UNIVERSITÀ DEGLI STUDI DI PALERMO

Dottorato di Ricerca in Scienze Molecolari e Biomolecolari (Industriale)-XXXIII CICLO
Dipartimento Scienze e Tecnologie Biologiche Chimiche e Farmaceutiche (STEBICEF) CHIM/08

RATIONAL DESIGN AND SYNTHESIS OF NOVEL INHIBITORS OF EPIGENETIC TARGET INVOLVED IN CANCER

IL DOTTORE
Dott.ssa JESSICA LOMBINO

IL COORDINATORE
Chiar.ma Prof.ssa PATRIZIA DIANA

IL TUTOR
Chiar.ma Prof.ssa PATRIZIA DIANA

IL TUTOR AZIENDALE
Dr. ALESSANDRO PADOVA

CICLO XXXIII
ANNO CONSEGUIMENTO TITOLO 2021

To Alessandro and Ugo for conveying to me their passion for research,
for teaching me to follow my instinct,
for believing in me,
for everything you taught me.
This work would not be possible without your guidance.

CONTENTS

CHAPTER 1

An Overview of The Molecular Basis of Epigenetics

1.1 Epigenetics: The Science of Change

Epigenetics is the study of heritable phenotype changes that are not linked to alterations in the DNA sequence.[1] Epigenetic changes play a key role in cellular plasticity that is responsible for cellular reprogramming and response to the environment.[2,3]

In eukaryotic cells, genomic material is highly organized: ~ 146 bp of DNA is wrapped around nucleosomes, which comprise the basic building blocks of chromatin and is tightly regulated to control gene transcription. The nucleosome is formed by a core containing histone octamer and each histone octamer contains two of each of the core histone proteins (H2A, H2B, H3, H4). Furthermore, in order to maintain higher order chromatin structure, various lengths of linker DNA separate each nucleosome by interactions with histone H1. [4,5] The histone protruding N- terminal ends are characterized by the presence of many well-conserved residues that are targets for a variety of post-translational modifications including acetylation, methylation, ubiquitination, phosphorylation. (Figure **1**) Altogether, changes in the epigenome include methylation of cytosines bases in DNA, and post-translational modification (PTMs) of the N-terminal tails of core histones.[6,7] These modifications have direct effect on the degree of chromatin compaction, affecting indirectly the accessibility of the underlying DNA template for the transcription machinery. Indeed, while DNA sequence encodes for genetic information, events as transcription recombination, DNA replication and DNA repair are regulated by "epigenome".[8] The chromatin basically exists in two interchangeable forms: heterochromatin, in the closed state, and euchromatin, in the opened state. The balance between these two states is regulated by epigenetic marks, which switch genes transcription "on" and "off" in response to extracellular signals. The heterochromatin form suppresses the activity of transcriptional factors, such as RNA polymerase etc. and this is mainly associated with gene silencing. On the other hand, when the chromatin is in the opened state, is accessible to transcriptional machinery and is associated with active transcription.[9] Although all cells in an organism contain the same genetic material, epigenetic regulators are responsible for different packaging of DNA and chromatin. Indeed these enzymes, organizing the genome into accessible and closed regions, regulate the re-modelling of the chromatin configuration, the accessibility to transcription factors and ensure the correct transcriptional program in a given cell type. Thus, epigenetic regulation is crucial for maintaining the unique physical characteristics and biological

functions of specific tissues and organs and is implicated in fundamental processes such as proliferation, development, differentiation and genome integrity.[10,11]

In the last decade, a considerable amount of studies provide evidence that aberrant patterns of gene expression are not only from the result of genetic mutations, but also determined by epigenetic changes. These alterations are known to have an important role in a number of diseases, such as cancer, but also inflammation, metabolic diseases and neurodegenerative disorders. In particular, epigenetic dysregulation are widespread in several human cancers.[11–14]

In contrast to genetic alterations, which are almost impossible to reverse, epigenetic modifications are potentially reversible, suggesting that they are of therapeutic relevance in the prevention and treatment of cancers.
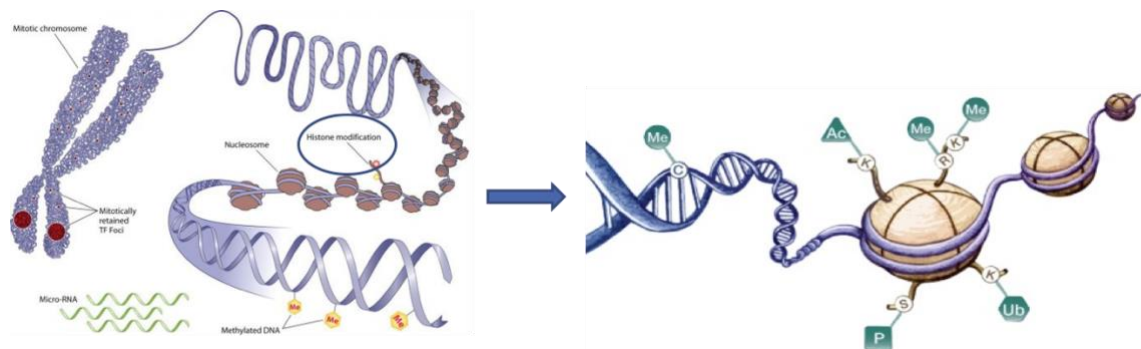
Figure 1. Chromatin structure and post-translational modifications.

1.2 Epigenetic modifications and the Histone code hypothesis

Epigenetic modification are broadly stable marks, resulting from covalent modifications to DNA and the histone proteins. Levels of histone modifications are dynamically regulated by the activities of histone-modifying enzymes that add or remove specific marks. Aberrant histone modification levels can result from an increasing or decreasing of these modifying enzymes, thus correcting the levels of a dysregulated enzyme could restore the physiological balance in the affected cells.[15] Protein responsible to modulate epigenetic marks can be divided into three categories based on their relative function including writers, that place the modification, erasers that remove the modification and readers that recognize the marks and then regulate gene expression, resulting in active or silent genes (Figure **2**).[16] The combination of how readers, writers and erasers modify and read the specific amino acids on the histones forms the basis of the histone code

hypothesis, which assumes that a single histone modification does not function alone but multiple histone modifications, on single or multiple tails, acting in a combinatorial or sequential manner, results in unique biological outcomes.[17,18]
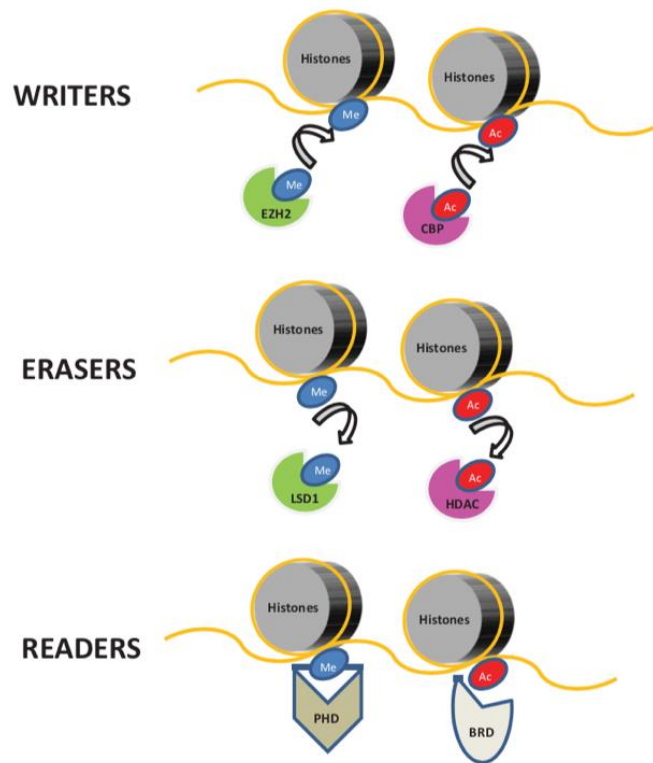
Figure 2. Classification of epigenetic modulators. Reproduced from "Epigenetic regulators as promising therapeutic targets in acute myeloid leukemia"[19]

At molecular level, epigenetic alterations result in a series of events that determine re-modelling of chromatin configuration and affect gene expression. Methylation of DNA is catalysed by DNA methyltransferases (DNMTs) and, using *S*-adenosyl methionine (SAM) as methyl donor, occurs at 5' position of cytosine residues of cytidine-phosphate-guanosine (CpG) dinucleotides. These regions are enriched in sites known as CpG islands, many of which are coincident with the promoter of protein encoding genes. Methylation of these promoter regions results in transcriptional repression.[20–22] Dynamic distribution of cytosine methylation plays a pivotal role in genomic imprinting, X-inactivation, genome stability and tissue-specific gene expression, especially during embryonic development.[23]

To date, several families of epigenetic enzymes in the human genome that catalyzed specific histone-post translational modifications, which are commonly referred to as histone marks, have been reported. Among these, there are protein methyl transferases which catalyze both lysine and arginine methylation, protein demethylases which remove the methyl groups from lysine and arginine residues, histone acetyltransferases (HATs) which add the acetyl group to lysine residues and that can be removed by the histone deacetylases (HDACs). In addition there are specific kinases which phosphorylate histone tails and ubiquitin ligases which add ubiquitin to specific lysine residues.[24] Histone de-acetylation of lysine residues is generally associated with closed chromatin state and gene silencing. Indeed, deacetylated lysine residues carry a positive charge that makes an ionic interaction with the negative charges of DNA, leading to condense of the nucleosomes and then, of the chromatin structure. Instead, acetylation of lysine residues results in a neutral charge, generating the open chromatin structure.[25,26]

Conversely, the interpretation of histone methylation is more complex. Different methylation states may occur on histone lysine residues: unmethylated, or mono-, di- or trimethylated (-me1, -me2,-me3). The histone methylation can result in different chromatin and transcription states depending on the extent of methylation and the particular residues in which occurs. Indeed, mono-methylation of H3K9, H3K27, and H3K79 histone proteins is associated with active transcription, whereas tri-methylation of these histones leads to transcription silencing.[7] In addition, di- and trimethylation of lysine 9 on histone H3 (H3K9me2/me3), trimethylation of lysine 27 on histone H3 (H3K27me3) and trimethylation of lysine 20 on histone H4 (H4K20me3) are associated to transcription inactivation whereas di- and trimethylation of lysine 4 (H3K4me2/ me3) and trimethylation of lysine 36 and 79 (H3K36me3, H3K79me3) result in active transcription.[27–29]

Finally, the third class of epigenetic proteins is formed by different nuclear factors containing specialized structural domains, such as chromodomains or bromodomains, that recognize and interact with the histone marks introduced by histone modification enzymes. The bromodomain and extra-terminal (BET) protein family is formed by bromodomain-containing protein BRD2, BRD3, BRD4 and bind acetylated lysine residues of histone, leading to transcriptional activation.[30–32] These proteins act as readers

of the epigenetic code and control the impact of histone modification on cellular phenotype, mediating its translation into a given functional effect.[33]

Hence, considering the pivotal role of epigenetic modifications in genome reprogramming, it is evident that their dynamic regulation is fundamental to maintain homeostasis in several biological process.

1.3 Epigenetic changes in cancer

As mentioned above, epigenetic modifications are characteristic of different cell types and play a key role for defining the cellular transcriptome at several levels. Aberrant changes in the pattern of epigenetic modifications may result in altered transcriptome, which may transform the identity of the cell. In the last decade, several researches linked mutations in gene encoding proteins that regulate epigenetic modification to human cancers. These mutations have been found to drive tumour initiation, immune evasion, metastasis and drug resistance and involve most of the major classes of epigenetic proteins including DNA methylation enzymes, histone modification enzymes and histone modification readers enzymes. In human cancer, apart from mutations, gene amplifications, translocation and indirect upregulation or inactivation of enzyme frequently occur. The recognition of abnormal DNA methylation was the first epigenetic alteration identified in cancer cells. DNA methylation is a dynamic process regulated by the balance between DNA methyltransferases DNMT1, DNMT3A and DNMT3B, which mediate active methylation and DNA demethylases, such as ten eleven translocation enzymes, TET1, TET2 and TET3, that catalyze demethylation reaction.[34–36] In tumor cells, aberrant DNA methylation events are associated with hypomethylation of the wide-genome, as well as hypermethylation of the promoter region.[37–39] Genome-wide hypomethylation has been linked to oncogene activation and genome instability whereas promoter hypermethylation has been associated with suppression of gene expression, including different genes encoding for tumor suppressors. [40–43]

Overexpression of the histone methyltransferase (HMT) EZH2, which acts on histone H3 lysine 27 (H3K27), is one of the major hallmark in several types of leukemia and in various solid tumors.[44] In addition, aberrant modifications that occur in HMT include translocation of MLL (also known as KMT2A, ALL1 or MLL1) that causes an improper recruitment of other epigenetic proteins, such as the methyltransferase DOT1-like protein

(DOT1L). This correlation was showed in tumor xenograft model contained MLL translocations, in which inhibition of DOT1L resulted in block of H3K79 methylation and selective MLL-rearranged cells death.[45,46] In several types of cancers were also identified mutations in gene encoding histone demethylases (HMDs). Most of HDMs mutations were found on KDM1 (also known as LSD1) which encodes H3K4 and H3K9 demethylases and in KDM4 subfamily, which encodes H3K9 and H3K36.[47,48]

Human cancers also showed altered dynamic equilibrium between histone acetylation and deacetylation. The imbalance of histone acetylation marks resulted in abnormally regulation of gene expression, leading to development of cancer. Histone acetyltransferases (HATs) family is composed by three groups: general control non depressible 5 Gcn5/PCAF, p300/CBP and MYST14 families, mutated in different forms of leukemia.[49–53] The histone deacetylated enzymes are amongst the most studied epigenetic target and due to their crucial role in tumorigenesis have attracted increasing attention.[54] Furthermore, it was also reported dysregulation in phosphorylase proteins, such as overexpression of Aurora kinase B, which phosphorylates serine 10 of histone 3 (H3S10) and blocks the interaction between heterochromatin protein 1 homologue-β (HP1β) with trimethylated H3K9 (H3K9me3) during mitosis and differentiation, leading to alteration of gene transcription and cell cycle control.[55,56]

Epigenetic reader proteins bind the histone modifications and modulate effector functions. Overexpression and translocation of the gene encoding the bromodomain and extra-terminal (BET) family protein bromodomain-containing protein 4 (BRD4), has been found implicated in dysregulation in many cancer type, such acute myeloid leukemia (AML).[19,32] BRD4 has also been identified as key player in modulation of oncogene expression in multiple myeloma and non-Hodgkin lymphoma.[57,58]

Considering the large number of mutations in epigenetic genes, acting as driver in all stage of carcinogenesis, it can be hypothesised that targeting epigenetic modulators could be a promising strategy to discover effective anticancer agents.

Furthermore, in contrast to genetic mutations, which are challenging to inhibit directly, epigenetic modifications are largely reversible and offer the potential to modulate undruggable genetic changes. Indeed to date, targeting transcription factors such as MYC, using small molecules, proved to be very difficult.[59] Dysregulation of the oncoprotein MYC is a common genetic alteration in human cancers and has been showed that it can

be modulate through the inhibition of BRD4.[59–62] In fact, BRD4, through the binding with acetylated histone, regulates chromatin-dependent signaling and transcription at MYC target loci. Inhibition of BRD4 protein resulted in suppression of transcription of the MYC gene, thus showing the potential of epigenetic therapy for targeting undruggable oncogenic transcription factors.[32,63]

1.4 Epi-drugs approved for clinical uses

Due to their dynamic and reversible nature, epigenetic alterations provide an attractive opportunity for therapeutic intervention in many diseases including cancer.

In the past few years, knowledge of the proteins involved in post-translational modifications of histone has grown greatly, leading to the discovery of a rich source of potential therapeutic targets. Inhibitors of DNA methylation and histone deacetylase (HDAC) inhibitors have been approved for clinical use in myelodysplastic syndrome (MDS) and cutaneous T cell lymphoma, respectively, and several epigenetic drugs are undergoing clinical trials, thus providing proof of concept in patients for epigenetic therapies.[64]

Based on their mechanism of action, DNMT inhibitors can be divided in two classes: nucleoside analogues, which act through incorporation into DNA, and non-nucleoside analogues which directly bind the catalytic site of DNMTs. To date, several molecules that inhibit DNMT activity and reactivate gene expression, including 5-azacytidine (*azacitidine*), 5-aza-2'-deoxycytidine (*decitabine*), have been clinically investigated. Both these nucleoside analogues were approved by Food and Drug Administration (FDA) for the treatment of myelodysplastic syndrome.[65,66] On the basis of their different mechanism of action, that includes cytotoxic activity due to incorporation into DNA, these agents are characterized by a significant toxicity and lower chemical stability. These issues were addressed, employing a strategy to develop a second-generation DNMT-inhibitors, which are mostly pro-drugs of azacitidine and decitabine, with improved pharmacokinetic and pharmacodynamic properties.[67–69] Therefore, many others DMNT inhibitors were investigated, leading to the discovery of new potent compounds with improved safety profile, such as *guadecitabine* (SGI-110), *zebularine*, CP-4200[70–74]

DNMT inhibitors are being tested in several clinical trials for leukemia as well as solid tumors, even if they showed lack single-agent efficacy in solid tumors.

Another successful example of epigenetic target is represented by HDAC. To date three HDAC inhibitors have been approved for clinical uses by FDA: *Vorinostat* (also known as *SAHA*) for the treatment of cutaneous T cell lymphoma (CTCL)[75], *Romidepsin* for the treatment peripheral T cell lymphoma (PTCL)[76] and CTCL and *Belinostat* for the treatment of PTCL.[77] As well as for DNMT1 inhibitors, the use of first generation HDAC inhibitors has been limited to hematological tumors and significative adverse effects were reported. Their toxicity is probably due to broad activity across HDAC isoforms and the lack of selectivity over HDAC family enzymes. A second generation of HDAC inhibitors was then developed and currently, many of these compounds are in clinical trials for both solid and hematological tumors.[78–88] Furthermore, several studies showed an improvement of their therapeutic efficacy in combination with other agents, such as DNMT inhibitors, resulting in greater re-expression of epigenetically silenced tumor suppressor genes and cell cycle regulators,[89] or immunomodulators agent, showing increase of the host antitumor response and improvement of immunotherapeutic agents efficacy.[82,90,91] To date, several agents targeting epigenetic proteins, including HMTs, KDMs, HATs, BETs etc., are under clinical investigation, highlighting the great therapeutic potential of this field. Overall, the results obtained by scientific community in the last decade showed that epigenetic cancer therapy may represent a good weapon in the battle against cancer.

1.5 Histone Methylation balance

Histone methylation is an epigenetic mark extensively investigated in recent years and the large amount of research allowed the discovery of different druggable proteins sensitive to modulation by small molecules. Despite several histone posttranslational modifications showed their important role in the epigenetic gene regulation and were linked to several type of cancers, in my work I focused on the mediators of methyl histone marks. The histone methylation and demethylation can occur, sequentially, on lysine and arginine residues and each enzyme family has a specificity for particular residues and degree of methylation. A more in depth description of these two classes of epigenetic proteins will be discussed later. As mentioned above, histone methylation is controlled by the interplay of methyltransferases and demethylases. The loss of function of a specific demethylase may be compensated through the inhibition of the methyltransferase

modulating the same histone mark, and *vice versa*. For example, several types of cancer including esophageal, pancreatic and bladder cancers, and multiple myeloma, medulloblastoma, acute lymphoblastic leukemia, showed loss-of-function of KDM6A, a demethylase protein that acts at lysine 27 of histone H3 (H3K27me3), counteracting the activity of EZH2, the enzyme that places this modification. The loss-of-function of KDM6A results in increase of H3K27me3 levels and then transcriptional repression of tumor suppressor genes. It was found that KDM6A-mutant myeloma cells are more dependent on EZH2 to maintain high H3K27me3 levels at specific loci, and are hypersensitive to EZH2 inhibitor GSK343.[92] Hence, inhibition of EZH2 reverses aberrations promoted by KDM6A loss, showing that restoring the balance between methyltransferases and demethylases may be usefully exploited to treat cancer.
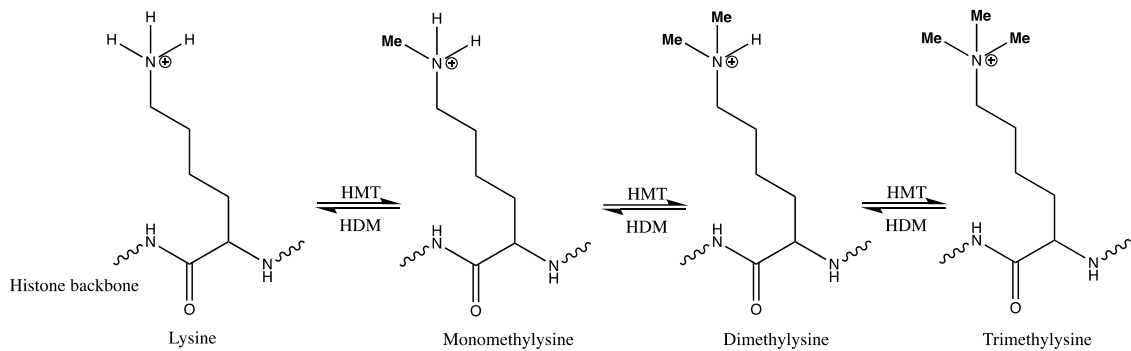

Figure 3. Sequential methylation and demethylation of lysine catalyzed by histone methyl transferases (HMTs) and histone demethylases (HDMs) respectively.

Bibliography:

1.   Allis, C. D. & Jenuwein, T. The molecular hallmarks of epigenetic control. *Nature Reviews Genetics* **17**, 487–500 (2016).

2.   Ruthenburg, A. J., Li, H., Patel, D. J. & David Allis, C. Multivalent engagement of chromatin modifications by linked binding modules. *Nature Reviews Molecular Cell Biology* **8**, 983–994 (2007).

3.   Fisher, A. G. Cellular identity and lineage choice. *Nature Reviews Immunology* **2**, 977–982 (2002).

4.   Luger, K. & Richmond, T. J. The histone tails of the nucleosome. *Current Opinion in Genetics and Development* **8**, 140–146 (1998).

5.   Kornberg, R. D. & Lorch, Y. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* **98**, 285–294 (1999).

6.   Senner, C. E. The role of DNA methylation in mammalian development. *Reproductive BioMedicine Online* **22**, 529–535 (2011).

7.   Kouzarides, T. Chromatin Modifications and Their Function. *Cell* **128**, 693–705 (2007).

8.   Yoo, C. B. & Jones, P. A. Epigenetic therapy of cancer: Past, present and future. *Nature Reviews Drug Discovery* **5**, 37–50 (2006).

9.   Li, B., Carey, M. & Workman, J. L. The Role of Chromatin during Transcription. *Cell* **128**, 707–719 (2007).

10.   Helin, K. & Dhanak, D. Chromatin proteins and modifications as drug targets. *Nature* **502**, 480–488 (2013).

11.   Worm, J. & Guldberg, P. DNA methylation: An epigenetic pathway to cancer and a promising target for anticancer therapy. *Journal of Oral Pathology and Medicine* **31**, 443–449 (2002).

12.   Kelly, T. K., De Carvalho, D. D. & Jones, P. A. Epigenetic modifications as therapeutic targets. *Nature Biotechnology* **28**, 1069–1078 (2010).

13.   Portela, A. & Esteller, M. Epigenetic modifications and human disease. *Nature Biotechnology* **28**, 1057–1068 (2010).

14.   Meaney, M. J. & Ferguson-Smith, A. C. Epigenetic regulation of the neural transcriptome: The meaning of the marks. *Nature Neuroscience* **13**, 1313–1318 (2010).

15.   Jones, P. A. & Baylin, S. B. The Epigenomics of Cancer. *Cell* **128**, 683–692 (2007).

16.   Schreiber, S. L. & Bernstein, B. E. Signaling network model of chromatin. *Cell* **111**, 771–778 (2002).

17.   Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403**, 41–45 (2000).

18.   Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **293**, 1074–1080 (2001).

19.   Gallipoli, P., Giotopoulos, G. & Huntly, B. J. P. Epigenetic regulators as promising therapeutic targets in acute myeloid leukemia. *Therapeutic Advances in Hematology* **6**, 103–119 (2015).

20.   Liss, P. S. & Slater, P. G. 5-Methylcytosine localised in mammalian constitutive heterochromatin. (1974).

21.   Keshet, I., Lieman-Hurwitz, J. & Cedar, H. DNA methylation affects the formation of active chromatin. *Cell* **44**, 535–543 (1986).

22.   Sharma, S., Kelly, T. K. & Jones, P. A. Epigenetics in cancer. *Carcinogenesis* **31**, 27–36 (2009).

23.   Hackett, J. A. & Azim Surani, M. DNA methylation dynamics during the mammalian life cycle. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368**, 1–8 (2013).

24.   Biswas, S. & Rao, C. M. Epigenetic tools (The Writers, The Readers and The Erasers) and their implications in cancer therapy. *European Journal of Pharmacology* **837**, 8–24 (2018).

25.   Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D. & Patel, D. J. How chromatin-binding modules interpret histone modifications: Lessons from professional pocket pickers. *Nature Structural and Molecular Biology* **14**, 1025–1040 (2007).

26.   Jones, P. A. & Baylin, S. B. The fundamental role of epigenetic events in cancer. *Nature Reviews Genetics* **3**, 415–428 (2002).

27.   Barski, A. *et al.* High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell* **129**, 823–837 (2007).

28.   Mikkelsen, T. S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).

29.   Bernstein, B. E. *et al.* A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* **125**, 315–326 (2006).

30.   Wu, S. Y. & Chiang, C. M. The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation. *Journal of Biological Chemistry* **282**, 13141–13145 (2007).

31.   Florence, B. & Faller, D. V. You bet-cha: a novel family of transcriptional regulators. *Frontiers in bioscience : a journal and virtual library* **6**, (2001).

32.   Dawson, M. A. *et al.* Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature* **478**, 529–533 (2011).

33.   De La Cruz, X., Lois, S., Sánchez-Molina, S. & Martínez-Balbás, M. A. Do protein motifs read the histone code? *BioEssays* **27**, 164–175 (2005).

34.   Mizuno, S. I. *et al.* Expression of DNA methyltransferases DNMT1, 3A, and 3B in normal hematopoiesis and in acute and chronic myelogenous leukemia. *Blood* **97**, 1172–1179 (2001).

35.   Abdel-Wahab, O. *et al.* Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. *Blood* **114**, 144–147 (2009).

36.   Tefferi, A. & Lim, K. Mutation in TET2 in Myeloid Cancers Age , Neuropathology , and Dementia Autologous Stem-Cell Transplantation for Multiple Myeloma. *Stem Cells* 10–11 (2009).

37.   Hypomethylation of DNA from Benign and Malignant Human Colon Neoplasms Author (s): Susan E . Goelz , Bert Vogelstein , Stanley R . Hamilton , Andrew P . Feinberg Reviewed work (s): Source : Science , New Series , Vol . 228 , No . 4696 (Apr . 12 , 198. **228**, 187–190 (2011).

38.   Baylin, S. B. & Herman, J. G. DNA hypermethylation in tumorigenesis: Epigenetics joins genetics. *Trends in Genetics* **16**, 168–174 (2000).

39.   Ehrlich, M. DNA methylation in cancer: Too much, but also too little. *Oncogene* **21**, 5400–5413 (2002).

40.   Rhee, I. *et al.* DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. *Nature* **416**, 552–556 (2002).

41.   Gaudet, F. *et al.* Induction of tumors in mice by genomic hypomethylation. *Science* **300**, 489–492 (2003).

42.   Howard, G., Eiges, R., Gaudet, F., Jaenisch, R. & Eden, A. Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. *Oncogene* **27**, 404–408 (2008).

43.   Clark, S. J. & Melki, J. DNA methylation and gene silencing in cancer: Which is the guilty party? *Oncogene* **21**, 5380–5387 (2002).

44.   Simon, J. A. & Lange, C. A. Roles of the EZH2 histone methyltransferase in cancer epigenetics. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis* **647**, 21–29 (2008).

45.   Okada, Y. *et al.* hDOT1L links histone methylation to leukemogenesis. *Cell* **121**, 167–178 (2005).

46.   Daigle, S. R. *et al.* Selective Killing of Mixed Lineage Leukemia Cells by a Potent Small-Molecule DOT1L Inhibitor. *Cancer Cell* **20**, 53–65 (2011).

47.   Magliulo, D., Bernardi, R. & Messina, S. Lysine-specific demethylase 1A as a promising target in acute myeloid leukemia. *Frontiers in Oncology* **8**, (2018).

48.   Berry, W. L. & Janknecht, R. KDM4/JMJD2 Histone demethylases: Epigenetic regulators in cancer cells. *Cancer Research* **73**, 2936–2942 (2013).

49.   Fraga, M. F. *et al.* Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nature Genetics* **37**, 391–400 (2005).

50.   Özdağ, H. *et al.* Mutation analysis of CBP and PCAF reveals rare inactivating mutations in cancer cell lines but not in primary tumours. *British Journal of Cancer* **87**, 1162–1165 (2002).

51.   Heery, D. M. & Fischer, P. M. Pharmacological targeting of lysine acetyltransferases in human disease: a progress report. *Drug Discovery Today* **12**, 88–99 (2007).

52.   Choudhary, C. *et al.* Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325**, 834–840 (2009).

53.   Di Cerbo, V. & Schneider, R. Cancers with wrong HATs: The impact of acetylation. *Briefings in Functional Genomics* **12**, 231–243 (2013).

54.   Falkenberg, K. J. & Johnstone, R. W. Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nature Reviews Drug Discovery* **13**, 673–691 (2014).

55.   Bischoff, J. R. *et al.* A homologue of Drosophila aurora kinase is oncogenic and amplified in human colorectal cancers. *EMBO Journal* **17**, 3052–3065 (1998).

56.   Hirota, T., Lipp, J. J., Toh, B. H. & Peters, J. M. Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature* **438**, 1176–1180 (2005).

57.   Lovén, J. *et al.* Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**, 320–334 (2013).

58.   Miyoshi, I. *et al.* BRD4 bromodomain gene rearrangement in aggressive carcinoma with translocation t(15;19). *American Journal of Pathology* **159**, 1987–1992 (2001).

59.   Darnell, J. E. Transcription factors as targets for cancer therapy. *Nature Reviews Cancer* **2**, 740–749 (2002).

60.   Meyer, N. & Penn, L. Z. Reflecting on 25 years with MYC. *Nature Reviews Cancer* **8**, 976–990 (2008).

61.   Beroukhim, R. *et al.* The landscape of somatic copy-number alteration across human cancers. *Nature* **463**, 899–905 (2010).

62.   Delmore, J. E. *et al.* BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell* **146**, 904–917 (2011).

63.   Zuber, J. *et al.* RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* **478**, 524–528 (2011).

64.   Mohammad, H. P., Barbash, O. & Creasy, C. L. Targeting epigenetic modifications in cancer therapy: erasing the roadmap to cancer. *Nature Medicine* **25**, 403–418 (2019).

65.   Hodges, E. *et al.* Directional DNA methylation changes and complex intermediate states accompany lineage specificity in the adult hematopoietic compartment. *Molecular Cell* **44**, 17–28 (2011).

66.   Agrawal, K., Das, V., Vyas, P. & Hajdúch, M. Nucleosidic DNA demethylating epigenetic drugs – A comprehensive review from discovery to clinic. *Pharmacology and Therapeutics* **188**, 45–79 (2018).

67.   Lübbert, M. *et al.* Decitabine improves progression-free survival in older high-risk MDS patients with multiple autosomal monosomies: results of a subgroup analysis

of the randomized phase III study 06011 of the EORTC Leukemia Cooperative Group and German MDS Study Group. *Annals of Hematology* **95**, 191–199 (2016).

68.   Fenaux, P. *et al.* Azacitidine prolongs overall survival compared with conventional care regimens in elderly patients with low bone marrow blast count acute myeloid leukemia. *Journal of Clinical Oncology* **28**, 562–569 (2010).

69.   Fenaux, P. *et al.* Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: a randomised, open-label, phase III study. *The Lancet Oncology* **10**, 223–232 (2009).

70.   Issa, J. P. J. *et al.* Safety and tolerability of guadecitabine (SGI-110) in patients with myelodysplastic syndrome and acute myeloid leukaemia: A multicentre, randomised, dose-escalation phase 1 study. *The Lancet Oncology* **16**, 1099–1110 (2015).

71.   Yoo, C. B. *et al.* Delivery of 5-aza-2′-deoxycytidine to cells using oligodeoxynucleotides. *Cancer Research* **67**, 6400–6408 (2007).

72.   Kantarjian, P. H. M. *et al.* Phase 2 study of guadecitabine (SGI-110), a novel hypomethylating agent, in treatment-naïve patients with acute myeloid leukaemia who are not candidates for intensive chemotherapy. **18**, 1317–1326 (2018).

73.   Zhou, L. *et al.* Zebularine: A novel DNA methylation inhibitor that forms a covalent complex with DNA methyltransferases. *Journal of Molecular Biology* **321**, 591–599 (2002).

74.   Hummel-Eisenbeiss, J. *et al.* The role of human equilibrative nucleoside transporter 1 on the cellular transport of the DNA methyltransferase inhibitors 5-azacytidine and CP-4200 in human leukemia cells. *Molecular Pharmacology* **84**, 438–450 (2013).

75.   Mann, B. S., Johnson, J. R., Cohen, M. H., Justice, R. & Pazdur, R. FDA Approval Summary: Vorinostat for Treatment of Advanced Primary Cutaneous T-Cell Lymphoma. *The Oncologist* **12**, 1247–1252 (2007).

76.   Bertino, E. M. & Otterson, G. A. Romidepsin: A novel histone deacetylase inhibitor for cancer. *Expert Opinion on Investigational Drugs* **20**, 1151–1158 (2011).

77.   Lee, H. Z. *et al.* FDA approval: Belinostat for the treatment of patients with relapsed or refractory peripheral T-cell lymphoma. *Clinical Cancer Research* **21**, 2666–2670 (2015).

78.   Evens, A. M. *et al.* A phase I/II multicenter, open-label study of the oral histone deacetylase inhibitor abexinostat in relapsed/refractory lymphoma. *Clinical Cancer Research* **22**, 1059–1066 (2016).

79.   Guzman, M. L. *et al.* Selective activity of the histone deacetylase inhibitor AR-42 against leukemia stem cells: A novel potential strategy in acute myelogenous leukemia. *Molecular Cancer Therapeutics* **13**, 1979–1990 (2014).

80.   Eigl, B. J. *et al.* A phase II study of the HDAC inhibitor SB939 in patients with castration resistant prostate cancer: NCIC clinical trials group study IND195. *Investigational New Drugs* **33**, 969–976 (2015).

81.   Qian, C. *et al.* Cancer network disruption by a single molecule inhibitor targeting both histone deacetylase activity and phosphatidylinositol 3-kinase signaling. *Clinical Cancer Research* **18**, 4104–4113 (2012).

82.   Knipstein, J. & Gore, L. Entinostat for treatment of solid tumors and hematologic malignancies. *Expert Opinion on Investigational Drugs* **20**, 1455–1467 (2011).

83.   Galli, M. *et al.* A phase II multiple dose clinical trial of histone deacetylase inhibitor ITF2357 in patients with relapsed or progressive multiple myeloma. *Annals of Hematology* **89**, 185–190 (2010).

84.   Furlan, A. *et al.* Pharmacokinetics, safety and inducible cytokine responses during a phase 1 trial of the oral histone deacetylase inhibitor IFT2357 (givinostat). *Molecular Medicine* **17**, 353–362 (2011).

85.   Younes, A. *et al.* Mocetinostat for relapsed classical Hodgkin's lymphoma: An open-label, single-arm, phase 2 trial. *The Lancet Oncology* **12**, 1222–1228 (2011).

86.   Brunetto, A. T. *et al.* First-in-human, pharmacokinetic and pharmacodynamic phase i study of resminostat, an oral histone deacetylase inhibitor, in patients with advanced solid tumors. *Clinical Cancer Research* **19**, 5494–5504 (2013).

87.   Santo, L. *et al.* Preclinical activity, pharmacodynamic, and pharmacokinetic properties of a selective HDAC6 inhibitor, ACY-1215, in combination with bortezomib in multiple myeloma. *Blood* **119**, 2579–2589 (2012).

88.   Leung, D. *et al.* Integrative analysis of haplotype-resolved epigenomes across human tissues. *Nature* **518**, 350–354 (2015).

89.   Thurn, K. T., Thomas, S., Moore, A. & Munster, P. N. Rational therapeutic combinations with histone deacetylase inhibitors for the treatment of cancer. *Future Oncology* **7**, 263–283 (2011).

90.   Héninger, E., Krueger, T. E. G. & Lang, J. M. Augmenting antitumor immune responses with epigenetic modifying agents. *Frontiers in Immunology* **6**, 1–14 (2015).

91.   Cruickshank, B. *et al.* Dying to be noticed: Epigenetic regulation of immunogenic cell death for cancer immunotherapy. *Frontiers in Immunology* **9**, 1–10 (2018).

92.   Ezponda, T. *et al.* Loss of the Histone Demethylase UTX Contributes to Multiple Myeloma and Sensitizes Cells to EZH2 Inhibitors. *Blood* **124**, 611 (2014).

Part I

Design and Synthesis of KDM4A Inhibitors

CHAPTER 2

# Histone Demethylases and KDM4s subfamily

2.1 Histone demethylases

Histone methylation was initially considered a stable and irreversible modification that could be only erased during histone exchange or DNA replication. However, only after the demonstration that lysine specific demethylase 1 (LSD1) catalyses the demethylation of H3K4me1/me2, it was recognised that this mark can be dynamically regulated, through the recruitment of methyltransferases and demethylases.[1] Following studies led to discovery of demethylase enzymes containing the catalytic Jumonji C (JmjC) domain.[2–5] To date, since the identification of LSD1, over 30 histone demethylases have been identified and have shown demethylase or hydroxylase activity.[6]

Depending on the sequence homology and structural similarities, the KDMs are clustered into seven subfamilies, (KDM1-7) while based on their catalytic mechanism these are grouped into two large families: KDM1 (LSDs) and KDM2-7.[7]

The first family includes lysine-specific demethylases: LSD1 (also called KDM1A, AOF2, BHC110 or KIAA0601)[1] and LSD2 (also called KDM1B or AOF1)[8] which are closely related to flavin adenine dinucleotide (FAD)-dependent monoamine oxidases (MAOs). The catalytic activity resides on the amine oxidase-like (AOL) domain, which is characteristic of several metabolic enzymes, and is formed by two lobes: the cofactor FAD-binding site and the substrate recognition site. The histone substrate site is positioned close to another lobe, the SWIRM domain, thus forming a hydrophobic cleft that allows LSD1 to accommodate histone H3 tail, and performs the demethylase activity.[9] The mechanism proposed involves the methyl group oxidation via hydride transfer from the N-methyl group onto FAD, forming an imine which is unstable to hydrolysis.[10] (Figure **1A**) This class of enzymes, according to their mechanism of action, is unable to remove trimethyl histone mark, since they require a lone pair of electrons on the lysine nitrogen atom.[1,11,12] Indeed, LSDs act on mono- an dimethyl histone lysine residues, specifically LSD1 demethylates mono- and dimethyl lysine residues at H3K4 and H3K9, while LSD2 acts only on H3K4. In addition, LSD1 has also been shown to

demethylate several non-histone substrates, as transcription factor p53 on the K370 residue, DNA methyltransferase 1 (DNMT1) on the K1096 residue, E2F1 on the K185 residue and MYPT1, that play a key role during gene expression. [13–16] Furthermore, several studies suggested that substrate recognition is regulated by the interaction with other protein partners. Depending on the substrate or protein partners, it has been demonstrated that LSD1 activity may results either in transcriptional repression or promotion. Indeed, LSD1 suppresses gene transcription as part of the CoREST complex whereas promotes transcriptional activation by binding to the androgen receptor (AR) or estrogen receptor (ER).[17] The second family of histone demethylases consists of α-ketoglutarate (α-KG) and Fe(II)-dependent oxygenases and is characterized by the presence of the Jumonji C (JmjC) domain. In humans, around 20 histone demethylases containing the JmjC domain have been identified and, based on the structural homology, this class can be divided into five subfamilies: KDM2/7, KDM3, KDM4, KDM5, and KDM6. [18,19] The JmjC domain-containing proteins act through an oxidative mechanism that requires two cofactors: Fe(II) and α-KG (also known as 2-oxoglutarate [2-OG]) which interact with the JmjC domain and demethylate lysine residues through N-methyl hydroxylation. In particular, the two cofactors react with dioxygen to form a highly reactive oxoferryl (Fe(IV)=O) species, leading to the incorporation of an oxygen atom (hydroxylation) on the methyl group of the methylated lysine substrate. The resulting hemiaminal product is unstable and spontaneously undergoes a hydrolysis reaction that yields a demethylated compound and formaldehyde. (Figure **1B**) Contrary to the first class, the mechanism of these enzymes does not require a protonable ε-ammonium group, allowing the removal of mono- di- and trimethyl histone lysine marks.[20]

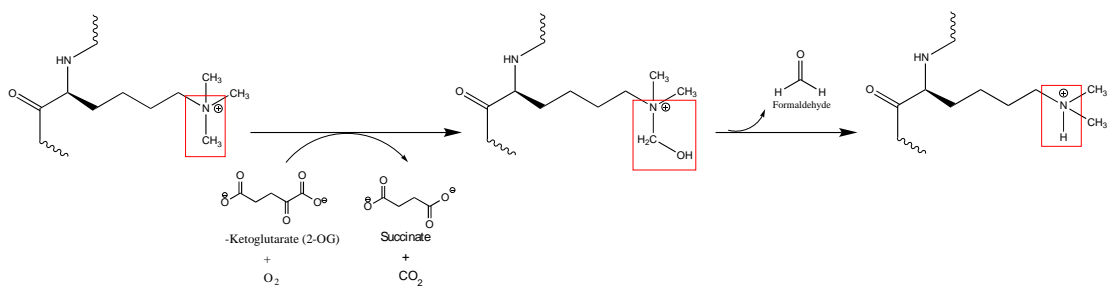**A   LSD1/KDM1 family:**
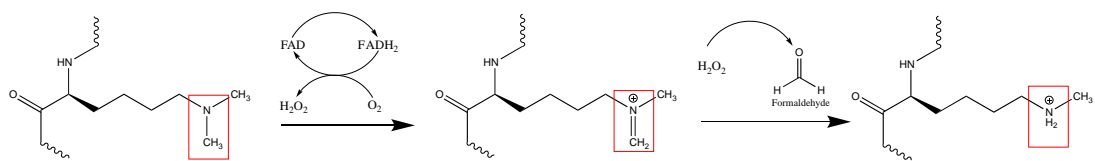
B   JmjC family:

Figure 1. Reaction mechanism of demethylation. Demethylation catalyzed by (A) LSD1/KDM1 family and (B) Jmjc family enzymes.

2.2 KDM4 subfamily

The human KDM4 subfamily is composed of four members KDM4A, KDM4B, KDM4C, KDM4D. In addition, KDM4D forms a cluster with two different genes, generating two further pseudogenes: KDM4E and KDM4F.[21] The enzymes KDM4A-C share more than 50% sequence identity and consist of a catalytic histone demethylase domains formed by Jumoji C (JmjC) domain and Jumonji N (JmjN) domain and a non-catalytic domains: two plant homeodomains (PHD) and two Tudor domains. Meanwhile, KDM4D contains only the catalytic domain and lacks PHD and Tudor domains. [22] (Figure **2**) The catalytic function is carried out by the JmjC domain, while the close JmjN domain, interacting extensively with it provides a structural integrity. Moreover, the non-catalytic domains perform different functions by determining substrate specificity and controlling enzyme activity. [23,24]
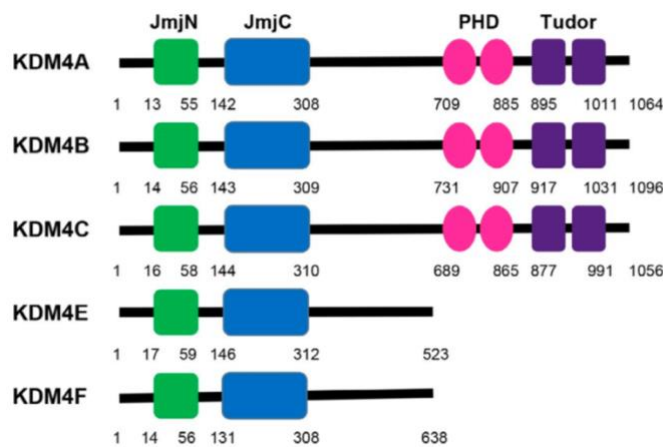

Figure 2. Schematic representation of KDM4s structure.

2.2.1 KDM4A structure

The catalytic domain JmjC is characterized by a double-stranded β-helix (DSBH) fold which is shared with the other members of 2-oxaglutarate oxygenases family. The β-sheets folding forms the active site pocket that bounds the two cofactors: Fe (II) and α-

KG. The Fe (II) coordination is mediated by three conserved residues, known as facial triad and two additional residues bound α-KG.[25,26]

The KDM4A crystal structure reveals a lysine-binding site in which substrates are bound in distinct bent conformations involving the Zn-binding site. Indeed, the Zinc ion exhibited a structural role in the formation of an active enzyme.[27]

Full-length KDM4A protein comprise 1.064 residues, for an overall surface area of 580 $Å^2$ and, as mentioned above, is characterized by a catalytic core formed by JmjN domain, JmjC domain and a zinc finger motif. In addition, the protein contains double non-catalytic domains as two plant homeodomains (PHD), and two Tudor domain. (Figure **3A**)
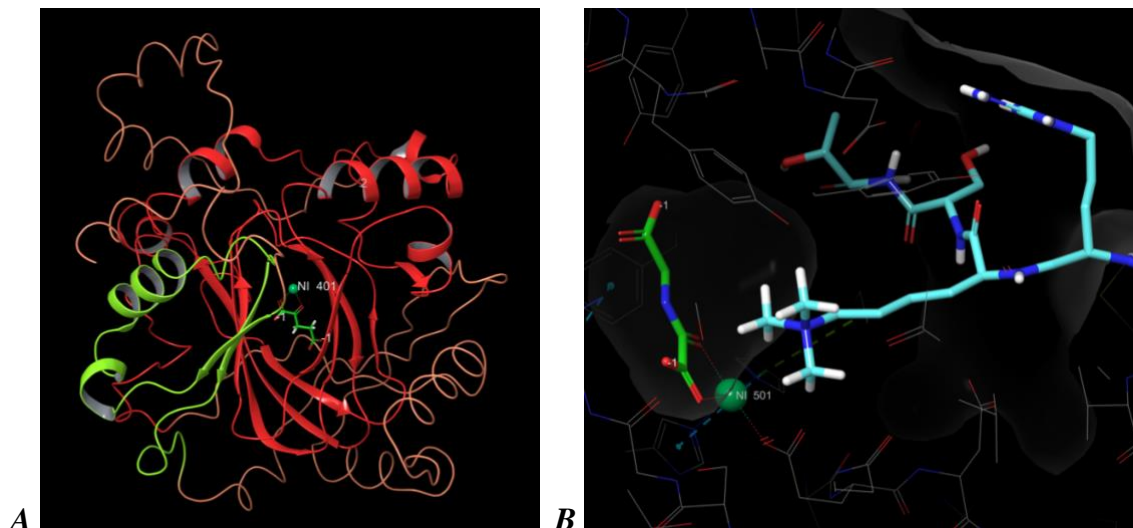

Figure 3. Structure of KDM4A (JMJD2A). (A) The JmjC domain (red) and JmjN (green) are depicted in cartoon representation, while the mixed region (orange) is depicted in thin tube representation, 2-OG showed in green stick and Nickel ion in green sphere. (PDB code: 5TVR)

(B) Trimethylated lysine side chain in catalytic region. (PDB code:2OQ6)2-OG showed in green stick, peptide in light blue stick and Nickel ion in green sphere.

The crystal structure of the histone demethylase KDM4A catalytic domain in complex with 2OG revealed the Fe(II) is coordinated by three residues: His188, Glu190 and

His276, and the co-factor 2OG in a bidentate manner. In several complexes the Ni(II) occupies the endogenous Fe(II) site and NOG replaces 2OG.[28]

The histone substrates are bound in a broad hydrophobic cleft and the methylated nitrogen of the lysine side chain is positioned in a deep catalytic pocket close to the metal cofactor (Figure **3B**). The binding site is characterized by a polar environmental, due to the presence of hydroxyl groups of residues Tyr177, Ser288 and three carbonyl groups of Gly170, Glu190 and Asn290. These residues establish electrostatic (Coulomb) interactions with methylated lysine nitrogen atom positively charged. The polar environmental in the binding pocket has been showed to be crucial for the catalytic activity because it may enable the methyl group to assume the correct orientation toward the Fe(II) ion, in order to facilitate the reaction. As shown in figure 4, in the catalytic core the 5-carboxylic acid group of 2OG forms a salt bridge with the basic residue Lys 206 and in parallel establishes H-bond interactions with the side-chain of two H-bond donor residues: Lys206 and Tyr132. In the binding site the histone methyl groups are coordinated by a network of CH---O hydrogen bonds.[28–31]
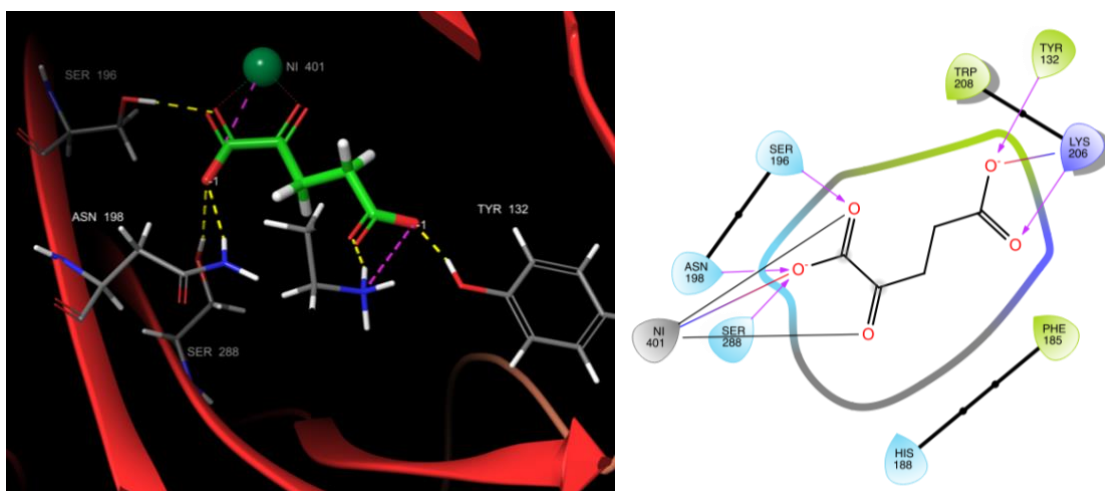

Figure 4. A detailed view of 2-OG binding mode in 3D and 2D.
2-OG shows green stick. The hydrogen bonds are depicted in yellow and electrostatic interaction in purple.

The structure folding of double Tudor domain was found to play a key role in the interaction with methylated histone tails. It was shown that this domain can directly bind

methylated H3K4, H3K9 and H4K20 through a specific chromatin-targeting module. In particular, a cluster of aromatic residues surrounded the Tudor domain, characterized by a negatively charged electrostatic potential, consistent with its supposed role in the binding of methyl-lysine substrate. [23]

2.2.2 Histone substrate specificity

Histone methylation can occur on the lysine residues of histone H3 and H4. KDMs catalyze the demethylation of di- and tri-methylated histone H3 on four lysine residues: K9 (H3K9me2/3)[5], K27 (H3K27me2/3)[32], K36 (H3K36me3)[2] K4 (H3K4), and histone H4 at lysine20 (H4K20).[33] Furthermore, the linker histone H1.4 can be methylated at lysine 26 (H1.4K26)[22]. The histone methylation plays a key role in modulating gene transcription and results in several functional outcomes depending on the specific methyl-substrate recognized by the protein or the number of methyl marks added or removed, leading to a different effect on the chromatin state. Overall, methylation at H3K4, and H3K36 is associated with transcriptionally active regions, while methylation at H3K9, H3K27, H4K20 and H1.4K26 is associated with gene silencing.[34–36]

The KDM4 subfamily showed functional diversity according to their differential substrate specificities. Indeed, despite a sequence identity higher than 75% within their catalytic domain, KDM4 enzymes exhibited a considerable variability in the substrate specificities. KDM4A-B-C can more efficiently demethylate H3K9me3, H3K36me3 than H3K9me2/H3K36me2 whereas KDM4D is more efficient at H3K9me2 than H3K9me3 but is unable to recognize H3K36me3.[37]

Structural comparison between the crystal structures of these homolog enzymes reveals that KDM4A-B-C conserved the residues that can suitably accommodate H3K36 substrate, while KDM4D contains divergent residues in the same region, displaying steric and electrostatic clashes as well as lack of efficacious interactions that hamper the recognition of H3K36me3 site.[28] In details, analyzing the crystal structure of KDM4A protein in complex with H3K9 (PDB ID: 2OQ6) and H3K36 (PDB ID:2OS2) (Figure **5**), it can be noticed that the two peptides adopt different conformation and orientation and H3K36 interacts with a loop (Arg309-Met312) that is well conserved also in KDM4B and KDM4C but is divergent in KDM4D. Furthermore, structural studies of KDM4

proteins reveal that KDM4A substrate specificity relies also on peculiar residues position surrounding the methylation site, thus explaining the demethylases activity with similar efficiency on H3K9 and H3K36, despite their low sequence similarity. KDM4A binds the histone peptide predominantly through main chain-main chain interactions. Specifically, the substrates is bound in the aromatic cage through cation-π interaction and is recognized through a network of CH---O hydrogen bonds. [28–30] In addition to the catalytic domain specificities, the Tudor and PHD domains have been found to play a key role in the recognition of methylated histone tail, discriminating among both site and degree of lysine methylation. Specifically, structural and biochemical studies have demonstrated that the TD of KDM4A are able to recognize and bind H3K20me3/me2 and H3K4me3/me2 marks. [23,24,38,39]

Every histone mark is related to the expression of specific target genes. Considering these broad substrate specificity in KDM4 subfamily, it's clear that a dis-regulation in KDM4s activity can affect the physiological functions, leading to the onset of a disease.[40]
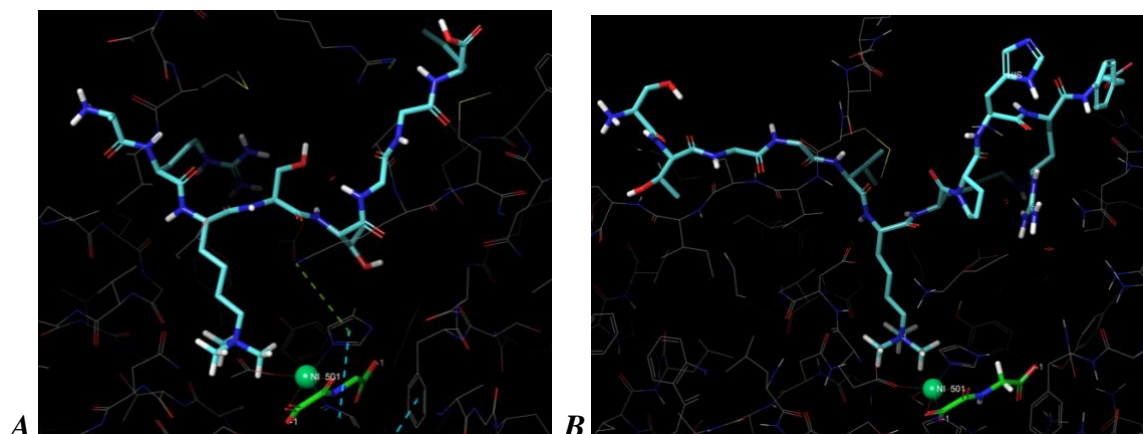

Figure 5. Substrate binding pose. *(A) KDM4A-H3K9me3 (PDB code: 2OQ6). (B) KDM4A-H3K36me3 (PDB code: 2OS2). 2-OG shows green stick. Peptide shows light blue stick and the green sphere is nickel ion.*

2.2.3 KDM4 role in cancer

Histone methylation and demethylation balance is essential to maintain cellular homeostasis. Aberrant activity of lysine demethylases is frequently observed in cancer and can results in different biological outcomes, depending on the specific gene expression pathway/network involved. Changes in expression levels has been found more frequently rather than mutations and translocation of genes encoding demethylases in several types of cancer, but also in cardiovascular diseases, mental retardation, and neurodegenerative disorders.[41,42] In cancer, deregulation of KDMs has been correlated directly or indirectly, to chromosomal stability change, tumor suppressor inactivation, promotion of oncogene expression, hormone receptor binding and downstream signaling.[43–45]

Genetic studies have shown that decrease of H3K9me3 is related to carcinogenesis in different mouse models. Indeed, H3K9me3 is usually an hallmark of stable heterochromatin regions of the genome that are transcriptionally inactive, and loss or reduction of this trimethylated histone level led to a conformational change from heterochromatin to euchromatin state, enabling the activation of genetic transcription. Therefore, overexpression of lysine demethylases targeting H3K9me3, such as KDM4 subfamily, can results in downregulation of H3K9me3, leading to genomic instability and aberrant gene activation.[46] To date, it has been demonstrated that the amplification and overexpression of histone demethylases are involved in tumor progression and in particular, deregulation of KDM4A-B-C has been linked to breast, prostate, esophageal, lung, and lymphatic cancers.[43,45,47] For instance, several studies reveal that KDM4C is overexpressed in squamous cell carcinoma, breast cancer and medulloblastoma,[47–49] whereas KDM4B has been showed to be involved in tumor proliferation and development as well as in breast cancer, colorectal cancer, prostate cancer and can promote cell growth through increased expression of cell cycle regulators.[50–54]

Although KDM4A-C share high sequence homology they display different subcellular localization in cancer cells: KDM4A is expressed similarly in both the nucleus and cytoplasm, KDM4B is mainly expressed in the nucleus and KDM4C is linked to chromatin. Thus, according to their different distribution their mis-regulation can results in distinct oncogenic function.

KDM4A-C were found overexpressed in prostate tumors and Pca cell lines.[5] The KDM4 subfamily is capable to activate the androgen receptor (AR) signaling that is known to be a key driver of prostate cancer (Pca) progression. In particular, KDM4A forms a complex with AR and acts as activator of prostate-specific antigen (PSA) gene by demethylating H3K9me3 at the PSA promoter.[55,56] Furthermore, KDM4A through the binding with ETS transcription factor ETV1 regulates the expression of the yes-associated protein 1 (YAP1) that is correlated to prostate tumor progression.[57,58]

In breast cancer the prevalent KDM4 isoforms overexpressed are KDM4A and KDM4D.[59] KDM4A is involved in breast carcinogenesis due to its interaction with the α estrogen receptor (ER). In fact, KDM4A forming a complex with αER, stimulate its activity and also promote αER-mediated transcription.[60] In addition KDM4A levels were found to be correlated to tumor stage, TNM and disease-free survival (DFS). KDM4A promotes metastasis and impacts tumor grade by silencing Sp-1 protein that is negatively associated with TNM stage.[61,62] Accordingly, inhibition of KDM4A resulted in reduction of proliferation, migration and invasion in breast cancer cell line MCF-7.[63]

Furthermore, KDM4A has been found overexpressed in non-small cell lung cancer (NSCLC) and has been negatively correlated with overall survival.[64] Several studies reported that through the regulation of H3K9me2/3 levels, is involved in lung carcinogenesis by increasing the expression of tumor-associated genes, such as ADAM12, CXCL5, and JAG1, and by acting as negative regulator of tumor suppressor genes.[65] In colon cancer cell line HCT-116 all KDM4A, KDM4B, KDM4C are overexpressed.[66] In this context the role of KDM4A seems to be associated with different functions in DNA damage pathway. Indeed, KDM4A act as a suppressor of p21 protein, that is a cell cycle inhibitor recruited in response to DNA damage, thus promoting tumorigenesis. Downregulation of KDM4A results in increase of p21 expression, then increase of apoptosis and consequently inhibition of CRC cell proliferation.[67] Researchers, identified KDM4A as promising therapeutic target also for pancreatic and gastric cancer treatment.[68–70] Decreased levels of KDM4A induces apoptosis through dual mechanism: activation of pro-apoptotic proteins and downregulation of antiapoptotic proteins.

KDM4A is further associated with proliferation, migration and metastasis in endometrial cancer cell lines RL95-2 , ISK and in glioma cell lines U87 and T98G.[71–73] In both studies

downregulation of KDM4A inhibits invasion and induces apoptosis, but in the latter case also reduces glioma cell survival by activating autophagy.[73] Moreover, KDM4A showed a crucial role in bladder cancer tissue, activating, always via H3K9 modulation, tumor-associated genes and AR-induced genes transcription.[74]

2.3.4 KDM4 inhibitors

As described in the section above, there are several evidences to suggest that overexpression and amplification of KDM4 proteins are involved in carcinogenesis and tumor progression, acting as both promoter of oncogenes and negative modulators of onco-suppressor. Therefore, KDM4 subfamily and in particular KDM4A may represent a suitable therapeutic target for cancer treatment.

To date, since the discovery of the demethylases proteins, several lysine demethylases inhibitors were identified.[75,76] These compounds can be divided into three categories depending on their mechanism of action. The first class is composed by 2-OG (compound **1**) with metal-chelated moiety and includes molecules that compete with the natural cofactor and bind the metal ion in the catalytic site (Figure **6**). The second class includes the metal cofactor disruptors that blocked the catalytic activity interacting with the metal ions in the binding pocket and then causing its structural ejection. The last class is represented by the histone substrate inhibitors, such methyl-lysine histone substrate mimics (Figure **7**).

To date, several chemotype of KDMs inhibitors have been identified and most of the published compounds act as 2-OG competitive inhibitors. Nevertheless, clinical compounds have not yet been reported for KDM4 subfamily.

- **2-OG cofactor mimics with metal-chelated moiety**

The research of histone demethylases JmjC domain containing the development of 2-OG cofactor competitive inhibitors represents the major strategy followed to date. This class of compounds competitively binds Fe (II) ion and occupies the key region inside the binding pocket, displacing the natural 2-OG required for the catalytic activity. [77]

The 2-OG analogues reported could be divided based on the different chemotype in a) oxalyl acid derivatives, b) hydroxamic acid derivatives, c) hydrazide derivatives, d) 8-

hydroxyquinoline derivatives, e) benzimidazole pyrazolone derivatives and f) pyridine derivatives.

1. Oxalyl acid derivatives

The *N-oxalylglycine* (NOG) (compound **2**) is an oxalyl acid derivative discovered initially as HIF prolyl hydroxylase PHD2 inhibitor and then characterized also as KDM4A and KDM4C inhibitor showing a 24 μM IC_{50} value for KDM4A.[77] The co-crystal structure of NOG inside KDM4A binding pocket showed that the oxalyl group of NOG binds to the Fe(II) ion, while the carboxyl group of glycine interacts with Lys241 and Tyr177 residues through hydrogen bonds pathway.[28] Despite some of oxalyl derivatives showed a potent inhibition of KDMs, their further development to deliver a lead compound is limited due to their high hydrophilicity and poor cell penetration ability.[19]

Structural study revealed that KDM4A possesses a sub-pocket adjacent to the active site that is not present in the PHD2 domain. Therefore a suitable approach to design inhibitors selective toward KDM4A could be the addition to the oxalyl acid scaffold of a big side chain that can be accommodated into this hydrophobic area. Moreover, depending on the different NOG side chain selectivity can be achieved toward the 2-OG dependent oxygenase subfamily.[78,79]

2. Hydroxamic acid and hydrazide derivatives

The hydroxamic acid (compound **3**) is a potent chelator of Fe (II) ion in KDM4 subfamily binding site. One of the first compounds tested on KDMs has been SAHA (compound **4**), an HDAC inhibitor with a hydroxamic acid group as metal-chelating moiety.[77] Compound **5** is a hydroxamic acid derivative that exhibited activity in the low micromolar range with $IC_{50} = 3$ μM and selectivity toward KDM4s over PDH2.[80] The replacement of dimethyl-amino group with a cyclopropyl ring lead to compound **6** that showed selectivity to KDM2/7 family.[81] Docking studies revealed that this selectivity can be due to the different environment in the binding site of KDM4 subfamily and KDM2/7. Indeed the KDM4A pocket is characterized by polar residues such as Tyr79 and Asp137, whereas KDM7A presents hydrophobic residues, as Phe250 and Phe359. The compound **7** is another potent KDM4A inhibitor with $IC_{50} = 1.7$ μM that contains two hydroxamic acid group. This compound showed antiproliferative effect against KYSE-150 and HL-60

cancer cell lines with $GI_{50} = 44\mu M$ and $53\mu M$ respectively.[82] Despite the wide variety of KDMs inhibitors with hydroxamic acid scaffold discovered, only few have reached clinical experimentation in Phase I and II. However to date there are four approved hydroxamic acid-based drugs.

