UNIVERSITÀ DEGLI STUDI DI SALERNO

DIPARTIMENTO DI SCIENZE FARMACEUTICHE
E BIOMEDICHE

Dottorato di ricerca in SCIENZE FARMAUCETICHE
X Ciclo – Nuova Serie

Design, virtual screening and structural studies of new molecules with potential antitumor and antiinflammatory activity

Tutor
Ch.mo Prof. Giuseppe Bifulco

Dottoranda
Maria Giovanna Chini

Coordinatore
Ch.ma Prof.ssa Nunziatina De Tommasi

Dedicated to the memory of my Father

Preface

My PhD three years course in Pharmaceutical Sciences at the Department of Pharmaceutical and Biomedical Sciences of Salerno University was started in November 2008 under the supervision of Prof. Giuseppe Bifulco.

My research activity was mainly focused onto study of ligand-receptor interactions and structural characterization by computational techniques and NMR spectroscopy in order to identify new antitumor and/or antiinflammatory molecules potentially utilizable in therapy.

These approaches were successfully applied to the characterization of novel inhibitors of Histone deacetilase (HDAC) Nicotinamide Phosphoribosyltransferase (NMPRTase or Nampt), microsomal prostaglandin E_2 synthase (mPGES-1), human synovial Phospholipases A_2, ($hsPLA_2$), human Farnesoid-X-Receptor (FXR), and agonist of human Pregnane-X-Receptor (PXR) and Bile Acid Receptor GPBAR-1 (TGR5).

The entire work was carried out under the direct supervision of Prof. Giuseppe Bifulco.

Furthermore, to improve my knowledge on methodologies for the stereostructural assignment, I moved to the Department of Chemistry of the Bristol University in 2011 (mid-July until mid-November 2011) under the supervision of the Dr. Craig Butts.

During this period in his research laboratory, my research work has included learning and conducting advanced NMR spectroscopic investigations of a number of natural products and synthetic compounds.

In addition to PhD course activities, I was involved in different side projects, mainly regarding the characterization of ligand-targets interactions of ligands on targets involved in other pathologies as e.g. Alzheimer and obesity.

List of publications related to the scientific activity performed during the three years PhD course in Pharmaceutical Sciences:

1. Terracciano, S.; Chini, M. G.; Riccio, R.; Bruno, I.; Bifulco, G. Design, Synthesis, and Biological Activity of Hydroxamic Tertiary Amines as Histone Deacetylase Inhibitors. *ChemMedChem* **2012**, DOI: 10.1002/cmdc.201100531.

2. Chini, M. G.; Jones, C. R.; Zampella, A.; D'Auria, M. V.; Renga, B.; Fiorucci, S.; Butts C. P.; and Bifulco, G. Quantitative NMR-Derived Interproton Distances Combined with Quantum Mechanical Calculations of ^{13}C Chemical Shifts in the Stereochemical Determination of Conicasterol F, a Nuclear Receptor Ligand from *Theonella swinhoei*. *J. Org. Chem.* **2012** dx.doi.org/10.1021/jo2023763.

3. De Marino, S.; Sepe, V.; D'Auria, M. V.; Chini, M. G.; Bifulco, G.; D'Amore, C.; Renga, B.; Petek, S.; Fiorucci, S.; Zampella, A. 4-methylenesterols from *Theonella swinhoei* sponge are natural PXR agonists and FXR antagonist that modulate liver detoxification genes and innate immunity. *Steroids* **2012**, doi:10.1016/j.steroids.2012.01.006

4. Renga, B; Mencarelli, A; D'Amore, C.; Cipriani, S.; D'Auria, M. V. Sepe, V.; Chini, M. G.; Bifulco, G.; Zampella, A.; Fiorucci. S.; Discovery of Theonellasterol as a highly selective FXR antagonis that protects against liver injury in cholestasis. *PLoS ONE* **2012**, *7*, e30443, 1-12.

5. Sepe, V.; Ummarino, R.; D'Auria, M. V.; Chini, M. G.; Bifulco, G.; Renga, B.; D'Amore, C.; Debitus, C.; Fiorucci, S.; Zampella, A. Conicasterol E, a small heterodimer partner sparing farnesoid-X-receptor modulator endowed with a pregnane-X-receptor agonistic activity, from the marine sponge *Theonella swinhoei. J Med Chem.* **2012**, *55*, 84-93.

6. Cipriani, S.; Mencarelli, A.; Chini, M. G.; Distrutti, E.; Renga, B.; Bifulco, G.; Baldelli, F.; Donidi, A.; Fiorucci, S. The Bile Acid Receptor GPBAR-1 (TGR5) Modulates Integrity of Intestinal Barrier and Immune Response to Experimental Colitis. *PloS One* **2011**, *6*, e25637, 1-11.

7. Monti, M. C.; Chini, M. G.; Margarucci, L.; Riccio, R.; Bifulco, G.; Casapullo, A. The Binding Mode of Cladocoran A to the Human Group IIA Phospholipase A_2. *Chembiochem* **2011**, *12*, 2686-2691.

8. De Marino, S.; Ummarino, R.; D'Auria, M. V.; Chini, M. G.; Bifulco, G.; Renga, B.; D'Amore, C.; Fiorucci, S.; Debitus, C.; Zampella, A. Theonellasterols and Conicasterols from *Theonella swinhoei*. Novel Marine Natural Ligands for Human Nuclear Receptors. *J. Med. Chem.* **2011**, *54*, 3065-3075.

9. De Simone, R.; Chini, M. G.; Bruno, I.; Riccio, R.; Mueller, D.; Werz, O.; Bifulco, G. Structure-Based Discovery of Inhibitors of Microsomal Prostaglandin E_2 Synthase-1, 5-Lipoxygenase and 5-Lipoxygenase-Activating Protein: Promising Hits for the Development of New Anti-inflammatory Agents. *J. Med. Chem.* **2011**, *54*, 1565-1575.

10. Chini, M. G.; Terracciano, S.; Riccio, R.; Bifulco, G.; Ciao, R.; Gaeta, C.; Troisi, F.; Neri, P. Conformationally Locked Calixarene-Based Histone Deacetylase Inhibitors. *Organic Letters* **2010**, *12*, 5382-5385.

11. Terracciano, S.; Chini, M. G.; Bifulco, G.; D'Amico, E.; Marzocco, S.; Riccio, R.; Bruno, I. Synthesis of new mono and bis amides projected as potential histone deacetylase (HDAC) inhibitors. *Tetrahedron* **2010**, *66*, 2520-2528.

12. Di Micco, S; Chini, M. G.; Riccio, R.; Bifulco, G. Quantum Mechanical Calculation of NMR Parameters in the Stereostructural Determination of Natural Products. *Eur. J. Org. Chem.* **2010**, *8*, 1411-1434.

13. Colombano, G.; Travelli, C.; Galli, U.; Caldarelli, A.; Chini, M. G.; Canonico, P. L.; Sorba, G.; Bifulco, G.; Tron, G. C.; Genazzani, A. A. Novel Potent Nicotinamide Phosphoribosyltransferase Inhibitor Synthesized via Click Chemistry. *J. Med. Chem.* **2010**, *53*, 616-623.

14. Monti, M. C.; Chini, M. G.; Margarucci, L.; Tosco, A.; Riccio, R.; Bifulco, G.; Casapullo, A. The molecular mechanism of human group IIA phospholipase A_2 inactivation by bolinaquinone. *J. Mol. Recognit.* **2009** *22*, 530-537.

15. Grolla, A. A.; Podesta, V.; Chini, M. G.; Di Micco, S.; Vallario, A.; Genazzani, A. A.; Canonico, P. L.; Bifulco, G.; Tron, G. C.; Sorba, G.; Pirali, T. Synthesis, biological evaluation, and molecular docking of Ugi products containing a zinc-chelating moiety as novel inhibitors of histone deacetylases. *J. Med. Chem.* **2009**, *52*, 2776-2785.

16. Chini, M. G.; Riccio, R.; Bifulco, G. DFT/NMR integrated approach: a valid support to the total synthesis of chiral molecules. *Magn. Reson. Chem.* **2008**, *46*, 962-968

Table of Contents

Abstract

Computational methodologies in combination with experimental techniques as Nuclear Magnetic Resonance (NMR) have become a crucial component in drug discovery process, from hit identification to lead optimization.

The study of ligand-macromolecule interactions, in fact, has a crucial role for the design and the development of new and more powerful drugs. In this project, different aspects of interaction and recognition processes between ligand and macromolecule, and streostructure assignment has been studied through this kind of combined approach with the aim to identify novel potential antitumor and/or antiinflammatory molecules.

In particular, because the strong interconnection between the tumoral and inflammatory pathology has led to the identification of new target utilizable for the therapy, in this project will be described proteins (Histone deacetilase, HDAC; Nicotinamide Phosphoribosyltransferase, NMPRTase or Nampt; microsomal prostaglandin E_2 synthase, mPGES-1; human synovial Phospholipases A_2, hsPLA_2; human Farnesoid-X-Receptor, FXR; human Pregnane-X-Receptor, PXR; Bile Acid Receptor GPBAR-1, TGR5) involved in essential cellular processes and acting at diverse levels and phases of the tumor and inflammation diseases.

The results obtained can be summarized in three main areas of activity, whose relative weight was varied according to the development of the overall project:

a) <u>Support in the design of original scaffolds for the generation of libraries potentially utilizable in therapy.</u> This work was exclusively conducted *in silico* by a molecular docking technique in order to direct the design of the new molecules basing on the analysis of ligand-target

interactions and the synthetic possibilities. This kind of approach was successfully applied leading to the identification of new potential inhibitors for HDAC enzymes with ciclic (mono and bis amides, paragraph 2.2; conformationally locked calixarenes, paragraph 2.4), and linear (hydroxamic tertiary amines, paragraph 2.3) structures, and isoform selective (paragraph 2.6), and of ligands for microsomal prostaglandin E_2 synthase (mPGES)-1 (two series of triazole-based compounds; paragraphs 4.2 and 4.3).

For each of this described studied, the good qualitative accordance between the calculated and experimental data has made possible the identifications of new lead compounds, rationalizing in this way the key features to the target inhibition.

b) Rationalization of the biological activity of compounds by the study of the drug-receptor interactions. Molecular docking was used for the detailed study of antiinflammatory and antitumoral compounds whose activities are known a priori. In fact, thanks to this procedure, in this thesis several rationalizations of binding modes were reported related to Ugi products derivatives of CHAP 1 (HDAC inhibitors, paragraph 2.5), new and potent inhibitor of NMPRTAse analogs of FK866 and CHS 828 (chapter 3), marine natural products as inhibitors of hsPLA_2 (BLQ and CLDA, chapter 5), 4-methylen sterols extracted from *Theonella swinhoei* as ligands of FXR and PXR (chapter 6), and known compounds as taurolitholic acid and ciprofloxacin (chapter 7), agonists of TGR5.

Through the *in silico* methodology the putative binding modes for the reported molecules was described offering a complete rationalization of docking results, evaluating the influence of the ligand target interactions (e.g. hydrophobic, hydrophilic, electrostatic contacts) on the biological activity.

c) <u>Determination of relative configuration of natural products.</u>

The complete comprehension of the three dimensional structure of synthetic or isolated molecules is fundamental to design and characterize new platform potentially utilizable in therapy. On this basis, the combined approach between the quantum mechanical (QM) calculation of NMR parameters and NMR spectroscopy was revealed a very useful mean to lead the total synthesis of natural product toward the right isomer avoiding waste of time and resources (paragraph 8.1).

Moreover, the stereostructure assignment of marine natural products conicasterol F and its analog thonellasterol I was reported in the paragraph 8.2. by a novel combined approach between the quantitative interproton distance determinations by ROE and quantum mechanical calculations of chemical shifts

-INTRODUCTION-

-CHAPTER 1-

Cancer-related inflammation

1.1 Inflammation and cancer

The link between inflammation and cancers, rather than a recent concern, was noticed ~150 years ago. As early as 1863, Virchow indicated that cancers tended to occur at sites of chronic inflammation.[1]

Although it is now clear that proliferation of cells alone does not cause cancer, sustained cell proliferation in an environment rich in inflammatory cells, growth factors, activated stroma, and DNA-damage-promoting agents, certainly potentiates and/or promotes neoplastic risk.

During tissue injury associated with wounding, cell proliferation is enhanced while the tissue regenerates; proliferation and inflammation subside after the assaulting agent is removed or the repair completed. In contrast, proliferating cells that sustain DNA damage and/or mutagenic assault (for example, initiated cells) continue to proliferate in microenvironments rich in inflammatory cells and growth/survival factors that support their growth. In a sense, tumors act as wounds that fail to heal.[2]

Today, the causal relationship between inflammation, innate immunity and cancer is more widely accepted; however, many of the molecular and cellular mechanisms mediating this relationship remain unresolved. Furthermore, tumor cells may usurp key mechanisms by which inflammation interfaces with cancers, to further their colonization of the host. Moreover, it was clear that the acquired immune response to cancer is intimately related to the inflammatory response.[3,4]

Here, the critical points and the pathways connections between these two kinds of pathologies will be described.

1.1.1 Inflammation: From Acute to Chronic

Inflammation is a physiologic process in response to tissue damage resulting from microbial pathogen infection, chemical irritation, and/or wounding.[5] At the very early stage of inflammation, neutrophils are the first cells to migrate to the inflammatory sites under the regulation of molecules produced by rapidly responding macrophages and mast cells prestationed in tissues.[6] As the inflammation progresses, various types of leukocytes, lymphocytes, and other inflammatory cells are activated and attracted to the inflamed site by a signaling network involving a great number of growth factors, cytokines, and chemokines.[6] All cells recruited to the inflammatory site contribute to tissue breakdown and are beneficial by strengthening and maintaining the defense against infection.[6a]

There are also mechanisms to prevent inflammation response from lasting too long.[7] A shift from antibacterial tissue damage to tissue repair occurs, involving both proinflammatory and anti-inflammatory molecules.[7] Prostaglandin E2,[8] transforming growth factor-β,[9] and reactive oxygen and nitrogen intermediates[6d] are among those molecules with a dual role in both promoting and suppressing inflammation. The resolution of inflammation also requires a rapid programmed clearance of inflammatory cells: neighboring macrophages, dendritic cells, and backup phagocytes do this job by inducing apoptosis and conducting phagocytosis.[10] The phagocytosis of apoptotic cells also promotes an anti-inflammatory response, such as enhancing the production of antiinflammatory mediator transforming growth factor-β.[11] However, if inflammation resolution is dysregulated, cellular response changes to the pattern of chronic inflammation. In chronic inflammation, the inflammatory foci are dominated by lymphocytes, plasma cells, and

macrophages with varying morphology.[5] Macrophages and other inflammatory cells generate a great amount of growth factors, cytokines, and reactive oxygen and nitrogen species that may cause DNA damage.[6a] If the macrophages are activated persistently; they may lead to continuous tissue damage.[12] A microenvironment constituted by all the above elements inhabits the sustained cell proliferation induced by continued tissue damage, thus predisposes chronic inflammation to neoplasia.[1]

1.1.2 Cancer Development: An Overview

Cancer defines malignant neoplasms characterized by metastatic growth. It may occur in almost every organ and tissue relating to a variety of etiologic factors, such as genomic instability and environmental stress.[5] A two-stage carcinogenesis model is first conceptualized in a mouse model of skin cancer.[13] In this model, carcinogenesis is initiated by carcinogen-triggered irreversible genetic alteration and then promoted by dysregulated gene expression of initiated cells that resulted from epigenetic mechanisms and host-selective pressure.[6a] Once the proliferation advantage is obtained, cancer cells enter the progression stage in which their population expands rapidly.[6b] This model was subjected to criticism because it oversimplifies and failed to apply to all types of cancer.[14]

However, cancer development is still accepted as a multistep process, during which genetic alterations confer specific types of growth advantage; therefore, it drives the progressive transformation from normal cells to malignant cancer cells.[15] Malignant growth is characterized by several key changes: self-sufficiency of growth signals, insensitivity to antigrowth signals, escaping from apoptosis, unregulated proliferation potential, enhanced angiogenesis, and metastasis.[15] Each of these shifts is complicated and

accomplished by combined efforts of various signaling processes, and moreover it will find out that inflammation may contribute to the formation of these cancer phenotypes.

1.1.3 Connecting inflammation and cancer

Common wisdom says ‘‘most things in life are a double-edged sword’’. While they are in our favor at one dose or under one condition; they may be disfavor at another dose or under another condition. Inflammation is a part of the host response to either internal or external environmental stimuli. This response serves to counteract the insult incurred by these stimuli to the host. This response can be pyrogenic, as indicated by fever. When acute inflammation or fever is manifested for a short period of time, it has a therapeutic consequence. However, when inflammation becomes chronic or lasts too long, it can prove harmful and may lead to disease. How is inflammation diagnosed and its biomarkers is not fully understood, however, the role of proinflammatory cytokines, chemokines, adhesion molecules and inflammatory enzymes have been linked with chronic inflammation (Figure 1.1). Chronic inflammation has been found to mediate a wide variety of diseases, including cardiovascular diseases, cancer, diabetes, arthritis, Alzheimer’s disease, pulmonary diseases, and autoimmune diseases.[16] Chronic inflammation has been linked to various steps involved in tumorigenesis, including cellular transformation, promotion, survival, proliferation, invasion, angiogenesis, and metastasis.[17,18] That inflammation is a risk factor formost type of cancers is now well recognized.[19]

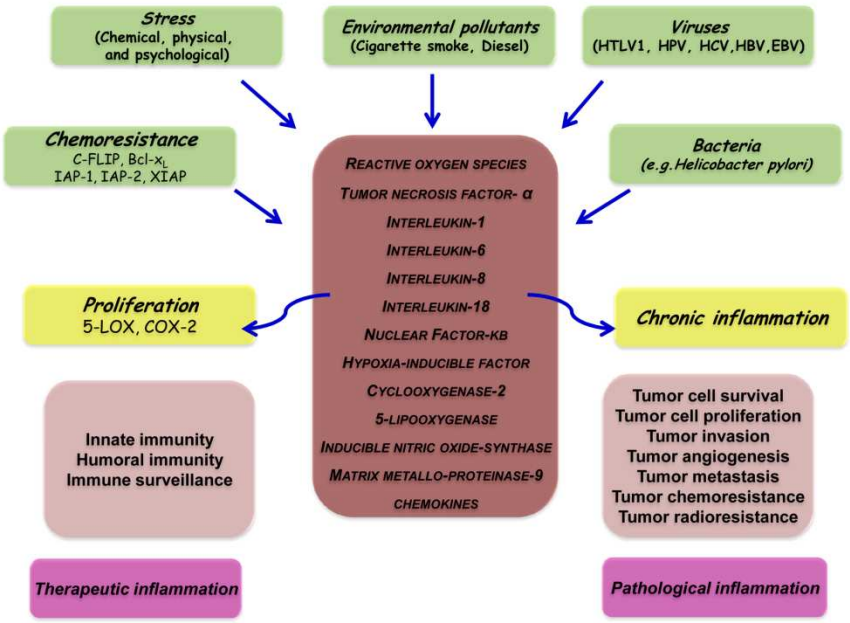


Figure 1. 1 Different faces of inflammation and its role in tumorigenesis.

Al already reported, the links between cancer and inflammation were first made in the nineteenth century, on the basis of observations that tumors often arose at sites of chronic inflammation and that inflammatory cells were present in biopsied samples from tumors,[1] but there has been a recent resurgence in interest.

Several lines of evidence[20] (Table 1.1) − based on a range of findings, from epidemiological studies of patients to molecular studies of genetically modified mice − have led to a general acceptance that inflammation and cancer are linked. Epidemiological studies have shown that chronic inflammation predisposes individuals to various types of cancer. It is estimated that underlying infections and inflammatory responses are linked to 15‑20% of all deaths from cancer worldwide.[1] There are many triggers of chronic inflammation that increase the risk of developing cancer. Such triggers include

microbial infections (for example, infection with *Helicobacter pylori* is associated with gastric cancer and gastric mucosal lymphoma), autoimmune diseases (for example, inflammatory bowel disease is associated with colon cancer) and inflammatory conditions of unknown origin (for example, prostatitis is associated with prostate cancer). Accordingly, treatment with non-steroidal anti-inflammatory agents decreases the incidence of, and the mortality that results from, several tumor types.[21]

Table 1. 1 The evidence that links cancer and inflammation

1	Inflammatory diseases increase the risk of developing many types of cancer (including bladder, cervical, gastric, intestinal, oesophageal, ovarian, prostate and thyroid cancer)
2	Non-steroidal anti-inflammatory drugs reduce the risk of developing certain cancers (such as colon and breast cancer) and reduce the mortality caused by these cancers.
3	Signaling pathways involved in inflammation operate downstream of oncogenic mutations (such as mutations in the genes encoding RAS, MYC and RET).
4	Inflammatory cells, chemokines, and cytokines are present in the microenvironment of all tumors in experimental animal models and humans from the earliest stages of development.
5	The targeting of inflammatory mediators (chemokines and cytokines, such as TNF-α and IL-1β), key transcription factors involved in inflammation (such as NF-κB and STAT3) or inflammatory cells decreases the incidence and spread of cancer.
6	Adoptive transfer of inflammatory cells or overexpression of inflammatory cytokines promotes the development of tumors.

The hallmarks of cancer-related inflammation include the presence of inflammatory cells and inflammatory mediators (for example, chemokines, cytokines and prostaglandins) in tumor tissues, tissue remodeling and

angiogenesis similar to that seen in chronic inflammatory responses, and tissue repair. These signs of 'smouldering' inflammation[20a] are also present in tumors for which a firm causal relationship to inflammation has not been established (for example, breast tumors). Indeed, inflammatory cells and mediators are present in the microenvironment of most, if not all, tumors, irrespective of the trigger for development.

In the tumor microenvironment, inflammatory cells and molecules influence almost every aspect of cancer progress, including the tumor cells' ability to metastasize.[22] Thus, whereas there were previously six recognized hallmarks of cancer − unlimited replicative potential, self-sufficiency in growth signals, insensitivity to growth inhibitors, evasion of programmed cell death, ability to develop blood vessels, and tissue invasion and metastasis[23] − cancer related inflammation now emerges as number seven (Figure 1.2). In 2000, Hanahan and Weinberg[23] proposed a model to define the six properties that a tumor acquires.

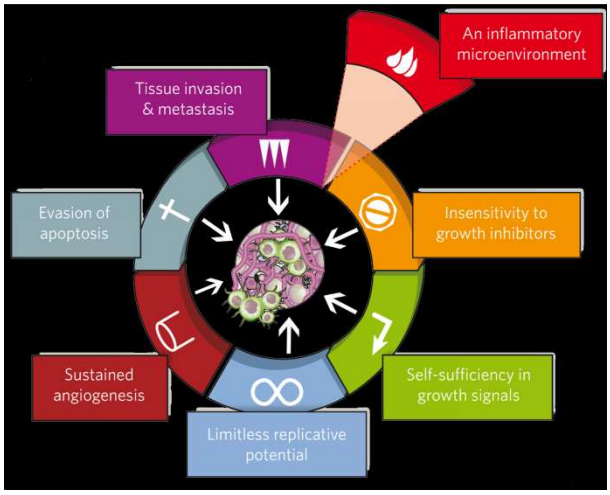


Figure 1. 2 The hallmarks of cancer.

These are unlimited replicative potential, ability to develop blood vessels (angiogenesis), evasion of programmed cell death (apoptosis), self-sufficiency in growth signals, insensitivity to inhibitors of growth, and tissue invasion and metastasis. Kim and colleagues' findings,[24] together with those of other studies,[22,18] indicate that this model should be revised to include cancer-related inflammation as an additional hallmark.[23]

The connection between inflammation and cancer can be viewed as consisting of two pathways: an extrinsic pathway, driven by inflammatory conditions that increase cancer risk (such as inflammatory bowel disease); and an intrinsic pathway, driven by genetic alterations that cause inflammation and neoplasia (such as oncogenes) (Figure 1.3).

The intrinsic pathway was uncovered when addressing why inflammatory cells and mediators are present in the microenvironment of most, if not all, tumors and therefore are present in cases for which there is no epidemiological basis for inflammation. This finding raised the question of whether the genetic events that cause neoplasia in these cases are responsible for generating an inflammatory environment. This question has been addressed only recently, by using preclinical and clinical settings in which various oncogenetic mechanisms can be assessed.

The intrinsic pathway is activated by genetic events that cause neoplasia. These events include the activation of various types of oncogene by mutation, chromosomal rearrangement or amplification, and the inactivation of tumor-suppressor genes. Cells that are transformed in this manner produce inflammatory mediators, thereby generating an inflammatory microenvironment in tumors for which there is no underlying inflammatory condition (for example, breast tumors). By contrast, in the extrinsic pathway,

inflammatory or infectious conditions augment the risk of developing cancer at certain anatomical sites (for example, the colon, prostate and pancreas).

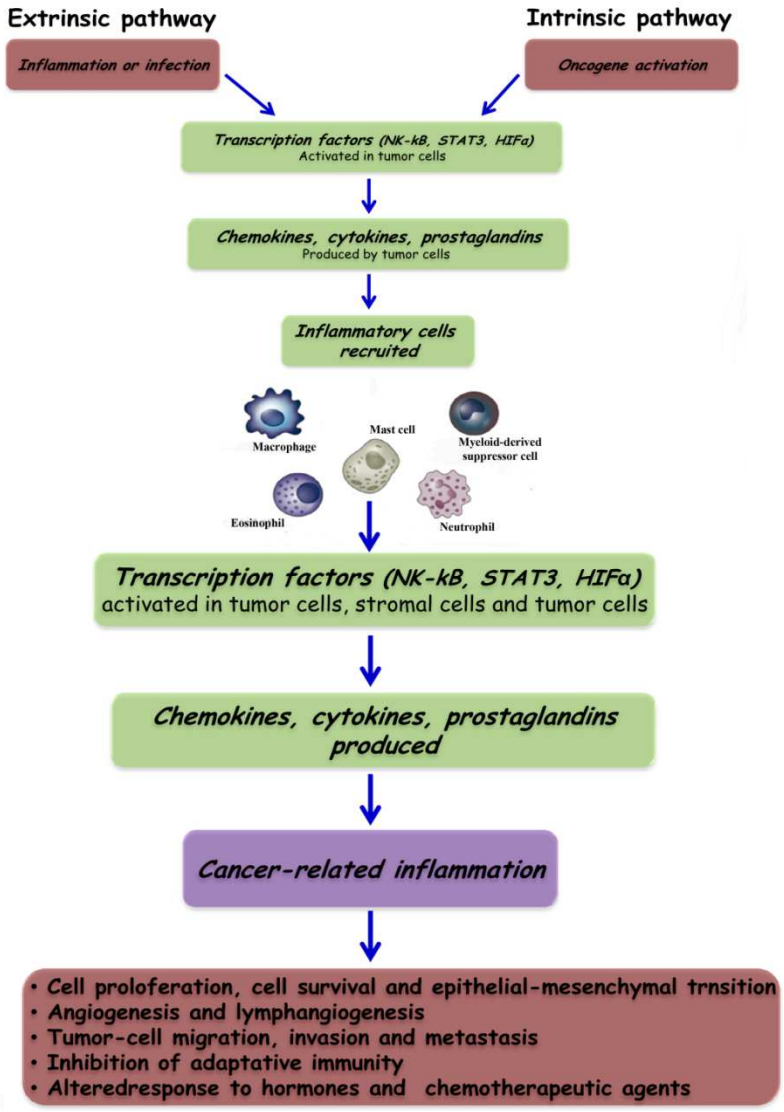


Figure 1. 3 Pathways that connect inflammation and cancer. Cancer and inflammation are connected by two pathways: the and the intrinsic extrinsic pathway.

The two pathways converge, resulting in the activation of transcription factors, mainly nuclear factor-kB (NF-kB), signal transducer and activator of transcription 3 (STAT3) and hypoxia-inducible factor 1 (HIF1α), in tumor cells. These transcription factors coordinate the production of inflammatory mediators, including cytokines and chemokines, as well as the production of cyclooxygenase 2 (COX2) (which, in turn, results in the production of prostaglandins). These factors recruit and activate various leukocytes, most notably cells of the myelomonocytic lineage. The cytokines activate the same key transcription factors in inflammatory cells, stromal cells and tumor cells, resulting in more inflammatory mediators being produced and a cancer-related inflammatory microenvironment being generated. Smouldering cancer-related inflammation has many tumor-promoting effects.

1.1.3.1 Mutagenic Potential of Inflammation

The chronic inflammation microenvironment is predominated by macrophages.[6] Those macrophages, together with other leukocytes, generate high levels of reactive oxygen and nitrogen species to fight infection.[25] However, in a setting of continuous tissue damage and cellular proliferation, the persistence of these infection-fighting agents is deleterious.[6b] They may produce mutagenic agents, such as peroxynitrite, which react with DNA and cause mutations in proliferating epithelial and stroma cells.[25,26] Macrophages and T lymphocytes may release tumor necrosis factor-α (TNF-α) and macrophage migration inhibitory factor to exacerbate DNA damage.[27] Migration inhibitory factor impairs p53-dependent protective responses, thus causing the accumulation of oncogenic mutations.[28] Migration inhibitory factor also contributes to tumorigenesis by interfering Rb-E2F pathway.[29]

Within an ileocolitis-associated mouse cancer model, the high susceptibility to inflammation and cancer in hydroperoxide-reducing enzyme-deficient mice suggested that intracellular hydroperoxides might also contribute to tumor initiation.[30]

1.1.3.2 Role of Inflammatory Cells in Tumor Development

Other than a single mutation, more genetic and epigenetic events are required to drive from initiated cells to malignant tumors.[23] Some of these events are also found to be related to chronic inflammation. For instance, angiogenesis, a critical process in tumor progression,[31] associates with chronic inflammation, such as psoriasis, rheumatoid arthritis, and fibrosis.[23] In addition, the tumor inflammatory microenvironment can facilitate the breakage of the basement membrane, a process required for the invasion and migration of tumor cells.[6a] A wide population of leukocytes and other types of immune cells infiltrate to the developing tumor site and establish the tumor inflammatory microenvironment.[6c] Macrophages, neutrophils, eosinophils, dendritic cells, mast cells, and lymphocytes are also found to be key components in the epithelial-originated tumors.[6c,12,32] The infiltration of immune cells to tumors may repress tumor growth.[33] However, the increasing concern is that inflammatory cells act as tumor promoters in inflammation-associated cancers.[6a,34,35] Accumulated mutations in epithelial cells lead to dysregulation of their growth and migration. These dysregulated epithelial cells may also signal to recruit leukocytes.[31] In addition, tumor cells may also produce cytokines and chemokines to attract immune cells to facilitate cancer development.[6a,c,31]

1.1.3.3 Key Molecular Players in Linking Inflammation to Cancer

To address the details of transition from inflammation to cancers and the further development of inflammation-associated cancers, it is necessary to investigate specific roles of key regulatory molecules involved in this process. In fact, in the panoply of molecules involved in cancer-related inflammation, key endogenous (intrinsic) factors can be identified. These include transcription factors (such as NF-kB and signal transducer and activator of transcription 3 (STAT3)) and major inflammatory cytokines (such as IL-1β, IL-6, IL-23 and TNF-α)[36,37,38] (Table 1.2).

Table 1. 2 Key Molecular Players Linking Cancer to Inflammation.

Potential linkers	**Functions in linking inflammation to cancer**
Cytokines	
IL-6	Promote tumor growth
TNF-α	Induce DNA damage and inhibit DNA repair Promote tumor growth Induce angiogenic factors
Chemokines	Promote tumor cell growth Facilitate invasion and metastasis by directing tumor cell migration and promoting basement membrane degradation
NF-κβ	Mediate inflammation progress, promoting chronic inflammation Promote the production of mutagenic reactive oxygen species Protect transformed cells from apoptosis Promote tumor invasion and metastasis Feedback loop between proinflammatory cytokines
iNOS	Downstream of NF-nB and proinflammatory cytokines

	Induce DNA damage and disrupt DNA damage response Regulate angiogenesis and metastasis
COX-2	Produce inflammation mediator prostaglandins Promote cell proliferation, antiapoptotic activity, angiogenesis, and metastasis
HIF-1α	Promote chronic inflammation Induced by proinflammatory cytokines through NF-nB Enhance the glycolytic activity of cancer cells Contribute to angiogenesis, tumor invasion, and metastasis by transactivating VEGF
STAT3	Activated by proinflammatory cytokines Promote proliferation, apoptosis resistance, and immune tolerance
Nrf2	Anti-inflammatory activity Protect against DNA damage
NFAT	Regulate proinflammatory cytokine expression Required in cell transformation

For sick of simplicity, between the molecular players involved in inflammatory networking cancer, the tumor necrosis factor (TNF-α) and NF-kB will be described. The TNF-α was first isolated as an anticancer cytokine than two decades ago.[39] Experience since then has indicated that when expressed locally by the cells of the immune system, TNF-α has a therapeutic role. However, when dysregulated and secreted in the circulation, TNF-a can mediate a wide variety of diseases, including cancer.[39] TNF-α has itself been shown to be one of the major mediators of inflammation.[40] Induced by a wide range of pathogenic stimuli, TNF-α induces other inflammatory mediators and proteases that orchestrate inflammatory responses. TNF-α is also produced by tumors and can act as an endogenous tumor promoter.[40] The role of TNF-α

has been linked to all steps involved in tumorigenesis, including cellular transformation, promotion, survival, proliferation, invasion, angiogenesis, and metastasis, as outlined below (Figure 1.4).

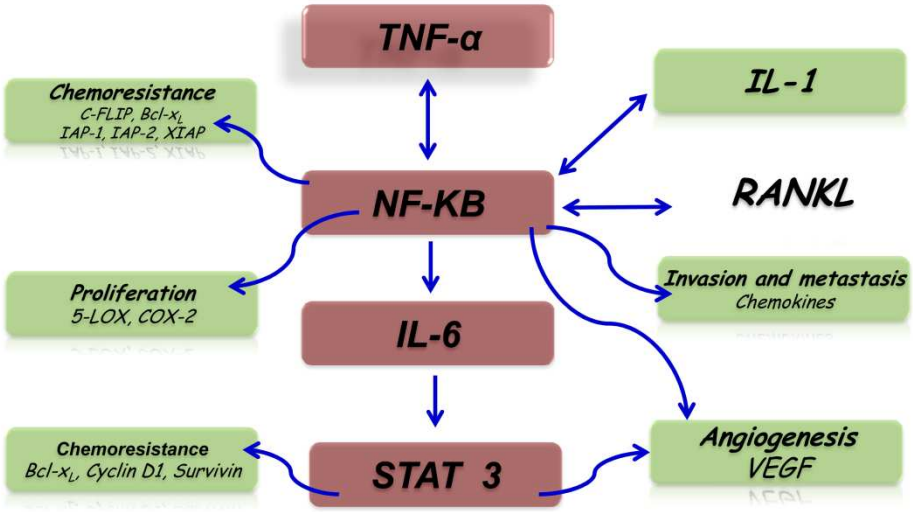


Figure 1. 4 Inflammatory networking in cancer.

On the other hand, NF-kB is a key coordinator of innate immunity and inflammation, and has emerged as an important endogenous tumor promoter.[36] NF-kB is crucial both in the context of tumor or potential tumor cells and in the context of inflammatory cells. In these cell types, NF-κB operates downstream of the sensing of microorganisms or tissue damage by the Toll-like receptor (TLR)‑MyD88 signaling pathway, and by signaling pathways mediated by the inflammatory cytokines TNF-α and IL-1β. In addition, NF-kB can be activated as a result of cell-autonomous genetic alterations (amplification, mutations or deletions)[41] in tumor cells. In tumor cells and epithelial cells at risk of transformation by carcinogens, as well as in inflammatory cells, NF-kB activates the expression of genes encoding

inflammatory cytokines, adhesion molecules, enzymes in the prostaglandin-synthesis pathway (such as COX2), inducible nitric oxide synthase (iNOS; also known as NOS2) and angiogenic factors.

In addition, one of the important functions of NF-κB in tumor cells or cells targeted by carcinogenic agents is promoting cell survival, by inducing the expression of anti-apoptotic genes (such as BCL2). There is also accumulating evidence of interconnections and compensatory pathways between the NF-κB and HIF1α systems,[42] linking innate immunity to the response to hypoxia. There is unequivocal evidence that NF-κB is involved in tumor initiation and progression in tissues in which cancer-related inflammation typically occurs (such as the gastrointestinal tract and the liver).[43] The NF-κB pathway is tightly controlled by inhibitors that function at various stages of the pathway. An example is TIR8 (also known as SIGIRR), a member of the IL-1-receptor family. TIR8 has a single immuno globulin domain, a long cytoplasmic tail, and a Toll/IL-1 receptor (TIR) domain that differs from that of other members of the IL-1-receptor family. Deficiency in the gene that encodes TIR8 is associated with increased susceptibility to intestinal inflammation and carcinogenesis.[44] Thus, the balance of inhibitors and activators tunes the extent to which the NF-κB pathway operates as an endogenous tumor promoter. Support for the connection between cancer and inflammation is further strengthened by studies of the role of NF-κB in tumor-infiltrating leukocytes. In established, advanced tumors, which typically have a microenvironment of smouldering inflammation,[20] tumor-associated macrophages (TAMs) have delayed and defective NF-κB activation.[45] Evidence suggests that homodimers of the p50 subunit of NF-κB (a negative regulator of the NF-κB pathway) are responsible for this sluggish activation of NF-κB in TAMs and for the protumor phenotype of these cells.[46] Thus, NF-κB seems to function as a

'rheostat' whose function can be tuned to different levels, a property that enables the extent of inflammation to be regulated. Such regulation allows the vigorous inflammation (for example, in inflammatory bowel disease) that predisposes individuals towards developing cancer to be sustained, and enables TAMs to sustain the smouldering inflammatory microenvironment present in established metastatic neoplasia.

Briefly, the mediators and cellular effectors of inflammation are important constituents of the local environment of tumors.

1.2 Scope and outline of this thesis

The study of ligand-macromolecule interactions has a fundamental role for the design and the development of new and more powerful drugs. In this project, different aspects of interaction and recognition processes between ligand and macromolecule has been studied through a combined approach based on computational chemistry techniques and NMR spectroscopy. In particular, these different aspects regard the employment and elaboration of screening methods, the analysis of structural determinants responsible of drug-macromolecule complex formation and the design of new potential bioactive compounds. Several and different proteins, involved in essential cellular processes, have been investigated as biological targets taking into account their implication in tumor and inflammation initiation and progress with the aim to identify and rationalize new molecules potentially utilizable in therapy.

As already reported, in some types of cancer, inflammatory conditions are present before a malignant change occurs. Conversely, in other types of cancer, an oncogenic change induces an inflammatory microenvironment that promotes the development of tumors. Regardless of its origin, inflammation in the tumor microenvironment has many tumor-promoting effects. It aids in the

proliferation and survival of malignant cells, promotes angiogenesis and metastasis, subverts adaptive immune responses, and alters responses to hormones and chemotherapeutic agents. The molecular pathways of this cancer-related inflammation are now being unraveled, resulting in the identification of new target molecules that could lead to improved diagnosis and treatment. Between them, in this project, the attention was focused on targets (Histone deacetilase, HDAC; Nicotinamide Phosphoribosyltransferase, NMPRTase or Nampt; microsomal prostaglandin E_2 synthase, mPGES-1; human synovial Phospholipases A_2, hsPLA$_2$; human Farnesoid-X-Receptor, FXR; human Pregnane-X-Receptor, PXR; Bile Acid Receptor GPBAR-1, TGR5) with different mechanisms of action involved in diverse levels and phases of tumor and inflammation process.

In particular, in the chapter 2 the results obtained for the design, in silico screening, and rationalization of binding modes of pan, selective cyclic and linear HDAC inhibitors are summarized.

In chapter 3 the analysis at atomic level of the interactions between NMPRTase and triazole-based analogs of APO866 and CHS2883 are reported.

Although some HDAC inhibitors are already showing therapeutic utility in animal models of inflammatory diseases (such as arthritis, inflammatory bowel diseases, septic shock, ischemia-reperfusion injury, asthma, ecc.),[47] and NMPRTase is able to control both cell viability and the inflammatory response,[48] by regulating NAD availability, it is important to underline that for all the compounds described in the chapters 2 and 3 only their antitumor activity was evaluated.

The chapters 4 and 5 are related to the design, *in silico* screening and rationalization of binding modes of mPGES-1 and hsPLA$_2$ inhibitors, key enzymes of arachidonic acid cascade and eicosanoid metabolism. In particular,

in the chapter 4 the design and *in silico* evaluation of two series of triazole-based compounds is described in detail; while in the chapter 5, the putative binding modes of two marine natural products to human synovial Phospholipases A_2 were obtained through molecular docking.

However, even if mPGES-1[49] is becoming a target for cancer suppression thanks to its inhibitory ability to suppress the PGE_2 synthesis offering the potential for therapeutic benefit without the potential toxicity associated with COX-2 inhibition, in the chapter 4, only the potential antiinflammatory activities of designed molecules were evaluated. In the chapter 5, the antiinflammatory activity of reported compounds was analyzed in detail, although the group II Group IIA Phospholipase A_2 (PLA_2-IIA) also plays a role in tumor progression in vivo, and inhibitors of PLA_2-IIA suppress the proliferative activity and invasiveness of prostate carcinoma cell lines.[50]

The chapters 6 and 7 refer to nuclear receptors FXR, PXR and TGR5 ligands. In particular, in the chapter 6, a detail rationalization of the 4-Methylen sterols antagonist and agonist activity on FXR are PXR respectively was reported. On the other hand, in the chapter 7, agonist activity of two well-known ligands to TGR5 is described. For these compounds will be analyzed and described their activity in the controlling of the bile acids metabolism and their involvement with dysfunctions connected with it; in fact the described nuclear receptors (FXR, PXR and TGR5) are important pharmacological targets for a number of diseases, including cancer and metabolic disorders.[51]

Finally in the chapter 8, it will be described the use of calculation at quantum mechanical level of the NMR parameters (e.g. δ, chemical shifts, and *J* coupling constant), as an efficient tool in the stereostructure assignment of natural products.

1.3 Methodologies employed

Before starting the discussion concerning the results obtained, it is appropriate briefly introduce the methodologies utilized to realize the project.

1.3.1 Molecular docking

Computational methodologies have become a crucial component in drug discovery, from virtual screening for hit identification to lead compound optimization. One key methodology is the molecular docking that consists in the prediction of ligand conformation and orientation within a targeted binding site. The molecular docking is based on the requirement that the 3D structure of the macromolecule is known. Many different programs have been developed, of which DOCK,[52] FlexX,[53] GOLD,[54]120 Autodock,[55,56,57] Autodock Vina,[58] and Glide122[59] are among the most popular. The mentioned tools are based on a range of different concepts, and each comes with its own set of strengths and weaknesses. One feature most docking programs share, however, is that they position a flexible ligand into a rigid binding site. Computational feasibility is the main reason for utilizing a rigid macromolecule in the docking calculations, as the number of freedom degrees that have to be considered grows exponentially with the number of accessible receptor conformations. Most molecular docking software have two key parts: (1) a search algorithm and (2) a scoring function.[60] For molecular docking to be useful in drug discovery, these key parts should be both fast and accurate. These two requirements are often in opposition to each other, requiring necessary compromises that commonly end in ambiguous results or failure.[61]

The search algorithm samples different ligand orientations and conformations fitting the macromolecular binding site. This step is

complicated by the number of freedom degrees contained in the small molecule, increasing the conformational space to sample. The search methods can be grouped in three categories: systematic methods, random or stochastic methods, and simulation methods. The systematic search algorithms try to explore all the degrees of freedom in a molecule, but they face the problem of huge number of generated conformations.[62] The random methods (often called stochastic methods) operate by making random changes to either a single ligand or a population of ligands. A newly obtained ligand is evaluated on the basis of a pre-defined probability function. Two popular random approaches are Monte Carlo and genetic algorithms. About simulations search methods, molecular dynamics is currently the most popular approach. However, molecular dynamics simulations are often unable to cross high-energy barriers within feasible simulation time periods, and therefore might only accommodate ligands in local minima of the energy surface.[63] Therefore, an attempt is often made to simulate different parts of a protein–ligand system at different temperatures.[64] Another strategy for addressing the local minima problem is starting molecular dynamics calculations from different ligand positions. In contrast to molecular dynamics, energy minimization methods are rarely used as stand-alone search techniques, as only local energy minima can be reached, but often complement other search methods. The scoring function aims to evaluate the results of the search algorithm predicting the affinity for the biological target. This evaluation is very difficult because the binding process is governed by enthalpic and entropic factors and one or of them can predominate. Other elements can affect the scoring method, such as limited resolution of crystallographic targets, inherent flexibility, induced fit or other conformational changes that occur on binding and the participation of water molecules in macromolecule–ligand interactions. Three classes of scoring

functions are currently applied: force field-based, empirical and knowledge-based scoring functions.

Molecular mechanics force fields usually quantify the sum of two energies, the macromolecule–ligand interaction energy and internal ligand energy (such as steric strain induced by binding). Most force field scoring functions only consider a single protein conformation, which makes it possible to omit the calculation of internal protein energy, which greatly simplifies scoring. The enthalpic contribution are essentially given by the electrostatic and Van der Waals terms, and is some software (AutoDock, Gold) take into account the hydrogen bond formation between drug and biological target.

The van der Waals potential energy for the general treatment of non-bonded interactions is often modeled by a Lennard–Jones 12–6 function (Equation 1.1):

$$E_{VdW}(r) = \sum_{j=1}^{N} \sum_{i=1}^{N} 4\varepsilon \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right]$$

Equation 1. 1

where ε is the well depth of the potential and σ is the collision diameter of the respective atoms *i* and *j*. The exp(12) term of the equation is responsible for small-distance repulsion, whereas the exp(6) provides an attractive term which approaches zero as the distance between the two atoms increases.

The Lennard–Jones 12–6 function is also used to describe the hydrogen bond in macromolecule-ligand complex, but compared to the Van der Waals function, is less smooth and angle dependent.

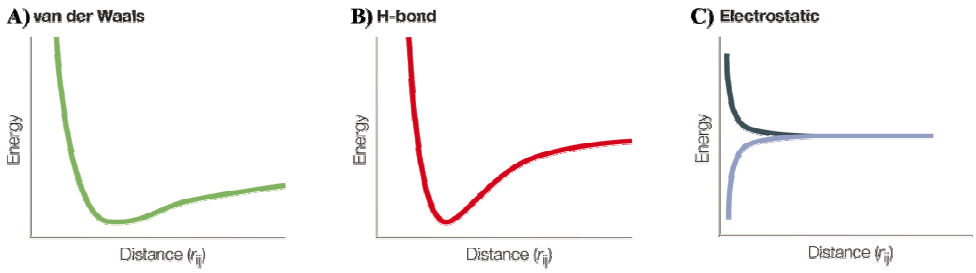


Figure 1. 5 Schematic representation of functions used to model pair-wise interactions that contribute to binding. Interactions are calculated as a function of the distance (*rij*) between two atoms *i* and *j*. a) van der Waals interaction given by a 12–6 Lennard–Jones potential (note the smoother attractive part of the potential compared to hydrogen bond term). B) hydrogen bond potential given by a 'harder' 12–10 Lennard–Jones potential. C) electrostatic potential for two like (blue) or opposite (black) charges of same magnitude calculated using a distance dependent dielectric constant.

The electrostatic potential energy is represented as a summation of Coulombic interactions, as described in equation 1.2:

$$E_{Coul}(r) = \sum_{j=1}^{N_A} \sum_{i=1}^{N_B} \frac{q_i \, q_j}{4\pi\varepsilon_0 \, r_{ij}}$$

Equation 1. 2

where N is the number of atoms in molecules A and B, respectively, and q the charge on each atom. The functional form of the internal ligand energy is typically very similar to the protein–ligand interaction energy, and also includes van der Waals contributions and/or electrostatic terms.

Empirical scoring functions work on the sum of several parameterized functions to reproduce experimental data. The design of empirical scoring functions is based on the idea that binding energies can be approximated by a sum of individual uncorrelated terms. The coefficients of the various terms are

obtained from regression analysis using experimentally determined binding energies and X-ray structural information.

By using the knowledge-based scoring functions protein–ligand complexes are modeled using relatively simple atomic interaction-pair potentials. A number of atom-type interactions are defined depending on their molecular environment.

1.3.1.2 Autodock: an Overview

There are numerous molecular docking software applications that utilize different searching and scoring algorithms and AutoDock is currently one of the most cited of these applications,[65] especially in a virtual screening of a compound libraries.[66] For the purposes of this project the software AutoDock 3.0.5,[55] 4.1,[56] 4.2,[57] and AutodockVina[58] have been used, where the differences between them are related to the speed, macromolecule sidechains flexibility, optimization of the free-energy scoring function based on a linear regression analysis, AMBER force field, larger set of diverse protein-ligand complexes with known inhibition constants; moreover the Lamarckian Genetic Algorithm (LGA) is a big improvement on the Genetic Algorithm, and both genetic methods are much more efficient and robust than SA in the new version of the software.

The best model obtained with the latest version AutoDock 4.2[57] in fact, was cross-validated with a separate set of HIV-1 protease complexes, and confirmed that the standard error is around 2.5 kcal/mol.

In AutoDock there are different available search methods, but the Lamarckian Genetic Algorithm (LGA) has been selected for the aim of this study, because it has demonstrated to give the best results compared to the other algorithms.[55]

The vast majority of genetic algorithms mimic the major characteristics of Darwinian evolution and apply Mendelian genetics. This is illustrated in Figure 1.7 note the oneway transfer of information from the genotype to the phenotype. However, in those cases where an inverse mapping function exists (i.e., one which yields a genotype from a given phenotype), it is possible to finish a local search by replacing the individual with the result of the local search; see the left-hand side of Figure 1.6.

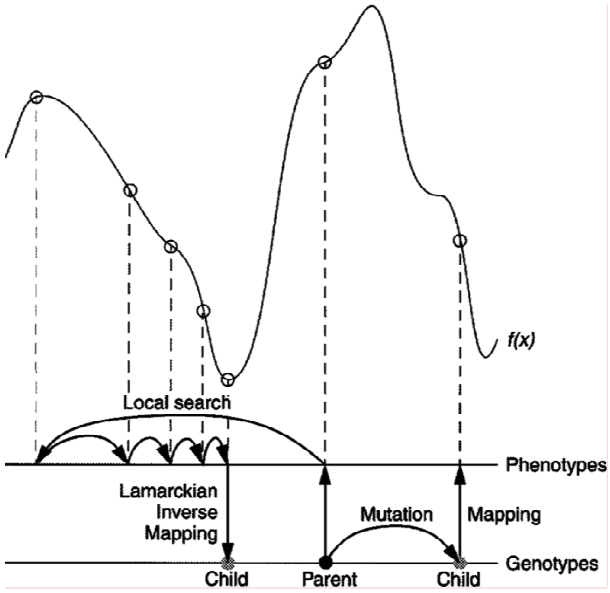


Figure 1. 6 This figure illustrates genotypic and phenotypic search, and contrasts Darwinian and Lamarckian search.[67] The space of the genotypes is represented by the lower horizontal line, and the space of the phenotypes is represented by the upper horizontal line. Genotypes are mapped to phenotypes by a developmental mapping function. The fitness function is $f(x)$. The result of applying the genotypic mutation operator to the parent's genotype is shown on the right-hand side of the diagram, and has the corresponding phenotype shown. Local search is shown on the left-hand side. It is normally performed in phenotypic space and employs information about the fitness landscape. Sufficient iterations of the local search arrive at a local minimum, and an inverse mapping function is used to convert from its phenotype to its

corresponding genotype. In the case of molecular docking, however, local search is performed by continuously converting from the genotype to the phenotype, so inverse mapping is not required. The genotype of the parent is replaced by the resulting genotype, however, in accordance with Lamarckian principles.

This is called the Lamarckian genetic algorithm (LGA), and is an allusion to Jean Batiste de Lamarck's (discredited) assertion that phenotypic characteristics acquired during an individual's lifetime can become heritable traits.[68]

The most important issues arising in hybrids (LGA) of Genetic Algorithm (GA) and the Local Search (LS) revolve around the *developmental mapping*, which transforms genotypic representations into phenotypic ones.

The genotypic space is defined in terms of the genetic operators mutation and crossover in our experiments by which parents of one generation are perturbed to form their children. The phenotypic space is defined directly by the problem, namely, the energy function being optimized. The local search operator is a useful extension of GA global optimization when there are local ''smooth-ness'' characteristics (continuity, correlation, etc.) of the fitness function that local search can exploit. In hybrid GA + LS optimizations, the result of the LS is always used to update the fitness associated with an individual in the GA selection algorithm. If, and only if, the developmental mapping function is *invertible*, will the Lamarckian option converting the phenotypic result of LS back into its corresponding genotype become possible. The fitness or energy is usually calculated from the ligand's coordinates, which together form its phenotype. The developmental mapping simply transforms a molecule's genotypic state variables into the corresponding set of atomic coordinates. A novel feature of this application of hybrid global-local

optimization is that the Solis and Wets LS operator searches through the genotypic space rather than the more typical phenotypic space. This means that the developmental mapping does not need to be inverted. Nonetheless, this molecular variation of the genetic algorithm still qualifies as Lamarckian, because any ‘‘environmental adaptations’’ of the ligand acquired during the local search will be inherited by its offspring. At each generation, it is possible to let a user defined fraction of the population undergo such a local search. The local search frequencies of just 0.06 have found improved efficiency of docking, although a frequency of 1.00 is not significantly more efficient.[67] Both the canonical and a slightly modified version of the Solis and Wets method have been implemented. In canonical Solis and Wets, the same step size would be used for every gene, but we have improved the local search efficiency by allowing the step size to be different for each type of gene: a change of 1 Å in a translation gene could be much more significant than a change of 1° in a rotational or torsional gene. In the Lamarckian genetic algorithm, genotypic mutation plays a somewhat different role than it does in traditional genetic algorithms. Traditionally, mutation plays the role of a local search operator, allowing small, refining moves that are not efficiently made by crossover and selection alone. With the explicit local search operator, however, this role becomes unnecessary, and is needed only for its role in replacing alleles that might have disappeared through selection. In LGA, mutation can take on a more exploratory role.

The LGA yields a maximum number of 256 potential bioactive conformations: run, whose number can be increased performing more docking calculations. Each conformational solution is the result of a selection. The GA, starting from the input geometry, gives rise to a group of n conformations or individuals (whose number can be set up) defining for them translational,

rotational and torsional variables. By the scoring function, each individual is labeled by the total interaction energy: fitness. Random pairs of individuals are mated using a process of crossover, in which new individuals inherit geometrical features from their parents leading to new generation of individuals. In addition, some offspring undergo random mutation, in which the translational, rotational and torsional variables are mutated randomly. Selection of the offspring of the current generation occurs based on the individual's fitness: thus the better solutions go on into the next generations, whereas conformations with a low fitness are discarded. This cycle of crossover, mutation to lead new generation is repeated until the better bioactive conformation (run) is given.

The LS perform an energy minimization of the current found conformation. In each generation a fraction of conformations population undergoes the geometry optimization, based on the local search frequency. Rapid energy evaluation is achieved by precalculating atomic affinity potentials (grid maps) for each atom type in the substrate molecule by grid method.[69]

These maps are calculated by AutoGrid. In this procedure the protein is embedded in a three dimensional grid and a probe atom is placed at each grid point (Figure 1.7). The energy of interaction of this single atom with the protein is assigned to the grid point.

An affinity grid is calculated for each type of atom in the substrate, typically carbon, oxygen, nitrogen and hydrogen, as well as a grid of electrostatic potential, either using a point charge of +1 as the probe, or using a Poisson-Boltzmann finite difference method, such as DELPHI.[70] The energetic of a particular substrate configuration is then found by tri-linear interpolation of affinity values of the eight grid points surrounding each of the atoms in the substrate.

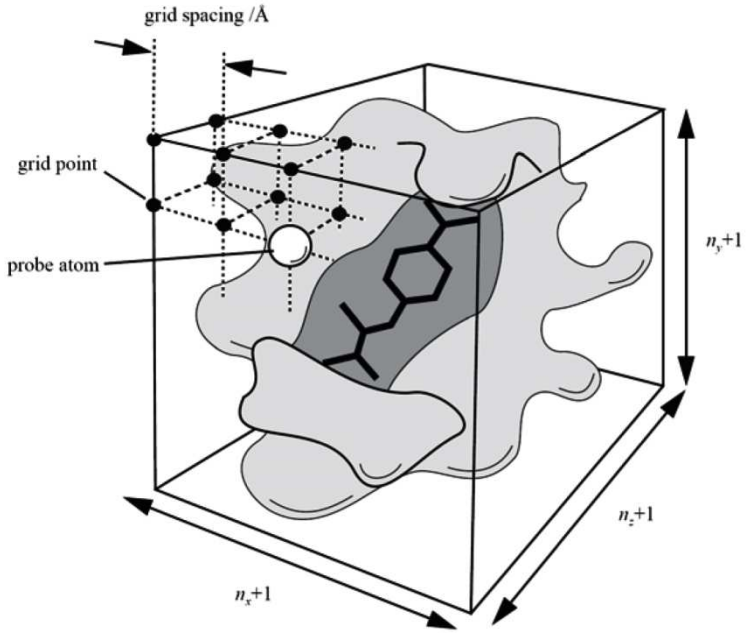


Figure 1. 7 Schematic representation of the grid map.

The electrostatic interaction is evaluated similarly, by interpolating the values of the electrostatic potential and multiplying by the charge on the atom (the electrostatic term is evaluated separately to allow finer control of the substrate atomic charges). The time to perform an energy calculation using the grids is proportional only to the number of atoms in the substrate, and is independent of the number of atoms in the protein. An estimated free energy of binding is used to evaluate the docked ligand conformations. This scoring function, based of force field Amber,[71] comprises terms above described (directional hydrogen bonding, electrostatics, Van der Waals, internal energy) and entropic contribution: desolvation and torsional entropy. The latter describes the loss of entropy upon interaction with macromolecule followed by immobilization in the active site. The desolvation belongs the displacement

of water molecules from the active site upon the binding of ligand to the macromolecular surface and the reorganization of solvent around the complex. The scoring function was implemented using the thermodynamic cycle of Wesson and Eisenberg. The function is:

$$\Delta G = \Delta G_{VdW} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) + \Delta G_{hbond} \sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) +$$

$$+ \Delta G_{elec} \sum_{i,j} \frac{q_i \, q_j}{\varepsilon(r_{ij}) r_{ij}} + \Delta G_{tor} N_{tor} + \Delta G_{solv} \sum_{i,j} (S_i V_j + S_i V_i) e^{\left({}^{-r_{ij}^2}/_{-2\sigma^2} \right)}$$

Equation 1. 3

where the five ΔG terms on the right hand side are coefficient empirically determined using a linear regression analysis from a set of protein-ligand complexes.[55]

For what concern AutoDock Vina,[58] this is a new open-source program for drug discovery, molecular docking and virtual screening, offering multi-core capability, high performance and enhanced accuracy and ease of use. Vina uses a sophisticated gradient optimization method in its local optimization procedure. The calculation of the gradient effectively gives the optimization algorithm a “sense of direction” from a single evaluation. In the spectrum of computational approaches to modeling receptor ligand binding molecular dynamics with explicit solvent, molecular dynamics and molecular mechanics with implicit solvent, molecular docking can be seen as making an increasing trade-off of the representational detail for computational speed.[72] Among the assumptions made by these approaches is the commitment to a particular protonation state of and charge distribution in the molecules that do not

change between, for example, their bound and unbound states. Additionally, docking generally assumes much or all of the receptor rigid, the covalent lengths, and angles constant, while considering a chosen set of covalent bonds freely rotatable (referred to as active rotatable bonds here). Importantly, although molecular dynamics directly deals with energies (referred to as force fields in chemistry), docking is ultimately interested in reproducing chemical potentials, which determine the bound conformation preference and the free energy of binding. It is a qualitatively different concept governed not only by the minima in the energy profile but also by the shape of the profile and the temperature.[73] Docking programs generally use a scoring function, which can be seen as an attempt to approximate the standard chemical potentials of the system. When the superficially physics-based terms like the 6–12 van der Waals interactions and Coulomb energies are used in the scoring function, they need to be significantly empirically weighted, in part, to account for this difference between energies and free energies.[73]

The afore mentioned considerations should make it rather unsurprising when such superficially physics-based scoring functions do not necessarily perform better than the alternatives. This approach was seen to the scoring function as more of “machine learning” than directly physics-based in its nature. It is ultimately justified by its performance on test problems rather than by theoretical considerations following some, possibly too strong, approximating assumptions

The general functional form of the conformation-dependent part of the scoring function AutoDock Vina is designed to work with is:

$$c = \sum_{i<j} f_{t_i t_j}\left(r_{ij}\right)$$

Equation 1. 4

where the summation is over all of the pairs of atoms that can move relative to each other, normally excluding 1–4 interactions, i.e., atoms separated by three consecutive covalent bonds. Here, each atom i is assigned a type t_i, and a symmetric set of interaction functions $f_{ti\text{-}tj}$ of the interatomic distance r_{ij} should be defined.

This value can be seen as a sum of intermolecular and intramolecular contributions:

$$c = c_{inter} + c_{intra}$$

Equation 1. 5

The optimization algorithm attempts to find the global minimum of *c* and other low-scoring conformations, which it then ranks.

The predicted free energy of binding is calculated from the intermolecular part of the lowest-scoring conformation, designated as 1:

$$s_1 = g\,(c_1 - c_{intra1}) = g(c_{inter1})$$

Equation 1. 6

where the function g can be an arbitrary strictly increasing smooth possibly nonlinear function.

In the output, other low-scoring conformations are also formally given *s* values, but, to preserve the ranking, using *c*intra of the best binding mode:

$$s_i = g(c_i - c_{intra1})$$

Equation 1. 7

For modularity reasons, much of the program does not rely on any particular functional form of f_{ti-tj} interactions or *g*. Essentially, these functions are passed as a parameter for the rest of the code.

In summary the evaluation of the speed and accuracy of Vina during flexible redocking of the 190 receptor-ligand complexes making up the AutoDock 4 training set showed approximately two orders of magnitude improvement in speed and a simultaneous significantly better accuracy of the binding mode prediction. In addition, Vina can achieve near-ideal speed-up by utilizing multiple CPU cores. However, AutodockVina does not provide very good weight of the energetic contribution derived from the hydrogen bond and electrostatic interactions, especially when the metal ions are presents.

It is important to underline in this phase of the study description that the molecular docking methodology was used for the design of new molecules with potential antitumoral and antiflammatory activities as HDAC (see paragraph 2.2-2.4, and 2.6) and mPGES-1 (see Chapter 4) inhibitors respectively. Alongside this application, in this results description, the molecular docking was also used to rationalize the binding modes and the mechanism of action of Ugi products CHAP 1 derivatives (paragraph 2.5) as HDAC inhibitors, of FX866 and CHS823analogs inhibitors of NMPRTase (Chapter 3), of marine natural products acting as inhibitors of hsPLA_2 (Bolinaquinone, BLQ; and Cladocoran A, CLD A), 4-methylensterols as ligands of nuclear receptors (Farnesoid-X-receptor (FXR) and the pregnane-X-receptor (PXR)), and of two molecules with already known pharmacology activivity as taurolitholic acid and ciprofloxacin (chapter 7) acting as agonists of TGR5.

