.; Martin, T. D.; Stratford, J. K.; Zand, T. P.; Reiner, D. J.; Der, C. J., *Genes & Cancer* **2011**, 2, 275.
50. Rawlings, J. S.; Rosler, K. M.; Harrison, D. A., *J. Cell Sci.* **2004**, 117, 1281.
51. McCubrey, J. A.; Steelman, L. S.; Chappell, W. H.; Abrams, S. L.; Wong, E. W. T.; Chang, F.; Lehmann, B.; Terrian, D. M.; Milella, M.; Tafuri, A.; Stivala, F.; Libra, M.; Basecke, J.; Evangelisti, C.; Martelli, A. M.; Franklin, R. A., *Biochim. Biophys. Acta* **2007**, 1773, 1263.
52. Wang, H. G.; Rapp, U. R.; Reed, J. C., *Blood* **1996**, 88, 1785.
53. Steelman, L. S.; Pohnert, S. C.; Shelton, J. G.; Franklin, R. A.; Bertrand, F. E.; McCubrey, J. A., *Leukemia* **2004**, 18, 189.
54. Fedorov, O.; Muller, S.; Knapp, S., *Nat. Chem. Biol.* **2010**, 6, 166.
55. Francavilla, C.; Lupia, M.; Tsafou, K.; Villa, A.; Kowalczyk, K.; Jersie-Christensen, R. R.; Bertalot, G.; Confalonieri, S.; Brunak, S.; Jensen, L. J.; Cavallaro, U.; Olsen, J. V., *Cell Rep.* **2017**, 18, 3242.
56. Wu, P.; Nielsen, T. E.; Clausen, M. H., *Drug Discov. Today* **2016**, 21, 5.
57. Mullard, A., *Nat. Rev. Drug Discov.* **2016**, 15, 73.
58. Mullard, A., *Nat. Rev. Drug Discov.* **2017**, 16, 73.
59. Simard, J. R.; Kluter, S.; Grutter, C.; Getlik, M.; Rabiller, M.; Rode, H. B.; Rauh, D., *Nat. Chem. Biol.* **2009**, 5, 394.
60. Fang, Z.; Grütter, C.; Rauh, D., *ACS Chemical Biology* **2013**, 8, 58.
61. Zhao, Z.; Liu, Q.; Bliven, S.; Xie, L.; Bourne, P. E., *J. Med. Chem.* **2017**, 60, 2879.
62. Cox, K. J.; Shomin, C. D.; Ghosh, I., *Future Med. Chem.* **2010**, 3, 29.
63. Druker, B. J.; Tamura, S.; Buchdunger, E.; Ohno, S.; Segal, G. M.; Fanning, S.; Zimmermann, J.; Lydon, N. B., *Nat. Med.* **1996**, 2, 561.
64. Davis, M. I.; Hunt, J. P.; Herrgard, S.; Ciceri, P.; Wodicka, L. M.; Pallares, G.; Hocker, M.; Treiber, D. K.; Zarrinkar, P. P., *Nat. Biotechnol.* **2011**, 29, 1046.
65. Gilmartin, A. G.; Bleam, M. R.; Groy, A.; Moss, K. G.; Minthorn, E. A.; Kulkarni, S. G.; Rominger, C. M.; Erskine, S.; Fisher, K. E.; Yang, J.; Zappacosta, F.; Annan, R.; Sutton, D.; Laquerre, S. G., *Clin. Cancer Res.* **2011**, 17, 989.
66. Shan, Y.; Kim, E. T.; Eastwood, M. P.; Dror, R. O.; Seeliger, M. A.; Shaw, D. E., *J. Am. Chem. Soc.* **2011**, 133, 9181.
67. Lombardo, L. J.; Lee, F. Y.; Chen, P.; Norris, D.; Barrish, J. C.; Behnia, K.; Castaneda, S.; Cornelius, L. A. M.; Das, J.; Doweyko, A. M.; Fairchild, C.; Hunt, J. T.;

Inigo, I.; Johnston, K.; Kamath, A.; Kan, D.; Klei, H.; Marathe, P.; Pang, S.; Peterson, R.; Pitt, S.; Schieven, G. L.; Schmidt, R. J.; Tokarski, J.; Wen, M.-L.; Wityak, J.; Borzilleri, R. M., *J. Med. Chem.* **2004**, *47*, 6658.

68. Wada, M.; Horinaka, M.; Yamazaki, T.; Katoh, N.; Sakai, T., *PLOS ONE* **2014**, *9*, e113217.

69. Okaniwa, M.; Hirose, M.; Arita, T.; Yabuki, M.; Nakamura, A.; Takagi, T.; Kawamoto, T.; Uchiyama, N.; Sumita, A.; Tsutsumi, S.; Tottori, T.; Inui, Y.; Sang, B.-C.; Yano, J.; Aertgeerts, K.; Yoshida, S.; Ishikawa, T., *J. Med. Chem.* **2013**, *56*, 6478.

70. De Palma, M.; Hanahan, D., *Mol. Oncol.* **2012**, *6*, 111.

71. Haberland, M.; Montgomery, R. L.; Olson, E. N., *Nat. Rev. Genet.* **2009**, *10*, 32.

72. Ciceri, P.; Müller, S.; O'Mahony, A.; Fedorov, O.; Filippakopoulos, P.; Hunt, J. P.; Lasater, E. A.; Pallares, G.; Picaud, S.; Wells, C.; Martin, S.; Wodicka, L. M.; Shah, N. P.; Treiber, D. K.; Knapp, S., *Nat. Chem. Biol.* **2014**, *10*, 305.

73. Sharma, P.; Allison, James P., *Cell* **2015**, *161*, 205.

74. Melero, I.; Berman, D. M.; Aznar, M. A.; Korman, A. J.; Gracia, J. L. P.; Haanen, J., *Nat. Rev. Cancer* **2015**, *15*, 457.

75. Simard, J. R.; Rauh, D., *Methods Enzymol.* **2014**, *548*, 147.

76. Ma, H.; Deacon, S.; Horiuchi, K., *Expert Opin. Drug Discov.* **2008**, *3*, 607.

77. Rudolf, A. F.; Skovgaard, T.; Knapp, S.; Jensen, L. J.; Berthelsen, J., *PLOS ONE* **2014**, *9*, e98800.

78. Gavrin, L. K.; Saiah, E., *MedChemComm* **2013**, *4*, 41.

79. Gao, Y.; Ye, D.-Y.; Zhou, W.-C.; Chu, Y., *Eur. J. Med. Chem.* **2017**, *135*, 370.

80. Williams, K. P.; Scott, J. E., *High Throughput Screening: Methods and Protocol, Second Edition* **2009**, 565, 107.

81. Smyth, L. A.; Collins, I., *J. Biol. Chem.* **2009**, *2*, 131.

82. Munoz, L., *Nat. Rev. Drug Discov.* **2017**, *16*, 424.

83. Smith, G. K.; Wood, E. R., *Drug Discov. Today Technol.* **2010**, *7*, e13.

84. Diez-Cecilia, E.; Kelly, B.; Perez, C.; Zisterer, D. M.; Nevin, D. K.; Lloyd, D. G.; Rozas, I., *Eur. J. Med. Chem.* **2014**, *81*, 427.

85. Diez-Cecilia, E.; Carson, R.; Kelly, B.; van Schaeybroeck, S.; Murray, J. T.; Rozas, I., *Bioorg. Med. Chem. Lett.* **2015**, *25*, 4287.

Chapter 2

Chapter 2 – Objectives

Previous work in Rozas' group found different structurally related compounds to be possible protein kinase inhibitors and, in particular, some of the compounds showed up to 99% inhibition at 10 μM of the C-Raf (Raf-1)/MEK-1 protein kinase pathway. Furthermore, a structurally related compound, lead compound **1** (Figure 2.1), was found to be able to inhibit B-Raf protein kinase at 20 μM in an *in vitro* assay against MEK-1 peptide and to downregulate several kinases and substrates downstream of EGFR (Src, AKT, B-Raf, MEK and ERK) in western blots studies. Moreover, lead compound **1** showed promising IC_{50} values of around 10 μM in a panel of different cancer cell lines and negative TR-FRET (time resolved fluorescence resonance energy transfer) B-Raf competitive binding assay excluded the possibility of a type I or type II inhibitor. *In silico* docking studies conducted in B-Raf, C-Raf and MEK showed how the lipophilic 4-chloro-3-trifluoromethylphenyl ring of lead compound **1** can occupy an hydrophobic pocket close to the ATP hinge region and how, when ATP was present (as in an active kinase), the top-ranked poses were the ones showing a guanidinium-phosphate electrostatic interaction (Figure 1.5.2.). Taking all this into account, it was possible to hypothesise that lead compound **1** could act as a type III allosteric inhibitor toward the inhibition of the Ras/Raf/MEK/ERK kinase pathway.

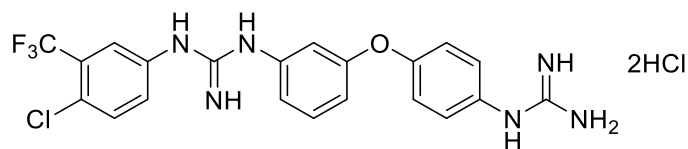


Figure 2.1. Chemical structure of lead compound **1**.

Considering the promising results of lead compound **1** as protein kinase inhibitor, several objectives (computational, synthetic and biological objectives) were proposed

to enhance the biological results obtained and improve the understanding of this lead compound mode of action and they are outlined in the next sections.

2.1. Computational Objectives

1) In previous research, Rozas' group had proposed that positively charged guanidinium-containing compounds could interact with negative ATP-phosphates present in all active protein kinases. Based on that and in a collaboration with Dr. Cristina Trujillo (School of Chemistry, TCD), we proposed to perform a theoretical study to understand the possible interaction modes between the phosphates in the ATP molecule and different guanidinium systems by means of Density Functional Theory (DFT) methods.

2) In a typical structure-based drug discovery approach, we need to know the structure of both target and ligand to carry out docking studies. Previous docking studies in Rozas' group had been conducted with lead compound **1** and, among the others, MEK protein kinase, the only protein kinase of the Ras/Raf/MEK/ERK pathway to be co-crystallised with ATP and Mg^{2+} , essential elements for the allosteric catalytic inhibition of active kinases. Given the fact lead compound **1** was found able to inhibit B-Raf *in vitro*, and considering that no crystal structure is present in the literature for this particular kinase including ATP and Mg^{2+} , we proposed to perform classical molecular dynamics (MD) simulations in collaboration with Dr. Trujillo and Prof. Eric Henon (University of Reims Champagne-Ardenne, Reims, France) to obtain a theoretical model of an active form of ATP-containing B-Raf protein kinase. Such a model facilitated the corresponding docking simulations with empirical and semi-empirical methods to investigate the mode of binding of compound **1** and other potentially improved derivatives. This collaborative study was supported by funds obtained from the Ulysses Scheme (Irish Research Council/French Embassy in Ireland).

2.2. Synthetic Objectives

Based on previous results and the computational study, new derivatives of compound **1** were designed and synthesized (see Figure 2.2.1) by rationally introducing changes at the three different moieties identified in this lead compound: 1) the hydrophobic moiety, 2) the link of the molecule and 3) the polar moiety.

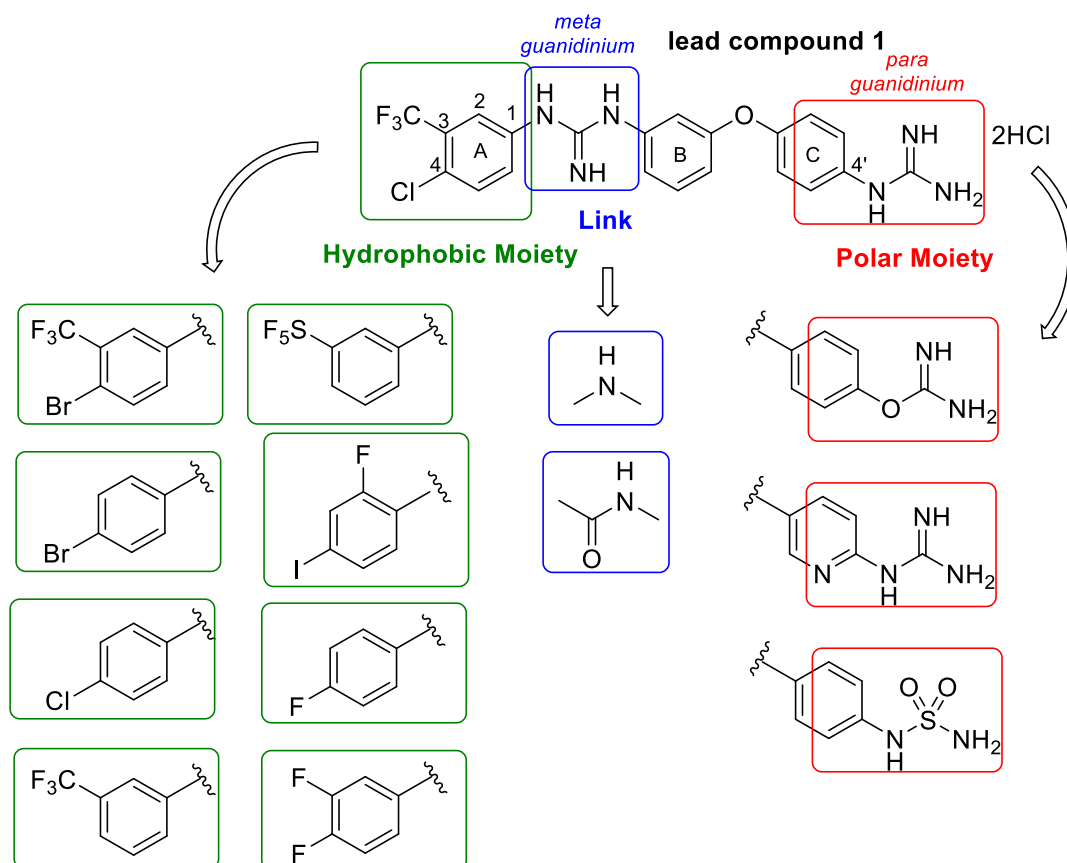


Figure 2.2.1. Proposed structural modifications of lead compound **1** investigated in this project and applied to the synthesised derivatives.

1) In order to probe the hydrophobic pocket highlighted in the docking studies and to find the structural requirements necessary to improve affinity, the hydrophobic moiety, which in lead compound **1** consists of the 4-chloro-3-trifluoromethylphenyl ring (aromatic ring A), was substituted with other halogen atoms/groups (Figure 2.2.1, green box).

2) Previous docking studies in Rozas' group indicated that shorter molecules would fit better into the binding site (of the MEK protein), hence the distance between phenyl rings A and B, and hence their chemical link, was modified. Thus, different *links* were explored changing the *meta* guanidinium present in lead compound **1** to a secondary amine or an amide, in order to investigate both the ideal length of the molecule and the subsequent right orientation of the hydrophobic moiety (Figure 2.2.1., blue box).

3) Previous docking studies in Rozas' group showed how the polar moiety, which in lead compound **1** comprises a phenylguanidinium cation (aromatic ring C), is responsible for forming ionic and hydrogen bond (HB) interactions with the phosphate groups of the ATP molecule present in all active protein kinases. By changing the phenyl ring into a 2-pyridinyl ring (Figure 2.2.1., red box), we aimed to create an intramolecular hydrogen bond (IMHB) between the pyridine nitrogen and the guanidinium cation. This locks the structural conformation resulting in a different orientation when interacting with ATP and helping the investigation of its preferred binding mode. Similarly, changing the phenylguanidinium system into phenylisouronium or sulfonamide functionalities will help to investigate the requirements for a better binding.

All new derivatives were synthesised by using a combination of synthetic methodology previously established within Rozas' group and also novel synthetic pathways. Properly functionalised amines were prepared to facilitate the introduction of the final Boc-protected guanidines utilising the established mercury (II) chloride chemistry developed in Rozas' group.

Introduction of different hydrophobic moieties was achieved through synthesis of *N*-aryl-*N'*-Boc-protected thiourea derivatives and standard mercury (II) chloride chemistry. 2-Amino pyridine derivatives were synthesised through S_NAr and both mercury (II) chloride and Buchwald-Hartwig chemistry. Similarly, isouronium and sulfonamide derivatives were accessed through properly functionalised phenols or amines and mercury (II) chloride chemistry or reaction with sulfamoyl chloride. All final Boc-protected derivatives were deprotected using different organic and inorganic solutions and the purity (>95%) of the final hydrochloride was assessed by HPLC

before biological assays. An in depth description of the design and synthesis of each compound will be discussed in detail in Chapter 4.

2.3 Biological Objectives

1) The ability of the new derivatives to affect human cancer cell proliferation was evaluated in collaboration with Prof. Daniela Zisterer (School of Biochemistry & Immunology, TCD) through alamarBlue[®] assay first in human promyelocytic leukemia cells (HL-60). This methodology allows for a fast screening to find the best structural requirements and to define Structural Activity Relationship (SAR) for a better cytotoxic activity. Some chosen derivatives from this assay were further investigated for cytotoxicity in a panel of different cancer cell lines, including breast cancer (MCF-7), cervical cancer (HeLa), colorectal cancer (HCT116 and HKH-2) and multiple myeloma (NCI-H929 and U266B1). Particularly, the use of Ras mutated cancer cell lines gave information about the cytotoxicity of our compounds in particularly poorly treatable cancers. Investigation of toxic effects of our derivatives in non-tumorigenic epithelial cell line MCF10A and human platelets helped define their biological profile.

2) Moreover, the mechanism by which the compounds prepared cause cell-death was evaluated and their abilities of inducing apoptosis were investigated through flow cytometry analysis and western blotting.

3) Finally, the molecular targets of our compounds were investigated by evaluating binding and inhibition of protein kinases of the Ras/Raf/MEK/ERK kinase pathway in collaboration with Prof. Stefan Knapp (SGC, Goethe-Universität Frankfurt University) and Prof. Walter Kolch (School of Medicine, Systems Biology Ireland, UCD).

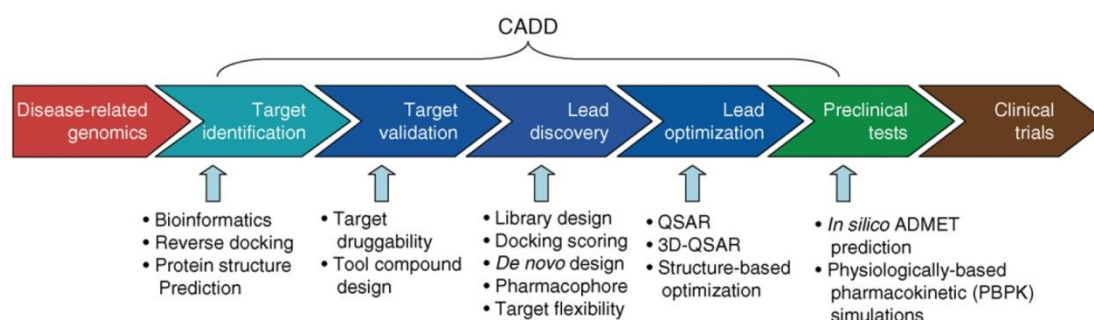
Chapter 3

Chapter 3 - Computational Modelling

3.1. Computational chemistry in drug discovery

Historically, bioactive compounds were identified in a typical phenotypic screening, i.e. assays where a given compound alters the phenotype of a cell or an organism in a desired manner. In the phenotypic screening approach for drug discovery, also known as ‘classical pharmacology’, the molecular mechanism and protein target can remain unknown even after the drug’s activity and efficacy are determined.¹ Even though the significant reduction in the number of approved drugs in the last decade has led to consider phenotypic screens as a renewed approach for drug discovery, nowadays therapeutics are discovered based on biological targets, genetic studies, transgenic animal models, molecular biology, gene technology, and protein science. Hence, the structure of the target biomolecule and its interactions with a bioactive ligand is of great practical interest.²

Currently, one of the most widely applied techniques in drug discovery is computational chemistry and molecular modelling.³ Computational tools find application at many different stages of the drug discovery and development process (Figure 3.1.1) and the central paradigm described previously in the Introduction section of structure-based drug discovery has allowed the widespread use of ‘chemistry competitive’ and reliable molecular modelling techniques.



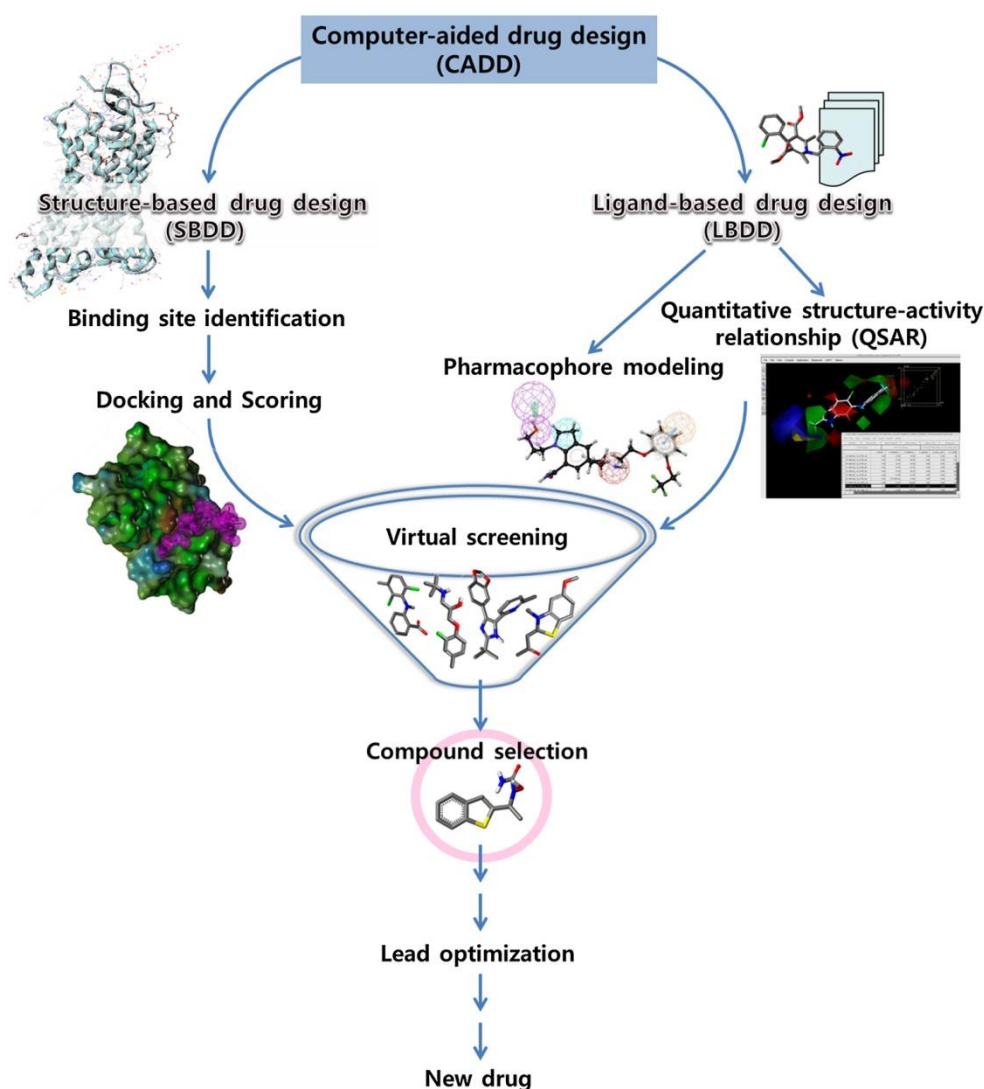


Figure 3.1.1. The drug discovery process and computer aided drug design (CADD) in drug discovery pipeline. Adapted from Macalino *et al*, 2015.⁴

A typical drug discovery and development pipeline nowadays consist of sequential steps where computer aided drug design (CADD) can fit at different time-points in order to help the finding of new lead candidates that will go into clinical trials.

First, a macromolecular target that plays a fundamental role in the pathology under study must be found, process known as *target identification*. Genomics has recorded the sequence of nucleic acid bases in many genomes, and continuous bioinformatics analyses are identifying the coding regions. Biology studies aimed at validation of such target, *target validation*, can be helped by target druggability predictions where

different methods consider shape, size, chemical character and flexibility as parameters in assessing the druggability of a binding site (Figure 3.1.1.).

Then, the search for biologically active small molecule binders for that particular target (*hits*) begins, i.e. *hit finding stage*. Once found, these small molecules must be optimized in order to achieve better in vitro activity, selectivity, pharmacodynamic and pharmacokinetic properties (*hit to lead optimisation stage*).³ Especially at this point, from *hit* finding until *lead* discovery, computational methods can be used to identify chemical structures with a greater probability of being active, and they can be mainly divided in two classes: ligand-based drug design (LBDD) and structure-based drug design (SBDD), Figure 3.1.1.

Ligand-based drug design (LBDD) consists of a series of techniques used for finding new compounds based solely on the structure of known biologically active small molecules. In contrast, structure-based drug design (SBDD) exploits the knowledge of the 3D structure of the aimed biological receptor or target.³

Finally, computational modelling is also helpful in the preclinical phase of the drug discovery pipeline, with *in silico* ADMET predictions or physiologically based pharmacokinetic (PBPK) simulations. A deep insight on how to change the chemistry to alter ADMET related parameters can be gained by simply knowing the binding mode of the candidate and exploiting it to, for example, functionalize the different scaffolds with solubilizing groups exposed to solvent.

3.2. Computational methods

All the different techniques used in both Structure-based and Ligand-based drug design that can guide the drug discovery process make use of particular computational methods. Overall we can distinguish between methods based on quantum mechanics (e.g. ‘*ab initio*’, Density Functional Theory – DFT-) and methods based on classical mechanics (e.g. Molecular Mechanics – MM-, Molecular Dynamics – MD-).

Quantum chemistry is primarily focused in the application of quantum mechanics (QM) in physical models and chemical systems. It applies quantum mechanical principles and

equations to the study of molecules. QM is used to solve the wavefunction of molecular systems in the Schrödinger's equation:

$$H\Psi = E\Psi$$

All properties of the system (molecule, or collection of molecules) may be calculated by the application of various mathematical operators on the wavefunction (Ψ). The operator of most interest is the energy operator or Hamiltonian, indicated in the Schrödinger's equation as H .

The Schrödinger's equation only can be completely solved in mono-electronic systems; hence, QM methods have to rely on a series of approximations such as the Born–Oppenheimer approximation, which considers that the movement of electrons and nuclei can be taken into account independently; it is then possible to study only the movement of the atomic centres assuming that the electron distribution is always in equilibrium.

In the most popular kind of QM calculations performed on molecules, each molecular spin orbital is expressed as a linear combination of atomic orbitals (the LCAO approach)⁵:

$$\Psi_j = \sum c_{ij} \phi$$

Gaussian functions are used to model atomic orbitals (AOs) and construct molecular orbitals (MOs) using the LCAO approach, the level of theory can be chosen to provide the best trade-off between quality of results and computational cost for each specific application.

If solutions are generated without reference to experimental data, the methods are usually called *ab initio* (latin: “from the beginning”).⁶ As *ab initio* methods can be computationally very demanding, semi-empirical approximations were introduced, which make use of experimentally determined data to avoid the calculation of certain terms.³ Density Functional Theory (DFT) provides a third class of QM methods, which consider that the energy and other properties of a system can be computed from its electron density (ρ).

A number of methods have been described to enable the extension of QM studies to large molecular systems, but their computational cost usually constrain them to relatively small organic molecules as drug-like inhibitors.⁷ The power and versatility of QM is anyway largely spread. QM can either be used to calculate energies and optimize structures or to calculate molecular properties.⁸ Very often the bioactive conformation of a ligand at the active site of an enzyme does not correspond to the absolute conformational minima, hence the internal energy of the ligand may play a major role in the structure-activity relationship (SAR) of a series.⁹ Additionally, QM can accurately calculate the energy associate with a configuration of a system, considering also the solvation effect.⁸ The study of the ligand can be useful also for reactivity indices, which may correlate with the biological activity to achieve the SARs for a series.³ Moreover, the computational procedures that are based on QM/MM, combine the strengths of both QM (accuracy) and molecular mechanics (MM) (efficiency) methods, and are widely employed to model chemical reactions and other electronic processes in biomolecular systems.⁸

Very successful and widely applied approaches in molecular modelling are also the parametric or molecular mechanics (MM) methods. They consist of using a set of molecules (represented by balls –atoms- and springs –bonds-) to derive rules and parameters of more general use, as force-fields. The parameters describing the force-field are derived from QM results and/or experimental data. Hence, QM is used routinely in optimizing geometries, fitting the torsion parameters, and deriving atomic charges for proteins, DNA, and in particular small molecules.⁸

The following equation describes a typical force field equation (AMBER¹⁰), which represents the potential energy function of the system:

$$V_{\text{AMBER}} = \sum_i^{n_{\text{bonds}}} b_i (r_i - r_{i,\text{eq}})^2 + \sum_i^{n_{\text{angles}}} a_i (\theta_i - \theta_{i,\text{eq}})^2 + \sum_i^{n_{\text{dihedrals}}} \sum_n^{n_{i,\text{max}}} (V_{i,n}/2) [1 + \cos(n\phi_i - \gamma_{i,n})] \\ + \sum_{i < j}^{n_{\text{atoms}}} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + \sum_{i < j}^{n_{\text{atoms}}} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}},$$

The bonded terms include the potentials due to bond distances, bond angles, dihedral angles and improper angles; the non-bonded potentials consists of Lennard–Jones and Coulombic terms for the dispersion–repulsion and electrostatic effects.

Utilising Newton's laws of movement it is possible to calculate a trajectory to develop a dynamic view of a molecular system (molecular dynamics –MD-) and force fields or other methods based on classical mechanics can be used to determine the energy of the system along the trajectory. MD allows to have an ensemble that can later on be used to calculate free energies of binding for a ligand–receptor complex¹¹, but also to sample protein stability, folding and molecular recognition.³

Computational methods described above and many others find applications in a series of drug design strategies, depending whether or not the target 3D structure is known, but these techniques will not be described in this thesis because they are outside its scope.

3.3. Theoretical studies proposed

As previously mentioned, during the last few years, the Rozas' group has reported a series of 3,4'-bis-guanidinium diaryl derivatives as potential inhibitors of the Ras/Raf/MEK/ERK pathway. Particularly, one of these bis-guanidine compounds (compound **1**, Figure 3.3.1) was found to inhibit wild-type B-Raf possibly through a type-III allosteric mechanism. Preliminary computational studies suggested that the guanidinium moiety of this ligand could interact with a phosphate of the ATP molecule present in the active kinase.^{12,13}

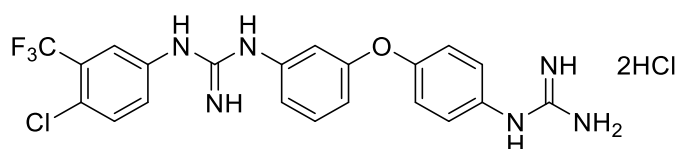


Figure 3.3.1. Structure of the bis-guanidine derivative compound **1**.

Structure-based drug design is regularly used to develop ATP-competitive inhibitors of B-Raf, since many crystallographic structures of this kinase complexed with competitive inhibitors have been reported in the last few years allowing optimization of

target-ligand interactions' requirements.^{14,15} However, a crystal structure of a B-Raf protein kinase complexed with ATP has yet to be resolved.¹⁶

For that reason, as anticipated in the Objective chapter, our study aimed at the investigation of the interactions between lead compound **1** and both an active wild type B-Raf protein kinase (containing ATP) and the ATP molecule itself. Therefore, we performed a long range MD study that allowed us to obtain a theoretical model of an active form of wild type B-Raf including ATP, which we then used to perform molecular docking with compound **1** in order to test our hypothesis of type III allosteric inhibition.¹⁷ Complementary, we also performed a QM study to theoretically analyse the interaction between the phosphates in the ATP molecule and the mentioned guanidinium ligand.¹⁸

3.4. Theoretical model of an active form of B-Raf protein kinase

Taking advantage of the large amount of crystallographic information on protein kinases reported in recent times, the understanding of their mechanism of regulation and activation has been advanced by the use of computational techniques such as molecular dynamic (MD) simulations.¹⁹ For example, recently published research has shown how phosphorylation of particular residues and ATP binding to protein kinases results in different structural changes.²⁰ Furthermore, Gosu *et al.* analysed the dynamic structural change of apo and ATP-bound IRAK4 kinase²¹ and recent MD studies on B-Raf gave insights into previously unexplored relationships between different types of B-Raf inhibitors.^{22,23} Lately, Kuzmanic *et al.* have reported enhanced MD simulations to study the canonical activation of p38a MAPK kinase.²⁴

Thus, considering that no crystal structure of ATP-bound B-Raf protein has ever been reported, we have performed a theoretical study to obtain a model of an active form of wild type B-Raf (comprising an ATP molecule and a phosphorylated residue), in collaboration with Prof. Eric Henon and his group (University of Reims Champagne-Ardenne, Reims, France). Taking into account that different hypotheses have been proposed on the B-Raf activation mechanism,^{25,26} in this study, we have:

- (i) chosen a related protein kinase, which crystal structure containing ATP is available in the Brookhaven Database, to model the actual location of ATP in the active site of B-Raf.
- (ii) focused on what has been defined as the major activation segment phosphorylation site in B-Raf, threonine Thr599 which is noted as Thp599, when phosphorylated.

A description of the most important structural features of B-Raf is necessary to understand the different aspects of this computational study and a brief synopsis has already been given in the Introduction for the structurally related PKA protein kinase. In Table 3.4.1 a summary of the discussed motifs with their function and an amino acid conversion between the previously mentioned PKA and B-Raf are presented. Additionally, Figure 3.4.1 illustrates the PKA residues mentioned in Table 3.4.1 to facilitate the reading.

Table 3.4.1. Relevant amino acid residues and their number conversion between PKA and B-Raf.

Acronym	Function	Relevant residues in protein	
		PKA	B-Raf
DFG	Contact with ATP via Mg^{2+} binding; catalysis, R spine formation, activation of kinase	Asp184, Phe185, Gly186	Asp594, Phe595, Gly596
HRD	Catalysis	Tyr164, Arg165, Asp166	His574, Arg575, Asp576
Catalytic loop	Catalysis	Lys168, Arg165	Lys578, Arg575
APE	Substrate binding	Ala206, Pro207, Glu208	Ala621, Pro622, Glu623
C-spine	Control activation of kinase	Val57, Ala70, Met128, Leu172, Leu173, Ile174, Leu227, Met231	Val471, Ala481, Leu537, Ile582, Phe583, Leu584, Val645, Leu649
R-spine	Control activation of kinase	Tyr164, Phe185, Leu95, Leu106	His574, Phe595, Leu505, Phe516
α F-helix	Hydrophobic spine, comprehends C-spine residues	Resid 215 - 235	Resid 635 - 653
α C-helix	Control activation of kinase, connection of different motifs	Glu91	Gln494, Phe498, Asn500, Glu501, Arg506, Lys507

β 3-sheet	Connection of different motifs	Lys72	Lys483
Phosphorylation site(s)	Control activation of kinase	Thr197	Thr599, Ser602, Ser605, Ser607, Ser614
RD pocket	Electrostatic interaction between phospho-residue	Arg165	Arg575, Arg603
Peptide binding region			Trp619, Tyr656, Pro655

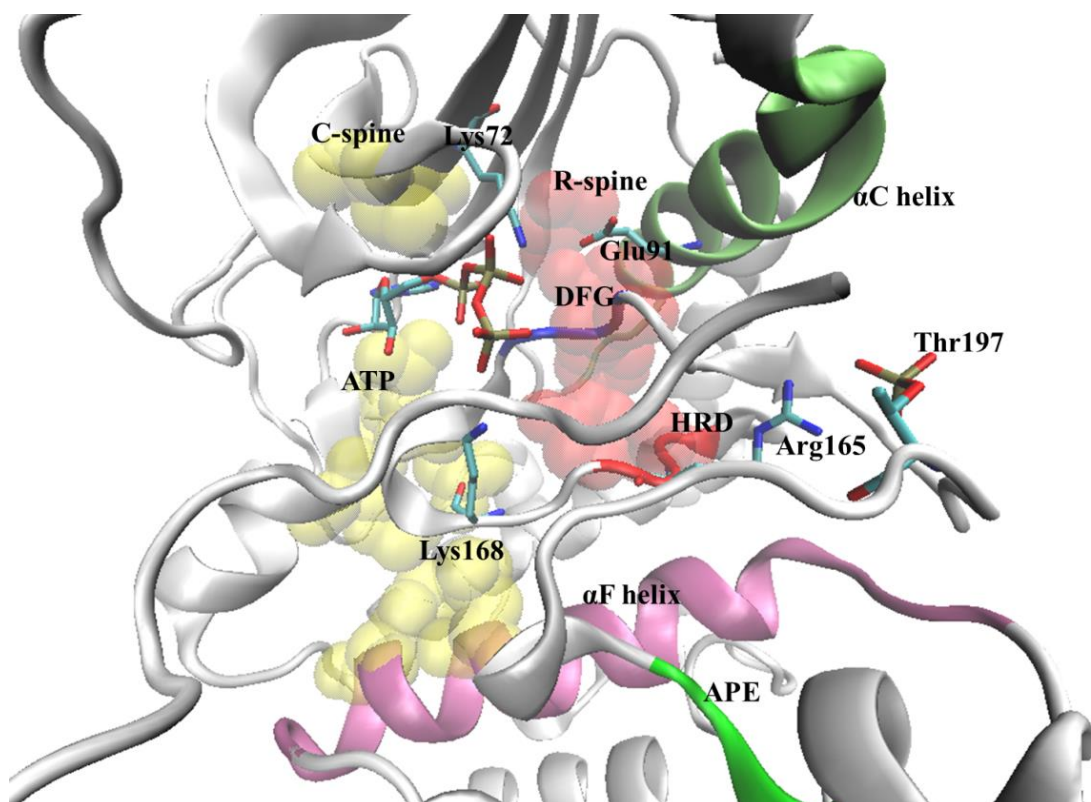


Figure 3.4.1. Structure of PKA (PDB: 1ATP); conserved structural parts and motifs are labelled according to the most recent literature. Yellow: C-spine; red: R-spine; lime: α C-helix; red: HRD; mauve: α F-helix, green: APE motif.^{25,27}

3.4.1. Generation of B-Raf systems

In order to obtain reliable and relevant information on how *ATP presence* and *phosphorylation* influence the active B-Raf structure, four different systems need to be taken into account (Table 3.4.2). Thus, **1uBRAFF_ATP** and **2pBRAFF_ATP** systems represent, respectively, B-Raf protein kinase in the presence of ATP with or without a phosphate at the phosphorylation site (Thr599). Additionally, **3uBRAFF** and **4pBRAFF**

systems represent B-Raf without an ATP molecule and without or with a phosphate at the phosphorylation site, respectively.

Since, as mentioned, no crystal structure of ATP-containing B-Raf has been reported, the initial ATP containing systems were constructed following a methodology used in the literature.²⁸ This methodology consists in the superposition of a homologous ATP-containing protein kinase (cAMP-dependent protein kinase complexed with Mn^{2+} and ATP; PDB: 1ATP) on a B-Raf structure (PDB: 3Q4C).²⁸ Thus, the ATP molecule was optimally located in the nucleotide-binding pocket of the B-Raf structure. Regarding the other aspect of the study, the activation of B-Raf by phosphorylating Thr599, the phosphorylated residue was added with the help of the Amber MD package.

Once these four systems (**1uBRAf_ATP**, **2pBRAf_ATP**, **3uBRAf** and **4pBRAf**) were built, they were subjected to 800 ns long range MD simulations and the resulting trajectories were carefully analysed (Table 3.4.2). Details of these MD simulations are reported in the Experimental Section.

Table 3.4.2. Summary of the trajectories under study

Trajectory	Colour code	Phosphorylation site	Binding Components	Solvent
1uBRAf_ATP	Cyan	-	ATP, 2× Mg^{2+}	Water
2pBRAf_ATP	Green	Thr599	ATP, 2× Mg^{2+}	Water
3uBRAf	Blue	-	-	Water
4pBRAf	Mauve	Thr599	-	Water

As part of a well-known and systematic procedure, backbone Root Mean Square Deviation (RMSD) values (Figure. 3.4.2) were computed to investigate both the structural changes occurring within each trajectory and the dynamic behaviour of the four different systems over the complete simulation. The four B-Raf systems studied reach stabilization at 100-150 ns, justifying the long simulations performed. From a conformational point of view, it is clear from these RMSD graphs that **1uBRAf_ATP** and **2pBRAf_ATP** are more stable than the systems without ATP (**3uBRAf** and

4pBRAf). As expected, the system lacking ATP or phosphorylation (**3uBRAf**) shows the most important conformational changes during the MD trajectory.

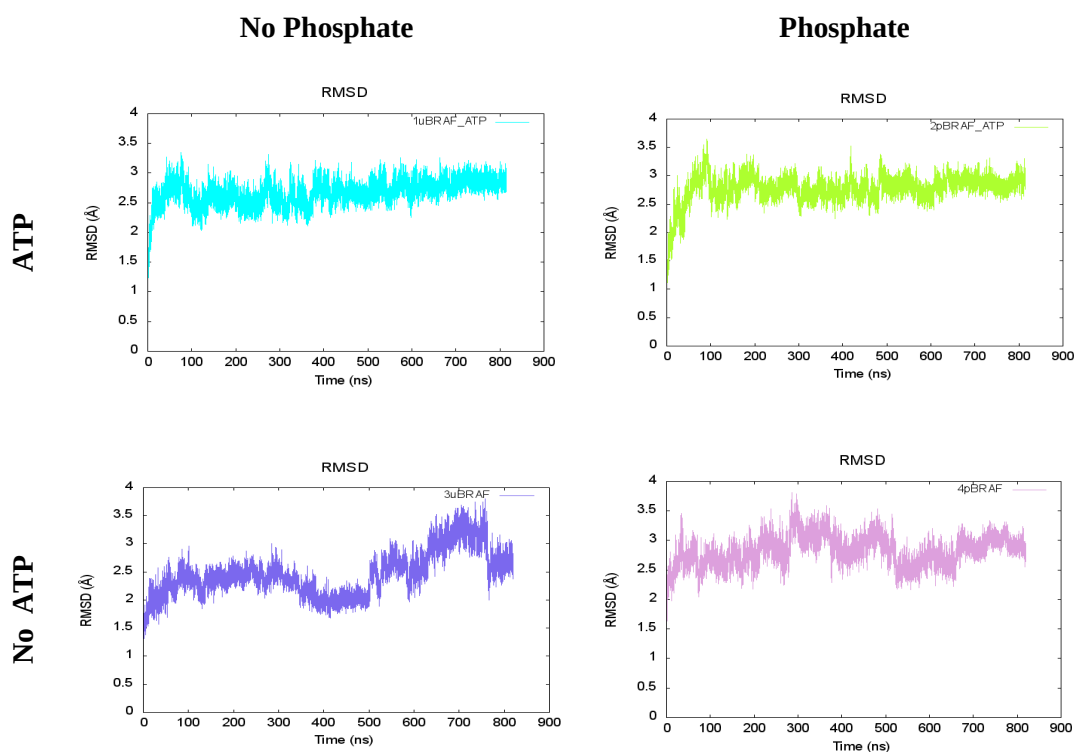


Figure 3.4.2. Backbone RMSD of the four B-Raf systems as a function of time along the MD simulation (800 ns); the RMSD is calculated between each successive structure and the first structure of the trajectory. **1uBRAf_ATP** in cyan, **2pBRAf_ATP** in green, **3uBRAf** in blue and **4pBRAf** in mauve.

Additionally, Root Mean Square Fluctuation (RMSF) functions have been computed over the last 200 ns of the four trajectories to identify the most flexible regions in each B-Raf model (Figure 3.4.3). In general, the highest RMSF peaks are found in the two phosphorylated trajectories, **2pBRAf_ATP** and **4pBRAf**. In particular, the **2pBRAf_ATP** trajectory (with a few exceptions) exhibits the highest fluctuations along all the important residues and it is the only one to show a very large mobility of residues 448-458 which are the ones involved in the dimerization of the Raf family, what is required for their activity and a sign of a fully active kinase.

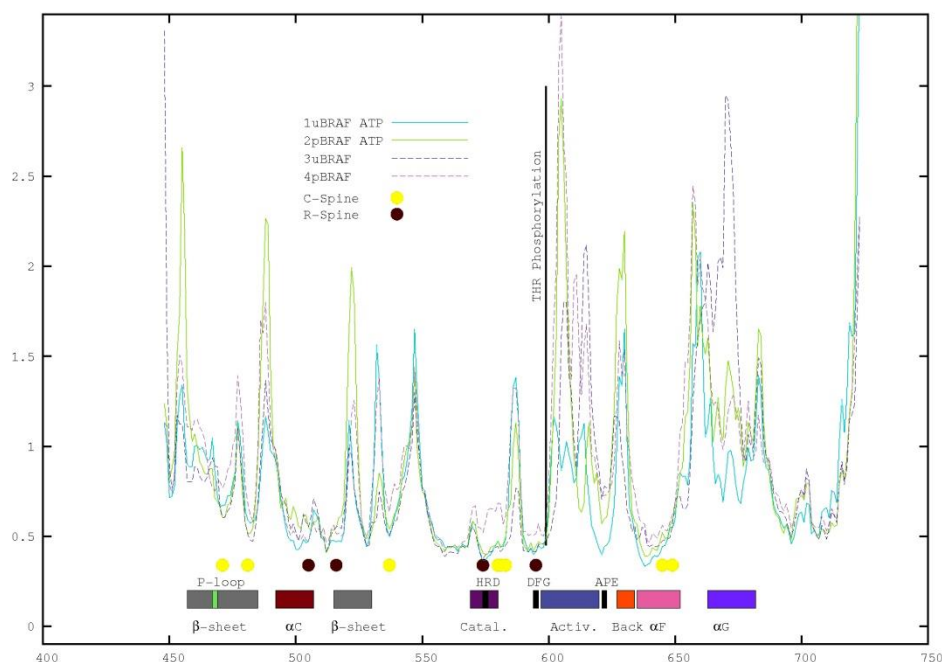


Figure 3.4.3. Backbone RMSF calculated for the four different B-Raf systems over the last 200 ns; specific domains of the protein are reported (coloured rectangles) as well as important motifs like HRD, DFG and APE; C- and R-spines are indicated by a circle.

3.4.2. Analysis of the trajectories obtained

3.4.2.1. DFG motif and Mg^{2+} binding loop in the optimised systems

The importance of the conformation of the DFG motif within the Mg^{2+} binding loop in the activation of kinases is well established and it is accepted that some known B-Raf mutations like V600E produce distortions in this DFG motif that affect its activity.²⁹ Therefore, to get a better understanding of the conformational diversity of this important motif within the four MD trajectories, we have constructed a one-dimensional dihedral free energy landscape (1D-FEL, see Experimental section for details) from a classical MD in a systematic approach. In this approach, the so-called Coarse-Grained Dihedral Angles (CGDA) acts as a local probe of the FEL along the primary sequence. This 1D-FEL analysis allows identifying possible local structure deformations in the DFG motif due to the presence of ATP or caused by phosphorylation. In Figure 3.4.4 is observed how the conformational space explored by the dynamical DFG backbone (dihedrals θ_{593} to θ_{596}) is slightly influenced by the presence (or absence) of the ATP/ Mg^{2+} system. More specifically, the lack of

ATP/Mg²⁺ (**3uBRAf** and **4pBRAf**) flattens the 1D-FEL of those dihedral torsions that mainly affect the position of Asp594 (θ_{593} and θ_{594}) because in the absence of Mg²⁺, this residue cannot longer be coordinated to the cation. Also, upon removing ATP, a slight shift by about 20° of the dihedrals θ_{594} and θ_{595} is observed. Besides, when ATP/Mg²⁺ is present (**1uBRAf_ATP** and **2pBRAf_ATP**), the influence of the phosphorylation is negligible, obtaining the same pattern for the DFG motif with and without a phosphate group.

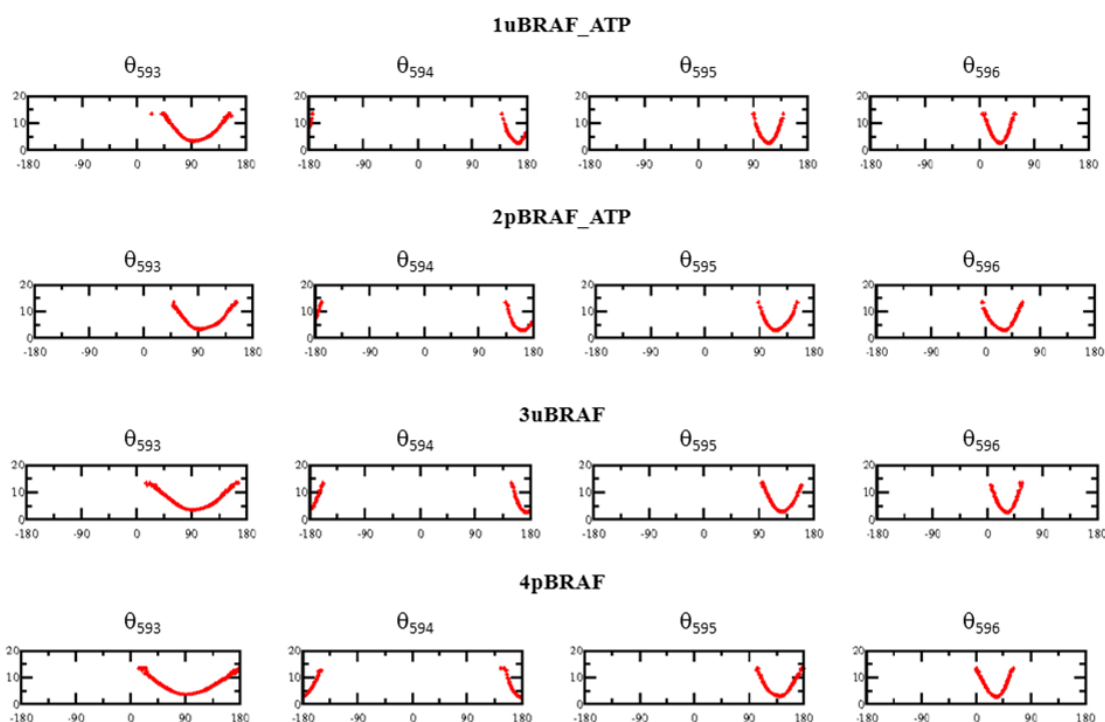


Figure 3.4.4. 1D-Free Energy Landscape (1D-FEL) of dihedral torsions affecting the position of the DFG motif for the four systems studied; units are $k_B T$ (energy) and degrees (angles).

If Asp594 and Phe595 role in the DFG motif are well known and described,²⁷ that of Gly596 is not clearly defined in the literature. Previous studies suggest that the conformation of Asp in the active kinase relies on the HB formed between one of its (C=)O and the NH group of the Gly and that disruption of this HB results in the possibility of DFG-flip.³⁰ To verify this hypothesis in the models here developed, the distances between one of the O atoms of the Asp594 carboxylate group and the N-H hydrogen atom of Gly596 were measured (Figure 3.4.5). Even though these distances

are too large for HB interactions, they still reveal a stabilising interaction for the Asp residue in the active state.²⁰ Thus, only the distance observed in the **2pBRAf_ATP system** (3.7 Å) is near to that expected for a HB³¹ while the rest of the systems exhibit much larger values (4.5-4.9 Å). This can be explained because, in **2pBRAf_ATP**, the presence of ATP and the phosphate group in the activation loop brings these two residues much closer together, resulting in a shorter contact.

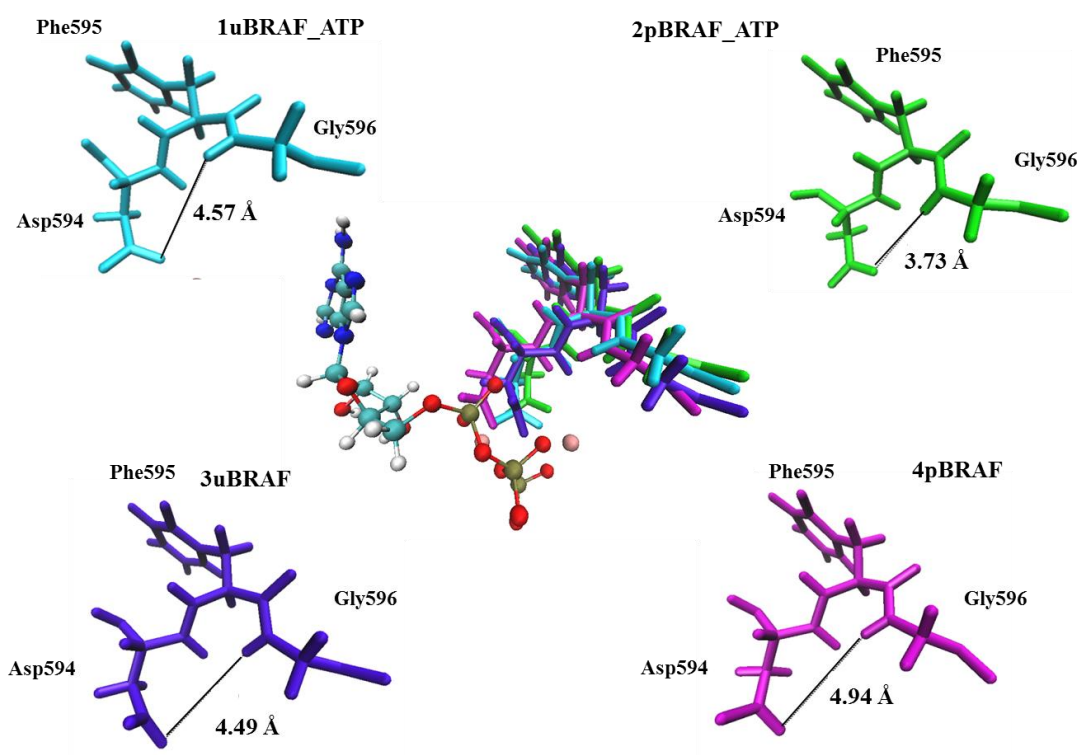


Figure 3.4.5. DFG motif orientation for the four different families under study. **1uBRAf_ATP** in cyan, **2pBRAf_ATP** in green, **3uBRAf** in blue and **4pBRAf** in mauve

3.4.2.2. Analysis of the Activation Loop and Phosphorylation Site

In agreement to previous studies,³² the central role of the phosphate group at Thp599 is evidenced in our study. Thus, by being able to interact with different important residues from several domains, Thp599 functions as a central core for the conformational changes that will allow ATP to bind.

The electrostatic interaction between the primary phosphate at the activation loop and a conserved basic pocket (RD pocket) has been previously shown to be critical for driving the conformational change in the activation loop.³² The transition of the activation segment to an active conformation is typically mediated by phosphorylation either by *cis* or *trans* auto-phosphorylation or by another activating kinase.^{27,30,33,34}

Simulations on the phosphorylated systems (**2pBRAF_ATP** and **4pBRAF**) showed that the phosphate group of Thp599 (phosphorylated Thr599) establishes two strong interactions with two positively charged residues, one with Arg575 (HRD motif in catalytic cleft) as previously suggested,^{25,32} and another with Arg603 (in the activation segment). Accordingly, the distances between Thr/Thp599 and these Arg residues have been monitored for the four systems studied and the corresponding plots are presented in Figure 3.4.6 and Figure 3.4.7.

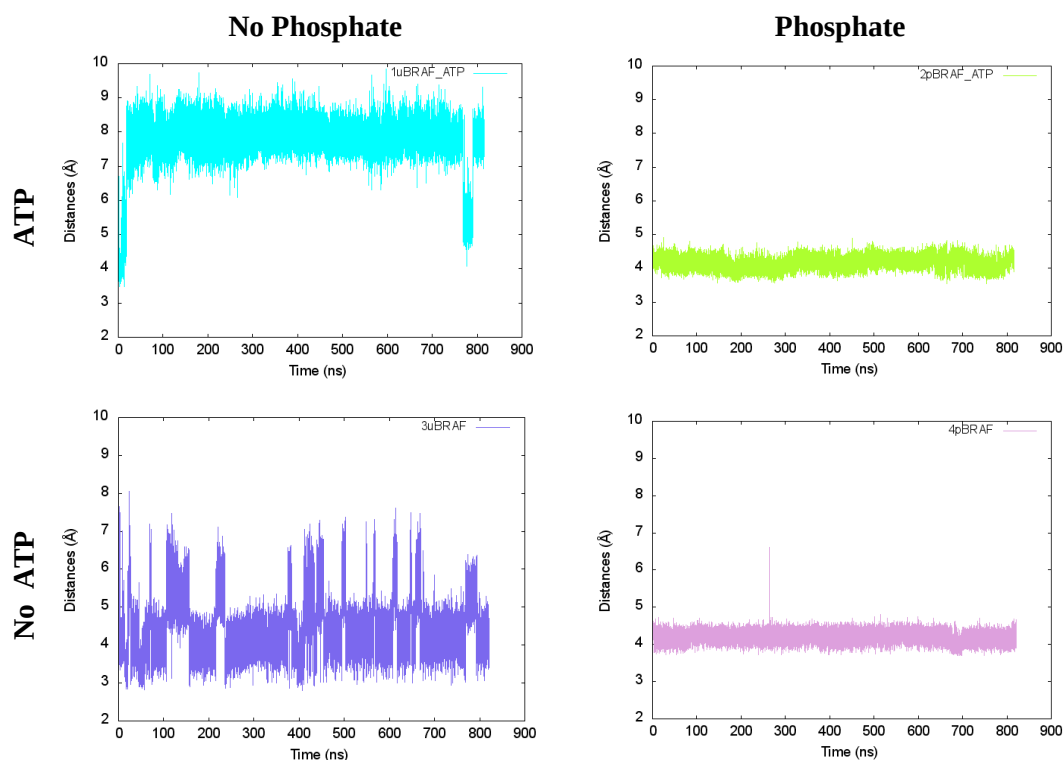


Figure 3.4.6. Distance between Arg575 and Thr/Thp599 during 800 ns simulation. **1uBRAf_ATP** in cyan, **2pBRAf_ATP** in green, **3uBRAf** in blue and **4pBRAf** in mauve

In the particular case of the phosphorylated systems **2pBRAf_ATP** and **4pBRAf**, the average distance between Thp599 and both Arg575 and Arg603 is in the range of 4.0-4.5 Å. This is in contrast with the larger Thr599-Arg575/Arg603 distances measured for the two un-phosphorylated trajectories (**1uBRAf_ATP** and **3uBRAf**). Both phosphorylated models show a stable behaviour along the trajectory (with almost no fluctuation). However, we can observe that while Arg575 approximates to the phosphorylated site in a very short time (interaction is formed during the equilibration of the system), it takes between 50 ns (**4pBRAf**) and 150 ns (**2pBRAf_ATP**) for Arg603 to electrostatically link Thp599.

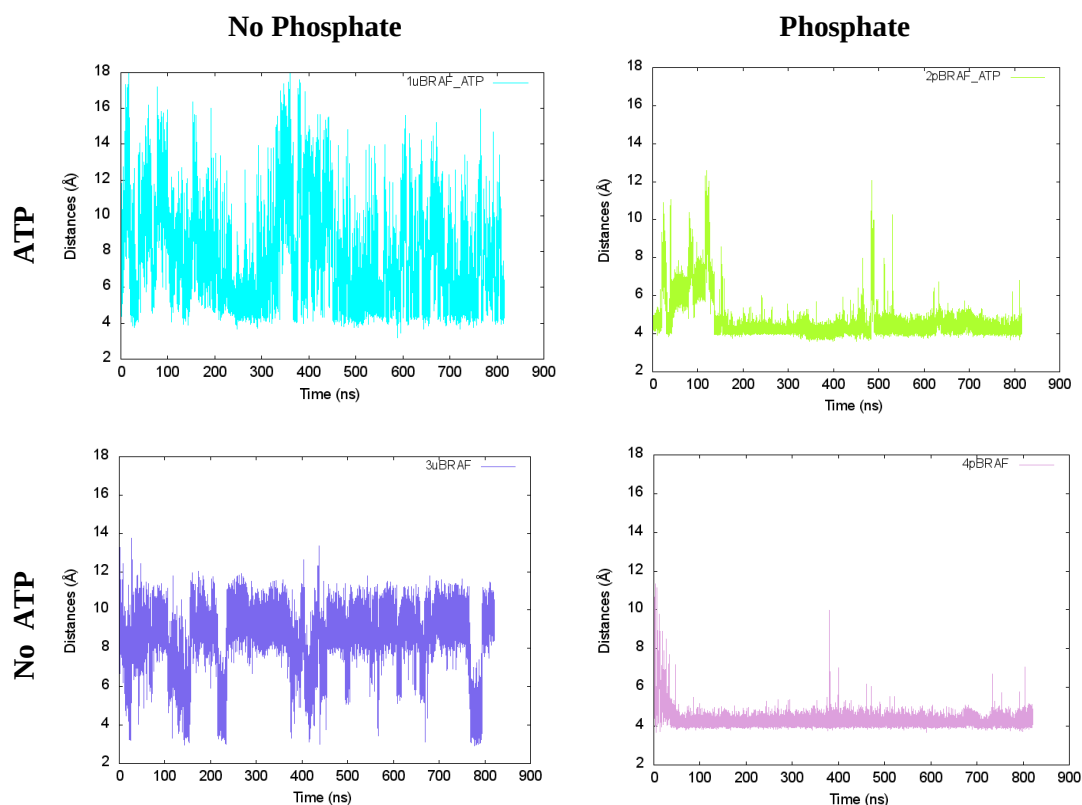


Figure 3.4.7. Distance between Arg603 and Thr/Thp599 during 800 ns simulation. **1uBRAf_ATP** in cyan, **2pBRAf_ATP** in green, **3uBRAf** in blue and **4pBRAf** in mauve

3.4.2.3. Analysis of the α C-helix region

Considering previous studies which proposed that hydrophobic packing along the entire α C-helix is a driving force for the regulation and assembly of all active kinases,²⁵ we have monitored some important motifs and residues surrounding the α C-helix to understand their arrangement in the four MD trajectories. The α C-helix and surrounding residues of **1uBRAf_ATP**, **2pBRAf_ATP**, **3uBRAf** and **4pBRAf** are shown in Figure 3.4.8.

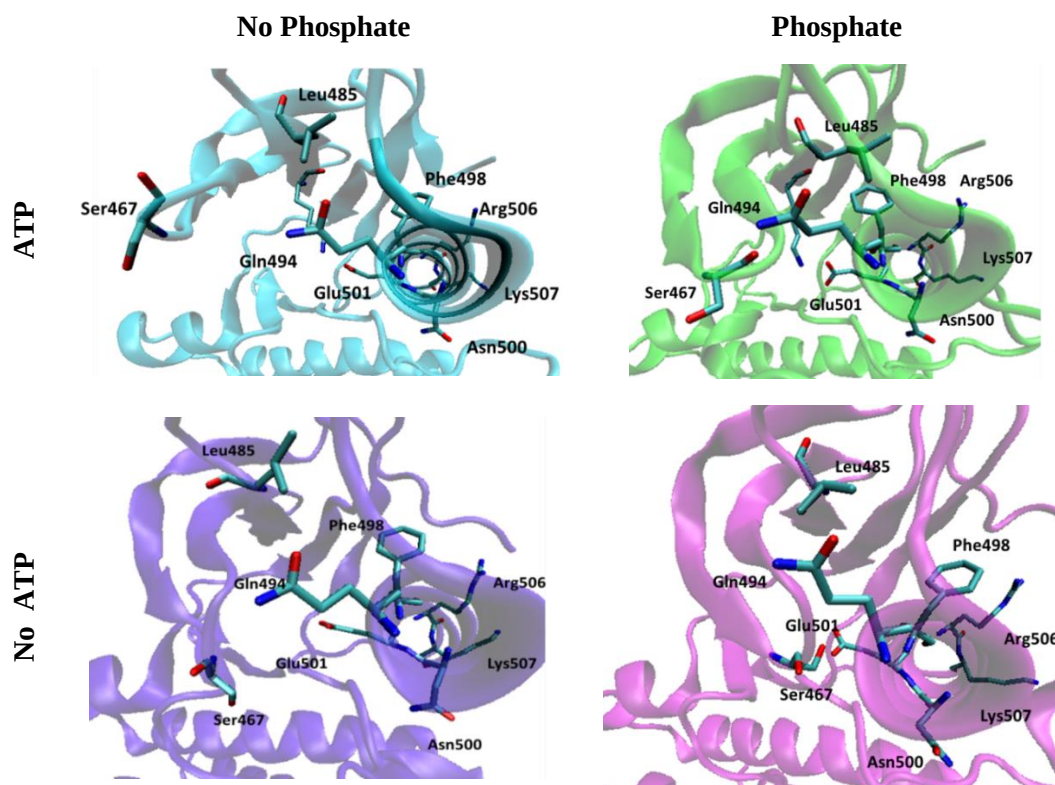


Figure 3.4.8. General view of the α C-helix and surrounding residues. **1uBRAFF_ATP** in cyan, **2pBRAFF_ATP** in green, **3uBRAFF** in blue and **4pBRAFF** in mauve

The HB interaction between Gln494 (in α C-helix) and Ser467 (in Gly-rich loop) was found to be very important in the **2pBRAFF_ATP** system (Figure 3.4.8). This HB between the Gln94 side-chain NH and Ser467 backbone (C=)O connects β -loops and α C-helix, helping to create space for ATP to bind. Regarding the **3uBRAFF** and **4pBRAFF** systems, the absence of ATP causes a collapse of the Gly-rich loop in the ATP region preventing this HB interaction. Thus, the $\text{NH}\cdots\text{O}$ distance changes from 1.83 Å in **2pBRAFF_ATP** to 5.35 and 6.45 Å in **3uBRAFF** and **4pBRAFF**, respectively (Figure 3.4.8.).

Results from previously reported molecular modelling studies suggest that an interaction between Leu485 and Phe498 would facilitate the movement of the α C-helix helping to position RS3 in the R-spine.²⁵ Accordingly, the Leu485-Phe498 distance was monitored in the four trajectories studied and the results indicate that after the first 200 ns, the average Leu485-Phe498 distance (α C-helix) is ~ 9 Å for **1uBRAFF_ATP**, **3uBRAFF** and **4pBRAFF**, and ~ 5.5 Å for **2pBRAFF_ATP**. This clearly indicates that the

simultaneous presence of both the phosphate and the ATP (as in **2pBRAf_ATP**) is required for this β_3 - α C helix attractive interaction.

Extensive studies on the structural basis for protein kinase activation²⁸ have established that the interaction between Glu501 (also in the α C-helix) and Lys483 is responsible for positioning the ATP's α - and β -phosphates for hydrolysis. Thus, the distance between the amino group of Lys483 and the carboxylate group of Glu501 was monitored throughout the simulations for both the apo and ATP-bound trajectories (Figure 3.4.9.). As expected from an active protein kinase state, this interaction remains relatively strong in all four models.

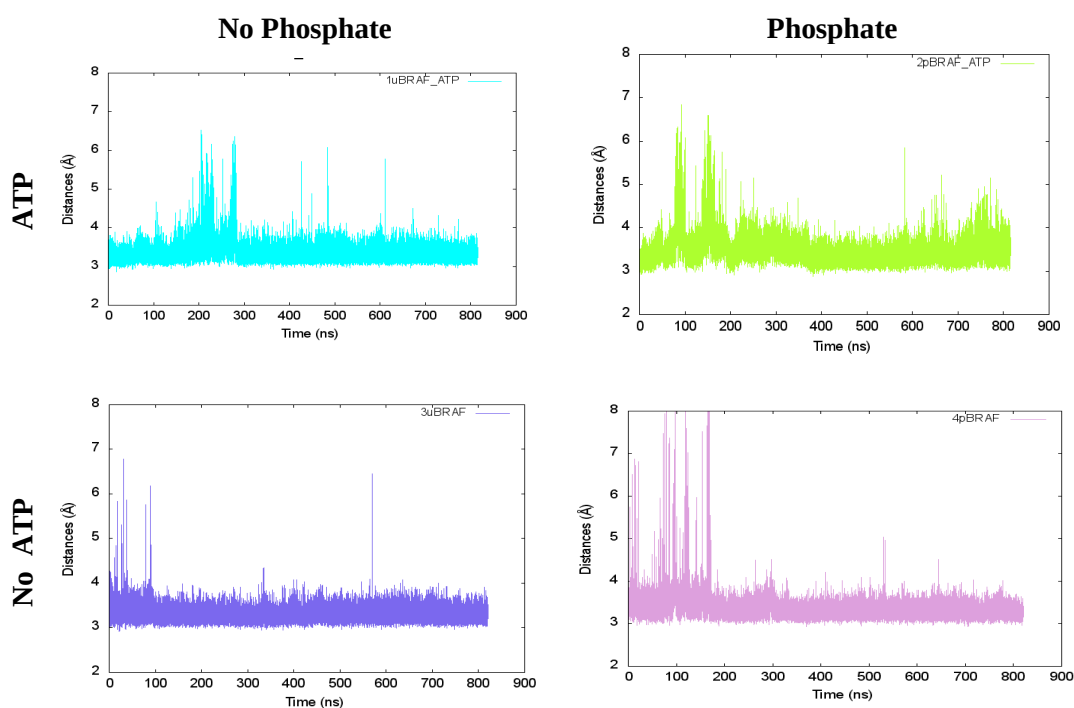


Figure 3.4.9. Distances analysis between Glu501 and Lys483 in **1uBRAf_ATP** in cyan, **2pBRAf_ATP** in green, **3uBRAf** in blue and **4pBRAf** in mauve

3.4.2.4. Structure and Activating Interactions of the 2pBRAF_ATP model

Once the aimed model of an active, phosphorylated and ATP-containing, B-Raf protein kinase (**2pBRAF_ATP**) was obtained, we proceed to understand and investigate its structural features. Hence, before going ahead with the ligand-target docking studies, the suitability of **2pBRAF_ATP** as a model for an activated form of the kinase should be analysed.

Despite the fact that the ATP region (conserved among different kinase families) is well described,²⁷ its interaction with the protein has never been characterised in detail since, as mentioned, no crystal structure of ATP-bound B-Raf is available.

A 2D-plot³⁵ and a Non-Covalent interaction Index (NCI) analysis^{36,37} of this region are presented in Figure 3.4.10. showing several important ATP contacts. The outcome of the NCI study shows the HB interaction between the SH group of Cys532 and the –NH₂ of adenine and the vdW interaction between the –NH of Cys532 and one of the –N= atoms of the adenine base. Moreover, the (C=)O atom of the backbone of Gln530 forms a HB interaction with the same –NH₂ group. Structure–function studies of protein kinases have identified a “gatekeeper” residue, which regulates binding of the nucleotide to small-molecule inhibitors.³⁸ This residue located on the conserved β 4 strand, which corresponds to Thr529 in B-Raf,³⁹ establishes a strong and well oriented interaction between its O(H) atom and the –NH₂ group of the adenine ring in our model.

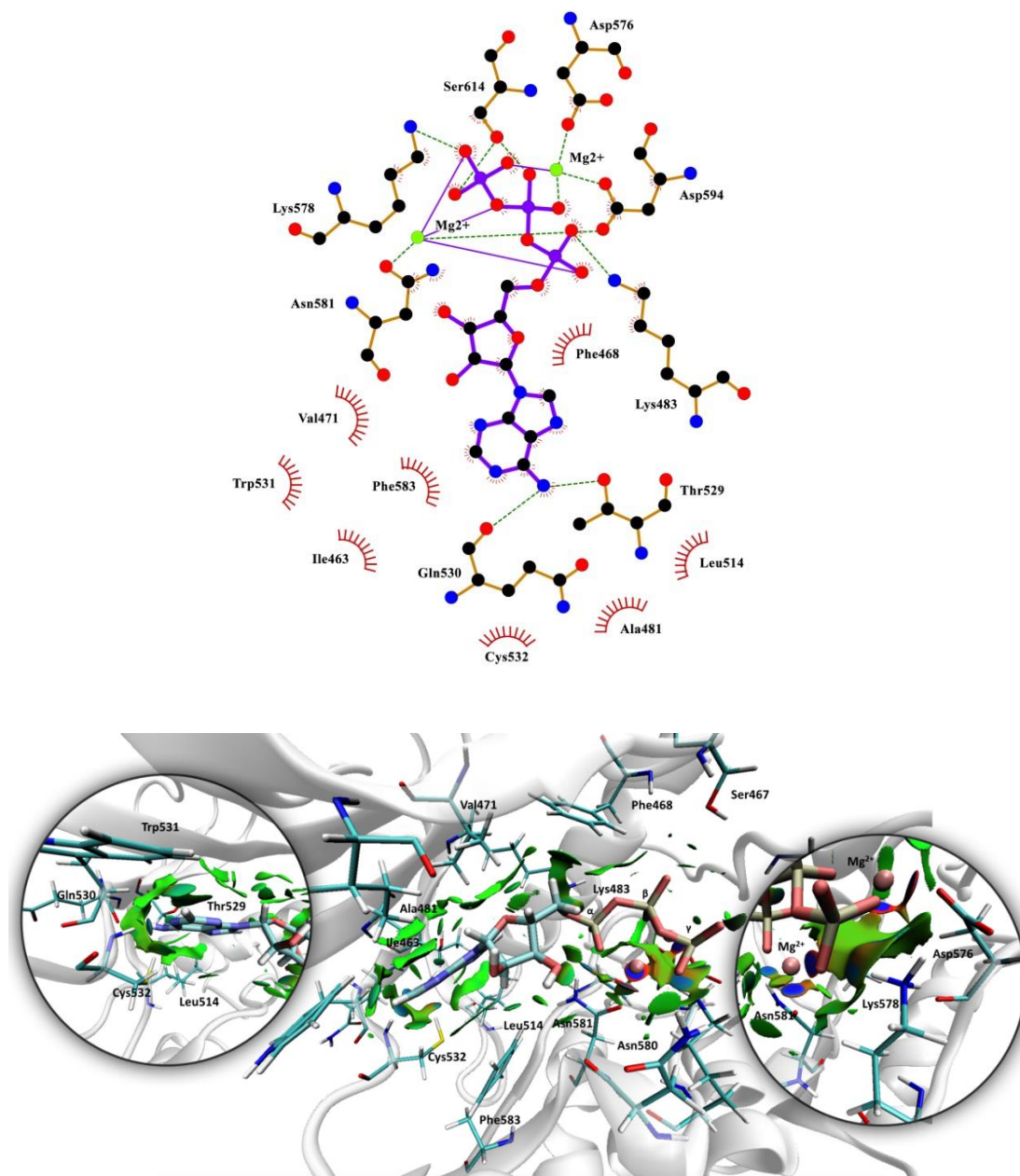


Figure 3.4.10. *Top:* Diagram of the interactions between ATP and the binding site of 2pBRAAF_ATP (created with LigPlot) which are indicated as follows: HBs by dashed lines; hydrophobic contacts by an arc with spikes radiating towards the ligand atoms they contact while the contacted atoms are shown with spikes radiating back. *Bottom:* Non-Covalent interactions Indices (NCI) plot of the non-covalent interactions formed in the protein-ATP complex. Blue areas correspond to strong NCI; green to weak attractive interactions or NCI and red to non-attractive.

Figure 3.4.10. shows how, in this model, the adenine heteroaromatic ring is very well positioned establishing non-covalent interactions mainly with the hydrophobic side

chain of Ala481 at the top side and with the aromatic ring of Phe583 at the bottom side, both residues form the C-spine, which is completed upon ATP binding.⁴⁰ The adenine of the nucleotide is also found to establish important vdW interactions with Leu514, Ile463 and the indole ring of Trp531. Moreover, the ATP's hydrophilic ribose ring establishes non-covalent interactions with Val471 at the C-spine and Phe468 in the Gly-rich loop. Also, one of the O atom at the α -phosphate, forms a HB with Lys483 (in the β 3).

In our model, the ATP's triphosphate chain establishes interactions with different residues that also help to neutralize its negative charged chain. In particular, Asn581, while coordinating with a Mg^{2+} ion, interacts through its NH_2 side chain group with an O atom of the α -phosphate. Interestingly, Asn581 also forms a HB with one of the two –OH groups in the ribose, although not very strong as seen from NCI analysis. On the back side of the ATP hinge region, one very important feature is that Asp594 (DFG motif) coordinates with the two Mg^{2+} ions, indirectly fixing the position of the β - and γ -phosphates. This coordination is supplemented by another one involving Asp576 (HRD motif). Together with Lys578 these residues are highly conserved in Ser/Thr protein kinases and can also be involved in the electrostatic stabilization of the phosphate transfer catalytic transition state because of the charged groups at their chain.⁴¹

The conformation of the activation segment (including DFG at the N-terminus and APE at the C-terminus) is of fundamental importance for the control of kinase activity and different studies have been performed to understand how it controls a unique set of signals to turn the kinase activity on or off.^{32,42} Thus, we have studied these features in the **2pBRA_F_ATP** model by means of a NCI analysis and the corresponding plot (Figure 3.4.11.) shows the cascade of non-covalent interactions connecting Thp599 to the DFG motif and α C-helix. For example, the HBs established between Thp599 and Arg575 (HRD motif) as well as with Arg603 indicates that the additional interaction with the phosphate orientates Arg575 in such a way that even His574 (HRD motif) can establish interactions with the DFG motif (through Phe595), thus assembling the R-spine, a key feature for an active kinase.²⁵

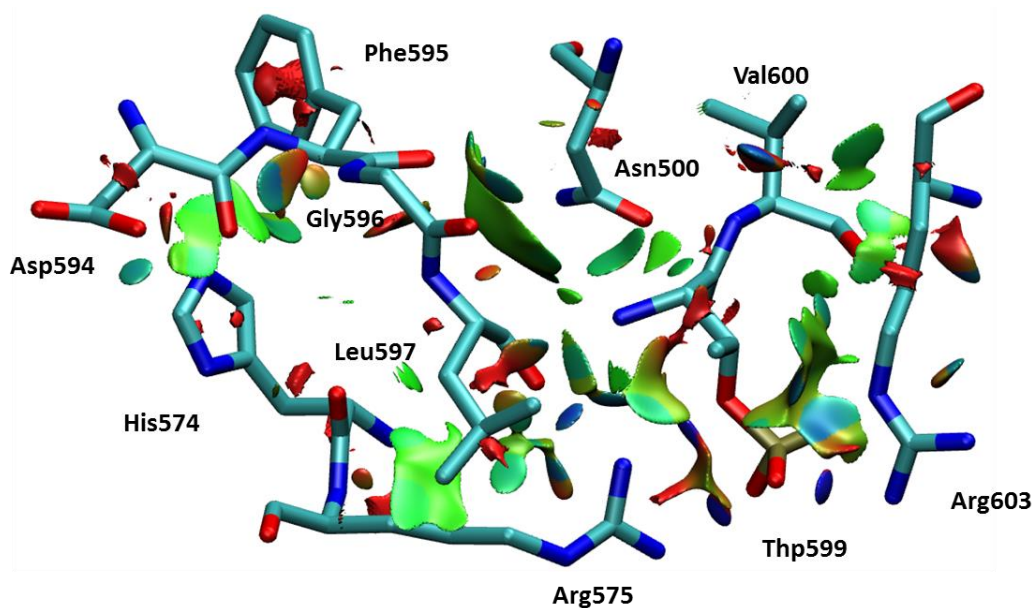


Figure 3.4.11. NCI plots of the non-covalent interactions established around phosphorylated Thr599 (Thp599) in the **2pBRAf_ATP** complex. Blue areas correspond to strong NCI; green to weak attractive interactions or NCI and red to non-attractive.

Also shown in Figure 3.4.10 is the key non-covalent interaction established by the hydrophobic side-chain of Leu597 (which is a point of interaction between the activation loop and the catalytic one) and the Arg575 side-chain. Furthermore, the point of contact between the activation segment and the α C-helix, Asn500, is shown to weakly interact through its side chain (C=)O group with the $-CH_3$ of Thp599, and by means of a HB through its $-NH_2$ group and the $-NH$ backbone of Thp599. Besides, Asn500 is also interacting with Gly596 in the DFG motif by means of its $-NH_2$ group and this Gly596 is establishing a HB with the (C=)O of Asp594 (DFG motif), affecting in this way the orientation of the Asp side-chain for ATP binding. Despite the variety of residues playing a role in the activation process, the optimised **2pBRAf_ATP** model shows similarities with previously reported studies suggesting that the geometrical arrangement of these HBs is a key feature that differentiate active kinases.³⁰

Continuing with the analysis of the **2pBRAf_ATP** model, we also explored the peptide binding region. Thus, we found a salt bridge between Glu208 and Arg280 (conserved feature of all eukaryote protein kinases⁴³), which is very relevant because Glu208 is part of the conserved APE motif¹⁶ that participates in the substrate binding area.⁴⁴

Moreover, in this region we found that Trp619 (two amino acids away from the APE motif) seems to be relevant for the pre-organization of peptide binding because in the **2pBRAf_ATP** model is positioned in a very different orientation than in **1uBRAf_ATP**, **3uBRAf** or **4pBRAf** systems (Figure 3.4.12., *above*) and this is supported by monitoring the distances of this residue and Lys578 along the four trajectories (Figure 3.4.12., *below*).

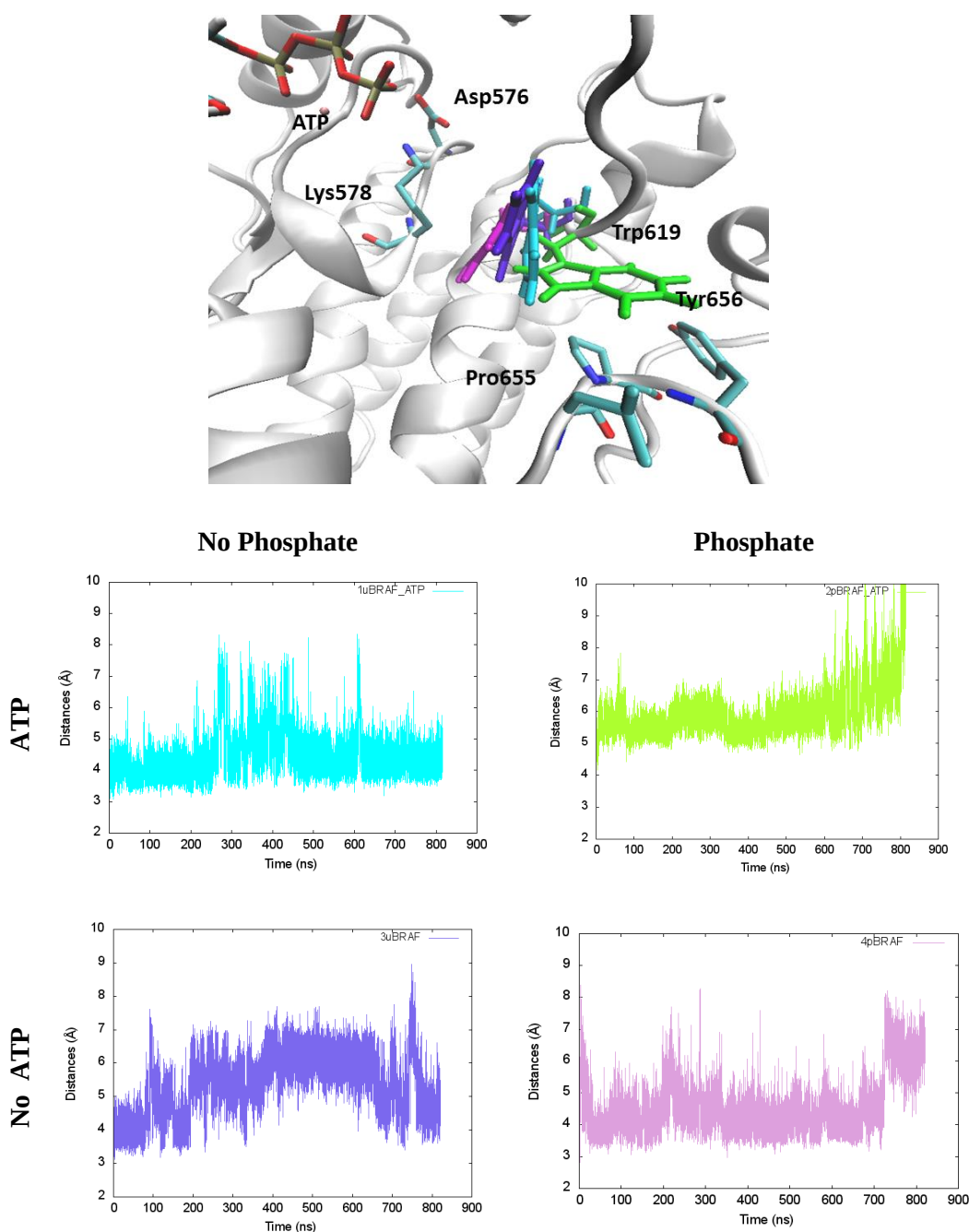


Figure 3.4.12. Orientation (*above*) and distance (*below*) between Trp619 and Lys578 within **1uBRAf_ATP** in **cyan**, **2pBRAf_ATP** in **green**, **3uBRAf** in **blue** and **4pBRAf** in **mauve**

It is known that the B-Raf peptide substrate binds in the C-lobe in such a way that it can reach not only Asp576 in the HRD region, but also the adjacent Lys578 responsible for the catalysis.⁴¹ Thus, we have found that at the very end of the MD simulations, Trp619 adopts this particular and unique orientation only in **2pBRAf_ATP** positioning its side chain 10 Å away from Lys578. This particular positioning of Trp619 also allows it on the one hand to form hydrophobic interactions with Pro655 and Tyr656 (Figure 3.4.12. *above*) and on the other hand to form an empty pocket adjacent to the ATP terminal phosphate (here referred to as BAP -Binding Allosteric Pocket-) enough for the peptide substrate or eventual allosteric inhibitors to bind.

Summarising, the model here developed for the active and ATP containing B-Raf protein kinase (**2pBRaf_ATP**) presents both all the distinguishing features of active kinases and other exclusive ones and hence, it can be used as a target for structure-based drug design.

3.5. Molecular Docking Study

With the aim of rationalizing our previously obtained results and, more importantly, to get insight into possible allosteric inhibition of B-Raf, the **2pBRAf_ATP** kinase model was subject to molecular docking experiments with the lead compound **1** (Figure 3.3.1). Our thorough full atomistic simulations clearly unveil the existence of an empty pocket (binding allosteric pocket, BAP) adjacent to the ATP terminal phosphate. The geometry of the protein was selected from the last 20 ns of the **2pBRAf_ATP** trajectory with the BAP pocket sufficiently open to host a ligand.

The program AlgoGen⁴⁵ was used for the molecular docking studies of the inhibitor using as a target the previously equilibrated structure of **2pBRAf_ATP**. Best poses were obtained using the Vina score (implemented in AlgoGen) for a fast exploration of the ligand-host interactions. Then, AlgoGen automatically refined the best poses at the Quantum Chemistry level of theory (using the semiempirical PM7 method⁴⁶ and the MOZYME function⁴⁷ available in MOPAC2016⁴⁸). Details of the docking experiments are given in the Experimental section. First, a QM rigid relaxation is performed; then, on the best pose for B-Raf binding, we refined the geometry by relaxing all valence and

dihedral angles of the ligand (flexible docking) as well as the valence and dihedral angles of the protein (protein flexibility).

One of the poses obtained by our molecular docking simulation (Figure 3.5.1.) shows that compound **1** can easily occupy the B-Raf peptide binding pocket in the optimised protein conformation. Two strong interactions can be highlighted in this particular pose: bifurcated HBs between the mono-substituted guanidinium cation and two negatively charged ATP phosphate groups (showing 1.6, 2.9 and 2.7 Å distances) and another parallel HB between the di-substituted guanidinium group and Glu648 (at 2.0 and 2.6 Å). Both sets of HBs are reinforced by ionic interactions due to the charged nature of guanidinium, phosphate and carboxylate groups.⁴⁹ In addition to stabilize the complex, these interactions favourably orientate the molecule inside the peptide binding pocket, BAP. Furthermore, a number of vdW contacts were identified involving, among others, with previously mentioned residues Trp619, Pro655.

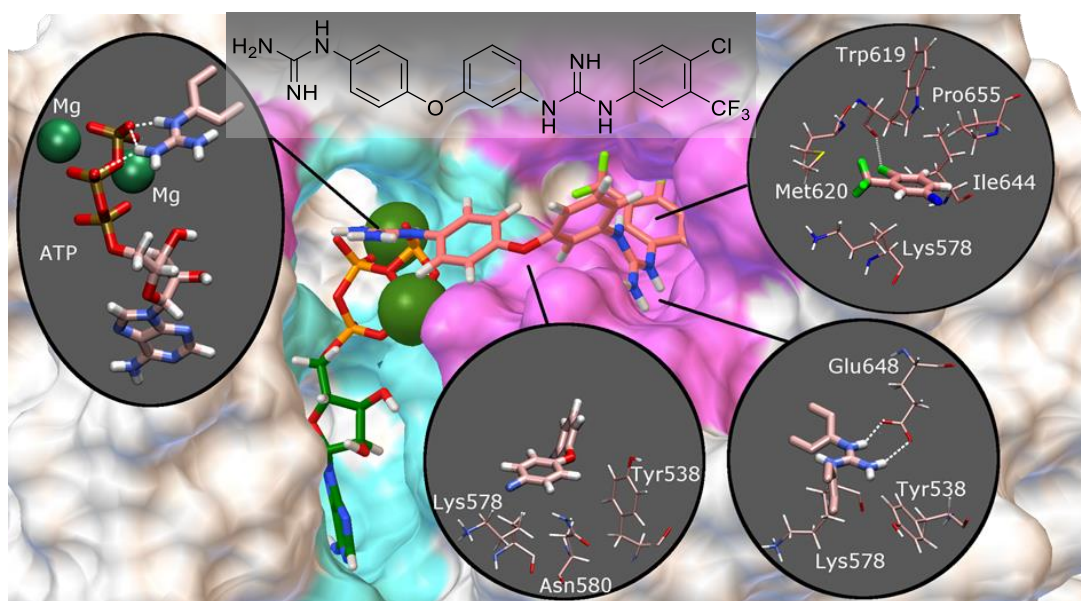


Figure 3.5.1. Best pose from the molecular docking study of compound **1** (structure is reported in the picture) in the system **2pBRAf_ATP** indicating the most relevant interactions established.

Interestingly, a NCI analysis with a quantitative aspect (Independent Gradient Model, IGM⁵⁰) indicates that the chlorine atom of compound **1** interacts with the (C=)O of the peptide bond of Trp619 by means of a halogen bond⁵¹ and the trifluoromethyl group points towards the methionine Met620 likely establishing hydrophobic interactions

(Figure 3.5.2.). Our model indicates that compound **1** orientates in such a way that prevents any other partner to bind B-Raf near to the ATP and to access the catalytic loop, thus probably inhibiting the phosphate transfer of the γ -phosphate involved in the phosphorylation mechanism of the kinase.

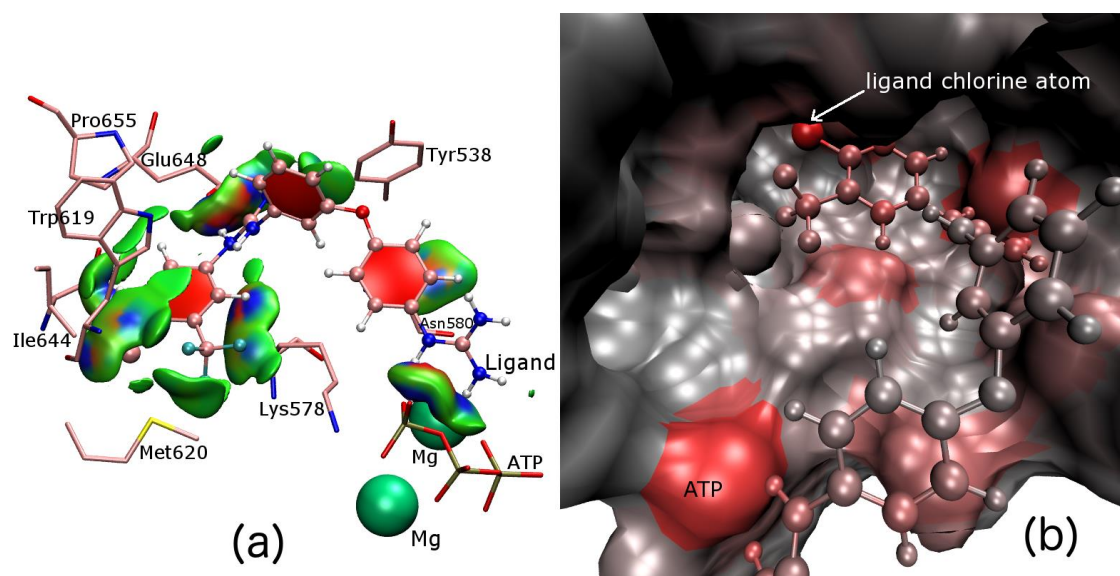


Figure 3.5.2. Analysis of non-covalent interactions between **2pBRAf_ATP** and compound **1** using the IGM method derived from NCI, (a) $\delta g^{\text{inter}} = 0.004$ a.u. isosurfaces coloured according to the BGR scheme over the electron density range $-0.05 < \text{sign}(\lambda_2)\rho < 0.05$ a.u. (b) Atoms of the complex coloured according to their contribution to the isosurfaces of panel (a) (summing over the δg^{inter} grid and using a grey-to-red colour scheme, with grey: no contribution to the NCI, red: largest relative contribution to the NCI).

3.6. Theoretical model of the interaction between phosphates in the ATP molecule and cationic systems

Continuing with our interest in modelling the potential allosteric interaction of compound **1** with B-Raf and the best pose obtained from the docking study as well as taking into account the specific guanidinium cation properties,⁵²⁻⁵⁷ we computationally studied the nature of the forces established between an arylguanidinium and the molecule of ATP.

Because of the large size of the complexes under study, a simplified model of compound **1** (Figure 3.3.1) was chosen to focus our study, the 4-phenyloxyphenyl-*N*-guanidinium (Figure 3.6.1). Moreover, to better understand the nature of the specific guanidinium-phosphate interactions, we first used the phosphate anion as a simplified model of ATP (results reported in Appendix A.1.1. Arylguanidinium – Phosphate complexes) and, next, we studied the complexes formed with the complete ATP molecule.

Despite its high basicity, the guanidinium cation has important features very useful for target binding such as forming strong HBs and ionic interaction.^{49,52,58} Interestingly, the related isouronium cation keeps these two binding features and, moreover, is less basic ($pK_{aH} = 9.8^{59}$); hence, this cation is very promising as a functional substitute of guanidinium derivatives with therapeutic potential.⁶⁰ Given the possibility for isouronium to overcome the bioavailability problems associated to guanidinium, we decided to explore both theoretically and experimentally (see Chapter 4 – Chemical Synthesis) the effect of substituting the guanidinium cation by an isouronium (arylisouronium, Figure 3.6.1) and analyse the interactions established with the ATP's phosphate chain of protein kinases.

The geometry of the monomers and the complexes has been fully optimized at DFT ω B97XD/6-31+G(d,p) level and then using MP2/6-311+G(d,p) to obtain more accurate binding energies. The bonding characteristics were analysed by means of the atoms in molecules (AIM) theory (for specifics see Computational details in Experimental section).

On the view of the results obtained for the interaction of the phosphate anion with arylguanidinium derivatives in complexes **Ia/b** – **II** – **III** – **IVa/b** (see Appendix A.1.1.

Arylguanidinium – Phosphate complexes), for which parallel HBs between two H atoms of the guanidinium moiety and two O atoms of the phosphate anion have been found to be the most stable interactions, only those parallel interactions that involve two H atoms of the guanidinium or isouronium moieties (H^2 and H^3 Figure 3.6.1) and two O atoms of the phosphate anion were considered when studying the ATP complexes.