3. *Hydrazide derivatives*

In general the hydrazide group is not a good choice as metal chelator because the amino nitrogen may be protonated and lose its lone pair electrons. Nevertheless, some successful and selective hydrazide derivatives were reported as KDMs inhibitors, with the hydrazide moiety coordinating the metal ion in a bidentate manner. For instance, compound **8** a tetrazolyl hydrazide compound showed activity with $IC_{50} = 1.7 \mu M$ against KDM4A. The tetrazole moiety occupies a similar position of the carboxylic group in 2-OG cofactor and contributes to maintain high binding affinity.[83]

4. *8-Hydroxyquinoline (8-HQ) derivatives*

The 8-Hydroxyquinoline (8-HQ) moiety was identified as successful scaffold with potent metal-chelator group. The 5-carboxy-8-hydroxyquinoline (compound **9**, IOX1) is a potent KDM4s inhibitor with IC_{50} of 0.6 μM against KDM4A. IOX1 binds the Fe(II) ion in a bidentate manner through nitrogen atom of quinoline moiety and phenol oxygen atom, while with the 5-carboxylic group interacts with Tyr 132 and Lys 206, mimicking the natural cofactor 2-OG. The 4-carboxy-8-hydroxyquinoline (4C8HQ), compound **10**, showed a similar binding mode to their isomer, but is less potent with IC_{50} of 2 μM, probably because IOX1 can determine the translocation of the Fe(II) ion in binding site.[84,85]

Rational approaches as structural based drug design (SBDD) and high throughput screening campaigns lead to the discovery of several 4-, 5-, 7-substituted 8-hydroxyquinoline derivatives. The compound **11** exhibited antiproliferative activities against HCT116, MCF-7 and A549 cancer cell lines with $IC_{50}= 1.5 \mu M$, 27.6μM, and 21.7 μM respectively.[86] The compound **12** is a potent and selective inhibitor of histone trimethyl-lysine demethylases KDM4 and blocks the growth of prostate cancer cells and PC3 xenograft model in vivo preventing the interaction between KDM4B and its promoter.[50]

5. Benzimidazole derivative

Starting from a HTS campaign of small molecules library, a new KDM4s inhibitor was identified with benzimidazole scaffold. Compound **13** showed activity with IC_{50} in the low micromolar range and antiproliferative effects in Pca cellular models, and was selected as starting point to an hit-to-lead optimization campaign.[87]

6. Pyridine derivatives

The pyridine moiety is one of the most exploited scaffold in KDM4 inhibitors. In the last few years several pyridine dicarboxylates (PDCA)-based inhibitors were identified. Compound **14**, the 2,4 PDCA, is a potent derivative in which the nitrogen atom in the pyridine moiety and the carboxylic acid group in 2 position chelate the metal ion in a bidentate manner and the carboxylic group in 4 position forms hydrogen bonds with Lys 206 and Tyr132. The introduction of a hydrophobic moiety on 3-position (e.g. compound **15**) lead to selectivity toward KDMs over PHD2.[79,88]

The compound **16** is a potent pyridine-based inhibitor against KDM4A with IC_{50}= 0.94 µM. The crystal structure of KDM4A protein in complex with compound **16** showed that the nitrogen atoms of pyridine binds the metal ion in a bidentate manner and the carboxylic group interacts with Tyr132 and Lys206. The secondary amine binds the carboxylate group of Glu190 and the terminal pyridine is inserted in a pocket formed by Tyr175, Val171, and Asp191.[89]

The compound **17** is formed by a 4-carboxylic-pyridine with a triazole moiety in 2-position and showed moderate activity toward KDM4 subfamily (pIC_{50}=5.5±0.33). The triazole moiety replaced the carboxylic acid group as chelator together with the pyridine nitrogen. Starting from an HTS campaign Bavetsias et al. identified N-substituted 4-(pyridine-2-yl)thiazole-2-amine derivatives as KDMs inhibitors.[90] An optimization strategy of compounds containing-pyridine scaffold yielded to compound **18** in which the carboxylic group on 4-position was isosterically replaced by pyrido[3,4-d]pyrimidin-4(3H)-one group, leading to a potent compounds with IC_{50} in the nanomolar range and an enhancement of physico-chemical properties and cellular permeability maintaining the binding affinity. Furthermore, docking studies of compound **18** revealed that the chlorobenzyl piperidine moiety is positioned into the histone substrate binding site, that could explain the selectivity toward KDM4s.[91,92]

Chen et al. discovered potent 4-carboxylic pyridine derivatives KDM4s inhibitor with antiproliferative activity and high cellular permeability. The researchers started with a fragment screening identifying compound **19** as active compound with IC$_{50}$=1.4 μM and ligand efficiency (LE)=0.73.[93] Like other compounds of this class, the fragment binds the metal ion with the nitrogen atom of pyridine and forms hydrogen bonds with Tyr132 and Lys206 with the carboxylic group in 4-position. Due to its high hydrophilicity compound **19** showed poor cellular permeability. Molecular modeling studies were performed and compound **20** (QC6352) was obtained with improved cellular potency and reasonable pharmacokinetics and bioavailability properties. QC6352 exhibited enzymatic activity against KDM4s with IC$_{50}$ in the nanomolar range and in vivo efficacy in a mouse xenograft model (EC$_{50}$=3.5nM in KYSE-150 cell line and 61% of tumor growth inhibition at dose level 50mg/kg).[94]

Another successful fragment-based strategy was reported by Korczynska et al. in which a fragment library was docked into the high-resolution KDM4A structure. Two of the most interesting compounds, 5-carboxypyridine and 5-aminosalicilate, that docked in distinct but overlapping orientation were chosen and linked in order to create a new potent scaffold containing a carboxy-pyridine and a phenol moieties. The new fragment were further derivatized and optimized to obtain compound **21** with a Ki=43nM and IC$_{50}$=0.064 μM on KDM4C.[95]

- **Metal cofactor disruptors**

Disruption of iron and zinc ions is another suitable strategy to inhibit KDM4A. Structural study revealed that the Zn (II) binding site (His220, Cys234, Cys306 and Cys308) is located very close to the catalytic domain and the substrate-recognition residues (Lys241, Arg309) in KDM4A playing a key role in substrate positioning process. Furthermore it is not present in other 2-OG-dependent oxygenase. Zinc-ejection compounds, including disulfiram (**22**) and ebselen (**23**), could locally release the metal ion and inhibit KDM4A activity with IC$_{50}$=3 μM and 10.6 μM. Although this strategy can be suitable for development of selective KDM4 inhibitors, there are a large number of human proteins containing zinc ion, therefore the selectivity of metal cofactor disruptors requires further investigations.[96]
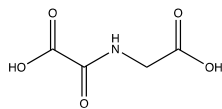
- **Histone substrate competitive inhibitors**

Truncated or modified substrate peptides were designed with the aim to develop substrate mimics inhibitors,. In this context, truncated H3K9me3 substrate H3(7-14)Kme$_3$ was identified and showed affinity to KDM4A with K_i=121μM and K_{cat}=0.01 min^{-1}.[97] Moreover, several methyl-lysine histone substrate mimics have been developed by introducing a metal-chelated moiety, such as uracil group, carboxylate group or oxalyl acid group, leading to compounds **24**, **25**, and **26**.[97,98]

JIB-04 (compound **27**) is one of the first reported KDM4 inhibitor that seems to act with mixed-mechanism of action, as 2-OG mimic and in part as histone substrate competitors. JIB-04 inhibits selectively KDM4A-KDM4B, KDM4C not affecting the activity of other 2-OG dependent enzymes and reduces cancer growth in vitro and in vivo in JIB-04 treated mice.[99] Furthermore structural data and docking studies indicates that JIB-04 interacts with Lys241 and Tyr177 through hydrogen bonds, thus disrupting the binding of O$_2$ and the substrate in the binding site.[100]
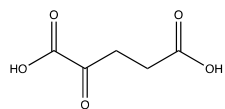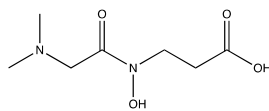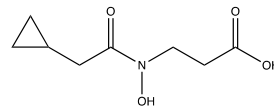
1 2-OG

2 NOG
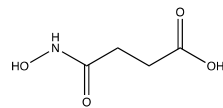
3 Hydroxamic acid
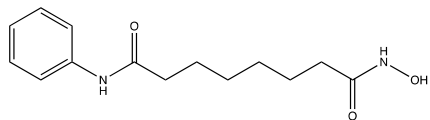
4 SAHA

5

6
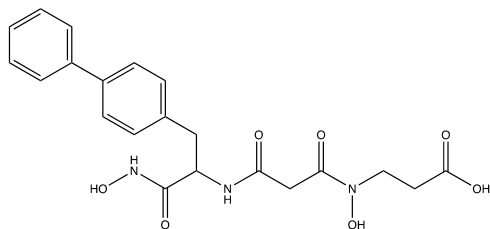
7
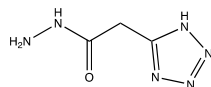
8
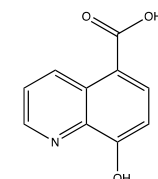
9 IOX1
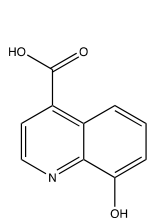
10
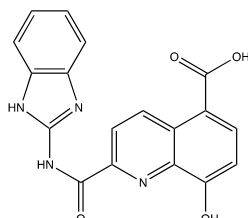
11
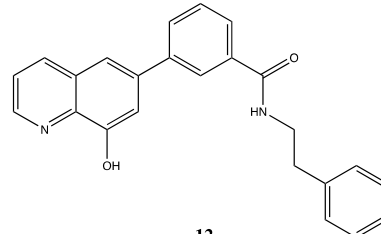
12
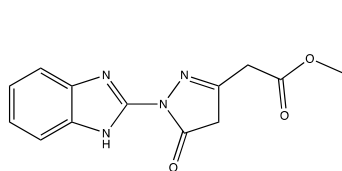
13
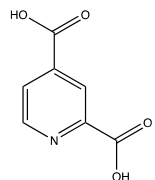
14 2,4-DPCA
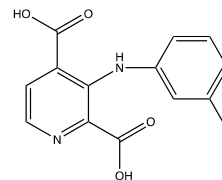
15
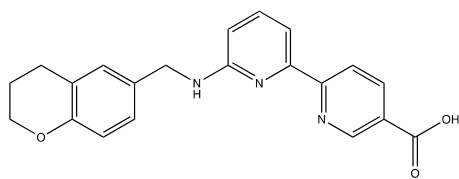
16
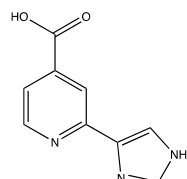
17
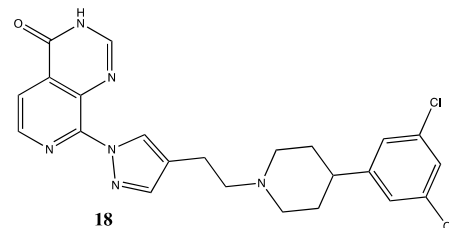
18
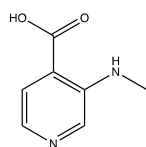
19
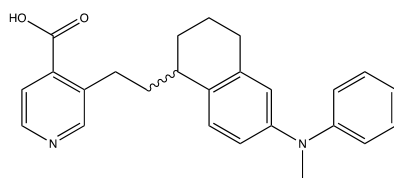
20 QC6352
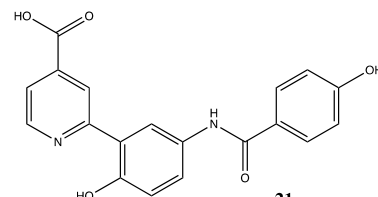
21

Figure 6. Structures of 2-OG cofactor mimics with metal-chelated moiety

A   **22** Disulfiram   **23** Ebselen

B   **24**   **25**   **26**

27 JIB-04
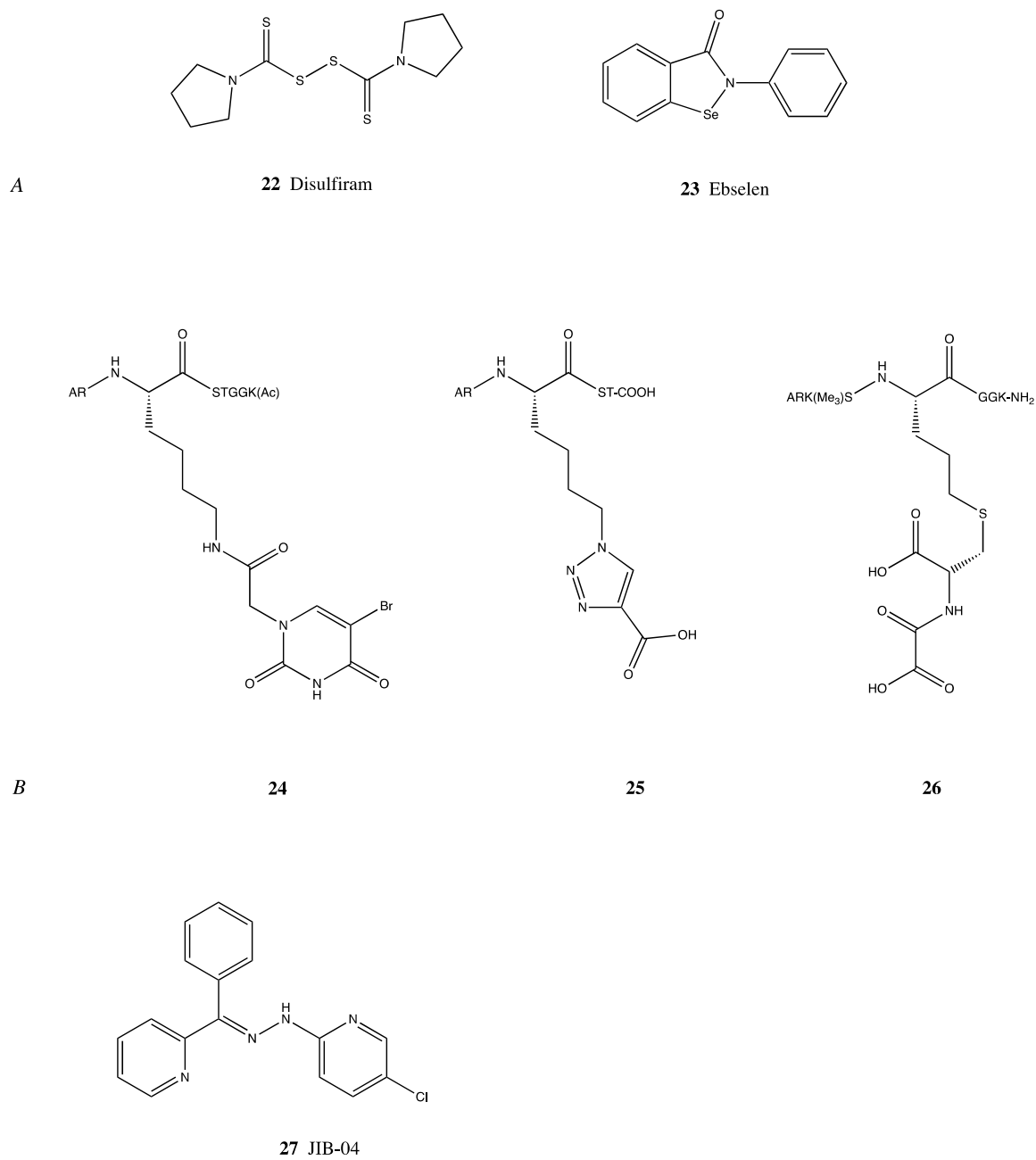
Figure 7. Structures of metal cofactor disruptors and histone substrate competitors.

Bibliography

1.   Shi, Y. *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941–953 (2004).

2.   Whetstine, J. R. *et al.* Reversal of Histone Lysine Trimethylation by the JMJD2 Family of Histone Demethylases. *Cell* **125**, 467–481 (2006).

3.   Fodor, B. D. *et al.* Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells. *Genes and Development* **20**, 1557–1562 (2006).

4.   Klose, R. J. *et al.* The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. *Nature* **442**, 312–316 (2006).

5.   Cloos, P. A. C. *et al.* The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3. *Nature* **442**, 307–311 (2006).

6.   Johansson, C. *et al.* The roles of Jumonji-type oxygenases in human disease. *Epigenomics* **6**, 89–120 (2014).

7.   Tsukada, Y. I. *et al.* Histone demethylation by a family of JmjC domain-containing proteins. *Nature* **439**, 811–816 (2006).

8.   Karytinos, A. *et al.* A novel mammalian flavin-dependent histone demethylase. *Journal of Biological Chemistry* **284**, 17775–17782 (2009).

9.   Burg, J. M., Gonzalez, J. J., Maksimchuk, K. R. & McCafferty, D. G. Lysine-Specific Demethylase 1A (KDM1A/LSD1): Product Recognition and Kinetic Analysis of Full-Length Histones. *Biochemistry* **55**, 1652–1662 (2016).

10.   Fitzpatrick, P. F. Oxidation of amines by flavoproteins. *Archives of Biochemistry and Biophysics* **493**, 13–25 (2010).

11.   Mohan, M. *et al.* The COMPASS Family of H3K4 Methylases in Drosophila. *Molecular and Cellular Biology* **31**, 4310–4318 (2011).

12.   Min, G. L. *et al.* Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science* **318**, 447–450 (2007).

13.   Huang, J. *et al.* p53 is regulated by the lysine demethylase LSD1. *Nature* **449**, 105–108 (2007).

14.   Wang, J. *et al.* The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. *Nature Genetics* **41**, 125–129 (2009).

15.   Kontaki, H. & Talianidis, I. Lysine Methylation Regulates E2F1-Induced Cell Death. *Molecular Cell* **39**, 152–160 (2010).

16.   Cho, H. S. *et al.* Demethylation of RB regulator MYPT1 by histone demethylase LSD1 promotes cell cycle progression in cancer cells. *Cancer Research* **71**, 655–660 (2011).

17.   Magliulo, D., Bernardi, R. & Messina, S. Lysine-specific demethylase 1A as a promising target in acute myeloid leukemia. *Frontiers in Oncology* **8**, (2018).

18.   Kooistra, S. M. & Helin, K. Post-translational modifications: Molecular mechanisms and potential functions of histone demethylases. *Nature Reviews Molecular Cell Biology* **13**, 297–311 (2012).

19.   Maes, T., Carceller, E., Salas, J., Ortega, A. & Buesa, C. Advances in the development of histone lysine demethylase inhibitors. *Current Opinion in Pharmacology* **23**, 52–60 (2015).

20.   Hausinger, R. P. *Fe(II)/α-ketoglutarate-dependent hydroxylases and related enzymes*. *Critical Reviews in Biochemistry and Molecular Biology* **39**, (2004).

21.   Katoh, M. & Katoh, M. Identification and characterization of JMJD2 family genes in silico. *Int J Oncol* **24**, 1623–1628 (2004).

22.   Berry, W. L. & Janknecht, R. KDM4/JMJD2 Histone demethylases: Epigenetic regulators in cancer cells. *Cancer Research* **73**, 2936–2942 (2013).

23.   Huang, Y., Fang, J., Bedford, M. T., Zhang, Y. & Xu, R. Domain of JMJD2A. *Science* **312**, 748–751 (2006).

24.   Kim, J. *et al.* Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Reports* **7**, 397–403 (2006).

25.   McDonough, M. A., Loenarz, C., Chowdhury, R., Clifton, I. J. & Schofield, C. J. Structural studies on human 2-oxoglutarate dependent oxygenases. *Current Opinion in Structural Biology* **20**, 659–672 (2010).

26.   Hegg, E. L. & Que, L. The 2-His-1-carboxylate facial triad - An emerging structural motif in mononuclear non-heme iron(II) enzymes. *European Journal of Biochemistry* **250**, 625–629 (1997).

27.   Klose, R. J., Kallin, E. M. & Zhang, Y. JmjC-domain-containing proteins. **7**, 715–727 (2006).

28.   Ng, S. S. *et al.* Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity. *Nature* **448**, 87–91 (2007).

29.   Chen, Z. *et al.* Structural basis of the recognition of a methylated histone tail by JMJD2A. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 10818–10823 (2007).

30.   Couture, J. F., Collazo, E., Ortiz-Tello, P. A., Brunzelle, J. S. & Trievel, R. C. Specificity and mechanism of JMJD2A, a trimethyllysine-specific histone demethylase. *Nature Structural and Molecular Biology* **14**, 689–695 (2007).

31.   Pilka, E. S., James, T. & Lisztwan, J. H. Structural definitions of Jumonji family demethylase selectivity. *Drug Discovery Today* **20**, 743–749 (2015).

32.   Krishnan, S. & Trievel, R. C. Structural and functional analysis of JMJD2D reveals molecular basis for site-specific demethylation among JMJD2 demethylases. *Structure* **21**, 98–108 (2013).

33.   Shi, Y. & Whetstine, J. R. Dynamic Regulation of Histone Lysine Methylation by Demethylases. *Molecular Cell* **25**, 1–14 (2007).

34.   Zhang, Z. & Pugh, B. F. High-resolution genome-wide mapping of the primary structure of chromatin. *Cell* **144**, 175–186 (2011).

35.   Henikoff, S. & Shilatifard, A. Histone modification: Cause or cog? *Trends in Genetics* **27**, 389–396 (2011).

36.   Tu, S. *et al.* Identification of histone demethylases in Saccharomyces cerevisiae. *Journal of Biological Chemistry* **282**, 14262–14271 (2007).

37.   Hillringhaus, L. *et al.* Structural and evolutionary basis for the dual substrate selectivity of human KDM4 histone demethylase family. *Journal of Biological Chemistry* **286**, 41616–41625 (2011).

38.   Fischle, W. *et al.* Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by polycomb and HP1 chromodomains. *Genes and Development* **17**, 1870–1881 (2003).

39.   Li, H. *et al.* Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* **442**, 91–95 (2006).

40.   Kooistra, S. M. & Helin, K. Molecular mechanisms and potential functions of histone demethylases. *Nature Reviews Molecular Cell Biology* **13**, 297–311 (2012).

41.   Shi, Y. Histone lysine demethylases: Emerging roles in development, physiology and disease. *Nature Reviews Genetics* **8**, 829–833 (2007).

42.   Greer, E. L. & Shi, Y. Histone methylation: A dynamic mark in health, disease and inheritance. *Nature Reviews Genetics* **13**, 343–357 (2012).

43.   Young, L. C. & Hendzel, M. J. The oncogenic potential of Jumonji D2 (JMJD2/KDM4) histone demethylase overexpression. *Biochemistry and Cell Biology* **91**, 369–377 (2013).

44.   Pfister, S. X. & Ashworth, A. Marked for death: Targeting epigenetic changes in cancer. *Nature Reviews Drug Discovery* **16**, 241–263 (2017).

45.   Black, J. C. *et al.* KDM4A lysine demethylase induces site-specific copy gain and rereplication of regions amplified in tumors. *Cell* **154**, 541–555 (2013).

46.   Peters, A. H. F. M. *et al.* Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**, 323–337 (2001).

47.   Liu, G. *et al.* Genomic amplification and oncogenic properties of the GASC1 histone demethylase gene in breast cancer. *Oncogene* **28**, 4491–4500 (2009).

48.   Yang, Z. Q. *et al.* Identification of a novel gene, GASC1, within an amplicon at 9p23-24 frequently detected in esophageal cancer cell lines. *Cancer Research* **60**, 4735–4739 (2000).

49.   Ehrbrecht, A. *et al.* Comprehensive genomic analysis of desmoplastic medulloblastomas: Identification of novel amplified genes and separate evaluation of the different histological components. *Journal of Pathology* **208**, 554–563 (2006).

50.   Duan, L. *et al.* KDM4/JMJD2 Histone Demethylase Inhibitors Block Prostate Tumor Growth by Suppressing the Expression of AR and BMYB-Regulated Genes. *Chemistry and Biology* **22**, 1185–1196 (2015).

51.   Fu, L. *et al.* HIF-1α-induced histone demethylase JMJD2B contributes to the malignant phenotype of colorectal cancer cells via an epigenetic mechanism. *Carcinogenesis* **33**, 1664–1673 (2012).

52.   Yang, J. *et al.* The histone demethylase JMJD2B is regulated by estrogen receptor α and hypoxia, and is a key mediator of estrogen induced growth. *Cancer Research* **70**, 6456–6466 (2010).

53.   Kawazu, M. *et al.* Histone demethylase JMJD2B functions as a co-factor of estrogen receptor in breast cancer proliferation and mammary gland development. *PLoS ONE* **6**, (2011).

54.   Shi, L. *et al.* Histone demethylase JMJD2B coordinates H3K4/H3K9 methylation and promotes hormonally responsive breast carcinogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 7541–7546 (2011).

55.   Yamane, K. *et al.* JHDM2A, a JmjC-Containing H3K9 Demethylase, Facilitates Transcription Activation by Androgen Receptor. *Cell* **125**, 483–495 (2006).

56.   Shin, S. & Janknecht, R. Activation of androgen receptor by histone demethylases JMJD2A and JMJD2D. *Biochemical and Biophysical Research Communications* **359**, 742–746 (2007).

57.   Kim, T. D. *et al.* Histone demethylase JMJD2A drives prostate tumorigenesis through transcription factor ETV1. *Journal of Clinical Investigation* **126**, 706–720 (2016).

58.   Kim, T. D., Shin, S. & Janknecht, R. ETS transcription factor ERG cooperates with histone demethylase KDM4A. *Oncology Reports* **35**, 3679–3688 (2016).

59.   Patani, N., Jiang, W. G., Newbold, R. F. & Mokbel, K. Histone-modifier gene expression profiles are associated with pathological and clinical outcomes in human breast cancer. *Anticancer Research* **31**, 4115–4126 (2011).

60.   Berry, W. L., Shin, S., Lightfoot, S. A. & Janknecht, R. Oncogenic features of the JMJD2A histone demethylase in breast cancer. *International Journal of Oncology* **41**, 1701–1706 (2012).

61.   Li, L. *et al.* JMJD2A-dependent silencing of Sp1 in advanced breast cancer promotes metastasis by downregulation of DIRAS3. *Breast Cancer Research and Treatment* **147**, 487–500 (2014).

62.   Li, L. L. *et al.* JMJD2A contributes to breast cancer progression through transcriptional repression of the tumor suppressor ARHI. *Breast Cancer Research* **16**, (2014).

63.   Li, B. X. *et al.* Effects of siRNA-mediated knockdown of jumonji domain containing 2A on proliferation, migration and invasion of the human breast cancer cell line MCF-7. *Experimental and Therapeutic Medicine* **4**, 755–761 (2012).

64.   Oncology, T. Lung cancer Lung cancer Lung cancer. *Conn's Current Therapy 2020* **2030**, 133–141 (2013).

65.   Xu, W., Jiang, K., Shen, M., Chen, Y. & Huang, H. Y. Jumonji domain containing 2A predicts prognosis and regulates cell growth in lung cancer depending on miR-150. *Oncology Reports* **35**, 352–358 (2016).

66.   Kim, T. D. *et al.* Pro-growth role of the JMJD2C histone demethylase in HCT-116 colon cancer cells and identification of curcuminoids as JMJD2 inhibitors. *American Journal of Translational Research* **6**, 236–247 (2014).

67.   Kim, T. D., Shin, S., Berry, W. L., Oh, S. & Janknecht, R. The JMJD2A demethylase regulates apoptosis and proliferation in colon cancer cells. *Journal of Cellular Biochemistry* **113**, 1368–1376 (2012).

68.   Li, Y., Wang, Y., Xie, Z. & Hu, H. JMJD2A facilitates growth and inhibits apoptosis of cervical cancer cells by downregulating tumor suppressor miR-491-5p. *Molecular Medicine Reports* **19**, 2489–2496 (2019).

69.   Neault, M., Mallette, F. A. & Richard, S. MiR-137 Modulates a Tumor Suppressor Network-Inducing Senescence in Pancreatic Cancer Cells. *Cell Reports* **14**, 1966–1978 (2016).

70.   Hu, C. E., Liu, Y. C., Zhang, H. D. & Huang, G. J. JMJD2A predicts prognosis and regulates cell growth in human gastric cancer. *Biochemical and Biophysical Research Communications* **449**, 1–7 (2014).

71.   Qiu, M. T. *et al.* KDM4B and KDM4A promote endometrial cancer progression by regulating androgen receptor, c-myc, and p27kip1. *Oncotarget* **6**, 31702–31720 (2015).

72.   Wang, H. L. *et al.* Expression and effects of JMJD2A histone demethylase in endometrial carcinoma. *Asian Pacific Journal of Cancer Prevention* **15**, 3051–3056 (2014).

73.   Wang, B. *et al.* Downregulation of KDM4A Suppresses the Survival of Glioma Cells by Promoting Autophagy. *Journal of Molecular Neuroscience* **60**, 137–144 (2016).

74.   Kim, S. *et al.* Deregulation of the histone lysine-specific demethylase 1 is involved in human hepatocellular carcinoma. *Biomolecules* **9**, 1–16 (2019).

75.   Chin, Y. & Han, S. KDM4 histone demethylase inhibitors for anti-cancer agents : a patent review. **3**, 1–10 (2014).

76. Lee, D. H. *et al.* Advances in histone demethylase KDM4 as cancer therapeutic targets. *FASEB Journal* **34**, 3461–3484 (2020).

77. Rose, N. R. *et al.* Inhibitor scaffolds for 2-oxoglutarate-dependent histone lysine demethylases. *Journal of Medicinal Chemistry* **51**, 7053–7056 (2008).

78. Rose, N. R. *et al.* Selective Inhibitors of the JMJD2 Histone Demethylases: Combined Nondenaturing Mass Spectrometric Screening and Crystallographic Approaches. *Journal of Medicinal Chemistry* **53**, 1810–1818 (2010).

79. Tavassoli, A., Hamilton, A. D. & Spring, D. R. Inhibition of 2- oxoglutarate dependent oxygenases. (2011).

80. Hamada, S. *et al.* Design, synthesis, enzyme-inhibitory activity, and effect on human cancer cells of a novel series of jumonji domain-containing protein 2 histone demethylase inhibitors. *Journal of Medicinal Chemistry* **53**, 5629–5638 (2010).

81. Rose, N. R. *et al.* Plant growth regulator daminozide is a selective inhibitor of human KDM2/7 histone demethylases. *Journal of Medicinal Chemistry* **55**, 6639–6643 (2012).

82. Morera, L. *et al.* 4-Biphenylalanine- and 3-Phenyltyrosine-Derived Hydroxamic Acids as Inhibitors of the JumonjiC-Domain-Containing Histone Demethylase KDM4A. *ChemMedChem* 2063–2083 (2016). doi:10.1002/cmdc.201600218

83. Małecki, P. H. *et al.* Structure-Based Screening of Tetrazolylhydrazide Inhibitors versus KDM4 Histone Demethylases. *ChemMedChem* **14**, 1828–1839 (2019).

84. King, O. N. F. *et al.* Quantitative high-throughput screening identifies 8-hydroxyquinolines as cell-active histone demethylase inhibitors. *PLoS ONE* **5**, (2010).

85. Hopkinson, R. J. *et al.* 5-Carboxy-8-hydroxyquinoline is a broad spectrum 2-oxoglutarate oxygenase inhibitor which causes iron translocation. *Chemical Science* **4**, 3110–3117 (2013).

86. Feng, T. *et al.* Novel 5-carboxy-8-HQ based histone demethylase JMJD2A inhibitors: Introduction of an additional carboxyl group at the C-2 position of quinoline. *European Journal of Medicinal Chemistry* **105**, 145–155 (2015).

87. Carter, D. M. *et al.* Identification of a Novel Benzimidazole Pyrazolone Scaffold That Inhibits KDM4 Lysine Demethylases and Reduces Proliferation of Prostate Cancer Cells. *SLAS Discovery* **22**, 801–812 (2017).

88.   Thalhammer, A. *et al.* Inhibition of the histone demethylase JMJD2E by 3-substituted pyridine 2,4-dicarboxylates. *Organic and Biomolecular Chemistry* **9**, 127–135 (2011).

89.   Roatsch, M. *et al.* Substituted 2-(2-aminopyrimidin-4-yl) pyridine-4-carboxylates as potent inhibitors of JumonjiC domain-containing histone demethylases. **8**, 1553–1571 (2016).

90.   England, K. S. *et al.* Optimisation of a triazolopyridine based histone demethylase inhibitor yields a potent and selective KDM2A (FBXL11) inhibitor. *MedChemComm* **5**, 1879–1886 (2014).

91.   Bavetsias, V. *et al.* 8-Substituted Pyrido[3,4-d]pyrimidin-4(3H)-one Derivatives As Potent, Cell Permeable, KDM4 (JMJD2) and KDM5 (JARID1) Histone Lysine Demethylase Inhibitors. *Journal of Medicinal Chemistry* **59**, 1388–1409 (2016).

92.   Le Bihan, Y. V. *et al.* C8-substituted pyrido[3,4-d]pyrimidin-4(3H)-ones: Studies towards the identification of potent, cell penetrant Jumonji C domain containing histone lysine demethylase 4 subfamily (KDM4)inhibitors, compound profiling in cell-based target engagement assays. *European Journal of Medicinal Chemistry* **177**, 316–337 (2019).

93.   Hopkins, A. L., Groom, C. R. & Alex, A. Ligand efficiency: A useful metric for lead selection. *Drug Discovery Today* **9**, 430–431 (2004).

94.   Chen, Y. K. *et al.* Design of KDM4 Inhibitors with Antiproliferative Effects in Cancer Models. *ACS Medicinal Chemistry Letters* **8**, 869–874 (2017).

95.   Korczynska, M. *et al.* Docking and Linking of Fragments to Discover Jumonji Histone Demethylase Inhibitors. *Journal of Medicinal Chemistry* **59**, 1580–1598 (2016).

96.   Sekirnik, R. *et al.* Inhibition of the histone lysine demethylase JMJD2A by ejection of structural Zn(ii). *Chemical Communications* 6376–6378 (2009). doi:10.1039/b916357c

97.   Lohse, B. *et al.* Targeting histone lysine demethylases by truncating the histone 3 tail to obtain selective substrate-based inhibitors. *Angewandte Chemie - International Edition* **50**, 9100–9103 (2011).

98.   Woon, E. C. Y. *et al.* Linking of 2-oxoglutarate and substrate binding sites enables potent and highly selective inhibition of JmjC histone demethylases. *Angewandte Chemie - International Edition* **51**, 1631–1634 (2012).

99.   Wang, L. *et al.* A small molecule modulates Jumonji histone demethylase activity and selectively inhibits cancer growth. *Nature Communications* **4**, (2013).

100.   Cascella, B., Lee, S. G., Singh, S., Jez, J. M. & Mirica, L. M. The small molecule JIB-04 disrupts O2 binding in the Fe-dependent histone demethylase KDM4A/JMJD2A. *Chemical Communications* **53**, 2174–2177 (2017).

CHAPTER 3

Design and synthesis of N-substituted-2-oxo-1,2-dihydroquinoline-4-carboxamide derivatives

Computer aided drug design is one of the most promising strategies to expedite the development of new therapeutics. In the last decade computational methodologies and high-performance computational resources were subjected to a rapid advancement, speeding-up the drug discovery process. To date, a multitude of *in silico* tools have been developed and have proven to be a useful method to quickly discover new bioactive hits specially in the early phases of drug discovery. In particular, due to the exploitation of 3D protein structure crystallography, structure-based virtual screening (SBVS) has emerged as powerful and convenient tool to identify novel hit and lead compounds. The virtual library screening is a valid approach that allows prompt access to millions of compounds which can be purchased and tested.

The recent availability of multiple crystal structures of KDM4A protein prompted us to employ a SBVS approach. Therefore in order to identify small molecules with potential inhibitory activity on KDM4A, a virtual screening campaign was performed with commercially available virtual library. Standard precision algorithm for molecular docking in GLIDE was used and, after visual inspection, the top-ranked 32 compounds were selected for biological assay. The primary assay was performed by Professor Altucci's team at Department of Precision Medicine of University Luigi Vanvitelli of Naples.

The assays results showed four potential hits: RIM141, RIM145, RIM148, RIM149. (Figure **1**)
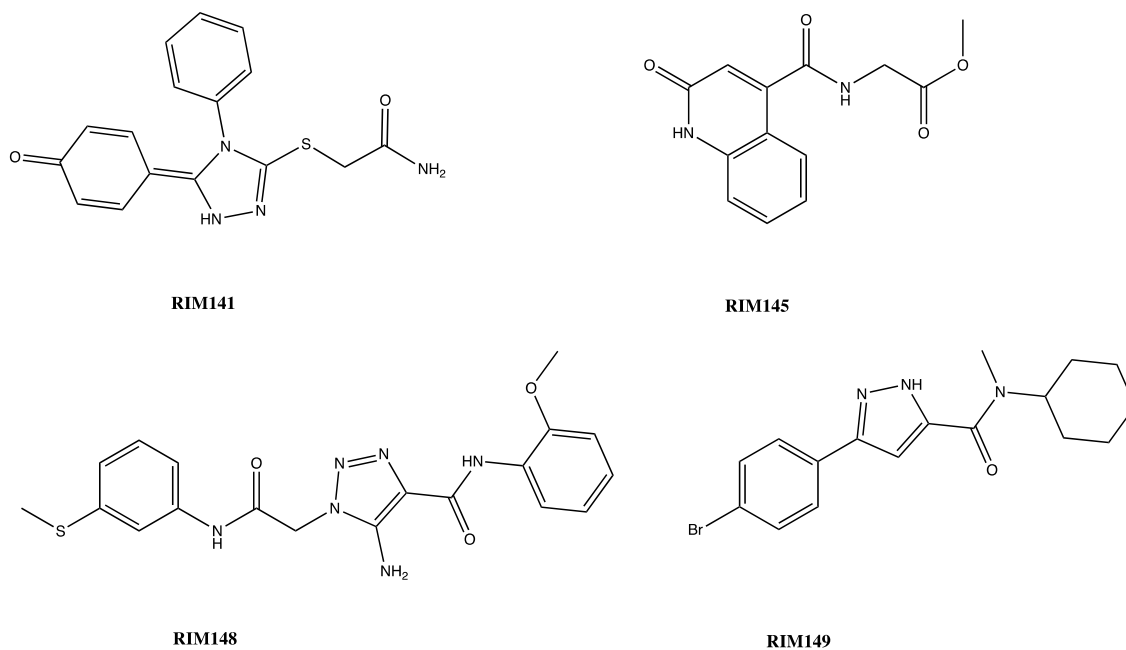
RIM141

RIM145

RIM148

RIM149

Figure 1. Structures of primary actives identified from the first virtual screening.

Among them, RIM145 (methyl (2-oxo-1,2-dihydroquinoline-4-carbonyl)glycinate) showed a relevant inhibitory activity with IC50 of 27,88 uM.(Figure **2**).
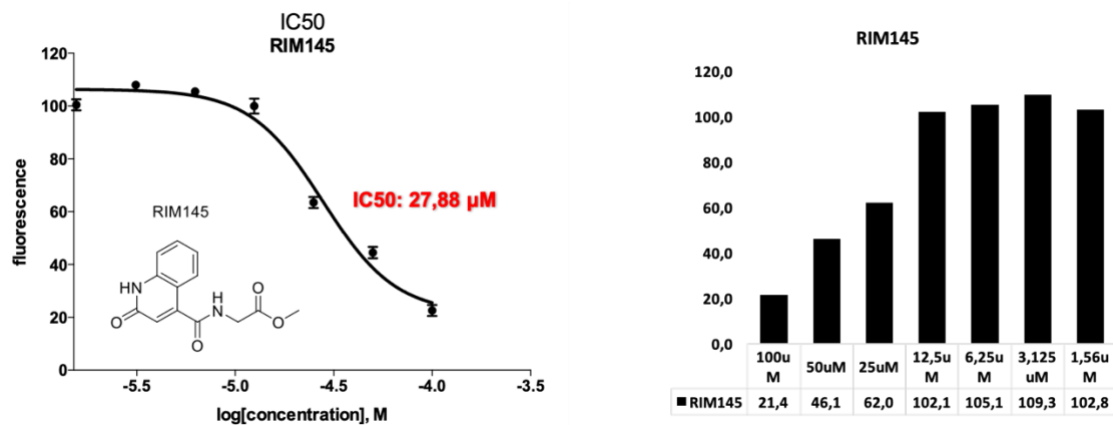

Figure 2. Dose-response curves of compound 145 inhibition with IC50 value.

3.1 Docking studies

In order to further investigate the RIM145 binding pose, docking studies were performed also in extra precision (XP) mode and compared with those of the previous virtual screening in standard precision (SP) mode. In the first virtual screening, the predicted binding mode of compound 145 revealed that the carbonyl group and the nitrogen atom of the quinolin-2(1*H*)-one core binds the Lys206 and Tyr132 residues, respectively, through a H-bond network, occupying the protein pocket in which is normally accommodated the natural cofactor 2-OG. Furthermore, the compound side chain establishes key interactions with the protein. Indeed, the cationic zinc ion interacts with the carbonyl group of the side chain amide, whereas the nitrogen atom of the side chain amide forms a hydrogen bond with residue Tyr177, which also establishes π-π stacking interaction with the pyridin-2(1*H*)-one scaffold (Figure **3A**).

In the binding pose retrieved from XP docking performance (Figure **3B**), the residue Asn198 replaces Lys206 in the hydrogen bond with the carbonyl group of the quinolin-2(1*H*)-one moiety and the residue Ser288 substitutes Tyr177 is acting as H-bond acceptor. The binding poses retrieved from XP and SP docking were superimposed, maintaining a similar orientation and interactions pattern (Figure **3C**). Overall, compound 145 binding mode, elucidated by this study, is consistent with the critical interactions for recognition reported in the literature.
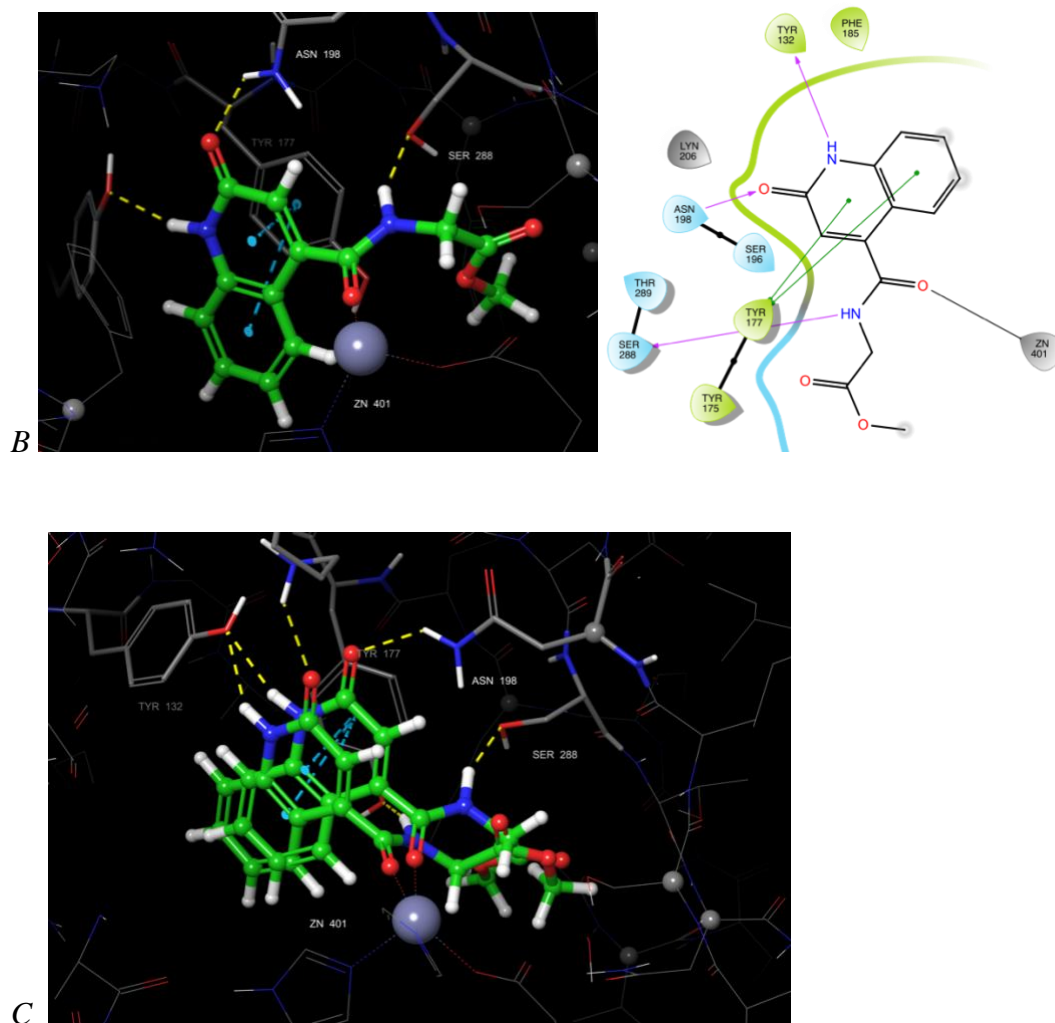

A
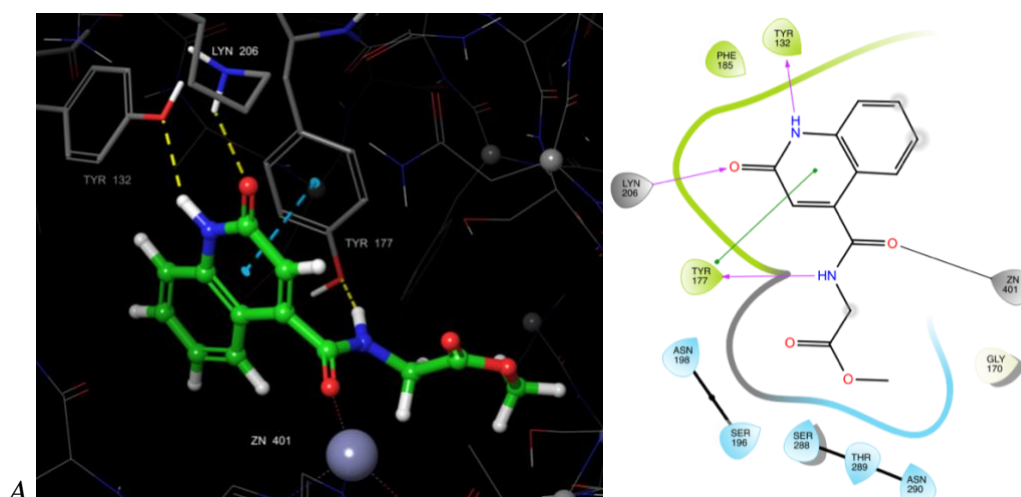
Figure 3. Binding poses compound 145, shown in green, in KDM4A protein (PDB ID: 5F32) A) Pose of compound RIM145 identified in the initial docking screens in SP mode and corresponding 2D ligand-protein interactions. B) Pose of the compound 145 retrieved from docking in XP mode and corresponding ligand-protein interactions. C) Superimposition of the two predicted compound 145 poses. The overlay shows that in both docking performance the molecule conformation is maintained.

3.2 Molecular dynamics

To investigate the complex stability and explore further potential key interactions, 50 ns molecular dynamics (MD) simulation was subsequently performed. The root-mean-square deviation (RMSD) plot graphs indicated that the system was stable. During the simulation small fluctuations were identified and only in the last part of the simulation medium fluctuations of around 1Å were reported. Analysing the simulation it can be deducted that compound 145 conformation was stable and the pattern interaction was comparable with the one retrieved from docking experiment. The H-bond networks between the pyridin-2(1*H*)-one moiety with Tyr132 and Lys206 residues were maintained consistently, whereas the nitrogen of the amide group in the compound side chain formed two H-bond with Ser288 and Asn198 via mediating water molecules. Furthermore, the simulation highlighted that the zinc ion is strongly bound in a bidentate manner. Indeed, it has been elucidated that both carbonyl groups in the compound side chain are involved in the chelation of the metal inside the binding pocket. The utility of using dynamic technics aims at exploring different ligand conformations inside the binding site, identifying all the ligand-protein interactions, otherwise hidden in a static molecular docking study. The structural information retrieved from MD simulation and XP docking provided guidance for hit optimization strategy.
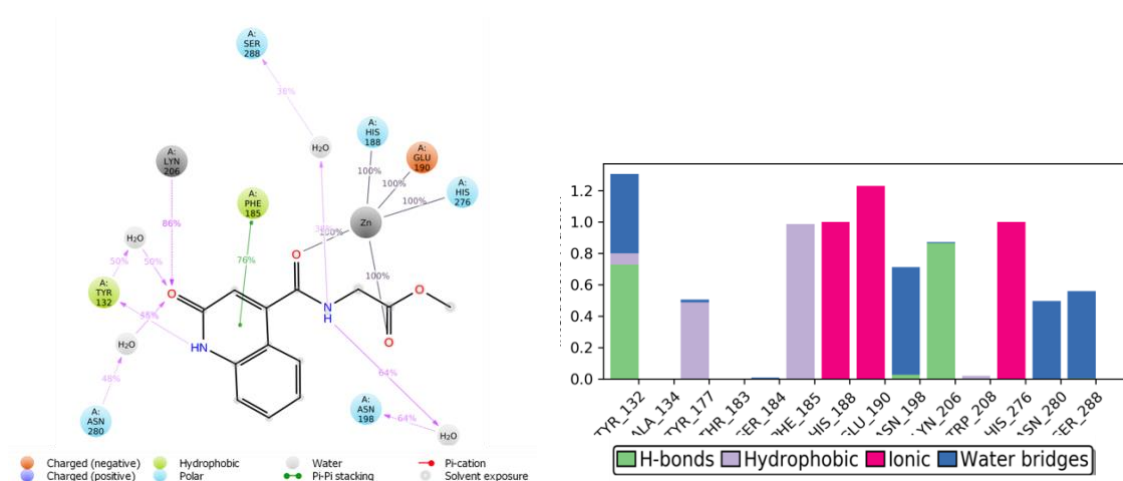

Figure 4. Ligand-Protein interaction diagram from MD simulation. A detailed diagram of ligand atom interaction with protein residues and division of protein-ligand contacts into four cluster.

3.3 Virtual combinational library

Recent studies about KDM4A inhibitors in the literature showed that the interactions with Tyr132 and Lys206 residues together with the metal ion binding are crucial features for protein inhibition. With the aim of validating and expanding the hit family and optimize compound RIM145 activity, it was decided to maintain the quinolin-2(1*H*)-one central core and substitute the 4-position side chain, in order to identify other potent zinc binder groups and to explore the pocket beside the interaction site of the natural cofactor. Therefore a virtual combinational library was created. In particular, it was decided to pursue two strategies. (Scheme **1**) Firstly, starting from the consideration that compound 145 side chain is a glycine-methyl ester, this moiety was strategically changed with other natural amino acid groups, in order to design new analogues and to generate a possible structure-activity relationship (SAR). Using Maestro tool "*Combinatorial library enumeration*", a combinatorial library was created, in which the quinolin-2(1*H*)-one central core was merged with different amino acids, generating all the possible combinations. Furthermore, a second strategy was based on the design of another combinatorial library in which, this time, the side chain was substituted with several zinc binding groups. The ZBG swapping is a common approach in molecular design to identify novel inhibitors that are able to interact with the zinc in the target protein. In addition, the ZBGs are characterized by different physicochemical properties, thus the ZBG replacement may be also useful to change and significantly improve the properties of the molecule.[1–3]
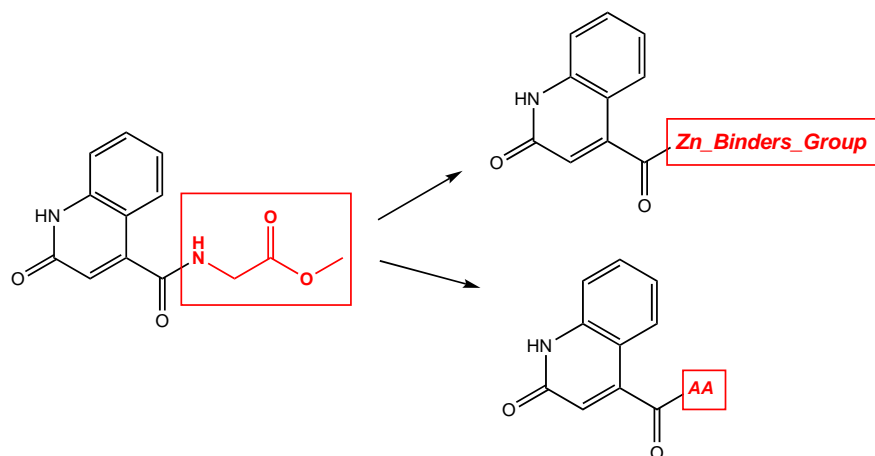
Scheme 1. Hit optimization strategy

The second library was created using the best-known zinc binding groups (ZBGs) in the literature. Kawai et al. published an analysis of zinc binding groups present in protein data bank (PDB). [4] In this study, using co-crystal structures in the PDB, the most common ZBGs were identified by calculating the distance between the zinc ion and the functional group of the ligand. Analysing all the metal-ligand interactions, it was found that the few heteroatoms, including oxygen, nitrogen and sulphur atoms are able to interact directly with zinc ion and that an ideal distance exists between the heteroatoms and the metal ion.[4,5]

In order to obtain an up-to-date list of ZBGs, it was done an exploration of the structures of most recent well-known inhibitors of proteins with a metal ion inside their binding site such as, HDAC or metalloproteinase inhibitors. [6–8]
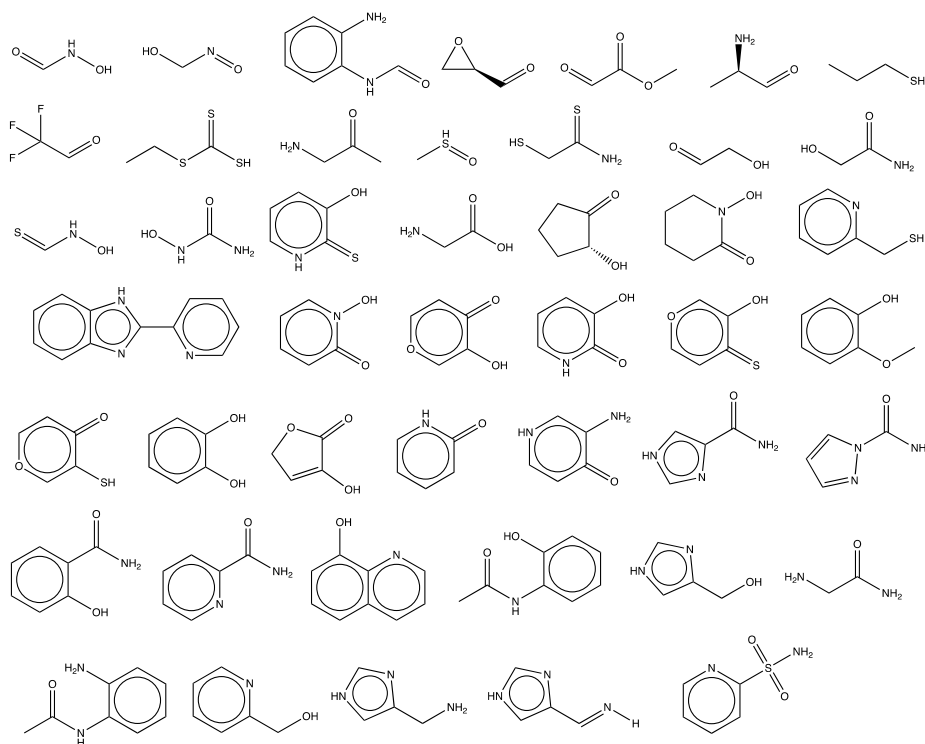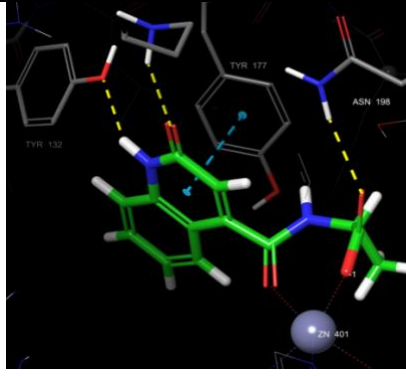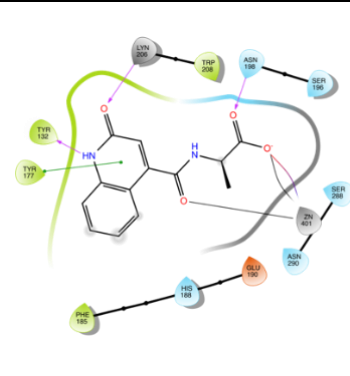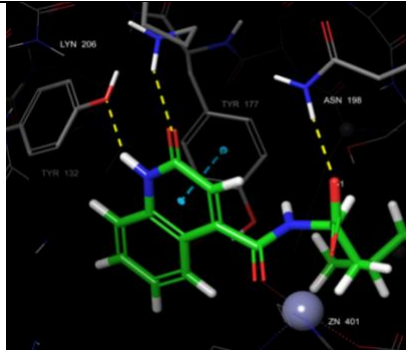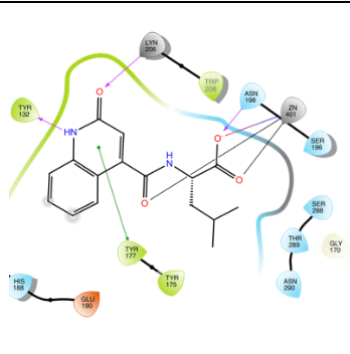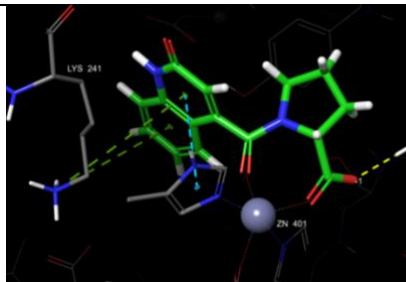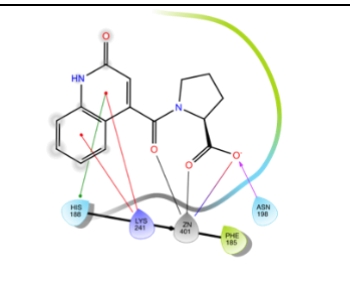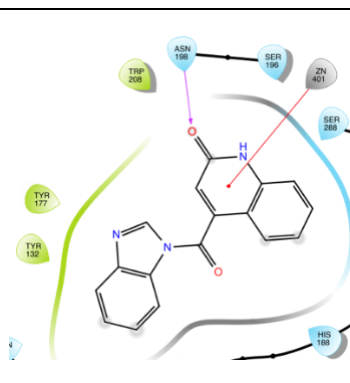

Figure 5. Example of frequently found ZBGs taken into account in our enumerate library

The enumerated virtual libraries created were screened on the docking model used in the first screening campaign and the best compounds, in terms of pose and docking score, were selected to be synthesized (Table **1**).

The best compounds, resulted from the virtual screening of the combinatorial library with the amino acid group, were those containing Alanine, Leucine and Proline side chains. Furthermore, it was decided to synthesize the new compounds using both enantiomers (-D and -L) either capped as the methyl ester of the amino acid or with the terminal carboxylic group, considering that the carboxylic group is a well-known strong metal binding group and could contribute to the binding with the zinc ion.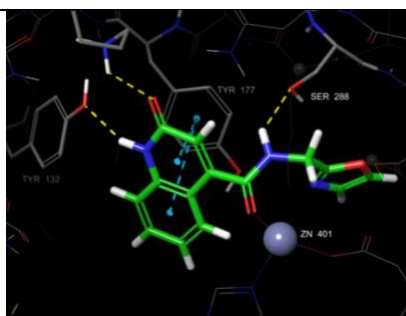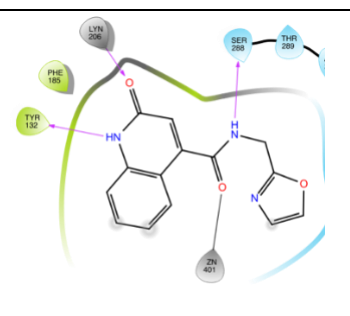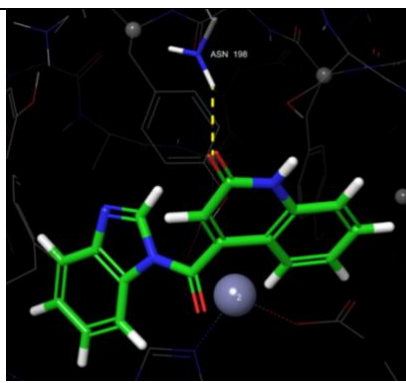 Indeed, as can be seen in Table 1, the best results in terms of docking score for the derivatives with the terminal carboxylic group supported the assumption that this group may contribute to improve the activity of these compounds. Whereas, the pyrazole, the oxazole and benzimidazole moiety were chosen as ZBGs replacing the methyl ester glycine in the starting compound side chain.

Title	3D pose	2D pose	Docking score	RO5
RIMJL1			-8.195	0
RIMJL2			-6.610	0

RIMJL3		-6.156	0
RIMJL4		-6.203	0
RIMJL7		-7.00	0
RIMJL8		-9.836	0
RIMJL9		-10.155	0

RIMJL10			-10.155	0
RIMJL11			-10.548	0
RIMJL12			-9.715	0
RIMJL33			-6.655	0
RIMJL36			-6.266	0
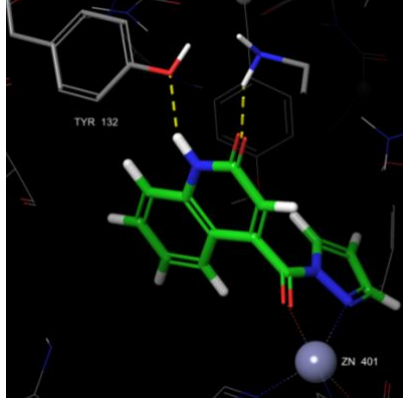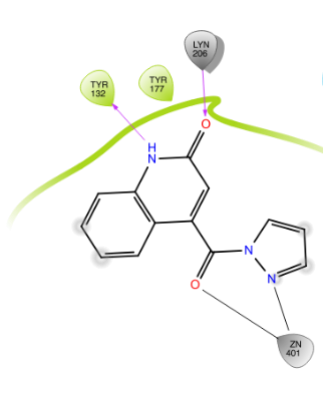
RIMJL41			-6.644	0

Table 1. Binding pose of compounds that were chosen to be synthesized (PDB ID: 5F32) and their corresponding ligand-protein 2D interactions. The docking score and Rule of five values are also indicated.

3.4 Chemistry

Synthesis of the *N*-substituted-2-oxo-1,2-dihydroquinoline-4-carboxamide derivatives **3a-f** was achieved by EDCI/HOBt-mediated coupling of commercially available 2-oxo-1,2-dihydroquinoline-4-carboxylic acid **1** with amino acid methyl ester derivatives **2 a-f** (Scheme **2**). This procedure has been frequently reported, for example for the synthesis of N-(4-(pyridine-2-yl)thiazol-2-yl)benzamides.[9] In addition to the new compounds designed, in order to validate the primary biological results and confirm its assigned chemical structure, compound 145 (compound 3a) was re-synthesized (Table**2**).
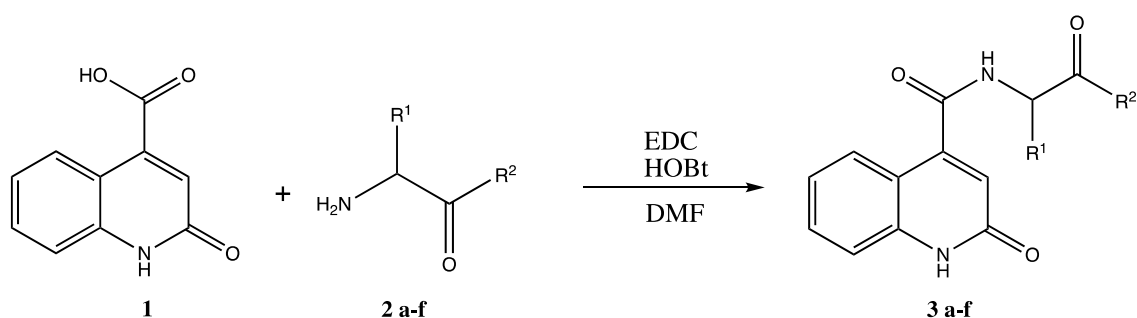
Scheme 2. Reagent and conditions: EDCI- HCl, 1-hydroxybenzotriazole, DMF, room temperature, 2h.
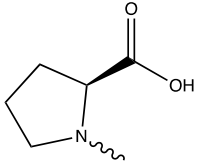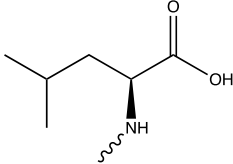
Compound	R¹	R²	Yield
3°	-H	-OCH$_3$	70%
3b	-L-CH$_3$	-OCH$_3$	60%
3c	-D-CH$_3$	-OCH$_3$	80%
3d		-OCH$_3$	60%
3e		-OCH$_3$	85%
3f	-H-	-NH$_2$	85%

Table 2. Structure of N-substituted-2-oxo-1,2-dihydroquinoline-4-carboxamide derivatives synthesized merging the 2-oxo-1,2-dihydroquinoline-4-carboxylic acid scaffold with the selected amino acid groups and the yield achieved.

Subsequently, the methyl (2-oxo-1,2-dihydroquinoline-4-carbonyl) derivatives **3a-e** were hydrolysed using an excess of lithium hydroxide, affording the corresponding compounds **4a-e** (Scheme **3**).

3 a-e → LiOH, THF/H$_2$O → **4 a-e**
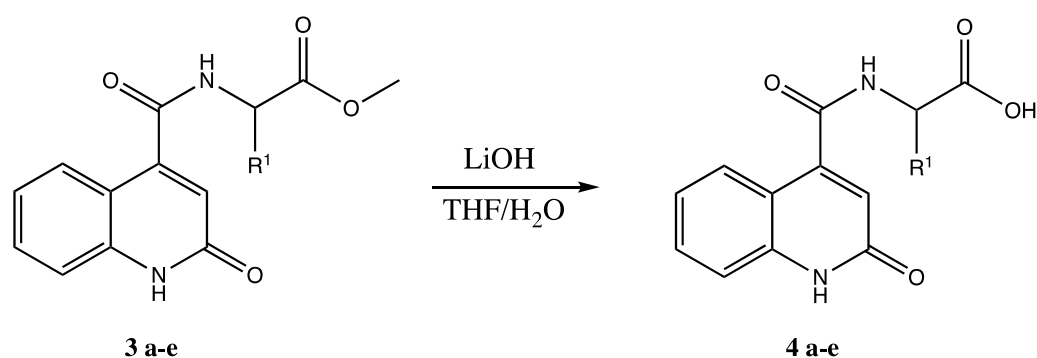
Scheme 3. Reagent and conditions: lithium hydroxide, THF, room temperature, 1-2h.
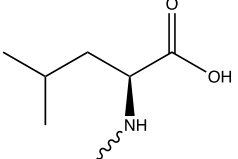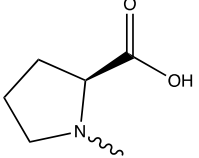
Compound	R^1	Yield
4°	-H	70%
4b	-L-CH$_3$	70%
4c	-D-CH$_3$	60%
4d		80%
4e		65%

Table 3. Structure of (2-oxo-1,2-dihydroquinoline-4-carbonyl) derivatives synthesized and the yield achieved.