1.3.2 Quantum Mechanical Calculation of NMR Parameters in the Stereostructural Determination of Natural Products.

Many molecular properties of organic compounds, such as chemical reactivity and catalytic, biological, and pharmacological activities, are critically affected not only by their functional groups but also by their spatial position. Thus, the disclosure of the relative configuration has a great impact in the full understanding of their chemical behaviours. Different approaches to determine the exact structure and/or configuration of organic products have been devised.[74,75,76] The total synthesis has played a primary role in the structural assignment and revision but its drawback is represented by the additional costs in terms of time and money. For these reasons, a series of new and more rapid methods, such as nuclear magnetic resonance (NMR), circular dichroism (CD), X-ray crystallography, and mass spectrometry (MS), that allow the preservation of the sample under investigation, have shown to be a valid alternative to the classical chemical approach.

In this field, NMR spectroscopy is one of the most employed tool, as some NMR parameters (coupling constant, chemical shift (cs)) can provide fundamental information on the configurational and conformational arrangement of organic molecules. For example, the $^{3}J_{H\text{-}H}$ coupling constants between protons separated by three bonds depend on the dihedral angles, following the well known Karplus equation.[77] Moreover, the Nuclear Overhauser Effect (NOE)[78] provides information of the 3D spatial arrangement of the nuclei, clarifying the geometrical information on the relative positions of the atoms in the analysed molecule. Thus, the evaluation of simple NMR parameters, such as proton-proton *J*-coupling values, chemical

shifts, and/or nuclear Overhauser effect intensities allows to determine the configuration of cyclic compounds with three- to six-membered rings presenting a predictable conformational behaviour. Polysubstitued opened chains and macrocycles, constitute a more difficult cases of relative configurational assignment, because the stereochemical analysis is complicated by the geometrical uncertainty of such types of flexible systems.

For the above situations, different NMR-based methods, such as the *J*-based analysis,[76,79] the Universal NMR Database (UDB),[76,80] and the quantum mechanical calculation of NMR parameters,[76,81,82] have been proposed for the relative (and/or absolute) configurational assignment of organic molecules. The *J*-based analysis was originally devised by Murata and co-workers and it has been shown to be very helpful for the relative configurational assignment of two adjacent (1,2) or alternate (1,3) stereocenters belonging to an acyclic carbon chain. Briefly, this *J*-based approach consists in considering the three main staggered rotamers (*anti*, g^+ and g^-) of the two possible relative configurations (*threo* and *erythro*) of a stereopair and assigning for each rotamers a predicted qualitative (small, medium and large) set of $^{3}J_{\text{H-H}}$ and $^{2,3}J_{\text{C-H}}$, based on the dependence of scalar couplings on the dihedral angles. The comparison of the experimental measured values of homo- and hetero-nuclear coupling constants with the rotamer predicted patterns allows to assign the relative configuration of two stereocenters.

The UDB method is based on the comparison of the proton and carbon chemical shifts of a structure having an unknown relative configuration with the resonance values of a molecular database formed by fragments of known compounds. In particular, the structure under studies could be divided in small fragments and its chemical shifts compared with an appropriate reference compound in the database.

In the last years, great advances have been made in developing QM methods of chemical interest able to predict molecular properties. In particular, the quantum mechanical calculations of NMR parameters have been used as an emerging strategy for the assignment of relative configuration of organic molecules, based on the high accuracy in the reproduction of experimental NMR properties achieved also at a low demanding level of theory.[83,84] For further details about theoretical concepts, applications and limitations of these NMR-based approaches we refer to our previous review on the determination of relative configuration of organic compounds.[76] It is noteworthy that, besides the development and application of QM approach for structural studies, fast empirical methods have been devised to predict NMR chemical shifts.[85] These empirical methods are based on fast calculation algorithms[86] that can generate a set of possible structural hypotheses with the average deviation between calculated and experimental chemical shifts equal to δ = 1.8 ppm for ^{13}C chemical shifts. Such empirical NMR chemical shift predictions could be useful with large-sized molecules or in presence of very flexible compounds for which different conformers have to be considered in the more time consuming QM calculations. Moreover, these empirical methods can be applied as filter to narrow the number of stereoisomers to be accurately verified by other methods such as x-ray, total synthesis, QM approaches.

The ^{13}C-based protocol (Figure 1.8), used in this project, consists of four fundamental steps: (a) conformational search and a preliminary geometry optimization of all the significantly populated conformers of each stereoisomer; (b) final geometry optimization of all the species at QM level; (c) GIAO (gauge including atomic orbital)[87] ^{13}C NMR calculations of all the so-obtained structures at QM level; (d) comparison of the Boltzmann averaged

NMR parameter calculated for each stereoisomer with those experimentally measured for the compound under examination. This protocol has been devised for flexible systems considering the importance of the contribution of all significant conformers to predict a chemical-physical property and the theory level used to calculate the energy of the single geometrical isomers.[81,82] Considering the simple case of a molecule with a couple of two adjacent stereocenters, the first step is to build two diastereoisomers by dedicated software.

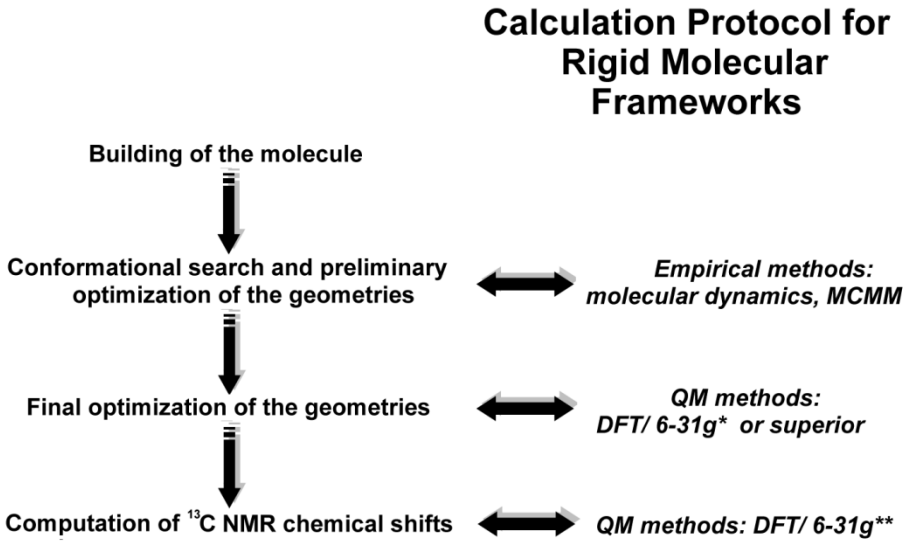


Figure 1. 8 Schematization of protocol used for the determination of relative configuration in organic compounds, based on ^{13}C calculation at QM level of theory.

The conformational sampling is performed at empirical theory level,[88] generally through molecular dynamics (MD) or by Monte Carlo Multiple Minimum methods (MCMM).[89]

A preliminary geometry optimization is run at empirical level (molecular mechanics, MM) or semi-empirical level (AM1,[90] PM3[91] or other) on all found

conformers for each diastereoisomer, followed by a QM optimization step. On the so obtained geometries the ^{13}C NMR chemical shift for each stereoisomer is calculated and the theoretical data are extrapolated taking into account the Boltzmann-weighted average derived from the energies of the single conformers. The calculated values are compared with the experimental NMR data and the relative (or absolute) configuration is determined based on the best fit between theoretical and experimental data set given by one of the two structural hypothesis.

Following the same key steps described for ^{13}C-based protocol, the calculation of homo- and heteronuclear coupling constants can be carried out for the conformational and configurational studies of organic molecules. In details, each global minimum conformer undergoes a full geometry optimization using the DFT theoretical level[92] and then, on the obtained geometries, the calculation of the *J* couplings is performed taking into account the contributions of the following interactions: Fermi contact (FC), paramagnetic spin-orbit (PSO), diamagnetic spin-orbit (DSO), and spin-dipole (SD). Based on the Boltzmann distribution of the conformers, the calculated *J*-coupling values are extrapolated and then compared to the experimental data set, suggesting the relative configuration of the examined compound. For large molecular systems, presenting many stereocenters, it is suggested that, given the prohibitive computational requirement for a simultaneous consideration of all combinations of the possible conformations and configurations, the molecule can be dissected into appropriately 2-C fragments prior to the *J*-coupling calculations,[93] as for the Murata's method.[79] Each reduced subsystem is treated like an entire molecule: a geometry optimization step, followed by $^{3}J_{\text{H-H}}$ and $^{2,3}J_{\text{C-H}}$ calculations, is performed for each staggered rotamer. It is only one of the six calculated data sets that should display a satisfactory

agreement with the experimental values. Differently from the original *J*-based approach proposed by Murata,[79] for which it is impossible to distinguish the *anti erythro* from the *anti threo* arrangement on the basis of the sole evaluation of the *J* coupling values, the quantitative analysis of the calculated *vs* the experimental data allows the relative configurational assignment for the right *anti* rotamer.

This methodology was applied for the structural studies reported in the chapter 8, where the QM/NMR was used as support of the total synthesis of natural products (kedarcidin chromophore and palau'amine, paragraph 8.1), and in combination with quantitative NMR-derived interproton distances for the assigment of the ralative configuration of marine natural sterol conicasterol F (see paragraph 8.2).

1.3.3 Quantitative Interproton Distances from Nuclear Overhauser Effect (NOE) Data

The accurate measurement of internuclear distances in organic molecules can be used to determine stereochemical and conformational detail of a structure. Internuclear distance information can be obtained through a number of NMR methods, including Residual Dipolar Couplings (RDC)[94] and cross-relaxation rates, using either inversion-recovery methods[95] or NOESY methods.

The nuclear Overhauser effect (NOE), described as the change in magnetisation of a spin when a neighbouring spin is irradiated as a result of dipolar coupling, can be used to measure such distances. However, the use of NOE data in distance measurements has mainly been applied in a semi-quantitative fashion to establish gross structural differences, rather than in a

quantitative fashion. Indeed in many cases, their quantitative use is advised against as NOE intensities can be perturbed by numerous factors other than internuclear distance. These relate to both experimental and molecular components including additional cross-relaxation pathways or spin diffusion, selective polarisation transfer, variation in effective τ_c (rotational correlation time) between spins, accuracy of signal integration and conformational flexibility.[96] Nevertheless, various NMR spectroscopy investigations have been reported on distance determination within small molecules using NOE data, principally utilising the full relaxation matrix analysis.[97,98]

Given the improvements in the experimental side of NOE measurements – new experimental methods (transient DPFGSE NOESY,[99,100,101] zero-quantum filtration[102] etc.) and hardware improvements – more accurate and clean NOE intensities can now be obtained. Not only this, but it was recently shown that many of the perturbing factors stated above do not contribute significantly if the molecule of interest is in the fast tumbling regime and if measurements are made within the initial rate approximation, (which states that at short mixing times only the cross relaxation rate, σ_{IS}, between two spins is responsible for the magnitude of the NOE enhancement, and that all spin pairs behave as if they were isolated two-spin systems).

Using the method described below, it has been shown that surprisingly accurate NOE-derived distances can be obtained.[103,104,105]

Various NMR spectroscopy investigations have been reported on distance determination within small molecules using NOE data, principally utilising the full relaxation matrix analysis[106] and using the Initial Rate Approximation, in particular here will be detailed described the second one.

1.3.3.1 The Initial Rate Approximation

The initial build up of NOE enhancement in all kinetic NOE experiments is approximately linear, and this build-up rate leads to the initial rate approximation. The initial rate approximation is used to analyse kinetic NOE data that is obtained through NOE build-up experiments (selective 1D and 2D) to extract internuclear distance information.

To understand the initial rate approximation, the Solomon equations[107] must first be revisited as in for a two-spin system (Equation 1.8)

$$dI_z/dt = -R_I(I_z - I_z^0) - \sigma_{IS}(S_z - S_z^0)$$

Equation 1. 8

where I_z and S_z correspond to the intensities of the I peak and S peak respectively at time t, and S_z^0 and I_z^0 correspond to the respective equilibrium intensities of these peaks. R_I is the total relaxation rate of spin I, and σ_{IS} is the cross relaxation rate between two spins I and S.

The initial conditions (after S_z inversion) for a NOESY experiment with a two-spin system are $I_z = I_z^0$ and $S_z = -S_z^0$, which yields the following initial rates of NOE growth as shown in Equations 1.9-10 for 1D and 2D[108] experiments respectively.

$$\frac{dI_Z}{dt} = 2\sigma_{IS}S_z^0$$

Equation 1. 9

$$\frac{d[a_{cross}(\tau_m)]}{dt} = -\sigma_{IS}a^0$$

Equation 1. 10

where $a_{cross}(\tau_m)$ is the intensity of the cross peak, and a^0 is the intensity of the diagonal peak at $\tau_m=0$.

The initial rate approximation assumes that only the cross-relaxation rate constant between spins I and S is responsible for the size of the NOE enhancement, i.e. other forms of relaxation have no effect, and that all spin pairs behave as if they were isolated two-spin systems (the Isolated Spin Pair Approximation (ISPA)). If more than two spins are considered, the subsequent evaluation gets increasingly more complicated. Cross-relaxation rates can thus be determined from the initial slope of the build-up curves, where the initial rate approximation assumes that there is a linear dependence of the NOE build-up on the mixing time of the experiment. The cross-relaxation rates can then be used to calculate internuclear distances as will be explained below. It has been found, however, that the initial rate approximation only holds true for the shortest mixing times when used in this fashion. In order to maximise the utility of NOE distance measurements, it is highly desirable that longer mixing times (and hence more intense NOEs) be used i.e. the initial rate approximation needs to be extended as much as possible.

However it can be seen to fail due to a number of factors. At short mixing times, cross-peak intensities in 2D experiments and peak intensities in 1D experiments are relatively small, and this is the region of lowest signal-noise, making accurate measurements more difficult, especially for medium to long range distances whose NOEs are fundamentally weaker. Unfortunately, these distances are often those of most interest, so the ability to include data from longer mixing times (where the signal-noise is better) in calculations is preferable. However, at longer mixing times, external relaxation sources other

than cross-relaxation (e.g. spin-rotation, chemical shift anisotropy, quadrupolar relaxation) can cause problems and so can the three spin effect/spin diffusion, all of which can affect the accuracy of quantitative distance measurements made.

It has been shown that this approximation can be extended to longer mixing times in 2D NOESY experiments by taking the ratio of cross-peak to the diagonal peak in the same experiment.[109,110] This compensates for the non-linearity of the NOE build-up rate due to external relaxation so that the initial rate approximation is valid and linear over a wider range of mixing times. Previous methods involved the scaling of each individual target resonance relative to experiments carried out at zero mixing time. Problems resulting from variations from experiment to experiment are reduced in this new method as scaling between the target and NOE resonances occur within each separate experiment. Hu and Krishnamurthy[110] then demonstrated that this practice could be applied to 1D NOE data, specifically to that obtained using the DPFGSE NOE pulse sequence. They refer to this method as peak amplitude normalisation for improved cross-relaxation (PANIC). They used the Solomon equations (Equation 1.8) as a basis to explain their process as follows.

After solving the Solomon equations for the conditions of a DPFGSE NOE experiment, Hu and Krishnamurthy showed that the NOE enhancement at spin I as a fraction of the equilibrium magnetisation can be displayed as in Equation 1.11:

$$\eta_{NOE}(\tau_m) = \frac{I_Z(\tau_m) - I_Z(\tau_m)_{ref}}{M_0} = \frac{2\kappa\sigma_{IS}}{D}\left(e^{-\lambda_2\tau_m} - e^{-\lambda_1\tau_m}\right)$$

Equation 1. 11

where $\boldsymbol{D = \sqrt{(\rho_I - \rho_S)^2 + 4\sigma_{IS}^2}}$ and $\boldsymbol{\lambda_1 = (\rho_I + \rho_S + D)/2}$ and $\boldsymbol{\lambda_2 = \frac{(\rho_I + \rho_S - D)}{2}}$ and κ is the scaling factor (0<<κ<<1) that accounts for signal loss from relaxation during the spin echo sequence of the DPFGSE. M_0 is the equilibrium z-magnetisation for spins I and S (assumed to be equal) and $(\tau_m)_{ref}$ corresponds to the mixing time of a reference spectrum, acquired with spin S aligned along the +z axis, so that spin I shows no NOE enhancement.

Equation 1.11 can be simplified by substituting the exponential terms with their Taylor series and ignoring all 2nd and higher order terms, which we assume to have little contribution, to give Equation 1.12:

$$\boldsymbol{\eta_{NOE}(\tau_m) \approx 2\kappa\sigma_{IS}\tau_m}$$

Equation 1. 12

This provides the basis for the method used prior to Hu and Krishnamurthy's extension. However, it involves acquiring a DPFGSE NOE experiment at zero mixing time and a simple 1D pulse-acquire experiment to determine the value of κ, and it only remains valid up until short mixing times of ca. 200ms. Hu and Krishnamurthy[110] went on to detail a more accurate method for determining cross-relaxation rates that also takes into account the irradiated peak intensity. This extends the initial rate approximation so that it can be used at longer mixing times. Much of the non-linearity at these times evolves from external relaxation, but by also considering the target magnetisation described in Equation 1.13, this effect is almost completely cancelled out.

$$\eta_{target}(\tau_m) = \frac{M_Z^S(\tau_m) - M_Z^S(\tau_m)_{ref}}{M_0}$$

$$\approx -\frac{2\kappa\sigma_{IS}}{D}\left(e^{-\lambda_1\tau_m} + e^{-\lambda_2\tau_m}\right)$$

Equation 1. 13

Dividing Equation 1.13 into the NOE enhancement, Equation 1.12 and simplifying by ignoring all 3rd order and higher terms, the effect of external relaxation can be seen to cancel to yield Equation 1.14.

$$\zeta_{NOE}(\tau_m) \approx \sigma_{IS}\tau_m$$

Equation 1. 14

where $\zeta_{NOE}(\tau_m)$ is the NOE enhancement normalised against the irradiated signal in the same spectrum.

Cross-relaxation rates can then be measured by a linear regression analysis. It is no longer necessary to determine the scaling factor, or to normalise against equilibrium magnetisation, leading to a simpler data analysis. As well as this, the initial rate approximation is extended to remain linear at longer mixing times. Hu and Krishnamurthy[110] have thus demonstrated that by taking the ratio of two NOE intensities, it is possible to correct for external relaxation.

Under conditions valid for the initial rate approximation, it is only the internuclear distance, r_{IS} between spins I and S that has an effect on the size of the NOE enhancement between the respective spins. This means that by

measuring the NOE enhancement between the two spins, a direct measure of σ_{IS} is achieved, and this value is directly proportional to the internuclear distance as described below in Equation 1.15.

$$\boldsymbol{\sigma_{IS} = \varphi r_{IS}^{-6}}$$

Equation 1. 15

where, $\boldsymbol{\varphi = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\hbar^2\gamma^4}{10}\left(\frac{6\tau_c}{1+4\omega^2\tau_c^2} - \tau_c\right)}$

In the above equation the only unknown is τ_c, the molecular correlation time, which is not the easiest of values to measure. By assuming that the molecule of interest is rigid and tumbling isotropically, it can be approximated that all components of the system will have the same correlation time. In doing so, a relative calibration can now be made using the NOE intensity from a pair of protons of known internuclear distance, e.g. methylene protons – the reference - and comparing this to the unknown distance's NOE intensity as in Equation 1.16.[111]

$$\boldsymbol{\frac{f_I\{S\}(t)}{f_{ref}(t)} = \left(\frac{r_{IS}}{r_{ref}}\right)^{-6}}$$

Equation 1. 16

where $f_I\{S\}(t)$ is the NOE intensity of interest at mixing time t, and $f_{ref}(t)$ is the reference NOE intensity at mixing time t.

Equation 1.16 can be rearranged to express the unknown internuclear distance as follows in Equation 1.17:

$$r_{IS} = r_{ref}\left[\frac{f_I\{S\}(t)}{f_{ref}(t)}\right]^{-1/6}$$

Equation 1. 17

Accurate distances can be obtained from this method as internuclear distance is thus proportional to the inverse sixth power root of the function, which means that the effect of experimental errors, spin diffusion and motional effects are greatly reduced. Not only this, but by taking the ratio of two NOE intensities, external relaxation is being corrected for, as described above, so that this method can be used for data collected at relatively long mixing times.

Wang et al.[112] determined interproton distances of three dolichodial-like diastereomers, each of which have three stereocenters sing 2D NOESY data sets obtained with a 2 second mixing time. They estimated interproton distances from NOE volumes using Equation 1.18, a form of the initial rate method as described above.

$$r = \left(\frac{r_{ref}^6 V_{ref}}{V}\right)^{1/6}$$

Equation 1. 18

where V and V_{ref} are the volumes of the unknown and reference (geminal) cross-peaks respectively and r and r_{ref} are the corresponding interproton distances.

Andersen et al.[113] also investigated the quantitative use of 2D NOESY data for small molecules. They reported the effects of preparatory delay (d_1) truncation on the accuracy of NOE distance data, and a method with which to

overcome this. Preparatory truncation delay, where the delay is below the normal value of 5 times T_1, means that not all signals will have fully relaxed back to their equilibrium states, and in 2D NOESY spectra, this can result in diagonal asymmetry, a problem for accurate data analysis.

Unfortunately, truncation is desirable as small molecules have long T_1 values, and so collection of 2D NOESY data can be a lengthy process if the normal pulse sequence is used. Andersen et al. estimated cross-relaxation rates in truncated NOESY data by taking the average of two diagonal cross-peaks after they have been normalised by their diagonal-peak as in Equation 1.19.

$$\frac{\boldsymbol{\eta_{IS}}^{+}/_{\boldsymbol{or}}\,\boldsymbol{\eta_{SI}}}{\boldsymbol{\eta_{II}}^{+}/_{\boldsymbol{or}}\,\boldsymbol{\eta_{SS}}} = \boldsymbol{\sigma_{IS}\tau_m}$$

Equation 1. 19

where η_{IS} and η_{SI} are the cross-peak cross-section intensities and η_{II} and η_{SS} are the diagonal peak cross-section intensities.

This normalisation accounts for cross-peak leakage, which is assumed to be well approximated by the net leakage at either of the corresponding diagonal peaks. In this case, the term leakage accounts for the reduction in the NOE caused by external relaxation. They found that cross-relaxation results with truncation agree with non-truncated experimental data, although they were neither as precise nor as accurate. By examining the differences between internuclear distance data derived from each spectrum dimension (f1 and f2 spectral projections) they found that f2 spectral projections were most useful for well-resolved resonances, as long as corrections were made for t_1 streaks. Quantitative determination around the critical point ($\omega\tau_c \approx 1.12$) is made more

difficult by the fact that NOE enhancements using the standard NOESY experiment are almost non-existent.

However, 1D ROESY data can also be used to produce quantitative distance data in these cases. Such an example is that of the natural diterpene derivatives investigated by Forgo et al.[114] They determined distances by obtaining the cross-relaxation rate for a pair of protons from the slope of the NOE build up curve over a range of mixing times. They then used Equation 1.20 to determine the corresponding distance. In order to do this, the rotational correlation time, τ_c is needed, and so this was determined from ^{13}C spin lattice relaxation data.

$$\sigma_{H,H} = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma_H^4 \hbar^2}{r^6} \left(\frac{3}{5}\left(\frac{\tau_c}{1+4\omega_H^2\tau_c^2}\right) - \frac{1}{10}\tau_c\right)$$

Equation 1. 20

These NOE observed distances were subsequently compared with solid-state X-ray crystallography distances for the equivalent compounds, and the correlation was observed to be very good, with differences of only 0.1-0.2Å between corresponding distances.

1.3.3.2 Method for quantitative interproton distance determinations

Under conditions valid for the initial rate approximation, and assuming that the molecule of interest is in the fast tumbling regime, it is only the internuclear distance, r_{IS}, between spins I and S that has an effect on the size of the NOE enhancement between the respective spins. This means that by measuring the NOE enhancement between the two spins, a direct measure of

σ_{IS} is achieved, and this value is directly proportional to the internuclear distance as described below in Equation 1.21.

$$\boldsymbol{\sigma_{IS} = k r_{IS}^{-6}}$$

Equation 1. 21

where, $\boldsymbol{k = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\hbar^2 \gamma^4}{10}\left(\frac{6\tau_c}{1+4\omega^2\tau_c^2} - \tau_c\right)}$

Assuming that the values of μ_0, γ (magnetogyric ratio) and ω (Larmor frequency) are essentially fixed for a given homonuclear experiment, and if we further assume that τ_c (rotational correlation time) is comparable for each spin pair in a given selective inversion experiment, then *k* can be assumed to be identical for each spin pair within a given selective inversion experiment. This leads to Equation 1.22, where the ratio of the intensities of a pair of NOE signals, η_{IS}:η_{ref}, can be assumed to be proportional to the ratio of the respective internuclear distances.

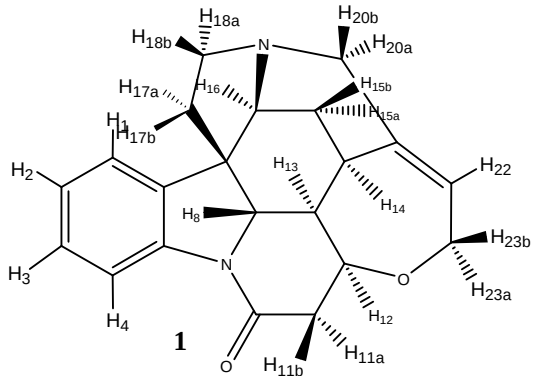
$$\frac{\boldsymbol{\eta_{IS}}}{\boldsymbol{\eta_{ref}}} = \left(\frac{\boldsymbol{r_{IS}}}{\boldsymbol{r_{ref}}}\right)^{-6}$$

Equation 1. 22

where η_{IS} is the NOE intensity of interest at mixing time t, and η_{ref} is the reference NOE intensity at mixing time t. Thus, by measuring η_{IS} and η_{ref}, one needs only to know one distance e.g. r_{ref}, in order to calculate the unknown distance, r_{IS}.

To prove this methodology, Butts[103,104,105] and co-workers had chosen the stryctinine molecule (Scheme 1.1) because it is a rigid molecule, with a well resolved ^{1}H spectrum, and its X-ray structure[115] as well as a computational

structure[116] data are available, so it is well characterised and internuclear distances are already known.



Scheme 1. 1 Molecular structure of strychnine (**1)**.

The initial reference distance used was the H15a-H15b NOE distance calibrated to 1.76Å (assumed intermethylene distance) and distances were propagated through the entire dataset by the following method in Figure 1.9.

The analysis employing equation 1.22 to determine interproton distances is illustrated for H15a of strychnine in $CDCl_3$.

The selective (1D) transient DPFGSE-NOESY spectrum of H15a is shown in Figure 1.10 and shows the clean, well-resolved NOE peaks with a very flat. For convenience, the absolute values for the NOE intensities were measured relative to the irradiated (negative) peak for which the integral was arbitrarily assigned a value of -1000.

The NOE intensity for the H15 methylene protons H15a-H15b (η = 216.2) corresponds to an estimated distance of ~1.76 Å, and hence the intensities of the remaining observed NOEs from H15a to H13 (η= 60.4), H14 (η = 25.8**)**, H16 (η = 31.7**)**, and H8 (η = 1.3) can be used to determine their corresponding distances to H15a to be 2.18, 2.50, 2.42, and 4.29 Å respectively.

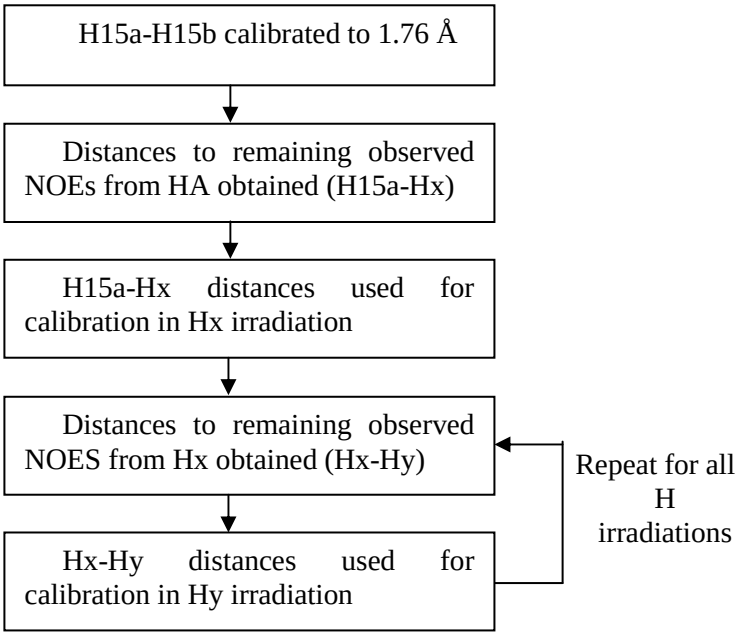


Figure 1. 9 Distance determination method.

In this way a list of interproton distances was obtaining analyzing both 1D and 2D NOESY/ROESY experiments. In this case, mean errors in interproton distances as low as 3% can be obtained from 1D-, 2D- NOESY or ROESY experiments when these are conducted in the fast-tumbling regime and inside the Initial Rate Approximation (IRA).

At distances >4 Å, signals from artifacts and NOEs become comparable in intensity and thus distance determinations become significantly less reliable under the conditions described. The interproton distances are most accurate when measured in a non-viscous solvent, but perturbations due to solvent viscosity and deviation from the Initial Rate Approximation were addressed by using a mixing time of around half of the mixing time observed to give maximum NOE intensity in a single 1D-NOESY build-up curve ‑ whether this rule-of-thumb can be applied generally to small molecules is unclear.

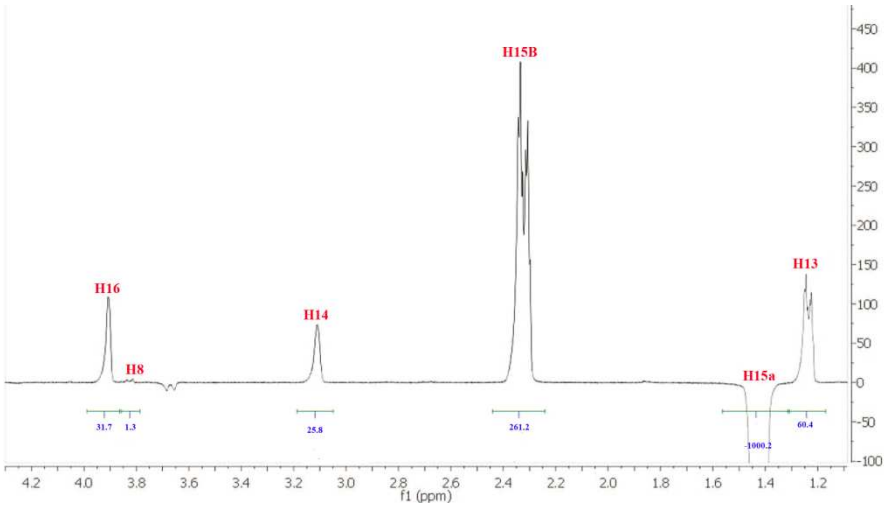


Figure 1. 10 1D-NOESY spectrum of H15a of strychnine in $CDCl_3$.

These so obtained datasets were combined and plotted against their computationally determined values (See an example in Figure 1.11). The slope of the line-of-best fit in Figure 1.13 is essentially unity (1.012), reflecting the accuracy of this NOE-based method in determining interproton distances across a range of separations in strychnine.

The precision is also surprisingly high, as the average absolute error is 3.3% (0.09Å) with a standard deviation of 3.1% (0.11 Å) for distances up to 4.5 Å. These error values compare very well with those obtained from X-ray crystallography, where mean interproton distance errors of 1.4‑4.3% (std 1.2%‑4.3%) arise from the numerous structure determinations of strychnine reported with R^2 values of 1.5%‑3.7%.

Where possible the use of 1D-NOE/ROE data is recommended, rather than 2D-data, as these generally have cleaner baselines, fewer artifacts and thus

reduce the level of post processing of spectra, which in turn minimizes perturbation of the weakest NOEs.

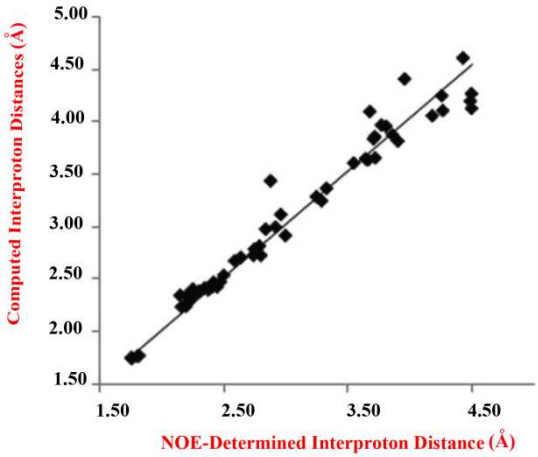


Figure 1. 11 Example of computed interproton distances *versus* those established by 1D-NOE measurements.

The most modern implementations of DPFGSE-based transient NOE experiments should be employed and incorporation of zero-quantum filtration is beneficial to the quality of spectra obtained, but in these experiments zero-quantum filtration did not substantially affect the distances determined. The effects of structural dynamics have not been considered, however the accuracy of the data obtained herein offers significant optimism for comparably accurate dynamical analysis of NOE measurements in flexible systems and we are currently investigating these.

In conclusion, it can be seen that the described method produces very accurate interproton distances in comparison to computed structures when experiments are collected in less viscous solvents such as $CDCl_3$ and d_6-benzene and perturbations due to solvent viscosity (deviation from IRA) were

addressed by reducing the mixing time to correspond to around half that of the mixing time observed to give maximum NOE intensity, or by increasing the temperature of the sample. The effects of spin diffusion are therefore negligible under the conditions used in these cases. Moreover, it is interesting to note that Neuhaus and Williamson[96] describe an r^{-3} rather than an r^{-6} relationship for internal motions faster than overall tumbling (methyl groups in particular) when considering protein structure determination.

This kind of approach in combination with the Quantum Mechanical Calculations of ^{13}C Chemical Shifts was used and described for the stereostructural determination of conicasterol F in the paragraph 8.2.

-CHAPTER 2-

Design, virtual screening and rationalization of potential HDAC inhibitors.

2.1 HDAC as drug target

Gene expression is highly regulated by post-translational modifications of histone proteins, including acetylation, methylation, phosphorylation, ubiquitylation, sumoylation, ADP-ribosylation, glycosylation, biotinylation and carbonylation.[117] In particular, histone acetyl groups from lysine or arginine residues[118] located on the amino-terminal tails of histones proteins (H2A, H2B, H3, and H4) are removed by metalloenzymes called Histone Deacetylases (HDACs)[117] (Figure 2.1)

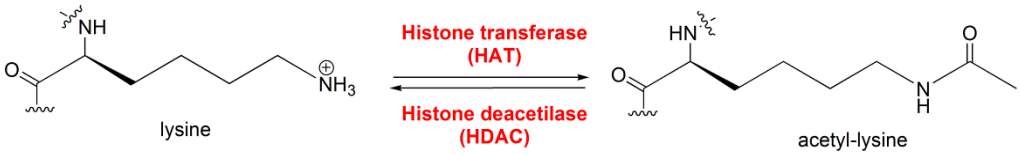


Figure 2. 1 The acetylation state of lysine amino acids are governed by the equilibrium activities of acetyltransferase enzymes and deacetylase enzymes. In the context of gene expression, the lysine residues of histone proteins are key substrates for acetylation.

Histones comprise nucleosomes, which are the basic packaging units of chromosomes.[119] By binding to genomic DNA, the accessibility of genes to transcriptional proteins is altered by histone lysine acetylation. As a result, HDAC proteins are generally associated with repression of transcription and reduced gene expression[120]. HDAC proteins comprise a family of 18 members in humans and are separated into four classes based on their size, cellular localization, number of catalytic active sites, and homology to yeast HDAC proteins. Class I includes HDAC1, HDAC2, HDAC3, and HDAC8. Class II consists of six HDAC proteins that are further divided into two subclasses. Class IIa includes HDAC4, HDAC5, HDAC7, and HDAC9, which each contains a single catalytic active site. Class IIb includes HDAC6 and

HDAC10, which both contain two active sites, although only HDAC6 has two catalytically competent active sites. HDAC11 is the sole member of class IV, based on phylogenetic analysis.[121] Class I, II, and IV HDAC proteins operate by a metal ion-dependent mechanism, as indicated by crystallographic analysis.[122] In contrast, class III HDAC proteins, referred to as sirtuins (SIRT1-7), operate by a NAD^+-dependent mechanism unrelated to the other HDAC proteins.[123] The metal-dependent HDAC proteins are the targets of the HDAC inhibitors discussed in this chapter.

Due to their fundamental role in gene expression, HDAC proteins have been associated with basic cellular events (transcriptional regulation,[124] intracellular transport,[125] metabolism, cell proliferation) and disease states, including cell growth, differentiation, and cancer formation[126] (activation of oncogenes and transcription deactivation of tumor suppressor genes.[127,128]). HDAC's, in fact, have been recently highlighted as promising targets in the epigenetic therapy as a consequence of their ability to influence transcriptional events for the treatment of several disorders included cancer.

The great potential of HDAC inhibitors as anticancer drugs seems to be related in fact, to the transcription and expression of oncogenes which are proved to be silent in cancer pathology.

Since this epigenetic event is associated with carcinogenesis and tumor progression, HDAC inhibitors (HDACi) have been considered promising anticancer agents,[129] and in fact it has been observed that overexpression of HDACs is correlated to cancerous pathologies,[130] and the different isoforms of HDAC are expressed in several tumor tissues with specific biological function (Table 2.1).[131]

Table 2. 1 Expression of HDACs in tumor tissues.

	isoform	Expression in tumor tissues
Class I	HDAC1	gastric, pancreatic, colorectal, prostate cancer; hepatocellular carcinoma
	HDAC2	colorectal cancer: upregulated in polyps; cervical carcinoma; gastric and prostate cancer: increased expression associated with advanced stage
	HDAC3	gastric, prostate, colorectal cancers
	HDAC8	childhood neuroblastoma
Class IIA	HDAC4	breast cancer
	HDAC7	colorectal cancer
Class IIA	HDAC6	oral squamous cell cancer; breast cancer

On the other hand, while individual members of class I and II HDAC proteins are linked to cancer formation, the role of each isoform in carcinogenesis is unclear. Particularly, the molecular mechanism connecting HDAC activity to cancer formation is not yet defined. Targeting of class I

HDACs 1, 2, and 3 inhibits cell cycle progression and promotes apoptosis. This phenotype in cancer cells is in line with the early embryonic lethal phenotype in knockout mice most likely caused by impaired cell cycle defects in early embryonal progenitor cells as has been shown for embryonic fibroblasts. HDAC8 and class II HDACs rather control specific functions such as differentiation, cell signalling, migration, cell adhesion, protein stability and function, and angiogenesis (Figure 2.2). It should be noted, however, that targeting of the same HDAC can have different biological effects depending on the cellular context. Given their association with cancer formation, class I II, and IV HDAC proteins have emerged as attractive targets for anticancer therapy; and the term "HDAC inhibitors is commonly used for compounds that target the "classical" class I, II, and IV HDACs and that are currently evaluated in clinical trials.

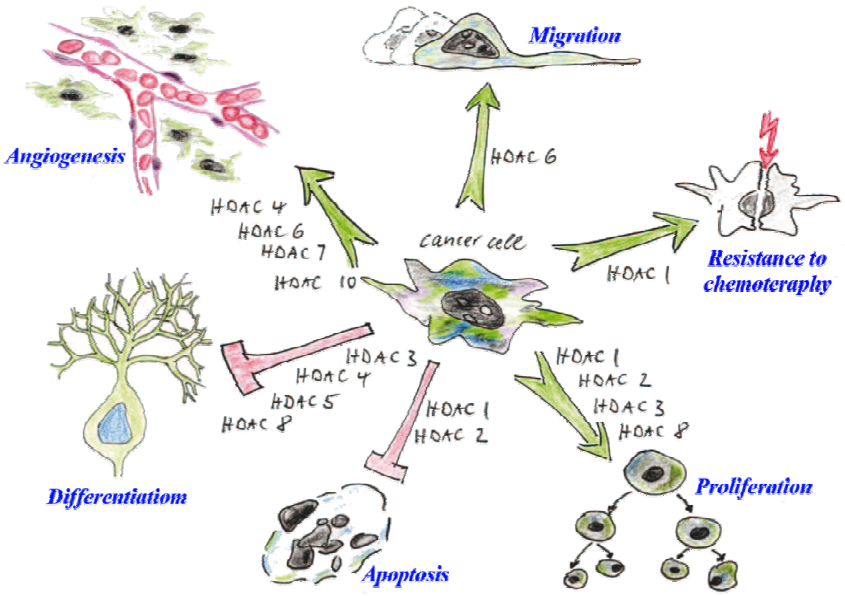


Figure 2. 2 HDAC family members control hallmarks of cancer cell biology.

HDAC inhibitors (HDACi) bear great potential as new antitumor drugs; indeed, they can induce differentiation, growth arrest, and apoptosis in transformed cell cultures. Consistent with their clinical effects, inhibitors of HDAC proteins suppress tumor cell proliferation, induce cell differentiation, and upregulate crucial genes associated with anti-cancer effects.[132]

Although a large number of HDACi have been obtained from both natural sources and through chemical synthesis, and despite the fact that some are in clinical trials—such as valproic acid (**4** in Figure 2.3),[133] MS-275 (**7** in Figure 2.3),[134]—there are only two HDAC inhibitors currently on the market: vorinostat (Zolinza, **17** in figure 2.4)[135] and depsipeptide FK228 (Romidepsin, **13** in Figure 2.3).[136] Both of these were approved by the US Food and Drug Administration (FDA) for the treatment of cutaneous T-cell lymphoma (CTCL).[137] Therefore, HDACi drugs represent a promising next generation of anti-cancer therapeutics.

However, the HDAC involvement and the therapeutic use of HDACi are not restricted to cancer, as several studies have also shown that these enzymes have a role in autoimmune diseases,[138] inflammatory regulation[139] central nervous system disorders[140] and development. Interestingly, most of the intracellular pathways involved in these processes share common intermediaries that are regulated by HDACs, suggesting a central role for these enzymes as regulators of seemingly unrelated physio-pathological conditions. Supporting this concept, HDACi have emerged as potential therapeutic tools for the treatment of autoimmune diseases[138,141] cystic fibrosis[142] and regulation of immune tolerance.[143] In contrast to the rapidly increasing knowledge of the role of HDACs in cancer and the use of HDACi in treating this and other pathological conditions, still little is known about the role of specific HDACs in immune cells and the functional consequences of their inhibition by

HDACi. Although the pro- and antiinflammatory profiles of HDAC regulation are still only in the very early stages of being resolved, it is starting to emerge that HDACi have some promising anti-inflammatory properties in a variety of animal models of inflammatory disease.[139,144] The relatively few HDACi that have so far been reported to influence inflammatory conditions are now summarized in Figure 2.3**.**

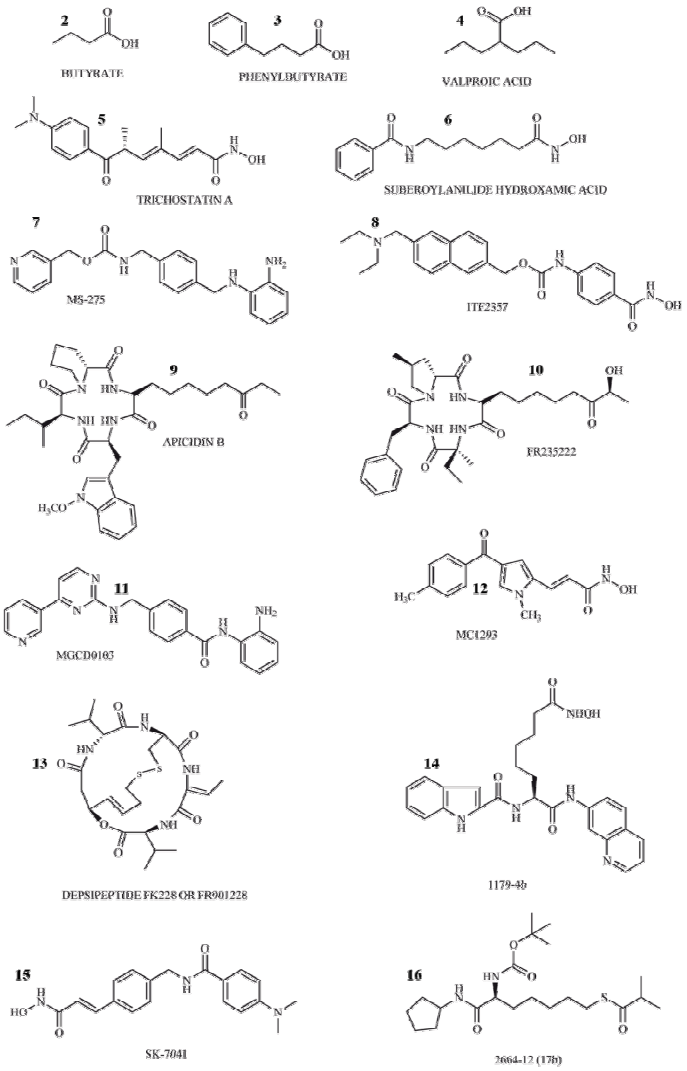


Figure 2. 3 Structures of HDACi with reported antiinflammatory activity.

In general, HDAC inhibitors have a standard, modular construction with structural similarities to the HDAC acetyllysine substrate (Figure 2.1).

HDAC inhibitors typically consist of a metal-binding moiety that coordinates to the catalytic metal atom (Zn^{2+}) within the HDAC active site and a capping group that interacts with the residues at the entrance of the active site (Figure 2.4).

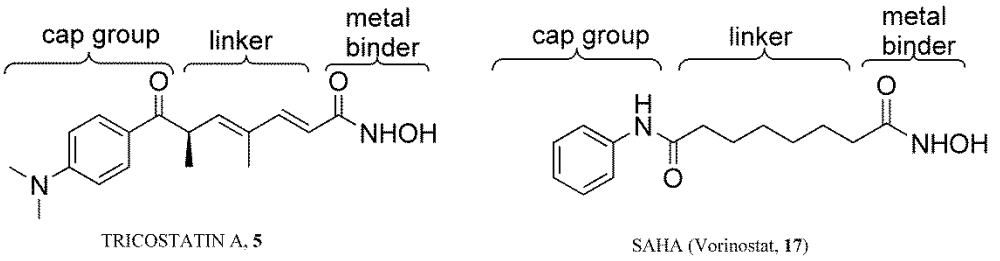


Figure 2. 4 Pan-inhibitors TSA and SAHA.

In addition, a linker that is structurally related to the carbon chain present in the acetyl-lysine substrate appropriately positions the metal-binding moiety and capping group for interactions in the active site. Crystallographic evidence[122] with TSA bound in the active site of a bacterial homologue of class I HDAC proteins (HDLP) confirms that the hydroxamic acid coordinates to the zinc atom at the bottom of the active site, the linker lies in a confined hydrophobic channel, and the aromatic capping group interacts with the amino acids surrounding the entrance of the active site (Figure 2.5)

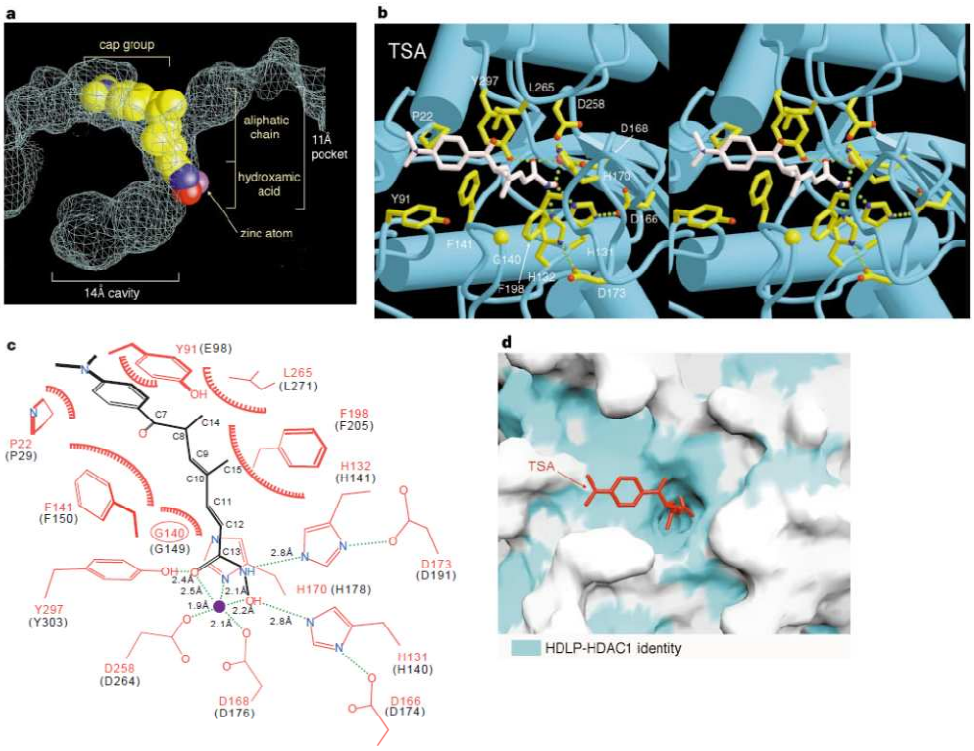


Figure 2. 5 TSA binds inside the pocket making contacts to residues at the rim, walls and bottom of the pocket. **a)** Space-filling representation of TSA in the active-site pocket; **b)** Closeup stereo view of the structure of the HDLP-Zn^{2+}-TSA complex TSA is in white; **c)** Schematic representation of HDLP-TSA interactions. TSA is in black and the protein is in red. HDLP residues are labelled in red with their counterparts in HDAC1 indicated in black; **d)** Surface representation of the HDLP-TSA interface in a similar orientation to **b**.

The majority of HDACi drugs in and out of clinical trials inhibit all HDAC isoforms nonspecifically (so called paninhibitors). SAHA and TSA are the canonical pan-inhibitors (Figure 2.4), influencing the activity of HDAC1–9 with roughly equivalent potency.[145] Selective HDAC inhibitors, which affect either a single HDAC isoform (isoform-selective HDACi) or several isoforms within a single class (class-selective HDACi), would be ideal chemical tools to elucidate the individual functions of each HDAC isoform. Specifically,

selective HDAC inhibitors would aid in defining the molecular mechanism connecting HDAC activity to cancer formation.[129a] In addition, it is possible that a class-selective or isoform-selective HDAC inhibitor would provide a more effective chemotherapy compared to pan-inhibitors.

In this chapter, the application of molecular docking will be described for the design, virtual screening and rationalization of binding modes of small libraries of compounds as potential HDAC inhibitors. In particular, this methodology was successfully applied for the identification of new potential cyclic (mono and bis amides, calixareni, see paragraph 2.2 and 2.4 respectively) and linear (hydroxamic tertiary amines, see paragraph 2.3) HDAC pan-inhibitors; and for the rationalization of the binding modes of Ugi products (see paragraph 2.5).

Because, clinical studies show that pan-HDAC inhibitors may also cause numerous side effects:[146] bone marrow depression, diarrhoea, weight loss, taste disturbances, electrolyte changes, disordered clotting, fatigue, and cardiac arrhythmias. Thus, the next step in the development of HDAC inhibitors is to target selectively individual HDAC isoforms, with the aim of interfering with critical oncogenic function in cancer cells and without affecting normal cells. On this basis, in the last paragraph the work-flow of the different stages involved in the structural characterization, design and synthesis of new selective HDAC inhibitors will be described (see paragraph 2.6)

All docking calculations described in this chapter, were performed using the software Autodock 3.05. For all molecular docking studies of pan-inhibitors (paragraphs 2.2-2.5) the histone deacetylase-like protein (HDLP) reported by Finnin[122] (Figure 2.5) and optimized by Maulucci[147] et al. was used as model receptor, where, in order to have an accurate weight of the

electrostatics, the partial charges of the zinc ion and of the amino acids involved in the catalytic center (Ala169, His170, Asp168, Asp258) have been calculated at DFT B3LYP level and 6-31G(d) basis set using the ChelpG[148] method for population analysis (Gaussian 03 Software Package).[149]

For what concern the design of potential HDAC selective inhibitor, the homology models for HDAC1, HDAC3 and HDAC6 (See Figure 2.6) were used in molecular docking calculations, along with the X-ray structures (See Figure 2.6) of HDAC2 (3MAX),[150] HDAC4 (2VQM),[151] HDAC7 (3C0Z)[152] and HDAC8 (3F0R,[153] and 1VKG);[154] also in this case, the partial charge of Zn^{2+} and of the amino acids constituting the catalytic center were derived by DFT calculations m05[15] level by the 6-31+G(d) basis set and ChelpG method[148] for population analysis (Gaussian 03 Software Package).[155]

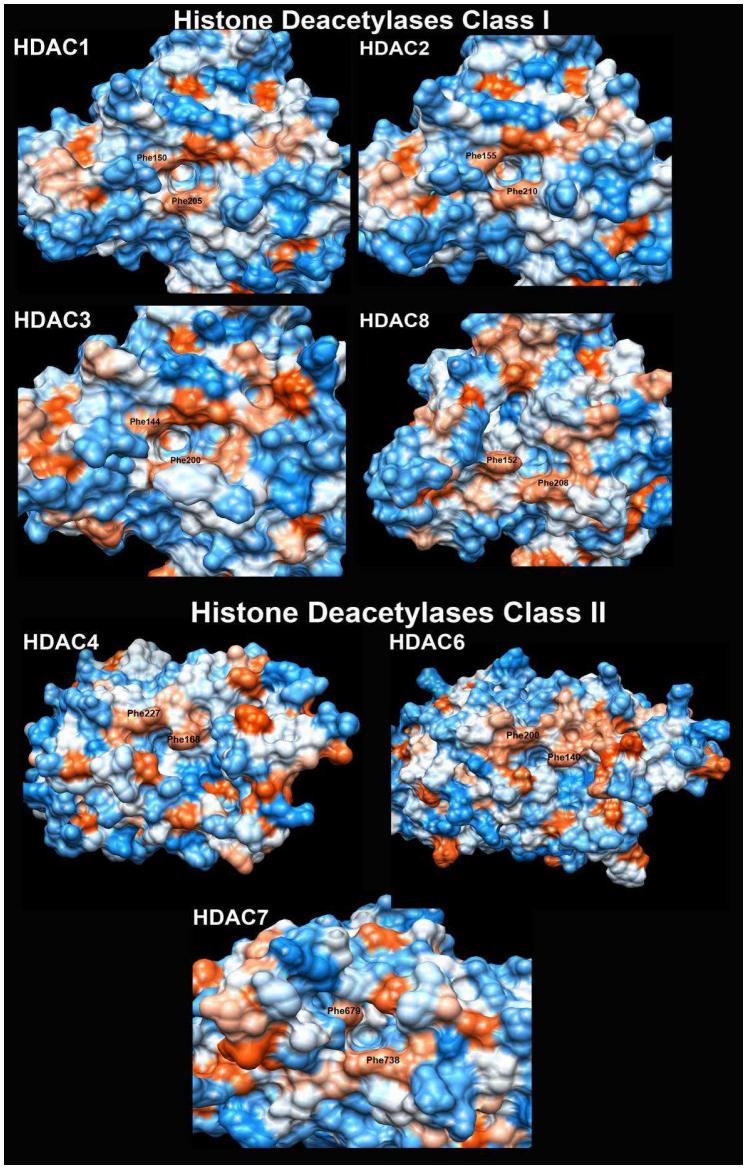


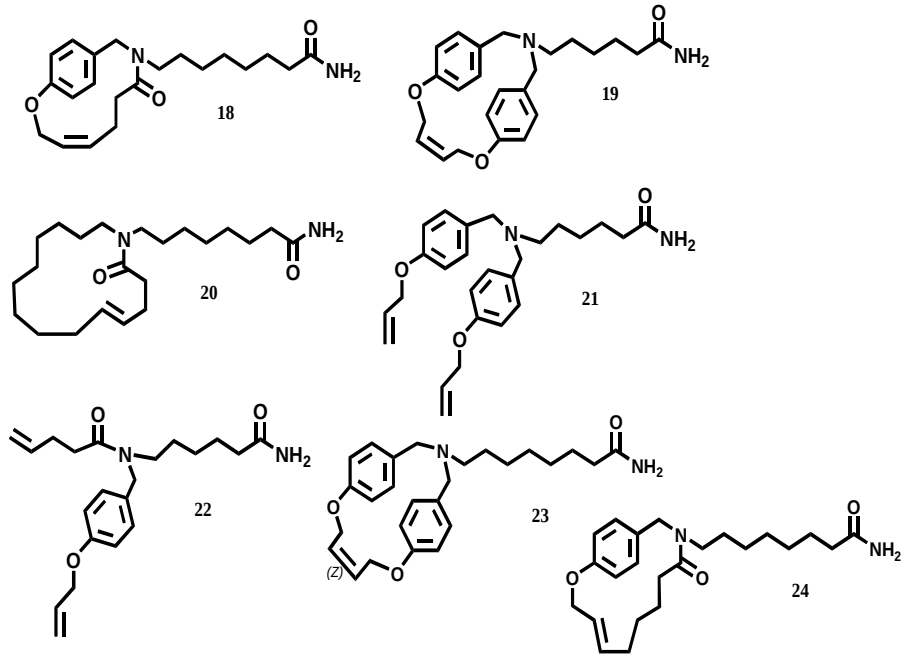
Figure 2. 6 Molecular surface of the Histone Deacetylase proteins (Class I and II) represented by molecular surface and colored according to the hydrophobicity (dodger blue=hydrophilic, orange red=hydrophobic). The figure highlights the two fundamental phenylalanine at the top of the catalytic site.

2.2 Synthesis of new mono and bis amides projected as potential histone deacetylase (HDAC) inhibitors

In this paragraph, it is reported the design, virtual screening, synthesis and biological evaluation of new collection of new potential HDAC inhibitors, whose cytotoxic properties are also described.[156] In particular with the aim to further increasing the in *vivo* stability of the potential HDAC binders, instead of the cyclopeptide framework, it was used a heterocyclic moieties or open chain tertiary amines, bearing a terminal functionalized aliphatic chain of appropriate length. As metal binder domain, it was selected the amide function present in other potent HDAC inhibitors of natural origin, like azumamides.[8] Concerning the retrosynthetic plan of the molecules, it was immediately recognized the possibility of applying, as key step to close the cycles, the RCM reaction, which has emerged as a powerful tool for the construction of carbocyclic and heterocyclic ring systems. On the basis of this assumption, and in consideration of the commercially availability of the building blocks, it were designed seven molecules (Scheme 2.1), five of which (**18**, **19**, **20**, **23**, **24**) showing, as cap group, cyclic structures of different size, and bearing variable length functional tails; the other two (**21**-**22**) mainly reproduced, in some extend, the open versions of the previous ones.

The first step was a docking study on the designed molecules to obtain a prediction of their histone deacetylases inhibitory activity, through virtual screening process. Prior to the docking calculations, a conformational search on the cyclic [157] compounds by means of molecular dynamics was performed at different temperatures (400 K, 600K and 800K) using the MMFFs[158] force field included in the MacroModel software package.[159] On the so obtained model, QM optimization of the energies and the geometries was performed in

vacuo at DFT B3LYP level, using the 6-31G(d) basis set (Gaussian 03 Software Package).[160] Subsequently, the charges of **1**-**7** were calculated with the ChelpG method[148] at the B3LYP/6-31G+(d) level.



Scheme 2. 1 Molecular structures of **18**-**24** compounds**.**

Docking studies were performed on optimized **18**-**24** compounds with the HDLP binding pocket,[122] using AutoDock 3.0.5 software,[55] which has been successfully used in the interpretation of the inhibitory activity of several HDAC ligands.[156,161] In addition to the compounds optimized as described above, we used as model receptor the HDLP active site refined at QM level,[147] in order to improve the calculations with the aim to obtain a good qualitative accordance between theoretical K_{Dcalc} and biological assays results. According to the receptor mapping, besides the binding channel (11 Å) there are four

hydrophobic cavities (A-D)[161a] acting as molecular recognition domain (Figure 2.7).

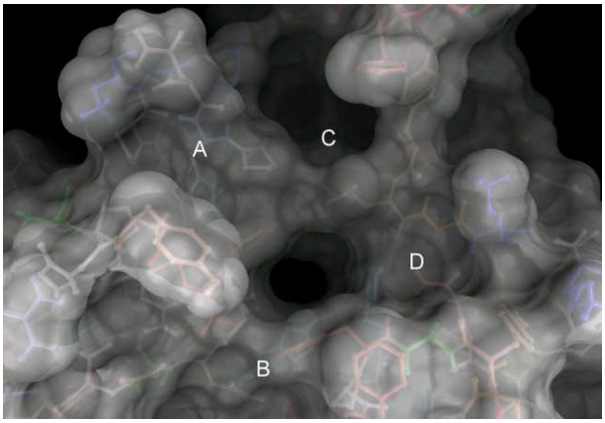


Figure 2. 7 3D model of the HDLP.

The results obtained, reported in Table 2.2, showed satisfactory K_{Dcalc} values for all compounds, even if among them, compounds **23** and **24** showed a little better binding properties for HDLP receptor surface.

Table 2. 2 Calculated (K_D) activities of **1-7** compounds.

	18	**19**	**20**	**21**	**22**	**23**	**24**
K_{Dcalc}	1.78×10^{-8}	4.09×10^{-8}	7.81×10^{-8}	2.74×10^{-8}	3.65×10^{-7}	8.88×10^{-9}	1.82×10^{-9}

Concerning all cyclic molecules, (**18-20, 23-24**), docking studies suggest strong interactions between the recognition binding domain, represented by the hetero cyclic framework, and the hydrophobic surface of HDLP active site. For sake of simplicity, here the detailed docking results for compounds **18** and **7** will be only described.

The docking studies indicate that the linker chain and the cap group of **24** and **18** fill equivalent spaces (hydrophobic pocket D: His170, Ala197, Leu265, Phe198, and Phe200, see Figure 2.8). Moreover, the linker chain exerts a set of interactions with the tubular hydrophobic pocket and the zinc- coordinating carboxylate group, forming hydrogen bonds with $H^{\varepsilon 2}$ of H131.

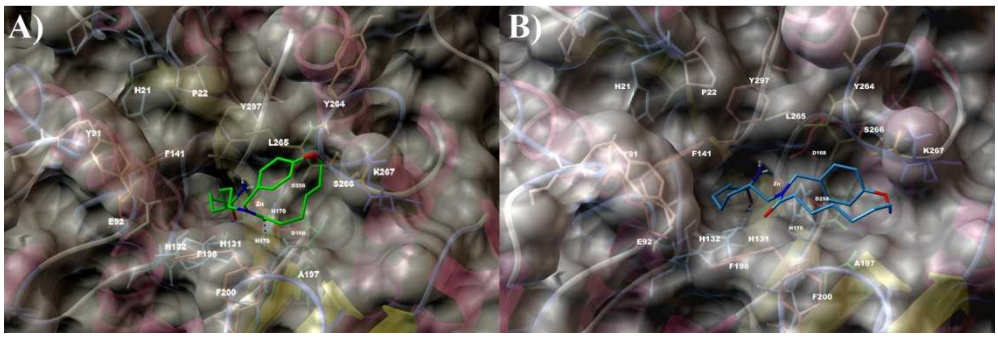


Figure 2. 8 (A) 3D model of the interaction between **18** and the HDLP binding site. The protein is represented by molecular surface and sticks and balls. **18** is depicted by sticks (by atom type: C green, polar H white, N dark blue, O red). (B) 3D model of the interaction between **24** and the HDLP binding site. The protein is represented by molecular surface and sticks and balls. **24** is depicted by sticks (by atom type: C sky blue, polar H white, N dark blue, O red).