The ATP molecule has three different phosphate groups that can interact with guanidinium or isouronium, so these three groups have been explored and identified as: [α] phosphate attached to the sugar and adenine system; [β] middle phosphate group; and [γ] mono-linked end phosphate (Figure 3.6.1). Regarding the interactions, for the [β] and [γ] phosphates, different approximations are possible: (i) for β and γ phosphates the interaction with the O atoms could be on the *right* or *left* side of the molecule ([β -R], [β -L], [γ -R] and [γ -L] complexes) and for γ phosphate, the interaction can occur also from the *bottom* side ([γ -B] complex). Besides, the position of the OPh group of the arylguanidinium or arylisouronium could present different orientations with respect to the ATP molecule upon complexation (Figure 3.6.1); since this orientation is different we have named **a** the more stable complexes and **b** the less stable ones.

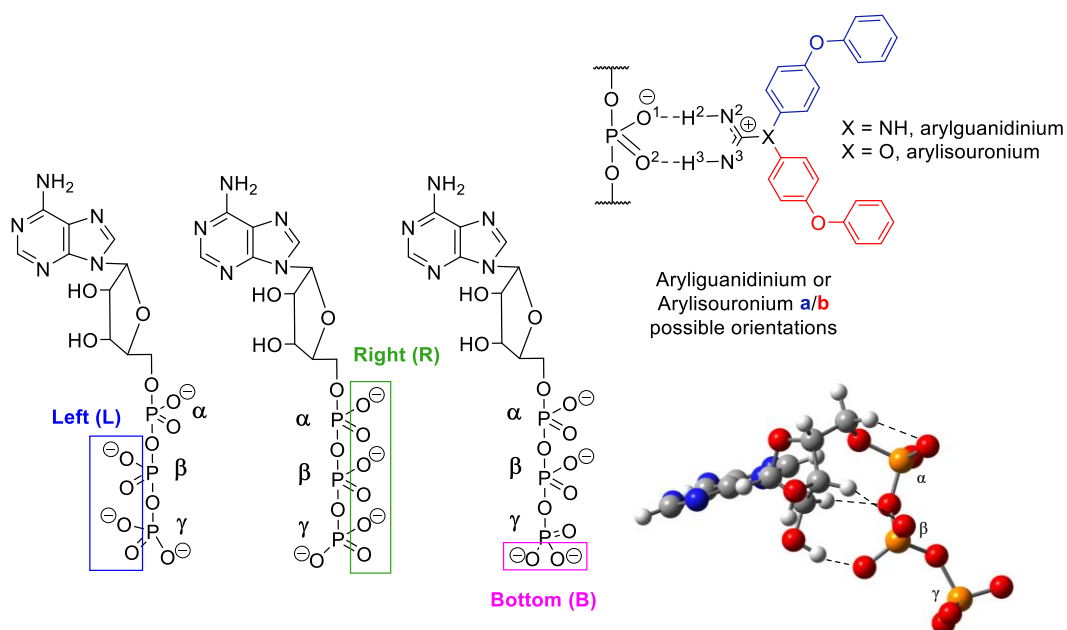


Figure 3.6.1. Different positions considered for the interactions between the 4-phenyloxyphenyl-*O*-guanidinium or –isouronium cation and the ATP molecule; ATP molecule optimised at ω B97XD/6-31+G(d,p) and PCM-water computational level.

3.6.1. Arylguanidinium – ATP interactions

3.6.1.1. Optimized geometries and interaction energies

In order to probe the phosphate groups in the ATP molecule with the arylguanidinium system, a total of 12 possible complexes have been considered in this study, named complexes **V**. Moreover, four extra configurations have been taken in consideration, in which a simultaneous interaction between the arylguanidinium and the three phosphates has been established, those dimers are named complexes **VI**. Interaction energies were calculated and the values reported in Table 3.6.1.

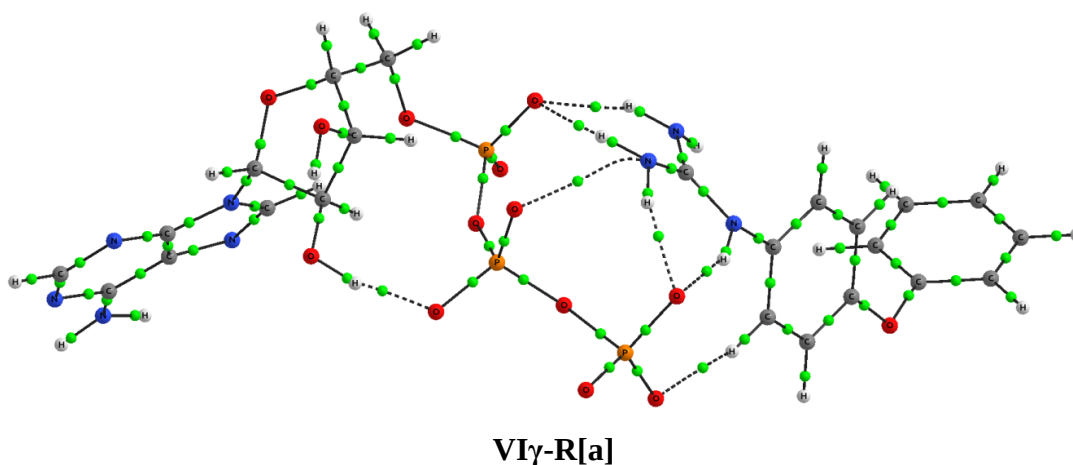
Table 3.6.1. Interaction energies (E_i , kJ mol⁻¹) for the different complexes arylguanidinium - ATP calculated at MP2/6-311+G(d,p) and PCM-water computational level.

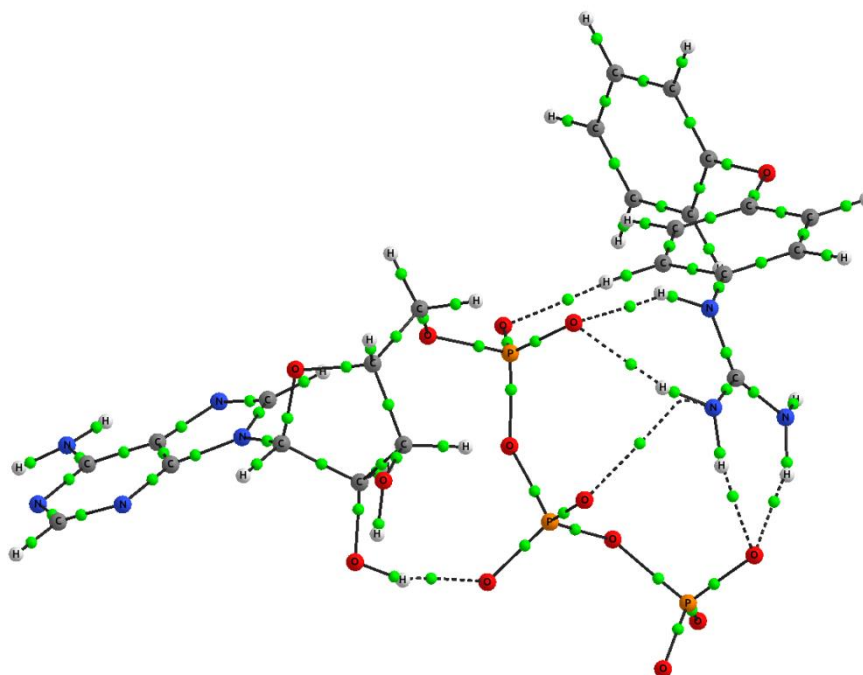
		V	VI
α -R	a	-105.9	-
	b	-105.8	-
β -R	a	-129.0	-
	b	-128.3	-
β -L	a	-95.4	-
	b	-94.5	-
γ -R	a	-120.6	-164.4
	b	-119.3	-160.9
γ -L	a	-124.0	-177.7
	b	-122.5	-169.3
γ -B	a	-163.2	-
	b	-116.2	-

The calculated E_i values range between -94.5 and -177.7 kJ mol⁻¹. In all cases the values obtained are larger than those calculated for the arylguanidinium and the single PO₄³⁻ molecule; therefore, the guanidinium-phosphate interaction is stronger when the phosphate is part of the ATP molecule. For the different complexes **V** it has been observed that when the interaction occurs by the *right* side of the ATP molecule (**R** complexes, Figure 3.6.1), **β** and **γ** complexes show stronger interaction energies than **α** complexes. Moreover, for dimers **V** the largest E_i value corresponds to the **γ-B[a]** complex, in which the interaction occurs on the *bottom* side of the end phosphate group of the ATP molecule. As it is shown in Table 3.6.1, complexes **VI**, in which multiple and simultaneous interactions with different phosphates occur, show larger E_i values than those calculated for complexes **V**.

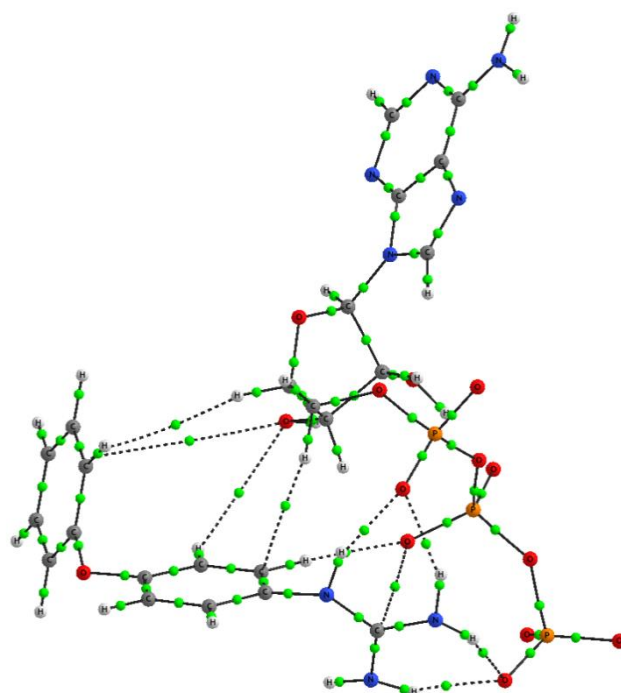
3.6.1.2. Electron density analysis

By means of the AIM (Atoms in Molecule) theory, the intermolecular interactions of all the different complexes (16 in total) were identified by the corresponding Bond Critical Points (BCPs) for each interaction. Thus, obtaining the electron density at each BCP (ρ_{BCP}) and its Laplacian ($\nabla^2\rho_{\text{BCP}}$) the nature of these interactions have been characterised and this was represented in the corresponding molecular graphs. For the sake of clarity, only the molecular graphs corresponding to complexes **VI** are shown in Figure 3.6.2 and the rest are gathered in the Appendix (Figure A.1.2.1.).

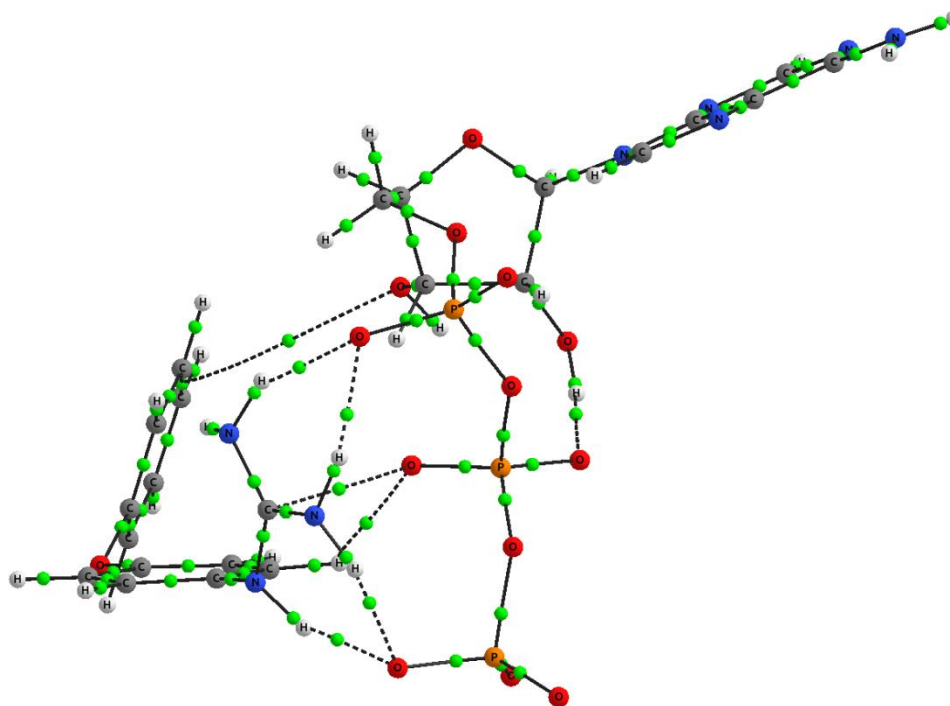




VIγ-R[b]



VIγ-L[a]



VIγ-L[b]

Figure 3.6.2. Molecular graphs corresponding to the complexes type **VI**. Green dots correspond to bond critical points (BCP).

In Tables 3.6.2 and 3.6.3, the ρ_{BCP} and $\nabla^2\rho_{\text{BCP}}$ values calculated for all complexes type **V** and **VI** are summarized. Values of the electron density ρ_{BCP} between 0.0107 and 0.0591 a.u. and positive Laplacians indicate that those interactions belong to the non-covalent regime and can be characterised as HBs.

Focusing in the values obtained for complexes **V** (Table 3.6.2), the evolution of the ρ_{BCP} follows the order $\gamma > \beta > \alpha$ in perfect agreement with the E_i values obtained. It is important to highlight that the interaction in complexes **V** occurs through parallel HBs (two O atoms of the phosphates of the ATP molecule and two H atoms of two different N atoms of the arylguanidinium).

Table 3.6.2. Electron density at the bond critical points (ρ_{BCP}) and the Laplacian of that electron density ($\nabla^2\rho_{\text{BCP}}$) calculated at $\omega\text{B97X/6-31+G(d,p)}$ and PCM-water for complexes **V**.

Complexes	$\text{O}^1\cdots\text{H}^2\text{N}^2$		$\text{O}^2\cdots\text{H}^3\text{N}^3$	
	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$
Vα-R[a]	0.0392	0.1145	0.0390	0.1152
Vα-R[b]	0.0390	0.1138	0.0390	0.1145
Vβ-R[a]	0.0433	0.1278	0.0469	0.1399
Vβ-R[b]	0.0460	0.1374	0.0441	0.1303
Vβ-L[a]	0.0503	0.1458	0.0468	0.1360
Vβ-L[b]	0.0502	0.1449	0.0472	0.1370
Vγ-R[a]	0.0531	0.1459	0.0567	0.1522
Vγ-R[b]	0.0556	0.1510	0.0543	0.1491
Vγ-L[a]	0.0538	0.1496	0.0557	0.1530
Vγ-L[b]	0.0543	0.1502	0.0538	0.1497
Vγ-B[a]	0.0508	0.1409	0.0591	0.1586
Vγ-B[b]	0.0558	0.1520	0.0543	0.1484

Regarding complexes **VI** (Table 3.6.3), the ρ_{BCP} values obtained for the HBs in those cases is lower than those calculated for complexes **V**. This could be explained because in all complexes **VI** the interaction between the phosphates in the ATP molecule and arylguanidinium occurs by two H atoms and only one oxygen atom i.e. three centre interaction. This is in total agreement with the finding that in the case of multiple HB systems, where parallel and bifurcated complexes can be compared, parallel HB interactions are stronger and preferred over the bifurcated ones.³¹ However, complexes **VI** show larger E_i values than complexes **V** because of the larger number of interactions established.

Table 3.6.3. Electron density at the bond critical points ($\rho_{\text{BCP}} \times 10^{-4}$) and its Laplacian ($\nabla^2 \rho_{\text{BCP}} \times 10^{-4}$) calculated at $\omega\text{B97X}/6\text{-}31\text{+G(d,p)}$ and PCM-water computational level for complexes **VI**.

	PαO$\cdots$HN^a		PαO$\cdots$HN^a		PβO$\cdots$HN^a		PγO$\cdots$HN^a		PγO$\cdots$HN^a		PO$\cdots$HC	
Comp	ρ	$\nabla^2 \rho$	ρ	$\nabla^2 \rho$	ρ	$\nabla^2 \rho$	ρ	$\nabla^2 \rho$	ρ	$\nabla^2 \rho$	ρ	$\nabla^2 \rho$
	BCP	BCP	BCP	BCP	BCP	BCP	BCP	BCP	BCP	BCP	BCP	BCP
VIγ-R[a]	258	729	283	849	107	414	302	857	461	1286	166	451
VIγ-R[b]	260	773	292	830	107	407	339	969	383	1088	154	441
VIγ-L[a]	270	787	315	980	108	409	298	855	387	1123	173	579
VIγ-L[b]	268	831	286	840	127	476	342	975	388	1126	124	415

^a Considering that in each complex the corresponding (P)O interacts with a different (H)N atom, the numbering followed so far cannot be used; to identify the groups involved in each interaction for each complex see Figure 3.6.2.

3.6.2. Arylisouronium – ATP interactions

3.6.2.1. Optimized geometries and interaction energies

Similar to the previous section on arylguanidinium-ATP dimers, a total of 16 possible complexes have been explored in this study, named complexes **VII** and **VIII**. Calculated interaction energies (E_i , kJ mol^{-1}) between ATP and arylisouronium cation are presented in Table 3.6.4, interaction energies with the arylguanidinium cation (in brackets) have been added for comparison purposes. As it happened for the arylguanidinium-ATP complexes, the E_i values follow the order $\gamma > \beta > \alpha$ and the largest ones are those corresponding to the interactions with the γ -phosphate (involved in the catalytic mechanism of protein kinases) as a result of more negative charges in this group.³⁰

Table 3.6.4. Interaction energies (E_i , kJ mol⁻¹) for the different complexes formed between arylisouronium (**VII** and **VIII**) or arylguanidinium (in brackets, **V** and **VI**) with ATP calculated at MP2/6-311+G(d,p) and PCM-water computational level.

		VII	(V)	VIII	(VI)
α -R	[a]	-117.2	(-105.9)	-	-
	[b]	-116.8	(-105.8)	-	-
β -R	[a]	-141.5	(-129.0)	-	-
	[b]	-140.7	(-128.3)	-	-
β -L	[a]	-101.7	(-95.4)	-	-
	[b]	-101.8	(-94.5)	-	-
γ -R	[a]	-137.6	(-120.6)	-148.6	(-164.4)
	[b]	-135.3	(-119.3)	-153.4	(-160.9)
γ -L	[a]	-139.3	(-124.0)	-175.8	(-177.7)
	[b]	-139.5	(-122.5)	-162.3	(-169.3)
γ -B	[a]	-181.7	(-163.2)	-	-
	[b]	-131.1	(-116.2)	-	-

The weaker interactions found for the β -L complexes **VII** ([a]: -101.7 and [b]: -101.8 kJ mol⁻¹) are an exception for these systems and this is probably a consequence of the disruption of the IMHBs' net existing in the ATP molecule (Figure 3.6.1) by the presence of the arylisouronium, resulting in a less favourable pose. When the interaction occurs by the *right* side (R) of the ATP molecule, allowing its natural IMHB net (Figure 3.6.1), both phosphates β and γ produce stronger E_i than with the α -phosphate. For complexes **VII**, the most stable E_i value corresponds to the γ -B[a] configuration in which the interaction occurs with the end phosphate of the ATP molecule. The E_i values for the arylisouronium-ATP complexes are larger than those of the arylguanidinium interactions previously reported (Figure 3.6.3); in particular, ΔE_i between isouronium and guanidinium complexes **VII** and **V** increases according to the E_i values increase.

This comparison can lead to the conclusion that, from a theoretical point of view, the interaction between arylisouronium and ATP is favoured with respect to that between arylguanidinium and ATP. However, an exception to this trend is found in complexes **VIII** which show slightly stronger E_i with the arylguanidinium than with the isouronium

analogue because the guanidinium cation has one extra H atom that may interact with the phosphate groups.

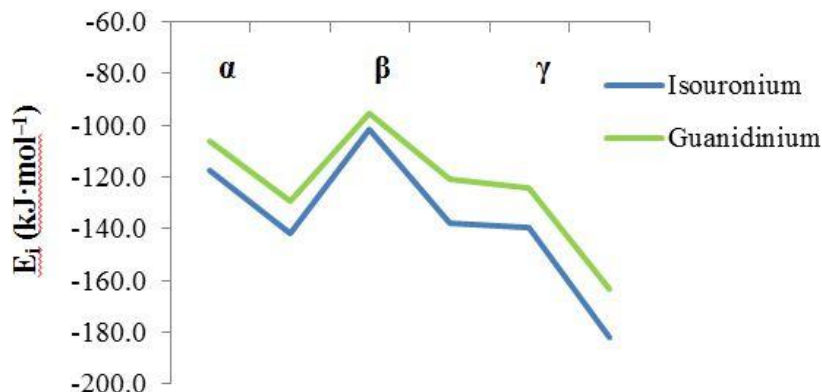
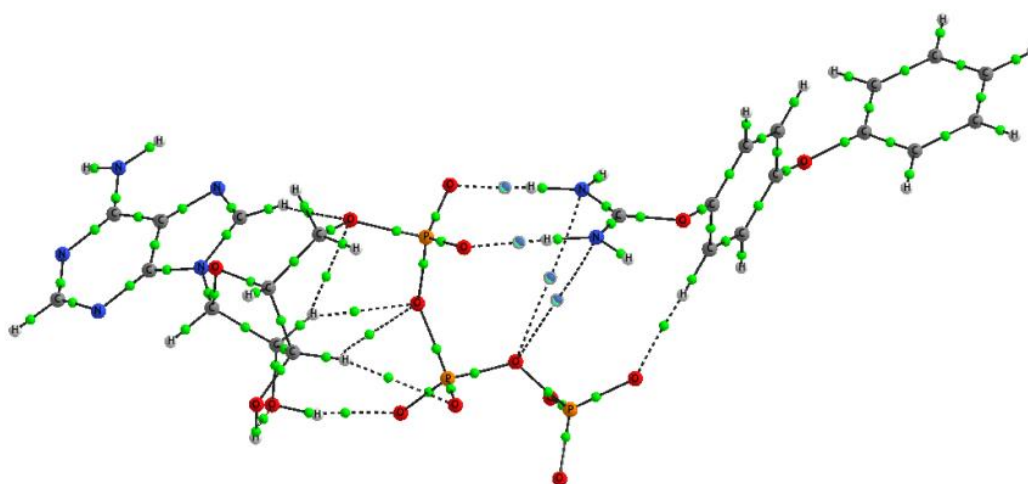


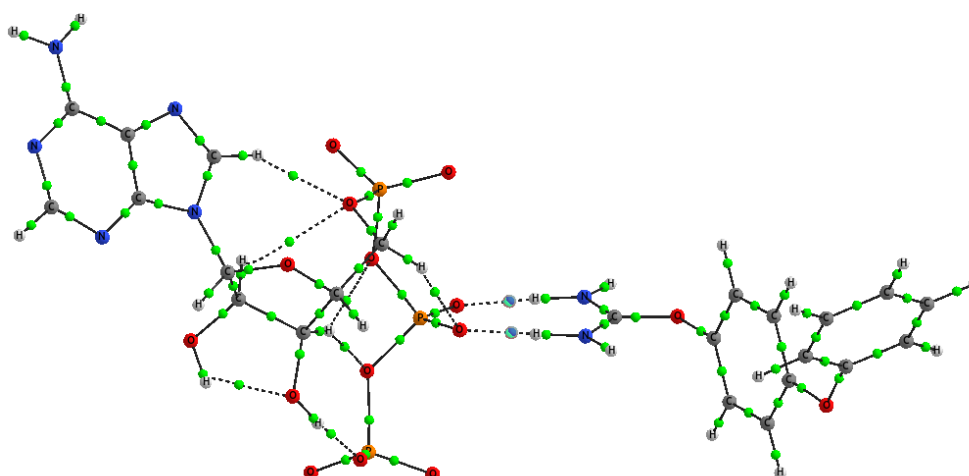
Figure 3.6.3. Line graph showing a comparison of interaction energies of arylisouronium (blue) and arylguanidinium (green) complexes **VII** and **V** with ATP.

3.6.2.2. Electron density analysis

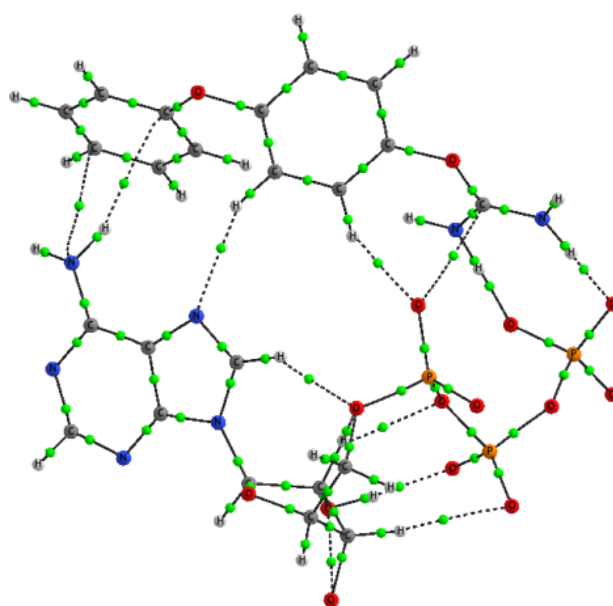
As previously reported for the arylguanidinium-ATP complexes, by means of the AIM theory, the characteristics of the BCPs found between the interacting moieties (i.e. ρ_{BCP} and $\nabla^2\rho_{\text{BCP}}$) were calculated and are presented in the Appendix (Table A.1.3.1. and Table A.1.3.2.). The calculated ρ_{BCP} values and the positive Laplacians indicate that, also in this case, all the interactions established within complexes **VII** and **VIII** are of non-covalent type and can be classified as HBs. For the sake of clarity, only the molecular graphs resulting from the AIM analysis corresponding to the most stable complexes **VII** interacting with the α -, β - and γ -phosphates are exhibited in Figure 3.6.4 and the rest are presented in the Appendix (Figure A.1.3.1.)



VII α -R[a]



VII β -R[a]



VII γ -B[a]

Figure 3.6.4. Molecular graphs corresponding to the most significant complexes **VII** for each α , β and γ phosphate. Green dots correspond to the bond critical points.

The results from the AIM theory analysis show that interactions in complexes **VII** occur by means of parallel HBs established between two O atoms of the phosphates in the ATP molecule and two H belonging to different N atoms of the arylisouronium. The ρ_{BCP} values for the arylisouronium-ATP interactions follow the order $\gamma > \beta > \alpha$, which is the same trend observed for the E_i values and, similarly, the ρ_{BCP} for the arylisouronium complexes are larger than for the guanidiniums.

It is interesting to mention that only for the complex **VII γ -B[a]** non-covalent interactions between the adenosine of the ATP molecule and both aryl groups of the isouronium are established. These extra interactions stabilise the complex and it is in completely agreement with the largest E_i of complex **VII γ -B[a]**.

Similar to the arylguanidium-ATP complexes, analysis of all complexes **VIII** indicates multiple interactions between the arylisouronium and the ATP phosphates as, for example, complex **VIII γ -R[a]** in which the isouronium moiety establishes HBs with the three phosphates of the ATP molecule (Appendix, Figure A.1.3.1.). As mentioned, complexes **VIII** present larger E_i than complexes **VII** because the interaction occurs

through multiple arylisouronium-ATP HBs as shown in the Appendix, Figure A.1.3.1.

As in complexes **VII**, the largest ρ_{BCP} value in complexes **VIII** corresponds to the interaction between the O atom of the γ phosphate and one of the H of the isouronium cation. In addition, an extra interaction between an O of the ATP molecule and an aromatic H atom of the arylisouronium is found in all complexes **VIII**; however, the low ρ_{BCP} values observed for these extra interactions indicate that they are very weak interactions (Appendix, Table A.1.3.2.)

3.7. Conclusions

In this chapter we have computationally studied and analysed the interactions between lead compound **1** and both an active, ATP-containing model of B-Raf and an isolated ATP molecule.

The best approach for designing new and more selective B-Raf allosteric inhibitors is the structure-based design and to pursue such endeavour a model of ATP containing active B-Raf is required since no crystal structure of such a system has been resolved. We have presented here a thoroughly constructed model of a phosphorylated, ATP-containing B-Raf protein (**2pBRAf_ATP**), which has been developed through a systematic MD approach (800 ns) taking into account the different options of presence/absence of both the phosphate group at the Thr599 site and the ATP molecule in the corresponding four systems. Different structural features have been analysed for each trajectory including the DFG motif and Mg^{2+} binding loop, activation loop and phosphorylation site as well as the α C-helix region. Moreover, the structure and activating interactions of the **2pBRAf_ATP** model have been investigated. Finally, the **2pBRAf_ATP** model has been used for a molecular docking study at semi-empirical level of theory with our previously developed lead compound **1**, finding a suitable docking pose that sees, among the others, also the interaction of the ATP γ -phosphate with the mono-substituted guanidinium of lead compound **1**, supporting previous hypothesis of type III allosteric inhibition.

Additionally, we have also theoretically studied the interaction of 4-phenyloxyphenyl-*N*-guanidinium (mimic of compound **1**) with the phosphate anion and with the isolated ATP molecule by means of DFT calculations. Guanidinium derivatives are able to form strong HB interactions not only with the O atoms of an isolated phosphate anion, but also with the different phosphate groups of the ATP molecule, where parallel interactions were found with one or two phosphate groups simultaneously. Binding energies and AIM analysis showed the complexes formed between ATP γ -phosphate and both guanidinium and isouronium to be the most stable ones. The formation of these interactions and their characterisation, indicates the suitability of guanidinium (or isouronium) derivatives as potential type III kinase's inhibitors.

3.8. References

1. Zheng, W.; Thorne, N.; McKew, J. C., *Drug Discov. Today* **2013**, 18, 1067.
2. Lounnas, V.; Ritschel, T.; Kelder, J.; McGuire, R.; Bywater, R. P.; Foloppe, N., *Comput. Struct. Biotechnol. J.* **2013**, 5, e201302011.
3. Hubbard, R. E., *Structure-based drug discovery: an overview*. Royal Society of Chemistry, **2006**.
4. Macalino, S.; Gosu, V.; Hong, S.; Choi, S., *Arch. Pharmacol. Res.* **2015**, 38, 1686.
5. Leach, A. R., *Molecular modelling: principles and applications*. Pearson education, **2001**.
6. Jensen, F., *Introduction to computational chemistry*. John Wiley & Sons, **2017**.
7. Morokuma, K., *Philos. Trans. R. Soc. London, Ser. A* **2002**, 360, 1149.
8. Zhou, T.; Huang, D.; Caflisch, A., *Curr. Top. Med. Chem.* **2010**, 33.
9. Nicklaus, M. C.; Wang, S.; Driscoll, J. S.; Milne, G. W. A., *Biorg. Med. Chem.* **1995**, 3, 411.
10. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A., *J. Am. Chem. Soc.* **1995**, 117, 5179.

11. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S. H.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., *Acc. Chem. Res.* **2000**, *33*, 889.
12. Diez-Cecilia, E.; Carson, R.; Kelly, B.; van Schaeybroeck, S.; Murray, J. T.; Rozas, I., *Bioorg. Med. Chem. Lett.* **2015**, *25*, 4287.
13. Diez-Cecilia, E.; Kelly, B.; Perez, C.; Zisterer, D. M.; Nevin, D. K.; Lloyd, D. G.; Rozas, I., *Eur. J. Med. Chem.* **2014**, *81*, 427.
14. Samatar, A. A.; Poulikakos, P. I., *Nat. Rev. Drug Discov.* **2014**, *13*, 928.
15. Nishiguchi, G. A.; Rico, A.; Tanner, H. R.; Aversa, R. J.; Taft, B. R.; Subramanian, S.; Setti, L.; Burger, M. T.; Wan, L.; Tamez, V. N., *J. Med. Chem.* **2017**.
16. Lavoie, H.; Therrien, M., *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 281.
17. Previtali, V.; Trujillo, C.; Boisson, J.; Khartabil, H.; Henon, E.; Rozas, I., *Submitted to Phys. Chem. Chem. Phys.* **2017**.
18. Trujillo, C.; Previtali, V.; Rozas, I., *Theor. Chem. Acc.* **2016**, *135*.
19. Leelananda, S. P.; Lindert, S., *Beilstein J. Org. Chem.* **2016**, *12*, 2694.
20. Gruszczyński, P.; Obuchowski, M.; Kaźmierkiewicz, R., *J. Comput. Aided Mol. Des.* **2010**, *24*, 733.
21. Gosu, V.; Choi, S., *Sci. Rep.* **2014**, *4*, 5748.
22. Verkhivker, G., *Mol. Biosyst.* **2016**, *12*, 3146.
23. Tang, H.-C.; Chen, Y.-C., *Int. J. Nanomedicine* **2015**, *10*, 3131.
24. Kuzmanic, A.; Sutto, L.; Saladino, G.; Nebreda, A. R.; Gervasio, F. L.; Orozco, M., *eLife* **2017**, *6*, e22175.
25. Hu, J.; Ahuja, L. G.; Meharena, H. S.; Kannan, N.; Kornev, A. P.; Taylor, S. S.; Shaw, A. S., *Mol. Cell. Biol.* **2015**, *35*, 264.
26. Jambrina, P. G.; Rauch, N.; Pilkington, R.; Rybakova, K.; Nguyen, L. K.; Kholodenko, B. N.; Buchete, N. V.; Kolch, W.; Rosta, E., *Angew. Chem. Int. Ed.* **2016**, *55*, 983.
27. Taylor, S. S.; Kornev, A. P., *Trends Biochem. Sci.* **2011**, *36*, 65.
28. Xie, P.; Streu, C.; Qin, J.; Bregman, H.; Pagano, N.; Meggers, E.; Marmorstein, R., *Biochemistry* **2009**, *48*, 5187.
29. Bollag, G.; Tsai, J.; Zhang, J. Z.; Zhang, C.; Ibrahim, P.; Nolop, K.; Hirth, P., *Nat. Rev. Drug Discov.* **2012**, *11*, 873.
30. Kornev, A. P.; Haste, N. M.; Taylor, S. S.; Ten Eyck, L. F., *Proc. Natl. Acad. Sci.* **2006**, *103*, 17783.

31. Rozas, I., *Phys. Chem. Chem. Phys.* **2007**, 9, 2782.
32. Nolen, B.; Taylor, S.; Ghosh, G., *Mol. cell* **2004**, 15, 661.
33. McClendon, C. L.; Kornev, A. P.; Gilson, M. K.; Taylor, S. S., *Proc. Natl. Acad. Sci.* **2014**, 111, 16973.
34. Davies, H.; Bignell, G. R.; Cox, C.; Stephens, P.; Edkins, S.; Clegg, S.; Teague, J.; Woffendin, H.; Garnett, M. J.; Bottomley, W.; Davis, N.; Dicks, E.; Ewing, R.; Floyd, Y.; Gray, K.; Hall, S.; Hawes, R.; Hughes, J.; Kosmidou, V.; Menzies, A.; Mould, C.; Parker, A.; Stevens, C.; Watt, S.; Hooper, S.; Wilson, R.; Jayatilake, H.; Gusterson, B. A.; Cooper, C.; Shipley, J.; Hargrave, D.; Pritchard-Jones, K.; Maitland, N.; Chenevix-Trench, G.; Riggins, G. J.; Bigner, D. D.; Palmieri, G.; Cossu, A.; Flanagan, A.; Nicholson, A.; Ho, J. W.; Leung, S. Y.; Yuen, S. T.; Weber, B. L.; Seigler, H. F.; Darrow, T. L.; Paterson, H.; Marais, R.; Marshall, C. J.; Wooster, R.; Stratton, M. R.; Futreal, P. A., *Nature* **2002**, 417, 949.
35. Wallace, A. C.; Laskowski, R. A.; Thornton, J. M., *Protein Eng. Des. Sel.* **1995**, 8, 127.
36. Rubez, G.; Etancelin, J. M.; Vigouroux, X.; Krajecki, M.; Boisson, J. C.; Hénon, E., *J. Comput. Chem.* **2017**, 38, 1071.
37. Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D. N.; Yang, W., *J. Chem. Theor. Comput.* **2011**, 7, 625.
38. Emrick, M. A.; Lee, T.; Starkey, P. J.; Mumby, M. C.; Resing, K. A.; Ahn, N. G., *Proc. Natl. Acad. Sci.* **2006**, 103, 18101.
39. Whittaker, S.; Kirk, R.; Hayward, R.; Zambon, A.; Viros, A.; Cantarino, N.; Affolter, A.; Nourry, A.; Niculescu-Duvaz, D.; Springer, C.; Marais, R., *Sci. Transl. Med.* **2010**, 2, 35ra41.
40. Taylor, S. S.; Keshwani, M. M.; Steichen, J. M.; Kornev, A. P., *Philos. Trans. R. Soc. London, Ser. B* **2012**, 367, 2517.
41. Madhusudan; Akamine, P.; Xuong, N.-H.; Taylor, S. S., *Nat. Struct. Mol. Biol.* **2002**, 9, 273.
42. Adams, J. A., *Biochemistry* **2003**, 42, 601.
43. Kannan, N.; Neuwald, A. F., *J. Mol. Biol.* **2005**, 351, 956.
44. Lahiry, P.; Torkamani, A.; Schork, N. J.; Hegele, R. A., *Nat. Rev. Genet.* **2010**, 11, 60.
45. Barberot, C.; Boisson, J.; Gerard, S.; Khartabil, H.; Thiriot, E.; Monard, G.; Henon, E., *Comput. Theor. Chem.* **2014**, 1028, 7.
46. Stewart, J. J. P., *J. Mol. Model.* **2013**, 19, 1.
47. Stewart, J. J., *Int. J. Quantum Chem.* **1996**, 58, 133.

48. Stewart, J. J. P., *Stewart Computational Chemistry, Colorado Springs, CO, USA, MOPAC* **2016**.
49. Rozas, I.; Alkorta, I.; Elguero, J., *Struct. Chem.* **2008**, *19*, 923.
50. Lefebvre, C.; Rubez, G.; Khartabil, H.; Boisson, J.-C.; Contreras-Garcia, J.; Henon, E., *Phys. Chem. Chem. Phys.* **2017**.
51. Cavallo, G.; Metrangolo, P.; Milani, R.; Pilati, T.; Priimagi, A.; Resnati, G.; Terraneo, G., *Chem. Rev.* **2016**, *116*, 2478.
52. Rozas, I.; Kruger, P. E., *J. Chem. Theory Comput.* **2005**, *1*, 1055.
53. Blanco, F.; Kelly, B.; Alkorta, I.; Rozas, I.; Elguero, J., *Chem. Phys. Lett.* **2011**, *511*, 129.
54. Kelly, B.; Sánchez-Sanz, G.; Blanco, F.; Rozas, I., *Comput. Theor. Chem.* **2012**, *998*, 64.
55. Rozas, I.; Sánchez-Sanz, G.; Alkorta, I.; Elguero, J., *J. Phys. Org. Chem.* **2013**, *26*, 378.
56. Blanco, F.; Kelly, B.; Sánchez-Sanz, G.; Trujillo, C.; Alkorta, I.; Elguero, J.; Rozas, I., *J. Phys. Chem. B* **2013**, *117*, 11608.
57. Marin-Luna, M.; Sanchez-Sanz, G.; O'Sullivan, P.; Rozas, I., *J. Phys. Chem. A* **2014**, *118*, 5540.
58. Rozas, I.; Sánchez-Sanz, G.; Alkorta, I.; Elguero, J., *J. Phys. Org. Chem.* **2013**, *26*, 378.
59. Ishikawa, T., *Superbases for organic synthesis: guanidines, amidines, phosphazenes and related organocatalysts*. John Wiley & Sons, **2009**.
60. Nagle, P. S.; Kahvedžić, A.; McCabe, T.; Rozas, I., *Struct. Chem.* **2012**, *23*, 315.

Chapter 4

Chapter 4 - Chemical Synthesis

4.1. Introduction

4.1.1. Guanidines and their therapeutic applications

The guanidine moiety is found in a good number of natural products.¹ It forms the nitrogenous backbone of guanine in DNA and RNA and it is also present in the amino acid arginine.² From a chemical point of view, guanidine is a very attractive functionality. It is considered as a superbases ($pK_{aH} = 13.6$ in water)³ and hence at physiological pH is found in its protonated guanidinium form. Given the presence of three N atoms with lone pairs, the guanidinium cation has three resonance form (Figure 4.1.1.) which confers a typical Y-shape planarity.⁴

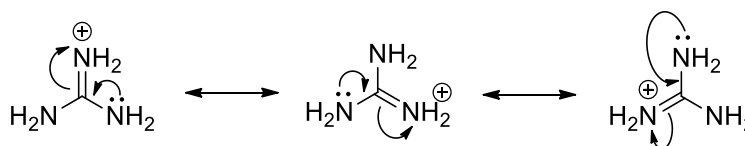


Figure 4.1.1. Resonance forms of the guanidinium cation.

Guanidinium cations can form ionic interactions with negatively charged amino acids such as aspartate or glutamate and HBs with accepting groups such as (C=)O, (C=)N or aromatic systems, which are present in some amino acids and in the DNA and RNA bases. Furthermore, due to its planarity, resonance, and positive charge, the guanidinium cation can form other types of interactions, more in particular, cation- π interactions.⁵

Given its exclusive characteristics and versatility, it is not a surprise that the guanidine moiety is found in different commercially available drugs.⁶ Few examples are

guanabenz and guanfacine (α_2 -adrenoceptor agonists used as an antihypertensive and in the treatment of attention deficit hyperactivity disorder, respectively),^{7,8} metformin for the treatment of type II diabetes⁹ or relenza (zanamivir) one of the first commercialized anti-influenza virus agents (Figure 4.1.2.).¹⁰

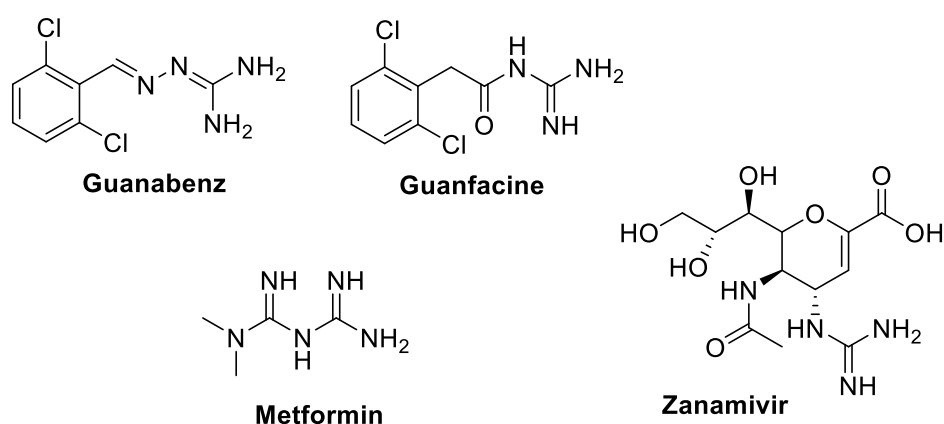


Figure 4.1.2. Biological substances containing guanidine-like moieties.

Our group has widely worked with guanidine and 2-aminoimidazoline derivatives and some of these compounds show high antagonistic affinity at the α_2 -AR,¹¹ binding to the DNA minor groove¹² or initiation of anticancer activity through the induction of cellular apoptosis.¹³

But guanidine's applications are not only widespread in the medicinal chemistry field, but they are also employed as base catalysts, chiral auxiliaries, and ligands in organic synthesis.³ Consequently, a large number of synthetic methods have been developed to install the guanidine unit and few of them are discussed in the next sections.¹⁴⁻¹⁶

4.1.2. Synthetic methods for the introduction of the guanidine functionality

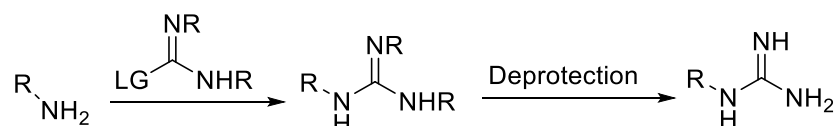
As described in the previous sections, a wide variety of guanidine's derivatives can be obtained both from nature and through synthetic procedures. Given their importance and relevance because of their biological activity, the synthesis of guanidines has come a long way and advances in the field have led to a myriad of methods for the

installation of the guanidine functional group.^{17,16,18,19} The following sections describe current methods used in the Rozas' group as well as few alternative guanidylation methods for the synthesis of acyclic mono-substituted guanidine and acyclic *N,N'*-disubstituted guanidines, given their relevance for the target products of this research.

4.1.2.1. Synthesis of acyclic mono-substituted guanidines

Classical methods for the synthesis of guanidines involve the reaction of an amine with an activated/protected guanidine precursor followed by the deprotection of this moiety to yield the corresponding free guanidine (Scheme 4.1.1.).

Scheme 4.1.1.



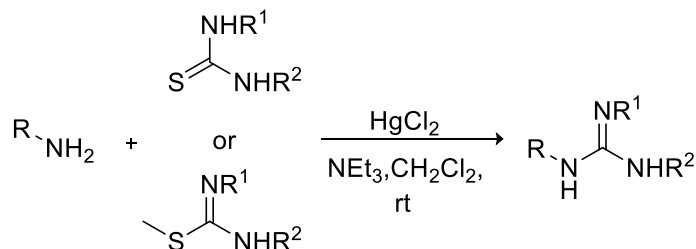
These methods usually utilise thiourea²⁰ or *S*-methyliothiourea² systems to react with an amine *via* the activation with a variety of reagents (Scheme 4.1.2.). In the case of anilines, Lewis acid promoted reactions have been the most successful while for aliphatic amines, there is a much wider scope of guanidine synthon available due to their greater nucleophilicity.

In particular, the use of mercury(II) chloride (HgCl₂) as a promotor for guanidylation is the most frequently employed approach and is the chosen method for the research conducted in this project. This guanidylation method, developed by Kim and Qian,²⁰ consists on the reaction of an amine with *N,N'*-bis-(*tert*-butoxycarbonyl)thiourea in the presence of HgCl₂ and triethylamine (NEt₃) to yield bis-Boc-protected guanidines, which upon deprotection will yield the appropriate guanidinium derivative (Scheme 4.1.2.).

However, the use of HgCl₂ is an obvious drawback in the above described methodology and it can prevent its larger scale applications. Recent findings in Rozas' group,²¹

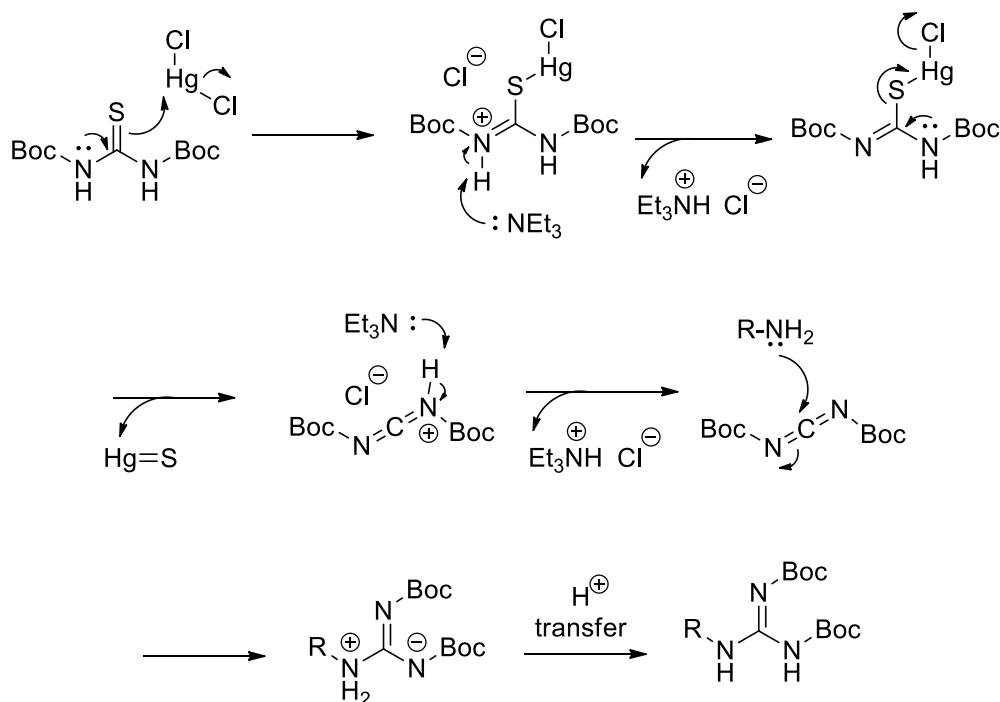
establish that CuCl_2 can be used as a Lewis acid with similar yields to those of the benchmark HgCl_2 reaction.

Scheme 4.1.2.



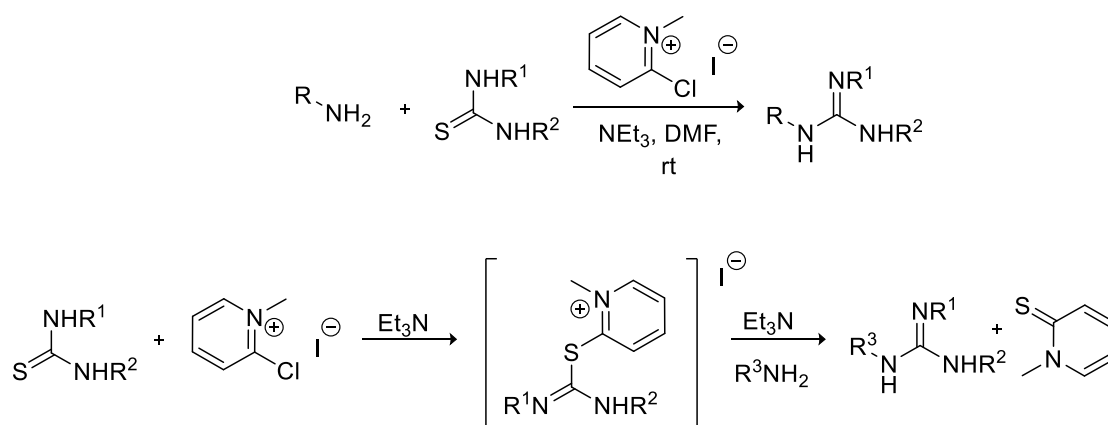
In the original article, it was suggested that the reaction is initiated by mercury-sulphur coordination, followed by trimethylamine deprotonation, leading to the electrophilic intermediates carbodiimide species. Next, the carbodiimide undergoes nucleophilic attack, even by deactivated amines, leading to the bis-Boc-protected guanidine after proton exchange (Scheme 4.1.3).

Scheme 4.1.3.



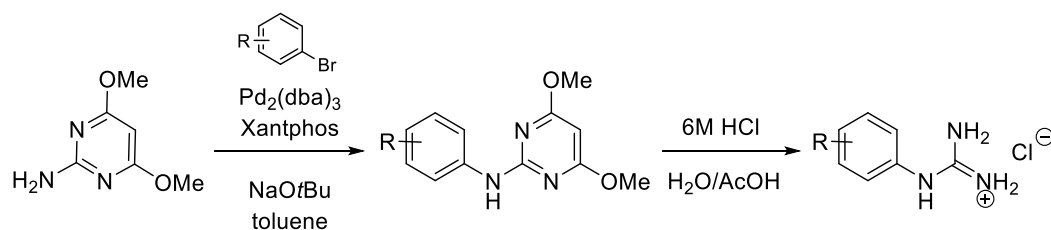
One alternative to the Kim and Qian mercury(II) chloride guanidylation method is the use of the Mukaiyama's reagent,²² which seems to be also based on the formation of an electrophilic carodiimide and its attack from a nucleophile amine (Scheme 4.1.4.).

Scheme 4.1.4.



Rozas' group has lately presented a novel method for the synthesis of aryl-substituted guanidines in good yield. Thus, 2-(arylamino)-4,6-dimethoxypyrimidines are prepared by coupling aryl bromides with 2-amino-4,6-dimethoxypyrimidine and then cleaved in strong acid conditions (6M HCl in H₂O/AcOH) to afford aryl-guanidines (Scheme 4.1.5.).²³

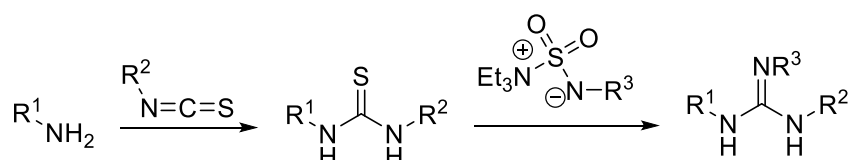
Scheme 4.1.5.



Recently, Maki *et al.* described another novel method for the synthesis of carbamate-protected guanidines from primary amines. Thus, in this method, a variety of thioureas

derived from primary amines and isothiocyanates react with the Burgess reagent²⁴ to give the corresponding guanidines. Selective deprotection of the products affords alkyl or aryl guanidines up to 86% yield (Scheme 4.1.6.).²⁵

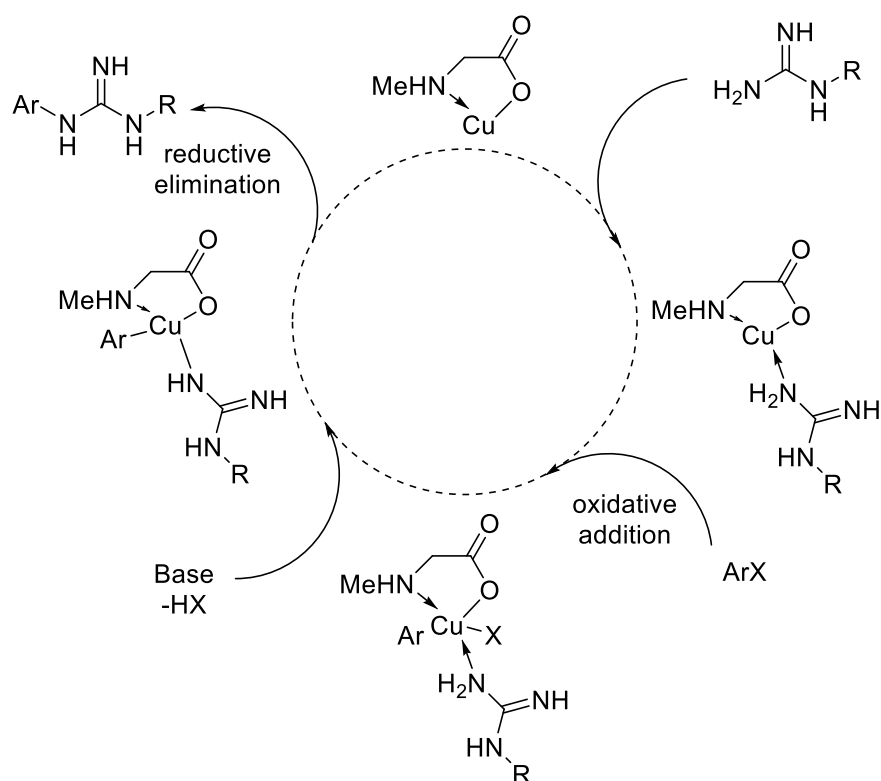
Scheme 4.1.6.



The choice of guanidine precursor depends on the reactivity of the amine involved and more recent developments in the area have seen the development of alternative methods to the classic $HgCl_2$ promoted reactions. In a number of cases the coupling reagent EDCI [1-ethyl-3-(3-dimethylaminopropyl)carbodiimide] has shown to promote guanidylation between thiourea derivatives and nucleophilic amines.¹⁸ Carbodiimide derivatives also can react with aniline derivatives in the presence of catalytic heterogeneous metal (as the nanocrystalline zinc oxide, ZnO).²⁶

Copper coupling chemistry has seen a renaissance in the past decade.^{2,27,28} If the previous methods all see nucleophilic anilines reacting with a guanidine precursor (Scheme 4.1.1.), copper catalysed cross coupling chemistry allows an aryl halide to be reacted with a nucleophilic guanidine source (Scheme 4.1.7.). This method skips the use of alkyl or aryl amines, overcoming poor nucleophilicity limitations.

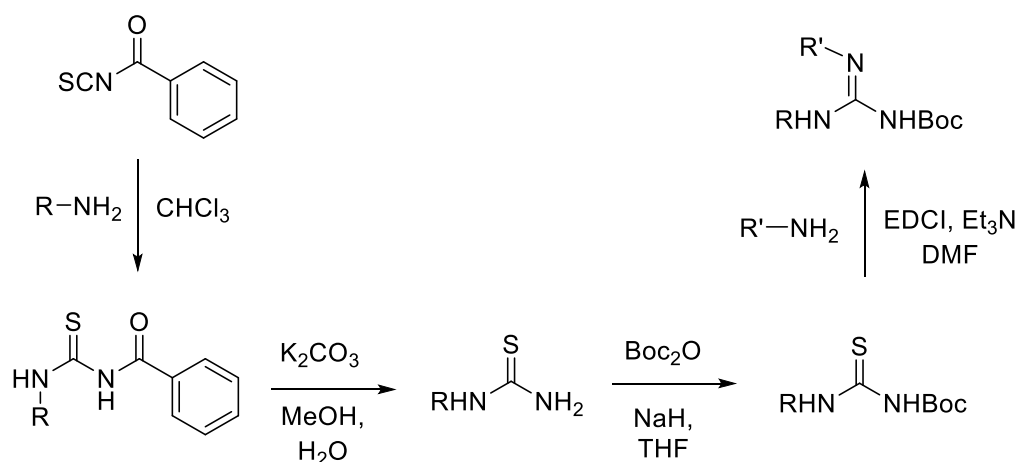
Scheme 4.1.7.



4.1.2.2. Synthesis of acyclic *N,N'*-disubstituted guanidines

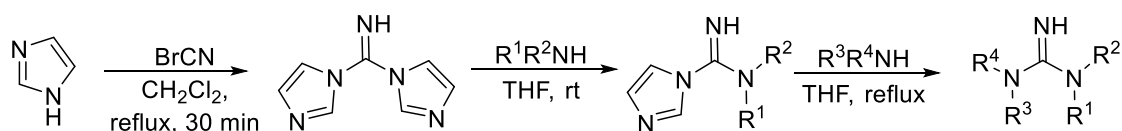
While Kim and Qian focused primarily on mono-substituted guanidine derivatives such as those discussed in the previous section, in 1992 Poss and Iwanowicz described procedures for the synthesis of *N,N'*-disubstituted guanidines from *N*-Boc protected *N'*-substituted thiourea derivatives. In one of these methods, *tert*-butyl carbamate is reacted with an alkyl or aryl isothiocyanate to give a Boc protected thiourea derivative, which is then subjected to guanidylation using their 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDCI) procedure (Scheme 4.1.8.).²⁹

Scheme 4.1.8.

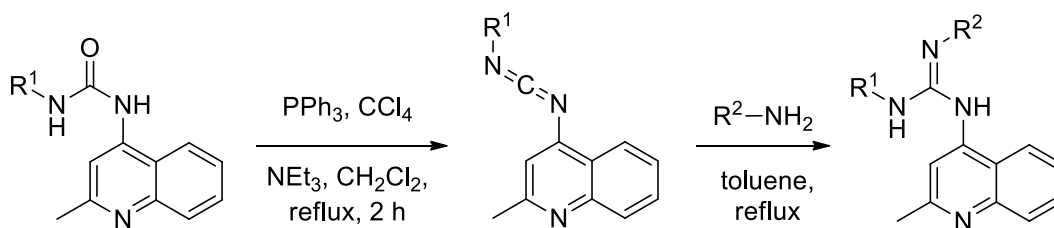


The use of di(imidazole-1-yl)methanimine has been reported to be effective for the synthesis of *N,N'*-disubstituted guanidines.^{2,30,31} Reaction of imidazole with cyanogen bromide, affords the desired *N,N'*-disubstituted guanidines. Exposure to nucleophilic amines allows the displacement of the imidazole leaving groups and the generation of structurally different guanidines (Scheme 4.1.9).

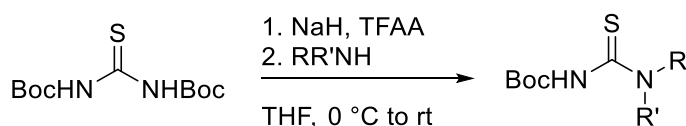
Scheme 4.1.9.



Guanidine has also been synthesized from urea using various dehydrating agents like thionyl chloride³² or typical Appel type conditions.³³ One such example is the synthesis of a library of quinoline reported by Rachlin *et al.* (Scheme 4.1.10).³⁴ Urea derivatised quinolines were exposed to PPh_3 , CCl_4 and NEt_3 in refluxing CH_2Cl_2 in Appel conditions to form crude carbodiimides which, after exposure to a number of amines, yielded *N,N',N''*-trisubstituted guanidines in varying yields (Scheme 4.1.10).

Scheme 4.1.10.

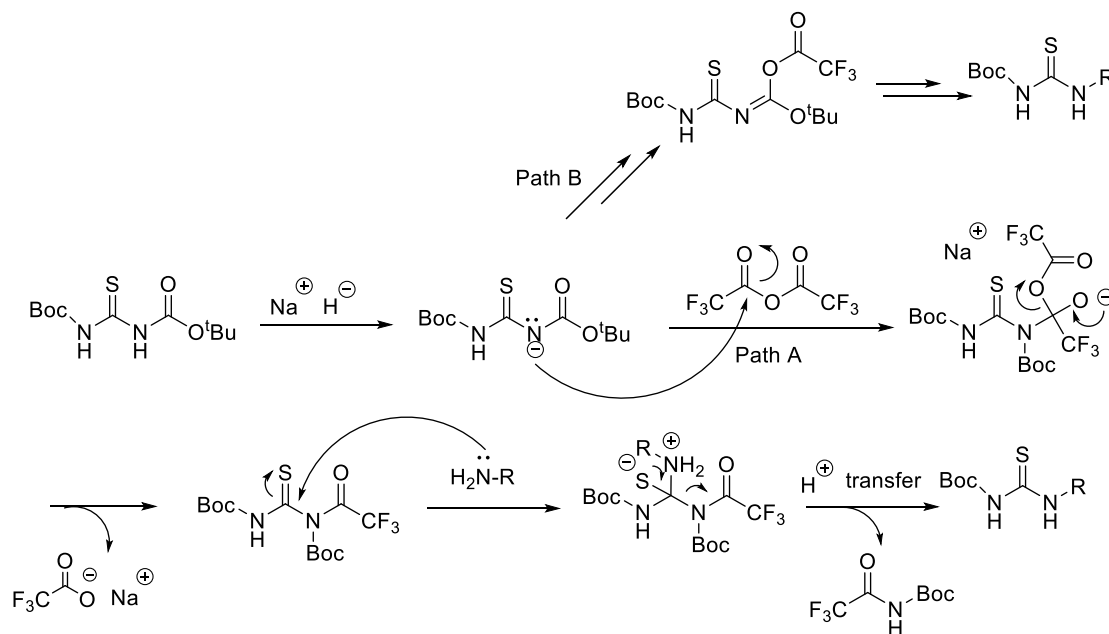
More recently, Yin *et al.* have described the preparation of *N*-Boc-*N'*-substituted thioureas by treatment of *N,N'*-di-Boc-substituted thiourea with sodium hydride (NaH) and trifluoroacetic anhydride (TFAA) in the presence of an aromatic or aliphatic amine (Scheme 4.1.11.). Next, previously described reaction with amine derivatives in presence of mercury(II) chloride, will afford the desired *N,N'*-disubstituted guanidines. This method was later optimized in Rozas' group for the synthesis of acyclic *N,N'*-disubstituted guanidines given high to excellent yields for both aromatic or aliphatic amines³⁵ and, therefore, it has been chosen for the present research.

Scheme 4.1.11.

Yin describes two possible mechanisms for this transformation.³⁶ On the one hand, it is proposed that the anion formed by deprotonation of *N,N'*-di-Boc-substituted thiourea is *N*-acylated to produce the *N*-Boc-*N*-trifluoroacetyl derivative, which then undergoes the nucleophilic attack by the amine, with elimination of the *N*-Boc-*N*-trifluoroacetyl anion (Scheme 4.1.12., path A). On the other hand, the authors suggested an alternative mechanism involving acylation at the carbonyl oxygen of the Boc group to produce the *O*-trifluoroacetyl derivative (Scheme 4.1.12., path B). Considering that the suggested intermediate (*tert*-butyl-trifluoroacetyl carbamate) has been previously isolated in the

Rozas' group, we propose the mechanism of Yin's reaction likely to be the former (path A), in which the reaction proceeds through the *N*-Boc-*N*-trifluoroacetyl derivative.

Scheme 4.1.12.



4.2. Synthesis of 3,4'-bis-guanidino phenyloxybenzene derivatives: exploring the hydrophobic moiety

4.2.1. General aspects

As mentioned in the Introduction chapter (Paragraph 1.5., Previous Work Leading Up to the Present Research), previous results showed how improvements in the cytotoxic activity of previously synthesised compounds were achieved when the (3- CF_3 , 4-Cl)Ph lipophilic group was inserted in the guanidinium moiety in the 3-position instead of the 4'-position of the 3,4'-dianiline core (as in lead compound **1**), and also that the link of this diaromatic core of the molecules does not really affect their cytotoxic activity.¹³

Therefore, for the preparation of the new 3,4'-bis-guanidino phenyloxybenzene derivatives proposed for this research, we chose the oxygen linker, as in the lead compound **1** (Figure 4.2.1.), because of the good biological results obtained in the previous series.¹³ Furthermore, the corresponding starting diamine, 3-(4-

aminophenoxy)aniline is commercially available, making the synthesis of different derivatives easily accessible.

Additionally, we introduced changes at the substituents of the lipophilic moiety (Figure 4.2.1.) to probe the lipophilic pocket of the target (B-Raf) and to gain a better understanding of the requirement for an optimal allosteric binding to protein kinases. We then propose to synthesise 3,4'-bis-guanidino phenyloxybenzene derivatives **2-6** with lipophilic substituents in the 3-position and some of the corresponding analogues with an aromatic group in the guanidine in the 4'-position **7-9** for the sake of comparison (Figure 4.2.1.).

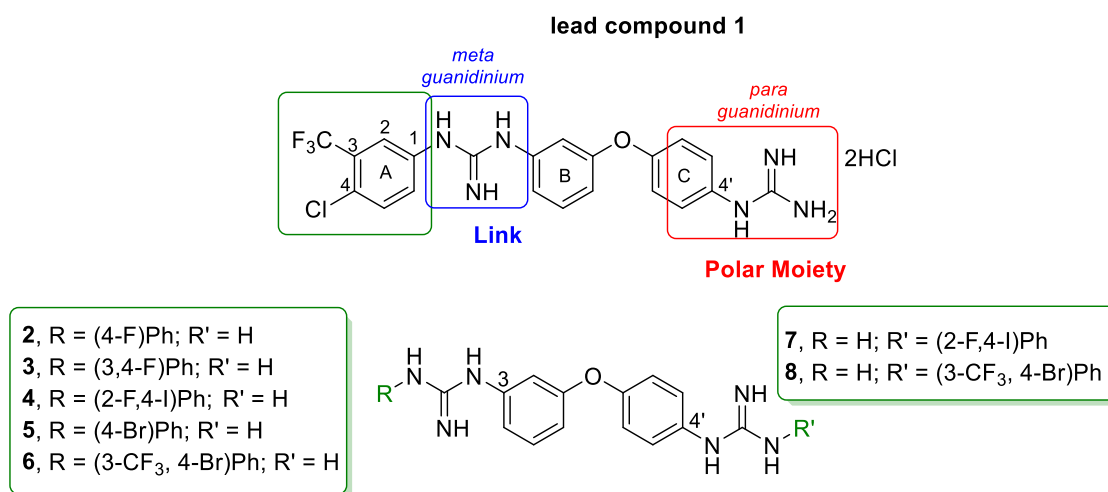


Figure 4.2.1. Lead compound **1** and proposed changing at the hydrophobic moiety at 3 and 4' position of the 3-(4-aminophenoxy)aniline.

Halogen substituted aryl groups as those chosen for 3,4'-bis-guanidino phenyloxybenzene derivatives **2-6** and **7-8** have been widely used in compounds binding at hydrophobic allosteric sites of some protein kinases, including MEK and Raf protein kinases.^{37,38} Compounds containing Cl⁻, Br⁻, or I⁻ can form direct close halogen contacts (R-X...Y-R'), where the halogen X acts as a Lewis acid and Y can be any electron donor moiety.³⁷ This type of interaction, referred as 'halogen bonding' since 1978³⁷ can be categorized as a non-covalent interaction (NCI) and is driven by the σ -hole, a positively charged region on the hind side of the halogen (X) along the R-X bond axis that is caused by an anisotropy of electron density on the halogen.³⁷ Upon

formation of a covalent carbon-halogen σ -bond, a valence electron of the two atoms pair up forming the carbon-halogen bond. As a result, the p_z of the halogen opposite to the σ -bond becomes depopulated, resulting in an electropositive crown (Figure 4.2.2., in blue), whereas the p_{xy} orbitals retain their complement electrons to account for the overall negative charge of the halogen.³⁹

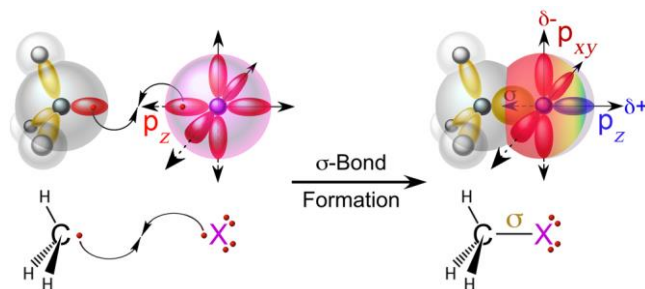


Figure 4.2.2. Scheme of the σ -hole formation, taken from Ford *et al.*³⁹

The existence of the σ -hole on the Valence Shell Charge Concentration (VSCC) of halogen atoms is a necessary condition in the formation of halogen bonds⁴⁰ but, as for other NCI (like the HB⁴¹), other geometrical features are fundamental for the establishment of such interaction. Figure 4.2.3. compares the geometries and types of donor and acceptor atoms for classic HBs and halogen bonds (XBs) seen in biomolecular systems.

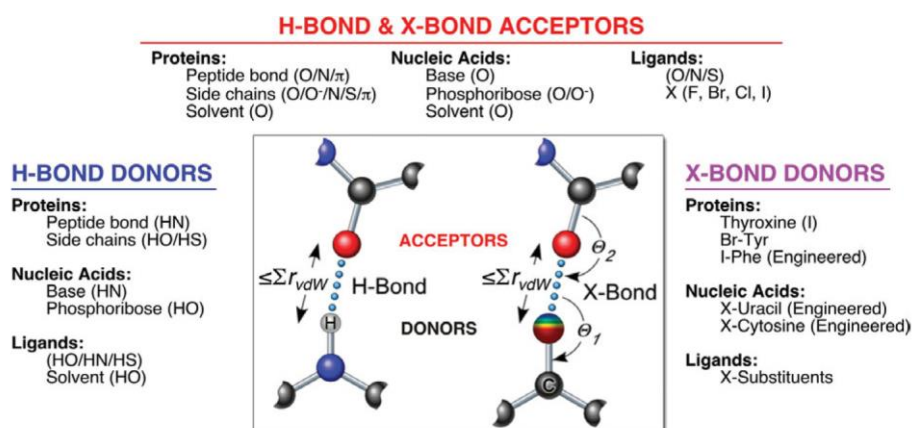


Figure 4.2.3. Hydrogen and halogen bonds geometries and requirements, taken from Scholfield *et al.*³⁸

Often, XBs formed by drug-like compounds binding to their biological target do not adopt optimal interaction geometries, given that their binding is modulated by the overall binding mode of the ligand and the protein environment; however, the distance between the halogen (X) and the Lewis base (LB) ($d_{X\cdots LB}$), the σ -hole-angle ($\alpha_{C(ar)-X\cdots LB}$) and the spatial orientation of the X with respect to the electron-rich regions of the LB should be considered to determine the quality of the interaction (Figure 4.2.3.). Thus, although the choice of X may not always increase affinity, it can still be relevant for inducing selectivity⁴² and XBs are proved to be a useful handle in drug-design.⁴³

4.2.2. Synthesis

The approach previously used in Rozas' group¹³ and briefly described in the last sections was used for the synthesis of the 3,4'-bis-guanidinium-like diphenyl derivatives in this study. In order to prepare these derivatives, the appropriate amines were synthesised, starting from the commercially available 3-(4-aminophenoxy)aniline and the in-house prepared *N,N'*-bis-(*tert*-butoxycarbonyl)thiourea. Then, substituted thioureas (carrying the lipophilic moiety) were prepared accordingly and reacted with the corresponding mono-guanidine diaromatic system (Figure 4.2.4.).

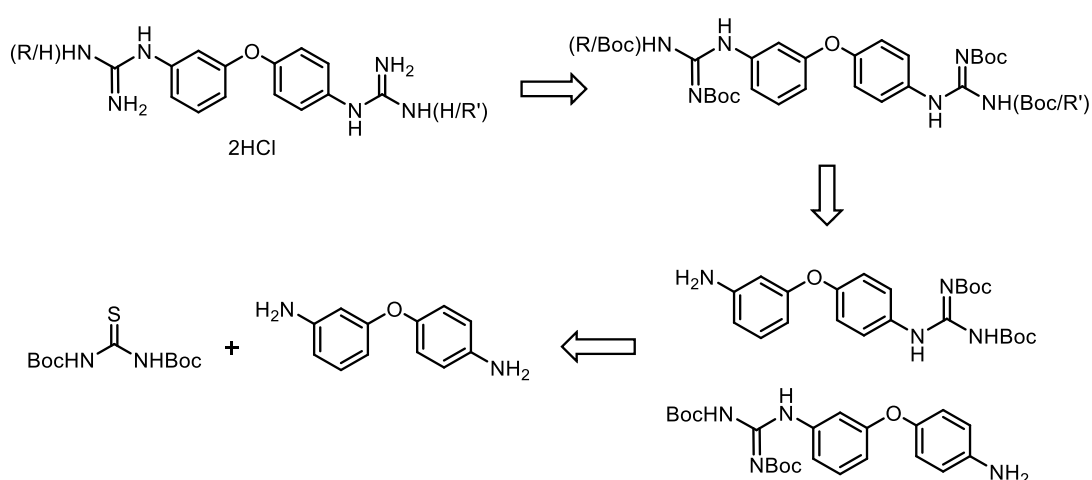
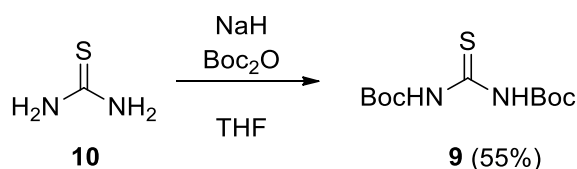


Figure 4.2.4. Retrosynthetic scheme for the preparation of 3,4'-bis-guanidino phenyloxybenzene derivatives.

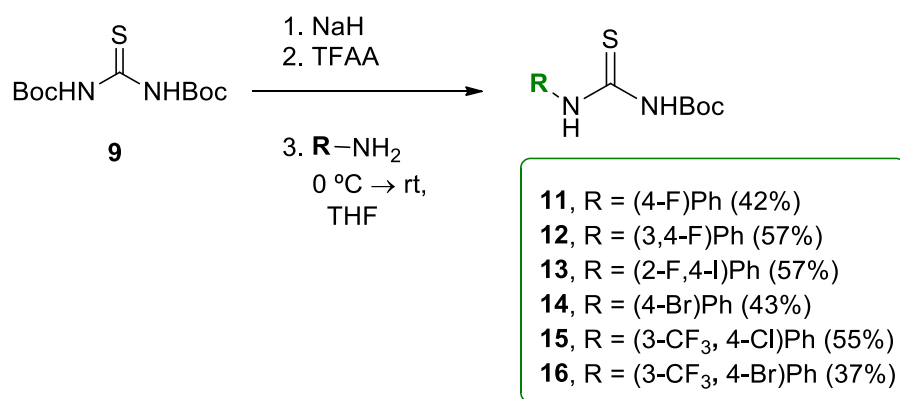
Taking into account that the preparation of the aimed 3,4'-bis-guanidino phenyloxybenzene derivatives first involves the introduction of a guanidine (Figure 4.2.4.), the first step consisted in the preparation of such “guanidylating” agent. Thus, *N,N'*-bis-(*tert*-butoxycarbonyl)thiourea **9**, was synthesised by the treatment of commercially available thiourea **10** with di-*tert*-butyldicarbonate (Boc₂O) and NaH in dry THF to form the Boc-activated ‘guanidylating’ agent in 55% yield (Scheme 4.2.1.).⁴⁴

Scheme 4.2.1.



The synthesis of the corresponding *N*-aryl-*N'*-Boc-protected thioureas **11-16** (Scheme 4.2.2.) was carried out following a previous procedure developed in the group¹³ and which mechanism has been reported in previous section 4.1.2.2. This reaction involves the mono deprotonation of previously synthesized *N,N'*-bis-(*tert*-butoxycarbonyl)thiourea **9** with NaH at 0 °C, following reaction with trifluoroacetic anhydride (TFAA) facilitates the nucleophilic addition of aromatic anilines to the carbon centre of the thiourea (Scheme 4.2.2.). Given the different electron-withdrawing effect of differently substituted commercially available anilines, different yields of the corresponding thioureas **11-16** were obtained (37-57%).

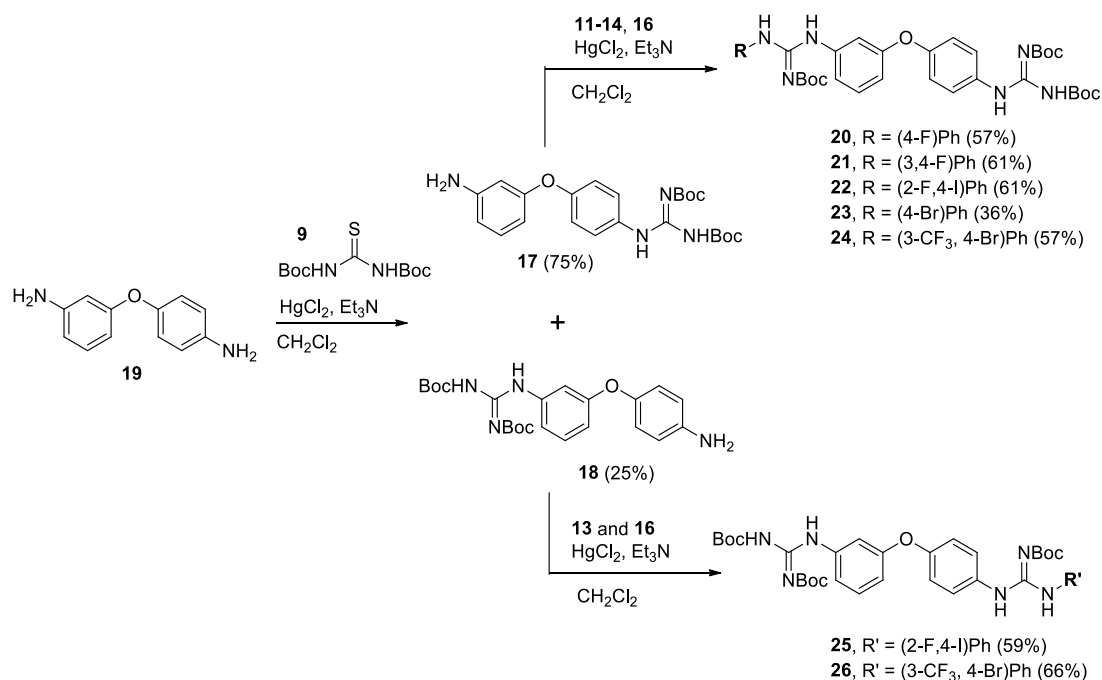
Scheme 4.2.2.



Once the synthesis of the *N*-aryl-*N'*-Boc-protected thioureas **11-16** was accomplished, the corresponding Boc-protected dipheyl mono-guanidines **17** and **18** were prepared by reaction of commercially available 3,4'-dianiline **19** and *N,N'*-bis-(*tert*-butoxycarbonyl)thiourea **9** in the presence of mercury(II) chloride and triethylamine yielding compounds **17** and **18** in moderate yield (Scheme 4.2.3.).

Considering the electron-donating effect of the oxygen linker, the nucleophilicity of the amine located in *para* position is increased and, therefore, the *para*-substituted derivative **17** was the major product compared to the *meta*-substituted one **18**.

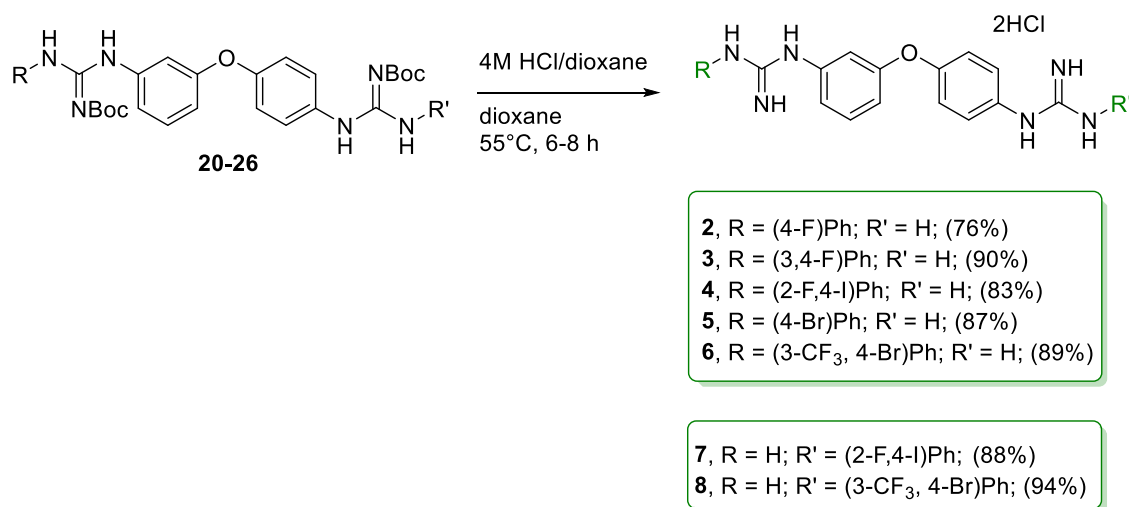
Scheme 4.2.3.



N-aryl-*N'*-Boc-protected thioureas **11-16** were then reacted with Boc-protected mono-guanidines **17** and **18** in similar conditions as for the preparation of Boc-protected mono-guanidines [mercury(II) chloride and trimethylamine, Scheme 4.1.2.] to obtain the corresponding 3-arylguanidine-4'-Boc-protected guanidine derivatives **20-24** and 4-arylguanidine-3'-Boc-protected guanidines **25** and **26**, as shown in Scheme 4.2.3.

Finally, *N,N'*-Bis-Boc-protected guanidine intermediates **20-26** were deprotected to yield hydrochloride salts **2-8** (Scheme 4.2.4.).¹³ Thus, treatment of the Boc-protected systems with 4M HCl in dioxane (6 eq. per Boc group) at 55 °C, until the reaction were complete by TLC (typically 8 h) afforded, after solvent and HCl evaporation under vacuum, a crude product that was purified by silica gel flash chromatography with a typical elution gradient of 95% CHCl₃ and 5% MeOH.

Scheme 4.2.4.



In a previous work in Rozas' group⁴⁵, it was found that intra-molecular HBs (IMHBs) could be formed between two N-H of the Boc protected guanidine (N_{1'} and N_{2'}) and two Boc CO groups as well as between the imine nitrogen N_{3'} and an aromatic hydrogen, rendering this molecule and derivatives planar (Figure 4.2.5.).

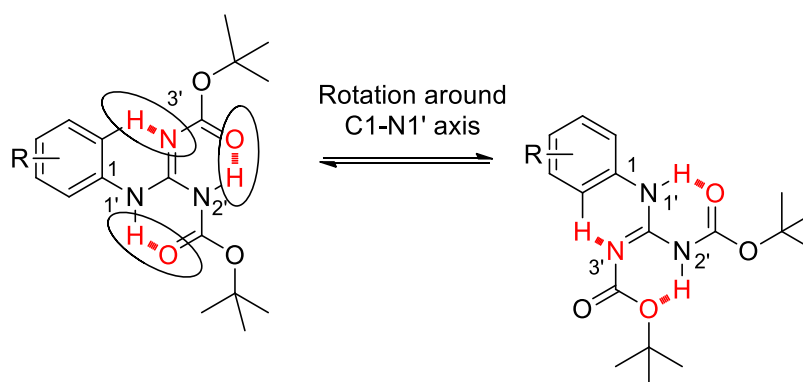


Figure 4.2.5. IMHBs formation in di-Boc-protected aryl guanidine.⁴⁵

This peculiarity helped the distinctive characterisation of compounds **17** and **18**. From the stack ¹H NMR spectra of compounds **17** and **18** (Figure 4.2.6.) the presence of the N_{3'}...H IMHB is clearly visible. Moreover, protons H-8 and H-8' can be found at a similar ppm value, around 7.00, both in compounds **18** (Figure 4.2.6., *top*) and **17** (Figure 4.2.6., *bottom*). Instead, protons H-9 and H-9' appear at very different chemical shifts for compounds **18** and **17**. This occurs because H-9 and H-9' in compound **17** form an IMHB with the Boc protected guanidine and, hence they are deshielded

appearing at around 7.60 ppm; on the contrary, in compound **18** the free NH₂ at the *para* position will shield H-9 and H-9' causing their shift upfield at around 6.70 ppm.

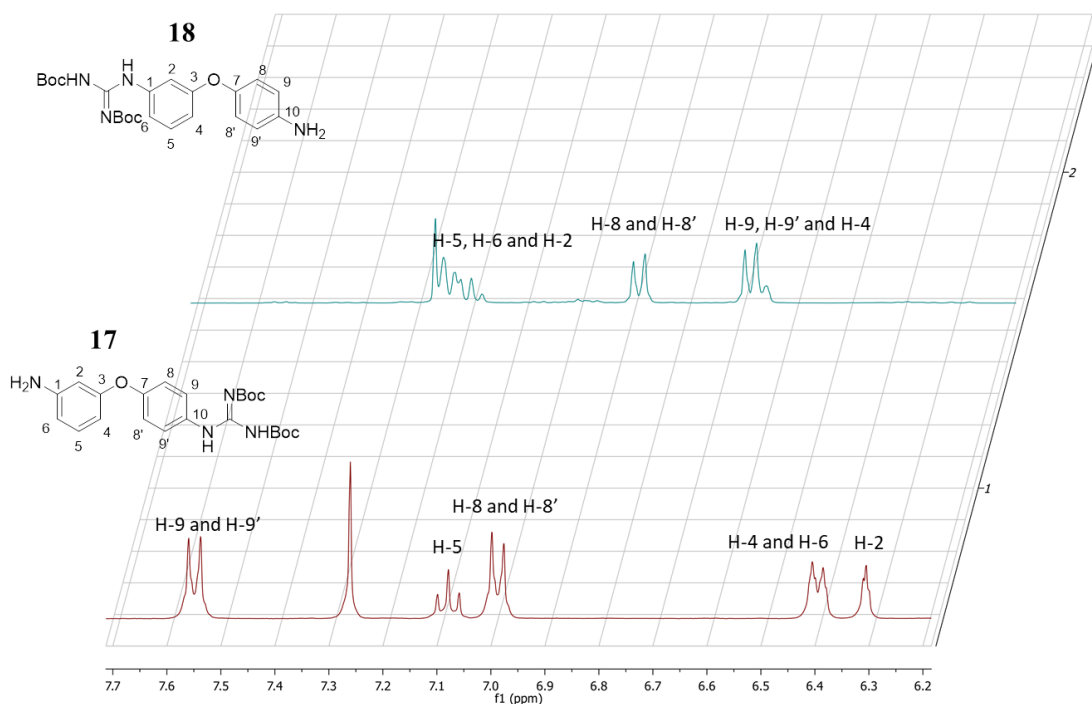


Figure 4.2.6. Stack ¹H NMR spectra of compounds **17** and **18**.

It has been previously observed in Rozas' group that the ¹H and ¹³C NMR spectra of different *N,N'*-bis-aryl-*N''*-Boc protected guanidines were always characterised by broad, poorly resolved signals. This phenomenon was attributed to the interconversion between different tautomers, which are induced by the sterically-congested environment surrounding the central guanidine moiety (Figure 4.2.7.).⁴⁶

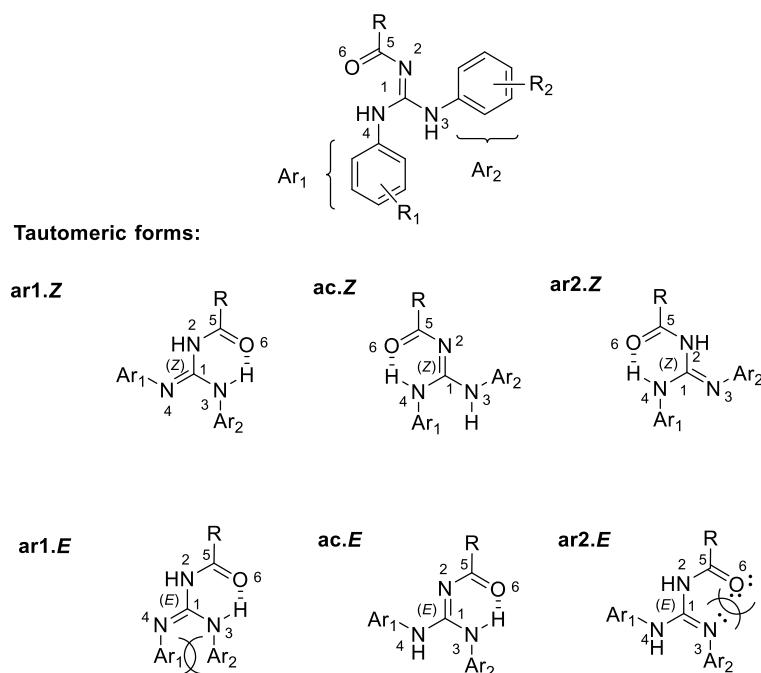


Figure 4.2.7. Different tautomeric forms of a general *N,N'*-bis-aryl-*N''*-Boc protected guanidine.⁴⁶

Due to this mixture of isomers and steric crowding between the two aromatic rings and the Boc group, the ^1H NMR and ^{13}C NMR signals for compounds **20-26** gave very broad and amorphous signals, making impossible the appropriate characterisation of these molecules. Hence, in some cases, 2D studies as HSQC, helped the correlation between protons and carbons, and the distinction between substituted carbons or quaternary ones. Figure 4.2.8. shows as example the ^1H NMR and ^{13}C NMR spectra of compound **20**.