Furthermore, the EDCI/HOBt-mediated coupling was also employed to afford compounds **5a-c** (Scheme **4**).
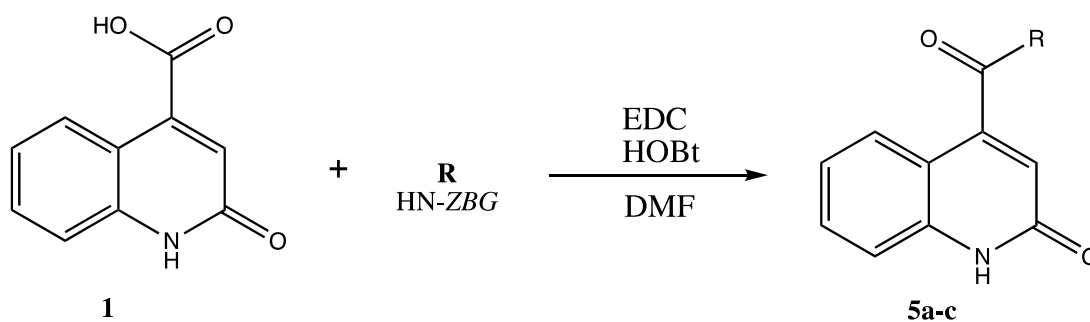
Scheme 4. Reagent and conditions: EDCI, 1-hydroxybenzotriazole, DMF, room temperature, 2h.
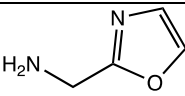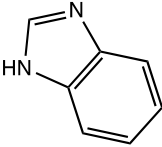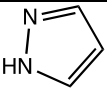
Compound	R	Yield
5°		40%
5b		30%
5c		40%

Table 4. Structure of N-substituted-2-oxo-1,2-dihydroquinoline-4-carboxamide derivatives synthesized merging the 2-oxo-1,2-dihydroquinoline-4-carboxylic acid scaffold with the selected zinc binders and the yield achieved.

Unfortunately, the synthesized compounds did not show good results in a enzymatic assay. All the compounds tested at 50μM reported an inhibition rate under 20%.

3.5 Material and methods

Molecular Docking workflow

- **Protein preparation:**

The crystal structure of the protein KDM4A in complex with a ligand with PDB ID: 5F32, and resolution of 2,05 Å, was chosen from the Protein Data Bank (PDB)[10] for the set-up of our model. The protein was prepared using *Protein Preparation Wizard*, [11] a tool from Schrödinger software adding bond orders and hydrogens to the crystal structure and deleting waters beyond 5.00 Å. *Epik* was used to generate het states at pH 7.4 ± 0.2 and then the protonation state of the protein and the ligand were optimized using *PropKa* at pH 7.4. [12,13]

- **Docking grid generation:**

The docking grid was generated using *Glide* software. [14] The scaling factor was set at 1.0 Å with a partial charge cut-off of 0.25, and the ligand was chosen to define the grid centroid.

- **Ligands preparation:**

LigPrep, a tool of MAESTRO software, was employed to prepare the ligands. OPLS3e was used as force field and *ionizer* generated all the possible states at pH: 7.0 ± 0.2. Desalt and generate tautomers were flagged on, the chiralities were retained and at most 32 conformers per ligand were generated.

- **Ligands docking:**

Molecular docking was carried out using *Glide* software[14] released by Schrödinger (release X). The simulation was performed both in standard precision (SP) and extra precision (XP) using OPLS3e as force field. [15] The ligands were considered as flexible, and Epik state penalties were included to docking score. The van der Waals (vdW) radii scaling factor was set as 0.8, with a partial charge cut off by 0.15. The searching algorithm on our model was tested using a cognate docking of the co-crystallized ligand, and we obtained an RMSD value of 0.578.

Molecular dynamics

The dynamics simulations were run for 50 ns using DESMOND [16] and OPLS3e as force field.[15,17]. The systems stability was evaluated checking the RMSD values of the aligned protein and ligand coordinate set calculated against the initial frame. The systems were built solvating the complex protein-ligand in orthorhombic boxes using the TIP3P water model. Then the box volumes were minimized and equilibrated at the temperature of 300 K and at 1.01325 bar pressure, and NPT ensemble was used. In order to analyze MD simulations, we used *simulation interactions diagram* that reported the graph of RMSD and RMSF values both for ligands and proteins, the protein secondary structure changes, ligand-protein interactions, including H-bonds, hydrophobic, ionic, and water-bridge contacts and the ligand torsion and properties.

Virtual combinatorial library

Firstly, the zinc binders group were prepared using *Reagent Preparation* a tool of MAESTRO. The ZBGs structures were designed and used to create functional groups. Then, the functional groups were ionize at pH: 7.0 ± 0.2 and tautomers were generated. Furthermore, 10 stereoisomers and 1 low-energy conformation for reagent were generated. Then, the combinatorial libraries were generated employing the tool *Combinatorial Library Enumeration,* incorporating the reagents list previously created. (Schrödinger Release 2020-4: CombiGlide, Schrödinger, LLC, New York, NY, 2020)

Chemistry

Commercially available starting materials, reagents, and anhydrous solvents were used as supplied. Flash column chromatography was performed using Merck silica gel 60 (0.025−0.04 mm). Thin layer chromatography was performed using Merck Millipore TLC silica gel 60 F254 sheets and visualized by UV (254 and 356 nm) iodine, and KmnO4.

General procedure 1: A solution of the starting material, 2-oxo-1,2-dihydroquinoline-4-carboxylic acid (1 equiv) in anhydrous DMF (3 ml) was cooled at 0°C, treated with EDCI hydrochloride (2 equiv) and 1-hydroxybenzotriazole hydrate (2 equiv) and was stirred for 30 minutes. Then, the amine (1,5 equiv) and DIPEA (3 equiv) were added and

the reaction mixture was stirred at room temperature for 1 or 2 h. After completion, the mixture was quenched with water, extracted with Ethyl acetate, dried with Na_2SO_4 and concentrated in vacuo.

Methyl (2-oxo-1,2-dihydroquinoline-4-carbonyl)glycinate (3a)*:*

According to general procedure 1, 2-oxo-1,2-dihydroquinoline-4-carboxylic acid (0.150 g, 0,79 mmol) and Glycine methyl ester hydrochloride (0.152 g, 1.20 mmol) were dissolved in DMF. After completion, the mixture was quenched with water (12 ml) extracted with ethyl acetate, dried with Na_2SO_4 and concentrated in vacuo. The crude obtained was washed with diethyl ether in order to provide a pure white solid. (0.145 g, 70%)

¹H NMR (200 MHz, DMSO) δ 11.99 (s, 1H), 9.24 (t, *J* = 5.9 Hz, 1H), 7.80 (d, *J* = 8.2 Hz, 1H), 7.62 – 7.50 (m, 1H), 7.35 (d, *J* = 8.2 Hz, 1H), 7.22 (t, *J* = 7.5 Hz, 1H), 6.51 (s, 1H), 4.06 (d, *J* = 5.9 Hz, 2H), 3.71 (s, 3H).

methyl (2-oxo-1,2-dihydroquinoline-4-carbonyl)-L-alaninate (3b)

According to general procedure 1, 2-oxo-1,2-dihydroquinoline-4-carboxylic acid (0.150 g, 0,79 mmol) and L-Alanine methyl ester hydrochloride (0.167 g, 1.20 mmol) were dissolved in DMF. After completion the mixture was quenched with water extracted with ethyl acetate, dried with Na_2SO_4 and concentrated in vacuo to provide the product as white solid. (0.132 g, 60%)

¹H NMR (200 MHz, DMSO) δ 11.98 (s, 1H), 9.21 (d, J = 6.8 Hz, 1H), 7.75 (d, J = 8.0, 1.3 Hz, 1H), 7.61 – 7.48 (m, 1H), 7.35 (d, J = 8.1 Hz, 1H), 7.29 – 7.16 (m, 1H), 6.50 (d, J = 1.4 Hz, 1H), 4.51 (d, J = 7.1 Hz, 1H), 3.71 (s, 3H), 1.38 (d, J = 7.3 Hz, 3H).

¹³C NMR (201 MHz, DMSO) δ 172.81, 165.97, 161.30, 145.94, 139.24, 131.01, 125.98, 122.17, 119.82, 116.23, 115.71, 52.18, 48.11, 16.54.

methyl (2-oxo-1,2-dihydroquinoline-4-carbonyl)-D-alaninate (3c)

According to general procedure 1, 2-oxo-1,2-dihydroquinoline-4-carboxylic acid (0.150 g, 0,79 mmol) and D-Alanine methyl ester hydrochloride (0.167 g, 1.20 mmol) were dissolved in DMF. After completion the mixture was diluted with water extracted with

ethyl acetate, dried with Na_2SO_4 and concentrated in vacuo to provide the product as white solid. (0.180 g, 80%)

¹H NMR (200 MHz, DMSO) δ 11.98 (s, 1H), 9.21 (d, J = 6.8 Hz, 1H), 7.75 (d, J = 7.9 Hz, 1H), 7.60 – 7.47 (m, 1H), 7.35 (d, J = 8.0 Hz, 1H), 7.22 (t, J = 7.4 Hz, 1H), 6.50 (s, 1H), 4.49 (t, J = 7.1 Hz, 1H), 3.71 (s, 3H), 1.38 (dd, J = 7.4, 3.0 Hz, 3H).

¹³C NMR (201 MHz, DMSO) δ 172.81, 165.97, 161.30, 145.94, 139.24, 131.01, 125.98, 122.17, 119.82, 116.23, 115.71, 52.18, 48.11, 16.54.

methyl (2-oxo-1,2-dihydroquinoline-4-carbonyl)-L-prolinate (3d)

According to general procedure 1, 2-oxo-1,2-dihydroquinoline-4-carboxylic acid (0.150 g, 0,79 mmol) and L-Proline methyl ester hydrochloride (0.2 g, 1.20 mmol) were dissolved in DMF. After completion the mixture was diluted with water extracted with ethyl acetate, dried with Na_2SO_4 and concentrated in vacuo. The crude product was purified by silica gel column chromatography eluting with 2%MeOH in DCM to provide the title compound as a white solid. (0.140 g, 60%)

¹H NMR (200 MHz, DMSO) δ 11.98 (s, 1H), 7.71 – 7.47 (m, 2H), 7.36 (d, J = 8.0 Hz, 1H), 7.23 (t, J = 7.7 Hz, 1H), 6.44 (s, 1H), 4.58 (dd, J = 8.9, 4.3 Hz, 1H), 3.74 (s, 3H), 2.31 (dd, J = 12.0, 7.6 Hz, 2H), 2.02 – 1.78 (m, 4H).

¹³C NMR (201 MHz, DMSO) δ 172.27, 164.96, 161.43, 146.82, 139.49, 131.38, 125.44, 122.57, 118.48, 115.98, 115.68, 58.27, 52.37, 48.47, 29.15, 24.57

methyl (2-oxo-1,2-dihydroquinoline-4-carbonyl)-D-leucinate (3e)

According to general procedure 1, 2-oxo-1,2-dihydroquinoline-4-carboxylic acid (0.150 g, 0,79 mmol) and L-Leucine methyl ester hydrochloride (0.218 g, 1.20 mmol) were dissolved in DMF. After completion the mixture was quenched with water extracted with ethyl acetate, dried with Na_2SO_4 and concentrated in vacuo. The crude product was purified by silica gel column chromatography eluting with 5%MeOH in DCM to provide the title compound as white powder. (0.215 g, 85%)

¹H NMR (200 MHz, DMSO) δ 12.00 (s, 1H), 9.18 (d, J = 7.4 Hz, 1H), 7.71 (dd, J = 8.2, 1.3 Hz, 1H), 7.65 – 7.47 (m, 1H), 7.37 (d, J = 8.1 Hz, 1H), 7.30 – 7.15 (m, 1H), 6.48 (d, J = 1.3 Hz, 1H), 4.58 – 4.42 (m, 1H), 3.71 (s, 3H), 1.69 (q, J = 8.4 Hz, 2H), 1.26 (d, J = 7.1 Hz, 1H), 0.93 (d, J = 6.0 Hz, 6H).

¹³C NMR (201 MHz, DMSO) δ 173.09, 166.67, 161.71, 146.36, 139.61, 131.46, 126.22, 122.65, 120.18, 116.59, 116.17, 52.60, 51.18, 24.95, 23.23, 21.57.

N-(2-amino-2-oxoethyl)-2-oxo-1,2-dihydroquinoline-4-carboxamide (3f)

According to general procedure 1, 2-oxo-1,2-dihydroquinoline-4-carboxylic acid (0.150 g, 0,79 mmol) and 2-aminoacetamide (0.133 g, 1.20 mmol) were dissolved in DMF. After completion the mixture was quenched with water and extracted with ethyl acetate, dried with Na₂SO₄ and concentrated in vacuo to provide the product as white solid. (0.172 g, 85%)

¹H NMR (200 MHz, DMSO) δ 11.95 (s, 1H), 8.94 (s, 1H), 8.00 – 7.76 (m, 1H), 7.57 – 7.08 (m, 5H), 6.62 (s, 1H), 3.83 (d, J = 6.0 Hz, 2H).

¹³C NMR (201 MHz, DMSO) δ 170.71, 166.59, 161.68, 146.48, 139.51, 131.14, 126.65, 122.29, 120.23, 116.57, 115.87, 42.32.

N-(oxazol-2-ylmethyl)-2-oxo-1,2-dihydroquinoline-4-carboxamide (5a)

According to general procedure 1, 2-oxo-1,2-dihydroquinoline-4-carboxylic acid (0.150 mg, 0,79 mmol) and oxazole-5-methanamine (0.12 mg, 1.20 mmol) were dissolved in DMF. After completion the mixture was diluted with water extracted with ethyl acetate, dried with Na₂SO₄ and concentrated in vacuo to provide the product as white solid. (0.08 g, 40%)

¹H NMR (800 MHz, DMSO) δ 11.94 (s, 1H), 9.30 (t, J = 5.7 Hz, 1H), 8.34 (s, 1H), 7.67 (dd, J = 8.1, 1.4 Hz, 1H), 7.54 (ddd, J = 8.4, 7.2, 1.4 Hz, 1H), 7.35 (dd, J = 8.3, 1.2 Hz, 1H), 7.20 (ddd, J = 8.2, 7.2, 1.2 Hz, 1H), 7.12 (s, 1H), 6.53 (d, J = 1.8 Hz, 1H), 4.57 (dd, J = 5.8, 1.0 Hz, 2H).

¹³C NMR (201 MHz, DMSO) δ 166.13, 161.47, 152.07, 149.19, 145.97, 139.53, 131.21, 126.05, 123.79, 122.39, 120.26, 116.33, 115.99, 33.90.

4-(1H-benzo[d]imidazole-1-carbonyl)quinolin-2(1H)-one (5b)

According to general procedure 1, 2-oxo-1,2-dihydroquinoline-4-carboxylic acid (0.150 g, 0,79 mmol) and benzimidazole (0.142 g, 1.20 mmol) were dissolved in DMF.. After completion the mixture was diluted with water extracted with ethyl acetate, dried with

Na$_2$SO$_4$ and concentrated in vacuo. The residue was crystallized with EtOH to provide the pure product as white solid. (0.068 g, 30%)

^{1}H NMR (800 MHz, DMSO) δ 12.16 (s, 1H), 8.65 (s, 1H), 8.29 (d, J = 8.0 Hz, 1H), 7.82 (d, J = 7.9 Hz, 1H), 7.59 (ddd, J = 8.4, 7.1, 1.3 Hz, 1H), 7.53 (td, J = 7.7, 1.2 Hz, 1H), 7.50 – 7.46 (m, 2H), 7.44 – 7.42 (m, 1H), 7.17 (ddd, J = 8.2, 7.1, 1.2 Hz, 1H), 7.05 (d, J = 1.9 Hz, 1H).

^{13}C NMR (201 MHz, DMSO) δ 164.61, 161.05, 144.85, 144.42, 142.55, 139.87, 131.68, 131.58, 126.13, 125.89, 125.35, 123.21, 122.76, 120.61, 116.26, 115.98, 115.80.

4-(1H-pyrazole-1-carbonyl)quinolin-2(1H)-one (5c)

According to general procedure 1, 2-oxo-1,2-dihydroquinoline-4-carboxylic acid (0.150 mg, 0,79 mmol) and pyrazole (81.69 mg, 1.20 mmol) were reacted in DMF. After completion the mixture was diluted with water extracted with ethyl acetate, dried with Na$_2$SO$_4$ and concentrated in vacuo to provide the product as white solid. (0.075 g, 40%)

^{1}H NMR (800 MHz, DMSO) δ 12.13 (s, 1H), 8.66 (d, *J* = 3.0 Hz, 1H), 7.91 (s, 1H), 7.56 (ddd, *J* = 8.3, 7.2, 1.4 Hz, 1H), 7.40 (dd, *J* = 8.4, 1.2 Hz, 1H), 7.25 (dd, *J* = 8.1, 1.4 Hz, 1H), 7.15 (ddd, *J* = 8.1, 7.1, 1.2 Hz, 1H), 6.82 (d, *J* = 1.9 Hz, 1H), 6.76 (dd, *J* = 3.0, 1.4 Hz, 1H).

^{13}C NMR (201 MHz, DMSO) δ 164.85, 160.91, 146.36, 143.57, 139.23, 131.46, 130.52, 125.33, 122.73, 122.15, 116.18, 116.15, 111.69.

General procedure 2:

Lithium hydroxide (3 equiv) was added to a solution of the SM (1 equiv) in 1:1 THF/H$_2$O. The reaction was stirred at RT for 1 h. After completion, the reaction mixture was neutralized with 6N aq HCl and extracted with EtOAc. The organic layer was dried with Na$_2$SO$_4$ and concentrated in vacuo.

(2-oxo-1,2-dihydroquinoline-4-carbonyl)glycine (4a)

According to general procedure 2, methyl (2-oxo-1,2-dihydroquinoline-4-carbonyl)glycinate (100 mg, 0.38 mmol) and LiOH (27.6 mg, 1.15 mmol) were reacted. Then, the reaction mixture was neutralized with 6N aq HCl and extracted with EtOAc.

The organic layer was dried with Na_2SO_4 and concentrated in vacuo, providing the title compound as white solid. (0.68 g, 70%)

¹H NMR (200 MHz, DMSO) δ 12.79 (s, 1H), 11.98 (s, 1H), 9.13 (t, J = 5.9 Hz, 1H), 7.82 (dd, J = 8.1, 1.4 Hz, 1H), 7.54 (ddd, J = 8.5, 7.1, 1.5 Hz, 1H), 7.44 – 7.30 (m, 1H), 7.20 (ddd, J = 8.2, 7.1, 1.2 Hz, 1H), 6.50 (s, 1H), 3.95 (d, J = 6.0 Hz, 2H).

¹³C NMR (201 MHz, DMSO) δ 170.87, 166.43, 161.34, 146.17, 139.19, 131.02, 126.11, 122.14, 119.65, 116.19, 115.68, 40.95.

(2-oxo-1,2-dihydroquinoline-4-carbonyl)-L-alanine (4b)

According to general procedure 2, methyl (2-oxo-1,2-dihydroquinoline-4-carbonyl)-L-alaninate (100 mg, 0.36 mmol) and LiOH (26.2 mg, 1.09 mmol) were reacted. Then, the reaction mixture was neutralized with 6N aq HCl and extracted with EtOAc. The organic layer was dried with Na_2SO_4 and concentrated in vacuo, providing the title compound as white solid. (0.66 g, 70%)

¹H NMR (200 MHz, DMSO) δ 12.73 (s, 1H), 11.96 (s, 1H), 9.07 (d, J = 7.0 Hz, 1H), 7.77 (dd, J = 8.1, 1.4 Hz, 1H), 7.59 – 7.46 (m, 1H), 7.44 – 7.28 (m, 1H), 7.25 – 7.09 (m, 1H), 6.49 (s, 1H), 4.40 (p, J = 7.1 Hz, 1H), 1.37 (d, J = 7.3 Hz, 3H).

¹³C NMR (201 MHz, DMSO) δ 173.81, 165.86, 161.31, 146.12, 139.22, 130.93, 126.09, 122.06, 119.78, 116.28, 115.65, 48.04, 16.65.

(2-oxo-1,2-dihydroquinoline-4-carbonyl)-D-alanine (4c)

According to general procedure 2, methyl (2-oxo-1,2-dihydroquinoline-4-carbonyl)-D-alaninate (100 mg, 0.36 mmol) and LiOH (26.2 mg, 1.09 mmol) were reacted. Then, the reaction mixture was neutralized with 6N aq HCl and extracted with EtOAc. The organic layer was dried with Na_2SO_4 and concentrated in vacuo, providing the title compound as white solid. (0.58 g, 60%)

¹H NMR (200 MHz, DMSO) δ 12.74 (s, 1H), 11.97 (s, 1H), 9.08 (d, J = 7.1 Hz, 1H), 7.78 (dd, J = 8.1, 1.4 Hz, 1H), 7.54 (ddd, J = 8.4, 7.0, 1.4 Hz, 1H), 7.35 (d, J = 8.1 Hz, 1H), 7.30 – 7.12 (m, 1H), 6.49 (s, 1H), 4.52 – 4.32 (m, 1H), 1.37 (d, J = 7.3 Hz, 3H).

¹³C NMR (201 MHz, DMSO) δ 172.97, 166.16, 161.53, 146.14, 139.30, 131.21, 126.09, 122.40, 119.84, 116.37, 115.88, 48.28, 16.64.

(2-oxo-1,2-dihydroquinoline-4-carbonyl)-D-leucine (4d)

According to general procedure 2, methyl (2-oxo-1,2-dihydroquinoline-4-carbonyl)-D-leucinate (100 mg, 0.28 mmol) and LiOH (20.36 mg, 0.85 mmol) were reacted. Then, the reaction mixture was neutralized with 6N aq HCl and extracted with EtOAc. The organic layer was dried with Na_2SO_4 and concentrated in vacuo, providing the title compound as white solid. (0.75 g, 80%)

¹H NMR (200 MHz, DMSO) δ 12.68 (s, 1H),11.98 (s, 1H), 9.20 (d, J = 7.4 Hz, 1H), 7.71 (dd, J = 8.2, 1.3 Hz, 1H), 7.65 – 7.47 (m, 1H), 7.37 (d, J = 8.1 Hz, 1H), 7.30 – 7.15 (m, 1H), 6.48 (d, J = 1.3 Hz, 1H), 4.58 – 4.42 (m, 1H), 1.69 (q, J = 8.4 Hz, 2H), 1.26 (d, J = 7.1 Hz, 1H), 0.93 (d, J = 6.0 Hz, 6H).

(2-oxo-1,2-dihydroquinoline-4-carbonyl)-L-proline (4e)

According to general procedure 2, methyl (2-oxo-1,2-dihydroquinoline-4-carbonyl)-L-prolinate (100 mg, 0.297 mmol) and LiOH (21.33 mg, 0.89 mmol) were reacted. Then, the reaction mixture was neutralized with 6N aq HCl and extracted with EtOAc. The organic layer was dried with Na_2SO_4 and concentrated in vacuo, providing the title compound as white solid. (0.60 g, 65%)

¹H NMR (200 MHz, DMSO) δ 12.20 (s, 1H), 11.98 (s, 1H), 7.71 – 7.47 (m, 2H), 7.36 (d, *J* = 8.0 Hz, 1H), 7.23 (t, *J* = 7.7 Hz, 1H), 6.44 (s, 1H), 4.58 (dd, *J* = 8.9, 4.3 Hz, 1H), 2.31 (dd, *J* = 12.0, 7.6 Hz, 2H), 2.02 – 1.78 (m, 4H).

Biochemical assay

Assay was performed using a homogeneous time-resolved fluorescence resonance energy transfer (TR-FRET, EpiGenetics https://www.cisbio.eu/media/asset/l/s/ls-tn-jmjd2c-histone-h3k9-mono-demethylation-me3-me2.pdf).

The principle of the assay relies on a europium cryptate -labelled antibody specific to the substrate's methylation state and a streptavidin acceptor molecule (SXL655) that binds the substrate biotinylated peptide. When the donor europium chelate is excited at 320 nm, energy in the form of FRET elicits fluorescence emission at 665 nm from the nearby streptavidin-tagged acceptor. HTRF signal is proportional to the concentration of demethylated H3(1-21) lysine 9 tri-methylated peptide.

The assay run in two steps: enzymatic reaction and detection. Optimal conditions for both steps were determined with manual experiments considering: time and temperature of incubation, substrate and DMSO concentration, purity of reagents, plates format.

Kdm4 enzyme concentration was set at 400 nm in buffer constituted by: Bovine Serum Albumin 0.01 % (BSA EMD Millipore 3458509), HEPES 50nM (EuroClone ECM0180D), Tween20 0,01 % (Sigma Aldrich P2287) and deionized water (Sigma Aldrich 38796). Enzyme substrates concentrations were 0.5 µM for α-Ketoglutaric Acid Monopotassium (2-OG, Sigma Aldrich K2000) and 300 nm for Histone H3(1-21) lysine 9 tri-methylated biotinylated peptide (AnaSpec AS-64360). 5µM Ammonium iron (II) sulfate hexa-hydrate (Sigma Aldrich 215406) and 1mM of Sodium ascorbate (Sigma Aldrich 1140) were added as co-factors.

16 µL of compounds were incubated for 1 h at 20°C with 8 µL of enzymatic solution and 16 µL of substrates solution in ½ area 96 well plates (ref. 6002290m, perkinHelmer). Dimethyl sulfoxide (DMSO Sigma-Aldrich D58791) final concentration was 1 %.

40 µl of antibodies solutions (Cisbio Bioassays H3K9 me2-Eu(K) Ab 61KB2KAE and Streptavidin XL-665 Cisbio Bioassays 610SAXLA, in detection buffer 62SDBRDD) were added to each well and incubated o. night at 20 °C. Fluorescence was measured with a TECAN Spark te-cool M-200 reader (Männedorf, Switzerland), excitation wavelength of 320 nm and emission wavelength of 620 and 665 nm. Automation was performed using Bravo (Agilent) platform. Robustness of the assay was proven by performing a Z' determination over 3 kinds of experiments: evaluation of max, min and mid signal by using 2,4-pyridinedicarboxylic acid monohydrate (2,4-PDCA, Sigma Aldrich P63395) at IC100 (10µM), IC50 (50 nM) and 0 µM; evaluation of dose -response curve with references compounds (PDCA and in-house); evaluation of single-dose experiments reproducibility. Dose response curves were performed by addressing 10 point of 1: 3 compound dilutions, departing from 100 µM.

Bibliography

1.   Wang, F. *et al.* Nє-acetyl lysine derivatives with zinc binding groups as novel HDAC inhibitors. *Royal Society Open Science* **6**, (2019).

2.   Goracci, L. *et al.* A Rational Approach for the Identification of Non-Hydroxamate HDAC6-Selective Inhibitors. *Scientific Reports* **6**, 1–12 (2016).

3.   Hsu, K. C. *et al.* Novel Class IIa-Selective Histone Deacetylase Inhibitors Discovered Using an in Silico Virtual Screening Approach. *Scientific Reports* **7**, 1–13 (2017).

4.   Kawai, K. & Nagata, N. Metal-ligand interactions: An analysis of zinc binding groups using the Protein Data Bank. *European Journal of Medicinal Chemistry* **51**, 271–276 (2012).

5.   Harding, M. M. Geometry of metal-ligand interactions in proteins. *Acta Crystallographica Section D: Biological Crystallography* **57**, 401–411 (2001).

6.   Zhang, L., Zhang, J., Jiang, Q., Zhang, L. & Song, W. Zinc binding groups for histone deacetylase inhibitors. *Journal of Enzyme Inhibition and Medicinal Chemistry* **33**, 714–721 (2018).

7.   Roche, J. & Bertrand, P. Inside HDACs with more selective HDAC inhibitors. *European Journal of Medicinal Chemistry* **121**, 451–483 (2016).

8.   Jacobsen, J. A., Major Jourden, J. L., Miller, M. T. & Cohen, S. M. To bind zinc or not to bind zinc: An examination of innovative approaches to improved metalloproteinase inhibition. *Biochimica et Biophysica Acta - Molecular Cell Research* **1803**, 72–94 (2010).

9.   Bavetsias, V. *et al.* 8-Substituted Pyrido[3,4-d]pyrimidin-4(3H)-one Derivatives As Potent, Cell Permeable, KDM4 (JMJD2) and KDM5 (JARID1) Histone Lysine Demethylase Inhibitors. *Journal of Medicinal Chemistry* **59**, 1388–1409 (2016).

10.   Berman, H. M. *et al.* The protein data bank. *Acta Crystallographica Section D: Biological Crystallography* **58**, 899–907 (2002).

11.   Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *Journal of Computer-Aided Molecular Design* **27**, 221–234 (2013).

12.   Shelley, J. C. *et al.* Epik: A software program for pKa prediction and protonation state generation for drug-like molecules. *Journal of Computer-Aided Molecular Design* **21**, 681–691 (2007).

13.   Greenwood, J. R., Calkins, D., Sullivan, A. P. & Shelley, J. C. Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. *Journal of Computer-Aided Molecular Design* **24**, 591–604 (2010).

14.   Friesner, R. A. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry* **47**, 1739–1749 (2004).

15.   Harder, E. *et al.* OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *Journal of Chemical Theory and Computation* **12**, 281–296 (2016).

16.   Bowers, K. J. *et al.* Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. *ACM/IEEE SC 2006 Conference (SC'06)* 43–43 (2007). doi:10.1109/sc.2006.54

17.   Shivakumar, D. *et al.* Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the opls force field. *Journal of Chemical Theory and Computation* **6**, 1509–1519 (2010).

CHAPTER 4

A Fragment-Based Approach to Identify and Optimize KDM4A inhibitors

4.1 Fragment-Based Drug-Discovery

In the last 20 years, Fragment-Based Drug Discovery (FBDD) has emerged as a successful approach for discovering new drug lead. To date, fragment-based approach has enabled the discovery of at least eighteen drugs that have entered the clinic, including PLX4032[1] and Zelboraf (vemurafenib)[2,3], is the latter being the first drug FDA-approved based on FBDD (http://practicalfragments.blogspot.com/).

The aim of FBDD is to design more efficient lead candidate starting from a smaller compound with less complex structure, that can be subsequently expanded or linked.[4] Over the past years the most common approach for hit identification has been high-throughput screening, with the possibility to screen several large diverse libraries[5] during the hit identification phase. The typical hit compound from HTS screening was a relatively big molecule with affinity in the low micromolar range with several non-optimised binding interactions. The rapid advance of technologies, especially in biophysical screening, has had a positive impact within important effort on trying to increase the success rate of the drug discovery process. Unfortunately, the success rate from HTS of combinatorial libraries has not been as high as expected.[6,7] As mentioned, the issue is due to the fact that the screening collections contains several highly complex compounds, with large and extensively hydrophobic groups, which are unsuitable to be used as viable hits and leads in the optimization process towards a clinical candidate.[8,9] Indeed, starting from a complex compound likely results in development of lead compounds that do not respect the drug-likeness parameters, as described by Lipinski rule[10] with consequent solubility and cell permeability issues. In general, HTS libraries are screened at 10 micromolar and sample a relatively large drug-like chemical space. That been said, considering the number of interactions during the recognition process, these compounds make several unfavourable interactions with the protein thus decreasing the possibility to promote a favourable recognition event and often delivering hits not suitable to enter lead optimization.[11,12]

Moreover, the HTS process is really expensive and time consuming with typical screening campaign in large pharma costing around 1 million euro and lasting about 6-12 months between the assay validation and library screening of 1-5 million molecules. According to the fragment-based approach, the discovery begins identifying weak binding ligands with a relatively simple structure and low molecular weight, that therefore are called fragments.[13] One of the many advantages in fragment-based screening is that the chemical space of fragment-size compound is smaller and then can be searched more efficiently.[14,15] Certainly, FBDD delivers low affinity (e.g., 100μM to 1 mM) fragment hit toward the target compared to a hit from HTS (< 10μM). In FBDD the affinity has not been considered as first aspect during drug discovery, because this parameter if often associated to molecular size. Indeed, the increment of affinity during optimization phases is translated in increase of molecular weight, typically by introducing lipophilic groups, which contribute to hydrophobic effect without the need of specific good interactions with the target. [8,16,17] It is now recognised that a suitable clinical candidate has to be characterized by favourable physicochemical properties.[18] In fact, to date, the drug discovery process is strongly influenced by the Lipinski's "Rule-of-five" , that is based on a statistical analysis of the physicochemical properties of orally bioavailable drugs in clinical Phase II. The rule predicts that poor absorption or permeation is more likely to occur when there are more than 5 H-bond donors, 10 H-bond acceptors, the molecular weight (MW) is greater than 500 and the calculated Log P (ClogP) is greater than 5.[10,13] Consistently, it was also shown that (ADMET) profile is negatively influenced by increasing MW and Log P.[13]

In order to avoid the bias discussed above, useful parameters that can be taken into account are the ligand efficiency scores. The concept of Ligand efficiency (LE) was developed by Hopkins et al., and correlates the binding affinity with the number of heavy atoms (non-hydrogen atoms), in order to normalize the potency and MW of a compound.[19]

$$LE = \frac{-\Delta G}{HA}$$

(-ΔG is the free energy of binding and HA is the number of non-hydrogen atom of the ligand)

The lipophilicity, as discussed above, plays a key role in achieving good ADMET properties, which in turn, are crucial in determining the fate of the compounds in vivo. Hence an additional parameter to consider is the ligand-lipophilicity efficiency (LLE) that evaluates the compound binding affinity related to its lipophilicity.[20]

$$LLE = pIC_{50} - clogP$$

Highly lipophilic compounds, tend to decrease their selectivity and therefore LLE parameter can be used during the design phase to synthesise molecules with higher potential to be selective compounds.

Overall, LE and LLE have proven to be useful tools in the selection of lead compounds and in the optimization process, and have become the most accepted metric for fragment hit selection. The starting fragment should show a high LE (≥0.3) considering that during optimization phases this value may decrease.

Whereas a clinical candidate should follow the Rule of five, a specifically rule was developed by researchers at Astex, for the fragments, namely the Rule of three.[21–23] Based on the Rule of three a suitable fragment should be characterized by ≤ 3 H-bond donors, ≤ 3 H-bond acceptors, the molecular weight (MW) < 300 Da, predicted Log P (ClogP) ≤ 3 and polar surface area (PSA) ≤60Å^2. The comparison of fragment-like and drug-like properties are summarized in table **1**.

Type of compound	*Fragment-like*	*Drug-like*
Rule	Rule of three	Rule of five
MW	<300	≤500
*c*LogP	≤3	≤5
H-bond donors	≤3	≤5
H-bond-acceptors	≤3	≤10
Rotatable bonds	≤3	≤10
PSA	≤60	≤110

Table 1. Comparison of fragment-like and drug-like compounds

The fragment hits are generally identified by biophysical screening technologies such as NMR, X-ray or Surface Plasmon Resonance and are subsequently optimized by a rational approach carefully adding functionalities to increase binding affinity while maintaining high LE in order to create a lead compound with improved potency, selectivity, physicochemical properties and novelty. The main followed approaches in fragments optimization are merging, linking and growing.

In fragment merging approach, the common structural features from fragments, substructure or known ligand were combined in order to create a hybrid molecule.[24–26]

The second approach involves the identification of fragments that bind adjacent sites in the protein, that are subsequently linked together to generate new chemical entities.[27,28] Fragment growing method includes the fragment building using the structure coordinates of the fragment crystallized into protein target. In general, the fragment is grown adding hydrogen bond donors or acceptors in a particular site on the fragment in order to form possible electrostatic interactions with the protein.[29] The most appropriate fragments substituents are chosen based on the visual inspection of the local environment around the binding pocket.[30]
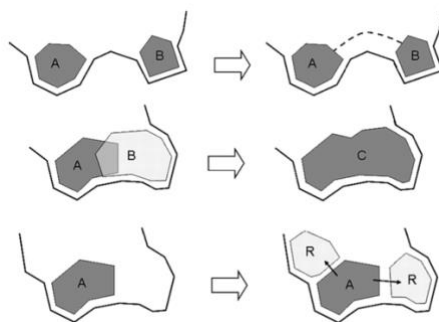

Figure1. Fragment development strategies. Top: fragment linking, where fragments binding in adjacent regions of the binding site are linked to create a larger, more potent compound. Middle: fragment fusion, where fragments in overlapping space are amalgamated to form a larger more potent compound. Bottom: fragment growth, where rational design is used to grow the core fragment into adjacent regions of the binding site. The picture was replied from "Fragment-Based Drug Discovery: A Practical Approach" edited by Edward R. Zartler and Michael J. Shapiro and published by Wiley.

FBDD was employed to discover several new compounds, including KDM4s inhibitors. Celgene's story is an example of a successful fragment growing approach.[31] Starting from a screening of known inhibitors of KDM4, compound 1 was identified (Figure **2**). Then, the growing of this fragment with the inclusion of tetrahydronaphthalene lead to a more potent compounds 2 and subsequently 3. Compound 3 was potent in the enzymatic assay but due to its carboxylic group, it showed a low cLogD and then poor cellular permeability. Therefore, it was decided to increase the lipophilicity of the molecule and was obtained the lead compound QC6352, a potent and active lead in breast and colon cancer PDX models. This study also highlighted that it is correct to avoid high lipophilicity of the molecule but also polar functional group could impact pharmacokinetic (PK) properties.
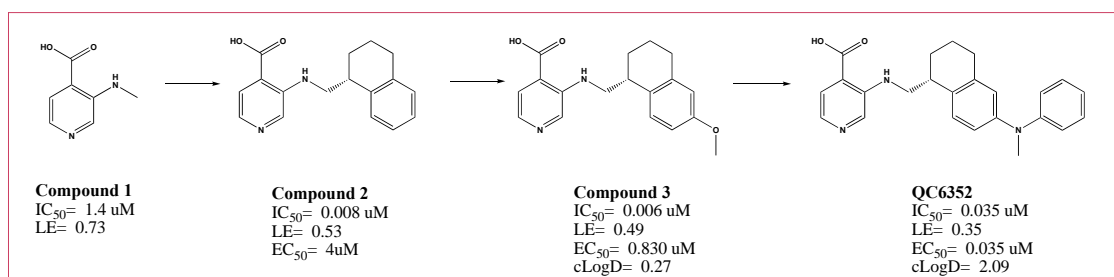

Compound 1
IC$_{50}$= 1.4 uM
LE= 0.73

Compound 2
IC$_{50}$= 0.008 uM
LE= 0.53
EC$_{50}$= 4uM

Compound 3
IC$_{50}$= 0.006 uM
LE= 0.49
EC$_{50}$= 0.830 uM
cLogD= 0.27

QC6352
IC$_{50}$= 0.035 uM
LE= 0.35
EC$_{50}$= 0.035 uM
cLogD= 2.09

Figure 2. Fragment elaboration by fragment growing.

In order to discover a potent and selective compound active against KDM4A, it was decided to start the lead discovery process adopting the fragment-based approach.

4.2 Virtual fragment library

In order to perform an *in silico* screening, a virtual fragment library was created. In the literature, there are several variants over the definition of fragments properties. In general, the most relevant fragment property profile is related to the rule of three. During the generation of fragment library, such rules for physicochemical properties can be applied to a list of commercially available compounds, in order to select and purchase the most suitable fragments. The fragments diversity is an important requisite, especially since chemical space for fragment compounds is expected to be significantly smaller than the chemical space of drug-like molecules. In general, the fragments show weak potency, often in the mM range, but during SAR development, increasing their MW the binding affinity can greatly increase.[32] Furthermore, it should be considered that chemical novelty is not an important condition for the inclusion in the fragments library and during selection of fragment hits considering the scaffold will be subjected to important modifications during hit to lead optimization. Since it has been noted that some marketed drugs shared a substructure with other marketed drugs, it has become common to include substructure drugs in the fragment libraries.[33] Indeed, unique composition for patent protection could be achieved during fragment elaboration and hit-to-lead optimization process.

4.2.1 Compound selection

In this study, compounds were selected from multiple virtual commercially available libraries, including Prestwick, Life Chemical, BioAscent, ChemDiv, Asinex, chemBridge, Discovery Chemistry and Enamine. All the compounds in the libraries were prepared in terms of conformation and protonation state using LigPrep.

4.2.2 Computational Filters

All the compounds in the libraries, after preparation using LigPrep, were further filtered off of all the Pan-assay interference compounds (PAINS) and reos compounds, to remove all the species that may lead to false positive due to the intrinsic reactivity or may cause assay interferences.[34] These chemical compounds furthermore tend to react non-specifically with multiple biological targets rather than specifically affecting one desired target. A molecular weight filter with a value in the range of $140 \leq MW \leq 300$ was applied

and only the compounds with at least one ring system were considered. In addition, all the compounds properties were evaluated using quikprop, a Schrodinger tool, and all the molecules that did not comply the rule of three parameters were discarded. Finally, a unique virtual fragment library with $\cong$ 250000 compounds was obtained.

4.2.3 Fragment Diversity

Diversity metric is an important parameter that should be used to prevent the inclusion of close analogues in the library. In order to evaluate the structural fragment variability in our fragment library, according to linear fingerprint, a similarity matrix was created. Using a sample of fragments, equivalent to 10% of the total, the average value was calculated, showing a sufficient structural difference in the fragment library.

4.3 Virtual screening campaign

The aim of this study was to identify fragment hits active as inhibitor on KDM4A, upon which apply an optimization strategy to improve potency and physicochemical properties. The virtual fragment library of 250000 compounds, was employed to perform a virtual screening campaign characterized by docking and pharmacophore approaches. The compounds were ranked by their docking and pharmacophore score and the best compounds, after a visual inspection, were selected to be purchased.

4.3.1 Pharmacophore screening

The fragment pharmacophore screening was performed using LigandScout software. The pharmacophore model was generated using the crystal structure of KDM4A protein in complex with an active inhibitor developed by Celgene QC6352 (PDB ID: 5VMP). The pharmacophore was characterized by five different features, and all of them were retrieved in the central core of molecule, constituted by carboxy-pyridine moiety, indicating that it was a preferred scaffold playing a crucial role in the binding activity (Figure **3**). The pharmacophore was characterized by two H-bond acceptor features, one positional feature in the aromatic pyridine ring, and a metal chelator feature formed by the nitrogen atom of the pyridine moiety. The screening was performed with 0 and 1 omitted features and the compounds with the best fit-pharmacophore score were selected.

These features are consistent with those retrieved in multiple crystal structure of KDM4A in complex with different inhibitors and also reported as important features in literature.
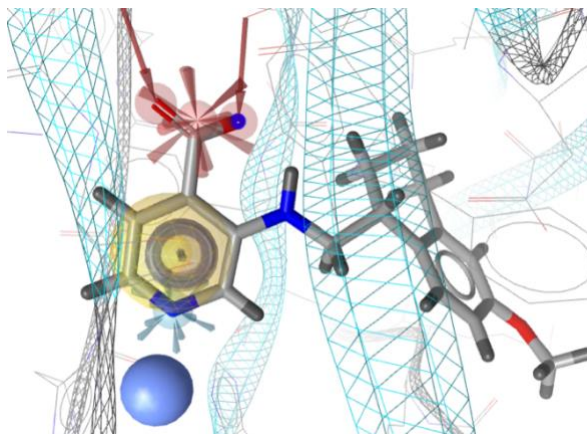

Figure 3. QC6352 pharmacophore. Red arrows: hydrogen-bond acceptor, yellow spheres: hydrophobic features, blue rings: aromatic features; red spikes: negative ionizable feature, blue spike: metal interaction. (PDB ID: 5VMP)

4.3.2 Docking screening

In the light of the unsuitable results of the previous virtual screening campaign, a new model for docking was developed. The crystal structure of KDM4A protein in complex with Celgene inhibitor (PDB ID: 5VGI) was used to create the docking grid. The pharmacophore features retrieved were used as constraints, including Tyr132 H-bond acceptor, Lys206 H-bond acceptor, positional constraint, and metal chelator binding. Firstly, the docking screening were performed in High-throughput virtual screening (HTVS) mode and only the first 10000 best ranked compounds were retrieved in order to decrease the number of molecules to screen with more precision. The following screening was carried out in standard precision (SP) mode and only compounds that matched at least 2 constraint features were retrieved.

Finally, a consensus approach by merging docking and pharmacophore screening results, lead to the identification of 280 compounds that were purchased to be tested in a primary biological assay.

4.4 Fragment screening results and hit series validation

The selected compounds from the virtual fragments screening were tested by conducting a biochemical assay to evaluate the KDM4A inhibition activity. The Celgene compound and other four known inhibitors were used as positive controls and the compounds were tested at 25 μM. From this primary assay eight fragments (Figure **4**) resulted active, showing a protein activity inhibition ≥ 30%
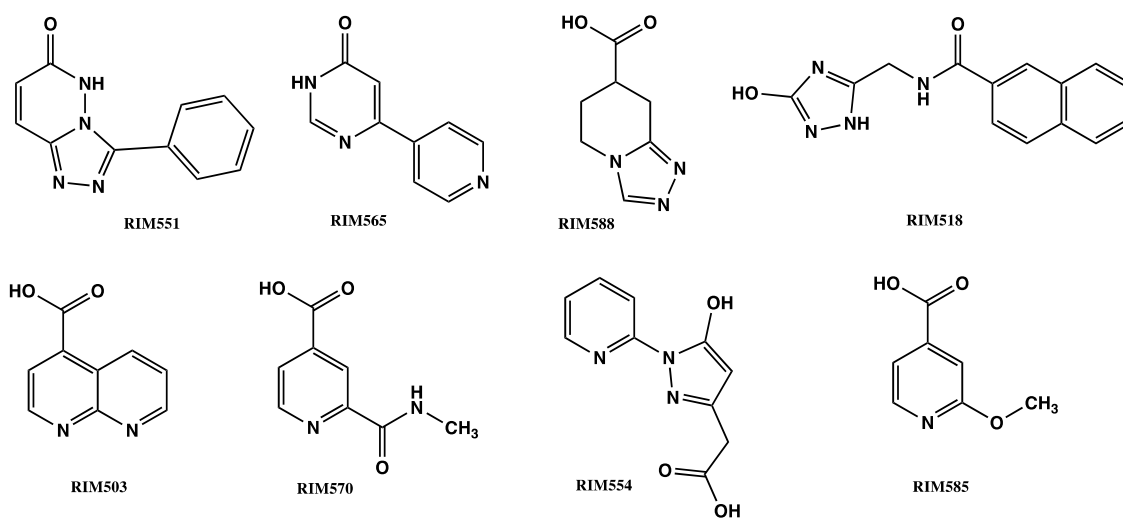

Figure 4. Structure of primary fragment hits result from the biochemical assay.

4.4.1 Series validation and analogues selection

The first step to validate the hits series was the selection of close analogues of preliminary active compounds. For each compound, a structural search was run using PubChem databases. Since analogues should be closely similar to their corresponding starting fragment, for the structural research using PubChem fingerprint, a Tanimoto similarity coefficient of 0.9 was applied and all the compounds retrieved were filtered by commercial availability. Then, the analogues were docked and, based on their poses, 46 compounds were selected and purchased.

4.4.2 Orthogonal assay: Biophysical assay results

In general, during the drug discovery process, an orthogonal assay is necessary to validate the primary screening results, while in FBDD an available biophysical assay becomes really crucial.[35–37] Indeed, due to the higher sensibility, the biophysical technics are the preferred screening method that have proved to be decisive for the success of this approach. Fragment molecules are characterized by low molecular weight and often show weak interactions. Therefore, the enzymatic assay could fail in detection of fragments activity or be plagued by high false positive rates, caused by aggregation phenomenon.[33] Bio-layer interferometry (BLI)[38] and Saturation transfer difference (STD)[39] techniques were used to detect protein-ligand interactions and a K_D value was calculated. (Table **2**) The primary fragment hits were repurchased, but unfortunately compound RIM518 was not anymore available and thus it was not possible to test it in the orthogonal assay. Among the remaining seven compounds, five of them confirmed their activity. In particular, compound RIM565 and RIM570 were the most potent molecules showing a K_D in the low micromolar range. Compound RIM585 binding was detected only by STD method and it required further studies to confirm its potential activity.
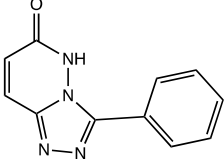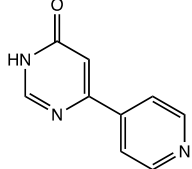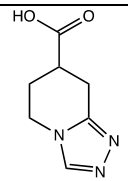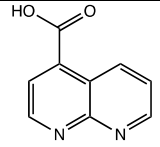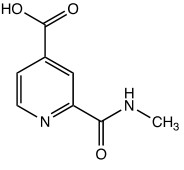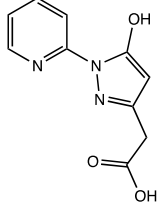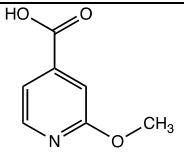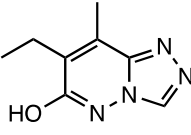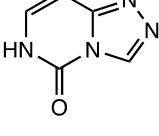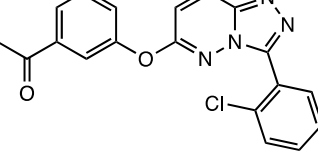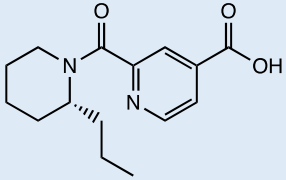
Compound RIM	Structure	K_D (μM)	Binding by BLI	Binding by STD
551		164 ± 4	Detected	Detected
565		9 ±0.12	Detected	Detected
588		64 ±1.2	Detected	Detected
503		33 ± 0.6	Detected	Detected
570		15 ± 0.16	Detected	Detected
554		NA	Not-detected	Not-detected
585		NA	Not-detected	Detected

Table 2. Biophysical assay results of primary fragment hits. In the table are indicated the ID number of the compound in our library, 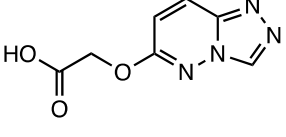the structure of primary fragment hits, the value of dissociation constant (NA means not applicable) and the presence or not of a binding process during the assays with the two different biophysical techniques)

In order to validate the compounds series and identify the best fragment to optimize, some pf the fragment hit analogues were also tested in the biophysical assays. (Table **3**) The compounds that did not show binding during BLI screening were not further tested with STD technic. Among RIM551 analogues, only compound RIM712 showed a K_D of 130 µM, better than the corresponding fragment hit but considering that RIM712 is a relatively large molecule and not really a fragment, this value cannot be consider satisfactory. The RIM570 analogues showed binding with the protein, confirming the series activity. Compound RIM692 is a close analogue of starting fragment (RIM570) but its K_D value is increased, whereas compound 694 showed a similar activity to RIM570 compound. Moreover, RIM565 analogues screening validated the series. Indeed, for all three analogues, (compounds RIM713, RIM714, RIM715) the binding with the protein were detected and in particular compound 715 showed an improvement of activity.

Primary Hit	Analogue	Structure	K_D (µM)	Binding by BLI	Binding by STD
551	681		NA	Not-detected	Not-tested
551	685		NA	Not-detected	Not-tested
551	696		>1000	Detected	Not-tested
551	712		130	Detected	Detected
570	692		41	Detected	Detected
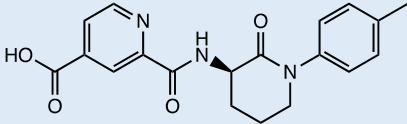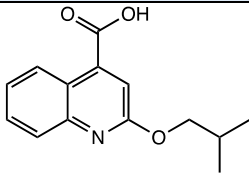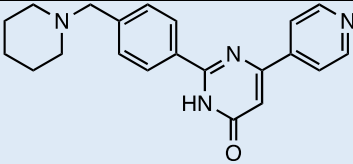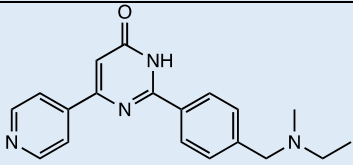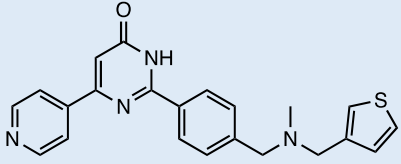
570	694		18.5	Detected	Detected
503	711		NA	Not-detected	Not-tested
565	713		40	Detected	Detected
565	714		38	Detected	Detected
565	715		6	Detected	Detected

Table 3. Biophysical assay results of primary fragment hits analogues. In the table are indicated the ID number of the compound in our library and the corresponding starting fragment hit, the structure of primary fragment hits, the value of dissociation constant (NA means not applicable) and the presence or not of binding during the assays with the two different biophysical techniques)

4. 5 Fragment hit optimization

The RIM565 and RIM570 series were validated. Then, a fragment growing approach was employed to enhance the fragment hits potency.

4.5.1 Induced-fit docking and design of 2-Carbamoylpyridine-4-carboxylic acid derivatives

The first step was to deepen computational study in order to rationally design compounds with improved activity using a structure-based rational drug design approach exploiting the local environment around the binding pocket. In order to exploit possible electrostatic interactions with the protein, docking screening in extra precision mode and induced-fit-docking were carried out. The studies were performed for all the 8 primary fragment hits but the focus has been maintained on compounds RIM565 and RIM570. The induced-fit docking is also known as flexible docking and diverges from classic docking studies since it allows the protein side chains as more flexible and enables the receptor to change his binding site based on the shape and the binding mode of the ligand.[40,41] The compound RIM565 induced-fit binding pose is different compared with the one retrieved in the classic docking mode. The 6-(pyridin-4-yl)pyrimidin-4(3*H*)-one moiety is involved in zinc ion binding and also establishes H-bond interactions with Glu190 and Thr270 residues, with the pyridine moiety forming π-π interaction with Tyr177 and H-bond interaction with Lys206 (Figure **5A**). In the classic docking, the position was reversed with the pyridine moiety involved with metal binding and 6-(pyridin-4-yl)pyrimidin-4(3*H*)-one moiety directed toward the 2-OG binding site. Whereas both binding poses retrieved with different docking protocols for compound RIM570 are convergent. The 4 position carboxylic group interacts with Tyr132 and Lys 206 inside the 2-OG binding site and the metal ion is bound in a bidentate manner by the nitrogen atom of pyridine and the carbonyl group of the side chain. (Figure **5B**)

Furthermore the compound RIM570 was overlapped with the crystal structure of Celgene lead compound and the structures were found perfectly superimposed. (Figure **5C**) Hence, we decided to substitute the methyl group in the carbamoyl moiety with larger substituents and to grow in this direction our fragment hit. (Figure **5D**) Several substituents were analyzed and, using as a guide the superimposition of compound RIM570 with QC6352, different combinations were designed. The new compounds

designed were docked with XP protocol and according to their binding mode, the most promising compounds were selected to be synthesized.
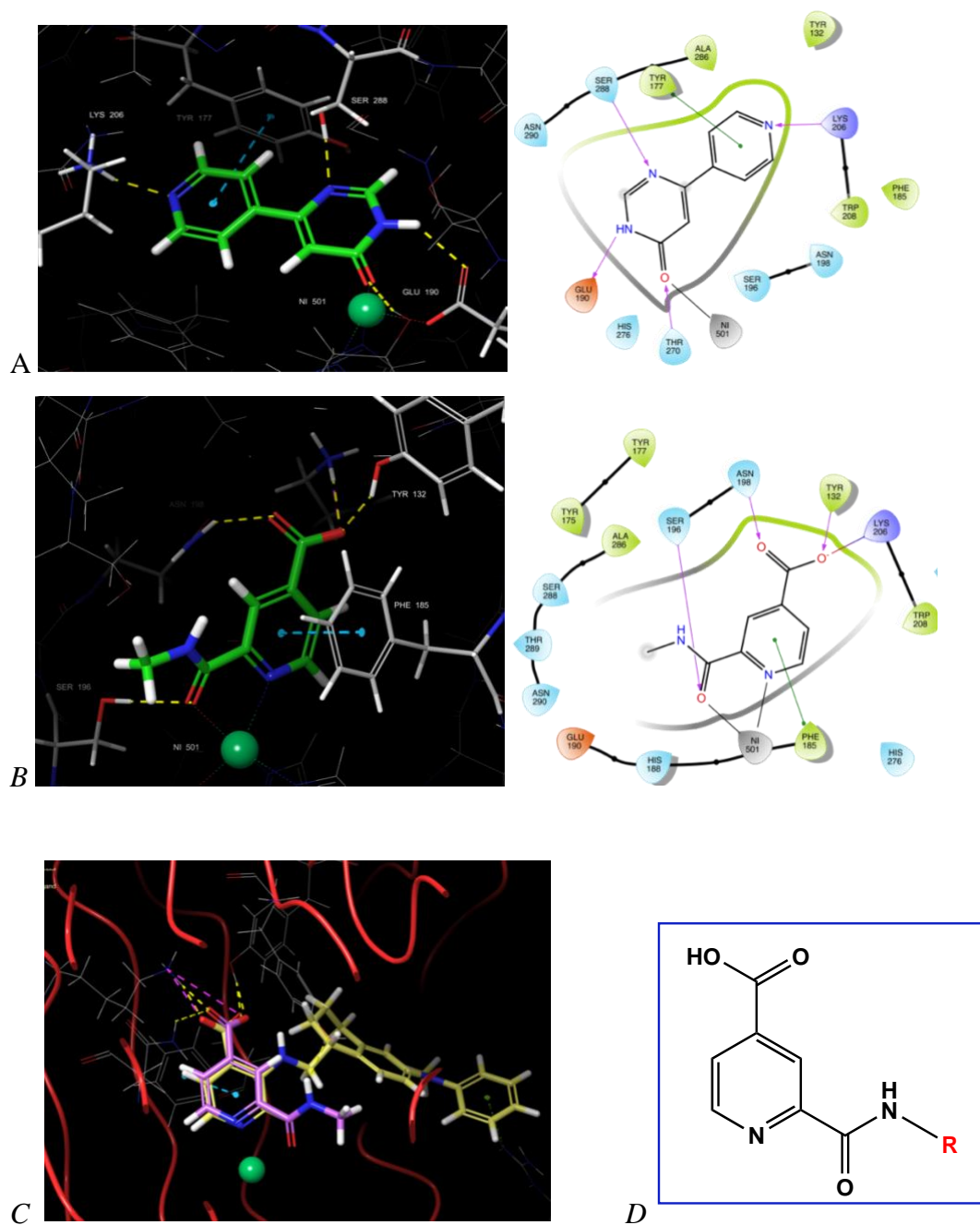
Figure 5. Compounds binding poses (A, compound 565) (B, compound 570) retrieved from Induced-fit docking with 2D ligand-protein interaction. C Superimposition of Celgene compound QC6352 in yellow and compound 570 in pink inside the protein KDM4A (PDB ID: 5VGI)

4.5.2 Synthesis of 2-Carbamoylpyridine-4-carboxylic acid derivatives

The 2-Carbamoylpyridine-4-carboxylic acid derivatives were prepared by the synthetic route shown in Scheme **1**. Starting from the 4-(methoxycarbonyl)pyridine-2-carboxylic, the coupling reaction with the corresponding amine was mediated by EDCI hydrochloride/ Hydroxy-benzotriazole reagents, in order to form the carbamoyl group. Subsequently, the methyl 2-carbamoylpyridine-4-carboxylate derivatives obtained were hydrolysed using lithium hydroxide, in order to obtain the carboxylic group in 4 position.
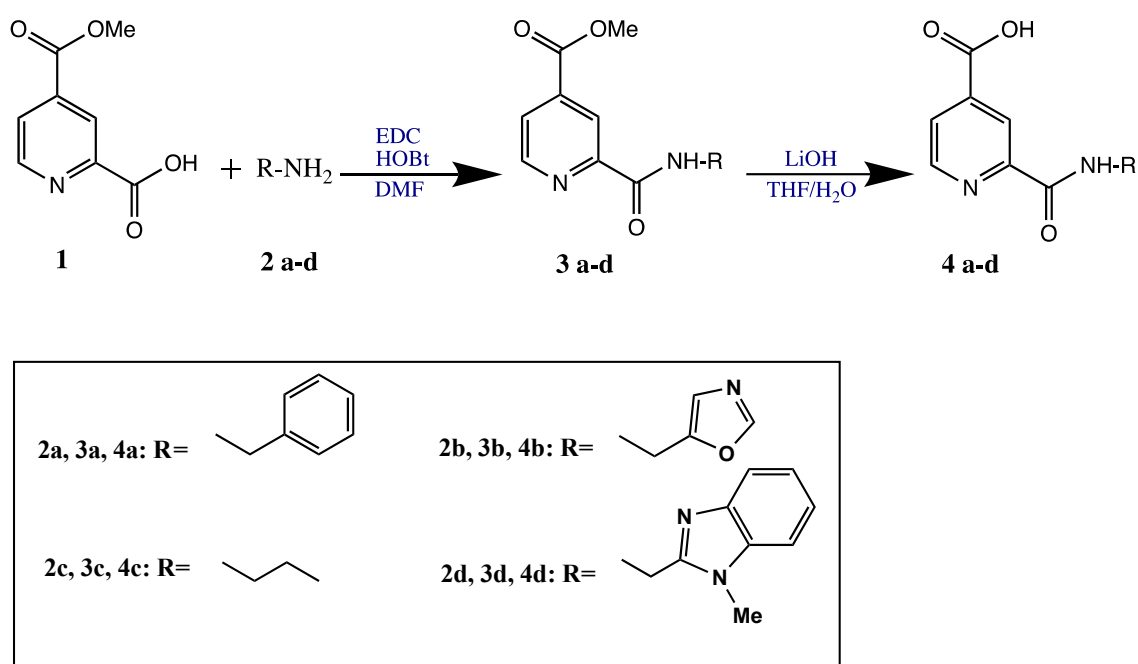
Scheme 1. Reagent and conditions: 1. EDCI- HCl, 1-hydroxybenzotriazole, DMF, room temperature, 3h. 2. Lithium hydroxide, THF, room temperature, 1h.

4.5.3 Biophysical and Biochemical assays

The synthesized compounds were tested in BLI and STD assays in order to evaluate if the activity of new fragments were improved. The results are shown in table **4**. The compound 3d with methyl ester group in 4 position showed weak activity compared to

the fragment hit and also considering the molecular weight increase, the LE and LLE values were not sufficient for further development. Compound 4a exhibited a weaker activity with K_D value of 28μM but its LE score was acceptable. The most potent compounds were 4b and 4c exhibiting an improved activity compared to compound RIM570. The LE and LLE scores calculated for compounds 4b and 4c are in the optimal range, indicating that these compounds are suitable for further optimization, with the aim to obtain a lead compound with activity in the nanomolar range.
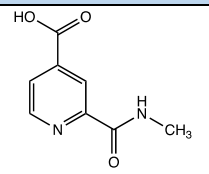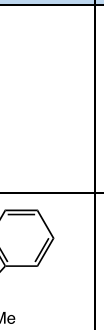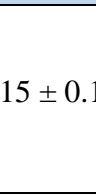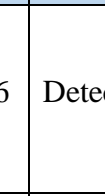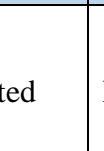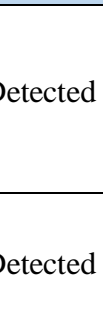
Compound	Structure	K_D (μM)	BLI binding	STD binding	LE	LLE
570		15 ± 0.16	Detected	Detected	0.51	4.59
3d		55.0 ± 2.0	Detected	Detected	0.24	1.88
4a		28.0 ± 1.0	Detected	Detected	0.33	2.49
4b		6.0 ± 0.08	Detected	Detected	0.39	4.50
4c		$6.0 \pm X$	Detected	Detected	0.48	4.18
4d		>50μM	Detected	Detected	ND	ND

Table 4. Biophysical assay results of the 2-Carbamoylpyridine-4-carboxylic acid derivatives synthesized

The compounds 4b and 4c were further tested on enzymatic assay.

Compounds 4b and 4c showed activity with $IC_{50} = 20$ μM and $IC_{50} = 40$ μM respectively.

A

B

C
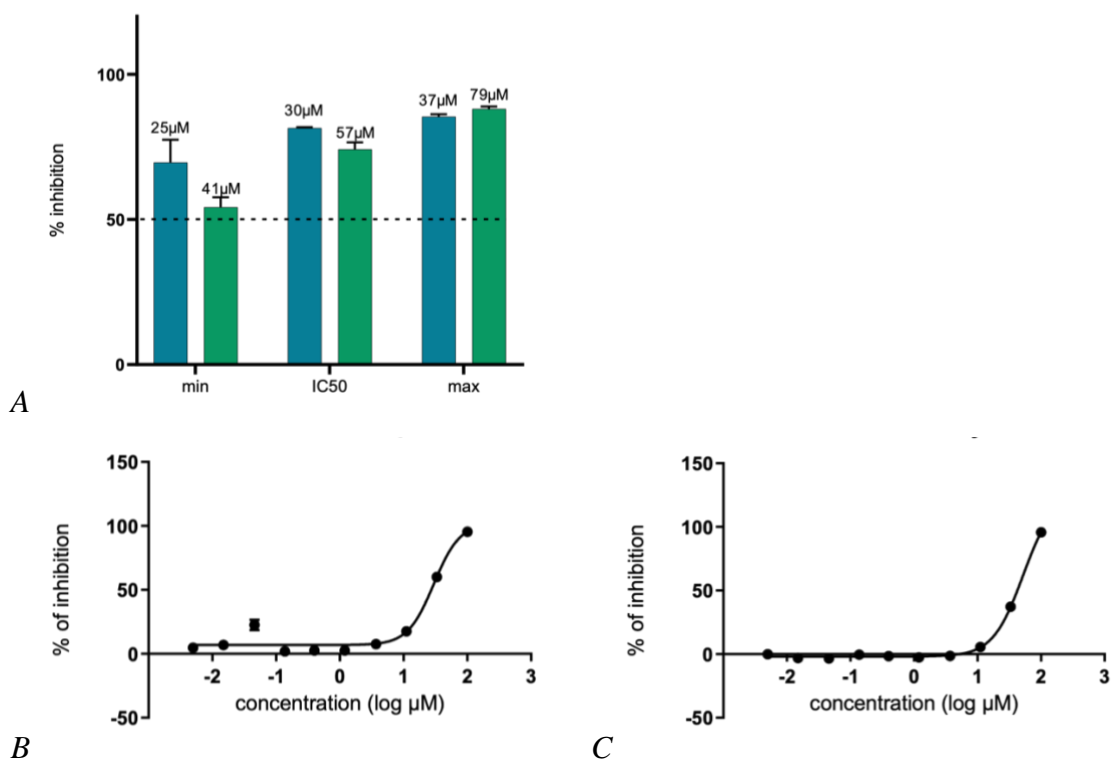
Figure 6. Biochemical assay results. A. Single dose test of compounds 4b (blue) and 4c (green). B. Compound 4b dose response curve. C. Compound 4c dose response curve.