The cap group portion of both compounds is accommodated in a shallow groove, establishing Van der Waals interactions and hydrogen bonds with the receptor counterpart, formed by His170, Aala197, Phe198 and Phe200 residues. The cap group extends its hydrophobic contacts thanks to the phenyl ring, which is accommodated in the deep pocket D (Tyr264, Leu265, Ser266, Lys267, see Figure 2.8); however, the macrocycle size seems to slightly modulate the activity, as emerged by directly comparing **18** with **24**; this last, in fact, in virtue of its higher dimension of the cycle (13 C atoms *vs* 15, respectively) showed to fit better with the enzyme binding domain. In fact, the

suboptimal hydrophobic interactions are responsible for a predicted increase in the binding affinity to the receptor of about 10-fold (K_{Dcalc} of **18** 1.78 x 10^{-8} *vs* K_{Dcalc} of **24** 1.82 x 10^{-9}).

Furthermore, comparing docking results of **19** and **23** (Figure 2.9), presenting the same cap group but differing for the linker length, it is possible to suppose that this last should be of 9 carbon atoms in order to have the optimal fit with receptor surface.

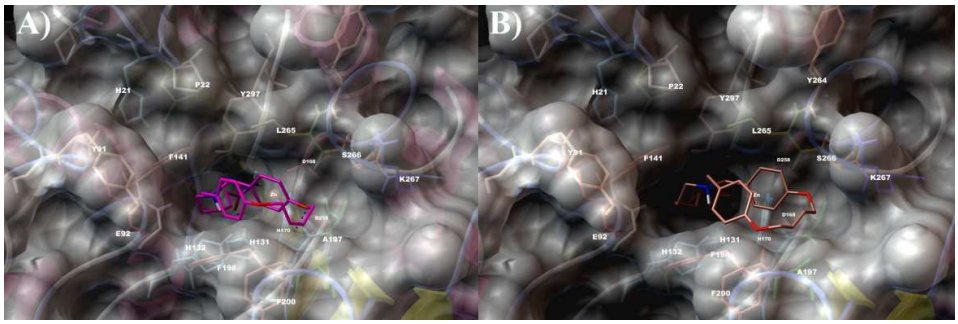


Figure 2. 9 (A) 3D model of the interaction between **19** and the HDLP binding site. The protein is represented by molecular surface and sticks and balls. **19** is depicted by sticks (by atom type: C violet, polar H white, N dark blue, O red). (B) 3D model of the interaction between **23** and the HDLP binding site. The protein is represented by molecular surface and sticks and balls. **23** is depicted by sticks (by atom type: C pink, polar H white, N dark blue, O red).

Finally, compound **20**, which presents the same linker of **23**, but an entirely aliphatic cap group (Figure 2.9), showed to be not able to establish strong hydrophobic interactions with the protein counterpart, being totally lacking of aromatic elements with respect to the others cyclic compounds.

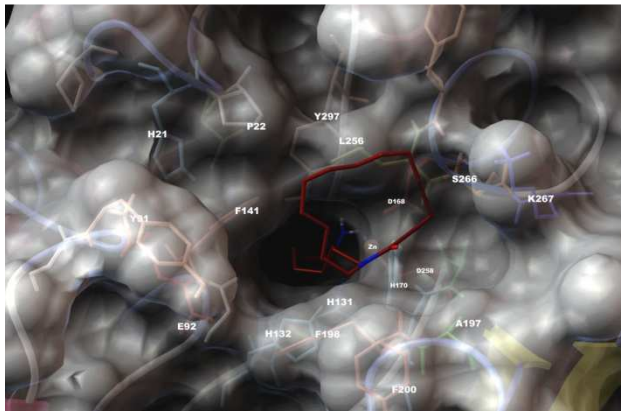


Figure 2. 10 3D model of the interaction between **20** and the HDLP binding site. The protein is represented by molecular surface and sticks and balls. **20** is depicted by sticks (by atom type: C dark red, polar H white, N dark blue, O red).

For the analysis of linear compounds **21** and **22**, the small difference of their K_{Dcalc} values, could be ascribed to their different aromatic rings content. In fact, the two aromatic rings, present in compound **21,** showed to correctly accommodate in the A and D hydrophobic pockets of the enzyme, increasing the stability of the drug-receptor complex. (Figure 2.11).

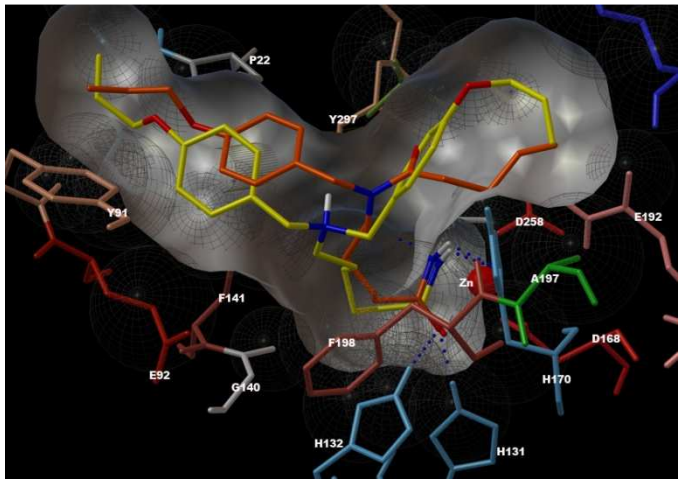


Figure 2. 11 21 and **22** superimposition in the zinc-binding site.

Prompted by the virtual screening results, the synthesis of the designed molecules was undertaken with the aim to verify the qualitative accordance between the theoretical and the experimental data, using a combination of solution and solid phase techniques to synthesize the compounds and, a ring closing metathesis (RCM) reaction as the key synthetic step to obtain the rings. A cytotoxicity assays was then performed on the synthesized molecules and the analysis of the results, reported in Table 2.3, was disappointing for some compounds but, at the same time, allowed to make some considerations on virtual screening outcome.

Table 2. 3 Biological (IC_{50}) activities of compounds **18-24** on HEK-293, J774A.1 and WEHI-164. Control cells viability was designated as 100%, and results were expressed as the concentration of tested compounds able to induce the 50% of mortality in cells (IC_{50}). Results are expressed as mean ± s.e.m. from at least three-independent experiments

	18	**19**	**20**	**21**	**22**	**23**	**24**
IC_{50} [M] HEK-293	$1.85 \times 10^{5} \pm 0.04$	n.d.	n.d.	$7.49 \times 10^{-5} \pm 0.72$	n.d.	$1.4 \times 10^{-5} \pm 0.03$	$1.0 \times 10^{-4} \pm 0.02$
IC_{50} [M] WEHI-164	$2.0 \times 10^{-5} \pm 0.05$	$3.0 \times 10^{-4} \pm 0.02$	n.d.	$3.2 \times 10^{-4} \pm 0.21$	n.d.	$2.4 \times 10^{-5} \pm 0.05$	$2.2 \times 10^{-5} \pm 0.08$
IC_{50}[M] J774A.1	$2.4 \times 10^{-5} \pm 0.18$	n.d.	n.d.	$7.38 \times 10^{-5} \pm 0.12$	n.d.	n.d.	$2.4 \times 10^{-4} \pm 0.13$

For example, the predicted negative response for some of the designed molecules (in good qualitative accordance with the experimental data) can be ascribed to unfavorable structural features, such as the lack of aromatic rings in compound **20**, while, the unsuitable size of the aliphatic chain spacer in **19** could account for the absence of cytotoxicity, in coherence with its slightly less satisfactory K_{Dcalc} value. On the contrary, as concerns compounds **23** and

24, the biological results were in disagreement with the virtual screening response. In fact these two compounds, despite the best affinity properties displayed in docking studies, did not exert the expected higher potency in antiproliferative assay, probably due to unfavorable pharmacokinetic parameters – such parameters are in fact not quantifiable in the computer simulation approach.

On summary in this paragraph it was reported the development of a new collection of potential HDAC inhibitors and a partial rationalization of their biological behaviour through molecular docking calculations. Four of the seven synthesized molecules showed a satisfying level of antiproliferative activity and all of them are currently investigating for their ability to inhibit HDAC enzyme.

2.1.1. Computational details

Molecular mechanics/dynamics (MM and MD) calculations were performed using the Macromodel 8.5 software package[159] and the MMFFs[162] force field at several temperatures (400, 600 and 800K). The solvent effects are simulated using the analytical Generalized-Born/Surface-Area[163] (GB/SA) model mimicking the presence of H_2O. All the structures were minimized using a Polak-Ribiere Conjugate Gradient (PRCG, 50000 steps, convergence threshold 0.005 kJ mol−1 Å−1). Autodock 3.0.5[55] was used for all docking calculations. HDLP[122] (histone deacetylase-like protein) is a metalloprotein, so a nonbonded model for metallic center according to the nonbonded Zn parameters of Stote[164] (Zinc Radius=1.10 Å, well depth=0.25 kcal/mol) was used. In order to have an accurate weight of the electrostatics, we derived the partial charge of Zn=1.175 and of the aminoacids involved in the catalytic center (Ala169, His170, Asp168, Asp258) by DFT calculations B3LYP level

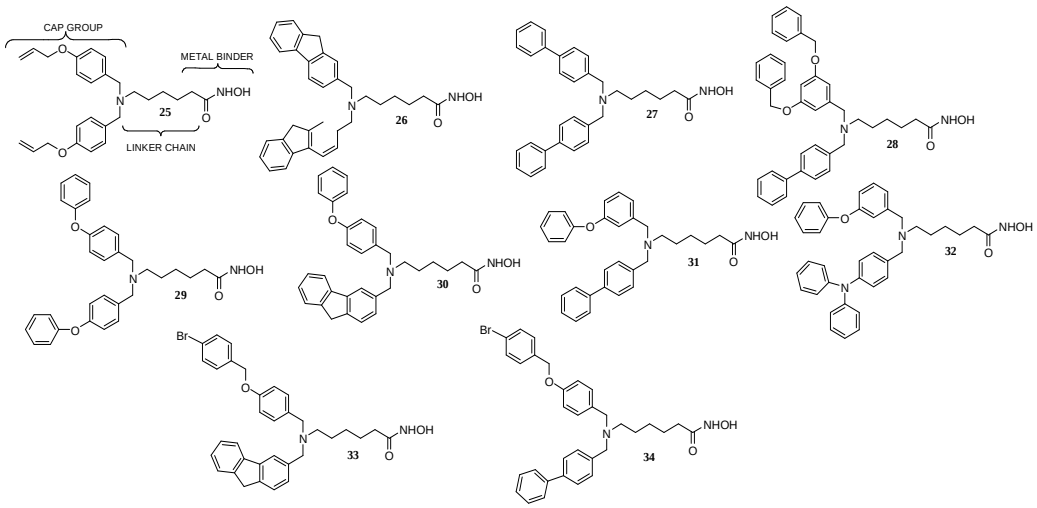
by the 6-31G(d) basis set and ChelpG method[148] for population analysis (Gaussian 03 Software Package).[160] For what concerns the ligands, the geometries were optimized atDFT B3LYP level using the 6-31G(d) basis set. Subsequently, the charges of compounds **1**-**7** were calculated with the ChelpG method[148] at the B3LYP/6-31G+(d) level. The above calculated charges were used for docking calculations. For all the docking calculations a grid box size of 66 x 64 x 48 with spacing of 0.375 Å between the grid points, centred between Zn^{2+} and His170 (x=49.75, y=5.0, z=101.491) and covering the catalytic centre surface of HDLP was used. For all the docked structures, all bonds were treated as active torsional bonds except the amide bonds. In order to achieve a representative conformational space during the docking calculations, six calculations consisting of 256 runs were performed, obtaining 1536 structures (256 x 6). The Lamarckian genetic algorithm was used for dockings. An initial population of 450 randomly placed individuals, a maximum number of 4.0×10^6 energy evaluations, and a maximum number of 3.0×10^6 generations were taken into account. A mutation rate of 0.02 and a crossover rate of 0.8 were used. Results differing by less than 3.5 Å in positional root-mean-square deviation (RMSD) were clustered together and represented by the result with the most favourable free energy of binding. All the 3D models were depicted using the Phyton software:[165] molecular surfaces are rendered using Maximal Speed Molecular Surface (MSMS) [166]

2.3 Design, synthesis, and biological activity of new hydroxamic tertiary amines as histone deacetylase (HDAC) inhibitors.

Following the results of a collection of new mono and bis amides reported in the previous paragraph, here, it will be described the design, docking studies, synthesis, and biological evaluation of new generation of hydroxamic tertiary amines.[167] With this second series of compunds reported below, it is explored the efficiency of non-peptidic linear compounds presenting the hydroxamic group as metal binder and two differently substituted arms as cap group.

In particular for the design of a small set of hydroxamic tertiary amines, it was respect the well established[168] structural features of a standard HDAC inhibitor: I) a hydrophobic region (cap group) involved in the molecular recognition process; II) a Zn^{II} chelating element (metal binder); and III) a five to seven-atoms spacer (linker) between the cap group and the metal binder. Specifically, the same linker chain, presenting the hydroxamic group, was conjugated with different aromatic moieties in order to optimize the size and the chemico-physical properties of the cap groups (**25**-**34**, Scheme 2.2).

For the modeling studies, Autodock 3.0.5 software[55] and the optimized HDLP model[147] (see previous sections pdb code 1C3R)[122] (Figure 2.12) as model enzyme, which has been successfully used in the interpretation of the inhibitory activity of several HDAC ligands,[161,169] were used.



Scheme 2. 2 Structures of compounds **25-34**.

The calculated affinity, expressed as the most favourable HDAC free energy of binding, and the three dimensional models of the complex between compounds **1-10** (Scheme 2.2) and the HDLP protein, are reported in Figure 2.13-15.

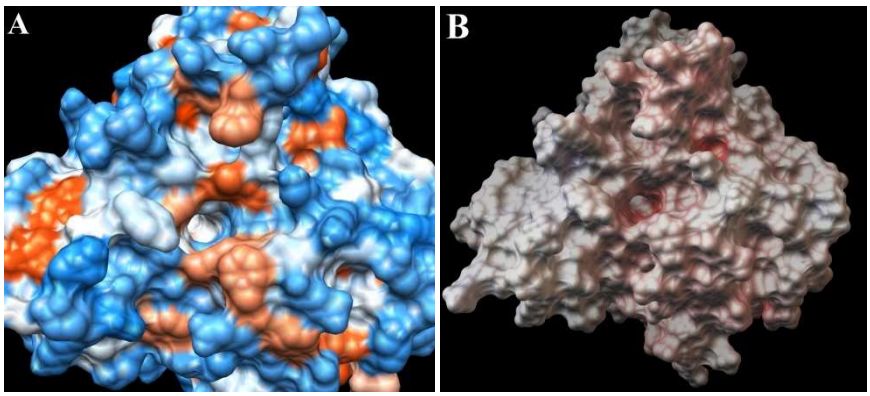


Figure 2. 12 Molecular surface of the Histone Deacetylase Like Protein represented by molecular surface colored according to the hydrophobicity (A) (dodger blue=hydrophilic, orange red=hydrophobic), and to the electrostatic (B) (positive potential=blue and negative potential=red).

The docking calculations point out that all the designed molecules fill the equivalent space on the surface protein, and that they are able to occupy the catalytic binding site coordinating the fundamental zinc ion (Figures 2.13 and 2.15), exhibiting thus a potential antagonist activity.

Concerning the molecules with the same arms (**25**-**27**, and **29**), predicted energy evaluation suggests a good ligand efficiency[170] (binding energy for heavy atom molecular ΔG/NHA) and strong interactions between the cap group and the hydrophobic surface of the HDLP active site (See Figure 2.13-2.15). In particular, **29** is the most promising compound among this subset of molecules.

In analogy with **25-27** it forms three hydrogen bonds with the aminoacids of the catalytic site (namely His131, 132 and 170) by the hydroxamic portion (See Figure 2.14), and, furthermore, it also projects the oxigen atoms of the arms in close contact with the NH of Phe198. Moreover, even if compound **27** establishes a cation-π interaction with Lys19, only compound **29** is able to generate π- π interactions with the Phe141 and Phe198 on the protein surface. On the other hand, the different arm present in **33** and **34** is not involved in further interactions with the macromolecule in comparison to **26** and **27**, and because the majority of the hydrophobic contacts are maintained (See Figure 2.14), these arms may be considered equivalent.

Moreover, as compound **31** is a hybrid of **27** and **29**, it exhibits as expected an intermediate value of calculated free energy of binding with respect to them, presenting different π-π interactions with the Phe200 and Tyr297 on the protein surface, and a different and less efficient pattern of hydrogen bonds (See Figure 2.14).

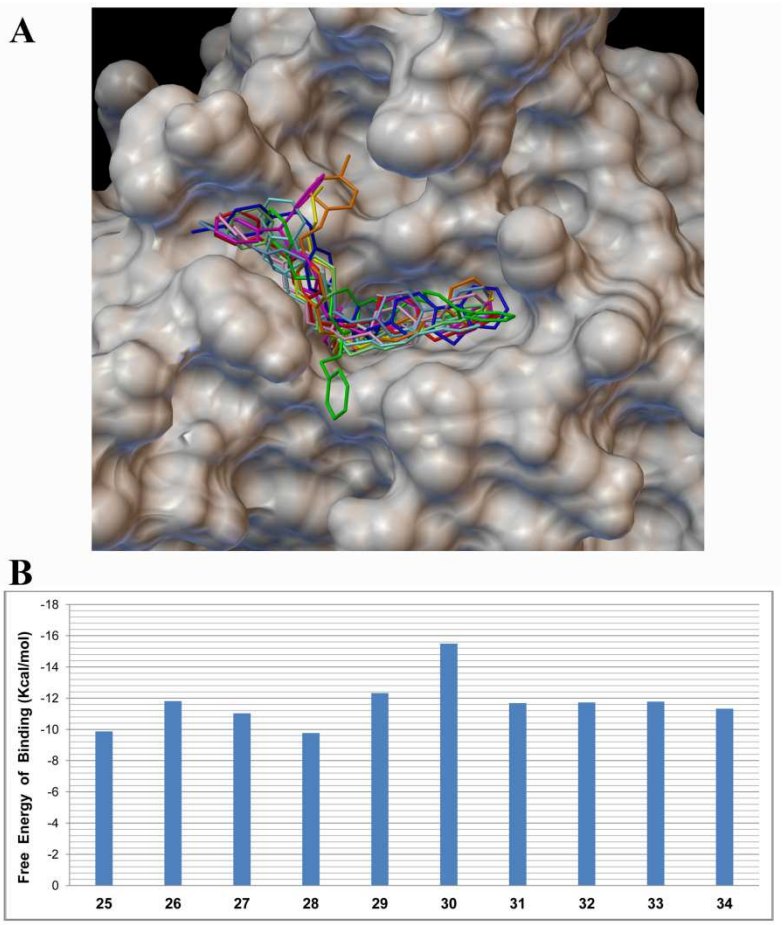


Figure 2. 13 (A) Superimposition of the compounds **25**-**34** in the HDLP binding site. The protein is represented by molecular surface and **25**-**24** are represented by tube, and coloured as: **25**, yellow; **26**, cyan; **27**, light blue; **28**, green; **29**, red; **20**, light green; **31**, pink; **32**, purple; **33**, blue; **34**, orange. (B) Calculated Free Energy of Binding expressed as kcal/mol of compounds **25**-**34** for HDLP.

On the contrary, compound **30** favorably combines the features of **26** and **29,** showing the best calculated free energy of binding among the all investigated molecules thanks to the efficient hydrophobic, hydrophilic and electrostatic contacts with the enzyme (See Figure 2.14).

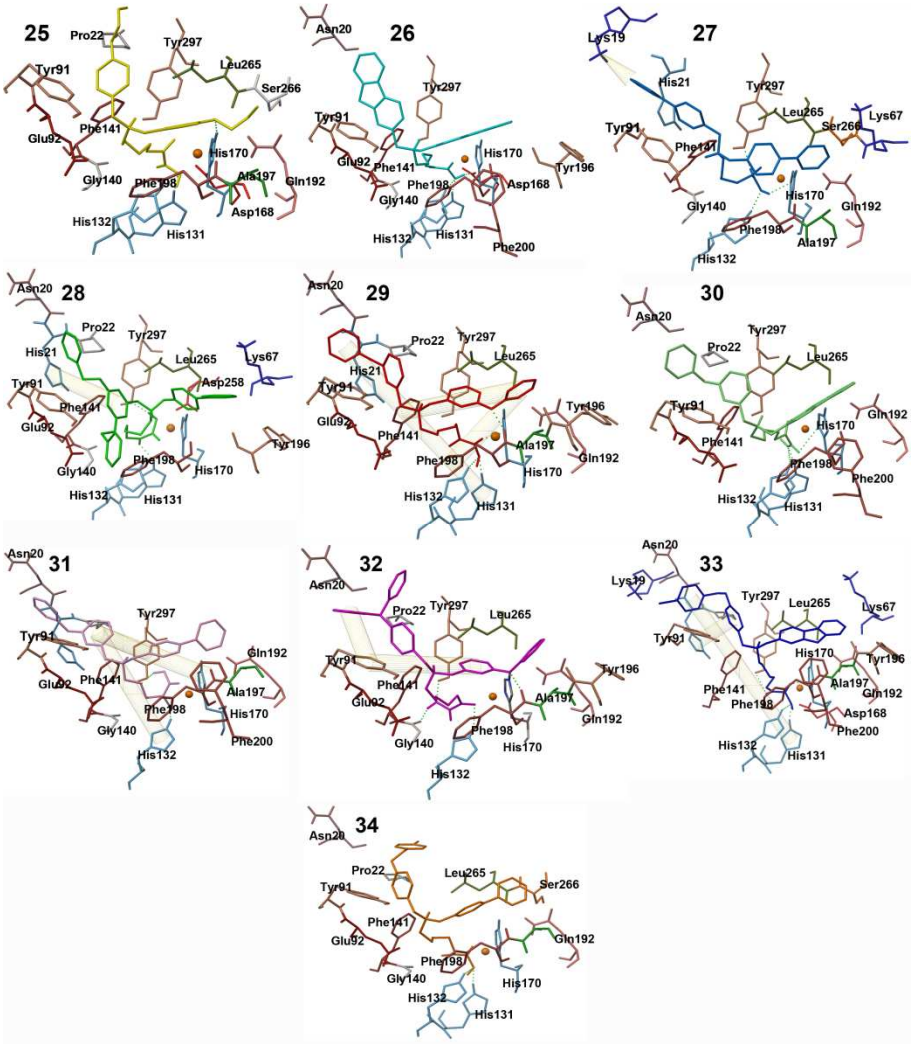


Figure 2. 14 Three-dimensional models of the complexes between compounds **25**-**34** and HDLP molecule. The side chains of amino acids and the compunds are represented by tube.

On the contrary, the **28** and **32**, presenting the most encumbering arms of our set of compounds show the worst ligand efficiency[170] values (Figure 2.15), because the bigger dimensions of the cap group does not balance the gain in energy derived from the further hydrophobic interactions with the target.

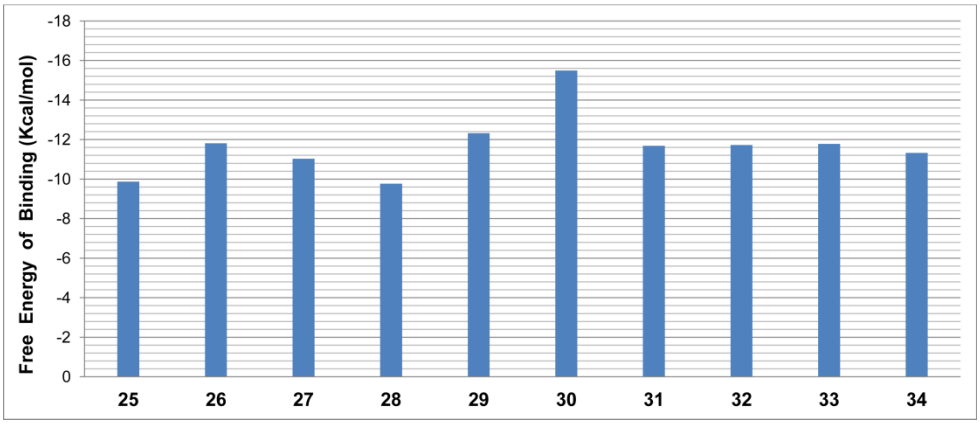


Figure 2. 15 Ligand efficiency of **25**-**34** with the HDLP enzyme.

On the basis of the above rationale, it was virtually evaluated the influence of the cap group dimension on the HDAC activity of a small set of hydroxamic tertiary amines, where the hydrophobic and/or the aromatic stacking interactions seem to be among the main driving forces of the target-ligand complexes formation. In fact, the molecular docking results revealed that the simultaneous and efficient interactions of **30** with both the enzyme surface and the tubular binding pocket, due to the proper selection of its arms, are critical for the potential antagonist activity. Prompted by the above *in silico* results, the synthesis of the designed molecules was undertaken in order to verify the qualitative accordance between the theoretical and the experimental data.

In particular for the synthesis of compounds **25**–**34**, the N-Fmoc-hydroxyamine 2-chlorotrityl resin was used for the production of the hydroxamic acids by solid-phase organic synthesis (SPOS).[171] Alkylation of the primary amine with the aldehydes (5 equiv) on solid phase was performed in a two-step procedure to minimize dialkylation,[172] employing $(MeO)_3CH$ (TMOF) as both solvent and dehydrating agent.[173] Finally, the desired

compounds were cleaved from the resin by a single-step treatment with a solution of TFA/TIS/CH_2Cl_2.

Inhibition of HDAC activity in HeLa nuclear extracts by compounds **25**-**34** and Trichostatin A (TSA, **5**) a well known HDAC inhibitor, was measured using a fluorescence-based assay. The IC_{50} values of compounds **25**-**34** are reported in Table 2.4.

Table 2. 4 In vitro HDAC-inhibitory activity (IC_{50} ± sd) of compounds **25**-**34** and TSA

Compound	IC_{50} (μM) [a]
25	1.00 (±0.08)
26	0.64 (±0.064)
27	0.19 (±0.017)
28	1.20 (±0.11)
29	0.47 (±0.037)
30	0.07 (±0.004)
31	0.58 (±0.047)
32	0.97 (±0.078)
33	0.18 (±0.016)
34	0.17 (±0.008)
TSA[b]	0.022 (±0.002)

[a] Mean values of at least two-independent assays are reported. Standard deviations are in parenthesis. [b] TSA was used as a control for these experiments.

On summary, differently decorated tertiary amines bearing hydroxamic acids as chelating element have been successfully developed as a new class of potential HDAC inhibitors, and a rationalization of their biological behavior has been pursued through molecular docking calculations. In particular, a good accordance between molecular modeling predictions and biological results

was found, and in fact, all the synthesized compounds displayed a considerable HDAC inhibition activity. In particular, as predicted by docking calculations, compound **30** showed the highest inhibitory activity in the nanomolar range (IC_{50} 0,07 μM **30** *vs* 0,022 μM TSA, **5**), using TSA as reference compound. Bearing in mind the biological results and the descriptions of the 3D models of **30** with HDLP discussed above, in this paragraph a complete a rationalization of the molecular docking results related to linear tertiary amines based inhibitors was offered. In particular, the presence of hydrophobic and bulky aromatic substituents as cap group, necessary for hydrophobic interactions, and π-π stacking, seems to be the driving force of the target-ligand complex efficiency. Alongside the intermolecular interactions, the *in silico* results have also revealed the critical role and influence of the nature and dimension of the arms on the potential HDAC activity. Such evidences are also confirmed by calculated and experimental **28** inactivity.

2.3.1. Computational Details

As previously reported, for docking calculations the optimized HDLP[147] DFT/B3LYP was used. Moreover, in the docking calculations the zinc ion was treated with the Stote[164] parameters: radius of 1.10 Å, and well depth of 0.25 kcal/mol. The molecular docking calculations were performed by Autodock 3.0.5[55] on quad-core Intel Xeon 3.4 GHz, using a grid box size of 66 x 64 x 48, with spacing of 0.375 Å between the grid points, and centered at 49.75(x), 5.0 (y), 101.491 (z), covering the active site of the HDLP To achieve a representative conformational space during the docking studies and for taking into account the variable number of active torsions, 6 calculations consisting of 256 runs were performed, obtaining 1536 structures for each ligand. The

Lamarckian genetic algorithm (LGA) was employed for docking calculations, choosing an initial population of 450 randomly placed individuals. The maximum number of energy evaluations and of generations was set up to 5 x 10^6 and to 4 x 10^6 respectively. Results differing by less than 3.5 Å in positional root-mean-square deviation (RMSD) were clustered together and represented by the most favorable free energy of binding. Illustrations of the 3D models were generated using Chimera[174] and the Python software.[165]

2.4. Conformationally Locked Calixarene-Based HDAC Inhibitors

In the last decade several examples of calixarene derivatives able to interact with molecules of biological interest, have been reported.[175] In particular, Hamilton has designed calixarene derivatives able to bind to protein surfaces and to block biologically important protein–protein interactions.[176] Interestingly, the treatment of nude mice bearing human tumors with peptidocalix[4]arene derivatives, able to selectively bind to platelet-derived growth factor (PDGF), resulted in a significant inhibition of tumor growth and angiogenesis.[176c] In addition, it was demonstrated that calixarene-based therapeutic agents do not show any toxic eff ect in mice tests.[176c,177]

Recently, it was demonstrated a surface recognition of tissue and microbial transglutaminases by peptidocalix[4]arene diversomers.[178] Larger calix[8]arene derivatives have shown competitive inhibition of recombinant human tryptase.[179] Water-soluble *p*-sulfonatocalixarenes have shown interesting biological activities, including anti-viral, anti-bacterial, and anti-thrombotic activity.[180] Recently, Ungaro and coworkers have reviewed the properties of calixarene-based multivalent ligands in lectin binding and inhibition, DNA condensation, and cell transfection.[181]

On this basis, it was decided to investigate the use of calix[4]arene scaffold to construct novel inhibitors of Histone deacetylase enzymes (HDACs) and, in this paragraph, it was reported the result of this study.[169] In this field, *Nature* provides a number of related cyclic scaffolds with HDAC inhibitory activity.[168]

Regarding the structural features of HDACi, Ghadiri has recently shown that an appropriate increased structural rigidity enhances the inhibitory activity.[182]

On the basis of these structural considerations,[183] we envisioned that preorganized calix[4]arenes[184] locked in the *cone* conformation, by propoxy groups at the *endo* rim, and endowed with the appropriate functional groups at the *exo* rim to bind to the receptor surface (Figure 2.16, top) should be ideal candidates as potential HDAC inhibitors.

Concerning the functional groups, first, a carbon aliphatic chain entering the binding channel E (channel linker, Figure 2.16, top) and bearing a metal binder for the Zn^{2+} coordination is required (Figure 2.16, bottom). Second, hydrophobic arms (cap groups, Figure 2.16, bottom) able to fit the four external hydrophobic pockets A-D (Figure 2.16, top) on the enzyme surface should be necessary.

To direct the synthesis toward derivatives with higher activity, it was performed a molecular docking study of a significant set of designed calix[4]arenes variously substituted at the *exo* rim with aliphatic or aromatic groups of different size and hydrophobicity (Figure 2.16).

In accordance with the current synthetic possibilities, the amide linkage[185] was selected to attach the above moieties at the upper rim of calixarene scaffold.

Molecular docking studies were articulated in the following steps: I) choice of the metal binder group and length of the linker chain, II) choice of the cap groups by gradually increasing the length of an amide aliphatic chain (from one to five carbon atom), and III) choice of aromatic cap groups by gradually increasing the size of the aromatic system (from one to four condensate rings).

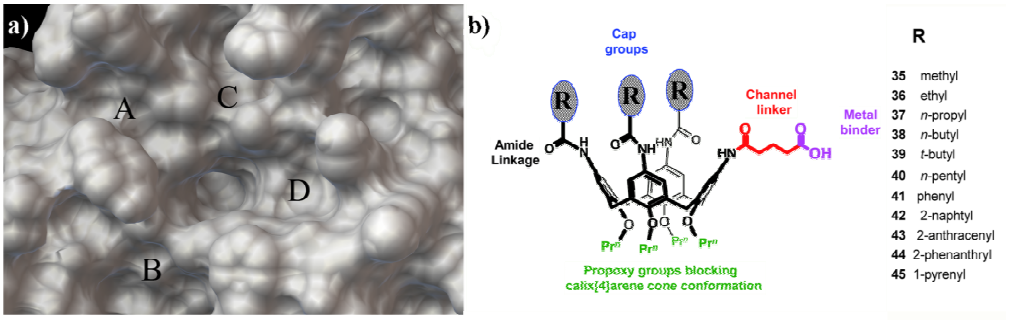


Figure 2. 16 (A) A-D: Hydrophobic pockets on the HDLP (Histone Deacetylase Like Protein) surface. E: Zn^{2+} binding channel. (B) Structural features of calix[4]arene derivatives candidate for HDAC inhibition.

To rationalize and to identify the structural features of each calixarene derivative, molecular docking studies was performed by using Autodock 3.0.5 program,[55] on the HDLP (Histone Deacetylase Like Protein, PDB code 1C3R)[122] binding site (Figure 2.16, top). For the linker chain and metal binder Class I selective HDAC inhibitor Azumamide E[186] was used as compound reference and then a six membered chain ending with the carboxy functionality was used. This latter is usually considered dissociated at physiological pH and, consequently, the deprotonated form was considered in the calculations. Subsequently, it was systematically conjugated the amide linker at the upper-rim with different alkylic and aromatic groups in order to optimize the size and the chemico-physical properties of the cap group. In Figure 2.17 we report the calculated affinity, expressed by a theoretical inhibition constant ($K_{Dcalc.}$) of molecules **35-45** (Figure 2.16) for HDLP target.

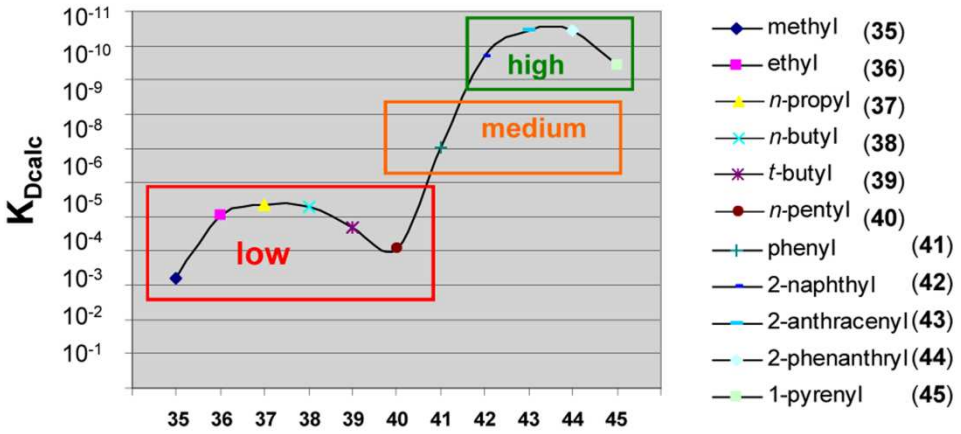


Figure 2. 17 Variation of the final calculated inhibition constant (K_{Dcalc}) as a function of different alkylic and aromatic groups connected at the amide linker at the upper rim of calix[4]arenes **35**-**45**.

As it is possible to observe in Figure 2.17, the K_{Dcalc} values are strictly dependent on the size of the arms. In particular, it is possible divide the curve in three different zones: low, medium, and high theoretical activity. Interestingly, the low zone corresponds to amidocalix[4]- arenes **35**-**40** bearing alkylic arms, whereas the medium and high zone of Figure 2.17 correspond to the series of amidocalixarenes **41**-**45** bearing aromatic arms of increasing size.

In the best arrangement of the low zone obtained with the propylic arms (derivative **37**, Figure 2.18c) it is important to note that the carboxylate group (metal binder) is of primary importance because of its network of interactions: it coordinates the zinc ion in a bidentate fashion and establishes hydrogen bonds with $H^{\varepsilon 2}$ of His131 and $H^{\varepsilon 2}$ of His132.

The six element chain (channel linker) makes stabilizing hydrophobic contacts with the zinc-containing tubular pocket. The aromatic rings of calix scaffold establish π stacking with Tyr196, Phe200 and Tyr297 and van der Waals interactions with the hydrophobic cavities B, C and D on the HDLP

surface (Figure 2.18c). The propylamido arms are correctly accommodated in the D and C hydrophobic pockets (constituted by Leu265, Ser266, Lys267, His170, Ala197, Phe198, Phe200, and by Pro22, Phe141, Tyr264, Tyr297, His298, Phe338, respectively). In the latter instance, a stabilizing H bond also occurs between the C=O of Tyr264 and the NH of propylamido arm.

Moreover, additional van der Waals interactions between one propoxy group at the lower rim and the macromolecular counterpart (pocket A), contribute to the calculated stability of calixarene-enzyme complex. Comparing the putative three-dimensional models of **35**-**40**, even if all the Zn^{2+} binding channel interactions are maintained, it appears clear that the different length of alkylic arms translates in suboptimal hydrophobic interactions with the enzyme surface leading to a decrease in the calculated binding affinity.

The linker chain of **35**-**45** fits into the 11 Å binding channel and the carboxylate moiety binds to the Zn^{2+} ion, at the bottom of the channel, in a bidentate fashion establishing hydrogen bonds with $H^{\varepsilon 2}$ of His132 for **41**-**45**, and $H^{\varepsilon 2}$ of His131 for **42** (see Figures 2.19-20).

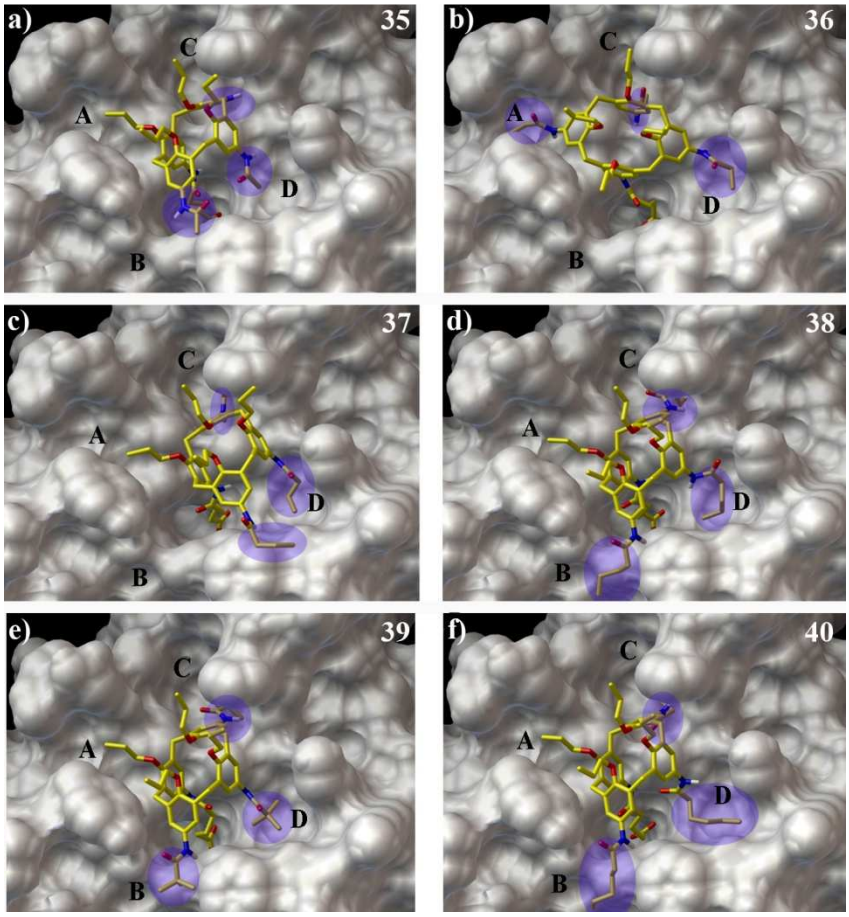


Figure 2. 18 3D model of the interaction between amidocalix[4]arenes **35**-**40** with alkylic arms and the HDLP binding site. The protein is represented by molecular surface, while compounds 1-6 are depicted by sticks (by atom type: C yellow, polar H white, N dark blue, O red). A-D: Hydrophobic pockets on the HDLP surface.

The aromatic arms of **41** and **42** occupy the pockets B, C, and D, while in the other models (**43**-**45**) the analogous arene moieties fill the pockets A, B, and C (Figures 2.19c-e). This different arrangement depends on the size of the aromatic group. As shown in Figure 2.19a and 2.19b the smaller benzene or naphthalene rings of **41** and **42** (Figures 2.19a and 2.19b) easily accommodate in the shallow groove on the enzyme surface formed by Leu265, Ser266,

Lys267, Ala197, Phe198 and Phe200, and establish van der Waals interactions with the protein counterpart (Figures 2.20a-d).

Moreover, in the case of compound **42** an amide function forms an additional hydrogen bond with NH of Leu265 (Figures 2.20c and 2.20d), while two similar additional hydrogen bonds with $N^{\varepsilon 1}$ of His170 and OH of Tyr91 (Figures 2.20e and 2.20f) are observed for compound **45**. On the other hand, the bulkier anthracene, phenanthrene or pyrene rings of the others ligand form a π-stacking interaction with Tyr91 (pocket A) for **43** and **44**; Phe134 (pocket B) for 10; Tyr264 (pocket C) for **43** and **44**; Tyr297 (pocket C) for 9 (see Figures 2.20e-f). Moreover, these aromatic rings establish cation-π interactions with Lys19 (pocket A) for **43** and **44** (Figure 2.20f).

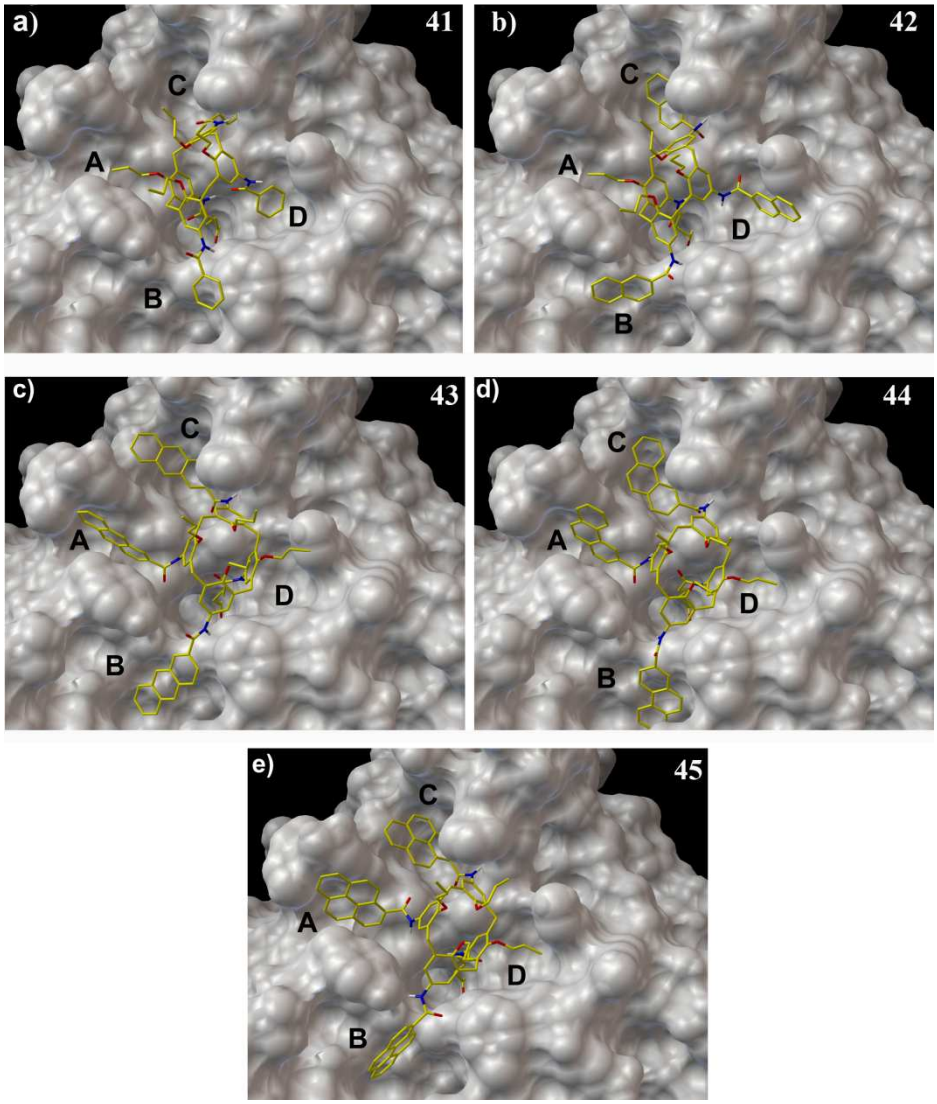


Figure 2. 19 3D model of the putative binding mode of amidocalix[4]arenes **41**-**45** with aromatic arms and the HDLP binding site. The protein is represented by molecular surface, while compounds **41**-**45** are depicted by sticks (by atom type: C yellow, polar H white, N dark blue, O red). A-D: Hydrophobic pockets on the HDLP surface.

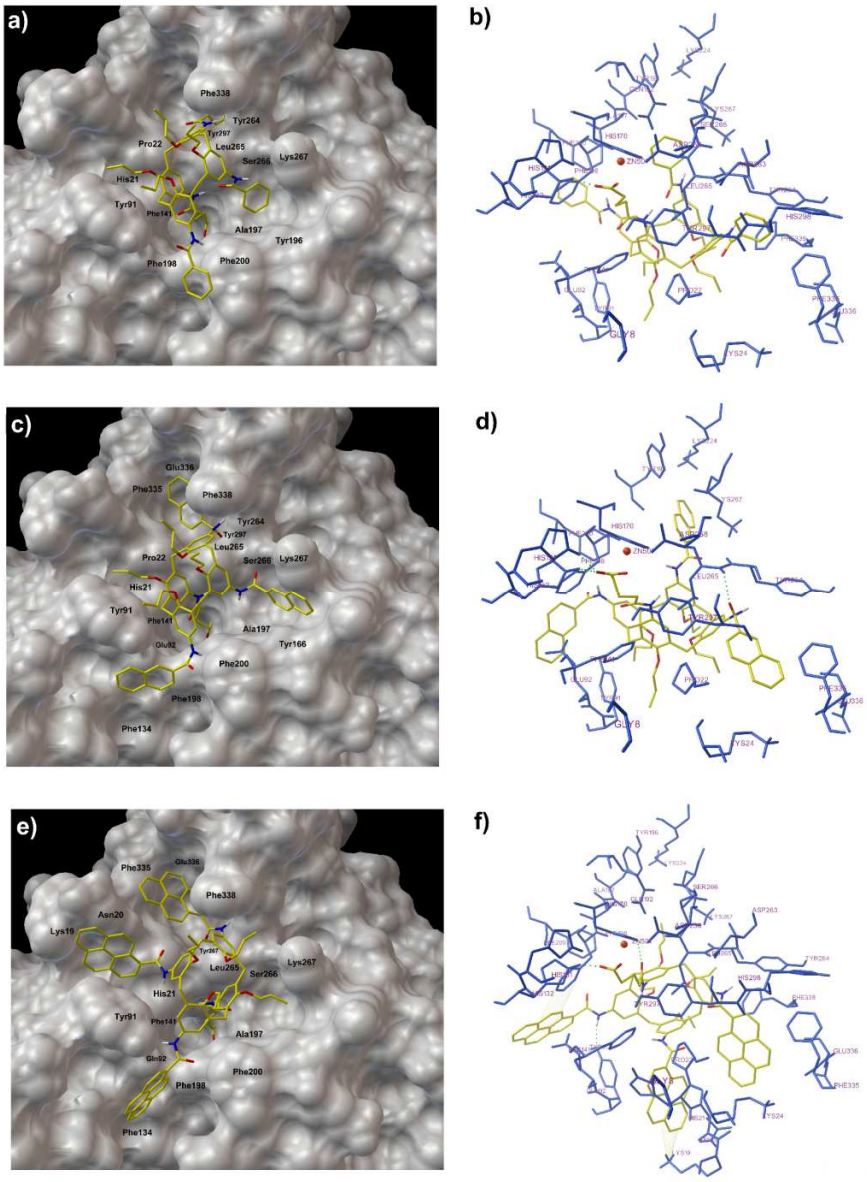


Figure 2. 20 Front view (A, C, and E) and back view (B, D, and F) of 3D model interactions between amidocalix[4]arenes **41**, **42** and **45** and the HDLP binding site. In the front view (A, C, and E), the protein is represented by molecular surface, while compounds **41**-**4**5 are depicted by sticks (by atom type: C yellow, polar H white, N dark blue, O red). In the back view (B, D, and F) the protein and compounds **41**-**45** are depicted by sticks coloured by blue and atom type: C yellow, polar H white, N dark blue, O red, respectively.

In particular, it was considered the most promising candidate **43** (Figure 2.21) to pin down the main features of new potential calixarene-based HDACi.

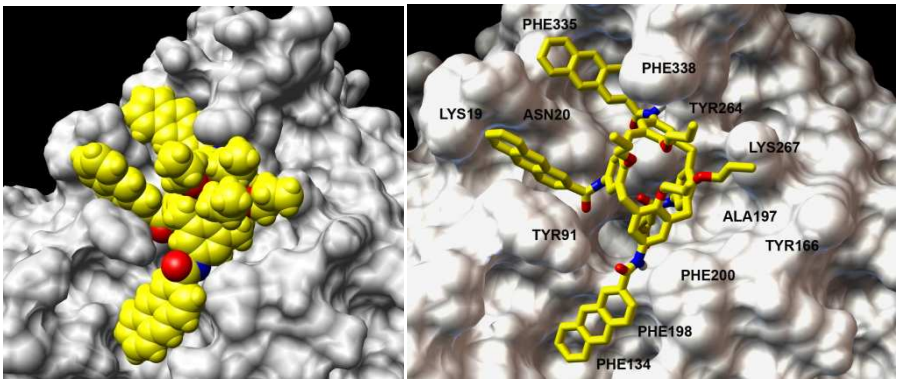


Figure 2. 21 3D model of the putative binding mode of amidocalix[4]arene **43** and the HDLP binding site. The protein is represented by its molecular surface (gray), while **43** is depicted as a CPK (top) or stick model (bottom) (colored by atom type: C, yellow, polar H, white; N, dark blue; O, red).

The linker chain of **43** fits into the 11 Å binding channel, and the carboxylate moiety binds to the Zn^{2+} ion, at the bottom of the channel, in a bidentate fashion establishing hydrogen bonds with $H^{\varepsilon 2}$ of His132; moreover, an amide function forms an additional hydrogen bond with $N^{\delta 1}$ of HIS170. The aromatic arms of **43** occupy the pockets A, B, and C establishing van der Waals interactions with the enzyme counterpart, a π-stacking interaction with TYR91 (pocket A) and TYR264 (pocket C), and a cation-π interaction with LYS19 (pocket A).

The docking results show that the influence of the groups at the upper rim is mainly related to hydrophobic or aromatic stacking interactions, which seem to be among the main driving forces of the target-ligand complexes.

To verify the above *in silico* results, it was decided to check the key points of the predicted curve represented in Figure 2.17: the most active candidate of the low zone (**37**), the intermediate (**41**), and three derivatives of the high zone, namely, **42**, **43**, and **45**. The five derivatives **37**, **41**, **42**, **43**, and **45** were synthesized easily obtained starting tetraaminocalix[4]arene, already locked in the cone conformation by the four propoxy groups at the endo rim according.

In vitro evaluation of inhibition of HDAC activity in HeLa nuclear extracts was performed by a fluorescence-based assay. The IC_{50} values of compounds **37**, **41**, **42**, **43**, **45,** and Trichostatin A (TSA, **5**), a well-known HDAC inhibitor, are reported in Table 2.5.

As predicted by the docking studies, alkyl derivative **3** was the less active compound (IC_{50} > 10 μM) followed by phenyl derivative **7** (IC_{50}) 5.10 μM). On the other hand, **8**, **9**, and **11**, bearing larger aromatic rings, displayed higher inhibitory activities (IC_{50}=0.14-0.86 μM), although with less pronounced differences with respect to the predicted ones.

Probably the calculated differences fall within the accuracy limit of the docking method. In any case, results confirm that topology, size, and hydrophobicity of the aromatic arms are the most important determinants for biological activity of this novel class of calix[4]arene inhibitors. In summary, a classic *in silico* screening of a new class of potential HDAC inhibitors was applied obtaining a good predictions of their inhibitory activity before proceeding to their synthesis.

In this way, it was possible to design a new class of amidocalix[4]arenes permanently locked in a cone conformation with convergently predisposed interacting moieties. The *in silico* evaluation of their binding ability toward the HDAC active site allowed us to direct the synthesis only to the most promising candidates, thus avoiding a useless waste of resources.

Table 2. 5 In vitro HDAC-inhibitory activity (IC_{50} ±sd)

Compound	IC_{50} (μM)[a]
37	>10
41	5.10 (±1.00)
42	0.14 (± 0.02)
43	0.14 (± 0.02)
45	0.86 (±0.10)
TSA	0.02 (±0.009)
[a]values are means of three independent experiments. Standard deviation values were < 20% and are reported in parenthesis.	

The subsequent synthesis and enzyme inhibition evaluation fully confirmed the theoretical prediction that arylamidocalix[4]- arenes bearing large aromatic arms constitute moderately active HDACi. Considering that the calixarene frameworks had shown no hint of toxicity in several in vivo biological tests,[176c,177] this work suggests an additional application of a scaffold already used in the fields of biomolecular recognition. Future work will be directed toward the *in silico* screening of nonsymmetrically substituted arylamidocalix[4]arenes that could give an even better fitting on the different enzyme hydrophobic A-D pockets. The influence of the "coneblocking" groups at the lower rim will also be evaluated as well as the possibility to introduce groups able to increase their water solubility.

2.4.1. Computational Details

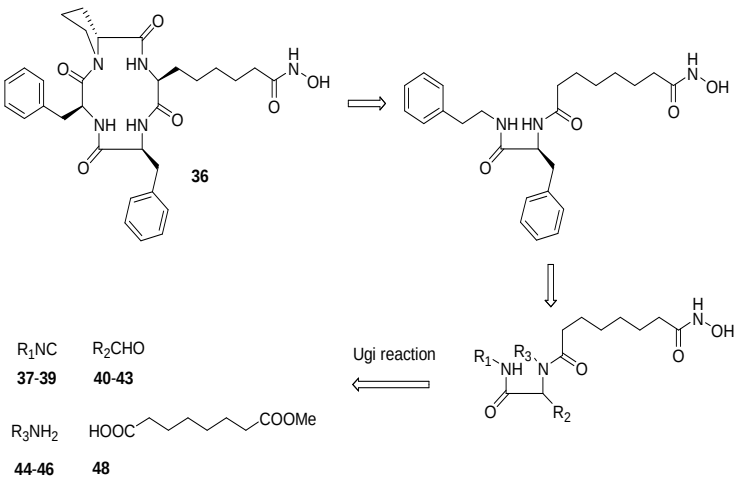
The molecular docking calculations were performed on a two dual-core Intel Xeon 3.4 GHz, using Autodock 3.0.5 software[55] using a already described QM optimized HDLP model. For all the docking studies a grid box size of 66 x 64 x 48 with spacing of 0.375 Å between the grid points, and centered at 49.75 (x), 5.0 (y) and 101.491 (z), covering the four hydrophobic pockets (A-D) on the HDLP surface was used. For all the docked structures, all bonds at the upper rim were treated as active torsional bonds. In order to achieve a representative conformational space during the docking calculations, from three to ten calculations consisting of 256 runs were performed, obtaining 768/2560 structures. The Lamarckian genetic algorithm was used for dockings. An initial population of 600 randomly placed individuals, a maximum number of 5.0×10^6 energy evaluations, and a maximum number of 6.0×10^6 generations were taken into account. A mutation rate of 0.02 and a crossover rate of 0.8 were used. Results differing by less than 3.5 Å in positional root-mean-square deviation (RMSD) were clustered together and represented by the result with the binding energies. All the 3D models were depicted using the Python software:[165] molecular surfaces are rendered using Maximal Speed Molecular Surface (MSMS).[166]

2.5. Synthesis, Biological Evaluation, and Molecular Docking of Ugi Products Containing a Zinc-Chelating Moiety as Novel Inhibitors of Histone Deacetylases

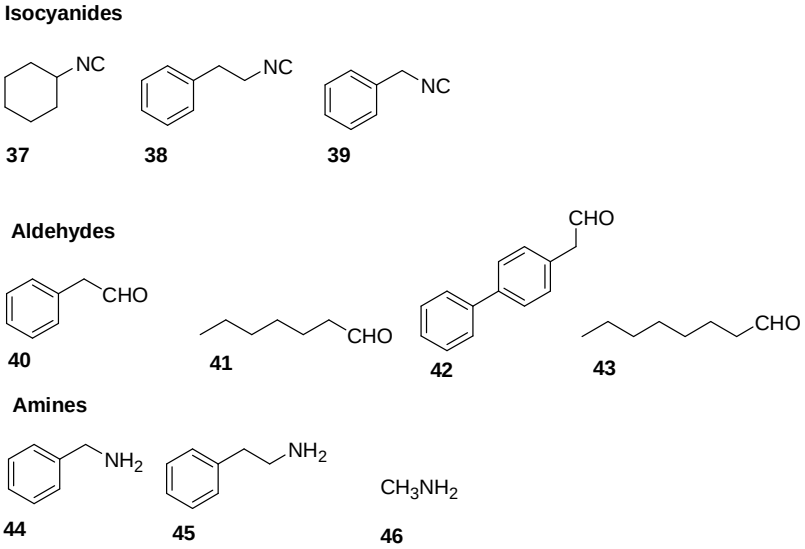
In this paragraph, it is reported the use of molecular docking to rationalize the binding mode of three enantiomers pairs of linear peptides[161a] HDAC inhibitors (**67**, **73**, and **74,** Scheme 2.5) derivatives of CHAP 1 (**36,** Scheme 2.3). **1** is an analogue of trapoxin B in which epoxyketone is replaced with hydroxamate, leading to a compound that reversibly inhibits HDAC at nanomolar concentrations, with a superior in vivo stability compared to trapoxin B.[187] In a SAR study on CHAP derivatives it has been shown that the presence of two hydrophobic amino acids, such as the bisphenylalanine moiety, is fundamental for the interaction with two lipophilic binding sites of HDACs.[188]

With the aim of probing binding interactions on the outer rim of HDAC enzymes, it was decided to focus on the bis-phenylalanine region markedly simplifying 1, generating a diamide scaffold by an Ugi reaction (Scheme 2.3).[189] This led to a peptidomimetic structure displaying a tertiary amide, making it possible to investigate the role of an additional substituent (R_3). Furthermore, it was also investigated different side chains (R_1, R_2), analyzing their influence on HDAC activity.

With this strategy in mind, different building blocks (isocyanides, **37**-**39**, aldehydes, **40**-**43**, amines, **44**-**46**, and carboxylic acids containing the alkyl chain and a methyl ester, **47** and **48** were chosen (Scheme 2.4).

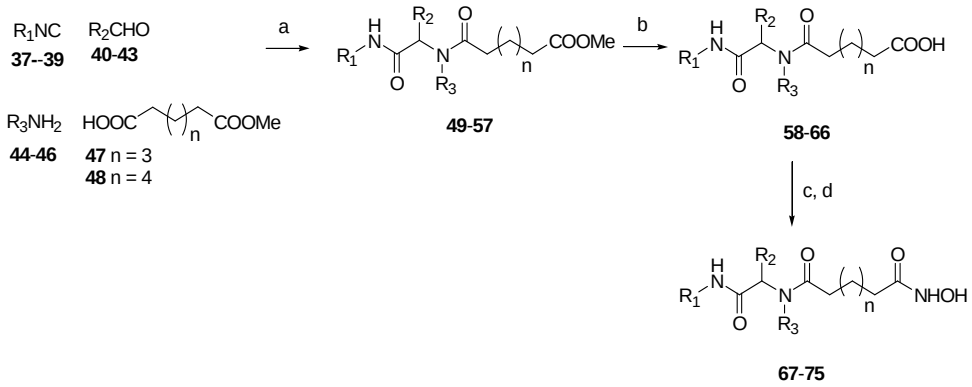


Scheme 2. 3 From CHAP 1 to the Ugi Products.



Scheme 2. 4 Building blocks.

The key reaction in the synthesis of the proposed HDACs inhibitors is the Ugi reaction, which leads to the R-aminoacylamide displaying an ester function. The methyl ester is then transformed into the corresponding hydroxamate passing through the carboxylic acid intermediate (Scheme 2.5).

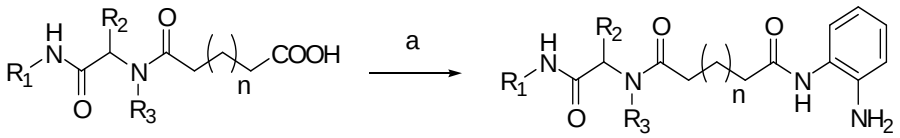


Scheme 2. 5 Reagents and conditions: (a) MeOH, rt, 17-44%; (b) LiOH, THF, H_2O, rt, 75-99%; (c) *o*-(*tert*-butyldimethylsilyl)hydroxylamine, TEA, EDCI, CH_2Cl_2, rt; (d) TBAF 1M in THF, THF, rt, 32-65% (two steps).

The Ugi reaction was performed in classical conditions (MeOH, 2 M, 48 h, r.t.), leading to the R-aminoacylamides in moderate yields (17-44%). Several attempts to optimize the transformation (**37**, **44**, **40**, **47**) were made; in detail, temperature (reflux), solvent (trifluoroethanol), reaction times (up to 7 days), and molar ratios (excess of **44** and **40**) were varied, alongside preforming the imine intermediate. Yet, none of these strategies proved better, in respect to the presence of starting materials, byproduct, and yields of the desired product. The hydrolysis of the methyl esters (**49**-**57**) to the corresponding carboxylic acids (**58**-**66**) was performed with LiOH in THF/water. The hydroxamic acids (**65**-**75**) were prepared by subsequent reaction with TBDMS-protected hydroxylamine, EDCI, and TEA in CH_2Cl_2, and final deprotection with TBAF in THF. To probe the structure-activity relationship of the metalchelating

group, analogues of **67** and **74** were also synthesized, replacing the hydroxamates with benzamides.

Benzamides were synthesized (**76** and **77**), coupling the corresponding carboxylic acids to o-phenylenediamine using standard peptide chemistry (DCC in THF; Scheme 2.6).



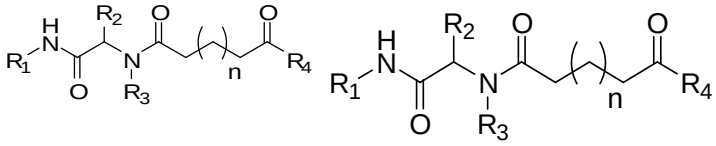
76,**77**

Scheme 2. 6 Reagents and conditions: (a) *o*-phenylene diamine, DCC, THF, rt, 23-24%.

To test the activity of the synthesized compounds, it was opted for a screening that evaluated cytotoxicity in a cell line previously reported to be sensitive to HDAC inhibitors.[190] For each compound viability was evaluated with the MTT assay (Table 2.6), which measures mitochondrial activity. Most compounds possessing the hydroxamates moiety displayed a cytotoxic activity comparable to that of SAHA (**17**), chosen as a reference compound. The only exception was represented by **38**, which was unable to induce a significant level of cell death at a concentration of 10 μM. To confirm the mechanism of action, a cellular HDAC activity assay was then performed in the same cell line concentrating on SAHA, **67** and **74**.