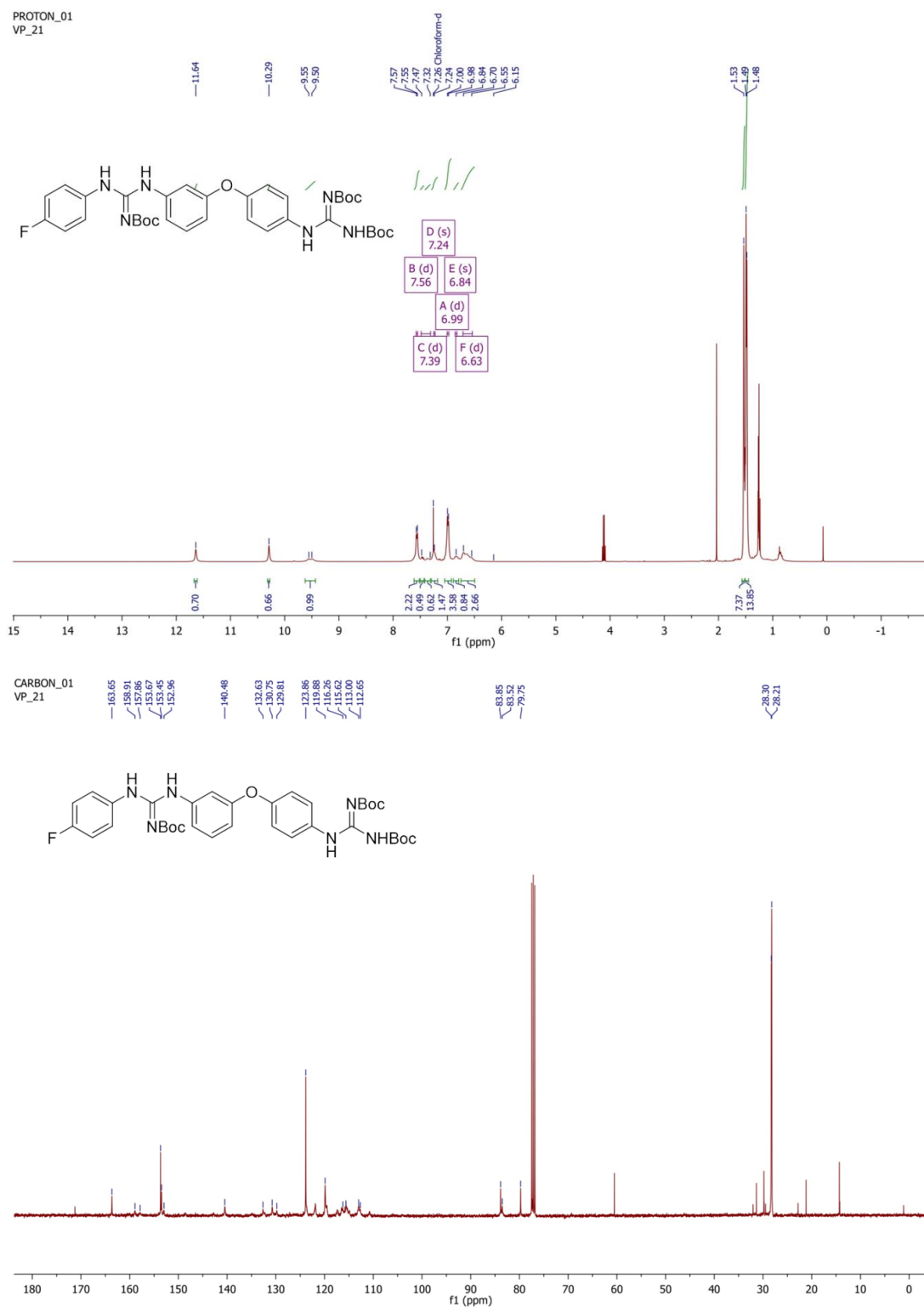
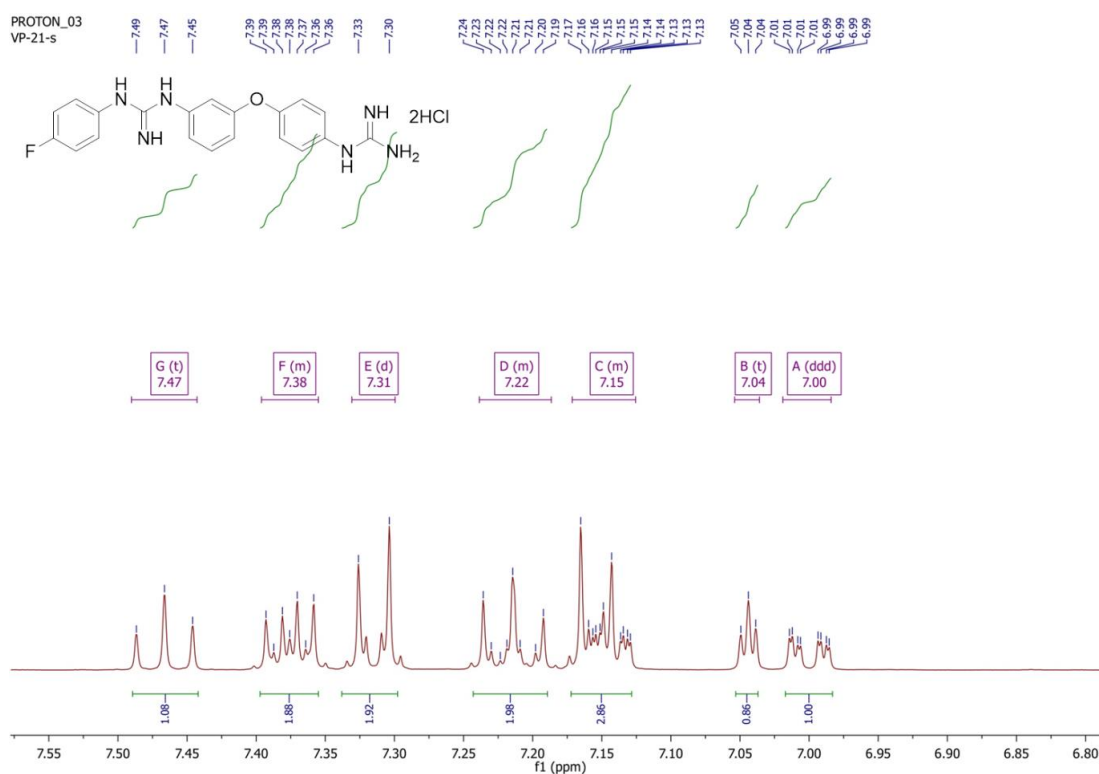


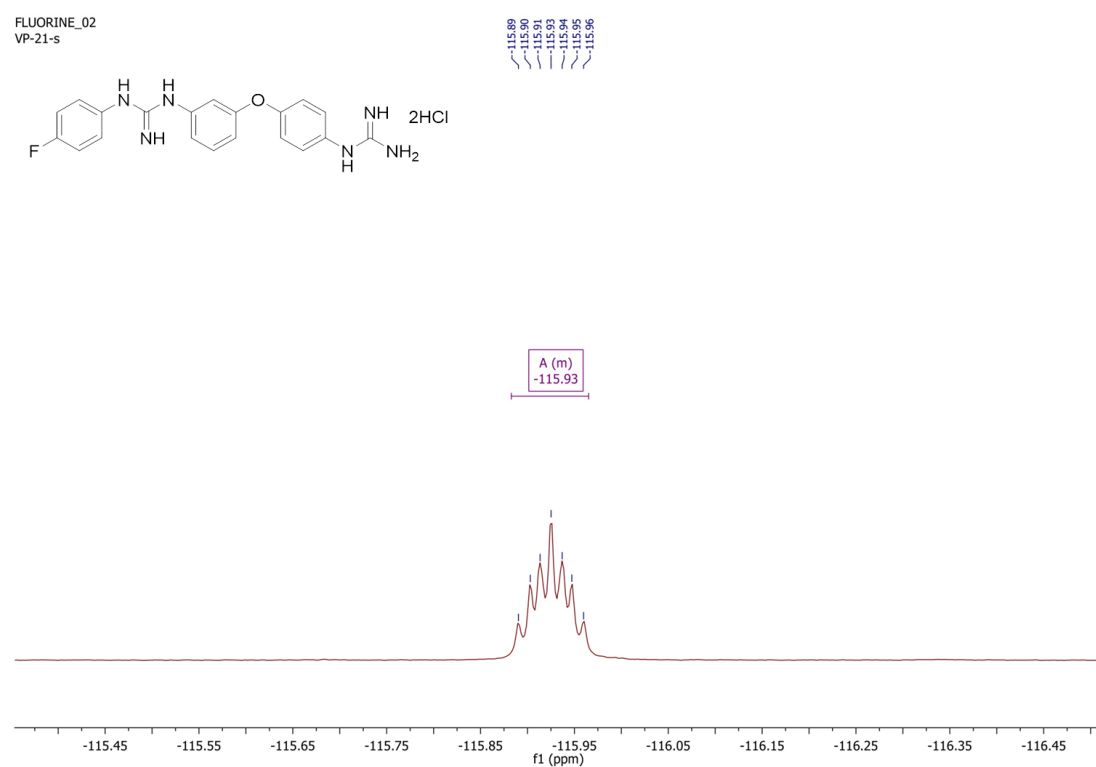
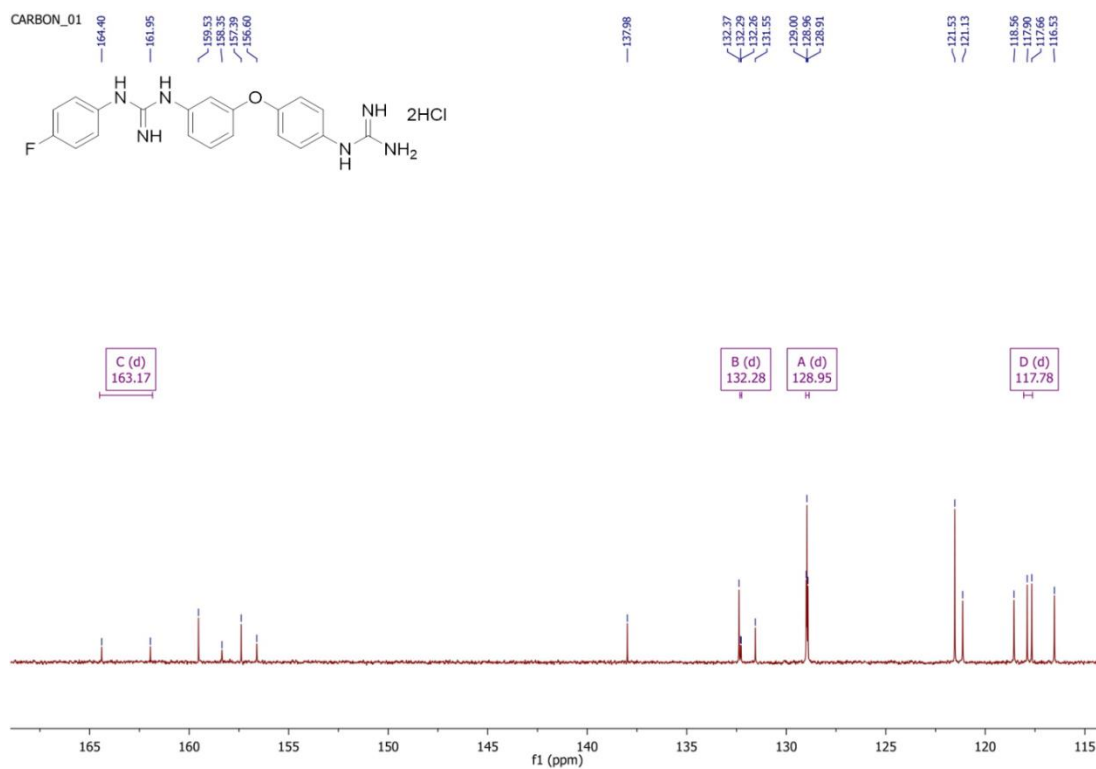
Figure 4.2.8. ^1H NMR and ^{13}C NMR spectra of compound 20.

In the case of fluorinated derivatives, the presence of one or more fluorine atoms, made the interpretation of the spectra even more complicated, due to the fluorine typical splitting with protons and carbons. Thankfully, deprotection of the Boc group,

alleviates the sterically-congested environment surrounding the central guanidine moiety making the ^1H NMR and ^{13}C NMR much clearer and defined for a full characterisation. Figure 4.2.9. shows as example ^1H NMR, ^{13}C NMR and ^{19}F NMR spectra of compound 2.

Coupling between H and F and between C and F are very strong, and characteristic, and this typical feature can be exploited for an easier characterisation of the ^1H NMR and ^{13}C NMR spectra. Thus, C-F couplings, slightly vary for *ortho*-, *meta*- or *para*-substituted fluorobenzenes, but typically $^1J_{\text{C-F}}$ is of 245 Hz, $^2J_{\text{C-F}}$ of 21 Hz and $^3J_{\text{C-F}}$ of 5 Hz. Hence, typical C-F couplings, as those shown in Figure 4.2.9., together with 2D HSQC studies can help the correct assignment of the molecule.





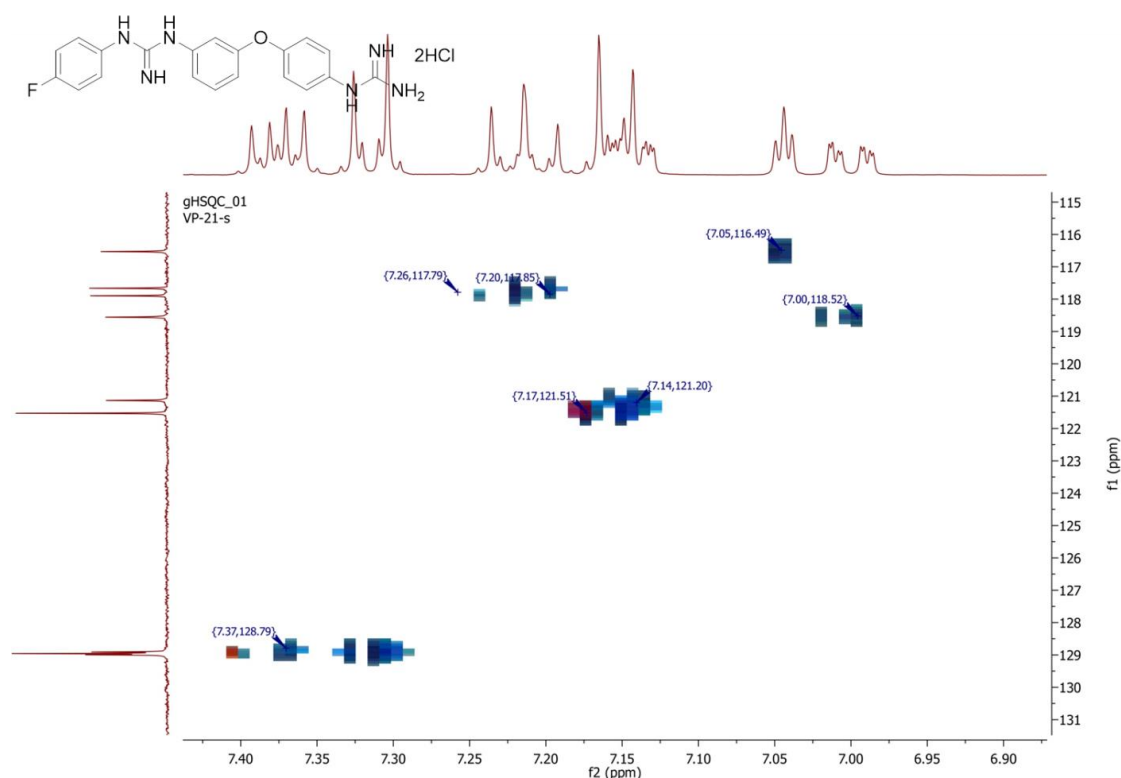


Figure 4.2.9. Details of ^1H NMR, ^{13}C NMR, ^{19}F NMR and HSQC spectra for compound 2.

4.3. Synthesis of 3-amino- or 3-amido-4'-guanidino phenyloxybenzene derivatives: exploring the link

4.3.1. General aspects

From the previously mentioned molecular modelling studies anticipated in the Objectives section, it was evident that these molecules were under conformational strain to fulfill both the guanidinium-phosphate and lipophilic interactions. Thus, shorter molecules would relieve this strain and fit better into the binding site of the biological target. For this reason, and in order to provide different structure combinations to be able to deduce a proper structure-activity relationship (SAR), we have prepared two new series of derivatives of lead compound **1**: arylamino **27-31** and arylamido **32** and **33** diphenyl mono-guanidines (Figure 4.3.1.).

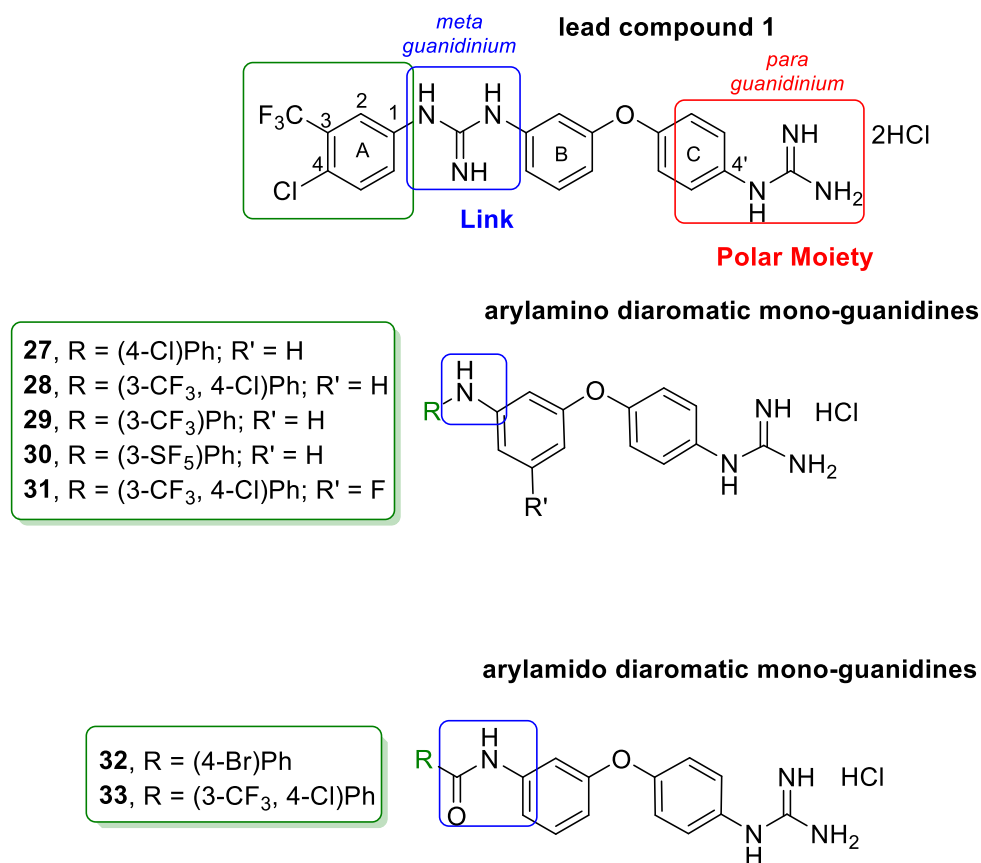


Figure 4.3.1. Lead compound **1** and proposed changing to obtain arylamino diaromatic mono-guanidines **27-31** and arylamido diaromatic mono-guanidines **32** and **33**.

To check the effect of the total length in the interactions with the target, the –NH– and –CONH– links that replaces the guanidinium cation in the arylamino or arylamido diaromatic mono-guanidines, were also combined with different hydrophobic moieties. Thus, a systematic approach was introduced to try to understand the requirement at the hydrophobic moiety. If compounds **28** and **33** explores just changing at the link of the molecule maintaining the (3-CF₃, 4-Cl)Ph moiety in ring A, other derivatives maintain only either the 4-chlorine as in compound **27** or the 3-trifluoromethyl as in compound **29**.

The –SF₅ group has been often referred to as a ‘supertrifluoromethyl group’ due to its similarity with the –CF₃ group.⁴⁷ The two groups share very similar features such as high electronegativity, steric bulk, thermal and chemical stability and lipophilicity, with the SF₅-group slightly prevailing over the CF₃ in all of these aspects.⁴⁷ Different studies, however, have shown how they are not biologically equivalent and SF₅ peculiar

characteristics make it very interesting for drug discovery/development applications.⁴⁷⁻⁴⁹ Hence, we aimed to replace the $-\text{CF}_3$ moiety in compound **29** with a $-\text{SF}_5$ substituent as in compound **30**, in order to probe perturbations of the electron density on the aryl ring and help to investigate the size of the hydrophobic pocket of the target.

Quite often, fluorine is used in medicinal chemistry and many marketed compounds contain this halogen.⁵⁰ The introduction of fluorine into a molecule can influence conformation, pK_a , intrinsic potency, membrane permeability, metabolic pathways and pharmacokinetic properties.⁵¹ Moreover, theoretical and experimental studies have presented several examples of what it is known as fluorine bonding, which has importance in protein-ligand interaction.⁵² There are different examples in which the introduction of a fluorine substituent leads to a slight enhancement of the binding affinity due to an increased lipophilicity of the molecule⁵⁰ and among these examples are some MEK inhibitors.⁵³ With these considerations in mind we synthesized compound **31**.

4.3.2. Synthesis

4.3.2.1. Synthesis of arylamino diaromatic mono-guanidines

The retrosynthetic approach for the synthesis of arylamino diaromatic mono-guanidines **27-30** is reported in Figure 4.3.2.

Starting diphenyl amine **34** was prepared by reaction of 3-aminophenol **35** and 1-fluoro-4-nitrobenzene **36** via a nucleophilic aromatic substitution ($\text{S}_{\text{N}}\text{Ar}$). Selective Buchwald-Hartwig reaction between diphenyl amine **34** and different bromine sources, allowed to obtain nitro derivatives **37-40** which were subjected to nitro reduction to yield the free amino derivatives. Next, these compounds were guanidylated using *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methylisothiourea and followed by Boc deprotection which afforded the aimed compounds **27-30**.

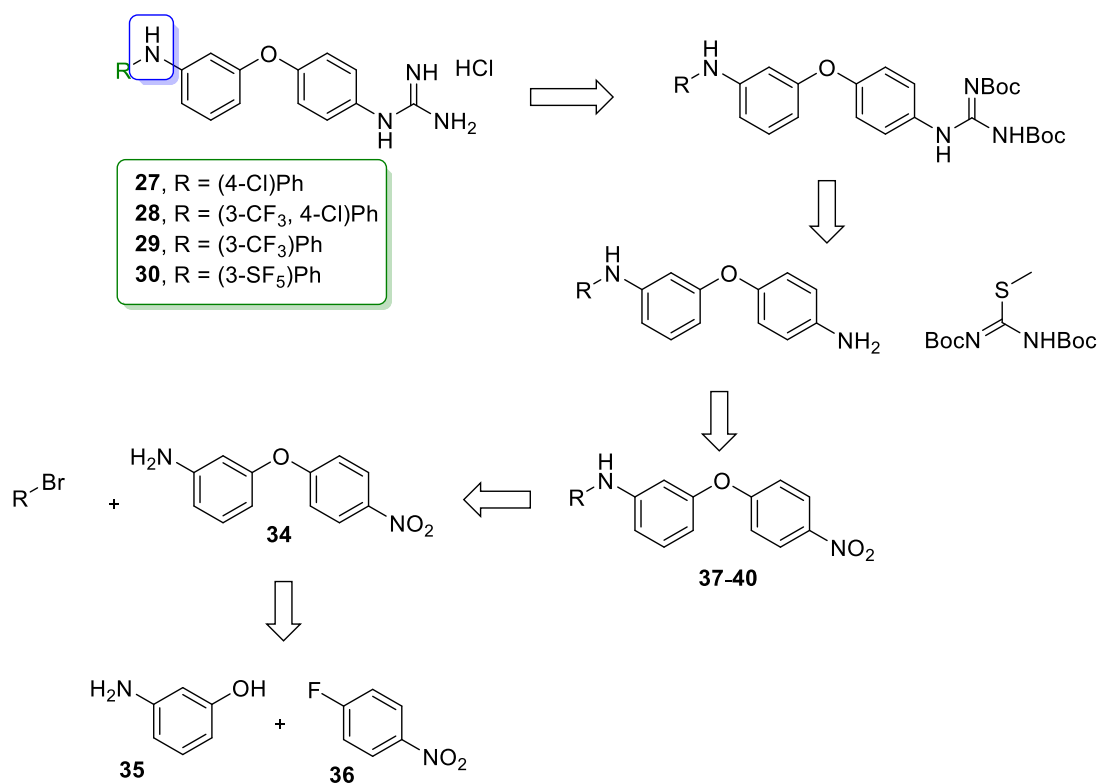
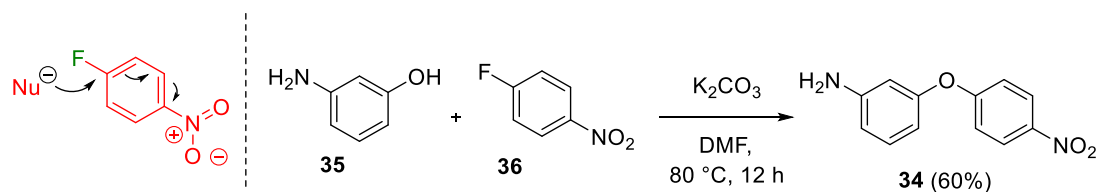


Figure 4.3.2. Retrosynthesis of arylamino diaromatic mono-guanidines **27-30**.

As mentioned, compound **34** was synthesised by a S_NAr reaction between commercially available 3-aminophenol **35** and the commercially available halo-aryl derivative 1-fluoro-4-nitrobenzene **36** (Scheme 4.3.1.).⁵⁴

Scheme 4.3.1.



An anion-stabilizing (electron-withdrawing) group, such as fluorine, *ortho* or *para* to a potential leaving group, as the nitro group in **36**, can be used to facilitate nucleophilic aromatic substitution on an otherwise not reactive benzene ring (Scheme 4.3.1.). The C-

F bond is the strongest of all the single bonds to carbon and it is difficult to break; despite this, fluorine is the halogen of choice for nucleophilic aromatic substitutions because it is the most electronegative of the elements and it stabilizes the anionic intermediate, assisting the acceptance of the electron by the benzene ring.⁵⁵

Regarding the synthesis of nitro derivatives **37-40**, even though their preparation is not accessible through a simple S_NAr substitution, the C-N bond formation can be achieved by the use of cross-coupling reactions. Palladium-catalysed coupling chemistry has gained widespread use in industrial and academic synthetic chemistry as a powerful methodology for the formation of C-C and C-heteroatom bonds. From a general background mechanism, several coupling reactions have been developed between general aryl-halides or triflates with different substrates: Suzuki-Miyaura (organoboron), Stille (organostannanes), Negishi (organo-zinc), Kumada (Grignard reagents), Hiyama (organosilanes), Sonogashira (aryl- and alkenyl-alkynes), Heck (olefins), Buchwald-Hartwig (amine nucleophiles) (Figure 4.3.3.).⁵⁶

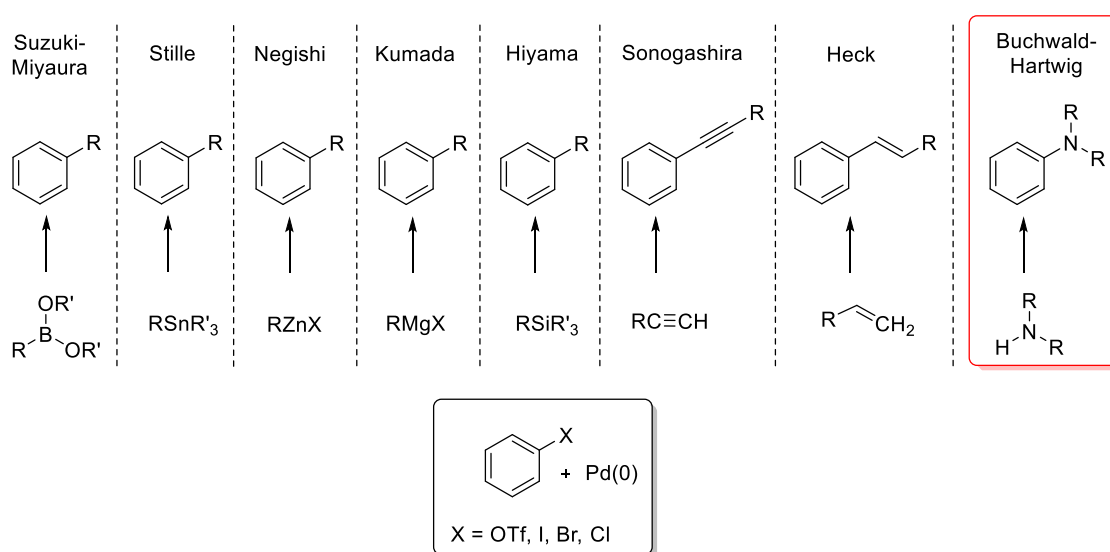
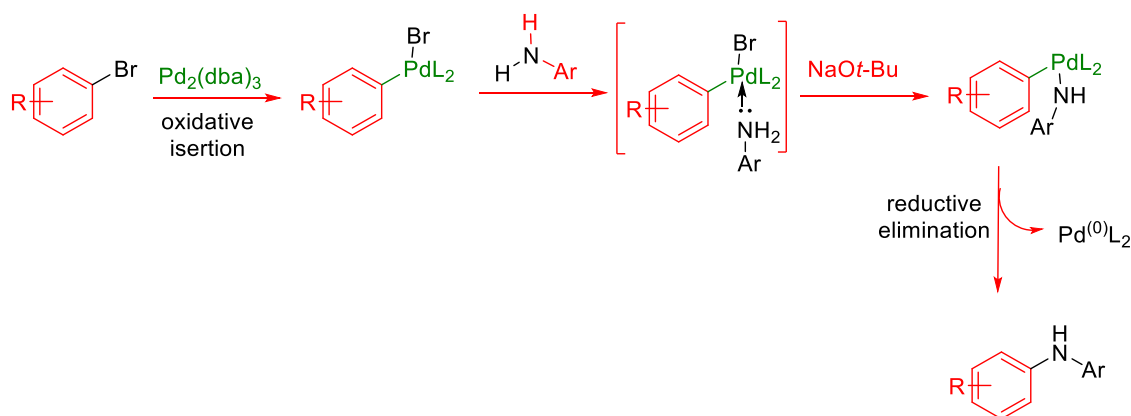


Figure 4.3.3. Most commonly utilized palladium-catalysed cross-coupling reactions.

Compounds **37-40** were synthesized choosing a Buchwald-Hartwig cross-coupling to afford the distinguishing C-N bond. The general reaction involves a Pd-catalyzed C-N and C-O bond formation between aryl halides (I, Br, Cl) or trifluoromethanesulfonates

(OTf) and amines (aromatic or aliphatic) or between aryl halides or triflates and alcohols in the presence of a stoichiometric amount of base (Scheme 4.3.2.).⁵⁷

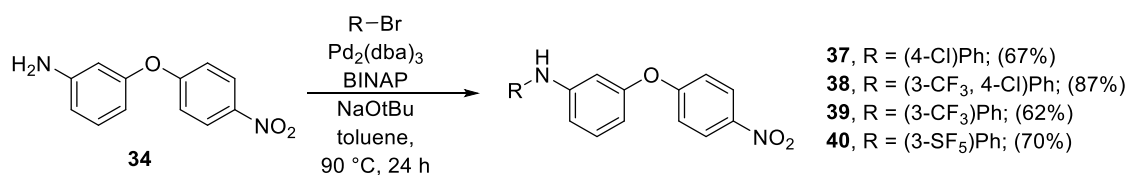
Scheme 4.3.2.



The first step is the oxidative insertion of Pd(0) into the aryl-halogen bond. The Pd(II) complex now adds the amine so that both coupling reagents find themselves bonded to the same palladium atom. The base eliminates the halogen from the complex and reductive elimination forms the Ar-N bond (Scheme 4.3.2.).

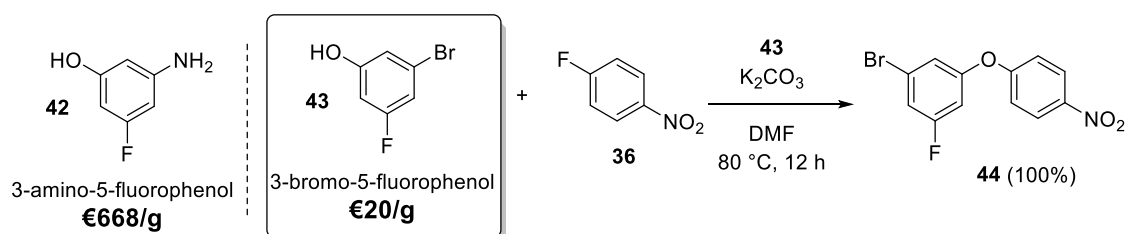
The success of this reaction often depends on different variables and parameters such as ligand, Pd source, base and solvent.⁵⁸ Fortunately, in our particular case the chosen starting conditions [Pd₂(dba)₃ 3 mol%, BINAP 3 mol%, NaOtBu 1.4 eq. in dry toluene (2 mL mmol⁻¹) at 90 °C) afforded aimed compounds **37-40** in high yield (62-87%), Scheme 4.3.3.

Scheme 4.3.3.



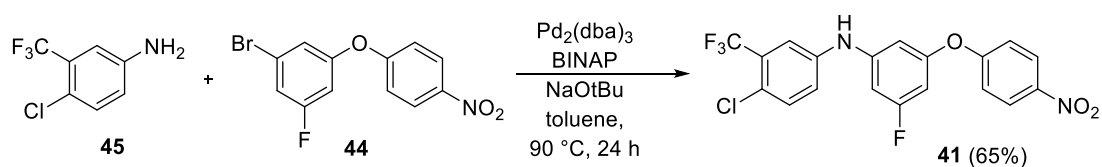
Preparation of compound **41** slightly differs from the previous synthetic scheme. Given the limiting price of the starting material 3-amino-5-fluorophenol **42** (€668 per g), we chose to use the cheaper 3-bromo-5-fluorophenol **43**, which in K_2CO_3 and DMF reacted with 4-fluoronitrobenzene **36**, by means of the previously described S_NAr , to afford compound **44** in quantitative yield (Scheme 4.3.4.).

Scheme 4.3.4.



Compound **44**, carrying the bromine source, is then used under the same conditions for the previously described Buchwald-Hartwig coupling with the commercially available 4-chloro-3-trifluoromethylamine **45** to afford compound **41** in good yield (Scheme 4.3.5.).

Scheme 4.3.5.

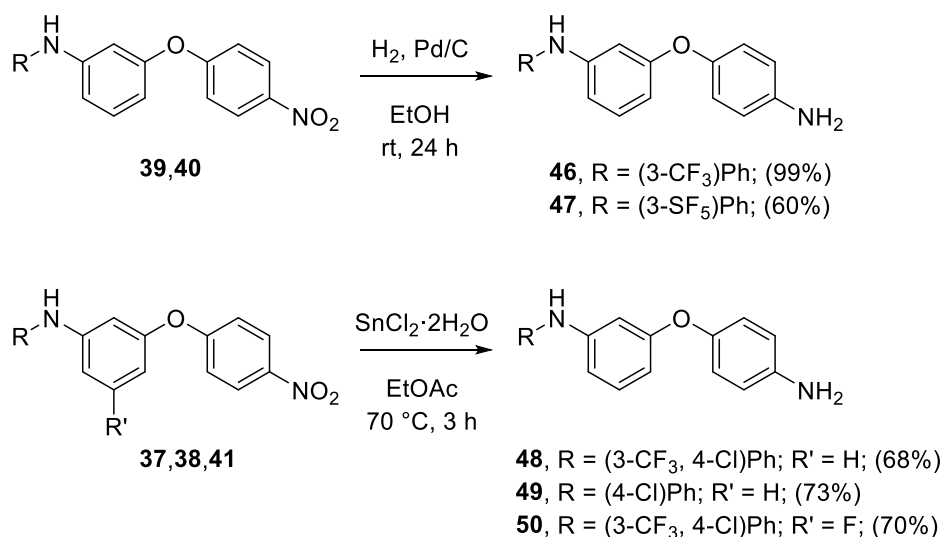


All nitro precursors **37-41** were then subjected to selective reduction of the nitro group to the amine that will serve as handle for the introduction of the guanidine moiety.

Organic halides are very important and versatile compounds with many applications as starting materials, solvents, reagents, and intermediates in synthetic organic chemistry.⁵⁹ It is well known that organic halides can suffer hydrodehalogenation under different conditions such as catalytic hydrogenation, reduction with metals and reduction with metal hydrides.⁵⁹ The ease of reductive hydrodehalogenation follows the general order $I > Br > Cl \gg F$ and it is related to the dissociation energy of carbon-halogen bonds (C-I, 53 kcal mol⁻¹; C-Br, 67 kcal mol⁻¹; C-Cl, 81 kcal mol⁻¹; C-F, 109 kcal mol⁻¹). Due to the strength of the C-F bond, fluorides are rather resistant to reduction.

For these reasons, nitro reduction of compounds **39** and **40** was achieved using catalytic hydrogenation with H₂ and Pd on carbon (10 mol%) to obtain compounds **46** and **47**; however, in the case of halo-compounds **37**, **38** and **41** selective reduction was achieved with the use of tin(II) chloride dihydrate (SnCl₂·2H₂O) to obtain compounds **48-50** (Scheme 4.3.6.).⁶⁰

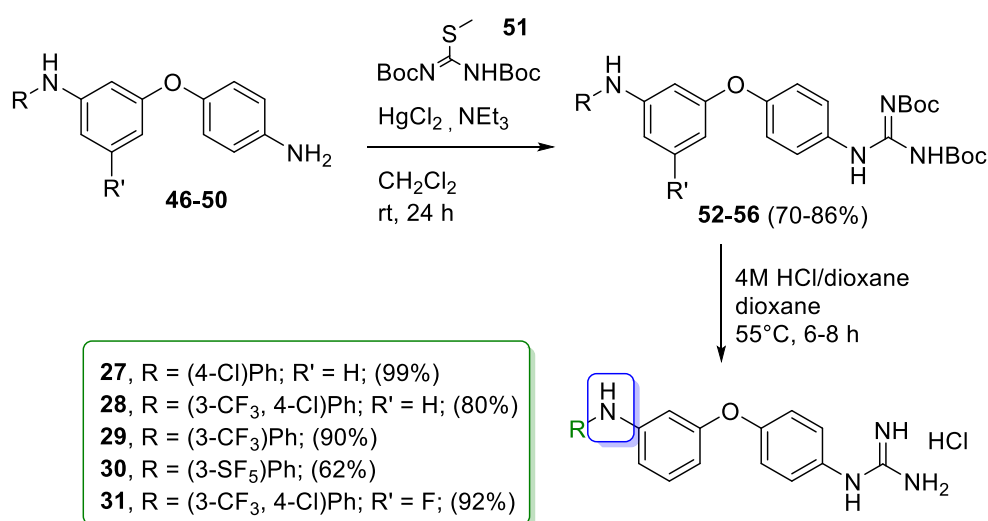
Scheme 4.3.6.



Although the catalytic hydrogenation gave overall better yields (60% and 99%), the reduction performed using tin(II) chloride achieved acceptable yields (68-73%) with a lower reaction time of three hours, compared with the 24 h for the catalytic hydrogenation.

Utilising previously described conditions, the guanidine moiety was introduced in compounds **46-50** using *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methyliothiourea **51**, mercury(II) chloride and trimethylamine to afford Boc-protected arylamino diaromatic mono-guanidines **52-56**. Usual deprotection with 4M HCl/dioxane at 55 °C for 6-8 h afforded compounds **27-31** in excellent yields (Scheme 4.3.7.).

Scheme 4.3.7.



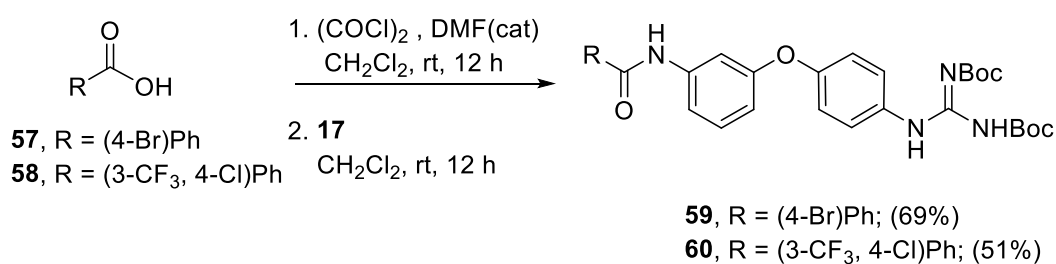
4.3.2.2. Synthesis of arylamido diaromatic mono-guanidines

Next, we considered the amido functionality as a connector because not only will allow to vary the total length of these potential kinase inhibitors, but also could fix the orientation of the lipophilic group in a different position compared to the one in compound **1** due to the strict conformation of the CO-NH *trans* arrangement.

Synthesis of compounds **32** and **33** required the conversion of benzoic acids **57** and **58** to the corresponding acid chlorides by reaction with oxalyl chloride in the presence of a

catalyst (DMF).⁶¹ Once the DMF and any remaining oxalyl chloride were removed, previously synthesized Boc-protected aniline **17** was added to obtain Boc-protected arylamido mono-guanidine derivatives **59** and **60** in good yields (Scheme 4.3.8.).

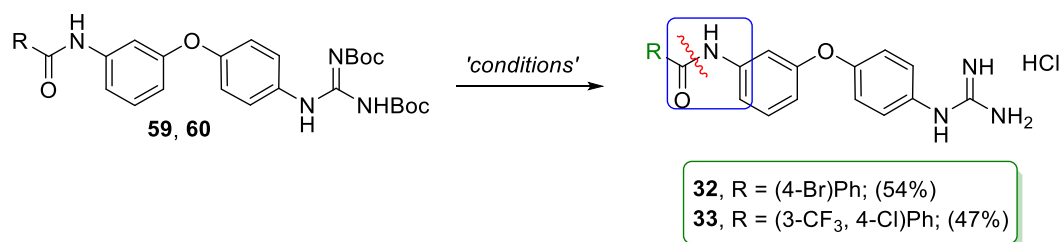
Scheme 4.3.8.



Boc-deprotection of compounds **59** and **60** to obtain final hydrochloride salts **32** and **33** following the procedure typically utilised by Rozas' group and described in the previous sections (HCl 4M solutions in dioxane, 0.2 M final concentration), lead to the hydrolysis of the amide bonds (as observed by TLC and confirmed by ¹H NMR and mass spectroscopy) resulting in the broken salt (Table 4.3.1, entry 1).

Although heat is often required to cleave amide bonds because they are the least reactive among the carboxylic acid derivatives, we believe that the amide link present in these compounds is particularly labile due to the electron withdrawing nature of the Ph-O-Ph oxygen atom and the electronegative halogen substituents present on the adjacent phenyl ring. These two elements lead to a decrease of the electron donating capabilities of the nitrogen in the amide bond, leaving the CO carbon atom more open to nucleophilic attack.

With such a labile amide bond in these molecules, alternative strategies were required. First, the experiment was repeated decreasing the acid concentration until reaching a solution ten times more dilute than the first one (0.02M final concentration); however, this did not improve the outcomes of the reaction (Table 4.3.1., entry 2). Dry conditions (Table 4.3.1., entry 3) and deprotection with 50:50 TFA/CH₂Cl₂ were also tried, with no success (Table 4.3.1., entry 4).^{62,63}

Table 4.3.1.

Entry	'conditions'	Cleavage	Yield (%)
1	HCl/dioxane, 0.2 M 55 °C, 6-8 h	yes	0
2	HCl/dioxane, 0.02 M 55 °C, 6-8 h	yes	0
3	HCl/dry dioxane , 0.2 M 55 °C, 6-8 h	yes	0
4	TFA/CH ₂ Cl ₂ 50:50 rt, 6-8 h	yes	0
5	SnCl ₄ , H ₂ O EtOAc, rt , 1 h	no	54%, 47%

In the search for an alternative procedure, a method described in the literature⁶⁴ utilising tin(IV) chloride (SnCl₄) was tested (Table 4.3.1, entry 5). Thus, compounds **59** and **60** were dissolved in EtOAc and then SnCl₄ was added. When the reaction was considered finished upon observation by TLC, EtOAc and SnCl₄ were removed under vacuum and the work-up was carried out with H₂O and CH₂Cl₂. The crude product of this reaction was analysed using mass spectroscopy and ¹H NMR and only the desired salts **32** and **33** were found.

This unusual Boc deprotection with SnCl₄ has also recently been used in the mild deprotection of Boc protected amino acids by Freeman and Gilon.⁶⁵ Under neutral conditions and using four equivalents of SnCl₄, removal of the two Boc groups was quickly and simultaneously achieved at room temperature with good yields and without isolation of any *mono*-Boc guanidine intermediate or cleaved final salts.

4.4. Synthesis of derivatives with modified cationic moieties: exploring the polar moiety

4.4.1. General aspects

As previously mentioned in the Objectives section of this thesis (Chapter 2), the introduction of different cations instead of the guanidinium at the 4' position of lead compound **1** (polar moiety), was considered in order to improve the interactions with the target and, hence, the activity (Figure 4.4.1.). Accordingly we proposed the synthesis of:

Pyridinoguanidines: (i) 3,4'-Bis-guanidino phenyloxypyridines, **61-65**, that carry a pyridinoguanidine instead of a phenylguanidine and in combination with different hydrophobic moieties, but maintaining a guanidine at position 3 as a link (Figure 4.4.1., a).

(ii) 3-Amino-4'-guanidino phenyloxypyridine **66** that carries the same pyridinoguanidine moiety as in compounds **61-65** and the hydrophobic moiety of the lead compound **1** [(3-CF₃,4-Cl)Ph], but with an amino group as link instead of the guanidine (Figure 4.4.1., b).

Phenylisouroniums: (i) 3-Guanidino-4'-O-isourea phenyloxybenzene derivative **67**, which maintains all the features of lead compound **1**, but carries an isouronium moiety instead of the guanidinium at the 4' position (Figure 4.4.1., c).

(ii) 3-Amino-4'-O-isourea phenyloxybenzene **68**, which maintains all the features of compound **28**, but carries an isouronium moiety instead of the guanidinium at the 4' position (Figure 4.4.1., d).

(iii) 3-Amido-4'-O-isourea phenyloxybenzene **69** that maintains all the features of compound **33**, but carries an isouronium moiety instead of the guanidinium at the 4' position (Figure 4.4.1., e).

Phenylsulfamides: 3-Guanidino-4'-sulfonamide phenyloxybenzene **70**, which maintains all the features of compound **28**, but carries a sulfamide moiety instead of the guanidine at the 4' position (Figure 4.4.1., f).

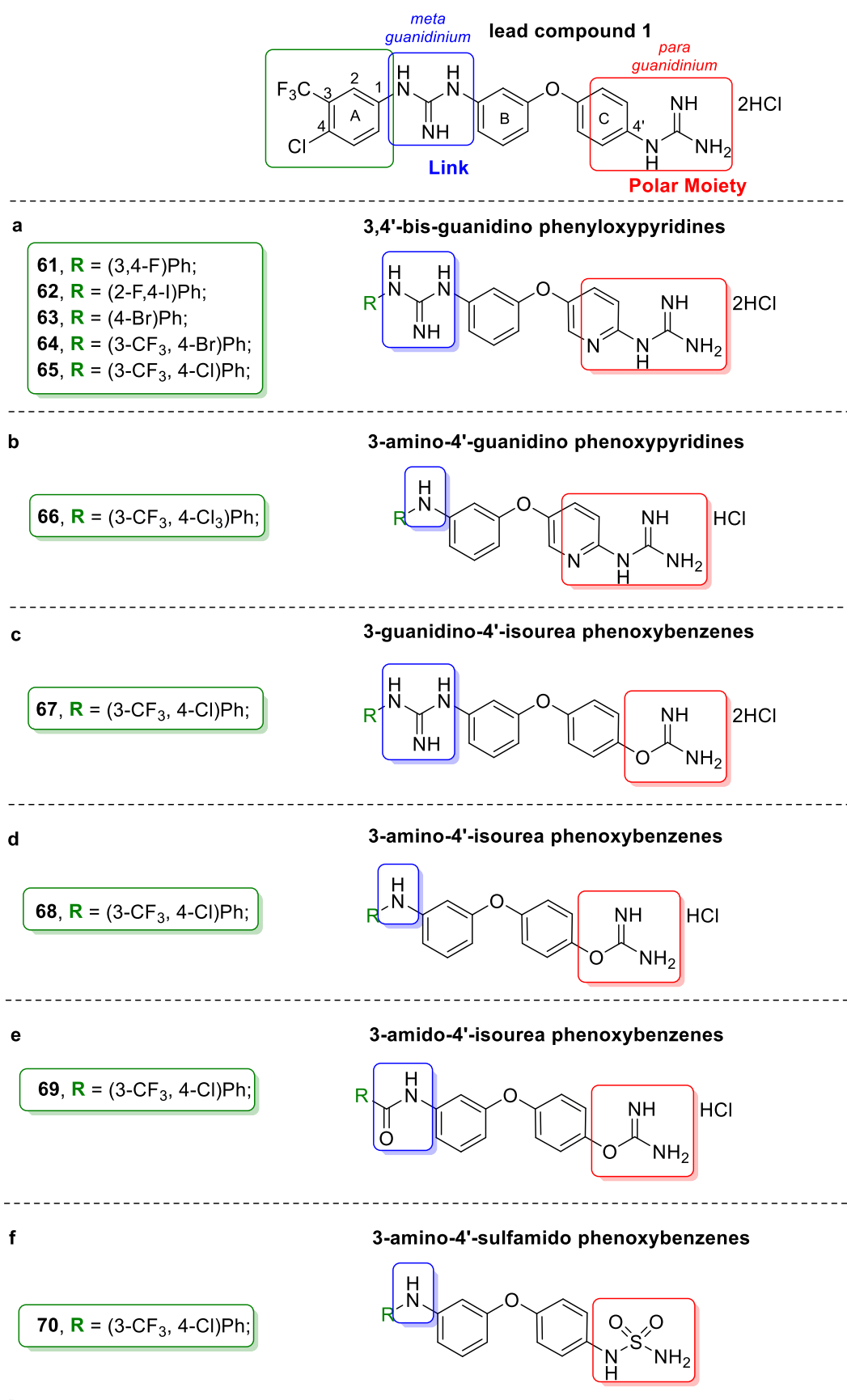


Figure 4.4.1. Lead compound **1** and proposed changes of the polar moiety.

4.4.2. Synthesis

4.4.2.1. Synthesis of 3,4'-bis-guanidino and 3-amino-4'-guanidino phenyloxypyridines

Previous studies conducted in our group on 2-pyridinoguanidines⁴⁶ showed how, considering the two possible *anti* and *syn* conformers, the Boc-protected derivatives exhibit a preference for the *anti* conformation while the 2-pyridinoguanidinium salts exist primarily in the *syn* conformation due to different intramolecular hydrogen bonds (IMHB) (Figure 4.4.2.). Accordingly, we propose that as a modification of the polar moiety of lead compound **1**, the introduction of a pyridine ring would change the orientation of the adjacent guanidinium and this will give indications of the best orientation of such group when interacting with the biological target.

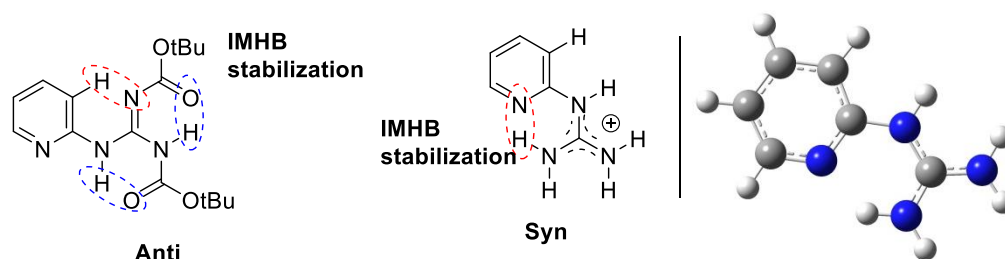
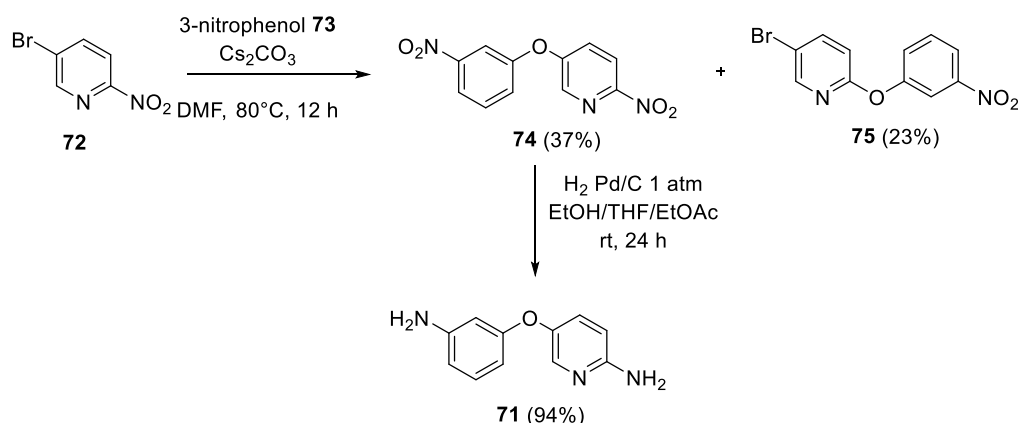


Figure 4.4.2. *Left:* anti and syn conformations for Boc-protected and Boc-deprotected 2-pyridinoguanidines, *Right:* Structure of 2-pyridinoguanidinium calculated at the B3LYP/6-31G* level of computation.

Accordingly, in this section we describe the preparation of 3,4'-bis-guanidino (**61-65**) and 3-amino-4'-guanidino (**66**) phenyloxypyridine derivatives to modify the polar moiety of the lead compound **1**.

In order to prepare the 3,4'-bis-guanidino phenyloxypyridines, the starting 5-(3-aminophenoxy)pyridin-2-amine **71** was synthesised by S_NAr between 5-bromo-2-nitropyridine **72** and 3-nitrophenol **73** to obtain intermediate **74**. Next, the hydrogenation of **74** gave the desired product **71** in good yield (Scheme 4.4.1.).^{54,66} Compound **75** was obtained together with **74** by S_NAr at the position 2 of the pyridine, as already described in the literature for similar starting material.⁶⁷

Scheme 4.4.1.



The N atom of the pyridine ring influences the chemical properties of this heterocycle to accommodate $\text{S}_{\text{N}}\text{Ar}$. Pyridine has a low electron density at position 2 and 4 of its ring, as seen from its resonance forms (Figure 4.4.3.); thus if these electrophilic positions are substituted with a potential leaving group, the pyridine ring becomes more active towards $\text{S}_{\text{N}}\text{Ar}$.

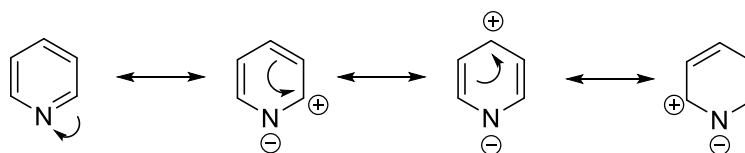
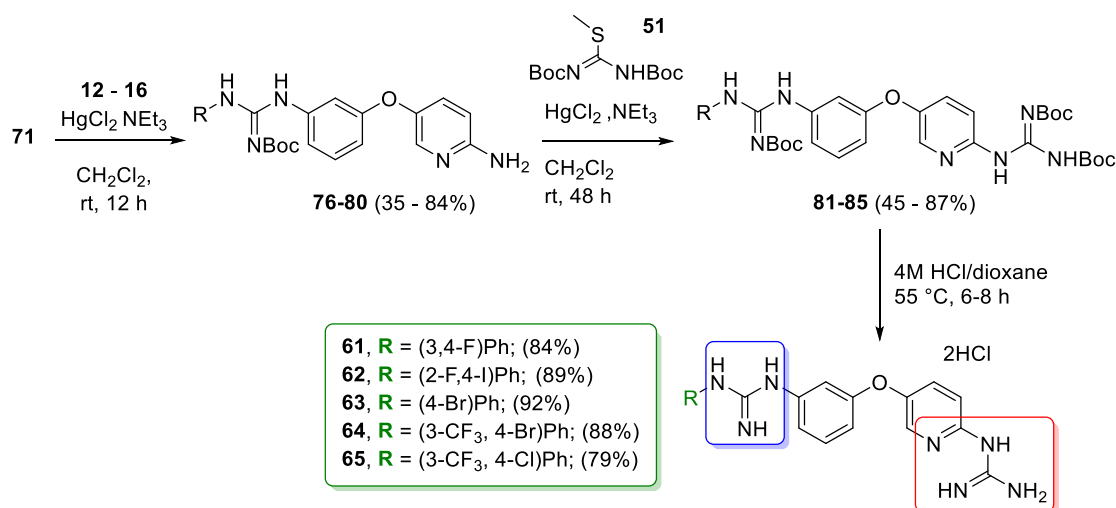


Figure 4.4.3. Resonance forms of pyridine.

Compound **71** was then reacted with *N*-aryl-*N'*-Boc-protected thioureas **12-16** following conditions previously described and in the presence of mercury(II) chloride and trimethylamine to obtain a substitution at the *meta* position in compounds **76-80**.¹³ Subsequent guanidinylation at position 2 of the pyridine ring to obtain compounds **81-85** was carried out with commercial *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methylisothiurea **51**. As this last step requires reaction times of over 48 hours and often a second addition of HgCl_2 , due to the low reactivity of this position,⁵⁵ the overall yield obtained in this process is low.

Removal of Boc protecting groups was achieved following the usual procedure that involves stirring of the appropriate Boc-protected compound in 4M HCl/dioxane (6 eq. per Boc group) at 55 °C for 6-8 h, which afforded the final hydrochloride salts **61-65** in excellent yields (Scheme 4.4.2.).

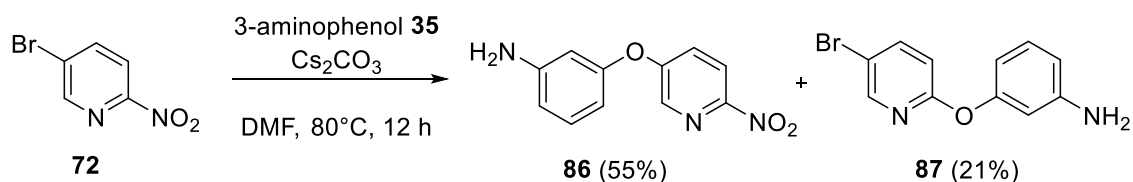
Scheme 4.4.2.



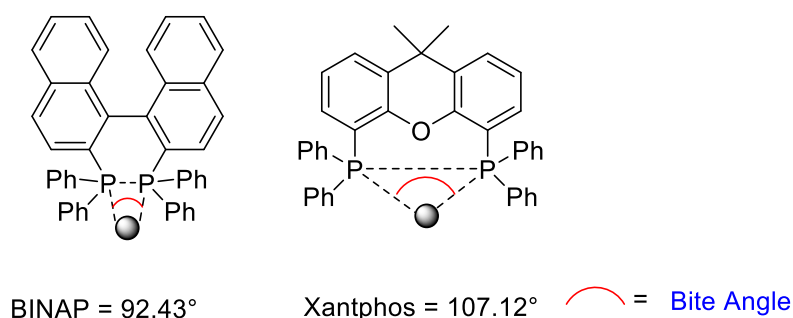
The isomerism previously reported for the Boc-protected structures of compounds **20-26** was also present in compounds **76-80** and **81-85**; hence the ^1H NMR and ^{13}C NMR spectra were also highly amorphous, making the assignment of the signals difficult. As previously presented in section 4.2.2, removal of Boc group gives much more defined spectra, which were then used for characterisation of the final compounds.

Similar to the synthesis of compound **28**, the preparation for compound **66** required the synthesis of a properly functionalised amine (through a $\text{S}_{\text{N}}\text{Ar}$) to access intermediate **86** starting from 5-bromo-2-nitropyridine **72** and 3-aminophenol **35**. This reaction also yielded the by-product of a $\text{S}_{\text{N}}\text{Ar}$ at position 2 of the pyridine, compound **87**, as similarly described for compound **74** (Scheme 4.4.3.).

Scheme 4.4.3.



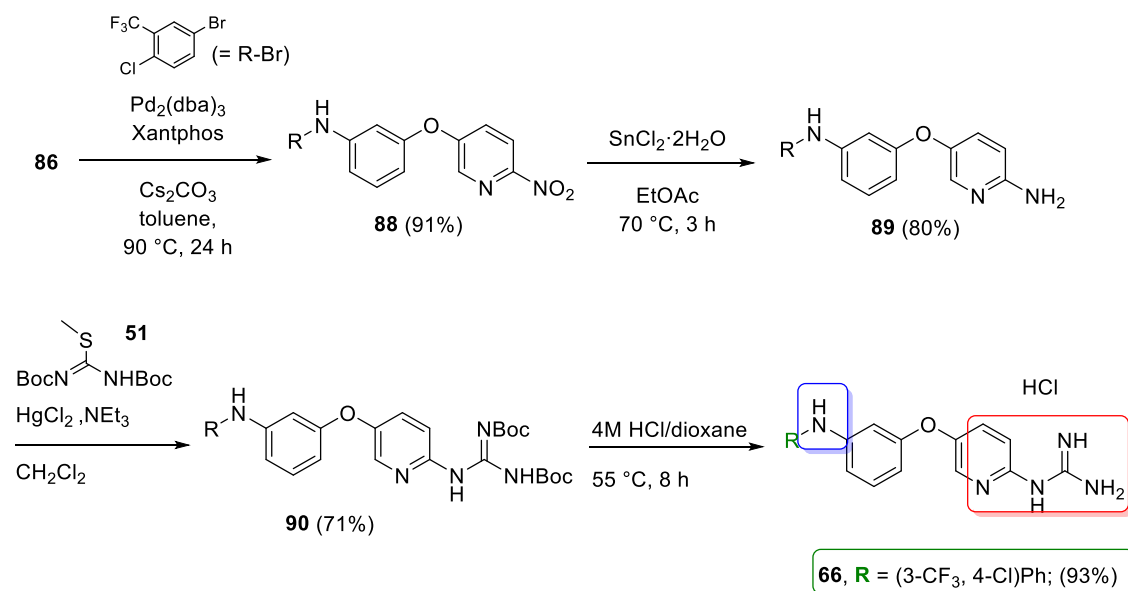
Although previous reaction conditions [$\text{Pd}_2(\text{dba})_3$ 3 mol%, BINAP 3 mol%, NaOtBu 1.4 eq. in dry toluene (2 mL mmol^{-1}) at 90°C] described for the synthesis of compounds **37-41** gave high yields, in the case of the pyridine analogue **88** the yield was much lower. This is probably due to the ability of the N atom of the pyridine ring to interact with the catalyst metal centre.^{68,69} In our case, utilisation of a bulky bidentate ligand such as Xantphos could prevent interactions between the pyridine N atom and the metal centre. The use of Xantphos has been applied for a variety of Pd catalysed transformations⁷⁰⁻⁷² given its ability to increase the electron density on the metal center (Pd) due to its electron rich phosphine substituents and by its ability to align the appropriate molecular orbitals of Pd for the oxidative step insertion into the aryl halide. Moreover, a correlation between the different bite angles in diphosphine ligands and reaction's selectivity and efficacy has been observed.⁷³ Xantphos large bite angle (107.1° vs. 92.4° of BINAP⁷³) facilitates also the reductive elimination step (Figure 4.4.4.).⁷⁴

Figure 4.4.4. Diphosphine ligands BINAP and Xantphos' bite angles.⁷³

In Pd-catalysed reaction coupling of amines with aryl halides, the use of NaOtBu in toluene as base remains the most versatile for this type of amination reactions with phosphine ligands. Unfortunately, because NaOtBu is a relatively strong base ($pK_a = 17.0$) it can participate in undesirable side reactions with various electrophilic functional groups and some aromatic heterocycles.⁵⁸ The scope of different bases under different reaction conditions has been the subject of different studies⁷⁵ and it is now consolidate that weak inorganic bases (i.e. Cs_2CO_3 , K_3PO_4 or K_2CO_3) instead of alkoxides, can provide more general conditions for substrates containing electrophilic functional groups such as ketones, esters and nitro aromatics.^{58,76}

After these considerations, the use of Xantphos and Cs_2CO_3 has expanded the scope of the reaction and in that way compound **88** was then achieved in good yield. Previously described $\text{SnCl}_2 \cdot 2\text{H}_2\text{O}$ reduction conditions afforded compound **89** that was then subjected to guanidylation using *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methyliothiourea **51** to obtain compound **90** that was Boc-deprotected under usual conditions to give final compound **66** in 93% yield (Scheme 4.4.4.).

Scheme 4.4.4.

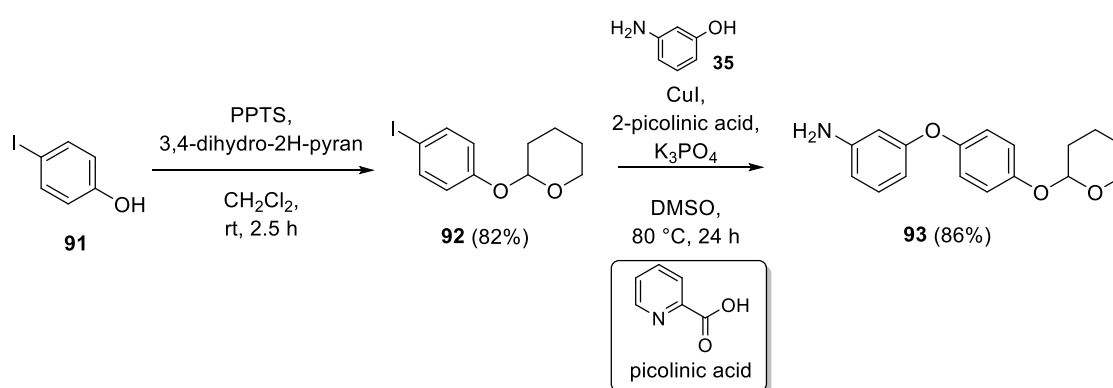


4.4.2.2. Synthesis of isouronium derivatives

Guanidine's high basicity (pK_{aH} in water is 13.6³) indicates that at physiological pH, it will be protonated forming the guanidinium cation. For this reason, guanidinium derivatives are sometimes discounted as useful drugs because of potential bioavailability problems, despite this functional group being present in many current therapies, as described in section 4.1.1. Regardless of its high basicity, the guanidinium cation has important features very useful for target binding such as forming strong HBs and ionic interactions.⁷⁷⁻⁷⁹ Interestingly, the related isouronium cation keeps these two binding features and, moreover, is less basic ($pK_{aH} = 9.8^3$); hence, this cation is very promising as a functional substitute of guanidinium derivatives with therapeutic potential.⁸⁰

Accordingly we have prepared a number of isouronium analogues of the guanidinium derivatives already synthesised. In the case of the 3-guanidine-4'-isourea phenyloxybenzene derivative **67**, its synthesis required the preliminary preparation of a conveniently functionalised phenol which was introduced by the initial protection of iodophenol **91** with pyridinium *p*-toluenesulfonate (PPTS) and 3,4-dihydro-2H-pyran in dry CH_2Cl_2 to yield compound **92** (Scheme 4.4.5).⁸¹ PPTS is the most common catalyst for the tetrahydropyranylation of alcohols, and seems to be superior to other catalysts such as hydrochloric acid, phosphoryl chloride, and boron trifluoride etherate.⁸²

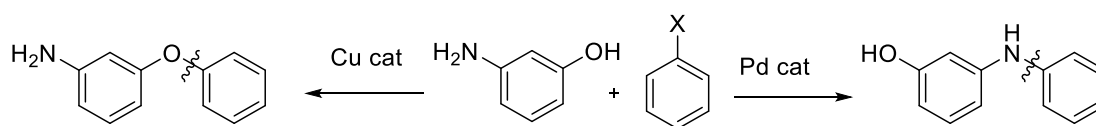
Scheme 4.4.5.



Compound **93** was accessed through a transition metal-catalysed carbon-heteroatom bond-forming reaction between the 3-aminophenol **35** and the halogen substrate **92**, as the previously described Buchwald-Hartwig coupling in section 4.3.2.1. Because of the

difference in acidity of PhOH ($pK_a \approx 18$ in DMSO) and PhNH₂ ($pK_a \approx 31$ in DMSO), either *O*- or *N*-arylation will be easily accessible upon variation of the reaction conditions. In the particular case of Pd catalysed reactions, the *N*-arylated product is obtained, whereas the Cu catalysed ones afford the *O*-arylated product (Scheme 4.4.6.).⁸³

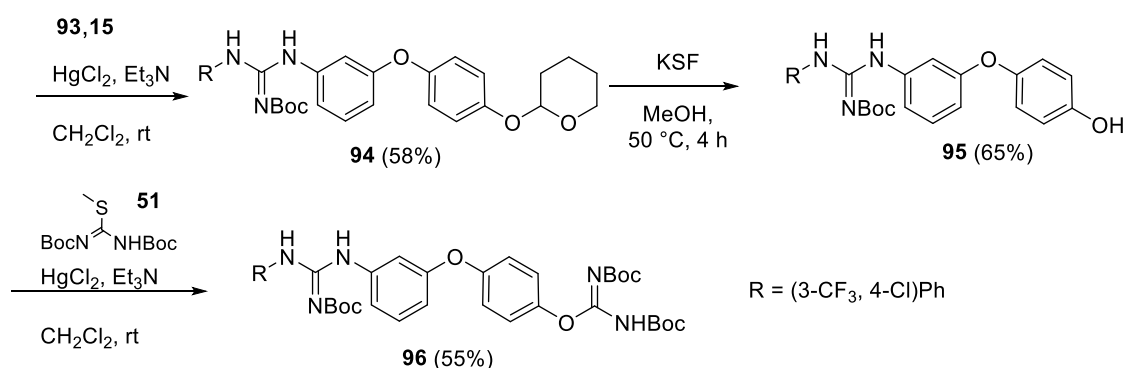
Scheme 4.4.6. Cu- and Pd-catalysed *O*- and *N*-arylation of aminophenol.⁸³



Therefore, in a Ullmann type of coupling, iodine source **92** was coupled with commercially available 3-aminophenol **35**, with copper (I) iodide as copper catalyst, picolinic acid and K₃PO₄ in dimethyl sulphoxide at 80 °C to obtain compound **93** (Scheme 4.4.5.).⁸³

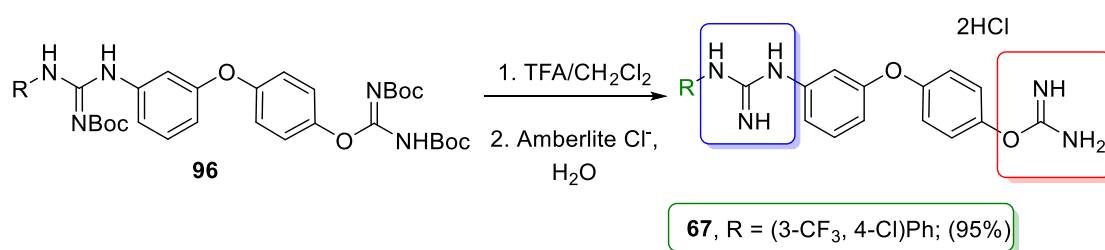
Following the procedure already described by our group¹³ and previously reported in section 4.1.2.1., coupling **93** with the previously synthesized thiourea derivative **15**, in presence of mercury(II) chloride and triethylamine, afforded **94** in good yield (58%). Compound **94** was then deprotected with montmorillonite KSF to obtain the corresponding phenol **95**. Montmorillonite clays have been extensively used as efficient catalysts for a variety of organic reactions and they are a mild and efficient alternative procedure for the deprotection of THPEs.⁸⁴ The OH free phenol derivative was then amidylated using standard conditions previously described by us,^{20,85} using *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methylisothiurea **51**, mercury(II) chloride and triethylamine to yield the corresponding Boc-protected isouronium **96** (Scheme 4.4.7.).

Scheme 4.4.7.



The Boc-deprotection for isourea **96** was carried out following standard conditions for the deprotection of isouronium derivatives within our group⁸⁵ to yield the final hydrochloride salt **67**. Thus, **96** was treated with TFA in dichloromethane, followed by chloride anion exchange, to afford the hydrochloride salt **67** in an excellent yield (Scheme 4.4.8.).

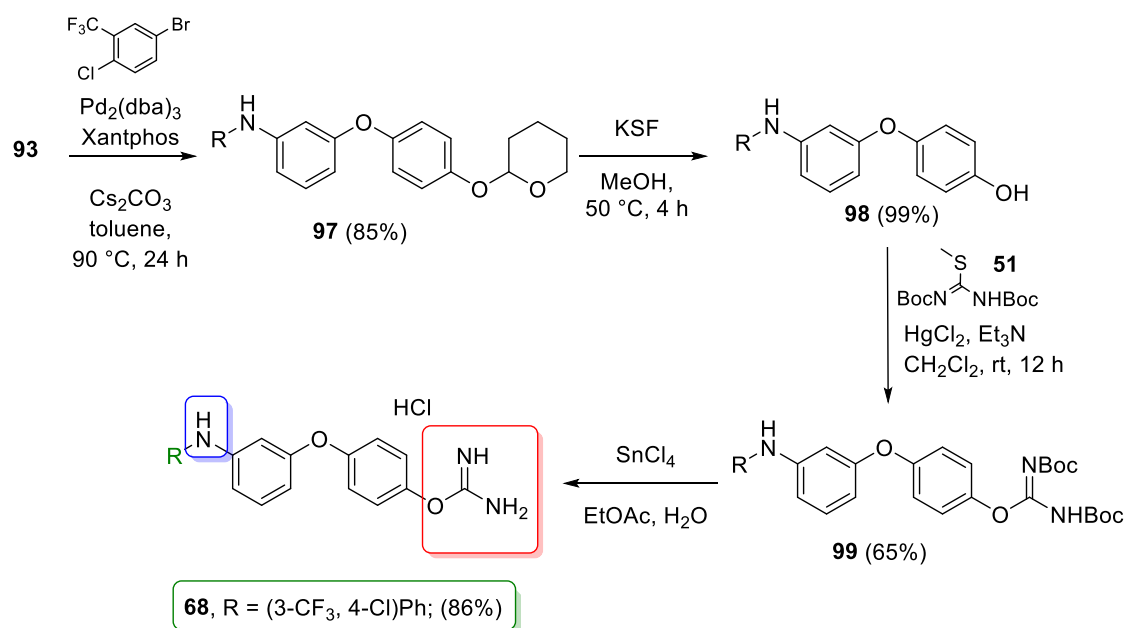
Scheme 4.4.8.



Preparation of the 3-amino-4'-*O*-isourea phenylozybenzene derivative **68** involved the use of the previously synthesised compound **93** as starting material for the synthesis of the intermediate **97** through the previously described conditions for a Buchwald-Hartwig coupling [$\text{Pd}_2(\text{dba})_3$, Xantphos, Cs_2CO_3 , toluene, 90 °C, 24 h]. Compound **97** was then deprotected with montmorillonite KSF to obtain the corresponding phenol **98**. The free phenolic OH of derivative **98** was then amidylated using standard conditions previously described by us using *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methylothiourea **51**, mercury(II) chloride and triethylamine to yield the corresponding Boc-protected isouronium **99**. Based on our experience for the preparation of compounds **32** and **33**,

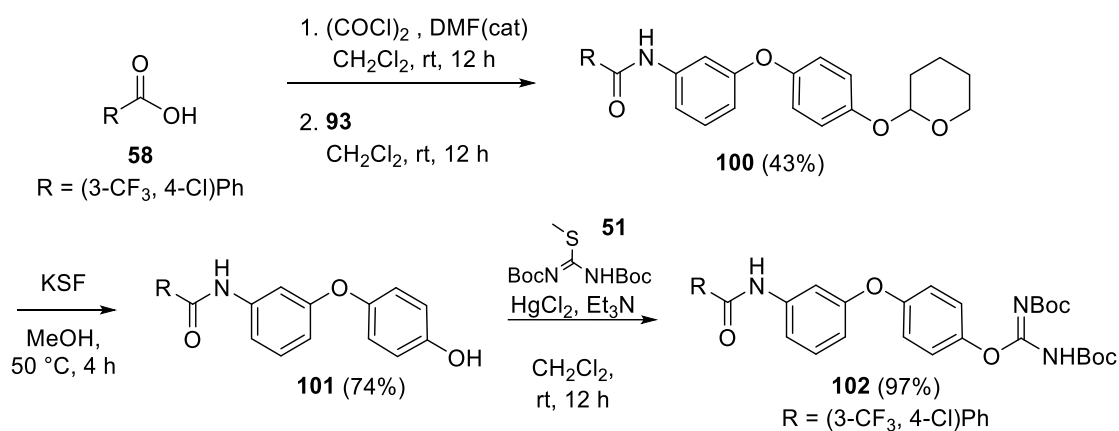
the use of HCl or TFA was not considered appropriate for such liable substrate and, hence, deprotection of **99** was carried out utilising SnCl_4 in EtOAc to yield the final hydrochloride salt **68** as previously described (Scheme 4.4.9.).

Scheme 4.4.9.



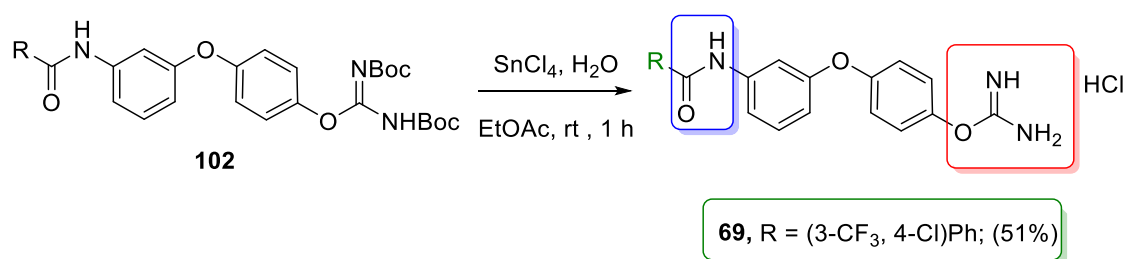
Finally, the synthesis of the 3-amido-4'-*O*-isourea phenylozybenzene derivative **69** required the preparation of the amide-linked phenol precursor, which was achieved with similar conditions as for the synthesis of the 3-amido-4'-guanidino phenylozybenzene derivative **33**. Thus, THP-protected compound **93** was reacted with benzoic acid **58**, properly converted in the acid chloride, to yield compound **100** which was deprotected, as previously described, using montmorillonite KSF to obtain compound **101**. This phenol was subsequently amidylated with *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methylisothiourea **51** to yield compound **102** in a very good yield (Scheme 4.4.10.).

Scheme 4.4.10.



As for the previous syntheses of **32** and **33**, the use of HCl or TFA was not considered appropriate for such liable substrate and, hence, deprotection of **102** was carried out with the use of SnCl_4 in EtOAc to obtain final hydrochloride salt **69** as previously described in section 4.3.2.2. (Scheme 4.4.11.).

Scheme 4.4.11.



4.4.2.3. Synthesis of the sulfamide derivative

Sulfonamides are an important category of pharmaceutical compounds with a broad spectrum of biological activities and they are present in different therapeutic agents such as the ultra-broad spectrum antibiotic doripenem (trade name Doribax),⁸⁶ and acetazolamide, an anti-glaucoma agent.^{87,88} Accordingly, we decided to incorporate this

particular functionality instead of the guanidinium group as a structural modification aiming at investigate the requirements for a better binding.

Hence, compound **70** was synthesised to prove the chemical requirement of the positively charged guanidinium moiety to exert its biochemical activity at the protein binding site. The sulfonamide moiety mimics the guanidinium in terms of geometry, but it differs in terms of position of the HBs donors/acceptors and in the charge (Figure 4.4.5.).

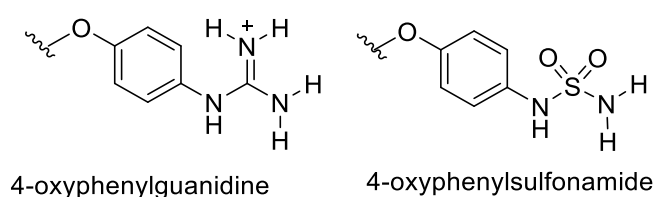
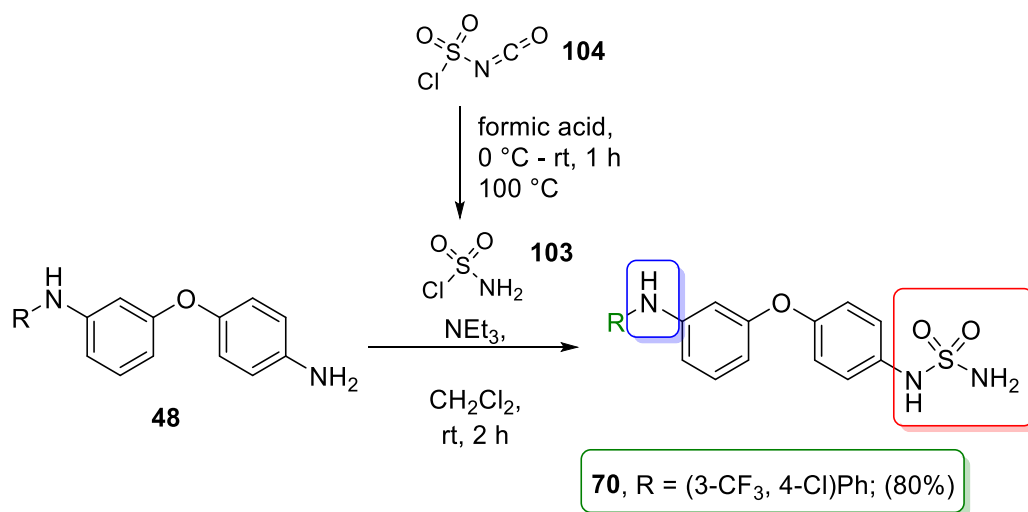


Figure 4.4.5. Structural comparison of guanidinium and sulfonamide moieties.

The most typical method for the preparation of sulfonamides, involves the sulfonation of alcohols and amines, in a basic environment like pyridine or trimethylamine.⁸⁹ Other typical and recent approaches involve their synthesis directly from sulfonic acids, or their sodium salts, under microwave irradiation and through their chloride intermediates.⁹⁰ Direct oxidation of thiol derivatives under combination of $\text{H}_2\text{O}_2/\text{SOCl}_2$ also affords a sulfonyl chloride that yields the corresponding sulfonamide upon reaction with amines.⁹¹ Hydrolysis of chlorosulfonyl isocyanate to give sulfamoyl chloride is also an alternative often used in the literature for a sulfonamide bond-forming reaction.^{92,93} Formic acid is used in this case as a gentler reagent for the formation of sulfamoyl chloride, given the high reactivity towards water of chlorosulfonyl isocyanate.

Following the last method mentioned, amine **48** was treated with sulfonamide chloride **103**, previously prepared from chlorosulfonyl isocyanate **104**^{92,93} and formic acid, to afford compound **70** in good yield (Scheme 4.4.12.).

Scheme 4.4.12.



The hydrochloride salt of this compound was not prepared because the electronic properties of the sulfonamide moiety do not allow the hosting of a positive charge.

4.5. Conclusions

This chapter presents the synthetic strategies explored in the preparation of novel types of potential PKI structurally related to the 3,4'-bis-guanidino phenoxybenzene lead compound **1**. Based on our previous studies we designed and synthesized new compounds by systematically changing the three different moieties identified aiming to probe the different target binding sites.

In order to prepare 3,4'-bis-guanidino phenoxybenzene derivatives **2-8**, the appropriate Boc-protected amines were synthesised, starting from the commercially available 3-(4-aminophenoxy)aniline **19** and the in-house prepared *N,N'*-di-Boc-protected thiourea **9**, following the Kim and Qian's procedure that has been further developed by Rozas' group involving the use of HgCl_2 and trimethylamine. Substituted thioureas **11-16** carrying the lipophilic moiety, which consists of phenyl rings differently substituted by halogens or trifluoromethyl, were prepared according the well-established procedure developed in Rozas' group and involving NaH /TFAA. Next, these thioureas were

reacted with the corresponding mono-guanidine diaromatic systems **17** and **18** to obtain Boc-protected compounds **20-24**, which were substituted at the 3 position of 3-(4-aminophenoxy)aniline, and Boc-protected compounds **25-26** substituted at the 4' position of 3-(4-aminophenoxy)aniline, for the sake of comparison.

Similarly preparation of 3,4'-bis-guanidino phenyloxypyridines **61-65** involved the synthesis of starting amine **71**, reaction with substituted thioureas **11-16** at its position 3, followed by the guanidylation at position 2 of the pyridine ring of compound **71**.

Then, 3-amino-4'-guanidino phenyloxybenzenes **27-30** and 3-amido-4'-guanidino phenyloxybenzenes **32-33** were synthesised to investigate the combination of changing the lipophilic moiety with a different link instead of the guanidinium at the 3 position of lead compound **1**.

Synthesis of 3-amino-4'-guanidino phenyloxybenzenes **27-30** required first the synthesis of starting aromatic amine **34** to then perform a selective Buchwald-Hartwig reaction to obtain derivatives **37-40**. Alternatively, synthesis of compound **41** required starting with compound **44** carrying the bromine source and prepared by a S_NAr reaction between 3-bromo-5-fluorophenol **43** and 1-fluoro-4-nitrobenzene **36**. Using the same conditions as the previously described Buchwald-Hartwig coupling with the commercially available 4-chloro-3-trifluormethyl-amine **45** afforded named compound **41**. Nitro intermediates **37-41** were subjected to nitro-group reduction to obtain the free amino derivatives **46-50** that were guanidylated using *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methylisothiourea **51** to obtain final Boc-protected compounds **52-56**.

Similar conditions were used for the preparation of compound **66** where the presence of the pyridine required starting from the 5-bromo-2-nitropyridine **72** and 3-aminophenol **35** and slightly different conditions for the following Buchwald-Hartwig coupling.

Synthesis of 3-amido-4'-guanidino phenyloxybenzenes **32** and **33** required the conversion of benzoic acids **57** and **58** to the corresponding acid chlorides by reaction with oxalyl chloride, followed by reaction with previously synthesized Boc-protected aniline **17** to afford final Boc-protected derivatives **59** and **60**.

In the same way, the synthesis of isouronium derivatives **67-69** required the preliminary preparation of the conveniently functionalised iodo-phenol **92** that was converted in starting amine **93** using Ullmann coupling type conditions with CuI and picolinic acid.

Compound **93** was then used as a starting point for the preparation of compound **94** using usual HgCl_2 /trimethylamine conditions; in the case of compound **97** Buchwald-Hartwig type conditions were used and for compound **100** proper functionalised acid chloride was chosen. Deprotection of the THP protecting group and further amidylation afforded final Boc-protected compounds **96**, **99** and **102**.

For the synthesis of sulfonamide **70**, amine **48** was treated with sulfonamide chloride **103**, previously prepared from chlorosulfonyl isocyanate **104** and formic acid.

All final Boc-protected compounds were deprotected using HCl 4M/dioxane to obtain compounds **2-8**, **27-31**, **61-66**, TFA/ CH_2Cl_2 to obtain compound **67**, and SnCl_4 to obtain compounds **32-33** and **68-69**.

The purity of all of the final hydrochloride salts was determined by HPLC, where a minimum purity of 95% was required before proceeding to biological testing.

4.6. References

1. Berlinck, R. G. S.; Burtoloso, A. C. B.; Trindade-Silva, A. E.; Romminger, S.; Morais, R. P.; Bandeira, K.; Mizuno, C. M., *Nat. Prod. Rep.* **2010**, 27, 1871.
2. Shaw, J. W.; Grayson, D. H.; Rozas, I., Synthesis of guanidines and some of their biological applications. In *Guanidines as Reagents and Catalysts I*, Springer, **2015**.
3. Ishikawa, T., *Superbases for organic synthesis: guanidines, amidines, phosphazenes and related organocatalysts*. John Wiley & Sons, **2009**.
4. Gund, P., *J. Chem. Educ.* **1972**, 49, 100.
5. Blanco, F.; Kelly, B.; Sánchez-Sanz, G.; Trujillo, C.; Alkorta, I.; Elguero, J.; Rozas, I., *J. Phys. Chem. B* **2013**, 117, 11608.
6. Greenhill, J. V.; Lue, P., *Prog. Med. Chem.* **1993**, 30, 203.
7. Bonham, A.; Trapani, A.; Portis, L.; Brody, M., *J. Hypertens. Suppl.* **1984**, 2, S543.
8. Cruz, M. P., *P&T* **2010**, 35, 448.

9. Maruthur, N. M.; Tseng, E.; Hutfless, S.; et al., *Ann. Intern. Med.* **2016**, 164, 740.
10. Michiels, B.; Van Puyenbroeck, K.; Verhoeven, V.; Vermeire, E.; Coenen, S., *PLOS ONE* **2013**, 8, e60348.
11. Kelly, B.; McMullan, M.; Muguruza, C.; Ortega, J. E.; Meana, J. J.; Callado, L. F.; Rozas, I., *J. Med. Chem.* **2015**, 58, 963.
12. O'Sullivan, P.; Rozas, I., *ChemMedChem* **2014**, 9, 2065.
13. Diez-Cecilia, E.; Kelly, B.; Perez, C.; Zisterer, D. M.; Nevin, D. K.; Lloyd, D. G.; Rozas, I., *Eur. J. Med. Chem.* **2014**, 81, 427.
14. Tahir, S.; Badshah, A.; Hussain, R. A., *Bioorg. Chem.* **2015**, 59, 39.
15. Berlinck, R. G. S.; Trindade-Silva, A. E.; Santos, M. F. C., *Nat. Prod. Rep.* **2012**, 29, 1382.
16. Katritzky, A. R.; Rogovoy, B. V., *Arkivoc* **2005**, 49.
17. Suhs, T.; Konig, B., *Mini-Rev. Org. Chem.* **2006**, 3, 315.
18. Manimala, J. C.; Anslyn, E. V., *Eur. J. Org. Chem.* **2002**, 3909.
19. Schneider, S. E.; Bishop, P. A.; Salazar, M. A.; Bishop, O. A.; Anslyn, E. V., *Tetrahedron* **1998**, 54, 15063.
20. Kim, K. S.; Qian, L., *Tetrahedron Lett.* **1993**, 34, 7677.
21. Kelly, B.; Rozas, I., *Tetrahedron Lett.* **2013**, 54, 3982.
22. Yong, Y. F.; Kowalski, J. A.; Lipton, M. A., *J. Org. Chem.* **1997**, 62, 1540.
23. Shaw, J. W.; Grayson, D. H.; Rozas, I., *Eur. J. Org. Chem.* **2014**, 2014, 3565.
24. Burgess, E. M.; Penton Jr, H. R.; Taylor, E., *J. Org. Chem.* **1973**, 38, 26.
25. Maki, T.; Tsuritani, T.; Yasukata, T., *Org. Lett.* **2014**, 16, 1868.
26. Kantam, M. L.; Priyadarshini, S.; Joseph, P. A.; Srinivas, P.; Vinu, A.; Klabunde, K.; Nishina, Y., *Tetrahedron* **2012**, 68, 5730.
27. Beletskaya, I. P.; Cheprakov, A. V., *Coord. Chem. Rev.* **2004**, 248, 2337.
28. Zhang, H.; Cai, Q.; Ma, D., *J. Org. Chem.* **2005**, 70, 5164.
29. Poss, M. A.; Iwanowicz, E.; Reid, J. A.; Lin, J.; Gu, Z., *Tetrahedron Lett.* **1992**, 33, 5933.
30. Wu, Y.-Q.; Hamilton, S. K.; Wilkinson, D. E.; Hamilton, G. S., *J. Org. Chem.* **2002**, 67, 7553.

31. Selig, P., *Guanidines as Reagents and Catalysts II*. Springer, **2017**.
32. Gagnon, P. E.; Boivin, J. L.; Dickson, J. H., *Can. J. Chem.* **1959**, 37, 520.
33. Wang, Z. D. Z., *Comprehensive organic name reactions and reagents*. John Wiley & Sons, **2009**.
34. Rachlin, S.; Bramm, E.; Ahnfelt-Rønne, I.; Arrigoni-Martelli, E., *J. Med. Chem.* **1980**, 23, 13.
35. O'Donovan, D. H.; Rozas, I., *Tetrahedron Lett.* **2011**, 52, 4117.
36. Yin, B.-L.; Liu, Z.-G.; Zhang, J.-C.; Li, Z.-R., *Synthesis* **2010**, 2010, 991.
37. Wilcken, R.; Zimmermann, M. O.; Lange, A.; Joerger, A. C.; Boeckler, F. M., *J. Med. Chem.* **2013**, 56, 1363.
38. Scholfield, M. R.; Zanden, C. M. V.; Carter, M.; Ho, P. S., *Protein Sci.* **2013**, 22, 139.
39. Ford, M. C.; Ho, P. S., *J. Med. Chem.* **2016**, 59, 1655.
40. Eskandari, K.; Zariny, H., *Chem. Phys. Lett.* **2010**, 492, 9.
41. Rozas, I., *Phys. Chem. Chem. Phys.* **2007**, 9, 2782.
42. Lange, A.; Günther, M.; Büttner, F. M.; Zimmermann, M. O.; Heidrich, J.; Hennig, S.; Zahn, S.; Schall, C.; Sievers-Engler, A.; Ansideri, F.; Koch, P.; Laemmerhofer, M.; Stehle, T.; Laufer, S. A.; Boeckler, F. M., *J. Am. Chem. Soc.* **2015**, 137, 14640.
43. Ho, P. S., *Future Med. Chem.* **2017**, 9, 637.
44. Dardonville, C.; Goya, P.; Rozas, I.; Alsasua, A.; Martín, M. I.; Borrego, M. J., *Biorg. Med. Chem.* **2000**, 8, 1567.
45. Kelly, B.; O'Donovan, D. H.; O'Brien, J.; McCabe, T.; Blanco, F.; Rozas, I., *J. Org. Chem.* **2011**, 76, 9216.
46. O'Donovan, D. H.; Kelly, B.; Diez-Cecilia, E.; Kitson, M.; Rozas, I., *New J. Chem.* **2013**, 37, 2408.
47. Altomonte, S.; Zanda, M., *J. Fluorine Chem.* **2012**, 143, 57.
48. Wipf, P.; Mo, T.; Geib, S. J.; Caridha, D.; Dow, G. S.; Gerena, L.; Roncal, N.; Milner, E. E., *Org. Biomol. Chem.* **2009**, 7, 4163.
49. Welch, J. T.; Lim, D. S., *Biorg. Med. Chem.* **2007**, 15, 6659.
50. Böhm, H. J.; Banner, D.; Bendels, S.; Kansy, M.; Kuhn, B.; Müller, K.; Obst-Sander, U.; Stahl, M., *ChemBioChem* **2004**, 5, 637.

51. Gillis, E. P.; Eastman, K. J.; Hill, M. D.; Donnelly, D. J.; Meanwell, N. A., *J. Med. Chem.* **2015**, *58*, 8315.
52. Zhou, P.; Zou, J.; Tian, F.; Shang, Z., *J. Chem. Inf. Model.* **2009**, *49*, 2344.
53. Sebolt-Leopold, J. S.; English, J. M., *Nature* **2006**, *441*, 457.
54. Okaniwa, M.; Hirose, M.; Imada, T.; Ohashi, T.; Hayashi, Y.; Miyazaki, T.; Arita, T.; Yabuki, M.; Kakoi, K.; Kato, J.; Takagi, T.; Kawamoto, T.; Yao, S.; Sumita, A.; Tsutsumi, S.; Tottori, T.; Oki, H.; Sang, B.-C.; Yano, J.; Aertgeerts, K.; Yoshida, S.; Ishikawa, T., *J. Med. Chem.* **2012**, *55*, 3452.
55. Clayden, J.; Greeves, N.; Warren, S.; Wothers, P., *Organic Chemistry* Oxford University Press, **2012**.
56. Nicolaou, K. C.; Bulger, P. G.; Sarlah, D., *Angew. Chem. Int. Ed.* **2005**, *44*, 4442.
57. Kurti, L.; Czako, B., *Strategic applications of named reactions in organic synthesis*. Elsevier, **2005**.
58. Surry, D. S.; Buchwald, S. L., *Chem. Sci.* **2011**, *2*, 27.
59. Alonso, F.; Beletskaya, I. P.; Yus, M., *Chem. Rev.* **2002**, *102*, 4009.
60. Bellamy, F. D.; Ou, K., *Tetrahedron Lett.* **1984**, *25*, 839.
61. Du, X.-J.; Bian, Q.; Wang, H.-X.; Yu, S.-J.; Kou, J.-J.; Wang, Z.-P.; Li, Z.-M.; Zhao, W.-G., *Org. Biomol. Chem.* **2014**, *12*, 5427.
62. Nagle, P. S.; Rodriguez, F.; Nguyen, B.; Wilson, W. D.; Rozas, I., *J. Med. Chem.* **2012**, *55*, 4397.
63. Rodríguez, F.; Rozas, I.; Kaiser, M.; Brun, R.; Nguyen, B.; Wilson, W. D.; García, R. N.; Dardonville, C., *J. Med. Chem.* **2008**, *51*, 909.
64. Miel, H.; Rault, S., *Tetrahedron Lett.* **1997**, *38*, 7865.
65. Freeman, N. S.; Gilon, C., *Synlett* **2009**, 2097.
66. Oguro, Y.; Cary, D. R.; Miyamoto, N.; Tawada, M.; Iwata, H.; Miki, H.; Hori, A.; Imamura, S., *Bior. Med. Chem.* **2013**, *21*, 4714.
67. Ishimoto, K.; Fukuda, N.; Nagata, T.; Sawai, Y.; Ikemoto, T., *Org. Process Res. Dev.* **2014**, *18*, 122.
68. Flapper, J.; Kooijman, H.; Lutz, M.; Spek, A. L.; van Leeuwen, P. W. N. M.; Elsevier, C. J.; Kamer, P. C. J., *Organometallics* **2009**, *28*, 3272.
69. Hartwig, J. F., *Angew. Chem. Int. Ed.* **1998**, *37*, 2046.
70. Alcaraz, L.; Bennion, C.; Morris, J.; Meghani, P.; Thom, S. M., *Org. Lett.* **2004**, *6*, 2705.

71. Yin, J. J.; Zhao, M. M.; Huffman, M. A.; McNamara, J. M., *Org. Lett.* **2002**, 4, 3481.
72. Birkholz, M. N.; Freixa, Z.; van Leeuwen, P. W. N. M., *Chem. Soc. Rev.* **2009**, 38, 1099.
73. Dierkes, P.; van Leeuwen, P. W. N. M., *J. Chem. Soc., Dalton Trans.* **1999**, 1519.
74. Kamer, P. C. J.; van Leeuwen, P. W. N.; Reek, J. N. H., *Acc. Chem. Res.* **2001**, 34, 895.
75. Sunesson, Y.; Limé, E.; Nilsson Lill, S. O.; Meadows, R. E.; Norrby, P.-O., *J. Org. Chem.* **2014**, 79, 11961.
76. Wolfe, J. P.; Buchwald, S. L., *Tetrahedron Lett.* **1997**, 38, 6359.
77. Rozas, I.; Sánchez-Sanz, G.; Alkorta, I.; Elguero, J., *J. Phys. Org. Chem.* **2013**, 26, 378.
78. Rozas, I.; Alkorta, I.; Elguero, J., *Struct. Chem.* **2008**, 19, 923.
79. Rozas, I.; Kruger, P. E., *J. Chem. Theor. Comput.* **2005**, 1, 1055.
80. Nagle, P. S.; Kahvedžić, A.; McCabe, T.; Rozas, I., *Struct. Chem.* **2012**, 23, 315.
81. Li, Y. N.; Urbas, A.; Li, Q., *J. Org. Chem.* **2011**, 76, 7148.
82. Miyashita, M.; Yoshikoshi, A.; Grieco, P. A., *J. Org. Chem.* **1977**, 42, 3772.
83. Maiti, D.; Buchwald, S. L., *J. Am. Chem. Soc.* **2009**, 131, 17423.
84. Li, T.-S.; Zhang, Z.-H.; Jin, T.-S., *Synth. Commun.* **1999**, 29, 181.
85. Kahvedžić-Seljubac, A.; Nathwani, S.-M.; Zisterer, D. M.; Rozas, I., *Eur. J. Med. Chem.* **2016**, 117, 269.
86. Anderson, R. J.; Groundwater, P. W.; Todd, A.; Worsley, A. J., Sulfonamide Antibacterial Agents. In *Antibacterial Agents*, John Wiley & Sons, **2012**.
87. Nissen, S., *Ophthalmologica* **1979**, 179, 286.
88. Supuran, C. T.; Scozzafava, A., *Expert Opin. Ther. Pat.* **2002**, 12, 217.
89. Rad, M. N. S.; Khalafi-Nezhad, A.; Asrari, Z.; Behrouz, S.; Amini, Z.; Behrouz, M., *Synthesis-Stuttgart* **2009**, 3983.
90. De Luca, L.; Giacomelli, G., *J. Org. Chem.* **2008**, 73, 3967.
91. Bahrami, K.; Khodaei, M. M.; Soheilzadt, M., *J. Org. Chem.* **2009**, 74, 9287.
92. Appel, R.; Berger, G., *Eur. J. Inorg. Chem.* **1958**, 91, 1339.