4.6 Material and methods

Creation of fragments virtual screening library

The virtual fragments library was created merging in a unique file different available virtual libraries: Prestwick, Life Chemical, BioAscent, ChemDiv, Asinex, chemBridge, Discovery Chemistry and Enamine. The libraries compounds were prepared using LigPrep, tool of Maestro program and their properties were calculated using Qikprop. Then, the compounds were filtered by MW and applying the Rule of three. Furthermore using Canvas[42,43], the library was filtered by PAINS and REOS in order to discard possible reactive compounds. In order to avoid close analogues in our library and ensure a sufficient chemical diversity, a similarity matrix, based on linear fingerprint, using Canvas tools, was created and medium value calculated is 0.025, indicating an acceptable diversity in the fragment library created.

Virtual screening campaign:

Pharmacophore creation and screening

LigandScout[44] software was employed to create the pharmacophore model and to perform the pharmacophore screening. The pharmacophore model was created, using the PDB coordinate of ligand-protein complex (PDB ID: 5VMP). In the screening the "pharmacophore fit-score" was used as scoring function and "match all query features" was chosen as screening mode. The selected retrieval mode was "get best matching conformation".

Docking screening

- **Protein preparation:**

The crystal structure of the protein KDM4A in complex with the ligand QC6362 with PDB ID: 5VGI, and resolution of 2,07 Å was used for the set-up of our model. The protein was prepared using *Protein Preparation Wizard*, [45] a tool from Schrödinger software adding bond orders and hydrogens to the crystal structure and deleting waters beyond 5.00 Å. *Epik* was used to generate het states at pH 7.4 ± 0.2 and then the protonation state of the protein and the ligand were optimized using *PropKa* at pH 7.4. [46,47]

- **Docking grid generation:**

The docking grid was generated using *Glide* software. [48] The scaling factor was set at 1.0 Å with a partial charge cut-off of 0.25, and the ligand was chosen to define the grid centroid. The key features retrieved from pharmacophore model, were imposed in the docking grid as constraints: H-bond acceptor on Tyr132, and Lys206, metal acceptor and positional constraint.

- **Ligands preparation:**

The ligands were prepared using *LigPrep*, a tool of MAESTRO software. OPLS3e was used as force field and *ionizer* generated all the possible states at pH: 7.0 ± 0.2. Desalt and generate tautomers were flagged on, the chiralities were retained and at most 32 conformers per ligand were generated.

- **Ligands docking:**

Molecular docking was carried out using *Glide* software[48] released by Schrödinger (release 2018-4). The docking screening of virtual fragments library was firstly run in HTVS (high-throughput virtual screening) mode, limiting the number of poses to report at 10000 compounds. After, these 10000 fragments were screened in standard precision (SP) mode. While during the optimization phase, in order to deepen the binding mode, the 8 primary fragment hits were docked using the extra precision (XP) protocol. OPLS3e was the force field employed for all the screening. [49] The ligands were considered as flexible, and Epik state penalties were included to docking score. The van der Waals (vdW) radii scaling factor was set as 0.8, with a partial charge cut off by 0.15.

Biochemical assay

Assay was performed using a homogeneous time-resolved fluorescence resonance energy transfer (TR-FRET, EpiGenetics https://www.cisbio.eu/media/asset/l/s/ls-tn-jmjd2c-histone-h3k9-mono-demethylation-me3-me2.pdf).

The principle of the assay relies on a europium cryptate -labelled antibody specific to the substrate's methylation state and a streptavidin acceptor molecule (SXL655) that binds

the substrate biotinylated peptide. When the donor europium chelate is excited at 320 nm, energy in the form of FRET elicits fluorescence emission at 665 nm from the nearby streptavidin-tagged acceptor. HTRF signal is proportional to the concentration of demethylated H3(1-21) lysine 9 tri-methylated peptide.

The assay run in two steps: enzymatic reaction and detection. Optimal conditions for both steps were determined with manual experiments considering: time and temperature of incubation, substrate and DMSO concentration, purity of reagents, plates format.

Kdm4 enzyme concentration was set at 400 nm in buffer constituted by: Bovine Serum Albumin 0.01 % (BSA EMD Millipore 3458509), HEPES 50nM (EuroClone ECM0180D), Tween20 0,01 % (Sigma Aldrich P2287) and deionized water (Sigma Aldrich 38796).

Enzyme substrates concentrations were 0.5 µM for α-Ketoglutaric Acid Monopotassium (2-OG, Sigma Aldrich K2000) and 300 nm for Histone H3(1-21) lysine 9 tri-methylated biotinylated peptide (AnaSpec AS-64360). 5µM Ammonium iron (II) sulfate hexa-hydrate (Sigma Aldrich 215406) and 1mM of Sodium ascorbate (Sigma Aldrich 1140) were added as co-factors.

16 µL of compounds were incubated for 1 h at 20°C with 8 µL of enzymatic solution and 16 µL of substrates solution in ½ area 96 well plates (ref. 6002290m, perkinHelmer). Dimethyl sulfoxide (DMSO Sigma-Aldrich D58791) final concentration was 1 %.

40 µl of antibodies solutions (Cisbio Bioassays H3K9 me2-Eu(K) Ab 61KB2KAE and Streptavidin XL-665 Cisbio Bioassays 610SAXLA, in detection buffer 62SDBRDD) were added to each well and incubated o. night at 20 °C.

Fluorescence was measured with a TECAN Spark te-cool M-200 reader (Männedorf, Switzerland), excitation wavelength of 320 nm and emission wavelength of 620 and 665 nm.

Automation was performed using Bravo (Agilent) platform. Robustness of the assay was proven by performing a Z' determination over 3 kinds of experiments: evaluation of max, min and mid signal by using 2,4-pyridinedicarboxylic acid monohydrate (2,4-PDCA, Sigma Aldrich P63395) at IC100 (10µM), IC50 (50 nM) and 0 µM; evaluation of dose - response curve with references compounds (PDCA and in-house); evaluation of single-dose experiments reproducibility. Dose response curves were performed by addressing 10 point of 1: 3 compound dilutions, departing from 100 µM.

Biophysical assay

STD experiment: Molar ratio RIM:Protein 50:1

Saturation Time: 1sec

Buffer: 20mM NaP pH 7.3, 150mM NaCl, 0.5mM TCEP, 5% DMSO

BLI experiment: 2/3 independent exp using single-use aliquots

Buffer: 10mM HEPES pH 7.3, 150mM NaCl, 0.5mM TCEP, 5% DMSO.

Tested concentrations of compound: 0-100 µM.

Analogues selection

With the aim of validating the activity of the primary fragment hits, analogues of these compounds were searched on PubChem database.[50] The objective was to identify close analogues to the corresponding fragment hits, thus the analogues search was carried out with Pubchem fingerprint and setting 0.9 as similarity coefficient of Tanimoto.

Induced-fit docking

The induced-fit docking was carried out using a standard protocol, generating up to 20 poses and sampling ring conformations with an energy window of 2.5Kcal/mol. The receptor and ligand van der Waals scaling factors were set at 0.5. Prime was used for the refinement and the redocking was performed in extra precision.

Chemistry

Commercially available starting materials, reagents, and anhydrous solvents were used as supplied. Flash column chromatography was performed using Merck silica gel 60 (0.025−0.04 mm). Thin layer chromatography was performed using Merck Millipore TLC silica gel 60 F254 sheets and visualized by UV (254 and 356 nm) iodine, and $KmnO_4$.

- **General procedure 1:**

A solution of the 4-methoxy-carbonyl pyridine-2-carboxylic acid (1 equ) in DMF was cooled at 0°C and treated with EDC-HCl (2 equ) and HOBt (2 equ). After 1 h, the amine (1.5 eq) and DIPEA (3 eq) was added and the reaction mixture was stirred at room temperature. After completion, the mixture was quenched with water and then extracted with Ethyl acetate. The organic layer was washed several times with brine, dried with Na_2SO_4 and concentrated in vacuo. The resulting material was purified by silica gel column chromatography.

Methyl 2-(benzylcarbamoyl)isonicotinate (3a)

According to general procedure 1, the 4-methoxy-carbonyl pyridine-2-carboxylic acid (0.100 g 0.55 mmol) and benzylamine (0.089 g 0.83 mmol) were reacted. The reaction mixture was stirred at RT for 3 h, then water was added and the mixture was extracted with Ethyl acetate. The organic layer was washed several times with brine, dried with Na_2SO_4 and concentrated in vacuo. The crude product was purified by silica gel column chromatography eluting with 40%EtOAc in DCM to provide the product as white powder. (0,110 g 75%)

^{1}H NMR (200 MHz, DMSO) δ 9.54 (s, 1H), 8.85 (d, J = 3.2 Hz, 1H), 8.45 (s, 1H), 8.06-8.04 (m, 1H), 7.35-7.25 (m, 5H), 4.51 (d, J = 5Hz, 2H), 3.93 (s, 3H).

^{13}C NMR (201 MHz, DMSO) δ 164.88, 163.81, 150.67, 148.23, 142.52, 138.92, 128.41, 127.42, 127.00, 124.84, 122.26, 52.65, 42.90.

methyl 2-((oxazol-5-ylmethyl)carbamoyl)isonicotinate (3b)

According to general procedure 1, the 4-methoxy-carbonyl pyridine-2-carboxylic acid (0.200 g 1.10 mmol) and 1,3-oxazol-5yl methanamine (0.162 g 1.65 mmol) were reacted. The reaction mixture was stirred at RT for 3 h, and then water was added and the mixture was extracted with Ethyl acetate. The organic layer was washed several times with brine, dried with Na_2SO_4 and concentrated in vacuo. The crude product was purified by silica gel column chromatography eluting with 5% MeOH in DCM to provide the product as white powder. (0.120 g 42%)

¹H NMR (200 MHz, DMSO) δ 9.53 (d, *J* = 4 Hz, 1H), 8.85 (d, *J* = 3.2 Hz, 1H), 8.46 (d, *J* = 3.4 Hz, 1H), 8.32 (d, *J* = 3.8 Hz, 1H), 8.00 (s, 1H), 7.11 ((d, *J* = 3 Hz, 1H), 4.58 (d, *J* = 3.6 Hz, 2H) 3.5 (s, 3H).

¹³C NMR (201 MHz, DMSO) δ 170.34, 164.80, 163.83, 151.72, 150.66, 142.05, 129.59, 124.82, 123.71, 122.24, 52.62, 33.97.

methyl 2-(propylcarbamoyl)isonicotinate (3c)

According to general procedure 1, the 4-methoxy-carbonyl pyridine-2-carboxylic acid (0.100 g 0.55 mmol) and Propilamine (0.05 g 0.83 mmol) were reacted. The reaction mixture was stirred at RT for 3 h, then water was added and the mixture was extracted with Ethyl acetate. The organic layer was washed several times with brine, dried with Na_2SO_4 and concentrated in vacuo. The crude product was purified by silica gel column chromatography eluting with 2% MeOH in DCM to provide the product as white oil. (0.03 g 55%)

¹H NMR (200 MHz, DMSO) δ 8.94 (t, *J* = 5.6 Hz, 1H), 8.84 (d, *J* = 4.9 Hz, 1H), 8.42 (d, *J* = 1.7 Hz, 1H), 7.97 (dd, *J* = 4.9, 1.7 Hz, 1H), 3.80 (s, 3 H), 3.25 (td, *J* = 7.2, 5.8 Hz, 2H), 1.55 (h, *J* = 7.3 Hz, 2H), 0.90 (t, *J* = 7.4 Hz, 3H).

¹³C NMR (201 MHz, DMSO) δ 170.35, 164.89, 150.55, 148.14, 142.83, 124.72, 122.17, 59.76, 52.60, 22.13, 11.44.

methyl 2-(((1-methyl-1H-benzo[d]imidazole-2-yl)methyl)carbamoyl)isonicotinate (3d)

According to general procedure 1, the 4-methoxy-carbonyl pyridine-2-carboxylic acid (0.300 g 1.65 mmol) and 1-Methyl-1-H-benzimidazol-2-yl methanamine (0.400 g 2.48 mmol) were reacted. The reaction mixture was stirred overnight at RT, then water was added and the mixture was extracted with Ethyl acetate. The organic layer was washed several times with brine, dried with Na_2SO_4 and concentrated in vacuo. The crude product was purified by silica gel column chromatography eluting with 2% MeOH in DCM to provide the product as white crystal. (0.110 g 20%)

¹H NMR (800 MHz, DMSO) δ 10.05 (d, *J* = 7.7 Hz, 1H), 8.92 – 8.90 (m, 1H), 8.53 (d, *J* = 1.7 Hz, 1H), 8.11 – 8.09 (m, 1H), 7.95 (d, *J* = 8.2 Hz, 1H), 7.79 (d, *J* = 8.0 Hz, 1H), 7.62 – 7.55 (m, 2H), 5.05 (d, *J* = 5.3 Hz, 2H), 4.05 (s, 3H), 3.93 (s, 3H),.

¹³C NMR (201 MHz, DMSO) δ 165.89, 165.10, 150.98, 150.56, 149.36, 141.47, 133.04, 126.49 – 125.14 (m), 124.86, 122.40, 114.67, 112.71, 51.60, 35.45, 31.38.

- **General procedure 2:**

Lithium hydroxide (3 equiv) was added to a solution of the starting material (compounds 3a-d) (1 equiv) in 1:1 THF/H2O. The reaction was stirred at RT for 1 h. After completion, the reaction mixture was neutralized with 6N aq HCl and extracted with EtOAc. The organic layer was dried with Na₂SO₄ and concentrated in vacuo.

2-(benzylcarbamoyl)isonicotinic acid (4a)

According to general procedure 2, methyl 2-(benzylcarbamoyl)isonicotinate (0.05 g 0.185 mmol) and LiOH (0.013 g, 0.55 mmol) were reacted. Then, the reaction mixture was neutralized with 6N aq HCl and extracted with EtOAc. The organic layer was dried with Na₂SO₄ and concentrated in vacuo, providing the title compound as white solid. (0.035 g 75%)

¹H NMR (200 MHz, DMSO) δ 9.54 (d, J = 2 Hz, 1H), 8.86 (d, J = 3.6 Hz, 1H), 8.48 (s, 1H), 8.02 (d, J = 3.2 Hz, 1H), 7.34-7.25 (m, 5H), 4.51 (d, J = 3.6 Hz, 2H).

¹³C NMR (201 MHz, DMSO) δ 164.88, 163.81, 150.67, 148.23, 142.52, 138.92, 128.41, 127.42, 127.00, 124.84, 122.26, 42.90.

2-((oxazol-5-ylmethyl)carbamoyl)isonicotinic acid (4b)

According to general procedure 2, methyl 2-((oxazol-5-ylmethyl)carbamoyl) isonicotinate (0.1 g 0.38 mmol) and LiOH (0.027 g, 1.15 mmol) were reacted. Then, the reaction mixture was neutralized with 6N aq HCl and extracted with EtOAc. The organic layer was dried with Na₂SO₄ and concentrated in vacuo. The crude obtained was washed with diethyl ether providing the product as white solid. (0.02 g 20%)

¹H NMR (200 MHz, DMSO) δ 9.53 (d, J = 4 Hz, 1H), 8.85 (d, J = 3.2 Hz, 1H), 8.46 (d, J = 3.4 Hz, 1H), 8.32 (d, J = 3.8 Hz, 1H), 8.00 (s, 1H), 7.11 ((d, J = 3 Hz, 1H), 4.58 (d, J = 3.6 Hz, 2H).

¹³C NMR (201 MHz, DMSO) δ 170.34, 164.80, 163.83, 151.72, 150.66, 142.05, 129.59, 124.82, 123.71, 122.24, 33.97.

2-(propylcarbamoyl)isonicotinic acid (4c)

According to general procedure 2, *methyl 2-(propylcarbamoyl)isonicotinate* (0.03 g 0.135 mmol) and LiOH (0.01 g, 0.40 mmol) were reacted. Then, the reaction mixture was neutralized with 6N aq HCl and extracted with EtOAc. The organic layer was dried with Na_2SO_4 and concentrated in vacuo. The crude obtained was crystallized with ethanol providing the product as white solid. (0.015 g 55%)

¹H NMR (800 MHz, DMSO) δ 8.94 (t, *J* = 5.6 Hz, 1H), 8.84 (d, *J* = 4.9 Hz, 1H), 8.42 (d, *J* = 1.7 Hz, 1H), 7.97 (dd, *J* = 4.9, 1.7 Hz, 1H), 3.25 (td, *J* = 7.2, 5.8 Hz, 2H), 1.55 (h, *J* = 7.3 Hz, 2H), 0.90 (t, *J* = 7.4 Hz, 3H).

¹³C NMR (201 MHz, DMSO) δ 166.18, 164.09, 150.55, 149.43, 143.12, 124.74, 122.28, 41.48, 22.44, 11.75.

2-(((1-methyl-1H-benzo[d]103midazole-2-yl)methyl)carbamoyl)isonicotinic acid (4d)

According to general procedure 2, methyl 2-(((1-methyl-1H-benzo[d]103midazole-2-yl)methyl)carbamoyl)isonicotinate (0.035 g 0.108 mmol) and LiOH (0.08 g, 0.324 mmol) were reacted. Then, the reaction mixture was neutralized with 6N aq HCl and extracted with EtOAc. The organic layer was dried with Na_2SO_4 and concentrated in vacuo. The crude obtained was washed with diethyl ether providing the product as white solid. (0.015 g 45%)

¹H NMR (800 MHz, DMSO) δ 10.05 (d, J = 7.7 Hz, 1H), 8.92 – 8.90 (m, 1H), 8.53 (d, J = 1.7 Hz, 1H), 8.11 – 8.09 (m, 1H), 7.95 (d, J = 8.2 Hz, 1H), 7.79 (d, J = 8.0 Hz, 1H), 7.62 – 7.55 (m, 2H), 5.05 (d, J = 5.3 Hz, 2H), 4.05 (s, 3H).

¹³C NMR (201 MHz, DMSO) δ 165.89, 165.10, 150.98, 150.56, 149.36, 141.47, 133.04, 126.49 – 125.14 (m), 124.86, 122.40, 114.67, 112.71, 35.45, 31.38.

Bibliography

1.   Flemming, A. Cancer: Targeting mutant BRAF in metastatic melanoma. *Nature Reviews Drug Discovery* **9**, 841 (2010).

2.   Kim, A. & Cohen, M. S. The discovery of vemurafenib for the treatment of BRAF-mutated metastatic melanoma. *Expert Opinion on Drug Discovery* **11**, 907–916 (2016).

3.   Bollag, G. *et al.* Vemurafenib: The first drug approved for BRAF-mutant cancer. *Nature Reviews Drug Discovery* **11**, 873–886 (2012).

4.   Murray, C. W. & Rees, D. C. The rise of fragment-based drug discovery. *Nature Chemistry* **1**, 187–192 (2009).

5.   Michne, W. F. Hit-to-lead chemistry: a key element in new lead generation. *Pharmaceut. News* **3**, 3 (1996).

6.   Brown, D. & Superti-Furga, G. Rediscovering the sweet spot in drug discovery. *Drug Discovery Today* **8**, 1067–1077 (2003).

7.   Hird, N. Isn't combinatorial chemistry just chemistry? *Drug Discovery Today* **5**, 307–308 (2000).

8.   Keserü, G. M. & Makara, G. M. The influence of lead discovery strategies on the properties of drug candidates. *Nature Reviews Drug Discovery* **8**, 203–212 (2009).

9.   Hann, M. M., Leach, A. R. & Harper, G. Molecular Complexity and Its Impact on the Probability of Finding Leads for Drug Discovery. *Journal of Chemical Information and Computer Sciences* **41**, 856–864 (2001).

10.   Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews* **64**, 4–17 (2012).

11.   Honma, T. Recent advances in De Novo design strategy for practical lead identification. *Medicinal Research Reviews* **23**, 606–632 (2003).

12.   Leach, A. R. & Hann, M. M. Molecular complexity and fragment-based drug discovery: Ten years on. *Current Opinion in Chemical Biology* **15**, 489–496 (2011).

13.   Kirsch, P., Hartman, A. M., Hirsch, A. K. H. & Empting, M. Concepts and core principles of fragment-based drug design. *Molecules* **24**, (2019).

14.   Fink, T., Bruggesser, H. & Reymond, J. L. Virtual exploration of the small-molecule chemical universe below 160 daltons. *Angewandte Chemie - International Edition* **44**, 1504–1508 (2005).

15.   Fink, T. & Raymond, J. L. Virtual exploration of the chemical universe up to 11 atoms of C, N, O, F: Assembly of 26.4 million structures (110.9 million stereoisomers) and analysis for new ring systems, stereochemistry, physicochemical properties, compound classes, and drug discove. *Journal of Chemical Information and Modeling* **47**, 342–353 (2007).

16.   Wang, Y. *et al.* Effects of Lipophilicity on the Affinity and Nonspecific Binding of Iodinated Benzothiazole Derivatives. *Journal of Molecular Neuroscience* **20**, 255–260 (2003).

17.   Parker, M. A., Kurrasch, D. M. & Nichols, D. E. The role of lipophilicity in determining binding affinity and functional activity for 5-HT2A receptor ligands. *Bioorganic and Medicinal Chemistry* **16**, 4661–4669 (2008).

18.   Wenlock, M. C., Austin, R. P., Barton, P., Davis, A. M. & Leeson, P. D. A comparison of physiochemical property profiles of development and marketed oral drugs. *Journal of Medicinal Chemistry* **46**, 1250–1256 (2003).

19.   Hopkins, A. L., Groom, C. R. & Alex, A. Ligand efficiency: A useful metric for lead selection. *Drug Discovery Today* **9**, 430–431 (2004).

20.   Leeson, P. D. & Springthorpe, B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nature Reviews Drug Discovery* **6**, 881–890 (2007).

21.   Schuffenhauer, A. *et al.* Library Design for Fragment Based Screening. *Current Topics in Medicinal Chemistry* **5**, 751–762 (2005).

22.   Någren, K. PET and knockout mice in drug discovery. *Drug Discovery Today* **8**, 876 (2003).

23.   Jhoti, H., Williams, G., Rees, D. C. & Murray, C. W. The 'rule of three' for fragment-based drug discovery: Where are we now? *Nature Reviews Drug Discovery* **12**, 644 (2013).

24.   Erlanson, D. A., McDowell, R. S. & O'Brien, T. Fragment-based drug discovery. *Journal of Medicinal Chemistry* **47**, 3463–3482 (2004).

25.   Edink, E. *et al.* Fragment growing induces conformational changes in acetylcholine-binding protein: A structural and thermodynamic analysis. *Journal of the American Chemical Society* **133**, 5363–5371 (2011).

26.   Miyake, Y. *et al.* Identification of novel lysine demethylase 5-selective inhibitors by inhibitor-based fragment merging strategy. *Bioorganic and Medicinal Chemistry* **27**, 1119–1129 (2019).

27.   Howard, N. *et al.* Application of fragment screening and fragment linking to the discovery of novel thrombin inhibitors. *Journal of Medicinal Chemistry* **49**, 1346–1355 (2006).

28.   Korczynska, M. *et al.* Docking and Linking of Fragments to Discover Jumonji Histone Demethylase Inhibitors. *Journal of Medicinal Chemistry* **59**, 1580–1598 (2016).

29.   Potter, A. *et al.* Discovery of cell-active phenyl-imidazole Pin1 inhibitors by structure-guided fragment evolution. *Bioorganic and Medicinal Chemistry Letters* **20**, 6483–6488 (2010).

30.   Verdonk, M. L. & Rees, D. C. Group efficiency: A guideline for hits-to-leads chemistry. *ChemMedChem* **3**, 1179–1180 (2008).

31.   Chen, Y. K. *et al.* Design of KDM4 Inhibitors with Antiproliferative Effects in Cancer Models. *ACS Medicinal Chemistry Letters* **8**, 869–874 (2017).

32.   Hajduk, P. J. Fragment-based drug design: How big is too big? *Journal of Medicinal Chemistry* **49**, 6972–6976 (2006).

33.   Siegel, M. G. & Vieth, M. Drugs in other drugs: a new look at drugs as fragments. *Drug Discovery Today* **12**, 71–79 (2007).

34.   Dahlin, J. L. *et al.* PAINS in the assay: Chemical mechanisms of assay interference and promiscuous enzymatic inhibition observed during a sulfhydryl-scavenging HTS. *Journal of Medicinal Chemistry* **58**, 2091–2113 (2015).

35.   Carr, R. A. E., Congreve, M., Murray, C. W. & Rees, D. C. Fragment-based lead discovery: Leads by design. *Drug Discovery Today* **10**, 987–992 (2005).

36.   Barker, J., Courtney, S., Hesterkamp, T., Ullmann, D. & Whittaker, M. Fragment screening by biochemical assay. *Expert Opinion on Drug Discovery* **1**, 225–236 (2006).

37.   Fattori, D., Squarcia, A. & Bartoli, S. Fragment-Based Approach to Drug Lead Discovery: Overview and Advances in Various Techniques. *Drugs in R and D* **9**, 217–227 (2008).

38.   Sultana, A. & Lee, J. E. Measuring protein-protein and protein-nucleic acid interactions by biolayer interferometry. *Current Protocols in Protein Science* **2015**, 19.25.1-19.25.26 (2015).

39.   Mayer, M. & Meyer, B. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angewandte Chemie - International Edition* **38**, 1784–1788 (1999).

40.   Zhong, H., Tran, L. M. & Stang, J. L. Induced-fit docking studies of the active and inactive states of protein tyrosine kinases. *Journal of Molecular Graphics and Modelling* **28**, 336–346 (2009).

41.   Xu, M. & Lill, M. A. Induced fit docking, and the use of QM/MM methods in docking. *Drug Discovery Today: Technologies* **10**, e411–e418 (2013).

42.   Sastry, M., Lowrie, J. F., Dixon, S. L. & Sherman, W. Large-scale systematic analysis of 2D fingerprint methods and parameters to improve virtual screening enrichments. *Journal of Chemical Information and Modeling* **50**, 771–784 (2010).

43.   Duan, J., Dixon, S. L., Lowrie, J. F. & Sherman, W. Analysis and comparison of 2D fingerprints: Insights into database screening performance using eight fingerprint methods. *Journal of Molecular Graphics and Modelling* **29**, 157–170 (2010).

44.   Wolber, G. & Langer, T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *Journal of Chemical Information and Modeling* **45**, 160–169 (2005).

45.   Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *Journal of Computer-Aided Molecular Design* **27**, 221–234 (2013).

46.   Shelley, J. C. *et al.* Epik: A software program for pKa prediction and protonation state generation for drug-like molecules. *Journal of Computer-Aided Molecular Design* **21**, 681–691 (2007).

47.   Greenwood, J. R., Calkins, D., Sullivan, A. P. & Shelley, J. C. Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. *Journal of Computer-Aided Molecular Design* **24**, 591–604 (2010).

48.   Friesner, R. A. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry* **47**, 1739–1749 (2004).

49.   Harder, E. *et al.* OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *Journal of Chemical Theory and Computation* **12**, 281–296 (2016).

50.   Kim, S. *et al.* PubChem 2019 update: Improved access to chemical data. *Nucleic Acids Research* **47**, D1102–D1109 (2019).

CHAPTER 5

A New Family of Jumonji C Domain-Containing KDM Inhibitors Inspired by Natural Product Purpurogallin

Recently, our group in collaboration with Professor Altucci, University Luigi Vanvitelli of Naples and Professor R de Lera, University of Vigo reported the identification of a new family of KDM4 subfamily inhibitors inspired by natural product Purpurogallin. In particular, starting from the structure of purpurogallin, a known natural KDM4A inhibitor, derivatives compounds were synthesized and showed improved inhibitory activity against KDM4A *in vitro* and in cell.[1] A previously high-throughput screening (HTS) campaign lead to the identification of purpurogallin as KDM4 inhibitor. Purpurogallin (**1**) is a natural product, member of the benzotropolone-containing family[2–7], and already displayed biological activity including antioxidant and antitumoral activity.[8–10] Furthermore, it has been reported its activity as inhibitor of Toll-like receptor ½ and modulators of mitogen-activated protein kinase 1/2 signaling pathway, inducing regression of esophageal squamous carcinoma.[11,12]

Therefore, in order to explore Purpurogallin scaffold to develop a more potent inhibitors of KDM4A, two sets of Purpurogallin analogues were synthesized and their activity were evaluated. (Table **1**)
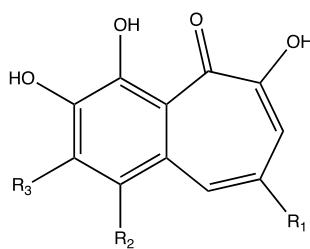
Compound	R$_1$	R$_2$	R$_3$
1	H	H	OH
2	H	Br	COOH
3	COOH	Br	H
4	COOH	Ph	H

Table 1. Most active compounds reported among Purpurogallin analogues.

Among the first set of analogues, compounds **2** and **3 were** tested in a biochemical assay at 50 μM and showed 50% and 80% inhibitory activity against KDM4A, modulating also methylation levels of lysine 9 and 36. To further characterize their activity, compounds 2 and 3 were tested at 25 μM for 24h on MCF7 cells. The results showed that compound 2 increased the methylation status of lysine 9 and 36, whereas compound 3 only affected weakly, the methylation level of lysine 9. The lower activity of compound 3 could be due to the presence of the carboxylic group that could limit its passive diffusion across the cell membrane. Despite these encouraging preliminary results, these compounds did not exhibit activity on cell cycle modification and cell death in cancer cell lines. Then, further modifications were applied on benzotropolone groups in order to improve their antitumoral effect. This modification lead to the discovery of compound **4** in which the benzotropolone scaffold incorporates an aromatic ring at position R_2 of the catechol moiety and a carboxylic group on the pyrogallol moiety. This compound tested at 50 μM, inhibited KDM4A by 80%, exhibiting IC_{50} of 24.37 μM in the primary enzymatic assay and IC_{50} of 10.1 μM in an orthogonal assay based on AlphaLISA technology. Treatment with compound **4** of MCF7 cells showed an increase in the levels of trimethyl-lysine 9 and 36. Furthermore, compound 4 exhibited antitumor activity inducing HCT-116 tumor cells death, block of cell cycle in G2/M phase and also cell death in a time- and dose-dependent manner in two leukemia lines NB4 and U397.

5.1 Molecular docking

In order to understand the structural basis of KDM4A inhibition by most active compounds in the purpurogallin derivatives series, computational studies were performed. Firstly, to deepen the putative binding mode of these compounds, docking studies were performed using Glide in extra precision mode (XP).[11,12] In the docking screening, compounds **4** and **3** were prioritized and showed a docking score of -9.820 and -8.302 Kcal/mol respectively. According to the predicted binding pose, **4** seems to act as 2-oxoglutarate (2-OG) mimic, the natural enzyme substrate. Indeed, as observed for other known KDM4 inhibitors belonging to the class of 2-OG cofactor mimics with metal-chelated moiety (i.e. QC6352 from Celgene),[13] compound **4** mimics the key 2-oxoglutarate interaction. (Figure **1**). In detail, the carboxylic group in the heptane ring forms two H-Bond with Tyr132 and Lys 206, crucial residues involved in the interactions

with the natural cofactor 2-OG, while the carbonyl group directly interact with the Ni metal ion (surrogate in the PDB of the natural Fe in the binding site). The binding pose is further stabilized by two H-bond acceptors involving hydroxyl groups of the catechol moiety as shown in figures **2** and **3**. Compounds **4** and **3** exhibited a similar binding mode. The enhanced potency could be explained considering that the benzene substitution of the *para*-bromine group probably did not affect the binding mode of the central core that maintained the pattern interactions, but contributed to increase the lipophilicity of the compound, allowing an improved cellular permeability.
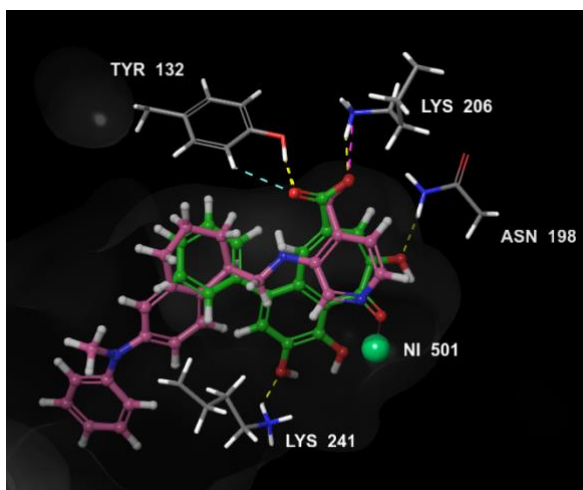

Figure 1. Superimposition of compound 4 (green) with QC6352 (pink), the co-crystallized ligand in the PDB used for our studies (5VGI). As shown in the picture, the carboxyl group of the 4- pyridine carboxylic acid (QC6352) is perfectly overlapped with that of heptane ring, and the both cores retrieved the same crucial interactions, suggesting that compound 4 could act as 2-oxyglutarate mimic.
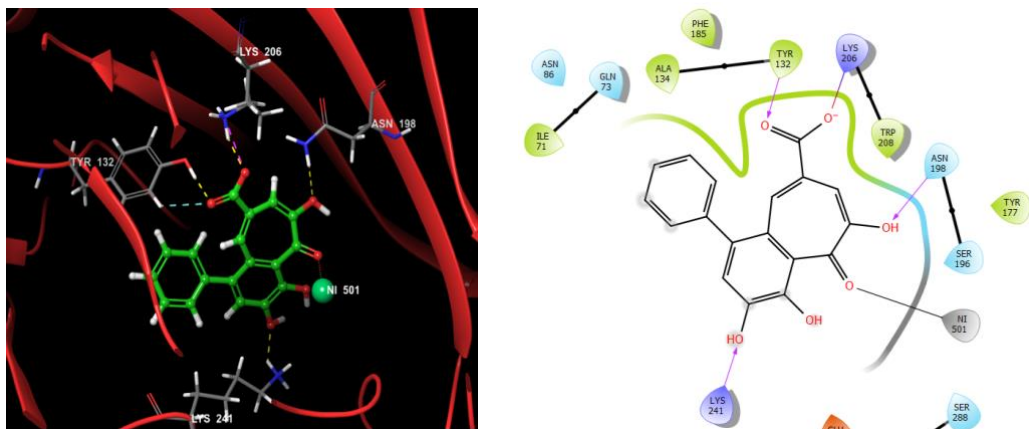
Figure 2. Compound 4 binding pose retrieved from XP docking simulation and 2D ligand interactions.
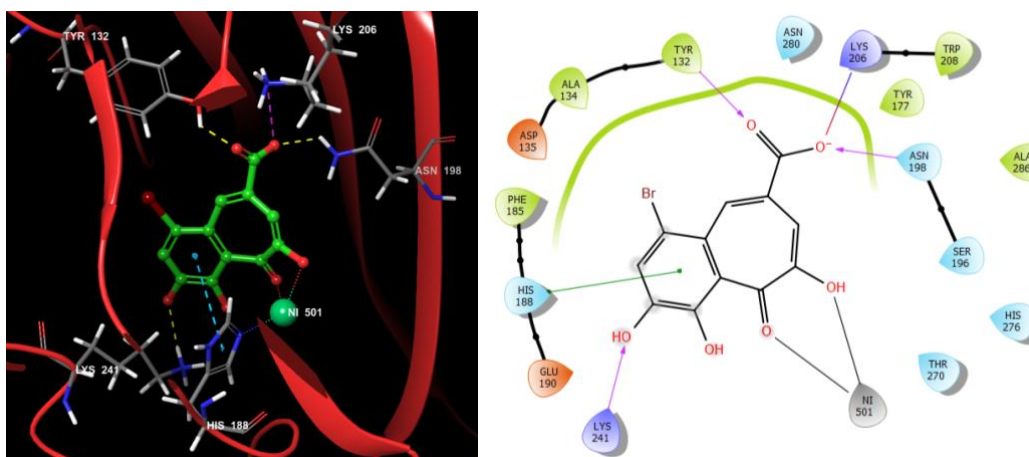

Figure 3. Compound 3 binding pose retrieved from XP docking simulation and 2D ligand interactions.

5.2 MM/GBSA Analysis

In order to evaluate the ligand-binding affinities and estimate accurately the binding free energy of the interaction, starting from XP docking output a MM/GBSA analysis was assessed. The *molecular mechanics generalized Born surface area* (MM/GBSA) is a common approach for binding free energy prediction that has been successfully used to study the binding affinity between small molecules and biological macromolecules. Since it has been proven to be more accurate than most scoring functions of molecular docking, it has been widely employed to improve the results of virtual screening and docking, and to rationalize experimental results. The MM/GBSA analysis confirmed the previous results and showed that also in this experiment compound **4** was prioritized with ΔG= - 27.47 Kcal/mol.[14–16]

5.3 Induced-fit docking

In standard virtual docking studies, ligands are docked into the binding site of a protein and, while the protein is maintained rigid, the ligand is free to move. The limited protein flexibility can generate misleading results since, as already demonstrated, the protein side chains are induced to adapt their position due to a dynamic interaction with the ligand.[17–20] Then, to strengthen the putative binding mode hypothesis, an Induced-fit docking was carried out, focusing on the Purpurogallin structure, compound **1**, and on the most actives compounds from this study, **3** and **4**. Induced-fit docking, also known as flexible docking, adapting atoms vdW radius and side chains position, allows the protein side chain or backbone movements and enables the receptor to change his binding site based on the shape and the binding mode of the ligand. From this experiment, XP binding mode of **4** and **3** were confirmed, we noticed the same key interactions with Tyr 132, Lys 206 and the metal binding with Ni, crucial for the protein activity. (Figure **4**)

Moreover, the induced-fit docking was performed to explore the binding mode of the natural product Purpurogallin (**1**), and the pose reported is different compared with **3** and **4**. (Figure **4C**) The absence of the carboxylic group in the heptane ring seems to lead a conformational change in the ligand pose. However, the interaction with Ty132 and the metal ion were retained, which could explain the purpurogallin inhibitory activity, even if weak. This modification, according to the assays results, could suggest that the

carboxylic group is crucial for the correct ligand orientation and interactions stability that can be translated in an activity modulation.
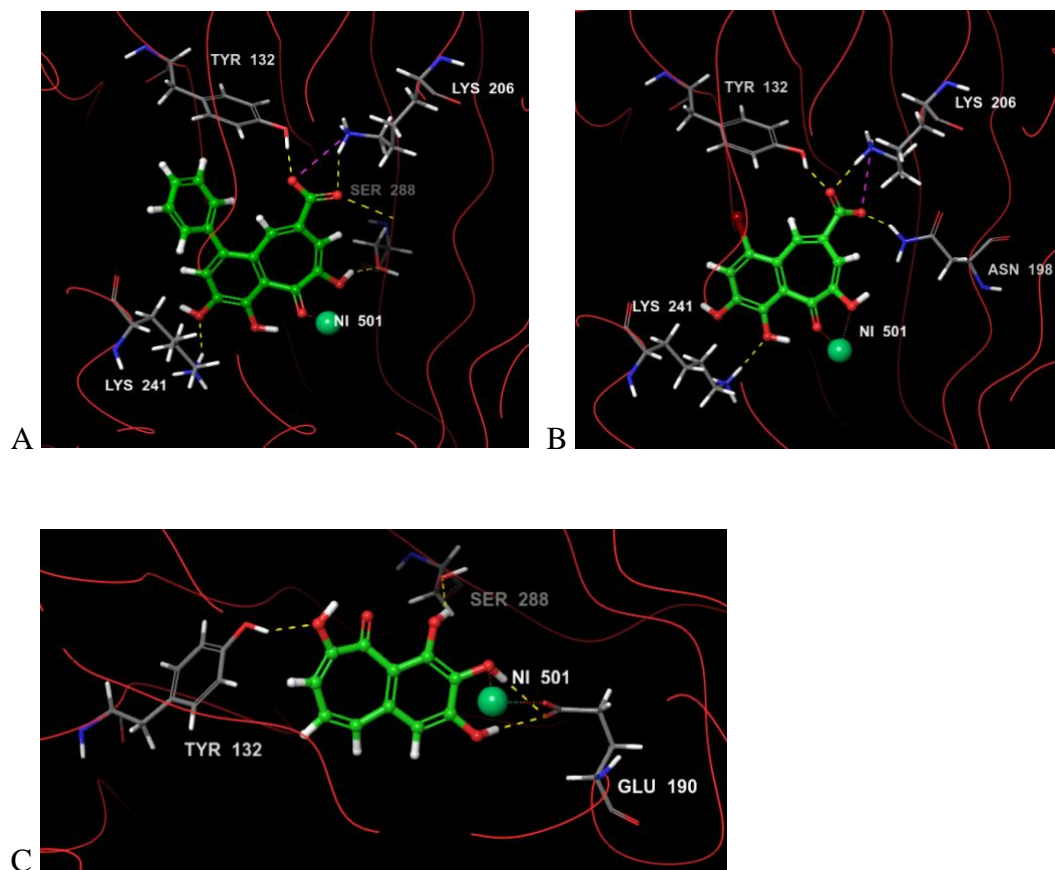

Figure 4. Induced-fit docking pose of compound 4 (A), compound 3 (B) and purpurogallin (C)

5.4 Material and methods

Ligands preparation

The ligands were prepared by LigPrep, a tool of MAESTRO software, using OPLS3e as force field[21–23] and ionizer to generate all the possible states at pH= 7.0 ± 0.2. Desalt and generate tautomers were flagged on, the chirality were retained and at most 32 conformers per ligand were generated.

Protein preparation

A crystal structure of the protein KDM4A in complex with a ligand (PDB ID: 5VGI, Resolution: 2,7 Å) was chosen from the Protein Data Bank (PDB)[24] for the set-up of our model. Protein was prepared via Protein Preparation Wizard Maestro tool.[10] Epik was used to generate het states at pH= 7.4 ± 0.2, deleting waters beyond 5.00 Å and then side chains protonation state was optimized using ProPKa3 at pH=7.4. [25,26]

Docking grid generation and ligands docking

The grid was generated with Glide, setting a scaling factor 1.0, and identifying the grid centroid onto the ligand. Molecular docking was carried out using Glide software released by Schrödinger (release 2018-4).[27] The simulation was performed in extra precision (XP), using OPLS3e as force field, considering ligands as Flexible and including Epik state penalties to docking score. In this step, the vdW radii scaling factor was set as 0.8 and the partial charge cut off was of 0.15. The searching algorithm on our model was tested using a cognate docking of the co-crystallized ligand, and we obtained an RMSD value of 0.5813. XP docking were performed both with constraints. As constraints 2 H-Bond on Tyr132 and Lys 206, a positional and a metal constraints were fixed, setting 2 of them as mandatory.

MM GBSA

The binding affinity was also investigated, starting from XP docking poses with Prime MM-GBSA by Schrodinger software, using VSGB as solvation model and OPLS3e as force field. [28]

Induced-fit docking

Induced-fit docking[28] was carried out. A standard protocol was used, generating up to 20 poses. (Sample ring conformations with an energy window of 2.5Kcal/mol). The receptor and ligand van der Waals scaling factors were set at 0.5. Prime was used for the refinement and the redocking was performed in extra precision.

Bibliography

1.   Souto, J. A. *et al.* A New Family of Jumonji C Domain-Containing KDM Inhibitors Inspired by Natural Product Purpurogallin. *Frontiers in Chemistry* **8**, 1–20 (2020).

2.   Nicholson, J. S. Bavltvop a!vd Nicholsovt : The Oxidatiovt Products of Pheutols. 30. The Oxidation Products. 116–120

3.   Matsuo, Y., Yoshida, A., Saito, Y. & Tanaka, T. Structural Revision and Biomimetic Synthesis of Goupiolone B. *Angewandte Chemie - International Edition* **56**, 11855–11859 (2017).

4.   Kerschensteiner, L., Löbermann, F., Steglich, W. & Trauner, D. Crocipodin, a benzotropolone pigment from the mushroom Leccinum crocipodium (Boletales). *Tetrahedron* **67**, 1536–1539 (2011).

5.   Klostermeyer, D., Knops, L., Sindlinger, T., Polborn, K. & Steglich, W. Novel benzotropolone and 2H-furo[3,2-b]benzopyran-2-one pigments from Tricholoma aurantium (Agaricales). *European Journal of Organic Chemistry* **13**, 603–609 (2000).

6.   Takino, Y., Imagawa, H., Horikawa, H. & Tanaka, A. Studies on the mechanism of the oxidation of tea leaf catechins. *Agricultural and Biological Chemistry* **28**, 64–71 (1964).

7.   Takino, Y. & Imagawa, H. Crystalline reddishi orange pigment of manufactured black tea. *Agricultural and Biological Chemistry* **28**, 255–256 (1964).

8.   Wu, T. W. *et al.* Molecular structure and antioxidant specificity of purpurogallin in three types of human cardiovascular cells. *Biochemical Pharmacology* **52**, 1073–1080 (1996).

9.   Kitada, S. *et al.* Discovery, characterization, and structure-activity relationships studies of proapoptotic polyphenols targeting B-cell lymphocyte/leukemia-2 proteins. *Journal of Medicinal Chemistry* **46**, 4259–4264 (2003).

10.   Leone, M. *et al.* Cancer Prevention by Tea Polyphenols Is Linked to Their Direct Inhibition of Antiapoptotic Bcl-2-Family Proteins. *Cancer Research* **63**, 8118–8121 (2003).

11.   Cheng, K., Wang, X., Zhang, S. & Yin, H. Discovery of small-molecule inhibitors of the TLR1/TLR2 complex. *Angewandte Chemie - International Edition* **51**, 12246–12249 (2012).

12.   Xie, X. *et al.* Purpurogallin is a novel mitogen-activated protein kinase kinase 1/2 inhibitor that suppresses esophageal squamous cell carcinoma growth in vitro and in vivo. *Molecular Carcinogenesis* **58**, 1248–1259 (2019).

13.   Chen, Y. K. *et al.* Design of KDM4 Inhibitors with Antiproliferative Effects in Cancer Models. *ACS Medicinal Chemistry Letters* **8**, 869–874 (2017).

14.   Genheden, S. & Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opinion on Drug Discovery* **10**, 449–461 (2015).

15.   Wang, E. *et al.* End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chemical Reviews* **119**, 9478–9508 (2019).

16.   Ylilauri, M. & Pentikäinen, O. T. MMGBSA as a tool to understand the binding affinities of filamin-peptide interactions. *Journal of Chemical Information and Modeling* **53**, 2626–2633 (2013).

17.   Barreca, M. L., Iraci, N., De Luca, L. & Chimirri, A. Induced-fit docking approach provides insight into the binding mode and mechanism of action of HIV-1 integrase inhibitors. *ChemMedChem* **4**, 1446–1456 (2009).

18.   Nabuurs, S. B., Wagener, M. & De Vlieg, J. A flexible approach to induced fit docking. *Journal of Medicinal Chemistry* **50**, 6507–6518 (2007).

19.   Bolia, A. & Ozkan, S. B. Adaptive BP-Dock: An Induced Fit Docking Approach for Full Receptor Flexibility. *Journal of Chemical Information and Modeling* **56**, 734–746 (2016).

20.   Xu, M. & Lill, M. A. Induced fit docking, and the use of QM/MM methods in docking. *Drug Discovery Today: Technologies* **10**, e411–e418 (2013).

21.   Jorgensen, W. L. & Tirado-Rives, J. The OPLS Potential Functions for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *Journal of the American Chemical Society* **110**, 1657–1666 (1988).

22.   Roos, K. *et al.* OPLS3e: Extending Force Field Coverage for Drug-Like Small Molecules. *Journal of Chemical Theory and Computation* **15**, 1863–1874 (2019).

23.   Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society* **118**, 11225–11236 (1996).

24.   Berman, H. M. *et al.* The protein data bank. *Acta Crystallographica Section D: Biological Crystallography* **58**, 899–907 (2002).

25.   Shelley, J. C. *et al.* Epik: A software program for pKa prediction and protonation state generation for drug-like molecules. *Journal of Computer-Aided Molecular Design* **21**, 681–691 (2007).

26.   Greenwood, J. R., Calkins, D., Sullivan, A. P. & Shelley, J. C. Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. *Journal of Computer-Aided Molecular Design* **24**, 591–604 (2010).

27.   Friesner, R. A. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry* **47**, 1739–1749 (2004).

28.   Li, J. *et al.* The VSGB 2.0 model: A next generation energy model for high resolution protein structure modeling. *Proteins: Structure, Function and Bioinformatics* **79**, 2794–2812 (2011).

Part II

Dynamic-Shared Pharmacophore Approach As Tool To Design New Allosteric PRC2 Inhibitors, Targeting EED Binding Pocket.

CHAPTER 6

Polycomb Repressive Complex 2

Histone methyl transferases (HMTs) catalyse the addition of methyl group donated from S-adenosylmethionine to histones.[1] Among HMTs three families of enzymes have been identified, the SET-domain- containing proteins[1] and DOT1-like proteins[2] that methylate lysine, and the protein arginine N-methyltransferase (PRMT) family[3] that acts methylating arginine residues.

The Polycomb Repressive Complex 2 (PRC2) is a lysine methyltransferase (KMT) that catalyses the three sequential methylation trimethylation of H3K27 and belongs to the Polycomb group proteins that are transcriptional repressors in eukaryotes. PRC2 complex comprised multiple subunits: Enhancer of zeste homolog 2 (EZH2), embryonic ectoderm development (EED), suppressor of zeste 12 (SUZ12), histone-binding protein RBBP4 (RbAp48) and Adipocyte Enhancer-Binding Protein 2 (AEBP2). (Figure **1**) EZH2 is the catalytic subunit of the complex and using *S*-adenosil-methionine (SAM) as the cofactor, acts as methyltransferase mainly through its C-terminal SET domain, which has been found in several KMTs.[4–7] Efficient histone methylation activity requires the cooperation of multiple PRC2 components, including EZH2, SUZ 12 and EED.[8–11] The RbAp48 and AEBP2 are not essential for methyltransferase activity but probably help to stabilize the overall architecture of the complex. PRC2 is involved in multiple biological processes, including cell cycle regulation[12], apoptosis[13], and regulation of cellular senescence.[14] Indeed, PRC2 deregulation has been linked to several diseases, including cancer.
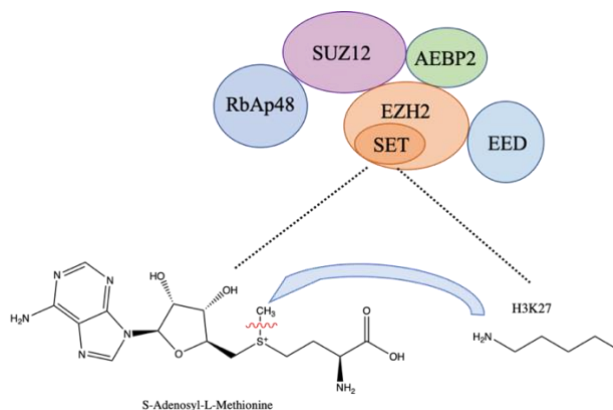

Figure 1. PRC2 subunits and mechanism of action

6.1 Structure of PRC2 complex

As shown from the resolved crystal structure depicted in Figure **2** (PDB ID: 5IJ7), the active complex is formed by EED subunit wrapped around by EZH2 and SUZ 12, positioned between EED and the SET domain of EZH2.
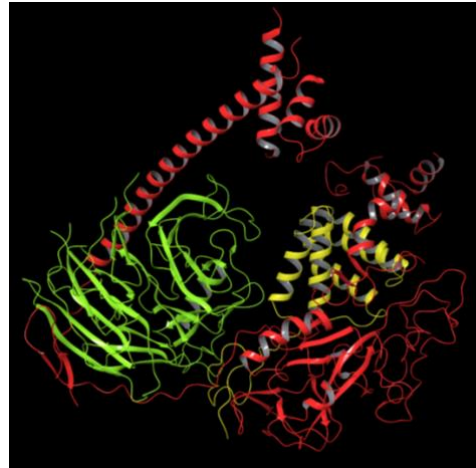

Figure2. PRC2 complex: EED in green; EZH2 in red and SUZ12 in yellow (PDB ID: 5IJ7)

EZH2, the catalytic subunit, is formed by multiple domains, including EED binding domain (EBD) and cysteine-rich domain (CXC domain) in the N-terminal region (residues 1-259), while suppressor of variegation 39, enhance of zeste and trithorax domain (SET domain) together with an inserted SET (iSET) and a C-terminal set (cSET) in the C-terminal region (residues 259-742). The N-terminal domain is not involved in the catalytic activity but is required for the structural integrity of the complex. [15,16] The cSET region in the isolated EZH2 subunit showed a disordered state distant from the cofactor binding site, that avoid the substrate binding and thus the catalytic activity.

The EED domain interacts with EZH2 through the EBD, that is composed by 30 -residues (residues 39-68).[17] The interactions between EED-EZH2 subunits are mainly represented by hydrogen bonds and van der Waals interactions. The recognition of H3K27Me3 by EED was found to be crucial for the PRC2 activity, enabling efficient methylation activity of the complex. Indeed, the peptide binding results in stabilization of stimulation-responsive motif (SRM) in EED that induces a conformational changes in the cSET of EZH2 leading to a repositioning of residues in an ordinated state and allowing the

catalytic activity.[15,17,18] Then, the H3K27 binding with EED mediates the allosteric activation of EZH2. Mutations in EED that prevent the recognition of its substrate results in inhibition of PRC2 activity.

6.2 Histone substrate specificity

The crystal structures resolved, indicated that EED specifically binds to histone methylation sites involved in gene silencing including, H1K26me3, H3K9me3, H3K27me3, and H4K20me3. Inside the EED subunit the trimethyl-lysine, is accommodated into an aromatic cage formed by Phe97, Tyr148, and Tyr365 via π-cation interactions. A comparison between the repressive histone methyl marks showed that H1K26me3, H3K9me3, and H3K27me3 exhibit an alanine residue two amino acids N-terminal to lysine, which fit into a small hydrophobic pocket on the surface of EED. This pocket is small and any group bulkier than alanine cannot bind it. While H4K20me3 bind EED through its leucine in position +2 of lysine that fit with a second hydrophobic pocket. The ability to bind one of these small pockets establishes the basis for EED recognition. Notably, the histone methyl marks associated to active transcription in position +2 and -2 contain a bulky groups, indicating that the selective binding of EED with repressive histone marks allosterically activate the PRC2 complex.[19]

Mutations have been reported on EZH2 sequence, such as tyrosine 641 (Tyr641) residue in the catalytic domain, including Tyr641F, Tyr641N, Tyr641S, Tyr641H, Tyr641C. The wild type have a substrate preference for unmethylated H3K27 and mono-methylated H3K27, while Tyr641 mutants showed a substrate specificity for di-methylated H3K27.[20,21]

6.3 PRC2 Role in cancer

EZH2 plays a pivotal role in cell cycle regulation, then dysregulation of its activity can lead to carcinogenesis and cancer development. Overexpression or mutations of EZH2 were found in several malignancies and its expression is correlated with aggressiveness, metastasis, and poor prognosis.[22]

For instance, overexpression of EZH2 was found in breast cancer and high levels of this enzyme has been linked to invasiveness and increased proliferation rate of breast carcinoma.[23,24] EZH2 is overexpressed in prostate cancer and also in this case is

associated with poor prognosis. High levels of EZH2 were reported in a majority of aggressive or metastatic prostate cancers and loss of its activity lead to tumour regression.[25,26] In gastric cancer overexpression of EZH2 is associated with tumour size, invasion, TNM stage, poor prognosis and poor disease-free survival.[26] Furthermore several studies linked EZH2 overexpression to additional types of cancer, such as endometrial carcinoma[27], lung cancer, lymphomas and oesophageal cancer[28]. EZH2 showed a fundamental role in cancer cells metastasis through regulation of epithelial-mesenchymal transition (EMT) tumour angiogenesis processes.[29,30] Knockdown of EZH2 in human lung cancer cell lines and in melanoma-positive model mice resulted in significant reduction of migration and invasion ability.[31,32]

Moreover, mutations of EZH2 were identified in hematologic malignancies, resulting in enhancement of EZH2 activity. The most frequent mutation occurs at Tyr641 and was found in different type of lymphoma including germinal centre B cell lymphoma, diffuse large B-cell lymphomas (DLBCL) and follicular lymphomas.[33–35]

Furthermore, EZH2 could suppress anti-cancer immunity in cancer cells.[36,37] Inhibition of EZH2 activity in T-cells resulted in enhancement of Th1, Th2 and cell plasticity[38], improvement of mature natural killer (NK) cells activity[39] and modification of microenvironment of tumour cells, enhancing the recruitment of T cell effectors, CD8+ and CD4+, leading to a robust anti-tumour immunity.[40]

6.4 PRC2 Inhibitors

Modulation of PRC2 activity became an attractive goal for cancer treatment, due to its broad involvement in cancer pathophysiology. The PRC2 methyltransferase activity can be inhibited through different mechanism of action, including acting as SAM, natural cofactor, competitive inhibitors, breaking the PRC2 structure disrupting interactions through EZH2 with EED or SUZ12, and suppressing EZH2 activity through triggering EZH2 degradation (Table **1**) Over the past years several PRC2 inhibitors have been discovered and several compounds are ongoing in clinical trials.[41] The 3-deazaneplanocin (DZNep) was the first PRC2 inhibitor reported, that increases the level of S-adenosyl-homocysteine (SAH) and inhibits the SAM-dependent methyltransferase activity.[42]

- **SAM-competitive inhibitors**

UNC1999 was the first oral SAM-competitors in both wild-type and Y641 mutant EZH2, that showed inhibition of MLL-rearranged leukemia in mice.[43,44]

GSK 126 (GSK2816126) is a potent and selective inhibitor of wild type and Tyr641(Y641) mutant EZH2 developed by Glaxo. The GSK126 efficacy, safety, pharmacokinetic and pharmacodynamic profile were evaluated in patients with advanced hematologic and solid tumours, in phase 1 clinical trial (NCT02082977). The trial was terminated showing insufficient evidence of clinical activity.[45,46]

EPZ005687, another potent SAM-competitor inhibitor, during clinical trial showed unsuitable pharmacokinetic profile that limited its clinical application.[47]

EI1 inhibits selectively EZH2 activity in diffused large B-cell lymphomas cells carrying the Y641 mutations, leading to decreasing of proliferation, cell cycle arrest, and apoptosis.[48]Glaxo Smith Kline developed other two potent EZH2 inhibitors, GSK343 and GSK926 that showed high affinity and selectivity in vitro but also substandard pharmacokinetics properties in rat model.[49] Epizime identified different potent EZH2 SAM-competitor inhibitors, including EPZ005687, EPZ011989 and EPZ6438 or Tazemtostat.[50–52] The latter compound, Tazemetostat, was recently approved by FDA for adult, paediatric and older patients with advance metastatic epithelioid sarcoma not eligible for complete resection. Meanwhile this compound is under evaluation in different clinical trials showing efficacy and safety profile in other hematologic tumour such relapsed/refractory follicular lymphoma.[53]
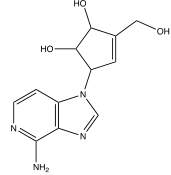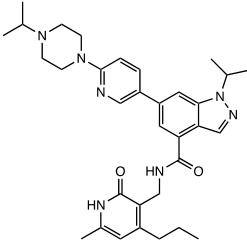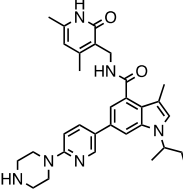
CIP-1205 is an indole-based compound and a potent and orally available EZH2 inhibitor, that was evaluated in phase 1 clinical trial for B-cell lymphoma. (NCT02395601).[54] This compound was optimized from CIP-169, that showed anti-tumour activity in a mouse xenograft model of a KARPAS-422 lymphoma but weak oral bioavailability.[55] Furthermore, CIP-1205 is undergoing phase1/2 clinical trial for advance solid tumours (NCT03525795) and phase1/2 clinical trial for metastatic castration-resistant prostate cancer (NCT03480646). ZLD1039, exhibited selectivity, oral bioavailability and regression of breast cancer and metastasis.[56] PF-06821497 is undergoing a phase 1 clinical trial in patients with relapsed/refractory small cell lung cancer (SCLC), castration-resistant prostate cancer (CRPC), and DLBCL (NCT03460977).[57]
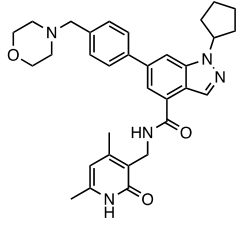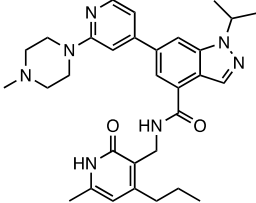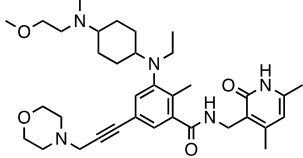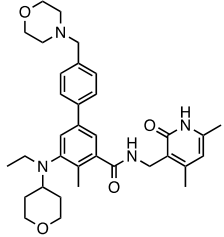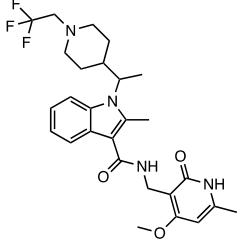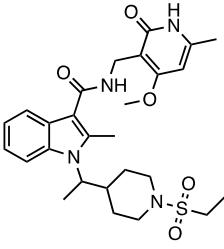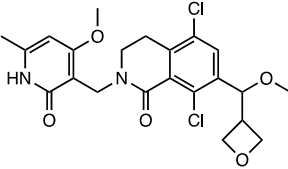
- **PRC2 disruptors**

An alternative strategy to targeting the enzyme catalytic domain is to disrupt the protein-protein interaction between EZH2 and the other two subunits EED or SUZ12.

SAH-EZH2 is a peptide that targets EED and disrupt its interaction with EZH2, showing an increase of H3K27 levels and inhibition of growth and differentiation of MLL-AF9 leukemic cells.[58] Astemizole, is a FDA approved antagonist of H1 receptor and exhibited capability to break the EED-EZH2 interaction, arresting the proliferation of PRC2-driven lymphoma.[59]

Novartis developed different EED inhibitors targeting the EED binding pocket that results in PRC2 allosteric inhibition.[60] Among these inhibitors EED226 (or MAK683) in currently undergoing phase ½ clinical trial for advanced malignancies including sarcoma, DLBCL, gastric cancer, ovarian cancer, nasopharyngeal carcinoma and prostate cancer (NCT02900651).[61]

Compound	Structure	Mechanism and activity
DZNep		SAM Competitor K_i=50 pm
UNC1999		SAM Competitor IC_{50}= 2 nM
GSK126		SAM Competitor IC_{50}= 9.9 nM

EPZ005687		SAM Competitor IC$_{50}$= 24 nM
GSK343		SAM Competitor IC$_{50}$= 4 nM
EPZ011989		SAM Competitor IC$_{50}$= <3 nM
EPZ6438/Tazemetostat		SAM Competitor IC$_{50}$= 2.5 nM
CIP-1205		SAM Competitor IC$_{50}$= 2 nM
CIP-169		SAM Competitor IC$_{50}$= 0.24 nM
PF-06821497		SAM Competitor IC$_{50}$= 17 nM
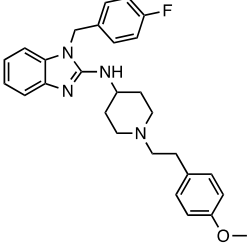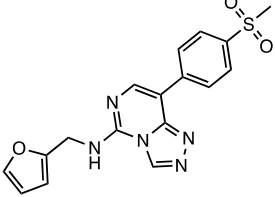
Astemizole		EZH2-EED binder IC$_{50}$= 93.80 uM
EED226/MAK 683		EED binder IC$_{50}$= 23.4 nM

Table 1 PRC2 complex inhibitors: structures activity and mechanism of action

Bibliography

1.   Jurkowska, R. Z. & Jeltsch, A. *Structure and Function of Human DNA Methyltransferases. Reference Module in Biomedical Sciences* (Elsevier, 2014). doi:10.1016/b978-0-12-801238-3.00093-3

2.   Rea, S. *et al.* Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**, 593–599 (2000).

3.   Feng, Q. *et al.* Methylation of H3-Lysine 79 Is Mediated by a New Family of HMTases without a SET Domain University of North Carolina at Chapel Hill. *Current* **12**, 1052–1058 (2002).

4.   Czermin, B. *et al.* Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* **111**, 185–196 (2002).

5.   Herz, H. M., Garruss, A. & Shilatifard, A. SET for life: Biochemical activities and biological functions of SET domain-containing proteins. *Trends in Biochemical Sciences* **38**, 621–639 (2013).

6.   Müller, J. *et al.* Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell* **111**, 197–208 (2002).

7.   Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P. & Reinberg, D. Histone methyltransferase activity associated with a human multiprotein complex containing the enhancer of zeste protein. *Genes and Development* **16**, 2893–2905 (2002).

8.   Cao, R. & Zhang, Y. SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. *Molecular Cell* **15**, 57–67 (2004).

9.   Cao, R. *et al.* Role of histone H3 lysine 27 methylation in polycomb-group silencing. *Science* **298**, 1039–1043 (2002).

10.   Montgomery, N. D. *et al.* The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. *Current Biology* **15**, 942–947 (2005).

11.   Montgomery, N. D., Yee, D., Montgomery, S. A. & Magnuson, T. Molecular and Functional Mapping of EED Motifs Required for PRC2-Dependent Histone Methylation. *Journal of Molecular Biology* **374**, 1145–1157 (2007).

12.   State Key Laboratory of Stem Cell and Reproductive Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing, 100101, China, 2 University of Chinese Academy of Sciences, Beijing, 100049, China. **12**, 947–960 (2019).

13.   Yao, Y. *et al.* Downregulation of enhancer of zeste homolog 2 (EZH2) is essential for the induction of autophagy and apoptosis in colorectal cancer cells. *Genes* **7**, (2016).

14.   Ito, T., Teo, Y. V., Evans, S. A., Neretti, N. & Sedivy, J. M. Regulation of Cellular Senescence by Polycomb Chromatin Modifiers through Distinct DNA Damage- and Histone Methylation-Dependent Pathways. *Cell Reports* **22**, 3480–3492 (2018).