All three compounds inhibited HDAC activity at a concentration of 10 μM, validating the mechanism of action. The rank order of potency of the three compounds was SAHA (IC_{50} 0.5 μM ± 0.1 μM) > **67** (1.2 μM ± 0.1 μM) > **39** (4.4 μM ± 0.5 μM). This rank order of potency mirrored the rank order of potency observed in the cytotoxicity assay.

Table 2. 6 Synthesized hydroxamates and benzamides and their cytotoxic activity.



no	R_1	R_2	R_3	R_4	Viability (% of control)
67	cyclohexyl	Bn	Bn	hydroxamic acid	27.4±0.6
68	phenethyl	Bn	phenethyl	hydroxamic acid	30.2±0.3
69	phenethyl	Bn	Bn	hydroxamic acid	30.8±1.8
70	cyclohexyl	hexyl	phenethyl	hydroxamic acid	35.0±2.2
71	Bn	Bn	phenethyl	hydroxamic acid	34.9±1.4
72	phenethyl	Bn	Bn	hydroxamic acid	53.1±1.0
73	phenethyl	Bn	methyl	hydroxamic acid	94.4±2.3
74	phenethyl	1,1'-diphenyl-4-methylen	Bn	hydroxamic acid	22.2±3.1
75	phenethyl	heptyl	Bn	hydroxamic acid	53.0±4.1
76	cyclohexyl	Bn	Bn	Benzamide	91.6±2.8
77	phenethyl	1,1'-diphenyl-4-methylen	Bn	Benzamide	74.0±4.2
SAHA					33.8±0.4

Viability is evaluated at a concentration of 10.0 μM and data are expressed as mean ± standard error.

To rationalize and to identify the structural features of the active molecules, we performed molecular docking studies on **67** and **74**, both active and structurally dissimilar, and on **73**, which does not display cytotoxic activity, with the HDLP binding pocket (PDB code 1C3R, Figure 2.22),[122] using AutoDock 3.0.5 software[55] as reported in the previous paragraph 2.2-2.4.

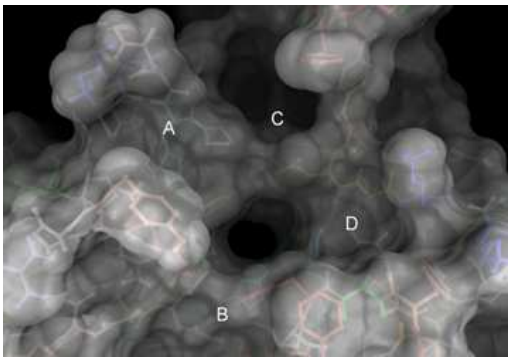


Figure 2. 22 3D model of the HDLP. The protein is represented by molecular surface and sticks and balls

For this docking calculations, with the aim to obtain a better accordance between theoretical K_{Dcalc} and biological assays results, in addition to the HDLP active site refined at QM level, the partial charges for the three ligands (**67**, **73**, and **74**) have been calculated at DFT B3LYP level and the Chelpg[148] method for population analysis (see the following section for details). The efficacy of such QM integration in the docking procedure has provided very promising results in previous studies on the HDAC inhibitors.[147,161b,f] According to the enzyme mapping, besides the binding channel of 11 Å, there are four hydrophobic cavities (A-D)[161f] acting as molecular recognition domain (Figure 2.22). The compounds were obtained in a racemic form; yet, this is one of the limits of the Ugi reaction. Docking studies were performed

on both possible configurations (*R* and *S*) for compounds **67**, **73**, and **74**, to evaluate if one or both enantiomers could be actually interacting with the biological target. In the three-dimensional model (Figure 2.23a) of the *S*-**67**, it is important to note that the hydroxamic moiety (*metal binder*) is of primary importance for its interactions network: it coordinates the zinc ion in a tridentate fashion and establishes hydrogen bonds with $H^{\varepsilon 2}$ of His132 and OH of Tyr297 (not shown). The heptanediamide chain (*linker*) makes stabilizing hydrophobic contacts with the zinc-containing tubular pocket. The *cap group*, formed by three rings, establishes noncovalent interactions with the hydrophobic cavities C and D on the HDLP surface (Figure 2.23a). In particular, the cyclohexylamino ring is correctly accommodated in the D hydrophobic pockets (Leu265, Ser266, Lys267, His170, Ala197, Phe198, and Phe200). The N^1-[1-(phenylmethyl)ethyl] is accommodated in a shallow groove, establishing van der Waals interactions with the macromolecular counterpart, formed by Ala197 and Phe200 residues, while the *N*1-(phenylmethyl) ring accommodates in the C hydrophobic pockets (Pro22, Phe141, His21) of the enzyme.

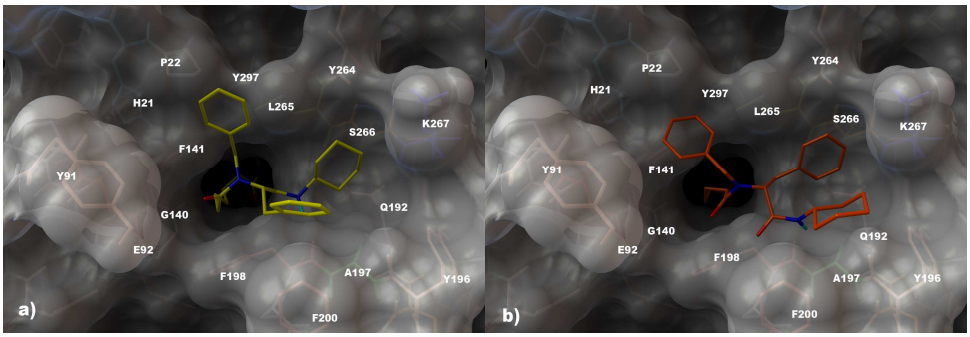


Figure 2. 23 (A) 3D model of the interaction between **67*S*** and the HDLP binding site. The protein is represented by molecular surface and sticks and balls. **67*S*** is depicted by sticks (by atom type: C yellow, polar H sky-blue, N dark blue, O red). (B) 3D model of the interaction

between **67*R*** and the HDLP binding site. The protein is represented by molecular surface and sticks and balls. **67*R*** is depicted by sticks (by atom type: C orange, polar H sky-blue, N dark-blue, O red).

Figure 2.23b depicts the 3D model of the interaction between *R*-**67** and the target; the *metal binder*, *linker chain*, and the *cap group* of the two docked enantiomers (*R*- and *S*-**67**) fill equivalent spaces, but the two rings (cyclohexylamino and N^1-[1-(phenylmethyl)ethyl]) of the *cap group* are in inverted positions on the HDLP surface with respect to *S*-**67** (Figure 2.23b). In particular, the cyclohexylamino ring interacts with the Ala197, and Phe200 residues, while the N^1-[1-(phenylmethyl)ethyl] is accommodated in D pocket (Leu265, Ser266, Lys267, His170, Ala197, Phe198, and Phe200); in the mean time, the N^1-(phenylmethyl) remains interacting with the aminoacid of the C hydrophobic pocket (Pro22, Phe141, His21). The different arrangement of the *R*-**67** and the suboptimal hydrophobic interactions are responsible for a predicted decrease in the binding affinity to the receptor of about 2.5-fold (K_{Dcalc} of *S*-**67** 4.45×10^{-8} *vs* K_{Dcalc} of the enantiomer 1.04×10^{-7}). The same approach was used for **74**. Bearing in mind the 3D model of the interactions between **67** and the HDLP binding site described above, the octanediamide chain of *S*-**74** fits into the 11 Å channel (Figure 2.24a), the NHOH group binds to the Zn^{2+} ion, at the bottom of the channel, in a tridentate fashion and establishes hydrogen bonds with $H^{\varepsilon 2}$ of H132, $O^{\delta 2}$ of Asp258, and OH of Tyr297 (not shown).

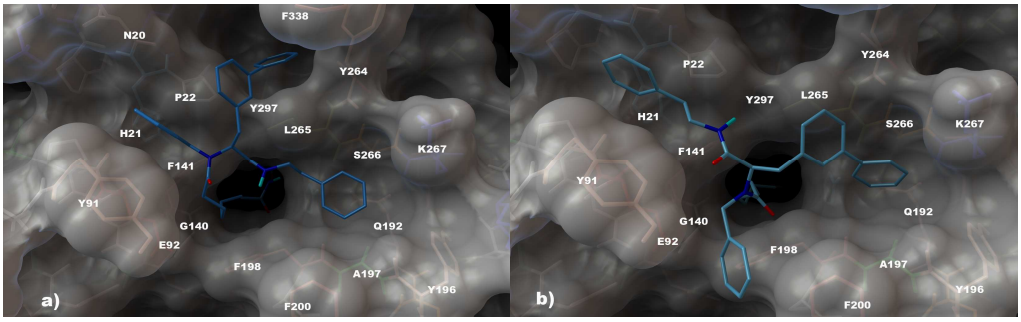


Figure 2. 24 (A) 3D model of the interaction between **74*S*** and the HDLP binding site. The protein is represented by molecular surface and sticks and balls. **74*S*** is depicted by sticks (by atom type: C blue, polar H sky-blue, N dark blue, O red). (A) 3D model of the interaction between **74*R*** and the HDLP binding site. The protein is represented by molecular surface and sticks and balls. **74*R*** is depicted by sticks (by atom type: C light-blue, polar H sky-blue, N dark blue, O red).

The *cap group*, formed by four aromatic rings, establishes hydrophobic interactions with the cavities (A, C, and D). The N^1-[1-([1,1'-biphenyl]-methyl)]ethyl] group accommodates in pocket C (Pro22, Phe141, Tyr264, Tyr297). The N^1-[(2-phenylethyl)amino]ethyl] ring accommodates in pocket D (Leu265, Ser266, Lys267, His170, Ala197, Phe198, and Phe200), while the N^1-(phenylmethyl) occupies pocket A (Pro22, Tyr91, Glu92, and Phe141). Figure 2.24b depicts the putative three-dimensional model of *R*-**74**-HDLP complex. Comparing this model with the S enantiomer binding mode, the only difference is the R_1-R_3 arrangement on the enzyme surface; in particular, the pockets A and D are occupied by the N^1-(phenylmethyl) and the N^1-[1-([1,1' -biphenyl]-methyl)]ethyl] moieties, respectively, while the N^1-[(2-phenylethyl)amino]ethyl] ring interacts with the Phe198, Phe200, and Tyr91 amino acids. The minor efficiency of such interactions are responsible of a small decrease in the binding affinity to the target (of about 5-fold: K_{Dcalc} of S-**74** 7.69 x 10^{-9} *vs* K_{Dcalc} of the enantiomer 3.86 x 10^{-8}).

In both the enantiomers R and S of **73**, the lack of a hydrophobic and bulky group on R_3 is responsible for a predicted decreased histone deacetylase inhibitory activity (K_{Dcalc} of *S*-**73** 2.44x 10^{-7} *vs* K_{Dcalc} of the enantiomer 3.87 x10^{-7}), because their expected noncovalent interactions are lost. In the first model (Figure 2.25a), the N^{δ}-[2-[(2-phenylethyl)amino], and N^{δ}-1-(phenylmethyl) ethyl groups are placed in D (Leu265, Ser266, Lys267, His170, Ala197, Phe198, and Phe200) hydrophobic cavity, but the hydroxamic moiety (metal binder) coordinates the zinc ion in a monodentate fashion. In the three-dimensional model of *R*-**73**, the R_1 and R_2 substituents are placed in the same hydrophobic cavity found for the S-enantiomer but with an inverted spatial arrangement, while the NHOH group binds to the Zn ion in a bidentate fashion.

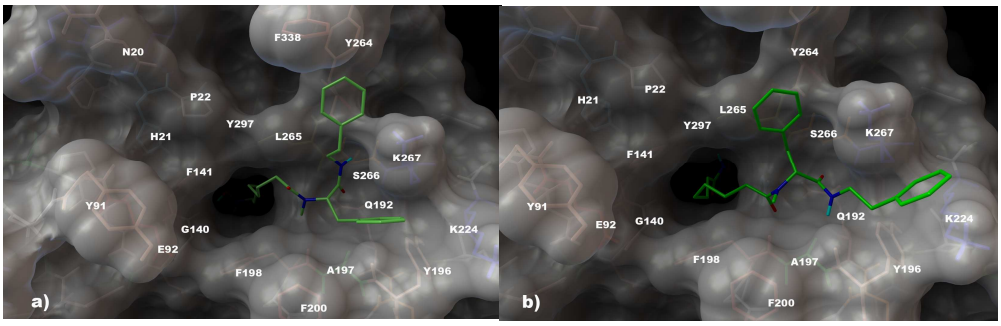


Figure 2. 25 (A) 3D model of the interaction between **73S** and the HDLP binding site. The protein is represented by molecular surface and sticks and balls. **73S** is depicted by sticks (by atom type: C light-green, polar H sky-blue, N dark blue, O red). (A) 3D model of the interaction between **73R** and the HDLP binding site. The protein is represented by molecular surface and sticks and balls. **73R** is depicted by sticks (by atom type: C green, polar H sky-blue, N dark blue, O red).

Keeping in mind the biological results and the descriptions of the 3D models of **67**, **73**, and **74** with HDLP discussed above, here a rationalization of

docking results related to compounds containing $NHR_1COCHR_2NR_3$ as cap group was offered.

In particular, the influence of the hydrophobic and bulky group on R_1, R_2, and R_3, in turn related especially to hydrophobic interactions, and to aromatic stacking, seem to be the driving forces of the target-ligand complexes. Such evidence is confirmed by calculated and experimental **73** inactivity. In light of the above results and thanks to the good qualitative accordance between the results of biological essays and the prediction of the molecular docking calculations, there is a complete rationalization of the putative binding mode for **67**, **73**, and **74** enantiomers pairs. In particular, a new scaffold was proven to be an efficient cap group model in the rational design of new linear HDAC inhibitors, and the critical features necessary for the optimal contact modes with HDLP binding pocket were determined.

2.5.1. Computational Details

Molecular docking calculations were performed on a two dual-core Intel Xeon 3.4 GHz, using Autodock 3.0.5 software[55] using the already described optimized HDLP model.

The charges of the ligands were optimized at DFT level using the B3LYP functional and the 6-31G+(d) basis set, as implemented in Gaussian 03 Package software.[160] For all the docking studies, a grid box size of 66 x 64 x 48 with spacing of 0.375 A between the grid points and centered at 49.75 (x), 5.0 (y), and 101.491 (z) covering the four hydrophobic pockets (A-D) on the HDLP surface was used. For all the docked structures, all bonds were treated as active torsional bonds. To achieve a representative conformational space during the docking calculations, 10 calculations consisting of 256 runs were performed, obtaining 2560 structures (256 × 10). The Lamarckian genetic

algorithm was used for dockings. An initial population of 450 randomly placed individuals, a maximum number of 10.0×10^6 energy evaluations, and a maximum number of 8.0×10^6 generations were taken into account. A mutation rate of 0.02 and a crossover rate of 0.8 were used. Results differing by less than 2.5 Å in positional root-mean-square deviation (rmsd) were clustered together and represented by the free energy of binding. All the 3D models were depicted using the Python software[165] molecular surfaces are rendered using Maximal Speed Molecular Surface (MSM).[166]

2.6. Structural Basis for the design and synthesis of selective HDAC inhibitors

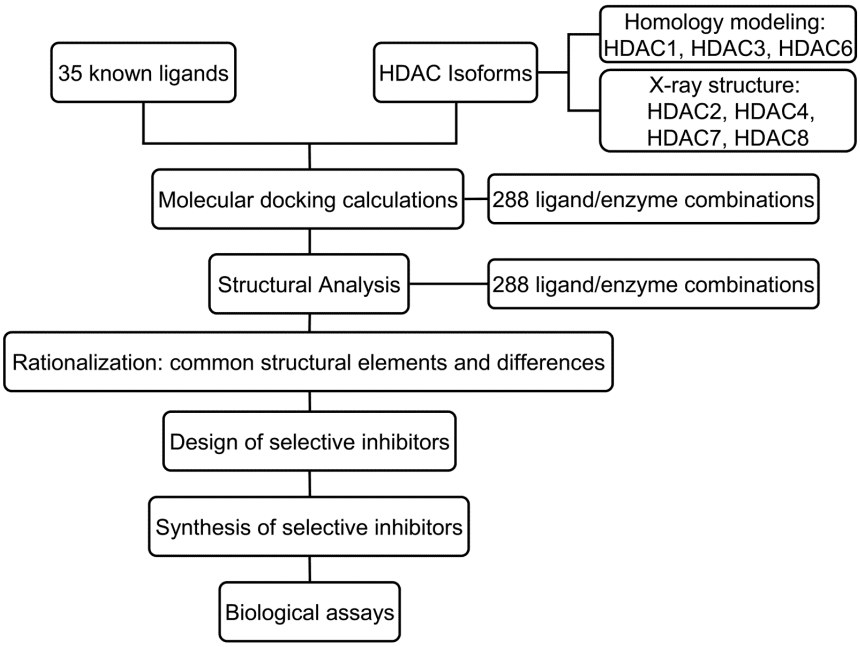
In this paragraph, it was traced out the structural elements responsible of selective binding for the therapeutically relevant different HDAC isoforms. The structural analysis has been carried out by molecular modeling, docking in the binding pockets of HDAC1-4 and HDAC6-8, thirty five inhibitors presenting a well defined selectivity for the different isoforms. As proof of evidence, it was reported the design, synthesis and biological evaluation of three selective inhibitors.[191] The experimental data suggest that the obtained structural guidelines can be useful tools for the rational design of inhibitors against selected HDAC isoforms.

In literature, studies devoted to develop selective ligands for specific isoforms are reported, even though such studies only consider a restricted number of isoforms.[192] Moreover, most of the published works focused on selective ligands are only based on biological profiles, lacking the structural investigation aimed to disclose the ligand and protein elements potentially responsible of class and isoform selectivity.[193]

On this basis, here, it was traced out the structural elements responsible of selective binding in the whole landscape of the therapeutically relevant HDAC isoforms (Scheme 2.7).

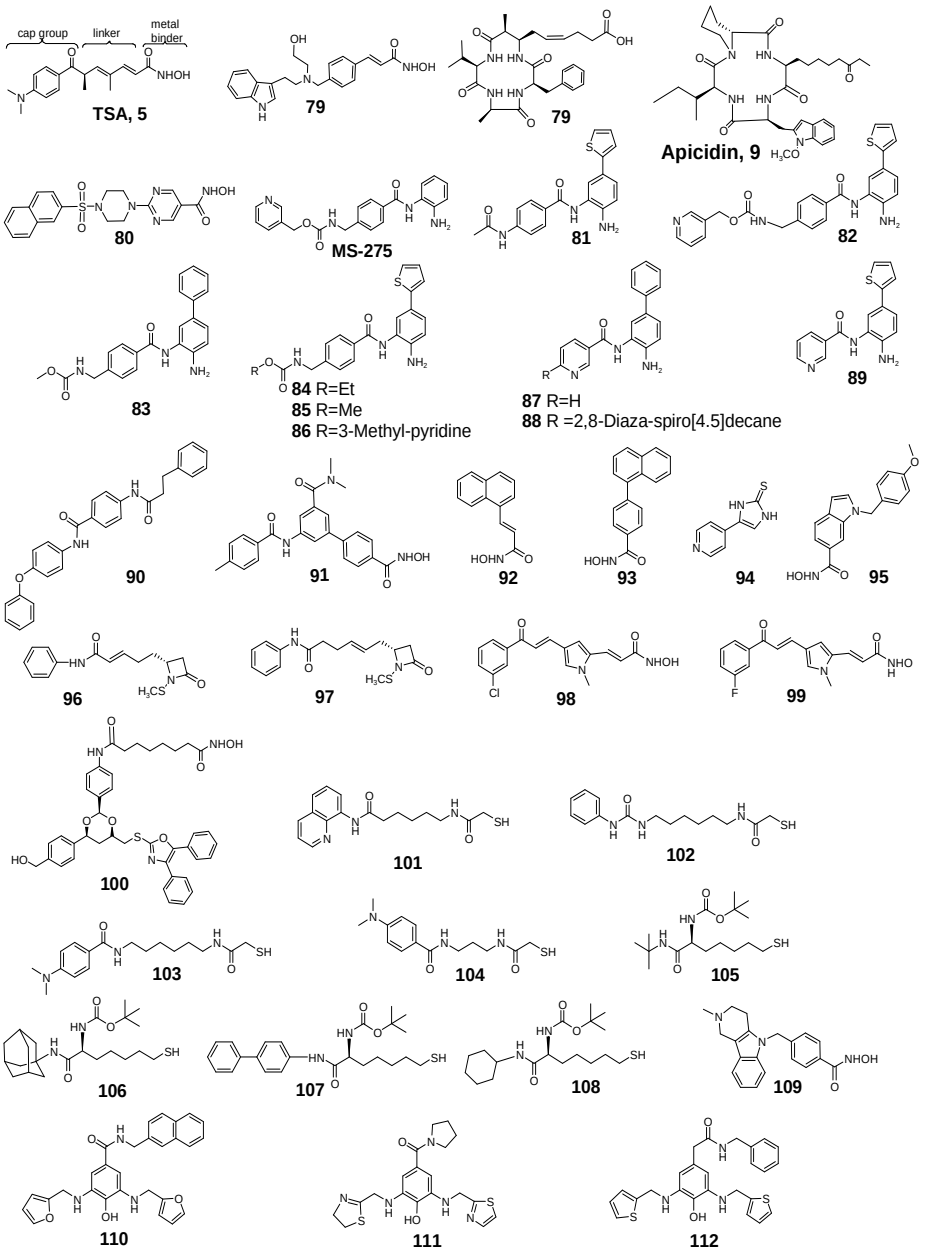
In particular, it was tried to rationalize a number of experimental observations and tried to systematically add new insights for a targeted design of selective inhibitors of the different HDAC isoforms, focusing our attention on HDAC1-4, and HDAC6-8. HDAC9-11, for which few information so far are available in literature on expression and function in tumor cells, and

HDAC5, missing a concrete ligand inhibitory profile,[129] have not been considered in this investigation.



Scheme 2. 7 Work-flow of the different stages involved in the structural characterization, design and synthesis of new selective HDAC inhibitors.

The structural analysis reported below was performed by molecular docking calculations, using as ligands pan and class selective HDAC inhibitors reported in the literature (**5**, **7**, **9**, and **78**-**109**, Scheme 2.8), presenting a well defined profile of HDACs inhibition. Based on the obtained structural guidelines, we designed (**110**-**112**, Scheme 2.8), synthesized and experimentally tested selective inhibitors for HDAC2.



Scheme 2. 8 Molecular structures of compounds **5**, **7**, **9**, and **78**-**112**. **5** and **78**, pan inhibitors. **9**, **7**, **79**-**80**, class I selective inhibitors. **81**-**90**, HDAC1,2 selective inhibitors. **91**-**97**, HDAC8 selective inhibitors. **98** and **99**, HDAC4 selective inhibitors. **100**-**109**, HDAC6 selective inhibitors. **110**-**112**, designed compounds (see Proof of concept).

2.6.1. Structural analysis

All the considered compounds in the present analysis were analyzed by molecular docking calculations on HDAC1-4 and HDAC6-8 (Scheme 2.8). Some enzyme calculation parameters related to the electrostatic and Van der Waals terms of binding energy were refined, following the successful strategy adopted for the structural studies on azumamide E and two stereochemical variants.[147] This strategy allowed to reach a good qualitative accordance between theoretical K_{Dcalc} and biological assays results, and it was also validated by its useful application to the study of other HDAC ligands.[161,169] In particular, for the electrostatic contribution, the partial charges of the zinc ion and the amino acids constituting the catalytic site of each isoforms were calculated at DFT/M05[194] theory level by using the 6-31+g(d) as basis set and the ChelpG method[148] for the population analysis and were used in the subsequent docking calculations. Similarly, for the van der Waals term the well depth and zinc radius proposed by Stote and Karplus were applied.[164]

2.6.1.1. Common structural features of all isoforms

As well defined by previous studies,[168,183] the general structure of HDAC inhibitors can be dissected (**5**, Scheme 2.5) in: a cap group involved in the molecular recognition process with surface amino acids; a linker, usually hydrophobic; a zinc-chelating group. Each structural moiety contributes to the binding event and biological activity of the small molecules. In particular, a fundamental structural element to inhibit these biological targets is the metal binder. Many chelating agents, such as hydroxamic acids, carboxylate, α-hydroxy-ketone were introduced, and theoretical and experimental evidence showed their ability to coordinate the prosthetic group of the investigated

enzymes. Class I and II proteins present a considerable sequence similarity in the catalytic site. By comparing homology and experimental models, it possible to observe the presence of two parallel phenylalanine units delineating the channel which accommodates the acetylated lysine of the histone (Table 2.7, and Figure 2.6).

Table 2. 7 Residue number of two phenylalanines constituting the hydrophobic channel, which accommodates the substrate.

	Phe	
HDAC1	150	205
HDAC2	155	210
HDAC3	144	200
HDAC4	168	227
HDAC6	140	200
HDAC7	679	738
HDAC8	152	208

As revealed by docking calculations on all considered HDACs, the switching from a linear carbon chain (**5**, **9**, **79**, **22**-**108**) to an aromatic linker (**7**, **78**, **80**-**95**, **109**) causes an increased affinity with the targets, thanks to favourable π-π interactions with the above mentioned two phenylalanine side chains.

2.6.1.2 General features and differences of Class I and II

In the reported analysis, it was observed that HDAC1-3 and HDAC8 present a CO of a glycine and side chain of a tyrosine pointing inside the 11 Å channel: Gly149 and Tyr303 for HDAC1, Gly154 and Tyr308 for HDAC2, Gly143 and Tyr298 for HDAC3, and Gly151 and Tyr306 for HDAC8. The class I isoform selective compounds **7, 81-89** present an amide bond between the linker and the metal binder (Figure 2.26). The NH of the amide functionality can establish a hydrogen bond with the carbonyl of the glycine. Moreover, the hydrophobic channel of class II proteins does not display acceptors and donors of hydrogen bonds. In this perspective, the linker could be modified inserting hydrogen bond acceptors and donors in order to give interactions with these two residues constituting the hydrophobic channel, sustaining selectivity for class I HDACs.

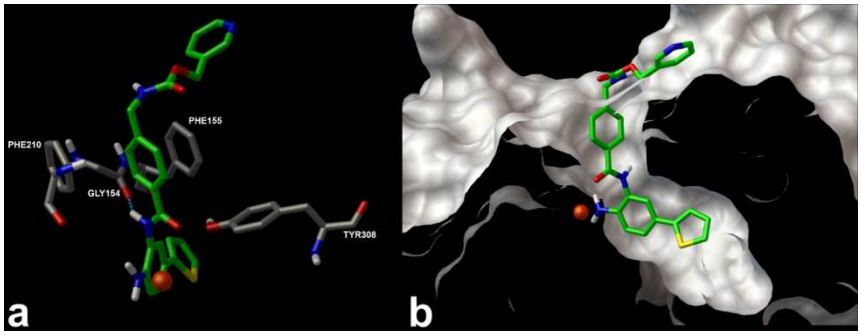


Figure 2. 26 Three-dimensional model of the interactions between **82** and HDAC2. The protein (a) and **82** are represented by tube and their atoms are coloured by atom type: C, gray; polar H, white; N, dark blue; O, red). For **82**, carbon and bonds are depicted in green. The zinc ion is represented in orange cpk. The figure highlights the hydrogen bond (dashed cyan line) between the NH of the amide functionality and the carbonyl of the Gly154; b) The protein is represented by molecular surface. The figure shows the coordination of zinc ion (represented in red cpk) and the accommodation of metal binder appendage in the internal cavity.

The analysis of crystal structures of bacterial homologues (HDLP) of class I[122] along with human HDAC2 and HDAC8 reveals the presence of a 14

Å internal cavity at the bottom of the 11 Å hydrophobic channel, close to the zinc active site.[161f] The high sequence similarity of HDAC1 and HDAC3 with HDLP, HDAC2 and HDAC8 confirm for all class I HDACs the presence of the 14 Å internal cavity is expected. The homology models obtained in this study of HDAC1 and HDAC3 in fact showed the presence of the internal channel observed in HDLP, HDAC2 and HDAC8, also in agreement with previously reported homology modelling studies.[195] With one exception (HDAC8), the residues forming the 14 Å internal cavities are identical across the different proteins of class I. Recently, 2-amino benzamides were proposed as metal binders, and compounds presenting this functionality showed selectivity for class I enzymes.[196] In particular, in agreement with reported experimental data, the analysis of docking results on the compounds presenting a benzamide as metal binder (**7**,[197] Scheme 2.8) revealed a preference for class I enzymes, in particular for HDAC1 and HDAC2. The coordination of the enzyme prosthetic group by the NH_2 of benzamide requires a side accommodation of phenyl ring, unlike the common metal binders, such as hydroxamic acid, carboxylate group, α-hydroxy-ketone. The consequence is the requirement of a side room at the bottom of the 11 Å channel, offered by the 14 Å internal cavity.

In literature it is reported that natural cyclopeptides[198] are selective inhibitors of class I proteins, and the most important of these ligands were considered in our studies. For example, the azumamide E[147] and apicidine[199] (**79** and **9**, Scheme 2.8) show selectivity for class I, with significant affinity for HDAC8 and superior affinity for HDAC1-3. From the reported theoretical analysis of cyclopeptides and the other ligands, it was found that the macrocycle can be accommodated on a shallow pocket located on the protein

surface at the entrance of 11 Å hydrophobic cavity, establishing Van der Waals interactions and hydrogen bonds (Figure 2.27).

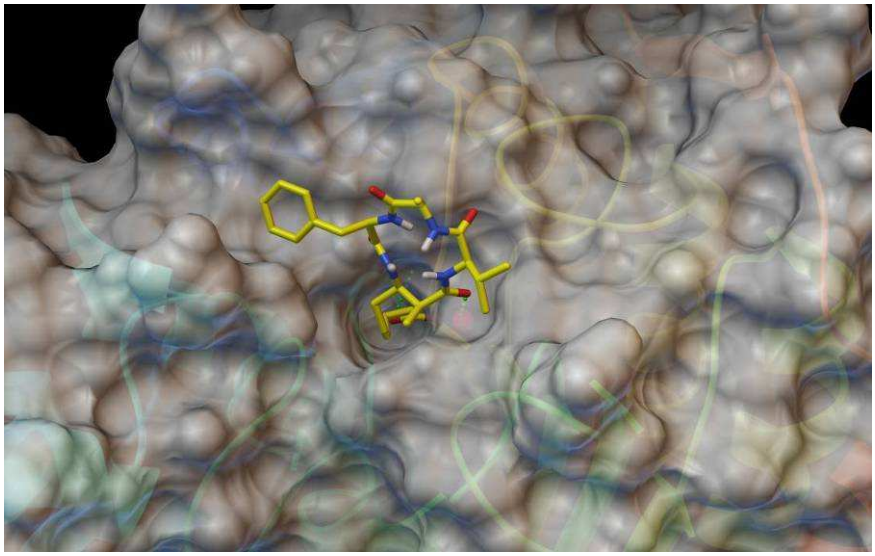


Figure 2. 27 Three-dimensional model of the interactions between **79** and HDAC1. The protein is represented by molecular surface and ribbons. **79** is depicted by tube. C atoms and bonds are shown in yellow and the remaining atoms are coloured by atom type: polar H, white; N, dark blue; O, red). The figure highlights that the phenyl group is located in a hydrophobic pocket and the tetrapeptide core interacts with a shallow cavity on the receptor surface. The green dashed line indicates a hydrogen bond.

These interactions contribute to the complex stability, favouring the binding for the HDACs of class I over the isoforms of class II. Such observation is suggested from the docking analysis of the binding mode of azumamide E and apicidine with HDACs of class II, showing that the macrocycle does not interact with macromolecular counterparts on the surface of HDAC4, HDAC6 and HDAC7, thus not contributing to the affinity for the protein. The pan inhibitors TSA (**5**) and NVP-LAQ824 (**78**),[200] for example, did not present this structural bulky cap group, highlighting the role of the peptidic macrocycle in dictating the selective class I binding. Compared to

other analyzed class I selective ligands, the cyclopeptides presented a wider cap group, which established more extended contacts with proteins surface. As reported in the previous section, the proteins of class II present two phenylalanines to form the channel leading to the active site (Table 2.7, and Figure 2.6), like the isoforms of class I. By this analysis of the homology and experimental models of class II isoforms, we observed that the distance of the aromatic side chains of these two residues is about 8.5 Å. This distance is larger than that one found in the class I isoforms (≈ 7 Å), thus the linker moiety could establish tighter interactions with hydrophobic channel of HDAC1-3 and HDAC8 with respect to HDAC4, HDAC6 and HDAC7.

2.6.1.3 HDAC1

The differences between HDAC1 and HDAC2 are very small due to an 85% sequence identity and 93% sequence similarity, which confer very similar shapes for all the protein regions. The comparison of the three-dimensional models of HDAC1 and HDAC2 reveals some small differences, which could be exploited to discriminate the recognition of these two isoforms. By this investigation it was observed a tighter access to the catalytic site of HDAC1 compared to HDAC2 (Figure 2.28), due to the different spatial arrangement of residues leading to the zinc ion and bordering the 11 Å channel. In particular, this diverse orientation reflects the replacement of Met233, Pro361 and Met364 in HDAC2 by Leu228, Asn356 and Leu359 in HDAC1.

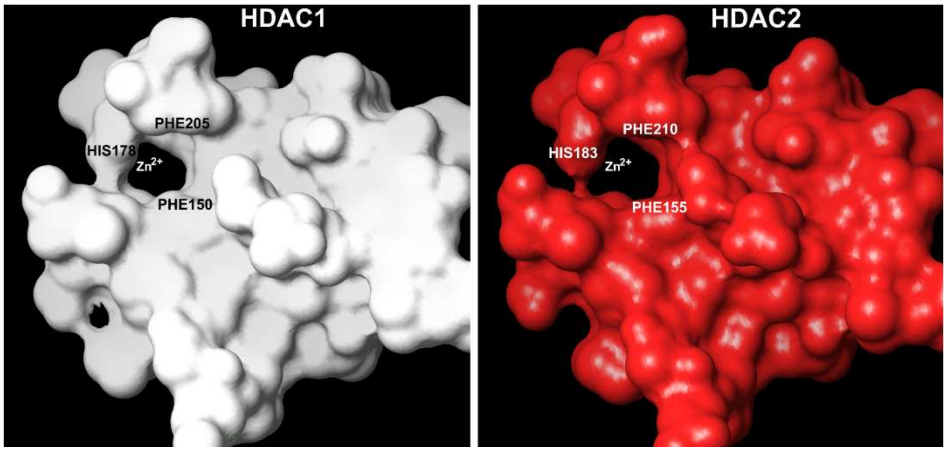


Figure 2. 28 View of the top of the ≈11 Å channel. The HDAC1 and HDAC2 are represented by molecular surface coloured in white and red, respectively. The figure highlights the difference in shape and dimensions of the tube like channel, leading to the zinc ion.

As reported above, all proteins of class I present an internal hydrophobic pocket at the bottom of the channel accommodating the substrate. It was observed that the amino acids constituting this internal cavity are identical or conservatively substituted, but differences in the surrounding amino acids of these internal cavities can be pointed out. In particular, bulkier residues in HDAC3 and HDAC8 (see below) prevent the accommodation of the chelating agent with an appendage. Indeed, in the case of HDAC1 and HDAC2 the docking results on **7** highlighted that the metal binder is well harboured and the zinc is coordinated by the CO and NH_2 functionalities. This docking studies show (Figure 2.29) that the phenyl ring interacts with the surrounding hydrophobic residues and along with the bidentate coordination increases the affinity for HDAC1 and HDAC2. These findings are in agreement with published biological assays,[197] which show the selectivity of MS275 (**7**) for HDAC1 with this preference: HDAC1 > HDAC3 and HDAC1 >> HDAC8, confirming the presence of the benzamide metal binder for directing isoform selectivity.[168,183,201]

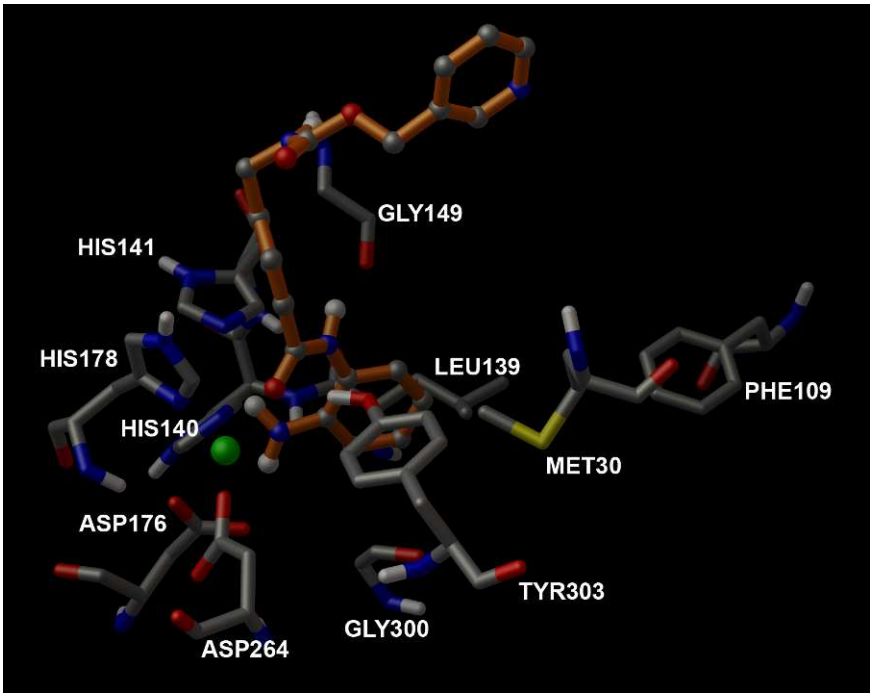


Figure 2. 29 Three-dimensional model of the interactions between **7** and HDAC1. The amino acids are represented by tube and coloured by atom type (C, grey; polar H, white; N, dark blue; O, red). **7** is depicted by stick (orange) and balls coloured as for the protein. The figure highlights that the phenyl group is located in the internal cavity, establishing hydrophobic interactions.

The insertion of a substituent on the benzamide leads to a further discrimination among HDACs of class I. Indeed, HDAC1 and HDAC2 were able to harbour a substituent on the benzamide (**81**-**89**, Scheme 2.8), such as thiophene or phenyl.[202] The reported docked pose (Figure 2.30) suggests the same considerations for compound **90**[203] where the phenethyl replaces the benzamide interacting with the internal cavity, and strengthening the coordination of the zinc to the carbonyl group.

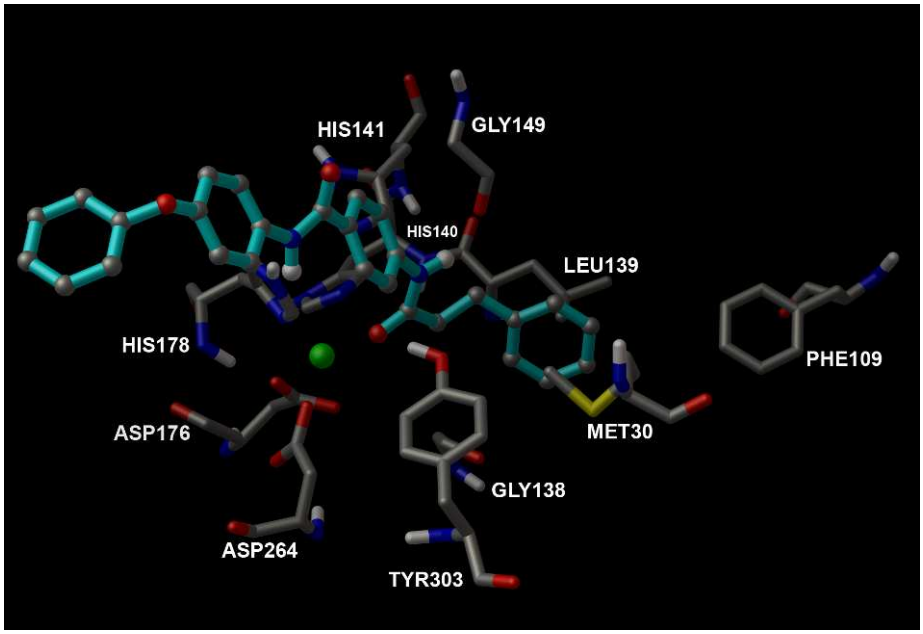


Figure 2. 30 Three-dimensional model of the interactions between **90** and HDAC1. The amino acids are represented by tube and coloured by atom type (C, grey; polar H, white; N, dark blue; O, red). **90** is depicted by stick (cyan) and balls coloured as for the protein atoms. The figure highlights that the phenyl group is located in the internal cavity, establishing hydrophobic interactions.

2.6.1.4 HDAC2

As outlined in the previous section, many structural elements of compounds having high affinity for HDAC1 are similar for the isoform 2. By comparing the models of HDAC1 and HDAC2, structural differences, even though small, are detectable. As described above, residues bordering the ≈ 11 Å channel, present a different spatial arrangement of Tyr204 in HDAC1 with respect to Tyr209 in HDAC2, in turn depending on the presence of Leu228, Asn356 and Leu359 in HDAC1 and Met233, Pro361 and Met364 in HDAC2. On the basis of this different arrangement, a deeper cavity is present for HDAC2 compared to the shallower cavity of isoform 1, formed by amino acids His183, Tyr209, Phe210 and Leu276. Together with this first discrimination, and even though

the amino acids constituting the 14 Å internal cavity are identical for HDAC1 and 2, differences could be found in the neighbouring residues. In details, it was observed that the Val19 (HDAC1) is substituted by the Ile24 in HDAC2, influencing the arrangement of surrounding residues. In particular, it was observed a different disposition of Met35, Phe114 and Leu144 giving rise to a larger room compared to HDAC1 (Figure 2.31).

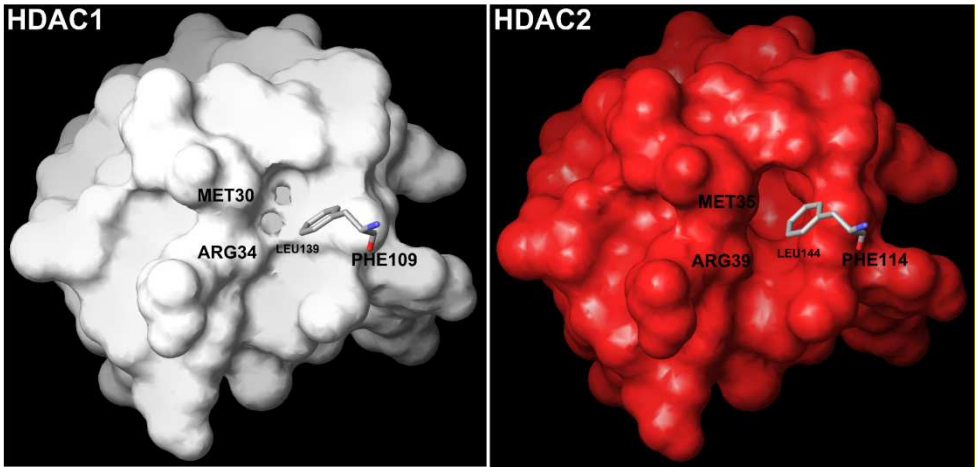


Figure 2. 31 View of the bottom of the side channel. The HDAC1 and HDAC2 are represented by molecular surface coloured in red and white, respectively. The figure highlights the difference in shape and dimensions of the internal channel, near to the catalytic site of the enzymes.

This can justify the observed slightly preference (about 10 fold)[183,203] of benzamides for HDAC1 *vs.* HDAC2, due to closer contacts with the internal cavity of isoform 1. On this basis, it was designed and tested new potential selective ligands for HDAC2 taking into account the combination of a metal binder able to interact with the 14 Å internal cavity and of an adapted capping moiety (see below).

2.6.1.5 HDAC3

As shown in the previous two sections, the relevant structural elements to discriminate isoforms of class I are appendages of metal binder, able to interact with the internal cavity at the bottom of the 11 Å channel. By the analysis of the amino acids surrounding the internal cavity close to the catalytic site, it was observed that in HDAC3 Tyr107 replaces Ser113 in HDAC1 and Ser114 in HDAC2. In the isoform 3 the presence of the bulkier side chain of Tyr107, forces the Leu133 to point towards the center of the internal cavity. This shift of Leu133 causes a steric clash preventing the accommodation of bulky metal binders. The reported docking results on **81-90** highlighted that the metal binder was not well accommodated in the internal cavity, as experimentally confirmed.[168,195,198,203] Thus, the design of an adapted chelating agent able to match the HDAC3 macromolecular counterparts and to give effective contacts, is necessary for gaining selectivity toward this isoform. The reported docking results allowed to appreciate differences on the protein surface near the catalytic site. In particular, it was observed the presence of Phe199 in HDAC3 and Tyr204 and Tyr209 in HDAC1 and HDAC2, respectively. In the isoforms 1 and 2 the OH group of tyrosine establishes a hydrogen bond with the CO of Leu271 in HDAC1 and Leu276 in HDAC2. This hydrogen bond is absent in HDAC3, giving rise to a larger and deeper hydrophobic cavity delimited by the amino acids Ile171, His172, Phe199, Phe200, Gly267 and Cys268 (Figure 2.32).

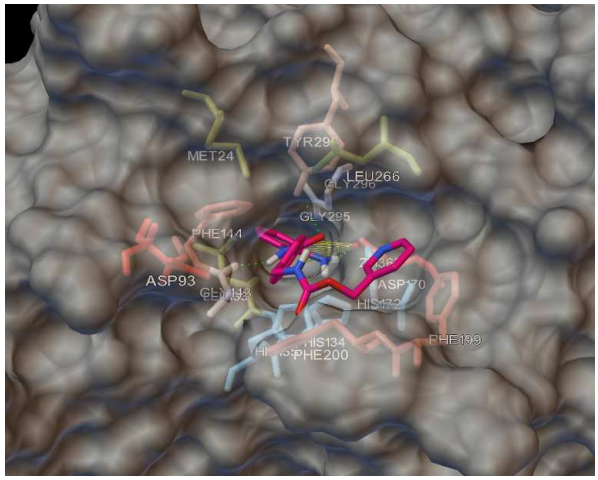


Figure 2. 32 Three dimensional model of the complex between **7** and HDAC3. The protein is represented by molecular surface and sticks. **7** is depicted by sticks. The C atoms and bonds are shown in pink and the remaining atoms are represented by atom type: polar H, white; N, dark blue; O, red.

This cavity can accommodate larger ligand moiety favouring the selectivity toward the isoform 3, in addition to the affinity. Moreover, it was found a pocket delimited by amino acids Phe88, Asp93, Gly143 and Phe200, as for HDAC1. In this regard, the insertion of an aromatic ring on the capping moiety in order to establish π-π interactions with Phe200, and anion-π contact with Asp93, a hydrogen bond donor or a positively charged group for further contacts with Asp93, may contribute to increase the affinity for HDAC3.

2.6.1.6 HDAC8

The resolved X-ray structures of HDAC8 complexed with **5**[204] and **91**[205] have suggested structural elements to selectivity bind this isoform. When bound to **91**, HDAC8 shows a shift from the normal position of Phe152, which is located along the hydrophobic 11 Å channel. This shift creates a sub pocket that can contain hydrophobic groups protruding from the linker moiety, extending Van der Waals contacts with the channel of HDAC8. These further

strong interactions contribute to the complex stability and favour a selective binding to HDAC8. Compounds **92**,[192b] **93**[192b] and **95**[206] were designed on the structural considerations obtained from **91** bound to HDAC8, and molecular docking calculations showed, as expected, high selectivity for HDAC8 thank to the interactions with the induced hydrophobic cavity. By comparing the amino acids constituting the internal cavity found in all isoforms of class I, it was observed that HDAC8 presents Trp141, in place of a leucine in HDAC1-3. The presence of the bulky side chain of Trp141 hinders the appropriate accommodation of chelating portions endowed with appendages, compared to HDAC1-3. Indeed, the docking results on **7** with HDAC8 showed that the Trp141 limited the accommodation of the benzamide, causing the NH_2 to coordinate the zinc ion in a monodentate manner (Figure 2.33). Moreover, for compounds **81**-**90** endowed of larger appendages, the prosthetic group of the enzyme is not coordinated.

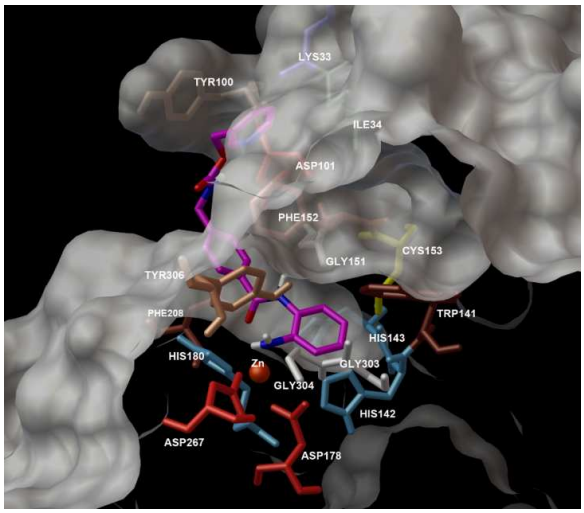


Figure 2. 33 Three dimensional model of the complex between **6** and HDAC8. The protein is represented by molecular surface and sticks. **6** is depicted by sticks. The C atoms and bonds are shown in pink and the remaining atoms are colored: polar H, white; N, dark blue; O, red).

On the other hand, well tailored chelating moieties can afford selectivity for isoform 8, as showed by compounds **96** and **97**.[192a] Indeed, the N-thiomethyl-azetidin-2-one contained in **96** and **97** interacts with Trp141, allowing a correct coordination of the zinc binding region of HDAC8 (Figure 2.34).

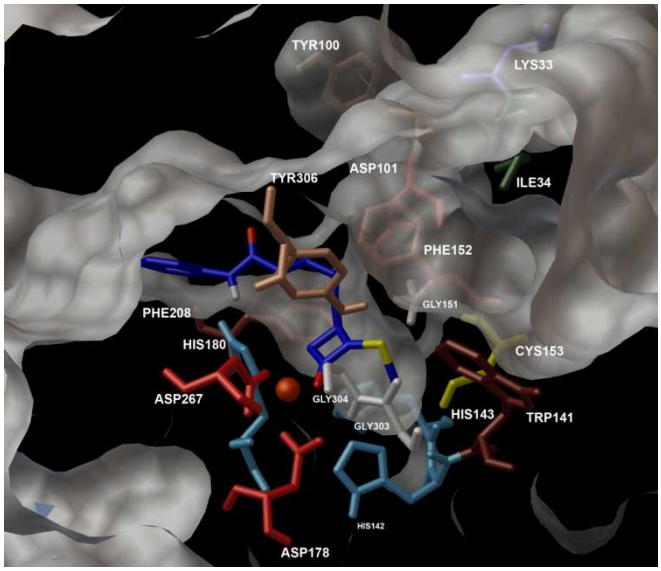


Figure 2. 34 Three-dimensional model of the interactions between **96** and HDAC8 (3FOR, pdb archive code). The amino acids are represented by tube, and coloured as: gly, white; asp, red; phe, brown; his, light blue; leu, green; cys, yellow; trp, dark brown; tyr, light brown. The zinc ion is represented in dark orange cpk. The **96** is depicted by blue tube and the atoms coloured as: C, blue; polar H, white; N, dark blue; O, red; S, yellow. The figure highlights the interactions of chelating moiety with the internal cavity.

2.6.1.7 HDAC4

Based on docked poses of all considered ligands (Scheme 2.8), we tried to trace out the structural elements responsible for the selectivity of HDAC4. In particular, aryl pyrrolyl hydroxamide (APHA, **98**, **99**, Scheme 2.8) compounds[207] are reported as selective HDAC4 inhibitors. These compounds show a lower general HDAC affinity compared to TSA, but present higher

selectivity for this enzyme. It was observed that substitution of chlorine with fluorine at phenyl C3 position improved the selectivity for HDAC4 from 78-fold to 176-fold over class I.[207] Moreover, non-halogenated or differently substituted APHA derivatives did not show any selectivity towards this isoform, highlighting the importance of the position of the halogen in the capping group. The described theoretical model showed that the halogen is involved in a hydrogen bond with the side chain of Tyr170 (Figure 2.35). Moreover, there is an interaction between NH group of Phe168 with the electron π system of the halogenated phenyl ring, along with a π-π interaction of C-halogen bond with CO of Phe168 that can contribute to the specific recognition for HDAC4. The reported theoretical investigation also suggested that the presence of an aromatic linker in **98** and **99** gives rise to π-π interactions with side chains of Phe168 and Phe227 (Figure 2.35).

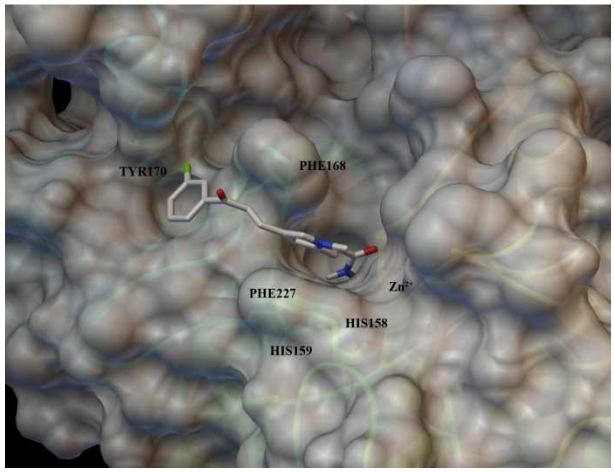


Figure 2. 35 Three-dimensional model of the interactions between **25** and HDAC4. The protein is represented by ribbon and the zinc ion is depicted in dark orange cpk. The side chains of His158, His159 and Tyr170 (light green), and **25** (white) are shown by tube. The atoms are coloured as: polar H, white; N, dark blue; O, red. The yellow dashed lines indicate hydrogen bonds.

It was also observed that the hydroxamic acid of **98** and **99** coordinated the zinc ion, but also established hydrogen bonds with the $N^{\varepsilon 2}$ of His158 and His159 by the OH and NH groups, respectively (Figure 2.35). In all other analyzed isoforms **98** and **99** did not show these two interactions, except one hydrogen bond with the catalytic site of HDAC3. As reported above,[207] a moderate biological activity has been shown for these HDAC4 selective inhibitors, thus structural modifications are required to increase the affinity of new candidate molecules. From this analysis, hydrophobic cavity delimited by residues His198, Phe226, Phe227, Leu299 may host a larger group, able to establish a hydrogen bond with the NH of His198, in replacement of the methyl group in the **98** and **99**. From the docked poses of **98** and **99** we observed the cap group near a small pocket formed by Pro165, Met166, Gly167, Tyr170, and Cys169.

Thus, along with a halogen, a hydrophobic group can be inserted to establish Van der Waals contacts with the described pocket on protein surface. The reported comparison of electrostatic potential maps of all considered isoforms on the surface around the channel leading to the zinc ion revealed that HDAC4 displayed a positive charged area, whereas the other proteins presented negative or neutral regions (Figure 2.36). Thus, the cap moiety can be elongated to favour electrostatic interactions with side chains of positively charged amino acids.

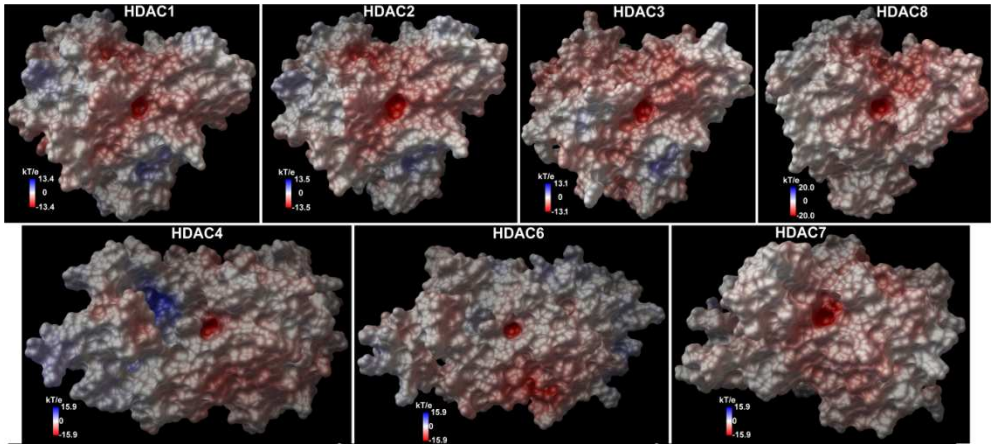


Figure 2. 36 Molecular surfaces of HDAC1-4 and 6-8, colored by electrostatic potential.

2.6.1.8 HDAC6

As already reported in previous studies on selective HDAC6 ligands,[192d,192e,193a] here it was confirmed through our analysis that the linker length is a crucial structural element for achieving selectivity. Upon zinc coordination, the linker length is responsible of correctly directing the extended interactions of the capping moiety with the macromolecular counterparts. HDAC6 presents a wider entrance of the binding pocket conducting to the prosthetic group, formed by several non-polar residues. The selective inhibitors (**100**-**108**) of HDAC6 follow the protein shape by their structural moieties and can assume an extended conformation interacting with the non-polar residues on protein surface (Figure 2.37).

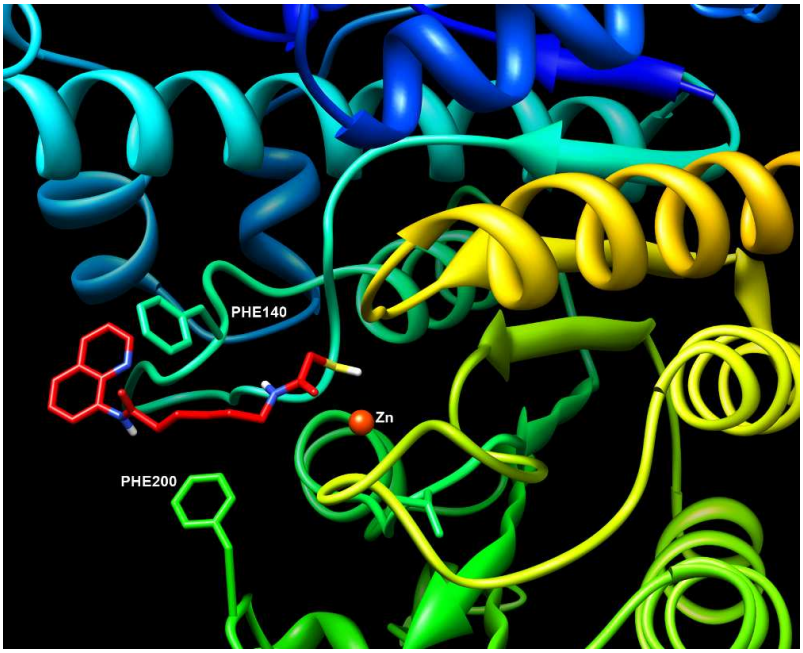


Figure 2. 37 Three dimensional model between **101** and HDAC6. The protein is represented by ribbon and the zinc ion is represented in dark orange cpk. The side chains of Phe140 and Phe200, and **101** are depicted in tube, respectively coloured in green and light green, and red. The atoms are coloured as: polar H, white; N, dark blue; O, red; S, yellow. The yellow dashed lines indicate hydrogen bonds.

In the case of other considered HDAC isoforms, due to the smaller entrance of the hydrophobic channel, the cap groups are not well accommodated in the surface cavities giving fewer contacts with amino acids, and suggesting a consequent lower contribution to the complex stability. Moreover, in the described docked poses the linker chain of **100-108** folds to allow contacts between cap group and amino acids on the surface of HDAC1-4, HDAC7 and HDAC8, but this entropic loss is not compensated by extended interactions with the macromolecular counterparts. These theoretical findings agree with experimental observations that all selective inhibitors for HDAC6, tubacin (**100**),[208] mercaptoacetamides (**101-104**),[193b] and their analogues (**105-108**)[209]presented longer spacers, differently from class I selective inhibitors

whose optimal linker length is six carbons.[168,183,147,161] As recently reported (**109**, Scheme 2.8),[210] the right combination of linker length with a large and rigid cap group can also dictate the selectivity for the isoform 6. The **109** presents a shorter linker length compared to the other selective HDAC6 compounds (**100**-**108**), but the tolyl linker combined with the tricycle confers a bent conformation to **109**, favouring tight interactions with the rim of hydrophobic channel of this isoform. On the contrary, the docked poses of **109** in the binding cavity of HDAC1-4, HDAC7 and HDAC8 shows steric clashes with amino acids on the proteins surface, leading to unfavourable ligand-enzyme binding.

2.6.1.9 HDAC7

Up to date, there are not selective inhibitors of HDAC7. Thus, the docking results of compounds **5**, **7**, **9**, and **78**-**108**, the structural features of the catalytic domain and of the surface pockets of HDAC7 may suggest interesting elements for designing selective inhibitors of this isoform. In detail, the unique sequence of HDAC7 gives rise to a novel zinc binding motif. This protein domain is formed by a β-hairpin positioned by two antiparallel β-strands (β3 and β4) and the loop between helices α1 and α2, which outlines a distinct and only groove contiguous to the opening of the active site channel.[211] This enlarged active site of HDAC7 could be able to harbour a well tailored metal binder, conferring selectivity and improving the affinity of new inhibitors for this enzyme. Selective inhibitors **81**-**90** presented a metal binder with an appendage able to interact with the 14 Å internal cavity of class I proteins. These docking results on HDAC7 showed that these selective inhibitors did not coordinate the prosthetic group of the enzyme and did not fill the enlarged active site due to the steric hindrance. The reported predicted

bioactive conformation of **7** is bound to the zinc ion in a monodentate fashion. On the contrary, the remainder docked inhibitors, presenting a classical chelating agent without decorative appendages, coordinate the zinc ion. Thus, in the hypothesis to design selective binders of HDAC7, new protuberances decorating the metal binder should be projected to selectively match the active site of HDAC7. It could be suggested to insert two flexible appendages flanking the metal binder of the putative ligand. It was observed a deep hydrophobic pocket near the Phe679, which is a constituting residue of hydrophobic channel harbouring the acetylated lysine. This cavity is delimited by the amino acids His531, His541, Pro542, Glu543, Ile628 and Phe679 (Figure 2.38), and it is a peculiarity of HDAC7. The docked poses of **5**, **7**, **9**, and **78**-**109** did not show interactions with this macromolecular counterpart by their cap groups. Thus, a capping moiety, able to establish contacts with this unique pocket on the surface, can be another structural element addressing selectivity for this enzyme.

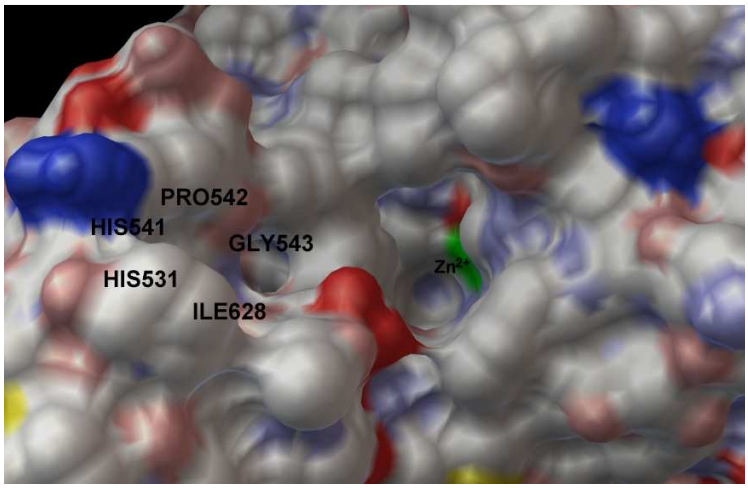


Figure 2. 38 Three dimensional model of HDAC6. The protein is represented by molecular surface, coloured by atom type (C, grey; polar H, white; N, dark blue; O, red; S, yellow; Zn^{2+},green).