93. Shahani, V. M.; Ball, D. P.; Ramos, A. V.; Li, Z. H.; Spagnuolo, P. A.; Haftchenary, S.; Schimmer, A. D.; Trudel, S.; Gunning, P. T., *Biorg. Med. Chem.* **2013**, *21*, 5618.

Chapter 5

Chapter 5 – Biochemical evaluation and Physicochemical Properties

5.1. Introduction

Previous research in the Rozas' group found that lead compound **1** has a broad spectrum of inhibition of kinases downstream of EGFR (Src, Akt, B-Raf, MEK, ERK). Like sorafenib, compound **1** (20 μ M) is capable of inhibiting B-Raf in an *in vitro* assay towards a MEK-1 substrate peptide. Moreover, a LanthaScreen assay suggested that compound **1**, unlike sorafenib could inhibit B-Raf through a type III allosteric mechanism, a hypothesis that was supported by docking studies.^{1,2} Previous studies show how compound **1** has an IC₅₀ value of around 10 μ M in the HL-60 cancer cell line which is a good starting point for a preclinical *hit* compound, but that still requires improvement in order to obtain more potent and active derivatives. Moreover, its spectrum of biochemical activity requires further elucidation to unmistakably identify the biological target and clearly define its mode of action.

This chapter describes the biochemical study of the synthesised derivatives of lead compound **1** in a wide range of cancer cell lines with the aim of improving the IC₅₀ values and defining the biochemical pathways by which these compounds execute their activity. Hence, in collaboration with Prof. Daniela Zisterer at the School of Biochemistry and Immunology at Trinity College Dublin, we have conducted cytotoxicity experiments of the novel synthesised compounds and we have studied their profile investigating the molecular mechanism(s) underlying this inhibitory activity.

5.2. Cytotoxicity Evaluation of the Novel Compounds

In vitro cell studies are particularly useful to screen large series of compounds in a fast and easily accessible way. Thus, in this project cell viability assays were used to

evaluate the cytotoxicity of all synthesised compounds in a range of different cancer cell lines that comprise both haematological and solid tumours types. Therefore, we have used the HL-60 cell line (human promyelocytic leukemia cells) as the starting point for the screening of all compounds. We were then able to deduce a provisional structure-activity relationship (SAR) that allowed us to define the structural requirements for improved binding at the active site. The most successful derivatives were further studied *in vitro* in MCF-7 (breast cancer), HeLa (cervical tumour), HCT116 and HKH-2 (colorectal) as well as NCI-H929 and U266B1 (multiple myeloma) cancer cell lines. The most promising compound was also chosen for apoptotic evaluation and its toxicity against MCF-10A (non-tumorigenic epithelial cell line) and human platelets was also studied.

There are various methods for the evaluation of cell viability like plasma membrane integrity, DNA synthesis, DNA content, enzyme activity, presence of ATP and cellular reducing condition.³ For the purposes of this project, we have chosen the alamarBlue[®] assay that uses the reducing power of living cells to quantitatively measure the proliferation of human and animal cells after treatment with different concentrations of a particular drug. Resazurin, the active component of alamarBlue[®] reagent, is a non-toxic, cell permeable compound that is blue in colour and virtually non-fluorescent. When cells are alive they maintain a reducing environment within the cytosol of the cell and, thus, upon entering cells, resazurin is reduced to resorufin, a red and highly fluorescent compound. Viable cells continuously convert resazurin to resorufin, increasing the overall fluorescence that is then read on a spectrophotometer (Figure 5.2.1).⁴

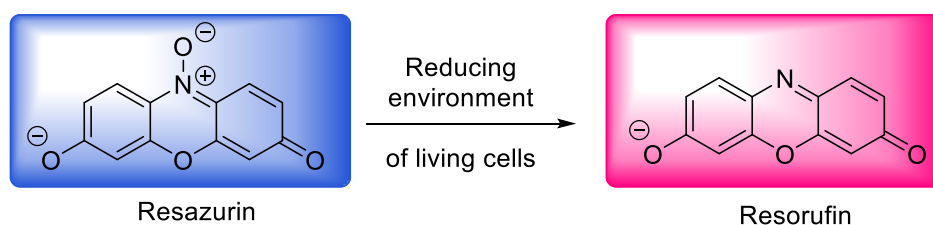


Figure 5.2.1. Resazurin, the active component of alamarBlue[®] reagent is reduced to the highly fluorescent resorufin.

5.2.1. Evaluation of the Cytotoxicity of Compounds in the HL-60 Cell Line

The HL-60 cell line is a human caucasian promyelocytic leukaemia. Peripheral blood leukocytes were first obtained by leukopheresis from a 36-year-old Caucasian female with acute promyelocytic leukemia.⁵ HL-60 cells grow in suspension culture with a doubling time that can vary from 20 h to 45 h, depending on the subline. This characteristic makes them very easy to handle and a robust cell line to use for cytotoxicity studies.⁶ All sublines display a myeloblastic/promyelocytic easy to recognise morphology: large, blast-like cells with characteristic large, rounded nuclei containing 2-4 distinct nucleoli (Figure 5.2.2).⁶

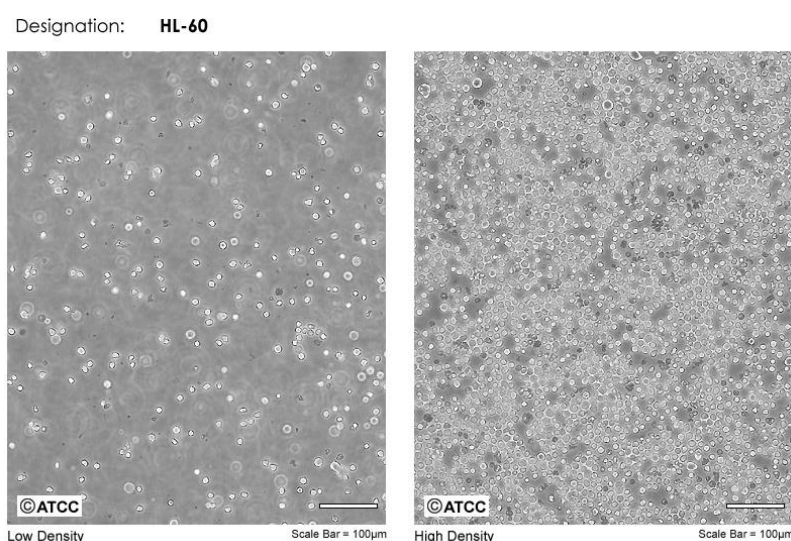


Figure 5.2.2. Morphology of HL-60 cancer cell line at low (*left*) and high (*right*) density.⁵

For these reasons, the Rozas' group has previously chosen them for a first general screen of the compounds synthesised allowing comparison of different families to identify active structural features and for the selection of derivatives to be evaluated in more resistant cancer cell lines.² Accordingly, we investigated the cytotoxicity of all the new compounds prepared and described in Chapter 4 (Figure 5.2.3) using the alamarBlue[®] viability assay with HL-60 cancer cells (see Experimental section for more details) and the results are reported in Table 5.2.1.

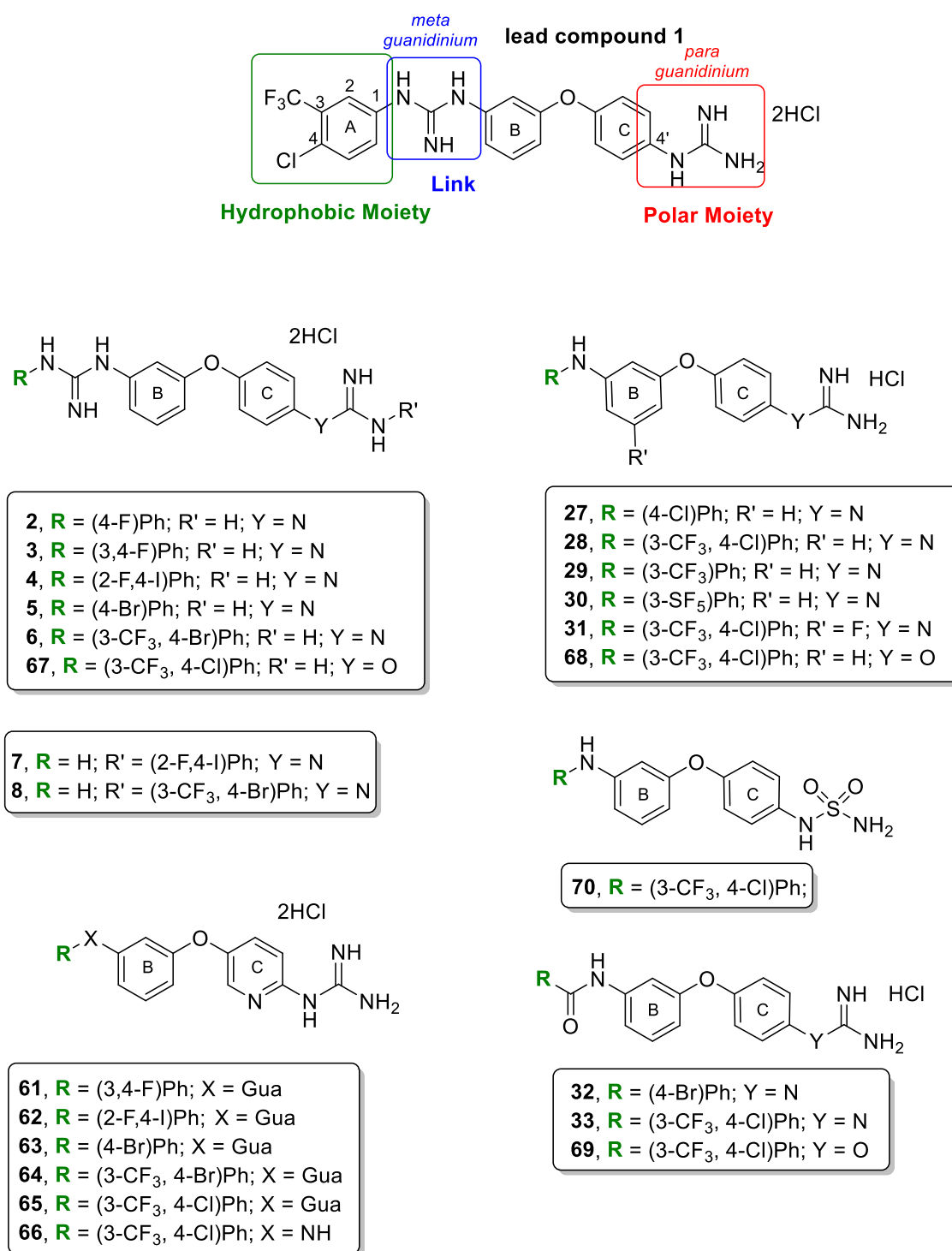


Figure 5.2.3. Structure of lead **1** and of the derivatives used for the HL-60 cancer cell viability study. The discussion in the text follows the legend names displayed in compound **1**.

Table 5.2.1. The effect of compounds **2-8**, **27-33**, **61-70**, lead compound **1** and sorafenib on the alamarBlue® assay for the HL-60 cell line.

Compound	HL-60 IC₅₀ ± SEM μM (72 h)	Compound	HL-60 IC₅₀ ± SEM μM (72 h)
Sorafenib	2.53 ± 0.68	32	26.70 ± 0.35
1	9.72 ± 0.9	33	26.70 ± 0.12
2	>100	61	11.61 ± 0.27
3	>100	62	8.33 ± 0.32
4	8.63 ± 0.51	63	9.37 ± 0.54
5	15.26 ± 1.98	64	1.53 ± 0.23
6	2.07 ± 0.33	65	2.36 ± 0.14
7	8.85 ± 0.55	66	3.48 ± 0.28
8	10.99 ± 0.64	67	5.49 ± 0.10
27	7.50 ± 0.05	68	4.22 ± 0.04
28	3.08 ± 0.15	69	35.60 ± 1.1
29	4.64 ± 0.83	70	9.14 ± 0.69
30	3.23 ± 0.36		
31	4.07 ± 0.10		

Note: Cells were seeded at a density of 2×10^5 cells/ mL in a 96-well plate and treated with different concentrations of the compounds dissolved in EtOH or DMSO (1% v/v and 0.1% v/v, respectively). Sorafenib was used as a reference and tested in the same manner. Once treated, cells were incubated for 72 hours at 37 °C after which they were treated with alamarBlue® and left in darkness in an incubator for 5 h. The resulting fluorescence was read using a plate reader from which percentage viability was calculated. IC₅₀ values were calculated using Prism GraphPad Prism software from at least three independent experiments performed in triplicate.

The results obtained for the 3,4'-bis-guanidine phenyloxybenzene derivatives **2-8** are, in general, better than those obtained for the previously tested compound **1**.² Despite that, compounds **2** and **3**, which are substituted with 3- and 3-,4- fluorine groups, respectively, give IC₅₀ values above 100 μM probably as a consequence of the small size of the fluorine atoms. This drop in activity compared to compound **1** is a clear indication of the importance of the bulky and lipophilic substituent on the phenyl ring. The 2-fluoro-4-iodo phenyl hydrophobic moiety (aryl ring A) is present in the orally bioavailable, synthetic organic molecule PD325901 targeting mitogen-activated protein kinase MEK kinase with potential antineoplastic activity.⁷ The presence of such

hydrophobic moiety in our derivative **4** maintains a good IC_{50} value (8.63 μM) compared to compound **1**, but it did not result in any dramatic increase of activity. Substitution of the 4-chloro presents in lead compound **1** with a 4-bromine as in compound **6**, gave a four-fold increase in activity, suggesting the important role played by the halogen in the *para* position of the phenyl ring at the hydrophobic moiety (aryl ring A). Hypothetically, the presence of a halogen-bond between such halogen and a Lewis base in the protein binding site could explain the improvement in activity of compound **6**. Similarly, compound **5** has a drop in activity because of the absence of the 3-trifluoromethyl substituent; this is an indication that a bigger and more polarized halogen atom as bromine in position 4 can result in a beneficial increment in activity only when accompanied by a bulky lipophilic substituent at position 3 as trifluoromethyl.

As anticipated in the Introduction chapter (section 1.5. 1.5. Previous Work Leading Up to the Present Research), when the guanidinium at the *meta* position is also substituted with the 3-trifluoromethyl-4-chloro phenyl moiety, the activity of the drug improves. Hence, and for the sake of comparison, we have prepared compounds **7** and **8** that are the equivalents of compounds **4** and **6** but substituted at the guanidinium in the *para* instead of the *meta* position. Biochemical results show that compound **7** maintains the same IC_{50} of around 8 μM as compound **4**; however, compound **8** has a drop in activity compared to related compound **6**.

Compound **28** represents the shorter version of compound **1**, with an -NH- link instead of a guanidinium at the *meta* position and this shorter derivative showed increased cytotoxicity with an IC_{50} value of 3.08 μM . Similar to what was observed for the 3,4'-bis-guanidine phenyloxybenzene derivatives, removal of the 3-trifluoromethyl substituent in aryl ring A caused a drop in activity in the shorter analogue **27** (7.50 μM). Interestingly, removal of the halogen at the 4 position did not affect the IC_{50} value of compounds **29** (4.64 μM) and **30** (3.23 μM) compared to the lead analogue **28**. This could be explained by the compensation of hydrophobicity and bulky lipophilic substituents when going from trifluoromethyl (-CF₃) to bulkier pentafluorosulphanyl (-SF₅) substitution.

When the *meta* guanidinium was replaced with an amide link, the hydrophobic moiety (aryl ring A) is forced to a different position compared to lead compound **1**. Such a

change causes a drop in activity as demonstrated for the larger IC₅₀ value of 26.70 μ M obtained for both compounds **32** and **33**. Their similar IC₅₀ values, despite their different substitution pattern, is a sign of a completely different orientation of aryl ring A, and hence, the ideal substitution patterns described so far are not relevant.

The 3,4'-bis-guanidine phenyloxypyridines **61-66** show, overall, better HL-60 cytotoxicity than the previously discussed compounds. Compound **65**, with a pyridine instead of a benzene, as in lead compound **1** (ring C), has an IC₅₀ value of 2.36 μ M, what represents a five-fold increment compared to compound **1**. Similarly, compound **64** with a 4-bromine instead of the 4-chlorine of compound **65**, shows a very good IC₅₀ of 1.53 μ M. Surprisingly, the introduction of a pyridinoguanidinium system as in compound **61** (11.61 μ M) results in a decrease of the IC₅₀ values of an inactive compounds such as **3** (>100 μ M). Even though this is not the best compound of the series, it is a clear indication of the importance of the pyridinoguanidinium system to improve cytotoxic activity.

Interestingly compound **66**, which includes both a shorter –NH– link and the pyridinoguanidinium moiety has a low IC₅₀ value of 3.48 μ M.

Compounds **67** and **68**, where an isouronium has been introduced instead of the *para* guanidinium, show similar cell viability as the guanidinium leads **1** and **28** with an IC₅₀ of 5.49 and 4.22 μ M, respectively. Isouronium derivative **69**, which is a shorter analogue of the lead **1** by means of the amide link, has a drop in activity compared to **1** showing very similar activity to the also short amide compounds **32** and **33**. From these results we can deduce that the isouronium moiety has a similar behaviour to the guanidium one, with slightly better IC₅₀ values.

Finally, compound **70**, where the *para* guanidinium is replaced by a sulfamide, shows a drop in activity to 9.14 μ M, indicating the importance of the guanidinium or isouronium positive charge for better cytotoxic activity.

Sorafenib was used as a positive control in the alamarBlue® viability assay and as expected from previous literature,² showed a low IC₅₀ of 2.53 μ M in HL-60; however, it was encouraging to see that some of our compounds have better IC₅₀ values than lead compound **1** and very similar or better than sorafenib, a compound that has been through several structural optimisations during development.

A graph representing the viability results of a translational comparison between compound **1** and its most similar derivatives is shown in Figure 5.2.4.

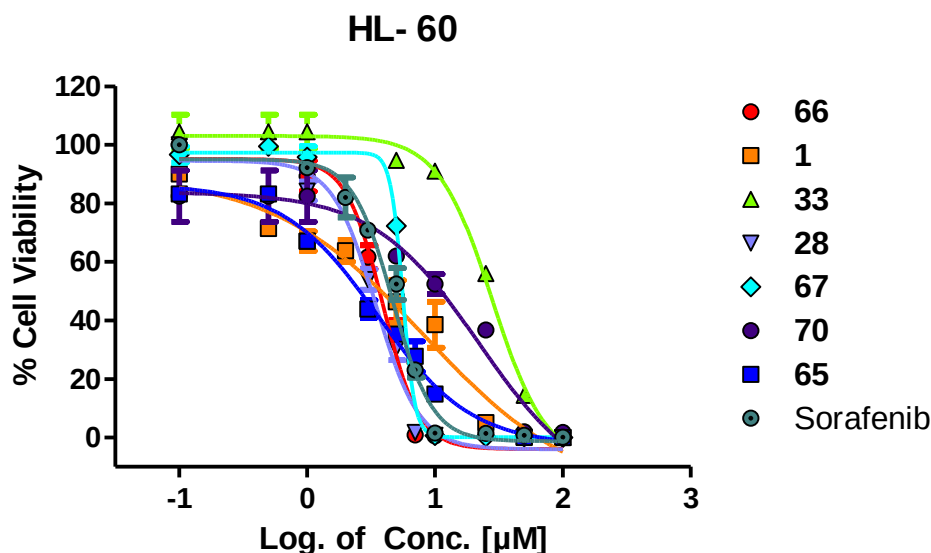


Figure 5.2.4. Compounds **66**, **33**, **28**, **67**, **70**, **65**, lead compound **1** and sorafenib reduce the viability of HL-60 cell line in a dose-dependent manner (same experimental details as described in legend from Table 5.2.1).

Furthermore, several SAR conclusions can be drawn from these cytotoxicity results and a graphical summary is reported in Figure 5.2.5.

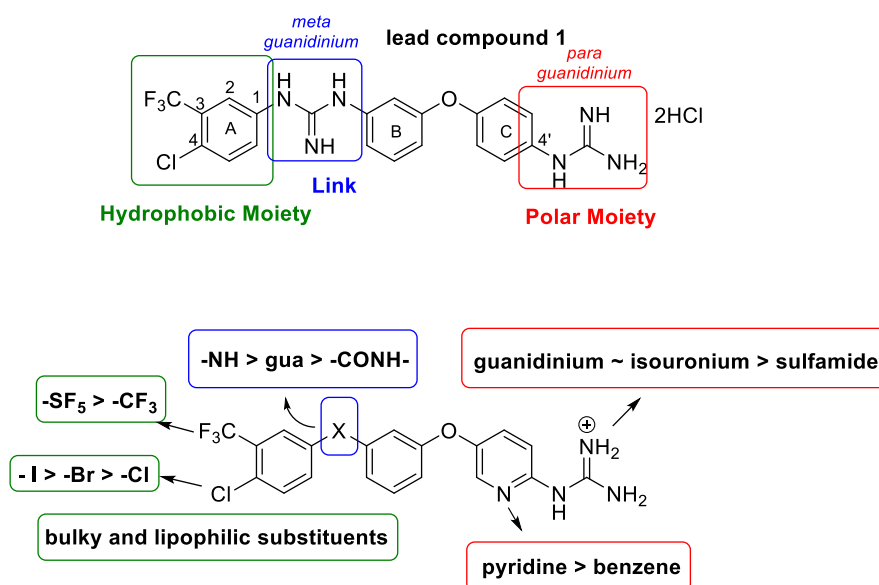


Figure 5.2.5. Structure Activity Relationship (SAR) deduced from the analysis of the HL-60 cytotoxicity results.

As represented in Figure 5.2.5: (i) the hydrophobic moiety (aryl ring A) present in lead compound **1** is necessary for the activity of new molecules. Particularly, bulky and lipophilic substituents improve the cytotoxicity of the compounds and I- or Br- substitution at the 4 position of ring A are better for the activity than Cl- substitution. This pattern of substitution has to coexist with a substituent at the 3 position of aryl ring A to maintain its efficacy. Particularly, the bulkier pentafluorosulphanyl group seems to increase the cytotoxic effect of the substituted derivatives.

(ii) The substitution of the *meta* guanidinium present in lead compound **1** with a shorter NH link results in compounds with better cytotoxicity. On the other side, amide substitution causes loss of activity.

(iii) A *para* guanidinium group as in lead compound **1** still results in the best cytotoxic derivatives, with similar results achieved for *para* isouronium derivatives. Moreover, insertion of a pyridine at aromatic ring C that helps to form an IMHB with the guanidinium cation, seems to affect positively the cytotoxicity potential of the synthesised derivatives.

5.2.2. Evaluation of the Cytotoxicity of Selected Compounds in the MCF-7 Cell Line

The MCF-7 (Michigan Cancer Foundation) cancer cell line was originally derived from a pleural effusion of a 69-year old patient with metastatic breast cancer.⁸ As previously described in the Introduction section, breast cancer has a major impact on the health of the population and its characteristics require further research in order to define its biochemical profile. The MCF-7 (Figure 5.2.6) line retains several characteristics of differentiated mammary epithelium including ability to process estradiol via cytoplasmic estrogen receptors.⁵ For this reason the MCF-7 cell line has been widely used as an *in vitro* model of breast cancer and is useful for the study of tumour response to endocrine therapy.

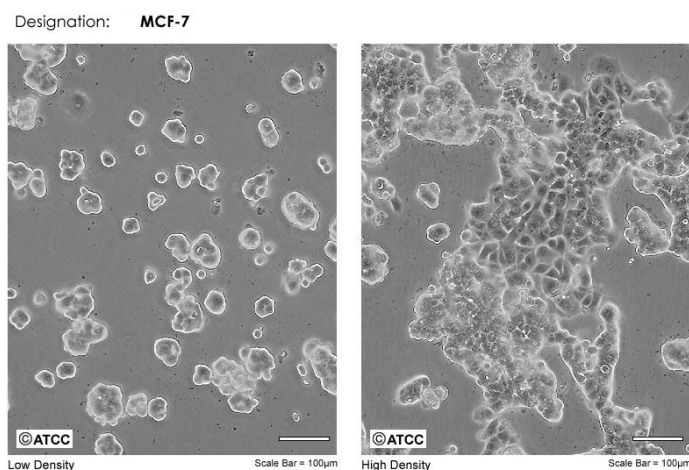
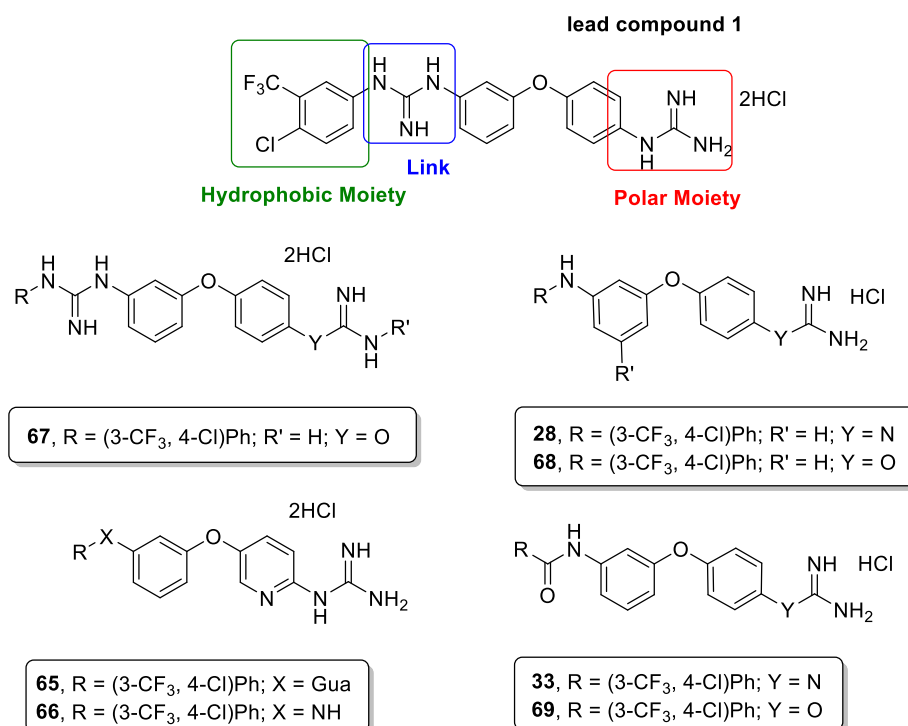


Figure 5.2.6. Morphology of MCF-7 cancer cell line at low (*left*) and high (*right*) density.⁵

Given the demonstrated cytotoxic effects of previously synthesised compounds in the Rozas' group against the MCF-7 cancer cell line,² we decided to investigate the effect of some of our most potent and significantly diverse compounds, in the MCF-7 cell line. Table 5.2.2 shows the compounds chosen for testing as well as the alamarBlue[®] viability assay results in terms of IC₅₀.

Table 5.2.2. The effect of compounds **67**, **28**, **68**, **65**, **66**, **33**, **69**, lead compound **1** and sorafenib on the alamarBlue® assay for the MCF-7 cell line.

Compound	MCF-7 IC ₅₀ ± SEM μM (72 h)
Sorafenib	3.07 ± 0.10
1	9.30 ± 1.87
28	2.02 ± 0.27
33	20.75 ± 2.3
65	4.91 ± 1.04
66	3.73 ± 0.35
67	4.68 ± 0.20
68	7.32 ± 0.35
69	28.72 ± 1.40

Note: Cells were seeded at a density of 25×10^3 cells/mL in a 96-well plate and treated with different concentrations of the compounds dissolved in EtOH or DMSO (1% v/v and 0.1% v/v, respectively). Sorafenib was used as a reference and tested in the same manner. Once treated, cells were incubated for 72 hours at 37 °C after which they were treated with alamarBlue® and left in darkness in an incubator for 5 h. The resulting fluorescence was read using a plate reader from which percentage viability was calculated. IC₅₀ values were calculated using Prism GraphPad Prism software from at least three independent experiments performed in triplicate.

The IC_{50} values obtained for the chosen compounds were very satisfactory as they were all in the low μM range, and, with few exceptions, highly resembled the results obtained for the HL-60 cell line. Compounds **28** and **66** with $IC_{50} = 2.02$ and 3.73 , respectively, are still the best compared to compound **1**, which maintains its IC_{50} at around $9.30 \mu M$. The pyridine ring present in compounds **66** and **65** ($4.91 \mu M$), still appears responsible for a gain in activity compared to compound **1**, even though if less accentuated than in the HL-60 cell line. Likewise, the isouronium derivative **67** maintains a similar IC_{50} of $4.68 \mu M$. Instead, the isouronium version of compound **28**, compound **68**, has a larger IC_{50} of $7.32 \mu M$ than the previous value for HL-60 ($4.22 \mu M$). As previously described for the HL-60 assays, the amide linked compounds **33** and **69** show large IC_{50} values of 20.75 and $28.72 \mu M$, respectively.

A graph representing the viability results of a translational comparison between compound **1** and its most similar derivatives is shown in Figure 5.2.7.

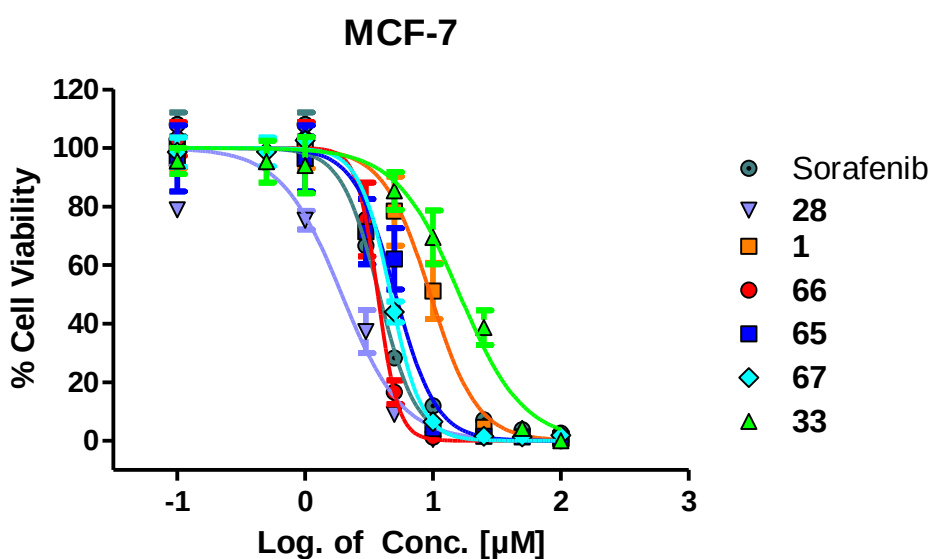


Figure 5.2.7. Compounds **28**, **66**, **65**, **67**, **33**, lead compound **1** and sorafenib reduce the viability of MCF-7 cell line in a dose-dependent manner (same experimental details as described in legend from Table 5.2.2).

5.2.3. Evaluation of the Cytotoxicity of Selected Compounds in the HeLa Cell Line

To further explore the biological profile of some chosen compounds, we decided to determine their cytotoxic effect in HeLa cell line. The HeLa (Henrietta Lacks) cell line is a cancer cell line derived from cervical cancer cells of Henrietta Lacks, a patient who died of her cancer on 1951.⁹ As for the MCF-7, the HeLa line is an adherent type of cancer cells (Figure 5.2.8), is the first immortal cell line used in scientific research and it was found to be remarkably durable and prolific which explains its extensive use in scientific research.

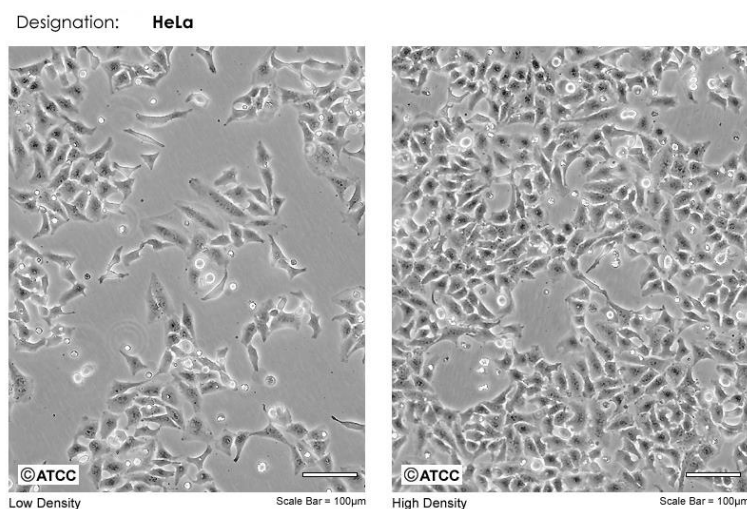
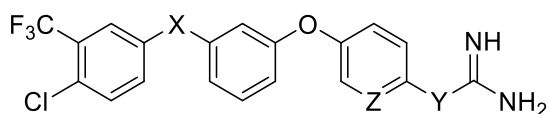


Figure 5.2.8. Morphology of HeLa cancer cell line at low (*left*) and high (*right*) density.⁵

The cytotoxicity of sorafenib had been determined previously in several publications using the HeLa cancer cell line.¹⁰ Thus, we have used sorafenib as a positive control, and evaluated the cytotoxicity of compounds **28**, **33**, **65**, **66**, **67**, **69** and lead compound **1** using the alamarBlue[®] viability assay and the results are reported in Table 5.2.3.

Table 5.2.3. The effect of compounds **28**, **33**, **65**, **66**, **67**, **69**, lead compound **1** and sorafenib on the alamarBlue® assay for HeLa cell line.

28, X = NH; Z = C; Y = N
33, X = CONH; Z = C; Y = N
65, X = Gua; Z = N; Y = N
66, X = NH; Z = N; Y = N
67, X = Gua; Z = C; Y = O
69, X = CONH; Z = C; Y = O

Compound **HeLa IC₅₀ ± SEM μM (72 h)**

Sorafenib	4.59 ± 0.45
1	9.48 ± 0.16
28	4.33 ± 0.54
33	11.28 ± 0.40
65	4.87 ± 0.27
66	1.33 ± 0.08
67	3.35 ± 0.30
69	83.57 ± 8.3

Note: Cells were seeded at a density of 25×10^3 cells/ mL in a 96-well plate and treated with different concentrations of the compounds dissolved in EtOH or DMSO (1% v/v and 0.1% v/v, respectively). Sorafenib was used as a reference and tested in the same manner. Once treated, cells were incubated for 72 hours at 37 °C after which they were treated with alamarBlue® and left in darkness in an incubator for 5 h. The resulting fluorescence was read using a plate reader from which percentage viability was calculated. IC₅₀ values were calculated using Prism GraphPad Prism software from at least three independent experiments performed in triplicate.

Analysis of the above results demonstrated that compound **66** distinguished itself from the others as having the lowest IC₅₀ value of 1.33 μM. The rest of the compounds maintained similar IC₅₀ values as those shown with HL-60 and MCF-7 cell lines, with the exception of compound **69** that loses its activity in HeLa cells with an IC₅₀ of above 80 μM and a cell viability of slightly below 40 μM even at 100 μM concentration. The dose-response curves of the compounds in HeLa cell line are illustrated in Figure 5.2.9.

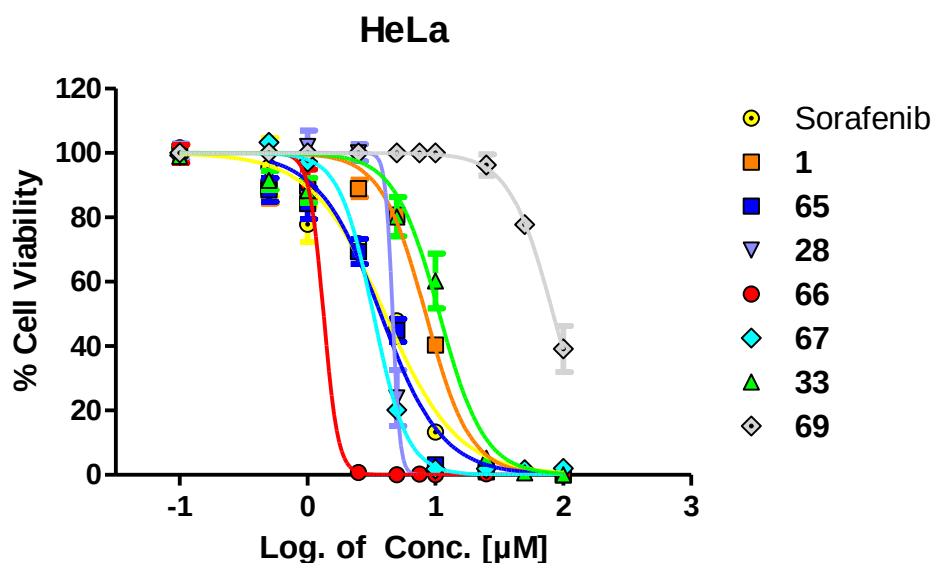


Figure 5.2.9. Compounds **65**, **28**, **66**, **67**, **33**, **69**, lead compound **1** and sorafenib reduce the viability of HeLa cell line in a dose-dependent manner (same experimental details as described in legend from Table 5.2.3).

5.2.4. Evaluation of the Cytotoxicity of Selected Compounds in the HKH-2 and HCT116 Cell Lines

Given the demonstrated cytotoxic effects of previously synthesised compounds in the Rozas' group against the RKO colorectal cancer cell line,² we decided to investigate the effect of compounds **28**, **65**, **66**, lead compound **1** and sorafenib in the HCT116 and HKH-2 cell lines, which are two different types of colorectal cancer (CRC). HCT116 cell line derives from a 48-year old male suffering from colorectal carcinoma originating in the ascending colon.¹¹ HCT116 is an adherent *KRAS*-G13D mutant cancer cell line and HKH-2 is its isogenic form with no *KRAS*-G13D mutation (Figure 5.2.10).

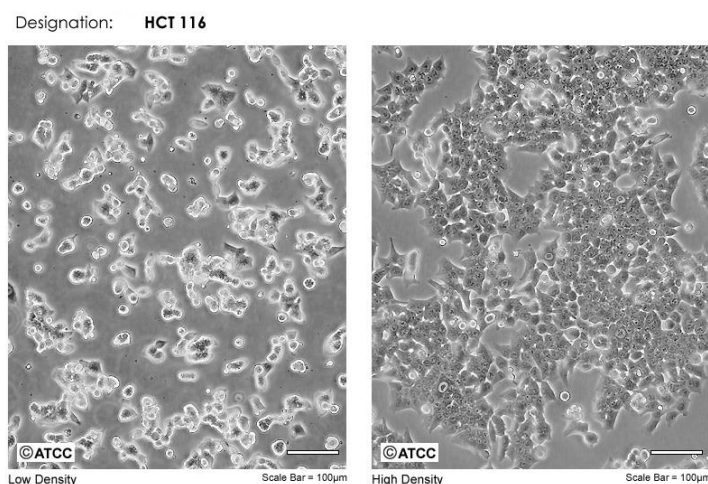
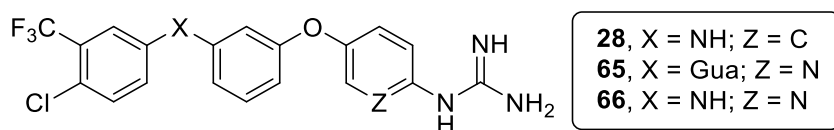


Figure 5.2.10. Morphology of HCT116 cancer cell line at low (*left*) and high (*right*) density.⁵

The three RAS genes (*HRAS*, *NRAS* and *KRAS*) constitute the most frequently mutated oncogene family in human cancers and generally RAS mutations are genetic events that occur early in tumour progression, defining the biology of the tumour.¹² In particular *KRAS* mutation is the dominant mutation in both colorectal adenocarcinoma (44.7%) and pancreatic ductal adenocarcinoma (97.7%).¹²

As previously described in the Introduction, sorafenib was originally developed as an inhibitor of the Ras effector Raf, and there are studies that show how Sorafenib enhances the therapeutic efficacy of rapamycin in colorectal cancers harbouring oncogenic *KRAS* and *PIK3CA*.¹³ Moreover the combination of Sorafenib with radiation^{14,15} or with 5-fluorouracil/leucovorin¹⁶ has demonstrated efficacy in solid tumours like CRC. Paradoxical activation of Ras effectors' pathway is very common when RAF inhibitors are evaluated in Ras-mutant cancer cells, but progress has been made recently on the development of small molecules that directly target RAS and decrease the viability of the *KRAS*-G13D mutant cancer cell line HCT116.¹⁷ Table 5.2.4 shows the compounds chosen for testing and alamarBlue® viability assay results on such cell line.

Table 5.2.4. The effect of compounds **28**, **65**, **66**, lead compound **1** and sorafenib on the alamarBlue® assay for HCT116 and HKH2 cell lines.

Compound	HCT116 IC ₅₀ ± SEM μM (72 h)	HKH-2 IC ₅₀ ± SEM μM (72 h)
Sorafenib	6.79 ± 0.18	4.43 ± 0.39
1	9.96 ± 0.60	19.18 ± 0.84
28	15.88 ± 2.73	10.34 ± 0.74
65	7.29 ± 0.53	8.09 ± 0.78
66	4.59 ± 0.40	2.88 ± 0.15

Note: Cells were seeded at a density of 1×10^5 cells/ mL in a 96-well plate and treated with different concentrations of the compounds dissolved in EtOH (1% v/v). Sorafenib was used as a reference and tested in the same manner. Once treated, cells were incubated for 72 hours at 37 °C after which they were treated with alamarBlue® and left in darkness in an incubator for 5 h. The resulting fluorescence was read using a plate reader from which percentage viability was calculated. IC₅₀ values were calculated using Prism GraphPad Prism software from at least three independent experiments performed in triplicate.

The results in Table 5.2.4 show that lead compound **1** has more potency in the HCT116 *KRAS* mutated cancer cell line (IC₅₀ = 9.96 μM) than in HKH-2 isogenic form (IC₅₀ = 19.18 μM). Remarkably, its cytotoxic effect is similar to that of sorafenib in HCT116 (6.79 μM), but not in HKH-2, where sorfenib has a much lower IC₅₀ of 4.43 μM. Similar activity in both cell lines is reached with compound **65** (around 7-8 μM) and compound **66** revealed to be the best one of the series with IC₅₀ of 4.59 μM in HCT116 and 2.88 μM in HKH-2. Interestingly, compound **28**, with a benzene instead of the pyridine of compounds **65** and **66**, loses activity with an IC₅₀ of 15.88 μM in HCT116 and 10.34 μM in HKH-2.

The dose-response curves of compounds in both HCT116 and HKH-2 cell lines are illustrated in Figure 5.2.11 where the good activity of compound **66** (red curve) is clearly showed.

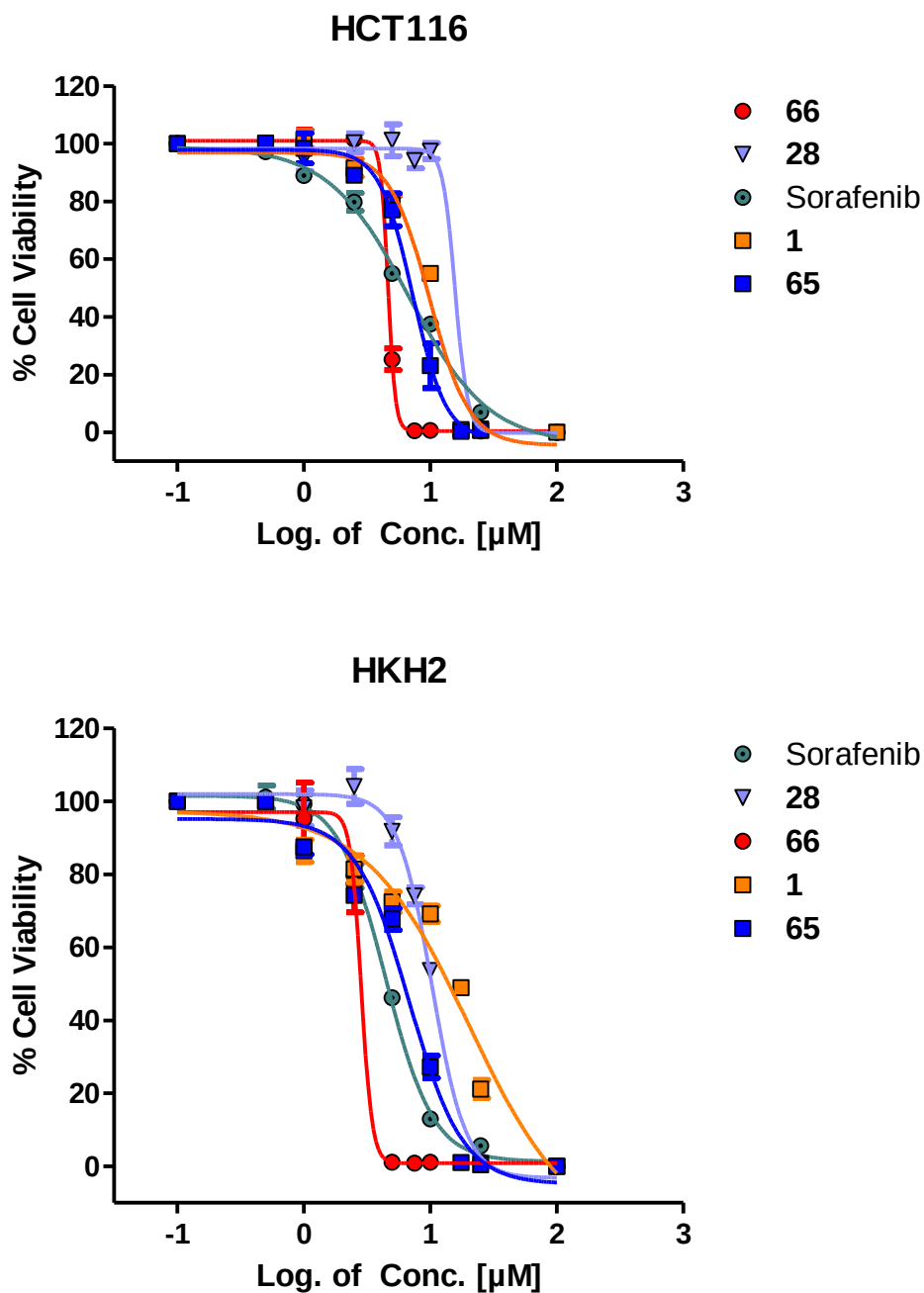


Figure 5.2.11. Compounds 28, 65, 66, lead compound 1 and sorafenib reduce the viability of HCT116 (*above*) and HKH2 (*below*) cell lines in a dose-dependent manner (same experimental details as described in legend from Table 5.2.4).

5.2.5. Evaluation of the Cytotoxicity of Compounds **1** and **66** in the NCI-H929 and U266B1 Cell Line

Following the exciting results for the inhibition of HCT116 and HKH-2 CRC cell lines and in collaboration with Prof. Daniela Zisterer at the School of Biochemistry and Immunology and Prof. Tony McElligott at the School of Medicine Trinity College Dublin, we continued the investigation of compound **66** and lead compound **1** in NCI-H929 and U266B1 cancer cell lines. NCI-H929 is a suspension cell line derived from bone marrow of type B lymphocytes that was established from a malignant effusion in a Caucasian female 62 years-old patient with myeloma (Figure 5.2.12).⁵ NCI-H929 also has an activated Ras allele.⁵

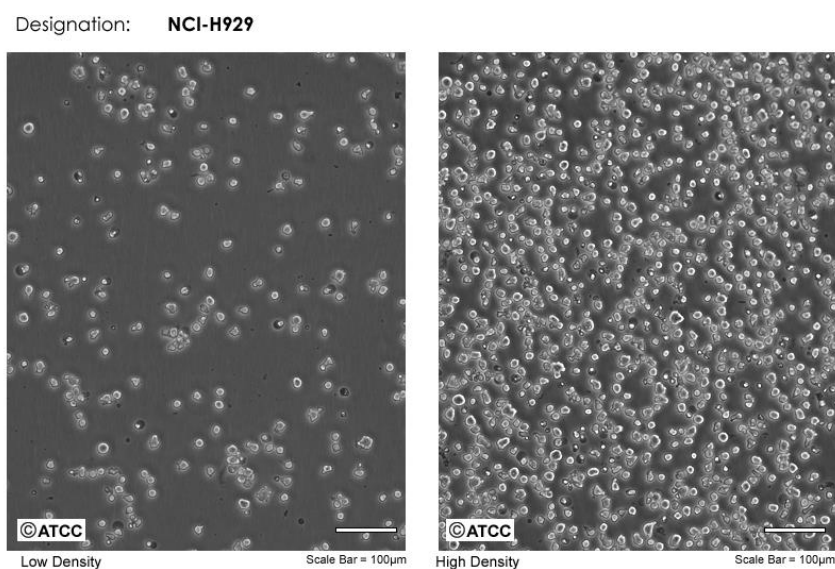


Figure 5.2.12. Morphology of NCI-H929 cancer cell line at low (*left*) and high (*right*) density.⁵

U266B1 is also a suspension cell line derived from peripheral blood of type B lymphocytes that was established from a malignant effusion in a 53 years-old male patient with myeloma (Figure 5.2.13)⁵

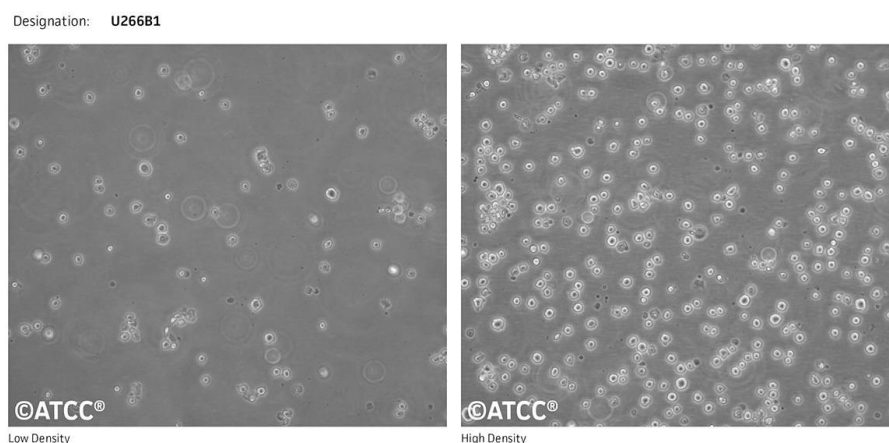


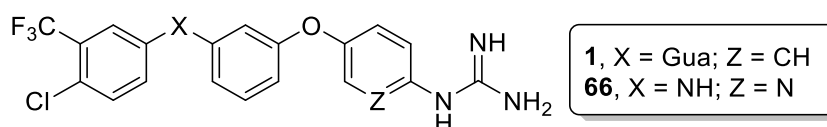
Figure 5.2.13. Morphology of U266B1 cancer cell line at low (*left*) and high (*right*) density.

Multiple myeloma (MM) is a haematological neoplastic plasma-cell disorder that is characterised by malignant proliferation of plasma cells in the bone marrow microenvironment.¹⁸ MM is an extremely genetically complex cancer type and there are a lot of subsequent mutations that contribute to disease aggression and progression.¹⁹ Multiple myeloma is the third type of cancer for which *KRAS* mutations account for the 22.8%¹² and *NRAS* and *BRAF* mutations are detectable in up to 50% of newly diagnosed MM patients.²⁰ As a consequence, the Ras effector pathways Ras-RAF-MEK-ERK²¹ and the PI3K-mTOR-Akt²² pathways are frequently activated in MM through Ras mutations. Many of the survival factors that contribute to MM progression and resistance are downstream of key signalling molecules and also involve the NF- κ B²³ and the JAK-STAT²⁴ pathways.

Over the past decade, the use of targeted therapy alone or in combination with proteasome inhibitors (e.g. bortezomib, carfilzomib), immunomodulatory drugs (e.g. lenalidomide, thalidomide, pomalidomide) and corticosteroids (e.g. prednisone and dexamethasone) have improved the outcomes for patients with MM, but more targeted and specific therapies still need to be developed.^{25,26} In 2015, panobinostat (pan-histone deacetylase inhibitor) and mono-clonal antibodies have been approved for the treatment of MM.²⁶ Sorafenib has shown anti-myeloma activity in both *in vitro* and *in vivo* model systems and is being investigated in combination with current MM therapies in preliminary studies.²⁷

Data in Table 5.2.5 show the alamarBlue® viability assay results for compounds **1** and **66**. This particular study was carried out by PhD student Rebecca Amet, in Prof. Daniela Zisterers' group at TCD.

Table 5.2.5. The effect of compounds **1** and **66** on the alamarBlue® assay for NCI-H929 and U266B1 cell lines.



Compound	NCI-H929 IC ₅₀ ± SEM μM (72 h)	U266B1 IC ₅₀ ± SEM μM (72 h)
1	4.85 ± 0.29	20.50 ± 3.33
66	3.04 ± 0.23	5.01 ± 0.36

Note: Cells were seeded at a density of 2.5×10^5 cells/mL in a 96-well plate and treated with different concentrations of the compounds dissolved in EtOH (1% v/v). Once treated, cells were incubated for 72 hours at 37 °C after which they were treated with alamarBlue® and left in darkness in an incubator for 5 h. The resulting fluorescence was read using a plate reader from which percentage viability was calculated. IC₅₀ values were calculated using Prism GraphPad Prism software from at least three independent experiments performed in triplicate.

Both compounds **1** and **66** showed good activity of 3-4 μM in the NCI-H929 cell line. Surprisingly, in the U266B1 cell line compound **1** lost efficacy with an IC₅₀ value of 20.50 μM; however compound **66** maintained an IC₅₀ value of 5.01 μM as in previously discussed cell lines.

The dose-response curves of compounds in both NCI-H929 and U266B1 cell lines are illustrated in Figure 5.2.14 where the increased activity of compound **66** (red curve) compared to lead compound **1** is clearly showed.

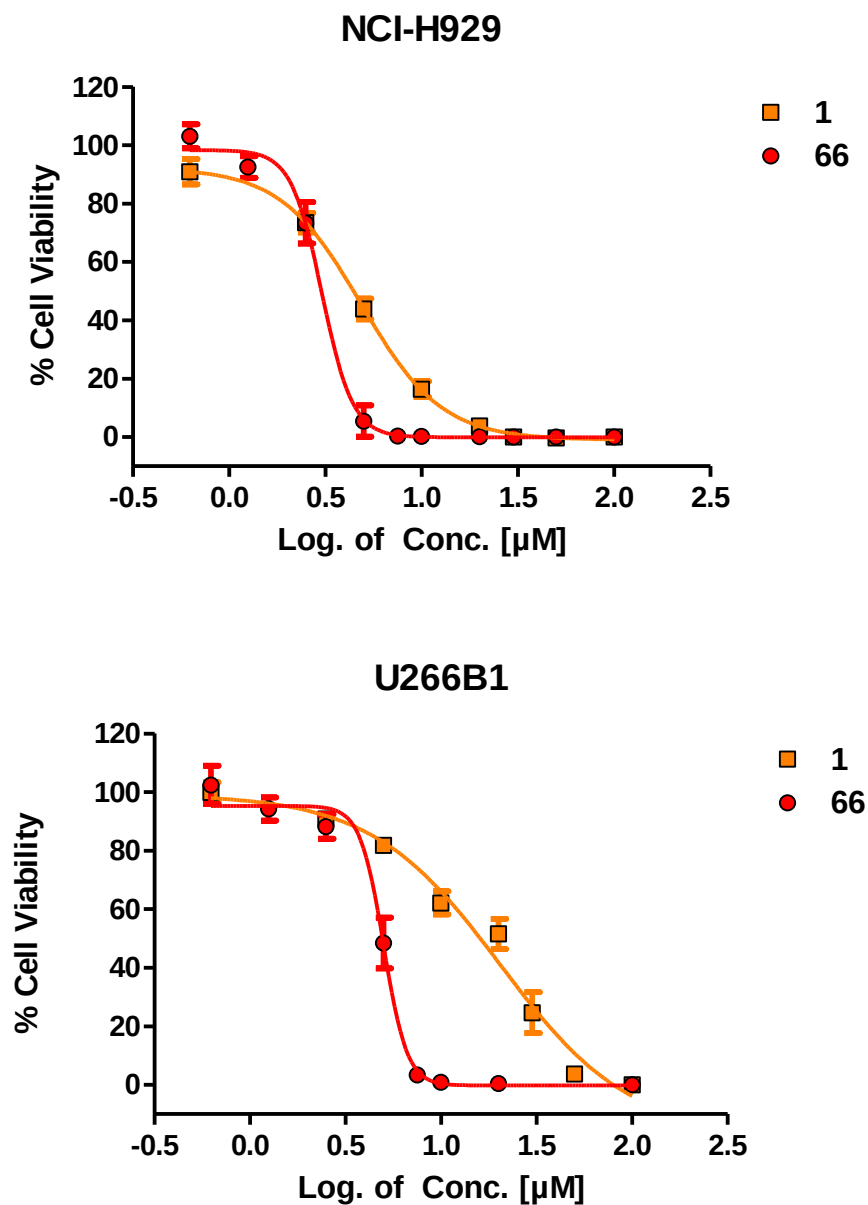


Figure 5.2.14. Compound **66** and lead compound **1** reduce the viability of NCI-H929 (*above*) and U266B1 (*below*) cell lines in a dose-dependent manner (same experimental details as described in legend from Table 5.2.5).

5.2.6. Flow cytometry

Flow cytometry is a technology that can quantitatively analyse physical and chemical characteristics of single cells and is being employed at all stages of the drug development process.²⁸ Current flow cytometers are rapid, use multiple lasers and can detect many different fluorochrome tags. Stained cells in suspension pass through a sheath fluid and they are hit by laser light sources of different choices, fluorescence is then emitted from stained cells, detected and analysed by a connected computer. This detection system measures cell size (forward scatter), cell complexity/granularity (side scatter) and fluorescence.²⁸ A variety of flow cytometric methods to analyse the cell cycle progression have been developed. Fluorescent stains such as propidium iodide are used to distinguish the cell cycle stage of a certain sample by quantitation of DNA content in a stoichiometric way, i.e. they bind in proportion to the amount of DNA present in the cell.

A cell cycle profile is a histogram representation of the distribution of DNA content of individual cells in a growing cell population. Figure 5.2.15 shows a typical cell cycle (*left*) and the related major subpopulations found in a standard DNA profile determined by the DNA content (*right*).

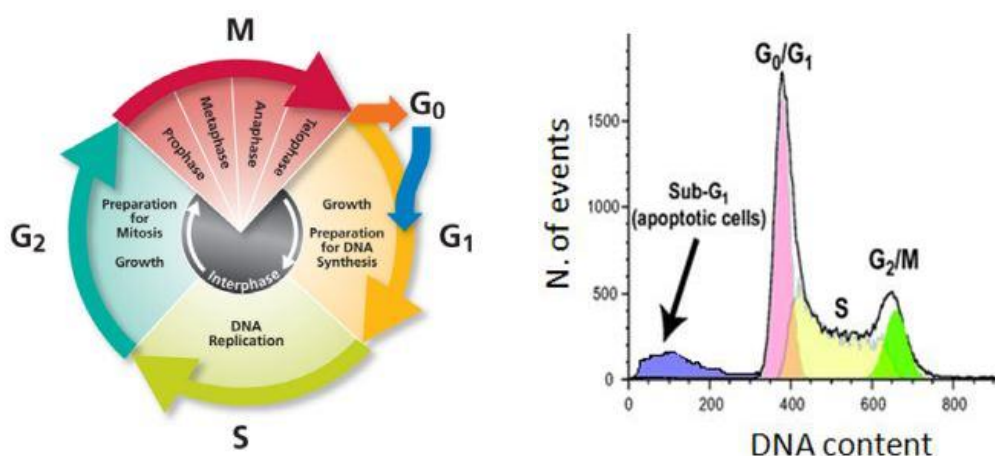


Figure 5.2.15. Cell-cycle representation (*left*) and DNA content found in each cell-cycle phase (*right*).²⁹

The pink peak in Figure 5.2.15 represents cells at rest (G_0 phase) and/or cells preparing to engage in division (G_1 phase), both of which have the same DNA content ($2N$, where N represents the haploid chromosome number). The second population (yellow peak, S phase) represents cells that have engaged in DNA replication and thus have a higher DNA content than the G_0/G_1 peak (higher than $2N$ and lower than $4N$). Finally, the third population includes cells that have doubled the DNA content but not engaged in mitosis ($4N$, green peak, G_2 phase) and cells already engaged in the preliminary steps of mitosis ($4N$, green peak, M phase). The group of cells in Figure 5.2.15 identified with a blue peak (Sub- G_1) are the apoptotic cells, with a DNA content lower than $2N$ ($< 2N$).

Propidium iodide (PI) is a popular red-fluorescent nuclear and chromosome counterstain that intercalates between the double strand of DNA. As PI cannot pass the cellular membrane, it will only reach the DNA if the cellular membrane is damaged, as in dead or dying cells. When bound to nucleic acids, it works as a dye with distinctive excitation wavelengths of 536 and 617 nm.³⁰ In addition, annexin V (AnV) is a cellular protein that binds to phosphatidylserine (PS), a phospholipid, which in normal viable cells, is located on the inner bilayer of the cell membrane. However, in the intermediate stages of apoptosis, PS is translocated from the inner to the outer leaflet of the membrane, exposing PS to the external cellular environment where it can be detected. Highly fluorescent AnV conjugates provide quick and reliable detection methods for studying the externalization of PS.³⁰

AnV staining combined with PI staining, allows quantitative data to be obtained on the level of apoptosis undergone in a cellular fraction: high levels of PI but low AnV indicates degraded and necrotic cells, low PI but high AnV indicates early apoptosis and high levels of both PI and AnV indicates late stage apoptosis.

Thus, given its good activity across all the previously discussed cancer cell lines, compound **66** was chosen for flow cytometry studies to assess its apoptotic inducing ability. BD Accuri C6 software was used to carry out the AnV staining analysis in NCI-H929 and U266B1 cell lines (multiple myeloma) after 24 h treatment with a range of concentrations (0.1 - 1 - 2.5 - 5- 7.5 - 10 μ M) of compound **66** (see Experimental section for more details) and the results are reported in Figure 5.2.16 (*above* NCI-H929

and *bottom* U266B1). This particular study was carried out by PhD student Rebecca Amet, in Prof. Daniela Zisterers' group at TCD.

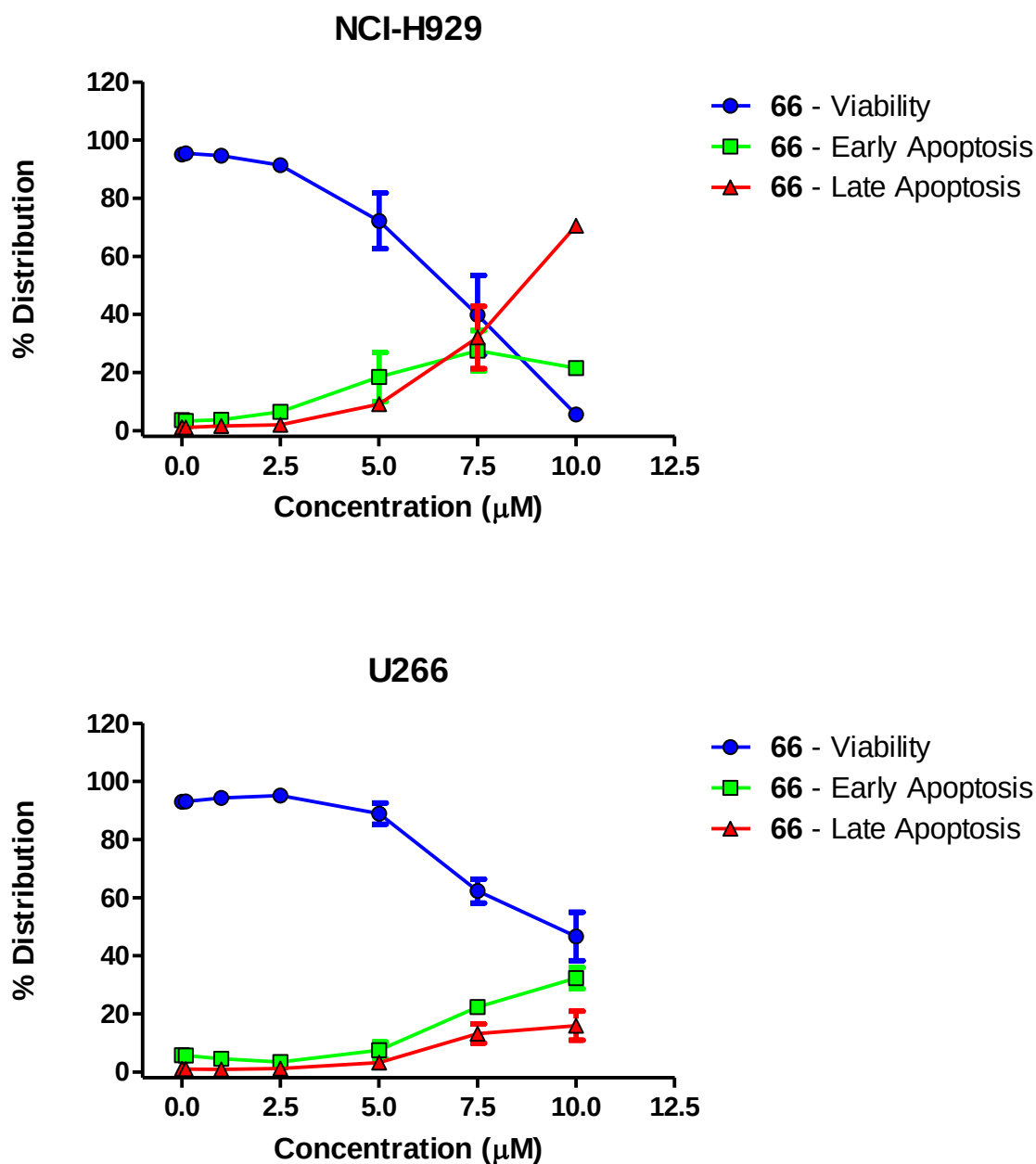


Figure 5.2.16. Compound **66** induces apoptosis in multiple myeloma cells. NCI-H929 and U266B1 cells were seeded at 3×10^5 cells/mL in a total volume of 2 mL per well in a 12-well plate, untreated, treated with vehicle EtOH (0.5% v/v), or treated with a range of concentration of compound **66** (2.5 – 10 μ M). Cells were incubated for 24 h at 37 °C and 5% CO₂. Samples were stained with AnV and PI, held on ice and in the dark, and analysed immediately. Data was gated to exclude cell debris and doublets. 10,000 cells were analysed using BD Accuri C6 software through the 630/22 nm band pass filter. Graphs were prepared using GraphPad Prism. Data represent the mean $\pm$ SEM for three separate experiments.

The early apoptotic events increased in a concentration dependent manner after treatment of both NCI-H929 and U266B1 cell lines with compound **66** for 24 h as can be observed in Figure 5.2.16. In particular, for the NCI-H929 cell line, 7.5 μM seems to be the threshold where cell viability, early and late apoptosis meet. At a higher concentration of compound **66** (10 μM), the cell viability is highly reduced and the late apoptosis events overcome the early apoptosis ones, resulting in cell damage and cell death. In the case of the U266B1 cell line, at 10 μM the cell viability is only slightly reduced from 60% and the early apoptotic events prevail over the late apoptosis ones, indicating the need of a higher concentration of around 10 μM to induce apoptosis.

AnV and PI staining analysis clearly indicated the ability of compound **66** to induce apoptotic cell death in NCI-H929 and U266 cancer cell lines at a concentration of around 8-10 μM after only 24 h.

5.2.7. Effect of compound 66 on the expression levels of the anti-apoptotic MCL-1 protein in NCI-H929 and U266 cancer cell lines

Flow cytometry results reported in the previous section demonstrate the ability of compound **66** to induce apoptosis. To further examine the biological mechanism behind this event, western blots were performed to probe for the anti-apoptotic Myeloid cell leukemia-1 (Mcl-1) protein.

As previously described in the Introduction section, Mcl-1 is a member of the Bcl-2 family of proteins responsible for the regulation of programmed cell death.³¹ Mcl-1 is an anti-apoptotic protein that functions by sequestering pro-apoptotic proteins to prevent apoptosis; hence, drug-induced inhibition of the Mcl-1 protein levels is a sign of induced apoptosis.³² For example, sorafenib induces apoptosis in HL-60 cells by inhibiting STAT3 phosphorylation, and by decreasing the expression of the anti-apoptotic proteins Mcl-1 and Bcl-2, (associated with increased apoptosis in HL-60 cells).³³

The biochemical technique used to probe for Mcl-1 levels was western blotting, which is one of the most common methods for protein detection in molecular biology.

Western blotting uses the high affinity relationship between certain antigens expressed on the proteins of interest and commercially available primary antibodies to detect down-regulation or up-regulation of those proteins. It is a multistep process consisting of protein extraction, quantification, separation, western blot analysis and protein detection in a qualitative or semi-quantitative way (see Experimental section for details).

Accordingly, NCI-H929 and U266B1 cancer cell lines were treated for 16 h with 10 μ M of compound **66** and Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was used as loading control. Figure 5.2.17 shows how in both cell lines compound **66** completely reduces the Mcl-1 protein levels, with similar effects to sorafenib for NCI-H929, but more effective than sorafenib in U266B1.



Figure 5.2.17. Compound **66** dramatically reduces Mcl-1 expression as determined by western blotting. NCI-H929 and U266B1 cells were treated with either vehicle (0.5% EtOH), sorafenib (10 μ M) or compound **66** (10 μ M) for 16 hours. Cells were lysed and 20 μ g protein was loaded and separated on 12% SDS-page gel. Membrane was probed with anti-Mcl-1 [1:1000] and anti-rabbit secondary antibody. Anti-GAPDH [1:2500] was used as a loading control. Results are representative of three separate experiments.

These western blot results correlates with the ability of compound **66** to efficiently induce cell death through apoptosis after only 16 h of treatment.

5.2.8. Thermal Shift analysis in B-Raf^{V600E} protein

The thermal shift assay is a biophysical method used to assess inhibitor's binding to protein. It operates on the principle that ligand binding alters thermal stability of proteins. A fluorescent dye based probe, which preferentially binds the hydrophobic

regions of a protein increasingly exposed during protein denaturation, is used. Fluorescent-based detection allows one to monitor protein denaturation upon heating. A ligand bound to a protein, e.g. to its active site, has the propensity to increase its thermal stability (and hence its melting, T_m) through newly formed ligand-protein interactions.³⁴⁻³⁶

In collaboration with Prof. Stefan Knapp at the University of Frankfurt, we employed a T_m shift assay to screen if lead compound **1** and chosen compounds **28** and **66** would bind to B-Raf^{V600E} (see Experimental Section for details). The results of the study are showed in Figure 5.2.18.

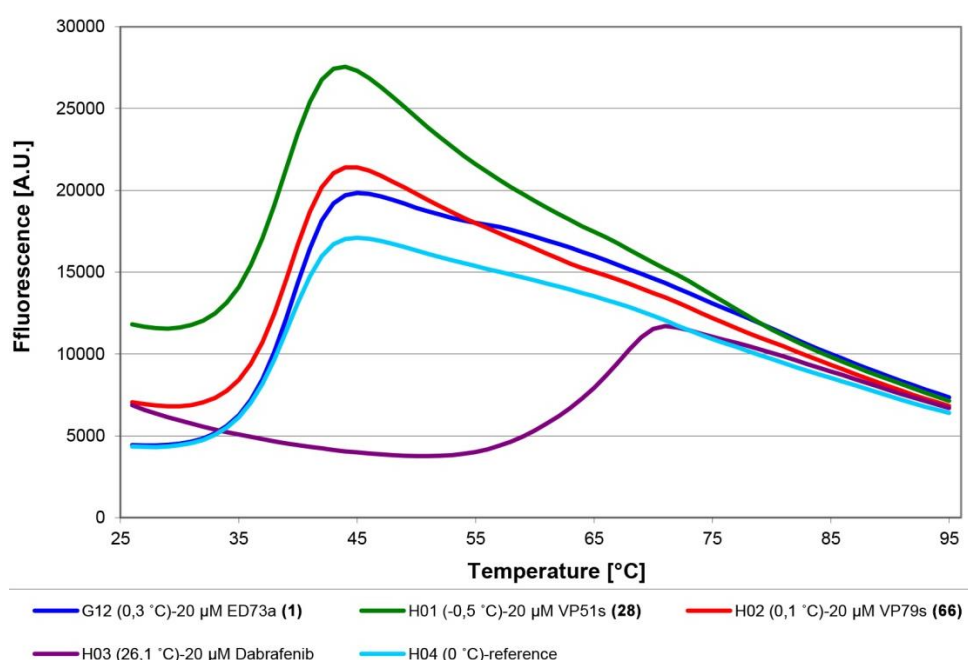


Figure 5.2.18. Thermal shift analysis in B-Raf^{V600E} for lead compound **1** (blue curve) and compounds **28** (green curve) and **66** (red curve) indicating the corresponding protein-ligand T_m . Dabrafenib (purple curve) was used as a positive control.

Surprisingly, none of the chosen inhibitors at a 20 μM concentration were shown to induce any melting temperature shift (Figure 5.2.18, green, red and blue curves) indicating no binding to the protein. Dabrafenib (commercially available inhibitor of B-Raf^{V600E}) was used as a positive control at the same concentration (Figure 5.2.18, purple curve) and showed as expected a pronounced melting temperature shift (26 °C).

These results came as a surprise given the fact that previous studies found compound **1** to inhibit B-Raf activity.¹ The negative results may be due to the fact the B-Raf chosen for the thermal shift assay is a mutated-type and previous results were conducted in wild-type B-Raf. Moreover, our hypothesis is that our compounds act as type III allosteric inhibitors. In principle, even though T_m shifts are expected to detect all binders including those that bind to cavities other than the active site, this has yet to be further investigated,³⁶ and this could be another reason for the negative results.

5.2.9. Investigation of the Cytotoxicity of compounds 1 and 66 on Healthy Cells and Human Platelets

Given the promising results for compound **66** as a potential anticancer agent and due to its higher cytotoxicity profile in cancer cell lines when compared to the lead compound **1**, the potential toxic effect of both compounds on a human breast epithelial cell line (MCF10A) and human platelets was also investigated.

5.2.9.1. Evaluation of the Cytotoxicity of compound 66 in MCF10A Cell Line

The MCF10A human mammary epithelial cell line is a widely used *in vitro* model to study normal breast cell function.³⁷ Given the promising IC_{50} values in the low micromolar range that we obtained with compound **66** against MCF-7 breast cancer cell line, we tested the cytotoxicity of this compound in the corresponding ‘normal’ cell line, MCF10A. Different concentrations (100, 10 and 1 μ M) of compound **66** were incubated for 24 h in a 96 well plate and then treated with alamarBlue[®]. Three experiments were performed in triplicate and average results are represented in Figure 5.2.19; due to lack of time, cell viability values and cytotoxicity percentage could not be properly determined, but, as shown in Figure 5.2.19, compound **66** only shows toxicity towards MCF10A at a high concentration of 100 μ M (blue dye). At lower concentrations of 10 and 1 μ M the compound does not appear to affect in any way cell viability (pink dye). This result is very promising and given that compound **66** has an IC_{50} value of 3.73 μ M in the previously tested MCF-7 cancer cell line, there is a wide therapeutic window available for this compound at which it may cause antiproliferative effects in breast tumour cells without significant toxicity to non-tumorigenic cells.

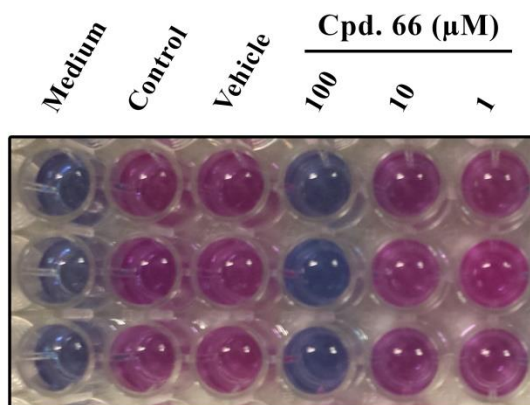


Figure 5.2.19. A 96 well plate, where an increase of compound concentration/cytotoxicity is observed due to the colour change of the alamarBlue[®] dye (blue-pink). Cells were seeded at a density of 25×10^3 cells/mL in a 96-well plate and treated with different concentrations of the compounds dissolved in EtOH (1% v/v). Once treated, cells were incubated for 24 hours at 37 °C after which they were treated with alamarBlue[®] and left in darkness in an incubator for 5 h.

5.2.9.2. Investigation of the Effect of compounds 1 and 66 on Platelet Function

The effect of compounds **1** and **66** on human platelet cytotoxicity and aggregation was investigated in collaboration with Prof. Maria Santos-Martinez's group from the School of Pharmacy & Pharmaceutical Sciences at Trinity College Dublin.

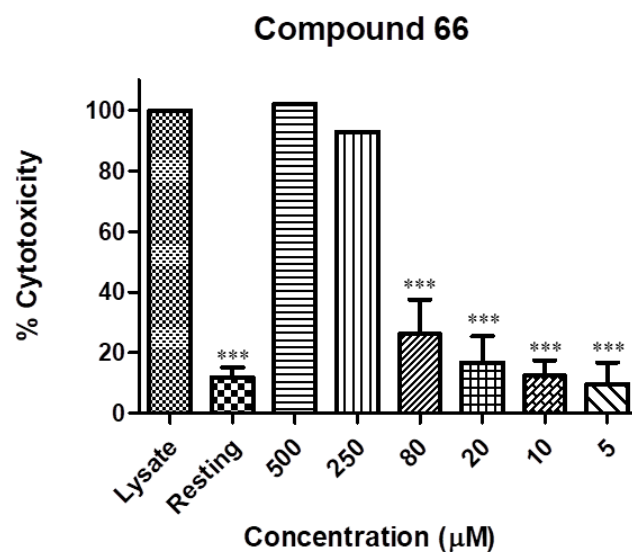
The administration of drugs represents one of the most common causes of platelet dysfunction in our society. While some drugs, like acetylsalicylic acid (aspirin) or adenosine diphosphate receptor antagonists (clopidogrel and prasugrel), are therapeutically used due to their antiplatelet effect, some others like nonsteroidal anti-inflammatory drugs, some antibiotics, lipid-lowering drugs or selective serotonin reuptake inhibitors, can impair platelet function.³⁸ In fact, it is well-known and reported in the literature that drug administration can affect platelet function leading to both, pro-thrombotic^{39,40} or thrombocytopenic^{41,42} states.

Drug-induced thrombocytopenia (DIT) or drug-induced thrombosis can be caused by different factors and different mechanisms, including those affecting the vessel wall, the blood flow, and/or different blood constituents.³⁹ Therefore, a direct evaluation of the potential effect of preclinical compounds on platelet function can be a starting point to avoid complications and side effects in future drug development.

5.2.9.2.1. Platelet Cytotoxicity

The cytotoxicity of compounds **1** and **66** on human platelets was assessed using the CytoTox-ONE Homogeneous Membrane Integrity Assay (Promega). Compounds were incubated with washed platelets (WP) from three different blood donors at increasing concentrations (see Experimental Section for details). The lysates from platelets from the same donors were taken as 100% of cytotoxicity for comparison. None of the compounds resulted in toxicity to platelets at the IC₅₀ range values calculated in previous experiments for cancer cells (9-10 μ M for compound **1** and 2-3 μ M for compound **66**). Even at higher concentrations (up to 200 μ M for compound **1** and 80 μ M for compound **66**) the drugs did not show any cytotoxic effect. In fact, excluding 500 and 250 μ M for compound **66**, the values from all the remaining concentrations tested were statistically significant when compared to the lysate and similar to the resting platelets (Figure 5.2.20).

(A)



(B)

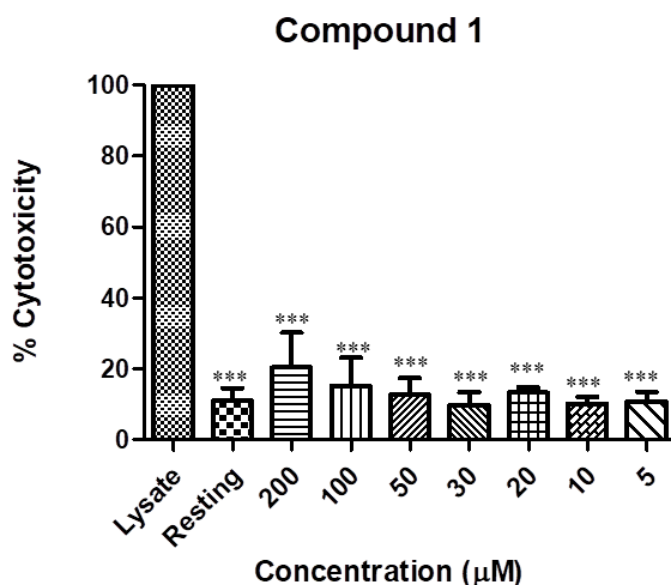


Figure 5.2.20. Statistical Analysis of the effect of compound **66** (A) and compound **1** (B) in human platelets. Bars represent mean $\pm$ Standard Deviation (SD). One-way analysis of variance (ANOVA) and Tukey post-test. *** $p < 0.001$ vs. lysate. The effect of ethanol on platelet function was also tested in three donors to prove that the vehicle did not interfere with the results obtained.

5.2.9.2.2. Platelet Aggregation

Light Transmission Aggregometry (LTA) was employed to investigate the potential pro-aggregatory effect of the compounds on human platelets and their effect

corroborated by optical microscopy (see Experimental Section for details). Human platelet-rich plasma (PRP) was incubated under continuous stirring in the presence or absence of the compounds at increasing concentrations. The normal function of platelets was verified for every donor using a platelet agonist (collagen $2\ \mu\text{g mL}^{-1}$) and corresponding aggregometry traces and microscopy pictures are reported in Figure 5.2.21. None of the compounds induced platelet aggregation at any of the concentrations tested when compared to stirred platelets. In fact, the statistical analysis ($n=3$) resulted in a p-value of 0.6213 for compound **66** and of 0.7924 for compound **1**, when one-way analysis of variance was applied (Figure 5.2.22).

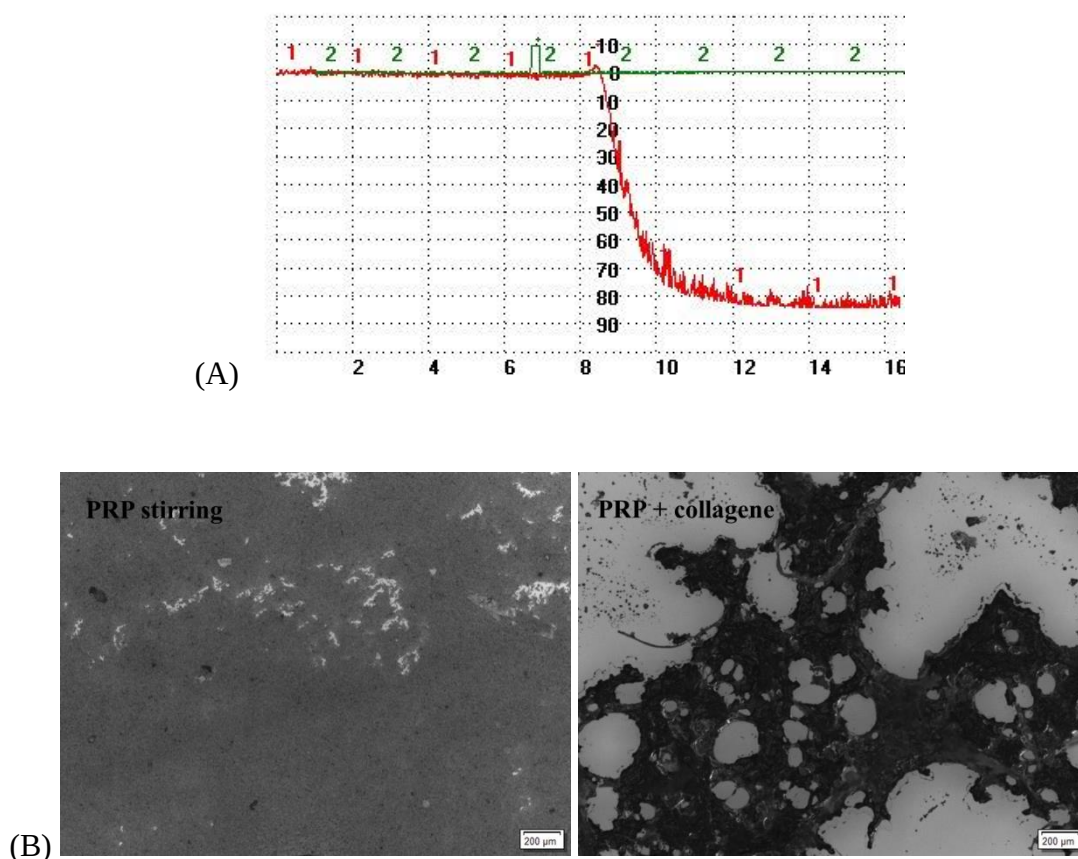
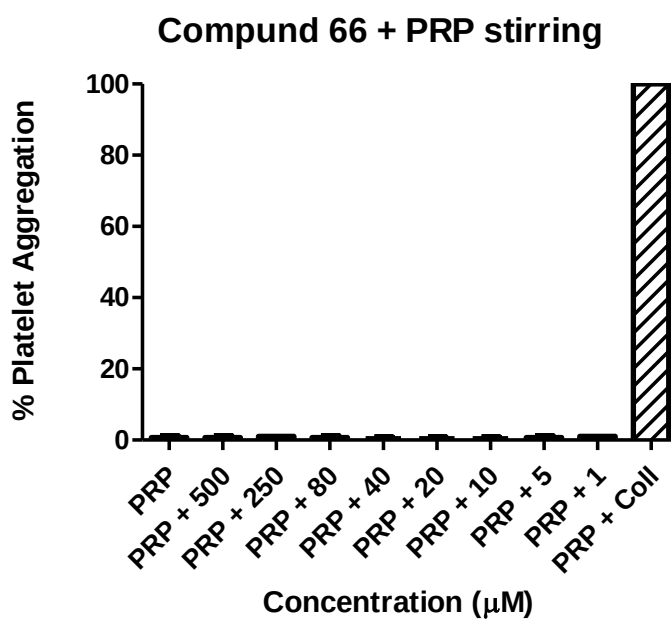


Figure 5.2.21. (A) Representative traces for PRP + collagen (trace 1 – red) and PRP alone (trace 2 – green). (B) Photomicrographic evidence of normal PRP (*left*, 10x) and PRP + collagen (*right*, 10x), with large aggregates shown in black.

(A)



(B)

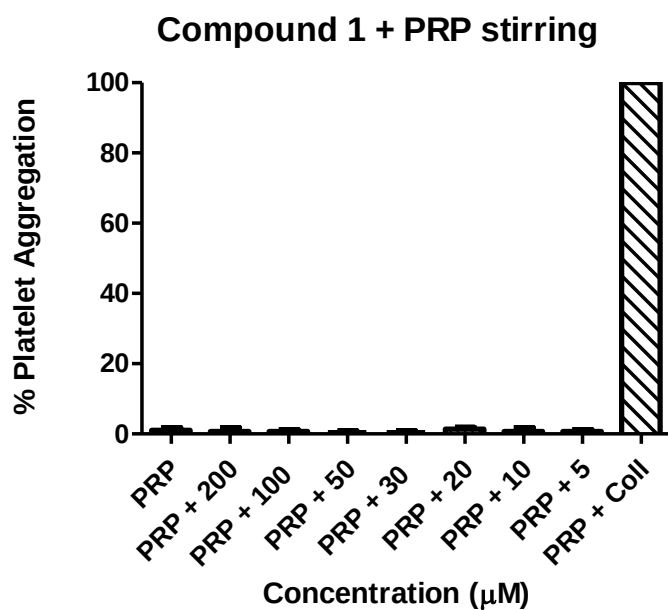


Figure 5.2.22. Statistical analysis of the effect for compound **66** (A) and lead compound **1** (B) at different concentrations on platelet aggregation (PRP). The normal function of platelets was verified for every donor using a platelet agonist (collagen $2 \mu\text{g mL}^{-1}$). Bars represent mean $\pm$ Standard Deviation (SD). One-way analysis of variance (ANOVA) and Tukey post-test ($n=3$). The effect of ethanol on platelet function was also tested in three donors to prove that the vehicle did not interfere with the results obtained.

5.3. Physicochemical Properties

Good activity and low toxicity are a necessary starting point in drug-development, but additionally, particular attention to the physicochemical properties of active molecules should be given if the end point for these molecules is the drug industry or further development. Different fundamental physicochemical properties are most often used in defining compound quality but, of these, log P, pK_a , log $D_{7.4}$, together with solubility and hydrogen bonding descriptors (HB donors and HB acceptors) are of critical importance.⁴³ The physicochemical properties of a molecule influences its solubility, binding, absorption, distribution, metabolism and excretion (ADME properties); they are therefore of big importance in determining the destiny of a drug. There is a large variety of *in silico* methods for predicting the ADME parameters from molecular structure⁴⁴ and certainly the pioneer work can be attributed to Lipinsky, who examined orally active compounds' characteristics to define physicochemical ranges for high probability of a compound to be an oral drug and gave the known *Rule-of-five* (Ro5) as a general guidance for drug-likeness.⁴⁵

We have chosen the freeware available tools SwissADME⁴⁶ and ChemAxon's Marvin⁴⁷ to computationally evaluate the physicochemical properties of our derivatives and the calculated results are reported in Table A.3.1. in the Appendix. Both SwissADME and Marvin are computer-aided drug design (CADD) tools that rely on several models, both physics based or statistical empirical models as the QSAR models.⁴³ CADD has been a pioneer field for the application of machine learning⁴⁸ technologies and QSAR analysis uses various machine learning statistical tools to predict correlations.⁴³

Next, we discuss the most relevant physicochemical properties and correlations found (for a complete list see Table A.3.1. in the Appendix), together with the related references for the computational methods used for the calculations, a thorough explanation of which is beyond the scope of this section. All synthesised molecules have a molecular weight (MW) of less than (but close to) 500 Da, except for 3,4'-bis-guanidine phenyloxybenzenes **4**, **6**, **7** and **8** or pyridine derivatives **62** and **64** which MW exceeds by only 7-10 units this threshold due to the presence of heavier bromine and iodines substituents (values calculated with OpenBabel,⁴⁹ version 2.3.0). Such small exceptions are considered acceptable, moreover considering that an analysis of

the FDA approved 31-small molecules protein kinases inhibitors on the market (SMKIs) showed that >30% of these approved SMKIs have a MW exceeding 500 Da.⁵⁰

SwissADME allows the calculation of the lipophilicity of molecules according to different methods, and in Table A.3.1. (Appendix) the consensus log $P_{o/w}$ for all synthesised compounds is reported. This consensus value is an average of all the values obtained with all the appropriate methods used (Figure 5.3.1).⁴⁶ Overall, all of the synthesised compounds have a log $P < 5$. In particular, 3,4'-bis-guanidine phenyloxypyridines **61-65** have the lowest values for log P (around 1.87 – 2.80). On the contrary, arylamino linked derivatives have the largest values of log P (around 4). Particularly, compounds **28**, **31** and **70** reach log P values of 4.44, 4.56 and 4.43, respectively. Interestingly, isouronium derivatives **67**, **68** and **69** have higher log P values than the corresponding guanidinium derivatives (3.69 for **67** vs. 3.31 for lead compound **1**; 4.65 for **68** vs. 4.44 for **28**; 4.23 for **69** vs. 3.90 for **33**); this is probably due to the exchange of the –NH of the *para* guanidinium (which is a HBD) for an oxygen in the isouronium moiety (which is a HBA).

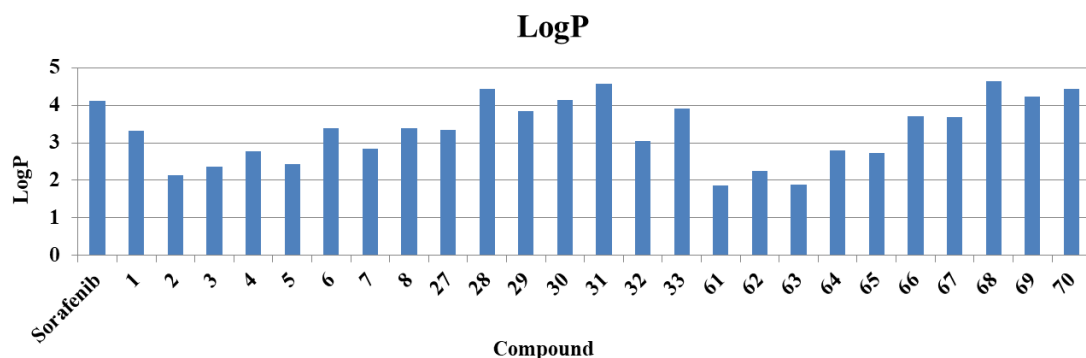


Figure 5.3.1. SwissADME calculated log P values of all synthesised compounds, lead compound **1** and sorafenib.⁴⁶

The scatter graph in Figure 5.3.2 represents a correlation between the calculated log P (x-axis) and the IC_{50} values (y-axis) in the HL-60 cell line for all synthesised compounds with the exception of the inactive ones (**2** and **3**, $IC_{50} > 100 \mu M$). The binding site in enzymes and receptors is often of hydrophobic nature and it has been shown that hydrophobic molecules prefer to minimise the amount of surface area

exposed to water favouring instead to self-associate or adopt appropriate conformations.⁴³ This phenomenon is known as the *hydrophobic effect* and it is entropically driven.⁴³ A similar type of correlation has also been found in our set of molecules, where a trend indicating how the IC_{50} decreases with the increase of the calculated log P of the molecules was found. This could also support our hypothesis that the hydrophobic moiety of our molecules could interact with a particular allosteric pocket in protein kinases, justifying bulkier and more lipophilic substituents (larger log P) for a better affinity (smaller IC_{50}).

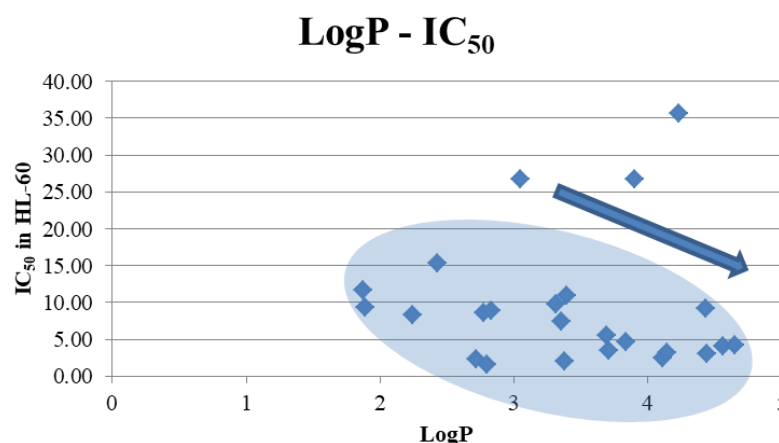


Figure 5.3.2. Scatter correlation found between the calculated log P (x-axis) and the IC_{50} values (y-axis) in the HL-60 cancer cell line for all of the synthesised compounds with the exception of the inactive ones (2 and 3, $IC_{50} > 100 \mu M$).

Topological Polar Surface Area (TPSA), which is an indicator of HB formation and a commonly used metric for the optimisation of a drug's ability to permeate cells, was also calculated for all the synthesised molecules. Compounds with a $TPSA > 140 \text{ \AA}^2$ tend to be poor at permeating cell membranes.⁵¹ All the synthesised molecules have a $TPSA < 140 \text{ \AA}^2$, which is the limit value suggested in the literature (values calculated using the fragmental technique reported in reference 50).⁴³ Specifically, shorter amino-linked derivatives have much lower TPSA (80-100 $\text{\AA}^2$) than derivatives 3,4'-bis-guanidinium or isouronium and 3,4'-bis-phenyloxypyridine guanidinium derivatives (120-135 $\text{\AA}^2$).