15.   Antonysamy, S. *et al.* Structural context of disease-associated mutations and putative mechanism of autoinhibition revealed by X-Ray crystallographic analysis of the EZH2-SET domain. *PLoS ONE* **8**, 1–15 (2013).

16.   Wu, H. *et al.* Structure of the catalytic domain of EZH2 reveals conformational plasticity in cofactor and substrate binding sites and explains oncogenic mutations. *PLoS ONE* **8**, 1–12 (2013).

17.   Han, Z. *et al.* Structural Basis of EZH2 Recognition by EED. *Structure* **15**, 1306–1315 (2007).

18.   Justin, N. *et al.* Structural basis of oncogenic histone H3K27M inhibition of human polycomb repressive complex 2. *Nature communications* **7**, 11316 (2016).

19.   Margueron, R. *et al.* Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* **461**, 762–767 (2009).

20.   Sneeringer, C. J. *et al.* Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 20980–20985 (2010).

21.   Yap, D. B. *et al.* Somatic mutations at EZH2 Y641 act dominantly through a mechanism of selectively altered PRC2 catalytic activity, to increase H3K27 trimethylation. *Blood* **117**, 2451–2459 (2011).

22.   Simon, J. A. & Lange, C. A. Roles of the EZH2 histone methyltransferase in cancer epigenetics. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis* **647**, 21–29 (2008).

23.   Kleer, C. G. *et al.* EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 11606–11611 (2003).

24.   Bachmann, I. M. *et al.* EZH2 expression is associated with high proliferation rate and aggressive tumor subgroups in cutaneous melanoma and cancers of the endometrium, prostate, and breast. *Journal of Clinical Oncology* **24**, 268–273 (2006).

25.   Varambally, S. *et al.* The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* **419**, 624–629 (2002).

26.   Wu, X. *et al.* Increased EZH2 expression in prostate cancer is associated with metastatic recurrence following external beam radiotherapy. *Prostate* **79**, 1079–1089 (2019).

27.   Krill, L. *et al.* Overexpression of enhance of Zeste homolog 2 (EZH2) in endometrial carcinoma: An NRG Oncology/Gynecologic Oncology Group Study. *Gynecologic Oncology* **156**, 423–429 (2020).

28.   Qiu, B. Q. *et al.* Long non-coding RNA PSMA3-AS1 promotes malignant phenotypes of esophageal cancer by modulating the miR-101/EZH2 axis as a ceRNA. *Aging* **12**, 1843–1856 (2020).

29.   Ma, J., Zhang, J., Weng, Y. C. & Wang, J. C. EZH2-mediated microRNA-139-5p regulates epithelial-mesenchymal transition and lymph node metastasis of pancreatic cancer. *Molecules and Cells* **41**, 868–880 (2018).

30.   Lu, C. *et al.* Regulation of tumor angiogenesis by EZH2. *Cancer Cell* **18**, 185–197 (2010).

31.   Zingg, D. *et al.* The epigenetic modifier EZH2 controls melanoma growth and metastasis through silencing of distinct tumour suppressors. *Nature Communications* **6**, (2015).

32.   Xia, L. *et al.* EZH2 enhances expression of CCL5 to promote recruitment of macrophages and invasion in lung cancer. *Biotechnology and Applied Biochemistry* 1–9 (2019). doi:10.1002/bab.1875

33.   Morin, R. D. *et al.* Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nature Genetics* **42**, 181–185 (2010).

34.   Van Kemenade, F. J. *et al.* Coexpression of BMI-1 and EZH2 polycomb-group proteins is associated with cycling cells and degree of malignancy in B-cell non-Hodgkin lymphoma. *Blood* **97**, 3896–3901 (2001).

35.   Morin, R. D. *et al.* Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* **476**, 298–303 (2011).

36.   Nutt, S. L., Keenan, C., Chopin, M. & Allan, R. S. EZH2 function in immune cell development. *Biological Chemistry* **401**, 933–943 (2020).

37.   Dong, H. *et al.* An allosteric PRC2 inhibitor targeting EED suppresses tumor progression by modulating the immune response. *Cancer Research* **79**, 5587–5596 (2019).

38.   Tumes, D. J. *et al.* The Polycomb protein Ezh2 regulates differentiation and plasticity of CD4+ T helper Type 1 and type 2 cells. *Immunity* **39**, 819–832 (2013).

39.   Yin, J. *et al.* Ezh2 regulates differentiation and function of natural killer cells through histone methyltransferase activity. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 15988–15993 (2015).

40.   Wang, D. *et al.* Targeting EZH2 Reprograms Intratumoral Regulatory T Cells to Enhance Cancer Immunity. *Cell Reports* **23**, 3262–3274 (2018).

41.   Duan, R., Du, W. & Guo, W. EZH2: A novel target for cancer treatment. *Journal of Hematology and Oncology* **13**, 1–12 (2020).

42.   Miranda, T. B. *et al.* DZNep is a global histone methylation inhibitor that reactivates developmental genes not silenced by DNA methylation. *Molecular Cancer Therapeutics* **8**, 1579–1588 (2009).

43.   Konze, K. D. *et al.* An orally bioavailable chemical probe of the lysine methyltransferases EZH2 and EZH1. *ACS Chemical Biology* **8**, 1324–1334 (2013).

44.   Xu, B. *et al.* Selective inhibition of EZH2 and EZH1 enzymatic activity by a small molecule suppresses MLL-rearranged leukemia. *Blood* **125**, 346–357 (2015).

45.   McCabe, M. T. *et al.* EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. *Nature* **492**, 108–112 (2012).

46.   Yap, T. A. *et al.* Phase I study of the novel enhancer of zeste homolog 2 (EZH2) inhibitor GSK2816126 in patients with advanced hematologic and solid tumors. *Clinical Cancer Research* **25**, 7331–7339 (2019).

47.   Knutson, S. K. *et al.* A selective inhibitor of EZH2 blocks H3K27 methylation and kills mutant lymphoma cells. *Nature Chemical Biology* **8**, 890–896 (2012).

48.   Qi, W. *et al.* Selective inhibition of Ezh2 by a small molecule inhibitor blocks tumor cells proliferation. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 21360–21365 (2012).

49.   Verma, S. K. *et al.* Identification of potent, selective, cell-Active inhibitors of the histone lysine methyltransferase EZH2. *ACS Medicinal Chemistry Letters* **3**, 1091–1096 (2012).

50.   Knutson, S. K. *et al.* Selective inhibition of EZH2 by EPZ-6438 leads to potent antitumor activity in EZH2-mutant non-Hodgkin lymphoma. *Molecular Cancer Therapeutics* **13**, 842–854 (2014).

51.   Campbell, J. E. *et al.* EPZ011989, A potent, orally-available EZH2 inhibitor with robust in vivo activity. *ACS Medicinal Chemistry Letters* **6**, 491–495 (2015).

52.   Knutson, S. K. *et al.* Durable tumor regression in genetically altered malignant rhabdoid tumors by inhibition of methyltransferase EZH2. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 7922–7927 (2013).

53.   Italiano, A. Targeting epigenetics in sarcomas through EZH2 inhibition. *Journal of Hematology and Oncology* **13**, 10–12 (2020).

54.   Vaswani, R. G. *et al.* Identification of ®-N-((4-Methoxy-6-methyl-2-oxo-1,2-dihydropyridin-3-yl)methyl)-2-methyl-1-(1-(1-(2,2,2-trifluoroethyl)piperidin-4-

yl)ethyl)-1H-indole-3-carboxamide (CPI-1205), a Potent and Selective Inhibitor of Histone Methyltransferase EZH2, Suitable for Phase i Clinical Trials for B-Cell Lymphomas. *Journal of Medicinal Chemistry* **59**, 9928–9941 (2016).

55.   Gehling, V. S. *et al.* Discovery, design, and synthesis of indole-based EZH2 inhibitors. *Bioorganic and Medicinal Chemistry Letters* **25**, 3644–3649 (2015).

56.   Song, X. *et al.* Selective inhibition of EZH2 by ZLD1039 blocks H3K27methylation and leads to potent anti-tumor activity in breast cancer. *Scientific Reports* **6**, (2016).

57.   Kung, P. P. *et al.* Optimization of Orally Bioavailable Enhancer of Zeste Homolog 2 (EZH2) Inhibitors Using Ligand and Property-Based Design Strategies: Identification of Development Candidate ®-5,8-Dichloro-7-(methoxy(oxetan-3-yl)methyl)-2-((4-methoxy-6-methyl-2-oxo-1,2-dihydropyridin-3-yl)methyl)-3,4-dihydroisoquinolin-1(2H)-one (PF-06821497). *Journal of Medicinal Chemistry* **61**, 650–665 (2018).

58.   Kim, W. *et al.* Targeted disruption of the EZH2-EED complex inhibits EZH2-dependent cancer. *Nature Chemical Biology* **9**, 643–650 (2013).

59.   Kong, X. *et al.* Astemizole arrests the proliferation of cancer cells by disrupting the EZH2-EED interaction of polycomb repressive complex 2. *Journal of Medicinal Chemistry* **57**, 9512–9521 (2014).

60.   Li, L. *et al.* Discovery and molecular basis of a diverse set of polycomb repressive complex 2 inhibitors recognition by EED. *PLoS ONE* **12**, 1–18 (2017).

61.   Qi, W. *et al.* An allosteric PRC2 inhibitor targeting the H3K27me3 binding pocket of EED. *Nature Chemical Biology* **13**, 381–388 (2017).

CHAPTER 7

Dynamic-Shared Pharmacophore Approach As Tool To Design New Allosteric PRC2 Inhibitors, Targeting EED Binding Pocket.

The EZH2 binding protein has been widely studied. The majority of SAM-competitive inhibitors contains a pyridone scaffold that fits perfectly in the catalytic core. Several inhibitors are currently undergoing clinical trials. However some of them resulted in drug resistance and presented several side effects, probably due to off-target activity, since the SET domain is shared with multiple human enzymes.[1–3] On the other hand, the allosteric inhibitors of PRC2 should be more selective because they act specifically modulating EED-EZH2 interaction. It has been shown that inhibition of H3K27 binding with the allosteric site EED prevents the EED capability to bind EZH2, resulting in inactivation of the whole PRC2. Then, the design of molecules acting on EED subunit would be an interesting approach in modulating PRC2 functions.[4,5] As shown in the crystal structures of EED in complex with the inhibitor EED226 (PDB ID: 5H15, Figure **1**), the EED binding pocket is located on the top surface, close to the solvent-exposed region and comprises a hydrophobic region composed by an aromatic cage involving residues such as Tyr365, Tyr148, Phe97 and Arg367. The hydrophobic residues in the aromatic cage of EED play a pivotal role for the binding to repressive histone marks and then for the PRC2 complex activity.[6]
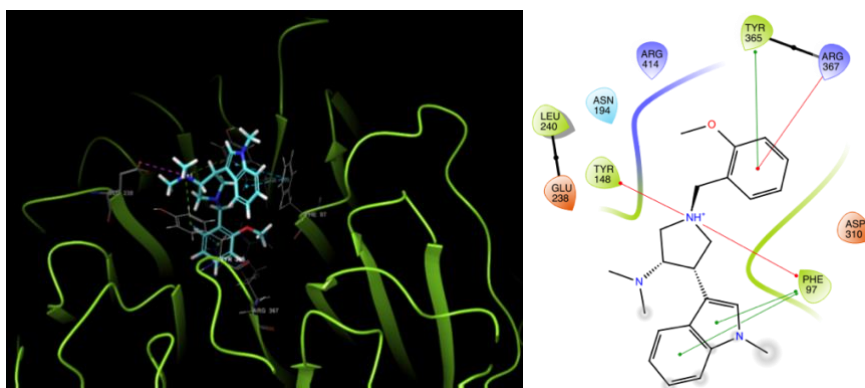

Figure 1. Known EED inhibitor (EED226) binding mode: EED226 in blue and EED in green (PDB ID: 5H15)

In my thesis work, a Computer-Aided Drug Design method based on the allosteric modulation of PRC2 was set-up using molecular dynamics and pharmacophore modeling to identify eventual hidden protein-ligand interactions. The use of this approach is increasingly applied in drug design. For instance in MYSHAPE (Molecular dYnamics SHAred PharmacophorE) workflow, the use of dynamic pharmacophore model has shown several advantages compared to classical methodologies.[7,8] The advantage of a dynamic system in drug design, relies on the possibility to overcome the bias generated from the consideration of the protein as a static structure.[9,10, 11] The concept of the protein as rigid complex has been superseded since the evidence that the protein is a flexible macromolecule that can incur into conformational changes and then affecting the ligand binding pose.[12,13] The molecular dynamic approach has proven to be useful to explore the protein conformational space and has been employed with other classic computational techniques, such as docking or pharmacophore modeling, in order to improve the performance in virtual screening process.[14,15]

In this study, multiple molecular dynamic simulations were used to identify the most significant ligand-protein interactions from multiple crystal structures of the same protein in complex with different ligands. The use of a single ligand could induce a bias in the side chains conformation, then the choice of different ligands aims at widely explore the active conformation of the binding site. From PDB database four complexes of EED with different known inhibitors (PDB ID: 5H13-5H14-5H15-5GSA) were chosen for the set-up of our workflow.

The analysis of ligands interaction pattern during the simulation lead to the creation of a "dynamic pharmacophore". A pharmacophore model is defined as an ensemble of steric and electronic features necessary for optimal binding in the protein binding pocket. In structure-based pharmacophore modeling the interactions derived from ligand-protein complex, obtained from crystallographic structure analysis or NMR experiment are used to create the pharmacophore features. [16] The static nature of the technique by which the crystal structures are resolved can result in protein-ligand systems in which several interactions are shown, but only a few of them play a key role in the binding process, and are potentially accountable for biological activity.[17,18] Hence, the aim of this approach is

to identify the crucial interactions that are predicted to be shared among the different complexes analysed, and use them to improve the virtual screening performance. [19,20]

The workflow followed in this approach are summarised in scheme **1**. In details, in order to discern the critical interactions from the occasional ones, three different MD simulations for each protein-ligand complex of EED with different ligands, were carried out. Then, the information derived from the simulations were analysed, and the common interactions of the three simulations were used to create a "dynamic" pharmacophore for each system.

A pharmacophore screening was carried out both on static pharmacophore, built with the classic method and on dynamic pharmacophores using a dataset of active and inactive molecules retrieved from ChEMBL database[21] in order to analyse and compare their performance in a retrospective way. The three dynamic pharmacophores created were aligned and their common observed features were used to create a new shared pharmacophore. Moreover, the shared dynamic pharmacophore features were then used to add constraints to the docking grid, and docking studies were performed both with and without constraints in order to evaluate a possible improvement in the docking screening performance.
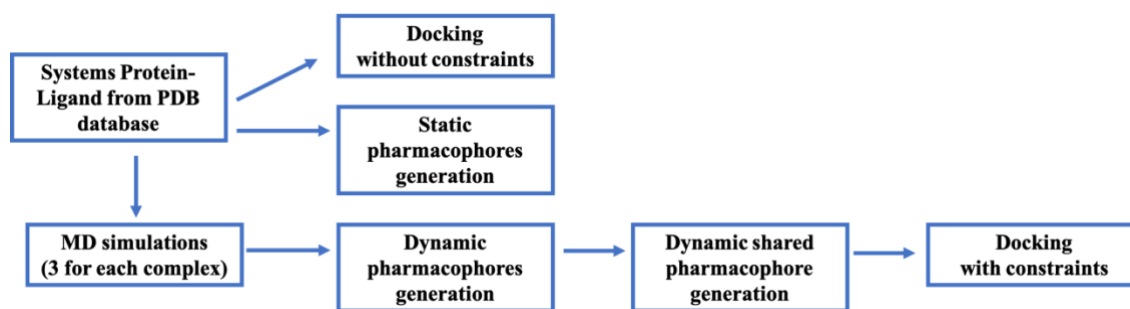

*Scheme 1. Workflow followed in our dynamic-shared pharmacophore approach*

7.1 Molecular Dynamics

The crystallized structures of the EED protein in complex with different ligands (PDB ID: 5H13-5H14-5H15-5GSA) chosen to set-up our workflow were acquired from PDB database.[22] Three different MD simulation of 25 ns were run for each complex, applying a random seed starting point. Instead of a unique long simulation was decided to run multiple short simulations, in order to better explore different conformations of the protein[23,24,25,26] For all the simulation, the analysis of the root-mean-square deviation (RMSD) plot graphs revealed only small medium fluctuations of around 1Å, both for protein and ligand (Figure **2**). The simulations were analysed and all the protein-ligand common interactions observed for the three simulations were collected. The comparison between the interactions retrieved from MD and those from static PDB structures showed contacts that were not visible in the crystal complexes, and some of them were observed most of the time throughout the simulation duration. New ligand-protein interactions were identified but these were not shared among the three simulations and, thus, for our approach, were not considered in the dynamic pharmacophore.
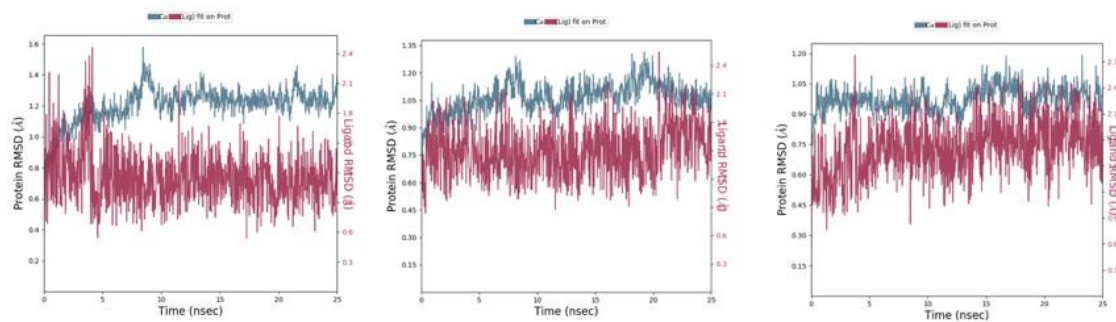
5H13
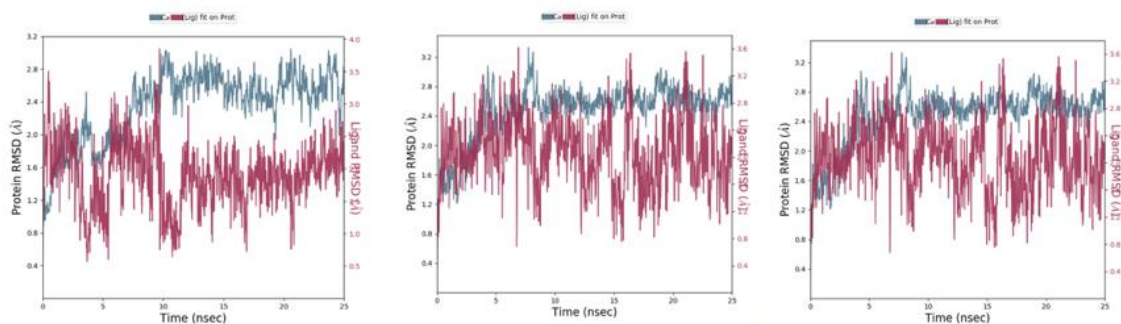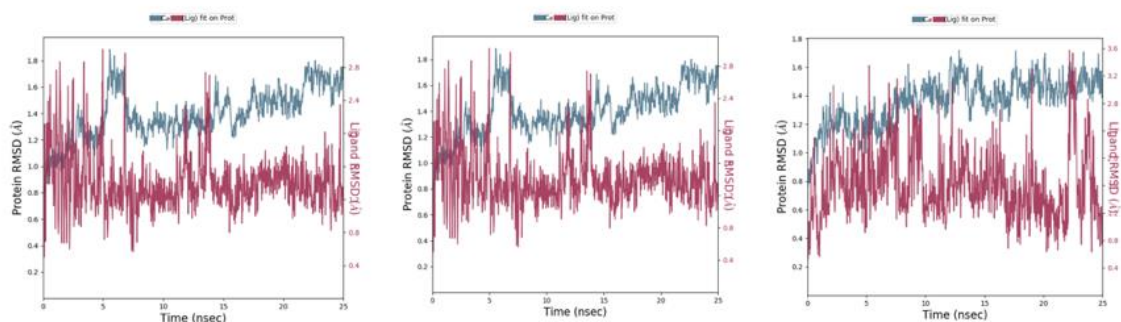

5H14
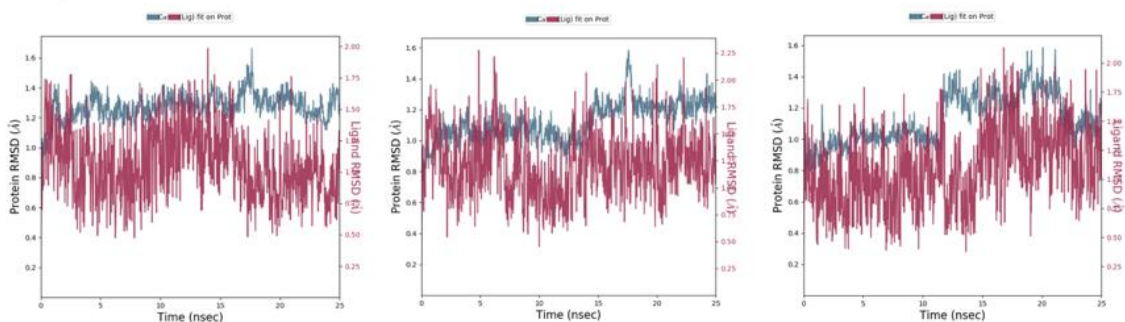
5H15

5GSA

Figure 2. RMSD Plot of three simulations for each protein-ligand complex.

The Root Mean Square Deviation (RMSD) is used to measure the average distance between atoms in a specific frame with respect to a reference frame

7.2 Protein-ligand interactions analysis and structure-based pharmacophores generation

LigandScout software[27] was used for generation of pharmacophore models. Starting from the PBD coordinates set, static pharmacophore models were created and, according to the information retrieved from MD simulations, dynamic pharmacophores were then generated. The 2D diagram of the common protein-ligand interactions resulted from MD simulations for each system analysed is shown in table **1**.

The first system evaluated was PDB 5H13. (Figure **3**) The interactions retrieved from PDB lead to a creation of pharmacophore with only two hydrophobic features. The MD simulations revealed several new interactions, such as π-π stacking interactions between the methoxy-benzene, the pyridine and benzo-dioxole component of the tetracyclic moiety, and Phe97, Tyr148, Tyr365 residues, respectively (Table **1**). These residues are involved also in π-cation interactions with the pyridine nitrogen atom, positively charged at physiological pH. Furthermore, the amine group in 2 position of pyridine ring formed H-bond donor contacts with Asn194 and Glu238 residues.
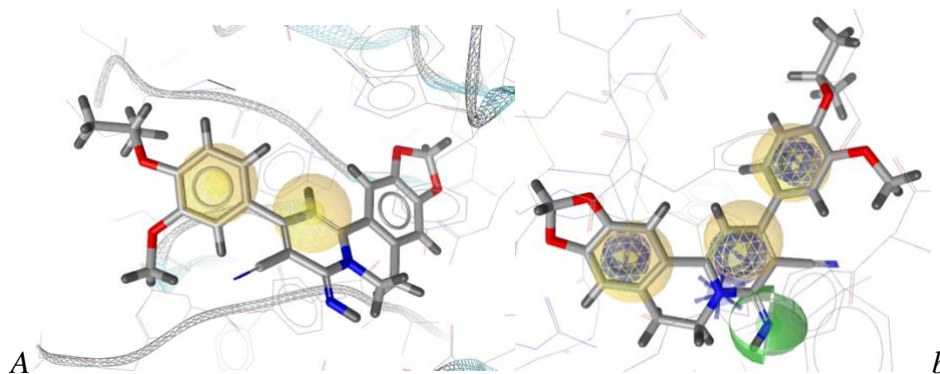

Figure3. 5H13 a. Static pharmacophore; b. Dynamic pharmacophore;
Green spheres: hydrogen-bond donor; yellow spheres: hydrophobic features; blue rings: aromatic features; blue spikes: positive ionisable features.

As it can be noted in the comparison between the static and dynamic pharmacophores (Figure **4**), multiple π-π stacking interactions were found during analysis of the protein-ligand interaction diagram of 5H14 dynamic simulations. The 5H14 ligand established

several π-π stacking interactions in the EED aromatic cage. The benzene moiety interacts with Phe97, Tyr148, and Tyr365 residues and the aromatic ring of the tetrahydroisoquinoline moiety binds Arg367 and Tyr365 residues. While the nitro group formed π-cation interaction with Tyr365 amino acid. Considerably, the interaction between the pyrazole moiety and Phe97 residue was not reported in the PDB. During a single experiment the pyrazole group established a H-bond contact with Tyr148, but this interaction was not observed in the other two simulations, then it was not indicated in the dynamic model. In addition, in the static pharmacophore, automatically generated by the software, a H-bond acceptor feature was found, but it was not reported neither in the PDB nor during MD simulation.
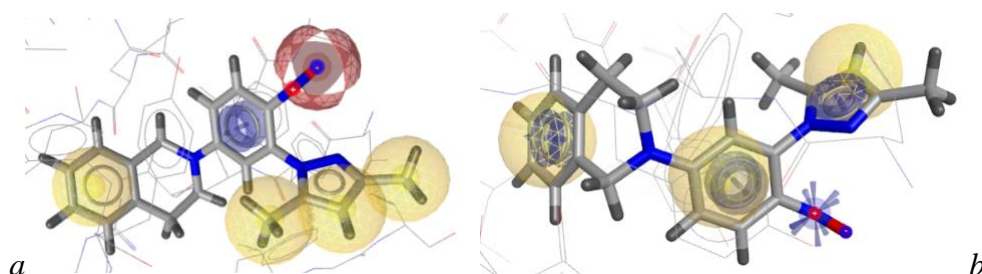

a   *b*

Figure 4. 5H14 a. Static pharmacophore. B. Dynamic pharmacophore.
Red spheres: hydrogen-bond acceptor, yellow spheres: hydrophobic features, blue rings: aromatic features; blue spikes: positive ionisable features.

In the dynamic simulations the anisole moiety of 5H15 ligand established π-π stacking interaction with Tyr365 amino acid and π-cation interaction with the protonated Arg367, and its indole moiety interacted through double π-π binding with Phe97 residue. Furthermore, the dimethyl-ammonium group binds Tyr365 and Tyr148 residues via π-cation interactions. The same group in a single simulation formed H-bond interaction with Asn194. Computational methods were employed to determine the predominant form at physiological pH of the two quaternary nitrogen atoms positively charged in the ligand-protein complex. It was found out that at pH=7.4, there are two forms in equilibrium, depending on whether the dimethyl-amino nitrogen or the pyrrolidine nitrogen was protonated (Figure **5c**). Then, two dynamic pharmacophores based on different state of ligand protonation, were created and screened but the one carrying the charge both on the

amine group and pyrrolidine moiety exhibited better results, in terms of ROC curve during the validation step, and so it was chosen to progress in our study (Figure **5**).
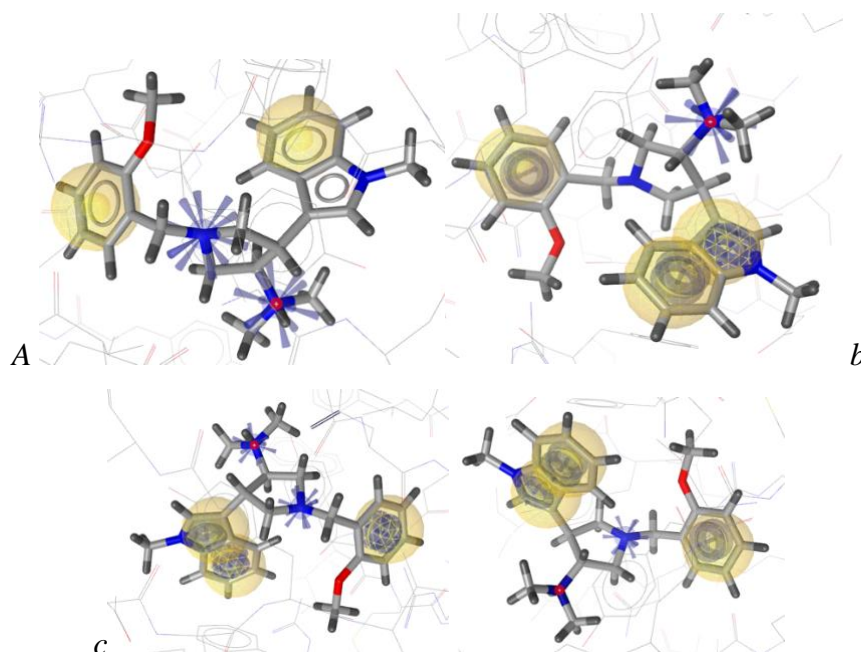

Figure 5. 5H15 a. Static pharmacophore. B. Dynamic pharmacophore. C. The dynamic pharmacophores with different protonation generated.

Yellow spheres: hydrophobic features, blue rings: aromatic features and smallest blue sphere: positive ionisable features.

During MD simulations of 5GSA system π-π stacking interactions were found between the residues Tyr148-Tyr365 and the furan and the triazole component of the bicycle. Furthermore, the benzene ring interacted *via* π-π stacking interaction with Phe97 and also in this case π-cation interaction with Arg414 was retrieved in a single simulation and then it was not reported in the dynamic model (Figure **6**). In addition, a hydrogen bond contacts were observed between the nitrogen atom of triazole moiety and Lys211, and between the aliphatic ammine group and Asn194 residue.
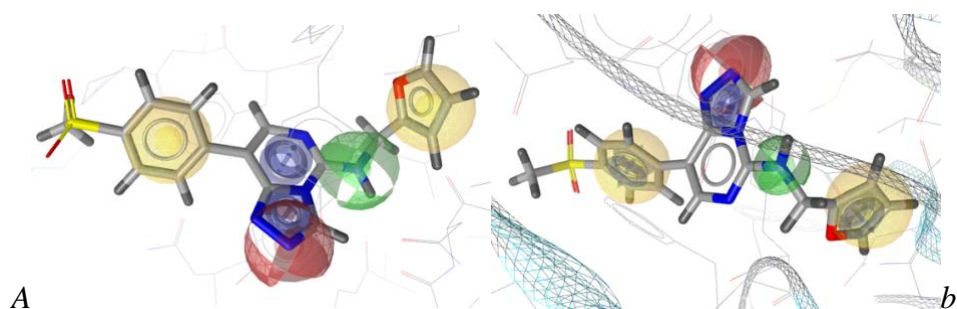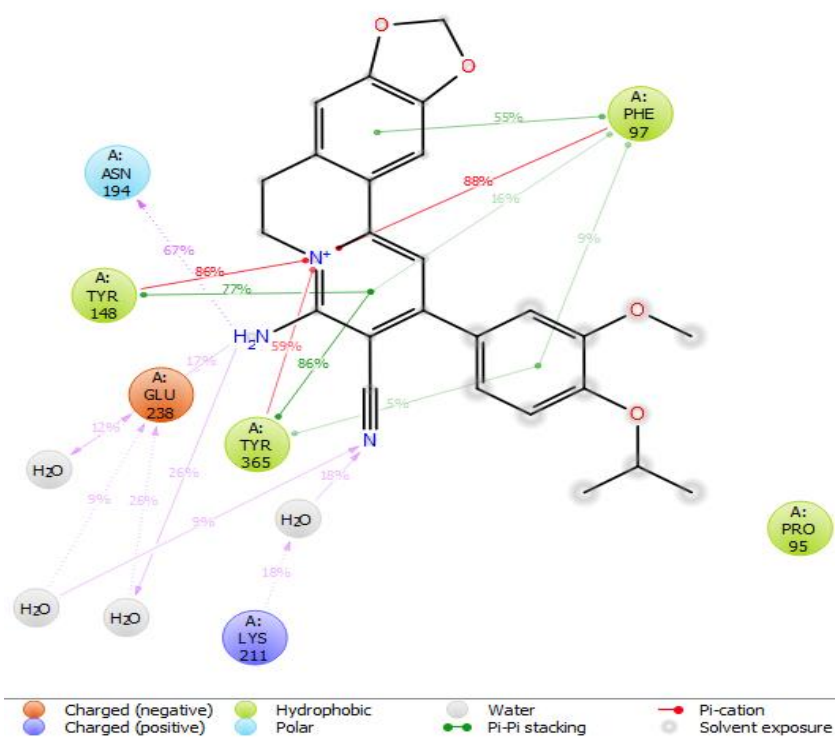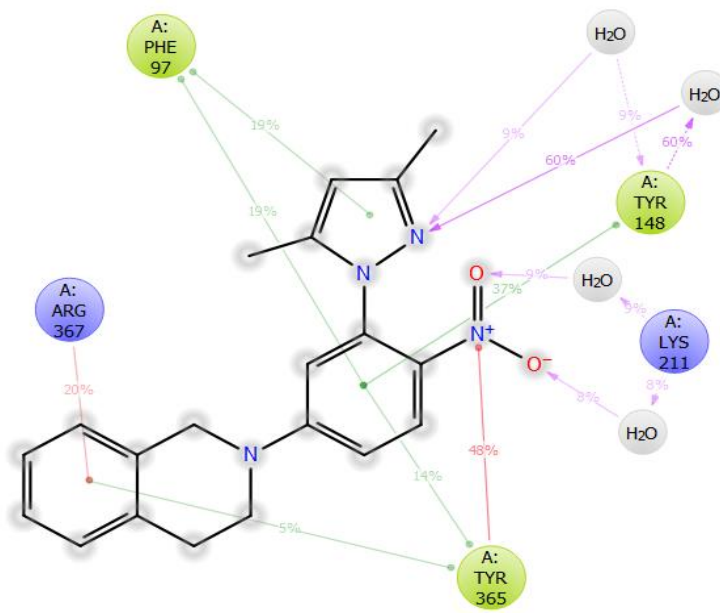
Figure 6. 5GSA a. Static pharmacophore. B. Dynamic pharmacophore.

Red spheres: hydrogen-bond acceptor, green spheres: hydrogen-bond donor yellow spheres: hydrophobic features and blue rings: aromatic features.

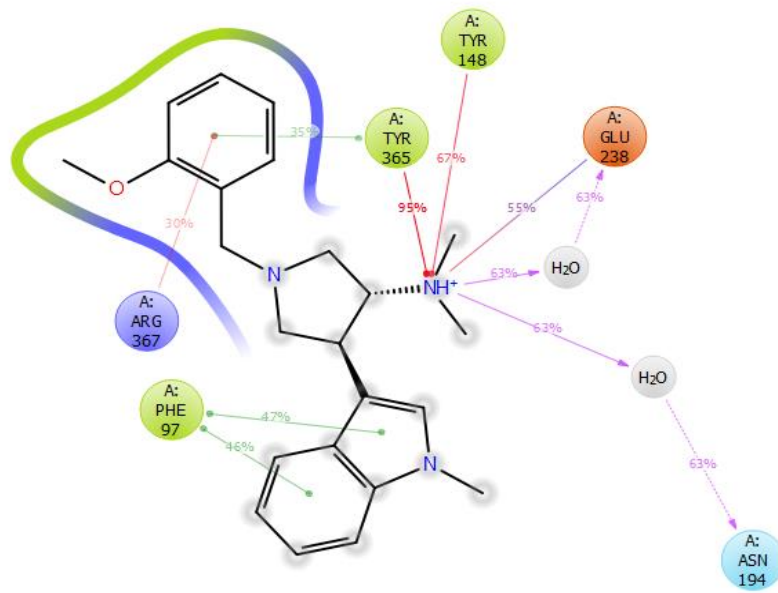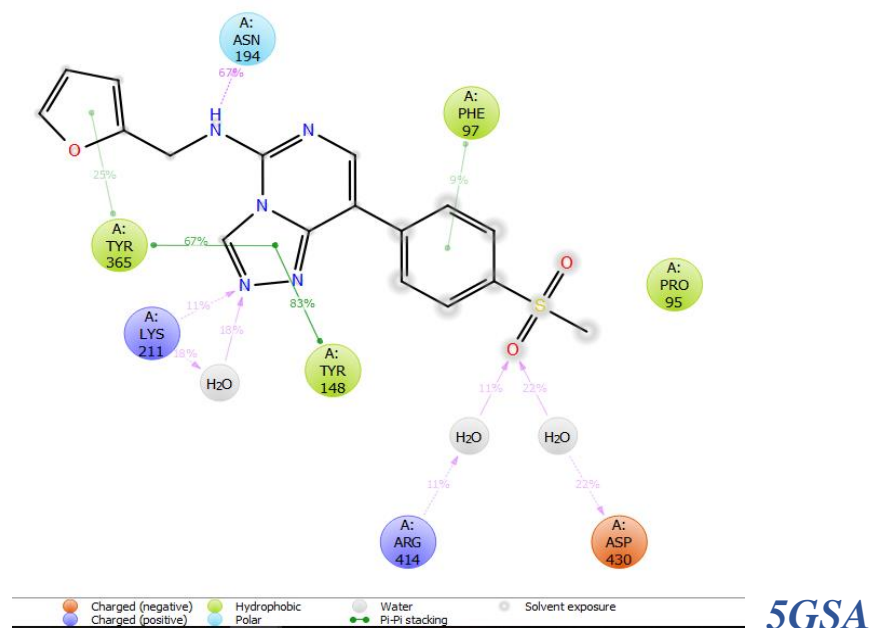*Table 1. Ligand-Protein interaction diagram from MDs*

5H14

5H15

5GSA

The study of systems analysed revealed that the pattern interaction of these ligands shared common features. Specifically, MD simulations revealed π-π interactions hidden in the interaction pattern retrieved from static crystal structures, and involving aromatic residues Ty148, Tyr365, Phe97 and Arg367, thus suggesting that these residues could be crucial for the correct alignment of the ligand inside the binding pocket.

7.3 Ligand-Based Pharmacophore generation

Considering the availability of active molecules targeting EED, a ligand- based pharmacophore was also developed. Ligand-based approaches exploit the structural information of active compounds that bind the biological target of interest. This approach was employed with the aim of evaluating if the use of the EED active dataset can provide useful information with the goal of improve the virtual screening process. Exploiting information from EED active compounds, the common structural features were used to create the ligand-based pharmacophore model.[28,29] However, the pharmacophore created using the active compounds available with the parameters reported in material section, was a unique hydrophobic feature (Figure **7**). Hence, the ligand-based approach could not be further used for the virtual screening, which ultimate makes a structure-based study indispensable.

Figure 7. Pharmacophore ligand-based. *Yellow sphere: hydrophobic feature*

7.4 Pharmacophore Screening

Moreover, the performances of dynamic pharmacophore models and static pharmacophores generated from the PDB coordinates set were compared in order to evaluate if the interactions retrieved from molecular dynamics can improve the virtual screening process. Then, a pharmacophore screening was carried out, using a dataset of active and inactive compounds retrieved from ChEMBL database[21], and the results were compared generating ROC curves and calculating the area under the curve (AUC) and the enrichment factor (EF) parameters. The pharmacophore screening was carried out with LigandScout software which requires at least three features in each model. The static pharmacophore retrieved from PDB 5H13 reported only two features, then was not suitable to perform standard pharmacophore screening. Conversely, the dynamic

pharmacophore obtained from 5H13 included six features, highlighting the usefulness of dynamic approach in such a case. However, considering that the screening of the static pharmacophore could not be carried out, it was not possible to make a comparison between the two models. In table **2** are reported the results of both pharmacophore dynamic and static screening considering 0, 1 and 2 omitted features.

The general trend of dynamic pharmacophores results showed similar values of AUC to the static ones and in some cases, an increase of EF values, indicating an improvement of discriminative capability to prioritize true active compounds, as showed in the ROC curves comparison (Table **2**).

The dynamic pharmacophore 5H14 enhanced the screening performance in the early recognition phase. With 0 omitted feature neither the static nor the dynamic model found any hit. Whereas during the screening with 1 and 2 omitted features both the AUC and EF values were significantly better in the dynamic pharmacophore (Figure **8**)

The 5H15 static pharmacophore displayed good results already in 0 omitted features, contrary to the dynamic pharmacophore that did not retrieve any hit. While in the screening with 1 and 2 omitted features, the AUC values were similar, and EF values were improved in the dynamic up to the 5% percent of the screening performance. (Figure **9**) The two models were different mainly in the number of retrieved hits. Indeed, the static retrieved more hits than the dynamic one. Probably the reason could lies in the different numbers of features that compose the two pharmacophores or in the kind of chemical structures of the hits retrieved. Indeed as it can be noted, the dynamic pharmacophore 5H15 was composed by 7 features against 4 in the static. Furthermore 3 features of the 5H15 dynamic model were π-π stacking, and were more selective compared to the two-hydrophobic observed in the corresponding static model. Moreover, analysing the chemical structures of the hits retrieved in the static model, it was found out a strong chemical similarity. This means that this pharmacophore model could be restrictive to a specific chemo-type. The screening on 5GSA models resulted in similar values of AUC both on static and dynamic pharmacophore but in better EF values on the dynamic pharmacophore. In details, the dynamic pharmacophore of 5GSA provided higher EF values during the screening with 0 and 1 omitted features, whereas with 2 omitted features these parameters were comparable (Figure **10**).

Table 2. Static and dynamic pharmacophores AUC and EF values at 1-5-10-100% for 5H14, 5H15, and 5GSA systems.

5H14						
	STATIC			DYNAMIC		
Omitted feature	0	1	2	0	1	2
AUC (1-5-10-100)	N.D.	0.36 0.65 0.58 0.51	0.36 0.43 0.59 0.66	N.D.	0.5 0.5 0.5 0.5	0.89 0.95 0.96 0.57
EF (1-5-10-100)	N.D.	3.4 3.4 3.4 3.4	2.7 0.5 3.4 2.5	N.D.	23.9 23.9 23.9 23.9	10.6 10.3 10.3 10.3
Hits retrieved	0	7	205	0	1	14

5H15						
	STATIC			DYNAMIC		
Omitted feature	0	1	2	0	1	2
AUC (1-5-10-100)	1.0	1.0	1.0		1.0	1.0
	1.0	1.0	1.0	N.D.	1.0	1.0
	1.0	1.0	1.0		1.0	1.0
	0.79	0.71	0.71		0.61	0.76
EF (1-5-10-100)	23.9	21.3	21.3		18.6	18.6
	13.4	11.7	11.7	N.D.	16.6	10.7
	13.4	5.9	5.9		16.6	6.5
	13.4	1.1	1.1		16.6	6.5
Hits retrieved	43	563	563	0	13	85

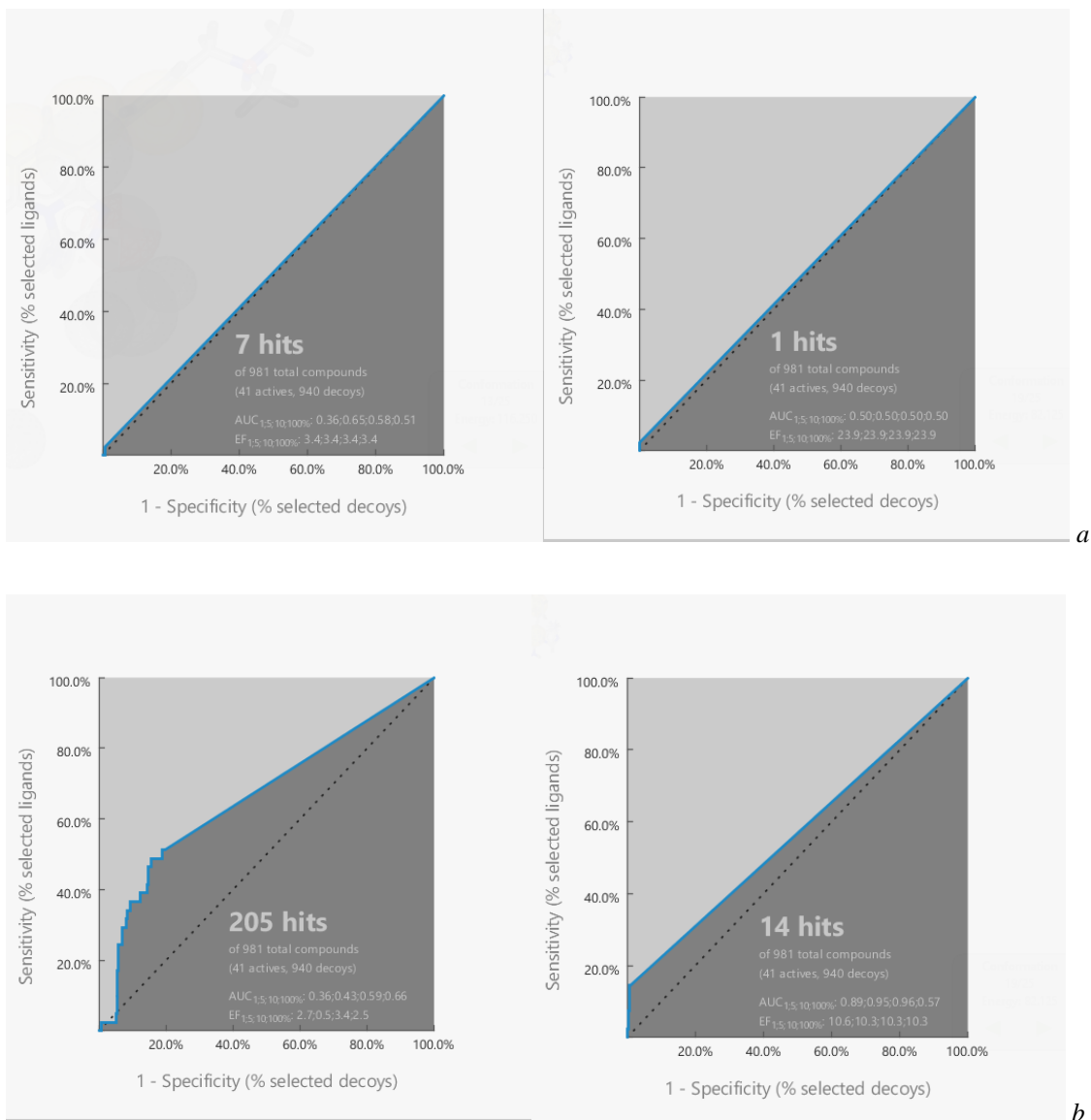| 5GSA						
	STATIC			DYNAMIC		
Omitted feature	0	1	2	0	1	2
AUC (1-5-10-100)	1.0	1.0	1.0	1.0	1.0	1.0
	1.0	1.0	1.0	1.0	1.0	1.0
	1.0	1.0	1.0	0.99	1.0	1.0
	0.60	0.58	0.72	0.55	0.58	0.59
EF (1-5-10-100)	21.3	21.3	21.3	23.9	20.9	21.3
	19.1	3.9	3.9	23.9	20.9	4.4
	19.1	3.9	2.4	23.9	20.9	4.4
	19.1	3.9	2.5	23.9	20.9	4.0
Hits retrieved	10	49	264	4	8	54

Figure. 8. ROC curves 5H14 static versus dynamic; results with 1(a) and 2 (b) omitted features.
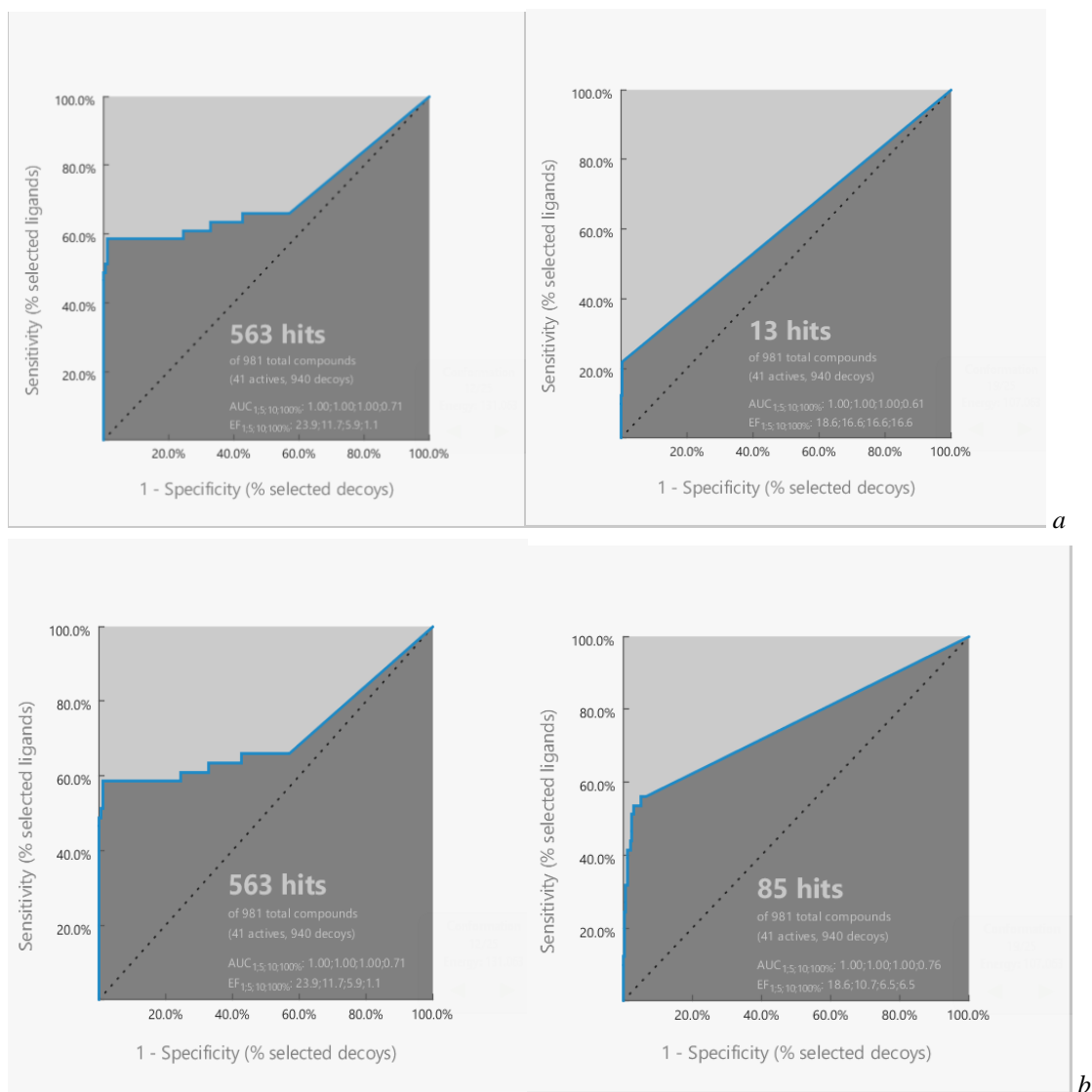
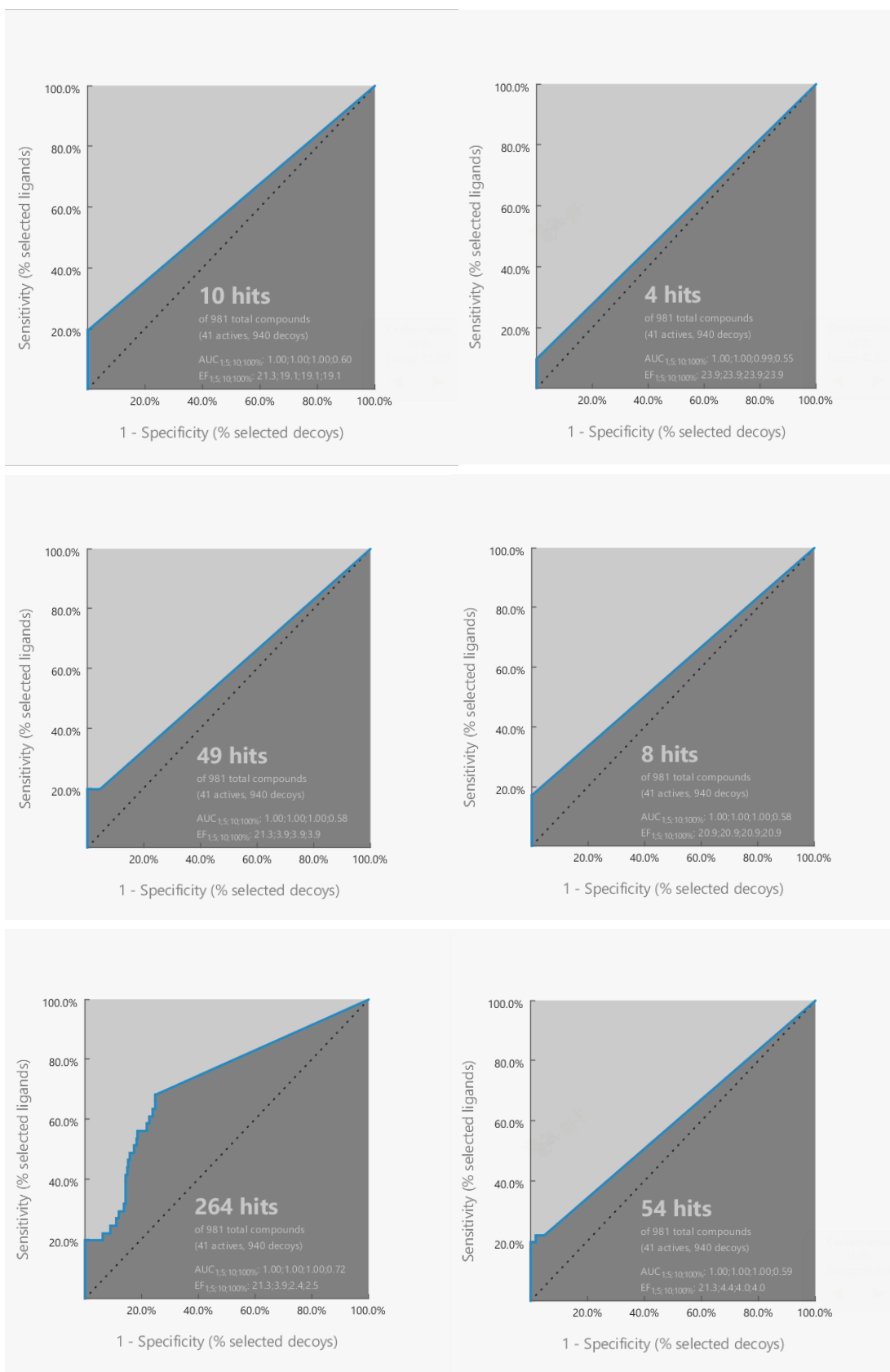***Figure. 9. ROC curves 5H15 static versus dynamic; results with 1(a) and 2 (b) omitted features***

Figure. 10. ROC curves 5GSA static versus dynamic; results with 0 (a), 1 (b) and 2 (b) omitted features

Thereafter, the three dynamic pharmacophore models were used to create a new, unique pharmacophore model, called "dynamic shared pharmacophore".

For this purpose, the three dynamic pharmacophores were aligned and the resulting overlapped features were interpolated. The dynamic shared pharmacophore resulted in an ensemble of 4 features, 2 π-π and 2 hydrophobic (Figure **11**)

Furthermore, in order to assess the capability of the new model to prioritize actives from inactive compounds, it was decided to screen the dataset on this new hybrid dynamic pharmacophore, which showed good values of AUC and EF, especially in the early recognition phase (Table **3**).

Figure 11. Dynamic shared pharmacophore. *Yellow spheres: hydrophobic features; blue rings: aromatic features.*

Table 3. ROC curve Dynamic shared pharmacophore results with 0, 1 and 2 omitted features. *AUC and EF values at 1-5-10-100%*

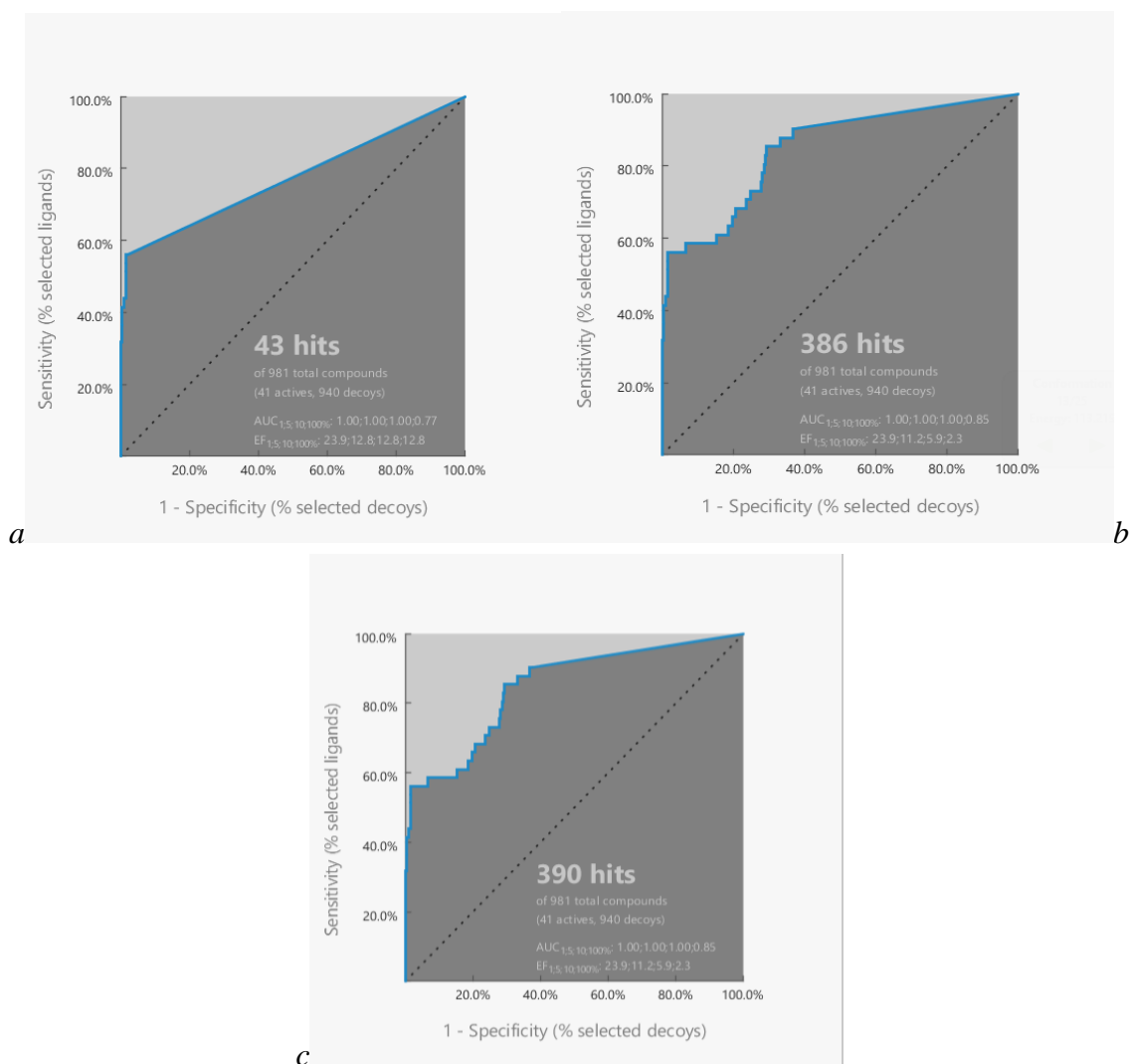| Dynamic-shared			
N.o omitted feature	0	1	2
AUC (1-5-10-100)	1.0 1.0 1.0 0.77	1.0 1.0 1.0 0.85	1.0 1.0 1.0 0.85
EF (1-5-10-100)	23.9 12.8 12.8 12.8	23.9 11.2 5.9 2.3	23.9 11.2 5.9 2.3
N.o hits retrieved	43	386	390

Figure. 12. ROC curves dynamic shared pharmacophore; results with 0(a), 1 (b) and 2 (c) omitted features

The shared dynamic pharmacophore exhibited a good performance, in terms of AUC and EF values (Figure **12**). With the aim of determining if the dynamic model had a bias and retrieved only one chemotype the hits retrieved during the screening were analysed. In the set of 43 hits retrieved different chemotypes were identified, showing an overall improvement of the performance compared to the previous 5H15 static model.

The results of the dynamic shared pharmacophore highlighted the pivotal role of the specifically ligand position inside the aromatic cage of the binding pocket. Then, the information retrieved from pharmacophore features could guide the creation of docking model with constraints to selectively discriminate actives from inactive compounds. Due to the low number of features in the dynamic shared model, we have considered it as a

starting point of exploration to be further improved using biological data of selected molecules.

7.5 Docking algorithm evaluation and Docking screening

The crystal structures of EED in complex with the ligands analysed in this work (PDB ID: 5H13-5H14-5H15-5GSA) were used to create a docking model for each system. The models validation was carried out firstly through cognate docking to mainly test the searching algorithm, and then by a retrospective virtual screening using dataset of active and inactive molecules, in order to stress the scoring function capability.

The virtual screening performance was evaluated through the ROC curves and the robust initial enrichment (RIE) parameters.

The dynamic shared pharmacophore, based on common interactions retrieved from MD simulations, showed two important π-π stacking features, which are used as constraints in the docking grid. In order to evaluate the contribute of these constraints on the screening capability molecular docking was carried out in standard precision mode, either with and without constraints. The screening performance was evaluated through ROC curves and Robust Initial Enrichment (RIE) values. The choice to use RIE values to validate the screening performance rather than EF, lies in the fact that the enrichment factor is influenced by the length of ranked list, while RIE is focused only on the early enrichment phase that is the most important in the context of the virtual screening in which thousands molecules were analysed.

The results reported in Table **4**, indicate that the addition of positional constraints significantly improves the screening performance. Specifically, the use of constraints influenced the RIE scores, exhibiting a performance enhancement in the early phase of the screening, with comparison to the models without constraints (Figure **13**)

Complex PDB ID	ROC	RIE
5H14		
Without Constraints	0.75	2.42
With Constraints	0.86	7.66
5H15		
Without Constraints	0.75	4.26
With Constraints	0.91	9.54
5GSA		
Without Constraints	0.60	2.95
With Constraints	0.72	7.88

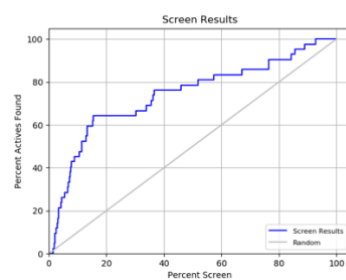*Table 4. ROC and RIE values calculated from docking screening, performed with and without constrains*
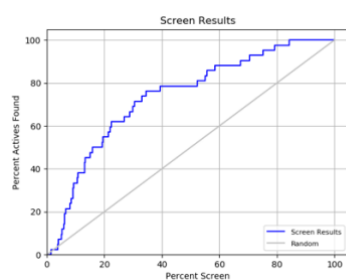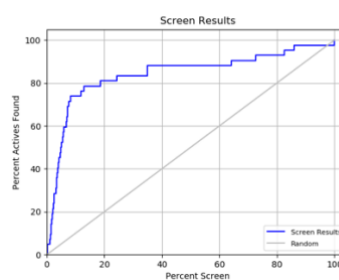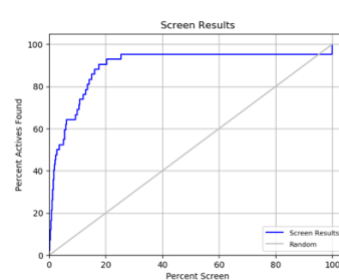

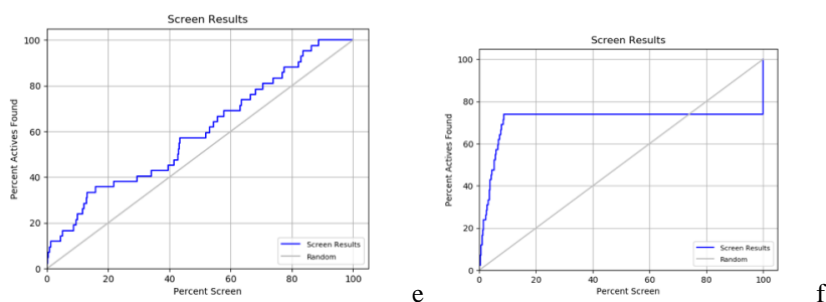*PDB ID 5H14 ROC graphs: a. without constraints, b. with constraints*

PDB ID 5H15 ROC graphs: c. without constraints, d. with constraints

PDB ID 5GSA ROC graphs: e. without constraints, f. with constraints

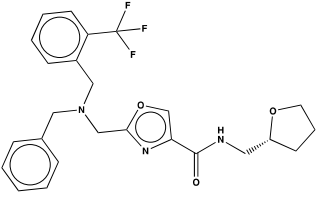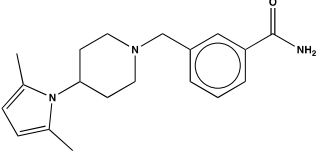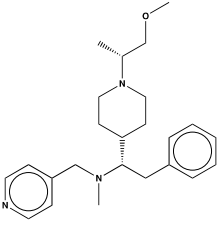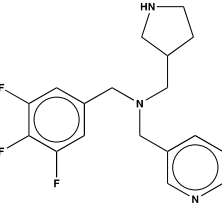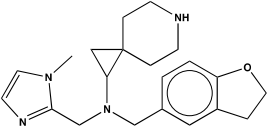**Figure 13. ROC graphs. Docking results with and without constraints.**
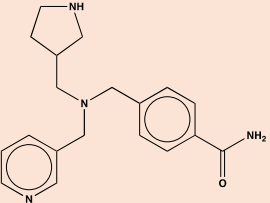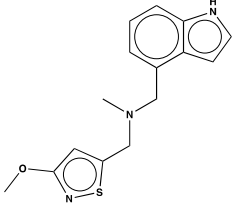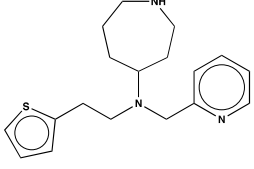
According to the results from this study, the common interaction pattern of known EED inhibitors involved π-π stacking contacts between aromatic moieties of ligands and the residues of the aromatic cage within the binding site, such as Tyr365, Tyr148, Phe97 and Arg367. The improvement of virtual screening results in our approaches could suggest that these aromatic interactions are pivotal for the discovery of new active compounds.

The Molport virtual library, composed by 1365148 compounds, was screened on the docking model built as result of the workflow followed in this approach, and 56 molecules were selected to be tested.