2.6.2. Proof of concept: Design, synthesis and biological evaluation of selective HDAC2 inhibitors

2.6.2.1 Design of 110-112

On the basis of the previous analysis aimed to find the structural elements responsible for a specific recognition of HDAC isoforms, three molecular probes (**110**-**112**, Scheme 2.8) were designed for the selective inhibition of HDACs. In particular, this design relied on the use of the weak carbonyl group as chelating agent in order to emphasize the contribution to selective binding by the other structural moieties. The design was focused on a metal binder with three different appendages in order to probe the influence on the class I and isoform selectivity. The three compounds (**110**-**112**, Scheme 2.8) were docked on all considered HDACs. As revealed by the theoretical investigation, the metal binder showed a Class I recognition preference. In fact, the structure of the metal binder presents an appendage able to establish contacts with the internal cavity near the catalytic site. Moreover, changing the structure of the decorative element of the metal binder, it was observed isoform selectivity as predicted by the investigation of **5**, **7**, **9**, and **78**-**109**. Indeed, compound **110** selectively reached the catalytic site of HDAC2. In all other isoforms, **110** was not able to coordinate the zinc ion, suggesting an exclusive binding to HDAC2 as confirmed by the biological assays (see below). In particular, the carbonyl moiety coordinates the zinc ion, whereas the NH of the amide group, as suggested by our analysis, establishes hydrogen bonds with the CO of Gly154 (Figure 2.39).

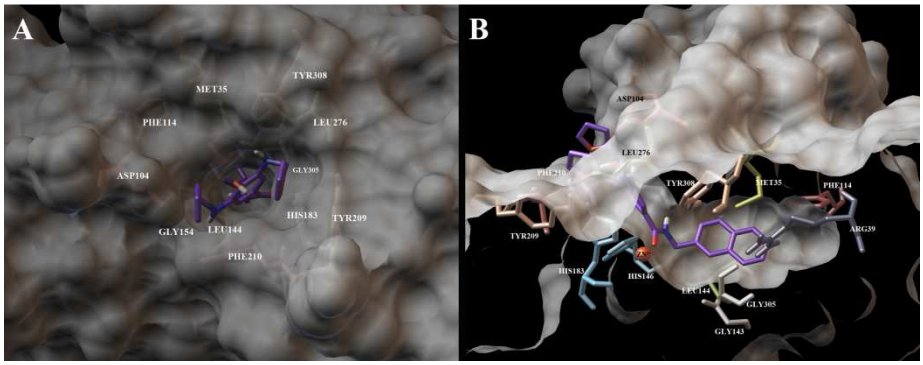


Figure 2. 39 Three-dimensional model of the interactions between **110** and HDAC2. The protein is represented by molecular surface and the zinc ion is depicted in dark orange cpk. The side chains of amino acids are represented by tube, and coloured as: gly, white; asp, red; phe, brown; his, light blue; leu, green; cys, yellow; arg, dark blue; met, olive. The **110** is depicted in purple tube and the atoms are coloured as: C, purple; polar H, white; N, dark blue; O, red.

The naphthalene was accommodated in the 14 Å internal cavity, giving hydrophobic interactions with the surrounding protein amino acids: Met35, Phe114, Gly143, Leu144, Gly305, Tyr308 (Figure 2.39). The aromatic linker interacted by π-π contacts with the hydrophobic channel conducting to the zinc ion. One of the two furans was accommodated in a shallow pocket delimited by residues His183, Tyr209, Phe210 and Leu276, and it also established π-π interactions with Tyr209 (Figure 2.39). The other furan ring established anion-π interaction with Asp104. The arrangement of these two aromatic groups of the capping moiety, induced by the rim shape of the channel, favoured the coordination of the zinc. It is noteworthy that we designed this cap group in order to discriminate between the isoforms of class I. Indeed, besides the HDAC1 and HDAC2 selectivity driven by the metal binder, it was observed that this cap group interacts differently with HDAC1 and HDAC2, leading to a preference towards isoform 2. In the theoretical model obtained by docking calculations, it was observed that the interactions of the cap group with the surface counterpart of HDAC1 prevented the right approaching of metal

binder to the zinc, predicting absence or at least scarce biological activity. Compound **111** and **112** presented a different appendage of the metal binder, which caused a different affinity and selectivity for the investigated isoforms. In compound **111**, the reduced dimensions of the metal binder caused fewer interactions with the internal cavity, justifying a lower affinity for the HDAC2. Moreover, the modification of metal binder structure gave rise to a possible binding to the HDAC8, as suggested by our calculations. The same considerations can be made for the **112**. Respect to **110** and **111,** compound **112** presents a methylene between the chelating agent and the linker. This structural modification does not favour the interactions given by the linker and the cap group with HDAC2. Indeed **112** showed a lower binding affinity with HDAC2, with respect to the other compounds.

These designed small molecules were synthesized and tested by biological assays on all considered HDAC isoforms.

Compounds **110**-**112** displayed significant selectivity in the in vitro inhibition tests against the HDAC1-4 and HDAC6-8 (Table 2.8). As expected, all compounds were not able to inhibit class II HDACs, and showed a selective isoform binding among proteins of class I. As predicted by the docking studies, **110** showed selectivity on HDAC2, thus confirming that the structure of cap group and metal binder is an important determinant for the biological activity towards isoform 2.

In particular, the capping moiety allowed to discriminate between HDAC1 and HDAC2. Compared to **110**, compound **111** presented a lower activity against HDAC2 in agreement to the docking results, due to the less extended interactions given by the pyrrolidine with the internal hydrophobic cavity. There was also an activity on HDAC8, probably due to the modified metal binder.

Table 2. 8 In vitro inhibitory activity of **110-112** against HDAC1-4 and HDAC6-8 (IC_{50}, M)[a]

	HDAC subtype						
compound	1	2	3	4	6	7	8
110	-	8.06×10^{-5}	-	-	-	-	-
111	-	> 100 μM	-	-	-	-	2.96×10^{-5}
112	-	-	4.58×10^{-5}	-	-	-	6.26×10^{-5}
TSA	1.01×10^{-8}	1.45×10^{-8}	1.72×10^{-8}	7.19×10^{-6}	1.71×10^{-9}	3.65×10^{-6}	2.52×10^{-7}

[a]Values are the means of three experiments. Compounds **110-112** were tested in 10-dose IC_{50} mode with 3-fold serial dilution starting at 100 μM. TSA was tested in a 10-dose IC_{50} with 3-fold serial dilution starting at 10 μM, and starting at 20 μM with Class2A substrate. IC_{50} values were extracted by curve-fitting the dose/response slopes. Screening was performed by Reaction Biology Corp. (www.reactionbiology.com/).

Concerning compound **112**, the experimental data confirmed the theoretically expected (see Proof of concept) lower/absent selectivity of binding to HDAC2 as theoretically foreseen. Moreover, thank to the cap group an inhibitory activity on HDAC1 was not detected. Compound **112** showed a comparable inhibition of HDAC3 and HDAC8.

In summary, **110** is a selective inhibitor of HDAC2, even though at modest potency. The evaluated inhibitory activity on these enzymes is in line with the theoretical findings, confirming the predicted structural observations.

2.6.3 Computational Details

2.6.3.1 Homology modeling

The amino acid sequences of Human HDAC1 (Genbank Accession Number Q13547,482aa), HDAC2 (Genbank Accession Number AAH-31055,488aa), HDAC3 (Genbank Accession Number AAH-00614,428aa), HDAC4

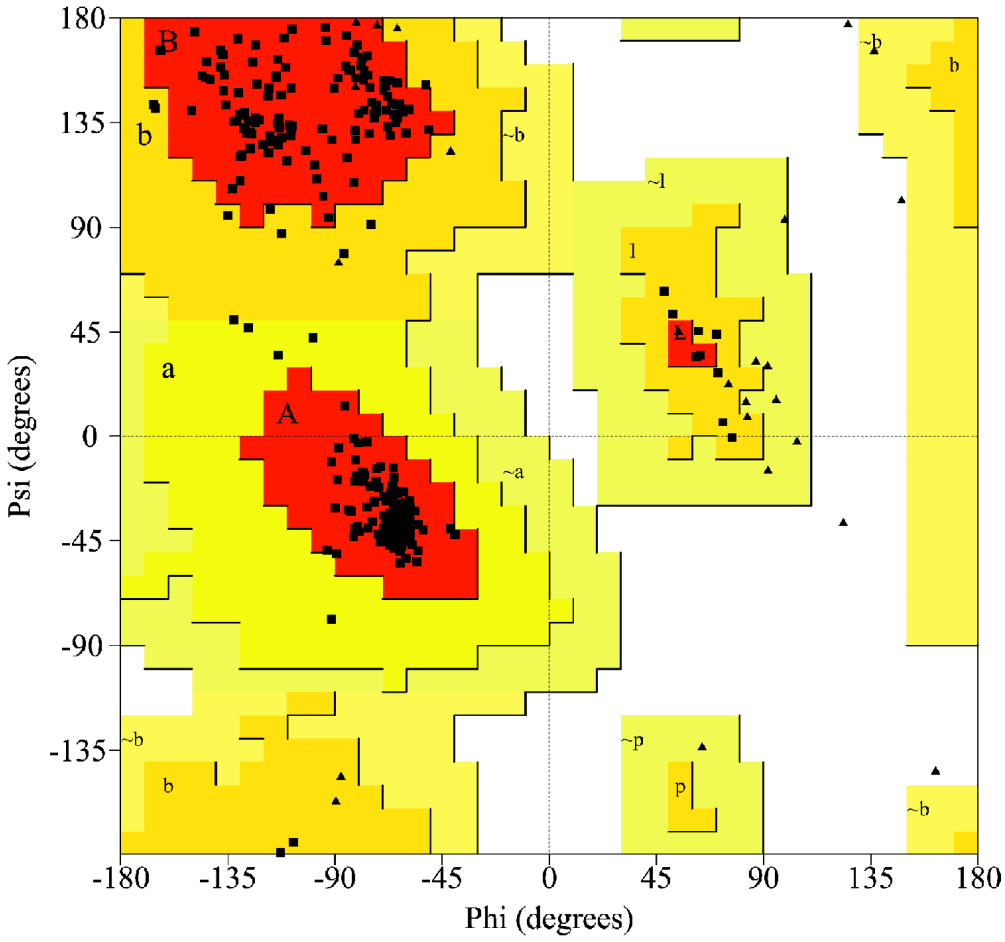
(Genbank Accession Number BAA22957,1097aa), and HDAC6 (Genbank Accession Number Q6NT75,1215aa), were extracted from the NCBI protein sequence database. The BLAST (Basic Local Alignment Search Tool)[212] search was performed to find homologous proteins in the PDB database, applying the BLOSUM62[213] (BLOcks of amino acid SUbstitution Matrix) matrix. The search of homologous proteins was run by the Chimera 1.5.3 package.[174] The resulting alignments were examined and modified manually. HDAC1 and HDAC3 have additional segments in their C-terminal domains that are about 50-110 amino acids long. When these portions of the sequences were subjected to a BLAST search, no alignment was possible, and no similar sequences (other than themselves) were found. Moreover, the function of these residues was proposed to recruit other enzymes to large protein complexes that may regulate their activities.[195,214] Thus, they may have less influence on the substrate/inhibitor binding. Due to the lack of structural information on these portions, they were omitted in the model building. The three-dimensional structure of HDAC2 chain A (PDB code: 3MAX)[150] and HDAC4 (PDB code: 2VQM)[151] were used as templates for human HDAC1 and HDAC3 and HDAC6 homology models building, respectively. HDAC6 differs from other HDACs, for the presence of two catalytic domains (HDAC6 CD I and HDCA6 CD II) sharing 46% sequence identity and 60% similarity. In general, CD1 and CD2 show the same relevant amino acid residues in the active site, whereas more differences can be observed in the loop regions. A recent study, using natural and synthetic substrates, showed that the second catalytic site is the major functional domain of HDAC6.[215] In particular, the authors demonstrated that the inhibition of HDAC6 can be solely ascribed to the interaction of ligands with the second binding domain. Moreover, recently, docking studies on HDAC6 have been performed using the second catalytic domain of

HDAC6.[210] The resulting alignments were used as input for the automated homology modeling program MODELER.[216] The number of generated loops was set to five along with high optimization level for models and loops. The generated models of HDAC1, HDAC3 and HDAC6, showing the lowest energy and minor number of restraint violations, were selected. On the obtained homology models, hydrogen atoms were added by using the graphical interface Maestro version 6.0, Schrödinger, LLC, New York, NY, 2003. The charges of side chains were assigned considering their pK_a at physiological pH of 7.4. The geometry of the added hydrogen atoms by OPLS force field[217] and steepest descent method (500 steps and convergence threshold 0.5 kJ mol^{-1} $Å^{-1}$) by using the MacroModel 8.5.[159] The quality of the obtained models for HDAC1, HDAC3 and HDAC6 were validated using the software PROCHECK (see below).[218]

Sequence alignment of HDAC1 and the template HDAC2, 93.46 % of identity. (*) indicates identity; (:) denotes strongly similar amino acids and (.) highlights weakly similar.

```
HDAC1       TRRKVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRP
HDAC2       AKKKVCYYYDGDIGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRP
            :::*********:************************************

HDAC1       HKANAEEMTKYHSDDYIKFLRSIRPDNMSEYSKQMQRFNVGEDCPVFDGL
HDAC2       HKATAEEMTKYHSDEYIKFLRSIRPDNMSEYSKQMQRFNVGEDCPVFDGL
            ***.**********:************************************

HDAC1       FEFCQLSTGGSVASAVKLNKQQTDIAVNWAGGLHHAKKSEASGFCYVNDI
HDAC2       FEFCQLSTGGSVAGAVKLNRQQTDMAVNWAGGLHHAKKSEASGFCYVNDI
            *************.*****:****:************************

HDAC1       VLAILELLKYHQRVLYIDIDIHHGDGVEEAFYTTDRVMTVSFHKYGEYFP
HDAC2       VLAILELLKYHQRVLYIDIDIHHGDGVEEAFYTTDRVMTVSFHKYGEYFP
            ************************************************

HDAC1       GTGDLRDIGAGKGKYYAVNYPLRDGIDDESYEAIFKPVMSKVMEMFQPSA
HDAC2       GTGDLRDIGAGKGKYYAVNFPMRDGIDDESYgqIFKPIISKVMEMYQPSA
            ******************:*:*********  ****::******:****
```

```
HDAC1       VVLQCGSDSLSGDRLGCFNLTIKGHAKCVEFVKSFNLPMLMLGGGGYTIR
HDAC2       VVLQCGADSLSGDRLGCFNLTVKGHAKCVEvVKTFNLPLLMLGGGGYTIR
            ******:**************:********.**:****:***********

HDAC1       NVARCWTYETAVALDTEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTN
HDAC2       NVARCWTYETAVALDcEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTp
            *************** **********************************

HDAC1       EYLEKIKQRLFENLRML
HDAC2       EYMEKIKQRLFENLRML
            **:**************
```



Plot statistics

Residues in most favoured regions [A,B,L]	304	93.3%
Residues in additional allowed regions [a,b,l,p]	22	6.7%
Residues in generously allowed regions [~a,~b,~l,~p]	0	0.0%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	326	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	31	
Number of proline residues	14	
Total number of residues	373	

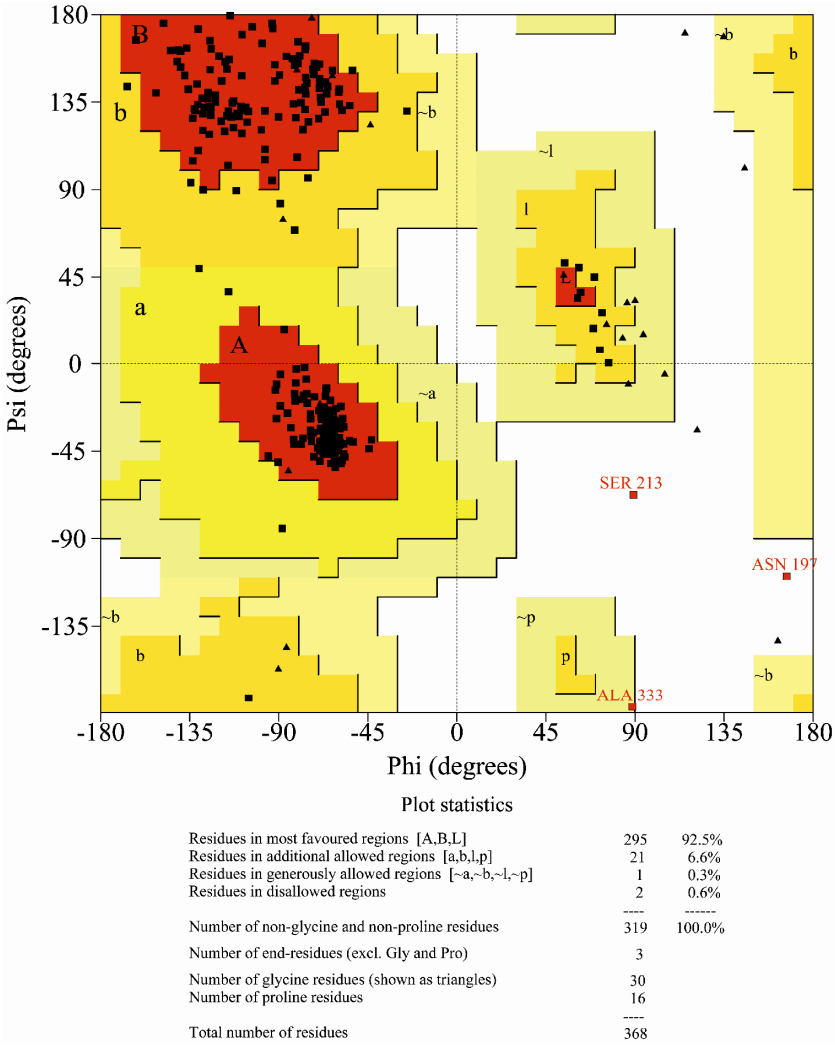
Figure 2. 40 Ramachandran plot of HDAC1 calculated by PROCHECK using a hypothetical resolution of 2 Å.

Figure 2. 41 Superimposition (rmsd = 0.1 Å, calculated by DaliLite 3.1[1]) of HDAC1 (green) and HDAC2 (yellow). The proteins are represented in ribbons.

Sequence alignment of HDAC3 and the template HDAC2, 63.54 % of identity. (*) indicates identity; (:) denotes strongly similar amino acids and (.) highlights weakly similar.

```
HDAC3       KTVAYFYDPDVGNFHYGAGHPMKPHRLALTHSLVLHYGLYKKMIVFKPYQ
HDAC2       KkVCYYYDgDIGNYYYGqGHPMKPHRIrMTHNLLLNYGLYRKMeIYRPHK
            *.*.*:** *:**::** ********: :**.*:*:****:** :::*::

HDAC3       ASQHDMCRFHSEDYIDFLQRVSPTNMQGFTKSLNAFNVGDDCPVFPGLFE
HDAC2       ATaEEMtKYHSDEYIkFLRsIrPdNMSeYSKQMQrFNVGEDCPVFdGLFE
            *: .:* ::**::**.**: : * **. ::*.:: ****:***** ****

HDAC3       FCSRYTGASLQGATQLNNKICDIAINWAGGLHHAKKFEASGFCYVNDIVI
HDAC2       FCQlsTGGSVaGAVKLNRQqtDMAVNWAGGLHHAKKsEASGFCYVNDIVL
            **.  **.*: **.:**.:  *:*:*********** ***********:

HDAC3       GILELLKYHPRVLYIDIDIHHGDGVQEAFYLTDRVMTVSFHKYGNYFFPG
HDAC2       AILELLKYHqRVLYIDIDIHHGDGVEEAFYtTDRVMTVSFHKYGEY-FPG
            .******** ***************:**** *************:* ***

HDAC3       TGDMYEVGAESGRYYCLNVPLRDGIDDQSYKHLFQPVINQVVDFYQPTCI
HDAC2       TGDLrDIGAgKGKYYAVNfPMRDGIDDESYgQIFKPIISKVMEMYQPSAV
            ***: ::** .*:**.:*.*:******:** ::*:*:*.:*:::***:.:

HDAC3       VLQCGADSLGCDRLGCFNLSIRGHGECVEYVKSFNIPLLVLGGGGYTVRN
```

```
HDAC2       VLQCGADSLSgDRLGCFNLTVKGHAKCVEvVKTFNLPLLMLGGGGYTIRN
            *********. ********:::**.:*** **:**:***:*******:**

HDAC3       VARCWTYETSLLVEEAISEELPYSEYFEYFAPDFTLHPDVSTRIENQNSR
HDAC2       VARCWTYETAVaLDceIpNELPYNDYFEYFGPDFkLHiSpSN-MtNQNTp
            *********:: ::  *.:****.:*****.***.** . *. : ***:

HDAC3       QYLDQIRQTIFENLKML
HDAC2       EYMEKIKQrLFENLRML
            :*:::*:* :****:*
```



Plot statistics

Residues in most favoured regions [A,B,L]	295	92.5%
Residues in additional allowed regions [a,b,l,p]	21	6.6%
Residues in generously allowed regions [~a,~b,~l,~p]	1	0.3%
Residues in disallowed regions	2	0.6%
Number of non-glycine and non-proline residues	319	100.0%
Number of end-residues (excl. Gly and Pro)	3	
Number of glycine residues (shown as triangles)	30	
Number of proline residues	16	
Total number of residues	368	

Figure 2. 42 Ramachandran plot of HDAC3 calculated by PROCHECK using a hypothetical resolution of 2 Å.

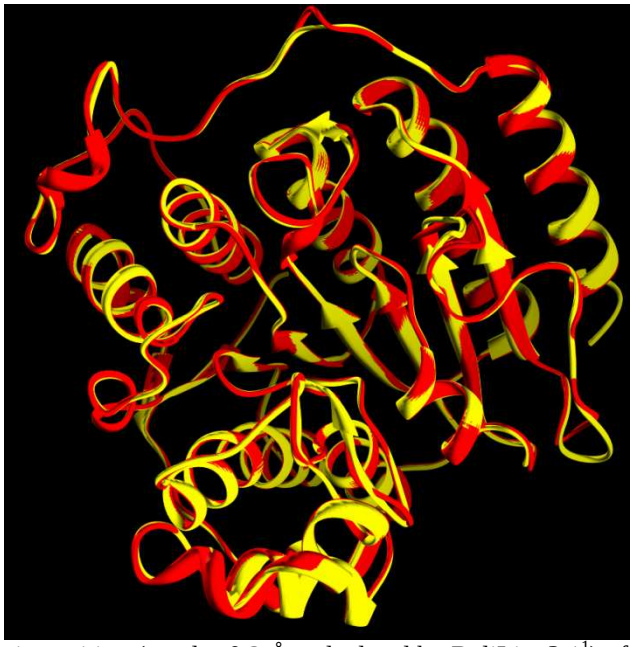
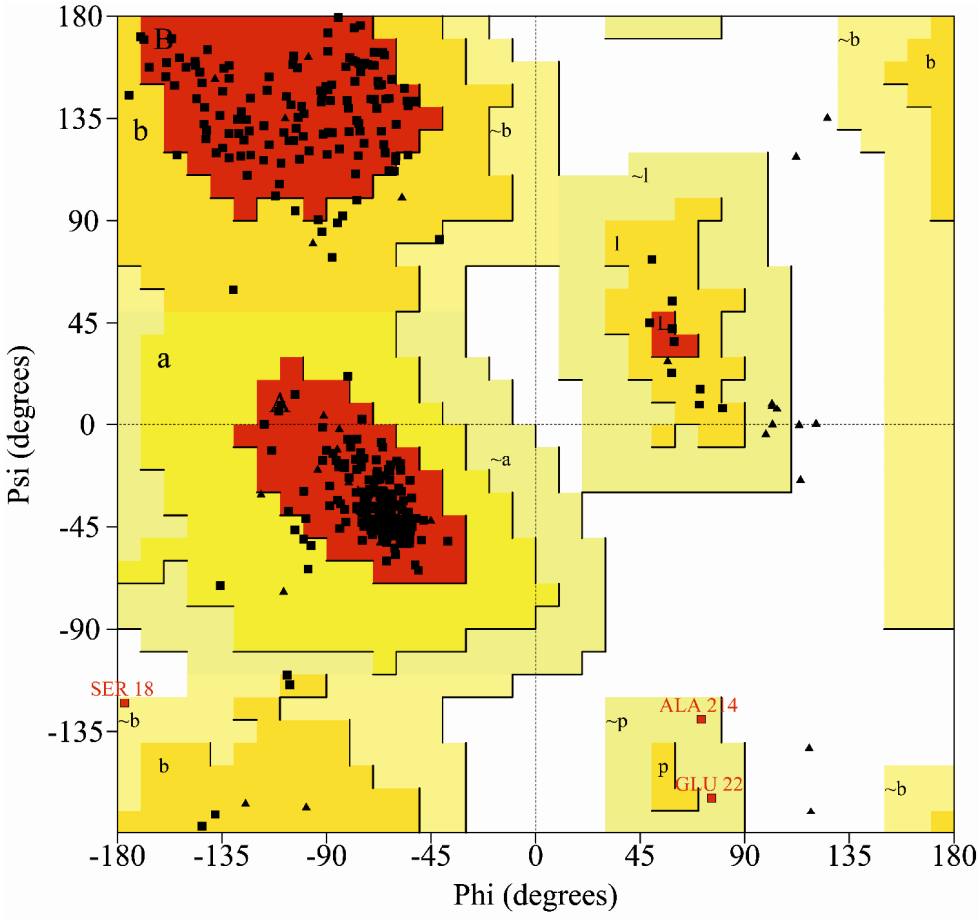


Figure 2. 43 Superimposition (rmsd = 0.2 Å, calculated by DaliLite 3.1[1]) of HDAC2 (yellow) and HDAC3 (red). The proteins are represented in ribbons.

Sequence alignment of catalytic domain II of HDAC6 and the template HDAC4, 47.76 % of identity. (*) indicates identity; (:) denotes strongly similar amino acids and (.) highlights weakly similar.

```
HDAC4          PRFTTGLVYDTLMLKHQCTCG/H----AGRIQSIWSRLQETGLRGKCECI
HDAC6_II       ----TGLVYDQNMMNHCNLWDSHHPEVPQRILRIMCRLEELGLAGRCLTL
                   ******  *::*    . *    . **  * .**:* ** *:*  :

HDAC4          RGRKATLEELQTVHSEAHTLLYGTNPLNRQKKLLGSLASVFVRLPCGGVG
HDAC6_II       TPRPATEAELLTCHSAEYVGHLRATEKMKT-RELHRESS----------
                 * **  ** * **  :.    :.   :  : *   :*

HDAC4          VDSDTIWNEVHSAGAARLAVGCVVELVFKVATGELKNGFAVVRPPGHHAE
HDAC6_II       -NFDSIYICPSTFACAQLATGAACRLVEAVLSGEVLNGAAVVRPPGHHAE
                : *:*:    : ..*:**.*.. .**  * :**: ** **********

HDAC4          ESTPMGFCYFNSVAVAAKLLQQR-LSVSKILIVDWDVHHGNGTQQAFYSD
HDAC6_II       QDAACGFCFFNSVAVAARHAQTISGHALRILIVDWDVHHGNGTQHMFEDD
               :.:. ***:********:  *     . :***************: * .*

HDAC4          PSVLYMSLHRYDDGNFFPGS--GAPDEVGTGPGVGFNVNMAFTGGLDPPM
HDAC6_II       PSVLYVSLHRYDHGTFFPMGDEGASSQIGRAAGTGFTVNVAWNG---PRM
               *****:******.*.***  .  **..::* ..*.**.**:*:.*   * *
```

```
HDAC4          GDAEYLAAFRTVVMPIASEFAPDVVLVSSGFDAVEGHPTPLGGYNLSARC
HDAC6_II       GDADYLAAWHRLVLPIAYEFNPELVLVSAGFDAARGD--PLGGCQVSPEG
               ***:****:: :*:*** ** *::****:****..*.  **** ::*..

HDAC4          FGYLTKQLMGLAGGRIVLALEGGHDLTAICDASEACVSALLGNELDPLPE
HDAC6_II       YAHLTHLLMGLASGRIILILEGGYNLTSISESMAACTRSLLG---DPPPL
               :.:**: *****.***:* ****::**:*.::  **. :***   ** *

HDAC4          KVLQQRPNANAVRSMEKVMEIHSKYWRCLQRTTS---------------
HDAC6_II       LTLPRPPLSGALASITETIQVHRRYWRSLRVMKVEDREGPSSSKLVTKKA
                .* : * :.*: *: :.:::* :***.*:  .

HDAC4          --TAGRSLIEAQTCENE--------------------------------
HDAC6_II       PQPAKPRLAERMTTREKKVLEAGMGKVTSASFGEESTPGQTNSETAVVAL
                 .*   * *  * .::

HDAC4          --------------------------------------------------
HDAC6_II       TQDQPSEAATGGATLAQTISEAAIGGAMLGQTTSEEAVGGATPDQTTSEE


HDAC4          --------------------------------------------------
HDAC6_II       TVGGAILDQTTSEDAVGGATLGQTTSEEAVGGATLAQTTSEAAMEGATLD


HDAC4          --------------------------------------------------
HDAC6_II       QTTSEEAPGGTELIQTPLASSTDHQTPPTSPVQGTTPQISPSTLIGSLRT


HDAC4          --------------------------------------------------
HDAC6_II       LELGSESQGASESQAPGEENLLGEAAGGQDMADSMLMQGSRGLTDQAIFY


HDAC4          --------------------------------------------------
HDAC6_II       AVTPLPWCPHLVAVCPIPAAGLDVTQPCGDCGTIQENWVCLSCYQVYCGR


HDAC4          --------------------------------------------------
HDAC6_II       YINGHMLQHHGNSGHPLVLSYIDLSAWCYYCQAYVHHQALLDVKNIAHQN


HDAC4          ----------
HDAC6_II       KFGEDMPHPH
```



Plot statistics

Residues in most favoured regions [A,B,L]	303	91.0%
Residues in additional allowed regions [a,b,l,p]	27	8.1%
Residues in generously allowed regions [~a,~b,~l,~p]	3	0.9%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	333	100.0%
Number of end-residues (excl. Gly and Pro)	3	
Number of glycine residues (shown as triangles)	32	
Number of proline residues	24	
Total number of residues	392	

Figure 2. 44 Ramachandran plot of HDAC6 calculated by PROCHECK using a hypothetical resolution of 2 Å.

Figure 2. 45 Superimposition (rmsd = 0.7 Å, calculated by DaliLite 3.1[1]) of HDAC4 (purple) and HDAC6 (white). The proteins are represented in ribbons.

2.6.3.2 Docking calculations

The homology models for HDAC1, HDAC3 and HDAC6 were used in molecular docking calculations, along with the X-ray structures of HDAC2 (3MAX),[150] HDAC4 (2VQM),[151] HDAC7 (3C0Z)[152] and HDAC8 (3F0R).[153] In particular, the crystal structure of HDAC8 complexed with CRAA-A (1VKG)[154] was also considered in the calculations, in order to explore the large sub-pocket created by the shift of Phe152 upon ligand binding, and located in the hydrophobic active site channel. Molecular docking studies were performed using AutoDock 3.0.5.[55] HDACs (histone deacetylase-like protein) are metalloproteins, so a non-bonded model for metallic center according to the non-bonded Zn^{2+} parameters of Stote[164] (Zinc Radius = 1.10 Å, well depth = 0.25 kcal/mol) were used. In order to have an accurate weight of the electrostatics, the partial charge of Zn^{2+} and of the amino acids constituting the

catalytic center by DFT calculations m05[194] were derived level by the 6-31+G(d) basis set and ChelpG method[148] for population analysis (Gaussian 03 Software Package).[155] By using the same theoretical level, the partial charges of **110**-**112** were achieved and were used in the subsequent docking calculations. All ligands structures were built using the graphical interface Maestro version 6.0, Schrödinger, LLC, New York, NY, 2003, and their geometries optimized through MacroModel 8.5[159] and using the MMFFs force field.[158] For the **109**, the tertiary amine on the tricyclic ring system was protonated and the two enantiomers were considered in our theoretical studies. Monte Carlo Multiple Minimum (MCMM) method (10000 steps) of the MacroModel module was used in order to allow a full exploration of the conformational space. The so obtained geometries were optimized using the Polak-Ribier conjugate gradient algorithm (maximum derivative less than 0.001 kcal/mol). A GB/SA (generalized Born/surface area) solvent treatment[163] was used, mimicking the presence of H_2O in the geometry optimization and in the conformational search steps.

.

-CHAPTER 3-

A Novel Potent Nicotinamide Phosphoribosyltransferase Inhibitor

3.1. A Novel Potent Nicotinamide Phosphoribosyltransferase Inhibitor Synthesized via Click Chemistry

The inhibition of NAD synthesis or salvage pathways has been proposed as a novel target for antitumoral drugs. Two molecules with this mechanism of action are at present undergoing clinical trials **113** (APO866)[219] and **114** (GMX1777) (Scheme 3.1),[220] In searching for similar novel molecules, the most promising triazole-based compound was identified between a library of 185 novel APO866 analogues[221] and, by molecular docking, a rationale of inhibition mechanism via was provide

While the pharmacological fight against cancer has made great advances in the last 20 years, novel molecules to fight this disease are still urgently needed. Many cancers still present unmet therapeutic needs, and chemoresistance is an importantphenomenon within the context of fast cell division and high mutation rates. Furthermore, side effects and safety are major concerns with antitumoral drugs. Although targeted biotechnology-based agents (e.g., monoclonal antibodies and vaccines) have and are being developed for a number of cancers, it is obvious that small-molecule drugs will result in lower costs and might be able to combat a wider range of tumors. Furthermore, the development of novel agents might also allow researchers to discover synergisms that will reduce the doses required for single agents used in combination, increasing efficacy while reducing side effects. For this to occur it is imperative that novel targets be exploited and that, for each target, a number of therapeutic agents (for example, with different pharmacokinetic profiles or different organotropisms) become available to the clinician.

In this context, it has recently been proposed that interfering with NAD(P) levels might lead to cell death of those cells that have a high usage rate of this

pyridine nucleotide, that is, tumoral cells with a high division rate.[222] While the general perception is that NAD(P) is mainly used as an enzymatic cofactor (and, as such, its depletion should be slow as it participates mainly in redox reactions), it is now accepted that a number of enzymes consume NAD(P). For example, NAD is the substrate for a specific subclass of histone deacetylases (sirtuins),[223,224,225] as well as mono- and poly-ADP ribosylating enzymes (e.g., PARPsa).[225,224] Furthermore, NAD-(P) is also the precursor for a number of Ca^{2+}-releasing second messengers (e.g., cADPR, NAADP) and as such is consumed by enzymes such as CD38.[226,227] Eukaryotic cells possess several mechanisms to replenish NAD levels, including a de novo pathway from tryptophan and at least two salvage/recycling pathways.[228,229] The most important of these latter two pathways relies on the enzyme nicotinamide phosphoribosyltransferase (NMPRTase or NAMPT), which converts nicotinamide into nicotinamide mononucleotide (NMN), which is subsequently converted to NAD by NMN adenyltransferase (NMNAT).

Beyond its well-described role in cellular metabolism, intracellular nicotinamide adenine dinucleotide (NAD) levels have been shown to affect the enzymatic activity of a series of NAD-dependent enzymes, influencing biological responses such as cell survival and inflammation. Nicotinamide phosphoribosyl transferase activity has been shown to be essential for maintaining adequate intracellular NAD levels, suggesting that this enzyme may in fact play a central role in modulating the activity of a wide range of NAD-dependent enzymes.[230] Several recent observations concur with this hypothesis and suggest that by regulating NAD availability, NMPRTase is able to control both cell viability and the inflammatory response. Nampt may thus represent a novel pharmacological target with valuable anti-inflammatory and antitumor properties (Figure 3.1).

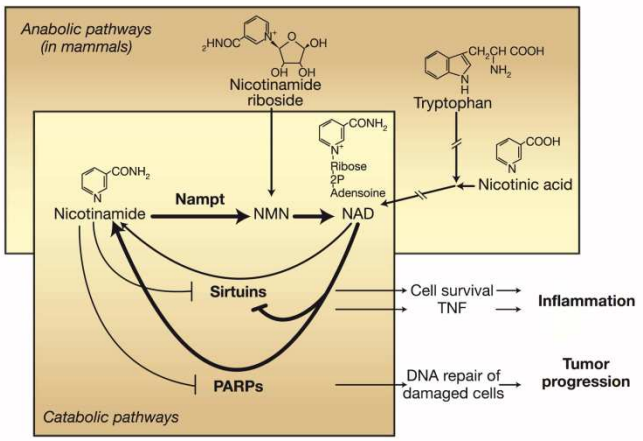
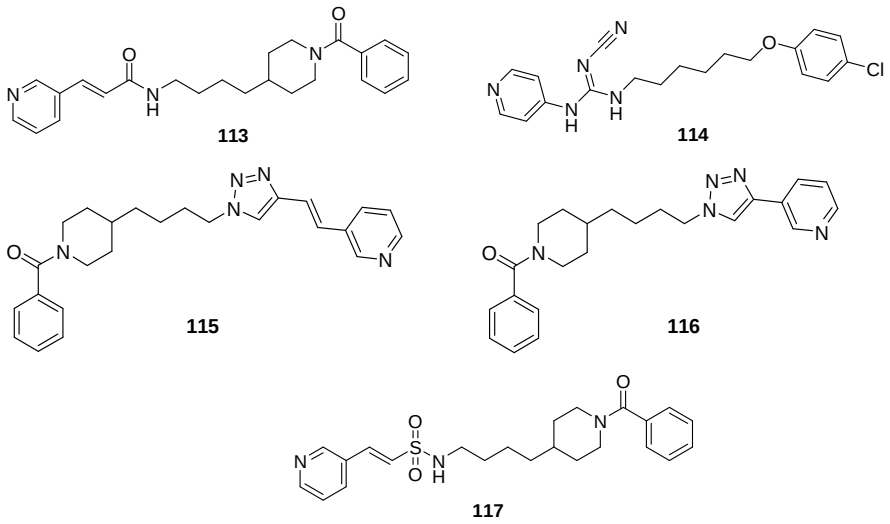


Figure 3. 1 NAD metabolism and regulation of cellular functions.

NMPRTase has been exploited as a target for developing a potential antitumoral drug (**113**, initially known as FK866 and now renamed APO866).[219] This compound has an IC_{50} for cytotoxicity of approximately 1nM against several cancer cell lines, and when evaluated preclinically it has been demonstrated to possess very promising antitumoral actions against both solid tumors and leukemia cells.[219,231] Phase I clinical trials have been completed, and this compound is at present in phase II alone or in combination with other antitumoral drugs.[232] A second compound, **114** (CHS 828, now renamed GMX1777) (Figure 3.1), is in early clinical development. This compound, originally screened as an antihypertensive drug,[233] showed very high cytotoxic activity.[234] Furthermore, its action, as for **113**, can be reverted by adding nicotinamide, the true substrate for NMPRTase.[219,234] It was recently attempted to replace the amide of **113** with a 1,4-disubstituted triazole using click chemistry, as it has been postulated that this substitution can generate a nonclassic bioisostere.[235] To surprise, the true bioisostere (**115**) was devoid of activity up to micromolar concentrations, while the triazolyl

pyridine derivative (**116**) maintained nanomolar potency even while being approximately 80-fold less potent than **113**.



Scheme 3. 1 NMPRTase inhibitors in clinical trials (**113** and **114**) and analogues of FK866 used to draw the original hypothesis (**115**, **116**, and **117**).

This observation suggests that the 1,4-disubstituted triazole ring is compatible with the binding pocket of the enzyme (Scheme 3.1). It was then performed molecular docking calculations on nicotinamide phosphoribosyltransferase (NMPRTase) and **113**, **115**, and **116**. Tong et al. in 2006 elucidated the X-ray crystal structure of the human NMPRTase-1 complex (PDB code: 2GVJ) with a 2.1 Å resolution.[236] The authors clarified the molecular mechanism for the substrate specificity of this enzyme, defining the binding mode of **113** and the structural basis for its specificity for NMPRTase. The structure revealed a tunnel at the interface between chains A and B that is a potential binding site for inhibitors. The first step of the reported molecular docking studies was the validation of the docking method.

In accordance with that reported by Tong et al.,[236] water molecule 645 was left in the NMPRTase active site during the molecular docking calculations with Autodock3.0.5 software.[55] Moreover, to improve the accuracy of calculated dissociation constant (K_{Dcalc}), it was performed energy and geometry optimizations of **113**, **115**, and **116** and computed the charges of the molecules at the quantum mechanical (QM) level (see computational details). In analogy with the rationalization of the NMPRTase inhibitor pharmacophore elucidated by Tong et al.,[236] the described calculated model for **113** maintains all of the principal interactions with the enzyme: (1) the pyridine ring of the inhibitor is sandwiched between the side chains of Phe193 and Tyr180, and it forms a cation-π with Arg196; (2) the carbonyl oxygen atom and the amide nitrogen of the amide bond form two hydrogen bonds with the hydroxyl of Ser275 and the water molecule, respectively. Moreover, the lead compound establishes hydrophobic interactions with the amino acids of the tunnel formed at the dimer interface, namely, Glu376, Asn377, Arg349, Ile378, Val330, Val350, Ala379, Ile351, and Ile309. In particular, the phenyl ring forms a π-stacking interaction with Tyr188 on the shallow groove of the NMPRTase surface. All of the above-mentioned interactions (Figure 3.2a) contribute to the 1-enzyme complex calculated stability (K_{Dcalc}=1.23x 10^{-9}M) and strengthen the reported model. Analyses of other compounds were conducted by taking into account the similarities and/or differences with the binding mode of **113**. The 1,4-disubstituted triazole rings of compounds **115** and **116** are able to form a hydrogen bond with the OH of Ser275 but cannot form it with the water molecule; this is a drawback that may be associated with the lower activities of both **115** and **116** with respect to the lead compound.

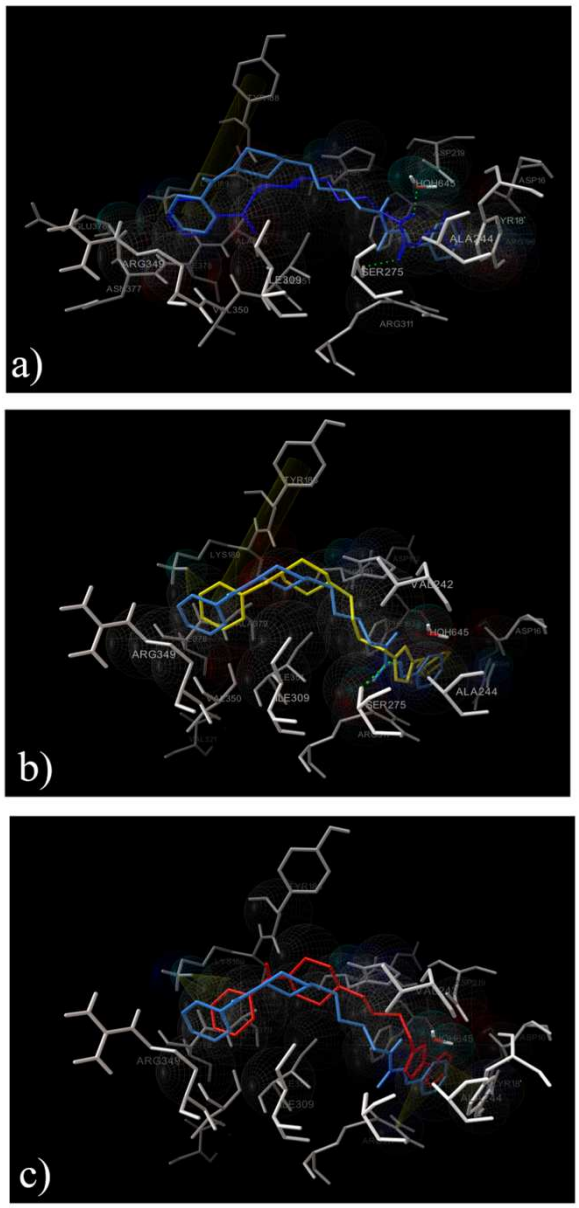


Figure 3. 2 Superimposition of the crystallized **113** structure and the structures calculated by molecular docking in the catalytic site of NMPRTase. The true crystallized **113** structure is present in all panels (light blue), superimposed with the calculated 1 (a, blue), the calculated 3 (b, yellow), and the calculated 4 (c, red). The NMPRTase A and B chains are represented by gray sticks.

The difference in activity between the true bioisostere (**115**) and the triazolyl pyridine derivative (**116**) is due to the different topological position of the pyridine ring (Figure 3.2b). The presence of the double bond between the triazole and pyridine ring causes the loss of π-π stacking with Phe193 and Tyr180, suggesting decreased activity. On the other hand, **116** has the same topological position of the pharmacophoric points of **113** cocrystallized with the enzyme (Figure 3.2c), and moreover the phenyl ring forms a cation-π interaction with Lys189 on the NMPRTase surface (K_{Dcalc}=1.89 x 10^{-9} M). In a set of unpublished data, it was also synthesized a compound bearing a sulfonamide group (**117**, Scheme 3.1**)** in place of the amide moiety of **113**.

For what concern **117**, the sulfonamide group should theoretically be more prone to form a hydrogen bond due to its higher acidity with respect to the amide group. Yet, this compound (either for inability of the drug to cross the plasma membrane or inability to enter the binding pocket) was inactive. This failure led it to exploit the compatibility of the triazolyl pyridine with the active site of NMPRTase to explore the possibility of generating more active analogues.

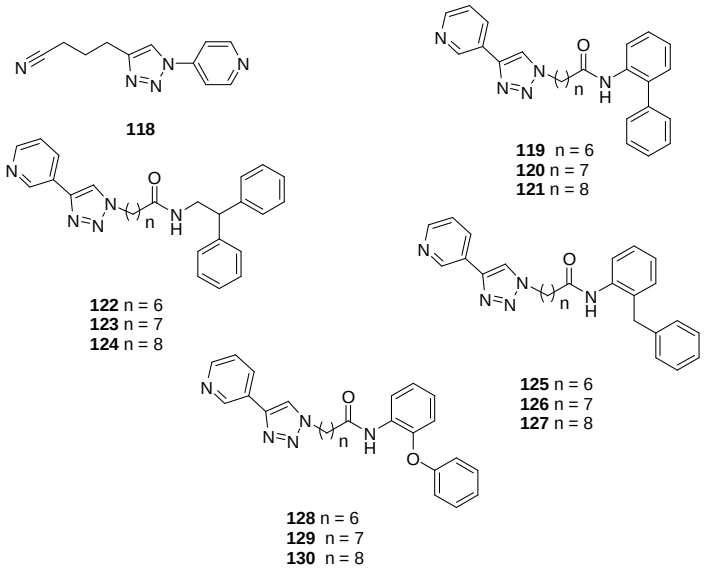
Indeed, the triazolyl pyridine also presents the great advantage of allowing us to capitalize on the simplicity of the click chemistry reaction.[237,238] Its amenability to solution phase parallel synthesis is also an advantage, as it allows the screening of a high number of analogues in a fast and reliable manner.[239,240,241] This, in turn, allows probing of the active site and the rim of the enzyme in more detail. On this basis, 185 triazolyl pyridines was synthesized, using a fast and versatile solution-phase parallel combinatorial synthesis via click chemistry. The most promising of these compounds

displays an IC_{50} for cytotoxicity in vitro of 3.8±0.3 nM and an IC_{50} for NAD depletion of 3.0±0.4 nM.

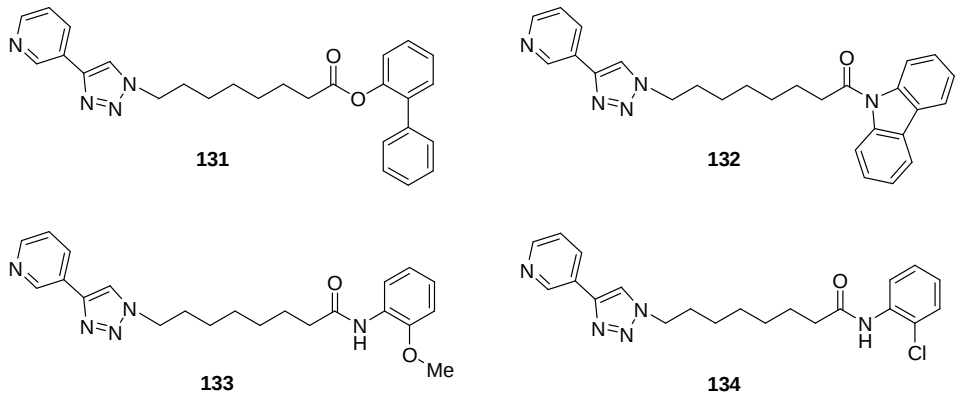
To investigate the activity of the synthesized compounds (Schemes 3.2 and 3.3), it was we decided to exploit a neuroblastoma cell line that had already been characterized as sensitive to **113**.[235] Full concentration response curves were generated, and most compounds were found to have IC_{50} values in the nanomolar range (See Table 3.1).

Compound **120** inhibited the growth of most cell lines tested, with nanomolar potency (GI_{50}) in cell lines derived from leukemia, lung, CNS, colon, melanoma, ovarian, renal, and prostate cancers. To surprise, this compound appeared truly cytotoxic in melanoma cell lines, while in the others it was mainly cytostatic. This might suggest potential areas of therapeutic use. Compound **120** was also further modified to gather additional structural information. In this context, the ester (**131**) as a possible bioisostere of the amide group and the carbazole (**132**), to rigidify the 2-aminobiphenyl scaffold, were synthesized and the phenyl ring at the 2-position of aniline (21, 22) with commercially available ortho-substituted anilines (Figure 3.4) was substituted to explore the role of the o-phenyl. Compound **132** (the rigidified analogue) displayed nanomolar potency (37.4±17 nM), albeit with a loss of potency of approximately 10-fold compared to **120**. Compound **134** displayed a decrease of approximately 50-fold (Table 3.1). Compound **131** did not display any activity. Yet, no conclusion can be drawn for this lack of efficacy as it may be also due to instability of the ester in the medium. The investigations on the induction NAD depletion have revelead that as expected, **113** depleted the cellular NAD pool with an IC_{50} of approximately 1 nM. Similarly, compound **120**, compound **132** (the rigidified analogue), and compounds **123** and **121** also depleted NAD pools with the expected rank order of potency (Table 3.1).

Although enzymatic assays have not been performed, biological data suggest that **120** shares the mechanism of action of **113**.



Scheme 3. 2 Compounds that emerged from the screening as cytotoxic at 1 μM.



Scheme 3. 3 Synthesized analogues of compound **8**.

Table 3. 1 Synoptic Biological Data of the Active Compounds That Emerged from the Screening[a]

compound	Viability (IC_{50} : nM)	NAD levels (% of control) at 1 µM	NAD levels (% of control) at 300 nM	NAD levels (% of control) IC_{50}
118	nd	31.0 ± 13.2	87.8 ± 4.6	/
119	56 ± 13	26.2 ± 12.0	88.2 ± 2.8	/
120	3.8 ± 0.3	11.6 ± 6.19	9.9 ± 7.1	3.0 ± 0.4
121	14.6 ± 3.6	13.0 ± 6.2	35.6 ± 16.7	84.4 ± 6.7
122	nd	16.8 ± 6.9	83.8 ± 2.0	/
123	97 ± 17	23.0 ±9.5	11.0 ± 5.6	49.2 ± 1.8
124	157 ± 50	32.3 ± 6.5	77.3 ± 14.4	/
125	nd	28.0 ±7.6	86.1 ± 6.2	/
126	72 ± 12	30.0 ± 9.7	75.2 ± 5.3	/
127	82 ± 8	13.1 ± 7.0	74.7 ± 7.3	/
128	120 ± 25	24.7 ± 6.0	81.3 ± 4.6	/
129	37 ± 7	17.7 ± 8.7	87.9 ± 1.4	/
130	78 ± 9	22.6 ± 8.6	79.5 ± 5.9	/
131	nd	/	74.7 ± 7.3	/
132	37.4 ± 17.5	/	12 ± 8.0	17.1 ± 0.1
133	nd	/	87.9 ± 1.4	/
134	205 ± 17.6	/	79.5 ± 5.9	/

[a]Viability was evaluated via the MTT method. Values are mean ± S.E.M. of 8-12 determinations. NAD levels were first screened at a fixed concentration (n=10) and, if the compound displayed activity at 300 nM a full concentration-response was performed. Values are mean ± S.E.M. n.d. not determined as full toxicity was not reached at concentrations up to 1 µM.

The main observations that lead to this statement are as follows: (1) both **113** and **120** deplete NAD levels and are cytotoxic at similar concentrations;

(2) both **113** and **120** induce autophagy; (3) the cytotoxic effects of **113** and **120** can be rescued by extracellular NAD application but cannot be rescued by nicotinamide. The metabolic instability of the pyridine moiety is one of themain problems of the two known inhibitors of NMPRTase in clinical trials (1 and 2), which leads to the need for continuous infusion.[242]

On this basis, molecular docking studies were performed on the compounds with nanomolar potency, **120** and its rigid analogue **132**, using the model described in the Introduction. The three-dimensional model of the interactions between NMPRTase and **120** highlights its ability in reproducing the binding mode of **113** (Figure 3.3a). Except for the loss of the hydrogen bond with the water molecule common to all the triazole derivatives, the other pharmacophoric interactions are maintained: π stacking of the pyridine ring with Phe193 and Tyr180, cation-π with Arg196, and hydrogen bonds with the hydroxyl of Ser275. The biphenyl group on **120** remarkably contributes to the stability of the drug-target complex ($K_{Dcalcd8}$ = 4.31 x 10^{-9} M) occupying a shallow groove on the NMPRTase surface formed by Gly185, Tyr188, and Lys189 and interacting with Ile378, Ala379, Val399, and Arg349. The alkyl chain interacts with His191, Val242, Ile350, and Arg31. In the calculated model, analogue **132**presents a binding mode very similar to that of **113** (Figure 3.3b), maintaining the interaction with the Phe193 and Tyr180 aromatic ring and the hydrogen bond with the OH of Ser275. Additionally, the loss of the cation-π interaction of the pyridine ring with Arg196 is balanced by the presence of the carbazole, which gives a cation-π interaction with Lys89 and π stacking with Tyr188 ($K_{Dcalcd20}$ = 3.07 x 10^{-9} M). Moreover, such a rigid group allows better hydrophobic interactions with Glu376, Asn377, Arg349, Lys189, Ala379, and Ile378, and a part of it places on the shallow groove on the enzyme surface (Figure 3.3c) (Gly185, Tyr188, and Lys189).

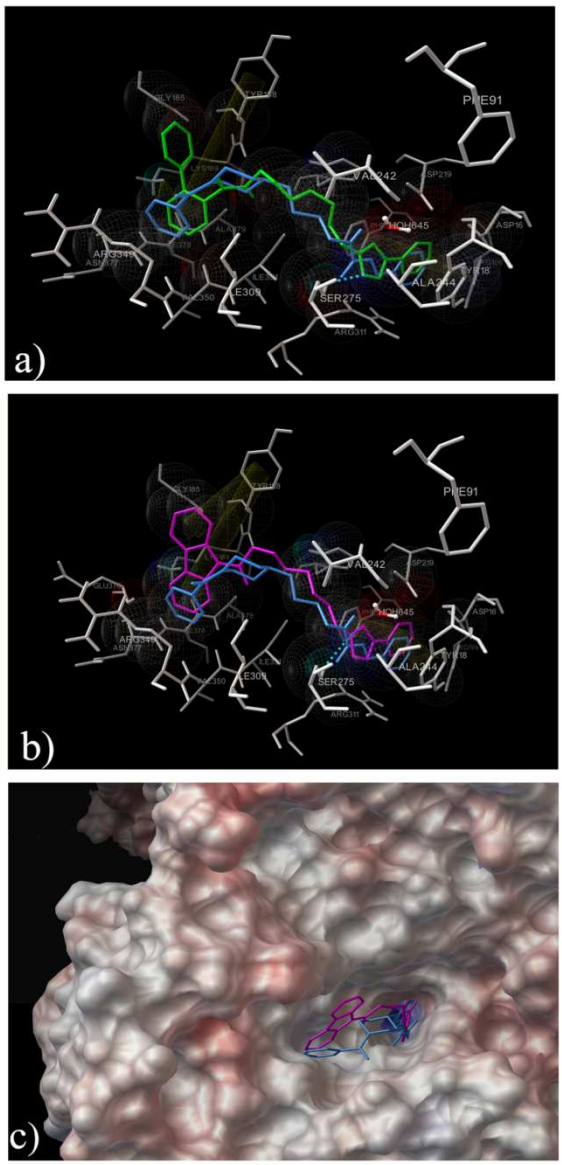


Figure 3. 3 Three-dimensional model of the superimposition in the catalytic site between the true crystallized **113** structure and the calculated structures for compound **120** (a) or compound **132** (b, c). The crystallized **113** structure, **120**, or **132** are depicted in light blue, green, and purple sticks, respectively. The NMPRTase A and B chains are represented by gray sticks (a, b) and by molecular surface colored according to electrostatic potential. Positive potential is shown in blue, and negative potential is in red (c).

The described docking calculations would therefore suggest that the triazole ring, the pyridine ring, and the hydrophobic tail are related to hydrogen bonds, hydrophobic interactions, and aromatic stacking that constitute the driving forces of the target-ligand complexes. Molecular docking calculations also suggest that the triazole group must be linked directly to the pyridine ring, as confirmed by the loss of activity of **115** and that the hydrophobic tail can be increased for a wider interaction with the amino acids on the target surface, as observed for **120** and **132**. Obviously, such speculations will require further analogues to be synthesized and tested to verify this hypothesis.

3.1.1 Computational Details

Molecular docking calculations were performed on a two dual-core Intel Xeon 3.4 GHz, using Autodock 3.0.5 software.[55] For all FK866 analogues a quantum mechanical optimization of the energies and the geometries was performed *in vacuo* at the DFT B3LYP level, using the 6-31G(d) basis set, and on so optimized structure, the charges were calculated at the DFT B3LYP level using the 6-31+G(d) basis set and ChelpG[148] method for population analysis, as implemented in Gaussian 03 Package software.[160] For all the docking studies a grid box size of 50 x 66 x 76 with spacing of 0.375 Å between the grid points, and centered at 14.93 (x), 5.56 (y) and 2.17 (z), covering the tunnel at interface of the dimer was used. For all the docked structures, all bonds were treated as active torsional bonds, except the amide bond. In order to achieve a representative conformational space during the docking calculations, ten calculations consisting of 256 runs were performed, obtaining 2560 structures (256 x 10). The Lamarckian genetic algorithm was used for docking calculations. An initial population of 650 randomly placed individuals, a maximum number of 6.0 x 10^6 energy evaluations, and a

maximum number of 5.0 x 10^6 generations were taken into account. A mutation rate of 0.02 and a crossover rate of 0.8 were used. Results differing by less than 2.5 Å in positional root-mean-square deviation (RMSD) were clustered together and represented by binding energy. All the 3D models were depicted using the Python software,[165] molecular surfaces are rendered using Maximal Speed Molecular Surface (MSMS).[166]

-CHAPTER 4-

Microsomal Prostaglandin E_2 Synthase-1: drug target in inflammation

4.1 Structure-based discovery of mPGES-1 inhibitors

In this chapter, it will be described the design and virtual screening of two series of triazole obtained for directing the click chemistry synthesis as potential inhibitors of microsomal prostaglandin E_2 synthase (mPGES)-1 enzyme. The reported results, in fact, prove the efficiency of the triazole group as a new scaffold useful in the rational design of new promising candidates as antiinflammatory drugs with potential action on other enzymes within the arachidonic acid cascade, such as 5-LO or FLAP. In fact, interference with 5-LO or FLAP, the key enzymes in the formation of leukotrienes (LTs) from arachidonic acid, is considered a valuable characteristic of a given mPGES-1 inhibitor, because dual suppression of PGE_2 and LT formation might be superior over single interference in terms of higher anti-inflammatory efficacy as well as in terms of reduced side effects.[243]

Non steroidal anti-inflammatory drugs (NSAIDs) represent so far the pivot of inflammation therapy as a consequence of their potent effect in the suppression of prostaglandins (PGs), prominent bioactive mediators involved in key physiological functions[244] and also implicated in several pathologic conditions like inflammation and tumorigenesis.[245]

However, especially for long-term treatments - like those required for chronic pathologies such as rheumatoid arthritis - their use comprises severe side effects; in particular NSAIDs are well known to be endowed with relevant gastric toxicity[246] due to the efficient suppression of constitutively generated PGE_2 involving the COX-1 pathway with gastro protection function. Not long ago, the introduction of coxibs in therapy was initially considered as a solution of all the problems connected with the use of NSAIDs, as these selective COX-2 inhibitors showed to exhibit potent anti-inflammatory activity without causing significant gastrointestinal injury. Unfortunately, several clinical

evidences indicated their implication in serious cardiovascular accidents.[247] In this perspective, there is an ever growing need for the research of safer anti-inflammatory drugs. Recently, great attention has been focused on the microsomal prostaglandin E_2 synthase (mPGES)-1 enzyme responsible for the conversion of the COX-derived unstable peroxide PGH_2 into PGE_2; this enzyme is over-expressed in several inflammatory disorders[248] as well as in many human tumors.[249,250,251]

Elevated levels of mPGES-1, in fact, are often observed concomitantly with COX-2 over-expression. In fact, in vitro studies have demonstrated that mPGES-1 is localized at the perinuclear membrane and endoplasmic reticulum and is in general functionally coupled with COX-2,[252] thereby enabling efficient generation of PGE^2 during inflammation.[252] Moreover, recent studies have shown that mPGES-1 expression can be specifically induced by lipopolysaccharide (LPS) in rat peritoneal macrophages,[253] interleukin-1b (IL-1b) and tumor necrosis factor (TNF)-a in a human lung carcinoma cell line, A549 with or without induction of COX-2.[254] However, studies with these diverse stimuli have clearly shown that mPGES-1 can also be functionally activated in the absence of induced COX-2 levels,[254] providing evidence that these two enzymes can be independently regulated. This latter observation is important from the standpoint of drug targeting. It suggests the possibility that the enzymatic activity of mPGES-1 can be pharmacologically targeted with resultant suppression of PGE_2 production by mechanisms that circumvent the toxicity associated with inhibition of COX-2 activity.

To date, a limited number of compounds have been described that inhibit mPGES-1 activity in vitro (Figure 4.1). None as yet have been developed as anti-cancer agents. There are several examples of compounds that were developed to target COX-2 but also found to inhibit mPGES-1 activity as well.

For example, NS-398 inhibits mPGES-1 with an IC50 of ~20 mM,[255] MK-866, ~3 nM,[256] ICMF-63 potently inhibited the human mPGES-1enzyme (IC50 of 1.3 nM), while Triclosan is a waek inhibitor.