An intuitive graphical classification model, also known as the BOILED egg representation, is shown in Figure 5.3.3. This representation consists of a correlation between calculated log P (y-axis) for lipophilicity and calculated topological polar surface area (TPSA) for apparent polarity and it is an intuitive method to predict simultaneously two key ADME parameters, *i.e.* the passive gastrointestinal absorption (human intestinal absorption, HIA) and brain access (blood brain barrier, BBB).^{46,52} Compounds that fall into the white part of the graph are likely to go through GI absorption and those that fall into the yellow part of the graph are likely to be brain permeant. Ultimately, all of our derivatives (except containing pentafluorosulphanyl compound **30** and sulfamide **70**) could have a passive gastrointestinal absorption, but none of them can pass the BBB (Figure 5.3.3).

Moreover, SwissADME enables the estimation for a chemical to be a substrate of the permeability glycoprotein (P-gp) applying the statistical method support vector machine algorithm (SVM) on large data-sets of known substrates/non-substrates or inhibitors/non-inhibitors.^{46,53} P-gp is suggested to be the most important member among ATP-binding cassette transporters or ABC-transporters and is responsible for an active efflux through biological membranes, for instance from the gastrointestinal wall to the lumen or exiting the brain.^{46,54} Thus, as reported in the legend of Figure 5.3.3, red dots indicate that all of our compounds, except compounds **61-63** (blue dots), are non-substrate of P-gp.

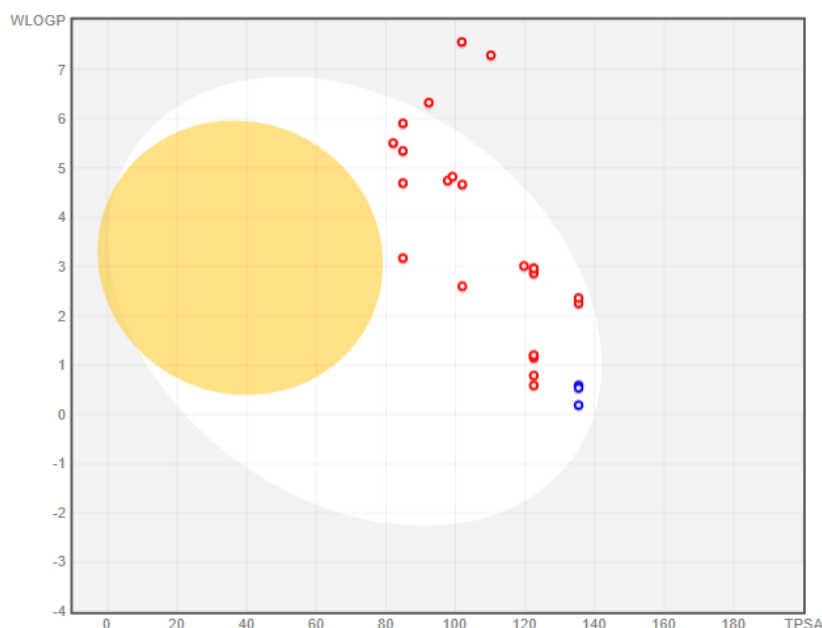


Figure 5.3.3. Legend: the white region is for high probability of passive absorption by the gastrointestinal tract, and the yellow region (yolk) is for high probability of brain penetration; the points are coloured in blue if predicted as actively effluxed by P-gp (PGP+) and in red if predicted as non-substrate of P-gp (PGP-).

The program Marvin was also used to calculate the pK_{aH} of the guanidinium or isouronium moieties at the *para* or *meta* position of all the synthesised compounds and calculated results are reported in Table A.3.1. in the Appendix. The table shows how changing the *para* guanidinium with an isouronium moiety results in a pK_{aH} decrease (less basic molecules). Compound **67** has a pK_{aH} of 8.34 compared to 8.71 for compound **1**, compound **68** has a pK_{aH} of 8.32 compared to 8.86 for compound **28** and compound **69** has a pK_{aH} of 8.32 compared to 8.71 for compound **33**. Interestingly, the pyridine derivatives **61-66** show lower pK_{aH} values (around 7.35 – 7.71), probably due to the IMHB formed between the *para* guanidinium and the pyridine ring.

Having a soluble molecule largely facilitates drug development and one of the main drawbacks of our molecules is their poor solubility in water.⁴³ Solubility is not an easy parameter to model *in silico*, but SwissADME gives an estimation of the solubility class based on three predictors, two topological and one fragmental method.⁴⁶ According to this program, all of our molecules are poorly soluble in aqueous medium. For a preclinical evaluation of our molecules, their solubility in EtOH/DMSO is still acceptable, but future work will have to be carried out to achieve full water solubility.

Alongside molecular weight and the mentioned physicochemical descriptors, many numeric descriptors are used including the number of HB acceptor and donor (HBA and HBD) groups or the number of rotatable bonds (RotB).⁴³

According to the Lipinski's Ro5, an orally active drug usually contains no more than 5 HBDs (the total number of nitrogen–hydrogen and oxygen–hydrogen bonds) and no more than 10 HBAs (all nitrogen or oxygen atoms) and many extensions to the Ro5 involve also 10 or less rotatable bonds.⁵⁵ The bar-graph presented in Figure 5.3.4 shows the number of rotatable bonds (blue), the number of HBAs (red) and the number of HBDs (green) for all the synthesised compounds plus the commercially available sorafenib and lead compound **1**. All compounds fulfil these requisites, except the number of HBDs (green), that lead compound **1** and 3,4-bis-guanidine phenyloxybenzene **2-8** or pyridine derivatives **61-65** exceed for one unit.

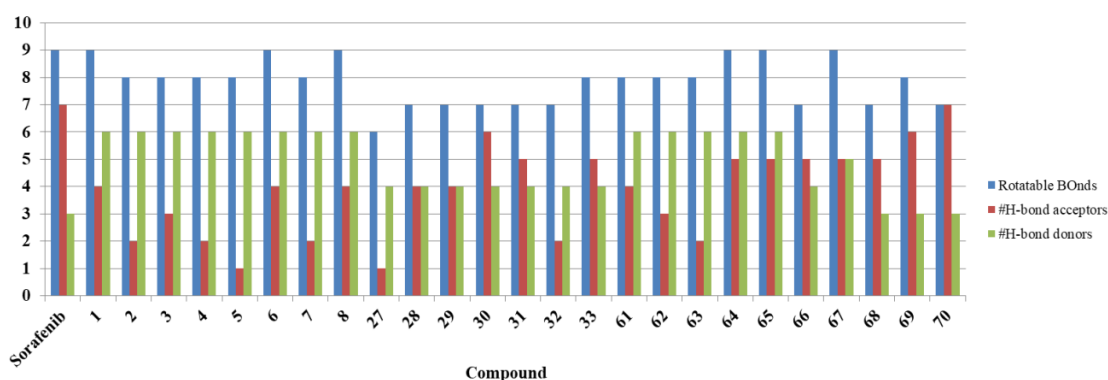


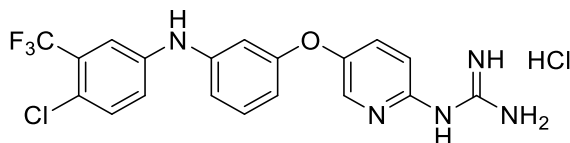
Figure 5.3.4. Graphical representation of the number of rotatable bonds (blue), HBAs (red), HBDs (green), calculated for all synthesised compounds, lead compound **1** and sorafenib.

5.4. Cancer Cell Cytotoxicity and Physicochemical Properties of Compound 66

Compound **66** (structure shown in Table 5.4.1) has proven to be the best among the all synthesised derivatives in terms of cancer cell cytotoxicity and structural properties.

Table 5.4.1 shows a summary of all the results reported in the previous sections in terms of IC_{50} values and cytotoxicity in ‘healthy’ cells for compound **66**.

Table 5.4.1. Summary of biochemical properties of compound **66**



$IC_{50} \pm SEM(\mu M)$						
HL-60	MCF-7	HeLa	NCI-H929	U266B1	HCT116	HKH2
3.48 ± 0.28	3.73 ± 0.35	1.33 ± 0.08	3.04 ± 0.23	5.01 ± 0.36	4.59 ± 0.40	2.88 ± 0.15
cytotoxicity in human platelets			cytotoxicity in MCF10A			
none (at active concentration)			none (at active concentration)			

Compound **66** shows an IC_{50} constant range of around 1-5 μM in all the cancer cell lines used for testing. The compound has shown activity in leukaemia, breast and cervical cancer cell lines. Moreover, it displays activity also in a colorectal *KRAS* mutated cancer cell line (and its isogenic form lacking the oncogenic *KRAS*) and in multiple myeloma *RAS* allele activated cancer cell line (and its isogenic form). Also, compound **66** was proven to be safe against ‘normal’ cell line MCF10A and human platelets.

Compound **66** was also proven to be a very strong apoptotic inducer as shown by Annexin V - PI analysis and confirmed by the inhibition of the anti-apoptotic Mcl-1 protein identified through western blot.

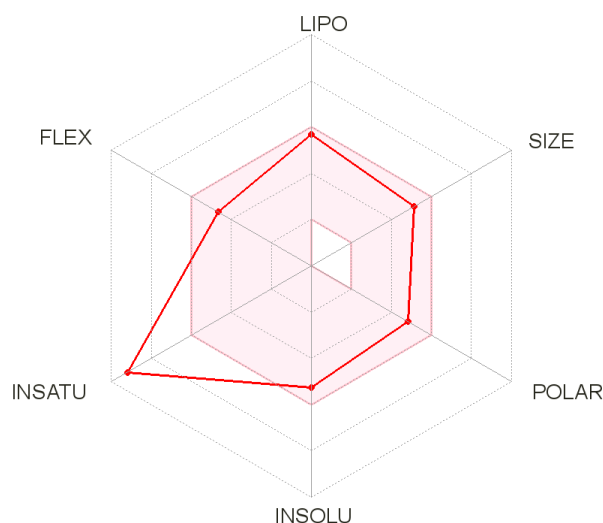
Given the good biological activity and safety profile for compound **66**, we have decided to summarise its more relevant pharmacokinetics properties in Table 5.4.2 as obtained with the previously mentioned SwissADME tool available online.

Table 5.4.2. Summary of physicochemical and pharmacokinetic properties of compound **66**

Physicochemical Properties		Pharmacokinetics	
Formula	C ₁₉ H ₁₆ ClF ₃ N ₅ O	GI absorption	High
Molecular weight	422.81 g/mol	BBB permeant	No
pKaH*	7.62	P-gp substrate	No
Num. heavy atoms	29	CYP1A2 inhibitor	Yes
Num. arom. heavy atoms	18	CYP2C19inhibitor	No
Fraction Csp ₃	0.05	CYP2C9 inhibitor	No
Num. rotatable bonds	7	CYP2D6 inhibitor	Yes
Num. H-bond acceptors	5	CYP3A4 inhibitor	Yes
Num. H-bond donors	4	Druglikeness	
TPSA	97.79 Å ²	Lipinski	Yes; 0 violation
Lipophilicity		PAINS	0 alert
Consensus Log P _{o/w}	3.71		
Water Solubility			
Class	Poorly soluble		

Properties calculated with SwissADME⁴⁶, except for *,calculated using ChemAxon's Marvin.⁴⁷

Moreover Figure 5.4.1 displays a graphical representation of compound **66** drug-likeness properties based on the data provided in Table 5.4.2.

**Figure 5.4.1.** Graphical representation of compound **66** drug-likeness properties obtained with SwissADME tool.

The calculated bioavailability radar graph represents in the pink area the optimal range for some of the reported properties. Lipophilicity: XLOGP3 between -0.7 and $+5.0$, size: MW between 150 and 500 Da, polarity: TPSA between 20 and $130 \text{ \AA}^2$, solubility: log S not higher than 6, saturation: fraction of carbons in the sp^3 hybridization not less than 0.25, and flexibility: no more than 9 rotatable bonds.⁴⁶ In our case compound **66** falls inside the Lipinski rule of five with 0 violations and also inside the range of rotatable bonds and polar surface area. Compound **66** has a poor solubility in water and a low fraction of Csp_3 of 0.05, but it has estimated high GI absorption and even though it is not BBB permeant, it is estimated not to be a substrate of P-gp glycoprotein. SwissADME enables the estimation of a chemical to be an inhibitor of cytochrome P450 (CYP) through the previously mentioned support vector machine-learning algorithm (SVM). The knowledge about interaction of molecules with cytochrome P450 (CYP) a key superfamily of isoenzymes in drug elimination through metabolic biotransformation, is essential for pre-clinical studies.⁵⁶ Five major isoforms can be identified (CYP1A2, CYP2C19, CYP2C9, CYP2D6, CYP3A4)^{46,57} and compound **66** was found to be a possible inhibitor of three of them, CYP1A2, CYP2D6, CYP3A4.

PAINS (pan assay interference compounds, known as frequent hitters or promiscuous compounds) are molecules containing substructures showing potent response in biological assays irrespective of the protein target;⁴⁶ compound **66** contains zero PAINS alerts, meaning that the biological results are trustable and attributable to the biochemical activity of the compound itself.

In conclusion compound **66** represents our new lead compound, with up to nine fold-increase in cytotoxicity towards cancer cell lines compared to compound **1**, with ideal physicochemical properties and with a modifiable scaffold for a hit-optimisation process.

5.5. Conclusions

In this chapter we have described the biochemical evaluation of the synthesised compounds as potential anti-cancer agents. This evaluation involved the study of their effects on cell viability and proliferation in a wide variety of cancer cell lines, their

apoptosis-inducing abilities, as well as the study of their molecular targets as potential protein kinase inhibitors.

Cell viability studies revealed that our compounds are able to inhibit cell proliferation of a variety of cancer cell lines in the low μM range of concentration, showing up to a nine fold increase in cytotoxicity compared to the previously prepared lead compound **1**. In particular, it can be concluded that our compounds are able to inhibit cell proliferation of leukaemia, breast cancer, cervical cancer, colorectal cancer and multiple myeloma.

Several conclusions can be made regarding SAR information from the evaluation of the cytotoxicity of all compounds in the HL-60 cell line. Bulky and lipophilic substituents are crucial to achieve high cancer cell cytotoxic activity. Compounds containing the guanidinium moiety show higher cancer cell cytotoxicity when this feature is in the presence of a pyridine ring with which an IMHB can be formed. The isouronium moiety provides similar cancer cell cytotoxicity as guanidinium, but with improved physicochemical properties. Moreover, shorter molecules (amino-linked) result in increased cancer cell cytotoxicity.

The potential toxic effect of both compound **66** and compound **1** on a human breast epithelial cell line (MCF10A) and human platelets was also investigated. Both compounds were proven to be safe in preliminary study against MCF10A cells and human platelets and to cause antiproliferative effects in tumorigenic cancer cell lines without any toxicity to non-tumorigenic cells or platelets.

In silico analysis of the physicochemical properties of all our derivatives helped to correlate optimal SAR information with drug-likeness requirements.

Compound **66**, which represents all the conclusions that can be drawn from the SAR analysis and physicochemical properties, was found to be the best compound in terms of cytotoxicity across all the cell lines tested. Specifically, it is effective in inducing cell cytotoxicity also in HCT116 colorectal cancer cell line, harbouring a *KRAS* mutated gene and resulting in a very promising agent for new possible therapies in poorly treated *RAS* mutated cancers. Moreover, compound **66** was also found to have IC_{50} values in the low μM range in multiple myeloma cell lines, a type of cancer that is poorly treatable.

Compound **66** was chosen as the lead compound and is representative of the new generation of molecules and it was used to investigate the cause of cell death induced by our compounds. Flow cytometry analysis (Annexin V analysis and PI staining) and Mcl-1 inhibition in NCI-H929 and U266B1 cell lines were performed to investigate its pro-apoptotic ability. Compound **66** is able to induce early apoptotic events only after 24 h incubation and at a low concentration of 10 μ M.

Due to previous studies that showed how the previously synthesised compounds could inhibit B-Raf protein kinases in a type III allosteric mechanism, we investigated if some of our derivatives could bind to B-Raf^{V600E} through thermal melting assays. The negative results of this study indicate that the target of our molecule is not the mutated B-Raf^{V600E}. Further investigations need to be made to explore the binding of our molecules in wild-type B-Raf through different binding assays that could discriminate for allosteric molecules.

Even though the exact molecular target(s) of our molecules still needs to be identified, the biological results obtained are very encouraging and our molecules, in particular compound **66**, can be considered excellent starting points for a phenotypic screening searching for new protein kinases inhibitors.

5.6. References

1. Diez-Cecilia, E.; Carson, R.; Kelly, B.; van Schaeybroeck, S.; Murray, J. T.; Rozas, I., *Bioorg. Med. Chem. Lett.* **2015**, 25, 4287.
2. Diez-Cecilia, E.; Kelly, B.; Perez, C.; Zisterer, D. M.; Nevin, D. K.; Lloyd, D. G.; Rozas, I., *Eur. J. Med. Chem.* **2014**, 81, 427.
3. Riss, T. L.; Moravec, R. A.; Niles, A. L.; Duellman, S.; Benink, H. A.; Worzella, T. J.; Minor, L., *Cell viability assays* **2016**.
4. ThermoFisher, *AlamarBlue Cell Viability Assay Protocol* **2015**.
5. <https://www.atcc.com>.
6. Birnie, G. D., *Br. J. Cancer* **1988**, 58, 41.

7. Barrett, S. D.; Bridges, A. J.; Dudley, D. T.; Saltiel, A. R.; Fergus, J. H.; Flamme, C. M.; Delaney, A. M.; Kaufman, M.; LePage, S.; Leopold, W. R.; Przybranowski, S. A.; Sebolt-Leopold, J.; Van Becelaere, K.; Doherty, A. M.; Kennedy, R. M.; Marston, D.; Howard, W. A.; Smith, Y.; Warmus, J. S.; Tecle, H., *Bioorg. Med. Chem. Lett.* **2008**, *18*, 6501.
8. Soule, H.; Vazquez, J.; Long, A.; Albert, S.; Brennan, M., *J. Natl. Cancer Inst.* **1973**, *51*, 1409.
9. Scherer, W. F.; Syverton, J. T.; Gey, G. O., *J. Exp. Med.* **1953**, *97*, 695.
10. Mao, W.-f.; Shao, M.-h.; Gao, P.-t.; Ma, J.; Li, H.-j.; Li, G.-l.; Han, B.-h.; Yuan, C.-g., *Acta Pharmacol. Sin.* **2012**, *33*, 1311.
11. Brattain, M. G.; Brattain, D. E.; Fine, W. D.; Khaled, F. M.; Marks, M. E.; Kimball, P. M.; Arcolano, L. A.; Danbury, B. H., *Tumour Biol.* **1981**, *2*, 355.
12. Cox, A. D.; Fesik, S. W.; Kimmelman, A. C.; Luo, J.; Der, C. J., *Nat. Rev. Drug Discov.* **2014**, *13*, 828.
13. Gulhati, P.; Zaytseva, Y. Y.; Valentino, J. D.; Stevens, P. D.; Kim, J. T.; Sasazuki, T.; Shirasawa, S.; Lee, E. Y.; Weiss, H. L.; Dong, J.; Gao, T.; Evers, B. M., *Carcinogenesis* **2012**, *33*, 1782.
14. Kuo, Y. C.; Lin, W. C.; Chiang, I. T.; Chang, Y. F.; Chen, C. W.; Su, S. H.; Chen, C. L.; Hwang, J. J., *Biomed. Pharmacother.* **2012**, *66*, 12.
15. Suen, A. W.; Galoforo, S.; Marples, B.; McGonagle, M.; Downing, L.; Martinez, A. A.; Robertson, J. M.; Wilson, G. D., *Int. J. Radiat. Oncol. Biol. Phys.* **2010**, *78*, 213.
16. Shacham-Shmueli, E.; Geva, R.; Figer, A.; Bulocinic, S.; Nalbandyan, K.; Shpigel, S.; Atsmon, J.; Brendel, E., *J. Clin. Pharmacol.* **2012**, *52*, 656.
17. Ostrem, J. M. L.; Shokat, K. M., *Nat. Rev. Drug Discov.* **2016**, *15*, 771.
18. Palumbo, A.; Anderson, K., *N. Engl. J. Med.* **2011**, *364*, 1046.
19. Manier, S.; Salem, K. Z.; Park, J.; Landau, D. A.; Getz, G.; Ghobrial, I. M., *Nat. Rev. Clin. Oncol.* **2017**, *14*, 100.
20. Walker, B. A.; Boyle, E. M.; Wardell, C. P.; Murison, A.; Begum, D. B.; Dahir, N. M.; Proszek, P. Z.; Johnson, D. C.; Kaiser, M. F.; Melchor, L.; Aronson, L. I.; Scales, M.; Pawlyn, C.; Mirabella, F.; Jones, J. R.; Brioli, A.; Mikulasova, A.; Cairns, D. A.; Gregory, W. M.; Quartilho, A.; Drayson, M. T.; Russell, N.; Cook, G.; Jackson, G. H.; Leleu, X.; Davies, F. E.; Morgan, G. J., *J. Clin. Oncol.* **2015**, *33*, 3911.
21. Xu, J.; Pfarr, N.; Endris, V.; Mai, E. K.; Md Hanafiah, N. H.; Lehnert, N.; Penzel, R.; Weichert, W.; Ho, A. D.; Schirmacher, P.; Goldschmidt, H.; Andrulis, M.; Raab, M. S., *Oncogenesis* **2017**, *6*, e337.

22. Hoang, B.; Frost, P.; Shi, Y.; Belanger, E.; Benavides, A.; Pezeshkpour, G.; Cappia, S.; Guglielmelli, T.; Gera, J.; Lichtenstein, A., *Blood* **2010**, *116*, 4560.
23. Demchenko, Y. N.; Kuehl, W. M., *Oncotarget* **2010**, *1*, 59.
24. Ramakrishnan, V.; D'Souza, A., *Curr. Hematol. Malig. Rep.* **2016**, *11*, 156.
25. Kumar, S. K.; Rajkumar, S. V.; Dispenzieri, A.; Lacy, M. Q.; Hayman, S. R.; Buadi, F. K.; Zeldenrust, S. R.; Dingli, D.; Russell, S. J.; Lust, J. A.; Greipp, P. R.; Kyle, R. A.; Gertz, M. A., *Blood* **2008**, *111*, 2516.
26. Rajan, A. M.; Kumar, S., *Blood Cancer J.* **2016**, *6*, e451.
27. Gentile, M.; Martino, M.; Recchia, A. G.; Vigna, E.; Morabito, L.; Morabito, F., *Expert Opin. Investig. Drugs* **2016**, *25*, 743.
28. Litwin, V.; Marder, P., *Flow cytometry in drug discovery and development*. John Wiley & Sons, **2011**.
29. <http://www.bdbiosciences.com/>
30. Spence, M. T. Z.; Johnson, I. D., *The molecular probes handbook : a guide to fluorescent probes and labeling technologies*. Live Technologies Corporation, **2010**.
31. Danial, N. N.; Korsmeyer, S. J., *Cell* **2004**, *116*, 205.
32. Adams, J. M.; Cory, S., *Oncogene* **2007**, *26*, 1324.
33. Zhao, W.; Zhang, T.; Qu, B. Q.; Wu, X. X.; Zhu, X.; Meng, F. Y.; Gu, Y. H.; Shu, Y. Q.; Shen, Y.; Sun, Y.; Xu, Q. A., *Anti-Cancer Drugs* **2011**, *22*, 79.
34. Fedorov, O.; Marsden, B.; Pogacic, V.; Rellos, P.; Muller, S.; Bullock, A. N.; Schwaller, J.; Sundstrom, M.; Knapp, S., *Proc. Natl. Acad. Sci.* **2007**, *104*, 20523.
35. Vedadi, M.; Niesen, F. H.; Allali-Hassani, A.; Fedorov, O. Y.; Finerty, P. J.; Wasney, G. A.; Yeung, R.; Arrowsmith, C.; Ball, L. J.; Berglund, H.; Hui, R.; Marsden, B. D.; Nordlund, P.; Sundstrom, M.; Weigelt, J.; Edwards, A. M., *Proc. Natl. Acad. Sci.* **2006**, *103*, 15835.
36. Bullock, A. N.; Debreczeni, J. É.; Fedorov, O. Y.; Nelson, A.; Marsden, B. D.; Knapp, S., *J. Med. Chem.* **2005**, *48*, 7604.
37. Qu, Y.; Han, B.; Yu, Y.; Yao, W.; Bose, S.; Karlan, B. Y.; Giuliano, A. E.; Cui, X., *PLOS ONE* **2015**, *10*, e0131285.
38. Scharf, R. E., *Semin. Thromb. Hemost.* **2012**, *38*, 865.
39. Ramot, Y.; Nyska, A.; Spectre, G., *Drug Safety* **2013**, *36*, 585.
40. Girolami, A.; Cosi, E.; Tasinato, V.; Santarossa, C.; Ferrari, S.; Girolami, B., *Clin. Appl. Thromb. Hemost.* **2016**, *1*.
41. Asvadi, P.; Ahmadi, Z.; Chong, B. H., *Blood* **2003**, *102*, 1670.

42. Visentin, G. P.; Liu, C. Y., *Hematol. Oncol. Clin. North Am.* **2007**, 21, 685.
43. Davis, A.; Ward, S. E., *The handbook of medicinal chemistry: principles and practice*. Royal Society of Chemistry, **2014**.
44. Tian, S.; Wang, J.; Li, Y.; Li, D.; Xu, L.; Hou, T., *Adv. Drug Deliv. Rev.* **2015**, 86, 2.
45. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., *Adv. Drug Deliv. Rev.* **2001**, 46, 3.
46. Daina, A.; Michielin, O.; Zoete, V., *Sci. Rep.* **2017**, 7.
47. *MarvinSketch 15.9.7 ChemAxon* **2015**.
48. Mitchell, J. B. O., *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, 4, 468.
49. O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R., *J. Cheminform.* **2011**, 3, 33.
50. Wu, P.; Nielsen, T. E.; Clausen, M. H., *Drug Discov. Today* **2016**, 21, 5.
51. Ertl, P.; Rohde, B.; Selzer, P., *J. Med. Chem.* **2000**, 43, 3714.
52. Daina, A.; Zoete, V., *ChemMedChem* **2016**, 11, 1117.
53. Cortes, C.; Vapnik, V., *Machine Learning* **1995**, 20, 273.
54. Montanari, F.; Ecker, G. F., *Adv. Drug Deliv. Rev.* **2015**, 86, 17.
55. Veber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D., *J. Med. Chem.* **2002**, 45, 2615.
56. Testa, B.; Krämer, S. D., *Chem. Biodivers.* **2008**, 5, 2171.
57. Di, L., *Expert Opin. Drug Metab. Toxicol.* **2014**, 10, 379.

Chapter 6

Chapter 6 – Conclusions and Future Work

6.1. Conclusions

During this project, computational modelling studies, such as MD and QM, have been performed to better understand the more likely mechanism of action of lead compound **1** and other proposed type III kinase inhibitors. Moreover, several novel derivatives were successfully synthesised and fully characterised. Additionally, several assays were conducted *in vitro*, involving the study of the effects of the compounds prepared on cell viability and proliferation in a wide variety of cancer cell lines, their apoptosis-inducing abilities, and the study of their molecular targets as potential protein kinase inhibitors. Particularly, 3-amino-4'-guanidino phenyloxypyridine **66** was selected as the new lead candidate given the improved overall biological activity and optimal pharmacokinetic profile. Based on the results obtained, some conclusions can be made.

6.1.1. Computational Modelling

A variety of modelling techniques have been utilised to understand and investigate the interactions between our derivatives and their biological target. Particularly, we have computationally studied and investigated the interactions between lead compound **1** and both an active wild type ATP-containing B-Raf protein kinase and ATP molecule.

Considering that no crystal structure is available for an ATP containing active B-Raf protein, thus hindering the possibility of structure-based drug design, a model of such a system has been computationally developed through molecular dynamics (MD). The final structure achieved represents the first model reported for a phosphorylated, active, ATP-containing B-Raf kinase and it would facilitate structure-based drug design of future allosteric inhibitors of this particular kinase. Next, the model has been used for molecular docking studies at semiempirical level of theory with our previously developed lead compound **1**. This study found a suitable docking pose that sees, among

others, the interaction of the ATP's γ -phosphate with the mono-substituted guanidinium of compound **1**, supporting previously hypothesis of type III allosteric inhibition.

Additionally, we have also theoretically studied more in-depth the interaction of 4-phenyloxyphenyl-*N*-guanidinium and –*O*-isouronium (as mimics of compound **1**) and ATP molecule by means of DFT calculations. Both amidinium derivatives are able to form strong HB interactions not only with the O atoms of an isolated phosphate anion, but also with the different phosphate groups of the ATP molecule, where parallel HB interactions were found. Binding energies and AIM analysis found the complexes formed between ATP γ -phosphate and both guanidinium and isouronium to be the most stable ones. The formation of these interactions and their characterisation, indicate the suitability of guanidinium (or isouronium) derivatives as potential type III kinase's inhibitors.

6.1.2. Synthesis

Based on our previous computational studies, different changes were introduced to formerly synthesised lead compound **1**, to systematically investigate the role of its three identified moieties: i) hydrophobic moiety; ii) link (*meta* guanidinium), which controls overall charge and length of the molecule; and iii) polar moiety (*para* guanidinium).

Synthesis of substituted thioureas **11-16** carrying different lipophilic moieties consisting in diverse halogens substitution and bulky trifluoromethyl substituent, were prepared according the well-established procedure developed in Rozas' group,¹ and, next, following the procedure described by Kim and Qian and developed by Rozas' group (HgCl₂ and trimethylamine),² they were utilised to obtain 3,4'-bis-guanidino phenyloxybenzene derivatives **2-8** and 3,4'-bis-guanidino phenoxy pyridines **61-65**.

Amino-linked diaromatic mono-guanidines (**27-30** and **66**) and amido-linked diaromatic mono-guanidines (**32-33**) were synthesised through Buchwald-Hartwig coupling and acyl chloride and amine reaction, respectively, to investigate the effect of different links in the orientation of a variety of lipophilic moieties in lead compound **1**.

When exploring the polar moiety, same conditions were used for the synthesis of isouronium derivatives (**67-69**), which required the preliminary preparation of the

conveniently functionalised starting phenol. Moreover, the synthesis of sulfonamide **70**, required the reaction of amine **48** with sulfonamide chloride **103**, previously prepared from chlorosulfonyl isocyanate and formic acid.

All final Boc protected guanidines or isoureas derivatives were synthesised through the methodology previously described by Kim and Qian, and posteriorly explored by Rozas and co-workers, using mercury (II) chloride, triethylamine and *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methylisothiurea **51**. All final Boc-protected compounds were deprotected using HCl 4M/dioxane to obtain compounds **2-8**, **27-31**, **61-66**, TFA/CH₂Cl₂ to obtain compound **67**, and SnCl₄ to obtain compounds **32-33** and **68-69**.

The signals in the ¹H and ¹³C NMR of Boc-protected 3,4'-bis-guanidino phenyloxybenzene and 3,4'-bis-guanidino phenyloxypyridines derivatives, were highly broad and amorphous due to the presence of isomerism of these tri-substituted guanidine derivatives. Significantly, when these Boc groups were removed, this isomerism disappeared and the ¹H and ¹³C NMR spectra of the hydrochloride salts could be resolved.

The purity of all of the final hydrochloride salts was determined by HPLC, where a minimum purity of 95% was required before proceeding to biological testing.

All modifications introduced to lead compound **1** generated SAR information which were used to understand in more detail the structural requirements for a more potent binding of the compounds to the biological target.

6.1.3. Biochemistry and physicochemical properties

A screening of the cytotoxicity of all compounds was performed in the promyelocytic leukaemia HL-60 cell line and the compounds with more potent activity in these assays were selected for further biological evaluation in several other tumour cell lines, such as MCF-7 (breast), HeLa (cervical), HCT116 and HKH-2 (colorectal) and NCI929 and U266B1 (multiple myeloma). The corresponding cell viability studies revealed that our compounds are able to inhibit cell proliferation of these cancer cell lines in the low μM range of concentration, showing up to a nine-fold increase in cytotoxicity compared to lead compound **1**. Thus, it can be concluded that our compounds are able to inhibit cell

proliferation of leukaemia, breast cancer, cervical cancer, colorectal cancer and multiple myeloma.

Several conclusions can be achieved regarding SAR information from the evaluation of the cytotoxicity of all compounds in HL-60 cell line. In the *hydrophobic moiety*, bulky and lipophilic substituents are crucial to achieve high cytotoxic activity. Regarding the *polar moiety*, positively charged guanidinium is essential for activity and it shows higher cytotoxicity when in the 2 position of a pyridine ring with which it can form an intra-molecular hydrogen bond. Moreover, the isouronium cation results in similar cytotoxicity as guanidinium, but with improved physicochemical properties. Finally, in terms of the *link* shorter molecules (amino-linked) increased the cytotoxicity of these derivatives.

Flow cytometry analysis (Annexin V analysis and PI staining) and Mcl-1 inhibition in NCI-H929 and U266B1 cell line were performed to investigate the pro-apoptotic ability of our new lead compound **66**. Compound **66** is able to induce early apoptotic events only after 24 h incubation and at a low concentration of 10 μ M. Cytotoxicity of compound **66** was also evaluated against human platelets and non-tumorigenic MCF10A and the results show that, at active concentration, compound **66** can cause antiproliferative effects in tumorigenic cancer cell lines without any toxicity to non-tumorigenic cells or platelets.

In silico physicochemical analysis of the properties of all our derivatives helped us in correlating optimal SAR information with drug-likeness requirements and even though the exact molecular target(s) of our molecules still needs to be confirmed, the biological results obtained are very encouraging and, in particular, compound **66** can be considered an excellent *hit* molecule in searching for new protein kinases' inhibitors.

6.2. Future Work

Throughout this project, some very interesting and promising results were obtained, which deserves further investigation.

6.2.1. Future Computational Modelling

The model of a phosphorylated ATP-containing active B-Raf protein kinase that has been theoretically obtained through MD simulations is a useful tool that can be used for further structure-based drug design studies with new derivatives. Moreover, the docking pose that was obtained using such model and compound **1**, deserves further investigation through MD simulations to assess the stability of such pose and its binding energy.

6.2.2. Future Protein Kinase Inhibitors

Considering the SAR conclusions found after the modelling, synthesis and biochemical evaluation of the compounds performed in this work and considering that compound **66** was selected to be the new lead compound, several ideas can be proposed to try to improve even further its potency.

During this work we identified how bulkier and more lipophilic substituents at the hydrophobic moiety, are responsible for a better cytotoxicity activity. Mortenson *et al.* have published an interesting paper in 2011 to shed a light of awareness of molecular efficiency.³ In their publication, they present a new ligand lipophilicity index (LLE_{AT}) that enables to estimate whether or not an initial *hit* is likely to be optimisable to a potent, drug-like *lead* combining lipophilicity, size and potency. To maintain the threshold value for LLE_{AT} of 0.3 kcal mol⁻¹ per heavy atom, the required fold improvement in binding energy for a range of common substituents will be different (Figure 6.2.1.). The green box of Figure 6.2.1. shows the best balance of required potency improvements and lipophilicity for common aromatic substituents.

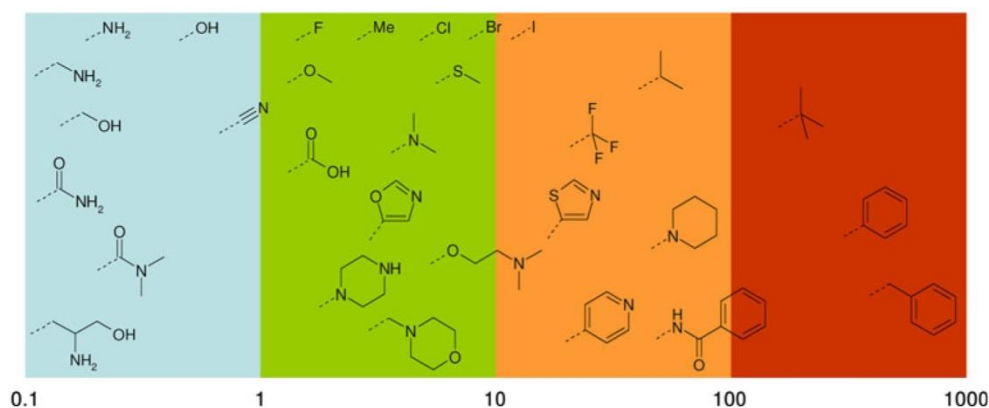


Figure 6.2.1. Fold improvements in potency required to maintain an LLE_{AT} of 0.3 for some common substituents, when attached to an aromatic carbon.³

Consequently, the (3- CF_3 -4-Cl)-Ph substituent of compound **66** can be substituted with (3- CF_3 -4-Br)-Ph, which bromine lies in the green box of Figure 6.2.1. and has already been proven to increase the activity of 3,4'-bis-guanidino phenyloxybenzenes **5,6** and **8** and 3,4'-bis-guanidino phenoxypyridines **63** and **64** (Figure 6.2.2.). Further hypothetic modifications could consist in the introduction of morpholine, oxazole or thiazole substituents (Figure 6.2.2.).

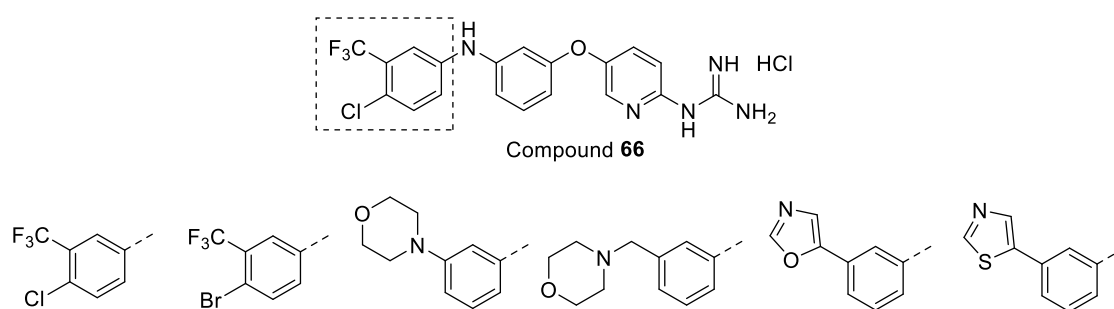


Figure 6.2.2. Suggested future compounds with modified *lipophilic moiety*.

Better cytotoxic activity of compound **66** and 3,4'-bis-guanidino phenoxypyridines **61-65** is postulated to be caused by intramolecular hydrogen bonding between the N of the pyridine ring and the positively charged guanidinium in position 2 of the ring, forcing the cationic moiety in a more specific conformation. Then, the synthesis of

conformationally restricted derivatives and its effect on cancer cell cytotoxicity could be an interesting feature to investigate (Figure 6.2.3.).

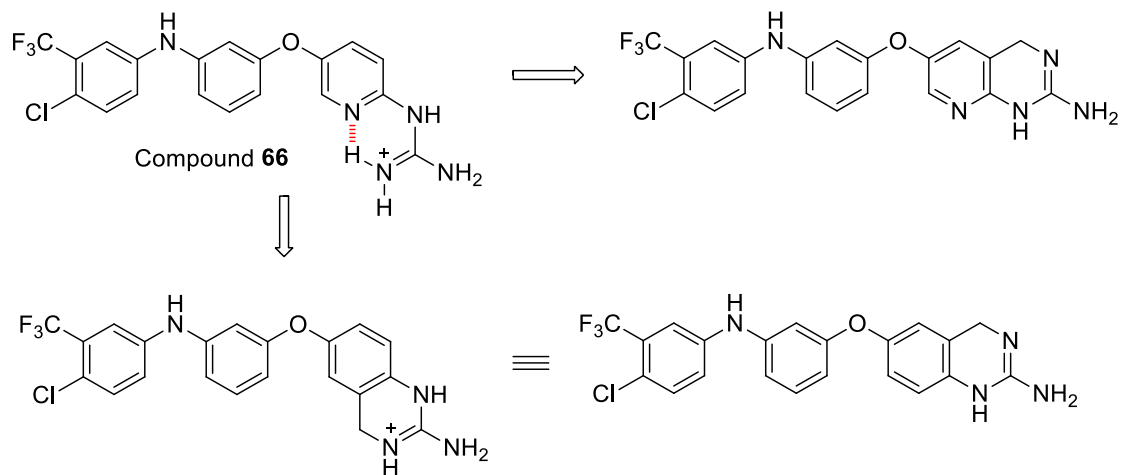


Figure 6.2.3. Suggested future compounds with modified *polar moiety*.

6.2.3. Future Biological Studies of our Compounds

Compound **66** was identified as the best compound among all the synthesised derivatives. Even though its cytotoxic activity is in the low micromolar range against all the studied cancer cell lines, its biological target has still to be identified. Compound **66** is a derivative of compound **1**, which was proven to be a B-Raf allosteric inhibitor. Even though such a behaviour is likely to be reproduced for compound **66**, this still has to be confirmed. The choice of the right cell line for such investigation is of extreme importance. For this reason, we have an ongoing collaboration with Prof. Walter Kolch's group (UCD, Ireland) to investigate the cell signalling pathways' inhibition pattern of compound **66** in the RKO T29 (B-Raf wild type) cancer cell line, a colorectal cancer cell line that was previously and positively used to investigate compound **1**. This investigation will be conducted through an enzyme-linked immunosorbent assay (ELISA) that allows to detect the presence of an antibody or an antigen in a cellular sample.

Compound **66** was found to induce apoptosis through the inhibition of anti-apoptotic protein Mcl-1 in multiple myeloma cell line. Signal transducer and activator of

transcription 3 (STAT3) is a transcription factor that controls, among others, the production of Mcl-1. Then, in collaboration with Prof. Daniela Zisterer's group (School of Biochemistry and Immunology, TCD, Ireland), we will investigate the effect of compound **66** against STAT3 and more in general the JAK/STAT pathway and its interconnections.

Compound **66** and compound **1** were proven to be safe in preliminary studies against MCF10A cells and human platelets. Besides further investigation in normal cell lines, the study of the effect of compound **66** in Tumour Cell-Induced Platelet Aggregation (TCIPA) should be considered. The formation of tumour cell-platelet aggregates facilitates metastasis and increased tumour survival.⁴ While a compound's effect on platelets is important to help establishing its toxic profile, investigating any potential application of a compound in the inhibition of TCIPA is of vast therapeutic interest. Then, in a collaboration with Prof. Maria Jose Santos-Martinez (School of Pharmacy, TCD, Ireland), we will continue investigating the effect of compound **66** and of other new compounds in the inhibition of TCIPA.

6.3. References

1. Diez-Cecilia, E.; Kelly, B.; Perez, C.; Zisterer, D. M.; Nevin, D. K.; Lloyd, D. G.; Rozas, I., *Eur. J. Med. Chem.* **2014**, *81*, 427.
2. Kim, K. S.; Qian, L., *Tetrahedron Lett.* **1993**, *34*, 7677.
3. Mortenson, P. N.; Murray, C. W., *J. Comput. Aided Mol. Des.* **2011**, *25*, 663.
4. Medina, C.; Jurasz, P.; Santos-Martinez, M. J.; Jeong, S. S.; Mitsky, T.; Chen, R.; Radomski, M. W., *J. Pharmacol. Exp. Ther.* **2006**, *317*, 739.

Chapter 7

Chapter 7 - Experimental

7.1. Molecular Modelling

7.1.1. Modelling of the target and ligand-target interaction study

7.1.1.1. *Starting B-Raf structures preparation*

Initial conformations for the molecular dynamics (MD) simulations were based on the PDB structure 3Q4C corresponding to the wild type B-Raf. The initial coordinates of the missing residues (601-612) were added using the program MODELLER 9.18.^{1,2} Next, we added an ATP molecule in the nucleotide-binding pocket of that B-Raf model. For ATP to fit in this 3D structure, we performed the superposition of the original B-Raf model on the structure of a very similar protein kinase containing ATP (PDB entry: 1ATP). Although 1ATP and 3Q4C have a low sequence homology (17%), the substituted structural homology is much larger (57%) facilitating the positioning of the ATP molecule inside the nucleotide pocket of 3Q4C by superimposing both structures.

7.1.1.2. *Molecular Dynamics simulation*

The preparation of the protein was achieved using the same protocol as in a previously reported study.³ We employed the program xLeap within the AMBER11 package⁴ to add the hydrogen atoms to the protein structure. The ff99SB force field was used for the protein. AMBER force field parameters to describe the phosphorylated state of threonine Thr599 (Thp599) were taken from Homeyer et al.⁵ Concerning the ATP molecule, force field and atomic charges parameters to be utilised with AMBER were taken from the R.E.DD.B repository.⁶ Each of these systems was fully explicitly solvated in a truncated octahedral box using TIP3P water molecules with a buffer distance of 7.0 Å and under periodic boundary conditions. Counterions (Cl⁻) were added to neutralize the unit system. The entire system consists of about 25,000 atoms (depending on the presence or not of ATP and/or threonine phosphorylation).

The MD simulations employed a 2 fs integration time step. The programs *sander* (preparation) and *pmemd* (for the production stage) of the AMBER11 package were employed. The particle mesh Ewald procedure was utilised to handle long-range electrostatic interactions. The value of 9.0 Å for the non-bonded cut-off was set to calculate van der Waals (vdW) and electrostatic energies. No switching function was used for vdW interactions. The system was then carefully prepared by first minimizing the energy of the entire system with 8,000 cycles of steepest descent followed by 500 cycles of conjugate gradient minimization. This process was repeated twice. Classical Langevin NVT simulations were then carried out with collision frequency of 2 ps⁻¹ and bonds involving hydrogen atoms were constrained. Firstly, the water molecules were heated from 0 to 300 K during 20 ps and then equilibrated for 20 ps. A force constant of 100 kcal mol⁻¹ Å⁻² was applied to restrain all other atoms during these steps. After cooling the solvent molecules to 0 K during 20 ps, this heating-equilibrating-cooling procedure was run for a second time for the entire system with no constraint at all. After a last heating-equilibrating process (20 + 20 ps), prior to production, our system was equilibrated using the NPT ensemble (isothermal-isobaric ensemble) at 300 K and 1 atm for 40 ps. A production final stage followed with 800 ns of data collection. Coordinates of the system were written every 1 ps.

For each of the four B-Raf models, to measure protein conformational changes during the simulation, a root mean squared deviation (RMSD) was computed between each successive trajectory structure (backbone selection) and the first structure taken as reference, and plotted as a function of time. In addition, the root mean square fluctuation (RMSF) of the protein was computed as a measure of the average atomic mobility and calculated on a residue basis using the fluctuations of the positions of the C α atoms of the protein during the last 200 ns of each simulation.

7.1.1.3. 1D Free Energy Landscapes (1D-FEL)

In order to analyse the influence of ATP-binding and Thr phosphorylation on the local conformational changes of the B-Raf protein backbone, we constructed Free Energy Landscapes (FEL)^{7,8} along the so-called Coarse-Grained Dihedral Angles (CGDA). Each CGDA is formed by 4 bonds joining 4 successive C α atoms along the main chain of the B-Raf protein and acts as a local probe of the FEL along the primary sequence.

The 276-residues of the B-Raf model have 273 CGDA (θ_i). From each of our 4 simulations we extracted the “trajectory” of each CGDA, $\theta_i(t)$. The 1D effective free-energy potential for each θ_i is then defined by the logarithmic relation: $\Delta G = -RT \ln P(\theta_i)$, (at $T = 300$ K) using the one-dimensional probability distribution function $P(\theta_i)$ derived from the simulations (R being the gas constant). It is to be noticed that this analysis was carried out on the last 200 ns of each trajectory. In order to identify which residues of the protein are influenced by phosphorylation and ATP binding, we compared each 1D-FEL $G(\theta_i)$ between trajectory pairs thanks to a similarity index that varies from 1 (similar) to 0 (dissimilar).⁹

7.1.1.4. NCI calculations

Calculations have been performed to describe non-covalent interactions (NCI) in ligand-protein complexes. The NCI method is based on the knowledge of the electron density (ρ) and its derivatives (gradient) to enable the identification of NCI in real space based on the peaks that appear at low ρ values in the reduced density gradient (RDG) plot $s(\rho)$ (see reference NCI¹⁰ for detailed information). RDG isosurfaces provide a 3D representation of interaction regions. The sign of the second eigenvalue of the ρ Hessian matrix is used to differentiate repulsive ($\lambda_2 > 0$) from attractive ($\lambda_2 < 0$) interactions. In the NCI pictures presented in this research, the nature of the interaction is colour-coded: blue for attractive interactions (e.g. HB), green for weakly repulsive attractive (e.g. vdW interactions) or red for repulsive situations (e.g. steric repulsion). The GPU version (cuNCI)¹¹ of the original program NCIPLOT¹² was employed throughout this study.

As a complement to this analysis, we performed IGM calculations,¹³ which utilize a new description (δ_g) that confers a quantitative aspect to the initial NCI approach. A very attractive feature of the IGM model is to provide an uncoupling scheme that automatically extracts interactions present between user selected fragments. In addition to the standard inter-fragment uncoupling scheme, an atomic decomposition scheme was used in the present study to estimate the relative contribution of atoms to the NCI regions and, accordingly, the program IGMPlot was employed (<http://www.lct.jussieu.fr/pagesperso/contrera/jcg-programs.html>). Furthermore, the

program LIGPLOT,¹⁴ which is a program to generate schematic diagrams of protein-ligand interactions, was used to obtain 2D schematic diagrams for the interactions between B-Raf and ATP.

7.1.1.5. Molecular Docking

Because strong electrostatic interactions are expected between the positively charged ligand (lead compound **1**) and the negatively charged ATP, the novel software AlgoGen¹⁵ was employed instead of standard molecular docking software like Autodock/Vina (based on molecular mechanics methods). AlgoGen combines a classical mechanics level of theory and quantum chemistry (QM): after exploring the best orientations of a ligand binding to a host, the best poses (as per classical mechanics methods) are automatically fully refined at the QM level of theory (i.e. semiempirical method PM7¹⁶ combined with the MOZYME linear-scaling algorithm¹⁷ proposed in the MOPAC2016 program¹⁸). First, a QM rigid relaxation is performed; then, on the best pose for B-Raf binding, the geometry is refined by relaxing all valence and dihedral angles of the ligand (flexible docking) as well as the protein.

7.1.2. Quantum Chemistry studies

7.1.2.1. Optimized geometries and binding energies

Regarding the ATP molecule, the starting conformation utilized in this study is that found for this molecule within the MEK-1 kinase crystal structure (3PP1). Moreover, the ionic state for the molecule of ATP (ATP^{4-}) is chosen considering that the pK_a of ATP is 6.5,¹⁹ at physiological pH (7.4) and, hence, almost 90% of ATP would be in this ionic form.

For the phosphate-guanidinium ligand dimers, the geometry of the monomers and the complexes was fully optimized at second-order Møller–Plesset (MP2)²⁰ level with 6-311+G(d,p) basis set.²¹ For the ATP-guanidinium ligand dimers, the geometry of the monomers and the complexes was fully optimized at DFT level (ω B97XD functional) with the 6-31+G(d,p) basis set and the binding energies were obtained from a single point calculation using MP2 with the 6-311+G(d,p) basis set. All of these calculations

were carried out using the Gaussian09 program.²² Frequency calculations were performed at the same computational level to confirm that the resulting optimized structures (monomers and dimers) were energetic minima. Binding (interaction) energies, E_i , were calculated as the difference between the total energy of the complex and the sum of the total energy of the isolated monomers. All the calculations were carried out in a solvent model (water) in order to obtain more realistic guanidinium binding energies because it is well known that these energies are over estimated when calculated in vacuum and that this is corrected in aqueous environment.^{23,24} The effect of water solvation was then accounted using the SCRF-PCM (Polarizable Continuum Model) approach implemented in the Gaussian09 package including dispersing, repulsing and cavitation energy terms of the solvent in the optimization. Considering that all the calculations were performed using PCM (which locates the molecules within a cavity) to model water solvent effects and that the inherent basis set superposition error (BSSE) cannot be calculated in this model, the interaction energy values were not BSSE corrected.

7.1.2.2. Bonding characteristics and electron density

The characteristics of the interactions established in the different complexes studied were analysed. Thus an accurate quantum chemical topology (QCT) technique such as the Atoms in Molecules (AIM) theory²⁵, was employed and by means of the AIMAll program²⁶ the most relevant bond critical points (BCP) were located and their electron density evaluated. Regarding the arylguanidinium: phosphate complexes, the natural bond orbital (NBO) theory was employed to evaluate atomic charges using the NBO-3 program.²⁷

7.2. Synthetic Chemistry

7.2.1. Materials and Methods

Reagents were purchased at the highest commercial quality and used without further purification, unless otherwise stated. Dry solvents were obtained from PureSolvTM

solvent purification system (Innovative Technology, Inc.) or otherwise purchased from Sigma-Aldrich. Solvents for synthesis purposes were used at GPR grade. Yields refer to chromatographically and spectroscopically (^1H NMR) homogeneous materials, unless otherwise stated. Reactions were monitored by thin layer chromatography (TLC) using Merck Kieselgel 60 F254 silica gel or aluminium oxide plates. Visualisation was achieved by UV light (254 nm) and basic aqueous potassium permanganate (KMnO_4) or ninhydrin and heat as developing agents. Concentration of organic solvents was performed on a rotary evaporator under reduced pressure followed by further evacuation using a two-stage mechanical pump.

Deuterated solvents for NMR use were purchased from Apollo and Euriso-top. Chromatographic columns were run using silica gel 60 (230–400 mesh ASTM). NMR spectra were recorded on Bruker DPX-600 and Bruker DPX-400 instruments operating at 400.13 MHz and 600.1 MHz for ^1H NMR and at 100.6 MHz and 150.9 MHz for ^{13}C NMR and they were calibrated using residual undeuterated solvent as an internal reference (CHCl_3 at δ 7.26 and 77.16 ppm for ^1H NMR and ^{13}C NMR, respectively; CH_3OH at δ 3.31 and 49.00 ppm for ^1H NMR and ^{13}C NMR, respectively; DMSO at δ 2.50 and 39.52 ppm for ^1H NMR and ^{13}C NMR, respectively). NMR data were processed using MestreNova software. The following abbreviations (or combinations thereof) were used in NMR characterisation: s = singlet, d = doublet, t = triplet, q = quartet, p = pentaplet, m = multiplet, bs = broad singlet, H = proton, CH Ar, aromatic carbon, qC = quaternary carbon. Compounds are assigned arbitrary numbers for assignment of the NMR spectrum and these numbers do not correspond to the naming of the compound.

High-resolution mass spectra (HRMS) were measured on a Micromass LCT electrospray TOF instrument with a WATERS 2690 autosampler and HPLC grade methanol as carrier solvent. Melting points were determined using a Stuart Scientific Melting Point SMP1 apparatus and are uncorrected. Infrared spectra were recorded on a Perkin Elmer Spectrum One FT-IR Spectrometer equipped with a Universal ATR sampling accessory.

7.2.2. Purity Assessment of Hydrochloride Salts

HPLC purity analysis was carried out using a Varian ProStar system equipped with a Varian Prostar 335 diode array detector and a manual injector (20 μ L). For purity assessment, UV detection was performed at 254 nm and peak purity was confirmed using a purity channel. The stationary phase consisted of an ACE 5 C18-AR column (150 $\times$ 4.6 mm), and the mobile phase used the following gradient system, eluting at 1 mL/min: aqueous formate buffer (30 mM, pH 3.0) for 10 min, linear ramp to 85% methanol buffered with the same system over 25 min, hold at 85% buffered methanol for 10 min. Minimum requirement for purity was set at 95.0%.

7.2.3. General Procedures

Method A: *General procedure for the deprotection of N-Boc protected guanidines using hydrochloric acid.*²⁸

To a solution of the corresponding Boc protected guanidine (1 eq.) in 1,4-dioxane, 4 M HCl in 1,4-dioxane (6 eq. per Boc group) was added to reach a guanidine final concentration of 0.2 M. The mixture was stirred at 55 °C until the reaction was complete (typically 6-8 h, adjudged by TLC). At the reaction endpoint, solvent and HCl were evaporated under vacuum and the crude salt was dissolved in the minimum volume of CHCl_3 . It was then purified by flash chromatography (silica gel, CHCl_3 :MeOH). The purified fractions were evaporated until dry to afford the pure hydrochloride salt.

Method B: *General procedure for the synthesis of N-aryl-N'-Boc-protected thiourea derivatives.*²⁸

To a solution of *N,N'*-bis-(*tert*-butoxycarbonyl)thiourea (1 eq.) in dry tetrahydrofuran (0.12 M) under argon at 0 °C, sodium hydride as a 60% suspension in mineral oil (1.5 eq.) were added. The reaction mixture was stirred at the same temperature for 1 h and

30 min, then trifluoroacetic anhydride (1.1 eq.) was added and the stirring continued for an additional 30 min. The corresponding aniline (1.15 eq.) was added and the reaction was stirred at 0 °C for 5 h. The reaction was cooled again to 0 °C and dropwise H₂O was added to quench the reaction, followed by extraction with EtOAc. The combined organic phases were washed with 80% brine and dried using anhydrous MgSO₄, followed by removal of solvents under vacuum. The compounds were purified by silica gel chromatography and/or recrystallization (as specified for each compound) to afford the desired product.

Method C: *General procedure for the synthesis of the asymmetric Boc-protected 3,4'-bis-guanidine derivatives.*²⁸

Mercury(II) chloride (HgCl₂) (1.2 eq.) was added to a solution of the corresponding mono Boc-protected guanidine (1 eq.), the corresponding thiourea (1.2 eq.) and triethylamine (3.1 eq.) in CH₂Cl₂ (0.19 M). The resulting mixture was stirred at 0 °C for 1 h and then at room temperature for 6-8 h or overnight depending (as specified for each compound). The crude was then diluted with EtOAc and filtered through a pad of Celite® in order to remove the mercury sulfide precipitate formed and the filter cake was rinsed with EtOAc. The organic phase was extracted with water, washed with brine, dried over anhydrous MgSO₄, and concentrated under vacuum to give a crude residue that was purified by flash chromatography on silica gel (eluting with a gradient of hexane:EtOAc).

Method D: *General procedure for N-Arylation.*²⁹

An oven-dried round bottom flask was charged with a magnetic stir bar, Pd₂(dba)₃ (3 mol%), organophosphorus ligand (3-5 mol%), aniline derivative (1 eq.), base (1.4 eq.) and bromine source (1 eq.). The flask was evacuated and refilled with argon (three times). In case of liquid bromine source, it was added by syringe under a counterflow of argon, followed by syringe addition of toluene (2 mL/mmol). The mixture was placed in a preheated oil bath at 90 °C and let it stirring for 24 h. The reaction mixture was then cooled to room temperature, diluted with EtOAc, filtered through a pad of Celite®, and washed with water. The organic layer was washed with brine, dried over MgSO₄,

concentrated under vacuum and purified by silica gel chromatography (hexanes:EtOAc) to obtain the desired product.

Method E: *General procedure for nitroreduction using Palladium on carbon (Pd/C).*

To a solution of appropriate nitro derivative (1 eq.) in specified solvent, was added 10% Pd/C mixture (10 mol%), and the reaction mixture was stirred at room temperature under hydrogen atmosphere (1.0 atm) for 24 h. The reaction mixture was filtered off, and the filtrate was concentrated in vacuo. The residue was purified by silica gel chromatography (hexanes:EtOAc), and the desired fractions were concentrated under reduced pressure to obtain the desired product.

Method F: *General procedure for nitroreduction using $\text{SnCl}_2 \cdot 2\text{H}_2\text{O}$.*³⁰

A mixture of appropriate nitro derivative (1 eq.) dissolved in absolute ethanol and $\text{SnCl}_2 \cdot 2\text{H}_2\text{O}$ (6 eq.) was stirred at 70 °C until the reaction was adjudged complete (typically 3 h, adjudged by TLC as specified for each compound). The solution was allowed to cool down and then poured into ice. The pH was made slightly basic (pH 7-8) by addition of 5% aqueous NaHCO_3 . The solution was washed with water, the organic layer extracted with EtOAc, washed with brine, dried over MgSO_4 , concentrated under vacuum and purified by silica gel chromatography (hexanes:EtOAc) to obtain the desired product.

Method G: *General procedure for the synthesis of 1-aryl-2,3-di(tert-butoxycarbonyl)guanidines derivatives.*³¹

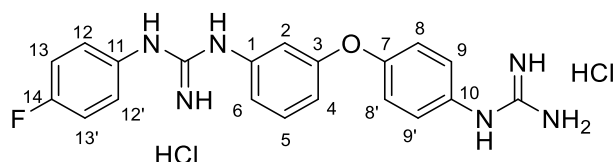
Mercury(II) chloride (HgCl_2) (1.04 eq.) was added to a solution of the corresponding aromatic amine (1 eq.), *N,N'*-bis-(tert-butoxycarbonyl)-*S*-methyliothiourea (1.04 eq.) and triethylamine (NEt_3) (3.5 eq.) in CH_2Cl_2 (0.19 M). The resulting mixture was stirred at 0 °C for 1 h and then 24 h or 48 h (as specified for each compound) at room temperature (reaction progress adjudged by TLC). The crude was then diluted with EtOAc and filtered through a pad of Celite[®] in order to remove the mercury sulfide

precipitate formed. The filter cake was rinsed with EtOAc. The organic phase was extracted with water, washed with brine, dried over anhydrous MgSO_4 , and concentrated under vacuum to give a residue that was purified by flash chromatography on silica gel as specified.

Method H: *General procedure for the synthesis of arylamido derivatives.*³²

The appropriate benzoic acid derivative (1.2 eq.) was placed under nitrogen atmosphere and dissolved in a 0.3 M solution of dry dichloromethane (CH_2Cl_2). Oxalyl chloride (2 eq.) was added, followed by dimethylformamide (DMF) (3 drops approx.). The solution was stirred overnight. Any remaining oxalyl chloride and DMF was removed under vacuum to obtain a yellow residue that was dissolved in CH_2Cl_2 (0.6 M) and cooled to 0 °C. Corresponding aniline (1 eq.) and triethylamine (NEt_3) (2.4 eq.) was also dissolved in CH_2Cl_2 and added dropwise to the cooled solution. After 12 h the solution was warmed to room temperature. The organic layer was extracted using water, washed with brine, dried over MgSO_4 and concentrated under vacuum to give a residue that was purified by silica gel chromatography (eluting with a gradient of hexane:EtOAc) to yield final compound.

7.2.4. Synthesis and Characterisation

1-(4-Fluorophenyl)-2-(3-(4-guanidinophenoxy)phenyl)guanidine dihydrochloride (2)

Following Method A, **20** (80 mg, 0.12 mmol) was dissolved in 4M HCl in dioxane (0.54 mL, 2.16 mmol) and in additional dioxane (0.07 mL) until a final concentration of 0.2 M was reached. After 8 h stirring at 55 °C, the reaction was adjudged complete (TLC), solvents were evaporated and the residue was purified by silica gel (CH₃Cl:MeOH) chromatography to afford the pure hydrochloride salt a white-yellow solid (35 mg, 76%). **Mp**: decomp. > 180 °C.

δ_{H} (400 MHz, CD₃OD): 7.00 (dd, J = 8.3, 2.4, 1H, H-4), 7.04 (t, J = 2.1 Hz, 1H, H-2), 7.13 - 7.17 (m, 3H, H-8 and H-8' and H-6), 7.19 – 7.24 (m, 2H, H-13 and H-13'), 7.31 (d, J = 8.9 Hz, 2H, H-9 and H-9'), 7.36 – 7.39 (m, 2H, H-12 and H-12'), 7.47 (t, J = 8.1 Hz, 1H, H-5).

δ_{C} (100 MHz, CD₃OD): 116.5 (CH Ar, C-2), 117.8 (d, J = 23.3 Hz, 2 CH Ar, C-13 and C-13'), 118.6 (CH Ar, C-4), 121.1 (CH Ar, C-6), 121.5 (2 CH Ar, C-8 and C-8'), 129.95 (d, J = 8.8 Hz, 2 CH Ar, C-12 and C-12'), 128.96 (2 CH Ar, C-9 and C-9'), 131.6 (qC), 132.3 (d, J = 3.1 Hz, qC, C-11), 132.4 (CH Ar, C-5), 138.0 (qC), 156.6 (qC), 157.4 (qC), 158.4 (qC), 159.5 (qC), 163.2 (d, J = 246.1 Hz, qC, C-14).

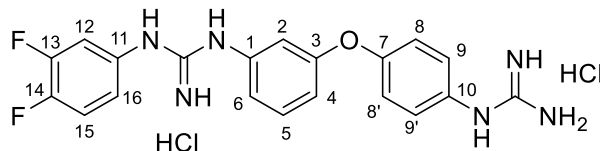
δ_{F} (376 MHz, CD₃OD): – 115.93 (m).

HRMS (m/z ESI⁺): Found: 379.1687 (M⁺ + H), C₂₀H₂₀N₆OF Requires: 379.1683.

ν_{max} (ATR)/cm⁻¹: 3110 (NH), 3052 (NH), 2922, 2330, 2134, 1655 (C=N), 1582 (C=N), 1505 (C-N), 1486, 1404, 1212 (C-O), 1066 (C-F), 834, 792, 552.

HPLC: 99.7% (t_R: 22.9 min).

1-(3,4-Di-fluorophenyl)-2-(3-(4-guanidinophenoxy)phenyl)guanidine dihydrochloride (3)



Following Method A, **21** (53 mg, 0.08 mmol) was dissolved in 4M HCl in dioxane (0.36 mL, 1.37 mmol) and in additional dioxane (0.06 mL) until a final concentration of 0.2M was reached. After 8 h stirring at 55 °C, the reaction was adjudged complete (TLC), solvents were evaporated and the residue was purified by silica gel chromatography (CHCl₃:MeOH) to afford the pure hydrochloride salt a white solid (34 mg, 90%). **Mp**: 158-160 °C

δ_{H} (400 MHz, CD₃OD): 7.00 (dd, J = 8.3, 2.4 Hz, 1H, H-4), 7.05 (t, J = 2.2 Hz, 1H, H-2), 7.13 – 7.20 (m, 4H, H-8 and H-8', H-6 and H-16), 7.32 (d, J = 8.9 Hz, 2H, H-9 and H-9'), 7.34 – 7.41 (m, 2H, H-12 and H-15), 7.47 (t, J = 8.1 Hz, 1H, H-5).

δ_{C} (100 MHz, CD₃OD): 116.3 (d, J = 19.7 Hz, C-12 or C-15), 116.5 (CH Ar, C-2), 118.6 (CH Ar, C-4), 119.5 (d, J = 18.8 Hz, C-12 or C-15), 121.1 (CH Ar, C-6), 121.5 (2 CH Ar, C-8 and C-8'), 123.4 (dd, J = 6.7, 3.7 Hz, C-16), 129.0 (2 CH Ar, C-9 and C-9'), 131.6 (qC), 132.4 (CH Ar, C-5), 133.0 (dd, J = 8.3, 3.6 Hz, qC, C-11), 137.9 (qC), 150.8 (dd, J = 247.8, 12.6 Hz, qC, C-13 or C-14), 151.8 (dd, J = 248.7, 13.7 Hz, qC, C-13 or C-14), 156.5 (qC), 157.4 (qC), 158.3 (qC), 159.5 (qC).

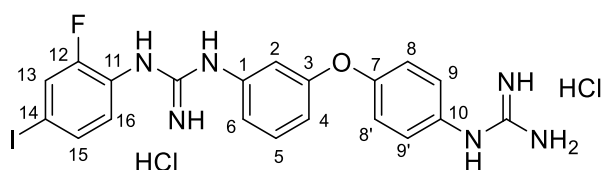
δ_{F} (376 MHz, CD₃OD): - 137.23 (m), - 141.12 (m).

HRMS (m/z ESI⁺): Found: 397.1598 (M⁺ + H), C₂₀H₁₉N₆OF₂ Requires: 397.1588.

ν_{max} (ATR)/cm⁻¹: 3228 (NH), 3040 (NH), 2923, 2853, 1655 (C=N), 1579 (C=N), 1505, 1485, 1401, 1259, 1211 (C-F), 1149, 972, 825, 771, 694, 649, 609, 587.

HPLC: 96.8% (tR: 23.2 min).

1-(2-Fluoro-4-iodophenyl)-2-(3-(4-guanidinophenoxy)phenyl)guanidine dihydrochloride (4)



Following Method A, **22** (357 mg, 0.44 mmol) was dissolved in 4M HCl in dioxane (2 mL, 7.92 mmol) and in additional dioxane (0.2 mL) until a final concentration of 0.2M was reached. After 8 h stirring at 55 °C, the reaction was adjudged complete (TLC), solvents were evaporated and the residue was purified by silica gel chromatography (CH₃Cl:MeOH) to afford the pure hydrochloride salt as a white solid (210 mg, 83%).
Mp: decomp. > 180 °C.

δ_H (400 MHz, CD₃OD): 6.99 – 7.01 (m, 2H, H-2 and H-4), 7.11 – 7.14 (m, 1H, H-6), 7.15 (d, *J* = 8.9 Hz, 2H, H-8 and H-8'), 7.20 (t, *J* = 8.2 Hz, 1H, H-5), 7.33 (d, *J* = 8.9 Hz, 2H, H-9 and H-9'), 7.44 – 7.49 (m, 1H, H-15), 7.64 – 7.66 (m, 1H, H-16), 7.69 (dd, *J* = 9.6, 1.8 Hz, 1H, H-13).

δ_C (100 MHz, CD₃OD): 93.4 (d, *J* = 7.5 Hz, qC, C-14), 116.4 (CH Ar, C-2), 118.6 (CH Ar, C-4), 121.0 (CH Ar, C-6), 121.5 (2 CH Ar, C-8 and C-8'), 124.1 (d, *J* = 12.4 Hz, qC, C-11), 127.3 (d, *J* = 22.2 Hz, CH Ar, C-13), 128.9 (2 CH Ar, C-9 and C-9'), 130.9 (CH Ar, C-5), 131.5 (qC), 132.4 (CH Ar, C-15), 136.0 (d, *J* = 3.9 Hz, CH Ar, C-16), 137.9 (qC), 156.4 (qC), 158.0 (d, *J* = 254.4 Hz, qC, C-12), 157.3 (qC), 158.3 (qC), 159.5 (qC).

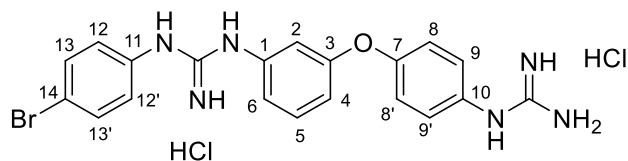
δ_F (376, CD₃OD): - 121.10 (t, *J* = 8.9 Hz).

HRMS (m/z ESI⁺): Found 505.0646 (M⁺ + H), C₂₀H₁₉N₆OFI Requires: 505.0649.

ν_{max} (ATR)/cm⁻¹: 3318 (NH), 3098 (NH), 2958, 1661 (C=N), 1620, 1579, 1485, 1214 (C-F), 1149, 625, 609, 576 (C-I), 566.

HPLC: 99.2% (t_R 25.3 min).

1-(4-Bromophenyl)-2-(3-(4-guanidinophenoxy)phenyl)guanidine dihydrochloride
(5)



Following Method A, **23** (212 mg, 0.29 mmol) was dissolved in 4M HCl in dioxane (1.30 mL, 5.22 mmol) and in additional dioxane (0.12 mL) until a final concentration of 0.2M was reached. After 8 h stirring at 55 °C, the reaction was adjudged complete (TLC), solvents were evaporated and the residue was purified by flash chromatography to afford the pure hydrochloride salt as a white-yellow solid (128 mg, 87%). **Mp**: decomp. > 110 °C.

δ_{H} (400 MHz, CD₃OD): 6.99 (dd, J = 8.2, 2.3 Hz, 1H, H-4), 7.03 (t, J = 2.2 Hz, 1H, H-2), 7.12 – 7.16 (m, 3H, H-6 and H-8 and H-8'), 7.27 (d, J = 8.7 Hz, 2H, H-12 and H-12' or H-13 and H-13'), 7.31 (d, J = 8.8 Hz, 2H, H-9 and H-9'), 7.46 (t, J = 8.1 Hz, 1H, H-5), 7.61 (d, J = 8.7 Hz, 2H, H-12 and H-12' or H-13 and H-13').

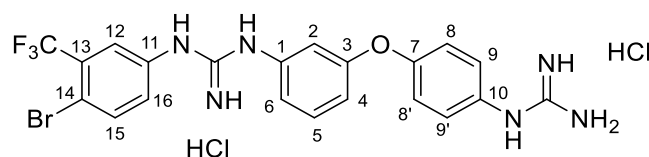
δ_{C} (100 MHz, CD₃OD): 116.3 (CH Ar, C-2), 118.5 (CH Ar, C-4), 120.9 (CH Ar, C-6), 121.5 (2 CH Ar, C-8 and C-8'), 121.7 (qC, C-14), 127.9 (2 CH Ar, C-12 and C-12' or C-13 and C-13'), 128.9 (2 CH Ar, C-9 and C-9'), 131.5 (qC), 132.4 (CH Ar, C-5), 134.1 (CH Ar, C-12 and C-12' or C-13 and C-13'), 135.8 (qC), 138.0 (qC), 156.2 (qC), 157.4 (qC), 158.3 (qC), 159.5 (qC).

HRMS (m/z APCI⁺): Found: 439.0857 ($M^+ + H$), C₂₀H₂₀BrN₆O Requires: 439.0876.

ν_{max} (ATR)/cm⁻¹: 3124 (NH), 3044 (NH), 1655, 1571 (C=N), 1504 (C=N), 1484, 1405, 1213 (C-O), 1070 (C-Br), 1010, 617-567.

HPLC: 99.8% (t_R 24.8 min).

1-(4-Bromo-3-(trifluoromethyl)phenyl)-2-(3-(4-guanidinophenoxy)phenyl)guanidine dihydrochloride (6)



Following Method A, **24** (566 mg, 0.70 mmol) was dissolved in 4M HCl in dioxane (3.15 mL, 12.6 mmol) and in additional dioxane (0.35 mL) until a final concentration of 0.2M was reached. After 8 h stirring at 55 °C, the reaction was adjudged complete (TLC), solvents were evaporated and the residue was purified by silica gel chromatography (CHCl₃:MeOH) to afford the pure hydrochloride salt as a white solid (361 mg, 89%). **Mp**: decomp. > 136 °C.

δ_{H} (400 MHz, CD₃OD): 6.99 (dd, J = 8.0, 1.9 Hz, 1H, H-4), 7.06 (t, J = 2.1 Hz, 1H, H-2), 7.14 – 7.16 (m, 3H, H-8 and H-8' and H-6), 7.32 (d, J = 8.9 Hz, 2H, H-9 and H-9'), 7.46 (t, J = 8.1 Hz, 1H, H-5), 7.51 (dd, J = 8.6, 2.4 Hz, 1H, H-16), 7.74 (d, J = 2.4 Hz, 1H, H-12), 7.88 (d, J = 8.6 Hz, 1H, H-15).

δ_{C} (100 MHz, CD₃OD): 116.2 (CH Ar, C-2), 118.3 (qC, C-14), 118.5 (CH Ar, C-4), 120.8 (CH Ar, C-6), 121.5 (2 CH Ar, C-8 and C-8'), 123.9 (d, J = 260.3 Hz, qCF₃), 125.3 (m, CH Ar, C-12), 128.9 (2 CH Ar, C-9 and C-9'), 130.5 (CH Ar, C-16), 131.6 (qC), 132.3 (d, J = 31.7 Hz, qC, C-13), 132.4 (CH Ar, C-5), 136.9 (qC), 137.8 (CH Ar, C-15), 138.0 (qC), 156.1 (qC), 157.3 (qC), 158.3 (qC), 159.6 (qC).

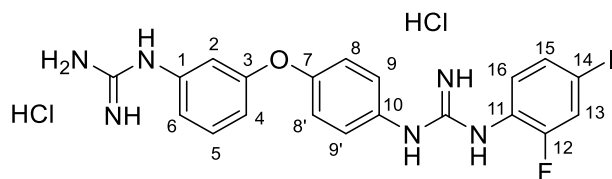
δ_{F} (376 MHz, CD₃OD): - 64.73 (s).

HRMS (m/z ESI⁺): Found 507.0766 (M^+ + H), C₂₁H₁₉N₆OF₃Br Requires: 507.0756.

ν_{max} (ATR)/cm⁻¹: 3119 (NH), 3053 (NH), 1663 (C=O), 1584 (C=N), 1478, 1412, 1320 (C-F), 1238, 1214, 1174, 1129 (CF₃), 1099 (C-Br), 1023, 828, 581 – 558.

HPLC: 99.9 % (t_R : 26.3 min).

1-(2-Fluoro-4-iodophenyl)-2-(4-(3-guanidinophenoxy)phenyl)guanidine dihydrochloride (7)



Following Method A, **25** (310 mg, 0.39 mmol) was dissolved in 4M HCl in dioxane (1.73 mL, 6.93 mmol) and in additional dioxane (0.17 mL) until a final concentration of 0.2M was reached. After 8 h stirring at 55 °C, the reaction was adjudged complete (TLC), solvents were evaporated and the residue was purified by flash chromatography to afford the pure hydrochloride salt as a white solid (198 mg, 88%). **Mp**: decomp. > 150 °C.

δ_{H} (400 MHz, CD₃OD): 6.97 (t, J = 2.1 Hz, 1H, H-2), 7.00 (dd, J = 8.2, 2.3 Hz, 1H, H-4), 7.08 (dd, J = 7.6, 1.5 Hz, 1H, H-6), 7.16 (d, J = 8.8 Hz, 2H, H-8 and H-8'), 7.22 (t, J = 8.2 Hz, 1H, H-5 or H-15), 7.36 (d, J = 8.8 Hz, 2H, H-9 and H-9'), 7.47 (t, J = 8.1 Hz, 1H, H-5 or H-15), 7.66 (d, J = 9.1 Hz, 1H, H-16), 7.70 (dd, J = 9.6, 1.7 Hz, 1H, H-13).

δ_{C} (100 MHz, CD₃OD): 93.5 (d, J = 7.5 Hz, qC, C-14), 116.6 (CH Ar, C-2), 118.6 (CH Ar, C-4), 121.3 (CH Ar, C-6), 121.6 (2 CH Ar, C-8 and C-8'), 124.0 (d, J = 12.5 Hz, qC, C-11), 127.3 (d, J = 22.2 Hz, C-13), 128.7 (2 CH Ar, C-9 and C-9'), 131.1 (CH Ar, C-5 or C-15), 131.6 (qC), 132.4 (CH Ar, C-5 or C-15), 136.0 (d, J = 3.9 Hz, C-16), 137.7 (qC), 156.8 (qC), 157.3 (qC), 157.9 (qC), 158.2 (d, J = 254.5 Hz, qC, C-12), 159.5 (qC).

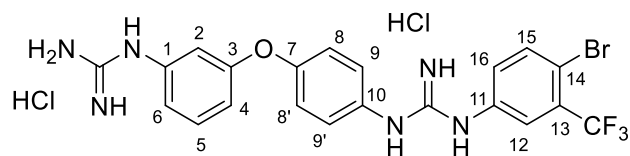
δ_{F} (376 MHz, CD₃OD): -121.04 (t, J = 8.8 Hz).

HRMS (m/z ESI⁺): Found: 505.0645 (M⁺ + H), C₂₀H₁₉N₆OFI Requires: 505.0649.

ν_{max} (ATR)/cm⁻¹: 3335 (NH), 3265 (NH), 3180 (NH), 3052, 2868, 2325, 1616 (C=N), 1560 (C=N), 1504, 1400, 1226 (C-F), 1162, 1109, 971, 875, 789, 684 – 573 (C-I).

HPLC: 99.9% (t_R: 25.9 min).

1-(4-Bromo-3-(trifluoromethyl)phenyl)-2-(4-(3-guanidinophenoxy)phenyl)guanidine dihydrochloride (8)



Following Method A, **26** (148 mg, 0.18 mmol) was dissolved in 4M HCl in dioxane (0.81 mL, 3.24 mmol) and in additional dioxane (0.10 mL) until a final concentration of 0.2M was reached. After 8 h stirring at 55 °C, the reaction was adjudged complete (TLC), solvents were evaporated and the residue was purified by flash chromatography to afford the pure hydrochloride salt as a white solid (100 mg, 94%). **Mp**: decomp. > 95 °C.

δ_{H} (400 MHz, CD₃OD): 6.98 – 7.02 (m, 2H, H-2 and H-4), 7.09 (ddd, 1H, J = 8.0, 1.9, 0.9 Hz, H-6), 7.17 (d, J = 8.9 Hz, 2H, H-8 and H-8'), 7.41 (d, J = 9.0 Hz, 2H, H-9 and H-9'), 7.49 (t, J = 8.0 Hz, 1H, H-5), 7.55 (dd, J = 8.5, 2.6 Hz, 1H, H-16), 7.77 (d, J = 2.6 Hz, 1H, H-12), 7.91 (d, J = 8.6 Hz, 1H, H-15).

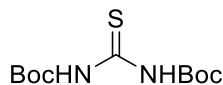
δ_{C} (100 MHz, CD₃OD): 116.6 (CH Ar, C-2 or C-4), 118.5 (qC, C-14), 118.6 (CH Ar, C-2 or C-4), 121.3 (CH Ar, C-6), 121.6 (2 CH Ar, C-8 and C-8'), 123.9 (d, J = 272.8 Hz, qCF₃), 125.5 (q, J = 5.6 Hz, CH Ar, C-12), 128.5 (2 CH Ar, C-9 and C-9'), 130.8 (CH Ar, C-16), 131.7 (qC), 132.3 (q, J = 31.6 Hz, qC, C-13), 132.4 (CH Ar, C-5), 136.7 (qC), 137.7 (qC), 137.8 (CH Ar, C-15), 156.5 (qC), 157.3 (qC), 157.9 (qC), 159.6 (qC).

δ_{F} (376 MHz, CD₃OD): - 64.30 (s).

HRMS (m/z ESI⁺): Found: 507.0750 (M⁺ + H), C₂₁H₁₉N₆OBrF₃ Requires: 507.0756.

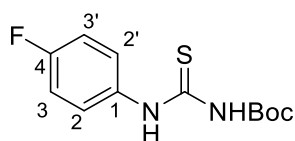
ν_{max} (ATR)/cm⁻¹: 3309 (NH), 3116 (NH), 3053 (NH), 2837, 2280, 1663 (C=N), 1577 (C=N), 1505, 1486, 1405, 1320, 1258 (C-O), 1214 (CF₃), 1129, 1023 (C-Br), 829, 595 - 575.

HPLC: 99.9% (tR: 26.6 min).

***N,N'*-Bis-(*tert*-butoxycarbonyl)thiourea (**9**)**³³

To a cooled solution of thiourea **10** (2.21 g, 24.04 mmol, 1 eq.) in dry tetrahydrofuran (600 mL) under argon, sodium hydride as a 60% suspension in mineral oil (8.12 g, 203.28 mmol, 7 eq.) was added. After 5 min, the ice-bath was removed and the reaction was stirred for 10 min at room temperature. The reaction mixture was again cooled to 0 °C and di-*tert*-butyldicarbonate (13.31 g, 60.98 mmol, 2.1 eq.) was added. After 30 min, the ice-bath was removed and the reaction mixture was stirred overnight at room temperature. The reaction was quenched by the dropwise addition of saturated NaHCO₃ solution and extracted with EtOAc. The organic phase was washed with brine, dried over MgSO₄, filtered, and concentrated under vacuum. The crude product was purified by silica gel chromatography (hexane:EtOAc) to afford **9** as a white solid (3.65 g, 55% yield). **Mp**: 136-139 °C (lit. 136-138 °C).³³

δ_H(400 MHz, CDCl₃): 1.51 (s, 18H, CH₃).

***N*-(*tert*-Butoxycarbonyl)-*N'*-(4-fluorobenzene)thiourea (11)**

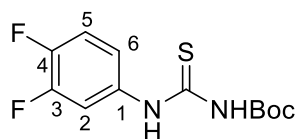
Following Method B, to a solution of **9** (422 mg, 1.53 mmol) in dry THF (12 mL) under argon at 0 °C, NaH as a 60% suspension in mineral oil (91 mg, 2.30 mmol) was added. The reaction mixture was stirred at the same temperature for 1.5 h, then 1.1 eq. of trifluoroacetic anhydride (0.20 mL, 1.68 mmol) were added, and the stirring continue for an additional 30 min. 1.15 eq. of 4-fluoroaniline (196 mg, 1.76 mmol) were then added and the reaction was stirred at 0 °C for 5 h. Usual work-up followed by silica gel chromatography (hexane:Et₂O) afforded the desired product as a white solid (173 mg, 42%). **Mp**: 132-134 °C.

δ_{H} (400 MHz, CDCl₃): 1.53 (s, 9H, (CH₃)₃), 7.04 – 7.12 (m, 2H, H-2 and H-2'), 7.54 – 7.61 (m, 2H, H-3 and H-3'), 7.97 (bs, 1H, NH), 11.42 (s, 1H, NH).

δ_{C} (100 MHz, CDCl₃): 28.0 ((CH₃)₃), 84.5 (qC, C(CH₃)₃), 115.7 (d, J = 22.9 Hz, 2 CH Ar, C-3 and C-3'), 126.5 (d, J = 8.2 Hz, 2 CH Ar, C-2 and C-2'), 133.7 (d, J = 2.9 Hz, qC, C-1), 151.9 (qC), 160.9 (d, J = 246.7 Hz, qC, C-4), 178.9 (qC).

HRMS (m/z ESI⁺): Found: 269.0756 (M^+ - H), C₁₂H₁₄FN₂O₂S Requires: 269.0766.

ν_{max} (ATR)/cm⁻¹: 3168 (NH), 2983, 2931, 1715 (C=O), 1525-1507 (C-F), 1249, 1146 (C=S), 838, 727.

***N*-(*tert*-Butoxycarbonyl)-*N'*-(3,4-difluoromethylphenyl)thiourea (12)**

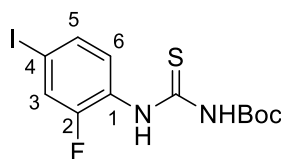
Following Method B, to a solution of **9** (1000 mg, 3.6 mmol) in dry THF (27 mL) under argon at 0 °C, NaH as a 60% suspension in mineral oil (216 mg, 5.4 mmol) was added. The reaction mixture was stirred at the same temperature for 1.5 h, then 1.1 eq. of trifluoroacetic anhydride (0.55 mL, 3.96 mmol) were added, and the stirring continue for an additional 30 min. 1.15 eq. of 3,4-difluoroaniline (0.41 mg, 4.14 mmol) were then added and the reaction was stirred at 0 °C for 5 h. Usual work-up followed by silica gel chromatography (hexane:Et₂O) afforded the desired product as a white solid (592 mg, 57%). **Mp**: 116-118 °C.

δ_{H} (400 MHz, CDCl₃): 1.53 (s, 9H, (CH₃)₃), 7.16 (dd, 1H, *J* = 18.1, 9.0 Hz, H-2), 7.25 - 7.27 (m, 1H, H-6), 7.71 - 7.76 (m, 1H, H-5), 7.96 (s, 1H, NH), 11.52 (s, 1H, NH).

δ_{C} (150 MHz, CDCl₃): 28.0 ((CH₃)₃), 84.7 (qC, C(CH₃)₃), 114.0 (d, *J* = 20.8 Hz, CH Ar, C-5), 117.1 (d, *J* = 18.3 Hz, CH Ar, C-2), 120.3 (dd, *J* = 6.1, 3.6 Hz, CH Ar, C-6), 134.1 (dd, *J* = 8.3, 3.4 Hz, qC, C-1), 148.5 (d, *J* = 248.6, 13.0 Hz, qC, C-3 or C-4), 149.82 (d, *J* = 248.6, 13.4 Hz, qC, C-3 or C-4) 151.9 (qC), 178.6 (qC).

HRMS (*m/z* ESI⁺): Found: 287.0661 (*M*⁺ - H), C₁₂H₁₃F₂N₂O₂S Requires: 287.0671.

ν_{max} (ATR)/cm⁻¹: 3183 (NH), 3081, 2978, 1713 (C=O), 1514 (C-F), 1251, 1112 (C=S), 1139, 723.

***N*-(*tert*-Butoxycarbonyl)-*N'*-(4-iodo-2-fluoromethylphenyl)thiourea (13)**

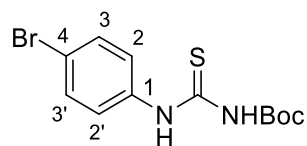
Following Method B, to a solution of **9** (1350 mg, 4.88 mmol) in dry THF (40 mL) under argon at 0 °C, NaH as a 60% suspension in mineral oil (293 mg, 7.32 mmol) was added. The reaction mixture was stirred at the same temperature for 1.5 h, then 1.1 eq. of trifluoroacetic anhydride (0.75 mL, 5.37 mmol) were added, and the stirring continue for an additional 30 min. 1.15 eq. of 2-fluoro-4-iodoaniline (1531 mg, 5.61 mmol) were then added and the reaction was stirred at 0 °C for 5 h. Usual work-up followed by silica gel chromatography (hexane:Et₂O) and recrystallization from boiling CH₂Cl₂, afforded the desired product as a white solid (1109 mg, 57%). **Mp**: 156-158 °C.

δ_H (400 MHz, CDCl₃): 1.54 (s, 9H, (CH₃)₃), 7.49 – 7.51 (m, 2H, H-3 and H-6), 8.02 (bs, 1H, NH), 8.17 – 8.21 (m, 1H, H-5), 11.63 (bs, 1H, NH).

δ_C (150 MHz, CDCl₃): 28.2 ((CH₃)₃), 84.9 (qC, C(CH₃)), 89.6(d, *J* = 7.5 Hz, qC, C-4), 124.9 (d, *J* = 21.8 Hz, CH Ar, C-3), 126.3 (d, *J* = 10.5 Hz, qC, C-1), 126.7 (CH Ar, C-5), 133.4 (d, *J* = 3.8 Hz, CH Ar, C-6), 151.9 (qC), 154.6 (d, *J* = 253.5 Hz, qC, C-2), 178.4 (qC).

HRMS (m/z ESI⁺): Found: 394.9703 (M⁺ - H), C₁₂H₁₃FIN₂O₂S Requires: 394.9732.

ν_{max} (ATR)/cm⁻¹: 3339 (NH), 2975, 1719 (C=O), 1587 (C-F), 1505, 1481, 1342, 1243, 1206, 1133 (C=S), 949, 853, 728 (C-I).

***N*-(*tert*-Butoxycarbonyl)-*N'*-(4-bromophenyl)thiourea (14)**

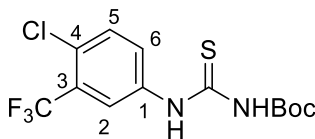
Following Method B, to a solution of **9** (2000 mg, 7.24 mmol) in dry THF (59 mL) under argon at 0 °C, NaH as a 60% suspension in mineral oil (434 mg, 10.86 mmol) was added. The reaction mixture was stirred at the same temperature for 1.5 h, then 1.1 eq. of trifluoroacetic anhydride (1.11 mL, 7.96 mmol) were added, and the stirring continue for an additional 30 min. 1.15 eq. of 4-bromoaniline(1432 mg, 8.33 mmol) were then added and the reaction was stirred at 0 °C for 5 h. Usual work-up followed by silica gel chromatography (hexane:Et₂O) and recrystallization from boiling CH₂Cl₂, afforded the desired product as a white solid (1031 mg, 43%). **Mp**: 150 - 153 °C.

δ_H (400 MHz, CDCl₃): 1.53 (s, 9H, CH₃), 7.50 (d, *J* = 8.8 Hz, 2H, H-2 and H-2'), 7.56 (d, *J* = 8.8 Hz, 2H, H-3 and H-3'), 7.96 (s, 1H, NH), 11.51 (s, 1H, NH).

δ_C (150 MHz, CDCl₃): 28.2 ((CH₃)₃), 84.7 (qC, C(CH₃)), 119.9 (qC, C-4), 125.9 (2 CH Ar, C-2 and C-2'), 132.0 (2 CH Ar, C-3 and C-3'), 136.9 (qC, C-1), 152.0 (q), 178.4 (qC).

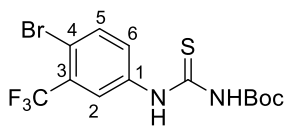
HRMS (*m/z* ESI⁺): Found: 328.9967 (M⁺ - H), C₁₂H₁₄BrN₂O₂S Requires: 328.9965.

ν_{max} (ATR)/cm⁻¹: 3415 (NH), 3200, 2982, 1715 (C=O), 1524, 1487, 1137 (C=S), 905 (C-Br), 726.

***N*-(*tert*-Butoxycarbonyl)-*N'*-[4-chloro-3-(trifluoromethyl)phenyl]thiourea (15)²⁸**

Following Method B, to a solution of **9** (2000 mg, 7.25 mmol) in dry THF (58 mL) under argon at 0 °C, NaH as a 60% suspension in mineral oil (435 mg, 10.9 mmol) was added. The reaction mixture was stirred at the same temperature for 1.5 h, then 1.1 eq. of trifluoroacetic anhydride (1.11 mL, 8.0 mmol), and the stirring continue for an additional 30 min. Then, 1.15 eq. of 4-chloro-3-(trifluoromethyl)aniline (1.63 g, 8.34 mmol) were added and the reaction was stirred at 0 °C for 5 h. Usual work-up followed by silica gel chromatography (eluting with a gradient of hexane:EtOAc) and recrystallization from boiling hexane, to afford the desired product as a white solid (1414 mg, 55%). **Mp**: 138 – 140 °C. (lit. 140-142 °C).²⁸

δ_{H} (400 MHz, CDCl_3): 1.54 (s, 9H, CH_3), 7.51 (d, 1H, $J = 8.7$ Hz, H-5), 7.91 (d, 1H, $J = 8.6$, 2.3 Hz, H-6), 7.96 – 8.00 (m, 2H, NH, H-2), 11.67 (bs, 1H, NH).

***N*-(*tert*-Butoxycarbonyl)-*N'*-[4-bromo-3-(trifluoromethyl)phenyl]thiourea (16)**

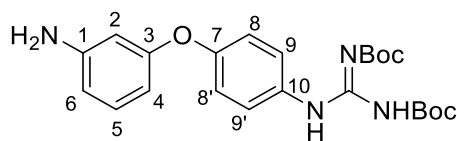
Following Method B, to a solution of **9** (925 mg, 3.38 mmol) in dry THF (27 mL) under argon at 0 °C, NaH as a 60% suspension in mineral oil (202 mg, 5.07 mmol) was added. The reaction mixture was stirred at the same temperature for 1.5 h, then 1.1 eq. of trifluoroacetic anhydride (0.52 mL, 3.72 mmol) were added, and the stirring continue for an additional 30 min. 1.15 eq. of 4-bromo-3-(trifluoromethyl)aniline(933 mg, 3.89 mmol) were then added and the reaction was stirred at 0 °C for 5 h. Usual work-up followed by silica gel chromatography (hexane:Et₂O) and recrystallization from boiling CH₂Cl₂, to afford the desired product as a white solid (500 mg, 37%). **Mp**: 132 - 134 °C.

δ_H (600 MHz, CDCl₃): 1.54 (s, 9H, CH₃), 7.71 (d, 1H, *J* = 8.6 Hz, H-5), 7.84 (dd, 1H, *J* = 8.6, 1.8 Hz, H-6), 7.98 – 8.02 (m, 2H, NH and H-2), 11.68 (bs, 1H, NH).

δ_C (150 MHz, CDCl₃): 28.2 ((CH₃)₃), 85.1 (qC, C(CH₃)), 117.0 (qC, C-4), 123.5 (q, *J* = 11.1 Hz, CH Ar, C-2), 124.5 (d, *J* = 273.6 Hz, qCF₃), 128.3 (CH Ar, C-6), 130.8 (d, *J* = 31.9 Hz, qC, C-3), 135.4 (CH Ar, C-5), 137.3 (qC), 152.1 (qC), 178.6 (qC).

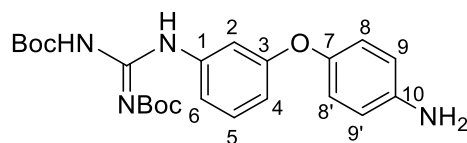
HRMS (*m/z* ESI⁺): Found: 396.9832 (M⁺ - H), C₁₃H₁₃N₂O₂BrF₃S Requires: 396.9839.