7.6 Identification of EZH2/EED binding inhibitors

An in vitro binding assay was used to determine the effects of fifty-six compounds on the binding of EZH2 to EED. The assay was performed by BPS Bioscience.

In this assay the capability to inhibit the protein-protein EED-EZH2 binding was evaluated. The compounds were tested in single dose of 25 μM and seven of them showed an inhibition rate higher than 30%. (Table **5**)

Table 5. Inhibition rate of compounds on EED-EZH2 binding

Compound	Structure	Inhibition (%)
EE07 (25 µM)		30
EE11 (25 µM)		36
EE28 (25 µM)		41
EE31 (25 µM)		43
EE38 (25 µM)		40
EE41 (25 µM)		54
EE54 (25 µM)		41
EE55 (25 µM)		40

As it can be noted from the structure of the active compounds, most of them shared a similar chemical structure, with a central tertiary amine and different aromatic moiety in the side chains (Figure **14**). The most active compounds were the analogues EE41 and EE3, which differed for the substituents group in the benzene ring. The putative binding mode indicated that the interaction with the binding pocket is mainly represented by π-π stacking and π-cation contacts. The central tertiary amine is protonated at physiological pH and interacts through π-cation contacts with Tyr365, Ty148 and Phe97, while the other aromatic rings establish π-π stacking interactions with Tyr 365 and Arg367 in EE31 compound, while π-cation contact with Arg414 in EE41. Furthermore, in both compounds binding poses, the nitrogen atom of pyrrole moiety is protonated and forms hydrogen bond with Glu238. (Figure **15**).

Further biological test will be carried out (e.g., IC$_{50}$ and evaluation on the methyl-histone level) on these compounds in order to validate the biological activity. In conclusion, compound EE41 seems to be a promising starting point to design a more potent EED-EZH2 binding inhibitor.
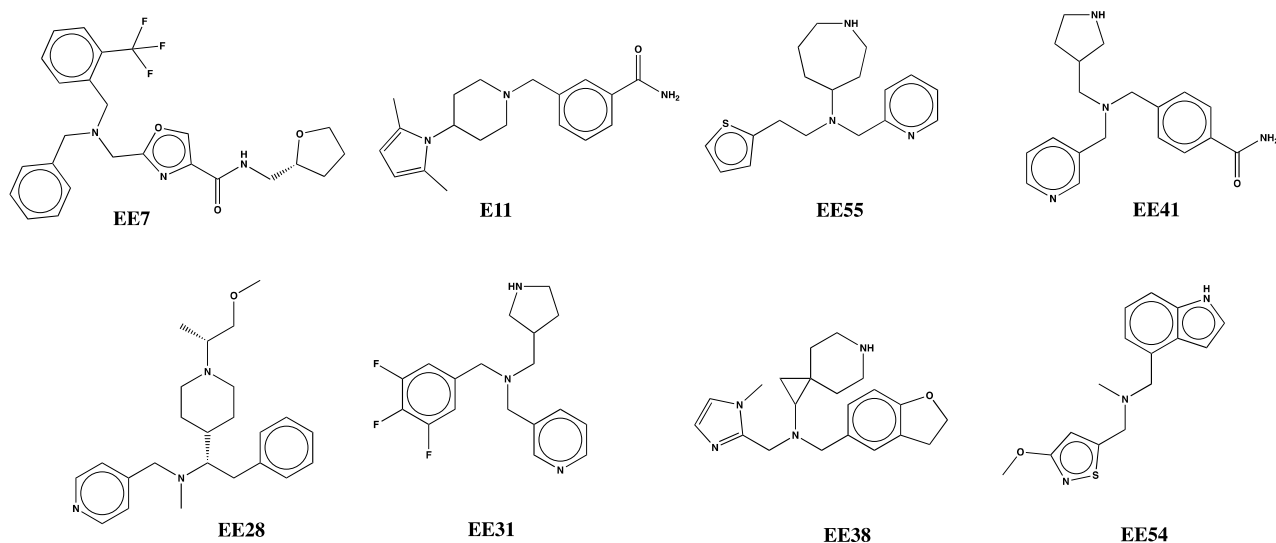

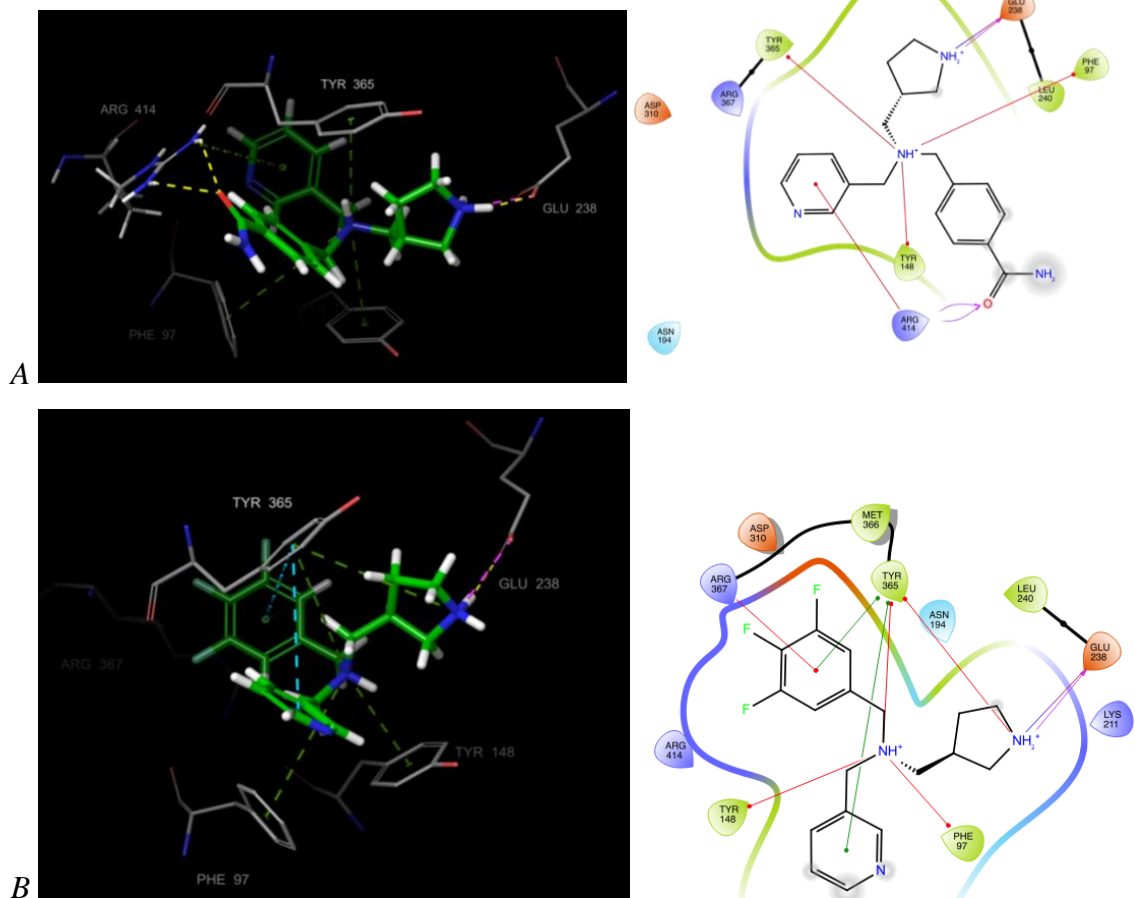*Figure 14. Most active compounds in inhibition of EED-EZH2 binding*

Figure 15. Binding poses and ligand interaction diagram of compound EE41 (A) and EE31 (B)

7.7 Material and Methods

System quality assessment and protein preparation

The crystal structures were selected from PDB database with resolution below 2.5 Å. Protein Preparation Wizard[30], a Schrödinger tool, was used to prepare the protein, adding bond orders and hydrogens to the crystal structure and deleting waters beyond 5.00 Å. Epik was used to generate het states at pH 7.4 ± 0.2 and subsequently the protonation state of the protein and the ligand were evaluated using PropKa at pH 7.4. [31,32]

Molecular Dynamics and common interactions evaluation

With the aim of better investigate eventual ligand-protein conformational changes it was chosen to perform multiple short simulations, using for each simulation random seeds as starting point. For each protein-ligand complex, three molecular dynamics simulations of 25 ns were run. The systems stability was evaluated checking the RMSD values of the aligned protein and ligand coordinate set calculated against the initial frame.

DESMOND[33] (version) was used to perform the simulations using OPLS3e force field.[34,35] The systems were built solvating the complex protein-ligand in orthorhombic boxes using the TIP3P water model. Then the box volumes were minimized and equilibrated at the temperature of 300 K and at 1.01325 bar pressure, and NPT ensemble was used. The trajectory coordinates were recorded every 25 ps, obtaining 1000 frames for each simulation. In order to analyse MD simulations, we used simulation interactions diagram that showed the graph of RMSD and RMSF values both for ligands and proteins, the protein secondary structure changes, ligand-protein interactions, including H-bonds, hydrophobic, ionic, and water-bridge contacts and the ligand torsion and properties. Starting from these data, the common interactions were evaluated. Only interactions present in the three simulation for each system were considered common.

Pharmacophore generation

LigandScout software[27] was used for the pharmacophore generation.

- **Structure-based (static and dynamic)**

For each ligand a static pharmacophore model was generated using the PDB coordinate of ligand-protein complex. Then the common interactions retrieved from the MD

simulations were used to create a new dynamic pharmacophore model for all three ligands. The dynamic pharmacophore models were generated importing the corresponding structure-based model and adding the common features identified from the simulations. Due to their proximity, in the dynamic 5H14 pharmacophore model the three hydrophobic features on the pyrazole moiety were interpolated and the tolerance was increased by 0.15 Å, in order to create a unique larger hydrophobic region. The created dynamic pharmacophores were then aligned and the overlapping features were interpolated.

- **Shared dynamic pharmacophore**

The "shared" dynamic pharmacophore was created only considering the common features retrieved from the 12 simulations (3 for each PDB chosen). In order to take into account the dynamicity of the interactions, the tolerance radius for overlapped feature was increased by 0.15 Å.

- **Ligand-based**

The active compounds (42) present in our dataset were separated randomly in training compounds (30) and test compounds (12). Conformers of these molecules were generated, using I-Con and applying the BEST setting. A max number of 200 conformers were generated with RMS threshold of 0.8, and an energy window for conformer selection of 20,0 Kcal/mol. Torsion driving, generate cords from CT, enumerate rings and nitrogen atoms were flagged on. Then, the ligand-set were clustered with pharmacophore-alignment score, as similarity measure, and setting a cluster average distance of 0.4. The ligand-based pharmacophore was created as "shared feature pharmacophore" using "pharmacophore-fit and atom overlap" as scoring function, creating excluded volumes, and setting a feature tolerance scale factor of 1 Å.

Docking grid generation

For every system, a docking grid was generated both with and without constraints, using Glide software.[36] The scaling factor was set at 1.0 Å with a partial charge cut-off of 0.25, and the ligand was chosen to define the grid centroid. Starting from the shared dynamic pharmacophore interaction pattern, constraints were set on the docking grid. For the

common π-π stacking and hydrophobic interactions positional constraints were imposed in line, with the GLIDE grid constraints panel workflow.

Database creation and preparation

The screening performance of both pharmacophore models and the docking grids was validated in a retrospective way, using a list of known active and enriched inactive compounds. The compounds were retrieved from ChEMBL database.[21]All compounds tested on the target EED and with different values of $IC_{50,}$ K_D, K_i were chosen and filtered to remove duplicates. All compounds tested on the target EED and with different values of $IC50$, K_D, Ki were chosen and filtered to remove duplicates. Compounds showing Ki ≥ 50000 nm and $IC50$ ≥100000 nm values were considered inactive. In order to enrich the inactive dataset, analogues of these inactive compounds were searched on PubChem database[37], with PubChem fingerprint and 0.9 as similarity coefficient of Tanimoto. Such a method does not provide analogues compounds because it is not based on substructure search. Duplicates were then filtered-off and, to choose representative molecules with different chemotypes, inactive compounds were furthermore clustered, using radial binary fingerprint and K-means clustering a tool provided by Canvas.[38,39] In order to avoid too strictly similar compounds in the inactive dataset a similarity matrix for all inactive compounds and for the single cluster was created. The average value for each class was calculated, showing a sufficient structural difference in the dataset. Moreover a hierarchical cluster analysis was assessed to further check chemical clusters.

Inactive compounds	Similarity matrix average
All inactive compounds	0.067
Cluster 1	0.081
Cluster 2	0.199
Cluster 3	0.122
Cluster 4	0.071

Table 6 Similarity matrix average calculated for the inactive compounds of the dataset

The final dataset was composed by 42 active and 940 inactive molecules. These compounds were then prepared by LigPrep, a tool of MAESTRO software, using OPLS3e as force field and EpiK to generate all the possible states at pH of 7.0 ± 0.2. Desalt and

generated tautomers were flagged on, the chirality was retained and at most 32 conformers per ligand were generated.[40]

Pharmacophore screening

The pharmacophore virtual screening was performed using LigandScout software.[27] The "pharmacophore fit-score" was used as scoring function and "match all query features" was chosen as screening mode. Furthermore, "get best matching conformation" was selected as retrieval mode. The screening performance was evaluated through Receiver Operating Characteristics (ROC) graphs, calculating Area Under the Curve (AUC) and enrichment factor (EF) parameters.

Molecular Docking screening

Molecular docking was carried out using Glide software released by Schrödinger[36] (release 2018-4). The simulation was performed in SP, with and without constraints, using OPLS3e as force field.[34] The ligands were considered as flexible, and Epik state penalties were included to docking score.[31] The van der Waals (vdW) radii scaling factor was set as 0.8, with a partial charge cut off by 0.15.

Assay Experimental Conditions

The enzymatic assay were performed by BPS Bioscience, San Diego, California.

- **Materials**

EZH2-EED Binding Assay Kit (BPS Bioscience, 52066)

EZH2 Inhibitor III (Sigma, 5083200001)

AlphaLISA Acceptor Beads Beads, 5 mg/ml (PerkinElmer, AL112C)

AlphaScreen Glutathione Donor Beads, 5 mg/ml (PerkinElmer, 6765300)

OptiPlate-384, White Opaque 384-well Microplate (PerkinElmer, 6007290)

- **Enzymes and Proteins**

Enzyme/Protein	Catalog #	Enzyme/Protein Lot #	Enzyme/Protein used (ng/reaction)
EZH2-GST	50279	201109-G	40
EED-FLAG	50280	201105	10

- **Assay Conditions**

EZH2-EED binding reactions were conducted at room temperature for 60 minutes in 10 µl mixtures containing Assay Buffer, EZH2-GST, EED-FLAG and test compounds. These reactions were carried out in wells of 384-well OptiPlate. Test compounds were first incubated with EZH2 for 30 minutes prior to addition of EED and the reaction mixtures were further incubated for another 30 minutes. After incubation for the binding step, 20 µl of anti-Flag acceptor beads (1:500) and GSH-donor beads (1:125) in Assay Buffer were added, and the mixtures were incubated for 60 minutes at room temperature and protected from light. The signal was then measured using an AlphaScreen microplate reader.

- **Data Analysis**

Binding assays were performed in duplicates. The AlphaLISA signal data were analyzed using the computer software, Graphpad Prism. In the absence of compound ("No compound"), the intensity was defined as 100% (C_e) activity. In the absence of enzyme ("Blank"), the intensity was defined as 0% (C_0) activity. The percent activity in the presence of each compound was calculated according to the following equation: % activity = $(C - C_0)/(C_e - C_0)$, where C = the intensity in the presence of compound (all values below zero were shown as zero in the table). The percentage activities were plotted on bar graphs.

Bibliography

1.   Gibaja, V. *et al.* Development of secondary mutations in wild-type and mutant EZH2 alleles cooperates to confer resistance to EZH2 inhibitors. *Oncogene* **35**, 558–566 (2016).

2.   Ferreira de Freitas, R., Ivanochko, D. & Schapira, M. Methyltransferase inhibitors: Competing with, or exploiting the bound cofactor. *Molecules* **24**, (2019).

3.   Bisserier, M. & Wajapeyee, N. Mechanisms of resistance to ezh2 inhibitors in diffuse large b-cell lymphomas. *Blood* **131**, 2125–2137 (2018).

4.   Qi, W. *et al.* An allosteric PRC2 inhibitor targeting the H3K27me3 binding pocket of EED. *Nature Chemical Biology* **13**, 381–388 (2017).

5.   Li, L. *et al.* Discovery and molecular basis of a diverse set of polycomb repressive complex 2 inhibitors recognition by EED. *PLoS ONE* **12**, 1–18 (2017).

6.   Margueron, R. *et al.* Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* **461**, 762–767 (2009).

7.   Perricone, U. *et al.* A Molecular Dynamics–Shared Pharmacophore Approach to Boost Early-Enrichment Virtual Screening: A Case Study on Peroxisome Proliferator-Activated Receptor α. *ChemMedChem* **12**, 1399–1407 (2017).

8.   Wieder, M., Perricone, U., Boresch, S., Seidel, T. & Langer, T. Evaluating the stability of pharmacophore features using molecular dynamics simulations. *Biochemical and Biophysical Research Communications* **470**, 685–689 (2016).

9.   Bowman, A. L. & Makriyannis, A. Approximating protein flexibility through dynamic pharmacophore models: Application to fatty acid amide hydrolase (FAAH). *Journal of Chemical Information and Modeling* **51**, 3247–3253 (2011).

10.   Xu, L. *et al.* Molecular modeling of the 3D structure of 5-HT1AR: Discovery of novel 5-HT1AR agonists via dynamic pharmacophore-based virtual screening. *Journal of Chemical Information and Modeling* **53**, 3202–3211 (2013).

11.   Erickson, J. A., Jalaie, M., Robertson, D. H., Lewis, R. A. & Vieth, M. Lessons in Molecular Recognition: The Effects of Ligand and Protein Flexibility on Molecular Docking Accuracy. *Journal of Medicinal Chemistry* **47**, 45–55 (2004).

12.   Cozzini, P. *et al.* Target Flexibility: *Journal of medicinal chemistry* **51**, 6237–6255 (2008).

13.   Cozzini, P., Dellafiora, L., Ginex, T. & Spyrakis, F. How Protein Flexibility Can Influence Docking/Scoring Simulations. *In Silico Drug Discovery and Design* 411–440 (2015). doi:10.1201/b18799-20

14.   Perricone, U. *et al.* An overview of recent molecular dynamics applications as medicinal chemistry tools for the undruggable site challenge. *MedChemComm* **9**, (2018).

15.   Hatmal, M. M., Jaber, S. & Taha, M. O. Combining molecular dynamics simulation and ligand-receptor contacts analysis as a new approach for pharmacophore modeling: beta-secretase 1 and check point kinase 1 as case studies. *Journal of Computer-Aided Molecular Design* **30**, 1149–1163 (2016).

16.   Leach, A. R., Gillet, V. J., Lewis, R. A. & Taylor, R. Three-dimensional pharmacophore methods in drug discovery. *Journal of Medicinal Chemistry* **53**, 539–558 (2010).

17.   cozzini2015.pdf.

18.   Habash, M., Abuhamdah, S., Younis, K. & Taha, M. O. Docking-based comparative intermolecular contacts analysis and in silico screening reveal new potent acetylcholinesterase inhibitors. *Medicinal Chemistry Research* **26**, 2768–2784 (2017).

19.   Sanders, M. P. A. *et al.* From the protein's perspective: The benefits and challenges of protein structure-based pharmacophore modeling. *MedChemComm* **3**, 28–38 (2012).

20.   Triballeau, N., Bertrand, H. O. & Acher, F. *Are You Sure You Have a Good Model? Pharmacophores and Pharmacophore Searches* (2006). doi:10.1002/3527609164.ch15

21.   Davies, M. *et al.* ChEMBL web services: Streamlining access to drug discovery data and utilities. *Nucleic Acids Research* **43**, W612–W620 (2015).

22.   Berman, H. M. *et al.* The protein data bank. *Acta Crystallographica Section D: Biological Crystallography* **58**, 899–907 (2002).

23.   Michalik, L. *et al.* Combined simulation and mutagenesis analyses reveal the involvement of key residues for peroxisome proliferator-activated receptor α helix 12 dynamic behavior. *Journal of Biological Chemistry* **282**, 9666–9677 (2007).

24.   Legge, F. S., Budi, A., Treutlein, H. & Yarovsky, I. Protein flexibility: Multiple molecular dynamics simulations of insulin chain B. *Biophysical Chemistry* **119**, 146–157 (2006).

25.   Buch, I., Giorgino, T. & De Fabritiis, G. Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 10184–10189 (2011).

26.   Perez, J. J., Tomas, M. S. & Rubio-Martinez, J. Assessment of the Sampling Performance of Multiple-Copy Dynamics versus a Unique Trajectory. *Journal of Chemical Information and Modeling* **56**, 1950–1962 (2016).

27.   Wolber, G. & Langer, T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *Journal of Chemical Information and Modeling* **45**, 160–169 (2005).

28.   Langer, T. & Hoffmann, R. D. Pharmacophore Modelling. *Encyclopedia of Molecular Pharmacology* 960–960 (2008). doi:10.1007/978-3-540-38918-7_6434

29.   Yang, S. Y. Pharmacophore modeling and applications in drug discovery: Challenges and recent advances. *Drug Discovery Today* **15**, 444–450 (2010).

30.   Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *Journal of Computer-Aided Molecular Design* **27**, 221–234 (2013).

31.   Shelley, J. C. *et al.* Epik: A software program for pKa prediction and protonation state generation for drug-like molecules. *Journal of Computer-Aided Molecular Design* **21**, 681–691 (2007).

32.   Greenwood, J. R., Calkins, D., Sullivan, A. P. & Shelley, J. C. Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. *Journal of Computer-Aided Molecular Design* **24**, 591–604 (2010).

33.   Bowers, K. J. *et al.* Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. *ACM/IEEE SC 2006 Conference (SC'06)* 43–43 (2007). doi:10.1109/sc.2006.54

34.   Harder, E. *et al.* OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *Journal of Chemical Theory and Computation* **12**, 281–296 (2016).

35.   Shivakumar, D. *et al.* Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the opls force field. *Journal of Chemical Theory and Computation* **6**, 1509–1519 (2010).

36.   Friesner, R. A. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry* **47**, 1739–1749 (2004).

37.   Kim, S. *et al.* PubChem 2019 update: Improved access to chemical data. *Nucleic Acids Research* **47**, D1102–D1109 (2019).

38.   Sastry, M., Lowrie, J. F., Dixon, S. L. & Sherman, W. Large-scale systematic analysis of 2D fingerprint methods and parameters to improve virtual screening enrichments. *Journal of Chemical Information and Modeling* **50**, 771–784 (2010).

39.   Duan, J., Dixon, S. L., Lowrie, J. F. & Sherman, W. Analysis and comparison of 2D fingerprints: Insights into database screening performance using eight fingerprint methods. *Journal of Molecular Graphics and Modelling* **29**, 157–170 (2010).

40.   Schrödinger. LigPrep 2.3 User Manual. *Schrödinger Press* 1 (2009).

APPENDIX

NEW LANDSCAPES IN HALOGEN-GUIDED HOMOLOGATION CHEMISTRY OF IMINE SURROGATES

SUPERVISOR
PROF. Dr. VITTORIO PACE

1. INTRODUCTION

Carbenoids are organometallic compounds containing a metal atom (e.g. Mg, Li) and at least one electronegative element (e.g. halogen). These reagents have proven to be useful homologating agents. The homologation reaction is a chemical operation that converts the reactant into the next member of the homologous series. The most common homologation reactions increase the number of methylene (-CH$_2$-) units in saturated chain within the molecule. The classic homologation reaction is the Arndt-Eistert reaction which allows the formation of homologated carboxylic acids or their derivatives by reacting the activated carboxylic acids with diazomethane, resulting in the formation of α-diazoketone.[1–3] The issue related to the safety profile of diazomethane, leaded to the development of alternative safer reactions, including Witting[4–6], Corey-Chaykovsky[7], Morita-Baylis-Hillman[8], and Köbrich[9] reactions (Scheme **1**).
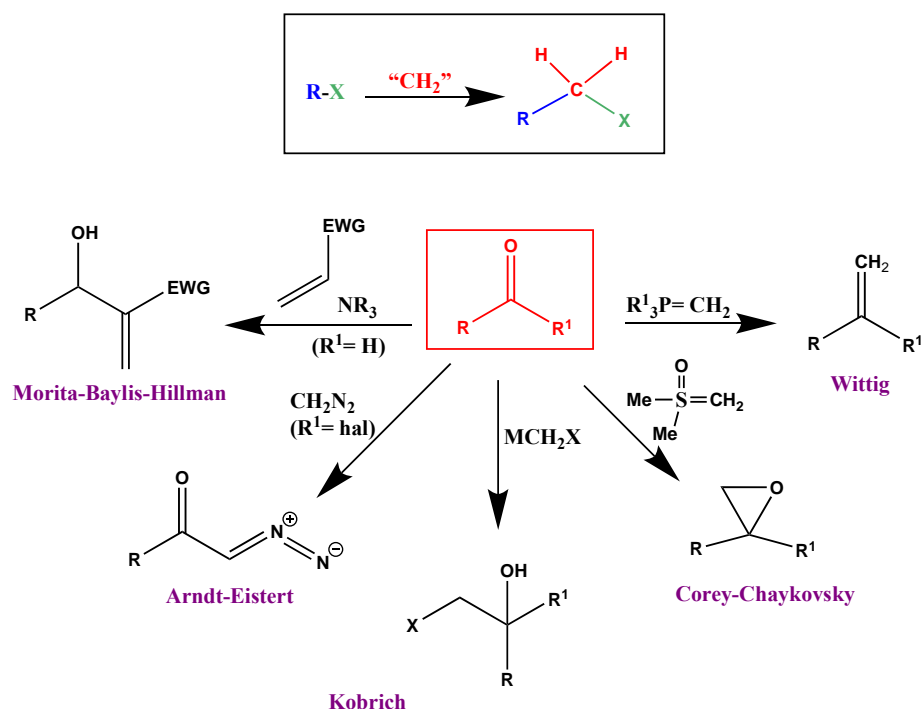

Scheme 1. Classical context of homologation in synthetic chemistry

Carbenoids are characterized by an amphiphilic profile due to the simultaneous presence of an electron-donating and electron-withdrawing substituent linked at the same carbon center. Indeed, they can react either as nucleophile or electrophile depending on reaction conditions, in particular on the nature of the metal and temperature.[9]

The presence of electropositive metal, such as lithium and magnesium induces a nucleophilic character. Carbenoids with less positive metal, such as zinc or rhodium act as electrophile. In addition, low temperature induces a prevalence of the nucleophilic behavior, while higher temperature facilitates the electrophilic one.

Due to their versatility, carbenoids present several advantages including their capability to insert, by one-step reaction, a functionalized carbon responsive for further modifications. Despite the high potential of these reagents, the main limitation is represented by the thermal instability leading to a degradation process through α-elimination (Scheme **2**). This is due to an internal coordination between the metal and the halogen resulting in a free carbene and a metal halide salt.[10,11] *De facto* the preparation of these reagents requires full chemo-control and a compromise between stability and reactivity.

The degradation issue can be solved conducting the reactions under Barbier-type conditions, at low temperatures (-78 °C), in the presence of lithium halide and ethereal solvents.[12–16]
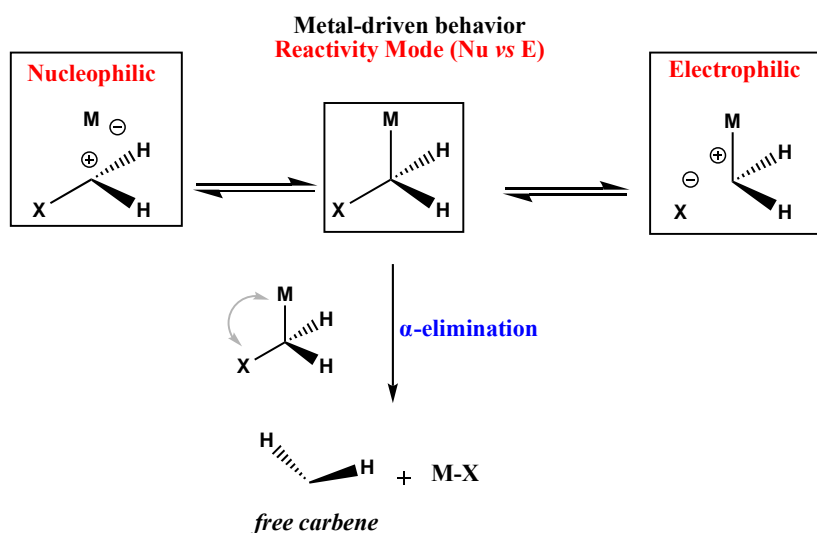

Scheme 2. Reactivity profile of carbenoids and their degradation process.

1.1 PREPARATION OF CARBENOIDS SPECIES

Carbenoids can be prepared *via* different methods: metal-halide exchange, metal-proton exchange (i.e., deprotonation), metal-sulfinyl exchange and metal-tin exchange.

- **Metal-halide exchange**

The lithium-halogen exchange is one of the most common preparative method, due to the wide availability of dihalomethane precursors. MeLi has proven to be useful source of lithium[17] but recently the MeLi-LiBr complex in Et_2O has become the reagent of choice for preparing chloromethyllithium ($LiCH_2Cl$) starting from chloroiodomethane (ICH_2Cl) (Scheme **3**).[2,18–20]
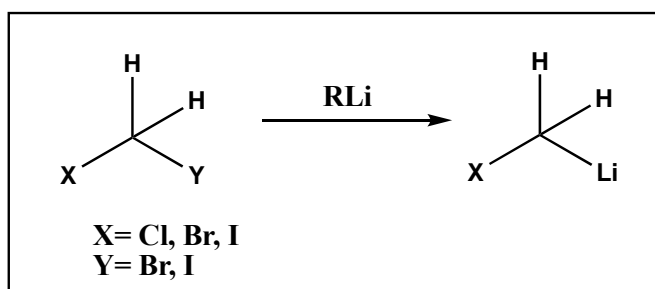
Scheme 3. Lithiation via lithium–halogen exchange.

The nature of the dihalomethane plays a crucial role in the implementation of the process, ICH_2Cl is the preferred precursor. Furthermore, the stoichiometry should also be considered. Despite the reaction between dihalomethane and organolithium reagents proceeds quantitatively and could be carried out in a ratio of 1:1, is preferred to use a small excess (0.2-0.4 equiv) of dihalomethane, considering the possibility of the competitive attack of the alkyl lithium to the electrophile or variations in its concentration.[21]

The process is also affected by the nature of the electrophiles, considering that the dihalomethane has to react firstly with the organolithium reagent rather than the electrophile.[2]

In addition, the magnesium halocarbenoids can be prepare using magnesium-iodine exchange in presence of ICH_2Cl. Contrary to the lithium halocarbenoids, the magnesium

ones are more stable and they can be easily generated at -78 °C, allowing the addition of the electrophile at the later stage of the reaction.[22–25]

However, the magnesium carbenoids could be a good choice only in the presence of a potent electrophile, while with weak electrophile species it fails.[20]

- **Metal-proton exchange**

The metal-proton exchange can be employed to generate dihalomethylcarbenoids, such as $LiCHCl_2$, $LiCHBr_2$, $LiCHI_2$, using the corresponding dihalomethane and lithium amide base, such as LDA, $LNCy_2$, LTMP and LiHDMS which removes the proton from the dihalomethane (Scheme **4**).[18,20,26,27]
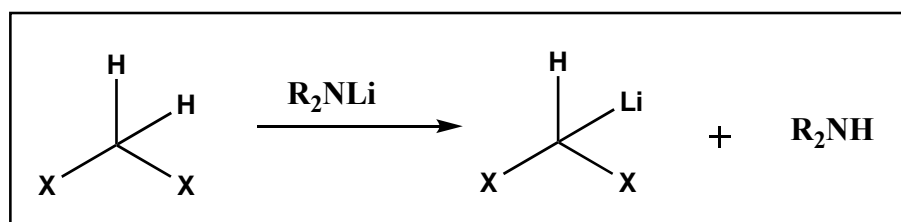
Scheme 4. Preparation of carbenoids via lithium-proton exchange.

- **Metal-sulfinyl exchange**

Hoffman[28] and Blakemore[29] reported an alternative method of magnesium and lithium carbenoids preparation respectively. The aryl group is attached on the sulfinyl group and the inversion of stereochemistry at sulfur occurs (Scheme **5**).
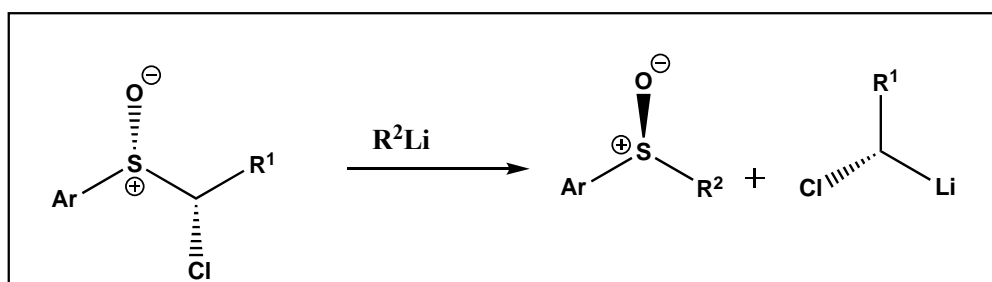

Scheme 5. Lithium carbenoids from halogenated arylsulfoxides.

Contrary to the dihalomethanes which are commercially available the starting α-halosulfoxides need to be prepared.[30,31]

- **Metal-tin exchange**

In this preparation method the nature of alkyl lithium affects the reaction outcome. The MeLi showed better performance compared to *n*-BuLi, which leads to the formation of undesired compounds (Scheme **6**).[2]
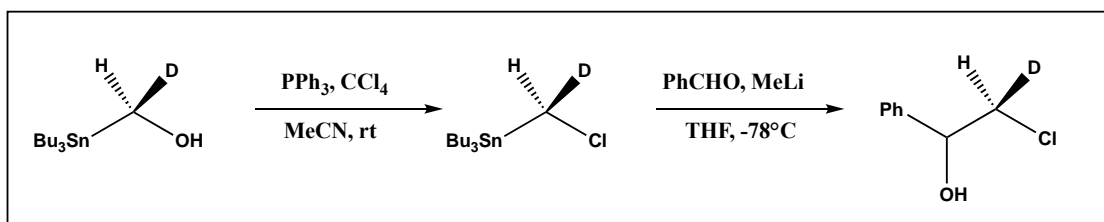

Scheme 6. Hammerschmidt's preparation of chiral chloromethyllithium via tin-lithium exchange.

## 1.2 Electrophilic partners for carbenoids

1.2.1 Carbonyl Adducts

The carbonyl compounds have proven to be optimal electrophilic counterparts on reactions with carbenoids. The addition of halomethyllithium to an aldehyde or ketone leads to the formation of β-halohydrin, which can be subsequently converted into an epoxide. The ring closure can take place through increase of temperature or through a base-assisted process. In this context Pace and co-workers reported the addition of chloromethyllithium to unsaturated cyclic ketones (Scheme **7**).[32]
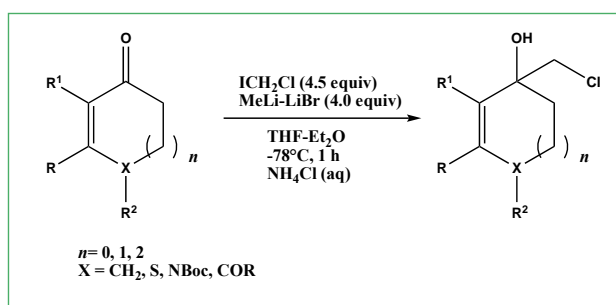
*Scheme 7. Addition of chloro-methyllithium to cyclic enones.*

Furthermore, another example of high electrophilic substrate for this reaction is the carbonyl group of isatins. Pace and co-workers also reported the formation of spiro-epoxindoles through LiCH$_2$Cl addition to isatins followed by base-mediated ring closure. (Scheme **8**)[33]
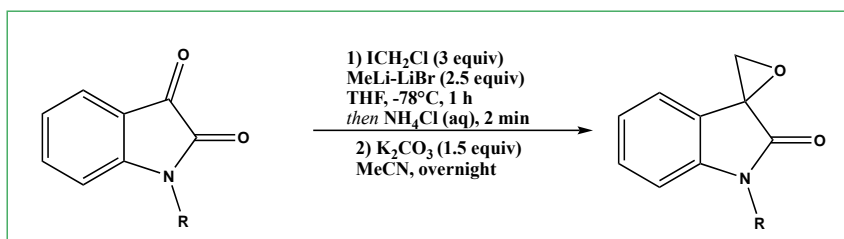
Scheme 8. Synthesis of spiro-epoxyoxindoles through LiCH$_2$Cl addition to isatins followed by base-mediated ring closure.

1.1.2 Imines

The addition of halomethyllithium to an imine generates the corresponding aziridine cycle by ring-closure. The use of carbenoids in this kind of reaction has been well explored. Savoia *et al.* reported the formation of aziridines starting from 2-pyridinecarboxaldehydes and (*S*)-valinol, in which the use of LiCH$_2$Cl leaded high stereocontrol and high yield. (Scheme **9**)[33]
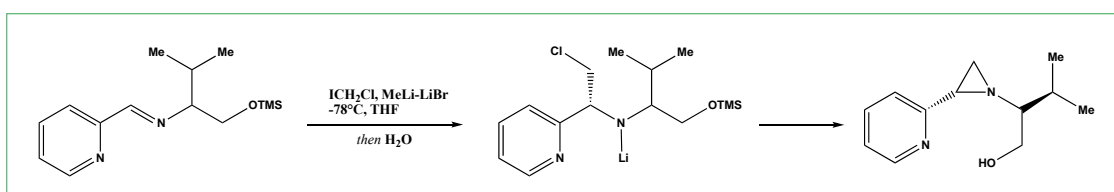
Scheme 9. Carbenoids-mediated 2-aziridination of 2-pyridinimines.

Aziridines were also generated through the addition of lithium iodocarbenoids to sulfonyl-protected imines, as reported by Concellón *et al.*[34,35]
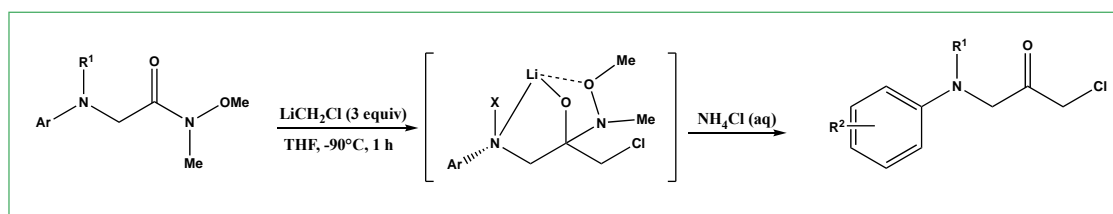
1.2.3 Weinreb Amides

The homologation reaction of Weinreb amides in presence of lithium halocarbenoids leads to generation of α-haloketones.[36]

Berluenga described the formation of α-haloketones *via* homologation of esters using carbenoids under Barbier-type conditions.[37] The intermediate reported is stable under reaction conditions, probably due to the presence of an electronegative halogen and an oxygen substituent preventing the elimination of alkoxide group.

However, this process does not occur in presence of amino or amido groups in which the nitrogen is linked with a hydrogen atom, because the carbenoids extract the acid proton rather than attacking the electrophile.

Subsequently, the use of Weinreb amide proved to be successful to overcome this issue, due to the higher stability of the tetrahedral intermediate formed during the reaction of Weinreb amides and carbenoids (Scheme **10**). Indeed, in the Weinreb amide intermediate the nitrogen group is chelated by the methoxy group and its basicity does not affect the stability of the intermediate.[37]


Scheme 10. Homologation of α-arylamino Weinreb amides.

Pace's group employed the homologation of Weinreb amide for the synthesis of several compounds, including the synthesis of Nelfinavir,[38] (an important HIV inhibitor, α-cyanoketones,[6] α-oxyketones,[39] β-oxothioethers,[40] α-aryl and α-alkyl-selenomethyl ketones.[41]

1.2.4. Disulfides and Diselenides

Pace and co-workers reported the homologation of disulfides to the corresponding dithioacetals using bromo-methyllithium carbenoid (Scheme **11**).[42]
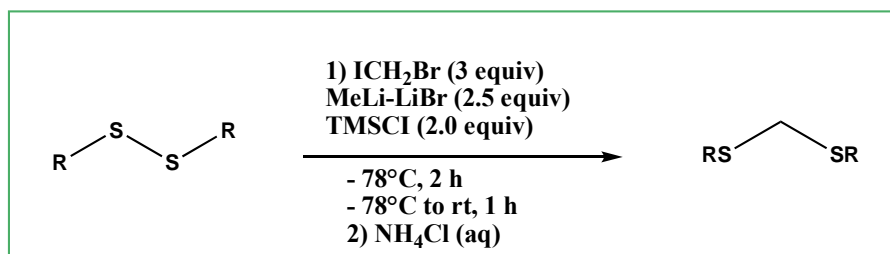
Scheme 11. Homologation of disulfides.

In this procedure chemoselectivity is fully maintained in presence of additional sensitive functionalities towards organolithium (e.g., aryl halides, heterocycles, esters and amides). In addition, the procedure could be exploited for homologating asymmetric disulfides and diselenides.

1.2.5. Heterocumulenes

Pace's group described the chemoselective synthesis of α-haloacetamides through the addition of halocarbenoids to isocyanates (Scheme **12**).[19]
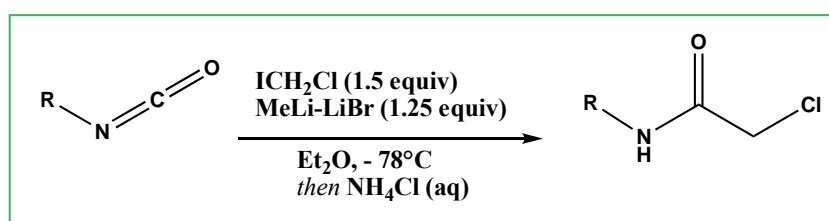
Scheme 12. N-substituted 2-Haloacetamides synthesis through isocyanates homologation.

1.3 Fluorocarbenoids

The introduction of fluoromethyl-group into organic skeleton would be highly desirable and could be achieved using fluorine-containing carbenoids. These carbenoids are highly instable and their preparation has been challenging. In 2017 Pace and co-workers proposed a new protocol for the preparation of fluorocarbenoids starting from the commercially available fluoroiodomethane. The optimization study highlighted the main conditions required for the success of the preparation:

a) Barbier-type conditions; b) MeLi·LiBr was used as the optimal lithiating agent; c) precise stoichiometry of the electrophile / MeLi·LiBr / ICH_2F = 1 : 1.5 : 2.0; d) use of THF: Et_2O 1:1 v/v mixture solvents (scheme **13**).[43]
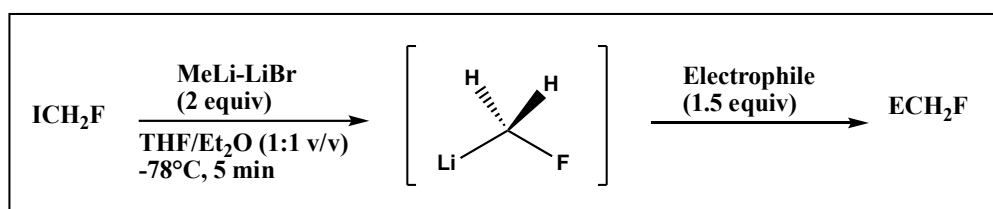
Scheme 13. Generation and use of fluoromethyllithium.

The employment of fluorocarbenoids proved to be successful in a series of homologation reactions. For instance, the addition of $LiCH_2F$ to α-, β-, or γ-halogeneted ketones followed by a base-triggered ring closure lead to epifluorohydrins and homologues (Scheme **14**).
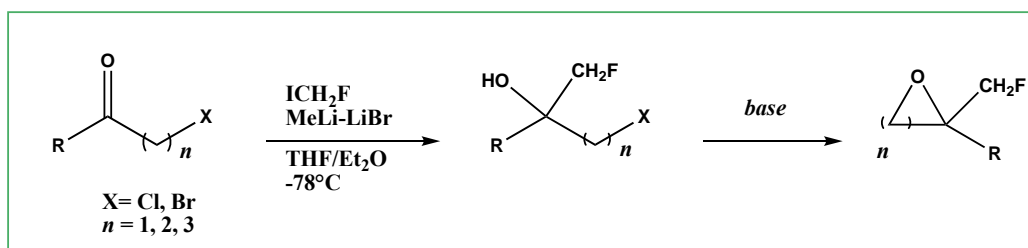
Scheme 14. Haloketone fluoromethylation/ring closure process.

In 2019 Pace and co-workers reported the employment of LiCH$_2$F for homologating ketones and imine followed by ring closure reaction, leading to the formation of α-fluoroepoxides and α-fluoroaziridines (Scheme **15**).[38,44–46]
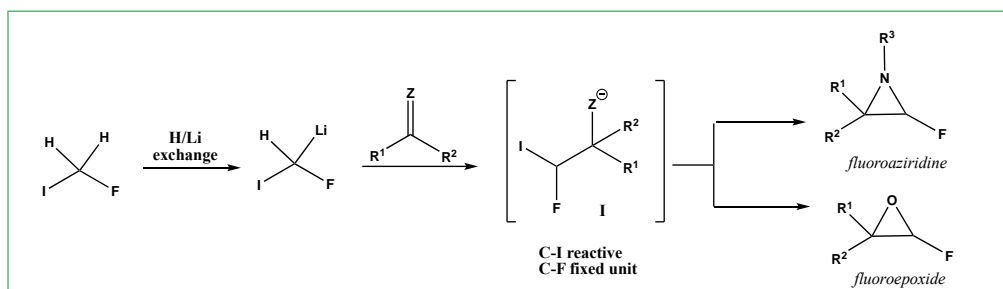
Scheme 15. Deprotonation of fluoroiodomethane for preparing fluoro-iodomethyllithium and reactions with (aza)carbonyls.

In this work benzophenone was chosen as model electrophile and LiN(*i*-Pr)Cy lithium amide base proved to be the best choice. The reaction showed good chemoselectivity in presence of sensitive groups in the molecular core, such as aldehyde, ketone, nitrile, alkene.[44]

2. RESULTS AND DISCUSSION

In 2019 Ielo *et al.* reported the telescoped homologation of imine surrogates into aziridines using carbenoids. In detail, α-chloro- and α-halomethyl-trifluoromethyl aziridine were synthesised *via* homologation of trifluoroacetimidoyl chlorides (TFAICs), which proved to be suitable electrophiles.

This work highlighted the interesting behavior of these electrophilic in reaction with carbenoids, enabling the insertion of either one or two nucleophilic elements by controlling the stoichiometry of the reaction. Indeed, by using 1.2 equivalent of $LiCH_2Cl$ a single C1 unit was introduced, obtaining the chloro-trifluoromethyl aziridine, whereas in presence of an excess of carbenoid (2.8 equivalent) a C1-C1 *bis* homologation was achieved, forming a series of chloro-methyl-trifluoromethyl-aziridines (Scheme **16**).
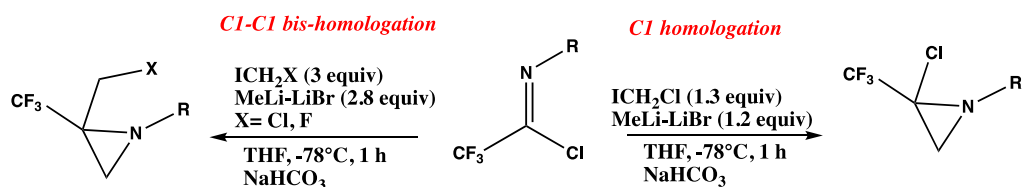
Scheme 16. Telescoped homologations of TFAICs to chloro-trifluoro- methylaziridines and chloromethyl-trifluoromethyl aziridines.

A plausible mechanism that can explain the chemoselective homologation involves the formation of a tetrahedral intermediate in which presumably there is a coordination between the lithium cation and the chlorine atom forming, after an internal rearrangement, the mono-homologated chloroaziridine. Whereas, in presence of an excess of homologating agent (both $LiCH_2Cl$ and $LiCH_2F$ were used), the intermediate chloroaziridine spontaneously forms the azirinium ion which, due to its high electrophilicity, reacts with a second equivalent of carbenoid furnishing the chloro-methyl aziridine (Scheme **17**).
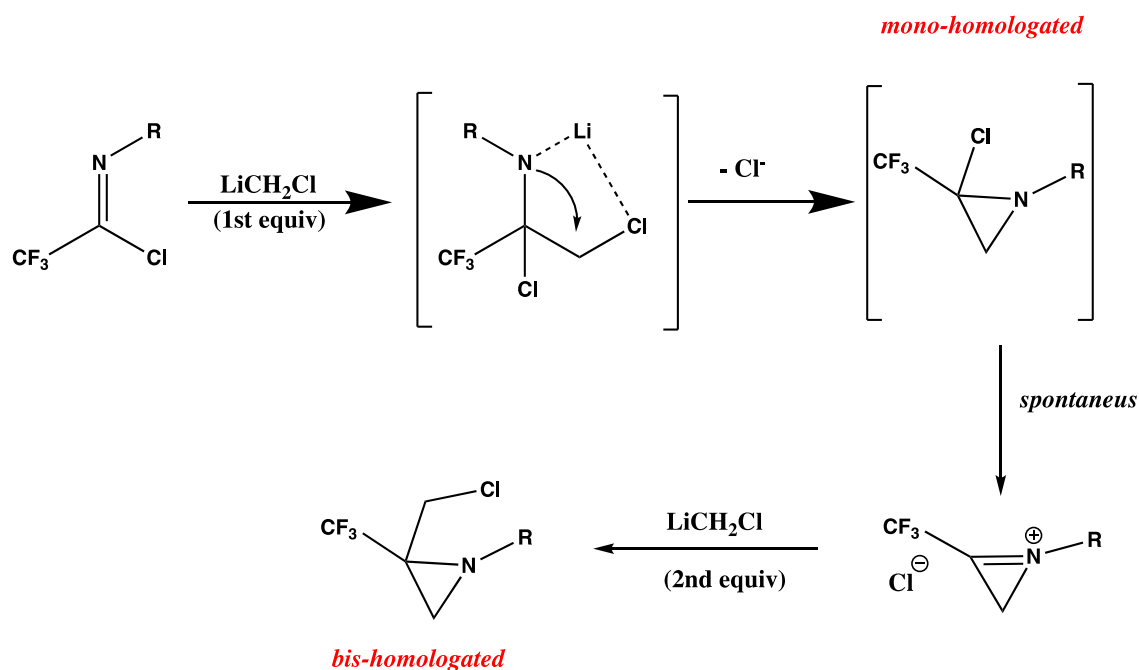
Scheme 17. Plausible mechanism governing C1 and C1-C1 homologations.

With these results in hand, we decided to investigate the reactivity of chloro-trifluoroimidates using LiCH₂I as homologating agent. What we obtained was a completely different reaction product, due to the formation of a dimeric derivative (Scheme **18**).
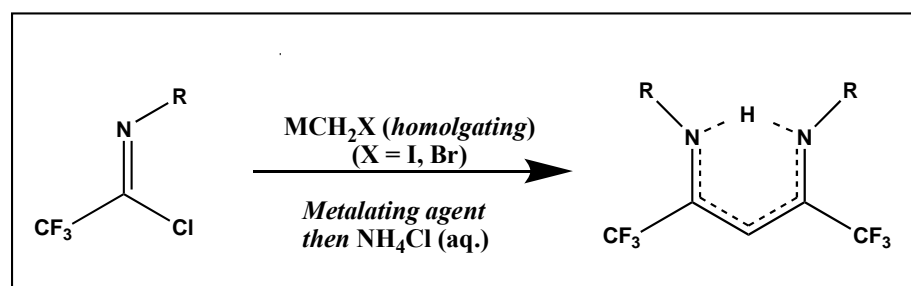
Scheme 18. Synthesis of bis- trifluoromethyl-β-diketiminates.

2.1 Reaction optimization

In order to afford the desired product possessing two TFAIC units, 2 equivalents of the starting TFAIC were used. When $LiCH_2I$ was used for the first homologation followed by the addition of MeLi-LiBr as metalating agent, compounds **2** and **3** were obtained in 76:24 ratio, while when LiBr-free MeLi was employed, an increase of compound **3** was noticed. The use of other organolithium derivatives for generating $LiCH_2I$, such as *s*-BuLi-LiBr or *t*-BuLi-LiBr did not significantly affect the reaction outcome. Whereas $LiCH_2Br$ exhibited a suitable reactivity as homologating agent affording compound **2** with a slight lower yield. Furthermore, variability in terms of yield and rate of conversion was reported when diethyl ether or toluene were used as solvent instead of THF. Afterwards, in order to maximize the ratio and conversion of the reaction, different equivalent of $LiCH_2I$ were used, identifying as best condition that one reported in entry 15, in which the desired compound **2** was obtained in 91% yield.

In details, the protocol includes:

a) The addition of the CH_2I_2 occurs in one-pot at the beginning of the reaction, followed by the addition of the minimum amount of MeLi-LiBr to form the $LiCH_2I$, acting as homologating agent.

b) The residual CH_2I_2 remains as *reservoir* until a second aliquot of MeLi-LiBr is added for generating $LiCH_2I$ that, this time, acts as metalating agent.
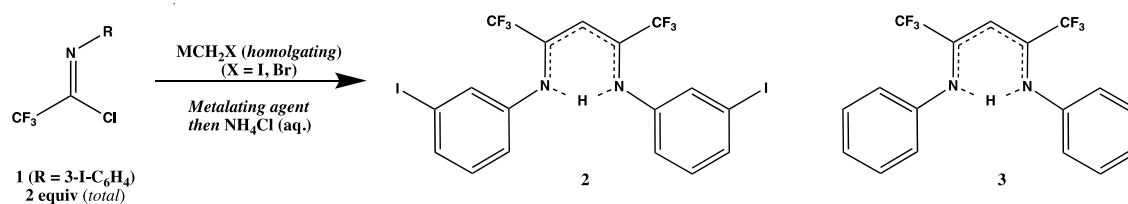
| Entry	Homologating agent (equiv)[a]	Metalating agent (equiv)	Conversion[b]	Ratio 2/3	Yield of 2 (%)[c]
1	LiCH$_2$I (1.2)	MeLi-LiBr (1.2)	41	76:24	26
2	LiCH$_2$I (1.2)	MeLi (1.2)	36	55:45	14
3	LiCH$_2$I (1.2)	n-BuLi-LiBr (1.2)	39	68:32	22
4[d]	LiCH$_2$I (1.2)	n-BuLi-LiBr (1.2)	35	41:59	11
5	LiCH$_2$I (1.2)	s-BuLi-LiBr (1.2)	31	63:37	15
6	LiCH$_2$I (1.2)	t-BuLi-LiBr (1.2)	36	71:29	23
7	LiCH$_2$Br (1.2)	MeLi-LiBr (1.2)	33	70:30	19
8	LiCH$_2$I (1.8)	MeLi-LiBr (1.8)	72	81:29	54
9	LiCH$_2$I (2.3)	MeLi-LiBr (2.3)	74	80:20	55
10[e]	LiCH$_2$I (1.8)	MeLi-LiBr (1.8)	18	83:17	10
11[f]	LiCH$_2$I (1.8)	MeLi-LiBr (1.8)	27	79:21	16
12	LiCH$_2$I (1.8)	LiCH$_2$I (1.8)	87	93:7	76
13	LiCH$_2$I (1.8)	LiCH$_2$I (1.3)	83	90:10	70
14	LiCH$_2$I (1.3)	LiCH$_2$I (1.8)	89	98:2	83
15	LiCH$_2$I (1.3)	LiCH$_2$I (1.5)	95	>99:1	91

Table 1. Reaction optimization. [a] *Unless stated otherwise (entries 2-6), LiCH2I was generated under Barbier conditions in THF starting from CH2I2 and MeLi-LIBr at -78 °C for 1 h. The metalating agent was added after 5 min from the end of the homologation step.* [b] *The ratio has been calculated by 1H-NMR analysis using 1,3,5- trimethylbenzene as internal standard.* [c] *Isolated yield.* d *TMEDA (1.2 equiv) was added.* [e] *Reaction run in Et2O.* [f] *Reaction run in toluene.*

2.2 Scope of the reaction

The developed methodology was applied on different halogen substituted TFAIC substrates. The *ortho*, meta, para positions of aromatic ring of TFAIC derivatives were explored with different halogen groups, obtaining the bromo analogues **4** and **5**, chlorinated derivatives (**6-8**), as well as monofluoro (**10**) and trifluoromethyl (**11**) derivatives. Furthermore, TFAICs with electrophilic groups were also employed, exhibiting a significant chemoselectivity. Compound **12,** possessing nitrile group was also obtained, and derivatives **13** and with **14** *t*-Bu and Et-ester groups, respectively. The naphtyl-derivative (**15**), was used for the X-ray analysis, furnishing important features of β-diketiminates structure. The use of TFAIC derivatives containing groups with different steric hindrance did not affect the reaction outcome. Indeed, compounds with *n*-Bu (**16**), 2,5-diMe (**17**) and *i*-Pr (**18**) groups, were obtained with high and similar yield. The methodology proved to be successful also for analogues with vinyl motif (**19**) and terminal alkine (**20**) that could undergo to deprotonation under the reaction condition. TFAICs with nitrogen-centered group reacted to afford compound **21** with morpholino moiety, and compounds **22** and **23** with nitro- and diazo- groups, respectively. Furthermore, the methodology was also explored with substrate presenting electro-donating groups, including ethers [3-ethoxy- (**24**), 2,5-dimethoxy- (**25**)], cyclic acetal [1,3-benzodioxolane (**26**)] and fluorinated ether [4-trifluoromethoxy- (**27**)]. Derivatives with sulfur substituents such as 4-sulfonamido-(**28**) and 4-thiomethyl (**29**) were also obtained. Moreover, compound **30** with 4-phenylseleno- substituent did not show Se/Li exchange obtaining the desired compound in 90% yield.
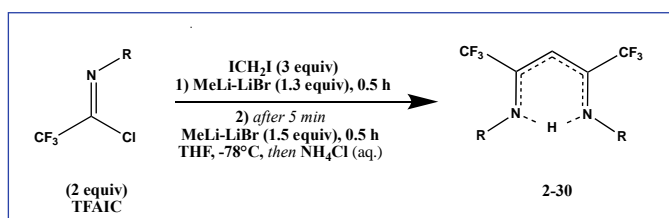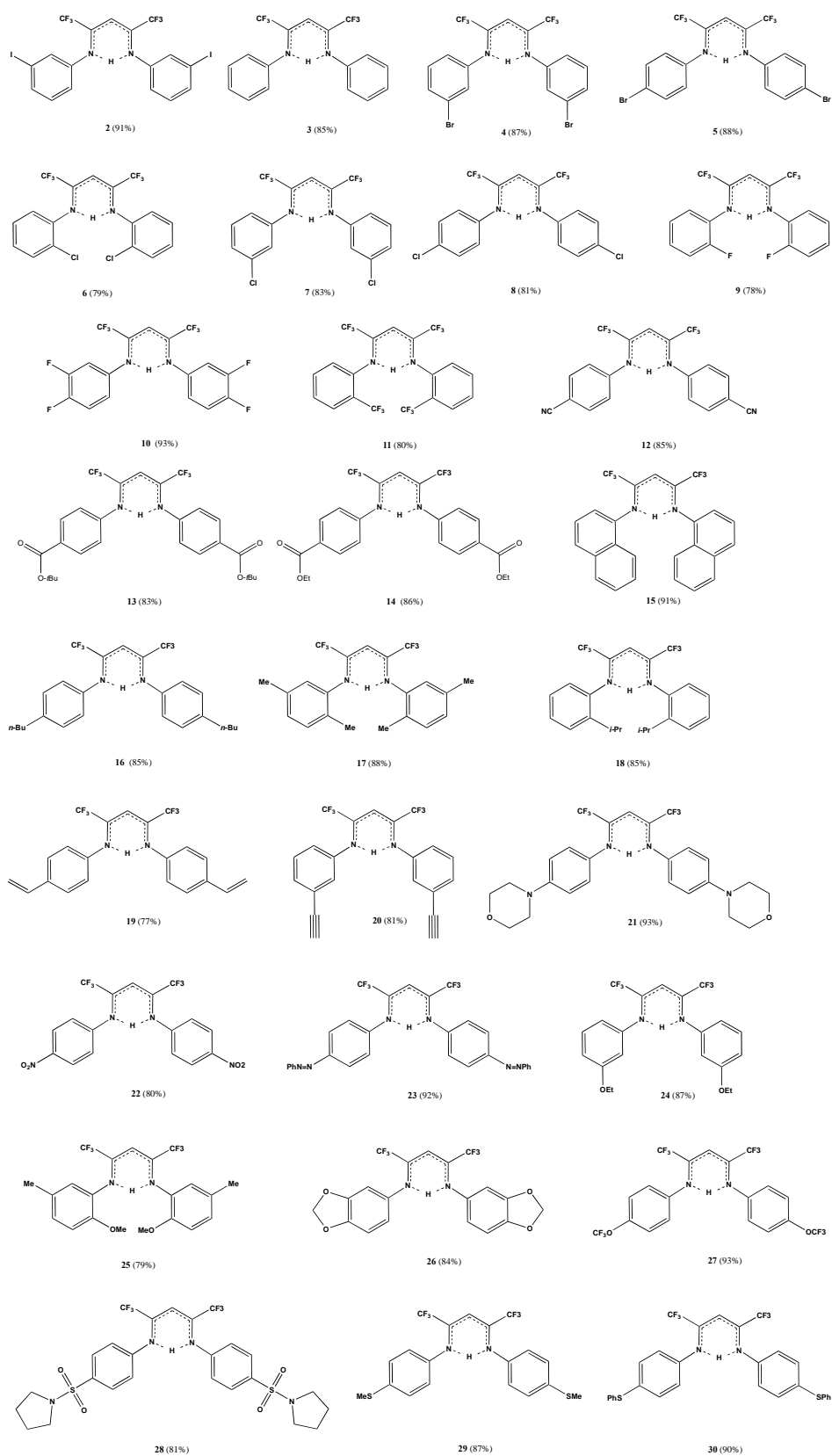
Scheme 19. Scope of the reaction.

In addition, the high level of chemocontrol exhibited during the reaction with ester derivatives, inspired us the attempt to use a TFAIC decorated with a Weinreb amide as substrate in order to investigate the possibility to obtain consecutive homologations. Indeed, by simply controlling the stoichiometry of the reaction, the mono- (**32**) and di- (**33**) functionalized compounds were obtained (Scheme **20**).
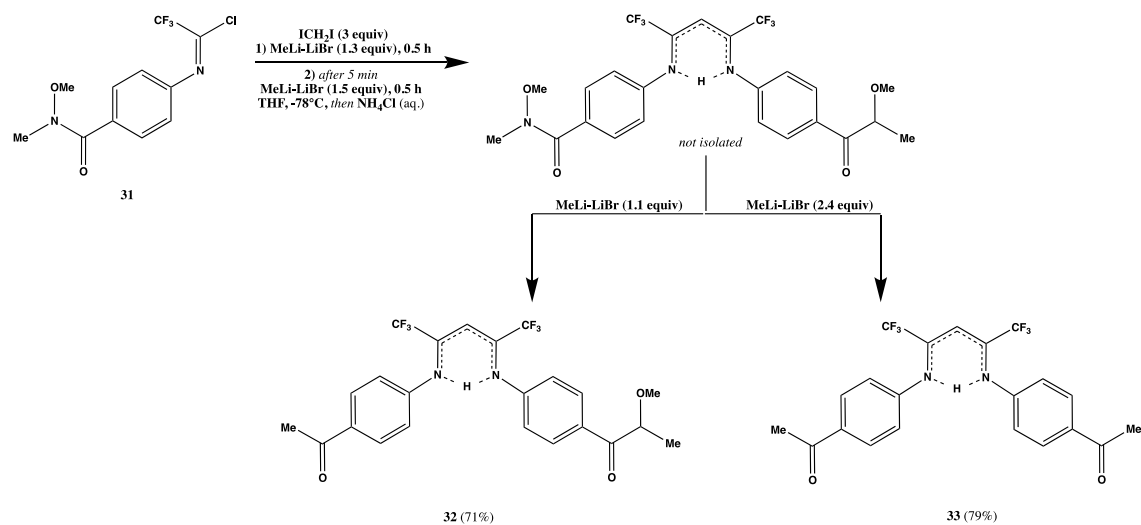
Scheme 20. Consecutive homologation-diversification of a Weinreb amide decorated TFAIC.

2.3 Plausible mechanism of the reaction

The plausible mechanism of the reaction involves the formation of a tetrahedral intermediate (B) subsequent to the addition of LiCH$_2$I to TFAIC (A), which trough the elimination of a chlorine atom, lead to the formation of α-iodomethylimine species (C). Subsequently, the lithiating capability of the reagent comes into play, thus providing the corresponding α-lithium imine (D) (via I/Li transmetalation) which attacks a second molecule of chlorotrifluoroimidates and, following a similar addition/elimination process (E) lead to the homologated adduct. This compound undergoes a tautomeric equilibrium that provides the final product in the G form (Scheme **21**).