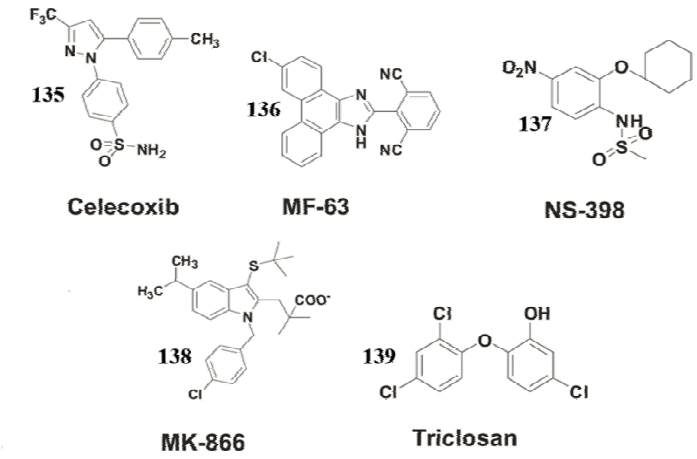


Figure 4. 1 Structures of known compounds that have mPGES-1 inhibitory action.

Inhibition of mPGES-1 has been proposed as a more promising approach for the development of safer drugs for cancer suppression[252a] and in inflammatory disorders[257,258] devoid of classical NSAID side effects, as this inducible enzyme affects the biosynthesis of massive PGE_2 generation as a response to inflammatory stimuli.[259] mPGES-1 is a glutathione-dependent trans-membrane enzyme belonging to the MAPEG (membrane associated proteins involved in eicosanoid and glutathione metabolism) family. This protein family consists of membrane-bound proteins with diverse functions like leukotriene C_4 synthase (LTCS), microsomal glutathione transferase-1 (MGST-1), and 5-lipoxygenase (5-LO)-activating protein (FLAP). Among the three isoforms so far identified for PGES, it is mPGES-1, functionally coupled with COX-2, that seems to be the isoform primarily involved in pathologies.[259] Even if Jegerschöld *et al.*[260] have recently elucidated the electron

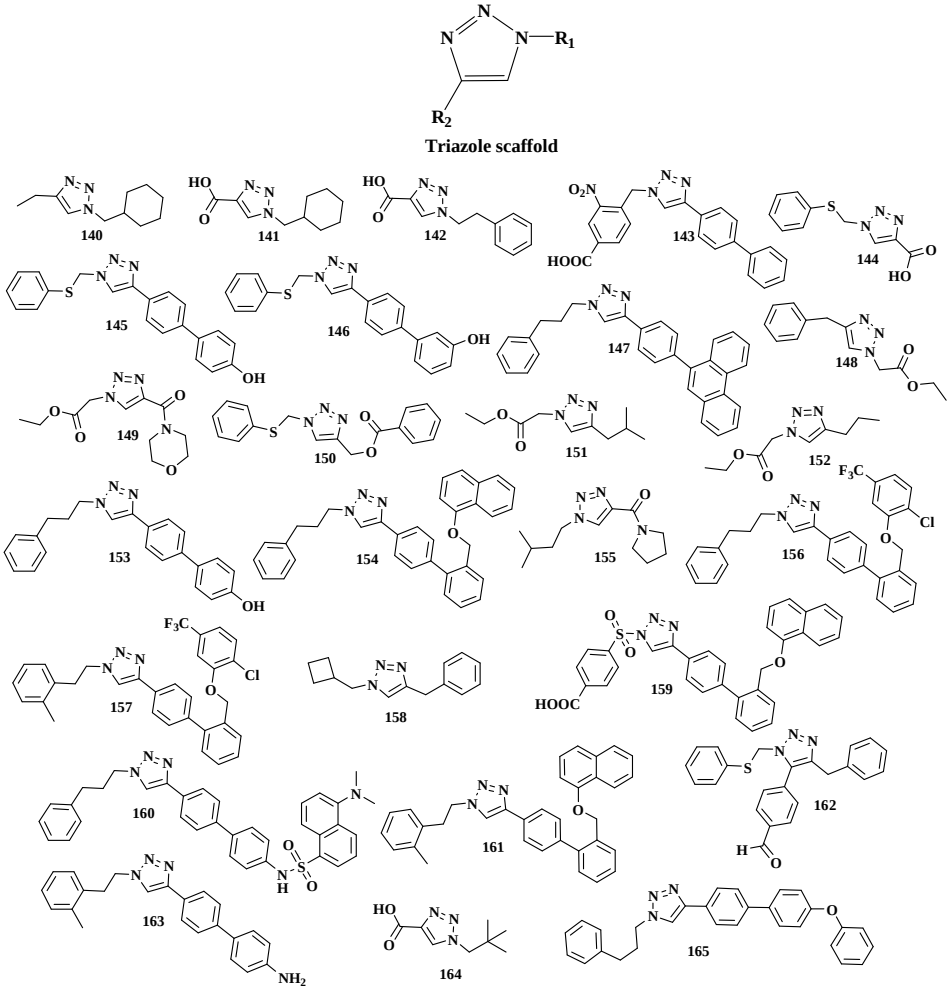
crystallographic structure of closed conformation of mPGES-1, the open form of the protein constitutes a model for the productive enzyme. Therefore the absence of three-dimensional (3D) X-ray crystal structure of open mPGES-1 conformation with a substrate or an inhibitor bound has represented the major difficulty for the rational design of new specific inhibitors, making the classical receptor-based approach quite challenging. In fact, despite many efforts spent in this area, only very few effective *in vivo* mPGES-1 inhibitors has been reported in the literature, therefore the discovery of potent inhibitors of this interesting target would be of great relevance for the development of a new generation of anti-inflammatory agents with potential safer profile. By means of an *in silico* screening, we describe here the development of fast synthetically accessible triazole-based[261,262,263,264,265,266,267,268] compounds, representing innovative scaffolds in this area as potential mPGES-1 inhibitors.

4.2 Structure-Based Discovery of Inhibitors of Microsomal Prostaglandin E_2 Synthase-1, 5-Lipoxygenase and 5-Lipoxygenase-Activating Protein: Promising Hits for the Development of New Anti-inflammatory Agents

In this paragraph, it will be described the results of the rapid *in silico* screening on a small set of triazole derivates that has disclosed three new potential anti-inflammatory drugs: (I) compound **143** displaying selectivity for mPGES-1 with an IC_{50} value of 3.2 μM, (II) compound **159** that dually inhibits 5-lipoxygenase and mPGES-1, and (III) compound **146** apparently acting as 5-lipoxygenase-activating protein inhibitor (IC50 = 0.4 μM). from for directing the click chemistry synthesis of the most promising mPGES-1 inhibitor **4**.[269]

The lack of a 3D X-ray crystal structure of open mPGES-1 conformation has stimulated many efforts for identifying the key characteristics of mPGES-1 inhibitors, based on QM calculations,[270] SAR[271,272] and 3D-QSAR analysis,[273,274,275,276] multistep ligand-based strategy,[277] high-throughput (HTS) screening,[278] molecular modeling and dynamics simulation[279] and site-direct mutagenesis studies.[280] As reported by Friesen *et al.*[258] these efforts have led to the identification of several classes of mPGES-1 inhibitors: fatty acids and PGH_2 analogues,[281] indole and **43** (MK-886, Scheme 4.2) analogues,[282] phenantrene imidazoles,[278] nonacidic agents,[277] and other inhibitors. Considering the well-known characteristic of indole-based agents—the simultaneous contributions to the inhibitory activity on mPGES-1 of hydrophobic and electrostatic effects—and the ring size of fatty acids and PGH2 analogues as starting point, we designed new triazole nucleus templates as potential scaffolds for anti-inflammatory drugs. We designed a small set of compounds (Scheme 4.1) decorating a disubstituted triazole ring, taking into

account both the synthetic accessibility and the compatibility of R_1 and R_2 groups with the binding requirements of the pocket situated in the region at the interface of the two mPGES-1 subunits.



Scheme 4. 1 Chemical structures of compounds **140**-**165** utilized for molecular docking Studies

In particular, it was gradually increased the length, size, and hydro- and lipophilicity of R_1 and R_2 with the aim to optimize their chemophysical properties. To identify the key structural features necessary for mPGES-1

inhibition, it was performed an *in silico* screening by molecular docking using AutoDock 3.0.5 software[55] of a small set of molecules. For reported docking calculation, it was used the MGST-1 structure solved by Holm et al. in 2006[283] in which significant amino acid conservation in comparison to mPGES-1[284] (38% of homology sequence) can be recognized. Recently, the structure based drug design targeting mPGES-1 was facilitated by the work of Hamza et al.,[279,280] who have described the PGH_2 binding to the mPGES-1-GSH complex. More precisely, as also demonstrated by site direct mutagenesis, the natural ligand at the interface of each mPGES-1 monomer establishes a strong salt bridge between its carboxylate group and the highly conserved Arg110 in the MAPEG family and interacts with Arg70, Asn74, Arg73, Glu77, Tyr117, Leu121, Arg122, Arg126, Thr129, Arg110, His72, Lys26, Leu69, and Ile125. Taking into account the considerations above, it was referred to the sequence alignment of these two MAPEG super family members for the rationalization of the small molecules binding mode (Table 4.1).[279]

Table 4. 1 List of the corresponding amino acids present both in mPGES-1 and MGST-1 catalytic sites.

mPGES-1	**MGST-1**
Arg110	Arg113
Arg70	Arg73
Asn74	Asn77
Tyr117	Tyr120
Leu121	Leu124
Arg126	Arg129
His72	His75
Lys26	Lys25
Glu66	Glu69

His113	His116
Arg73	Leu76
Arg122	Pro125
Thr129	Ala132
Leu69	Arg72
Ile125	Asn128
Thr78	Asn81
Tyr130	Phe133

The data shown in Figure 4.2 indicate the best calculated affinities for compounds presenting one H-bond acceptor group and a lipophilic substituent of adequate dimensions.[170d] For the sake of simplicity, it was report the most promising candidates derived from the *in silico* screening, **143** and **20** (Figure 4.3), to trace the features of new potential anti-inflammatory drugs. From the analysis of the results, both compounds disclosed a similar binding mode at the interface of the target monomer. The reported proposed poses are in agreement with the model reported by Hamza et al.[279] In fact, the compounds interact with residues considered critical for PGH_2 binding, such as the hydrogen bonds with the carboxy group in **143** and **159** with the highly conserved Arg113 in MGST-1 ($Arg110_{mPGES-1}$), guaranteeing, at least in theory, the enzyme binding specificity, as well as van der Waals and other interactions with residues of the active site—the cation-π interactionwith Lys67, Arg72, and Arg196 for 4 and with Lys67 and Arg196 for 20. The above *in silico* results suggested the synthesis of the molecules **143**, **145**-**147**, **150**, **153**, **154**, **156**, **157**, **159**-**163**, and **165**, all within the lowest free energy of binding and a good ligand efficiency ($E_{binding}$ < 9 kcal/mol and ΔG/NHA deeper than -0.24 kcal/mol, Figure 4.2) as the starting point for obtaining preliminary experimental results.

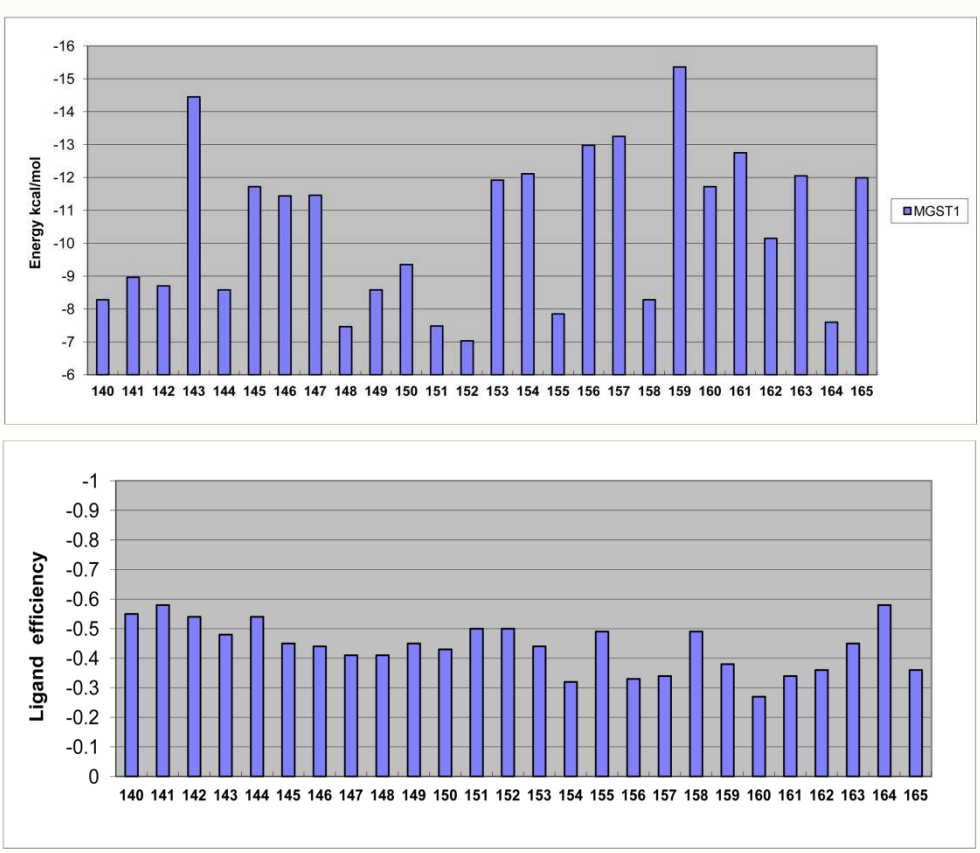


Figure 4. 2 Calculated affinities and ligand efficiency of compounds **140**-**165** for MGST-1.

The evaluation of the bioactivity of this small set of compounds might be helpful for the comprehension of the key features of new mPGES-1 triazole based inhibitors.

For the synthesis of compounds **143**, **145-147**, **150**, **153**, **154**, **156**, **157**, **159-163**, and **165**, synthetic procedure based on the copper-catalyzed 3 + 2 Huisgen's reaction (click chemistry) was used, where the triazoles intermediates were generated through the condensation between the appropriate terminal alkynes and the azides.

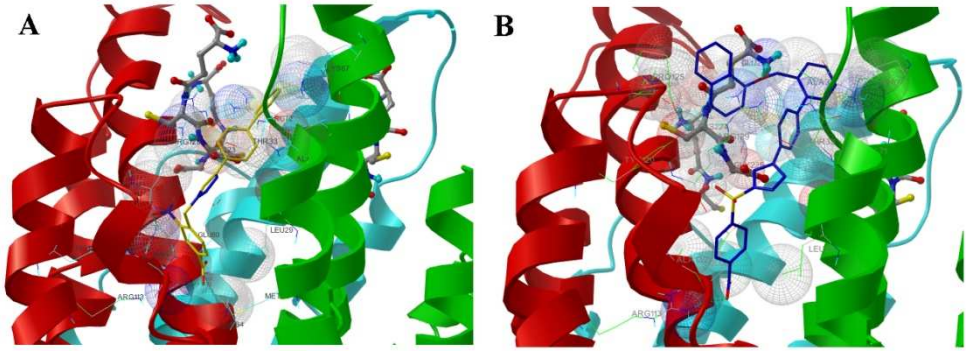


Figure 4. 3 Three-dimensional model of interactions of **143** (A) and **159** (B) with the MGST-1 binding site. The protein is represented by secondary structure, by CPK, and by lines colored by atom type (C, gray; polar H, sky blue; N, blue; and O, red). Compound **143** (A) is depicted by sticks and balls (by atom type: C, yellow; O, red; and N, blue). Compound **159** (B) is depicted by sticks and balls (by atom type: C, blue; O, red; N, dark blue; and S, yellow).

To assess the ability of the selected compounds **143**, **145-147**, **150**, **153**, **154**, **156**, **157**, **159-163**, and **165**, to interfere with the activity of mPGES-1, a cell-free assay using the microsomal fractions of interleukin-1β (IL-1β)-stimulated A549 cells (as source for mPGES-1) was applied. As shown in Table 4.2, compounds **143**, **146**, **159**, and **162** significantly inhibited mPGES-1 activity, whereas all other derivatives were not significantly active at a concentration of 30 μM. Interestingly, these data confirm the results from the docking studies favoring **143** and **159** as mPGES-1 inhibitors. As reported above, mPGES-1 inhibitors often interact also with other enzymes within the arachidonic acid cascade, such as 5-LO or FLAP and the dual suppression of PGE_2 and LT formation might reduce side effects.[243] Thus, it was performed a further analysis of the test compounds (30 μM, each) for inhibition of 5-LO activity in a cell-free assay using isolated human recombinant 5-LO as the enzyme source.

Table 4. 2 Effect of test compounds on the activity of mPGES-1[a]

comp.	mPGES-1 activity remaining activity IC $_{50}$ [µM at 30 µM	
143	3.2	12.0% ± 3.7**
145	> 30	89.0% ± 2.9
146	> 30	60.1% ± 4.3**
147	> 30	78.2% ± 12.8
150	> 30	76.1% ± 7.9
153	> 30	73.6% ± 8.0
154	> 30	96.8% ± 0.5
156	> 30	91.9% ± 8.1
158	> 30	87.9% ± 8.4
159	> 30	59.8% ± 7.2**
160	> 30	85.0% ± 9.3
161	> 30	90.0% ± 6.8
162	> 30	72.1% ± 4.2*
163	> 30	98.2% ± 7.3
165	> 30	88.7% ± 2.7
[a]Data are given as mean +/- S.E., n=4-6. *p < 0.05, **p < 0.01		

Intriguingly, among the test compounds, **159** was the most active derivative with IC_{50} = 0.8 µM, followed by **160**, **143**, **156**, and **146** (Table 4.2, IC_{50} = 4.1, 6.7, 8.8, and 27 µM, respectively), which all inhibited 5-LO activity in a concentration-dependent manner. 5-LO inhibition was reversible, as demonstrated by wash-out experiments (not shown). Also, **145**, **153**, **162**, and **165** significantly inhibited 5-LO at a concentration of 30 µM, but the magnitude of inhibition did not exceed 50% (Table 4.3), and thus, IC_{50} values could not be determined. Because FLAP inhibitors do not inhibit 5-LO activity in cell-free assays but only LT formation in intact cells,[285] it was assessed a

potential inhibitory effect on FLAP in human neutrophils activated by ionophore A23187. Compound 43 (IC_{50} for FLAP in neutrophils approximately 70 nM)51 served as the control, and DMSO (0.3%, v/v) was used as the vehicle control. Compounds **143**, **145**, **146**, **150**, **159**, **160**, **163**, and **164** reduced 5-LO product formation at 30 μM by more than 50% in a concentration-related manner with IC_{50} values in the range of 0.4-9.3 μM (Table 4.3). For **143** and **159**, the IC50 values were determined at 8.8 and 0.6 μM, respectively, which fits well with the activities in cell-free 5-LO assays, and also, **160** was similarly efficient (IC_{50} = 2.8 μM) as for isolated 5-LO.

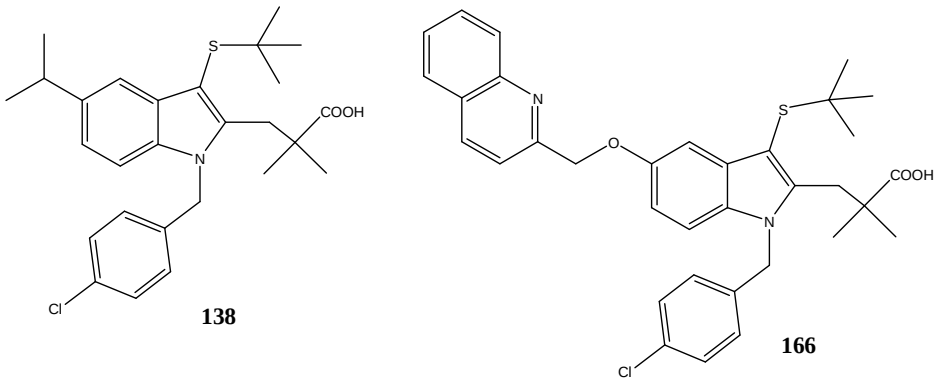
Table 4. 3 Effect of Test Compounds on the Activity of 5-LO in Cell-Free and Cell-Based (Intact Neutrophils) Assays[a]

comp.	**5-LO activity; cell-free**		**5-LO activity; intact neutrophils**	
	IC_{50} [μM] activity at 30 μM (%)		**IC_{50} [μM] remaining activity at 30 μM (%)**	
143	6.7	20.0% ± 0.9**	9.2	20.1% ± 11.0**
145	> 30	62.3% ± 1.4**	0.9	34.8% ± 7.0** [b)]
146	27	48.8% ± 0.4**	0.4	1.1% ± 0.3** [b)]
147	> 30	82.9% ± 0.9	> 30	85.7% ± 3.5
150	> 30	80.8% ± 5.3	9.3	14.6% ± 5.2**
153	> 30	58.4% ± 12.7*	> 30	70.2% ± 8.5
154	> 30	77.4% ± 0.9	> 30	79.9% ± 10.7
156	8.8	10.1% ± 4.6**	> 30	52.7% ± 15.0*
157	> 30	82.5% ± 4.4	> 30	92.1% ± 8.1
159	0.8	13.6% ± 2.8** [b)]	0.6	3.5% ± 2.5** [b)]
160	4.1	5.1% ± 0.8**	2.8	21.2% ± 4.3** [b)]
161	> 30	78.4% ± 10.3	> 30	84.5% ± 3.4
162	> 30	57.3% ± 1.2**	6.0	17.3% ± 2.5** [b)]
163	> 30	60.7% ± 10.0	1.7	22.0% ± 2.6** [b)]
165	> 30	59.2% ± 6.4*	> 30	70.1% ± 9.0

[a]Data are given as means (SEs, n = 4-6. *p < 0.05, and **p < 0.01. [b] Remaining activity at 10 μM

All in all, on the basis of the outcomes of the biological activity data, 4 is the most efficient inhibitor of mPGES-1, **146** might act as a FLAP inhibitor, while **159** might be a potent direct 5-LO inhibitor, besides moderate inhibition of mPGES-1. Hence, we aimed to rationalize the results through molecular modeling studies. As preliminarily remarked, it should be put in evidence that compounds **143** and **146**, inhibiting the two MAPEG family members, showed quite similar chemical features; on the contrary, the more encumbering ligand **159** seems to target no structurally related MAPEG enzymes.

For the reported calculations, it was used the 3D structure of FLAP in complex with the inhibitor **166** (MK-591, Scheme 4.2)[286] solved by Ferguson et al.[287] in 2007 [protein data bank (PDB) ID code 2Q7M]. Because of the lack of crystal structure information on 5-LO, we used a 15-LO[288] (PDB ID code 1LOX) enzyme, presenting the highest sequence similarity (38% identity with human 5-LO) among the dioxygenase family (8-, 9-, 11-, and 12-LO). Taking into account the considerations reported above for the MGST-1 enzyme, also in the case of FLAP, the binding specificity was conferred by the H-bond with the Lys116.



Scheme 4. 2 Chemical Structures of **138** and **166**

In the reported proposed pose, **146** (Figure 4.4) not only interacts with the fundamental amino acids but also adopts the equivalent spatial disposition of the cocrystallized inhibitor,[288] maintaining the same interactions network. Moreover, the phenyl group in R_1 forms a π-π stacking with Phe25.

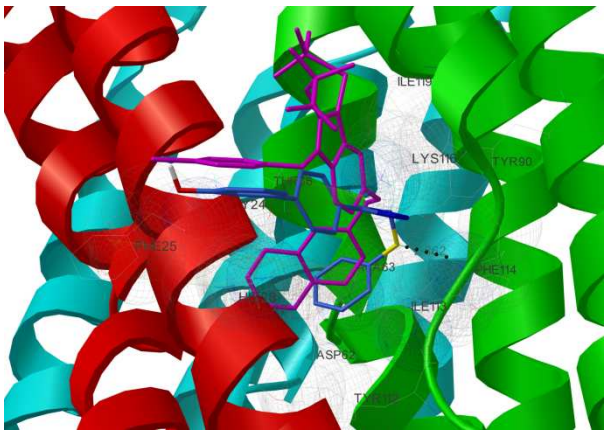


Figure 4. 4 Three-dimensional model of interactions between **146** and FLAP. The protein is represented by secondary structure, by CPK, and by lines colored by atom type (by atom type: C, gray; polar H, sky blue; N, blue; O, red; and S, yellow). Compound 7 is depicted by sticks and balls (by atom type: C, sky blue; O, red; N, dark blue; S, yellow; and H, white). Compound 44 is depicted by sticks and balls (C, O, N, S, and Cl dark pink). The figure highlights similar interactions for both 7 and 44with arachidonic acid-binding site.

Three different classes of inhibitors can be generally considered for 5-LO:[289] (1) antioxidant agents interfering with the redox catalytic cycle of the enzyme, (2) iron-chelating agents, and (3) nonredox type inhibitors, which compete with arachidonic acid for the binding to the enzyme.[285] In the described docking studies, it was supposed that **159** acts as nonredox type LO inhibitor. As described for mPGES-1, the rationalization of the 5-LO binding mode was obtained considering the fundamental amino acids in the active site

of the enzyme as reported by Charlier et al.[290] (see Table 4.4), taking into consideration the specific polar interaction of the carboxylate moiety of arachidonic acid with $Lys409_{5\text{-}LO}$ ($Arg403_{15\text{-}LO}$).

Table 4. 4 List of the corresponding amino acids present both in 5 and 15-LO catalytic sites.

5-LO	**15-LO**
Ala424	Ile418
Asn425	Met419
Ala603	Ile593
Arg411	Arg405
Gln363	Glu357
Gln557	Gln548
Ile406	Ile400
Leu368	Leu362
Leu420	Ile414
Leu607	Leu597
Lys409	Arg403
Leu414	Leu408
Phe177	Phe175
Phe421	Phe415
Phe359	Phe353
Ser608	Gly598
Trp599	Leu589
His372	His366
His367	His361

For compound **159**, it was obtained two different conformation families, accounting for two independent high affinity binding modes (Figure 4.4 A,B). In both of the conformations, the specific interaction with Arg403 was maintained.

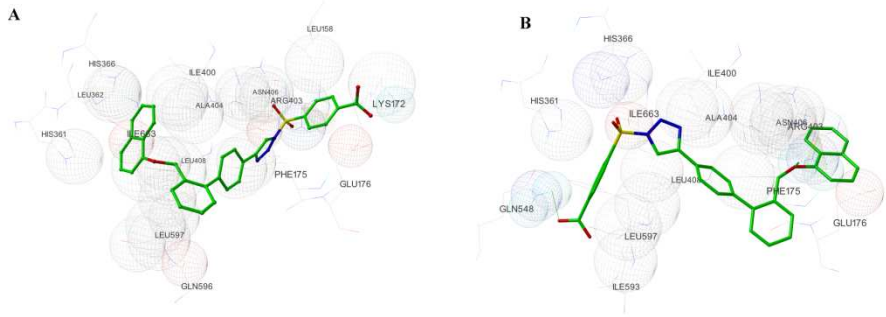


Figure 4. 5 Three-dimensional model of interactions of **159** and 15-lipoxygenase enzyme active site. The protein is represented by CPK and by lines colored by atom type (by atom type: C, gray; polar H, sky blue; N, blue; O, red; and S, yellow). Two different conformations (A and B) of the complex are depicted. Compound **159** is depicted by sticks and balls (by atom type: C, green; O, red; N, dark blue; and S, yellow).

In particular, in the first conformation (Figure 4.4A), the phenyl group in R_1 shows a cation-π interaction, while in the second conformation (Figure 4.4B), the same cation-π interaction with Arg403 was exerted by the naphtyl group in R_2 present in the alternative conformation. In the latter case, the oxygen atom in R_2 forms an additional H-bond with the positively charged (Arg403) residue. Even if R_1 and R_2 are located on the opposite sites of the target pocket, and the other interactions with the receptor counterpart remain unmodified and are in accordance with the structural requirements indicated by Charlier et al.,[290] that is, two hydrophobic groups, an aromatic ring, and two hydrogen bond acceptors.

In summary, it was applied a rapid *in silico* screening on a small set of triazole derivates for directing the click chemistry synthesis of the most promising mPGES-1 inhibitor **143**. In light of the good qualitative accordance between the results from the biological assays and the prediction of the molecular docking calculations, a satisfactory explanation of the putative binding mode for the new triazole based compounds was provided. Biological assays disclosed three different benchmark compounds **143**, **146** and **159** as inhibitor of mPGES-1, FLAP and 5-LO, respectively.

4.2.1 Computational Details

Molecular docking calculations were performed on a two dual-core Intel Xeon 3.4 GHz, using Autodock 3.0.5 software.[55] For all the docking studies on MGST1, FLAP and 15-LO a grid box size of 98 x 98 x 126, 70 x 126 x 78, and 68 x 74 x 96, with spacing of 0.375 Å between the grid points, and centered at -43.667 (x), 33.493 (y), 2.656 (z), 30.35 (x), 21.836 (y), 55.454 (z), and -29.186 (x), 151.802 (y), 51.228 (z) was used respectively. For all the docked structures, all bonds were treated as active torsional bonds. In order to achieve a representative conformational space during the docking calculations, from four to six calculations consisting of 256 runs were performed, obtaining 1024/1536 structures. The Lamarckian genetic algorithm was used for dockings. An initial population of 600 randomly placed individuals, a maximum number of 5.0×10^6 energy evaluations, and a maximum number of 6.0×10^6 generations were taken into account. A mutation rate of 0.02 and a crossover rate of 0.8 were used. Results differing by less than 2.5/3.5 Å in positional root-mean-square deviation (RMSD) were clustered together and

represented by the result with free energy of binding ($E_{binding}$). All the 3D models were depicted using the Python software.[165]

4.3 Design and Synthesis of a Second series of Triazole-based compounds as potent dual mPGES-1 and 5-Lipoxygenase inhibitors.

Continuing the studies on small molecules able to block mPGES-1 activity described in a previous paragraph, some interesting molecules able to inhibit mPGES-1 as well as other key enzymes within the arachidonic acid cascade such as 5-lipoxygenase (5-LO) and 5-LO-activating protein (FLAP) were identified. The full coherence found between the biological results and the predictions of molecular docking calculations provided a satisfactory binding mode model of the triazoles reported in a previous paragraph with their respective biological targets (i.e., mPGES-1, 5-LO, and FLAP), gaining more information to envisage further structural optimizations of the basic template in accordance with the catalytic pocket requirements of the enzyme. In more detail, the previous investigations disclosed a basic structure **143**,[269] able to efficiently interact with key aminoacid residues in the catalytic site, which actually needed to be further optimized especially concerning the substitution pattern on the biphenyl moiety in order to provide a better binding affinity with the target enzyme.

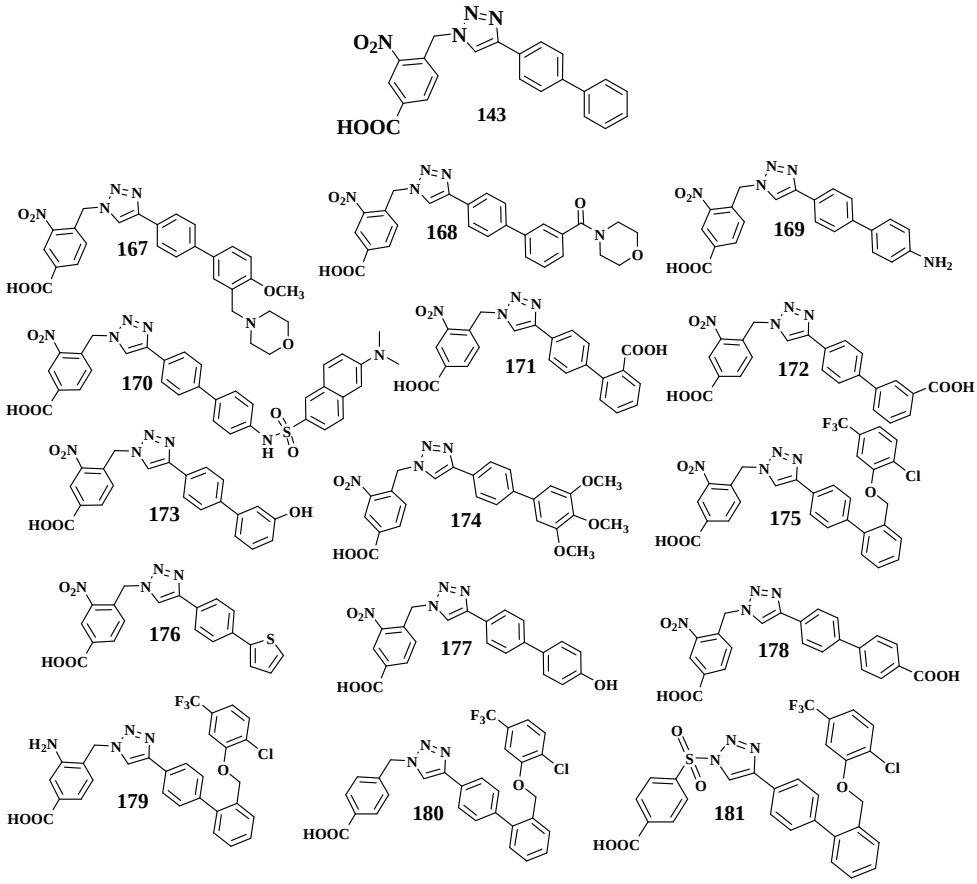
In line with these considerations, it was decided to leave the 2-nitro-4-carboxybenzyl moiety (Scheme 4.3) unchanged and to variously decorate the right hand portion of the molecule, coming to the identification of compound 11, which retained mPGES-1 inhibition ability, and compound **175**, emerging as potent dual inhibitor of mPGES-1 and 5-LO enzymes. Building on these results, it was decided to further refine the structure of compound **175**, leaving unmodified the right hand portion of the molecule and performing some minor modifications on the left hand moiety, including the reduction of the nitro

group (compound **179**), the full elimination of the ortho-substituent on the phenyl ring (compound **180**) and finally the insertion of a sulfonamide function between the phenyl and the triazole ring (compound **181**) by analogy with the structural backbone showed by a triazole-based compound previously identified by us as potent dual inhibitor of mPGES-1 and 5-LO enzymes.[269]

On the basis of the encouraging results regarding the rationalization of the inflammation processes related to natural products[291,292,293,294,295] and the design of new synthetic molecules,[269,296] here we report the design and the theoretical evaluation of a new significant set of molecules obtained by revisiting the lead compound **143** (See Scheme 4.3).[297] In particular, the scaffold of **143** was substituted with various groups of different size and hydrophobicity with the aim to improve potency and to trace a more accurate SAR profile of this kind of potential anti-inflammatory drugs.

As suggested from our previous work,[269] the 2-nitro-4-carboxybenzyl moiety remarkably influences the activity of the triazole molecules, so this moiety was not altered for the molecules **167-178** (Scheme 4.3) combining it with small hydrophilic groups (COOH, OH, NH_2) at different positions (**169**, **171-173**, **177-178**, Scheme 4.3), or with hydrophobic groups with increasing size (**168-169**, **170**, **174-176** Scheme 4.3) to understand their potential influence on biological profiles.

Owing to the lack of the experimental structure of the mPGES-1 open active form, for the reported docking calculations by Autodock4.2 software the microsomal glutathione transferase 1 (MGST-1)[283] was used as model receptor; [269] which shows 38% sequence identity with the human mPGES-1.[284]



Scheme 4. 3 Structures of lead compound **143** and its derivatives.

The obtained results point out two different docking poses for these triazole-based potential mPGES-1 inhibitors: (a) the first one includes molecules with ortho- and/or meta-substituted biphenyl groups (**168**, **171**-**173**, and **175**, Figure 4.6A), or with a smaller aromatic ring (**176**) as arms at position 4; (b) the second family relates to compounds containing the biphenyl portion with at least one substituent in the para position (**168**, **169**-**170**, **174**, **177**-**178**, Figure 4.6B). However, all the designed molecules accommodate in the ligand pocket situated in the region at the interface of the two mPGES-1 subunits (Figure 4.6) interacting with the highly conserved Arg113 in MGST-

1 (Arg110 mPGES-1), which guarantees at least in theory the mPGES-1 binding specificity.

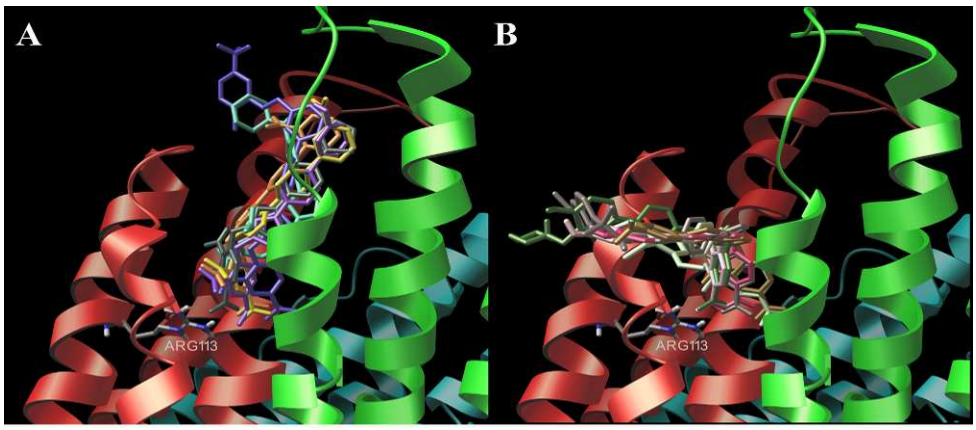


Figure 4. 6 Two different spatial arrangements of triazole-based compounds **167**-**178** in the binding site of MGST-1. (A) Superposition of compounds **168** (emerald), **171** yellow, **172** (orange), **173** (purple), **175** (blue) and **176** (dark green). (B) Superposition of **167** (green), **169** (pink), **170** (grey), **174** (light green), **177** (light pink) and **178** (light brown).

On this basis, the substitutions at ortho, meta and para of the biphenyl arm with small hydrophilc groups cause a different pattern of hydrogen bonds with MGST-1 among **169**, **171**-**173**, **177**-**178** (See Figure 4.7A). In particular, even though **171** and **172** have similar docking poses and both molecules are able to establish a π-cation interaction with Arg129, the ortho COOH of **6** forms a hydrogen bond with Arg129 while the meta COOH of **7** interacts with Arg37 (Figure 4.7A). On the other hand, the ortho carboxybiphenyl arm of **178** interacts with amino acids located at the bottom part of the ligand binding pocket establishing a π-π interaction with the Phe133, and moreover, its 2-nitro-4-carboxybenzyl moiety interacts with Asn81 and Lys25 (Figure 4.7A). The same hydrogen bond pattern is also observed for the molecules **169**, **173**

and **177** (See Figure 4.7B), where the OH (**177**) and/or NH_2 (**169**) group at para-position of biphenyl arm do not allow the hydrogen bonds with Arg37 and/or Arg129, but allow the π-π interaction with Phe133 (See Figure 4.7B); whereas for **173**, as like for **172**, the hydrogen bond with Arg37 and π-cation interactions with Lys67 and Arg129 are observed.

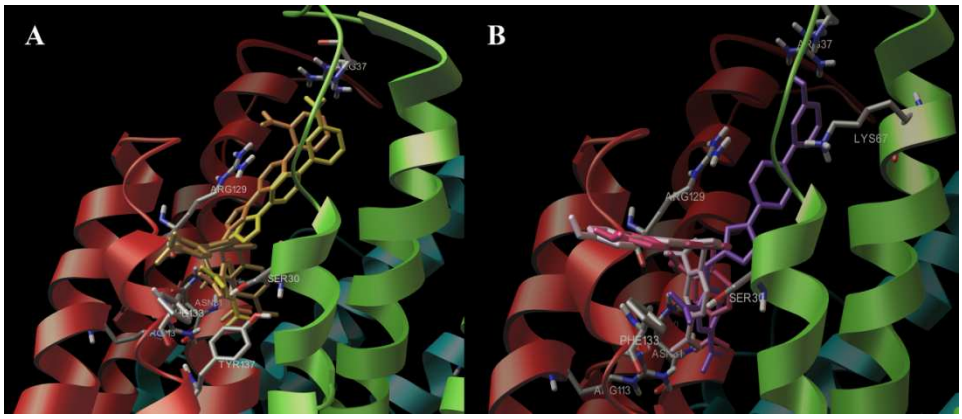


Figure 4. 7 (A) Three dimensional model of the different hydrogen bond pattern of **171** (yellow), **172** (orange) and **178** (light brown) with MGST-1 binding site. (B) Three dimensional model of the different hydrogen bond pattern of **173** (purple), **169** (pink) and **177** (light pink) with MGST-1 binding site. In both the figures the crucial aminoacids of MGST-1 receptor are depicted by stick and balls (by atom type: C, gray; O, red; N, dark blue; and S, yellow).

Regarding the molecules with bulky groups (**168-169**, **170**, **174**, **175-176**, Scheme 4.3), depicted in Figure 4.8, the different steric hindrance of the arms influences the positioning in the ligand binding site. In particular, the molecules with para (**170**) groups, or combined with meta (**167**) and/or ortho (**174**) substituents at biphenyl portion are in close contact with Gly130, Phe134 and establish π-π interactions with Phe133. On the other hand, even though **176** and compounds with bulkier arms (i.e. **168** and **175**) have similar

docking poses, the thiophene ring of **176** allows peculiar van-der-Waals contacts with Ala32 and Val68, and a further π-cation interaction with Lys67; while **175** with respect to **168** establishes additional π-cation interactions with Arg129 and Arg37 and hydrophobic contact with Pro125.

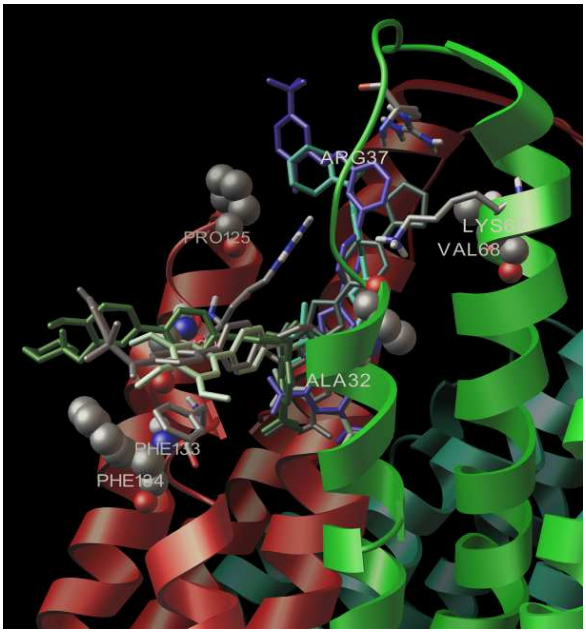


Figure 4. 8 Three dimensional models of the different hydrogen bond pattern and peculiar hydrophobic interactions of **167** (green), **168** (emerald)**, 170** (grey), **174** (light green), **175** (blue), and **176** (dark green) with the MGST-1 binding site. The crucial aminoacids of MGST-1 are depicted by stick and balls and CPK (by atom type: C, gray; O, red; N, dark blue; and S, yellow).

From the above considerations, it was possible to identify the differences between these two families of molecules where the substituents position on the arms influences the pattern of hydrophobic interactions and/or the hydrogen bonding. For these reasons, it was decided to undertake the synthesis of the designed molecules (all within the lowest free energy of binding ($E_{binding}$) < -7

kcal/mol) as the starting point for obtaining preliminary experimental results based on the evaluation of the bioactivity of this small set of compounds with the aim of a deeper understanding of the key features of new mPGES-1 triazole-based inhibitors. On the basis of the biological results, it was have decided to further improve and refine the structure of the most active molecule **175** that was recognized as new lead compound (see below) for the design of triazole based mPGES-1 inhibitors. In particular, with the aim to improve the biological activity of new lead compound **175**, it was modify its 2-nitro-4-carboxybenzyl moiety (**179**-**181**) taking into account the reported previous considerations.[269] As be seen from the Figure 4.8, the molecule **180** assumes the same docking pose with respect to **175**, where the elimination of the NO_2 group only influence the strength of the H-bonding with the Arg113, while the π-cation interactions with Arg129 and Arg37, and most of the hydrophobic interactions are maintained. On the other hand, the chemical reduction of the nitro group to NH_2 for **179**, and the insertion of a sulfonamide function between the phenyl and the triazole ring at **143** yielding **181** causes a different positioning of the compounds compared to **175**, and a shifting of the bulky arm at C-4 of the triazole ring in the MGST-1 ligand pocket. In particular, the NH_2 group of **179** establishes two additional hydrogen bonds acting as H-bond donor with the carbonyl oxygen of Gly223 and $N^{\delta 1}$ of His75, making also possible the hydrogen bond between the triazole ring with the hydroxyl of Thr33, and hydrophobic interactions of the halogenated benzene with Gly130, Pro127 and Gln126. On the other hand, the sulfonamide function of **181** only forms a hydrogen bond with the OH moiety of Thr33.

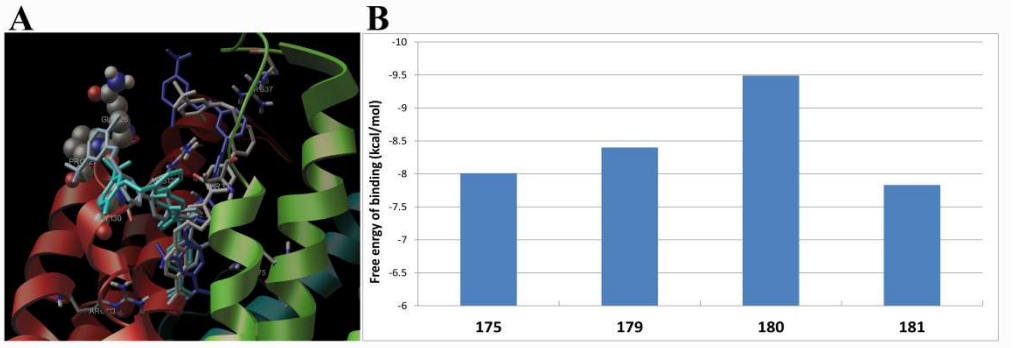


Figure 4. 9 (A) Three dimensional models of the different hydrogen bond pattern and peculiar hydrophobic interactions of **174** (blue), **179** (light blue)**, 180** (white), and **181** (cyan) with MGST-1 binding site. (B) Calculated free energy of binding expressed as kcal/mol of compounds **174** and **179-181** for MGST-1. The crucial amino acids of MGST-1 receptor are depicted by stick and balls and CPK (colored by atom type: C, gray; O, red; N, dark blue; and S, yellow).

As shown in Figure 4.9B, only **179** and **180** showed an improvement of the calculated affinity compared to the lead compound **175**, suggesting the critical role of the hydrogen bonding and hydrophobic contacts for the potential antagonist activity; as a consequence, these compounds were synthesized and their biological activity was assessed to prove the discussed theoretical findings.

Concerning the synthesis of analogues **167-181**, the retro-synthetic approach suggest to obtain the triazole ring through the copper-catalyzed 3+2 Huisgen's cycloaddition between the appropriate terminal alkynes and azides,[298] while to generate the differently decorated biphenyl systems we relied on the Suzuki cross-coupling reaction.[299]

In order to assessed the ability of the test compounds to interfere with mPGES-1 activity, a well-established cell-free assay[257,300] was applied where microsomes of interleukin-1β-stimulated A549 cells were used as source for

human mPGES-1 enzyme and PGH_2 (20 µM) as substrate for mPGES-1. The mPGES-1 inhibitor MK-886 (IC_{50} = 2.4 µM) was used a reference compound[248] that inhibited the enzymatic reaction by 82±6.4% at a concentration of 10 µM (not shown). The results of the assessment of the mPGES-1-inhibitory activities of compounds **143, 167-181** are given in Table 4.5.

Table 4. 5 Inhibition of mPGES-1 and 5-LO by the test compounds.

Comp.	**mPGES-1** (IC_{50} [µM]; remaining activity at 10 µM)	**5-LO** (IC_{50} [µM]; remaining activity at 10 µM)	
		cell-free	cell-based
143	3.2	6.7	9.2
167	> 10 (87.4 ± 9.1%)	> 10 (84.9 ± 11.1%)	> 10 (111.1 ± 8.4%)
168	> 10 (88.9 ± 4.2%)	> 10 (62.7 ± 12.2%)	> 10 (125.0 ± 12.0%)
169	> 10 (65.5 ± 5.1%)	> 10 (63.5 ± 11.3%)	> 10 (100.1 ± 5.3%)
170	> 10 (72.5 ± 4.3%)	6.1	> 10 (96.5 ± 7.3%)
171	> 10 (94.1 ± 8.5%)	> 10 (75.3 ± 13.1%)	> 10 (65.9 ± 4.7%)

172	> 10 (82.1 ± 2.5%)	> 10 (74.7 ± 9.6%)	> 10 (123.8 ± 5.5%)
173	> 10 (55.6 ± 0.8%)	> 10 (80.2 ± 2.7%)	> 10 (113.9 ± 10.1%)
174	> 10 (54.3 ± 1.5%)	> 10 (64.3 ± 11.8%)	> 10 (90.2 ± 9.3%)
175	1.2	2.0	5.3
176	5.5	> 10 (52.2 ± 9.6%)	> 10 (89.7 ± 2.8%)
177	> 10 (74.2 ± 1.8%)	6.8	> 10 (108.3 ± 1.1%)
178	> 10 (93.0 ± 21.2%)	> 5 (62.7 ± 12.7%) [5 µM]	> 5 (90.9 ± 8.6%) [5 µM]
179	0.68	0.9	2.1
180	1.0	1.47	2.0
181	> 10 (78.6 ± 9.6%)	1.96	1.2

For compound **143,** carrying a biphenyl, an IC_{50} value of 3.2 µM was obtained, which is in agreement with previous reported finding.[269] Bioisosteric replacement of the 4-phenyl within the biphenyl of **143** by thiophene, yielding

176, was essentially tolerated (IC_{50} = 5.5 µM). However, substitution of the 4-phenyl in **167**-**168**, **170**-**174**, **178**-**177**, show clearly detrimental with IC_{50} values > 10 µM. In contrast, when the bulky, halogenated phenylether was inserted in 2´-position (compound **175**), a significant improvement in the potency was obtained, with an IC_{50} value = 1.2 µM.

Next, the influence of the 2-nitro group within the 4-carboxybenzyl moiety on mPGES-1 inhibition was investigated. Omission of the 2-nitro moiety (resulting in compound **170**) slightly improved mPGES-1 inhibition (IC_{50} = 1.0 µM). Replacement of the 2-nitro function by an amino moiety further increased the efficiency and the respective compound **179** exhibited an IC_{50} value of 0.68 µM, representing the most potent mPGES-1 inhibitor within this study. Exchange of the methylene bridge next to the triazole of **180** by a sulfone moiety (compound **181**) again abolished the gain of potency in agreement with reduced free energy of mPGES-1 binding.

Therefore, and based on the previous observation that compound **143** also efficiently inhibits 5-LO,[269] we assessed the potential of the above-mentioned compounds for inhibition of human 5-LO in well-established cell-free and a cell-based models.[269,300,301] Interestingly, inhibition of 5-LO in both the cell-free and the cell-based test systems essentially correlated with the interference of the compounds with mPGES-1. Thus, compounds **167**-**174**, **177** and **178** that all failed or hardly inhibited mPGES-1 were also modest 5-LO inhibitors (IC_{50} > 10 µM), except the sulfonamide **170** and the 4´-OH biphenyl analogue **177** of compound **143** that suppressed 5-LO in the cell-free (but not in the intact cell) assay with IC_{50} = 6.1 and 6.8 µM, respectively. On the other hand, the potent mPGES-1 inhibitors **175**, **179** and **180** efficiently inhibited 5-LO activity with IC_{50} values in the range of 0.9 to 2 µM in the cell-free and 2.0 to 5.3 µM in the cell-based assay, respectively. One exception was the sulfone

181, which was inactive for mPGES-1 but a potent 5-LO inhibitor with IC_{50} = 1.9 and 1.2 μM in cell-free and cell-based systems, respectively. Obviously, the sulfone or sulfonamide moieties (in compounds **181** and **170**, respectively) govern 5-LO inhibition but are detrimental for interference with mPGES-1. Taken together, compound **179** turned out to be the most potent dual mPGES-1/5-LO inhibitor out of these series with 4- and 7-fold lower IC_{50} values *versus* the parental lead compound **143**.

To rationalize the inhibitory activity of **170**, **175**, **177** and **179**-**181** on 5-LO by means of molecular docking, it was used the crystal structure of stable 5-LO recently reported by Newcomer and co-workers.[302] It was considered the molecules as possible nonredox-type 5-LO inhibitors[285,289] since it does not show any features of a redox-active agent or properties of an iron-ligand. Because the 5-LO was crystallized without ligand,[302] the model interactions described by Wouters *et al.*[290] was used as reference which was already successfully used for rationalize the binding mode of nonredox-type 5-LO inhibitors.[269,303,304]

In the reported models, in fact, all the molecules interact with a binding site formed by Arg411, Ile406, Phe177, Lys409, Tyr181, Leu607, Leu414, Leu420, Trp599, Asn425, Gln363, Phe421, and Leu368, however, the different size of the second arms influence their binding mode on the protein molecular surface. As depicted in Figure 4.10, for all six potential inhibitors, the -COOH group and the triazole ring of 2-nitro-4-carboxybenzyl form a hydrogen bond with Arg596 and a π-cation interaction with Lys243, respectively. On the other hand, the $-NH_2$ group of **179** allows a further hydrogen bond with the carbonyl oxygen of Gly419 and the carboxylic group in the side chain of Asp422, while the oxygen of the halogenated arms of **175, 179**-**181** act as H-bond acceptors with the $NH^{\varepsilon 2}$ of Gln413, and moreover, they

are able to establish electrostatic interactions with Lys409 and Lys173 by the chlorine atom. Furthermore, even though **170** shows a different binding mode compared to the other compounds, it establishes hydrophobic contacts with Arg411, Asn148, Cys159, Glu417, Ile415, Lys158, Trp147.

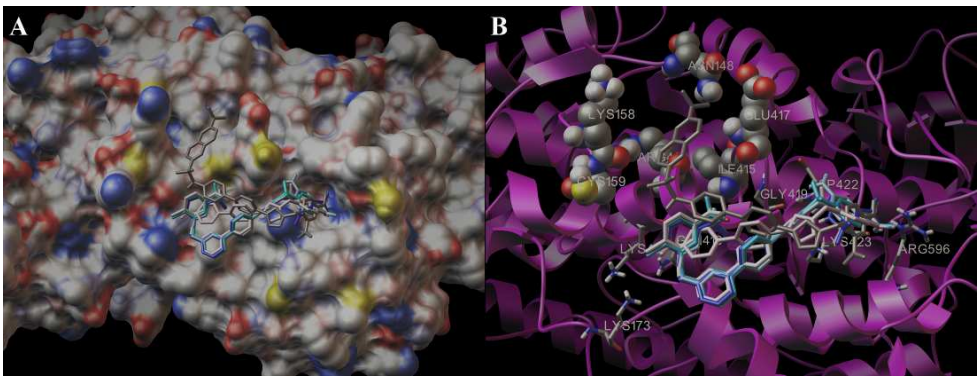


Figure 4. 10 Three dimensional model of the **170** (grey), **175** (blue), **177** (light pink) **179** (light blue)**, 180** (white), and **181** (cyan) with 5-LOX. The 5-LOX is represented by molecular surface (A) (colored by atom type: C, gray; O, red; N, dark blue; and S, yellow) and purple ribbon (B). The crucial aminoacids of 5-LOX are depicted by stick and balls and CPK (colored by atom type: C, gray; O, red; N, dark blue; and S, yellow).

All these interactions provide complexes with an increased predicted stability that is fully compatible with the results of the experimental biological assays, giving a good rationalization of the 5-LO inhibitory activity of **170**, **175, 177**, **179-181**.

4.3.1 Computational details

The molecular docking calculation were performed using Autodock 4.2 software[56] on 4x AMD 486 Opteron SixCore 2.4 GHz. For all the docking studies on MGST1 (pdb code: 2H8A)[283] a grid box size of 98 x 98 x 126 and

30 x 28 x 28, with spacing of 0.375 Å between the grid points, and centered at -43.667 (x), 33.493 (y) and 2.656 (z) was used covering the active site of the receptor. For all the docked structures, all bonds were treated as active torsional bonds. To achieve a representative conformational space during the docking studies and for taking into account the variable number of active torsions, 10 calculations consisting of 256 runs were performed, obtaining 2560 structures for each ligand. The Lamarckian genetic algorithm (LGA) was employed for docking calculations, choosing an initial population of 600 randomly placed individuals. The maximum number of energy evaluations and of generations was set up to 5 x 10^6 and to 6 x 10^6 respectively. Results differing by less than 3.5 Å in positional root-mean-square deviation (RMSD) were clustered together and represented by the result with the most favorable free energy of binding. On the other hand, for the rationalization of the binding mode **170**, **175, 177**, **179-181** with 5LO (pdb code: 3O8Y)[302] we used a Autodock Vina software choosing a grid box size of 30 x 28 x 28, with spacing of 1.000 Å between the grid points, and centered -11.146 (x), 66.57 (y) and 0.523 (z) covering the active site of the receptor. For the docking studies, we used an exhaustiveness of 8 with maximum energy difference of 3kcal/mol between the best binding mode and the worst one displayed. Illustrations of the 3D models were generated using the Python software.[165]

-CHAPTER 5-

Marine natural products as hsPLA$_2$ inhibitors

5.1. Human Group IIA Phospholipase A_2: an Overview

In this chapter, it will be described the rationalization by molecular docking of the marine natural products binding mode and mechanism of action, which act as ihibitors of human synovial Phospholipases A_2: Bolinaquinone (BLQ, *Dysidea* sp., **182** Scheme 5.1)[291] and Cladocoran A (CLD A, Mediterranean coral Cladocora cespitosa, **183** Scheme 5.2).[292]

Phospholipases A_2 (PLA_2) belong to the esterase enzymes group and acts in the release of the arachidonic acid from the membrane phospholipids[305], which in turn leads to the synthesis of key signaling molecules in the inflammation responses, such as prostaglandins, leucotrienes and platelet-activating factors through the so-called 'arachidonic acid cascade'.[306] PLA_2 enzymes are classified as cytosolic PLA_2 ($cPLA_2$), intracellular PLA_2 ($iPLA_2$) and secretory PLA_2 ($sPLA_2$) on the basis of their cellular location.[306] The $sPLA2_s$ are small secreted proteins (14–18 kDa) usually containing 5 to 8 disulfide bonds (Table 5.1).

Table 5. 1 Secreted phospholipases A_2 ($sPLA_2$)

Group Source Molecular mass	(kDa)	Disulfide	bonds
IA	Cobras and Kraits	13–15	7
IB	Human/porcine pancreas	13–15	7
IIA	Rattlesnakes; human synovial	13–15	7
IIB	Gaboon viper	13–15	6
IIC	Rat/murine testis	15	8
IID	Human/murine	14–15	7

	pancreas/spleen		
IIE	Human/murine brain/heart/uterus	14–15	7
IIF	Human/murine testis/embryo	16–17	6
III	Human/murine/lizard /bee	15–18 55 (human/murine)	8
V	Human/murine heart/lung/macrophage	14	6
IX	Snail venom (conodipine-M)	14	6
X	Human spleen/thymus/leukocyte	14	8
XIA	Green rice shoots (PLA2-I)	12.4	6
XIB	Green rice shoots (PLA2-II)	12.9	6
XII	Human/murine	19	7
XIII	Parvovirus	<10	0
XIV	Symbiotic fungus/ bacteria	13–19	2

This group of enzymes uses an active site histidine and absolutely requires μM levels of Ca^{2+} for catalysis. In close proximity to the catalytic histidine there is a conserved aspartate. The two amino acids form a His/Asp dyad (Figure 5.1), not showing a distinct preference for particular fatty acids. The sPLA$_{2s}$ are found in plants,[307] insects,[308] molluscs,[309] reptiles[310] and mammals.[311] Mammalian sPL$_{A2}$ enzymes possess a larger molecular mass of around 55 kDa due to N-terminal and C-terminal extensions.[312]

Among the calcium dependent sPLA$_2$'s, type IIA PLA$_2$ (sPLA$_2$-IIA), an isoform isolated from human synovial fluids, is well known to be implicated in the pathogenesis of inflammation;[313] high levels of this enzyme have been recovered in serum tissues affected by a variety of inflammatory pathologies, including rheumatoid arthritis, septic shock, psoriasis and asthma.[314]

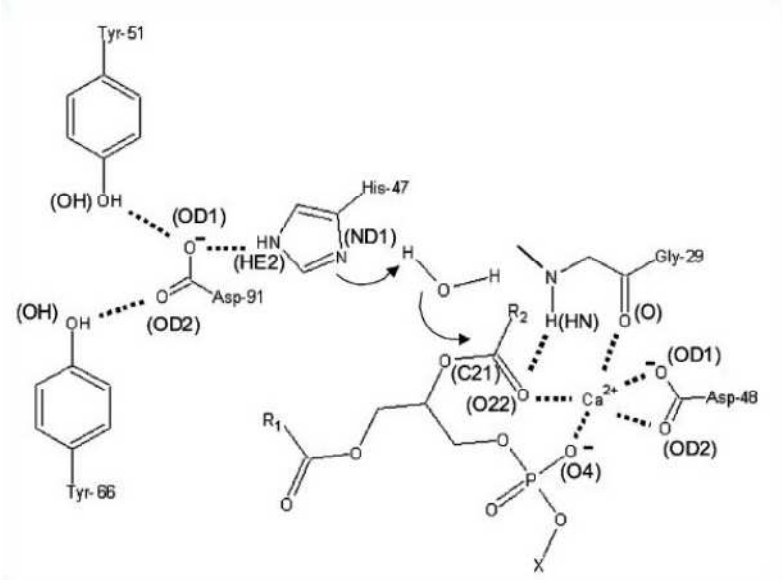


Figure 5. 1 Schematic rappresentation of catalytic mechenism of sPLA$_2$.

Moreover, the latest findings on the oncogenic action of PLA$_2$ in prostate cancer and its association with neurological disease offers further motivation for finding new modulators of its function.[315,316]

The marine environment is a rich source of molecules endowed with anti-PLA$_2$ properties, a number of which have shown potent and selective in vitro and in vivo anti-sPLA$_2$ activity.[317,318,319] One of the most widespread groups of constituents found in sessile organisms with a typical selective profile against secretory PLA$_2$ are the sesquiterpenes bearing a γ-hydroxybutenolide ring,

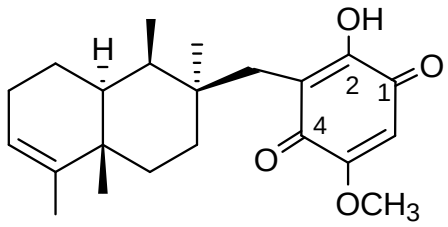
whose reference compound is manoalide,[319,320,321,322] Since then, many other related molecules have been isolated, such as seco-manoalide,[323] luffariellolides,[324] luffariellins,[325] luffolide,[326] cacospongionolides,[327] and recently petrosaspongiolides M-R,[328] all of which are capable of irreversible inhibition of PLA$_2$. Among these compounds, petrosaspongiolide M (PM, 2) has been the subject of detailed in vitro and in vivo pharmacological investigation, and the results suggested a double interaction of PM with two enzyme molecules, one of them covalently modified and the other contacting the inhibitor through its active site.[329]

Therefore, the discovery of selective sPLA$_2$ inhibitors and the assessment of their molecular mechanism of enzyme inactivation may have a relevant impact in the fields of chemical biology, medicinal chemistry and pharmacology.