ν_{max} (ATR)/cm⁻¹: 3159 (NH), 1706 (C=O), 1592, 1533, 1474, 1324, 1237 (C=S), 1128 (C-F), 1016 (C-Br), 818, 688.

2,3-(di-*tert*-butoxycarbonyl)-1-(4-(3-aminophenoxy)phenyl)guanidine (17)²⁸

HgCl₂ (1200 mg, 4.42 mmol, 1.2 eq.) was added over a solution of **19** (2170 mg, 10.86 mmol, 3 eq.), **9** (1000 mg, 3.62 mmol, 1 eq.) and triethylamine (1.56 mL, 11.22 mmol, 3.1 eq.) in CH₂Cl₂ (18 mL) at 0 °C. The resulting mixture was stirred at 0 °C for 1 h and then overnight at room temperature. When the reaction was adjudged complete by TLC, the organic phase was washed with brine, dried over MgSO₄, filtered, and concentrated under vacuum. The crude product was purified by silica gel chromatography (hexane:EtOAc) to afford the title product as a whitesolid (1201 mg, 75%). **Mp**: 78-79 °C (lit. 78-79 °C).²⁸

δ_H (400 MHz, CDCl₃): 1.50 (s, 9H, (CH₃)₃), 1.55 (s, 9H, (CH₃)₃), 3.68 (bs, 2H, NH₂), 6.30 (s, 1H, H-2), 6.38 – 6.40 (m, 2H, H-4 and H-6), 6.99 (d, 2H, *J* = 8.9, H-8 and H-8'), 7.08 (t, 1H, *J* = 8.0 Hz, H-5), 7.55 (d, 2H, *J* = 8.9 Hz, H-9 and H-9'), 10.28 (bs, 1H, NH), 11.64 (bs, 1H, NH).

2,3-(di-*tert*-butoxycarbonyl)-1-(3-(4-aminophenoxy)phenyl)guanidine (18**)**

HgCl₂ (1200 mg, 4.42 mmol, 1.2 eq.) was added over a solution of **19** (2170 mg, 10.86 mmol, 3 eq.), **9** (1000 mg, 3.62 mmol, 1 eq.) and triethylamine (1.56 mL, 11.22 mmol, 3.1 eq.) in CH₂Cl₂ (18 mL) at 0 °C. The resulting mixture was stirred at 0 °C for 1 h and then overnight at room temperature. When the reaction was adjudged complete by TLC, the organic phase was washed with brine, dried over MgSO₄, filtered, and concentrated under vacuum. The crude product was purified by silica gel chromatography (hexane:EtOAc) to afford the title product as a yellow oil (400 mg, 25%).

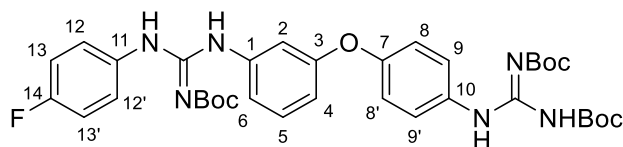
δ_{H} (400 MHz, CDCl₃): 1.49 (s, 9H, (CH₃)₃), 1.52 (s, 9H, (CH₃)₃), 3.58 (bs, 2H, NH₂), 6.65 - 6.69 (m, 3H, H-9 and H-9' and H-4), 6.88 (d, 2H, *J* = 8.6, H-8 and H-8'), 7.17 - 7.24 (m, 3H, H-5, H-6 and H-2), 10.25 (bs, 1H, NH), 11.59 (bs, 1H, NH).

δ_{C} (100 MHz, CDCl₃): 28.2 ((CH₃)₃), 28.3 ((CH₃)₃), 79.7 (qC, C(CH₃)₃), 83.9 (qC, C(CH₃)₃), 111.4 (CH Ar, C-2), 113.5 (CH Ar, C-4), 116.1 (CH Ar, C-6), 116.4 (2 CH Ar, C-9 and C-9'), 121.5 (2 CH Ar, C-8 and C-8'), 129.7 (CH Ar, C-5), 138.1 (qC), 142.9 (qC), 148.4 (qC), 153.4 (qC), 153.6 (qC), 159.4 (qC), 163.6 (qC).

HRMS (*m/z* - ESI⁺): Found: 443.2295 (M⁺ + H), C₂₃H₃₁N₄O₅ Requires: 443.2289.

ν_{max} (ATR)/cm⁻¹: 3474 (NH), 2976, 1719 (C=O), 1632 (C=N), 1507, 1410, 1213, 1145 (C-O).

1-(4-Fluorophenyl)-2-(*tert*-butoxycarbonyl)-3-{3-[4-(1',2'-di-*tert*-butoxycarbonyl)guanidinophenoxy]phenyl}guanidine (20**)**



Following Method C, HgCl_2 (92 mg, 0.34 mmol) was added over a solution of **17** (124 mg, 0.28 mmol), thiourea derivative **11** (93 mg, 0.34 mmol) and NEt_3 (0.12 mL, 0.87 mmol) in CH_2Cl_2 (1.5 mL). After 12 h workup and silica gel chromatography (hexanes:EtOAc) afforded compound **20** as a yellow oil (108 mg, 57%).

δ_{H} (400 MHz, CDCl_3): 1.48 (s, 9H, $(\text{CH}_3)_3$), 1.49 ((s, 9H, $(\text{CH}_3)_3$), 1.53 (s, 9H, $(\text{CH}_3)_3$), 6.55-6.70 (m, 3H, CH Ar), 6.84 (bs, 1H, CH Ar), 6.98-6.70 (m, 4H, CH Ar), 7.24 (bs, 1H, CH Ar), 7.32-7.47 (bs, 1H, CH Ar), 7.56 (m, 2H, CH Ar), 9.50-9.55 (m, NH), 10.29 (bs, NH), 11.64 (bs, NH).

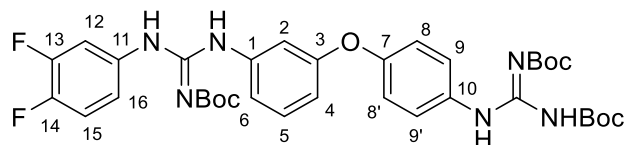
δ_{C} (100 MHz, CDCl_3): 28.2 ($2(\text{CH}_3)_3$), 28.3 ($(\text{CH}_3)_3$), 79.8 (qC, $\underline{\underline{\text{C}}}(\text{CH}_3)_3$), 83.5 (qC, $\underline{\underline{\text{C}}}(\text{CH}_3)_3$), 83.9 (qC, $\underline{\underline{\text{C}}}(\text{CH}_3)_3$), 112.7 (CH Ar), 113.0 (CH Ar), 115.6 (CH Ar), 116.3 (CH Ar), 119.9 (CH Ar), 123.9 (CH Ar), 129.8 (CH Ar), 130.8 (CH Ar), 132.6 (qC), 140.5 (qC), 153.0 (qC), 153.5 (qC), 153.7 (qC), 157.9 (qC), 158.9 (qC), 163.7 (qC).

Note: Aromatic protons and carbons, quaternary carbons and fluorine peaks could not be fully assigned and characterised because the peaks were very broad.

HRMS (m/z ESI^+): Found: 679.3253 ($\text{M}^+ + \text{H}$), $\text{C}_{35}\text{H}_{44}\text{N}_6\text{O}_7\text{F}$ Requires: 679.3256.

ν_{max} (ATR)/ cm^{-1} : 3412 (NH), 3308 (NH), 3263, 2979, 2931, 1719 (C=O), 1628, 1596 (C=N), 1234 (C-N), 1210, 1121 (C-O), 1145 (C-F), 1095, 832, 804, 771, 731, 687.

1-(3,4-Difluorophenyl)-2-(*tert*-butoxycarbonyl)-3-{3-[4-(1',2'-di-*tert*-butoxycarbonyl) guanidinophenoxy]phenyl}guanidine (21**)**



Following Method C, HgCl_2 (92 mg, 0.34 mmol) was added over a solution of **17** (124 mg, 0.28 mmol), thiourea derivative **12** (98 mg, 0.34 mmol) and NEt_3 (0.12 mL, 0.87 mmol) in CH_2Cl_2 (1.5 mL). After 12 h workup and silica gel chromatography (hexanes:EtOAc) afforded compound **21** as a white-yellow amorphous solid (120 mg, 61%). **Mp**: 125 – 127 °C

δ_{H} (400 MHz, CDCl_3): 1.48 (s, 9H, $(\text{CH}_3)_3$), 1.49 (s, 9H, $(\text{CH}_3)_3$), 1.54 (s, 9H, $(\text{CH}_3)_3$), 6.54 - 6.71 (m, 4H, CH Ar and NH), 7.00 (bs, 2H, CH Ar), 7.07 - 7.18 (m, 2H, CH Ar), 7.28 – 7.44 (bs, 1H, CH Ar), 7.56 - 7.58 (bs, 2H, CH Ar), 7.82 (bs, 1H, CH Ar), 9.62 (bs, NH), 10.30 (bs, NH), 11.64 (bs, NH).

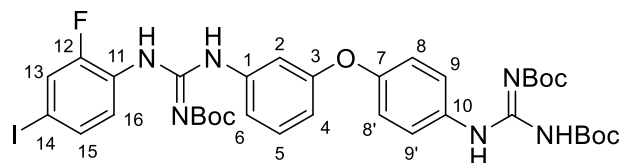
δ_{C} (100 MHz, CDCl_3): 28.2 ($2(\text{CH}_3)_3$), 28.3 ($(\text{CH}_3)_3$), 79.8 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 83.8 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 84.0 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 109.7 (CH Ar), 112.5 (CH Ar), 113.3 (CH Ar), 115.6 (CH Ar), 117.1 (CH Ar), 118.2 (CH Ar), 119.9 (2 CH Ar), 123.9 (2 CH Ar), 130.9 (CH Ar), 132.7 (qC), 140.1 (qC), 148.1 (qC), 153.0 (qC), 153.5 (qC), 153.7 (qC), 159.1 (qC), 163.7 (qC).

Note: Aromatic protons and carbons, quaternary carbons and fluorine peaks could not be fully assigned and characterised because the peaks were very broad.

HRMS (m/z ESI⁺): Found: 697.3157 ($\text{M}^+ + \text{H}$), $\text{C}_{35}\text{H}_{43}\text{F}_2\text{N}_6\text{O}_7$ Requires: 697.3161.

ν_{max} (ATR)/ cm^{-1} : 3413 (NH), 3304, 3264 (NH), 2979, 2928, 2856, 1719 (C=O), 1627, 1596 (C=N), 1565, 1505, 1480, 1461, 1233 (C-N), 1210 (C-O), 1145 (C-F), 803, 773, 731.

1-(2-Fluoro-4-iodophenyl)-2-(*tert*-butoxycarbonyl)-3-{3-[4-(1',2'-di-*tert*-butoxycarbonyl) guanidinophenoxy]phenyl}guanidine (22)



Following Method C, HgCl_2 (147 mg, 0.54 mmol) was added over a solution of **17** (200 mg, 0.45 mmol), thiourea derivative **13** (214 mg, 0.54 mmol) and NEt_3 (0.19 mL, 1.40 mmol) in CH_2Cl_2 (2.5 mL). After 12 h workup and silica gel chromatography (hexanes:EtOAc) afforded compound **22** as a white-pale yellow amorphous solid (220 mg, 61%). **Mp**: 79-81 °C

δ_{H} (400 MHz, CDCl_3): 1.49 (s, 18H, $(\text{CH}_3)_3$), 1.54 (s, 9H, $(\text{CH}_3)_3$), 6.54 - 6.71 (m, 3H, CH Ar), 6.99 (bs, 2H, CH Ar), 7.22 - 7.28 (bs, 2H, CH Ar), 7.38 - 7.42 (m, 1H, CH Ar), 7.56 - 7.58 (m, 2H, CH Ar), 8.32 (bs, 1H, CH Ar), 9.69 (m, NH), 9.90 (m, NH), 10.30 (bs, NH), 11.64 (bs, NH).

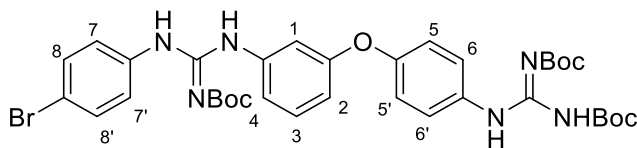
δ_{C} (100 MHz, CDCl_3): 28.2 ($(\text{CH}_3)_3$), 28.3 ($(\text{CH}_3)_3$), 79.8 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 83.9 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 112.4 (CH Ar), 113.3 (CH Ar), 116.9 (CH Ar), 120.0 (2 CH Ar), 123.1 (CH Ar), 123.9 (2 CH Ar), 130.0 (CH Ar), 131.0 (CH Ar), 133.5 (CH Ar), 134.1 (qC), 152.9 (qC), 153.5 (qC), 153.7 (qC), 159.2 (qC), 163.7 (qC).

Note: Aromatic protons and carbons, quaternary carbons and fluorine peaks could not be fully assigned and characterised because the peaks were very broad.

HRMS (m/z ESI^+): Found: 805.2213 ($\text{M}^+ + \text{H}$), $\text{C}_{35}\text{H}_{43}\text{FIN}_6\text{O}_7$ Requires: 805.2216.

ν_{max} (ATR)/ cm^{-1} : 3450 (NH), 2977, 2935, 1718 (C=O), 1591 (C=N), 1407, 1367, 1234 (C-N), 1144 (C-F), 1112 (C-O), 1056, 772 (C-I).

1-(4-Bromophenyl)-2-(*tert*-butoxycarbonyl)-3-{3-[4-(1',2'-di-*tert*-butoxycarbonyl)guanidinophenoxy]phenyl}guanidine (23)



Following Method C, HgCl_2 (209 mg, 0.77 mmol) was added over a solution of **17** (283 mg, 0.64 mmol), thiourea derivative **14** (255 mg, 0.77 mmol) and NEt_3 (0.27 mL, 1.98 mmol) in CH_2Cl_2 (3.5 mL). After 12 h workup and silica gel chromatography (hexanes:EtOAc) afforded compound **23** as a white amorphous solid (170 mg, 36%).

Mp: 116 - 119 °C

δ_{H} (600 MHz, CDCl_3): 1.48 (bs, 9H, $(\text{CH}_3)_3$), 1.49 (bs, 9H, $(\text{CH}_3)_3$), 1.54 (bs, 9H, $(\text{CH}_3)_3$), 6.55 - 6.78 (m, 4H, CH Ar), 6.99 – 7.01 (m, 2H, CH Ar), 7.39 – 7.43 (m, 3H, CH Ar), 7.57 (bs, 3H, CH Ar), 9.60 (bs, 1H, NH), 10.29 (bs, 1H, NH), 11.64 (bs, 1H, NH).

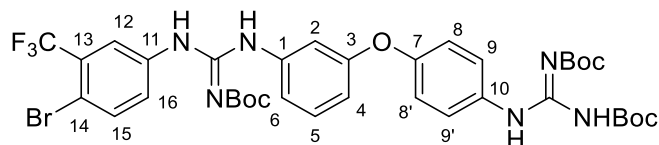
δ_{C} (150 MHz, CDCl_3): 28.2 ($(\text{CH}_3)_3$), 28.3 ($(\text{CH}_3)_3$), 79.8 (qC, $\text{C}(\text{CH}_3)_3$), 83.9 (qC, $\text{C}(\text{CH}_3)_3$), 112.6 (CH Ar), 113.2 (CH Ar), 117.2 (CH Ar), 119.6 (CH Ar), 119.9 (CH Ar), 121.7 (CH Ar), 121.8 (CH Ar), 124.0 (CH Ar), 124.5 (CH Ar), 131.8 (CH Ar), 132.1 (CH Ar), 132.8 (CH Ar), 140.2 (qC), 145.7 (qC), 153.2 (qC), 153.5 (qC), 158.8 (qC), 163.4 (qC).

Note: Aromatic protons and carbons and quaternary carbons peaks could not be fully assigned and characterised because the peaks were very broad.

ν_{max} (ATR)/ cm^{-1} : 3412, 3260 (NH), 2980, 2933, 1719 (C=O), 1625, 1599 (C=N), 1551, 1481, 1459, 1481, 1408, 1368, 136, 1304, 1231, 1211, 1145 (C-O), 1113 (C-Br), 1097, 1057, 909, 731.

HRMS (m/z ESI⁺): Found: 739.2458 ($\text{M}^+ + \text{H}$), $\text{C}_{35}\text{H}_{44}\text{N}_6\text{O}_7\text{Br}$ Requires: 739.2455.

1-[4-Bromo-3-(trifluoromethyl)phenyl]-2-(*tert*-butoxycarbonyl)-3-{3-[4-(1',2'-di-*tert*-butoxycarbonyl)guanidinophenoxy]phenyl}guanidine (24**)**



Following Method C, HgCl_2 (147 mg, 0.54 mmol) was added over a solution of **17** (200 mg, 0.45 mmol), thiourea derivative **16** (217 mg, 0.54 mmol) and NEt_3 (0.19 mL, 1.40 mmol) in CH_2Cl_2 (3 mL). After 12 h workup and silica gel chromatography (hexanes:EtOAc) afforded compound **24** as a white-pale yellow amorphous solid (207 mg, 57%). **Mp**: 92-94 °C

δ_{H} (400 MHz, CDCl_3): 1.49 (s, 18H, $(\text{CH}_3)_3$), 1.54 (s, 9H, $(\text{CH}_3)_3$), 6.54-6.72 (m, 3H, CH Ar and NH), 7.01 (bs, 3H, CH Ar), 7.26 – 7.29 (m, 2H, CH Ar), 7.57 - 7.59 (m, 3H, CH Ar), 7.87-7.92 (m, 1H, CH Ar), 9.67-9.83 (bs, NH), 10.31 (bs, NH), 11.64 (bs, NH).

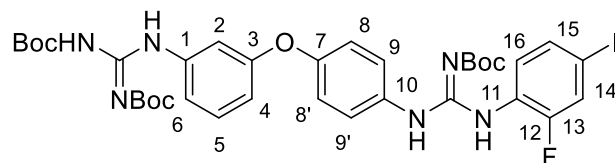
δ_{C} (100 MHz, CDCl_3): 28.3 (3 $(\text{CH}_3)_3$), 79.8 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 83.9 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 84.1 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 112.3 (CH Ar), 113.3 (CH Ar), 117.0 (CH Ar), 119.7 (CH Ar), 120.0 (CH Ar), 123.9 (2 CH Ar), 124.1 (CH Ar), 124.3 (qC), 127.0 (CH Ar), 130.0 (CH Ar), 130.1 (CH Ar), 132.8 (qC), 135.2 (qC), 136.1 (qC), 138.5 (qC), 140.0 (qC), 147.9 (qC), 153.1 (qC), 153.5 (qC), 153.7 (qC), 158.1 (qC), 159.2 (qC), 163.7 (qC).

Note: Aromatic protons and carbons, quaternary carbons and fluorine peaks could not be fully assigned and characterised because the peaks were very broad.

HRMS (m/z ESI^+): Found: 807.2333 ($\text{M}^+ + \text{H}$), $\text{C}_{36}\text{H}_{43}\text{BrF}_3\text{N}_6\text{O}_7$ Requires: 807.2323.

ν_{max} (ATR)/ cm^{-1} : 3410 (NH), 3305 (NH), 2980, 2940, 1720 (C=O), 1589 (C=N), 1554 (C=N), 1477, 1407, 1236 (C-N), 1211 (C-O), 1144 (CF_3), 1112, 1094 (C-Br), 771, 832.

1-(4-Iodo-2-fluorophenyl)-2-(*tert*-butoxycarbonyl)-3-{4-[3-(1',2'-di-*tert*-butoxycarbonyl) guanidinophenoxy]phenyl}guanidine (25)



Following Method C, HgCl_2 (206 mg, 0.76 mmol) was added over a solution of **18** (278 mg, 0.63 mmol), thioureaderivative **13** (300 mg, 0.76 mmol) and NEt_3 (0.27 mL, 1.95 mmol) in CH_2Cl_2 (4 mL). After 12 h workup and silica gel chromatography (hexanes:EtOAc) afforded compound **25** as a yellow amorphous solid (299 mg, 59%).

Mp: 87 – 88 °C

δ_{H} (400 MHz, CDCl_3): 1.50 (bs, 18H, $(\text{CH}_3)_3$), 1.53 (bs, 9H, $(\text{CH}_3)_3$), 6.71 – 6.75 (m, 1H, CH Ar), 6.88 – 6.90 (m, 1H, CH Ar), 7.00 – 7.07 (m, 2H, CH Ar), 7.22 – 7.24 (m, 1H, CH Ar), 7.28 – 7.32 (m, 1H, CH Ar), 7.37 – 7.46 (m, 2H, CH Ar), 7.52 – 7.65 (m, 1H, CH Ar), 7.97 – 8.01 (m, 1H, CH Ar), 8.36 – 8.40 (m, 1H, CH Ar), 9.64 (bs, 1H, NH), 9.91 (bs, 1H, NH), 10.31 (bs, 1H, NH), 11.60 (bs, 1H, NH).

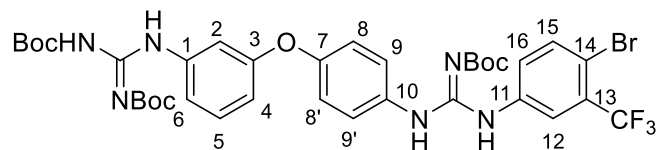
δ_{C} (100 MHz, CDCl_3): 28.2 ($(\text{CH}_3)_3$), 28.3 ($(\text{CH}_3)_3$), 79.8 (qC, $\text{C}(\text{CH}_3)_3$), 83.9 (qC, $\text{C}(\text{CH}_3)_3$), 84.1 (qC, $\text{C}(\text{CH}_3)_3$), 112.3, 114.4, 116.8, 120.1, 121.2, 123.0, 123.6, 124.1, 129.9, 133.8, 134.1, 138.3, 150.1, 153.4, 153.6, 158.4, 163.6.

Note: Aromatic protons and carbons, quaternary carbons and fluorine peaks could not be fully assigned and characterised because the peaks were very broad.

HRMS (m/z – ESI^+): Found: 805.2226 ($\text{M}^+ + \text{H}$), $\text{C}_{35}\text{H}_{43}\text{N}_6\text{O}_7\text{FI}$ Requires: 805.2222.

ν_{max} (ATR)/ cm^{-1} : 3411 (NH), 2976, 1719 (C=O), 1638, 1599 (C=N), 1535 (C=N), 1499, 1466, 1411, 1367, 1240 (C-O), 1153 (C-F), 1140, 1110, 1058, 816, 762 (C-I).

1-[4-Bromo-3-(trifluoromethyl)phenyl]-2-(*tert*-butoxycarbonyl)-3-{4-[3-(1',2'-di-*tert*-butoxycarbonyl)guanidinophenoxy]phenyl}guanidine (26**)**



Following Method C, HgCl_2 (152 mg, 0.56 mmol) was added over a solution of **18** (207 mg, 0.47 mmol), thiourea derivative **16** (224 mg, 0.56 mmol) and NEt_3 (0.20 mL, 1.46 mmol) in CH_2Cl_2 (3 mL). After 12 h workup and silica gel chromatography (hexanes:EtOAc) afforded compound **26** as a yellow oil (250 mg, 66%).

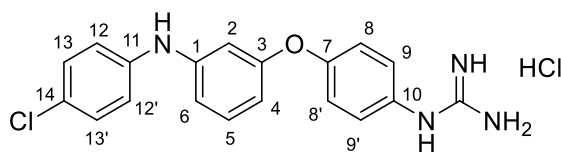
δ_{H} (400 MHz, CDCl_3): 1.49 (bs, 9H, $(\text{CH}_3)_3$), 1.53 (bs, 9H, $(\text{CH}_3)_3$), 6.71 – 6.76 (m, 1H, CH Ar), 6.88 – 6.90 (m, 1H, CH Ar), 7.00 – 7.08 (m, 2H, CH Ar), 7.22 – 7.38 (m, 3H, CH Ar), 7.57 – 7.63 (m, 2H, CH Ar), 7.88 – 7.90 (m, 1H, CH Ar), 8.00 (bs, 1H, CH Ar), 9.65 (bs, 1H NH), 9.83 (bs, 1H NH), 10.30 (bs, 1H NH), 11.60 (bs, 1H NH).

δ_{C} (100 MHz, CDCl_3): 28.2 ($(\text{CH}_3)_3$), 28.3 ($(\text{CH}_3)_3$), 86.2 (qC, $\text{C}(\text{CH}_3)_3$), 112.4 (CH Ar), 114.0 (CH Ar), 114.4 (CH Ar), 116.8 (CH Ar), 118.0 (CH Ar), 120.1 (CH Ar), 121.1 (CH Ar), 123.9 (CH Ar), 130.0 (CH Ar), 134.5 (qC), 135.2 (qC), 135.0 (qC), 136.5 (qC), 137.9 (qC), 138.0 (qC), 138.3 (qC), 152.7 (qC), 152.8 (qC), 153.6 (qC).

Note: Aromatic protons and carbons, quaternary carbons and fluorine peaks could not be fully assigned and characterised because the peaks were very broad.

HRMS (m/z – ESI^+): Found: 807.2329 ($\text{M}^+ + \text{H}$), $\text{C}_{36}\text{H}_{43}\text{BrF}_3\text{N}_6\text{O}_7$ Requires: 807.2323.

ν_{max} (ATR)/ cm^{-1} : 3260 (NH), 2980, 2929, 1718 (C=O), 1630, 1595 (C=N), 1555, 1238, 1140 (CF_3), 1105, 1055 (C-Br), 835, 773.

1-{4-[3-(4-Chlorophenylamino)phenoxy]phenyl}guanidine hydrochloride (27)

Following Method A, **55** (97 mg, 0.18 mmol) was dissolved in 4M HCl in dioxane (0.53 mL, 2.10 mmol) and in additional dioxane (0.9 mL) until a final concentration of 0.2M was reached. After 8 h stirring at 55 °C, the reaction was adjudged complete (TLC), solvents were evaporated and the residue was purified by silica gel chromatography (CHCl₃:MeOH) to afford the pure hydrochloride salt as a purple solid (69 mg, 99%). **Mp**: 50 - 52 °C

δ_H (400 MHz, CD₃OD): 6.51 (dd, J = 8.3, 1.9 Hz, 1H, H-4), 6.73 (t, J = 2.2 Hz, 1H, H-2), 6.85 (dd, J = 7.8, 1.8 Hz, 1H, H-6), 7.05 (d, J = 8.9 Hz, 2H, H-12 and H-12'), 7.08 (d, J = 8.9 Hz, 2H, H-8 and H-8'), 7.18 (d, J = 8.9 Hz, 2H, H-13 and H-13'), 7.22 (t, J = 8.2 Hz, 1H, H-5), 7.27 (d, J = 8.9 Hz, 2H, H-9 and H-9').

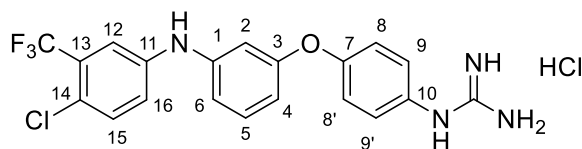
δ_C (100 MHz, CD₃OD): 108.6 (CH Ar, C-2), 111.8 (CH Ar, C-4), 113.7 (CH Ar, C-6), 120.0 (2 CH Ar, C-12 and C-12'), 120.7 (2 CH Ar, C-8 and C-8'), 126.1 (qC, C-14), 128.8 (2 CH Ar, C-9 and C-9'), 130.1 (2 CH Ar, C-13 and C-13'), 130.6 (qC), 131.5 (CH Ar, C-5), 143.4 (qC), 146.7 (qC), 158.4 (qC), 158.5 (qC), 158.9 (qC).

HRMS (m/z ESI⁺): Found 353.1177 (M^+ + H), C₁₉H₁₈N₄OCl Requires: 353.1169.

ν_{max} (ATR)/cm⁻¹: 3297 (N-H), 3126 (N-H), 1688, 1586 (C=N), 1502, 1485 (C-N), 1325, 1216 (C-O), 1142 (C-Cl), 997, 972, 823, 770, 689, 604, 588, 570.

HPLC: 98.0 % (t_R : 32.3 min).

1-(4-{3-[4-Chloro-3-(trifluoromethyl)phenylamino]phenoxy}phenyl)guanidine hydrochloride (28)



Following Method A, **54** (112 mg, 0.18 mmol) was dissolved in 4M HCl in dioxane (0.54 mL, 2.16 mmol) and in additional dioxane (0.36 mL) until a final concentration of 0.2M was reached. After 8 h stirring at 55 °C, the reaction was adjudged complete (TLC), solvents were evaporated and the residue was purified by flash chromatography to afford the pure hydrochloride salt as a white-brown solid (59 mg, 80%). **Mp**: 58 - 60 °C

δ_{H} (400 MHz, CD₃OD): 6.62 (dd, J = 8.7, 2.3 Hz, 1H, H-4), 6.78 (t, J = 2.2 Hz, 1H, H-2), 6.90 (dd, J = 8.1, 2.1 Hz, 1H, H-6), 7.11 (d, J = 8.9 Hz, 2H, H-8 and H-8'), 7.24 (dd, J = 8.8, 2.8 Hz, 1H, H-16), 7.27 - 7.31 (m, 3H, H-9, H-9' and H-5) 7.36 - 7.39 (m, 2H, H-12, H-15).

δ_{C} (100 MHz, CD₃OD): 109.8 (CH Ar, C-2), 113.1 (CH Ar, C-4), 114.8 (CH Ar, C-6), 116.1 (q, J = 5.6 Hz, CH Ar, C-12), 121.0 (2 CH Ar, C-8 and C-8'), 121.5 (CH Ar, C-16), 122.1 (qC, C-14), 124.4 (d, J = 272.5 Hz, qCF₃), 129.0 (2 CH Ar, C-9 and C-9'), 129.6 (q, J = 31.0, qC, C-13), 131.0 (qC), 131.8 (CH Ar, C-5), 133.4 (CH Ar, C-15), 144.4 (qC), 145.3 (qC), 158.2 (qC), 158.4 (qC), 159.2 (qC).

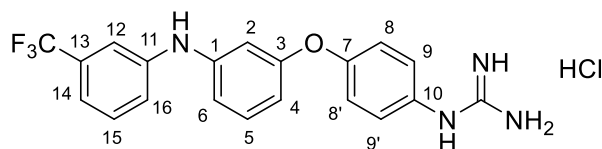
δ_{F} (376 MHz, CD₃OD): - 64.18 (s).

HRMS (m/z ESI⁺): Found 421.1044 (M⁺ + H), C₂₀H₁₇ClF₃N₄O Requires: 421.1043.

ν_{max} (ATR)/cm⁻¹: 3295 (NH), 3163, 2923, 2853, 2400, 1664 (C=O), 1595 (C=N), 1504, 1482, 1441, 1333, 1258, 1217 (CF₃), 1127, 1112 (C-Cl), 1027, 999, 977, 825.

HPLC: 97.8% (tR 32.9 min).

1-(4-{3-[3-(Trifluoromethyl)phenylamino]phenoxy}phenyl)guanidine hydrochloride (29)



Following Method A, **52** (371 mg, 0.63 mmol) was dissolved in 4M HCl in dioxane (1.90 mL, 7.56 mmol) and in additional dioxane (1.25 mL) until a final concentration of 0.2M was reached. After 6 h stirring at 55 °C, the reaction was adjudged complete (TLC), solvents were evaporated and the residue was purified by flash chromatography to afford the pure hydrochloride salt as a white solid (242 mg, 90%). **Mp**: 93 - 95 °C

δ_{H} (400 MHz, CD₃OD): 6.59 (dd, J = 8.1, 2.3 Hz, 1H, H-4), 6.78 (t, J = 2.2 Hz, 1H, H-2), 6.90 (dd, J = 7.9, 1.9 Hz, 1H, H-6), 7.07 – 7.13 (m, 3H, H-8 and 8' and H-12 or H-14), 7.26 – 7.30 (m, 5H, H-9 and 9', H-12 or H-14, H-16 and H-5 or H-15), 7.37 (t, J = 8.3 Hz, 1H, H-5 or H-15).

δ_{C} (100 MHz, CD₃OD): 109.4 (CH Ar, C-2), 112.6 (CH Ar, C-4), 113.9 (q, J = 4.0 Hz, CH Ar, C-12 or C-14), 114.6 (CH Ar, C-6), 117.3 (q, J = 4.0 Hz, CH Ar, C-12 or C-14), 119.9 (d, J = 280.7 Hz, qCF₃), 120.9 (2 CH Ar, C-8 and C-8'), 121.2 (CH Ar, C-16), 128.9 (2 CH Ar, C-9 and C-9'), 130.8 (qC), 131.1 (CH Ar, C-5 or C-15), 131.7 (CH Ar, C-5 or C-15), 132.6 (d, J = 31.9 Hz, qC, C-13), 145.7 (qC), 145.9 (qC), 158.3 (qC), 158.4 (qC), 159.1 (qC).

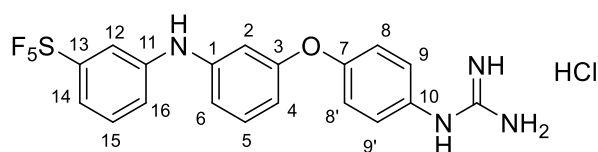
δ_{F} (376 MHz, CD₃OD): – 64.42 (s).

HRMS (m/z ESI⁺): Found 387.1438 ($M^+ + H$), C₂₀H₁₈N₄OF₃ Requires: 387.1433.

ν_{max} (ATR)/cm⁻¹: 3301 (NH), 3135 (NH), 1665, 1587 (C=N), 1490, 1486, 1335 (C-N), 1216 (C-O), 1161, 1116 (CF₃), 1067 (C-Cl), 976, 836, 785, 689.

HPLC: 97.5 % (t_{R} : 31.7 min).

1-(4-{3-[3-(Pentafluorosulphanyl)phenylamino]phenoxy}phenyl)guanidine hydrochloride (30)



Following Method A, **53** (272 mg, 0.42 mmol) was dissolved in 4M HCl in dioxane (1.27 mL, 5.06 mmol) and in additional dioxane (0.83 mL) until a final concentration of 0.2M was reached. After 6 h stirring at 55 °C, the reaction was adjudged complete (TLC), solvents were evaporated and the residue was purified by silica gel chromatography (CHCl₃:MeOH) to afford the pure hydrochloride salt as an orange solid (127 mg, 62%). **Mp**: 104-106 °C

δ_H (400 MHz, CD₃OD): 6.61 (dd, *J* = 8.2, 2.3 Hz, 1H, H-4), 6.78 (t, *J* = 2.2 Hz, 1H, H-2), 6.90 (dd, *J* = 8.1, 2.1 Hz, 1H, H-6), 7.11 (d, *J* = 8.9 Hz, H-8 and H-8'), 7.22 – 7.31 (m, 5H, H-9 and H-9', H-5 or H-15, H-14 and H-16), 7.34 – 7.38 (m, 1H, H-5 or H-15), 7.44 (t, *J* = 2.2 Hz, 1H, H-12).

δ_C (100 MHz, CD₃OD): 109.6 (CH Ar, C-2), 112.9 (CH Ar, C-4), 114.6 (CH Ar, C-6), 114.9 (p, *J* = 4.6 Hz, CH Ar, C-12), 118.0 (p, *J* = 4.7 Hz, CH Ar, C-14), 120.8 (CH Ar, C-16), 120.9 (2 CH Ar, C-8 and C-8'), 128.8 (2 CH Ar, C-9 and C-9'), 130.6 (CH Ar, C-5 or C-15), 130.8 (qC), 131.8 (CH Ar, C-5 or C-15), 145.6 (qC), 145.7 (qC), 155.9 (p, *J* = 16.4 Hz, qC, C-13), 158.2 (qC), 158.3 (qC), 159.2 (qC).

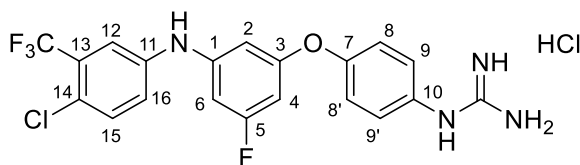
δ_F (376 MHz, CD₃OD): – 64.34 (s).

HRMS (m/z ESI⁺): Found 445.1124 (M⁺ + H), C₁₉H₁₈N₄OSF₅ Requires: 445.1121.

ν_{max} (ATR)/cm⁻¹: 3273 (NH), 3150 (NH), 1669, 1593 (C=N), 1487 (C-N), 1218 (C-O), 834 (SF₅), 567.

HPLC: 95.4 % (t_R: 32.3 min).

1-(4-{3-[4-Chloro-3-(trifluoromethyl)phenylamino]-5-fluorophenoxy}phenyl)guanidine hydrochloride (31)



Following Method A, **56** (166 mg, 0.26 mmol) was dissolved in 4M HCl in dioxane (0.78 mL, 3.12 mmol) and in additional dioxane (0.51 mL) until a final concentration of 0.2M was reached. After 8 h stirring at 55 °C, the reaction was adjudged complete (TLC), solvents were evaporated and the residue was purified by silica gel chromatography (CHCl₃:MeOH) to afford the pure hydrochloride salt as a white-brown solid (114 mg, 92%). **Mp**: 92-94 °C

δ_H (600 MHz, CD₃OD): 6.31 (dt, *J* = 9.9, 2.1 Hz, 1H, H-4), 6.55 (s, 1H, H-2), 6.58 (dt, *J* = 10.7, 2.0 Hz, 1H, H-6), 7.16 (d, *J* = 8.8 Hz, 2H, H-8 and H-8'), 7.29 (dd, *J* = 8.8, 2.7 Hz, 1H, H-16), 7.32 (d, *J* = 8.8 Hz, 2H, H-9 and H-9'), 7.40 (d, *J* = 2.6 Hz, 1H, H-12), 7.43 (d, *J* = 8.7 Hz, 1H, H-15).

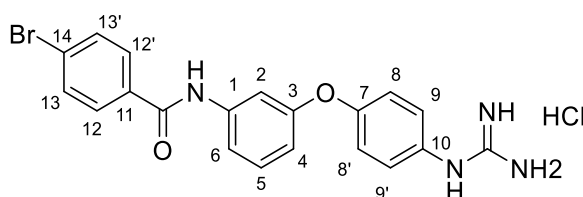
δ_C (150 MHz, CD₃OD): 99.4 (d, *J* = 25.7 Hz, CH Ar, C-4), 100.3 (d, *J* = 25.5 Hz, CH Ar, C-6), 103.8 (d, *J* = 2.6 Hz, CH Ar, C-2), 117.4 (q, *J* = 5.5 Hz, CH Ar, C-12), 121.8 (2 CH Ar, C-8 and C-8'), 122.7 (CH Ar, C-16), 123.4 (qC, C-14), 124.3 (d, *J* = 272.2 Hz, qCF₃), 128.9 (2 CH Ar, C-9 and C-9'), 129.8 (q, *J* = 31.0, qC, C-13), 131.7 (qC), 133.5 (CH Ar, C-15), 143.4 (qC), 146.7 (d, *J* = 13.2 Hz, qC, C-1 or C-3), 157.1 (qC), 158.4 (qC), 160.7 (d, *J* = 13.7 Hz, qC, C-1 or C-3), 165.8 (d, *J* = 243.4 Hz, qC, C-5).

δ_F (376 MHz, CD₃OD): - 64.67 (s), - 112.53 (s).

HRMS (*m/z* ESI⁺): Found 439.0945 (M⁺ + H), C₂₀H₁₆N₄OF₄Cl Requires: 439.0943.

ν_{max} (ATR)/cm⁻¹: 3285 (NH), 3139 (NH), 1666, 1601 (C=N), 1504, 1476, 1323 (CF₃), 1216 (C-O), 1112 (C-F), 1020 (C-Cl), 994, 823, 660.

HPLC: 95.7 % (*t*R: 33.1 min).

1-{4-[3-(4-bromobenzamido)phenoxy]phenyl}guanidine hydrochloride (32)³⁴

To a stirred solution of **59** (100 mg, 0.16 mmol) in EtOAc was added SnCl_4 (0.07 mL, 0.64 mmol). After 1 h of stirring at room temperature, the solvent and the excess of SnCl_4 were evaporated *in vacuo*. The product was purified by flash chromatography and the purified fraction was then dissolved in MeOH and recrystallized from Et_2O to obtain **32** as a pale yellow/white solid (40 mg, 54%). **Mp**: decomp. > 200 °C

δ_{H} (400 MHz, CD_3OD): 6.79 (d, $J=7.0$ Hz, 1H, H-4), 7.07 (d, $J = 8.2$ Hz, 2H, H-8 and H-8'), 7.18 (s, 1H, NH), 7.25 (d, $J = 8.3$ Hz, 2H, H-9 and H-9'), 7.30 – 7.38 (m, 2H, H-5, H-6), 7.55 (s, 1H, H-2), 7.62 (d, $J = 8.2$ Hz, 2H, H-13 and H-13'), 7.79 (d, $J = 8.2$ Hz, 2H, H-12 and H-12'), 9.31 (bs, 1H, NH), 10.16 (bs, 1H, NH).

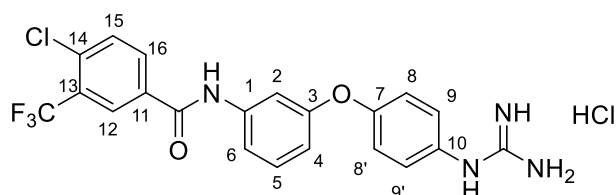
δ_{C} (100 MHz, CD_3OD): 112.8 (CH Ar, C-2), 116.0 (CH Ar, C-4), 117.2 (CH Ar, C-6), 121.0 (2 CH Ar, C-8 and C-8'), 127.3 (qC), 128.8 (2 CH Ar, C-9 and C-9'), 130.5 (2 CH Ar, C-12 and C-12'), 131.0 (CH Ar, C-5), 132.8 (2 CH Ar, C-13 and C-13'), 134.2 (qC), 135.0 (qC), 136.4 (qC), 147.6 (qC), 158.0 (qC), 158.3 (qC), 162.5 (qC).

HRMS (m/z ESI^+): Found 425.0613 ($\text{M}^+ + \text{H}$), $\text{C}_{20}\text{H}_{18}\text{N}_4\text{O}_2\text{Br}$ Requires 425.0613.

ν_{max} (ATR)/ cm^{-1} : 3432 (NH), 3332 (NH), 2508, 1658 (C=O), 1589 (C=N), 1565, 1547, 1507, 1486, 1250, 1218 (C-O), 1008 (C-Br), 838, 743.

HPLC: 99.9 % (t_{R} 31.08 min).

1-(4-{3-[4-Chloro-3-(trifluoromethyl)benzamido]phenoxy}phenyl)guanidine hydrochloride (33**)³⁴**



To a stirred solution of **60** (90 mg, 0.14 mmol) in EtOAc was added SnCl_4 (0.06 mL, 0.56 mmol). After 1 h of stirring at room temperature, the solvent and the excess of SnCl_4 were evaporated *in vacuo*. The product was purified by flash chromatography (CHCl_3 :MeOH) and the purified fraction was then dissolved in MeOH and recrystallized from Et_2O to obtain **33** as a white solid (32 mg, 47%). **Mp**: decomp. over 120 °C.

δ_{H} (400 MHz, CD_3OD): 6.86 (d, $J = 7.8$ Hz, 1H, H-4), 7.12 (d, $J = 8.6$ Hz, 2H, H-8 and H-8'), 7.30 (d, $J = 8.7$ Hz, 2H, H-9 and H-9'), 7.35 – 7.43 (m, 2H, H-5 and H-6), 7.62 (s, 1H, H-2), 7.77 (d, $J = 8.3$ Hz, 1H, H-15), 8.16 (dd, $J = 8.3, 1.5$ Hz, 1H, H-16), 8.32 (d, $J = 1.2$ Hz, 1H, H-12).

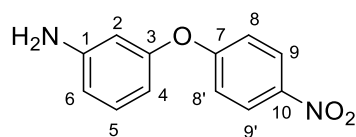
δ_{C} (100 MHz, CD_3OD): 112.9 (CH Ar, C-2), 116.3 (CH Ar, C-4), 117.3 (CH Ar, C-6), 121.0 (2 CH Ar, C-8 and C-8'), 124.0 (d, $J = 272.5$ Hz, qCF_3), 128.2 (q, $J = 5.5$ Hz, CH Ar, C-12), 128.9 (2 CH Ar, C-9 and C-9'), 129.4 (d, $J = 31.6$, qC, C-13), 130.9 (qC, C-14), 131.2 (CH Ar, C-5), 133.1 (CH Ar, C-15), 133.6 (CH Ar, C-16), 135.3 (qC), 136.5 (qC), 141.2 (qC), 158.1 (qC), 158.3 (qC), 158.4 (qC), 166.0 (qC).

δ_{F} (376 MHz, CD_3OD): – 64.11 (s).

HRMS (m/z ESI⁺): Found 449.0987 ($\text{M}^+ + \text{H}$), $\text{C}_{21}\text{H}_{17}\text{N}_4\text{O}_2\text{ClF}_3$ Requires: 449.0992.

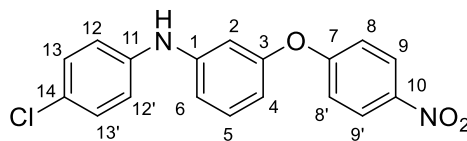
ν_{max} (ATR)/ cm^{-1} : 3413 (NH), 3343 (NH), 1602 (C=O), 1504 (C=N), 1450, 1412, 1261 (C-N), 1216 (C-O), 1134 (CF_3), 1035 (C-Cl), 924, 845.

HPLC: 98.9% (t_{R} 32.25 min).

3-(4-Nitrophenoxy)aniline (34)³⁵

1-Fluoro-4-nitrobenzene **36** (100 mg, 0.71 mmol, 1 eq.), 3-aminophenol **35** (117 mg, 1.07 mmol, 1.5 eq.) and K₂CO₃ (1148 mg, 1.07 mmol, 1.5 eq.) were dissolved in DMF (1.5 mL) and stirred at 80 °C for 12 h. The mixture was cooled to room temperature, washed with water and the organic layer extracted with EtOAc, washed with brine, dried over MgSO₄, concentrated under vacuum and purified by flash chromatography (silica gel, hexanes:EtOAc) to get **34** as a yellow solid (98 mg, 60%). **Mp**: 76-79 °C (lit. 79 °C)³⁵

δ_H (400 MHz, CDCl₃): 3.79 (bs, 2H, NH₂), 6.39 (app. t, *J* = 2.2 Hz, 1H, H-2), 6.45 (dd, *J* = 8.0, 2.2 Hz, 1H, H-4 or H-6), 6.55 (dd, *J* = 8.1, 2.2 Hz, 1H, H-4 or H-6), 7.02 (d, *J* = 9.2, 2H, H-8 and H-8'), 7.18 (t, *J* = 8.0 Hz, 1H, H-5), 8.19 (d, *J* = 9.2, 1H, H-9 and H-9').

4-Chloro-*N*-[3-(4-nitrophenoxy)phenyl]aniline (37)

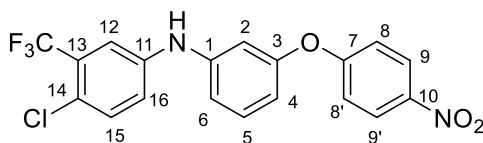
Following Method D, Pd₂(dba)₃ (3 mol%, 55 mg), BINAP (3 mol%, 37 mg), compound **34** (460 mg, 2 mmol), NaOtBu (269 mg, 2.8 mmol) and remaining 1-bromo-4-chlorobenzene (383 mg, 2 mmol) were mixed, followed syringe addition of toluene (4 mL). The mixture was heated at 90 °C for 24 h and usual work up and flash chromatography afforded **37** as a yellow-brown oil (457 mg, 67%).

δ_{H} (400 MHz, CDCl₃): 5.80 (s, 1H, NH), 6.60 (dd, J = 8.1, 1.8 Hz, 1H, H-4), 6.74 (t, J = 2.2 Hz, 1H, H-2), 6.87 (dd, J = 8.1, 2.1 Hz, 1H, H-6), 7.02 – 7.05 (m, 4H, H-8 and H-8', H-12 and H-12'), 7.23 – 7.30 (m, 3H, H-13, H-13' and H-5), 8.20 (d, J = 9.2 Hz, 2H, H-9 and H-9').

δ_{C} (100 MHz, CDCl₃): 108.8 (CH Ar, C-2), 112.6 (CH Ar, C-4), 114.0 (CH Ar, C-6), 117.3 (2 CH Ar, C-8 and C-8' or C-12 and C-12'), 120.3 (2 CH Ar, C-8 and C-8' or C-12 and C-12'), 126.1 (2 CH Ar, C-9 and C-9'), 127.0 (qC, C-14), 129.6 (2 CH Ar, C-13 and C-13'), 131.2 (CH Ar, C-5), 140.7 (qC), 142.8 (qC), 145.4 (qC), 156.0 (qC), 163.3 (qC).

HRMS (m/z ESI⁺): Found: 339.0544 (M⁺ - H), C₁₈H₁₂ClN₂O₃ Requires: 339.0536.

ν_{max} (ATR)/cm⁻¹: 3386 (NH), 3078, 2924, 2853, 1724, 1578, 1511, 1482 (NO₂), 1339 (NO₂), 1234 (CF₃), 1164, 1141, 1110, 1091 (C-Cl), 998, 847, 749.

4-Chloro-N-[3-(4-nitrophenoxy)phenyl]-3-(trifluoromethyl)aniline (38)

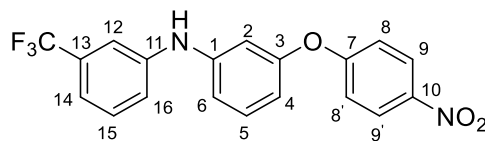
Following Method D, Pd₂(dba)₃ (3mol%, 55 mg), BINAP (3 mol%, 37 mg), compound **34** (460 mg, 2 mmol), NaOtBu (269 mg, 2.8 mmol) and remaining liquid 4-bromo-1-chloro-2-(trifluoromethyl)benzene (0.28 mL, 2 mmol) were mixed, followed syringe addition of toluene (4 mL). The mixture was heated at 90 °C for 24 h and usual work up and flash chromatography afforded **38** as a yellow solid (713 mg, 87%). **Mp**: 103-105 °C

δ_{H} (600 MHz, CDCl₃): 5.99 (s, 1H, NH), 6.70 (dd, J = 8.1, 2.1 Hz, 1H, H-4), 6.77 (s, 1H, H-2), 6.92 (dd, J = 8.1, 1.5 Hz, 1H, H-6), 7.05 (d, J = 9.3 Hz, 2H, H-8 and H-8'), 7.16 (dd, J = 8.7, 2.6 Hz, 1H, H-16), 7.33 (t, J = 8.1 Hz, 1H, H-5), 7.35 – 7.37 (m, 2H, H-12 and H-15), 8.21 (d, J = 9.3 Hz, 2H, H-9 and H-9').

δ_{C} (150 MHz, CDCl₃): 109.8 (CH Ar, C-2), 113.7 (CH Ar, C-4), 114.8 (CH Ar, C-6), 116.7 (q, J = 5.5 Hz, CH Ar, C-12), 117.4 (2 CH Ar, C-8 and C-8'), 121.5 (CH Ar, C-16), 122.6 (d, J = 273.2 Hz, qCF₃), 123.6 (qC, C-14), 126.0 (2 CH Ar, C-9 and C-9'), 129.2 (d, J = 31.2 Hz, qC, C-13), 131.3 (CH Ar, C-5), 132.4 (CH Ar, C-15), 141.3 (qC), 142.9 (qC), 143.9 (qC), 156.1 (qC), 162.9 (qC).

HRMS (m/z ESI⁺): Found: 407.0422 (M⁺ - H), C₁₉H₁₁ClF₃N₂O₃ Requires: 407.0416.

ν_{max} (ATR)/cm⁻¹: 3381 (NH), 1481 (NO₂), 1338, 1328, 1236 (C-O), 1131 (CF₃), 1110 (C-N), 847 (C-Cl), 749.

N-[3-(4-Nitrophenoxy)phenyl]-3-(trifluoromethyl)aniline (39)

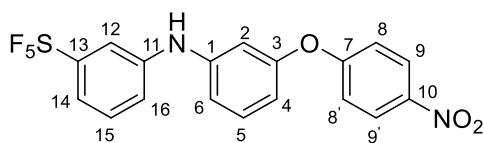
Following Method D, Pd₂(dba)₃ (3 mol%, 55 mg), BINAP (3 mol%, 37 mg), compound **34** (460 mg, 2 mmol), NaOtBu (269 mg, 2.8 mmol) and remaining liquid 1-bromo-3-(trifluoromethyl)benzene (0.28 mL, 2 mmol) were mixed, followed syringe addition of toluene (4 mL). The mixture was heated at 90 °C for 24 h and usual work up and flash chromatography afforded **39** as an orange solid (464 mg, 62%). **Mp**: 111-113 °C

δ_H (600 MHz, CDCl₃): 5.93 (bs, 1H, NH), 6.69 (dd, *J* = 8.1, 2.1 Hz, 1H, H-4), 6.80 (s, 1H, H-2), 6.94 (dd, *J* = 8.1, 1.8, 1H, H-6), 7.06 (d, *J* = 9.1 Hz, 2H, H-8 and H-8'), 7.19 (d, *J* = 7.8 Hz, 1H, H-14), 7.24 (d, *J* = 8.0 Hz, 1H, H-16), 7.30 (s, 1H, H-12), 7.33 (t, *J* = 8.1 Hz, 1H, H-5 or H-15), 7.38 (t, *J* = 7.9 Hz, 1H, H-5 or H-15), 8.21 (d, *J* = 9.1 Hz, H-9 and H-9').

δ_C (150 MHz, CDCl₃): 109.8 (CH Ar, C-2), 113.4 (CH Ar, C-4), 114.6 (q, *J* = 7.6 Hz, CH Ar, C-12), 114.8 (CH Ar, C-6), 117.5 (2 CH Ar, C-8 and C-8'), 118.3 (q, *J* = 7.5 Hz, CH Ar, C-14), 121.2 (CH Ar, C-16), 124.1 (d, *J* = 272.4 Hz, qCF₃), 126.1 (2 CH Ar, C-9 and C-9'), 130.2 (CH Ar, C-5 or C-15), 131.3 (CH Ar, C-5 or C-15), 132.1 (q, *J* = 32.2 Hz, qC, C-13), 142.96 (qC), 142.97 (qC), 144.4 (qC), 156.2 (qC), 163.2 (qC).

HRMS (m/z ESI⁺): Found 373.0816 (M⁺ - H), C₁₉H₁₂F₃N₂O₃ Requires: 373.0806.

ν_{max} (ATR)/cm⁻¹: 3377 (NH), 1598 (NO₂), 1528, 1476 (NO₂), 1413, 1340 (CF₃), 1243 (C-O), 1165, 1122 (C-N), 978, 848, 793.

N-[3-(4-Nitrophenoxy)phenyl]-3-(pentafluorosulfanyl)aniline (40)

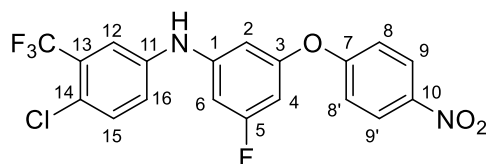
Following Method D, $\text{Pd}_2(\text{dba})_3$ (3 mol%, 14 mg), BINAP (3 mol%, 9 mg), compound **34** (115 mg, 0.5 mmol), NaOtBu (67 mg, 0.7 mmol) and remaining liquid (3-bromophenyl)sulfur pentafluoride (0.08 mL, 0.5 mmol) were mixed, followed syringe addition of toluene (1 mL). The mixture was heated at 90 °C for 24 h and usual work up and flash chromatography afforded **40** as a yellow solid (151 mg, 70%). **Mp**: 114-116 °C

δ_{H} (400 MHz, CDCl_3): 6.00 (s, 1H, NH), 6.70 (dd, $J = 8.1, 2.3$, Hz, 1H, H-4), 6.79 (t, $J = 2.2$ Hz, 1H, H-2), 6.93 (dd, $J = 8.1, 2.1$ Hz 1H, H-6), 7.06 (d, $J = 9.3$ Hz, 2H, H-8 and H-8'), 7.18 – 7.22 (m, 1H, H-16), 7.29 – 7.37 (m, 3H, H-5, H-15 and H-14) 7.44 (t, 1H, H-12), 8.21 (d, $J = 9.3$ Hz, 2H, H-9 and H-9').

δ_{C} (100 MHz, CDCl_3): 109.8 (CH Ar, C-2), 113.7 (CH Ar, C-4), 114.8 (CH Ar, C-6), 115.6 (p, $J = 4.8$ Hz, CH Ar, C-12), 117.5 (2 CH Ar, C-8 and C-8'), 118.9 (p, $J = 4.5$ Hz, CH Ar, C-14), 120.8 (CH Ar, C-16), 126.1 (2 CH Ar, C-9 and C-9'), 129.7 (CH Ar, C-15 or C-5), 131.4 (CH Ar, C-15 or C-5), 142.93 (qC), 142.95 (qC), 144.2 (qC), 155.0 (m, qC, C-13), 156.2 (qC), 163.1 (qC).

HRMS (m/z ESI⁺): Found 431.0488 ($\text{M}^+ - \text{H}$), $\text{C}_{18}\text{H}_{12}\text{F}_5\text{N}_2\text{O}_3\text{S}$ Requires: 431.0489.

ν_{max} (ATR)/ cm^{-1} : 3383 (NH), 3079, 2927, 1597 (NO_2), 1580 (NO_2), 1483, 1340 (C-N), 1236 (C-O), 821 (SF_5), 833 (SF_5), 804 (SF_5).

4-Chloro-*N*-[3-fluoro-5-(4-nitrophenoxy)phenyl]-3-(trifluoromethyl)aniline (41)

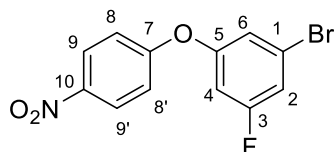
Following a modification of Method D, an oven-dried round bottom flask was charged with a magnetic stir bar, Pd₂(dba)₃ (3 mol%, 14 mg), BINAP (3 mol%, 9 mg), 4-chloro-3-(trifluoromethyl)aniline **45** (98 mg, 0.5 mmol, 1 eq.) and NaOtBu (67 mg, 0.7 mmol, 1.4 eq.). The tube was evacuated and refilled with argon (three times). Under a counterflow of argon, remaining **44** (156 g, 0.5 mmol, 1 eq.) in toluene (1 mL) was added dropwise to the mixture, followed by additional toluene (1 mL) by syringe. The tube was placed in a preheated oil bath at 90 °C and stirred overnight. The reaction mixture was then cooled to room temperature, diluted with EtOAc, filtered through a pad of Celite[®], and washed with water. The organic layer was washed with brine, dried over MgSO₄, concentrated under vacuum and purified by flash chromatography (silica gel, hexanes:EtOAc) to obtain **41** as a yellow solid (139 mg, 65%). **Mp**: 133-137 °C

δ_H (400 MHz, CDCl₃): 5.93 (bs, 1H NH), 6.39 (dt, *J* = 9.2, 2.1 Hz, 1H, H-4), 6.50 (s, 1H, H-2), 6.60 (dt, *J* = 10.2, 2.1 Hz, 1H, H-6), 7.10 (d, *J* = 9.2 Hz, 2H, H-8 and H-8'), 7.21 (dd, *J* = 8.6, 2.7 Hz, 1H, H-16), 7.39 (d, *J* = 2.6 Hz, 1H, H-12), 7.41 (d, *J* = 8.6 Hz, H-15), 8.24 (d, *J* = 9.2 Hz, 2H, H-9 and H-9').

δ_C (100 MHz, CDCl₃): 100.9 (d, *J* = 25.0 Hz, CH Ar, C-4), 100.91 (d, *J* = 25.0 Hz, CH Ar, C-6), 104.4 (d, *J* = 3.0 Hz, CH Ar, C-2), 118.1 (2 CH Ar, C-8 and C-8'), 118.2 (q, *J* = 5.5 Hz, CH Ar, C-12), 122.7 (d, *J* = 273.7, qCF₃), 123.0 (CH Ar, C-16), 125.1 (qC, C-14), 126.2 (2 CH Ar, C-9 and C-9'), 132.7 (CH Ar, C-15), 140.3 (qC), 143.5 (qC), 145.3 (d, *J* = 12.6 Hz, qC, C-1 or C-3), 157.4 (d, *J* = 13.4 Hz, qC, C-1 or C-3), 162.2 (qC), 164.6 (d, *J* = 247.1 Hz, qC, C-5).

HRMS (*m/z* ESI⁻): Found: 425.0319 (*M*⁻ - H), C₁₉H₁₀N₂O₃ClF₄ Requires: 425.0316.

ν_{max} (ATR)/cm⁻¹: 3359 (NH), 3085, 1592 (NO₂), 1480 (NO₂), 1340 (C-F), 1238 (CF₃), 1180 (C-O), 1156, 1143, 1131 (C-N), 1128, 1020 (C-Cl), 845, 828, 661.

1-Bromo-3-fluoro-5-(4-nitrophenoxy)benzene (44)

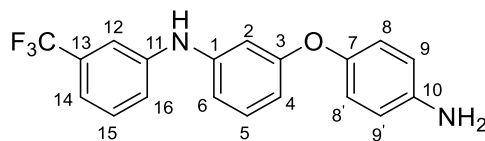
1-Fluoro-4-nitrobenzene **36** (2.00 g, 14.17 mmol, 1 eq.), 3-bromo-5-fluorophenol **43** (2.71 g, 14.17 mmol, 1 eq.) and K_2CO_3 (2.35 g, 17.00 mmol, 1.2 eq.) were dissolved in DMF (30 mL) and stirred at 80 °C for 12 h. The mixture was cooled to room temperature, washed with water and the organic layer extracted with EtOAc, washed with brine, dried over $MgSO_4$, concentrated under vacuum and purified by flash chromatography (hexanes:EtOAc) to get **44** as a yellow solid (4.42 g, 100%). **Mp**: 65-67 °C

δ_H (400 MHz, $CDCl_3$): 6.77 (dt, 1H, $J = 9.2, 2.2$, H-4), 7.04 (dd, 1H, $J = 2.9, 1.8$, H-6), 7.07-7.11 (d, $J = 9.2$ Hz, 2H, H-8 and H-8'), 7.12 – 7.15 (m, 1H, H-2), 8.26 (d, $J = 9.2$ Hz, 2H, H-9' and H-9').

δ_C (100 MHz, $CDCl_3$): 107.3 (d, $J = 24.4$, CH Ar, C-4), 116.1 (d, $J = 24.4$, CH Ar, C-2), 118.4 (2 CH Ar, C-8 and C-8'), 119.3 (d, $J = 3.6$, CH Ar, C-6), 123.6 (d, $J = 11.7$, qC, C-1), 126.3 (2 CH Ar, C-9 and C-9'), 143.9 (qC, C-7), 156.9 (d, $J = 11.7$, qC, C-5), 161.6 (qC, C-10), 163.3 (d, $J = 252.8$, qC, C-3).

HRMS (m/z APCI⁺): Found: 309.9526 ($M^+ - H$), $C_{12}H_6BrFNO_3$ Requires: 309.9521.

ν_{max} (ATR)/ cm^{-1} : 3080, 2922, 1580 (NO_2), 1342 (NO_2), 1237 (C-F), 1221 (C-O), 1122 (C-Br), 838, 848.

***N*-[3-(4-Aminophenoxy)phenyl]-3-(trifluoromethyl)aniline (46)**

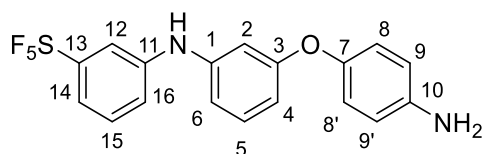
Following Method E, to a solution of nitro derivative **39** (409 mg, 1.1 mmol) in absolute ethanol (20 mL) was added 10% Pd/C mixture (10 mol%, 117 mg). After 24 h, usual work up followed by silica gel chromatography afforded **46** as brown oil (375 mg, 99 %).

δ_{H} (400 MHz, CDCl_3): 3.61 (bs, NH_2 , 2H), 5.84 (bs, 1H, NH), 6.57 (dd, $J = 8.0, 1.9$ Hz, 1H, H-4), 6.64 (t, $J = 2.2$ Hz, 1H, H-2), 6.68 (d, $J = 8.7$, 2H, H-9 and H-9'), 6.73 (dd, $J = 8.0, 1.5$ Hz, 1H, H-6), 6.89 (d, $J = 8.7$ Hz, 2H, H-8 and H-8'), 7.11 – 7.21 (m, 3H, H-5 or H-15, H-14 and H-16), 7.25 (s, 1H, H-12), 7.32 (t, $J = 7.9$ Hz, H-5 or H-15).

δ_{C} (100 MHz, CDCl_3): 107.2 (CH Ar, C-2), 110.9 (CH Ar, C-4), 112.2 (CH Ar, C-6), 113.7 (q, $J = 3.9$ Hz, CH Ar, C-12), 116.4 (2 CH Ar, C-9 and C-9'), 117.3 (q, $J = 3.9$ Hz, CH Ar, C-14), 120.4 (CH Ar, C-16), 121.5 (2 CH Ar, C-8 and C-8'), 124.2 (d, $J = 272.6$ Hz, qCF_3), 130.0 (CH Ar, C-5 or C-15), 130.5 (CH Ar, C-5 or C-15), 131.8 (d, $J = 32$ Hz, qC, C-13), 143.1 (qC), 143.3 (qC), 143.7 (qC), 148.2 (qC), 160.4 (qC).

HRMS (m/z ESI^+): Found 345.1211 ($\text{M}^+ + \text{H}$), $\text{C}_{19}\text{H}_{16}\text{F}_3\text{N}_2\text{O}$ Requires: 345.1209.

ν_{max} (ATR)/ cm^{-1} : 3316 (NH), 1594 (C-C), 1505, 1333 (C-O), 1217, 1173 (C-N), 1124 (C-Cl), 702.

***N*-[3-(4-Aminophenoxy)phenyl]-3-(pentafluorosulfanyl)aniline (47)**

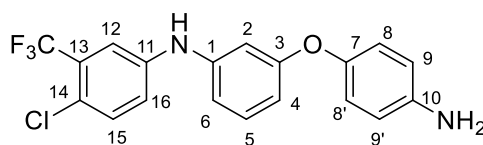
Following Method E, to a solution of nitro derivative **40** (75 mg, 0.17 mmol) in absolute ethanol (6 mL) was added 10% Pd/C mixture (10 mol%, 18 mg). After 24 h, usual work up followed by silica gel chromatography afforded **47** as a brown oil (41 mg, 60%).

δ_{H} (400 MHz, CDCl_3): 3.60 (s, 2H, NH_2), 5.86 (s, 1H, NH), 6.58 (dd, $J = 8.2, 2.2$ Hz, 1H, H-4), 6.62 (t, $J = 2.2$ Hz, 1H, H-2), 6.68 (d, $J = 8.7$ Hz, 1H, H-9 and H-9'), 6.72 (dd, $J = 8.0, 1.9$ Hz, 1H, H-6), 6.89 (d, $J = 8.7$ Hz, 2H, H-8 and H-8'), 7.12 – 7.15 (m, 1H, H-16), 7.19 (t, $J = 8.1$ Hz, 1H, H-5 or H-15), 7.23 – 7.31 (m, 2H, H-15 or H-5, and H-14), 7.39 (t, $J = 2.0$ Hz, 1H, H-12).

δ_{C} (100 MHz, CDCl_3): 107.3 (CH Ar, C-2), 111.2 (CH Ar, C-4), 112.2 (CH Ar, C-6), 114.7 (m, CH Ar, C-12), 116.4 (2 CH Ar, C-9 and C-9'), 118.0 (m, CH Ar, C-14), 120.0 (CH Ar, C-16), 121.5 (2 CH Ar, C-8 and C-8'), 129.5 (CH Ar, C-15 or C-5), 130.6 (CH Ar, C-15 or C-5), 143.1 (qC), 143.2 (qC), 143.7 (qC), 148.2 (qC), 155.1 (m, qC, C-13), 160.5 (qC).

HRMS (m/z ESI^+): Found 403.0909 ($\text{M}^+ + \text{H}$), $\text{C}_{18}\text{H}_{16}\text{F}_5\text{N}_2\text{OS}$ Requires: 403.0904.

ν_{max} (ATR)/ cm^{-1} : 3491 (NH_2), 3404 (NH_2), 2924, 1595, 1488 (C-N), 1212 (C-O), 829 (SF_5), 806 (SF_5), 769, 681, 667.

***N*-[3-(4-Aminophenoxy)phenyl]-4-chloro-3-(trifluoromethyl)aniline (**48**)**

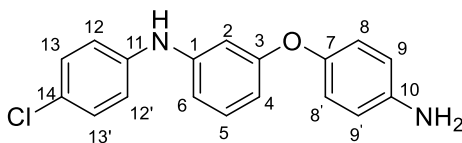
Following Method F, a mixture of **38** (570 mg, 1.32mmol) and $\text{SnCl}_2 \cdot 2\text{H}_2\text{O}$ (1890 mg, 8.37 mmol) in 3 mL of absolute ethanol and few drops of EtOAc to increase solubility, were heated at 70 °C for 3 h. Usual work up and flash chromatography afforded **48** as a brown solid (340 mg, 68%). **Mp**: 102-104 °C

δ_{H} (600 MHz, CDCl_3): 3.58 (s, 2H, NH_2), 5.80 (s, 1H, NH), 6.58 – 6.61 (m, 2H, H-4 and H-2), 6.67 – 6.71 (m, 3H, H-9, H-9' and H-6), 6.89 (d, $J = 8.7$ Hz, 2H, H-8 and H-8'), 7.09 (dd, $J = 8.6, 2.6$ Hz, 1H, H-16), 7.20 (t, $J = 8.1$ Hz, 1H, H-5), 7.30 – 7.32 (m, 2H, H-12 and H-15).

δ_{C} (150 MHz, CDCl_3): 107.4 (CH Ar, C-2), 111.3 (CH Ar, C-4), 112.4 (CH Ar, C-6), 116.0 (q, $J = 5.4$ Hz, CH Ar, C-12), 116.4 (2 CH Ar, C-9 and C-9'), 120.9 (CH Ar, C-16), 122.0 (2 CH Ar, C-8 and C-8'), 122.7 (qC C-14), 122.9 (q, $J = 273.4$ Hz, qCF_3), 129.1 (q, $J = 31.2$ Hz, qC, C-13), 130.7 (CH Ar, C-5), 132.4 (CH Ar, C-15), 142.2 (qC, C-11), 142.9 (qC, C-1), 143.2 (qC, C-10), 148.1 (qC, C-7), 160.5 (qC, C-3).

HRMS (m/z ESI^+): Found: 379.0824 ($\text{M}^+ + \text{H}$), $\text{C}_{19}\text{H}_{15}\text{ClF}_3\text{N}_2\text{O}$ Requires: 379.0820.

ν_{max} (ATR)/ cm^{-1} : 3383 (NH), 2926, 1725, 1594, 1504 (C-N), 1482 (C-N), 1329, 1250, 1172 (C-Cl), 1130 (C-O), 1112 (CF_3), 825, 681, 666.

***N*-[3-(4-Aminophenoxy)phenyl]-4-chloroaniline (**49**)**

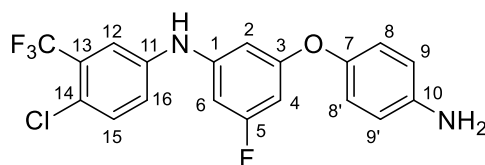
Following Method F, a mixture of **37** (208 mg, 0.61 mmol) and $\text{SnCl}_2 \cdot 2\text{H}_2\text{O}$ (826 mg, 3.7 mmol) in 2 mL of absolute ethanol were heated at 70 °C for 3 h. Usual work up and flash chromatography afforded **49** as a brown solid (139 mg, 73%). **Mp**: 119-121 °C

δ_{H} (400 MHz, CDCl_3): 3.58 (bs, 2H, NH_2), 5.65 (bs, 1H, NH), 6.48 (dd, $J = 8.2, 2.3$ Hz, 1H, H-4), 6.61 (t, $J = 2.2$ Hz, 1H, H-2), 6.66 – 6.70 (m, 3H, H-6, H-9 and H-9'), 6.88 (d, $J = 8.8$ Hz, 2H, H-8 and H-8'), 6.98 (d, $J = 8.8$ Hz, 2H, H-12 and H-12'), 7.15 (t, $J = 8.1$ Hz, 1H, H-5), 7.20 (d, $J = 8.8$ Hz, 2H, H-13 and H-13').

δ_{C} (100 MHz, CDCl_3): 106.6 (CH Ar, C-2), 110.0 (CH Ar, C-4), 111.4 (CH Ar, C-6), 116.3 (2 CH Ar, C-9 and C-9'), 119.4 (2 CH Ar, C-12 and C-12'), 121.4 (2 CH Ar, C-8 and C-8'), 126.0 (qC, C-14), 129.4 (2 CH Ar, C-13 and C-13'), 130.4 (CH Ar, C-5), 141.5 (qC), 143.0 (qC), 144.3 (qC), 148.4 (qC), 160.3 (qC).

HRMS (m/z ESI^+): Found 311.0954 ($\text{M}^+ + \text{H}$), $\text{C}_{18}\text{H}_{16}\text{ClN}_2\text{O}$ Requires: 311.0946.

ν_{max} (ATR)/ cm^{-1} : 3409 (NH), 3286 (NH), 3377 (NH), 1588, 1527, 1484 (C-O), 1324 (C-N), 1250, 1206, 1139 (C-Cl), 994, 972, 760.

***N*-(3-(4-Aminophenoxy)-5-fluorophenyl)-4-chloro-3-(trifluoromethyl)aniline (50)**

Following Method F, a mixture of **41** (139 mg, 0.33 mmol) and $\text{SnCl}_2 \cdot 2\text{H}_2\text{O}$ (441 mg, 1.95 mmol) in 3 mL of absolute ethanol, were heated at 70 °C for 3 h. Usual work up and flash chromatography afforded **50** as a brown solid (92 mg, 70%). **Mp**: 113-115 °C

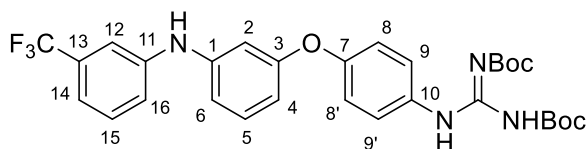
δ_{H} (400 MHz, CDCl_3): 3.63 (bs, 1H, NH_2), 5.83 (bs, 1H, NH), 6.25 (dt, $J = 10.2$, 2.1 Hz, 1H, H-4), 6.34 (s, 1H, H-2), 6.39 (dt, $J = 10.2$, 2.0 Hz, 1H, H-6), 6.69 (d, $J = 8.7$ Hz, 2H, H-9 and H-9'), 6.88 (d, $J = 8.7$ Hz, 2H, H-8 and H-8'), 7.14 (dd, $J = 8.6$, 2.7 Hz, 1H, H-16), 7.33 (d, $J = 2.6$ Hz, 1H, H-12), 7.35 (d, $J = 8.7$ Hz, H-15).

δ_{C} (100 MHz, CDCl_3): 98.2 (d, $J = 25.6$ Hz, CH Ar, C-4), 98.5 (d, $J = 25.4$ Hz, CH Ar, C-6), 101.8 (d, $J = 2.1$ Hz, CH Ar, C-2), 116.4 (2 CH Ar, C-9 and C-9'), 117.2 (q, $J = 5.4$ Hz, CH Ar, C-12), 121.8 (2 CH Ar, C-8 and C-8'), 122.0 (CH Ar, C-16), 122.8 (d, $J = 273.3$ Hz, qCF_3), 123.9 (qC, C-14), 129.3 (q, $J = 31.2$ Hz, qC, C-13), 132.5 (CH Ar, C-15), 141.2 (qC), 143.7 (qC), 144.1 (d, $J = 12.7$ Hz, qC, C-1 or C-3), 147.4 (qC), 161.8 (d, $J = 13.3$ Hz, qC, C-1 or C-3), 164.5 (d, $J = 244.3$ Hz, qC, C-5).

HRMS (m/z ESI⁺): Found: 395.0573 ($\text{M}^+ - \text{H}$), $\text{C}_{19}\text{H}_{12}\text{N}_2\text{OClF}_4$ Requires: 395.0574.

ν_{max} (ATR)/ cm^{-1} : 3361 (NH), 3256 (NH), 1629, 1601, 1578, 1505, 1474, 1322 (CF_3), 1259, 1110 (C-F), 1004 (C-Cl), 822, 748, 670.

1,2-Di-(*tert*-butoxycarbonyl)-3-(4-{3-[3-(trifluoromethyl)phenylamino]phenoxy}phenyl) guanidine (52**)**



Following Method G, HgCl_2 (22 mg, 0.08 mmol) was added over a solution of **46** (25 mg, 0.07 mmol), *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methyliothiourea **51** (23 mg, 0.08 mmol), and NEt_3 (0.03 mL, 0.25 mmol) in CH_2Cl_2 (1 mL). The reaction was stirred at room temperature for 24 h, then work up and silica gel chromatography (hexanes:EtOAc) afforded **52** as a yellow solid (30 mg, 73%). **Mp**: 60-63 °C

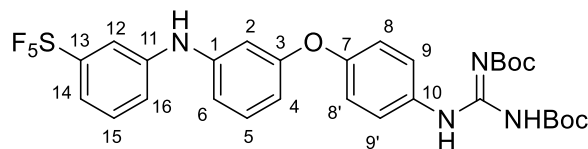
δ_{H} (400 MHz, CDCl_3): 1.48 (s, 9H, $(\text{CH}_3)_3$), 1.54 (s, 9H, $(\text{CH}_3)_3$), 5.99 (bs, 1H, NH), 6.58 (dd, $J = 8.1, 2.3$ Hz, 1H, H-4), 6.67 (t, $J = 2.2$ Hz, 1H, H-2), 6.80 (dd, $J = 8.0, 1.4$ Hz, 1H, H-6), 6.98 (d, $J = 8.9$ Hz, 2H, H-8 and H-8'), 7.12 (d, $J = 7.6$ Hz, 1H, H-14), 7.18 – 7.22 (m, 2H, H-5 or H-15 and H-16), 7.25 (s, 1H, H-12), 7.33 (t, $J = 7.9$ Hz, 1H, H-5 or H-15), 7.54 (d, $J = 8.9$ Hz, 2H, H-9 and H-9'), 10.28 (bs, NH), 11.64 (bs, NH).

δ_{C} (100 MHz, CDCl_3): 28.2 ($(\text{CH}_3)_3$), 28.3 ($(\text{CH}_3)_3$), 79.8 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 83.9 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 108.4 (CH Ar, C-2), 111.8 (CH Ar, C-4), 112.8 (CH Ar, C-6), 114.0 (q, $J = 3.9$ Hz, CH Ar, C-12), 117.4 (q, $J = 3.8$ Hz, CH Ar, C-14), 119.8 (2 CH Ar, C-8 and C-8'), 120.4 (CH Ar, C-16), 124.1 (2 CH Ar, C-9 and C-9'), 124.2 (d, $J = 272.4$ Hz, qCF₃), 129.9 (CH Ar, C-5 or C-15), 130.6 (CH Ar, C-5 or C-15), 131.9 (q, $J = 32.1$ Hz, qC, C-13), 132.6 (qC), 143.6 (qC), 143.7 (qC), 153.5 (qC), 153.6 (qC), 153.8 (qC), 159.0 (qC), 163.6 (qC).

HRMS (m/z ESI⁺): Found 587.2485 ($\text{M}^+ + \text{H}$), $\text{C}_{30}\text{H}_{34}\text{N}_4\text{O}_5\text{F}_3$ Requires: 587.2481.

ν_{max} (ATR)/ cm^{-1} : 3295 (NH), 3261 (NH), 2979, 1719 (C=O), 1594 (C=N), 1407, 1325, 1304, 1231, 1145 (CF₃), 1110 (C-O), 1056 (C-Cl), 1028, 998, 978, 774, 656.

1,2-Di-(*tert*-butoxycarbonyl)-3-(4-{3-[3-(pentafluorosulphanylmethyl)phenylamino] phenoxy}phenyl)guanidine (53)



Following Method G, HgCl_2 (28 mg, 0.11 mmol) was added over a solution of **47** (40 mg, 0.1 mmol), *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methylisothiourea **51** (30 mg, 0.11 mmol), and NEt_3 (0.05 mL, 0.35 mmol) in CH_2Cl_2 (1 mL). The reaction was stirred at room temperature for 24 h, then work up and silica gel chromatography (hexanes:EtOAc) afforded **53** as a yellow solid (48 mg, 75%). **Mp**: 70-73 °C

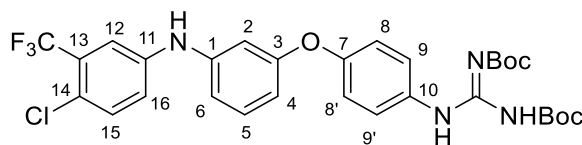
δ_{H} (600 MHz, CDCl_3): 1.48 (s, 9H, $(\text{CH}_3)_3$), 1.54 (s, 9H, $(\text{CH}_3)_3$), 6.03 (bs, 1H, NH), 6.60 (dd, $J = 8.1, 2.1$ Hz, 1H, H-4), 6.66 (t, $J = 1.9$ Hz, 1H, 1H, H-2), 6.79 (dd, $J = 7.9, 1.7$ Hz, 1H, H-6), 6.97 (d, $J = 8.8$ Hz, 2H, H-8 and H-8'), 7.16 (d, $J = 7.6$ Hz, 1H, H-16), 7.21 (t, $J = 8.1$ Hz, 1H, H-5 or H-15), 7.24 – 7.26 (m, 1H, H-14), 7.30 (t, $J = 8.1$ Hz, H-5 or H-15), 7.39 (s, 1H, H-12), 7.54 (d, $J = 8.8$ Hz, 2H, H-9 and H-9'), 10.28 (bs, NH), 11.63 (bs, NH).

δ_{C} (150 MHz, CDCl_3): 28.0 ($(\text{CH}_3)_3$), 28.3 ($(\text{CH}_3)_3$), 80.0 (q, $\underline{\text{C}}(\text{CH}_3)_3$), 83.9 (q, $\underline{\text{C}}(\text{CH}_3)_3$), 108.5 (CH Ar, C-2), 112.1 (CH Ar, C-4), 112.8 (CH Ar, C-6), 115.1 (m, CH Ar, C-12), 118.1 (m, CH Ar, C-14), 119.9 (2 CH Ar, C-8 and C-8'), 120.0 (CH Ar, C-16), 124.2 (2 CH Ar, C-9 and C-9'), 129.5 (CH Ar, C-5 or C-15), 130.7 (CH Ar, C-5 or C-15), 132.6 (qC), 143.5 (qC), 143.6 (qC), 153.5 (qC), 153.7 (qC), 153.9 (qC), 155.1 (m, qC, C-13), 159.0 (qC), 163.6 (qC).

HRMS (m/z ESI⁺): Found 645.2171 ($\text{M}^+ + \text{H}$), $\text{C}_{29}\text{H}_{34}\text{F}_5\text{N}_4\text{O}_5\text{S}$ Requires: 645.2165.

ν_{max} (ATR)/ cm^{-1} : 3303 (NH), 2979, 1719 (C=O), 1594 (C=N), 1486, 1486, 1407, 1303, 1293, 1213, 1146 (C-N), 1110 (C-O), 839 (SF_5), 806 (SF_5), 772 (SF_5).

1,2-Di-(*tert*-butoxycarbonyl)-3-(4-{3-[4-chloro-3-(trifluoromethyl)phenylamino]phenoxy} phenyl)guanidine (54**)**



Following Method G, HgCl_2 (206 mg, 0.76 mmol) was added over a solution of **48** (277 mg, 0.73 mmol), *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methylisothiurea **51** (221 mg, 0.76 mmol) and NEt_3 (0.35 mL, 2.56 mmol) in CH_2Cl_2 (4 mL). The reaction was stirred at room temperature for 24 h, then work up and silica gel chromatography (hexanes:EtOAc) afforded **54** (363 mg, 80%) as a white solid. **Mp**: 92-94 °C

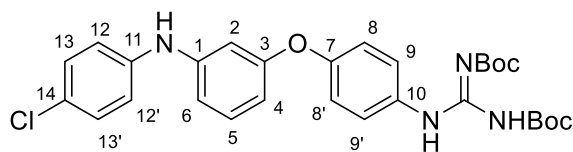
δ_{H} (600 MHz, CDCl_3): 1.48 (s, 9H, $(\text{CH}_3)_3$), 1.54 (s, 9H, $(\text{CH}_3)_3$), 6.01 (s, 1H, NH), 6.61 (dd, $J = 8.2, 1.5$ Hz, 1H, H-4), 6.65 (s, 1H, H-2), 6.78 (dd, $J = 8.0, 1.2$ Hz, 1H, H-6), 6.98 (d, $J = 8.8$ Hz, 2H, H-8 and H-8'), 7.12 (dd, $J = 8.6, 2.5$ Hz, 1H, H-16), 7.22 (t, $J = 8.1$ Hz, 1H, H-5), 7.31-7.33 (m, 2H, H-12 and H-15), 7.54 (d, $J = 8.7$ Hz, 2H, H-9 and H-9'), 10.29 (s, 1H, NH), 11.65 (s, 1H, NH).

δ_{C} (150 MHz, CDCl_3): 28.2 ($(\text{CH}_3)_3$), 28.3 ($(\text{CH}_3)_3$), 78.0 (qC, $\text{C}(\text{CH}_3)_3$), 84.0 (qC, $\text{C}(\text{CH}_3)_3$), 108.6 (CH Ar, C-2), 112.3 (CH Ar, C-4), 113.0 (CH Ar, C-6), 116.3 (d, $J = 5.1$ Hz, CH Ar, C-12), 120.0 (2 CH Ar, C-8 and C-8'), 120.9 (CH Ar, C-16), 122.8 (CH Ar, C-14), 122.9 (q, $J = 273.6$ Hz, qCF_3), 124.1 (2 CH Ar, C-9 and C-9'), 129.2 (q, $J = 31.2$, qC, C-13), 130.7 (CH Ar, C-5), 132.4 (CH Ar, C-15), 132.6 (qC), 142.1 (qC), 143.3 (qC), 153.5 (qC), 153.6 (qC), 153.8 (qC), 159.1 (qC), 163.4 (qC).

HRMS (m/z ESI^+): Found: 621.2102 ($\text{M}^+ + \text{H}$), $\text{C}_{30}\text{H}_{33}\text{ClF}_3\text{N}_4\text{O}_5$ Requires: 621.2086.

ν_{max} (ATR)/ cm^{-1} : 3260 (NH), 2979, 2930, 1719 (C=O), 1596 (C=N), 1482, 1406, 1304, 1233 (C-O), 1213, 1141 (CF_3), 1109 (C-Cl), 1027, 818, 772.

1,2-Di-(*tert*-butoxycarbonyl)-3-(4-{3-[4-chlorophenylamino]phenoxy}phenyl)guanidine (55)



Following Method G, HgCl_2 (41 mg, 0.15 mmol) was added over a solution of **49** (45 mg, 0.14 mmol), *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methylothiourea **51** (44 mg, 0.15 mmol), and NEt_3 (0.07 mL, 0.49 mmol) in CH_2Cl_2 (2 mL). The reaction was stirred at room temperature for 24 h, then work up and silica gel chromatography (hexanes:EtOAc) afforded **55** as an orange solid (95 mg, 86%). **Mp**: 120 °C

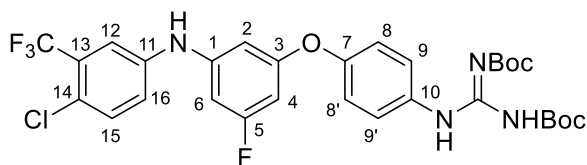
δ_{H} (400 MHz, CDCl_3): 1.49 (s, 9H, $(\text{CH}_3)_3$), 1.54 (s, 9H, $(\text{CH}_3)_3$), 5.73 (bs, 1H, NH), 6.54 (dd, $J = 8.1, 2.3$ Hz, 1H, H-4), 6.65 (t, $J = 2.2$ Hz, 1H, H-2), 6.75 (dd, $J = 8.1, 2.1$ Hz, 1H, H-6), 6.98 – 7.01 (m, 4H, H-12, H-12', H-8 and H-8'), 7.16 – 7.22 (m, 3H, H-13, H-13' and H-5), 7.56 (d, $J = 9.0$ Hz, 2H, H-9 and H-9'), 10.29 (bs, 1H, NH), 11.64 (bs, 1H, NH).

δ_{C} (100 MHz, CDCl_3): 28.2 ($(\text{CH}_3)_3$), 28.3 ($(\text{CH}_3)_3$), 79.8 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 83.9 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 107.6 (CH Ar, C-2), 111.1 (CH Ar, C-4), 112.0 (CH Ar, C-6), 119.6 (2 CH Ar, C-12 and C-12' or C-8 and C-8'), 119.8 (2 CH Ar, C-12 and C-12' or C-8 and C-8'), 123.9 (2 CH Ar, C-9 and C-9'), 126.2 (qC, C-14), 129.5 (2 CH Ar, C-13 and C-13'), 130.5 (CH Ar, C-5), 132.5 (qC), 141.3 (qC), 144.6 (qC), 153.5 (qC), 153.73 (qC), 153.74 (qC), 158.9 (qC), 163.7 (qC).

HRMS (m/z ESI^+): Found 553.2190 ($\text{M}^+ + \text{H}$), $\text{C}_{29}\text{H}_{34}\text{ClN}_4\text{O}_5$ Requires: 553.2212.

ν_{max} (ATR)/ cm^{-1} : 3296 (NH), 2925, 1718 (C=O), 1649, 1628 (C=O), 1589 (C=N), 1485, 1390, 1367, 1330, 1310, 1258, 1148 (C-O), 1113 (C-N), 1096 (C-Cl), 829, 765.

1,2-Di-(*tert*-butoxycarbonyl)-3-(4-{3-[4-chloro-3-(trifluoromethyl)phenylamino]-5-fluorophenoxy}phenyl)guanidine (56**)**



Following Method G, HgCl_2 (429 mg, 1.58 mmol) was added over a solution of **50** (604 mg, 1.52 mmol), *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methylisothiurea **51** (460 mg, 1.58 mmol), and NEt_3 (0.74 mL, 5.32 mmol) in CH_2Cl_2 (8 mL). The reaction was stirred at room temperature for 24 h, then work up and silica gel chromatography (hexanes:EtOAc), afforded **56** as a yellow solid (680 mg, 70%). **Mp**: 75 - 77 °C

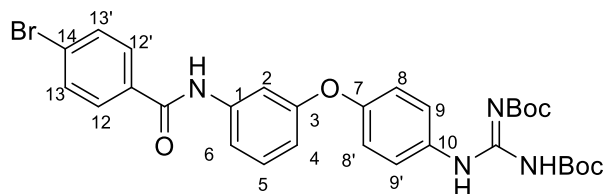
δ_{H} (400 MHz, CDCl_3): 1.47 (s, 9H, $(\text{CH}_3)_3$), 1.54 (s, 9H, $(\text{CH}_3)_3$), 6.09 (s, 1H, NH), 6.27 (dt, $J = 9.9, 2.1$ Hz, 1H, H-4), 6.36 (s, 1H, H-2), 6.44 (dt, $J = 10.3, 2.1$ Hz, 1H, H-6), 6.92 (d, $J = 8.9$ Hz, 2H, H-8 and H-8'), 7.14 (dd, $J = 8.7, 2.7$ Hz, 1H, H-16), 7.31 (d, $J = 2.7$ Hz, 1H, H-12), 7.35 (d, $J = 8.7$ Hz, 1H, H-15), 7.56 (d, $J = 8.9$ Hz, 2H, H-9 and H-9'), 10.32 (s, 1H, NH), 11.65 (s, 1H, NH).

δ_{C} (100 MHz, CDCl_3): 28.2 ($(\text{CH}_3)_3$), 28.3 ($(\text{CH}_3)_3$), 80.0 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 84.0 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 98.9 (CH Ar, d, $J = 25.5$ Hz, C-4), 99.0 (CH Ar, d, $J = 25.5$ Hz, C-6), 102.8 (CH Ar, d, $J = 2.7$ Hz, C-2), 117.3 (q, $J = 5.5$ Hz, CH Ar, C-12), 120.4 (2 CH Ar, C-8 and C-8'), 122.0 (CH Ar, C-16), 122.8 (d, $J = 273.6$ Hz, qCF_3), 123.9 (qC, C-14), 124.4 (2 CH Ar, C-9 and C-9'), 129.3 (q, $J = 31.2$ Hz, qC, C-13), 132.5 (CH Ar, C-15), 133.2 (qC), 141.0 (qC), 144.4 (d, $J = 12.9$ Hz, qC, C-1 or C-3), 152.7 (qC), 153.5 (qC), 153.9 (qC), 160.3 (d, $J = 13.5$ Hz, qC, C-1 or C-3), 164.4 (d, $J = 244.8$ Hz, qC, C-5), 163.6 (qC).

HRMS (m/z ESI^+): Found: 639.1996 ($\text{M}^+ + \text{H}$), $\text{C}_{30}\text{H}_{32}\text{N}_4\text{O}_5\text{ClF}_4$ Requires: 639.1997.

ν_{max} (ATR)/ cm^{-1} : 3257 (NH), 2978, 1719 (C=O), 1598 (C=O), 1476, 1401, 1368, 1320 (CF_3), 1145 (C-F), 1109 (C-Cl), 808, 768.

1,2-Di-(*tert*-butoxycarbonyl)-3-(4-{3-[4-bromobenzamido]phenoxy}phenyl)guanidine (59)



Following Method H, to a solution of 4-bromobenzoic acid **57** (217 mg, 1.08 mmol) in CH_2Cl_2 (3 mL), oxalyl chloride (0.15 mL, 1.81 mmol) and DMF (3 drops) were added and the reaction was stirred overnight. Any remaining oxalyl chloride and DMF were removed under vacuum to obtain a yellow residue that was dissolved in CH_2Cl_2 and cooled to 0 °C. Following, compound **17** (400 mg, 0.90 mmol) in CH_2Cl_2 (1.5 mL) and NEt_3 (0.30 mL, 2.16 mmol) were added dropwise. After 12 h, workup and silica gel chromatography afforded **59** (389 mg, 69%) as a white solid. **Mp**: 218 - 223 °C.

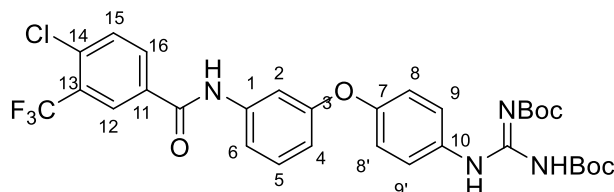
δ_{H} (600 MHz, CDCl_3): 1.55 (s, 9H, $(\text{CH}_3)_3$), 1.56 (s, 9H, $(\text{CH}_3)_3$), 6.81 (dd, $J = 8.2, 2.2$ Hz, 1H, H-4), 6.96 (d, $J = 8.8$ Hz, 2H, H-8 and H-8'), 7.03 (s, 1H, H-2), 7.30 (t, $J = 8.2$ Hz, 1H, H-5), 7.50 (d, $J = 8.8$ Hz, 2H, H-9 and H-9'), 7.55 (d, $J = 7.9$ Hz, 1H, H-6), 7.60 (d, $J = 8.5$ Hz, 2H, H-13 and H-13'), 7.75 (d, $J = 8.4$ Hz, 2H, H-12 and H-12'), 8.10 (s, 1H, NH), 10.26 (s, 1H, NH), 11.64 (s, NH).

δ_{C} (150 MHz, CDCl_3): 28.2 ($(\text{CH}_3)_3$), 28.3 ($(\text{CH}_3)_3$), 80.0 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 84.0 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 109.8 (CH Ar, C-2), 114.7 (CH Ar, C-4), 115.0 (CH Ar, C-6), 120.2 (2 CH Ar, C-8 and C-8'), 124.4 (2 CH Ar, C-9 and C-9'), 126.8 (qC), 129.0 (2 CH Ar, C-12 and C-12'), 130.4 (CH Ar, C-5), 132.1 (2 CH Ar, C-13 and C-13'), 132.6 (qC), 133.8 (qC), 139.3 (qC), 153.5 (qC), 153.6 (qC), 154.0 (qC), 158.5 (qC), 163.6 (qC), 164.9 (qC).