The presence of CH$_2$I$_2$ in the crude mixture indicates that subsequently it is consumed in the homologation step, but it is then reformed during the metalating step. Indeed, when

MeLi, n-BuLi and LiCH₂Br were used as metalating agents, the GC-MS analysis reveals the peak of MeI, n-BuI and Br CH₂I, providing experimental evidence of the mechanism hypothesized above.
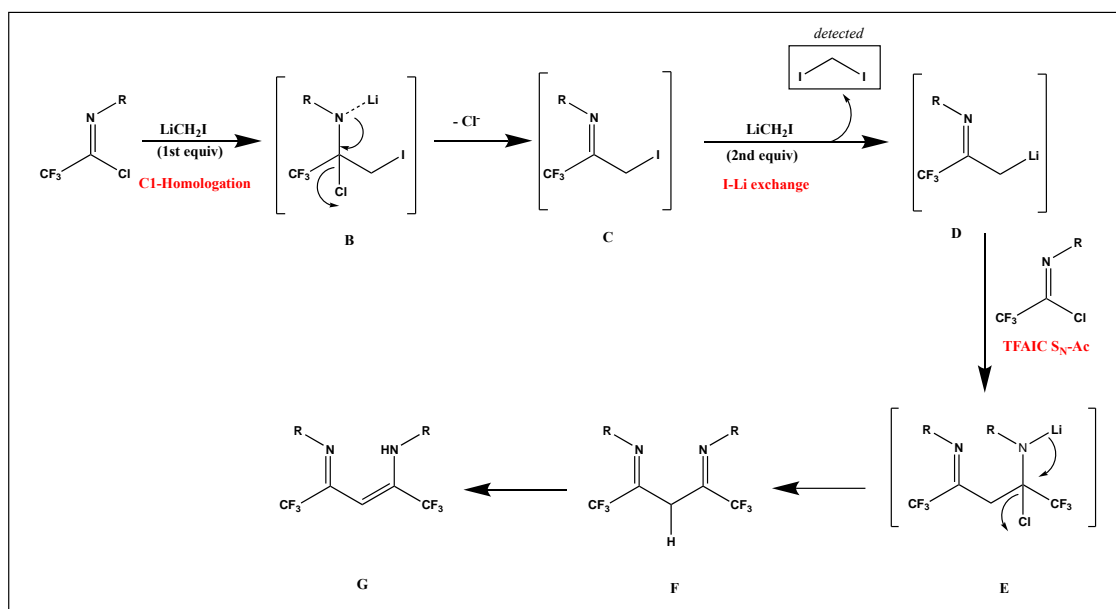
Scheme 21. Plausible mechanism of the reaction.

2.4 X-ray analysis

Definitive proof of the structures was deducted from NMR experiments and X-ray crystallography. The X-ray structural analysis of compound **15** (Figure **1**) was carried out in collaboration with Dr. A. Roller (Institute of Inorganic Chemistry, University of Vienna).
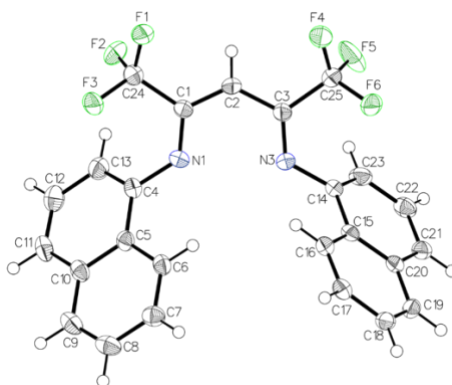
Figure 1. X-ray structural analysis of compound 15.

2.5 CONCLUSIONS

In summary, we reported the synthesis of new fluorinated β-diketiminates which are promising candidates to be used as ligands. The reaction was carried out starting from TFAICs and LiCH$_2$I carbenoids. The new developed methodology exploits the capability of iodomethyllithium to act contemporaneously as homologating (C1 source) and metalating agent. The excellent chemocontrol was highlighted by the several sensitive substrates employed. The structural similarity with the *Nac-Nac* ligands, could suggest an employment of trifluoromethyl-β-diketimites as metal-chelator agents.

2.6 MATERIAL AND METHODS

- **General Methods**

Melting points were determined on a Reichert–Kofler hot-stage microscope and are uncorrected. Mass spectra were obtained on a Shimadzu QP 1000 instrument (EI, 70 eV) and on a Bruker maXis 4G instrument (ESI-TOF, HRMS). ^{1}H, ^{13}C, ^{19}F, ^{15}N and ^{77}Se NMR spectra were recorded with a Bruker Avance III 400 spectrometer (400 MHz for ^{1}H, 100 MHz for ^{13}C, 376 MHz for ^{19}F, 40 MHz for ^{15}N and 76 MHz for ^{77}Se) at 298 K using a directly detecting broadband observe (BBFO) probe. The center of the (residual) solvent signal was used as an internal standard which was related to TMS with δ 7.26 ppm (^{1}H in CDCl$_3$) and δ 77.0 ppm (^{13}C in CDCl$_3$). ^{19}F NMR spectra were referenced via the Ξ ratio (absolute referencing). ^{15}N NMR spectra (gs-HMBC, gs-HSQC) were referenced against neat, external nitromethane. ^{77}Se NMR spectra were referenced against diphenyldiselane (δ Ph$_2$Se$_2$ 463 ppm). Spin-spin coupling constants (J) are given in Hz. In nearly all cases, full and unambiguous assignment of all resonances was performed by combined application of standard NMR techniques, such as APT, HSQC, HMBC, HSQCTOCSY, COSY and NOESY experiments.

THF was distilled over Na/benzophenone. Chemicals were purchased from Sigma-Aldrich, Acros, Alfa Aesar, Fluorochem and TCI Europe, otherwise specified. Organolithium reagents were provided by Albemarle Corporation. Organolithium reagents were titrated immediately before the use according to established literature procedure. Solutions were evaporated under reduced pressure with a rotary evaporator. TLC was carried out on aluminium sheets precoated with silica gel 60F254 (Macherey-Nagel, Merck); the spots were visualized under UV light (λ = 254 nm) and/or KMnO$_4$ (aq.) was used as revealing system. Neutral Aluminium Oxide – Brockmann grade 2 (Alox-BG2) for chromatographic purifications was prepared as we previously reported.

- **General Procedures**

General Procedure 1. - Preparation of trifluoroacetimidoyl chlorides - To a solution of Ph$_3$P (3.0 equiv) in DCE was added CCl$_4$ (4.0 equiv), Et$_3$N (1.2 equiv) and TFA (1.0 equiv) at 0 °C and the mixture was stirred for 10 minutes. After the solution was cooled to room temperature, suitable aniline (1.0 equiv) was added. The mixture was then refluxed overnight. Solvent was removed under reduce pressure, and the residue was diluted and washed with *n*-hexane several times and filtered. The filtrate was concentrated under reduced pressure, the so-obtained crude mixture was subjected to chromatography (silica gel) to afford pure compounds.

General Procedure 2. - Preparation of bis-CF$_3$-β-diketiminates - To a cooled (-78 °C) solution of trifluoromethyl chloroimidate (1.0 equiv) in dry THF was added diiodomethane (3.0 equiv). After 2 min, an ethereal solution of MeLi-LiBr (1.3 equiv, 1.5 M) was added dropwise, using a syringe pump (flow: 0.200 mL/min). The resulting solution was stirred for 30 minutes at -78 °C. After 5 min from the end of the addition of the first aliquot of MeLi-LiBr, the second portion of trifluoromethyl chloroimidate (1.0 equiv) dissolved in dry THF was added. Soon after, it was commenced the addition of the second portion of MeLi-LiBr (1.5 equiv, 1.5 M) *via* syringe pump (flow: 0.200 mL/min) and, the resulting reaction mixture was stirred for further 30 minutes at -78 °C. A saturated solution of NH$_4$Cl was added (2 mL/mmol substrate), then was extracted with Et$_2$O (2 x 5 mL) and washed with water (5 mL) and brine (10 mL). The organic phase was dried over anhydrous Na$_2$SO$_4$, filtered and, after removal of the solvent under reduced pressure, the so-obtained crude mixture was subjected to chromatography (silica gel, Alox-BG2 or reverse phase silica gel) to afford pure compounds.

## • **Spectral and Characterization Data**

4-(Phenylselanyl)aniline
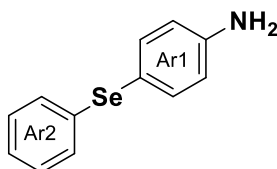
To a round-bottomed flask containing diphenyl diselenide (3339 mg, 10.7 mmol, 1.0 equiv), aniline (2000 mg, 1.96 mL, 21.5 mmol, 2.0 equiv) and CuI (123 mg, 0.65 mmol, 3 mol%), DMSO (5 mL) was added. The reaction mixture was stirred at 110 °C for 2 hours. After that time, the solution was cooled to room temperature, diluted with ethyl acetate (20 mL), and washed with water (3 x 20 mL). The organic phase was separated, dried over anhydrous Na_2SO_4 and concentrated under vacuum. The desired product was obtained in 80% yield (4113 mg) as yellow solid (m.p.: 89-91 °C, 87-91 °C) after chromatography on silica gel (50:50 v/v, n-hexane/dichloromethane).

¹H NMR (400 MHz, CDCl₃) δ: 7.41 (m, 2H, Ar1 H-3,5), 7.29 (m, 2H, Ar2 H-2,6), 7.20 (m, 2H, Ar2 H-3,5), 7.15 (m, 1H, Ar2 H-4), 6.65 (m, 2H, Ar1 H-2,6), 3.84 (brs, 2H, NH₂).

¹³C NMR (100 MHz, CDCl₃) δ: 146.7 (Ar1 C-1), 137.1(Ar1 C-3,5), 134.0 (Ar2 C-1), 130.1 (Ar2 C-2,6), 129.0 (Ar2 C-3,5), 126.0 (Ar2 C-4), 116.6 (Ar1 C-4), 116.1 (Ar1 C-2,6).

HRMS (ESI), m/z: calcd. for $C_{12}H_{12}NSe^+$: 250.0129 $[M+H]^+$; found: 250.0127.

(1Z)-2,2,2-Trifluoro-*N*-[4-(phenylselanyl)phenyl]ethanimidoyl chloride
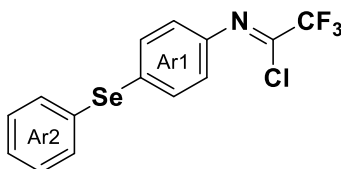
By following the General Procedure **1**, starting from 4-(phenylselanyl)aniline (550 mg, 2.2 mmol, 1.0 equiv), TFA (303 mg, 0.2 mL, 2.2 mmol, 1.0 equiv), Ph₃P (1744 mg, 6.65

mmol, 3.0 equiv), Et$_3$N (269 mg, 0.4 mL, 2.6 mmol, 1.2 equiv), CCl$_4$ (1363 mg, 0.85 mL, 8.9 mmol, 4.0 equiv) and DCE (10 mL), the desired product was obtained in 78% yield (1030 mg) as a yellow oil after chromatography on silica gel (98:02 *v/v, n*-hexane/dichloromethane).

¹H NMR (400 MHz, CDCl$_3$) δ: 7.53 (m, 2H, Ar2 H-2,6), 7.48 (m, 2H, Ar1 H-3,5), 7.33 (m, 1H, Ar2 H-4), 7.32 (m, 2H, Ar2 H-3,5), 7.04 (m, 2H, Ar1 H-2,6).
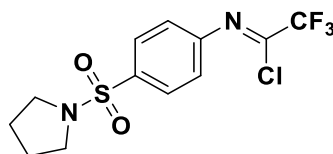
¹³C NMR (100 MHz, CDCl$_3$) δ: 142.0 (Ar1 C-1), 133.8 (Ar2 C-2,6), 132.7 (Ar1 C-3,5), 131.8 (q, $^2J_{C,F}$ = 43.2 Hz, $\underline{C}$CF$_3$), 131.5 (Ar1 C-4), 130.0 (Ar2 C-1), 129.5 (Ar2 C-3,5), 127.9 (Ar2 C-4), 121.9 (Ar1 C-2,6), 116.8 (q, $^1J_{C,F}$ = 277.1 Hz, CF$_3$).

¹⁹F NMR (376 MHz, CDCl$_3$) δ: -71.5 (s, CF$_3$).

⁷⁷Se NMR (76 MHz, CDCl$_3$) δ: 419.8 (s, Se).

HRMS (ESI), *m/z*: calcd. for C$_{14}$H$_{10}$ClF$_3$NSe$^+$: 363.9614 [M+H]$^+$; found: 363.9612.

(1Z)-2,2,2-Trifluoro-*N*-[4-(1-pyrrolidinylsulfonyl)phenyl]ethanimidoyl chloride

By following the General Procedure **1**, starting from 4-(pyrrolidin-1-ylsulfonyl)aniline (1000 mg, 4.4 mmol, 1.0 equiv), TFA (504 mg, 0.3 mL, 4.4 mmol, 1.0 equiv), Ph$_3$P (3480 mg, 13.3 mmol, 3.0 equiv), Et$_3$N (537 mg, 0.74 mL, 5.3 mmol, 1.2 equiv), CCl$_4$ (2718 mg, 1.70 mL, 17.7 mmol, 4.0 equiv) and DCE (20 mL), the desired product was obtained in 72% yield (1079 mg) as a yellow oil after chromatography on silica gel (60:40 *v/v, n*-hexane/dichloromethane).

¹H NMR (400 MHz, CDCl$_3$) δ: 7.90 (m, 2H, Ph H-3,5), 7.15 (m, 2H, Ph H-2,6), 3.26 (m, 4H, NC$\underline{H}_2$CH$_2$), 1.78 (m, 4H, NCH$_2$C$\underline{H}_2$).

¹³C NMR (100 MHz, CDCl$_3$) δ: 147.1 (Ph C-1), 135.6 (Ph C-4), 134.8 (q, $^2J_{C,F}$ = 43.6 Hz, $\underline{C}$CF$_3$), 128.7 (Ph C-3,5), 120.4 (Ph C-2,6), 116.5 (q, $^1J_{C,F}$ = 277.6 Hz, CF$_3$), 47.9 (N$\underline{C}$H$_2$CH$_2$), 25.2 (NCH$_2$$\underline{C}H_2$).

19F NMR (376 MHz, CDCl$_3$) δ: -71.7 (s, CF$_3$).

15N NMR (40 MHz, CDCl$_3$) δ: -281.6 (NCH$_2$CH$_2$), -53.1 (N=C).

HRMS (ESI), *m/z*: calcd. for C$_{12}$H$_{13}$ClF$_3$N$_2$O$_2$S$^+$: 341.0333 [M+H]$^+$; found: 341.0335.

N,N´-[(2Z,4Z)-1,1,1,5,5,5-Hexafloruro-2-penten-4-yl-4-ylidene]bis(3-iodoaniline) (2)

By following the General Procedure **2**, the homologation event was conducted starting from (1*Z*)-*N*-(3-iodophenyl)-2,2,2-trifluoroethanimidoyl chloride (570 mg, 1.76 mmol, 1.0 equiv), CH$_2$I$_2$ (1400 mg, 0.43 mL, 5.28 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 1.53 mL, 2.29 mmol, 1.3 equiv) and THF (10 mL). Then, the second aliquots of (1*Z*)-*N*-(3-iodophenyl)-2,2,2-trifluoroethanimidoyl chloride (570 mg, 1.76 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 1.76 mL, 2.64 mmol, 1.5 equiv) were added. The desired product was obtained in 91% yield (977 mg) as yellow oil after chromatography on silica gel (95:5 *v/v*, *n*-hexane/chloroform).

1H NMR (400 MHz, CDCl$_3$) δ: 11.83 (s, 1H, NH), 7.55 (m, 2H, Ph H-4), 7.42 (m, 2H, Ph H-2), 7.07 (m, 2H, Ph H-5), 7.02 (m, 2H, Ph H-6), 5.90 (s, 1H, CH).

13C NMR (100 MHz, CDCl$_3$) δ: 149.3 (q, $^2J_{C,F}$ = 30.3 Hz, <u>C</u>CF$_3$), 143.5 (Ph C-1), 134.9 (Ph C-4), 131.5 (Ph C-2), 130.2 (Ph C-5), 122.1 (Ph C-6), 118.9 (q, $^1J_{C,F}$ = 283.2 Hz, CF$_3$), 93.7 (Ph C-3), 89.3 (CH).

19F NMR (376 MHz, CDCl$_3$) δ: -62.8 (s, CF$_3$).

15N NMR (40 MHz, CDCl$_3$) δ: -178.0 (N).

HRMS (ESI), *m/z*: calcd. for C$_{17}$H$_{11}$F$_6$I$_2$N$_2$$^+$: 610.8910 [M+H]$^+$; found: 610.8906.
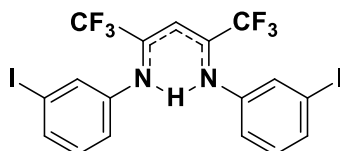
N,N'-[(2Z,4Z)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]dianiline (3)
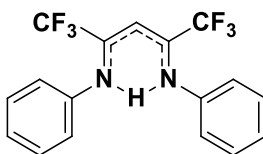
By following the General Procedure **2**, the homologation event was conducted starting from *N*-phenyl-2,2,2-trifluoroethanimidoyl chloride (150 mg, 0.70 mmol, 1.0 equiv), CH$_2$I$_2$ (578 mg, 0.2 mL, 2.2 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.61 mL, 0.91 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of *N*-phenyl-2,2,2-trifluoroethanimidoyl chloride (150 mg, 0.70 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 0.70 mL, 1.05 mmol, 1.5 equiv) were added. The desired product was obtained in 85% yield (213 mg) as a yellow oil after chromatography on silica gel (90:10 *v/v*, *n*-hexane/chloroform).

¹H NMR (400 MHz, CDCl$_3$) δ: 12.08 (s, 1H, NH), 7.37 (m, 4H, Ph H-3,5), 7.22 (m, 2H, Ph H-4), 7.08 (m, 4H, Ph H-2,6), 5.92 (s, CH).

¹³C NMR (100 MHz, CDCl$_3$) δ: 149.1 (q, $^2J_{C,F}$ = 29.8 Hz, <u>C</u>CF$_3$), 142.6 (Ph C-1), 128.8 (Ph C-3,5), 125.7 (Ph C-4), 122.7 (Ph C-2,6), 119.1 (q, $^1J_{C,F}$ = 283.6 Hz, CF$_3$), 88.8 (sept, $^3J_{C,F}$ = 4.8 Hz, CH).

¹⁹F NMR (376 MHz, CDCl$_3$) δ: -62.8 (s, CF$_3$).

¹⁵N NMR (40 MHz, CDCl$_3$) δ: -175.8 (N).

HRMS (ESI), *m/z*: calcd. for C$_{17}$H$_{13}$F$_6$N$_2$$^+$: 359.0977 [M+H]$^+$; found: 359.0975.

***N,N´*-[(2Z,4Z)-1,1,1,5,5,5-Hexafloruro-2-penten-1-yl-4-ylidene]bis(3-bromoaniline) (4)**

By following the General Procedure **2**, the homologation event was conducted starting from (1*Z*)-*N*-(3-bromophenyl)-2,2,2-trifluoroethanimidoyl chloride (300 mg, 1.05 mmol, 1.0 equiv), CH_2I_2 (845 mg, 0.27 mL, 3.50 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.91 mL, 1.37 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1*Z*)-*N*-(3-bromophenyl)-2,2,2-trifluoroethanimidoyl chloride (300 mg, 1.05 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 1.05 mL, 1.58 mmol, 1.5 equiv) were added. The desired product was obtained in 87% yield (472 mg) as a colourless oil after chromatography on silica gel (90:10 *v/v*, *n*-hexane/diethyl ether).

¹H NMR (400 MHz, C_6D_6) δ: 11.67 (s, 1H, NH), 7.03 (m, 4H, Ph H), 6.60 (m, 4H, Ph H), 5.83 (s, 1H, CH).

¹³C NMR (100 MHz, $CDCl_3$) δ: 149.5 (q, $^2J_{C,F}$ = 30.3 Hz, $\underline{C}CF_3$), 143.6 (Ph C-1), 130.1 (Ph C-5), 129.0 (Ph C-4), 125.8 (Ph C-2), 122.4 (Ph C-3), 121.4 (Ph C-6), 118.8 (q, $^1J_{C,F}$ = 283.3 Hz, CF_3), 89.3 (sept, $^3J_{C,F}$ = 4.7 Hz, CH).

¹⁹F NMR (376 MHz, $CDCl_3$) δ: -62.8 (s, CF_3).

¹⁵N NMR (40 MHz, $CDCl_3$) δ: -178.0 (N).

HRMS (ESI), *m/z*: calcd. for $C_{17}H_{11}Br_2F_6N_2^+$: 514.9188 [M+H]⁺; found: 514.9187.

***N,N´*-[(2Z,4Z)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis(4-bromoaniline) (5)**

By following the General Procedure **2**, the homologation event was conducted starting from (1*Z*)-*N*-(4-bromophenyl)-2,2,2-trifluoroethanimidoyl chloride (250 mg, 0.87 mmol, 1.0 equiv), CH_2I_2 (701 mg, 0.21 mL, 2.61 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.75 mL, 1.13 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1*Z*)-*N*-(4-bromophenyl)-2,2,2-trifluoroethanimidoyl chloride (250 mg, 0.87 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 0.87 mL, 1.31 mmol, 1.5 equiv) were added. The desired product was obtained in 88% yield (325 mg) as a yellow oil, after chromatography on silica gel (90:10 v/v *n*-hexane/dichloromethane).

^{1}H NMR (400 MHz, CDCl$_3$) δ: 11.87 (s, 1H, NH), 7.46 (m, 4H, Ph H-3,5), 6.93 (m, 4H, Ph H-2,6), 5.91 (s, 1H, CH).

^{13}C NMR (100 MHz, CDCl$_3$) δ: 149.4 (q, $^2J_{C,F}$ = 30.0 Hz, $\underline{C}$CF$_3$), 141.5 (Ph C-1), 131.9 (Ph C-3,5), 124.4 (q, $J_{C,F}$ = 1.7 Hz, Ph C-2,6), 119.3 (Ph C-4), 118.9 (q, $^1J_{C,F}$ = 283.5 Hz, CF$_3$), 89.3 (sept, $^3J_{C,F}$ = 4.8 Hz, CH).

^{19}F NMR (376 MHz, CDCl$_3$) δ: -62.8 (s, CF$_3$).

^{15}N NMR (40 MHz, CDCl$_3$) δ: -178.2 (N).

HRMS (ESI), *m/z*: calcd. for C$_{17}$H$_{10}$Br$_2$F$_6$N$_2$Na$^+$: 536.9007 [M+Na]$^+$; found: 536.9009.

***N,N´*-[(2Z, 4Z)-1,1,1,5,5,5-hexafluoro-2-penten-2-yl-4-ylidene]bis(2-chloroaniline) (6)**

By following the General Procedure **2**, the homologation event was conducted starting from (1*Z*)-*N*-(2-chlorophenyl)-2,2,2-trifluoroethanimidoyl chloride (250 mg, 1.03 mmol, 1.0 equiv), CH_2I_2 (830 mg, 0.25 mL, 3.10 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.89 mL, 1.34 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1*Z*)-*N*-(2-chlorophenyl)-2,2,2-trifluoroethanimidoyl chloride (250 mg, 1.03 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 1.03 mL, 1.55 mmol, 1.5 equiv) were added. The desired product was obtained in 79% yield (348 mg) as a yellow oil, after chromatography on silica gel (90:10 v/v, *n*-hexane/diethyl ether).

¹H NMR (400 MHz, CDCl₃) δ: 11.94 (s, 1H, NH), 7.42 (m, 2H, Ph H-3), 7.25 (m, 2H, Ph H-5), 7.15 (m, 2H, Ph H-4), 7.13 (m, 2H, Ph H-6), 5.98 (s, 1H, CH).
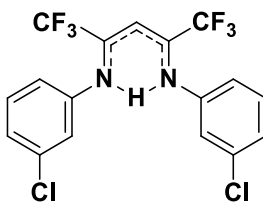
¹³C NMR (100 MHz, CDCl₃) δ: 149.9 (q, $^2J_{C,F}$ = 30.4 Hz, $\underline{C}CF_3$), 140.0 (Ph C-1), 129.6 (Ph C-3), 128.0 (Ph C-2), 127.01 (Ph C-5), 126.98 (Ph C-4), 124.2 (q, $J_{C,F}$ = 2.6 Hz, Ph C-6), 118.8 (q, $^1J_{C,F}$ = 283.5 Hz, CF₃), 89.8 (sept., $^3J_{C,F}$ = 4.8 Hz, CH).

¹⁹F NMR (376 MHz, CDCl₃) δ: -64.3 (s, CF₃).

¹⁵N NMR (40 MHz, CDCl₃) δ: -179.7 (N).

HRMS (ESI), *m/z*: calcd. for $C_{17}H_{10}Cl_2F_6N_2Na^+$: 449.0017 [M+Na]⁺; found: 449.0006.

***N,N'*-[(2*Z*,4*Z*)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis(3-chloroaniline) (7)**

By following the General Procedure **2**, the homologation event was conducted starting from (1*Z*)-*N*-(3-chlorophenyl)-2,2,2-trifluoroethanimidoyl chloride (250 mg, 1.03 mmol, 1.0 equiv), CH_2I_2 (830 mg, 0.25 mL, 3.10 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.89 mL, 1.34 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1*Z*)-*N*-(3-chlorophenyl)-2,2,2-trifluoroethanimidoyl chloride (250 mg, 1.03 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 1.03 mL, 1.55 mmol, 1.5 equiv) were

added. The desired product was obtained in 83% yield (365 mg) as a yellow solid (m.p.: 79-81 °C) after chromatography on silica gel (90:10 v/v, n-hexane/diethyl ether).

¹H NMR (400 MHz, CDCl₃) δ: 11.86 (s, 1H, NH), 7.28 (m, 2H, Ph H-5), 7.20 (m, 2H, Ph H-4), 7.07 (m, 2H, Ph H-2), 6.94 (m, 2H, Ph H-6), 5.92 (s, 1H, CH).
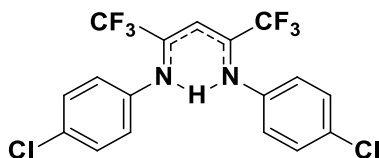
¹³C NMR (100 MHz, CDCl₃) δ: 149.5 (q, $^2J_{C,F}$ = 30.1 Hz, $\underline{C}$CF₃), 143.5 (Ph C-1), 134.5 (Ph C-3), 129.9 (Ph C-5), 126.0 (Ph C-4), 122.9 (q, $J_{C,F}$ = 1.8 Hz, Ph C-2), 120.9 (q, $J_{C,F}$ = 2.0 Hz, Ph C-6), 118.9 (q, $^1J_{C,F}$ = 283.4 Hz, CF₃), 89.4 (m, CH).

¹⁹F NMR (376 MHz, CDCl₃) δ: -62.9 (s, CF₃).

¹⁵N NMR (40 MHz, CDCl₃) δ: -177.9 (N).

HRMS (ESI), m/z: calcd. for $C_{17}H_{11}Cl_2F_6N_2^+$: 427.0198 [M+H]⁺; found: 427.0192.

_N,N'_-[(2_Z_,4_Z_)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis(4-chloroaniline) (8)

By following the General Procedure **2**, the homologation event was conducted starting from (1Z)-N-(4-chlorophenyl)-2,2,2-trifluoroethanimidoyl chloride (300 mg, 1.2 mmol, 1.0 equiv), CH₂I₂ (821 mg, 0.3 mL, 3.7 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 1.04 mL, 1.56 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1Z)-N-(4-chlorophenyl)-2,2,2-trifluoroethanimidoyl chloride (300 mg, 1.2 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 1.20 mL, 1.8 mmol, 1.5 equiv) were added. The desired product was obtained in 81% yield (415 mg) as a yellow oil after chromatography on silica gel (90:10 v/v, n-hexane/chloroform).

¹H NMR (400 MHz, CDCl₃) δ: 11.89 (s, 1H, NH), 7.31 (m, 4H, Ph H-3,5), 6.98 (m, 4H, Ph H-2,6), 5.90 (s, 1H, CH).
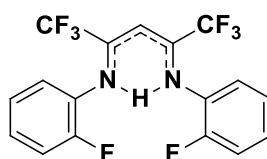
¹³C NMR (100 MHz, CDCl₃) δ: 149.5 (q, $^2J_{C,F}$ = 30.0 Hz, $\underline{C}$CF₃), 141.0 (Ph C-1), 131.6 (Ph C-4), 129.0 (Ph C-3,5), 124.1 (Ph C-2,6), 118.9 (q, $^1J_{C,F}$ = 283.6 Hz, CF₃), 89.2 (m, CH).

^{19}F NMR (376 MHz, CDCl$_3$) δ: -62.9 (s, CF$_3$).

^{15}N NMR (40 MHz, CDCl$_3$) δ: -178.1 (N).

HRMS (ESI), *m/z*: calcd. for C$_{17}$H$_{11}$Cl$_2$F$_6$N$_2$$^+$: 427.0198 [M+H]$^+$; found: 427.0199.

N,N´-**[(2Z,4Z)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis(2-fluoroaniline) (9)**

By following the General Procedure **2**, the homologation event was conducted starting from (1*Z)*-2,2,2-trifluoro-*N*-(2-fluorophenyl)ethanmidoyl chloride (250 mg, 1.10 mmol, 1.0 equiv), CH$_2$I$_2$ (890 mg, 0.27 mL, 3.32 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.95 mL, 1.43 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1*Z)*-2,2,2-trifluoro-*N*-(2-fluorophenyl)ethanmidoyl chloride (250 mg, 1.10 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 1.10 mL, 1.65 mmol, 1.5 equiv) were added. The desired product was obtained in 78% yield (338 mg) as a yellow solid (m.p.: 56 °C) after chromatography on silica gel (*n*-hexane).

^{1}H NMR (400 MHz, CDCl$_3$) δ: 11.85 (s, 1H, NH), 7.19 (m, 2H, Ph H-4), 7.13 (m, 4H, Ph H-5,6), 7.12 (m, 2H, Ph H-3), 5.99 (s, 1H, CH).

^{13}C NMR (100 MHz, CDCl$_3$) δ: 154.7 (d, $^{1}J_{C,F}$ = 247.6 Hz, Ph C-2), 150.6 (q, $^{2}J_{C,F}$ = 30.3 Hz, <u>C</u>CF$_3$), 130.4 (d, $^{2}J_{C,F}$ = 13.0 Hz, Ph C-1), 127.3 (d, $^{3}J_{C,F}$ = 7.6 Hz, Ph C-4), 125.0 (q, $^{3}J_{C,F}$ = 2.2 Hz, Ph C-6), 124.0 (d, $^{4}J_{C,F}$ = 3.9 Hz, Ph C-5), 118.9 (q, $^{1}J_{C,F}$ = 283.2 Hz, CF$_3$), 115.8 (d, $^{2}J_{C,F}$ = 20.0 Hz, Ph C-3), 89.8 (sept, $^{3}J_{C,F}$ = 4.7 Hz, CH).

^{19}F NMR (376 MHz, CDCl$_3$) δ: -123.7 (m, Ph F-2), -65.0 (d, *J* = 3.0 Hz, CF$_3$).

^{15}N NMR (40 MHz, CDCl$_3$) δ: -189.4 (N).

HRMS (ESI), *m/z*: calcd. for C$_{17}$H$_{10}$F$_8$N$_2$Na$^+$: 417.0608 [M+Na]$^+$; found: 417.0612.

N,N´-**[(2Z,4Z)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis(3,4-**

difluoroaniline) (10)
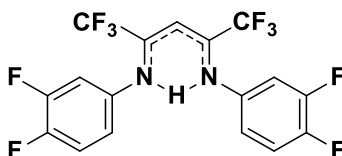
By following the General Procedure **2**, the homologation event was conducted starting from (1Z)-*N*-(3,4-difluorophenyl)-2,2,2-trifluoroethanimidoyl chloride (250 mg, 1.03 mmol, 1.0 equiv), CH_2I_2 (825 mg, 0.25 mL, 3.07 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.89 mL, 1.34 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1Z)-*N*-(3,4-difluorophenyl)-2,2,2-trifluoroethanimidoyl chloride (250 mg, 1.03 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 1.03 mL, 1.55 mmol, 1.5 equiv) were added. The desired product was obtained in 93% yield (412 mg) as a yellow solid (m.p.: 42 °C), after chromatography on silica gel (*n*-hexane).

^{1}H NMR (400 MHz, CDCl$_3$) δ: 11.76 (s, 1H, NH), 7.14 (m, 2H, Ph H-5), 6.91 (m, 2H, Ph H-2), 6.78 (m, 2H, Ph H-6), 5.91 (s, 1H, CH).

^{13}C NMR (100 MHz, CDCl$_3$) δ: 150.0 (q, $^2J_{C,F}$ = 29.5 Hz, $\underline{C}$CF$_3$), 150.0 (dd, $^1J_{C,F}$ = 250.3, $^2J_{C,F}$ = 13.7 Hz, Ph C-3), 148.6 (dd, $^1J_{C,F}$ = 248.0, $^2J_{C,F}$ = 12.6 Hz, Ph C-4), 138.6 (dd, $^3J_{C,F}$ = 7.6, $^4J_{C,F}$ = 3.6 Hz, Ph C-1), 119.1 (m, Ph C-6), 118.8 (q, $^1J_{C,F}$ = 283.5 Hz, CF$_3$), 117.3 (dd, $^2J_{C,F}$ = 18.3, $^3J_{C,F}$ = 1.3 Hz, Ph C-5), 112.6 (m, Ph C-2), 89.2 (sept, $^3J_{C,F}$ = 4.7 Hz, CH).

^{19}F NMR (376 MHz, CDCl$_3$) δ: -140.5 (m, Ph F), -135.8 (m, Ph F), -63.0 (s, CF$_3$).

^{15}N NMR (40 MHz, CDCl$_3$) δ: -179.7 (N).

HRMS (ESI), *m/z*: calcd. for $C_{17}H_9F_{10}N_2$: 431.0601 [M+H]$^+$; found: 431.0603.

N,N´-[(2Z,4Z)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis[2-(trifluoromethyl)aniline (11)**
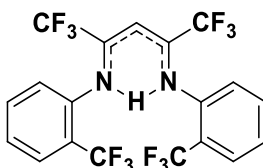
By following the General Procedure **2**, the homologation event was conducted starting from (1*Z*)-2,2,2-trifluoro-*N*-[2-(trifluoromethyl)phenyl]ethanimidoyl chloride (250 mg, 0.91 mmol, 1.0 equiv), CH$_2$I$_2$ (729 mg, 0.22 mL, 2.72 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.79 mL, 1.18 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1*Z*)-2,2,2-trifluoro-*N*-[2-(trifluoromethyl)phenyl]ethanimidoyl chloride (250 mg, 0.91 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 0.91 mL, 1.37 mmol, 1.5 equiv) were added. The desired product was obtained in 80% yield (360 mg) as a yellow solid (m.p.: 50 °C) after chromatography on silica gel (*n*-hexane).

^{1}H NMR (400 MHz, CDCl$_3$) δ: 11.80 (s, 1H, NH), 7.64 (m, 2H, Ph H-3), 7.53 (m, 2H, Ph H-5), 7.31 (m, 2H, Ph H-4), 7.15 (m, 2H, Ph H-6), 5.97 (s, 1H, CH).

^{13}C NMR (100 MHz, CDCl$_3$) δ: 149.7 (q, $^2J_{C,F}$ = 30.6 Hz, $\underline{C}$CF$_3$), 140.5 (Ph C-1), 132.4 (Ph C-5), 126.3 (q, $^3J_{C,F}$ = 4.8 Hz, Ph C-3), 125.9 (Ph C-4), 124.5 (q, $^4J_{C,F}$ = 2.6 Hz, Ph C-6), 123.5 (q, $^1J_{C,F}$ = 273.1 Hz, Ph-$\underline{C}$F$_3$), 123.3 (q, $^2J_{C,F}$ = 31.0 Hz, Ph C-2), 118.9 (q, $^1J_{C,F}$ = 283.5 Hz, CF$_3$), 90.8 (sept, $^3J_{C,F}$ = 4.4 Hz, CH).

^{19}F NMR (376 MHz, CDCl$_3$) δ: -63.4 (s, CF$_3$), -61.4 (s, Ph-CF$_3$)

^{15}N NMR (40 MHz, CDCl$_3$) δ: -184.2 (N).

HRMS (ESI), *m/z*: calcd. for C$_{19}$H$_{11}$F$_{12}$N$_2^+$: 495.0725 [M+H]$^+$; found: 495.0735.

4-({(2*Z*,3*Z*)-4-[(4-Cyanophenyl)amino]-1,1,1,5,5,5-hexafluoro-2-penten-2-ylidene}amino) benzonitrile (12)
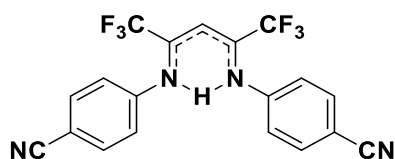
By following the General Procedure **2**, the homologation event was conducted starting from (1*Z*)-*N*-(4-cyanophenyl)-2,2,2-trifluoroethanimidoyl chloride (250 mg, 1.10 mmol, 1.0 equiv), CH$_2$I$_2$ (857 mg, 0.26 mL, 3.20 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.95 mL, 1.43 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1*Z*)-*N*-(4-cyanophenyl)-2,2,2-trifluoroethanimidoyl chloride (250 mg, 1.10 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 1.10 mL, 1.65 mmol, 1.5 equiv) were

added. The desired product was obtained in 85% yield (382 mg) as a colorless oil after chromatography on silica gel (90:10 v/v *n*-hexane/diethyl ether).

^{1}H NMR (400 MHz, CDCl$_3$) δ: 11.80 (s, 1H, NH), 7.65 (m, 4H, Ph H-3,5), 7.13 (m, 4H, Ph H-2,6), 6.03 (s, 1H, CH).

^{13}C NMR (100 MHz, CDCl$_3$) δ: 149.2 (q, $^2J_{C,F}$ = 30.9 Hz, $\underline{C}$CF$_3$), 146.1 (Ph C-1), 133.1 (Ph C-3,5), 122.7 (q, $J_{C,F}$ = 1.9 Hz, Ph C-2,6), 118.7 (q, $^1J_{C,F}$ = 283.3 Hz, CF$_3$), 118.3 (C≡N), 109.5 (Ph C-4), 91.0 (sept, $^3J_{C,F}$ = 4.9 Hz, CH).

^{19}F NMR (376 MHz, CDCl$_3$) δ: -62.8 (s, CF$_3$).

^{15}N NMR (40 MHz, CDCl$_3$) δ: -176.5 (N).

HRMS (ESI), *m/z*: calcd. for C$_{19}$H$_{10}$F$_6$N$_4$Na$^+$: 431.0696 [M+Na]$^+$; found: 431.0708.

tert-Butyl4-{[(2Z,3Z)-4-{[4-(_tert_-butoxycarbonyl)phenyl]amino}-1,1,1,5,5,5-hexafluoropent-3-en-2-ylidene] amino} benzoate (13)
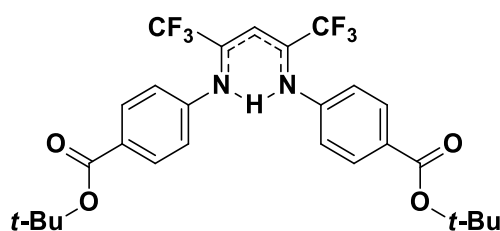
By following the General Procedure **2**, the homologation event was conducted starting from *tert*-butyl 4-[(*Z*)-(1-chloro-2,2,2-trifluoroethylidene)amino]benzoate (250 mg, 0.81 mmol, 1.0 equiv), CH$_2$I$_2$ (651 mg, 0.20 mL, 2.43 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.70 mL, 1.05 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of *tert*-butyl 4-[(*Z*)-(1-chloro-2,2,2-trifluoroethylidene)amino]benzoate (250 mg, 0.81 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 0.81 mL, 1.22 mmol, 1.5 equiv) were added. The desired product was obtained in 83% yield (375 mg) as a yellow solid (m.p.: 48 °C) after chromatography on silica gel (70:30 v/v *n*-hexane/dichloromethane).

^{1}H NMR (400 MHz, CDCl$_3$) δ: 11.90 (brs, 1H, NH), 7.97 (m, 4H, Ph H-3,5), 7.07 (m, 4H, Ph H-2,6), 5.96 (s, 1H, CH), 1.59 (s, 18H, CCH$_3$).

^{13}**C NMR** (100 MHz, CDCl$_3$) δ: 165.0 (C=O), 148.9 (q, $^2J_{C,F}$ = 29.9 Hz, $\underline{C}$CF$_3$), 146.1 (Ph C-1), 130.4 (Ph C-3,5), 129.3 (Ph C-4), 121.8 (Ph C-2,6), 118.9 (q, $^1J_{C,F}$ = 283.2 Hz, CF$_3$), 90.1 (m, CH), 81.2 ($\underline{C}$CH$_3$), 28.2 (C$\underline{C}$H$_3$).

^{19}**F NMR** (376 MHz, CDCl$_3$) δ: -62.8 (s, CF$_3$).

HRMS (ESI), *m/z*: calcd. For C$_{27}$H$_{28}$F$_6$N$_2$O$_4$Na$^+$: 581.1845 [M+Na]$^+$; found: 581.1843.

Ethyl4-{[(2Z,3Z)-4-{[4-(ethoxycarbonyl)phenyl]amino}-1,1,1,5,5,5-hexafluoro-3-penten-2-ylidene]amino} benzoate (14)
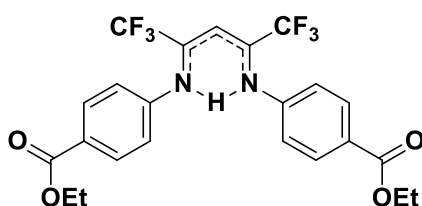
By following the General Procedure **2**, the homologation event was conducted starting from ethyl 4-[(Z)-(1-chloro-2,2,2-trifluoroethylidene)amino]benzoate (250 mg, 0.89 mmol, 1.0 equiv), CH$_2$I$_2$ (718 mg, 0.22 mL, 2.68 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.77 mL, 1.16 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of ethyl 4-[(Z)-(1-chloro-2,2,2-trifluoroethylidene)amino]benzoate (250 mg, 0.89 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 0.89 mL, 1.34 mmol, 1.5 equiv) were added. The desired product was obtained in 86% yield (385 g) as a colorless oil after chromatography on silica gel (90:10 v/v *n*-hexane/ethyl acetate).

^{1}**H NMR** (400 MHz, CDCl$_3$) δ: 11.94 (s, 1H, NH), 8.03 (m, 4H, Ph H-2,6), 7.09 (m, 4H, Ph H-3,5), 5.97 (s, 1H, CH), 4.38 (q, ^3J = 7.1 Hz, 4H, OC$\underline{H}_2$CH$_3$), 1.39 (t, ^3J = 7.1 Hz, 6H, OCH$_2$C$\underline{H}_3$).

^{13}**C NMR** (100 MHz, CDCl$_3$) δ: 165.9 (C=O), 149.0 (q, $^1J_{C,F}$ = 30.7 Hz, $\underline{C}$CF$_3$), 146.4 (Ph C-4), 130.5 (Ph C-2,6), 127.8 (Ph C-1), 121.9 (q, $J_{C,F}$ = 1.9 Hz, Ph C-3,5), 118.9 (q, $^1J_{C,F}$ = 283.5 Hz, CF$_3$), 90.2 (m, CH), 61.0 (O$\underline{C}$H$_2$CH$_3$), 14.3 (OCH$_2$$\underline{C}H_3$).

^{19}**F NMR** (376 MHz, CDCl$_3$) δ: -62.8 (s, CF$_3$).

^{15}**N NMR** (40 MHz, CDCl$_3$) δ: -175.6 (N).

HRMS (ESI), *m/z*: calcd. for C$_{23}$H$_{20}$F$_6$N$_2$O$_4$Na$^+$: 525.1219 [M+Na]$^+$; found: 525.1234.

***N,N´*-[(2Z,4Z)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]di(1-naphtalenamine) (15)**

By following the General Procedure **2**, the homologation event was conducted starting from (1*Z*)-2,2,2-trifluoro-*N*-(naphtalen-1-yl)ethanimidoyl chloride (250 mg, 0.97 mmol, 1.0 equiv), CH_2I_2 (779 mg, 0.23 mL, 2.91 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.84 mL, 1.26 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1*Z*)-2,2,2-trifluoro-*N*-(naphtalen-1-yl)ethanimidoyl chloride (250 mg, 0.97 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 0.97 mL, 1.46 mmol, 1.5 equiv) were added. The desired product was obtained in 91% yield (405 mg) as an orange solid (m.p.: 128 °C) after chromatography on silica gel (90:10 v/v *n*-hexane/diethyl ether).

Scaling-up of the reaction (20 mmol) - By following the General Procedure **2**, the homologation event was conducted starting from (1*Z*)-2,2,2-trifluoro-*N*-(naphtalen-1-yl)ethanimidoyl chloride (2576 mg, 10.0 mmol, 1.0 equiv), CH_2I_2 (8035 mg, 2.41 mL, 30.0 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 8.7 mL, 13.0 mmol, 1.3 equiv) and THF (25 mL). Then, the second aliquots of (1*Z*)-2,2,2-trifluoro-*N*-(naphtalen-1-yl)ethanimidoyl chloride (2576 mg, 10.0 mmol, 1.0 equiv) in THF (10 mL) and MeLi-LiBr complex (1.5 M, 10.0 mL, 15.0 mmol, 1.5 equiv) were added. The desired product was obtained in 91% yield (405 mg) as an orange solid (m.p.: 128 °C) after chromatography on silica gel (90:10 v/v *n*-hexane/diethyl ether). *Spectroscopic and spectrometric data match with those reported for the 1.51 mmol scale reaction.*

¹H NMR (400 MHz, CDCl₃) δ: 12.56 (s, 1H, NH), 7.98 (m, 2H, Naph H-8), 7.85 (m, 2H, Naph H-5), 7.73 (m, 2H, Naph H-4), 7.55-7.48 (m, 4H, Naph H-6,7), 7.45 (m, 2H, Naph H-3), 7.23 (m, 2H, Naph H-2), 6.15 (s, 1H, CH).

¹³C NMR (100 MHz, CDCl₃) δ: 150.6 (q, $^2J_{C,F}$ = 29.9 Hz, $\underline{C}CF_3$), 138.9 (Naph C-1), 133.9 (Naph C-4a), 128.2 (Naph C-8a), 128.1 (Naph C-5), 126.5 (Naph C-4,6), 126.4
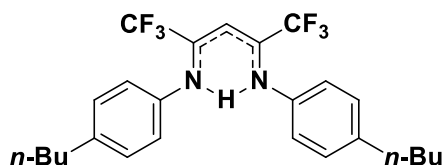
(Naph C-7), 125.3 (Naph C-3), 122.8 (Naph C-8),119.3 (Naph C-2), 119.1 (q, $^1J_{C,F}$ = 283.6 Hz, CF$_3$), 89.2 (sept., $^2J_{C,F}$ = 4.7 Hz, CH).

^{19}F NMR (376 MHz, CDCl$_3$) δ: -63.5 (s, CF$_3$).

^{15}N NMR (40 MHz, CDCl$_3$) δ: -180.6 (N).

HRMS (ESI), *m/z*: calcd. for C$_{25}$H$_{16}$F$_6$N$_2$Na$^+$: 481.1104 [M+Na]$^+$ found: 481.1118.

N,N´-[(2Z,4Z)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis(4-butylaniline) (16)

By following the General Procedure **2**, the homologation event was conducted starting from (1*Z*)-*N*-(4-butylphenyl)-2,2,2-trifluoroethanimidoyl chloride (250 mg, 0.95 mmol, 1.0 equiv), CH$_2$I$_2$ (761 mg, 0.29 mL, 2.8 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.82 mL, 1.24 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1*Z*)-*N*-(4-butylphenyl)-2,2,2-trifluoroethanimidoyl chloride (250 mg, 0.95 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 0.95 mL, 1.43 mmol, 1.5 equiv) were added. The desired product was obtained in 83% yield (371 mg) as a colorless oil after chromatography on reverse phase silica gel (acetonitrile).

^{1}H NMR (400 MHz, CDCl$_3$) δ: 12.00 (brs, 1H, NH), 7.13 (m, 4H, Ph H-3,5), 6.94 (m, 4H, Ph H-2,6), 5.84 (s, 1H, CH), 2.59 (m, 4H, C$\underline{H}_2$CH$_2$CH$_2$CH$_3$), 1.58 (m, 4H, CH$_2$C$\underline{H}_2$CH$_2$CH$_3$), 1.34 (m, 4H, CH$_2$CH$_2$C$\underline{H}_2$CH$_3$), 0.92 (t, ^3J = 7.4 Hz, 6H, CH$_2$CH$_2$CH$_2$C$\underline{H}_3$).

^{13}C NMR (100 MHz, CDCl$_3$) δ: 140.5 (Ph C-4), 140.2 (Ph C-1), 128.6 (Ph C-3,5), 122.6 (q, $J_{C,F}$ = 1.7 Hz, Ph C-2,6), 119.1 (q, $^1J_{C,F}$ = 283.6 Hz, CF$_3$), 88.5 (m, CH), 35.1 ($\underline{C}$H$_2$CH$_2$CH$_2$CH$_3$), 33.6 (CH$_2$$\underline{C}H_2CH_2CH_3$), 22.3 (CH$_2CH_2$$\underline{C}H_2CH_3$), 13.9 (CH$_2CH_2CH_2$$\underline{C}H_3$).

^{19}F NMR (376 MHz, CDCl$_3$) δ: -62.9 (s, CF$_3$).

^{15}N NMR (40 MHz, CDCl$_3$) δ: -176.2 (N).

HRMS (ESI), *m/z*: calcd for $C_{25}H_{29}F_6N_2$: 471.2229 [M+H]$^+$; found: 471.2227.

N,N´-[(2*Z*,4*Z*)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis(2,5-dimethylaniline) (17)
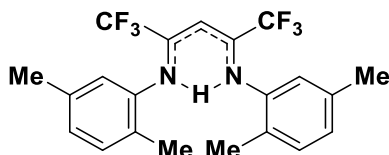
By following the General Procedure **2**, the homologation event was conducted starting from (1*Z*)-*N*-(2,5-dimethylphenyl)-2,2,2-trifluoroethanimidoyl chloride (235 mg, 1.0 mmol, 1.0 equiv), CH_2I_2 (804 mg, 0.24 mL, 3.0 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.87 mL, 1.3 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1*Z*)-*N*-(2,5-dimethylphenyl)-2,2,2-trifluoroethanimidoyl chloride (235 mg, 1.0 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 1.0 mL, 1.5 mmol, 1.5 equiv) were added. The desired product was obtained in 88% yield (367 g) as a yellow oil after chromatography on silica gel (*n*-hexane).

^{1}H NMR (400 MHz, CDCl$_3$) δ: 11.91 (s, 1H, NH), 7.07 (d, ^3J = 7.7 Hz, 2H, Ph H-3), 6.92 (d, ^3J = 7.7 Hz, 2H, Ph H-4), 6.78 (s, 2H, Ph H-6), 5.87 (s, 1H, CH), 2.30 (s, 6H, 5-CH$_3$), 2.14 (s, 6H, 2-CH$_3$).
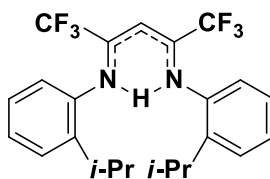
^{13}C NMR (100 MHz, CDCl$_3$) δ: 149.6 (q, $^2J_{C,F}$ = 29.4 Hz, <u>C</u>CF$_3$), 141.4 (Ph C-1), 135.9 (Ph C-5), 130.0 (Ph C-3), 127.9 (Ph C-2), 126.7 (Ph C-4), 123.6 (Ph C-6), 119.0 (q, $^1J_{C,F}$ = 283.8 Hz, CF$_3$), 87.9 (sept, $^3J_{C,F}$ = 4.5 Hz, CH), 20.9 (5-CH$_3$), 17.6 (2-CH$_3$).

^{19}F NMR (376 MHz, CDCl$_3$) δ: -64.4 (s, CF$_3$).

^{15}N NMR (40 MHz, CDCl$_3$) δ: -176.5 (N).

HRMS (ESI), *m/z*: calcd. for $C_{21}H_{21}F_6N_2^+$: 415.1603 [M+H]$^+$; found: 415.1601.

N,N´-[(2*Z*,4*Z*)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis(2-isopropylaniline) (18)

By following the General Procedure **2**, the homologation event was conducted starting from (1Z)-2,2,2-trifluoro-N-[2-(propan-2yl)phenyl)]ethanimidoyl chloride (250 mg, 1.0 mmol, 1.0 equiv), CH_2I_2 (529 mg, 0.22 mL, 3.0 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.67 mL, 1.0 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1Z)-2,2,2-trifluoro-N-[2-(propan-2yl)phenyl)]ethanimidoyl chloride (250 mg, 1.0 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 1.0 mL, 1.5 mmol, 1.5 equiv) were added. The desired product was obtained in 85% yield (376 mg) as a yellow solid (m.p.: 56-58 °C) after chromatography on reverse phase silica gel (acetonitrile).

¹H NMR (400 MHz, $CDCl_3$) δ: 11.89 (brs, 1H, NH), 7.29 (m, 2H, Ph H-3), 7.21 (m, 2H, Ph H-4), 7.15 (m, 2 H, Ph H-5), 6.94 (m, 2H, Ph H-6), 5.91 (s, 1H, CH), 3.08 (sept, ^3J = 6.9 Hz, 2H, C$\underline{H}$CH₃), 1.17 (d, ^3J = 6.9 Hz, 12H, CHC$\underline{H}_3$).

¹³C NMR (100 MHz, $CDCl_3$) δ: 149.6 (q, $^2J_{C,F}$ = 29.5 Hz, $\underline{C}CF_3$), 141.6 (Ph C-2), 140.3 (Ph C-1), 126.6 (Ph C-4), 125.9 (Ph C-5), 125.7 (Ph C-3), 123.3 (q, $J_{C,F}$ = 2.4 Hz, Ph C-6), 119.1 (q, $^1J_{C,F}$ = 283.8 Hz, CF_3), 88.8 (sept, $^3J_{C,F}$ = 4.8 Hz, CH), 28.4 ($\underline{C}HCH_3$), 23.1 (CH$\underline{C}H_3$).

¹⁹F NMR (376 MHz, $CDCl_3$) δ: -63.4 (s, CF_3).

¹⁵N NMR (40 MHz, $CDCl_3$) δ: -178.7 (N).

HRMS (ESI), m/z: calcd for $C_{23}H_{25}F_6N_2$: 443.1916 [M+H]⁺; found: 443.1914.

N,N´-[(2Z,4Z)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis(4-vinylaniline) (19)
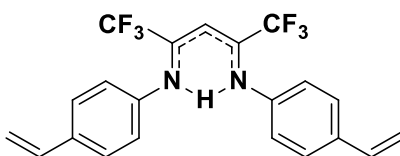
By following the General Procedure **2**, the homologation event was conducted starting

from (1Z)-N-(4-ethenylphenyl)-2,2,2-trifluoroethanimidoyl chloride (150 mg, 0.64 mmol, 1.0 equiv), CH_2I_2 (516 mg, 0.16 mL, 1.92 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.55 mL, 0.83 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1Z)-N-(4-ethenylphenyl)-2,2,2-trifluoroethanimidoyl chloride (150 mg, 0.64 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 0.64 mL, 0.96 mmol, 1.5 equiv) were added. The desired product was obtained in 77% yield (202 g) as a yellow oil after chromatography on silica gel (n-hexane).

^{1}H NMR (400 MHz, CDCl$_3$) δ: 12.06 (s, 1H, NH), 7.38 (m, 4H, Ph H-3,5), 7.01 (m, 4H, Ph H-2,6), 6.69 (dd, $^3J_{trans}$ = 17.6 Hz, $^3J_{cis}$ = 10.9 Hz, 2H, CH=CH$_2$), 5.89 (s, 1H, CH), 5.72 (dd, $^3J_{trans}$ = 17.6, ^2J = 0.8 Hz, 2H, CH=CH$_2$ trans), 5.24 (dd, $^3J_{cis}$ = 10.9, ^2J = 0.8 Hz, 2H, CH=CH$_2$ cis).
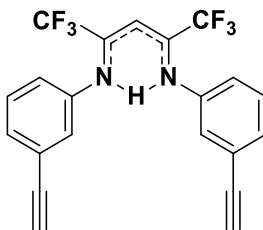
^{13}C NMR (100 MHz, CDCl$_3$) δ: 148.9 (q, $^2J_{C,F}$ = 29.9 Hz, CCF$_3$), 142.1 (Ph C-1), 135.9 (CH=CH$_2$), 135.2 (Ph C-4), 126.6 (Ph C-3,5), 122.7 (q, $J_{C,F}$ = 1.7 Hz, Ph C-2,6), 119.1 (q, $^1J_{C,F}$ = 283.4 Hz, CF$_3$), 113.8 (CH=CH$_2$), 89.1 (sept, $^3J_{C,F}$ = 4.7 Hz, CH).

^{19}F NMR (376 MHz, CDCl$_3$) δ: -62.8 (s, CF$_3$).

^{15}N NMR (40 MHz, CDCl$_3$) δ: -175.9 (N).

HRMS (ESI), m/z: calcd. for $C_{21}H_{16}F_6N_2Na^+$: 433.1110 [M+Na]$^+$; found: 433.1113.

N,N´-[(2Z,4Z)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis(3-ethynylaniline) (20)

By following the General Procedure **2**, the homologation event was conducted starting from (1Z)-N-(3-ethynylphenyl)-2,2,2-trifluoroethanimidoyl chloride (250 mg, 1.08 mmol, 1.0 equiv), CH_2I_2 (867 mg, 0.26 mL, 3.23 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.94 mL, 1.4 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1Z)-N-(3-ethynylphenyl)-2,2,2-trifluoroethanimidoyl chloride (250 mg, 1.08 mmol, 1.0

equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 1.08 mL, 1.62 mmol, 1.5 equiv) were added. The desired product was obtained in 81% yield (355 mg) as a yellow oil after chromatography on silica gel (70:30 v/v, n-hexane/dichloromethane).

¹H NMR (400 MHz, CDCl₃) δ: 11.91 (s, 1H, NH), 7.34 (m, 2H, Ph H-4), 7.30 (m, 2H, Ph H-5), 7.18 (s, 2H, Ph H-2), 7.04 (m, 2H, Ph H-6), 5.91 (s, 1H, CH), 3.09 (s, 2H, C≡CH).

¹³C NMR (100 MHz, CDCl₃) δ: 149.4 (q, $^2J_{C,F}$ = 30.0 Hz, $\underline{C}$CF₃), 142.5 (Ph C-1), 129.6 (Ph C-4), 128.9 (Ph C-5), 126.2 (q, $J_{C,F}$ = 1.6 Hz, Ph C-2), 123.3 (q, $J_{C,F}$ = 1.9 Hz, Ph C-6), 122.9 (Ph C-3), 118.9 (q, $^1J_{C,F}$ = 283.5 Hz, CF₃), 89.1 (sept., $^3J_{C,F}$ = 4.8 Hz, CH), 82.7 ($\underline{C}$≡CH), 77.9 (C≡$\underline{C}$H).

¹⁹F NMR (376 MHz, CDCl₃) δ: -62.8 (s, CF₃).

¹⁵N NMR (40 MHz, CDCl₃) δ: -177.6 (N).

HRMS (ESI), *m/z*: calcd for $C_{21}H_{12}F_6N_2Na^+$: 429.0797 [M+Na]⁺ ; found: 429.0788.

N,N´-[(2Z,4Z)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis[4-(4-morpholinyl)aniline] (21)
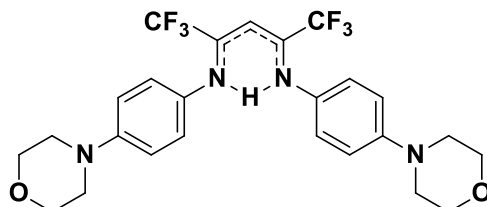
By following the General Procedure **2**, the homologation event was conducted starting from (1*Z*)-2,2,2-trifluoro-*N*-[4-(morpholin-4-yl)phenyl]ethanimidoyl chloride (250 mg, 0.85 mmol, 1.0 equiv), CH₂I₂ (452 mg, 0.19 mL, 2.56 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.74 mL, 1.11 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1*Z*)-2,2,2-trifluoro-*N*-[4-(morpholin-4-yl)phenyl]ethanimidoyl chloride (250 mg, 0.85 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 0.85 mL, 1.28 mmol, 1.5 equiv) were added. The desired product was obtained in 93% yield (418 mg) as an orange oil after chromatography on reverse phase silica gel (acetonitrile).

¹H NMR (400 MHz, CDCl₃) δ: 12.06 (s, 1H, NH), 6.99 (m, 4H, Ph H-2,6), 6.87 (m, 4H,

Ph H-3,5), 5.83 (s, 1H, CH), 3.86 (m, 8H, morph H-2,6), 3.15 (m, 8H, morph H-3,5).

13C NMR (100 MHz, CDCl$_3$) δ: 149.2 (Ph C-4), 148.6 (q, $^2J_{C,F}$ = 29.3 Hz, <u>C</u>CF$_3$), 135.0 (Ph C-1), 124.0 (q, $J_{C,F}$ = 1.8 Hz, Ph C-2,6), 119.3 (q, $^1J_{C,F}$ = 283.5 Hz, CF$_3$), 115.6 (Ph C-3,5), 88.3 (m, CH), 66.8 (morph C-2,6), 49.3 (morph C-3,5).

19F NMR (376 MHz, CDCl$_3$) δ: -62.9 (s, CF$_3$).

15N NMR (40 MHz, CDCl$_3$) δ: -177.4 (N).

HRMS (ESI), *m/z*: calcd for C$_{25}$H$_{27}$F$_6$N$_4$O$_2$: 529.2033 [M+H]$^+$; found: 529.2035.

N,N´-[(2Z,4Z)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis(4-nitroaniline) (22)
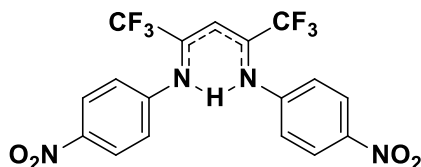
By following the General Procedure **2**, the homologation event was conducted starting from (1Z)-2,2,2-trifluoro-*N*-(4-(nitrophenyl)ethanimidoyl chloride (250 mg, 0.56 mmol, 1.0 equiv), CH$_2$I$_2$ (295 mg, 0.12 mL, 1.67 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.49 mL, 0.73 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1Z)-2,2,2-trifluoro-*N*-(4-(nitrophenyl)ethanimidoyl chloride (250 mg, 0.56 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 0.56 mL, 0.84 mmol, 1.5 equiv) were added. The desired product was obtained in 80% yield (201 mg) as a yellow oil after chromatography on reverse phase silica gel (85:15 v/v, acetonitrile/water).

1H NMR (400 MHz, CDCl$_3$) δ: 11.81 (brs, 1H, NH), 8.25 (m, 4H, Ph H-3,5), 7.17 (m, 4H, Ph H-2,6), 6.08 (s, 1H, CH).
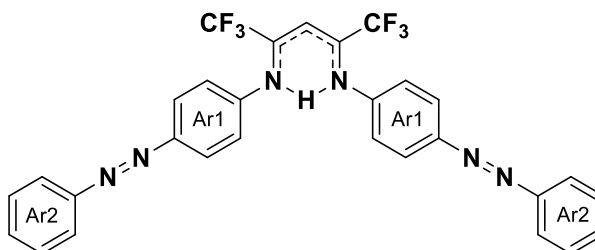
13C NMR (100 MHz, CDCl$_3$) δ: 149.3 (q, $^2J_{C,F}$ = 31.2 Hz, <u>C</u>CF$_3$), 147.8 (Ph C-1), 145.4 (Ph C-4), 124.9 (Ph C-3,5), 122.4 (Ph C-2,6), 118.7 (q, $^1J_{C,F}$ = 283.3 Hz, CF$_3$), 91.5 (m, CH).

19F NMR (376 MHz, CDCl$_3$) δ: -62.8 (s, CF$_3$).

15N NMR (40 MHz, CDCl$_3$) δ: 15.2 (NO$_2$), -176.7 (N).

HRMS (ESI), *m/z*: calcd for $C_{17}H_{11}F_6N_4O_4$: 449.0679 [M+H]$^+$; found: 449.0677.

N,N´-[(2Z,4Z)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis{4-[(E)-phenyldiazenyl]aniline} (23)

By following the General Procedure **2**, the homologation event was conducted starting from (1*Z*)-2,2,2-trifluoro-*N*-{4-[(*Z*)-phenyldiazenyl]phenyl}ethanimidoyl chloride (600 mg, 1.94 mmol, 1.0 equiv), CH_2I_2 (1550 mg, 0.47 mL, 5.79 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 1.68 mL, 2.52 mmol, 1.3 equiv) and THF (10 mL). Then, the second aliquots of (1*Z*)-2,2,2-trifluoro-*N*-{4-[(*Z*)-phenyldiazenyl]phenyl}ethanimidoyl chloride (600 mg, 1.94 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 1.94 mL, 2.91 mmol, 1.5 equiv) were added. The desired product was obtained in 92% yield (1011 mg) as orange oil after chromatography on silica gel (98:2 *v/v*, *n*-hexane/diethyl ether).

^{1}H NMR (400 MHz, CDCl$_3$) δ: 12.10 (brs, 1H, NH), 7.95 (m, 4H, Ar1 H-3,5), 7.91 (m, 4H, Ar2 H-2,6), 7.51 (m, 4H, Ar2 H-3,5), 7.49 (m, 2H, Ar2 H-4), 7.21 (m, 4H, Ar1 H-2,6), 6.00 (s, 1H, CH).
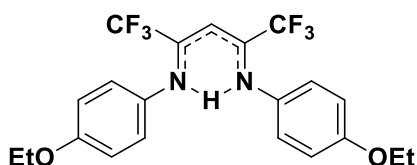
^{13}C NMR (100 MHz, CDCl$_3$) δ: 152.6 (Ar2 C-1), 150.3 (Ar1 C-4), 149.0 (q, $^2J_{C,F}$ = 30.4 Hz, $\underline{C}$CF$_3$), 144.9 (Ar1 C-1), 131.0 (Ar2 C-4), 129.1 (Ar2 C-3,5), 123.7 (Ar1 C-3,5), 122.9 (Ar1 C-2,6), 122.8 (Ar2 C-2,6), 119.0 (q, $^1J_{C,F}$ = 283.2 Hz, CF$_3$), 90.0 (m, CH).

^{19}F NMR (376 MHz, CDCl$_3$) δ: -62.8 (s, CF$_3$).

^{15}N NMR (40 MHz, CDCl$_3$) δ: -175.5 (N).

HRMS (ESI), *m/z*: calcd. for $C_{29}H_{20}F_6N_6Na^+$: 589.1546 [M+Na]$^+$; found: 589.1547.