5.2 The molecular mechanism of human group IIA phospholipase A2 inactivation by bolinaquinone

It does not come as a surprise that natural kingdoms, and particularly the marine environment, have proven to be a rich source of molecules exhibiting anti-PLA$_2$ properties; for instance, a number of sponge metabolites have shown potent activity in vitro and in vivo against sPLA$_2$, as opposed to their rather poor effects on cPLA$_2$.[317,319,330] Numerous marine natural products (MNPs) featuring a quinone and/or hydroxyquinone nucleus elicited sPLA$_2$ inactivation with good potency.[331,332a] Among MNPs,[332] Bolinaquinone (BLQ, Scheme 5.1),[333] one of the most active metabolites of this class, inactivates human synovial sPLA$_2$-IIA, bee venom PLA$_2$ and porcine pancreatic group I PLA$_2$, with IC$_{50}$ values in the mM range, demonstrating its higher activity with respect to other known PLA$_2$ natural inhibitors, such as manoalide[334,335,336] and petrosaspongiolide M.[337,321,322,329]

BLQ is a sesquiterpene metabolite isolated from the marine sponge Dysidea sp. in 1998 by Ireland and co-workers,[333] possessing a hydroxy-p-quinone moiety connected to a trans-decalin terpene unit in a rearranged drimane skeleton.[333,338] Its significant anti-inflammatory properties were recognized and thoroughly examined both in vitro and in vivo.[338] In particular, BLQ was tested as inhibitor of the production of mediators in acute and chronic inflammatory processes and as modulator of the oxidative stress parameters in 2,4,6-trinitrobenzenesulphonicacid (TNBS)-induced colitis in mice. BLQ has been shown its ability to significantly reduce nitrotyrosine immunodetection and colonic superoxide anion production and, furthermore, to decrease the extension of apoptosis, suggesting potential protective actions in intestinal inflammatory diseases.[339]



BLQ , **182**
M_W: 358,21 Da

Scheme 5. 1 Molecular structure of BLQ.

Following on-going extensive investigation on the PLA2 inactivation mechanism by anti-inflammatory marine natural products,[321,322,329,340] and attracted by the perspective of a rational design of novel inhibitors as new potential lead compounds against inflammation-related diseases,[341,342] the molecular mechanism of sPLA$_2$-IIA inhibition by BLQ was investigated. The opportunity of exploring the reactivity of the p-quinone systems in this biological context and the latest findings of an oncogenic action of PLA$_2$ in prostate cancer[343] offered further motivations for this project.

Herein, the reported results suggest that the inhibition mechanism is mainly ruled by a non-covalent event with a key role of the catalytic Ca^{2+} ion, anchoring the BLQ hydroxyl-quinone moiety inside the active site.

The reported results on the elucidation of the sPLA$_2$-IIA inactivation mechanism by BLQ consisted of the following steps: (a) analysis of the sPLA2-IIA inhibition profile; (b) structural analysis of the protein-inhibitor complex; (c) *in silico* generation of a 3D model of the protein–inhibitor complex. A combination of spectroscopic techniques (UV and Surface Plasmon Resonance (SPR)), classical protein chemistry protocols (proteolytic digestions, RP-HPLC), mass spectrometry (MS) and molecular modelling was

employed. The inhibition of sPLA$_2$-IIA activity by BLQ was monitored measuring the enzymatic activity of the native protein in the considered experimental conditions. sPLA$_2$-IIA enzymes are inactive on isolated phospholipid substrates, but undergoes 'interfacial activation' after the correct arrangement on the phospholipid layers.

In the attempt to ascertain the mechanism of the inhibition, the BLQ-sPLA2-IIA affinity by SPR biosensor analysis was monitored.[344] On the basis of these experiments, the experimental dissociation constant (K_{Dexp}) of the BLQ-sPLA$_2$-IIA complex was then calculated through the bioinformatics analysis of the sensogram traces. In absence of Ca^{2+} ions, K_{Dexp} was estimated at $1.13\pm0.75x10^{-6}$ M, whereas its value significantly decreased in presence of Ca^{2+} ions ($1.87\pm2.93x10^{-8}$ M). Thus, the inactivation mechanism appears to be mainly induced by a direct interaction between BLQ and sPLA$_2$-IIA. Furthermore, these preliminary results also suggest a key role of the calcium ion in the enhancement of BLQ affinity to the enzyme. As already done in previous studies, in which the bee venom PLA$_2$ was shown to be affected by both covalent and non-covalent interactions when incubated with several natural inhibitors,[321,322,329,340] it was applied a mass spectrometric approach in the analysis of the sPLA$_2$-IIA-BLQ complexes. A selective histidine alkylating agent pNBr (MW¼244 Da), well known for being able to specifically modify the catalytic His residues of different sPLA$_2$ (DM¼164 Da)[345] was used. Comparative pNBr modification experiments in presence and absence of a given inhibitor hence provides valuable evidence on the type of inhibition mechanism, (if competitive or not).

From the result analysis, all evidences suggest that BLQ acts as a competitive sPLA$_2$-IIA inhibitor, even though a mixed inhibition process cannot be excluded. Moreover, in the attempt to establish whether BLQ promoted

enzymatic inactivation in terms of pure non-covalent association, covalent modification of the enzyme or both, a RP-HPLC-MS analysis of the incubation mixture was performed.

The detection of a considerable amount of a sPLA$_2$-IIABLQ covalent adduct was obtained (Figure 5.2).

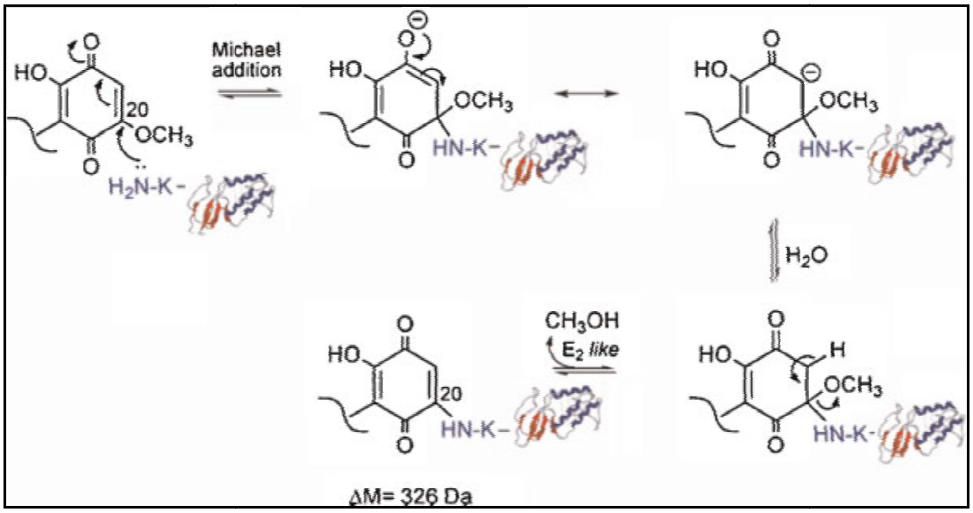


Figure 5. 2 Reaction mechanism for the covalent inactivation of sPLA$_2$-IIA by BLQ. BLQ reaction site has been pointed on C-20.[340]

This evidence suggests the formation of an irreversible complex between the protein and the inhibitor with 1:1 stoichiometry, due to a Michael addition of a sPL$_{A2}$-IIA nucleophilic group (e.g. an e-amino group of Lys) onto the p-quinone ring of BLQ, followed by a *b*-elimination of one molecule of methanol, as already reported in a previous study regarding the interaction between BLQ and bee venom PLA$_2$.[340] However, since the covalent reactivity of BLQ occurred at pH values far from physiological conditions, this reaction was considered irrelevant in the inactivation mechanism of the enzyme, and

consequently the identification of the reactive site(s) on the sPLA$_2$-IIA was left unresolved.

As a final step, in order to shed more light on the relevance of the non-covalent interaction in the enzyme inhibition process, it was resorted to a molecular docking approach by building a 3D model of the sPLA$_2$-IIA-BLQ complex. The human sPLA$_2$-IIA belongs to the calcium-dependent class of sPLA$_2$[346] in which the Ca^{2+} ion is anchored into the active site through the contacts with the carboxylate oxygen atoms of Asp48, the backbone oxygen of His27, Gly 29 and Gly 31 (See Figure 5.1), and the two oxygen of the substrate, if present. The crystal structure of inhibited sPLA$_2$-IIA from inflammatory exudates solved by Sigler and co-workers in 1991[347] (PDB code: 1POE) was used as molecular target for docking calculations (Figure 5.3).

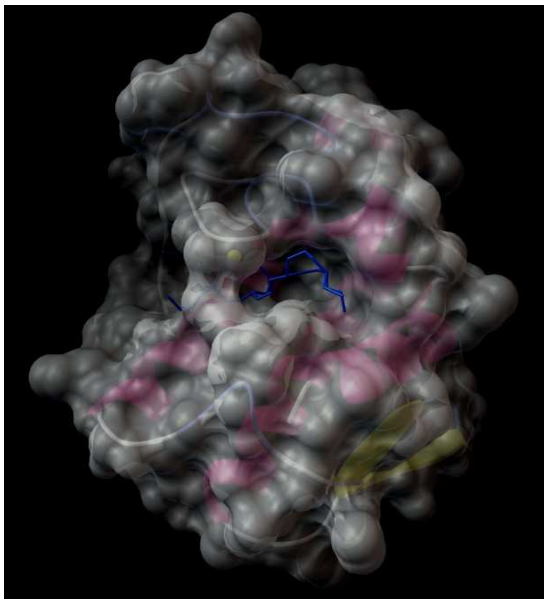


Figure 5. 3 Three-dimensional model of crystal structure solved by Sigler in 1991.

The first step of our studies was the refinement at density functional theory (DFT), B3LYP level of the sPLA$_2$-IIA catalytic site. In particular, in order to obviate for the lack of appropriate force field parameter to model the metal/ligand interactions, the charges of the Ca^{2+} ion and of the amino acids involved in the metal coordination, namely His27, Gly29, Gly30 and Asp48 (Figure 5.4) were calculated using Gaussian 03 Software Package.[160]

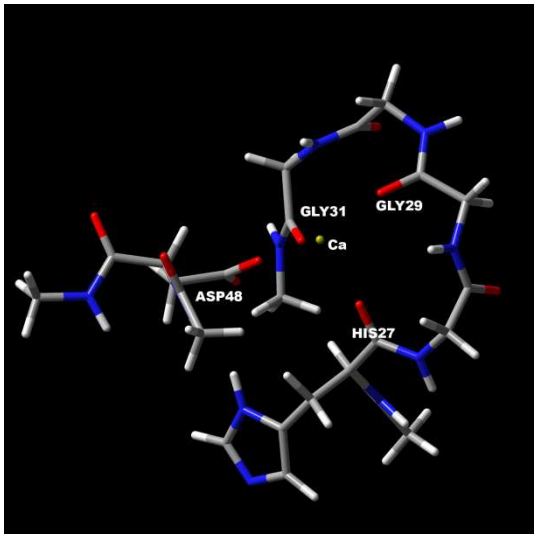


Figure 5. 4 Three-dimensional model of the Ca^{2+} ion anchored into the active site of sPLA$_2$-IIA.

Such QM supported procedure highly improved the accuracy of the molecular docking results, as already established in the docking studies of other metal enzymes.[147] The hydroxyl group of the BLQ quinone ring as dissociated at physiological pH and, consequently, it was considered deprotonated[348] in the molecular docking calculations by Autodock 4.0.1 software. In the three-dimensional model of the sPLA$_2$-IIA-BLQ complex, BLQ interacts in the enzyme active site with the same set of residues observed in the transition-

state analogue reported in literature.[347] Such analogy in the binding of sPLA$_2$ is depicted in Figure 5.5a, and strongly supports the significance of the obtained docking results. Mainly, the oxygen atoms of hydroxy–quinone system are of primary importance for their interactions network; namely, both the C-1 carbonyl group and the deprotonated C-2 hydroxyl residue coordinate the calcium ion in a bidentate fashion, while only the C-2 hydroxyl establishes hydrogen bonds with NH of Gly29 (Figure 5.5b).

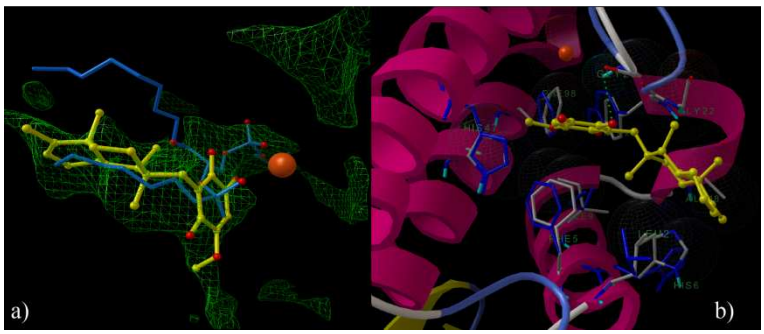


Figure 5. 5 (a) BLQ and transition-state analogue superimposition in the sPLA2-IIA binding site. The green mesh represents the hydrophobic pocket of the protein. The Ca^{2+} is represented in orange. BLQ is depicted by sticks (yellow) and balls (C, yellow; O, red); the transition state analogue is shown by sticks (sky blue) and balls (C, sky blue;O, red). The figure highlights the similar interactions with the shallow hydrophobic groove on the enzyme surface. (b) The flexible protein is represented by CPK and sticks and balls coloured by atom type (by atom type: C, gray; polar H, sky blue; N, blue; O, red), the rigid protein is represented by dark blue sticks and balls. BLQ is depicted by sticks and balls (by atom type: C yellow, O red).

Moreover, it is important to note that the flexible amino acids Leu2, Phe5, His6, Ile9, Tyr21, His47, Phe98 rearrange around the BLQ (see relative position of the amino acids in the rigid—coloured in dark blue—and flex target-coloured by atom type-) generating multiple Van der Waals interactions with the target. Moreover the entire hydroxy–quinone ring interacts with

Gly22, Gly29, Val30 and the bicyclic sesquiterpene moiety creates a number of hydrophobic contacts with Ala17, Ala18, Gly22, Cys28 and Gly29. All the above-mentioned contacts contribute to the final theoretical dissociation constant (K_{Dcalc}) of $6.92x10^{-6}$ M. Indeed, both the non-covalent interactions with large hydrophobic surfaces of the active site and the chelation of the essential calcium ion by BLQ are responsible of the competitive inhibition process.

Besides, in order to validate and to obtain a term of quantitative comparison for the results described above, it was re-investigated by molecular docking (Autodock 3.5 software)[55] the bee venom PLA_2 inhibition by BLQ whose main features had been previously investigated by molecular docking.[340]

Also for the bvPLA_2, its binding site was refined at Quantum Mechanical (QM) level. The charges of Ca^{2+} ion 501 and of Trp8, Gly10, Gly12, Asp35 were calculated at the DFT B3LYP level using the 6-31+G(d) basis set and ChelpG[148] method for population analysis (Figure 5.6) using Gaussian 03 Software Package.[160]

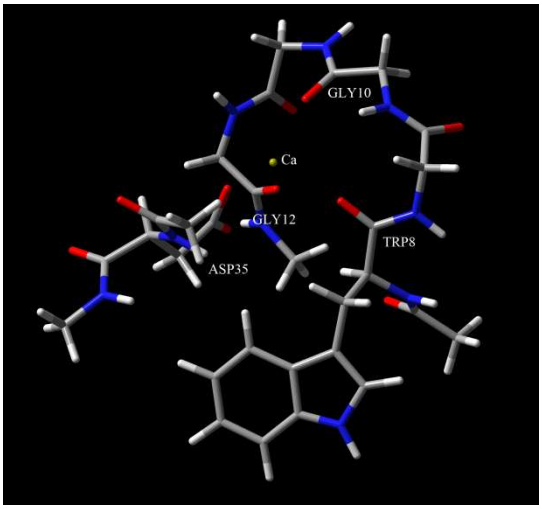


Figure 5. 6 3D model of the Ca^{2+} ion anchored into the active site of *bv*PLA_2.

This refined structure of target with the BLQ optimized, as described above, were used as input for docking calculations. The analysis of the complex between *bv*PLA$_2$ and BLQ (Figure 5.7) revealed that the oxygen atoms of hydroxy-quinone system bind to the calcium ion in a bidentate fashion and establish hydrogen bonds with the amino acids of the catalytic center, in particular, C-4 carbonyl group with NH of Gly12, C-5 methoxyl group with NH of Gly10, and C-1 carbonyl group with NH of Thr57. Moreover, the hydroxy-quinone system interacts with the hydrophobic aminoacids of the catalytic site: Gly10, His11, Gly32, His56, Thr57.

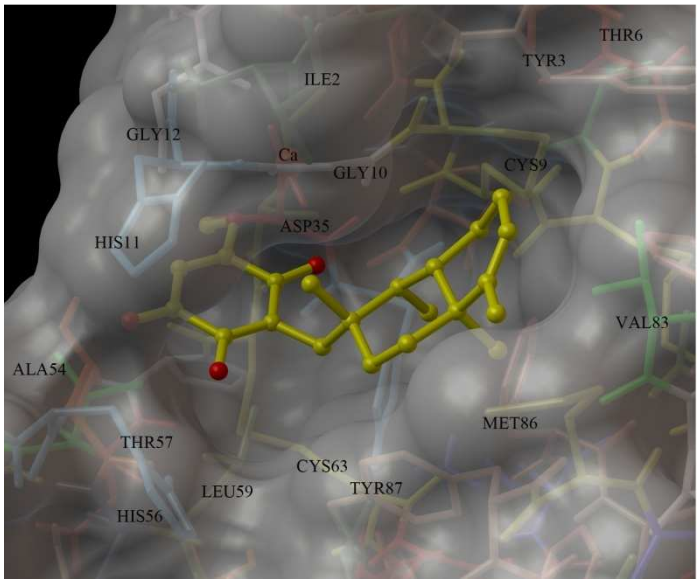


Figure 5. 7 3D model of the interaction between BLQ and the *bv*PLA$_2$ binding site.

The bicyclic sesquiterpene moiety establishes appropriate hydrophobic contacts with the Tyr3, Cys9, Gly10, Val83, Met86, Tyr87 playing an essential role in the inhibition mechanism. In summary, the docking calculations show a slight different arrangement of the BLQ in the bee venom

PLA_2 catalytic centre with respect to the complex with the $sPLA_2$-IIA. (K_{Dcalc} of the BLQ-bee venom PLA_2= $41.13x10^{-6}$ M). Such results are in good qualitative agreement with the biological assays,[349] where BLQ inhibits $sPLA_2$-IIA, and bee venom PLA_2 with IC_{50} values of 0.2 and 0.1 μM, respectively.

Moreover, on the basis of the mechanism of histidine alkylanting agent Para-Nitrophenacyl-Bromide (pNBr) the three dimensional model of the *hs*PLA2 with the modified His47 was built. The modified active site was refined at the quantum mechanical (QM) level calculating the partial charges of the Ca^{2+} ion and of the amino acids involved in the metal coordination (His27, Gly29, Gly30, Asp48), and of the His47 covalent linked to pNBr (Figure 5.8) at the DFT B3LYP level using the 6-31+G(d) basis set and ChelpG[148] method for population analysis using Gaussian 03 Software Package.[160]

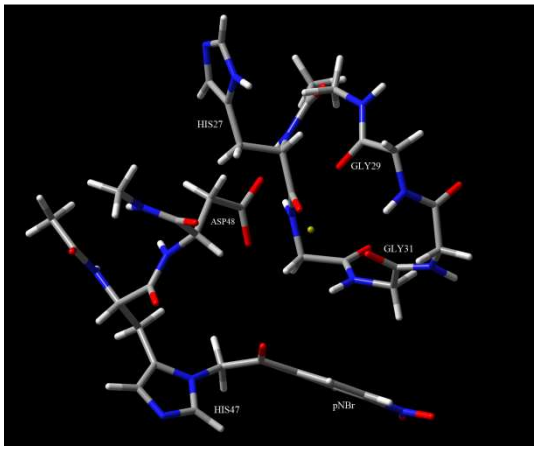


Figure 5. 8 3D model of the Ca^{2+} ion anchored into the modified active site of $sPLA_2$-IIA.

Docking calculation by Autodock 3.0.5[55] were performed on the mutated $sPLA_2$-IIA and the refined BLQ structure. From the analysis of the calculations it emerged that, in the modified complex the pNBr fills equivalent

space of the BLQ complexed with the *hs*PLA2, as described in the main text (Figure 5.5), on consequence in the three dimensional model the hydroxyquinone marine terpenoid is obliged to arrange at the opposite side of the pNBr (Figure 5.9).

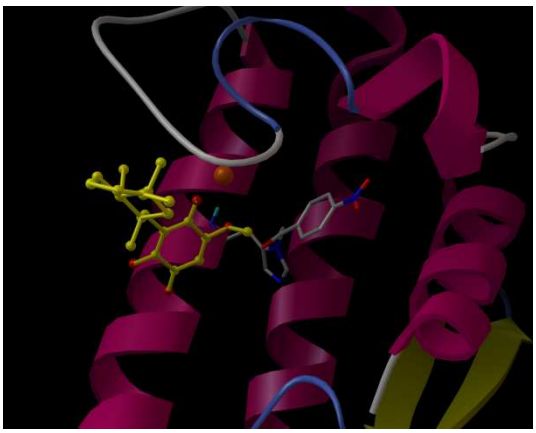


Figure 5. 9 3D model of the BLQ and the pNBr into the active site of *hs*PLA2. The protein is represented by secondary structure. BLQ is depicted by sticks and balls (colored by atom type: C yellow, O red). The pNBr is depicted by sticks colored by atom type: C grey, O red, N dark blue.

The BLQ coordinates the calcium ion in the bidentate fashion, and the C-4 carbonyl group establishes hydrogen bonds with NH of Gly31 (not displayed), and it interacts with the Gly31, Val30, Asp48, Gly125, Lys52, Gly125, Tyr51, Tyr126. All the above-mentioned contacts contribute to the final theoretical dissociation constant (K_{Dcalc}) of $2.3x10^{-5}$ M. This result suggests that alkylation of His47 with pNBr do not completely block, but undoubtedly obstruct the BLQ binding suggesting that the suboptimal hydrophobic interactions are responsible for a predicted decrease in the binding affinity to the receptor of about 1000-fold (K_{Dcalcr} of BLQ $1.55x10^{-8}$ *vs* K_{Dcalc} of BLQ in

presence of pNBr 2.3x10^{-5}), confirming that the non-covalent interactions with large hydrophobic surfaces of the active site are the driving force in the inhibition processes.

On summary, it is possible to affirm confident that the inactivation mechanism is completely due to a non-covalent event at physiologically relevant conditions. Thus, the covalent reactivity of BLQ detected in more drastic experimental setting, at high pH values, must be considered irrelevant in terms of enzyme inhibition. The comprehension of the molecular mechanism of the sPLA$_2$-IIA inactivation by BLQ could be helpful for the design and synthesis of a new chemical class of PLA$_2$ inhibitors, able to specifically target the enzyme active site without being endowed with covalent reactivity.

5.1.2 Computational details

All the QM and molecular docking calculations were performed on a two dual-core Intel Xeon 3.4 GHz. The charges of the sPLA$_2$-IIA chain A (Ca^{2+} ion 801 and of Gly29, Gly31, His27, Asp48) catalytic centre were calculated at the DFT B3LYP level using the 6–31+G(d) basis set and ChelpG[148] method for population analysis. The standard AutoDock 4 force field parameters were used for the docking calculations of the metallo-protein sPLA$_2$-IIA: calcium radius=1.98Å and well depth=0.55 kcal/mol. A flexible docking involving the amino acids of the shallow groove on the enzyme surface, namely Leu2, Phe5, His6, Ile9, Tyr21, His47, Phe98 was performed. A grid box size of 50 x 50 x 50 with spacing of 0.375 Å between the grid points, and centered at 23.0 (x), 5.5 (y) and 64.0 (z), covering the active site on the sPLA$_2$-IIA surface was used. The implicit salvation method of Autodock 4.0.1 was used in the

calculation. In particular, the water was considered as a continuous medium, and the desolvation term was evaluated considering: (a) an atomic solvation parameter for each atom type, corresponding to an estimate of the energy needed to transfer the atom between a fully hydrated state and a fully buried state, and (b) an estimate of the amount of desolvation when the ligand is docked. The Lamarckian genetic algorithm with an initial population of 150 randomly placed individuals, a maximum number of $2.5x10^5$ energy evaluations, and a maximum number of $2.7x10^4$ generations were taken into account was used for dockings by Autodock 4.0.1 software.[56] A mutation rate of 0.02 and a crossover rate of 0.8 were used. Results differing by less than 2.0 Å in positional root-mean-square deviation (RMSD) were clustered together and represented by the result with the most favorable free energy of binding.

For what concern *bv*PLA$_2$ (PDB code 1POC),[350] it being a metal enzymes and so a non-bonded model for metal center (Calcium Radius=1.98 Å, well depth=0.55 kcal/mol) was used. For all the docking calculations a grid box size of 100 x 96 x 115 with spacing of 0.375 Å between the grid points, and centered at 46.6 (x), 30.6 (y) and 26.0 (z), covering the active site on the bvPLA$_2$ surface was used. In order to achieve a representative conformational space during the docking calculations, six calculations consisting of 256 runs were performed, obtaining 1536 structures (256 x 6). The Lamarckian genetic algorithm was used for dockings. An initial population of 150 randomly placed individuals, a maximum number of 2.5×10^5 energy evaluations, and a maximum number of 2.7×10^4 generations were taken into account. A mutation rate of 0.02 and a crossover rate of 0.8 were used. Results differing by less than 2.0 Å in positional root-mean-square deviation (RMSD) were clustered together and represented by the result with the most favorable free energy of binding.

For what concern a modified model sPLA$_2$-IIA with pNBr, as described above, a grid box of 48 x 60 x 66 with spacing of 0.375 Å between the grid points, and centered at 23.0 (x), 6.6 (y) and 65.0 (z), covering the active site on the sPLA$_2$-IIA surface was used. In order to achieve a representative conformational space during the docking calculations, six calculations consisting of 256 runs were performed, obtaining 1536 structures (256 x 6). The Lamarckian genetic algorithm was used for dockings. An initial population of 150 randomly placed individuals, a maximum number of 2.5 x 10^5 energy evaluations, and a maximum number of 2.7 x 10^4 generations were taken into account. Results differing by less than 3 Å in positional root-mean-square deviation (RMSD) were clustered together and represented by the result with the most favourable free energy of binding.

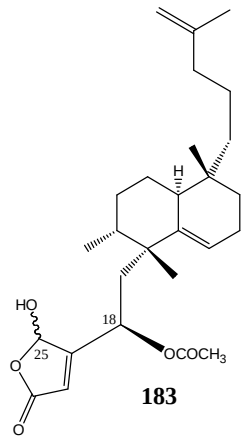
All the 3D models were depicted using the Python[165] software, molecular surfaces are rendered using maximal speed molecular surface (MSMS).[166]

5.3 The Binding Mode of Cladocoran A to the Human Group IIA Phospholipase A2

Cladocoran A (CLD A, Scheme 5.2) is a member of the sesterterpenoid ghydroxybutenolides with a unique carbon skeleton, isolated from the Mediterranean coral Cladocora cespitosa by Fontana et al., together with its C-18 deacetylated congener (CLD B).[351,352,353] CLD A is endowed with a high sPLA$_2$ inhibitory effect, comparable to that of manoalide, with an IC$_{50}$ of 0.78(±0.06) μm.[351] On this basis, it would be interesting to disclose the CLD A mechanism of action on sPLA2-IIA enzyme.

The elucidation of the sPLA$_2$-IIA interaction mechanism by CLD A was performed by a detailed structural analysis of the covalent and noncovalent contributions to the protein–ligand complex, through a combination of classical protein chemistry protocols (proteolytic digestion, RP-HPLC), nano-ESI and MALDI mass spectrometry (MS). Furthermore, an *in silico* 3D model was generated by molecular docking calculations to provide a more detailed picture of the protein–ligand adduct. The results described in this paragraph prove that the interaction mechanism is mainly ruled by a no covalent competitive occupancy of sPLA2-IIA active site by CLD A, together with a selective covalent modification of the enzyme.

As already carried out in previous studies on sPLA$_2$,[319-322,329,340,291,354] the first step is represented by the analysis of sPLA2 potential covalent modification by CLD A through liquid chromatography-nano-ESI MS (LC-nano-ESI-MS).



183

Scheme 5. 2 Molecular structure of CLD A.

In the first pathway, two reaction steps can be postulated. In the first pathway (Figure 5.10A), in which an aspecific esterase activity of the enzyme gives rise, after hydrolysis at the C-18 acetate function, to a secondary alcohol and a Schiff base formation between a nucleophilic amine residue on the sPLA$_2$-IIA and the masked aldehyde of the γ-hydroxybutenolide ring (total mass increment of 384.2 Da) occurs. In a second hypothesis (Figure 5.10B) nucleophilic attack of a sPLA$_2$-IIA residue on CLD A C-18 causes the loss of acetic acid (total mass increment of 384.35 Da).

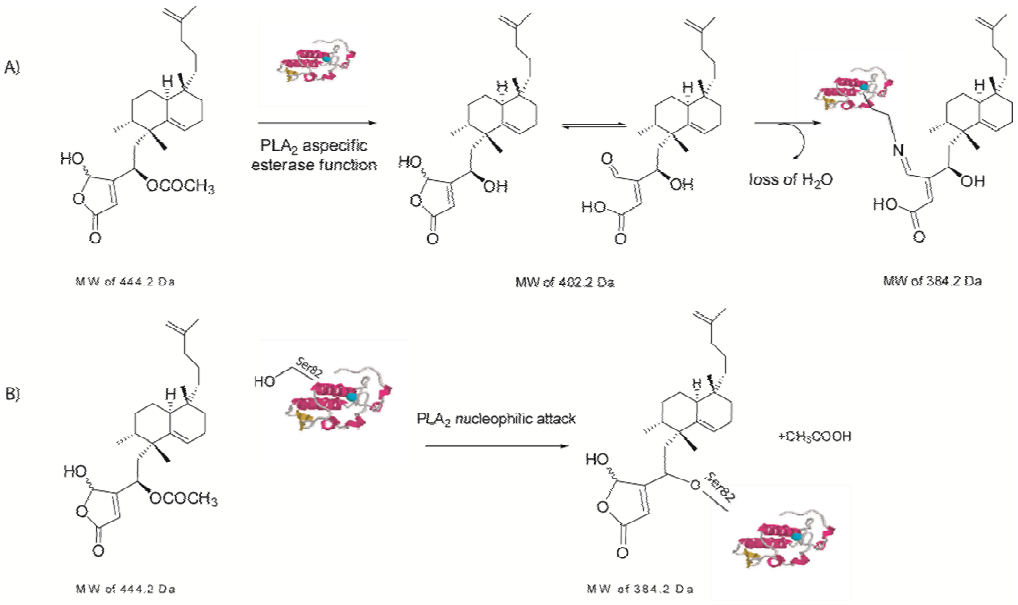


Figure 5. 10 Covalent interaction pathways between sPLA$_2$-IIA and CLD A. (a) Aspecific esterase activity of sPLA$_2$-IIA, which hydrolyzed the acetate on the C-18 of CLD A, followed by a Schiff base formation between a sPLA$_2$-IIA aminic residue and CLD A masked aldehyde of the g-hydroxybutenolide ring. (b) Formation of the covalent adduct is thought to occur by nucleophilic attack of sPLA$_2$-IIA Ser82 residue on CLD A C-18 followed by acetic acid loss.

Thanks to a combination of classical protein chemistry protocols and MALDI-MS techniques, residue(s) involved in the covalent binding with CLD A were unequivocally determined. In particular, Ser 82 was identifies as the selective site of covalent modification of the enzyme on the protein surface.

Since Ser82 is located far from the sPLA$_2$-IIA active site, its covalent modification cannot be easily correlated with enzyme inhibition. In order to determine the role of the noncovalent interactions in sPLA$_2$-IIA inactivation by CLD A, the p-nitrophenacyl bromide (pNBr, MW=244 Da), a well-known specific ligand of the catalytic His residues of several PLA$_2$ was used.[329,354,291,355] The results analysis show as a low reactivity of sPLA$_2$-IIA

towards pNBr result of specific accommodation of CLD A into the enzyme active site.

In order to rationalize the binding mode of CLD A on sPLA2-IIA enzyme (PDB ID: 1POE)[347] molecular docking calculations using Autodock 4.2 software[356] were performed. As reported in the paragraph 4.2,[291] sPLA2-IIA structure refined at quantum mechanical level, in which the charge of the amino acids of the catalytic site (namely His27, Gly29, Gly30 and Asp48 and calcium ion) were calculated at density functional theory (DFT) B3LYP level, was used as model receptor. On the other hand, the CLD A γ-hydroxybutenolide ring was considered whether opened and deprotonated at physiological pH. Prior to the docking calculations, CLD A conformational search by means of molecular dynamics at different temperatures (300 and 750 K) with a Monte Carlo conformational search using the MMFFs force field included in the MacroModel[159] software package were performed.

As shown in Figure 5.11, CLD A fits into the sPLA$_2$-IIA active-site pocket, establishing extensive van der Waals and hydrophobic interactions with protein side-chains and chelating the essential calcium ion with its carboxylate moiety. Mainly, it is able to interact with Leu2, Phe5, Ile9, Ala17, Gly22, Cys28, Val30, Phe98, and it forms hydrogen bonds with Lys62 and Gly29 (Figure 5.11 B). All the above-mentioned contacts contribute to the final theoretical dissociation constant (K_{Dcalc}) of 0.38 μm.

Thus, the results of docking calculations support the view that CLD A preferentially interacts with the enzyme active site, as previously disclosed by MS analysis.

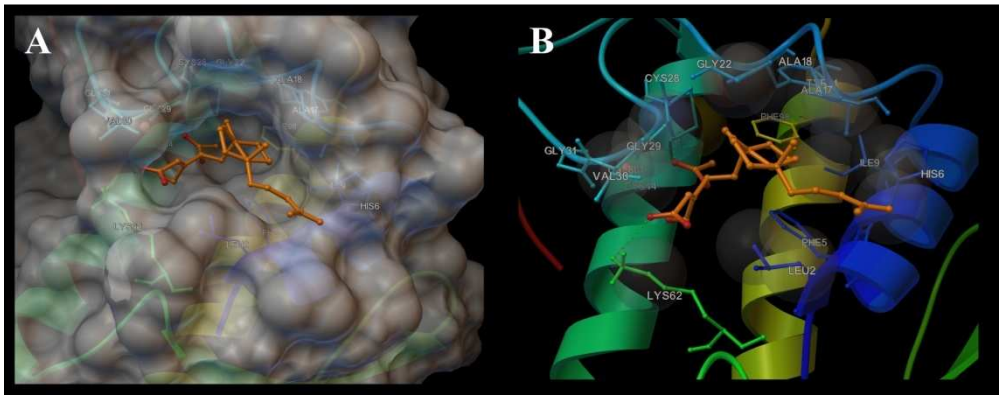


Figure 5. 11: a) Three dimensional model of the complex between CLD A and sPLA2-IIA; b) detailed interactions of CLD A with the sPLA2-IIA active site.

Moreover, since MS data clearly show also the existence of covalent binding of CLD A on sPLA2-IIA surface, a second molecule of CLD A on sPLA2-IIA–CLD A noncovalent complex was docked, in order to establish if the Ser82 environment can be recognized by the marine bioactive sesterterpenoid. The reported docking calculations showed that CLD A is able to interact with external part of the enzyme presenting amino acid Ser82 (Figure 5.12). In particular, the ligand is able to form a hydrogen bond with Ser82 NH by the CLD A carbonyl at C-18, and with Arg57 and Cys83 side-chains by CLD A carboxyl and carbonyl groups on the g-hydroxybutenolide ring. Moreover, CLD A is in close contact with Gly58, Cys59, and Asp91 on the protein surface.

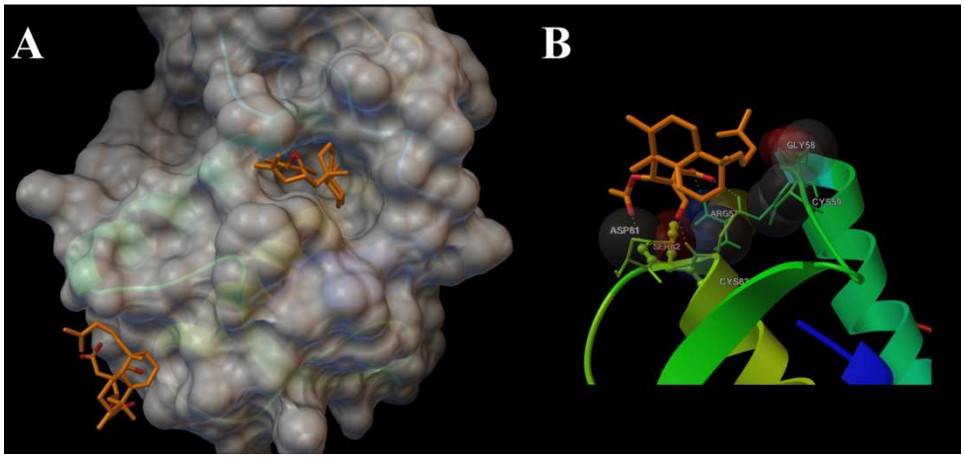


Figure 5. 12: a) Three dimensional model of the complex between two molecules of CLD A and sPLA$_2$-IIA. b) Detailed interactions of the second molecule of CLD A with the external part of sPLA$_2$-IIA.

This pattern of hydrophilic and hydrophobic interactions of CLD A with sPLA$_2$-IIA causes the projection of the reactive center C-18 bearing an acetyl function towards the hydroxyl group of sPLA$_2$-IIA Ser82. However, the minor efficiency of such interactions is responsible for a low calculated binding affinity to the target (K_{Dcalc} of 194.32 μm).

The reported results, collected by a combination of biochemical approaches, advanced mass spectrometry and molecular modeling, suggest a competitive inhibition mechanism guided by a noncovalent molecular recognition event, and disclose the key role of the CLD A γ-hydroxybutenolide ring in the chelation of the catalytic calcium ion inside the enzyme active site. Moreover, CLD A is able to react selectively with Ser82, although this covalent event seems to play a secondary role in terms of enzyme inhibition.

5.3.1 Computational details

The sPLA$_2$-IIA model optimzed at QM level as described in the section 4.2.1 was used for molecular docking calculations using Autodock 4.2 software[356] on a two dual-core Intel Xeon 3.4 GHz. For the docking studies, a grid box size of 126 x 126 x 126 and centered at 23.912 (x), 1.212 (y), 67.788 (z) in the first calculation (1:1, sPLA2-IIA/CLD A), and of 50 x 72 x 46 and centered at 41.225 (x), 2.823 (y), 59.81 (z) in the second case (1:2, sPLA2-IIA–CLD A/CLD A) was used. The spacing between the grid points was 0.375 Å, and the active site and the external part of sPLA$_2$-IIA were covered in both calculations. To achieve a representative conformational space during the docking studies and to take into account the variable number of active torsions, ten calculations, in both cases consisting of 256 runs, were performed, and 2560 structures were obtained for CLD A. The Lamarckian genetic algorithm (LGA) was employed for docking calculations, choosing an initial population of 600 randomly placed individuals. The maximum number of energy evaluations and of generations was set to $5x10^6$ and to $6x10^6$. Results differing by less than 3.5 Å in positional root-mean-square deviation (RMSD) were clustered together and represented by the most favorable free energy of binding. Illustrations of the 3D models were generated the Python software.[165]

-CHAPTER 6-

4-Methylensterols as ligands of Human Farnesoid-X-Receptor (FXR) and Human Pregnane-X-Receptor (PXR)

6.1. The Farnesoid X Receptor (FXR) and Human Pregnane-X-Receptor (PXR) as Modulator of Bile Acid Metabolism

In this chapter it will be described, by mean of molecular docking, the binding modes of the marine natural products extract from *Theonella swinhoei*, acting as ligands of two foundamental bile acid receptors: human farnesoid-X-receptor (FXR) and human pregnane-X-receptor (PXR). In particular it will be rationalize at atomic level the antagonist and/or modulator activity of Conicasterols (B-D, G-K), Theonellasterols (B-H, J), and Swinosterol B on FXR, and their agonist activity on PXR.[294,357,358,359]

Nuclear receptors (NRs) are ligand-dependent transcription factors that control many biological functions, such as cell growth, differentiation, embryonic development, and metabolism. Upon activation by binding of small lipophilic molecules, such as steroids and thyroid hormones, retinoids, vitamin D, and dietary and endogenous lipids, NRs interact with coactivators to modulate directly the expression of responsive genes involved in development, reproduction, and metabolism.[360] Pharmaceutical control of the activity of NRs with synthetic ligands having agonistic or antagonistic activity is a powerful tool for the management of various clinical conditions, including several forms of cancer, type 2 diabetes, and metabolic syndrome.[361]

Hepatobiliary transport of bile acids is mediated by specific transporters expressed at the canalicular membrane of the hepatocyte. Impaired function of these transporters leads to reduced bile formation or cholestasis and mutations in these genes are associated with a variety of hereditary cholestatic syndromes.[362,,364] At the transcriptional level, these transporters and the phases I and II metabolizing enzymes involved in processing of their substrates are coordinately regulated by members of the nuclear receptor family of ligand-

modulated transcription factors.[365] In the last decade, two nuclear receptor, the farnesoid-X-receptor (FXR) and the pregnane-X-receptor (PXR), with both receptors functioning as bile acid activated receptors, have emerged as the main receptors involved in regulating bile acid synthesis, detoxification and excretion in the liver and gastro-intestinal tract.[366,367,368369] However, since FXR agonists inhibit MRP-4, a basolateral transporter that regulate bile acids excretion from basolateral membrane of hepatocytes,[370] and FXR gene ablation protects against liver injury caused by bile duct ligation,[371] it has been suggested that FXR antagonists, rather than FXR agonists, might hold promise in the treatment of cholestatic disorders.[372],[293] In addition to FXR, PXR is a recognized target for treatment of cholestasis.[373]

The pregnane X receptor (PXR; also known as NR1I2), a member of the nuclear receptor family of ligand-activated transcription factors, is a key regulator of cytochrome P450-3A (*CYP3A*) gene expression in mammalian liver and small intestine,[374] CYP2B6,[375] and CYP2C9 as well as many other enzymes and transporters, such as P-glycoprotein (ABCB1),[376] and proteins involved in the transport, metabolism, and biosynthesis of bile acids.[377] These discoveries have shown how drugs may regulate not only their own metabolism but potentially also their own efflux, as demonstrated for paclitaxel.[378] Overall, there is an incredibly broad structural diversity in the molecules that bind to human PXR in vitro. In particular, PXR is activated by a broad range of compounds as e.g. bile salts,[378,379] cholesterol and its metabolites,[380] statins,[381] endocrine disruptors,[382] imidazoles,[383] biphenyles,[384] azoles,[385] synthetic peptide bond mimetics,[386] anticancer compounds,[387] herbal components and plant extracts,[388,389,390] carotenoids,[391] vitamins,[392] HIV protease inhibitors,[393] calcium channel modulators,[394] steroids,[395] plasticizers,[396] pesticides,[397] peroxisome proliferator-activated receptor-γ

antagonists,[398] as well as other diverse xenobiotics and endobiotics,[399,400,401,402,403,404] including agonists for additional nuclear receptors[405] (See Figure 6.1 for some examples).

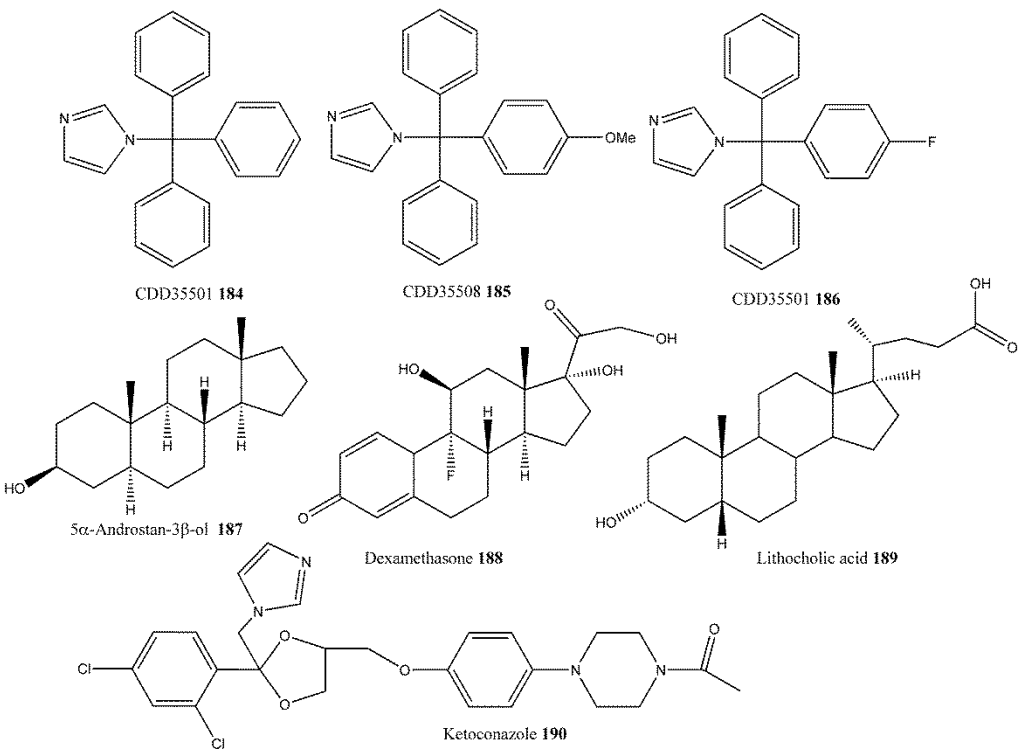


Figure 6. 1 Structures of selected PXR agonists and antagonists.

Although these diverse interactions imply promiscuity, PXR also exhibits specificity, as evidenced by the differences in the pharmacologic activation profile of PXR across species. For instance, human PXR is activated by rifampicin and the cholesterol-lowering drug SR12813,[406] whereas mouse PXR is not;[412] mouse PXR is activated by the synthetic steroid 5-pregnen-3b-ol-20-one-16a-carbonitrile (PCN), whereas the human receptor is not. Thus,

by binding diverse but precise arrays of compounds, PXR exhibits directed promiscuity.

PXR acts as a transcription factor.[407] Once activated, PXR heterodimerizes with the retinoid-X-receptor (RXR) (Figure 6.2), binds to regulatory DNA sequences in the nucleus, and modulates transcription of genes involved in the oxidation, conjugation, and export of compounds from cells. PXR mediates the detoxification of endogenous and exogenous compounds and prevents toxic accumulation of metabolites within cells. While PXR is primarily expressed in the gastrointestinal tract and liver, recent reports showed that PXR is expressed in immune cells, including $CD4^+$, $CD8^+$ T lymphocytes, $CD19^+$ B lymphocytes, and $CD14^+$ monocytes in humans.[408,410]

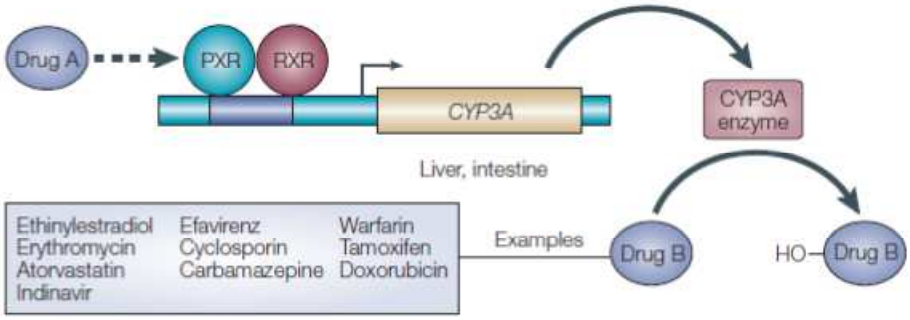


Figure 6. 2 Drug–drug interactions. The molecular basis of a drug–drug interaction. The orphan nuclear receptor PXR is a transcription factor that regulates the expression of the *CYP3A* gene (yellow) in the liver and intestine. It functions as a heterodimer with the nuclear receptor RXR. Drug A binds to PXR and induces expression of the CYP3A enzyme (pink), accelerating the metabolism of drug B, which is a substrate for CYP3A. CYP, cytochrome P450; OH, hydroxyl group; PXR, pregnane X receptor; RXR, retinoid X receptor.

Several studies have shown that rifampicin, an human PXR activator,[408,409] used for treating pruritus in cholestasis, suppresses both humoral and cellular

immunity and identified rifampicin as a powerful immunosuppressive drug. Moreover, reciprocal repression between PXR and NF-κB was shown. In the colon, PXR-mediated repression of NF-κB target genes appears to be a critical mechanism by which PXR activation decreases inflammation[408,409] The *CYP3A* gene products are heme-containing proteins that metabolize a wide variety of chemicals, including >50% of all prescription drugs.[411] PXR is activated by most of the xenobiotics (exogenous chemicals) that are known to induce *CYP3A* gene expression, including the commonly used antibiotic rifampicin, the glucocorticoid dexamethasone, and the herbal antidepressant St. John's wort.[374,412] Like other nuclear receptors, PXR contains both a DNA-binding domain and a ligand-binding domain. PXR binds to the xenobiotic DNA response elements in the regulatory regions of *CYP3A* genes as a heterodimer with the 9-cis retinoic acid receptor, also known as the retinoid X receptor (RXR).[374] PXR can mediate dangerous drug-drug interactions. For example, hyperforin,[413] a constituent of St. John's wort, activates PXR and upregulates *CYP3A* expression, which leads to the metabolism of vital drugs including the antiretroviral drug indinavir and the immunosuppressant compound cyclosporin.[414,415,416] Cholestasis is a severe hepatic disorder that is characterized by the accumulation of toxic bile acids in the liver.[417] It is a chronic incurable disease that ultimately leads to fatal hepatic failure. *Pxr* $^{-/-}$ mice show defective bile-acid excretion[418,419] and the herb St John's wort — a PXR activator[420,421]— has been used as a folk remedy for cholestasis.[422] So, the use of PXR agonists as therapeutic agents might be warranted in these patients, despite the likelihood of inducing deleterious metabolism of other drugs.

Among nuclear receptors, farnesoid X receptor (FXR; also known as NR1H4) has emerged as a valuable pharmacological target[423] because of its

role in regulating bile acids (BAs), lipid, and glucose homeostasis. Activation of FXR, highly expressed in the liver, intestine, kidney, and adrenals, leads to complex responses, the most relevant of which is the inhibition of bile acids synthesis through the indirect repression of the expression of cytochrome 7A1 (CYP7A1), the rate limiting enzyme of this pathway.

FXR is activated by cholesterol metabolism end-products, bile acid derivatives, such as primary bile acids and secondary bile acids, and synthetic ligands.[424] Primary bile acids include chenodeoxycholic acid (CDCA) and cholic acid (CA), while secondary bile acids include deoxycholic acid (DC) and lithocholic acid (LCA). FXR heterodimerizes with another nuclear receptor, retinoid X receptor (RXR) (Figure 6.3), and the heterodimer regulates gene expression by binding to a specific consensus DNA sequence, termed farnesoid X responsive element (FXRE), which is an inverted repeat of the hexameric AGGTCA recognition motif separated by single nucleotide (IR-1), located in the promoter region of FXR target genes.[425]

FXR regulates the expression of genes encoding proteins involved in cholesterol homeostasis, in bile acid biosynthesis (CYP7A1[426] and CYP8B1[427]) transport and disposition (BSEP,[428] IBABP,[428] and NTCP[429]) FXR regulates the expression of small heterodimer partner (SHP).[430] SHP attenuates the expression of CYP7A1 by inhibiting the activity of liver receptor homologue 1 (LRH-1), which is known to augment CYP7A1 expression.[430] FXR also decreases the expression of CYP8B1, which is the enzyme catalyzing hydroxylation of CDCA at the 12a position to produce CA. FXR induces bile salt export pump (BSEP), which transports bile acids from hepatocytes to bile canaliculi, and induces the expression of the intestinal bile acid binding protein (IBABP), which shuttles bile acids from the apical to the basolateral side of the enterocytes during their absorption.

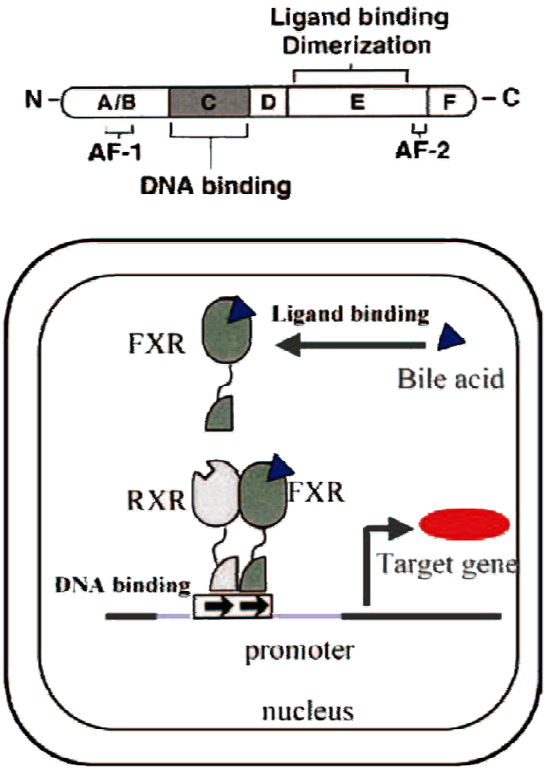


Figure 6. 3 Schematic representation of FXR structure (top) and FXR/RXR heterodimer formation upon activation by their respective ligands.

It down-regulates the expression of Na^+ taurocholate cotransporting polypeptide (NCTP), which takes up bile salts into the liver after intestinal absorption.

As already mentioned above, cholestasis is a liver disorder that occurs primarily in the context of genetic mutation of basolateral or apical membrane transporters in hepatocytes. Cholestasis represents the main biochemical feature of primary biliary cirrhosis[431,432] (PBC) and sclerosing cholangitis (PSC), two immune-mediated disorders characterized by progressive bile duct destruction for which medical therapy is still poorly effective and investigations are ongoing to identify novel therapeutic approaches.[431,432] In

addition to PSC and PBC, an obstructive form of cholestasis occurs in patients suffering from biliary stones or biliary and pancreatic tumors.[431] Theoretically, because PBC and PSC are characterized by bile duct destruction, therapy should be aimed at activating bile acid secretion from the basolateral membrane of hepatocytes, while stimulation of bile acid secretion from the apical membrane is likely to worsens liver injury due to the obstruction of bile flow.[433] FXR is a bile acid sensor that regulates bile acid synthesis and excretion. While activation of FXR favours bile acid detoxification by hepatocytes and FXR ligands have been proposed in the treatment of PBC patients,[431] results from models of obstructive cholestasis in $FXR^{-/-}$ mice have shown that FXR gene ablation protects against liver injury caused by ligation of common bile duct (BDL).[433] Molecular decoding of BDL model has lead to the demonstration that FXR functions as a negative regulator of multidrug resistance-associated protein (MRP)-4, a gene mediating basolateral secretion of bile acids. Thus, while $FXR^{-/-}$ mice adapt to bile duct obstruction by an ≈20 fold induction in the expression of MRP-4 mRNA, these changes are not reproduced in wild type mice. Because, the induction of MRP-4 represents an adaptive response to bile duct obstruction and protects the liver from accumulation of toxic bile acids during cholestasis by facilitating their efflux into blood for ultimate renal excretion, and MRP-4-knockout mice are sensitised to liver injury induced by BDL,[434] regulation of this basolateral transporter exerts an essential role in orchestrating the adaptive changes under conditions of impaired canalicular bile salt excretion.[432,435a,357]

After its deorphanization[436,437,438] a number of nonsteroidal[439] and steroidal compounds,[440] have been shown to interact with the ligand binding domain (LBD) of the receptor and to promote FXR mediated gene transcription (Figure 6.4).[441]

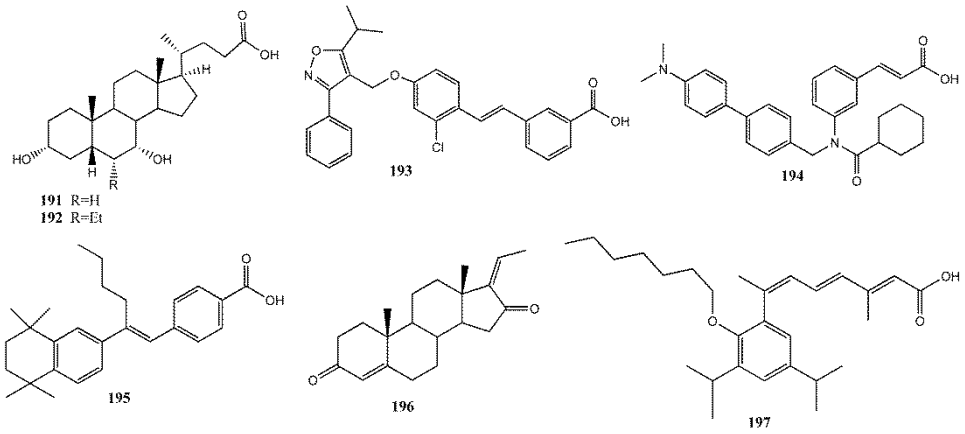


Figure 6. 4 Structures of FXR agonists CDCA (**191**), 6-ECDCA (**192**), GW-4064 (**193**), fexeramine (**194**), and AGN-31 (**195**): and of FXR antagonists guggulsterone (**196**) and AGN-34 (**197**).

Among these, 6-ECDCA (**192**) has emerged as a potent, orally bioavailable FXR agonist,[442] and ongoing clinical trials have shown its utility in the treatment of type 2 diabetes.[443] In this scenario the discovery of FXR modulators represents an important answer to the urgent demand of new drugs for the treatment of relevant human diseases including dyslipidemia, cholestasis, nonalcoholic steatohepatitis (NASH), and type 2 diabetes.

Nevertheless the use of potent FXR ligands holds some potential risk. Indeed, it has been shown that FXR activation in mammalian cells and tissues inhibits biosynthesis of endogenous bile acids by indirect transrepression of cholesterol 7α-hydroxylase (CYP7A1), a gene encoding for the first and rate limiting enzyme involved in their biosynthesis. This effect is indirect and mediated by activation of SHP, small heterodimer partner, an atypical nuclear receptor that lacks the DNA binding domain and that binds to liver X receptor (LXR), causing its displacement from a positive regulatory element in the CYP7A1 promoter.[423] Despite the effect of SHP, which has been shown to be dispensable in some settings, it is well recognized that SHP activation

amplifies the effects of FXR on bile acids uptake and biosynthesis, strongly suggesting that identification of SHP-sparing FXR modulators might have the potential to promote bile acid detoxification without interfering with the biosynthesis (Figure 6.5).

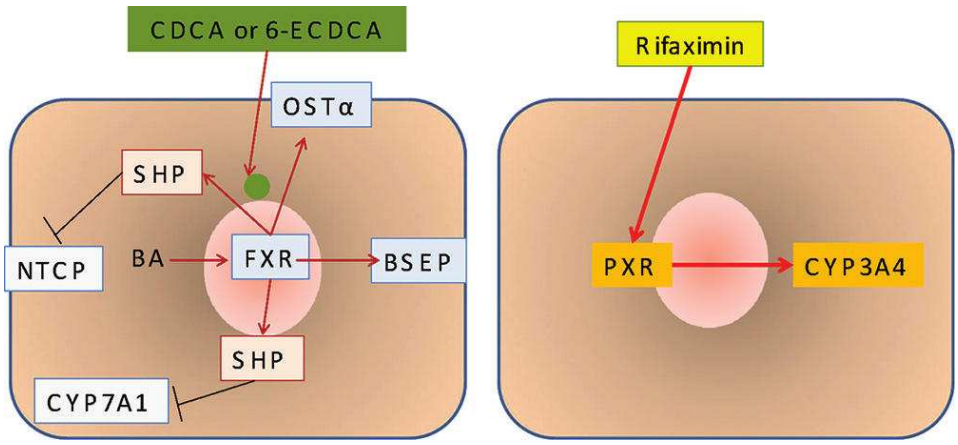


Figure 6. 5 Schematic representation of the activity of FXR and PXR on target genes.

As the results of enzymatic reactions, natural products have an intrinsic capacity to recognize and bind macromolecules, perturb their activity, and modulate biological processes. Besides their potential use as pharmaceutical drugs, natural products have and will continue to play critical roles as biological probes, essential component of today's research arsenal and useful to dissect complex biological processes and ultimately, to identify novel therapeutic targets. The observation that ≈40% of modern pharmaceuticals are derived from biological sources[444] highlights the incredible biomedical potential represented by the chemical analysis of the biodiversity of natural organisms.[444]

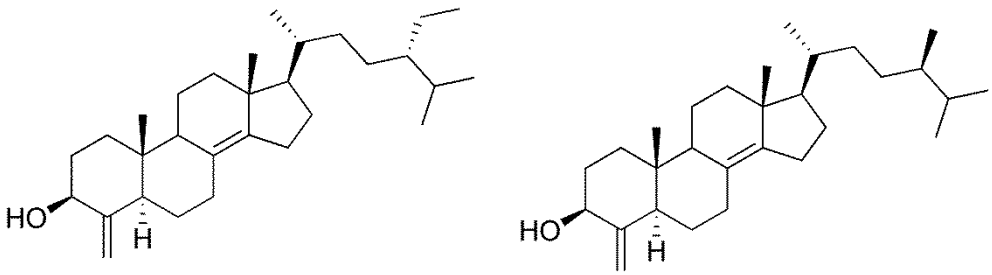
Besides, in fact, the significant contribution derived from highthroughput screenings of chemical libraries and chemical strategy based on extensive

modifications of the BAs body and side chain, only few natural FXR modulators have been described. Guggulsterone (**196** Figure 6.4) the active component of the resin extract of the tree *Commiphora mukul*,[445]and xanthohumol,[446] the principal prenylated chalcone from beer hops, are two well characterized examples of FXR modulators isolated from the vegetal realm. Recently the marine environment has also emerged as a source of human nuclear receptor ligands, and several molecules, including scalarane sesterterpenes,[447] isoprenoids,[448] and polyhydroxylated sulfated steroids,[293] have been shown to act as FXR antagonists, whereas to the best of the knowledge, no examples of marine derived FXR agonists are known.

Among natural sources, marine environment, with its vast pool of plants, animal and micro-organisms, represents a greater promise to provide original molecules for treatment of human diseases. Sponges of the genus *Theonella* have attracted the interest from the scientific community for the impressive variety of bioactive secondary metabolites with unusual structures and powerful biological effects. Representative compounds include non-ribosomal peptides exemplified by the antifungal theonellamides, a new class of sterol-binding molecules that induce membrane damage and activate Rho1-mediated 1,3βD-glucan synthesis[449] and complex polyketides such as the actin-bounding macrolide swinholide A.[450] In addition, sponges of *Theonella* genus are distinctive in producing biosynthetically unique sterols. Decodification of these non conventional steroids has allowed the identification of 24-ethylsterols endowed with potent activity towards mammalian nuclear receptors including the FXR and pregnane-X-receptor (PXR).[294]

Steroids bearing a 4-methylene group are relatively rare metabolites. They were exclusively isolated from sponges of the genus Theonella, mainly from *T. swinhoei*, and were unaccompanied by conventional steroids and are

proposed as ideal taxonomic markers for sponges of the genus Theonella.[451] Since the isolation, by Djerassi et al., of conicasterol and theonellasterol (Figure 6.6) from T. conica and T. swinhoei,[452] respectively, about 20 new 4-methylene-steroids were isolated from Theonella sponges.[451,453,454,455,456] Common structural features are a 24-methyl and/or 24-ethyl side chain, the presence of oxygenated functions at C(3), C(7), or C(15), and of a $\Delta^{8,14}$ double bond rarely replaced by a 8(14)- seco-skeleton.