HRMS (m/z ESI⁻): Found 623.1497 ($\text{M}^- - \text{H}$), $\text{C}_{30}\text{H}_{32}\text{BrN}_4\text{O}_6$ Requires: 623.1505.

ν_{max} (ATR)/ cm^{-1} : 3304 (NH), 3258 (NH), 2980, 2927, 2854, 1719, 1641 (C=O), 1596 (C=N), 1542, 1505, 1487, 1408, 1302, 1146 (C-O), 1108 (C-Br), 1058, 729.

1,2-Di-(*tert*-butoxycarbonyl)-3-(4-{3-[4-chloro-3-(trifluoromethyl)benzamido]phenoxy} phenyl)guanidine (60)



Following Method H, to a solution of 4-chloro-3-trifluoromethylbenzoic acid **58** (183 mg, 0.82 mmol) in CH₂Cl₂ (3 mL), oxalyl chloride (0.1 mL, 1.36 mmol) and DMF (3 drops) were added and the reaction was stirred overnight. Any remaining oxalyl chloride and DMF were removed under vacuum to obtain a yellow residue that was dissolved in CH₂Cl₂ and cooled to 0 °C. Following, compound **17** (300 mg, 0.68 mmol) dissolved in CH₂Cl₂ (1.13 mL) and NEt₃ (0.23 mL, 1.63 mmol) were added dropwise. After 12 h, workup and silica gel chromatography afforded **60** (225 mg, 51%) as a yellow oil.

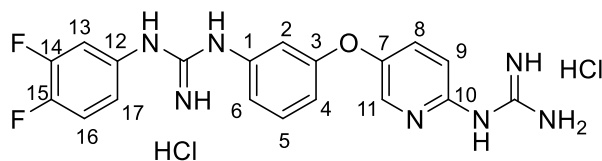
δ_{H} (600 MHz, CDCl₃): 1.55 (s, 9H, (CH₃)₃), 1.57 (s, 9H, (CH₃)₃), 6.81 (dd, J = 8.2, 1.7 Hz, 1H, H-4), 6.86 (d, J = 8.8 Hz, 2H, H-8 and H-8' or H-9 and H-9'), 6.96 (s, 1H, H-2), 7.28 (t, J = 8.2 Hz, 1H, H-5), 7.36 (d, J = 8.8 Hz, 2H, H-8 and H-8' or H-9 and H-9'), 7.55 - 7.58 (m, 2H, H-15, H-6), 7.98 (dd, J = 8.3, 1.8 Hz, H-16), 8.25 (d, J = 1.6 Hz, 1H, H-12), 8.74 (s, 1H, NH), 10.19 (s, 1H, NH), 11.62 (s, 1H, NH).

δ_{C} (150 MHz, CDCl₃): 28.2 (2(CH₃)₃), 84.4 (qC, 2 C(CH₃)₃), 109.9 (CH Ar, C-2), 114.9 (CH Ar, C-4), 115.2 (CH Ar, C-6), 120.6 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 122.7 (d, J = 273.9 Hz, qCF₃), 124.3 (qC), 125.1 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 127.1 (m, CH Ar, C-12), 130.4 (CH Ar, C-5), 131.91 (CH Ar, C-15), 131.95 (CH Ar, C-16), 132.0 (qC), 133.7 (qC), 136.0 (qC), 139.3 (qC), 153.3 (qC), 154.2 (qC), 154.3 (qC), 158.4 (qC), 158.5 (qC), 163.7 (qC).

HRMS (m/z ESI⁺): Found 649.2042 (M^+ + H), C₃₁H₃₃ClF₃N₄O₆ Requires: 649.2041.

ν_{max} (ATR)/cm⁻¹: 3253 (NH), 2979, 2932, 1720 (C=O), 1648 (C=O), 1598, 1406, 1324, 1240, 1213, 1143 (CF₃), 1108 (C-Cl), 1056, 1035, 909, 730.

1-(3,4-Di-fluorophenyl)-2-[3-(4-guanidinopyridin-3-yloxy)phenyl]guanidine dihydrochloride (61)



Following Method A, **81** (176 mg, 0.25 mmol) was dissolved in 4M HCl in dioxane (1.14 mL, 4.54 mmol) and in additional dioxane (0.11 mL) until a final concentration of 0.2 M was reached. After 8 h stirring at 55 °C, the reaction was adjudged complete (TLC), solvents were evaporated and the residue was purified by silica gel chromatography (CHCl₃:MeOH) to afford the pure hydrochloride salt as a white solid (101 mg, 84%). **Mp**: decomp. > 110 °C

δ_{H} (400 MHz, CD₃OD): 6.62 (dd, J = 7.5, 2.4 Hz, 1H, H-4), 6.77 – 6.81 (m, 1H, H-17), 6.93 (dd, J = 8.0, 1.8 Hz, 1H, H-6), 6.99 (d, J = 8.9 Hz, 1H, H-9), 7.00 – 7.13 (m, 3H, H-2, H-13 and H-16), 7.25 (t, J = 8.1 Hz, 1H, H-5), 7.49 (dd, J = 8.9, 2.9 Hz, 1H, H-8), 8.05 (d, J = 2.9 Hz, 1H, H-11).

δ_{C} (100 MHz, CD₃OD): 112.0 (d, J = 18.8 Hz, CH Ar, C-13 or C-16), 112.4 (CH Ar, C-2), 113.0 (CH Ar, C-4), 116.4 (CH Ar, C-9), 118.0 (CH Ar, C-6), 118.1 (d, J = 17.9 Hz, CH Ar, C-13 or C-16), 119.1 (dd, J = 5.6, 3.1 Hz, CH Ar, C-17), 131.3 (CH Ar, C-5), 131.4 (CH Ar, C-8), 138.5 (CH Ar, C-11), 143.9 (dd, J = 7.4, 2.2 Hz, qC, C-12), 146.6 (qC), 147.2 (dd, J = 240.4, 12.9 Hz, qC, C-14 or C-15), 150.1 (qC), 151.1 (qC), 151.4 (dd, J = 245.1, 13.4 Hz, qC, C-14 or C-15), 152.5 (qC), 157.4 (qC), 159.0 (qC).

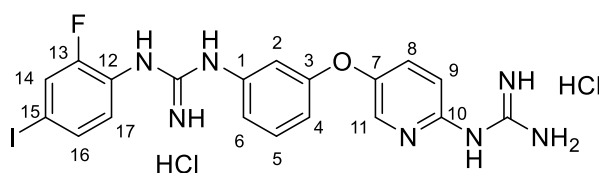
δ_{F} (376 MHz, CD₃OD): – 139.72 (m), - 149.19 (m).

HRMS (m/z ESI⁺): Found 398.1548 (M^+ + H), C₁₉H₁₈N₇OF₂ Requires: 398.1541.

ν_{max} (ATR)/cm⁻¹: 3313 (NH), 3154 (NH), 2922, 2861, 1682, 1625 (C=N), 1507, 1473 (C-F), 1375, 1228 (C-F), 1166, 1146 (C-O), 866, 830, 770, 570, 557.

HPLC: 99.8% (t_{R} 23.6 min).

1-(2-Fluoro-4-iodophenyl)-2-[3-(4-guanidinopyridin-3-yloxy)phenyl]guanidine dihydrochloride (62)



Following Method A, **82** (104 mg, 0.13 mmol) was dissolved in 4M HCl in dioxane (0.58 mL, 2.32 mmol) and in additional dioxane (0.10 mL) until a final concentration of 0.2M was reached. After 8 h stirring at 55 °C, the reaction was adjudged complete (TLC), solvents were evaporated and the residue was purified by silica gel chromatography (CHCl₃:MeOH) to afford the pure hydrochloride salt as a white solid (67 mg, 89%). **Mp**: decomp. > 120 °C

δ_{H} (400 MHz, CD₃OD): 6.98 (ddd, J = 8.3, 2.4, 0.7 Hz, 1H, H-4), 7.02 (t, J = 2.2 Hz, 1H, H-2), 7.11 – 7.18 (m, 3H, H-5 or H-16, H-6 and H-9), 7.46 (t, J = 8.1 Hz, 1H, H-5 or H-16), 7.61 – 7.64 (m, 2H, H-17 and H-8), 7.67 (dd, J = 9.7, 1.8 Hz, 1H, H-14), 8.16 (d, J = 2.9 Hz, 1H, H-11).

δ_{C} (100 MHz, CD₃OD): 92.7 (qC, C-15), 115.5 (CH Ar, C-2), 115.6 (CH Ar, C-9), 117.6 (CH Ar, C-4), 120.9 (CH Ar, C-6), 127.2 (d, J = 22.2 Hz, CH Ar, C-14), 127.3 (d, J = 14.1 Hz, qC, C-12), 130.6 (CH Ar, C-5 or C-16), 132.1 (CH Ar, C-8), 132.5 (CH Ar, C-5 or C-16), 136.0 (CH Ar, J = 3.9 Hz, C-17), 138.7 (qC), 139.1 (CH Ar, C-11), 149.2 (qC), 151.2 (qC), 156.2 (qC), 156.9 (qC), 157.9 (qAr, J = 253.9 Hz, C-13), 159.4 (qC).

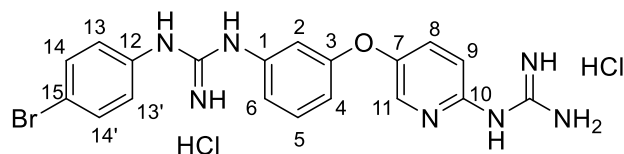
δ_{F} (376 MHz CD₃OD): - 121.55 (t, J = 8.3 Hz).

HRMS (m/z ESI⁺): Found 506.0609 (M⁺ + H), C₁₉H₁₈N₇OFI Requires: 506.0602.

ν_{max} (ATR)/cm⁻¹: 3277 (NH), 3122 (NH), 2923, 2849, 1680, 1660, 1623 – 1570 (C=N), 1474 (C-F), 1227 (C-O), 1160, 1026, 945, 600 (C-I).

HPLC: 98.1% (t_R 25.7 min).

1-(4-Bromophenyl)-2-[3-(4-guanidinopyridin-3-yloxy)phenyl]guanidine dihydrochloride (63)



Following Method A, **83** (184 mg, 0.24 mmol) was dissolved in 4M HCl in dioxane (1.08 mL, 4.4 mmol) and in additional dioxane (0.12 mL) until a final concentration of 0.2M was reached. After 8 h stirring at 55 °C, the reaction was adjudged complete (TLC), solvents were evaporated and the residue was purified by silica gel chromatography (CHCl₃:MeOH) to afford the pure hydrochloride salt as a white solid (113 mg, 92%). **Mp**: decomp. > 124 °C

δ_H (400 MHz, CD₃OD): 6.99 – 7.01 (m, 1H, H-4), 7.04 (bs, 1H, H-2), 7.14 – 7.16 (m, 2H, H-6 and H-9), 7.28 (d, *J* = 8.5 Hz, 2H, H-13 and H-13' or H-14 and H-14'), 7.48 (t, *J* = 8.1 Hz, 1H, H-5), 7.60 – 7.63 (m, 3H, H-13 and H-13' or H-14 and H-14' and H-8), 8.16 (d, *J* = 2.6 Hz, 1H, H-11).

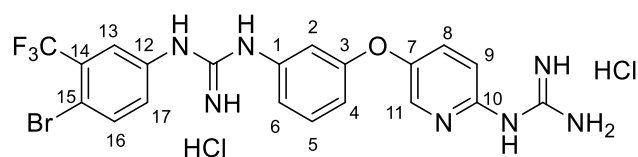
δ_C (100 MHz, CD₃OD): 115.6 (CH Ar, C-6 or C-9), 115.8 (CH Ar, C-2), 118.0 (CH Ar, C-4), 121.1 (CH Ar, C-6 or C-9), 121.7 (qC, C-15), 127.9 (CH Ar, C-13 and C-13' or C-14 and C-14'), 132.1 (CH Ar, C-8), 132.5 (CH Ar, C-5), 134.1 (CH Ar, C-13 and C-13' or C-14 and C-14'), 135.7 (qC), 138.2 (qC), 139.0 (CH Ar, C-11), 149.2 (qC), 151.2 (qC), 156.1 (qC), 156.8 (qC), 159.4 (qC).

HRMS (m/z ESI⁺): Found 440.0837 (M⁺ + H), C₁₉H₁₉N₇OBr Requires: 440.0834.

ν_{max}(ATR)/cm⁻¹: 3256 (NH), 3114 (NH), 2971, 1680, 1660, 1619 (C=N), 1566 (C=N), 1474, 1376 (C-N), 1226 (C-O), 1069 (C-Br), 1010, 832, 637 – 584.

HPLC: 97.5% (*t*_R 25.2 min).

1-[4-Bromo-3-(trifluoromethyl)phenyl]-2-[3-(4-guanidinopyridin-3-yloxy)phenyl]guanidine dihydrochloride (64)



Following Method A, **84** (358 mg, 0.44 mmol) was dissolved in 4M HCl in dioxane (2 mL, 7.97 mmol) and in additional dioxane (0.2 mL) until a final concentration of 0.2M was reached. After 8 h stirring at 55 °C, the reaction was adjudged complete (TLC), solvents were evaporated and the residue was purified by silica gel chromatography (CHCl₃:MeOH) to afford the pure hydrochloride salt as a white solid (225 mg, 88%).

Mp: decomp. > 170 °C

δ_{H} (400 MHz, CD₃OD): 6.66 (dd, J = 8.1, 2.4 Hz, 1H, H-4), 6.96 (dd, J = 8.0, 1.9 Hz, 1H, H-6), 7.04 (d, J = 8.9 Hz, 1H, H-9), 7.08 (t, J = 2.1 Hz, 1H, H-2), 7.20 (dd, J = 8.6, 2.5 Hz, 1H, H-17), 7.28 (t, J = 8.1 Hz, 1H, H-5), 7.48 (d, J = 2.5, 1H, H-13), 7.56 (dd, J = 8.9, 2.9 Hz, 1H, H-8), 7.60 (d, J = 8.6 Hz, 1H, H-16), 8.10 (d, J = 2.7 Hz, 1H, H-11).

δ_{C} (100 MHz, CD₃OD): 111.8 (d, J = 1.8 Hz, qC, C-15), 112.4 (CH Ar, C-2), 113.3 (CH Ar, C-4), 115.5 (CH Ar, C-9), 117.8 (CH Ar, C-6), 122.4 (q, J = 5.5 Hz, CH Ar, C-13), 124.4 (d, J = 272.6 Hz, qCF₃), 127.7 (CH Ar, C-17), 130.9 (d, J = 30.8 Hz, qC, C-14), 131.4 (CH Ar, C-5), 131.6 (CH Ar, C-8), 136.6 (CH Ar, C-16), 138.6 (CH Ar, C-11), 146.1 (qC), 147.2 (qC), 148.9 (qC), 151.8 (qC), 152.4 (qC), 156.9 (qC), 159.0 (qC).

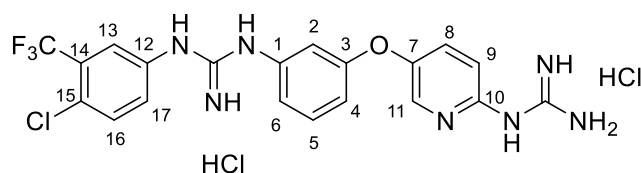
δ_{F} (376 MHz, CD₃OD): - 63.96 (s).

HRMS (m/z ESI⁺): Found 508.0710 (M⁺ + H), C₂₀H₁₈N₇OF₃Br Requires: 508.0708.

ν_{max} (ATR)/cm⁻¹: 3281 (NH), 3142 (NH), 2922, 2849, 1680, 1625 (C=N), 1566 (C=N), 1473, 1319, 1227 (CF₃), 1129 (C-Br), 1023, 832, 592-583.

HPLC: 99.8% (tR 27.5 min).

1-[4-Chloro-3-(trifluoromethyl)phenyl]-2-[3-(4-guanidinopyridin-3-yloxy)phenyl]guanidine dihydrochloride (65)



Following Method A, **85** (113 mg, 0.15 mmol) was dissolved in 4M HCl in dioxane (0.67 mL, 2.70 mmol) and in additional dioxane (0.10 mL) until a final concentration of 0.2M was reached. After 6 h stirring at 55 °C, the reaction was adjudged complete (TLC), solvents were evaporated and the residue was purified by flash chromatography to afford the pure hydrochloride salt as a white solid (63 mg, 79%). **Mp**: 169 – 171 °C

δ_{H} (400 MHz, CD₃OD): 6.93 (dd, J = 8.3, 2.3 Hz, 1H, H-4), 7.06 (t, J = 2.1 Hz, 1H, H-2), 7.11 – 7.13 (m, 2H, H-6 and H-9), 7.44 (t, J = 8.2 Hz, 1H, H-5), 7.53 (dd, J = 8.6, 2.5 Hz, 1H, H-17), 7.60 – 7.65 (m, 2H, H-16 and H-8), 7.69 (d, J = 2.4 Hz, 1H, H-13), 8.15 (d, J = 2.9 Hz, 1H, H-11).

δ_{C} (100 MHz, CD₃OD): 115.1 (CH Ar, C-2), 115.6 (CH Ar, C-9), 117.1 (CH Ar, C-4), 120.5 (CH Ar, C-6), 123.9 (d, J = 272.5 Hz, qCF₃) 124.5 (q, J = 5.5 Hz, CH Ar, C-13), 129.9 (qC, C-15), 130.1 (CH Ar, C-17), 130.2 (d, J = 34.7 Hz, qC, C-14), 132.0 (CH Ar, C-8 or C-16), 132.3 (CH Ar, C-5), 134.0 (CH Ar, C-8 or C-16), 138.3 (qC), 139.0 (CH Ar, C-11), 139.7 (qC), 149.1 (qC), 151.3 (qC), 155.5 (qC), 156.9 (qC), 159.3 (qC).

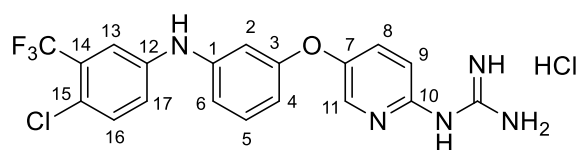
δ_{F} (376 MHz, CD₃OD): - 64.24 (s).

HRMS (m/z ESI⁺): Found: 464.1222 (M^+ + H), C₂₀H₁₈N₇OF₃Cl Requires: 464.1213.

ν_{max} (ATR)/cm⁻¹: 3297 (NH), 3121 (NH), 2923, 2854, 1625 (C=N), 1581 (C=N), 1474 (CF₃), 1320, 1227 (C-O), 1130 (C-Cl), 1032, 832, 589, 557.

HPLC: 96.9% (t_{R} : 26.8 min).

1-(4-{3-[4-Chloro-3-(trifluoromethyl)phenylamino]phenoxy}pyridin-2-yl)guanidine hydrochloride (66)



Following Method A, **90** (200 mg, 0.32 mmol) was dissolved in 4M HCl in dioxane (0.96 mL, 3.86 mmol) and in additional dioxane (0.65 mL) until a final concentration of 0.2M was reached. After 8 h stirring at 55 °C, the reaction was adjudged complete (TLC), solvents were evaporated and the residue was purified by silica gel chromatography (CHCl₃:MeOH) to afford the pure hydrochloride salt as a white solid (136 mg, 93%). **Mp**: 89 - 91 °C

δ_H (400 MHz, CD₃OD): 6.60 (dd, *J* = 8.2, 2.3 Hz, 1H, H-4), 6.75 (t, *J* = 2.2 Hz, 1H, H-2), 6.90 (dd, *J* = 8.1, 2.0 Hz, 1H, H-6), 7.08 (d, *J* = 8.9 Hz, 1H, H-9), 7.24 (dd, *J* = 8.7, 2.7 Hz, 1H, H-17), 7.29 (t, *J* = 8.1 Hz, 1H, H-5), 7.37 – 7.39 (m, 2H, H-13 and H-16), 7.59 (dd, *J* = 8.9, 2.9 Hz, 1H, H-8), 8.13 (d, *J* = 2.9 Hz, 1H, H-11).

δ_C (100 MHz, CD₃OD): 108.7 (CH Ar, C-2), 112.2 (CH Ar, C-4), 114.7 (CH Ar, C-6), 115.5 (CH Ar, C-9), 116.2 (q, *J* = 5.5 Hz, CH Ar, C-13), 121.8 (CH Ar, C-17), 122.4 (qC, C-15), 124.3 (d, *J* = 272.4 Hz, qCF₃), 129.6 (d, *J* = 31.0 Hz, qC, C-14), 131.8 (CH Ar, C-8), 132.0 (CH Ar, C-5), 133.4 (CH Ar, C-16), 138.7 (CH Ar, C-11), 144.2 (qC), 145.6 (qC), 148.8 (qC), 151.8 (qC), 156.9 (qC), 159.4 (qC).

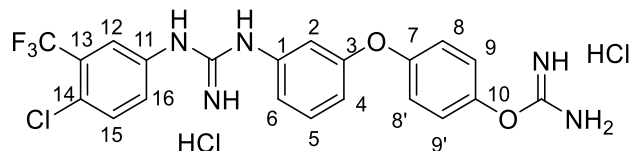
δ_F (376 MHz, CD₃OD): – 64.15 (s).

HRMS (m/z ESI⁺): Found 422.0987 (M⁺ + H), C₁₉H₁₆N₅OClF₃ Requires: 422.0995.

ν_{max} (ATR)/cm⁻¹: 3264 (NH), 2923, 2863, 1684, 1629, 1595, 1474 (C=N), 1400, 1332, 1229 (C-O), 1129 (CF₃), 1115 (C-Cl), 977, 998.

HPLC: 98.6% (*t*R: 33.3 min).

O-(4-{3-[4-Chloro-3-(trifluoromethyl)phenylguanidino]phenoxy}phenyl)carbamimidate dihydrochloride (67**)³⁶**



A solution of the corresponding Boc-protected precursor **96** (100 mg, 0.13 mmol) in 10 mL of a 50% v/v solution of TFA and CH₂Cl₂ was stirred at room temperature. Once disappearance of starting material was observed by TLC (3 h), the excess of TFA and CH₂Cl₂ were removed under vacuum to generate the trifluoroacetate salt. The residue was dissolved in water (10 mL) and 500 mg of IRA400 Amberlite resin in its chloride-activated form was added, allowing the reaction to proceed overnight. The resin was then removed by filtration and the aqueous solution washed with CH₂Cl₂. Removal of water under vacuum yielded the final isouronium hydrochloride **67** as a white/yellow solid (67 mg, 95%). Absence of the trifluoroacetate salt was confirmed by ¹⁹F NMR. **Mp**: decomp. > 120 °C.

δ_H (400 MHz, CD₃OD): 7.02 (dd, *J* = 8.2, 2.0 Hz, 1H, H-4), 7.09 (t, *J* = 2.1 Hz, 1H, H-2), 7.17 (dd, *J* = 8.0, 1.3 Hz, 1H, H-6), 7.20 (d, *J* = 9.0 Hz, 2H, H-8 and H-8' or H-9 and H-9'), 7.35 (d, *J* = 9.0 Hz, 2H, H-8 and H-8' or H-9 and H-9'), 7.47 (t, *J* = 8.1 Hz, 1H, H-5), 7.60 (dd, *J* = 8.6, 2.4 Hz, 1H, H-16), 7.70 (d, *J* = 8.6 Hz, 1H, H-15), 7.75 (d, *J* = 2.3 Hz, 1H, H-12).

δ_C (100 MHz, CD₃OD): 116.5 (CH Ar, C-2), 118.6 (CH Ar, C-4), 121.1 (CH Ar, C-6), 122.2 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 123.8 (d, *J* = 271.0 Hz, qCF₃), 124.0 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 125.3 (q, *J* = 4.3 Hz, CH Ar, CH Ar, C-12), 130.5 (d, *J* = 31.8 Hz, qC, C-13), 130.8 (CH Ar, C-16), 131.3 (qC, C-14), 132.5 (CH Ar, C-5), 134.3 (CH Ar, C-15), 136.1 (qC), 137.9 (qC), 146.8 (qC), 156.3 (qC), 157.5 (qC), 159.5 (qC), 163.3 (qC).

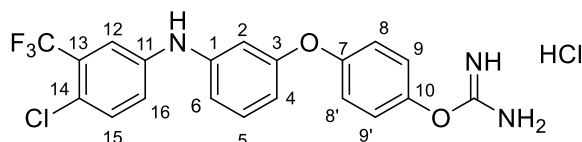
δ_F (376 MHz, CD₃OD): -64.33 (s).

HRMS (*m/z* ESI⁺): Found: 464.1101 (M⁺ + H), C₂₁H₁₈N₅O₂ClF₃ Requires: 464.1101.

$\nu_{\max}$ (ATR)/ cm^{-1} : 3183 (NH), 2963 (CH), 1574 (C=N), 1477, 1250 (CF₃), 1096, 1030 (C-O), 1012 (C-Cl), 785, 667.

HPLC: 98.4% (t_R 27.01 min).

O-(4-{3-[4-Chloro-3-(trifluoromethyl)phenylamino]phenoxy}phenyl)carbamimidate hydrochloride (68)³⁴



To a stirred solution of **99** (205 mg, 0.33 mmol, 1eq) in EtOAc was added SnCl₄ (0.15 mL, 1.32 mmol, 4 eq). After 2 h of stirring at room temperature, the solvent and the excess of SnCl₄ were evaporated in vacuo. The remaining liquid was purified by silica gel chromatography (CHCl₃:Acetone) to afford the pure hydrochloride salt (130 mg, 86%) as a colourless gum.

δ_{H} (400 MHz, DMSO-*d*₆): 6.61 (dd, *J* = 7.6, 2.3 Hz, 1H, H-4), 6.76 (t, *J* = 2.2 Hz, 1H, H-2), 6.92 (dd, *J* = 7.8, 1.7 Hz, 1H, H-6), 7.16 (d, *J* = 9.1 Hz, 2H, H-8 and H-8' or H-9 and H-9'), 7.30 – 7.34 (m, 2H, H-5 and H-16), 7.37 (d, *J* = 9.0 Hz, 2H, H-8 and H-8' or H-9 and H-9'), 7.40 (d, *J* = 2.7 Hz, 1H, H-12), 7.50 (d, *J* = 8.8 Hz, 1H, H-15), 8.61 (bs, 4H, NH), 8.88 (bs, NH amide).

δ_{C} (100 MHz, DMSO-*d*₆): 117.3 (CH Ar, C-2), 120.7 (CH Ar, C-4), 122.6 (CH Ar, C-6), 124.2 (q, *J* = 5.5 Hz, CH Ar, C-12), 128.8 (qC, C-14), 129.8 (CH Ar, C-5 or C-16), 130.0 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 131.5 (d, *J* = 227.5 Hz, qCF₃), 132.6 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 136.6 (d, *J* = 30.7 Hz, qC, C-13), 140.4 (CH Ar, C-5 or C-16), 141.9 (CH Ar, C-15), 152.2 (qC), 152.4 (qC), 153.0 (qC), 154.7 (qC), 164.7 (qC), 167.1 (qC).

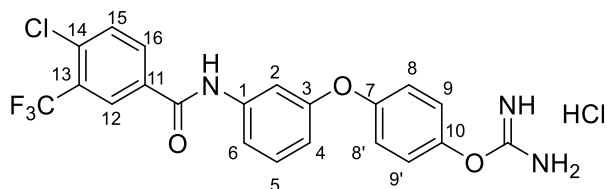
δ_{F} (376 MHz, CD₃OD): – 61.59 (s).

HRMS (m/z ESI⁺): Found: 422.0883 (M⁺ + H), C₂₀H₁₆N₃O₂ClF₃ Requires: 422.0883.

$\nu_{\max}$ (ATR)/ cm^{-1} : 3285 (NH), 2924, 2854, 1693, 1655, 1593 (C=N), 1481 (C-N), 1400, 1258, 1231, 1195, 1175 (C-O), 1128 (CF₃), 1111 (C-Cl), 1027, 824, 681, 665.

HPLC: 96.2% (*t*R 32.2 min).

O-(4-{3-[4-Chloro-3-(trifluoromethyl)benzamido]phenoxy}phenyl)carbamimidate hydrochloride (69**)³⁴**



To a stirred solution of **102** (47 mg, 0.07 mmol) in EtOAc was added SnCl₄ (0.033 mL, 0.28 mmol). After 1h of stirring at room temperature, the solvent and the excess of SnCl₄ were evaporated *in vacuo*. The product was purified by flash chromatography (CHCl₃:Acetone) to afford the final hydrochloride salt **69** (17 mg, 51%) as a white/yellow gum.

δ_H (400 MHz, CD₃OD): 6.84 (dd, *J* = 8.0, 1.8 Hz, 1H, H-4), 7.13 (d, *J* = 8.8 Hz, 2H, H-8 and H-8' or H-9 and H-9'), 7.30 (d, *J* = 8.9 Hz, 2H, H-8 and H-8' or H-9 and H-9'), 7.35 (t, *J* = 8.1 Hz, 1H, H-5), 7.44 (d, *J* = 8.0 Hz, 1H, H-6), 7.60 (s, 1H, H-2), 7.74 (d, *J* = 8.3 Hz, 1H, H-15), 8.17 (dd, *J* = 8.3, 1.6 Hz, 1H, H-16), 8.29 (d, *J* = 1.5 Hz, 1H, H-12), 8.63 (bs, NH).

δ_C (150 MHz, CD₃OD): 113.0 (CH Ar, C-2), 116.3 (CH Ar, C-4); 117.5 (CH Ar, C-6), 121.6 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 122.2 (qC, C-14), 124.1 (q, *J* = 272.5 Hz, qCF₃), 123.8 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 128.2 (q, *J* = 5.4 Hz, CH Ar, C-12), 129.5 (q, *J* = 31.7 Hz, qC, C-13), 131.3 (CH Ar, C-5), 133.1 (CH Ar, C-15), 133.7 (CH Ar, C-16), 135.4 (qC), 136.6 (qC), 141.3 (qC), 146.4 (qC), 158.2 (qC), 158.4 (qC), 166.1 (qC).

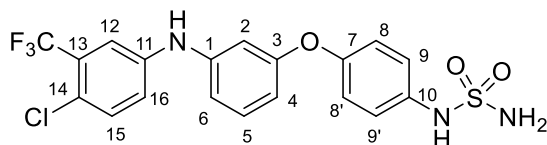
δ_F (376 MHz, CD₃OD): - 62.22 (s).

HRMS (*m/z* ESI⁺): Found 450.0835 (M⁺ + H), C₂₁H₁₆N₃O₃ClF₃, Requires: 450.0832.

ν_{max} (ATR)/cm⁻¹: 3329 (NH), 1695, 1648, 1602 (C=O), 1482 (C=N), 1249 (C-O), 1192 (CF₃), 1131, 1111 (C-Cl), 780, 660.

HPLC: 97.69% (*t*R 31.72 min).

***N*-(4-{3-[4-Chloro-3-(trifluoromethyl)phenylamino]phenoxy}phenyl)sulfonamide (70)³⁷**



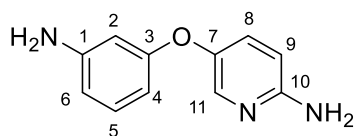
Compound **48** (100 mg, 0.26 mmol, 1 eq.), **103** (30 mg, 0.26 mmol, 1 eq.) and NEt₃ (0.05 mL, 0.29 mmol, 1.1 eq.) were dissolved in CH₂Cl₂ (2 mL) and stirred at overnight at room temperature. The mixture was then washed with water and the organic layer extracted with EtOAc, washed with brine, dried over MgSO₄, concentrated under vacuum and purified by silica gel chromatography (hexanes:EtOAc) to get **70** as a light brown solid (95 mg, 80%). **Mp**: 124-126°C

δ_H (400 MHz, CD₃OD): 6.54 (dd, *J* = 8.2, 2.3 Hz, 1H, H-4), 6.69 (t, *J* = 2.2 Hz, 1H, H-2), 6.84 (dd, *J* = 8.1, 2.1 Hz, 1H, H-6), 6.99 (d, *J* = 9.0 Hz, 2H, H-8 and H-8'), 7.21 – 7.26 (m, 4H, H-9 and H-9', H-16, H-5), 7.34 – 7.38 (m, 2H, H-12, H-15).

δ_C (100 MHz, CD₃OD): 109.1 (CH Ar, C-2), 112.2 (CH Ar, C-4), 113.8 (CH Ar, C-6), 116.1 (q, *J* = 5.5 Hz, C-12, CH Ar), 121.1 (2 CH Ar, C-8 and C-8'), 121.2 (CH Ar, C-16), 121.9 (qC, C-14), 123.2 (2 CH Ar, C-9 and C-9'), 124.4 (d, *J* = 272.5 Hz, qCF₃), 129.7 (d, *J* = 30.9 Hz, qC, C-13), 131.5 (CH Ar, C-5), 133.3 (CH Ar, C-15), 136.1 (qC), 144.6 (qC), 145.1 (qC), 154.4 (qC), 160.5 (qC).

HRMS (m/z ESI⁺): Found: 456.0397 (M⁺ - H), C₁₉H₁₄N₃O₃SClF₃ Requires: 456.0397.

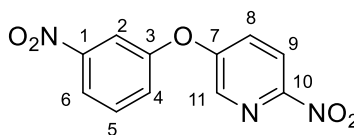
ν_{max} (ATR)/cm⁻¹: 3404 (NH), 3279 (NH), 1596 (S=O), 1489 (S=O), 1143 (C-O), 1153 (CF₃), 1125 (C-N), 830 (C-Cl), 821.

2-Amino-5-(3-aminophenoxy)pyridine (71)^{34,37}

Following Method E, to a solution of nitro derivative **74** (781 mg, 2.99 mmol) in methanol/ tetrahydrofuran/ethyl acetate (5:1:1, 78 mL) was added 10% Pd/C mixture (10 mol%, 318 mg). After 24 h, usual work up followed by silica gel chromatography afforded **71** as a brown solid (566 mg, 94%). **Mp**: 80 °C.^{35,38}

δ_{H} (400 MHz, CDCl_3): 3.68 (s, 2H, NH_2), 4.37 (s, 2H, NH_2), 6.24 (s, 1H, H-2), 6.30 (d, $J = 8.1$ Hz, 1H, H-4 or H-6), 6.36 (d, $J = 7.9$ Hz, 1H, H-4 or H-6), 6.51 (d, $J = 8.7$ Hz, 1H, H-9), 7.05 (t, $J = 8.1$ Hz, 1H, H-5), 7.20 (dd, $J = 8.8, 2.7$ Hz, 1H, H-8), 7.91 (d, $J = 2.4$ Hz, 1H, H-11).

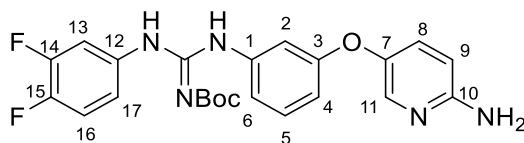
HRMS (m/z ESI^+): Found: 202.0964 ($\text{M}^+ + \text{H}$), $\text{C}_{11}\text{H}_{12}\text{N}_3\text{O}$ Requires: 202.0975.

2-Nitro-5-(3-nitrophenoxy)pyridine (74)^{35,38}

To a mixture of 5-bromo-2-nitropyridine **72** (100 mg, 0.49 mmol, 1 eq.) and Cs_2CO_3 (241 mg, 0.74 mmol, 1.5 eq.) in DMF (10 mL) was added dropwise a solution of 3-nitrophenol **73** (75 mg, 0.54 mmol, 1.1 eq.) in DMF (5 mL) and the reaction mixture was stirred at room temperature for 12 h. The reaction was then diluted with water and extracted with ethyl acetate. The organic layer was washed with 5% aqueous NaHCO_3 and brine and dried over anhydrous MgSO_4 . The residue was purified by basic silica gel chromatography (hexanes:EtOAc), and the desired fractions were concentrated under reduced pressure to give the product **74** as a yellow solid (47 mg, 37%). **Mp**: 102-103 °C (lit. 113-114 °C)^{35,38}

δ_{H} (400 MHz, CDCl_3): 7.47 (dd, $J = 8.1, 2.0$ Hz, 1H, H-4), 7.54 (dd, $J = 8.9, 2.7$ Hz, 1H, H-8), 7.67 (t, $J = 8.2$ Hz, 1H, H-5), 7.98 (s, 1H, H-2), 8.16 (d, $J = 8.2$ Hz, 1H, H-6), 8.32 (d, $J = 8.9$ Hz, 1H, H-9), 8.39 (d, $J = 2.7$ Hz, 1H, H-11).

1-(3,4-Di-fluorophenyl)-2-(*tert*-butoxycarbonyl)-3-[3-(6-aminopyridin-3-yloxy)phenyl]guanidine (76)



Following a modification of Method C, HgCl_2 (404 mg, 1.49 mmol) was added over a solution of **71** (250 mg, 1.24 mmol), thiourea derivative **12** (430 mg, 1.49 mmol) and NEt_3 (0.5 mL, 3.84 mmol) in CH_2Cl_2 (8 mL). After 6 h workup and silica gel chromatography (hexanes:EtOAc) afforded compound **76** as a yellow amorphous solid (254 mg, 45%). **Mp**: 119 - 121 °C

δ_{H} (600 MHz, CDCl_3): 1.48 (s, 9H, $(\text{CH}_3)_3$), 4.38 (s, 2H, NH_2), 6.47 – 6.67 (m, 4H, CH Ar), 7.09 (m, 2H, CH Ar), 7.05 – 7.13 (m, 1H, CH Ar), 7.22 (bs, 1H, CH Ar), 7.43 (bs, 1H, NH), 7.83 (bs, 1H, CH Ar), 7.92 (bs, 1H, CH Ar), 9.62 (s, 1H, NH).

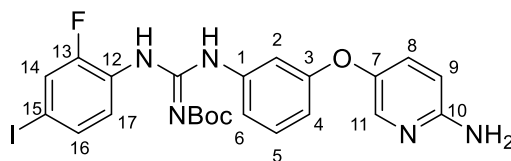
δ_{C} (150 MHz, CDCl_3): 28.2($(\text{CH}_3)_3$), 83.9 (qC, $(\text{CH}(\text{CH}_3)_3$), 109.4 (CH Ar), 111.1 (CH Ar), 111.4 (CH, Ar), 111.9 (CH, Ar), 115.5 (CH, Ar), 116.6 (CH, Ar), 117.0 (CH, Ar), 118.0 (CH, Ar), 129.9 (CH, Ar), 140.4 (CH Ar), 143.5 (qC), 145.3 (qC), 148.2 (qC), 149.2 (qC), 151.0 (qC), 153.1 (qC), 155.3 (qC), 159.1 (qC), 160.1 (qC).

Note: Aromatic protons and carbons, quaternary carbons and fluorine peaks could not be fully assigned and characterised because the peaks were very broad.

HRMS (m/z ESI⁺): Found: 456.1855 ($\text{M}^+ + \text{H}$), $\text{C}_{23}\text{H}_{24}\text{F}_2\text{N}_5\text{O}_3$ Requires: 456.1842.

ν_{max} (ATR)/ cm^{-1} : 3422 (NH), 3304 (NH), 2978, 2932, 1723 (C=O), 1662, 1560, 1594 (C=N), 1482, 1226 (C-F), 1144 (C-F), 1085.

1-(2-Fluoro-4-iodophenyl)-2-(*tert*-butoxycarbonyl)-3-[3-(6-aminopyridin-3-yloxy)phenyl]guanidine (77)



Following a modification of Method C, HgCl_2 (307 mg, 1.13 mmol) was added over a solution of **71** (190 mg, 0.94 mmol), thiourea derivative **13** (448 mg, 1.13 mmol) and NEt_3 (0.40 mL, 2.9 mmol) in CH_2Cl_2 (5 mL). After 6 h workup and silica gel chromatography (hexanes:EtOAc) afforded compound **77** as a yellow oil (185 mg, 35%).

δ_{H} (400 MHz, $\text{DMSO}-d_6$): 1.27 (s, 9H, $(\text{CH}_3)_3$), 5.84 (s, 2H, NH_2), 6.46 – 6.49 (m, 2H, CH Ar and H-9), 6.66 (t, $J = 8.5$ Hz, 1H, H-5), 7.16 – 7.19 (m, 2H, CH Ar), 7.25 – 7.27 (m, 2H, CH Ar), 7.36 – 7.38 (m, 1H, CH Ar), 7.46 – 7.49 (m, 1H, H-14), 7.72 (d, $J = 2.4$ Hz, 1H, H-11), 9.08 (bs, 1H, NH), 9.31 (bs, 1H, NH).

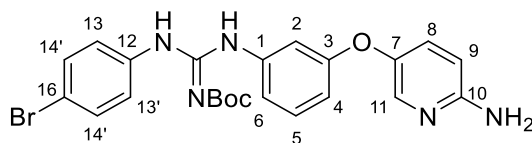
δ_{C} (150 MHz, $\text{DMSO}-d_6$): 27.7 ($(\text{CH}_3)_3$), 59.7 (qC, $\text{C}(\text{CH}_3)_3$), 80.1 (qAr, C-15), 107.3 (CH Ar), 108.6 (CH Ar, C-9), 110.0 (CH Ar), 113.0 (CH Ar), 124.0 (d, $J = 21.5$ Hz, CH Ar, C-14), 126.1 (CH Ar, C-5), 129.5 (CH Ar), 130.4 (CH Ar), 132.9 (CH Ar), 139.6 (CH Ar, C-11), 141.3 (qC), 142.7 (qC), 142.9 (qC), 151.9 (qC), 152.7 (d, $J = 239.8$ Hz, qAr, C-13), 156.9 (qC), 158.9 (qC), 170.3 (qC).

Note: Aromatic protons and carbons, quaternary carbons and fluorine peaks could not be fully assigned and characterised because the peaks were very broad.

HRMS (m/z ESI^+): Found: 564.0908 ($\text{M}^+ + \text{H}$), $\text{C}_{23}\text{H}_{24}\text{FIN}_5\text{O}_3$ Requires: 564.0902.

ν_{max} (ATR)/ cm^{-1} : 3489 (NH), 3413 (NH), 3299, 3203 (NH), 2970, 2931, 1720 (C=O), 1661 (C=N), 1482 (C-F), 1462, 1224 (C-N), 1147 (C-O), 687 (C-I).

1-(4-bromophenyl)-2-(*tert*-butoxycarbonyl)-3-[3-(6-aminopyridin-3-yloxy)phenyl]guanidine (78**)**



Following a modification of Method C, HgCl_2 (406 mg, 1.49 mmol) was added over a solution of **71** (250 mg, 1.24 mmol), thiourea derivative **14** (493 mg, 1.49 mmol) and NEt_3 (0.5 mL, 3.84 mmol) in CH_2Cl_2 (8 mL). After 6 h workup and silica gel chromatography (hexanes:EtOAc) afforded compound **78** as a yellow-brown amorphous solid (517 mg, 84%). **Mp**: 141 - 143 °C

δ_{H} (400 MHz, CDCl_3): 1.48 (s, 9H, $(\text{CH}_3)_3$), 6.53 – 6.57 (m, 2H, CH Ar and H-9), 6.61 – 6.68 (m, 2H, CH Ar), 7.20 - 7.24 (m, 3H, CH Ar), 7.37 - 7.42 (m, 3H, CH Ar), 7.87 (d, $J = 2.7$ Hz, 1H, H-11).

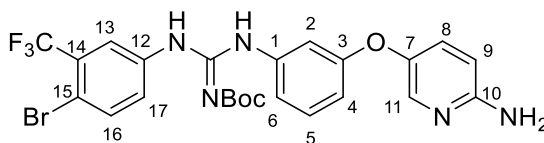
δ_{C} (150 MHz, CDCl_3): 28.2 ($(\text{CH}_3)_3$), 83.7 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 109.4 (CH Ar), 111.1 (CH Ar), 111.8 (CH Ar), 114.4 (CH Ar), 115.3 (CH Ar), 116.6 (q), 121.7 (CH Ar), 124.5 (CH Ar), 131.0 (CH Ar), 131.8 (CH Ar), 132.7 (CH Ar), 138.1 (CH Ar), 140.5 (qC), 145.3 (qC), 148.4 (qC), 153.1 (qC), 155.3 (qC), 159.1 (qC), 160.1 (qC).

Note: Aromatic protons and carbons and quaternary carbons peaks could not be fully assigned and characterised because the peaks were very broad.

HRMS (m/z ESI $^+$): Found: 498.1141 ($\text{M}^+ + \text{H}$), $\text{C}_{23}\text{H}_{25}\text{BrN}_5\text{O}_3$ Requires: 498.1135).

ν_{max} (ATR)/ cm^{-1} : 3477 (NH), 3415 (NH), 3296 (NH), 3166, 2965, 1729 (C=O), 1655 (C=N), 1592, 1580, 1549, 1477, 1398 (C-N), 1145 (C-Br), 1067 (C-O), 823, 770.

1-[4-Bromo-3-(trifluoromethyl)phenyl]-2-(*tert*-butoxycarbonyl)-3-[3-(6-aminopyridin-3-yloxy)phenyl]guanidine (79)



Following a modification of Method C, HgCl_2 (988 mg, 3.64 mmol) was added over a solution of **71** (610 mg, 3.03 mmol), thiourea derivative **16** (1455 mg, 3.64 mmol) and NEt_3 (1.30 mL, 9.39 mmol) in CH_2Cl_2 (16 mL). After 6 h workup and silica gel chromatography (hexanes:EtOAc) afforded compound **79** as a yellow oil (789 mg, 46%).

δ_{H} (600 MHz, $\text{DMSO}-d_6$): 1.25 (s, 9H, $(\text{CH}_3)_3$), 5.84 (s, 2H, NH_2), 6.47 – 6.49 (m, 2H, CH Ar), 6.94 (bs, 1H, CH Ar), 7.17 – 7.18 (m, 3H, CH Ar), 7.28 (bs, 1H, CH Ar), 7.68 – 7.83 (m, 3H, CH Ar), 9.12 (bs, 1H, NH), 9.36 (bs, 1H, NH).

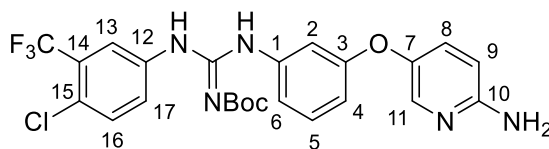
δ_{C} (100 MHz, $\text{DMSO}-d_6$): 27.7 ($(\text{CH}_3)_3$), 79.0 (qC, $\text{C}(\text{CH}_3)_3$), 80.3 (qC, C-15), 107.4 (CH Ar), 108.8 (CH Ar), 110.0 (CH Ar), 113.1 (CH Ar), 124.2 (CH Ar), 126.2 (CH Ar), 129.7 (CH Ar), 130.5 (CH Ar), 133.0 (CH Ar), 139.6 (CH Ar), 141.31 (qC), 141.33 (qC), 142.8 (qC), 143.0 (qC), 152.0 (qC), 157.0 (qC), 159.0 (qC), 170.5 (qC).

Note: Aromatic protons and carbons, quaternary carbons and fluorine peaks could not be fully assigned and characterised because the peaks were very broad.

HRMS (m/z ESI^+): Found: 566.1014 ($\text{M}^+ + \text{H}$), $\text{C}_{24}\text{H}_{24}\text{BrF}_3\text{N}_5\text{O}_3$ Requires: 566.1015.

ν_{max} (ATR)/ cm^{-1} : 3476 (NH), 3413 (NH), 3297, 2981, 2935, 1720 (C=O), 1659 (C=N), 1588 (C=N), 1553 (C=N), 1476, 1230 (C-O), 1134 (CF_3), 1053 (C-Br), 826, 771.

1-[4-Chloro-3-(trifluoromethyl)phenyl]-2-(*tert*-butoxycarbonyl)-3-[3-(6-aminopyridin-3-yloxy)phenyl]guanidine (80**)**



Following a modification of Method C, HgCl_2 (548 mg, 2.02 mmol) was added over a solution of **71** (338 mg, 1.68 mmol), thiourea derivative **15** (715 mg, 2.02 mmol) and NEt_3 (0.72 mL, 5.21 mmol) in CH_2Cl_2 (9 mL). After 6 h workup and silica gel chromatography (hexanes:EtOAc) afforded compound **80** as a yellow oil (368 mg, 42%).

δ_{H} (400 MHz, CDCl_3): 1.49 (s, 9H, $(\text{CH}_3)_3$), 4.38 (bs, 2H, NH_2), 6.47 – 6.72 (m, 4H), 7.20 – 7.22 (m, 2H), 7.38 – 7.52 (m, 1H), 7.92 (bs, 2H), 9.66 (bs, 1H, NH), 9.80 (bs, 1H, NH).

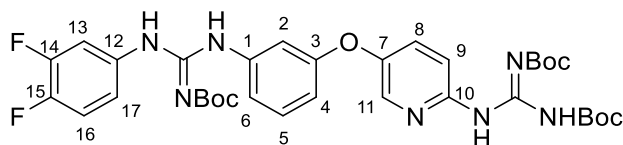
δ_{C} (400 MHz, CDCl_3): 28.2 ($(\text{CH}_3)_3$), 84.1 (qC, $\text{C}(\text{CH}_3)_3$), 109.4, 111.0, 111.9, 116.4, 121.5, 123.9, 124.0, 130.9, 131.0, 132.6, 140.1, 140.4, 140.5, 155.3, 155.4.

Note: Aromatic protons and carbons, quaternary carbons and fluorine peaks could not be fully assigned and characterised because the peaks were very broad.

HRMS (m/z APCI): Found: 520.1391 ($\text{M}^+ - \text{H}$), $\text{C}_{24}\text{H}_{22}\text{N}_5\text{O}_3\text{ClF}_3$ Requires: 520.1369.

ν_{max} (ATR)/ cm^{-1} : 3406 (NH), 3299 (NH), 3191 (NH), 2981, 2936, 1720 (C=O), 1659 (C=N), 1590, 1478, 1320, 1231 (C-O), 1135 (CF_3), 1088 (C-Cl), 1044, 826, 771, 666.

1-(3,4-Di-fluorophenyl)-2-(*tert*-butoxycarbonyl)-3-{3-[6-(2',3'-di-*tert*-butoxycarbonyl)guanidinopyridin-3-yloxy]phenyl}guanidine (81**)**



Following Method G, HgCl_2 (130 mg, 0.48 mmol, 1.04 eq.) was added over a solution of **76** (217 mg, 0.46 mmol), *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methyliothiourea **51** (139 mg, 0.48 mmol), and NEt_3 (0.22 mL, 1.61 mmol) in CH_2Cl_2 (2.5 mL). The reaction was stirred at room temperature for 48 h, then work up and silica gel chromatography (hexanes:EtOAc) afforded **81** (280 mg, 87%) as a white amorphous solid. **Mp**: 135-137 °C.

δ_{H} (400 MHz, CDCl_3): 1.48 (s, 9H, $((\text{CH}_3)_3)$), 1.51 (9H, (s, $(\text{CH}_3)_3$), 1.53 (9H, $((\text{CH}_3)_3)$), 6.52 - 6.72 (m, 3H, CH Ar), 7.04 - 7.14 (m, 2H, CH Ar), 7.29 (bs, 1H, CH Ar), 7.40 - 7.49 (bs, 1H, CH Ar), 7.80 (bs, 1H, CH Ar), 8.11 (bs, 1H, CH Ar), 8.38 (bs, 1H, CH Ar), 9.36 (bs, NH), 9.92 - 10.14 (m, NH), 10.87 (bs, NH), 11.52 (bs, NH).

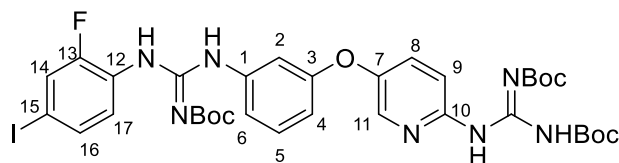
δ_{C} (100 MHz, CDCl_3): 28.16 $((\text{CH}_3)_3)$, 28.2 $((\text{CH}_3)_3)$, 28.3 $((\text{CH}_3)_3)$, 80.1 (qC, $\underline{\underline{\text{C}}}(\text{CH}_3)_3$), 83.9 (qC, $\underline{\underline{\text{C}}}(\text{CH}_3)_3$), 84.1 (qC, $\underline{\underline{\text{C}}}(\text{CH}_3)_3$), 109.8 (CH Ar), 110.4 (CH Ar), 112.1 (CH Ar), 112.8 (CH Ar), 116.9 (CH Ar), 117.6 (CH Ar), 118.2 (CH Ar), 128.8 (CH Ar), 129.2 (CH Ar), 130.0 (CH Ar), 131.0 (CH Ar), 140.0 (CH Ar), 145 - 160, 146.3 (qC), 146.6 (qC), 149.8 (qC), 152.9 (qC), 153.0 (qC), 157.7 (qC), 158.8 (qC), 163.3 (qC).

Note: Aromatic protons and carbons, quaternary carbons and fluorine peaks could not be fully assigned and characterised because the peaks were very broad.

HRMS (m/z ESI^+): Found: 698.3110 ($\text{M}^+ + \text{H}$), $\text{C}_{34}\text{H}_{42}\text{F}_2\text{N}_7\text{O}_7$ Requires: 698.3114.

ν_{max} (ATR)/ cm^{-1} : 3412 (NH), 3258 (NH), 2979, 2930, 1720 (C=O), 1624 (C=N), 1596 (C=N), 1560, 1514, 1463, 1368, 1292, 1224, 1123-1104 (C-F), 771.

1-(2-Fluoro-4-iodophenyl)-2-(*tert*-butoxycarbonyl)-3-{3-[6-(2',3'-di-*tert*-butoxycarbonyl) guanidinopyridin-3-yloxy]phenyl}guanidine (82**)**



Following Method G, HgCl_2 (261 mg, 0.96 mmol) was added over a solution of **77** (520 mg, 0.92 mmol), *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methylisothiurea **51** (279 mg, 0.96 mmol), and triethylamine (0.45 mL, 3.22 mmol) in CH_2Cl_2 (5 mL). The reaction was stirred at room temperature for 48 h, then work up and silica gel chromatography (hexanes:EtOAc) afforded **82** as a white amorphous solid (500 mg, 67%). **Mp**: decomp. > 116 °C

δ_{H} (400 MHz, CDCl_3): 1.35 (s, 9H, $(\text{CH}_3)_3$), 1.39 (s, 18H, $(\text{CH}_3)_3$), 6.39 (bs, 1H, CH Ar), 6.52 - 6.58 (m, 3H, CH Ar), 7.06 - 7.20 (m, 2H, CH Ar), 7.24 - 7.28 (m, 2H, CH Ar), 7.97 (bs, 1H, CH Ar), 8.16 (bs, NH), 8.25 (bs, 1H CH Ar), 9.00 - 9.78 (m, NH), 10.74 (bs, NH), 11.39 (bs, NH).

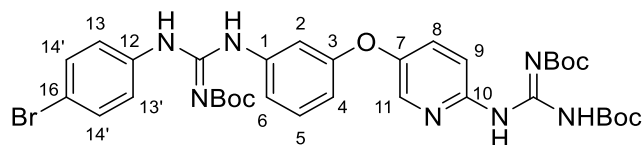
δ_{C} (100 MHz, CDCl_3): 28.2 ($(\text{CH}_3)_3$), 28.3 ($(\text{CH}_3)_3$), 80.0 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 83.9 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 84.1 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 112.1 (CH Ar), 112.9 (CH Ar), 116.9 (CH Ar), 117.4 (CH Ar), 123.1 (CH Ar), 128.9 (CH Ar), 130.0 (CH Ar), 131.1 (CH Ar), 133.5 (CH Ar), 139.9 (CH Ar), 145 - 160, 146.5 (qC), 148.3 (qC), 149.8 (qC), 152.8 (qC), 153.0 (qC), 157.7 (qC), 158.8 (qC), 163.3 (q).

Note: Aromatic protons and carbons, quaternary carbons and fluorine peaks could not be fully assigned and characterised because the peaks were very broad.

HRMS (m/z ESI⁺): Found: 806.2176 ($\text{M}^+ + \text{H}$), $\text{C}_{34}\text{H}_{42}\text{FIN}_7\text{O}_7$ Requires: 806.2169.

ν_{max} (ATR)/ cm^{-1} : 3257 (NH), 2979, 2930, 1721 (C=O), 1626 (C=N), 1591 (C=N), 1465, 1368, 1223 (C-O), 1146 - 1104 (C-F), 1056, 852, 770 (C-I).

1-(4-Bromophenyl)-2-(*tert*-butoxycarbonyl)-3-{3-[6-(2',3'-di-*tert*-butoxycarbonyl)guanidinopyridin-3-yloxy]phenyl}guanidine (83**)**



Following Method G, HgCl_2 (185 mg, 0.68 mmol) was added over a solution of **78** (327 mg, 0.66 mmol), *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methylisothiurea **51** (197 mg, 0.68 mmol), and triethylamine (0.32 mL, 2.31 mmol) in CH_2Cl_2 (4 mL). The reaction was stirred at room temperature for 48 h, then work up and silica gel chromatography (hexanes:EtOAc) afforded **83** (244, 50%) as a white amorphous solid. **Mp**: 104-107 °C.

δ_{H} (600 MHz, CDCl_3): 1.48 (s, 9H, $((\text{CH}_3)_3)$), 1.52 (s, 9H, $((\text{CH}_3)_3)$), 1.54 (9H, $((\text{CH}_3)_3)$), 6.53 – 6.78 (m, 4H, CH Ar and NH), 7.34 – 7.42 (m, 4H, CH Ar), 7.51 – 7.55 (m, 1H, CH Ar), 8.12 (bs, 1H, CH Ar), 8.39 (bs, 1H, CH Ar), 9.62 (bs, NH), 10.87 (bs, NH), 11.51 (bs, NH).

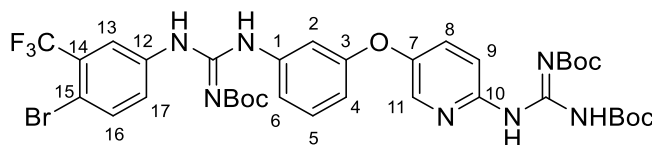
δ_{C} (100 MHz, CDCl_3): 28.0 $((\text{CH}_3)_3)$, 28.1 $((\text{CH}_3)_3)$, 28.2 $((\text{CH}_3)_3)$, 79.9 (qC, $\underline{\underline{\text{C}}}(\text{CH}_3)_3$), 83.6 (qC, $\underline{\underline{\text{C}}}(\text{CH}_3)_3$), 83.9 (qC, $\underline{\underline{\text{C}}}(\text{CH}_3)_3$), 112.4 (CH Ar), 116.8 (CH Ar), 117.5 (CH Ar), 121.5 (CH Ar), 124.3 (CH Ar), 129.0 (CH Ar), 129.8 (CH Ar), 130.8 (CH Ar), 131.6 (CH Ar), 132.6 (CH Ar), 139.8 (CH Ar), 140 – 152, 152.8 (qC), 163.2 (qC).

Note: Aromatic protons and carbons and quaternary carbons peaks could not be fully assigned and characterised because the peaks were very broad.

HRMS (m/z ESI^+): Found: 740.2402 ($\text{M}^+ + \text{H}$), $\text{C}_{34}\text{H}_{43}\text{BrN}_7\text{O}_7$ Requires: 740.2407.

ν_{max} (ATR)/ cm^{-1} : 3253 (NH), 2979, 2932, 1719 (C=O), 1625, 1586 (C=N), 1551 (C=N), 1387, 1323, 1223 (C-N), 1145 (C-O), 1123 (C-Br), 1103, 806 – 687.

1-[4-Bromo-3-(trifluoromethyl)phenyl]-2-(*tert*-butoxycarbonyl)-3-{3-[6-(2',3'-di-*tert*-butoxycarbonyl)guanidinopyridin-3-yloxy]phenyl}guanidine (84**)**



Following Method G, HgCl_2 (255 mg, 0.94 mmol) was added over a solution of **79** (507 mg, 0.90 mmol), *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methylisothiurea **51** (273 mg, 0.94 mmol), and NEt_3 (0.44 mL, 3.15 mmol) in CH_2Cl_2 (5 mL). The reaction was stirred at room temperature for 48 h, then work up and silica gel chromatography (hexanes:EtOAc) afforded **84** as a white amorphous solid (324 mg, 45%). **Mp**: 142–144 °C

δ_{H} (400 MHz, CDCl_3): 1.50 (s, 9H, $(\text{CH}_3)_3$), 1.53 (s, 18H, $(\text{CH}_3)_3$), 6.52 (bs, 1H, CH Ar), 6.66 – 6.69 (bs, 2H, CH Ar), 7.22 (bs, 1H, CH Ar), 7.37 (bs, 2H, CH Ar), 7.58 – 7.60 (bs, 1H, CH Ar), 7.85 – 7.92 (bs, 1H, CH Ar), 8.10 (bs, 1H, CH Ar), 8.39 (bs, 1H, CH Ar), 9.82 (bs, NH), 10.88 (bs, NH), 11.52 (bs, NH).

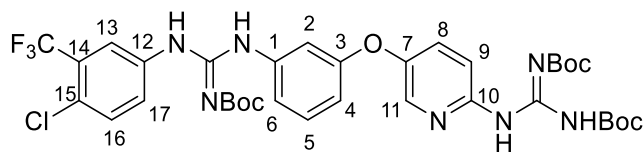
δ_{C} (100 MHz, CDCl_3): 28.19 ($(\text{CH}_3)_3$), 28.21 ($(\text{CH}_3)_3$), 80.1 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 84.1 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 84.2 (q, $\underline{\text{C}}(\text{CH}_3)_3$), 112.0 (CH Ar), 112.9 (CH Ar), 112.94 (CH Ar), 117.0 (CH Ar), 119.3 (CH Ar), 128.9 (CH Ar), 129.0 (CH Ar), 131.2 (CH Ar), 135.2 (CH Ar), 140.0 (CH Ar), 146.6 (qC), 148.1 (qC), 150.1 (qC), 152.9 (qC), 153.1 (qC), 163.3 (qC).

Note: Aromatic protons and carbons, quaternary carbons and fluorine peaks could not be fully assigned and characterised because the peaks were very broad.

HRMS (m/z ESI^+): Found: 808.2288 ($\text{M}^+ + \text{H}$), $\text{C}_{35}\text{H}_{42}\text{BrF}_3\text{N}_7\text{O}_7$ Requires: 808.2276.

ν_{max} (ATR)/ cm^{-1} : 3410 (NH), 3251 (NH), 2985 (CH), 2921, 1720 (C=O), 1645, 1631, 1583 (C=N), 1561 (C=N), 1386, 1326, 1233 (C-N), 1144 (C-O), 1128 (CF_3), 1108 (C-Br), 1081, 1059, 1020, 842, 772, 753, 550.

1-[4-Chloro-3-(trifluoromethyl)phenyl]-2-(*tert*-butoxycarbonyl)-3-{3-[6-(2',3'-di-*tert*-butoxycarbonyl)guanidinopyridin-3-yloxy]phenyl}guanidine (85**)**



Following Method G, HgCl_2 (214 mg, 0.79 mmol) was added over a solution of **80** (398 mg, 0.76 mmol), *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methylisothiurea **51** (230 mg, 0.79 mmol), and NEt_3 (0.37 mL, 2.66 mmol) in CH_2Cl_2 (4 mL). The reaction was stirred at room temperature for 48 h, then work up and silica gel chromatography (hexanes:EtOAc) afforded **85** (354 mg, 61%) as a white-yellow amorphous solid. **Mp**: 84 - 86 °C

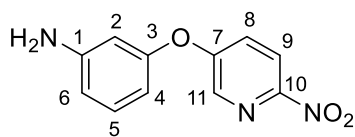
δ_{H} (400 MHz, CDCl_3): 1.49 (s, 9H, $(\text{CH}_3)_3$), 1.51 (s, 9H, $(\text{CH}_3)_3$), 1.53 (s, 9H, $(\text{CH}_3)_3$), 6.51 – 6.71 (m, 2H, CH Ar), 6.99 (bs, 1H, CH Ar), 7.23 – 7.30 (m, 1H, CH Ar), 7.40 – 7.63 (m, 3H, CH Ar), 7.83 – 7.92 (m, 1H, CH Ar), 8.11 (bs, 1H, CH Ar), 8.38 – 8.60 (m, 1H, CH Ar), 9.72 (bs, 1H, NH), 9.81 (bs, 1H, NH), 10.87 (bs, 1H, NH), 11.52 (bs, 1H, NH).

δ_{C} (100 MHz, CDCl_3): 28.1 ($(\text{CH}_3)_3$), 28.2 ($(\text{CH}_3)_3$), 80.0 (qC, $\text{C}(\text{CH}_3)_3$), 84.0 (qC, $\text{C}(\text{CH}_3)_3$), 112.9, 116.9, 119.5, 121.4, 124.1, 129.1, 130.0, 131.0, 132.5, 139.9, 146.7, 148.0, 149.7, 150.1, 152.8, 152.9, 157.7, 158.8, 163.2.

Note: Aromatic protons and carbons, quaternary carbons and fluorine peaks could not be fully assigned and characterised because the peaks were very broad.

HRMS (m/z ESI^+): Found: 764.2784 ($\text{M}^+ + \text{H}$), $\text{C}_{35}\text{H}_{42}\text{ClF}_3\text{N}_7\text{O}_7$ Requires: 764.2786.

ν_{max} (ATR)/ cm^{-1} : 3256 (NH), 2925, 2855, 1720 (C=O), 1628 (C=N), 1592 (C=N), 1557, 1465, 1412, 1387, 1229 (C-N), 1143 (CF_3), 1125, (C-O), 1105 (C-Cl), 1057, 1105, 848, 804, 772, 688.

5-(3-Aminophenoxy)-2-nitropyridine (86)³⁹

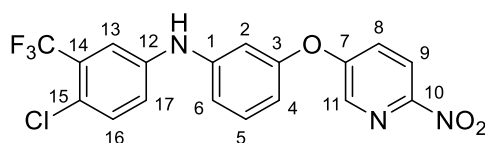
5-Bromo-2-nitropyridine **72** (500 mg, 2.45 mmol, 1eq.), 3-aminophenol **35** (295 mg, 2.7 mmol, 1.1 eq.) and Cs_2CO_3 (1.2 mg, 3.7 mmol, 1.5 eq.) were dissolved in DMF (5 mL) and stirred at 80 °C for 12 h. The mixture was cooled to room temperature, washed with water and the organic layer extracted with EtOAc, washed with brine, dried over MgSO_4 , concentrated under vacuum and purified by flash chromatography (hexanes:EtOAc) to get **86** as a yellow solid (311 mg, 55 %). **Mp**: 118-120 °C

δ_{H} (400 MHz, CDCl_3): 6.40 (s, 1H, H-2), 6.45 (d, $J = 8.0$ Hz, H-4), 6.58 (d, $J = 8.1$ Hz, H-6), 7.20 (t, $J = 8.1$ Hz, H-5), 7.42 (dd, $J = 8.9, 2.8$ Hz, H-8), 8.22 (d, $J = 8.6$ Hz, 1H, H-9), 8.32 (d, $J = 2.3$ Hz, 1H, H-11).

δ_{C} (150 MHz, CDCl_3): 106.6 (CH Ar, C-2), 109.7 (CH Ar, C-4), 112.7 (CH Ar, C-6), 120.0 (CH Ar, C-9), 125.7 (CH Ar, C-8), 131.3 (CH Ar, C-5), 139.0 (CH Ar, C-11), 149.0 (qC), 155.4 (qC), 159.1 (qC), 171.3 (qC).

HRMS ($m/z - \text{ESI}^+$): Found: 232.0715 ($\text{M}^+ + \text{H}$), $\text{C}_{11}\text{H}_{10}\text{N}_3\text{O}_3$ Requires: 232.0717.

ν_{max} (ATR)/ cm^{-1} : 3473 (NH), 3378 (NH), 3219, 3054, 292, 1624, 1604, 1521 (NO_2), 1488, (NO_2), 1455, 1344 (C-N), 1298, 1238 (C-O), 1142, 1114, 995, 875, 850, 686, 674.

5-{3-[4-Chloro-3-(trifluoromethyl)phenyl]aminophenoxy}-2-nitropyridine (88**)**

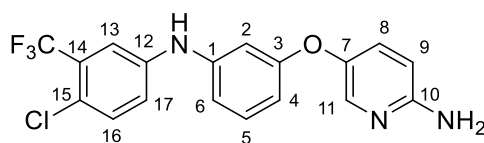
Following Method D, Pd₂(dba)₃ (3 mol%, 6 mg), Xantphos (3 mol%, 4 mg), compound **86** (50 mg, 0.22 mmol), Cs₂CO₃ (100 mg, 0.31 mmol) and remaining liquid 4-bromo-1-chloro-2-(trifluoromethyl)benzene (0.03 mL, 0.22 mmol) were mixed, followed syringe addition of toluene (0.5 mL). The mixture was heated at 90 °C for 24 h and usual work up and flash chromatography afforded **88** as a yellow oil (82 mg, 91%).

δ_{H} (400 MHz, CDCl₃): 6.01 (bs, 1H, NH), 6.70 (dd, J = 8.1, 2.2 Hz, 1H, H-4), 6.79 (t, J = 2.2 Hz, 1H, H-2), 6.96 (dd, J = 7.9, 1.8 Hz, 1H, H-6), 7.18 (dd, J = 8.5, 2.8 Hz, 1H, H-17), 7.34 – 7.39 (m, 3H, H-13, H-16 and H-5), 7.46 (dd, J = 8.9, 2.8 Hz, 1H, H-8), 8.24 (d, J = 8.9 Hz, 1H, H-9), 8.33 (d, J = 2.8 Hz, 1H, H-11).

δ_{C} (100 MHz, CDCl₃): 109.4 (CH Ar, C-2), 113.2 (CH Ar, C-4), 115.3 (CH Ar, C-6), 117.2 (q, J = 5.5 Hz, CH Ar, C-13), 119.9 (CH Ar, C-9), 122.0 (CH Ar, C-17), 122.7 (d, J = 273.4 Hz, qCF₃), 124.1 (qC, C-15), 126.0 (CH Ar, C-8), 129.4 (d, J = 31.2 Hz, qC, C-14), 131.7 (CH Ar, C-5 or C-16), 132.6 (CH Ar, C-5 or C-16), 138.8 (CH Ar, C-11), 141.1 (qC), 144.5 (qC), 151.3 (qC), 155.5 (qC), 158.8 (qC).

HRMS (m/z ESI⁺): Found: 410.0520 (M^+ + H), C₁₈H₁₂N₃O₃ClF₃ Requires: 410.0519.

ν_{max} (ATR)/cm⁻¹: 3385 (NH), 3064, 2924, 1724, 1599, 1567, 1531 (NO₂), 1483 (NO₂), 1459, 1347, 1332, 1236 (CF₃), 1130 (C-O), 1111 (C-Cl), 1028, 999, 975, 868, 823, 786.

2-Amino-5-{3-[4-chloro-3-(trifluoromethyl)phenyl]aminophenoxy}pyridine (89)

Following Method F, a mixture of **88** (66 mg, 0.16 mmol) and $\text{SnCl}_2 \cdot 2\text{H}_2\text{O}$ (218 mg, 0.97 mmol) in 1 mL of absolute ethanol were heated at 70°C for 3 h. Usual work up and flash chromatography afforded **89** as a brown oil (48 mg, 80%).

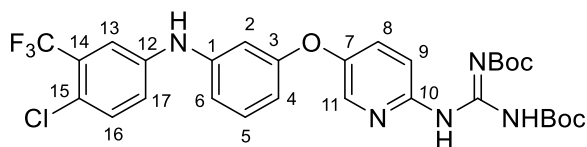
δ_{H} (400 MHz, CDCl_3): 4.41 (s, 2H, NH_2), 5.88 (s, 1H, NH), 6.53 (d, $J = 8.8$ Hz, 1H, H-9), 6.56 (dd, $J = 8.3, 2.2$ Hz, 1H, H-4), 6.61 (t, $J = 2.0$ Hz, 1H, H-2), 6.73 (dd, $J = 8.0, 1.7$ Hz, 1H, H-6), 7.11 (dd, $J = 8.6, 2.7$ Hz, 1H, H-17), 7.19 – 7.23 (m, 2H, H-5 and H-8), 7.32 – 7.34 (m, 2H, H-13 and H-16), 7.91 (d, $J = 2.1$ Hz, 1H, H-11).

δ_{C} (100 MHz, CDCl_3): 107.0 (CH Ar, C-2), 109.4 (CH Ar, C-9), 110.8 (CH Ar, C-4), 112.6 (CH Ar, C-6), 116.2 (q, $J = 5.4$ Hz, CH Ar, C-13), 121.1 (CH Ar, C-17), 122.8 (d, $J = 273.3$ Hz, qCF_3), 123.0 (qC, C-15), 129.2 (d, $J = 31.1$ Hz, CH Ar, C-14), 130.8 (CH Ar, C-5 or C-8), 131.1 (CH Ar, C-5 or C-8), 132.4 (CH Ar, C-16), 140.5 (CH Ar, C-11), 142.0 (qC), 143.2 (qC), 145.1 (qC), 155.4 (qC), 160.1 (qC).

HRMS (m/z ESI⁺): Found: 378.0629 ($\text{M}^+ - \text{H}$), $\text{C}_{18}\text{H}_{12}\text{N}_3\text{OF}_3\text{Cl}$ Requires: 378.0621.

ν_{max} (ATR)/ cm^{-1} : 3332 (NH), 2923 (C-H), 2847 (C-H), 1595, 1480, 1327, 1258, 1225, 1172, 1125 (CF_3), 1111 (C-Cl), 820, 711, 665.

1,2-Di-(*tert*-butoxycarbonyl)-3-(4-{3-[4-chloro-3-(trifluoromethyl)phenylamino]phenoxy}pyridin-2-yl)guanidine (90)



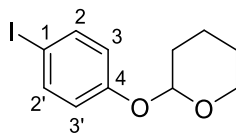
Following Method G, HgCl₂ (337 mg, 1.24 mmol) was added over a solution of **89** (451 mg, 1.19 mmol), *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methylothiourea **51** (359 mg, 1.24 mmol) and NEt₃ (0.69 mL, 4.17 mmol) in CH₂Cl₂ (6 mL). The reaction was stirred at room temperature for 24 h, then work up and silica gel chromatography (hexanes:EtOAc) afforded **90** (525 mg, 71%) as a white solid. **Mp**: 68-70 °C

δ_{H} (400 MHz, CDCl₃): 1.52 (s, 9H, (CH₃)₃), 1.54 (s, 9H, (CH₃)₃), 5.88 (s, 1H, NH), 6.59 (dd, *J* = 8.4, 2.1 Hz, 1H, H-4), 6.66 (t, *J* = 2.1 Hz, 1H, H-2), 6.80 (dd, *J* = 7.9, 1.8 Hz, H-6), 7.14 (dd, *J* = 8.7, 2.7 Hz, 1H, H-17), 7.24 (t, *J* = 8.1 Hz, 1H, H-5), 7.32 – 7.36 (m, 2H, H-13 and H-16), 7.40 (dd, *J* = 9.0, 2.8 Hz, 1H, H-8), 8.11 (d, *J* = 2.7 Hz, 1H, H-11), 8.39 (d, *J* = 8.6 Hz, 1H, H-9), 10.88 (s, 1H, NH), 11.52 (s, 1H, NH).

δ_{C} (100 MHz, CDCl₃): 28.2 ((CH₃)₃), 28.3 ((CH₃)₃), 80.1 (qC, C(CH₃)₃), 84.2 (qC, C(CH₃)₃), 108.1 (CH Ar, C-2), 111.7 (CH Ar, C-4), 113.3 (CH Ar, C-6), 116.6 (q, *J* = 5.7 Hz, CH Ar, C-13), 117.1 (CH Ar, C-9), 121.2 (CH Ar, C-17), 122.8 (d, *J* = 273.4 Hz, qCF₃), 123.3 (qC, C-15), 129.1 (CH Ar, C-8), 129.3 (d, *J* = 27.0 Hz, qC, C-14), 131.0 (CH Ar, C-5), 132.5 (CH Ar, C-16), 140.0 (CH Ar, C-11), 141.8 (qC), 143.5 (qC), 146.6 (qC), 143.8 (qC), 152.3 (qC), 153.1 (qC), 158.8 (qC), 163.3 (qC).

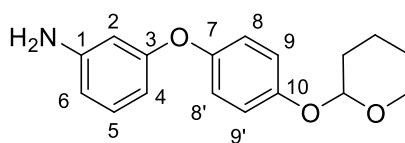
HRMS (*m/z* ESI⁺): Found: 622.2054 (M⁺ + H), C₂₉H₃₂N₅O₅ClF₃ Requires: 622.2044.

ν_{max} (ATR)/cm⁻¹: 3253 (N-H), 2980 (C-H), 1721 (C=O), 1597 (C=N), 1561, 1477, 1386, 1368, 1290 (C-O), 1122 (CF₃), 1140 (C-N), 1110, 1056 (C-Cl), 1027, 998, 978, 764, 664.

2-(4-Iodophenoxy)tetrahydro-2H-pyran (92**)**⁴⁰

To a solution of 4-iodophenol **91** (1.10 g, 5 mmol, 1 eq.) and pyridinium p-toluenesulfonate (PPTS) (126 mg, 0.5 mmol, 10 mol%) in dry CH₂Cl₂ (25 mL) was added 3,4- dihydro-2H-pyran (0.7 mL, 7.5 mmol, 1.5 eq.). The resultant mixture was stirred for 2.5 h at room temperature. The reaction mixture was diluted with diethyl ether and washed once with saturated brine. The organic layer was dried over anhydrous Na₂SO₄, filtered, and concentrated under reduced pressure. Purification by flash column chromatography over silica gel (hexane-EtOAc) gave the product **92** as a white solid (1.2 g, 82%). **Mp**: 63-64 °C (lit. 64-65 °C).⁴⁰

δ_H (400 MHz, CDCl₃): 1.56 – 1.73 (m, 3H, CH₂), 1.82 – 1.86 (m, 2H, CH₂), 1.92 – 2.03 (m, 1H, CH₂), 3.55 – 3.61 (m, 1H, CH₂), 3.82 – 3.88 (m, 1H, CH₂), 5.36 (t, *J* = 3.2 Hz, 1H, CH), 6.82 (dd, *J* = 8.4, 3.7 Hz, 2H, H-3 and H-3'), 7.54 (dd, *J* = 8.4, 3.7 Hz, 2H, H-2 and H-2').

3-{4-[(Tetrahydro-2H-pyran-2-yl)oxy]phenoxy}aniline (93**)⁴¹**

An oven-dried screw cap test tube was charged with a magnetic stirbar, copper(I) iodide (20.8 mg, 0.11 mmol, 5 mol%), 2-picolinic acid (26.8 mg, 0.22 mmol, 10 mol%), aryl iodide **92** (661.8 mg, 2.18 mmol), 3-aminophenol **35** (285.9 mg, 2.62 mmol) and K_3PO_4 (926 mg, 4.36 mmol). The tube was then evacuated and back-filled with argon (three times). Under a counterflow of argon, DMSO (4.0 mL) was added by syringe. The tube was placed in a preheated oil bath at 80 °C and the reaction mixture was stirred vigorously for 24 h. The reaction mixture was cooled to room temperature and diluted with EtOAc and H_2O . The organic layer was separated and the aqueous layer was extracted twice more with EtOAc. Combined organic layers were dried over Na_2SO_4 and the filtrate was concentrated and the resulting residue was purified via silica gel chromatography (hexane-EtOAc) to give **93** (534 mg, 86%) as a yellow liquid.

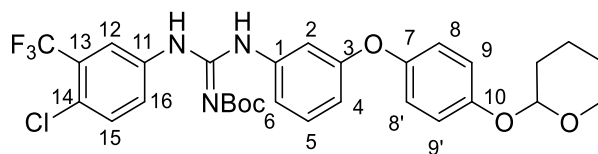
δ_H (400 MHz, $CDCl_3$): 1.58 – 1.73 (m, 3H, CH_2), 1.84 – 1.88 (m, 2H, CH_2), 1.96 – 2.04 (m, 1H, CH_2), 3.59 – 3.64 (m, 1H, CH_2), 3.90 – 3.98 (m, 1H, CH_2), 5.35 (t, $J = 3.4$ Hz, 1H, CH), 6.33 (t, $J = 2.2$ Hz, 1H, H-2), 6.38 (dd, $J = 8.2, 2.3$ Hz, 1H, H-4), 6.42 (dd, $J = 7.3, 2.1$ Hz, 1H, H-6), 6.96 (d, $J = 9.2$ Hz, 2H, H-8 and H-8' or H-9 and H-9'), 7.03 (d, $J = 9.1$ Hz, 2H, H-9 and H-9' or H-8 and H-8'), 7.07 (t, $J = 8.1$ Hz, 1H, H-5).

δ_C (150 MHz, $CDCl_3$): 19.0 (CH_2), 25.4 (CH_2), 30.6 (CH_2), 62.3 (CH_2), 97.4 (CH), 105.1 (CH Ar, C-2), 108.6 (CH Ar, C-4), 110.0 (CH Ar, C-6), 117.8 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 121.0 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 130.4 (CH Ar, C-5), 147.0 (qC), 150.9 (qC), 153.4 (qC), 159.7 (qC).

HRMS (m/z ESI⁺): Found 286.1454 ($M^+ + H$), $C_{17}H_{20}NO_3$, Requires: 286.1443.

ν_{max} (ATR)/ cm^{-1} : 3331 (NH), 2944 (CH), 2872 (CH), 1654, 1601, 1499, 1486, 1242, 1197 (C-O), 1180, 1034, 961, 835, 776, 660, 685.

1-[4-Chloro-3-(trifluoromethyl)phenyl]-2-(*tert*-butoxycarbonyl)-3-(3-{4-[(tetrahydro-2*H*-pyran-2-yl)oxy]phenyloxy}phenyl)guanidine (94**)**



Following a modification of Method C, HgCl_2 (261 mg, 0.96 mmol) was added over a solution of **93** (228 mg, 0.80 mmol), thiourea derivative **15** (341 mg, 0.96 mmol) and NEt_3 (0.34 mL, 2.48 mmol) in CH_2Cl_2 (5 mL). After 12 h workup and silica gel chromatography (hexanes:EtOAc) afforded compound **94** as a white amorphous solid (282 mg, 58%). **Mp**: 48-51 °C

δ_{H} (400 MHz, CDCl_3): 1.72 (s, 9H, $(\text{CH}_3)_3$), 1.82 – 1.96 (m, 3H, CH_2), 2.08 – 2.10 (m, 2H, CH_2), 2.19 – 2.24 (m, 1H, CH_2), 3.82 – 3.86 (m, 1H, CH_2), 4.13 – 4.19 (m, 1H, CH_2), 5.58 (s, 1H, CH), 6.74 – 6.99 (m, 4H, CH Ar), 7.20 – 7.28 (m, 4H, CH Ar), 7.48 – 7.50 (m, 2H, CH Ar), 7.62 – 7.64 (m, 1H, CH Ar), 8.17 (s, 1H, NH), 10.04 (s, 1H, NH).

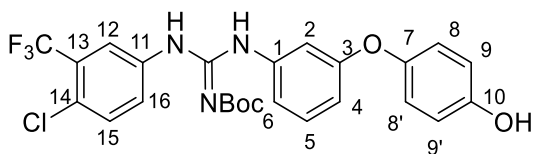
δ_{C} (100 MHz, CDCl_3): 19.0 (CH_2), 25.3 (CH_2), 28.2 ($(\text{CH}_3)_3$), 30.5 (CH_2), 62.2 (CH_2), 83.9 (qC, $\text{C}(\text{CH}_3)_3$), 97.1 (CH), 111.6, 112.8, 116.4, 117.8, 118.8, 120.9, 121.5, 122.1, 123.9, 124.2, 126.8, 129.9, 130.8, 131.8, 132.5, 137.9, 139.9, 150.6, 153.1, 153.6, 171.2.

Note: Aromatic protons and carbons, quaternary carbons and fluorine peaks could not be fully assigned and characterised because the peaks were very broad.

HRMS (m/z ESI⁺): Found: 604.1823 ($\text{M}^+ - \text{H}$), $\text{C}_{30}\text{H}_{30}\text{N}_3\text{O}_5\text{ClF}_3$ Requires: 604.1826.

ν_{max} (ATR)/ cm^{-1} : 3410 (NH), 2916 (CH), 2848 (CH), 1722 (C=O), 1664, 1590, 1556, 1500 (C=N), 1478, 1465, 1322, 1238, 1209, 1198, 1198, 1141 (CF_3), 1110 (C-Cl), 906, 731, 666.

1-[4-Chloro-3-(trifluoromethyl)phenyl]-2-(*tert*-butoxycarbonyl)-3-[3-(4-hydroxyphenyloxy)phenyl]guanidine (95**)⁴²**



A mixture of **94** (244 mg, 0.40 mmol), methanol (4 mL) and montmorillonite KSF (230 mg) was stirred at 50 °C for 4 h. After completion of the reaction, the catalyst was removed by filtration and washed with EtOAc. Evaporation of the solvent gave the corresponding crude product that was further purified by silica gel chromatography (hexane:EtOAc), to afford the title product as a white amorphous solid (135 mg, 65%).

Mp: 52-55 °C

δ_{H} (400 MHz, CDCl_3): 1.48 (s, 9H, $(\text{CH}_3)_3$), 6.41 – 6.73 (m, 3H, CHAr), 6.79 – 6.81 (m, 3H, CHAr), 6.91 – 7.00 (m, 3H, CHAr), 7.15 – 7.24 (m, 1H, CHAr), 7.38 – 7.40 (m, 1H, CHAr), 7.90 (bs), 9.81 (bs).

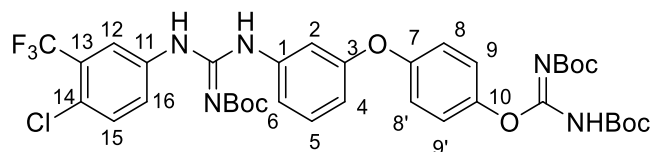
δ_{C} (100 MHz, CDCl_3): 28.2 ($(\text{CH}_3)_3$), 84.1 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 111.5, 112.6, 112.7, 116.5, 118.7, 119.2, 121.4, 124.2, 126.9, 128.9, 130.2, 130.8, 131.9, 132.2, 137.7, 140.3, 149.8, 152.3, 153.1, 153.2, 160.1.

Note: Aromatic protons and carbons, quaternary carbons and fluorine peaks could not be fully assigned and characterised because the peaks were very broad.

HRMS (m/z ESI): Found: 520.1246 ($\text{M}^- - \text{H}$), $\text{C}_{25}\text{H}_{22}\text{N}_3\text{O}_4\text{ClF}_3$ Requires: 520.1251.

ν_{max} (ATR)/ cm^{-1} : 3407 (NH), 3295 (OH), 2983, 1722 (C=O), 1658, 1589, 1556, 1505, 1461, 1203, 1140 (CF_3), 1092 (C-Cl), 835, 772.

***N,N'*-(Di-*tert*-butoxycarbonyl)-O-[4-(3-{1-[4-chloro-3-(trifluoromethyl)phenyl]-2-(di-*tert*-butoxycarbonyl)guanidino}phenoxy)phenyl]carbamimidate (96)**



Following a modification of Method G, HgCl_2 (114 mg, 0.42 mmol) was added over a solution of **95** (207 mg, 0.40 mmol), *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methylisothiurea **51** (122 mg, 0.42 mmol) and NEt_3 (0.20 mL, 1.4 mmol) in CH_2Cl_2 (4 mL). The reaction was stirred at room temperature for 24 h, then work up and silica gel chromatography (hexanes:EtOAc) afforded **96** (168 mg, 55%) as a white-yellow amorphous solid. **Mp**: 63-65 °C.

δ_{H} (400 MHz, CDCl_3): 1.33 – 1.65 (m, 27H, $(\text{CH}_3)_3$), 6.56 – 6.75 (m, 3H, CH Ar), 6.88 – 7.02 (m, 3H, CH Ar), 7.14 – 7.23 (m, 2H, CH Ar), 7.39 – 7.49 (m, 2H, CH Ar), 7.84 – 7.92 (m, 1H, CH Ar), 9.82 (bs, NH), 10.64 (bs, NH), 11.61 (bs, NH).

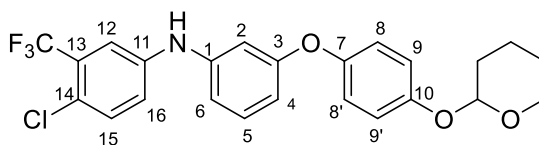
δ_{C} (100 MHz, CDCl_3): 28.1 ($(\text{CH}_3)_3$), 28.2 ($(\text{CH}_3)_3$), 84.0 (qC, $\text{C}(\text{CH}_3)_3$), 112.7, 113.6, 116.5, 117.3, 119.0, 120.0, 121.3, 121.5, 122.1, 123.1, 124.2, 126.9, 130.1, 131.0, 131.8, 132.3, 132.8, 137.8, 140.1, 146.9, 148.0, 153.1, 154.5, 158.9.

Note: Aromatic protons and carbons, quaternary carbons and fluorine peaks could not be fully assigned and characterised because the peaks were very broad.

HRMS (m/z ESI⁺): Found: 764.2676 ($\text{M}^+ + \text{H}$), $\text{C}_{36}\text{H}_{42}\text{N}_5\text{O}_8\text{ClF}_3$ Requires: 764.2674.

ν_{max} (ATR)/ cm^{-1} : 3415 (NH), 3268 (NH), 2981, 2933, 1769 (C=O), 1722 (C=O), 1659 (C=N), 1565 (C=N), 1410, 1297, 1238, 1132 (CF_3), 1088 (C-Cl), 957, 829, 770.

N-[4-Chloro-3-(trifluoromethyl)phenyl]-3-{4-[(tetrahydro-2H-pyran-2-yl)oxy]phenoxy}aniline (97)



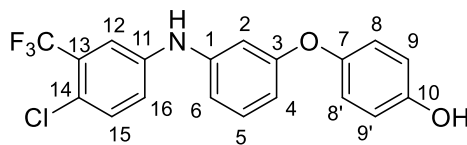
Following Method D, $\text{Pd}_2(\text{dba})_3$ (3 mol%, 37 mg), Xantphos (5 mol%, 39 mg), Cs_2CO_3 (615.8 mg, 1.89 mmol), compound **93** (385 mg, 1.35 mmol) and remaining liquid 4-bromo-1-chloro-2-(trifluoromethyl)benzene (0.20 mL, 1.35 mmol) were mixed, followed syringe addition of toluene (2.8 mL). The mixture was heated at 90 °C for 24 h and usual work up and flash chromatography afforded **97** as a yellow oil (532 mg, 85%).

δ_{H} (400 MHz, CDCl_3): 1.51 – 1.88 (m, 4H, CH_2), 1.96 – 2.05 (m, 2H, CH_2), 3.50 – 3.64 (m, 1H, CH_2), 3.86 – 4.04 (m, 1H, CH_2), 5.36 (t, $J = 3.3$ Hz, 1H, CH), 5.83 (s, 1H, NH), 6.59 (dd, $J = 8.1, 1.8$ Hz, 1H, H-4), 6.62 (t, $J = 2.1$ Hz, 1H, H-2), 6.72 (dd, $J = 7.9, 1.8$ Hz, 1H, H-6), 6.82 (d, $J = 8.9$, 2H, H-8 and H-8' or H-9 and H-9'), 6.94 (d, $J = 8.9$ Hz, 2H, H-8 and H-8' or H-9 and H-9'), 7.09 (dd, $J = 8.7, 2.6$ Hz, 1H, H-16), 7.21 (t, $J = 8.1$, H-5), 7.30 – 7.33 (m, 2H, H-12 and H-15).

δ_{C} (100 MHz, CDCl_3): 25.3 (CH_2), 30.6 (CH_2), 62.2 (CH_2), 62.4 (CH_2), 97.2 (CH), 107.6 (CH Ar, C-2), 111.5 (CH Ar, C-4), 112.6 (CH Ar, C-6), 116.0 (q, $J = 5.3$ Hz, qC, C-12), 116.5 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 121.0 (CH Ar, C-16), 121.5 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 122.9 (d, $J = 273.3$ Hz, q CF_3), 129.1 (d, $J = 31.2$ Hz, qC, C-13), 130.7 (CH Ar, C-5), 132.4 (CH Ar, C-15), 142.1 (qC), 143.0 (qC), 149.6 (qC), 152.4 (qC), 160.1 (qC).

HRMS (m/z ESI⁺): Found 462.1085 ($\text{M}^+ - \text{H}$), $\text{C}_{24}\text{H}_{20}\text{NO}_3\text{ClF}_3$ Requires: 462.1084.

ν_{max} (ATR)/ cm^{-1} : 3333 (NH), 3059, 3028, 2928, 1649, 1617, 1593, 1484, 1448, 1336 (C-N), 1177 (C-O), 1133 (CF_3), 1108 (C-Cl), 1133, 977, 764, 694.

4-(3-{{4-chloro-3-(trifluoromethyl)phenyl}amino}phenoxy)phenol (98)⁴²

A mixture of **97** (684 mg, 1.47 mmol), methanol (12mL) and montmorillonite KSF (700 mg) was stirred at 50 °C for 4 h. After completion of the reaction, the catalyst was removed by filtration and washed with EtOAc. Evaporation of the solvent gave the corresponding crude product that was further purified by silica gel chromatography (hexane:EtOAc) to afford the title product as a yellow liquid (558 mg, 99%).

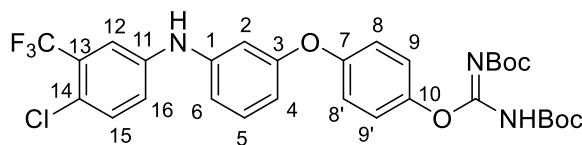
δ_{H} (400 MHz, CDCl_3): 6.59 (dd, $J = 7.9, 2.0$ Hz, 1H, H-4), 6.62 (app. t, $J = 2.2$ Hz, 1H, H-2), 6.72 (dd, $J = 8.0, 1.5$ Hz, 1H, H-6), 6.83 (d, $J = 8.9$ Hz, 2H, H-9 and H-9'), 6.95 (d, $J = 8.9$ Hz, 2H, H-8 and H-8'), 7.09 (dd, $J = 8.7, 2.7$ Hz, 1H, H-16), 7.21 (t, $J = 8.1$ Hz, 1H, H-5), 7.31 – 7.33 (m, 2H, H-12 and H-15).

δ_{C} (100 MHz, CDCl_3): 107.6 (CH Ar, C-2), 111.5 (CH Ar, C-4), 112.6 (CH Ar, C-6), 116.1 (q, $J = 5.5$ Hz, CH Ar, C-12), 116.5 (2 CH Ar, C-9 and C-9'), 120.9 (CH Ar, C-16), 121.5 (2 CH Ar, C-8 and C-8'), 122.8 (qC, C-14), 122.9 (d, $J = 273.5$ Hz, qCF_3), 129.1 (d, $J = 31.2$ Hz, qC, C-13), 130.7 (CH Ar, C-5), 132.4 (CH Ar, C-15), 140.1 (qC), 143.0 (qC), 149.6 (qC), 152.4 (qC), 160.1 (qC).

HRMS (m/z ESI^+): Found 380.0661 ($\text{M}^+ + \text{H}$), $\text{C}_{19}\text{H}_{14}\text{NO}_2\text{ClF}_3$ Requires: 380.0665.

ν_{max} (ATR)/ cm^{-1} : 3368 (OH), 2984 (C-H), 1704, 1596, 1503, 1484, 1442, 1331 (C-N), 1258, 1232, 1205 (C-O), 1138 (CF_3), 1114 (C-Cl), 1043, 1029, 837, 780, 683.

***N,N'*-(Di-*tert*-butoxycarbonyl)-*O*-(4-{3-[4-chloro-3-(trifluoromethyl)phenylamino]phenoxy} phenyl)carbamidate (**99**)**



Following a modification of Method G, HgCl_2 (251 mg, 0.93 mmol) was added over a solution of **98** (338 mg, 0.89 mmol), *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methylisothiurea **51** (269 mg, 0.93 mmol), and NEt_3 (0.43 mL, 3.11 mmol) in CH_2Cl_2 (5 mL). The reaction was stirred at room temperature for 24 h, then work up and silica gel chromatography (hexanes:EtOAc) afforded **99** as a yellow oil (364 mg, 65%).