***N,N´*-[(2Z,4Z)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis(4-ethoxyaniline) (24)**

By following the General Procedure **2**, the homologation event was conducted starting from (1*Z*)-*N*-(4-ethoxyphenyl)-2,2,2-trifluoroethanimidoyl chloride (550 mg, 2.18 mmol, 1.0 equiv), CH_2I_2 (1750 mg, 0.52 mL, 6.54 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 1.89 mL, 2.83 mmol, 1.3 equiv) and THF (10 mL). Then, the second aliquots of (1*Z*)-*N*-(4-ethoxyphenyl)-2,2,2-trifluoroethanimidoyl chloride (550 mg, 2.18 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 2.18 mL, 3.27 mmol, 1.5 equiv) were added. The desired product was obtained in 87% yield (847 mg) as orange oil after chromatography on silica gel (95:5 *v/v*, *n*-hexane/diethyl ether).

^{1}H NMR (400 MHz, CDCl$_3$) δ: 11.99 (brs, 1H, NH), 6.98 (m, 4H, Ph H-2,6), 6.85 (m, 4H, Ph H-3,5), 5.83 (s, 1H, CH), 4.03 (q, ^3J = 7.0 Hz, 4H, OC<u>H</u>$_2$CH$_3$), 1.41 (t, ^3J = 7.0 Hz, 6H, OCH$_2$C<u>H</u>$_3$).
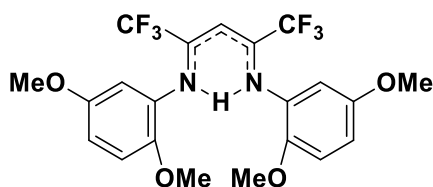
^{13}C NMR (100 MHz, CDCl$_3$) δ: 157.0 (Ph C-4), 149.1 (q, $^2J_{C,F}$ = 29.4 Hz, <u>C</u>CF$_3$), 135.5 (Ph C-1), 124.3 (Ph C-2,6), 119.1 (q, $^1J_{C,F}$ = 283.3 Hz, CF$_3$), 114.5 (Ph C-3,5), 88.2 (sept., $^3J_{C,F}$ = 4.8 Hz, CH), 63.7 (O<u>C</u>H$_2$CH$_3$), 14.8 (OCH$_2$<u>C</u>H$_3$).

^{19}F NMR (376 MHz, CDCl$_3$) δ: -62.9 (s, CF$_3$).

^{15}N NMR (40 MHz, CDCl$_3$) δ: -177.4 (N).

HRMS (ESI), *m/z*: calcd. for C$_{21}$H$_{21}$F$_6$N$_2$O$_2^+$: 447.1502 [M+H]$^+$; found: 447.1500.

***N,N'*-[(2Z,4Z)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis(2,5-dimethoxyaniline) (25)**

By following the General Procedure **2**, the homologation event was conducted starting from (1*Z*)-*N*-(2,5-dimethoxyphenyl)-2,2,2- trifluoroethanimidoyl chloride (500 mg, 1.87 mmol, 1.0 equiv), CH_2I_2 (1501 mg, 0.45 mL, 5.6 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 1.62 mL, 2.43 mmol, 1.3 equiv) and THF (10 mL). Then, the second aliquots of (1*Z*)-*N*-(2,5-dimethoxyphenyl)-2,2,2- trifluoroethanimidoyl chloride (500 mg, 1.87 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 1.87 mL, 2.81 mmol, 1.5 equiv) were added. The desired product was obtained in 79% yield (707 mg) as an orange oil after chromatography on silica gel (90:10 v/v, *n*-hexane/acetone).

¹H NMR (400 MHz, CDCl₃) δ: 12.10 (brs, 1H, NH), 6.81 (d, ^3J = 8.4 Hz, 2H, Ph H-3), 6.67 (s, 2H, Ph H-6), 6.66 (m, 2H, Ph H-4), 5.90 (s, 1H, CH), 3.76 (s, 6H, 5-OCH₃), 3.71 (s, 6H, 2-OCH₃).

¹³C NMR (100 MHz, CDCl₃) δ: 153.4 (Ph C-5), 148.8 (q, $^2J_{C,F}$ = 30.2 Hz, <u>C</u>CF₃), 145.3 (Ph C-2), 132.9 (Ph C-1), 119.2 (q, $^1J_{C,F}$ = 283.2 Hz, C<u>C</u>F₃), 112.0 (Ph C-3), 110.4   (Ph C-4), 109.1 (Ph C-6), 90.3 (m, CH), 56.3 (2-OCH₃), 55.7 (5-OCH₃).

¹⁹F NMR (376 MHz, CDCl₃) δ: -64.8 (s, CF₃).

HRMS (ESI), *m/z*: calcd for $C_{21}H_{21}F_6N_2O_4$: 479.1400 [M+H]⁺; found: 479.1403.

N,N´-**[(2Z,4Z)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis(1,3-benzodioxol-5-amine) (26)**
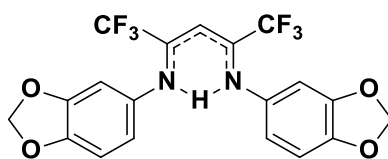
By following the General Procedure **2**, the homologation event was conducted starting from (1*Z*)-*N*-(2*H*-1,3-benzodioxol-5-yl)-2,2,2-trifluoroethanimidoyl chloride (250 mg, 0.99 mmol, 1.0 equiv), CH_2I_2 (798 mg, 0.24 mL, 2.98 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 1.06 mL, 1.58 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1*Z*)-*N*-(2*H*-1,3-benzodioxol-5-yl)-2,2,2-trifluoroethanimidoyl chloride (250 mg, 0.99 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 0.99 mL, 1.49 mmol, 1.5 equiv) were added. The desired product was obtained in 84% yield (371 mg) as a yellow oil after chromatography on silica gel (90:10 v/v *n*-hexane/diethyl ether).

¹H NMR (400 MHz, CDCl₃) δ: 11.87 (brs, 1H, NH), 6.75 (d, ^3J = 8.2 Hz, 2H, Ph-H 7), 6.58 (d, ^4J = 2.1 Hz, 2H, Ph-H 4), 6.51 (dd, ^3J = 8.2 Hz, ^4J = 2.1 Hz, 2H, Ph-H 6), 5.98 (s, 4H, OCH₂O), 5.82 (s, 1H, CH).

¹³C NMR (100 MHz, CDCl₃) δ: 149.3 (q, $^2J_{C,F}$ = 29.4 Hz, $\underline{C}$CF₃), 147.8 (Ph C-3a), 145.7 (Ph C-7a), 136.7 (Ph C-5), 119.0 (q, $^1J_{C,F}$ = 283.1 Hz, CF₃), 116.1 (q, $J_{C,F}$ = 1.8 Hz, Ph C-6), 107.9 (Ph C-7), 104.8 (q, $J_{C,F}$ = 1.7 Hz, Ph C-4), 101.5 (OCH₂O), 88.5 (m, CH).

¹⁹F NMR (376 MHz, CDCl₃) δ: -62.9 (s, CF₃).

HRMS (ESI), *m/z*: calcd for C₁₉H₁₂F₆N₂O₄Na⁺: 469.0593 [M+Na]⁺; found: 469.0598.

N,N´-[(2Z,4Z)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis[4-(trifluoromethoxy)aniline (27)
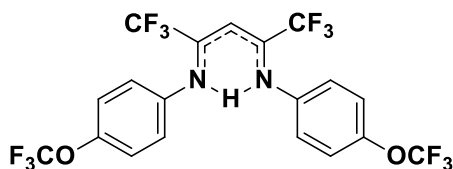
By following the General Procedure **2**, the homologation event was conducted starting from (1Z)-2,2,2-trifluoro-*N*-[4-(trifluoromethoxy)phenyl]ethanimidoyl chloride (150 mg, 0.51 mmol, 1.0 equiv), CH₂I₂ (415 mg, 0.12 mL, 1.55 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.44 mL, 0.66 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1Z)-2,2,2-trifluoro-*N*-[4-(trifluoromethoxy)phenyl]ethanimidoyl chloride (150 mg, 0.51 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 0.51 mL, 0.77 mmol, 1.5 equiv) were added. The desired product was obtained in 93% yield (250 g) as a yellow oil after chromatography on reverse phase silica gel (85:15 v/v, acetonitrile/water).

¹H NMR (400 MHz, CDCl₃) δ: 11.90 (brs, 1H, NH), 7.20 (m, 4H, Ph H-3,5), 7.07 (m, 4H, Ph H-2,6), 5.92 (s, 1H, CH).

¹³C NMR (100 MHz, CDCl₃) δ: 149.6 (q, $^2J_{C,F}$ = 30.0 Hz, $\underline{C}$CF₃), 147.0 (q, $^3J_{C,F}$ = 2.0 Hz, Ph C-4), 140.9 (Ph C-1), 124.1 (q, $J_{C,F}$ = 1.6 Hz, Ph C-2,6), 121.5 (Ph C-3,5), 120.4 (q, $^1J_{C,F}$ = 257.4 Hz, OCF₃), 118.9 (q, $^1J_{C,F}$ = 283.6 Hz, CF₃), 89.2 (m, CH).

¹⁹F NMR (376 MHz, CDCl₃) δ: -62.9 (s, CF₃).

^{15}N NMR (40 MHz, CDCl$_3$) δ: -178.6 (N).

HRMS (ESI), *m/z*: calcd for C$_{19}$H$_{11}$F$_{12}$N$_2$O$_2$: 527.0623 [M+H]$^+$; found: 527.0621.

_N_,_N_´-[(2_Z_,4_Z_)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis[4-(pyrrolidinylsulfonyl) aniline) (28)
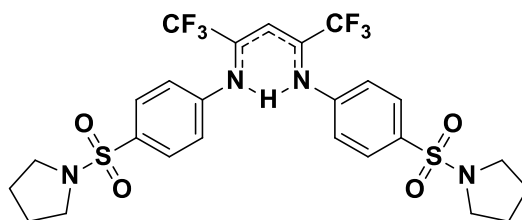
By following the General Procedure **2**, the homologation event was conducted starting from (1*Z*)-2,2,2-trifluoro-*N*-[4-(pyrrolidine-1-sulfonyl)phenyl]ethanimidoyl chloride (120 mg, 0.35 mmol, 1.0 equiv), CH$_2$I$_2$ (283 mg, 0.08 mL, 1.05 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.30 mL, 0.46 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1*Z*)-2,2,2-trifluoro-*N*-[4-(pyrrolidine-1-sulfonyl)phenyl]ethanimidoyl chloride (120 mg, 0.35 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 0.35 mL, 0.53 mmol, 1.5 equiv) were added. The desired product was obtained in 81% yield (177 mg) as a yellow oil, after chromatography on Alox-BG2 (60:40 v/v, dichloromethane/n-hexane).

^{1}H NMR (400 MHz, CDCl$_3$) δ: 11.86 (brs, 1H, NH), 7.82 (m, 4H, Ph H-3,5), 7.18 (m, 4H, Ph H-2,6), 6.01 (s, 1H, CH), 3.25 (m, 8H, NC$\underline{H}_2$CH$_2$), 1.76 (m, 8H, NCH$_2$C$\underline{H}_2$).

^{13}C NMR (100 MHz, CDCl$_3$) δ: 149.4 (q, $^2J_{C,F}$ = 30.6 Hz, $\underline{C}$CF$_3$), 146.1 (Ph C-1), 134.4 (Ph C-4), 128.5 (Ph C-3,5), 122.6 (q, $J_{C,F}$ = 1.8 Hz, Ph C-2,6), 118.8 (q, $^1J_{C,F}$ = 283.4 Hz, CF$_3$), 90.5 (m, CH), 48.0 (N$\underline{C}$H$_2$CH$_2$), 25.3 (NCH$_2\underline{C}$H$_2$).

^{19}F NMR (376 MHz, CDCl$_3$) δ: -62.8 (s, CF$_3$).

^{15}N NMR (40 MHz, CDCl$_3$) δ: -177.1 (N).

HRMS (ESI), *m/z*: calcd for C$_{25}$H$_{27}$F$_6$N$_4$O$_4$S$_2$: 625.1372 [M+H]$^+$; found: 625.1374.

_N_,_N_´-[(2_Z_,4_Z_)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis[4-

(methylsulfanyl)aniline] (29)
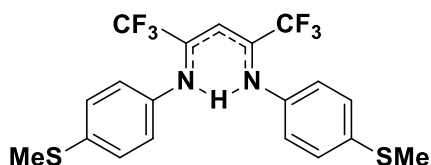
By following the General Procedure **2**, the homologation event was conducted starting from (1*Z*)-2,2,2-trifluoro-*N*-[4-(methylsulfanyl)phenyl]ethanimidoyl chloride (250 mg, 0.99 mmol, 1.0 equiv), CH_2I_2 (792 mg, 0.24 mL, 2.96 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.86 mL, 1.29 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1*Z*)-2,2,2-trifluoro-*N*-[4-(methylsulfanyl)phenyl]ethanimidoyl chloride (250 mg, 0.99 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 0.99 mL, 1.49 mmol, 1.5 equiv) were added. The desired product was obtained in 87% yield (388 mg) as a yellow solid (m.p.:112 °C) after chromatography on silica gel (90:10 v/v, *n*-hexane/diethyl ether).

¹H NMR (400 MHz, CDCl₃) δ: 12.00 (s, 1H, NH), 7.23 (m, 4H, Ph H-3,5), 6.98 (m, 4H, Ph H-2,6), 5.87 (s, 1H, CH), 2.48 (s, 6H, SCH₃).
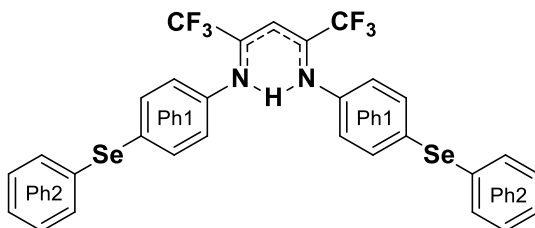
¹³C NMR (100 MHz, CDCl₃) δ: 149.0 (q, $^2J_{C,F}$ = 29.8 Hz, $\underline{C}CF_3$), 139.8 (Ph C-1), 135.9 (Ph C-4), 127.1 (Ph C-3,5), 123.3 (Ph C-2,6), 119.0 (q, $^1J_{C,F}$ = 283.5 Hz, CF₃), 88.9 (sept, $^3J_{C,F}$ = 4.9 Hz, CH), 16.2 (SCH₃).

¹⁹F NMR (376 MHz, CDCl₃) δ: -62.8 (s, CF₃).

¹⁵N NMR (40 MHz, CDCl₃) δ: -177.2 (N).

HRMS (ESI), *m/z*: calcd for $C_{19}H_{17}F_6N_2S_2$: 451.0732 [M+H]⁺; found: 451.0729.

N,N´-[(2Z,4Z)-1,1,1,5,5,5-Hexafluoro-2-penten-2-yl-4-ylidene]bis[4-(phenylselanyl)aniline] (30)

By following the General Procedure **2**, the homologation event was conducted starting from (1*Z*)-2,2,2-trifluoro-*N*-[4-(phenylselanyl)phenyl]ethanimidoyl chloride (120 mg, 0.33 mmol, 1.0 equiv), CH_2I_2 (266 mg, 0.08 mL, 0.99 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 0.29 mL, 0.43 mmol, 1.3 equiv) and THF (5 mL). Then, the second aliquots of (1*Z*)-2,2,2-trifluoro-*N*-[4-(phenylselanyl)phenyl]ethanimidoyl chloride (120 mg, 0.33 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 0.33 mL, 0.50 mmol, 1.5 equiv) were added. The desired product was obtained in 90% yield (198 mg) as a yellow oil after chromatography on reverse phase silica gel (acetonitrile).

^{1}H NMR (400 MHz, $CDCl_3$) δ: 11.97 (s, 1H, NH), 7.45 (m, 4H, Ph2 H-2,6), 7.43 (m, 4H, Ph1 H-3,5), 7.28 (m, 4H, Ph2 H-3,5), 7.27 (m, 2H, Ph2 H-4), 6.95 (m, 4H, Ph1 H-2,6), 5.88 (s, 1H, CH).
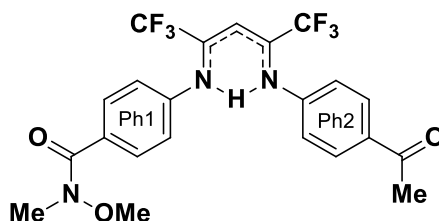
^{13}C NMR (100 MHz, $CDCl_3$) δ: 149.1 (q, $^2J_{C,F}$ = 29.5 Hz, $\underline{C}CF_3$), 141.9 (Ph1 C-1), 133.5 (Ph1 C-3,5), 132.9 (Ph2 C-2,6), 131.0 (Ph2 C-1), 129.4 (Ph2 C-3,5), 128.2 (Ph1 C-4), 127.4 (Ph2 C-4), 123.5 (q, $J_{C,F}$ = 1.7 Hz, Ph1 C 2,6), 119.0 (q, $^1J_{C,F}$ = 283.0 Hz, CF_3), 89.2 (m, CH).

^{19}F NMR (376 MHz, $CDCl_3$) δ: -62.8 (s, CF_3).

^{15}N NMR (40 MHz, $CDCl_3$) δ: -177.1 (N).

HRMS (ESI), *m/z*: calcd for $C_{29}H_{21}F_6N_2Se_2$: 670.9934 [M+H]$^+$; found: 670.9936.

4-({(2Z,3Z)-4-[(4-Acetylphenyl)amino]-1,1,1,5,5,5-hexafluoro-3-penten-2-ylidene}amino)*N*-methoxy-*N*-methylbenzamide (32)

By following the General Procedure **2**, the homologation event was conducted starting from (1*E*)-2,2,2-trifluoro-*N*-{4-[methoxy(methyl)carbamoyl]phenyl}ethanimidoyl chloride (500 mg, 1.70 mmol, 1.0 equiv), CH_2I_2 (1363 mg, 0.41 mL, 5.09 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 1.47 mL, 2.21 mmol, 1.3 equiv) and THF (10 mL).

Then, the second aliquots of (1*E*)-2,2,2-trifluoro-*N*-{4-[methoxy(methyl)carbamoyl]phenyl}ethanimidoyl chloride (500 mg, 1.70 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 1.70 mL, 2.55 mmol, 1.5 equiv) were added. However, no aqueous quenching was realized after the conclusion of the homologation-metalation sequence. Indeed, additional MeLi-LiBr complex (2.2 M, 0.85 mL, 1.87 mmol, 1.1 equiv) was added to the reaction mixture kept at -78 °C. After 30 min, a saturated solution of NH_4Cl was added (2 mL/mmol substrate), then was extracted with Et_2O (2 x 5 mL) and washed with water (5 mL) and brine (10 mL). The organic phase was dried over anhydrous Na_2SO_4, filtered and the solvent was removed under reduced pressure. The desired compound was obtained in 71% yield (588 mg) as a yellow oil, after chromatography on Alox-BG2 (70:30 v/v, dichloromethane/*n*-hexane).

^{1}H NMR (400 MHz, CDCl₃) δ: 11.95 (s, 1H, NH), 7.95 (m, 2H, Ph2 H-3,5), 7.72 (m, 2H, Ph1 H-2,6), 7.11 (m, 2H, Ph1 H-3,5), 7.08 (m, 2H, Ph2 H-2,6), 5.97 (s, 1H, CH), 3.54 (s, 3H, OCH₃), 3.36 (s, 3H, NCH₃), 2.59 (s, 3H, COCH₃).
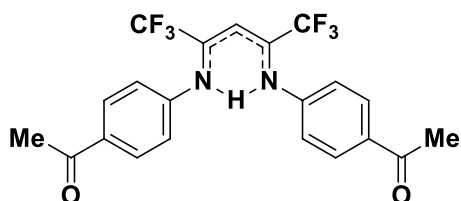
^{13}C NMR (100 MHz, CDCl₃) δ: 196.8 (C=O), 168.8 (NC=O), 149.7 (q, $^2J_{C,F}$ = 30.0 Hz, <u>C</u>CF₃), 147.8 (Ph2 C-1), 143.4 (Ph1 C-4),134.2 (Ph2 C-4), 131.6 (Ph1 C-1), 129.40 (Ph1 C-2,6), 129.36 (Ph2 C-3,5), 122.4 (q, $J_{C,F}$ = 1.8 Hz, Ph C-3,5), 121.5 (q, $J_{C,F}$ = 1.9 Hz, Ph C-2,6), 90.0 (m, CH), 61.0 (OCH₃), 33.6 (br, NCH₃), 26.5 (CO<u>C</u>H₃), C-2 and C-4 not unambiguously identified.

^{19}F NMR (376 MHz, CDCl₃) δ: -62.9 (s, CF₃), -62.8 (s, CF₃).

^{15}N NMR (40 MHz, CDCl₃) δ: -178.1 (N), -173.9 (N).

HRMS (ESI), *m/z*: calcd for $C_{22}H_{20}F_6N_3O_3$: 488.1403 [M+H]⁺; found: 488.1405.

1-[4-({(2Z,3Z)-4-[(4-Acetylphenyl)amino]-1,1,1,5,5,5-hexafluoro-3-penten-2-ylidene}amino) phenyl]ethanone (33)

By following the General Procedure **2**, the homologation event was conducted starting

from (1*E*)-2,2,2-trifluoro-*N*-{4-[methoxy(methyl)carbamoyl]phenyl}ethanimidoyl chloride (500 mg, 1.70 mmol, 1.0 equiv), CH_2I_2 (1363 mg, 0.41 mL, 5.09 mmol, 3.0 equiv), MeLi-LiBr complex (1.5 M, 1.47 mL, 2.21 mmol, 1.3 equiv) and THF (10 mL). Then, the second aliquots of (1*E*)-2,2,2-trifluoro-*N*-{4-[methoxy(methyl)carbamoyl]phenyl}ethanimidoyl chloride (500 mg, 1.70 mmol, 1.0 equiv) in THF (3 mL) and MeLi-LiBr complex (1.5 M, 1.70 mL, 2.55 mmol, 1.5 equiv) were added. However, no aqueous quenching was realized after the conclusion of the homologation-metalation sequence. Indeed, additional MeLi-LiBr complex (2.2 M, 1.85 mL, 4.08 mmol, 2.4 equiv) was added to the reaction mixture kept at -78 °C. After 30 min, a saturated solution of NH_4Cl was added (2 mL/mmol substrate), then was extracted with Et_2O (2 x 5 mL) and washed with water (5 mL) and brine (10 mL). The organic phase was dried over anhydrous Na_2SO_4, filtered and the solvent was removed under reduced pressure. The desired compound was obtained in 79% yield (594 mg) as a yellow oil, after chromatography on Alox-BG2 (70:30 v/v, dichloromethane/*n*-hexane).

1H NMR (400 MHz, $CDCl_3$) δ: 11.93 (brs, 1H, NH), 7.96 (m, 4H, Ph H-3,5), 7.12 (m, 4H, Ph H-2,6), 6.00 (s, 1H, CH), 2.60 (s, 3H, CH_3).

13C NMR (100 MHz, $CDCl_3$) δ: 196.8 (C=O), 148.9 (q, $^2J_{C,F}$ = 30.5 Hz, $\underline{C}CF_3$), 146.6 (Ph C-1), 134.5 (Ph C-4), 129.4 (Ph C-3,5), 122.0 (q, $J_{C,F}$ = 1.7 Hz, Ph C-2,6), 118.9 (q, $^1J_{C,F}$ = 283.6 Hz, CF_3), 90.5 (m, CH), 26.5 (CH_3).
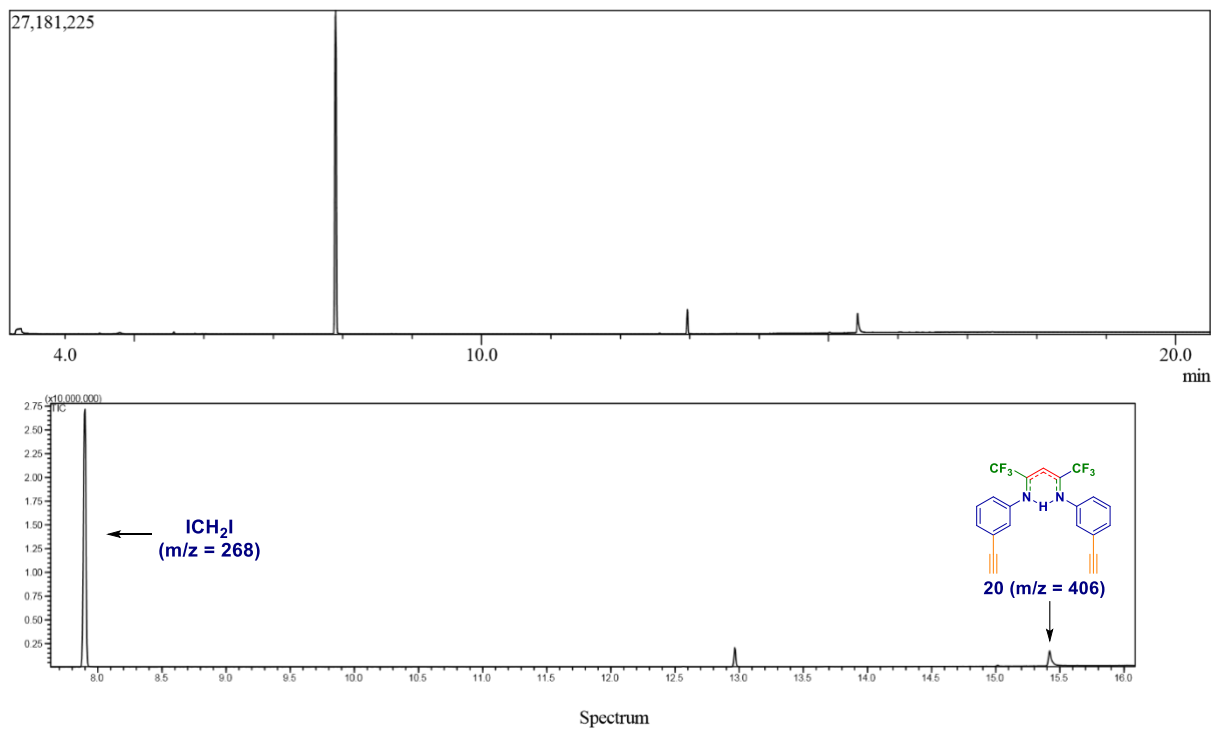
19F NMR (376 MHz, $CDCl_3$) δ: -62.8 (s, CF_3).

15N NMR (40 MHz, $CDCl_3$) δ: -175.6 (N).

HRMS (ESI), *m/z*: calcd for $C_{21}H_{17}F_6N_2O_2$: 443.1189 [M+H]$^+$; found: 443.1187.
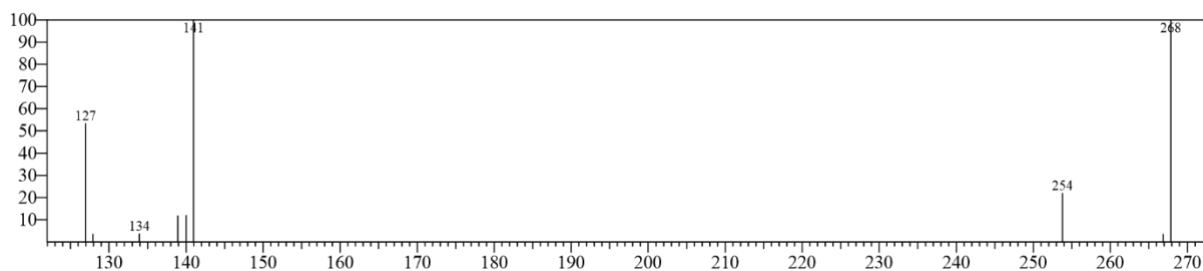
- **GC-MS Study of the Mechanistic Evidence for Compound 20**

Scheme a – Use of LiCH₂I as metalating agent leading to ICH₂I as exchange product

Line#:1 R.Time:7.895(Scan#:940) MassPeaks:9

RawMode:Single 7.895(940) BasePeak:141(8421955) BG Mode:None Group 1 - Event ICH₂I (m/z = 268)
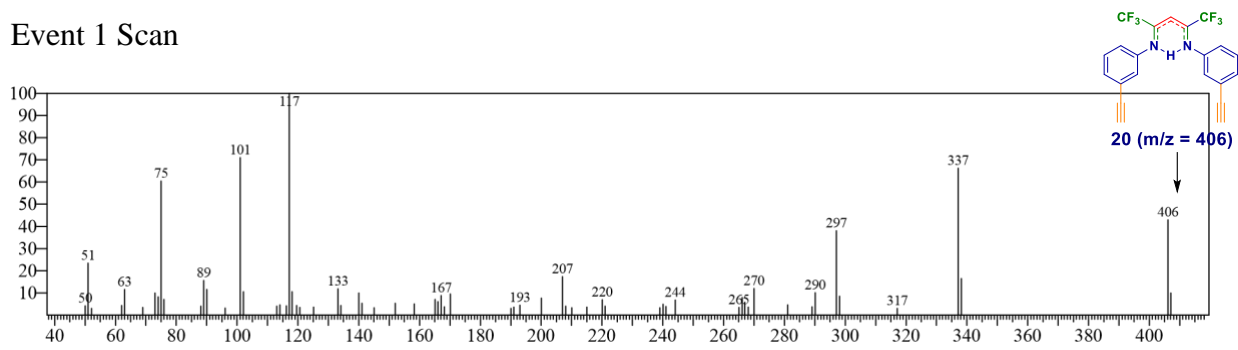
1 Scan

m/z

Line#:3 R.Time:15.425(Scan#:2446)

MassPeaks:65

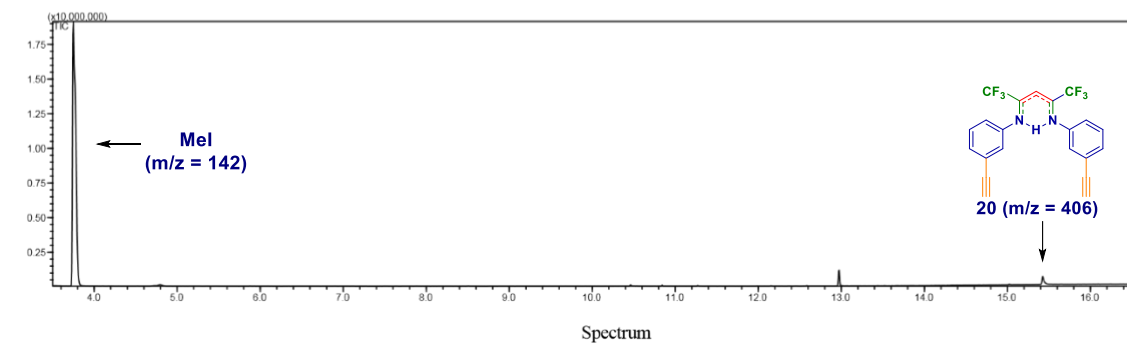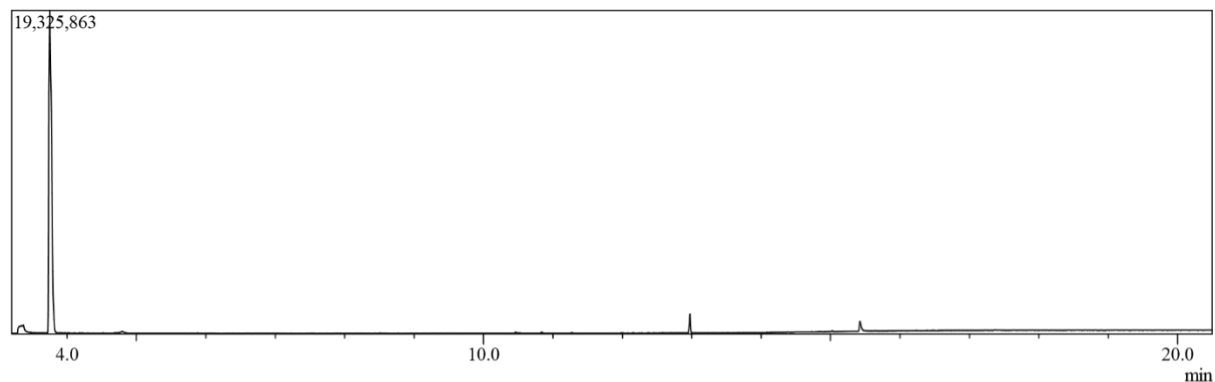RawMode:Single 15.425(2446) BasePeak:117(139145) BG Mode:None Group 1 -

Event 1 Scan

m/z

Scheme b – Use of MeLi as metalating agent leading to CH₃I as exchange product
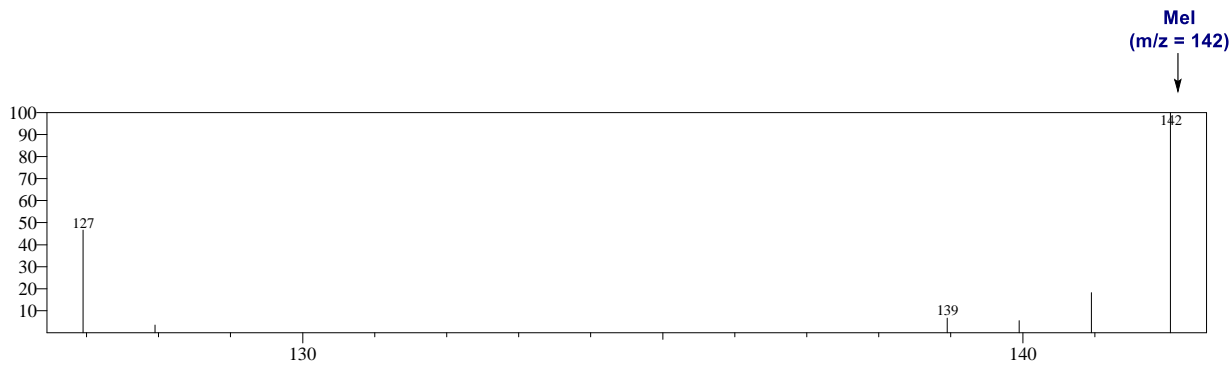
TIC

Line#:1   R.Time:3.750(Scan#:111)

MassPeaks:6

RawMode:Averaged 3.735-3.865(108-134)   BasePeak:142(3395652) BG Mode:None
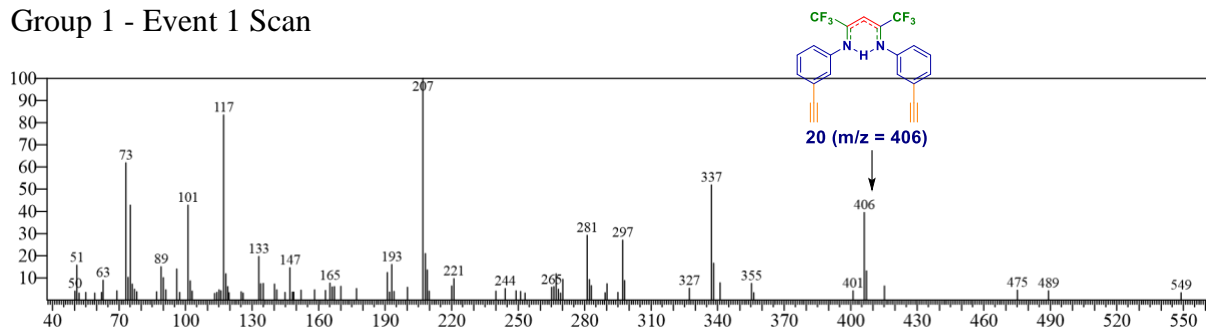
Group 1 - Event 1 Scan

m/z

Line#:3   R.Time:15.425(Scan#:2446)

MassPeaks:93

RawMode:Averaged 15.385-15.475(2438-2456) BasePeak:207(24766) BG Mode:None

Group 1 - Event 1 Scan

m/z

*Scheme c – Use of n-**BuLi** as metalating agent leading to n-**BuI** as exchange product*

TIC

Line#:1   R.Time:6.695(Scan#:700)

MassPeaks:13

RawMode:Averaged 6.660-6.730(693-707)   BasePeak:57(4034785) BG Mode:None

Group 1 - Event 1 Scan

m/z

Line#:4   R.Time:15.425(Scan#:2446)

MassPeaks:93

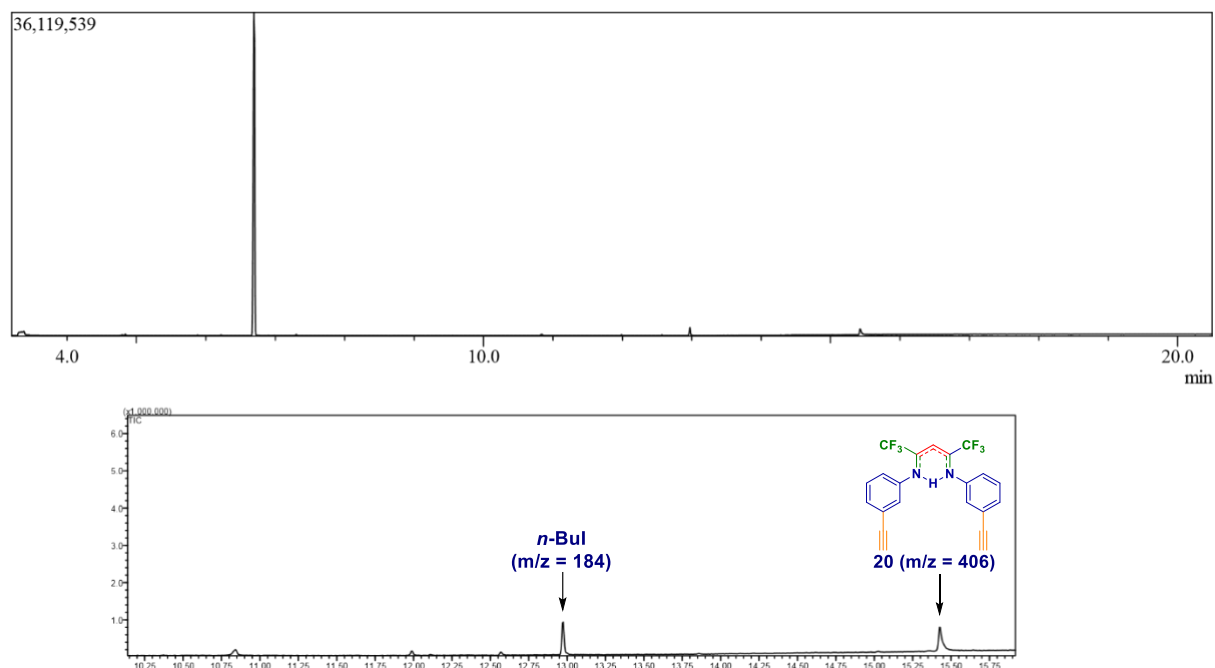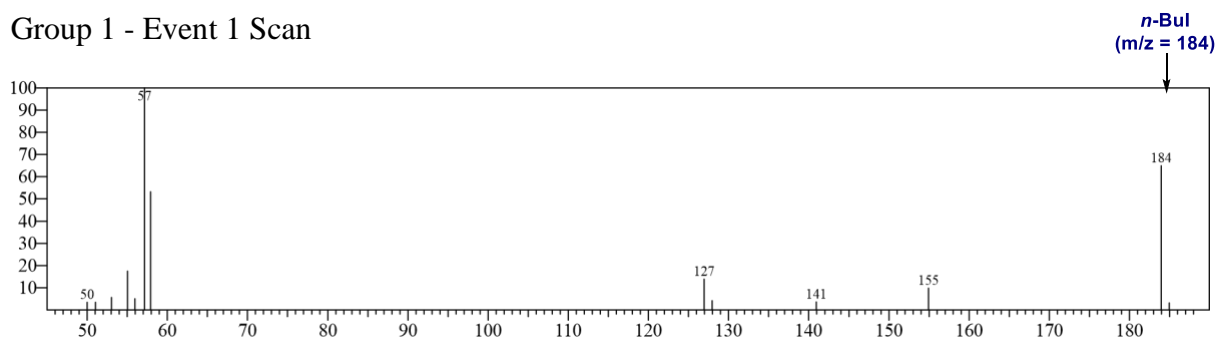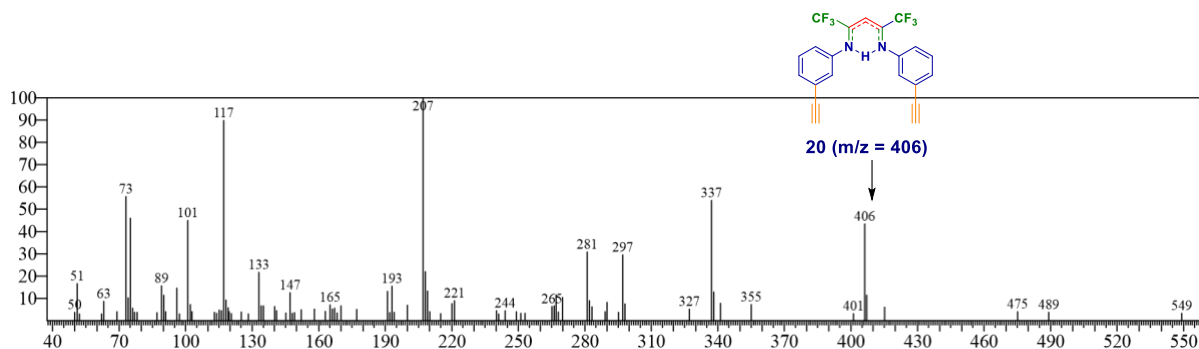RawMode:Averaged 15.395-15.495(2440-2460) BasePeak:207(25255) BG Mode:None

Group 1 - Event 1 Scan

m/z

Additional Detection of 1-Iodobutane and characterization after chromatographic purification

To a cooled (-78 °C) solution of (1Z)-N-(3-ethynylphenyl)-2,2,2-trifluoroethanimidoyl chloride (250 mg, 1.08 mmol, 1.0 equiv) in dry THF (5 mL) was added diiodomethane (376 mg, 0.11 mL, 1.4 mmol, 1.3 equiv). After 2 min, an ethereal solution of MeLi-LiBr (1.5 M, 0.94 mL, 1.40 mmol, 1.3 equiv) was added dropwise, using a syringe pump (flow: 0.200 mL/min). The resulting solution was stirred for 30 minutes at -78 °C. After that, n-BuLi (1.6 M in hexanes, 1.5 equiv, 1.0 mL, 1.62 mmol) was added dropwise, using a syringe pump (flow: 0.200 mL/min) and the reaction mixture was stirred for a further 30 minutes at -78 °C. A saturated solution of NH$_4$Cl was added (2 mL/mmol substrate), then was extracted with Et$_2$O (2 x 5 mL) and washed with water (5 mL) and brine (10 mL). The organic phase was dried over anhydrous Na$_2$SO$_4$, filtered and, after removal of the solvent under reduced pressure, the so-obtained crude mixture was subjected to chromatography (silica gel, Alox-BG2 or reverse phase silica gel using 90:10 v/v, n-hexane/dichloromethane as the eluent) to afford compound **20** (136 mg, 31% yield) and 1-iodobutane (50 mg, 25% yield). *Spectral data of compound* **20** *match with those reported above. For all the experiments subjected to GC-MS analysis (Scheme 3), the corresponding aliquots were extracted after 3 min from the beginning of the addition of the metalating agent.*

224

Me⌢⌢⌢I
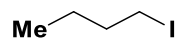
¹H NMR (400 MHz, CDCl₃) δ: 3.19 (t, J = 7.0 Hz, 2H, CH₂CH₂C$\underline{H}$₂I), 1.80 (m, 2H, CH₂C$\underline{H}$₂CH₂I), 1.42 (m, 2H, C$\underline{H}$₂CH₂CH₂I), 0.92 (t, J = 7.0 Hz, 3H, CH₃).

¹³C NMR (100 MHz, CDCl₃) δ: 35.5 (CH₂$\underline{C}$H₂CH₂I), 23.6 ($\underline{C}$H₂CH₂CH₂I), 12.9 (CH₃), 6.9 (CH₂CH₂$\underline{C}$H₂I).

HRMS (ESI), *m/z*: calcd for C₄H₁₀I: 184.9822 [M+H]⁺; found: 184.9823.

Scheme d – Use of LiCH₂Br as metalating agent leading to ICH₂Br as exchange product

TIC
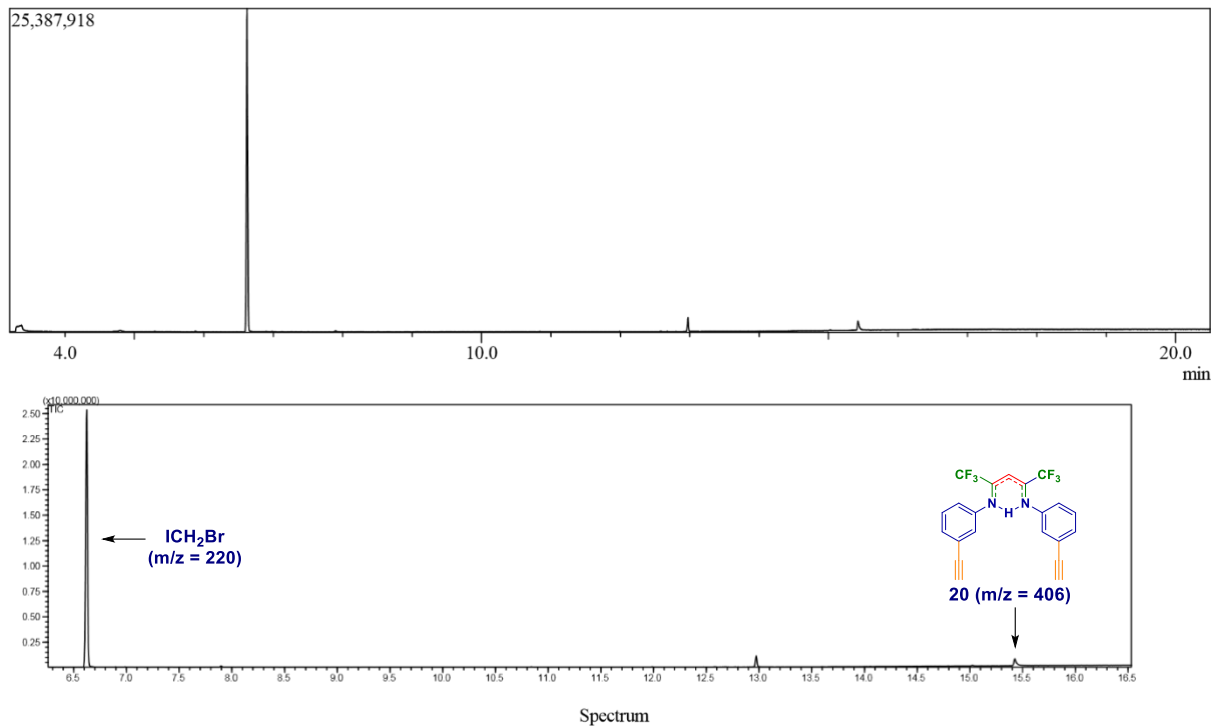
Line#:1   R.Time:6.625(Scan#:686)

MassPeaks:17

RawMode:Averaged 6.570-6.680(675-697)   BasePeak:93(888969) BG Mode:None

Group 1 - Event 1 Scan

m/z

Line#:3   R.Time:15.425(Scan#:2446)

MassPeaks:98

RawMode:Averaged 15.405-15.455(2442-2452) BasePeak:117(34158) BG Mode:None

Group 1 - Event 1 Scan

Bibliography

1. Pace, V., Verniest, G., Sinisterra, J. V., Alcántara, A. R. & De Kimpe, N. Improved arndt-eistert synthesis of α-diazoketones requiring minimal diazomethane in the presence of calcium oxide as acid scavenger. *Journal of Organic Chemistry* **75**, 5760–5763 (2010).
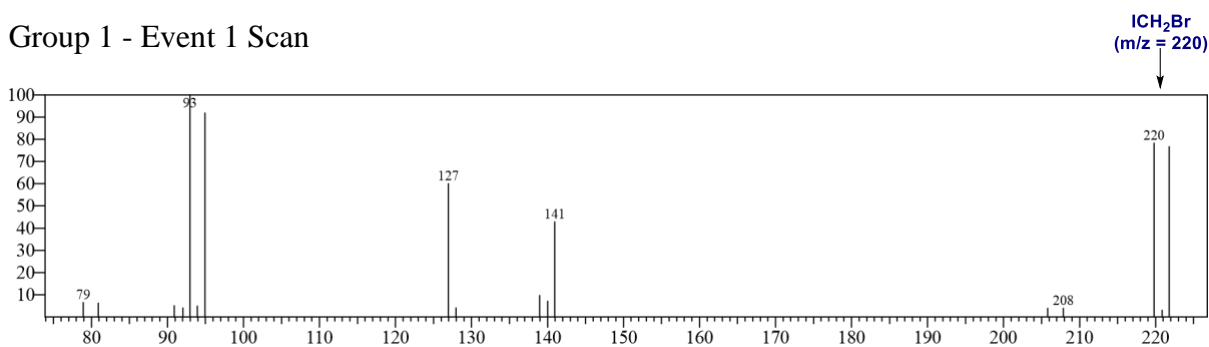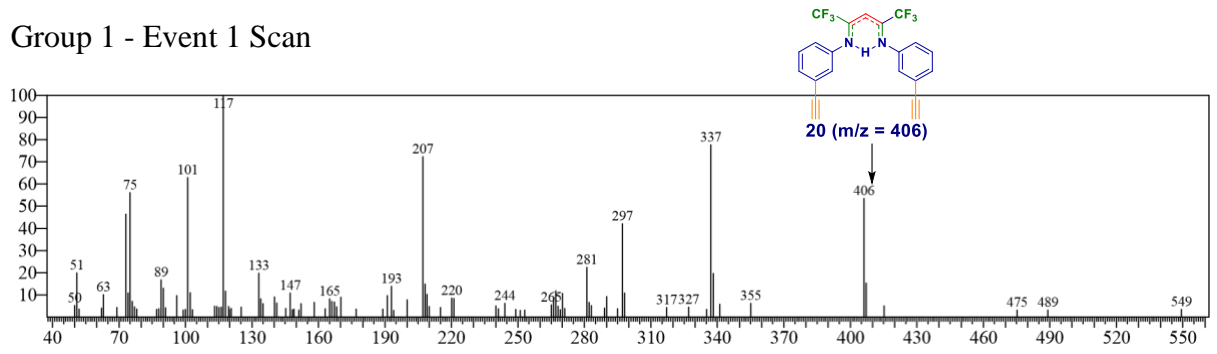
2. Pace, V., Holzer, W. & De Kimpe, N. Lithium Halomethylcarbenoids: Preparation and Use in the Homologation of Carbon Electrophiles. *Chemical Record* 2061–2076 (2016). doi:10.1002/tcr.201600011

3. Maas, G. New syntheses of diazo compounds. *Angewandte Chemie - International Edition* **48**, 8186–8195 (2009).

4. Maryanoff, B. E. & Reitz, A. B. The Wittig Olefination Reaction and Modifications Involving Phosphoryl-Stabilized Carbanions. Stereochemistry, Mechanism, and Selected Synthetic Aspects. *Chemical Reviews* **89**, 863–927 (1989).

5. Barluenga, J., Fernandez-simon, J. L., Concellon, J. M. & Yus, M. Methylenation of Carbonyl Compounds Using Chloromethyl-lithium; a New Method for Terminal and Exocyclic Olefins. 2721 (1986).

6. Castagnolo, D., Degennaro, L., Luisi, R. & Clayden, J. Enantioselective carbolithiation of S-alkenyl-N-aryl thiocarbamates: Kinetic and thermodynamic control. *Organic and Biomolecular Chemistry* **13**, 2330–2340 (2015).

7. Corey, E. J. & Chaykovsky, M. Dimethylsulfoxonium Methylide. *Journal of the American Chemical Society* **84**, 867–868 (1962).

8. Basavaiah, D., Rao, A. J. & Satyanarayana, T. Recent advances in the Baylis - Hillman reaction and applications. *Chemical Reviews* **103**, 811–891 (2003).

9. Köbrich, G. The Chemistry of Carbenoids and Other Thermolabile Organolithium Compounds. *Angewandte Chemie International Edition in English* **11**, 473–485 (1972).

10. Castoldi, L., Monticelli, S., Senatore, R., Ielo, L. & Pace, V. Homologation chemistry with nucleophilic α-substituted organometallic reagents: Chemocontrol, new concepts and (solved) challenges. *Chemical Communications*

54, 6692–6704 (2018).

11. Boche, G. & Lohrenz, J. C. W. The electrophilic nature of carbenoids, nitrenoids, and oxenoids. *Chemical Reviews* **101**, 697–756 (2001).

12. Okazaki, R. & West, R. Chemistry of Stable Disilenes. *Advances in Organometallic Chemistry* **39**, 231–273 (1996).

13. Sadhu, K. M. & Matteson, D. S. (Chloromethyl) lithium in an efficient conversion of carbonyl compounds to chlorohydrins or oxiranes. *Tetrahedron Letters* **27**, 795–798 (1986).

14. Barluenga, J., Baragaiia, B. & Concellón, J. M. Highly Diastereoselective Synthesis of Threo or Erythro Aminoalkyl Epoxides from α-Amino Acids. *Journal of Organic Chemistry* **60**, 6696–6699 (1995).

15. Barluenga, J., Llavona, L., Yus, M. & Concellón, J. M. Reactivity of in situ generated dihalomethyllithium towards dicarboxylic acid diesters and lactones: Synthetic applications. *Tetrahedron* **47**, 7875–7886 (1991).

16. Tarhouni, R., Kirschleger, B., Rambaud, M. & Villieras, J. Monohalomethyllithium XCH2Li: stabilization of a potential synthetic reagent. *Tetrahedron Letters* **25**, 835–838 (1984).

17. Stymiest, J. L., Dutheuil, G., Mahmood, A. & Aggarwal, V. K. Lithiated carbamates: Chiral carbenoids for iterative homologation of boranes and boronic esters. *Angewandte Chemie - International Edition* **46**, 7491–7494 (2007).

18. Pace, V., Castoldi, L., Mamuye, A. D. & Holzer, W. Homologation of isocyanates with lithium carbenoids: A straightforward access to α-halomethyl- and α,α-dihalomethylamides. *Synthesis (Germany)* **46**, 2897–2909 (2014).

19. Pace, V., Castoldi, L. & Holzer, W. Addition of Lithium Carbenoids to Isocyanates: a Direct Access to Synthetically Useful N-substituted 2-Haloacetamides. *Chemical Communications* **49**, 8383 (2013).

20. Pace, V., Castoldi, L. & Holzer, W. Synthesis of α,β-unsaturated α′-haloketones through the chemoselective addition of halomethyllithiums to Weinreb amides. *Journal of Organic Chemistry* **78**, 7764–7770 (2013).

21. Pace, V. Halomethyllithium carbenoids: Versatile reagents for the homologation of electrophilic carbon units. *Australian Journal of Chemistry* **67**, 311–313 (2014).

22. Knochel, P. *et al.* Highly functionalized organomagnesium reagents prepared through halogen-metal exchange. *Angewandte Chemie - International Edition* **42**, 4302–4320 (2003).

23. Nishimura, R. H. V., Murie, V. E., Soldi, R. A. & Clososki, G. C. (Chloromethyl)magnesium chloride-lithium chloride: A chemoselective reagent for the synthesis of functionalized aromatic chlorohydrins. *Synthesis (Germany)* **47**, 1455–1460 (2015).

24. Bozzini, L. A., Batista, J. H. C., de Mello, M. B. M., Vessecchi, R. & Clososki, G. C. Selective functionalization of cyano-phenyl-2-oxazolines using TMPMgCl·LiCl. *Tetrahedron Letters* **58**, 4186–4190 (2017).

25. Avolio, S., Malan, C., Marek, I. & Knochel, P. Preparation and reactions of functionalized magnesium carbenoids. *Synlett* 1820–1822 (1999). doi:10.1055/s-1999-2935

26. Boultwood, T., Affron, D. P., Trowbridge, A. D. & Bull, J. A. Synthesis of cis-C-iodo-N-tosyl-aziridines using diiodomethyllithium: Reaction optimization, product scope and stability, and a protocol for selection of stationary phase for chromatography. *Journal of Organic Chemistry* **78**, 6632–6647 (2013).

27. Bull, J. A., Boultwood, T. & Taylor, T. A. Highly cis-selective synthesis of iodo-aziridines using diiodomethyllithium and in situ generated N-Boc-imines. *Chemical Communications* **48**, 12246–12248 (2012).

28. Hoffmann, R. W. The quest for chiral Grignard reagents. *Chemical Society Reviews* **32**, 225–230 (2003).

29. Blakemore, P. R., Marsden, S. P. & Vater, H. D. Reagent-controlled asymmetric homologation of boronic esters by enantioenriched main-group chiral carbenoids. *Organic Letters* **8**, 773–776 (2006).

30. Blakemore, P. R. & Burge, M. S. Iterative stereospecific reagent-controlled homologation of pinacol boronates by enantioenriched α-chloroalkyllithium reagents. *Journal of the American Chemical Society* **129**, 3068–3069 (2007).

31. Cogan, D. A., Liu, G., Kim, K., Backes, B. J. & Ellman, J. A. Catalytic asymmetric oxidation of tert-butyl disulfide. Synthesis of tert-butanesulfinamides, tert-butyl sulfoxides, and tert-butanesulfinimines. *Journal of the American Chemical Society* **120**, 8011–8019 (1998).

32. Pace, V., Castoldi, L. & Holzer, W. Chemoselective additions of chloromethyllithium carbenoid to cyclic enones: A direct access to chloromethyl allylic alcohols. *Advanced Synthesis and Catalysis* **356**, 1761–1766 (2014).

33. Mamuye, A. D., Monticelli, S., Castoldi, L., Holzer, W. & Pace, V. Eco-friendly chemoselective N-functionalization of isatins mediated by supported KF in 2-MeTHF. *Green Chemistry* **17**, 4194–4197 (2015).

34. Concellón, J. M., Rodríguez-Solla, H. & Simal, C. Addition reactions of iodomethyllithium to imines. A direct and efficient synthesis of aziridines and enantiopure amino aziridines. *Organic Letters* **10**, 4457–4460 (2008).

35. Concellón, J. M., Rodríguez-Solla, H., Bernad, P. L. & Simal, C. Addition reactions of chloro- Or iodomethyllithium to imines. Synthesis of enantiopure aziridines and β-chloroamines. *Journal of Organic Chemistry* **74**, 2452–2459 (2009).

36. Pace, V., Castoldi, L., Monticelli, S., Rui, M. & Collina, S. New Perspectives in Lithium Carbenoid Mediated Homologations. *Synlett* **28**, 879–888 (2017).

37. Barluenga, J., Baragaña, B., Alonso, A. & Concellón, J. M. The first direct preparation of chiral functionalised ketones and their synthetic uses. *Journal of the Chemical Society, Chemical Communications* 969–970 (1994). doi:10.1039/C39940000969

38. Singh, G. S., D'hooghe, M. & De Kimpe, N. Synthesis and reactivity of C-heteroatom-subtituted aziridines. *Chemical Reviews* **107**, 2080–2135 (2007).

39. Pace, V., Murgia, I., Westermayer, S., Langer, T. & Holzer, W. Highly efficient synthesis of functionalized α-oxyketones: Via Weinreb amides homologation with α-oxygenated organolithiums. *Chemical Communications* **52**, 7584–7587 (2016).

40. Senatore, R., Ielo, L., Urban, E., Holzer, W. & Pace, V. Substituted α-Sulfur Methyl Carbanions: Effective Homologating Agents for the Chemoselective Preparation of β-Oxo Thioethers from Weinreb Amides. *European Journal of Organic Chemistry* **2018**, 2466–2470 (2018).

41. Senatore, R., Castoldi, L., Ielo, L., Holzer, W. & Pace, V. Expeditious and Chemoselective Synthesis of α-Aryl and α-Alkyl Selenomethylketones via Homologation Chemistry. *Organic Letters* **20**, 2685–2688 (2018).

42. Pace, V. *et al.* Bromomethyllithium-mediated chemoselective homologation of disulfides to dithioacetals. *Chemical Communications* **52**, 2639–2642 (2016).

43. Parisi, G. *et al.* Exploiting a 'beast' in carbenoid chemistry: Development of a straightforward direct nucleophilic fluoromethylation strategy. *Journal of the American Chemical Society* **139**, 13648–13651 (2017).

44. Monticelli, S. *et al.* Modular and Chemoselective Strategy for the Direct Access to α-Fluoroepoxides and Aziridines via the Addition of Fluoroiodomethyllithium to Carbonyl-Like Compounds. *Organic Letters* **21**, 584–588 (2019).

45. Degennaro, L., Trinchera, P. & Luisi, R. Recent advances in the stereoselective synthesis of aziridines. *Chemical Reviews* **114**, 7881–7929 (2014).

46. Verniest, G., Colpaert, F., Van Hende, E. & De Kimpe, N. Synthesis and reactivity of 1-substituted 2-fluoro- and 2,2-difluoroaziridines. *Journal of Organic Chemistry* **72**, 8569–8572 (2007).

PUBLICATIONS

- **An overview of recent Molecular Dynamics applications as medicinal chemistry tool for undruggable sites challenge.**
  U. Perricone, M.R. Gulotta, **J. Lombino**, B. Parrino, S. Cascioferro, P. Diana, G. Cirrincione, A. Padova *MedChemComm* **9**, 920–936 (2018)..

- **In Silico Insights towards the Identification of NLRP3 Druggable Hot Spots.**
  N. Mekni, M. De Rosa, C. Cipollina, M.R. Gulotta, G. De Simone, **J. Lombino**, A. Padova, U. Perricone. *International Journal of Molecular Sciences* **20**, 4974 (2019).

- **A New Family of Jumonji C Domain-Containing KDM Inhibitors Inspired by Natural Product Purpurogallin.**
  J.A Souto, F. Sarno, A. Nebbioso, C. Papulino, R. Alvarez, J. Lombino, U. Perricone, A. Padova, L. Altucci, A. de Lera. *Frontiers in Chemistry* **8**, 1–20 (2020).

- **Dynamic-shared pharmacophore approach as tool to design new allosteric PRC inhibitors, targeting EED binding pocket.**
  J. Lombino, M.R Gulotta, G. De Simone, N. Mekni, M. De Rosa, D. Carbone, B. Parrino, S.M Cascioferro, P. Diana, A. Padova, U. Perricone *Molecular Informatics* (2020). doi:10.1002/minf.202000148

- **Targeting SARS-CoV-2 RBD Interface: a Supervised Computational Data-Driven Approach to Identify Potential Modulators.**
  M.R Gulotta*, **J. Lombino***, U. Perricone, G. De Simone, N. Mekni, M. De Rosa, P. Diana, A. Padova. *ChemMedChem* **15**, 1921–1931 (2020).

- **Halogen-Imparted Reactivity in Lithium Carbenoid Mediated Homologations of Imine Surrogates: Direct Assembly of bis-Trifluoromethyl-β-Diketiminates and the Dual Role of LiCH2I**
  L. Ielo, L. Castoldi, S. Touqeer, J. Lombino, A. Roller, C. Prandi, W. Holzer, V. Pace *Angewandte Chemie* **132**, 21038–21043 (2020).

CONFERENCE ABSTRACT AND POSTER PRESENTATION

- **PRC2 allosteric modulation an alternative strategy in drug discovery for epigenetic disease**
  J. Lombino, S. Cascio, M.R Gulotta, U.Perricone, S. Cascioferro, B. Parrino, P. Diana, A. Padova
  . "MedChemSicily 2018" Università degli studi di Palermo, Palermo 17-20 luglio 2018.

- **A Molecular Dynamics-based virtual screening framework for the identification of EED inhibitors in the PRC2 modulation**
  J. Lombino, M.R Gulotta, U. Perricone, P.Diana, A. Padova
  "European Workshop in Drug Design" 2019, 19-24 Maggio, 2019, Siena, Italy

- **KDM4A inhibitors screening: case study for biophysical screening setup for the drug discovery process**
  R. Vallone, M.A Morando, R. Sabbatella, F. Sarno, **J. Lombino**, M. De Rosa, C. Cipollina, U. Perricone, C. Alfano.
  "Organ insuffiency change it or fix it, 13th Ri.MED scientific symposium". 25 Ottobre 2019. Palermo, Italia.

- **A computational study to explore the molecular mechanisms behind the antiproliferative activity of Nortopsentin derivatives**
  M.R. Gulotta, **J. Lombino**, B. Parrino, S.M. Cascioferro, P. Diana, A. Padova, G. Cirrincione, U. Perricone. Presentato al MedChemSicily 2018, 17-20 luglio 2018, Palermo, Italia

- **A computational structure-based approach to address hit identification modulating TOM20/α-Synuclein Interaction**
  M.R. Gulotta, R. Di Maio, **J. Lombino**, G. De Simone, A. Padova, G. Cirrincione, U. Perricone. Presentato all'EMBO Practical Course "Computational analysis of protein-protein interactions: Sequences, networks and diseases", 5-10 novembre 2018, Roma, Italia

- **A combined pharmacophore and MD approach to design putative modulators of a protein-protein interaction involved in Parkinson's Disease**
  M.R. Gulotta, R. Di Maio, N. Mekni, **J. Lombino**, G. De Simone, P. Diana, U. Perricone. Presentato al XII European Workshop in Drug Design 2019, 19-24 maggio 2019, Siena, Italia

- **A structure-based computational approach to design protein-protein inhibitors for Parkinson's Disease**
  M.R. Gulotta, N. Mekni, **J. Lombino**, M. De Rosa, U. Perricone. Presentato al Ri.MED Research Retreat 2019, 23 ottobre 2019, Palermo, Italia

Acknowledgments

The writing of this dissertation is the final round of a long process of education. Since it required high effort and commitment, I would like to thank everyone who helped me and supported me in the last three years.

Thanks to the RiMED foundation to give me the opportunity to work in their great team, especially to the medicinal chemistry group, the biophysics group and the biologics group, without whom would not be possible carry on my research project. A special thanks to the molecular informatic team, Ugo, Maria Rita, Giada and Nedra, with whom I shared many beautiful moments during these years.

I would like to thank Professor Vittorio Pace and his research group at University of Vienna, especially my colleague and friend Laura Ielo for inspiring me the interest for organometallic chemistry, for everything you taught me and for your precious friendship.

I also would like to thank my supervisor Professor Patrizia Diana and the doctors Barbara Parrino and Stella Cascioferro for their support and encouragement.

And last but not least, a special thanks to my whole family, my friends and especially to Francesco for having supported and endured during every moment of these years. Your love and your support gave me the chance to follow my passion and the strength to overcame difficulties.

This PhD has been a special journey, thanks to you.