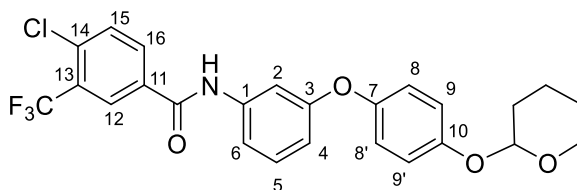
δ_{H} (400 MHz, CDCl_3): 1.43 (s, 9H, $(\text{CH}_3)_3$), 1.54 (s, 9H, $(\text{CH}_3)_3$), 5.92 (bs, 1H, NH), 6.62 (dd, $J = 8.4, 2.0$ Hz, 1H, H-4), 6.67 (t, $J = 2.2$ Hz, 1H, H-2), 6.78 (dd, $J = 7.9, 1.8$ Hz, 1H, H-6), 6.99 (d, $J = 9.0$ Hz, 1H, H-8 and H-8' or H-9 and H-9'), 7.11-7.16 (m, 3H, H-8 and H-8' or H-9 and H-9' and H-16), 7.23 (t, $J = 8.1$ Hz, 1H, H-5), 7.30 (d, $J = 2.8$ Hz, 1H, H-12), 7.33 (d, $J = 8.7$ Hz, 1H, H-15), 10.66 (bs, 1H, NH).

δ_{C} (100 MHz, CDCl_3): 28.1 ($(\text{CH}_3)_3$), 28.2 ($(\text{CH}_3)_3$), 81.4 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 83.2 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 108.9 (CH Ar, C-2), 112.4 (CH Ar, C-4), 113.3 (CH Ar, C-6), 116.3 (q, $J = 5.6$ Hz, CH Ar, C-12), 120.0 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 120.8 (qC, C-14), 120.9 (CH Ar, C-16), 122.9 (d, $J = 273.5$ Hz, q CF_3), 123.1 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 129.2 (d, $J = 31.2$ Hz, qC, C-13), 130.8 (CH Ar, C-5), 132.4 (CH Ar, C-15), 142.0 (qC), 143.3 (qC), 146.9 (qC), 148.4 (qC), 154.5 (qC), 158.68 (qC), 158.74 (qC), 162.1 (qC).

HRMS (m/z ESI^+): Found 644.1747 ($\text{M}^+ + \text{Na}$), $\text{C}_{30}\text{H}_{31}\text{N}_3\text{O}_6\text{ClF}_3$ Requires: 644.1751.

ν_{max} (ATR)/ cm^{-1} : 3359(NH), 2934, 1739 (C=O), 1770 (C=O), 1739 (C=O), 1625 (C=N), 1593 (C=N), 1483 (C-N), 1131 (CF_3), 1114, (C-Cl), 1090 (C-O), 876, 853, 818, 768.

4-Chloro-3-trifluoromethyl-N-(3-{4-[(tetrahydro-2H-pyran-2-yl)oxy]phenoxy}phenyl) benzamide (100)



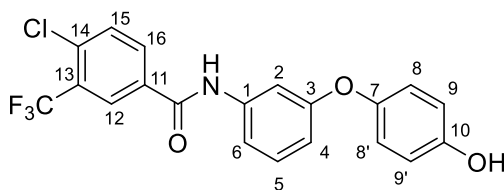
Following Method H, to a solution of 4-chloro-3-trifluoromethylbenzoic acid **58** (321 mg, 1.43 mmol) in CH_2Cl_2 (5 mL), oxalyl chloride (0.1 mL, 2.4 mmol) and DMF (3 drops) were added and the reaction was stirred overnight. Any remaining oxalyl chloride and DMF were removed under vacuum to a yellow residue that was dissolved in CH_2Cl_2 and cooled to 0 °C. Following, compound **93** (340 mg, 1.20 mmol) dissolved in CH_2Cl_2 (2 mL) and NEt_3 (0.40 mL, 2.88 mmol) were added dropwise. After 12 h, workup and silica gel chromatography afforded **100** (253 mg, 43%) as a white gum.

δ_{H} (400 MHz, CDCl_3): 1.64 – 1.70 (m, 3H, CH_2), 1.84 – 1.87 (m, 2H, CH_2), 1.95 – 2.00 (m, 1H, CH_2), 3.58 – 3.63 (m, 1H, CH_2), 3.90 – 3.96 (m, 1H, CH_2), 5.33 (t, $J = 3.3$ Hz, 1H, CH), 6.77 (dd, $J = 8.1, 1.9$ Hz, 1H, H-4), 6.95 (d, $J = 9.1$ Hz, 2H, H-8 and H-8' or H-9 and H-9'), 7.00 (d, $J = 9.1$ Hz, 2H, H-8 and H-8' or H-9 and H-9'), 7.14 (s, 1H, H-2), 7.27 (t, $J = 8.1$ Hz, 1H, H-5), 7.37 (d, $J = 8.2$ Hz, 1H, H-6), 7.59 (d, $J = 8.3$ Hz, 1H, H-15), 7.94 (dd, $J = 8.3, 1.9$ Hz, 1H, H-16), 7.99 (bs, NH), 8.16 (d, $J = 1.5$ Hz, 1H, H-12).

δ_{C} (100 MHz, CDCl_3): 21.2 (CH_2), 25.3 (CH_2), 30.6 (CH_2), 62.5 (CH_2), 64.0 (CH_2), 97.2 (CH), 109.7 (CH Ar, C-2), 114.6 (CH Ar, C-6), 116.6 (CH Ar, C-4), 117.9 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 121.1 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 121.5 (qC), 121.54 (d, $J = 273.8$ Hz, qCF_3), 126.6 (m, CH Ar, C-12), 129.1 (d, $J = 21.9$ Hz, qC, C-13), 130.3 (CH Ar, C-5), 131.6 (CH Ar, C-16), 132.1 (CH Ar, C-15), 133.7 (qC), 138.7 (qC), 150.5 (qC), 153.7 (qC), 159.3 (qC), 163.6 (qC).

HRMS (m/z ESI⁻): Found 490.1032 ($\text{M}^- - \text{H}$), $\text{C}_{25}\text{H}_{20}\text{NO}_4\text{ClF}_3$ Requires: 490.1033.

ν_{max} (ATR)/ cm^{-1} : 3310 (NH), 2945 (CH), 2873 (CH), 1738, 1652, 1601 (C=O), 1545, 1499, 1485, 1209, 1198 (CF_3), 1109 (C-Cl), 963, 919, 838, 871, 778 (C-Cl), 685, 660.

4-Chloro-3-trifluoromethyl-N-[3-(4-hydroxyphenoxy)phenyl]benzamide (101)⁴²

A mixture of **100** (240 mg, 0.49 mmol), methanol (4 mL) and montmorillonite KSF (250 mg) was stirred at 50 °C for 4 h. After completion of the reaction, the catalyst was removed by filtration and washed with EtOAc. Evaporation of the solvent gave the corresponding crude product that was further purified by silica gel chromatography (hexane:EtOAc) to afford the title product as a white gum (148 mg, 74%).

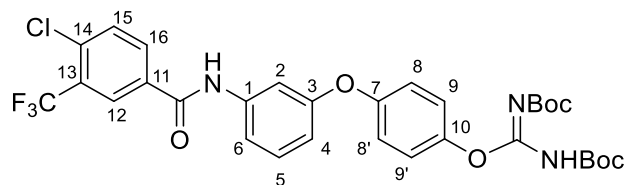
δ_{H} (400 MHz, CDCl_3): 5.10 (bs, 1H, OH), 6.76 – 6.80 (m, 3H, H-8 and H-8' or H-9 and H-9' and H-4), 6.82 (d, $J = 8.9$ Hz, 2H, H-8 and H-8' or H-9 and H-9'), 6.95 (d, $J = 8.9$ Hz, 2H, H-8 and H-8' or H-9 and H-9'), 7.23 (bs, 1H, H-2), 7.28 – 7.31 (m, 2H, H-6 and H-5), 7.62 (d, $J = 8.3$ Hz, 1H, H-15), 7.85 (bs, 1H, NH), 7.95 (dd, $J = 8.3, 2.1$ Hz, 1H, H-16), 8.14 (d, $J = 2.0$ Hz, 1H, H-12).

δ_{C} (100 MHz, CDCl_3): 109.6 (CH Ar, C-2), 114.4 (CH Ar, C-4 or C-6), 114.5 (CH Ar, C-4 or C-6), 116.6 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 119.8 (d, $J = 273.7$ Hz, qCF_3), 121.5 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 123.9 (qC), 126.6 (m, CH Ar, C-12), 129.2 (d, $J = 32.3$ Hz, qC, C-13), 130.3 (CH Ar, C-5), 131.5 (CH Ar, C-16), 132.2 (CH Ar, C-15), 133.7 (qC), 138.6 (qC), 149.7 (qC), 152.4 (qC), 159.5 (qC), 163.6 (qC).

HRMS (m/z ESI⁺): Found 406.0462 ($\text{M}^+ - \text{H}$), $\text{C}_{20}\text{H}_{12}\text{NO}_3\text{ClF}_3$ Requires: 406.0458.

ν_{max} (ATR)/ cm^{-1} : 3322 (OH), 2973 (CH), 1736 (C=O), 1659, 1602, 1505, 1486, 1444, 1243 (CF_3), 1142 (C-Cl), 1038, 915, 843, 753, 701.

***N,N'*-(Di-*tert*-butoxycarbonyl)-O-(4-{3-[4-chloro-3-(trifluoromethyl)benzamido]phenoxy} phenyl)carbamimidate (**102**)**



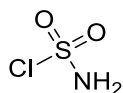
Following a modification of Method G, HgCl_2 (103 mg, 0.38 mmol) was added over a solution of **101** (150 mg, 0.37 mmol), *N,N'*-bis-(*tert*-butoxycarbonyl)-*S*-methylisothiurea **51** (111 mg, 0.38 mmol) and NEt_3 (0.18 mL, 1.30 mmol) in CH_2Cl_2 (2 mL). The reaction was stirred at room temperature for 24 h, then work up and silica gel chromatography (hexanes:EtOAc) afforded **102** (233 mg, 97%) as a white gum.

δ_{H} (400 MHz, CDCl_3): 1.43 (bs, 18H, $2(\text{CH}_3)_3$), 6.74 – 6.77 (m, 3H, H-8 and H-8' or H-9 and H-9' and H-4), 6.90 (d, $J = 9.0$ Hz, 2H, H-8 and H-8' or H-9 and H-9'), 6.98 (t, $J = 2.0$, 1H, H-2), 7.24 (t, $J = 8.1$ Hz, 1H, H-5), 7.50 (dd, $J = 8.2$, 1.0 Hz, H-6), 7.56 (d, $J = 8.4$ Hz, 1H, H-15), 8.03 (dd, $J = 8.4$, 1.9 Hz, 1H, H-16), 8.27 (d, $J = 1.7$ Hz, 1H, H-12), 9.08 (bs, 1H, NH).

δ_{C} (100 MHz, CDCl_3): 28.0 (s, 9H, $(\text{CH}_3)_3$), 28.1 (s, 9H, $(\text{CH}_3)_3$), 83.4 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 83.5 (qC, $\underline{\text{C}}(\text{CH}_3)_3$), 110.9 (CH Ar, C-2), 115.4 (CH Ar, C-4 or C-6), 115.9 (CH Ar, C-4 or C-6), 119.5 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 122.7 (d, $J = 273.7$ Hz, qCF₃), 122.9 (2 CH Ar, C-8 and C-8' or C-9 and C-9'), 126.8 (qC, C-14), 126.9 (q, $J = 5.1$ Hz, CH Ar, C-12), 128.5 (q, $J = 31.9$ Hz, qC, C-13), 130.3 (CH Ar, C-5), 131.0 (CH Ar, C-15), 132.5 (CH Ar, C-16), 133.6 (qC), 136.0 (qC), 139.5 (qC), 146.3 (qC), 148.9 (qC), 154.9 (qC), 157.3 (qC), 159.1 (qC), 163.7 (qC).

HRMS (m/z ESI⁺): Found 648.1721 ($\text{M}^+ - \text{H}$), $\text{C}_{31}\text{H}_{30}\text{N}_3\text{O}_7\text{ClF}_3$, Requires: 648.1724.

ν_{max} (ATR)/ cm^{-1} : 3412 (NH), 3314 (NH), 2981 (CH), 1769 (C=O), 1737, 1663, 1630, 1620, (C=N)1484, 1369, 1293, 1245, 1131 (CF₃), 1089 (C-Cl), 913, 874, 822, 752.

Sulfamoyl chloride (103**)**⁴³

Formic acid (0.93 mL, 24.6 mmol), was added dropwise at 0 °C and under stirring to chlorosulfonyl isocyanate **104** (2.14 mL, 24.6 mmol). When the evolution of gases had ceased, normally after 1 h, the resulting solid was dissolved in toluene, the mixture was filtered and the filtrate evaporated at reduced pressure. Product **103** was obtained without any further purification as a brown solid (2.84 g, 100 %). **Mp**: 33 – 35 °C (lit. 40 °C).⁴³

δ_{H} (400 MHz, DMSO-*d*₆): 8.45 – 9.67 (bs, NH₂).

7.3. Biochemical Procedures

7.3.1. Materials

All reagents were purchased from Sigma Aldrich or Gibco (Life Technologies) and sterile when required. Sorafenib was purchased from LC laboratories. Buffers were prepared when needed according to the following recipes: 20X Annexin Binding Buffer (pH 7.4 – 50 mL: 2.60 g HEPES 10.9 mM, 8.18 g NaCl 140 mM, 0.28 g CaCl_2 2.5 mM); 1X Annexin Binding Buffer (40 mL: 2 mL 20X Binding Buffer, 38 mL PBS); Annexin V (600 μL : 18 μL Annexin V, 582 μL 1X Annexin Binding Buffer); Propidium Iodide (3 μL propidium iodide, 6 mL 1X Annexin Binding Buffer); Laemmli buffer [Tris-HCl 50 mM (pH 6.7), glycerol 10% (w/v), sodium dodecyl sulphate 2% (w/v), bromophenol blue 0.02% (w/v)]; Radioimmunoprecipitation assay buffer (RIPA buffer) was purchased from SigmaAldrich and protease inhibitors freshly added before use; running buffer 10 X for wester blotting (250 mM Tris base, 2 M glycine, 5% v/v SDS); transfer buffer 10 X for wester blotting (250 mM Tris base, 2 M glycine).

7.3.2. General Procedures

7.3.2.1. Cell Storage/Cryopreservation

When an aliquot of cells was required, the cells were quickly removed from the liquid nitrogen and thawed at 37 °C for 2 min. Just before the samples had fully thawed, their contents were gently pipetted into a sterile 20 mL tube containing growing medium. The cells were then centrifuged at 300 x *g* for 5 min. The supernatant was discarded and the pellet was resuspended in 5 mL of medium. This solution was then added to a T25 flask and the cells monitored closely over the next few days.

The growth medium was stored in the fridge at 4 °C and heated to 37 °C prior to culture work. Cells were grown at 37 °C in a humidified environment maintained at 95% O_2 and 5% CO_2 and passaged at least twice weekly depending on their levels of confluency. When required for sub-culturing, if in suspension, cells were transferred to a sterile tube and centrifuged at 300 x *g* for 5 min. Otherwise, if adherent cells, cells were trypsinised and transferred to a sterile tube and centrifuged at 300 x *g* for 5 min.

The supernatant was discarded and the cell pellet was resuspended in 15 mL of fresh medium.

7.3.2.2. Use of Haemocytometer

A haemocytometer is a modified microscope slide containing an accurately sub-divided grid that enables the counting of cells in suspension. Each grid (there are two) consists of nine large squares, each measuring 1 mm^2 . Each of these primary squares contains 16 medium squares, each measuring 0.04 mm^2 . When a particular seeding density of cells was required, a clean cover slip was placed on the haemocytometer; $10\text{ }\mu\text{L}$ of the cell suspension was then pipetted into the groove at the end of the plane. The cell suspension was then drawn across the grid by capillary action. The number of cells within the four outer primary squares was counted. Those that touched the sides of the grid were counted and those that were outside the four squares were not counted. As each medium square is 0.04 mm^3 , the total volume per primary square is $1 \times 10^{-4}\text{ mL}$. This implies that the total cell concentration in the original suspension (cells/mL) is = average cell count/primary square $\times 1 \times 10^4$.

7.3.2.3. Preparation of the Drugs' Solutions

Stock solutions (10 mM) of the compounds were prepared in sterile EtOH or DMSO and were then sterile filtered ($0.2\text{ }\mu\text{M}$ filters). Required concentration ranges (10-0.01 mM) of each drug were prepared in sterile EtOH or DMSO/ddH₂O and stored at $-20\text{ }^{\circ}\text{C}$ until required. Sorafenib was purchased from LC laboratories.

7.3.3. Growth and Maintenance of the Cell Lines

7.3.3.1. HL-60 Cell Line

The HL-60 (human caucasian promyelocytic leukemia) cell line was obtained from European Collection of Cell Cultures (Porton Down, Wiltshire, U.K.). It was maintained between 200,000 and 2,000,000 cells/mL in Roswell Park Memorial Institute (RPMI) 1640 medium with stable glutamate (GlutaMax I) supplemented with 10% (v/v) fetal bovine serum (FBS) and 1% penicillin/ streptomycin (pen-strep).

7.3.3.2. MCF-7 Cell Line

The MCF-7 (human breast tumour) cell line was obtained from Prof. Mary Meegan, School of Pharmacy, TCD. It was maintained in Roswell Park Memorial Institute (RPMI) 1640 medium with stable glutamate (GlutaMax I) supplemented with 10% (v/v) fetal bovine serum (FBS), 50 $\mu\text{g mL}^{-1}$ penicillin/ streptomycin (pen-strep) and 2 mM L-glutamine.

7.3.3.3. HeLa Cell Line

The HeLa (human epitheloid cervix carcinoma) cell line was obtained from Prof. Thorfinnur Gunnlaugsson, School of Chemistry TCD. It was maintained in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% (v/v) fetal bovine serum (FBS) and 50 $\mu\text{g mL}^{-1}$ penicillin/ streptomycin (pen-strep).

7.3.3.4. HCT116 and HKH-2 Cell Lines

The HCT116 and HKH-2 (human colon cancer cells) cell lines were obtained from Prof. James Murray, School of Biochemistry and Immunology TCD. They were maintained in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% (v/v) fetal bovine serum (FBS), 1% (v/v) penicillin/ streptomycin (pen-strep), 1% (v/v) sodium pyruvate (NaPyr), 1% (v/v) L-glutamine (L-Glu) and, only for HKH-2, 0.6 mg mL^{-1} geneticine (G418).

7.3.3.5. NCI-H929 and U266B1 Cell Line

The NCI-H929 and U266B1 (human bone marrow myeloma) cell lines were cultured by PhD student Rebecca Amet in Prof. Daniela Zisterer's group (School of Biochemistry and Immunology, TCD). They were maintained in Roswell Park Memorial Institute (RPMI) 1640 medium with stable glutamate (GlutaMax I) supplemented with 10% (v/v) fetal bovine serum (FBS) and 1% penicillin/ streptomycin (pen-strep).

7.3.3.6. MCF-10A

MCF-10a (human breast epithelial) cell line was obtained from Prof. Mary Meegan, School of Pharmacy, TCD. It was cultured in Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F12; Gibco) supplemented with 5% horse serum (Invitrogen), 20 ng mL⁻¹ epidermal growth factor (Merck Millipore), 0.5 µg mL⁻¹ hydrocortisone (Sigma), 100 ng mL⁻¹ cholera toxin (Sigma), 10 µg/mL insulin (Sigma), and penicillin/streptomycin (1% v/v) (Gibco).

7.3.4. AlamarBlue® Cell Viability Assay

Cells were counted and seeded at a density of 2×10^5 cells/mL for HL-60, 2.5×10^4 cells/mL for MCF-7, MCF10A and HeLa, 1×10^5 cells/mL for HCT116 and HKH-2, 2.5×10^5 cells/mL for NCI-H929 and U266B, all of them in complete medium. The 96 well plates were then treated with a 1:100 dilution of stock concentrations of drugs or EtOH (1% v/v)/DMSO (0.1% v/v) as vehicle control in triplicate. Three blank wells containing 200 µL RPMI with no cells were also set-up as blanks. After a 72 h incubation, 20 µL of AlamarBlue was added to each well. The plates were incubated in darkness at 37 °C for 4-5 h. Using a Molecular Devices microplate reader, the fluorescence (F) was then read at an excitation wavelength of 544 nm and an emission wavelength of 590 nm. Cell viability was then determined by subtracting the mean blank fluorescence (F_b) from the treated sample fluorescence (F_s) and expressing this as a percentage of the fluorescence of the blanked vehicle control (F_c). This is demonstrated in the equation below. The results were then plotted as a nonlinear regression, sigmoidal dose-response curves on Prism, from which the IC₅₀ value for each drug was determined.

$$\frac{F_s - F_b}{F_c - F_b} \times 100 = \% \text{ Cell Viability}$$

7.3.5. Effect on Human Platelets

Platelet evaluation was carried out by Prof. Maria Jose Santos-Martinez's group (School of Pharmacy and Pharmaceutical Sciences, TCD).

7.3.5.1. Preparation of Platelets

This study was approved by the School of Pharmacy and Pharmaceutical Sciences Research Ethics Committee (2015-06-01 MS). Signed informed consent was obtained from all donors before donation. Blood was collected from healthy volunteers who had abstained from taking any medications known to affect platelet function in the 14 days prior to testing. Blood was taken from a forearm vein using tubing, a syringe and a butterfly needle (21 gauge) with minimal or no venostasis.⁴⁴ Platelet-rich plasma (PRP) and washed platelet (WP) suspensions were prepared as described before and adjusted to 2.5×10^8 platelets/mL using Tyrode's salt solution.⁴⁵

7.3.5.2. Light Transmission Aggregometry (LTA)

Platelet function was investigated using an eight channel PAP 8 aggregometer (Bio/Data Corporation, Ireland). This technique is based on the principle and technique first described by Born.⁴⁶ Briefly, platelet aggregation is determined by the percentage of light passing through a suspension of platelets (PRP), where its autologous PPP is considered as 100% of aggregation. PRP was incubated under continuous stirring (900 rpm) at 37°C in the presence or absence of the vehicle (ethanol) and compounds at different concentrations and changes in light transmission recorded by the device. Platelet aggregation induced by collagen (2µg/mL) was tested in every donor as internal control of platelet function.

7.3.5.3. Optical Microscopy

To corroborate the effect of the compounds on platelet function, samples from LTA were taken and studied using an Olympus Microscope (Mason Technology, Ireland). Platelet samples were fixed with 2% paraformaldehyde for 30 minutes at 37°C. Afterwards, mounted on slides using a Cytospin™ 4 Cytocentrifuge (Thermo Scientific, Ireland) at 1,500 rpm for 30 minutes and visualized under the microscope. Photomicrographs were captured with a digital camera attached to the microscope.

7.3.5.4. CytoTox-ONE Homogeneous Membrane Integrity Assay (Promega G7890)

The potential toxic effect of the compounds on platelets was assessed using a cytotoxicity assay based on the detection of lactate dehydrogenase (LDH). Lactate dehydrogenase is an intracellular enzyme responsible for the conversion of lactate to pyruvate. Using this assay, the amount of released LDH from damaged cells, in this case platelets, can be measured by fluorescence. The assay was performed in duplicate for each compound with at least three different blood donors.

The percentage of cytotoxicity was calculated following the manufacturer's instructions and using the following equation:

$$\frac{\text{experimental} - \text{culture medium background}}{\text{maximal LDH release} - \text{culture medium background}} \times 100 = \% \text{ Cytotoxicity}$$

Where the readings obtained corresponded to: Experimental (washed platelets incubated in the presence of the compounds to be tested); Culture medium background (Tyrode's salt solution); Maximal LDH release (normalised as 100%, WP in the presence of lysis solution supplied with the assay).

7.3.5.5. Statistical analysis

Data from at least three independent experiments were analyzed using GraphPad Prism 5 software (GraphPad Software, La Jolla, CA, USA). All means are reported with standard deviation. One-way analysis of variance (ANOVA) and Tukey multiple comparisons post-test were performed as appropriate. Statistical significance was considered at $p < 0.05$.

7.3.6. Flow Cytometric Analysis

Flow cytometric analysis was carried out by PhD student Rebecca Amet in Prof. Daniela Zisterer's group (School of Biochemistry and Immunology, TCD).

NCI-H929 and U266 cells were seeded at 3×10^5 cells/mL in a total volume of 2 mL in a 12-well plate, untreated, treated with the vehicle (EtOH, 0.5% v/v) and with serial dilutions of the compound of interest. Cells were incubated for 24 h. Cell suspension from each well were then transferred to a 20 mL falcon tubes and adherent cells remaining were washed from the base of the wells with 1 mL PBD and added to the cell suspension in the falcon tubes. Cells were then centrifuged at 1600 g for 10 min and supernatant was removed. Cells pellets were rinsed with 1X Annexin V binding buffer (500 μ L) and transferred to Eppendorf for centrifugation at 600 g for 10 minutes at 4 °C. The supernatant was removed leaving pellets for staining. For AnV staining, pellets were re-suspended in AnV (50 μ L) and vortexed. Samples were then left on ice for 20 min in darkness. 1X Annexin V binding buffer (500 μ L) was added to the samples and then centrifuged at 600 g for 5 min. For PI staining, pellets were re-suspended in PI (500 μ L) and kept in darkness. Following staining, samples were kept on ice and analysed immediately. Data was gated to exclude cell debris and doublets. 10,000 cells were analysed using BD Accuri C6 software through the 630/22 nm band pass filter. Graphs were prepared using GraphPad Prism.

7.3.7. Western Blotting

Western Blotting study was carried out by PhD student Rebecca Amet in Prof. Daniela Zisterer's group (School of Biochemistry and Immunology, TCD).

NCI-H929 and U266B1 cells were seeded at 3×10^5 cells/mL in a total volume of 2 mL in a T25 flask. They were then treated with either vehicle (0.5% EtOH), Sorafenib (10 μ M) or compound of interest (10 μ M) for 16 hours. Following incubation, cell suspension was transferred to a 20 mL tube and centrifuged at 1200 rpm for 5 min, the supernatant was discarded and tubes were inverted to remove remaining drops. Cells pellets were re-suspended in 1 mL of PBS and transferred to an Eppendorf and centrifuged at 600 x g for 10 min. Again, supernatant was removed and pellets were kept on ice.

Radioimmunoprecipitation assay buffer (RIPA buffer, 100 μ L or 200 μ L depending on the size of the pellet), supplemented with 1% v/v phosphatase inhibitor cocktail 2 and 3 and 10% of protease inhibitor, was added to each sample to induce cell lysis while preventing protein degradation. Sample were stored on ice for 30 min, and then

centrifuged in a fast-cool centrifuge at 12000 x g for 10 min at 4 °C. Supernatant was transferred to new Eppendorf and kept on ice and the pellets were discarded.

The bicinchoninic acid (BCA) assay was then used to determine the amount of protein required to make up 100 μL of a 100 mg mL^{-1} protein sample. Absorbance of protein was measured at 562 nm using a Spectramax Plus 384 Microplate Reader. Protein concentration for each sample was determined by interpolation from a standard curve generated using serial dilutions of known concentrations of bovine serum albumin (BSA).

Following, the appropriate volume of protein was transferred to new Eppendorf. Laemmli sample buffer (20 μL) was added and each sample was made up to 100 μL with dH_2O . Samples were boiled at 90 °C for 3 min and 50 mM dithiothreitol (DTT) was added (5 μL).

Then, 20 μg protein was loaded and separated on 12% SDS-page gel. Separated proteins were transferred on apolyvinylidene difluoride (PVDF) membrane and probed with anti-Mcl-1[1:1000] and anti-rabbit secondaries. Anti-GAPDH[1:2500] was used as a loading control. For protein detection, the membrane was soaked in electrochemiluminescence (ECL) mixture for 5 min at room temperature and developed in a BioRad Image Gel Doc system.

7.3.8. Thermal Melting B-Raf^{V600E}

Thermal stability-shift assays were carried out by Dr. Apirat Chaikuad in Prof. Stefan Knapp's group at the University of Frankfurt using, with a Real-Time PCR Mx 3005 p machine (Stratagene). In brief, B-Raf^{V600E} was mixed with 20 μM kinase inhibitors. The experiments and data evaluation for the melting temperatures were performed according to previously described protocols.⁴⁷

7.4. References

1. Šali, A.; Blundell, T. L., *J. Mol. Biol.* **1993**, 234, 779.
2. Fiser, A.; Do, R. K.; Sali, A., *Protein Science : A Publication of the Protein Society* **2000**, 9, 1753.
3. Hénon, E.; Dauchez, M.; Haudrechy, A.; Banchet, A., *Tetrahedron* **2008**, 64, 9480.
4. Case, D.; Darden, T.; Cheatham III, T.; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Walker, R.; Zhang, W.; Merz, K., *University of California, San Francisco* **2010**, 5.
5. Homeyer, N.; Horn, A. H. C.; Lanig, H.; Sticht, H., *J. Mol. Model.* **2006**, 12, 281.
6. Dupradeau, F.-Y.; Cézard, C.; Lelong, R.; Stanislawiak, É.; Pêcher, J.; Delepine, J. C.; Cieplak, P., *Nucleic Acids Res.* **2008**, 36, D360.
7. Nicolai, A.; Delarue, P.; Senet, P., *PLoS Comp. Biol.* **2013**, 9, e1003379.
8. Wales, D., *Energy landscapes: Applications to clusters, biomolecules and glasses*. Cambridge University Press, 2003.
9. Cote, Y.; Senet, P.; Delarue, P.; Maisuradze, G. G.; Scheraga, H. A., *Proc. Natl. Acad. Sci.* **2012**, 109, 10346.
10. Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W., *J. Am. Chem. Soc.* **2010**, 132, 6498.
11. Rubez, G.; Etancelin, J.-M.; Vigouroux, X.; Krajecki, M.; Boisson, J.-C.; Hénon, E., *J. Comput. Chem.* **2017**, 38, 1071.
12. Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D. N.; Yang, W., *J. Chem. Theor. Comput.* **2011**, 7, 625.
13. Lefebvre, C.; Rubez, G.; Khartabil, H.; Boisson, J.-C.; Contreras-Garcia, J.; Henon, E., *Phys. Chem. Chem. Phys.* **2017**.
14. Wallace, A. C.; Laskowski, R. A.; Thornton, J. M., *Protein Eng.* **1995**, 8, 127.
15. Barberot, C.; Boisson, J.; Gerard, S.; Khartabil, H.; Thiriot, E.; Monard, G.; Henon, E., *Comput. Theor. Chem.* **2014**, 1028, 7.
16. Stewart, J. J. P., *J. Mol. Model.* **2013**, 19, 1.
17. Stewart, J. J., *Int. J. Quantum Chem.* **1996**, 58, 133.

18. Stewart, J. J. P., *Stewart Computational Chemistry, Colorado Springs, CO, USA, MOPAC 2016*.
19. Stockbridge, R. B.; Wolfenden, R., *J. Biol. Chem.* **2009**, *284*, 22747.
20. Møller, C.; Plesset, M. S., *Phys. Rev.* **1934**, *46*, 618.
21. Frisch, M. J.; Pople, J. A.; Binkley, J. S., *J. Chem. Phys.* **1984**, *80*, 3265.
22. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Gaussian, Inc.: Wallingford CT, 2009.
23. Rozas, I.; Kruger, P. E., *J. Chem. Theor. Comput.* **2005**, *1*, 1055.
24. Rozas, I.; Sánchez-Sanz, G.; Alkorta, I.; Elguero, J., *J. Phys. Org. Chem.* **2013**, *26*, 378.
25. Bader, R.; *Molecules*, A. I., *A Quantum Theory*, Clarendon. Oxford, England: 1990.
26. Keith, T. A. *AIMAll, TK Gristmill Software*, Version 17.01.25; 2017.
27. Reed, A. E.; Curtiss, L. A.; Weinhold, F., *Chem. Rev.* **1988**, *88*, 899.
28. Diez-Cecilia, E.; Kelly, B.; Perez, C.; Zisterer, D. M.; Nevin, D. K.; Lloyd, D. G.; Rozas, I., *Eur. J. Med. Chem.* **2014**, *81*, 427.
29. Surry, D. S.; Buchwald, S. L., *Chem. Sci.* **2011**, *2*, 27.
30. Bellamy, F. D.; Ou, K., *Tetrahedron Lett.* **1984**, *25*, 839.
31. Kelly, B.; McMullan, M.; Muguruza, C.; Ortega, J. E.; Meana, J. J.; Callado, L. F.; Rozas, I., *J. Med. Chem.* **2015**, *58*, 963.
32. Kurti, L.; Czakó, B., *Strategic applications of named reactions in organic synthesis*. Elsevier, **2005**.
33. Dardonville, C.; Goya, P.; Rozas, I.; Alsasua, A.; Martín, M. I.; Borrego, M. J., *Biorg. Med. Chem.* **2000**, *8*, 1567.
34. Miel, H.; Rault, S., *Tetrahedron Lett.* **1997**, *38*, 7865.

35. Okaniwa, M.; Hirose, M.; Imada, T.; Ohashi, T.; Hayashi, Y.; Miyazaki, T.; Arita, T.; Yabuki, M.; Kakoi, K.; Kato, J.; Takagi, T.; Kawamoto, T.; Yao, S.; Sumita, A.; Tsutsumi, S.; Tottori, T.; Oki, H.; Sang, B.-C.; Yano, J.; Aertgeerts, K.; Yoshida, S.; Ishikawa, T., *J. Med. Chem.* **2012**, *55*, 3452.
36. Kahvedzic-Seljubac, A.; Nathwani, S. M.; Zisterer, D. M.; Rozas, I., *Eur. J. Med. Chem.* **2016**, *117*, 269.
37. Cohen, E.; Klarberg, B., *J. Am. Chem. Soc.* **1962**, *84*, 1994.
38. Oguro, Y.; Cary, D. R.; Miyamoto, N.; Tawada, M.; Iwata, H.; Miki, H.; Hori, A.; Imamura, S., *Biorg. Med. Chem.* **2013**, *21*, 4714.
39. Patent number: WO2008/46216, April 2008.
40. Li, Y. N.; Urbas, A.; Li, Q., *J. Org. Chem.* **2011**, *76*, 7148.
41. Maiti, D.; Buchwald, S. L., *J. Am. Chem. Soc.* **2009**, *131*, 17423.
42. Li, T.-S.; Zhang, Z.-H.; Jin, T.-S., *Synth. Commun.* **1999**, *29*, 181.
43. Shahani, V. M.; Ball, D. P.; Ramos, A. V.; Li, Z. H.; Spagnuolo, P. A.; Haftchenary, S.; Schimmer, A. D.; Trudel, S.; Gunning, P. T., *Biorg. Med. Chem.* **2013**, *21*, 5618.
44. Cattaneo, M.; Cerletti, C.; Harrison, P.; Hayward, C. P. M.; Kenny, D.; Nugent, D.; Nurden, P.; Rao, A. K.; Schmaier, A. H.; Watson, S. P.; Lussana, F.; Pugliano, M. T.; Michelson, A. D., *J. Thromb. Haemost.* **2013**, *11*, 1183.
45. Radomski, M.; Moncada, S., *Thromb. Res.* **1983**, *30*, 383.
46. Born, G. V. R., *Nature* **1962**, *194*, 927.
47. Vedadi, M.; Niesen, F. H.; Allali-Hassani, A.; Fedorov, O. Y.; Finerty, P. J.; Wasney, G. A.; Yeung, R.; Arrowsmith, C.; Ball, L. J.; Berglund, H.; Hui, R.; Marsden, B. D.; Nordlund, P.; Sundstrom, M.; Weigelt, J.; Edwards, A. M., *Proc. Natl. Acad. Sci.* **2006**, *103*, 15835.

Appendix

A.1. Computational Modelling

A.1.1. Arylguanidinium – Phosphate Complexes

The interactions between the phenylguanidinium cation and phosphate have been studied by means of MP2/6-311+G(d,p)//6-31+G(d,p) using the PCM-water solvent model. In order to analyse the influence in the interactions of different substituents in the *para* position of the phenyl ring, five groups have been chosen, R= COH, NH₂, OH, OPh and CH₃. Since the guanidinium group presents a formal positive charge, interactions with anions are expected to be strong and favourable.

A.1.1.1. Optimized Structures and Interaction Energies

Optimized structures of the different complexes obtained are represented in Figure A.1.1.1 and their structural parameters summarized in Table A.1.1.1. Four different complexes have been identified for all the aryl substituents considered, with the exception of R= COH which presents five complexes. As shown in Figure A.1.1.1, the interaction between both molecules presents two different possible contacts, the first one involves the simultaneous interaction between two H atoms of the guanidinium moiety (H² and H³) and two O atoms of the phosphate anion (O¹ and O²) and is present in complexes **Ia,b**, **II** and **III** (Figure A.1.1.1.). The second complexation pattern involves an interaction between the same hydrogens of the guanidinium and a single oxygen atom of phosphate and is present in complexes **IVa,b** (Figure A.1.1.1.). In the particular case of R= COH (complex **Ib**), an additional contact is observed where the phosphate moiety forms an extra interaction between the O atom not involved in the interactions shown in complexes **Ia**, **II** and **III** and a H atom of the phenyl ring.

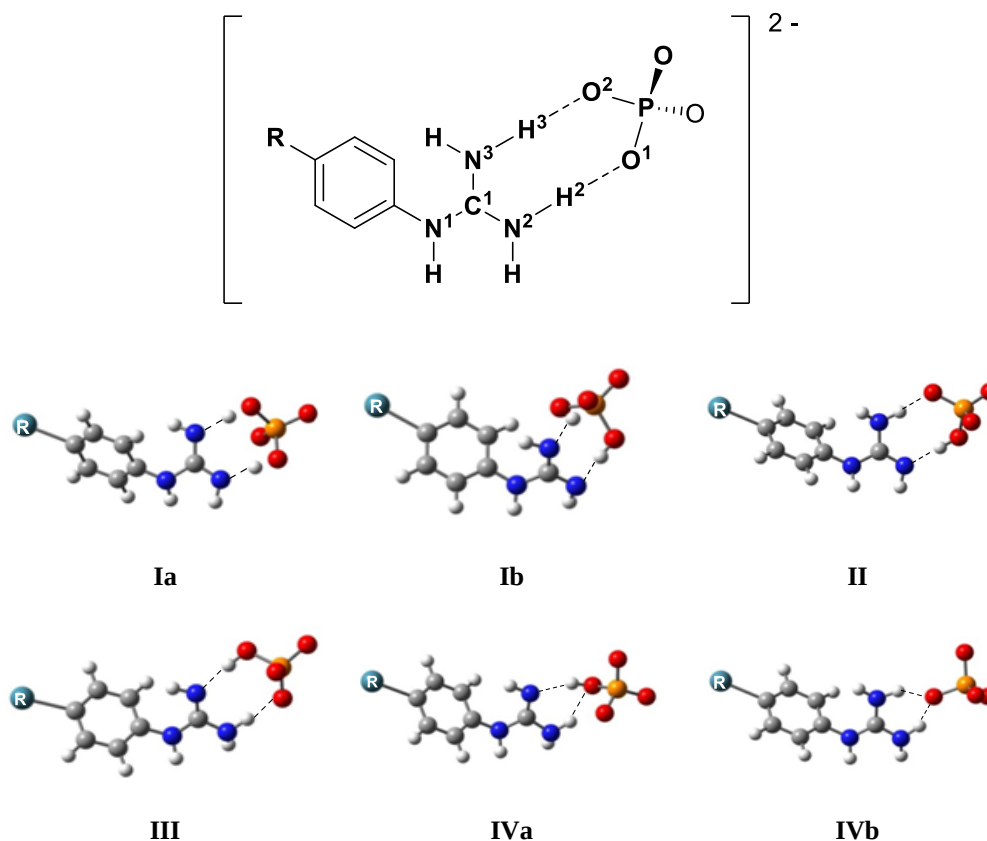


Figure A.1.1.1. Above: schematic description of the interaction between guanidinium cation and phosphate anion with labelled atoms. Below: optimized structures of the arylguanidinium-phosphate complexes at MP2/6-31+G(d,p) and PCM-water computational level for the different substituents, R= COH, NH₂, OH, OPh and CH₃. Complex **IVa** is obtained in all the cases except for R= COH for which complex **IVb** is found; complex **Ib** is only found for this COH substituent.

In complex **Ia**, both hydrogen atoms (H² and H³) are shared between the guanidinium and the oxygen atoms of the phosphate molecule. The extra O¹⋯H interaction observed in complex **Ib** only happens because the electro-withdrawing effect of the COH substituent deactivates the aromatic ring. Regarding complexes **II** and **III**, one of the H atoms of the guanidinium cation is attached to one of the nitrogen atoms (N² or N³), while the other one is transferred to an oxygen atom of the phosphate anion. Finally, in complexes **IVa,b** different interactions take place; thus, while in most of the complexes structure **IVa** is found (in which the hydrogen atom H³ is transferred from N³ to O² which also interact with H²), in the system with R= COH only complex **IVb** is obtained, with both hydrogens (H² and H³) kept in the guanidinium giving rise to a three centre interaction with O² (Figure A.1.1.1.).

Table A.1.1.1. Bond length (Å) calculated for all the derivatives at MP2/6-31+G(d,p) and PCM-water computational level.

Cmplx.	N ² -H ²					N ³ -H ³				
	COH	NH ₂	OH	OPh	CH ₃	COH	NH ₂	OH	OPh	CH ₃
Ia	1.599	1.583	1.586	1.582	1.574	1.618	1.600	1.603	1.611	1.605
Ib	1.709	-	-	-	-	1.606	-	-	-	-
II	1.847	1.817	1.819	1.843	1.826	1.041	1.038	1.038	1.040	1.039
III	1.038	1.037	1.037	1.037	1.038	1.849	1.832	1.844	1.838	1.834
IVa- IVb	1.058	1.017	1.017	1.017	1.018	1.050	1.904	1.926	1.900	1.936
Cmplx.	O ¹ -H ²					O ²⁽¹⁾ -H ³				
	COH	NH ₂	OH	OPh	CH ₃	COH	NH ₂	OH	OPh	CH ₃
Ia	1.034	1.040	1.038	1.039	1.037	1.029	1.034	1.033	1.031	1.034
Ib^a	1.012	-	-	-	-	1.031	-	-	-	-
II	0.990	0.993	0.992	0.990	0.992	1.720	1.747	1.744	1.735	1.742
III	1.752	1.768	1.768	1.760	1.740	1.030	0.991	0.990	0.990	1.017
IVa- IVb	1.622	2.102	2.062	2.122	2.056	1.655 ^b	0.985 ^b	0.984 ^b	0.985 ^b	0.983 ^b

^a Extra interaction found O³-H(Ph) 2.319 Å. ^b These distances are from the O¹

It is important to notice that for the complexes with R= NH₂, OH, OPh and CH₃, the intermolecular distances for the different complexes are very similar; indeed, the standard deviation of the average for those distances ranges between 0.001 and 0.017. In the case of the complexes with the COH substituent more divergence in these distances can be found; for instance, complexes **I**, **II** and **III** with the COH substituent present the largest N²⋯H distances.

The interaction energies of all studied complexes range between -115.0 and -29.2 kJ mol⁻¹ and are gathered in Table A.1.1.2. The most negative values, considering all the different phenylguanidiniums, are those obtained for complexes type **Ia** in which two H

atoms are shared between the guanidinium and phosphate groups, establishing two simultaneous interactions. Only an exception occurs with complexes with the COH substituent in which adduct **IVb** has the largest interaction energy. Curiously, for the rest of the complexes, dimer type **IVa** show the smallest interaction energy, maybe because the stabilising three centre interaction formed between H², H³ and a phosphate O atom is formed only in complex **IVb** (see Figure A.1.1.1.).

Table A.1.1.2. Interaction energies (E_i) for the different complexes calculated at MP2/6-311+G(d,p) and PCM-water computational level . All values are in kJ mol⁻¹.

	COH	NH ₂	OH	OPh	CH ₃
Ia	-90.6	-94.8	-94.6	-94.1	-98.9
Ib	-90.2	-	-	-	-
II	-60.1	-54.4	-54.6	-55.4	-56.4
III	-54.3	-53.1	-52.3	-53.4	-54.3
IVa-IVb	-115.0	-29.3	-29.2	-29.4	-38.3

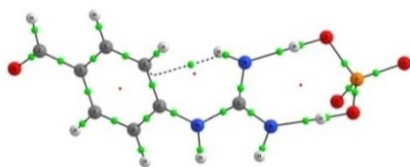
The fact that, in general, the most stable complexes found (type **Ia**) are those where two H atoms of the guanidinium and two O atoms of the phosphate interact, is in total agreement with our previous docking studies of different guanidinium containing compounds and a kinase protein in the presence of ATP. In this study it was shown that when ATP is present, a guanidinium-phosphate interaction between two guanidinium H atoms and two O atoms of ATP is formed.

A.1.1.2. Electron density and molecular orbital analyses

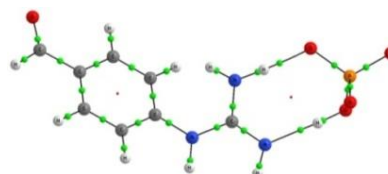
The electron density of the different phenylguanidinium-phosphate complexes has been analysed by means of the Atoms in Molecules (AIM) theory; accordingly molecular graphs, showing all the interactions established within the complexes, were generated and are shown in Figure A.1.2. In all the 20 complexes considered, HBs

have been found as indicated by a Bond Critical Point (BCP) between the interacting atoms (N, O and H) (Figure A.1.1.2.).

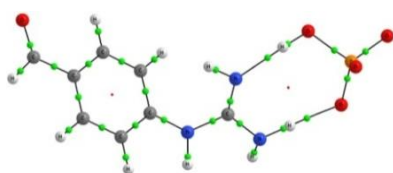
R= COH



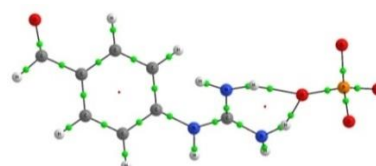
I



II

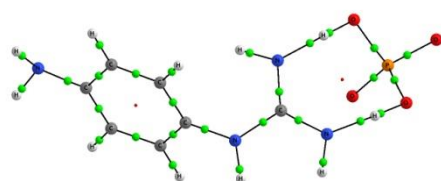


III

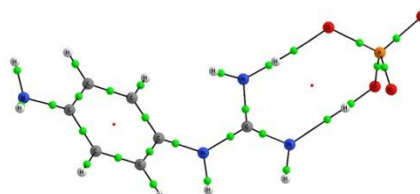


IVb

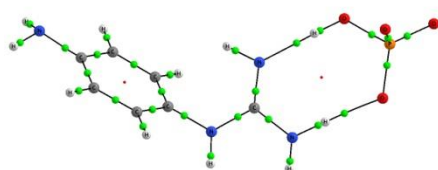
R= NH₂



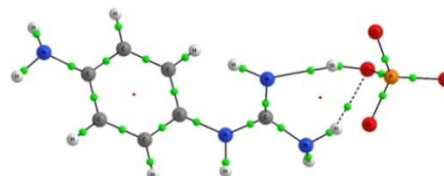
I



II

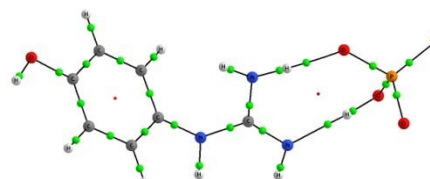
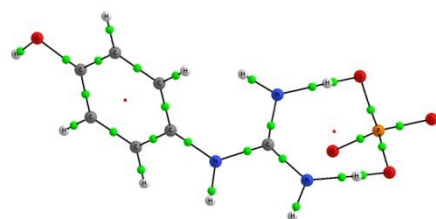


III



IVa

R= OH



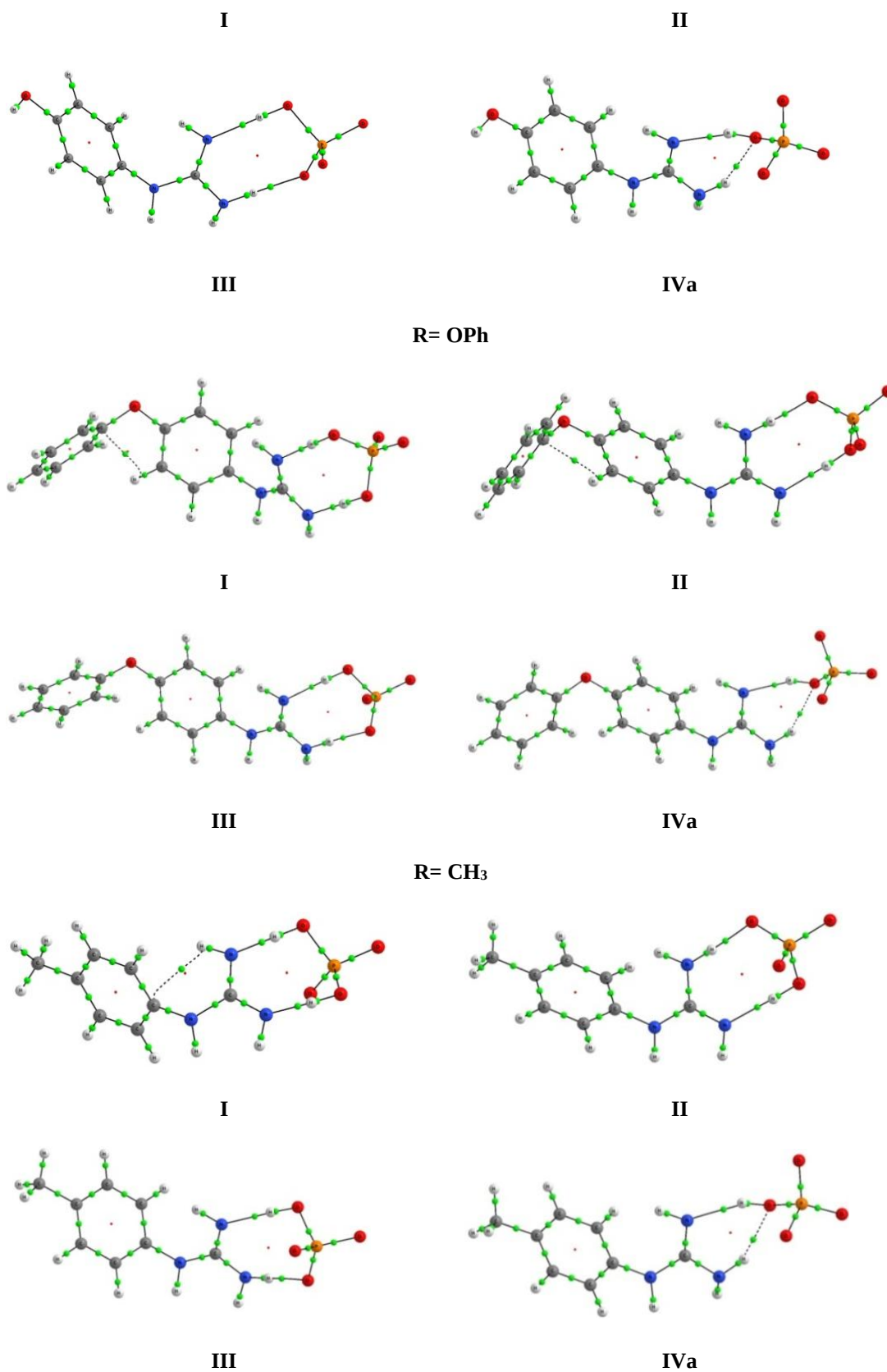


Figure A.1.1.2. Molecular graphs showing the interactions formed between guanidinium and phosphate. Green and red dots correspond to bond and ring critical points, respectively.

The electron density on the BCPs (ρ_{BCP}) ranges from 0.019 to 0.071 a.u. and the corresponding Laplacian ($\nabla^2\rho_{\text{BCP}}$) give always positive values from 0.056 to 0.157 a.u. (Table A.1.1.3.). Based on this data, the interactions established within all these complexes correspond to closed-shell interactions, and the ρ_{BCP} and $\nabla^2\rho_{\text{BCP}}$ values are in agreement with HBs of different strength.

Table A.1.1.3. Electron density at the bond critical points (ρ_{BCP}) and the Laplacian of that electron density ($\nabla^2\rho_{\text{BCP}}$) calculated at MP2/6-31+G(d,p) and PCM-water computational level for all the complexes studied.

$\text{N}^2\cdots\text{H}^2\text{O}^1$										
	COH		NH ₂		OH		OPh		CH ₃	
	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$
Ia	0.067	0.106	0.069	0.103	0.069	0.103	0.070	0.103	0.071	0.100
Ib^a	0.050	0.104	-	-	-	-	-	-	-	-
II	0.036	0.091	0.039	0.096	0.039	0.096	0.037	0.093	0.038	0.094
$\text{N}^3\cdots\text{H}^3\text{O}^2$										
	COH		NH ₂		OH		OPh		CH ₃	
	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$
Ia	0.063	0.109	0.066	0.106	0.066	0.107	0.065	0.107	0.066	0.107
Ib	0.065	0.108	-	-	-	-	-	-	-	-
III	0.036	0.091	0.037	0.094	0.036	0.092	0.037	0.093	0.037	0.093
IVa	-	-	0.032	0.084 ^b	0.031	0.081 ^b	0.032	0.085 ^b	0.030	0.079 ^b
$\text{O}^2\cdots\text{H}^3\text{N}^3$										
	COH		NH ₂		OH		OPh		CH ₃	
	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$
II	0.043	0.126	0.041	0.117	0.040	0.118	0.040	0.114	0.041	0.119
IVb	0.050	0.152 ^b	-	-	-	-	-	-	-	-
$\text{O}^1\cdots\text{H}^2\text{N}^2$										

	COH		NH ₂		OH		OPh		CH ₃	
	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$
III	0.040	0.116	0.039	0.112	0.039	0.111	0.039	0.114	0.041	0.120
IVa/b	0.055	0.157	0.020	0.058	0.021	0.062	0.019	0.056	0.022	0.063

^a Extra interaction found O3-H(Ph) ($\rho_{\text{BCP}} = 0.013$ and $\nabla^2\rho_{\text{BCP}} = 0.039$). ^b These distances are from the O¹

The largest value of ρ_{BCP} in any of the HB formed has been found in all complexes **Ia** (Table A.1.1.3.). This result is in agreement with the largest interaction energy computed for these complexes that was also found in complex **Ib**. The largest the electron density values, the largest the interaction energy, and consequently, complex **IVa** has the lowest value of the density at BCP for the HBs found in this set of complexes. Exponential relationships have been found between the electron density at the BCPs and the corresponding interatomic distances in all complexes studied (Figures A.1.1.3.); similar relationships have been previously described not only for HBs^{1,2}, but also for other weak interactions.²⁻⁵

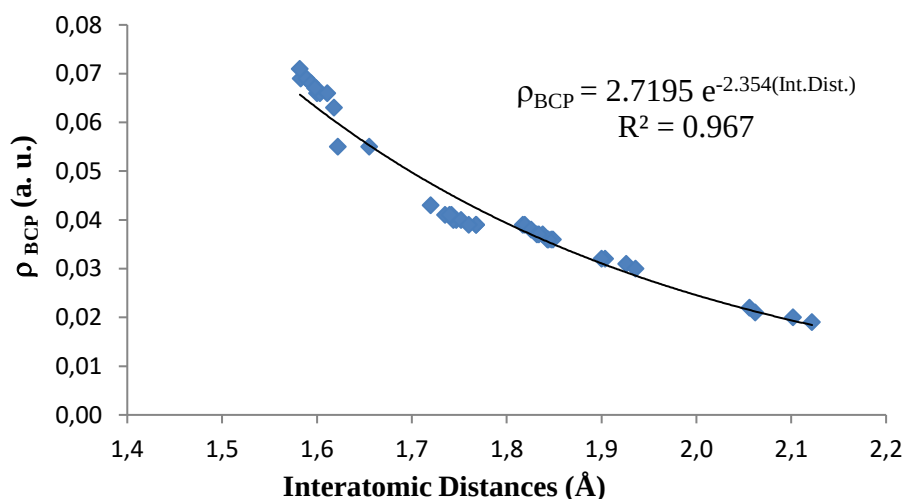


Figure A.1.1.3. Relationship between the interatomic distance (Å) and the value of the electron density at BCP (a.u.) for H $\cdots$ X (X=O, N) interactions in all complexes studied.

Additionally, electron density shift maps (EDS) have been obtained in order to provide with an insight of the changes on the electron density of the monomers upon complexation. Only complexes **Ia** and **III** have been studied and only for R= OPh and CH₃. All the EDS maps have been obtained following the procedure reported by Sánchez-Sanz *et al.*⁶ and the intermolecular HB donor and acceptor density shift pattern is shown in Figure A.1.1.4. for the dimers. On the one hand, a blue region is observed surrounding the H corresponding to the HB donor, which represents a decrease of the electron density on this area. On the other hand, there is an increase in the intermolecular density on the area located between that H atom and the N or O atom that acts as a HB acceptor. This increase on the density is shown as a positive yellow region.

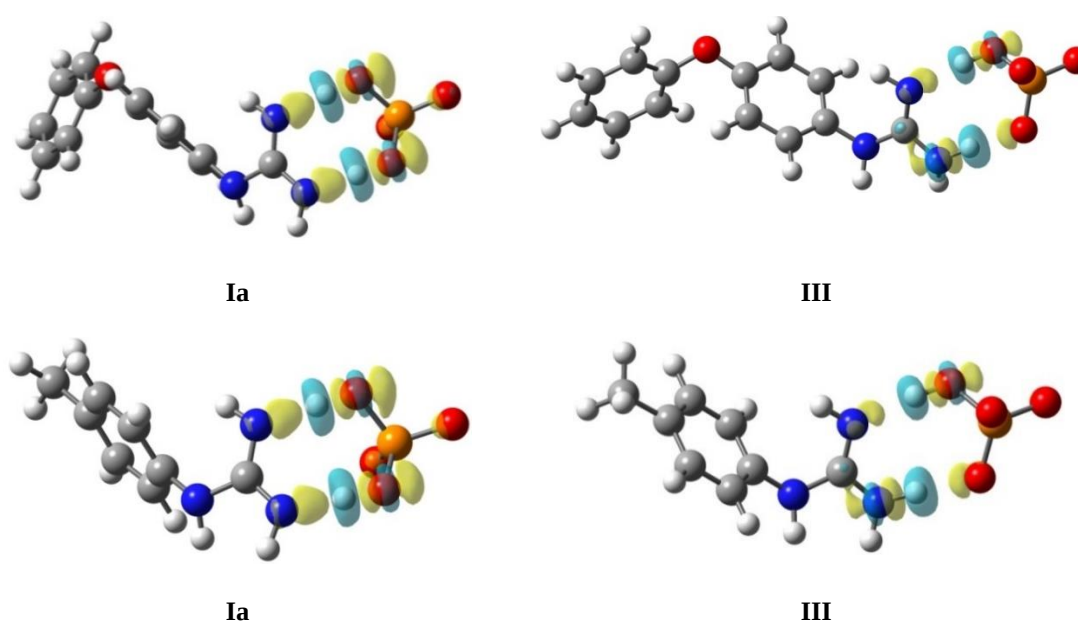


Figure A.1.1.4. Electron Density Shifts (EDS) at 0.0005 a.u. calculated for complexes **Ia** and **III** with arylguanidiniums R= OPh and CH₃ at MP2/6-31+g(d,p) and PCM-water computational level. Yellow and blue areas represent positive (increase) and negative (decrease) electron density regions respectively.

NBO analysis has also been used to identify intermolecular charge transfer between occupied molecular orbitals to empty ones upon complexation. The results of such analysis at B3LYP/6-31+G(d,p) have been summarized in Table A.1.1.4. All the

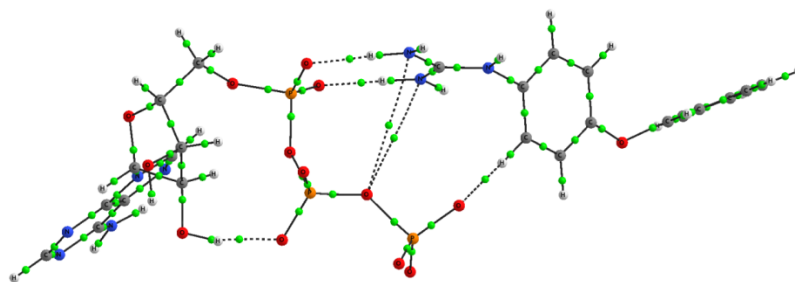
molecular orbital interactions observed are established between the lone pair of an O or an N atom (O^1 , O^2 , N^2 and N^3) and an empty N–H or O–H σ -bond (H^2-N^2 , H^3-N^3 , H^2-O^1 and H^3-O^2), in agreement with HB interactions (Table A.1.1.4.). The largest second order orbital energies contribution $E(2)$ is found in type **Ia** complexes, in agreement with the AIM results, since these compounds also present the largest electron density at their BCPs.

Table A.1.1.4. NBO $E(2)$ orbital interaction energies (kJ mol^{-1}) calculated at B3LYP/6-31+G(d,p) and PCM-water computational level for all the complexes studied.

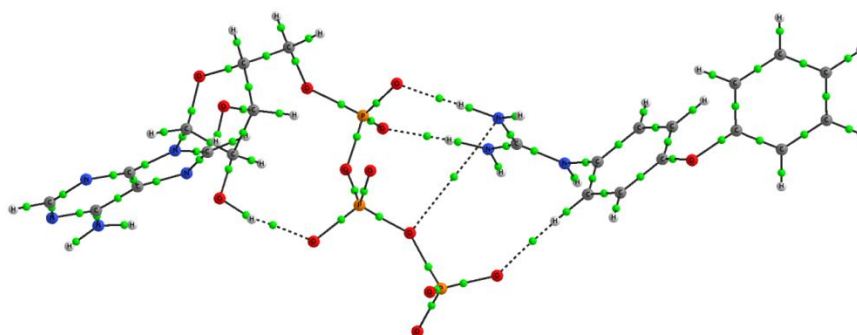
Complex	LP $N^2 \rightarrow \sigma^*H^2-O^1$					LP $N^3 \rightarrow \sigma^*H^3-O^2$				
	COH	NH ₂	OH	OPh	CH ₃	COH	NH ₂	OH	OPh	CH ₃
Ia	251.2	267.6	259.7	427.7	435.5	227.1	217.7	209.3	369.3	361.3
Ib^a	92.1	-	-	-	-	239.3	-	-	-	-
II	99.5	112.4	112.1	98.4	119.9	-	-	-	-	-
III	-	-	-	-	-	91.9	104.3	102.7	98.03	112.3
IVa/b	-	-	-	-	-	-	77.4 ^b	70.5 ^b	77.5 ^b	68.9 ^b
Complex	LP $O^1 \rightarrow \sigma^*H^2-N^2$					LP $O^2 \rightarrow \sigma^*H^3-N^3$				
	COH	NH ₂	OH	OPh	CH ₃	COH	NH ₂	OH	OPh	CH ₃
II	-	-	-	-	-	156.1	143.7	145.8	162.0	152.8
III	142.1	140.7	132.7	145.1	146.5	-	-	-	-	-
IVa/b	177.4	27.4	35.0	31.9	36.9	121.9 ^b	-	-	-	-

^a Extra interaction found $O^3-H(\text{Ph})$ LP $O^3 \rightarrow \sigma^*HC$ 21.6 kJ mol^{-1} . ^b These values are from the O^1

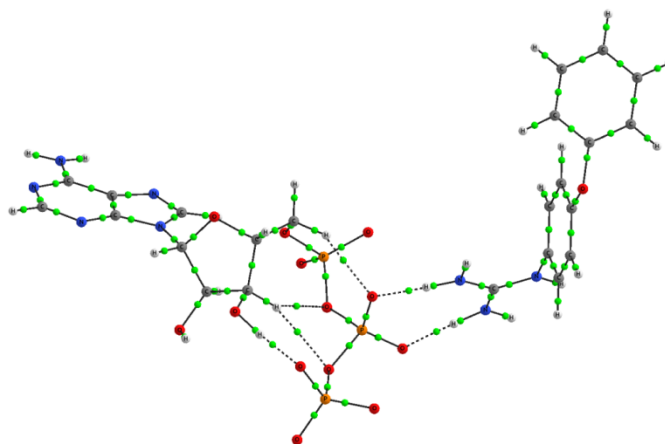
A.1.2. Arylguanidinium – ATP Complexes



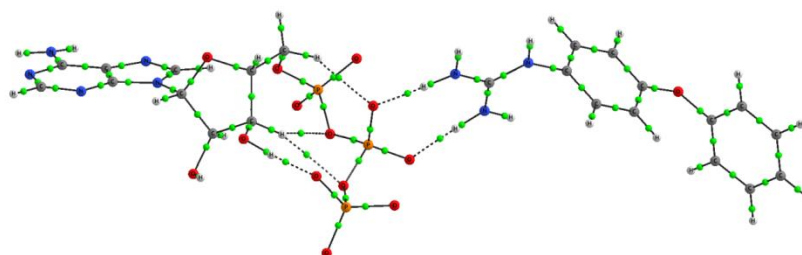
V α -R[a]



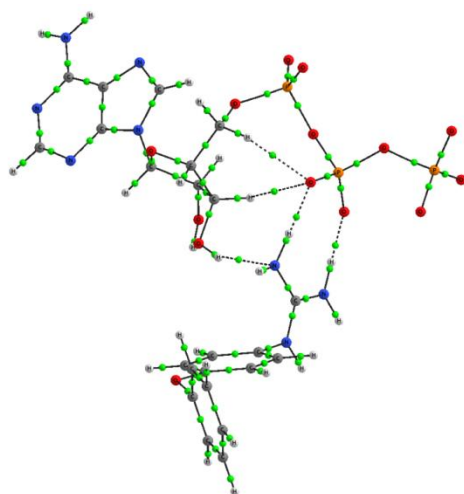
V α -R[b]



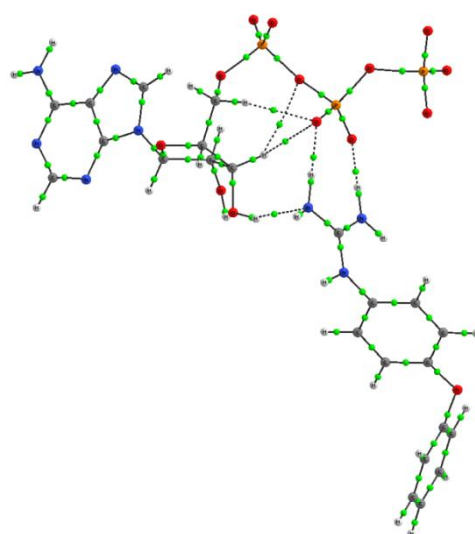
V β -R[a]



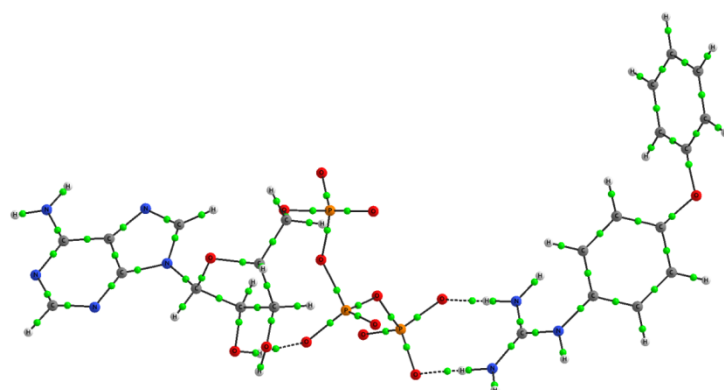
V β -R[b]



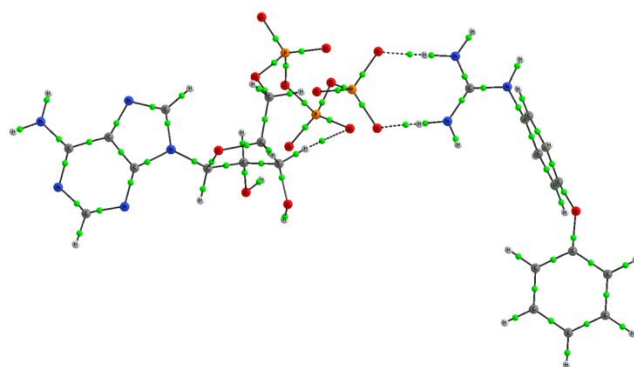
Vβ-L[a]



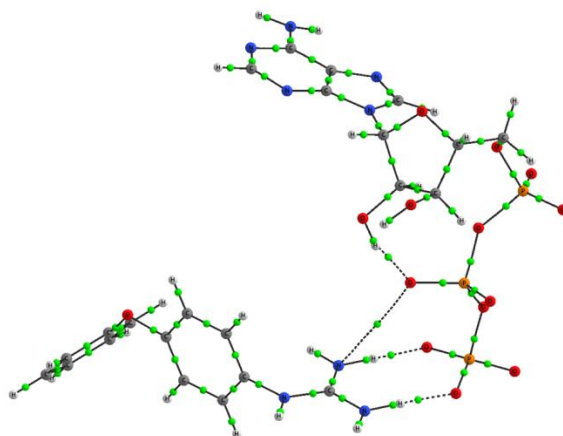
Vβ-L[b]



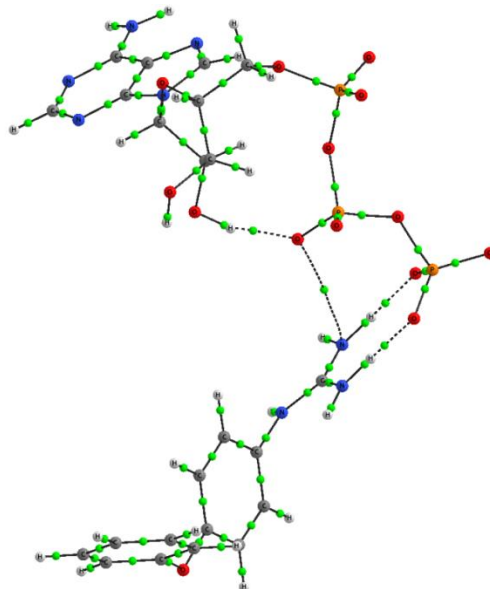
Vγ-R[a]



V γ -R[b]



V γ -L[a]



V γ -L[b]

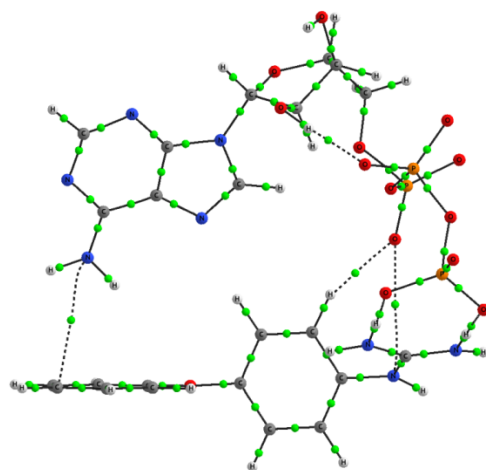
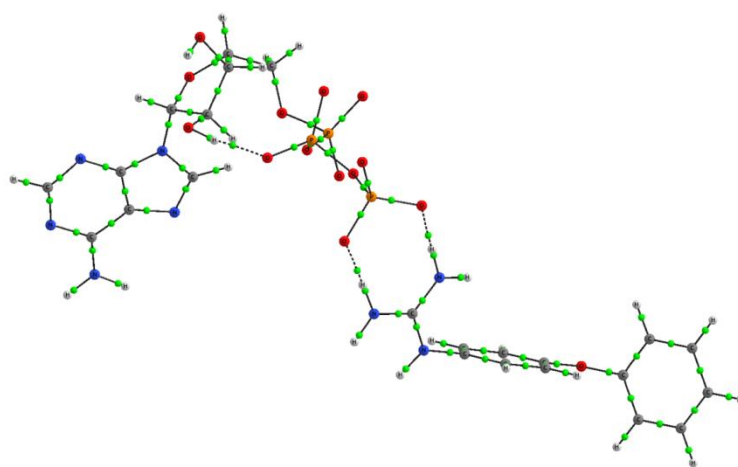
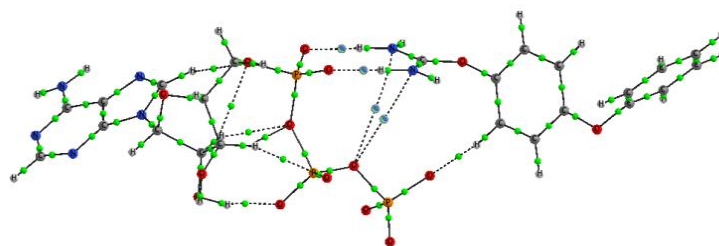
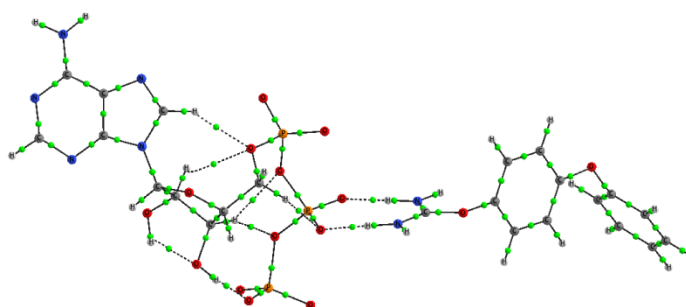
**V γ -B[a]****V γ -B[b]**

Figure A.1.2.1. Molecular graphs corresponding to complexes **V** between arylguanidinium and ATP molecule. Green dots correspond to bond and ring critical points, respectively.

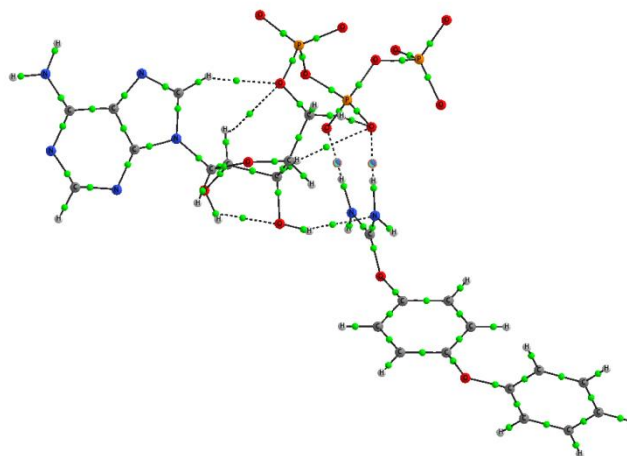
A.1.3. Arylisouronium – ATP Complexes



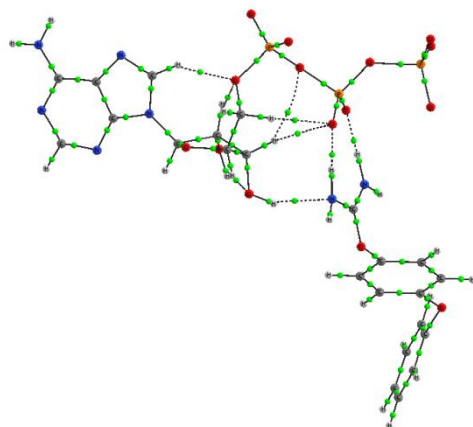
VII α -R[b]



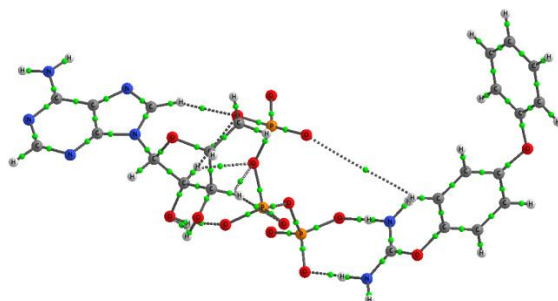
VII β -R[b]



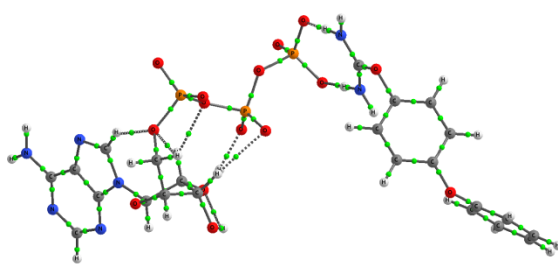
VII β -L[a]



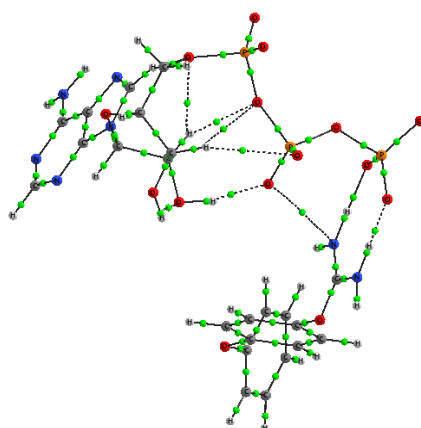
VIIβ-L[b]



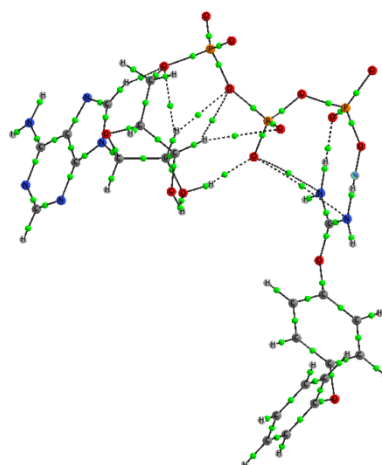
VIIγ-R[a]



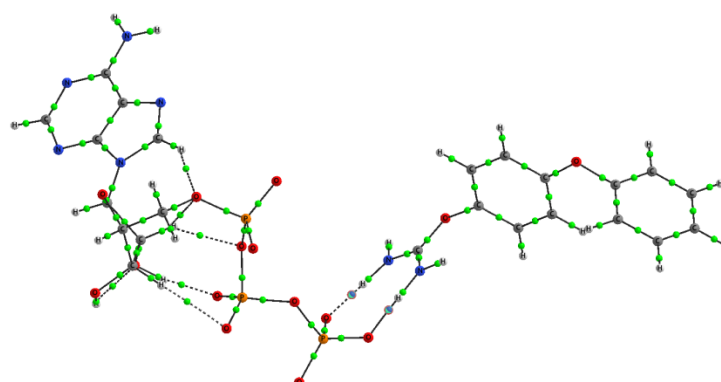
VIIγ-R[b]



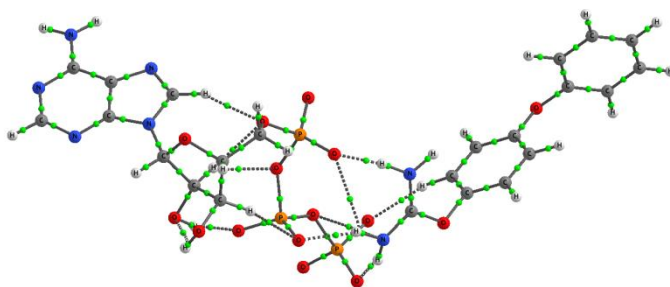
VII γ -L[a]



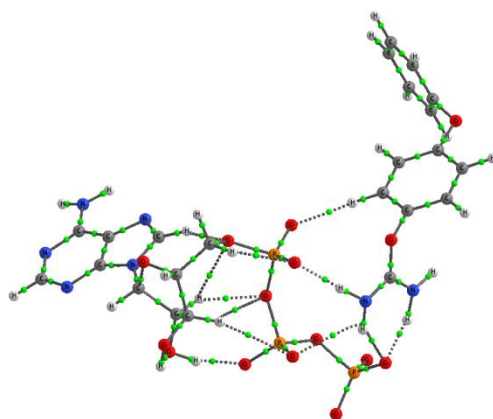
VII γ -L[b]



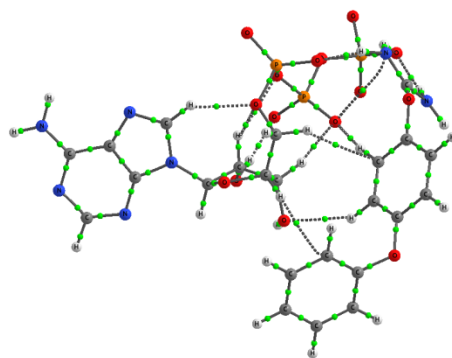
VII γ -B[b]



VIII γ -R[a]



VIII γ -R[b]



VIII γ -L[a]

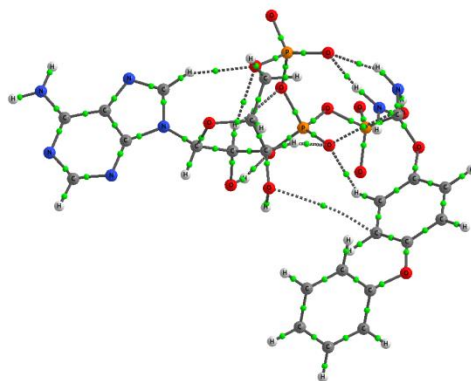
VIII γ -L[b]

Figure A.1.3.1. Molecular graphs corresponding to complexes **VII** and **VIII** between arylisouronium and the phosphates of the ATP molecule. Green dots correspond to bond critical points.

Table A.1.3.1. Electron density at the bond critical points (ρ_{BCP} , a.u.), the Laplacian of that electron density ($\nabla^2\rho_{\text{BCP}}$, a.u.) and interatomic distances ($\text{\AA}$) calculated at $\omega\text{B97XD/6-31+G(d,p)}$ and PCM-water computational level for all complexes **VII**.

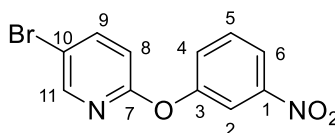
	O1 $\cdots$ H2N2			O2 $\cdots$ H3N3		
Complex	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$	$\text{\AA}$	ρ_{BCP}	$\nabla^2\rho_{\text{BCP}}$	$\text{\AA}$
VII α -R[a]	0.046	0.133	1.688	0.041	0.121	1.730
VII α -R[b]	0.041	0.120	1.733	0.044	0.130	1.701
VII β -R[a]	0.048	0.141	1.661	0.054	0.155	1.612
VII β -R[b]	0.052	0.148	1.632	0.050	0.147	1.644
VII β -L[a]	0.055	0.154	1.612	0.054	0.152	1.615
VII β -L[b]	0.059	0.159	1.582	0.052	0.149	1.631
VII γ -R[a]	0.070	0.160	1.527	0.057	0.151	1.604
VII γ -R[b]	0.067	0.160	1.542	0.063	0.158	1.569
VII γ -L[a]	0.068	0.162	1.532	0.060	0.157	1.583
VII γ -L[b]	0.062	0.160	1.567	0.067	0.163	1.537
VII γ -B[a]	0.070	0.164	1.525	0.057	0.151	1.608
VII γ -B[b]	0.063	0.160	1.562	0.067	0.160	1.542

Table A.1.3.2. Electron density at the bond critical points (ρ_{BCP} , a.u.), the Laplacian of that electron density ($\nabla^2\rho_{\text{BCP}}$, a.u.) calculated at $\omega\text{B97XD}/6\text{-}31\text{+G(d,p)}$ and PCM-water computational level for all complexes **VIII**.

	$\text{P}\alpha\text{O}\cdots\text{HN}$	$\text{P}\alpha\text{O}\cdots\text{HN}$	$\text{P}\beta\text{O}\cdots\text{N}$	$\text{P}\beta\text{O}\cdots\text{N}$	$\text{P}\gamma\text{O}\cdots\text{HN}$	$\text{P}\gamma\text{O}\cdots\text{HN}$	$\text{PO}\cdots\text{HC}$
Complex	ρ_{BCP} $\nabla^2\rho_{\text{BCP}}$	ρ_{BCP} $\nabla^2\rho_{\text{BCP}}$	ρ_{BCP} $\nabla^2\rho_{\text{BCP}}$	ρ_{BCP} $\nabla^2\rho_{\text{BCP}}$	ρ_{BCP} $\nabla^2\rho_{\text{BCP}}$	ρ_{BCP} $\nabla^2\rho_{\text{BCP}}$	ρ_{BCP} $\nabla^2\rho_{\text{BCP}}$
VIIIγ-L[b]	0.0359	0.0227	-	-	0.0597	-	0.0174
	0.1138	0.0669	-	-	0.1615	-	0.0573
VIIIγ-R[a]	0.0469	0.0057	0.0302	0.0078	0.0542	-	0.0167
	0.1446	0.0241	0.0886	0.0267	0.1488	-	0.0473
VIIIγ-L[a]	0.0400	-	-	-	0.0367	0.1054	0.0239
	0.1267	-	-	-	0.1054	0.1097	0.0764
VIIIγ-R[b]	0.0378	-	0.0094	-	0.0244	0.0569	0.0154
	0.1133	-	0.0353	-	0.0680	0.1553	0.0454

A.2. Chemical Synthesis

5-bromo-2-(3-nitrophenoxy)pyridine (**75**)



Following procedure for the synthesis of **74**, side product **75**, was obtained as an orange solid side product, 23%. **Mp**: 68 – 70 °C

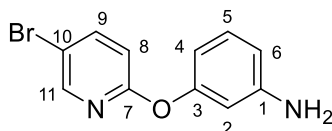
δ_{H} (400 MHz, CDCl_3): 6.95 (d, $J = 8.7$ Hz, 1H, H-8), 7.48 (dd, $J = 8.2, 2.2$ Hz, 1H, H-4), 7.57 (t, $J = 8.2$ Hz, 1H, H-5), 7.85 (dd, $J = 8.7, 2.5$ Hz, 1H, H-9), 8.02 (t, $J = 8.2, 2.1$ Hz, 1H, H-2), 8.08 (dd, $J = 8.2, 2.1$ Hz, 1H, H-6), 8.21 (d, $J = 2.4$ Hz, 1H, H-11)

δ_{C} (100 MHz, CDCl_3): 113.8 (CH Ar, C-8), 114.8 (qC, C-10), 116.8 (CH Ar, C-2), 119.9 (CH Ar, C-6), 127.6 (CH Ar, C-4), 130.3 (CH Ar, C-5), 142.6 (CH Ar, C-9), 148.4 (CH Ar, C-11), 149.2 (qC), 154.3 (qC), 161.5 (qC).

HRMS (m/z APCI $^+$): 294.9716 ($\text{M}^+ + \text{H}$), $\text{C}_{11}\text{H}_8\text{N}_2\text{OBr}$ Requires: 294.9713.

ν_{max} (ATR)/ cm^{-1} : 3105, 2920, 2850, 1580, 1525 (NO_2), 1468, 1341 (NO_2), 1281, 1264 (C-O), 1243, 1090 (C-Br), 833, 800, 742, 680.

3-((5-bromopyridin-2-yl)oxy)aniline (**87**)



Following procedure for the synthesis of **86**, side product **87**, was obtained as an orange brown oil side product, 21%.

δ_{H} (600 MHz, CDCl_3): 3.74 (bs, 2H, NH_2), 6.44 (t, $J = 2.2$ Hz, 1H, H-2), 6.49 (dd, $J = 8.1, 1.6$ Hz, 1H, H-4 or H-6), 6.52 (dd, $J = 8.1, 2.0$ Hz, 1H, H-4 or H-6), 6.80 (d, $J = 8.7, 1\text{H}$, H-8), 7.16 (t, $J = 8.0, 1\text{H}$, H-5), 7.74 (dd, $J = 8.7, 2.5$ Hz, 1H, H-9), 8.24 (d, $J = 2.5$ Hz, 1H, H-11).

δ_c (150 MHz, $CDCl_3$): 107.8 (CH Ar, C-2), 111.0 (CH Ar, C-4 or C-6), 112.1 (CH Ar, C-4 or C-6), 113.1 (CH Ar, C-8), 113.5 (qC), 130.5 (CH Ar, C-5), 142.0 (CH Ar, C-9), 148.2 (qC), 148.7 (CH Ar, C-11), 155.1 (qC), 162.8 (qC).

HRMS (m/z ESI^+): Found: 264.9979 ($M^+ + H$), $C_{11}H_{10}N_2OBr$ Requires: 264.9976.

ν_{max} (ATR)/ cm^{-1} : 3450 (NH), 3349 (NH), 3220, 3052, 1606, 1572, 1450 (CN), 1364, 1238-1144 (C-O), 1144, 1089, 998, 959, 823 (C-Br), 754.

A.3. Pharmacokinetic Properties

Table A.3.1. Calculated physicochemical properties for all synthesised final hydrochloride salts.

Code	MW (g/mol)	pKaH*	#Heavy atoms	Fraction Csp3	#Rotatable bonds	#H-bond acceptors	#H- bond donors
Sorafenib	464.82		32	0.1	9	7	3
1	464.87	8.71/ 7.02	32	0.05	9	4	6
2	380.42	8.72/7.40	28	0	8	2	6
3	398.41	8.70/6.93	29	0	8	3	6
4	506.32	8.56/6.29	29	0	8	2	6
5	441.32	8.72/7.36	28	0	8	1	6
6	509.32	8.71/7.10	32	0.05	9	4	6
7	506.32	6.47/8.40	29	0	8	2	6
8	509.32	7.27/8.55	32	0.05	9	4	6
27	353.83	8.86	25	0	6	1	4
28	421.82	8.86	29	0.05	7	4	4
29	387.38	9	28	0.05	7	4	4
30	445.43	8.17	30	0	7	6	4
31	439.81	8.62	30	0.05	7	5	4
32	426.29	8.71	27	0	7	2	4
33	449.83	8.71	31	0.05	8	5	4
61	399.4	7.56/6.79	29	0	8	4	6
62	507.3	7.35/6.22	29	0	8	3	6
63	442.31	7.71/7.09	28	0	8	2	6
64	510.31	7.61/6.92	32	0.05	9	5	6
65	465.86	7.59/6.86	32	0.05	9	5	6
66	422.81	7.62	29	0.05	7	5	4
67	465.86	8.34 / 7.03	32	0.05	9	5	5
68	422.81	8.32	29	0.05	7	5	3
69	450.82	8.32	31	0.05	8	6	3
70	457.85		30	0.05	7	7	3

Properties calculated with SwissADME, except for *, calculated using ChemAxon's Marvin (references in the text).

Table A.3.1. *following*

Code	TPSA (Å ²)	XLOGP3	Consensus Log P (average of all methods)	Ali Log S	Ali Sol. (mg/mL)	Ali Sol. (μmol/L)	Ali Class
Sorafenib	92.35	4.07	4.11	-5.71	8.98E-04	1.93	Mod. sol.
1	122.52	4.31	3.31	-6.6	1.18E-04	2.53	Poorly sol.
2	122.52	2.9	2.13	-5.13	2.80E-03	7.35	Mod. sol.
3	122.52	3	2.37	-5.24	2.31E-03	5.79	Mod. sol.
4	122.52	3.55	2.77	-5.81	7.88E-04	1.56	Mod. sol.
5	122.52	3.49	2.43	-5.75	7.93E-04	1.80	Mod. sol.
6	122.52	4.37	3.38	-6.66	1.12E-04	2.19	Poorly sol.
7	122.52	3.55	2.83	-5.81	7.88E-04	1.56	Mod. sol.
8	122.52	4.37	3.39	-6.66	1.12E-04	2.19	Poorly sol.
27	84.9	4.31	3.35	-5.81	5.52E-04	1.56	Mod. sol.
28	84.9	5.2	4.44	-6.73	7.85E-05	1.86	Poorly sol.
29	84.9	4.57	3.84	-6.08	3.25E-04	8.39	Poorly sol.
30	110.2	6.96	4.14	-9.09	3.64E-07	8.17	Poorly sol.
31	84.9	5.3	4.56	-6.83	6.45E-05	1.47	Poorly sol.
32	101.97	3.67	3.05	-5.5	1.35E-03	3.16	Mod. sol.
33	101.97	4.49	3.9	-6.35	2.00E-04	4.45	Poorly sol.
61	135.41	2.26	1.87	-4.74	7.27E-03	1.82	Mod. sol.
62	135.41	2.81	2.24	-5.31	2.48E-03	4.89	Mod. sol.
63	135.41	2.75	1.89	-5.25	2.50E-03	5.64	Mod. sol.
64	135.41	3.63	2.8	-6.16	3.52E-04	6.89	Poorly sol.
65	135.41	3.57	2.72	-6.1	3.71E-04	7.95	Poorly sol.
66	97.79	4.46	3.71	-6.23	2.47E-04	5.85	Poorly sol.
67	119.72	4.88	3.69	-7.13	3.46E-05	7.43	Poorly sol.
68	82.1	5.77	4.65	-7.26	2.31E-05	5.46	Poorly sol.
69	99.17	5.07	4.23	-6.89	5.74E-05	1.27	Poorly sol.
70	101.83	4.8	4.43	-6.67	9.78E-05	2.14	Poorly sol.

Properties calculated with SwissADME, except for *, calculated using ChemAxon's Marvin (references in the text).

Table A.3.1. *following*

Code	HIA	BBB perm.	Pgp sub.	CYP1A2 inh.	CYP2C19 inh.	CYP2C9 inh.	CYP2D6 inh.	CYP3A4 inh.
Sorafenib	Low	No	No	Yes	Yes	Yes	Yes	Yes
1	High	No	No	No	No	No	No	No
2	High	No	No	No	No	No	No	No
3	High	No	No	No	No	No	No	No
4	High	No	No	No	No	No	No	No
5	High	No	No	No	No	No	No	No
6	High	No	No	No	No	No	No	No
7	High	No	No	No	No	No	No	No
8	High	No	No	No	No	No	No	No
27	High	No	No	Yes	Yes	Yes	Yes	Yes
28	High	No	No	No	Yes	Yes	Yes	No
29	High	No	No	No	No	Yes	Yes	No
30	Low	No	No	No	Yes	Yes	No	No
31	High	No	No	No	Yes	Yes	Yes	No
32	High	No	No	Yes	No	Yes	Yes	Yes
33	High	No	No	No	Yes	Yes	No	Yes
61	High	No	Yes	Yes	No	No	No	No
62	High	No	Yes	Yes	No	No	No	Yes
63	High	No	Yes	Yes	No	No	No	Yes
64	High	No	No	Yes	No	No	No	Yes
65	High	No	No	Yes	No	No	No	Yes
66	High	No	No	Yes	No	No	Yes	Yes
67	High	No	No	No	No	No	No	No
68	High	No	No	No	Yes	Yes	Yes	No
69	High	No	No	No	Yes	Yes	No	No
70	Low	No	No	No	Yes	Yes	No	No

Properties calculated with SwissADME, except for *, calculated using ChemAxon's Marvin (references in the text).

A.4. References

1. Sánchez-Sanz, G.; Alkorta, I.; Elguero, J., *Mol. Phys.* **2011**, *109*, 2543.
2. Sanchez-Sanz, G.; Trujillo, C.; Alkorta, I.; Elguero, J., *ChemPhysChem* **2012**, *13*, 496.
3. Sanchez-Sanz, G.; Trujillo, C.; Alkorta, I.; Elguero, J., *Phys. Chem. Chem. Phys.* **2014**, *16*, 15900.
4. Bauzá, A.; Alkorta, I.; Frontera, A.; Elguero, J., *J. Chem. Theory Comput.* **2013**, *9*, 5201.
5. Alkorta, I.; Elguero, J.; Solimannejad, M., *J. Phys. Chem. A* **2014**, *118*, 947.
6. Sánchez-Sanz, G.; Trujillo, C.; Alkorta, I.; Elguero, J., *Comput. Theor. Chem.* **2012**, *991*, 124.

A.5. Publications

- Trujillo, C.; Previtali, V.; Rozas, I., *Theor. Chem. Acc.* **2016**, 135.
- Previtali, V.; Trujillo, C.; Boisson, J.; Khartabil, H.; Henon, E.; Rozas, I., *Phys. Chem. Chem. Phys.* **2017**, 19, 31177.