Theonellasterol **198** Conicasterol **199**

Figure 6. 6 Theonellasterol and conicasterol previously isolated from Theonella species.

As a part of a systematic study on the chemical diversity and bioactivity of secondary metabolites from marine organisms collected at Solomon Islands,[457] it was found a single specimen of the sponge *Theonella swinhoei* as an extraordinary source of new metabolites. Analysis of the polar extracts afforded anti-inflammatory perthamides C-D,[458,459] solomonamides A-B,[460] and solomonsterols A-B.[295] *Theonellasterol* and Conicasterol contains a relatively rare 8(14) double bond and a biosynthetically unusual 4-methylene functionality. This unusual functional group has been proposed to biogenetically arise from a shunt in the oxidative demethylation of the 4-methyl series, through the dehydration of the primary alcohol formed in the

first oxidation of the methyl group.[452] The biological function of these 4-methylenesteroids is unknown. However, the relative abundance of these steroids in the apolar extract of *Theonella swinhoei* as well as their ability to fit in the ligand binding domain of FXR and PXR, emphasize a plausible role as putative ligands for ancestral sponge nuclear receptor(s).

6.2 Theonellasterols and Conicasterols from *Theonella swinhoei*. Novel Marine Natural Ligands for Human Nuclear Receptors

In this paragraph, it will be report the results of the chemical analysis of the less polar extracts, which resulted in the isolation and identification of theonellasterol[452] together with 10 new polyoxygenated steroids, which were named theonellasterols B-H (**200**-**206**) and conicasterols B-D (**207**-**209**) (Figures 6.7 and 6.8). These marine steroids are endowed with potent agonistic activity on the human pregnane-X-receptor (PXR) while antagonize the effect of natural ligands for the human farnesoid-X-receptor (FXR). Exploiting these properties, the theonellasterol G (**205**) was identified as the first example of PXR agonist and FXR modulator from marine origin that might have utility in treating liver disorders.

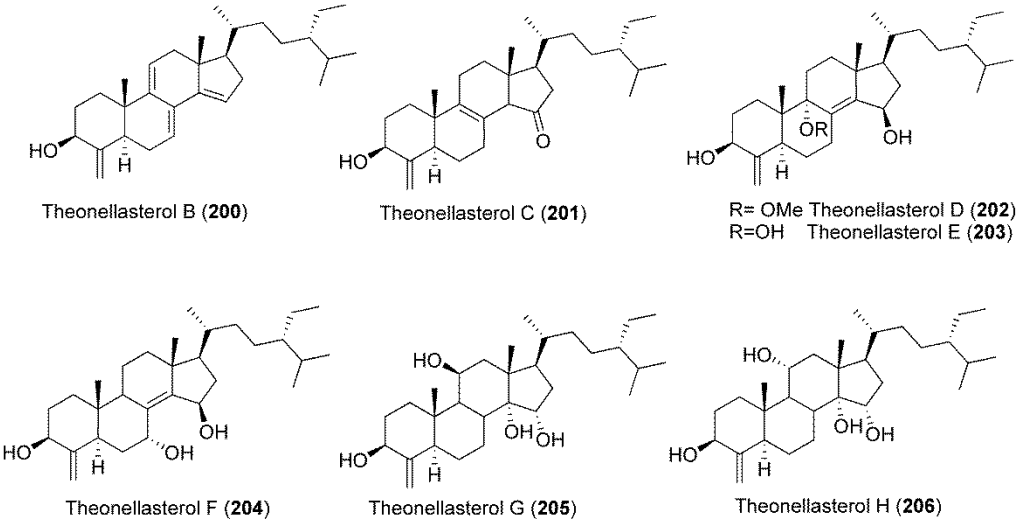
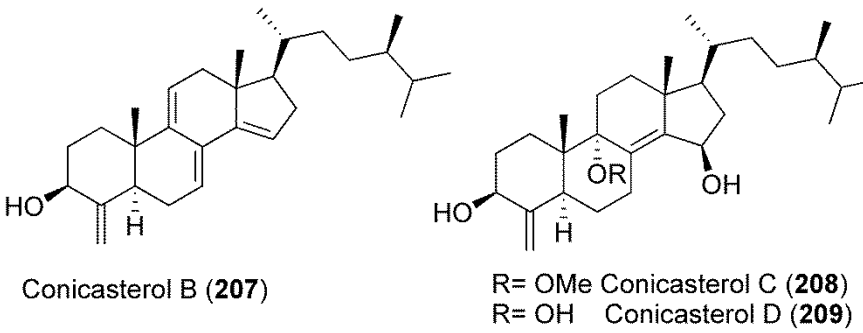


Figure 6. 7 Theonellasterols from *Theonella swinhoei*.

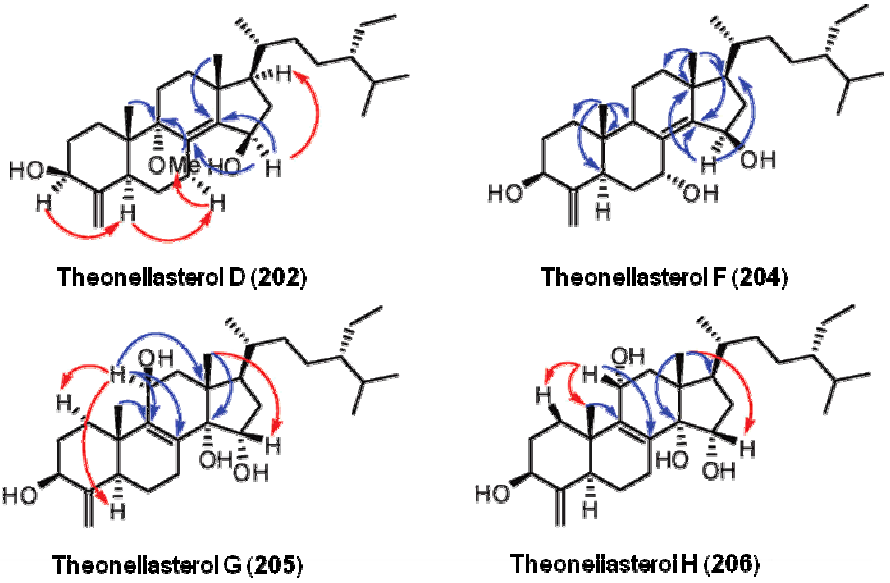


Conicasterol B (**207**)

R= OMe Conicasterol C (**208**)
R= OH Conicasterol D (**209**)

Figure 6. 8 Conicasterols from *Theonella swinhoei*.

The structures of compounds **200-209** were determined on the basis of extensive spectroscopic data (MS, ^{1}H and ^{13}C NMR, COSY, HSQC, HMBC, and ROESY) analysis (See Figure 6.9).



Theonellasterol D (**202**)

Theonellasterol F (**204**)

Theonellasterol G (**205**)

Theonellasterol H (**206**)

Figure 6. 9 Key HMBC (blue arrows) and ROESY (red arrows) correlations for theonellasterols D (**202**), F (**204**), G (**205**), and H (**206**).

As a part of continuing research directed toward the discovery of marine natural NRs ligands,[293] it was recently demonstrated that the marine sponge

Theonella swinohei produces steroids that act as NRs ligands. Solomonsterols A and B,[295] isolated from its polar extracts, are potent PXR agonists and new pharmacological anti-inflammatory leads. With this background in mind, it was investigated whether this family of hydroxylated sterols might act as modulators of two well characterized nuclear receptors, the farnesoid-X-receptor (FXR) and pregnane-X-receptor (PXR), highly expressed in the mammalian livers. For this purpose, compounds **200**-**207**[461] and compound **209** were challenged in a reporter gene assay using HepG2 cells, a human hepatoma cell line. As illustrated in Figure 6-10A, all compounds, except theonellasterol G (**205**), that partially activated FXR, failed to activate FXR at the concentrations of 10 μM. By contrast, all compounds, with the exception of theonellasterols D (**202**) and H (**207**), effectively antagonized FXR transactivation induced by CDCA (**191**, Figure 6.4) (Figure 6-10B).

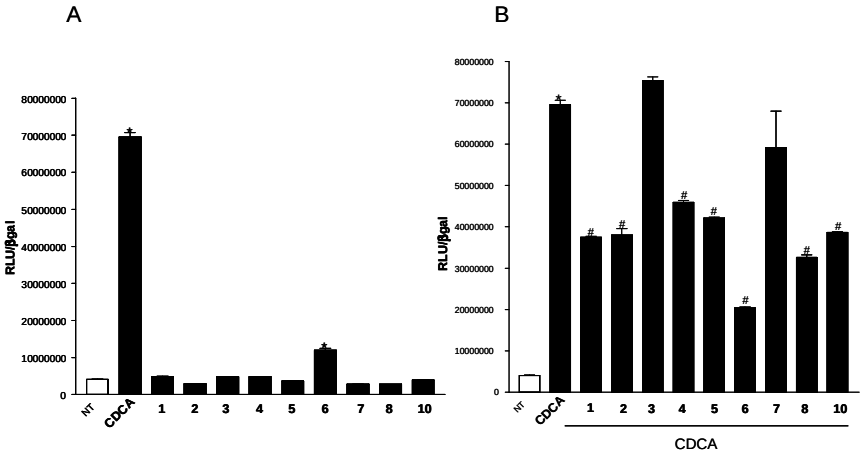


Figure 6. 10 Luciferase reporter assay performed in HepG2 transiently transfected with pSG5-FXR, pSG5-RXR, pCMV-bgal, and p(hsp27)TKLUC vectors and stimulated 18 h with (A) CDCA, 10 μM, and compounds 1-8 and compound 10, 10 μM. (B) CDCA, 10 μM, alone or in combination with compounds **200-207** and compound **209**, 50 μM. *P < 0.05 *versus* not treated (NT). #P < 0.05 *versus* CDCA (n = 4).

It is noteworthy that theonellasterol G (**205**) behaves as an antagonist in the presence of CDCA but is able to partially transactivate FXR, indicating that this agent is an FXR modulator. As shown in Figure 6.11A, at the concentration of 10 μM, all these compounds were PXR agonists.

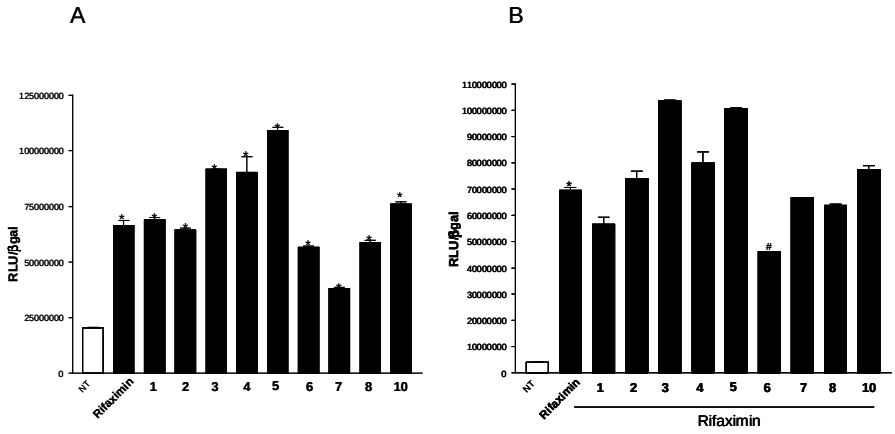


Figure 6. 11 Luciferase reporter assay performed in HepG2 transiently transfected with pSG5-PXR, pSG5-RXR, pCMV-bgal, and p(cyp3a4)TKLUC vectors and stimulated 18 h with (A) rifaximin, 10 μM, and of compounds **200**-**207** and compound **209**, 10 μM. (B) Rifaximin, 10 μM, alone or in combination with compounds 1-8 and compound 10, 50 μM. *P < 0.05 *versus* not treated (NT). #P < 0.05 *versus* rifaximin (n = 4).

Interestingly, also theonellasterol G (**205**) acted as a PXR ligand, suggesting that this compound might be considered the first FXR modulator and PXR ligand so far identified.

On the basis of the above-mentioned results, it was then analyzed, by means of molecular docking calculations, the interactions of polyhydroxylated steroids **200**-**207** and **209** with human FXR and PXR, in order to generate a structure-activity relationship and obtain information on their binding mode at atomic level. All the calculations were run by Autodock4.2 software.[56] As

shown in the Figure 6.12, compounds **200**-**207** and **209** adopt the same positioning in the FXR binding site when compared to 6-ECDCA (**192** Figure 6.4). Even if the junction between A/B rings is trans and the OH group at position 3 is in the β position with respect to the cocrystallized molecule, for all the complex models relative to compounds **200**-**207** and **209**, the fundamental hydrogen bond contacts with the two amino acids of the catalytic triad[462] (namely Tyr358 in helix 7, His444 in helix 10/11) are maintained.

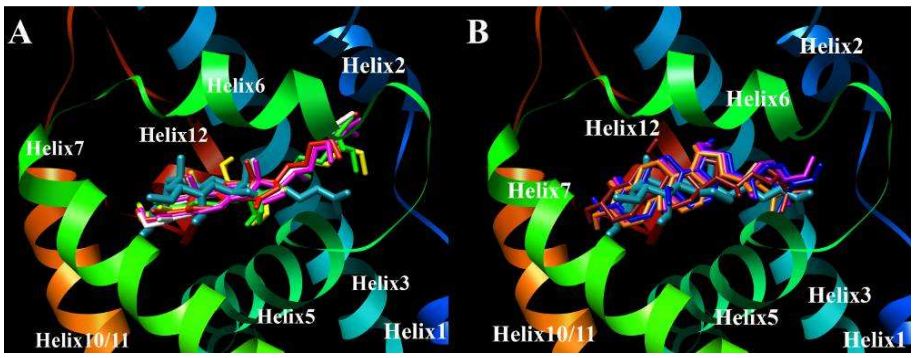


Figure 6. 12 Two different spatial arrangements of polyhydroxylated steroids **200**-**207** and **209** in the binding site of FXR (chain A of crystal structure 1OSV). (A) Superposition of 6-ECDCA (**192**, light blue) with compounds **200** (pink), **201** (red), **205** (yellow), **206** (green), and **207** (white). (B) Superposition of 6-ECDCA (**192**, light blue) with compounds **202** (blue), **203** (purple), **204** (dark red), and **209** (orange).

Consequently, it is possible to single out two different docking poses for these molecules: (a) the first one regards compounds containing the OH group at position 3 and/or the OH group at position 15 in a trans relationship (**200**, **201**, **205**-**207**, Figure 6.12A), (b) the second family includes compounds having the 3- and 15- OHs in a relative cis arrangement (**202**-**205**, and **209**, Figure 6.12B). These results are also in agreement with the biological activity of our polyhydroxylated steroids (see above). Nevertheless, for all compounds

200-**207** and **209**, the steroid nucleus is able to accommodate in the ligand binding site, establishing hydrophobic interactions with the cavity pocket formed between the Helix 2, 3, 5-7, and 10/11 (Figure 6.12).[462] Concerning the interactions regarding other positions, the ROH at position 9 of **203** and **209** forms a hydrogen bond with Ser329 (Helix 5) and this interaction is lost when the hydroxyl group is replaced by a methoxy group, as in compound **202**. Moreover, OH groups at 14α and 15α positions (compounds **205** and **206**, respectively) are able to establish hydrogen bonds with Ser329 (Helix 5), whereas the OH group at position 15β (compounds **202**-**204**, and **209**) interacts with Leu284 (Helix 3). The carbonyl group at position 15 (compound **201**) and the R-OH group at position 7 (compound **204**) do not exert further polar interactions with the FXR ligand binding site. Regarding the side chain, the presence of a methyl group at position 24 for **207** and **209** allows stronger interactions with amino acids Met287 (Helix 3), Met262 (Coil 2), and His291 (Helix 3), present on the external part of the target molecular surface, with respect to compounds **200** and **203** bearing an ethyl group with different configuration (Figure 6.13C).

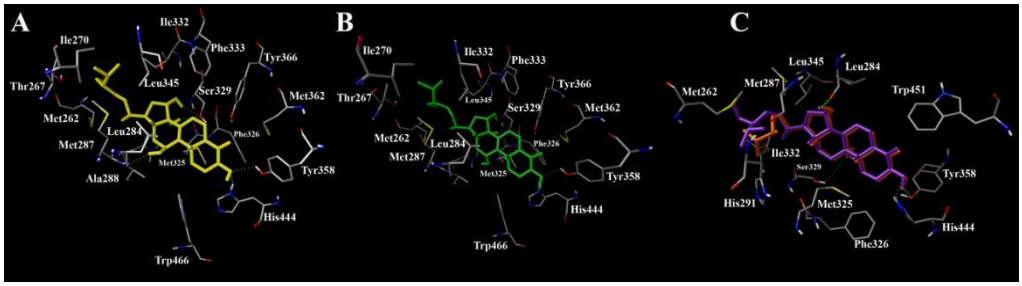


Figure 6. 13 Three dimensional models of docking pose of **205** (A, yellow), **206** (B, green), **203** (C, purple) and **209** (C, orange) with FXR (see text for details).

As previously reported,[463,464,465,466] the activation of FXR receptor by bile acids and bile acid analogues is affected by simultaneous presence of two R-OH at position 3 and 7 and by a good balance between hydrophobic and hydrophilic substituents at the R and β face of the nucleus. Moreover, Fujino et al.[463] have demonstrated that the introduction of β-alkyl and/or β-hydroxyl groups at 3 and/or 7 positions decreases the ligand potency. This kind of interactions are missing for the described compounds, and in fact all compounds, except theonellasterol G (**205**), failed in the activation of FXR at the concentration of 50 μM (see above). In particular, as depicted in Figure 6.13A, the β-OH group at position 11 of compound **205** is involved in an additional hydrogen bond with Leu284 (Helix 3) with respect to its epimer **206** (Figure 6.13C), where this interaction is lost. Moreover, even if also **203** and **209** are able to interact with Ser329 (Helix 5) and Leu284 (Helix 3), they present a different spatial arrangement (Figures 6.12 and 6.13C) with respect to theonellasterol G (**205**), lacking some hydrophobic interactions with the amino acids of FXR ligand binding pocket (e.g., Trp466, Ile270, Thr267, Leu345, Figure 6.13). So, the described docking calculations point out that the simultaneous interactions[467] of 6 with Helix 7 (Tyr358) and Helix 3 (Leu284) and its optimal hydrophobic contacts with the ligand binding domain compared to compounds **206** (only interacting with Helix 7) and **203** and **209**, respectively (Figures 6.12 and 6.13), cause a great difference in activity between these sterols, suggesting that a correct orientation of the OH group at position 11 and the hydrophobic contacts with the receptor are critical for the FXR modulator activity.

For what concerns the PXR[468] agonist activity, the first observation suggested by molecular docking analysis regards the positioning of compounds **200**-**207** and **209** in occupying in the region of PXR's expansive

ligand binding pocket. As already reported,[468,469,470,471,472] this region is formed by hydrophobic (Phe251, Phe288, Phe429, Cys284, Leu206, Leu209, Leu324, Leu411, Met243, Met425, Trp299), polar, and charged (Ser247, Ser208, His407, Asp205, Arg410, Gln285) amino acids; particularly crucial is the interaction of the reported sterols with the Ser247, previously reported to be involved in key interactions for the PXR agonist activity.[470,469,468,471,472] In fact, this interaction is associated to the activity modulation observed in this series of compounds that is relative to substitution pattern of steroidal nucleus. It is noteworthy that 15α-OH, 15β-OH, and 15-keto substitutions are all possible H-bond acceptor sites interacting with H-bond donor Ser247. The further interactions of OH group at position 7 with Ser247 for compound **204**, of OH group at position 9 with Gln285 for **203**, and of OMe group at position 9 with Cys284 for **202**, may be responsible for the increased activity observed for these derivatives (Figure 6.14).

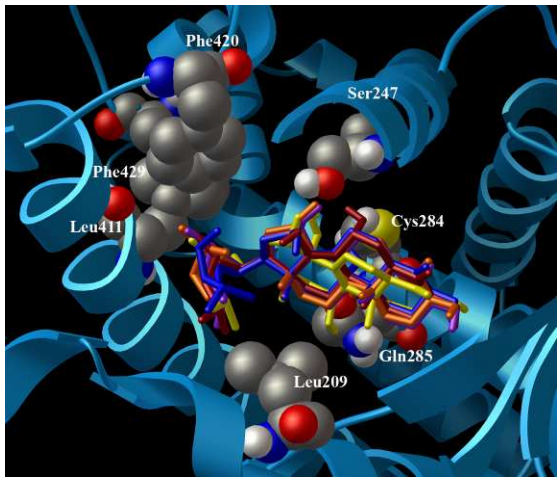


Figure 6. 14 Three-dimensional model of the most representative polyhydroxylated steroids **202** (blue), **203** (purple), **204** (dark red), **205** (yellow), and **209** (orange) with the PXR binding pocket.

Finally, compounds **207** and **209**, featuring a methyl group at position 24, are endowed with a weaker agonist activity with respect to **200** and **203**, bearing an ethyl group with different configuration. Even though compounds **207** and **209** are able to interact by their side chain with Leu209, the obtained models suggest that their decreased activity is due to lacking interactions with Phe420, Leu411, and Phe429 (Figure 6.14).

In summary, here it was described a novel class of FXR/PXR modulators of marine origin, and the above results pave the way to design new selective and potent modulators for human nuclear receptors. In particular, this series of marine sterols are endowed with potent PXR agonist activity and further studies are ongoing to define their potential applications in biomedical settings, such as liver toxicity induced by hepatotoxic drugs or sepsis. Of relevance, because both FXR and PXR exert antiinflammatory effects in the intestine, a dual ligand holds potential in the treatment of inflammatory bowel diseases. Moreover, docking analysis validated the experimental biological results and allowed rationalization of the peculiarity activity of theonellasterol G (**205**).

All together, these results suggested that 4-methylene sterols, proposed as ideal taxonomic markers for Theonella sponges, are FXR antagonists and PXR agonists.

6.2.1 Computational Details

Molecular docking calculations were performed by Autodock4.2 software[56] on quad-core Intel Xeon 3.4 GHz. A grid box size of 94 x 96 x 68 for FXR receptor, and of 90 x 106 x 92 for PXR rceptor with spacing of 0.375 Å between the grid points and centered for FXR at 20.689 (x), 39.478 (y), and 10.921 (z), and for PXR at 14.282 (x), 74.983 (y), and 0.974 (z) covering the

active site on the two targets surface was used. The Lamarckian genetic algorithm with an initial population of 600 randomly placed individuals, a maximum number of 5.0 $x10^6$ energy evaluations, and a maximum number of 6.0x 10^6 generations were taken into account for dockings by Autodock4.2 software. A mutation rate of 0.02 and a crossover rate of 0.8 were used. Results differing by less than 3.5 Å in positional rootmean-square deviation (rmsd) were clustered together and represented by the result with the most favorable free energy of binding. Illustrations of the 3D models were generated using the Chimera[174] and the Python software.[165]

6.3 Conicasterol E, a Small Heterodimer Partner Sparing Farnesoid X Receptor Modulator Endowed with a Pregnane X Receptor Agonistic Activity, from the Marine Sponge *Theonella swinhoei*

Secondary metabolites from marine organisms collected at Solomon Islands,[457] were found a single specimen of the sponge *Theonella swinhoei* as an extraordinary source of NRs steroidal ligands (Figure 6.15). Analysis of the polar extracts afforded solomonsterols A and B,[295] two potent PXR agonists and new leads in the treatment of immune-driven inflammatory bowel diseases,[473] whereas analysis of the apolar extracts allowed the isolation of a small library of 4-methylene steroids.[294,474]

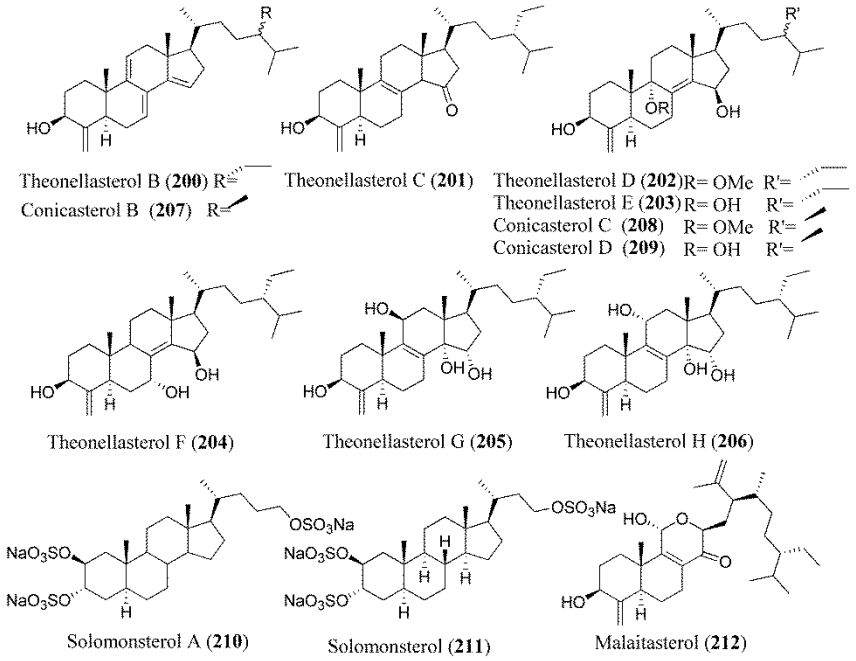
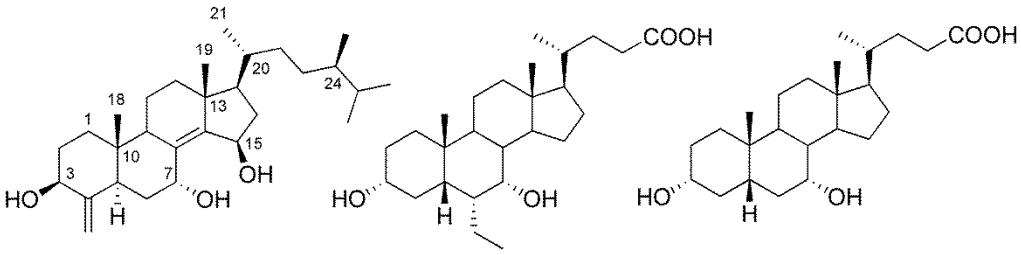


Figure 6. 15 Nuclear receptors ligands previously isolated from the marine sponge *Theonella swinhoei*.

Importantly, even if the junction between A/B rings is trans and the OH group at position 3 is in the β position with respect to the natural ligand CDCA, the described docking studies demonstrated that all these compounds could be accommodated in the ligand binding domain of FXR, establishing hydrophobic and hydrogen bond contacts with the catalytic triad.[462] Furthermore, within this series, it was have demonstrated that a methyl group at position 24 (conicasterols B−D in Figure 6.15) allows stronger interactions with a shallow groove on the FXR molecular surface with respect to the ethyl group in theonellasterol-like compounds (theonellasterols B−H in Figure 6.15).

Here it will be described the rationalization at atomic level of the putative binding mode of conicasterol E (**213**),[357] a 7α,15β-dihydroxyconicasterol analogue (Figure 6.16), as the first example of an SHP sparing FXR modulator endowed with PXR agonistic activity from the same specimens of *Theonella swinhoei*.

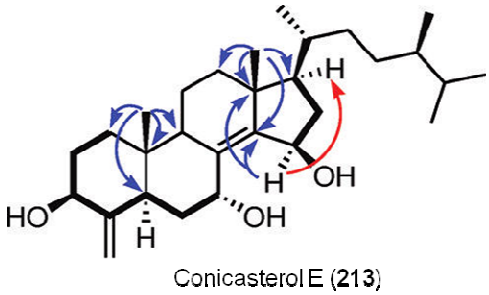


Conicasterol E (**213**) 6-ECDCA (**192**) CDCA (**191**)

Figure 6. 16 Conicasterol E (**213**), the first example of SHP-sparing marine FXR modulator.

The structural characterization of conicasterol E was obtained from a detailde analysis of the COSY, TOCSY, ROESY, HSQC, and HMBC experiments. The configuration at C-24 was determined by comparison of ^{13}C

NMR data with literature data for epimeric steroidal side chains.[475] Retrospective analysis of NMR data of theonellasterol F (Figure 6.15), previously isolated from the same sponge,[475] indicated a strong resemblance with conicasterol E (Figure 6.17).



Conicasterol E (**213**)

Figure 6. 17 COSY connectivities (bold bonds) and HMBC (blue arrows) and ROESY correlations (red arrows) for conicasterol E (**213**).

Conicasterol E (**213**) was tested in vitro using an hepatocarcinoma cell line (HepG2 cells) transfected with FXR, RXR, β-galactosidase expression vectors (pSG5FXR, pSG5RXR, and pCMV-βgal) and with p(hsp27)TKLUC reporter vector that contains the promoter of the FXR target gene heat shock protein 27 (hsp27) cloned upstream the luciferase gene. As shown in Figure 6.18, conicasterol E (**213**) activates FXR in transactivation assay. However, in contrast to CDCA (**191**), the results of these experiments demonstrate that conicasterol E (1) activates FXR with a bell shaped concentration−response curve, the agonistic activity being partially reduced at 50 μM. At 10 μM, 1 was as effective as CDCA (**191**) in transactivation assay but significantly less potent than the synthetic FXR ligand 6-ECDCA (**192**) and GW4064 (data not shown, **193**, Figure 6.4).

It was then tested whether conicasterol E (**213**) exerted any antagonistic activity against FXR. Because the above-mentioned results revealed a bell-shaped curve in the concentration−response effect of compound 1 in transactivating FXR, the effects was tested of this agent at 10 μM (data not shown) and 50 μM and found that compound 1 was devoid of any antagonistic activity when coadministered with CDCA (**191**) (Figure 5B) and 6-ECDCA (data not shown) to HepG2 cells.

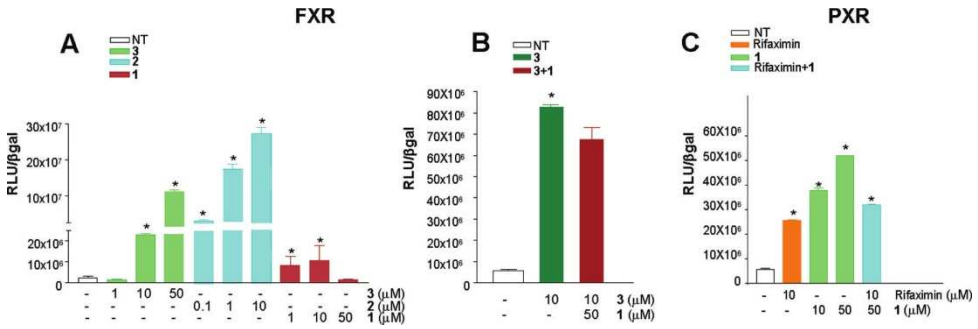


Figure 6. 18 (A) Relative potency of FXR activation by CDCA (**191**), 10 μM, conicasterol E (1), 10 μM, and 6-ECDCA (2), 1 μM, as measured by transactivation assay in HepG2 cells. (B) Conicasterol E (1), 50 μM, does not revert the effect of CDCA (**191**), 10 μM, on FXR transactivation in HepG2 cells. (C) Relative potency of PXR activation by rifaximin, 10 μM, and conicasterol E (1) alone, 10 μM, or in combination, 50 μM. Data are the mean ± SE of four experiments: (∗) $P < 0.05$ *versus* untreated cells (NT).

In addition to an FXR agonistic activity, conicasterol E (**213**) effectively induced PXR expression, being as effective as rifaximin in inducing PXR transactivation (Figure 6.18C). Thus, conicasterol E (**213**) is a dual FXR and PXR agonist. To further characterize the biological activity of the conicasterol E (**213**), the effect of this agent on the expression of canonical FXR and PXR target genes in hepatocytes was examined, and as shown in Figure 6.19, it was found that exposure to conicasterol E slightly increased the expression of

OSTα and BSEP mRNAs (two FXR regulated genes) and the expression of CYP3A4 mRNA (a PXR-regulated gene), while no effect was observed on SHP mRNA expression. In addition, in contrast to CDCA, **213** failed to repress CYP7A1. Thus, while the expression of this gene was reduced by 30% by CDCA, exposure to conicasterol E (**213**) increased CYP7A1 mRNA by 2- to 3-fold. These data are further evidence that in HepG2 cells repression of CYP7A1 by FXR is indirect and requires induction of SHP.

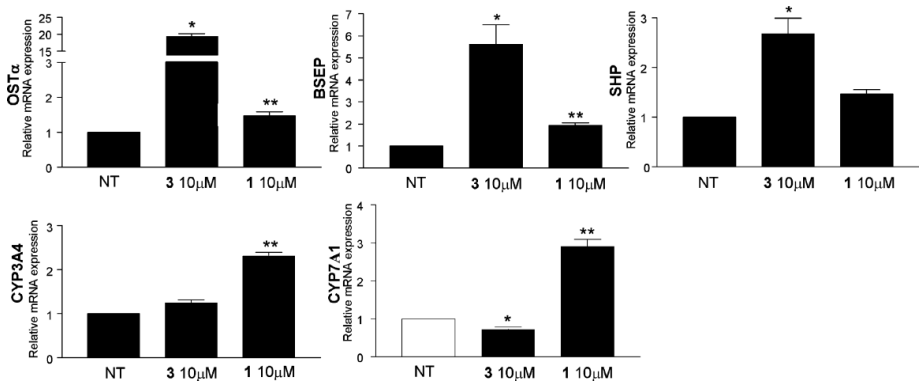


Figure 6. 19 RT-PCR analysis of effects of CDCA (**191**), 10 μM, and conicasterol E (**213**), 10 μM, on expression of FXR and PXR-regulated genes in HepG2 cells. Conicasterol E (**213**) does not induce SHP, whereas it induces the expression of CYP7A1. Data are the mean ± SE of four experiments: (∗) $P < 0.05$ *versus* untreated cells (NT); (∗∗) $P < 0.05$ *versus* CDCA alone.

Further on, when administered in combination with a concentration of CDCA of 10 μM, conicasterol E exerted an additive effect with CDCA on the expression of OSTα and BSEP while no further changes were observed in the expression of SHP (Figure 6.20).

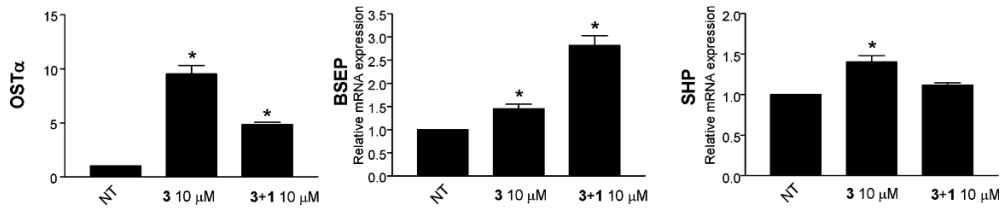


Figure 6. 20 RT-PCR analysis of effects of CDCA (**191**), 10 μM, alone or in combination with conicasterol E (**213**), 10 μM, on expression of FXR-regulated genes in HepG2 cells. Conicasterol E (**213**) does not induce SHP even when cells were co-incubated with 3, while the association of the two agents partially attenuated the expression of OSTα but increased the expression of BSEP. Data are the mean ± SE of four experiments: (∗) $P < 0.05$ *versus* untreated (NT).

Taken together, these data highlight that conicasterol E (**213**) is a FXR modulator whose potency on selective target genes is very close to that of the endogenous mammalian ligand CDCA (**191**) and lower than that of the synthetic agonist 6-ECDCA (**192**). Interestingly, conicasterol E (**213**) failed to stimulate SHP even when coadministered in combination with CDCA (**191**). Finally, analysis of CYP3A4 expression, shown in Figure 6.21, demonstrated that conicasterol E (**213**) has no antagonistic effects on expression of CYP3A4 mRNA induced by rifaximin, a potent PXR agonist.

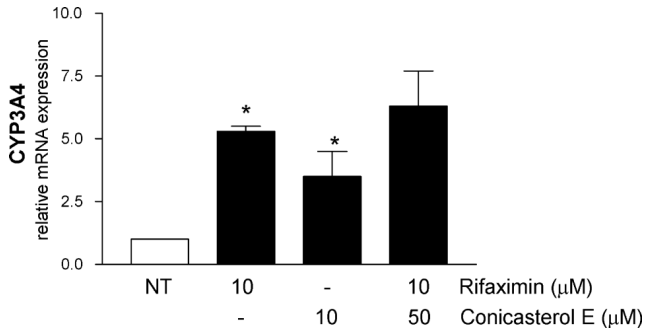


Figure 6. 21 Activation of CYP3A4 by the PXR agonist rifaximin, 10 μM, is not modulated by conicasterol E (**213**), 50 μM. Data are the mean ± SE of four experiments: (∗) $P < 0.05$ *versus* untreated (NT).

As reported in the previous paragraph,[294] 4-methylene sterols isolated from *Theonella swinhoei* are able to modulate in different ways the FXR activity depending on the steroid skeleton substitutions. On this basis and in order to describe at atomic level the interactions of **213** with FXR macromolecule, molecular docking calculations were performed using Autodock 4.2 software.[56] As shown in Figure 6.22A, the FXR binding site, located between helixes 2, 3, 5−7, and 10/11, is occupied by **213**, and as previously reported,[462] the *β*-OH groups at positions 3 and 15 and the trans junction between A/B rings cause a different positioning with respect to the cocrystallized molecule 6-ECDCA (**192**). In particular (Figure 6.22B) conicasterol E (**213**), compared to the synthetic agonist 6-ECDCA (**192**), is able to interact with two amino acids of the catalytic triad formed by Tyr358 in helix 7, His444 in helix 10/11, Trp466 helix 12, responsible for the activation of FXR.33

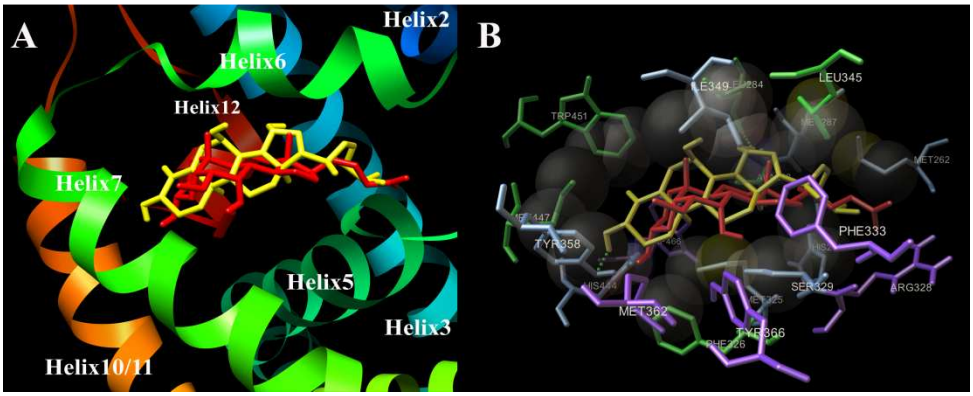


Figure 6. 22 (A) Superimposition of **213** (yellow) with 6-ECDCA (**192**) (red) in the binding pocket of FXR (PDB code 1OSV). (B) Amino acids interacting with 6-ECDCA (red) are depicted in purple. Amino acids interacting with **213** (yellow) are depicted in green, and amino acids interacting with both molecules are depicted in light blue.

Specifically, the 3-OH group at β position forms a hydrogen bond with Tyr358 in helix 7, while the trans junction between the A/B ring allows a hydrophobic interaction with His444 (helix 10/11). The influence of the side chain on the FXR binding, it was already described,[294] and in fact, the methyl at position 24 of conicasterol E (**213**) (yellow, Figure 6.22B) relating to 6-ECDCA (**192**) (red, Figure 6.22B) is able to simultaneously interact with the Met262 (coil 2), His291 (helix 3), and Met287 (helix 3) present on the shallow groove of the FXR molecular surface protruding toward the solvent. Moreover, the OH at position 15β in **213** forms an additional hydrogen bond with the CO of Leu284 (helix 3), and the steroid skeleton is in close contact with Leu345, Ala288, Met447, Phe326, and Trp451 relating to the 6-ECDCA (**192**). On the other hand, the OH at 7α position does not seem to exert further polar interactions with the FXR binding site. In summary, conicasterol E (**213**) presents a different spatial arrangement relating to the cocrystallized molecule 6-ECDCA; however, the reoprted docking calculations point out that its simultaneous and efficient hydrophobic and hydrophilic additional interactions with the receptor binding site might be responsible for its agonist activity on FXR. Further on, the exclusive amino acid interactions exerted by conicasterol E (**213**) might support the notion that the compound is an FXR modulator endowed with the ability to activate OSTα and BSEP without effect on SHP expression.

6.3.1 Computational Details

Molecular docking calculations were performed by Autodock 4.2 software[56] on quad-core Intel Xeon 3.4 GHz, using a grid box size of 94 × 96 × 68, with spacing of 0.375 Å between the grid points and centered at 20.689 (x), 39.478 (y), 10.921 (z), covering the active site of the FXR.[462] To achieve a

representative conformational space during the docking studies and for taking into account the variable number of active torsions, 10 calculations the ligand. The Lamarckian genetic algorithm (LGA) was employed for docking calculations, choosing an initial population of 600 randomly placed individuals. The maximum number of energy evaluations and of generations was set up to 5×10^6 and to 6×10^6, respectively. Results differing by less than 3.5 Å in positional rootmean-square deviation (rmsd) were clustered together and represented by the result with the most favorable free energy of binding. Illustrations of the 3D models were generated using the Chimera[174] and the Python software.[165]

6.4 4-Methylenesterols from *Theonella swinhoei* Sponge are Natural Pregnane-X-Receptor Agonists and Farnesoid-X-Receptor Antagonists that Modulate Innate Immunity

Marine sponges of the genus *Theonella* have attracted a great interest from the scientific community for the impressive variety of bioactive secondary metabolites with unusual structures and powerful biological activity. In a recent report Crews *et al.*[476] pointed on the existence of at least three phenotypes of *Theonella swinhoei*,[477] and observed that the morphology of the sponge has some influence on the chemical composition.

Studying the biodiversity associated to the marine organisms collected at Solomon Islands, three *Theonella swinhoei* specimens was analyzed, two of them, whose macroscopic morphology could be ascribed to phenotype I, have been subjected by extensive chemical investigation that disclosed those specimens as invaluable sources of new secondary metabolites. These two sponges were found to contain swinholides,[478] theonellamide A, antiinflammatory new perthamide derivatives,[458,459] a new class of cyclic peptides, solomonamides,[460] and new truncated-chain sulfated steroids, solomonsterols.[295,473]

With surprise, none of the above metabolites was found in the third taxonomic voucher, ascribable to phenotype III, even if field collection data (date, site, depth) were very similar. On the other hand, when the apolar extracts was investigated, in all three specimens, a great variety of polyhydroxysteroids all characterized by a 4-methylene functionality was found.

The isolation of several polyoxygenated 4-methylenesteroids[293,294] from the two “PKS and NRPS producing” specimens (phenotype I) was recently reported, mainly possessing a 24-ethyl side chain (Figure 6.23) and their

pharmacological evaluation as modulators of two well known nuclear receptors, FXR and PXR.

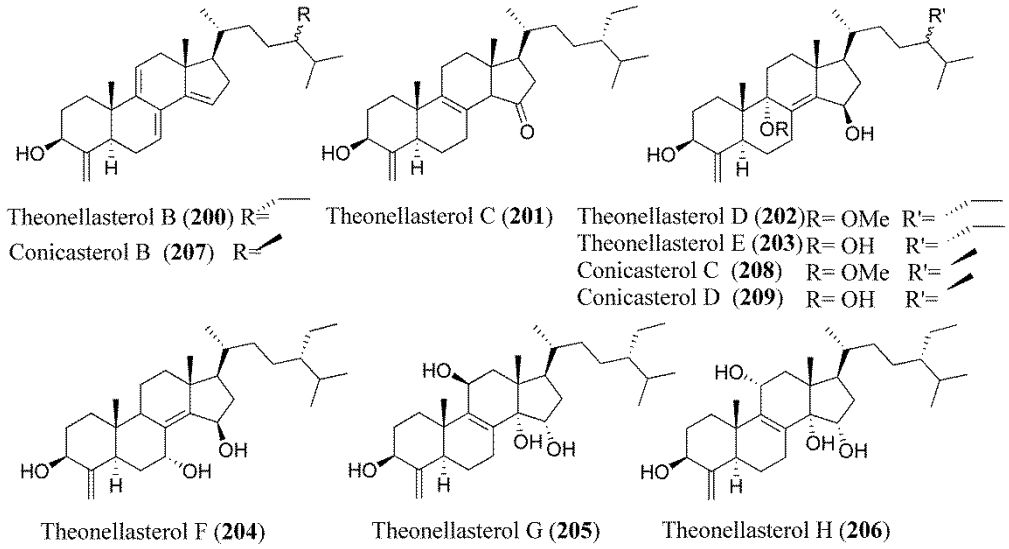
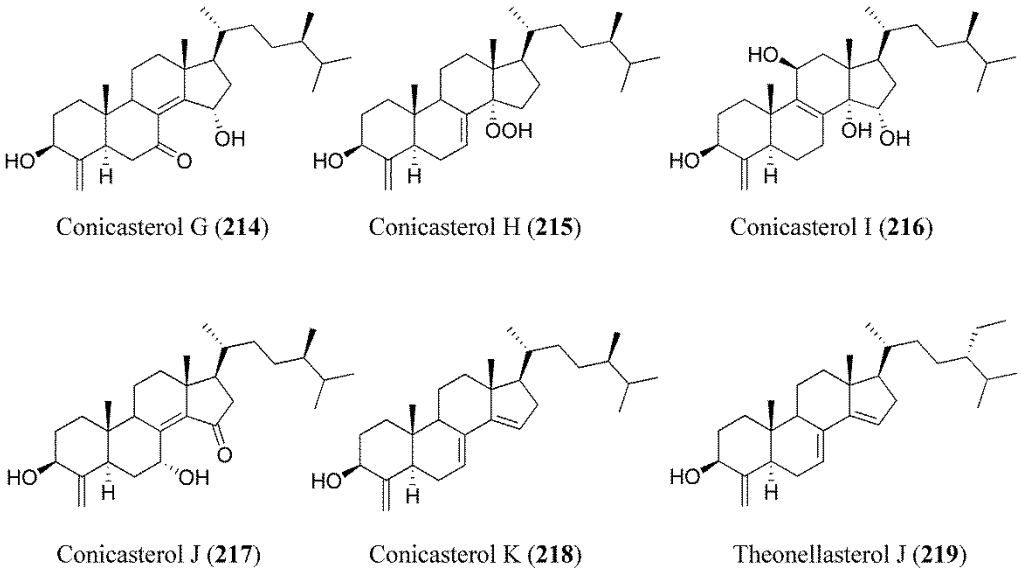


Figure 6. 23 Theonellasterols and conicasterols previously isolated from *Theonella swinhoei*.

The analysis of the apolar extract of the third specimen evidenced again the presence of a family of polyhydroxy steroids but in this case most of them featuring a 24-methyl side chain (Figure 6.24, conicasterols G-K, **214**-**218**), including known conicasterol (**199**)[479] with its 7- and 15-hydroxy derivatives (**221** and **222**),[480] dehydroconicasterol (**220**),[481] and the 8-14-seco, swinhosterol B (**223**)[482] (Figure 6.25). Trace of a new theonellasterol-like derivative, theonellasterol J (**219**), was also isolated.

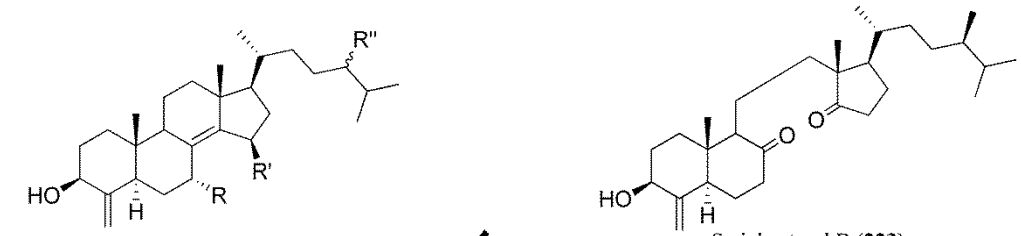
The structural characterizations of compounds were determined by detailde analysis of the COSY, TOCSY, ROESY, HSQC, and HMBC experiments (See Figure 6.26). The 24*R* stereochemistry of the side chains of conicasterols G-K (**214**-**219**) and 24*S* of theonellasterol J (**219**) were determined by comparison of ^{1}H [483] and ^{13}C chemical shifts.[484]



Conicasterol G (**214**) Conicasterol H (**215**) Conicasterol I (**216**)

Conicasterol J (**217**) Conicasterol K (**218**) Theonellasterol J (**219**)

Figure 6. 24 New 4-methylene-sterols isolated from *Theonella swinhoei*, phenotype III.



Conicasterol (**199**)	R=H	R'=H	R"=
Dehydroconicasterol (**220**)	R=H	R'=H	R"=
7α-hydroxyconicasterol (**221**)	R=OH	R'=H	R"=
15β-hydroxyconicasterol (**222**)	R=H	R'=OH	R"=

Swinhosterol B (**223**)

Figure 6. 25 Known 4-methylene-sterols isolated from *Theonella swinhoei*, phenotype III.

The structures of compounds **199, 220-223** as conicasterol, dehydroconicasterol, 7α-hydroxyconicasterol, 15β-hydroxyconicasterol, and swinhosterol B, respectively, were deduced by NMR and mass analysis and by direct comparison of their chemical shift data with existing literature values.[479-482]

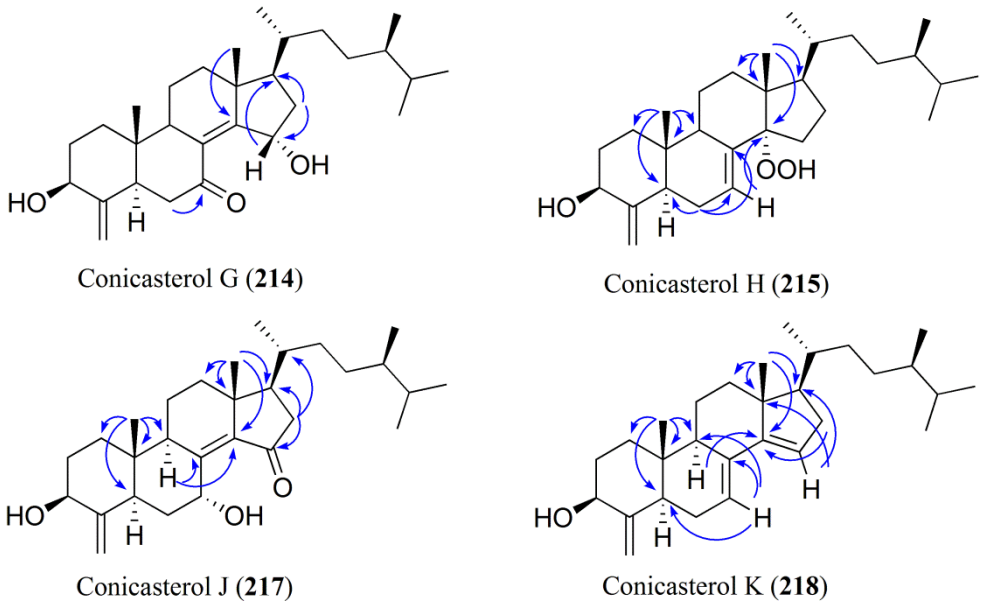


Figure 6. 26 Key HMBC correlations for conicasterols G (**214**), H (**215**), J (**216**) and K (**218**).

As mentioned before, a deep investigation of steroidal composition of *Theonella swinhoei* phenotype I has demonstrated that these specimens produce almost exclusively 24-ethylsterols (Figure 6.23), which were already reported, and most of them endowed with potent activity towards to well know nuclear receptors, FXR and PXR.[294,474] Within this library of compounds (Figure 6.23), it was also established that the presence of a methyl group at position 24 (conicasterols B-D) allows stronger interactions in the external part of FXR molecular surface, with respect to compounds bearing an ethyl group (theonellasterols B-H).

Therefore, following this recent acquisition, all compounds isolated from *Theonella swinhoei* phenotype III (Figures 6.24-25) were tested in vitro, using an hepatocarcinoma cell line (HepG2 cells) transfected with FXR, RXR, β-galactosidase expression vectors (pSG5FXR, pSG5RXR and pCMV-βgal),

and with p(hsp27)TKLUC reporter vector that contains the promoter of the FXR target gene heat shock protein 27 (hsp27) cloned upstream the Luciferase gene (Figure 6.27). HepG2 cells were stimulated with 10 µM of compounds 1-11 with or without 10 µM CDCA. As shown in Figure 6.27, even if none of these compounds appears to be an FXR agonist in the transactivation assay, several compounds, tested at the concentration of 50 µM, showed a slight inhibitory activity against FXR transactivation induced by 10 µM of CDCA with conicasterol H (**215**), conicasterol J (**214**), swinhosterol B (**223**) and the parent conicasterol (**199**) the most potent of this series.

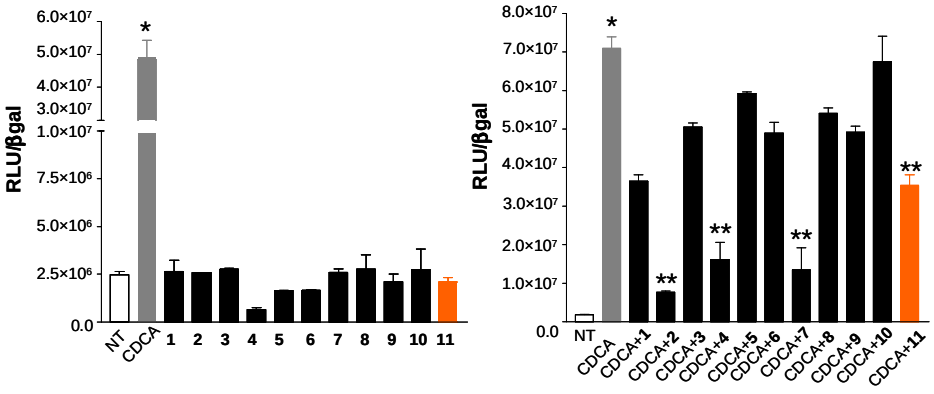


Figure 6. 27 Luciferase reporter assay performed in HepG2 transiently transfected with pSG5-FXR, pSG5-RXR, pCMV-βgalactosidase, and p(HSP27)-TK-Luc vectors and stimulated 18 h with (A) CDCA (10 µM) and compounds **214-223** (10 µM). (B) CDCA (10 µM) alone or in combination with compounds 1-11 (50 µM). *P < 0.05 *versus* not treated (NT). **P < 0.05 *versus* CDCA (n = 4).

In addition to an FXR antagonistic activity, many of these steroids effectively induced PXR expression with compounds **199** and **218-219** and compounds **221-223** being as effective as rifaximin in inducing PXR transactivation (Figure 6.28).

Because these data indicated that 4-methylenesterols from Theonella hold the potential to act as PXR agonists and FXR antagonists, and they hold potential in treating human disorders, the activity was further characterized of a select sample against expression of a whole family of nuclear receptors. Among all molecules showing this dual behavior, swinhosterol B (**223**) was selected because it is a potent PXR agonist endowed with a robust FXR antagonism whereas conicasterol (**199**) was excluded due to the observed citotoxicity on HepG2 cells.

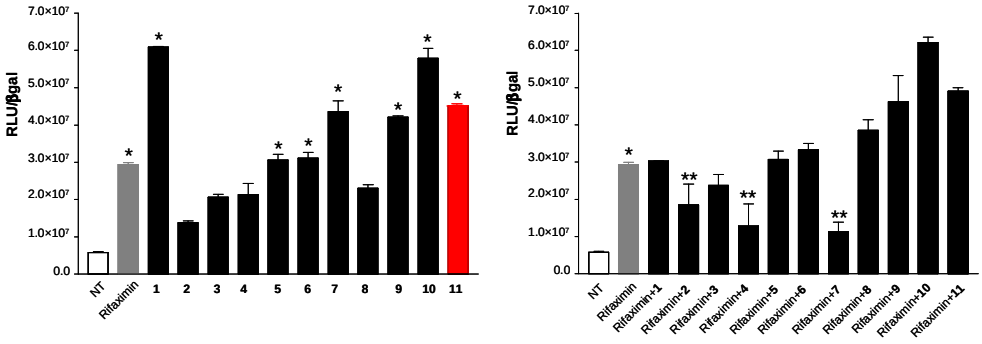


Figure 6. 28 Luciferase reporter assay performed in HepG2 transiently transfected with pSG5-PXR, pSG5-RXR, pCMV-βgalactosidase, and p(CYP3A4)-TK-Luc vectors and stimulated 18 h with (A) rifaximin (10 μM) and with 1-11 (10 μM). (B) Rifaximin (10 μM) alone or in combination with 1-11 (50 μM). *P < 0.05 *versus* not treated (NT).**P< 0.05 *versus* rifaximin (n = 4).

By profiling the expression of 86 genes using a microarray system (Figure 6.29), it was found that exposure of human HepG2 cells to 10 μM of swinhosterol B (**223**) had no effect on the expression of the waste majority of these genes.

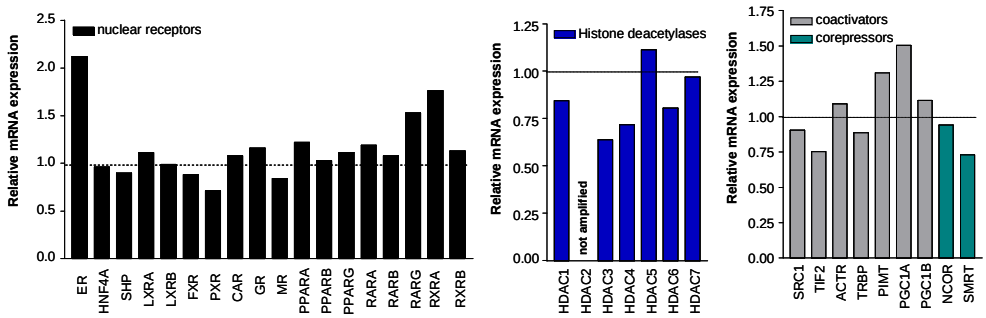


Figure 6. 29 RT^2 profile PCR array analysis showing the relative mRNA expression of (A) various nuclear receptors, (B) histone deacetylases and (C) transcriptional co-regulators (co-activators and co-repressors) following stimulation of HepG2 cells with 10 μM swinhosterol B (11). Data are the mean ± S.E. of three experiments.

However, **223** induced the expression of ESR1, NC0A6, NR1D2, NR2C1, PPARGC1A, PSMC3, PSMC5, RARG, RORA and RXRA. None of these genes were known target for FXR or PXR. Because the above mentioned data suggested that swinhosterol B (**223**) is a PXR agonist and FXR antagonist, it was then investigated whether this agent modulate immune response of macrophages. Both PXR and FXR have been shown to exert immunomodulatory effects on macrophages. Results of these experiments demonstrate that swinhosterol B (**223**) effectively counteracts stimulation of hPXR-macrophages caused by LPS. Indeed, at the concentration of 10 μM, swinhosterol B (**223**) causes a robust attenuation of TNFα, IL-1β and IL-6 generation induced by LPS.

To further investigate if the FXR antagonistic profile was maintained in macrophages expressing the murine PXR, we have challenged spleen-derived monocytes with swinhosterol B (**223**) in the presence of LPS. In this context swinhosterol B (**223**) fails to inhibit cytokine generation caused by LPS indicating that this compound activates selectively the human PXR and fails to

recognize the murine PXR. It was investigated whether these immunological effects extended to cells of adaptive immunity. For these purposes CD4+ T cells were prepared from the spleen on transgenic mice expressing the hPXR and wild type mice expressing the murine PXR. Data shown similarly to solomonsterol A, a sponge steroid endowed with potent PXR agonistic activity, swinhosterol B (**223**) had no effect on generation of IFNγ, but potently stimulated the production of IL-10, an anti-inflammatory cytokine from cells isolated from hPXR transgenic mice. The relative potency of the effect of swinhosterol B (**223**) at the concentration of 10 μM was comparable to that of solomonsterol A, 10 μM, and to that of T cell activator and mitogen concanavallin A, 2 μg/ml.

Of interest, the ability to induce IL-10 mRNA was lost in CD4+ T cells prepared from wild type mice, i.e. mice expressing the murine PXR. Thus swinhosterol B (**223**) induces generation of anti-inflammatory IL-10 via induction of PXR.

The FXR antagonistic effect of swinhosterol B (**223**) was finally assayed in HepG2 cells stimulated with the FXR agonist CDCA. Stimulation of HepG2 with **223** was itself sufficient to inhibit FXR target genes such as OSTα, BSEP and SHP. In addition, when **223** was combined with CDCA, a robust down-regulation of FXR target genes mediated by CDCA was reported. The antagonistic activity of **223** was maintained also for CYP7A1. Indeed, this gene was down-regulated by CDCA and this inhibition was significantly reversed by **223** co-treatment.

In order to rationalize the binding mode of the 4-methylenesterols isolated from *Theonella swinhoei* sponge on FXR receptor, molecular docking experiments were performed using Autodock 4.2 software.[56] In particular, here it was report the detailed binding mode analysis of sterols with new

substitution patterns (**214** and **217**) and/or with different or original nuclei (**199**-**223**) to obtain useful information for tracing a detailed and accurate profile of new potential steroid-based FXR antagonists.

Considering the possible hydrophilic interactions with the FXR ligand binding domain, the attention on **214**, **217** and **223** was focused presenting both carbonyl and hydroxyl groups in different position of tetracyclic nucleus. In this way, comparing docking results, it is possible to rationalize the influence of the different H-bond donor and acceptor pattern on steroids biological activity. For these reasons, it was excluded from the reported discussion the detailed analysis of compounds with only one OH group at C-7 as in compound 9 or at C-15 as in compound **222** or, as the case of conicasterol I (**216**), presenting the same nucleus of theonellasterol G, already described.[294]

For what concern the other compounds, comparing the diverse rigidity of the nucleus of conicasterol (**199**) *vs* conicasterol K (**218**), and swinhosterol B (**223**) combined with the different substitution at the C-24 (**219**, **199** and **220**), it was tried to rationalize the influence of the hydrophobic interactions on FXR antagonist behavior of marine steroids. On this basis, the compounds of the series (**214**, **199**, **217**-**220**, and **223**) are able to interact by the OH at C-3 with FXR catalytic triad (namely Tyr358 in Helix 7, His444 in Helix 10/11, Trp466 Helix12)[462] (Figure 6.30A), that, as reported in previous studies[293,294,474] represents one of the principal factor responsible of activity on the nuclear receptor. In particular conicasterol (**199**) and swinhosterol B (**223**), both able to antagonize CDCA, form hydrogen bonds with Tyr358 (Helix 7) and His444 (Helix 10/11); on the other hand conicasterol G (**214**), conicasterol J (**217**), theonellasterol J (**219**), and dehydroconicasterol (**220**) interact only with the hydroxyl group of Tyr358 (Helix 7). Comparing the docking pose of the series (Figure 6.30A) with respect to the co-crystallized agonist 6-ECDCA and the

isomers E and Z of gugglusterone (**196,** Figure 6.4), it is evident that all the molecules bind the FXR binding pocket formed between the Helix 2, 3, 5-7, and 10/11.

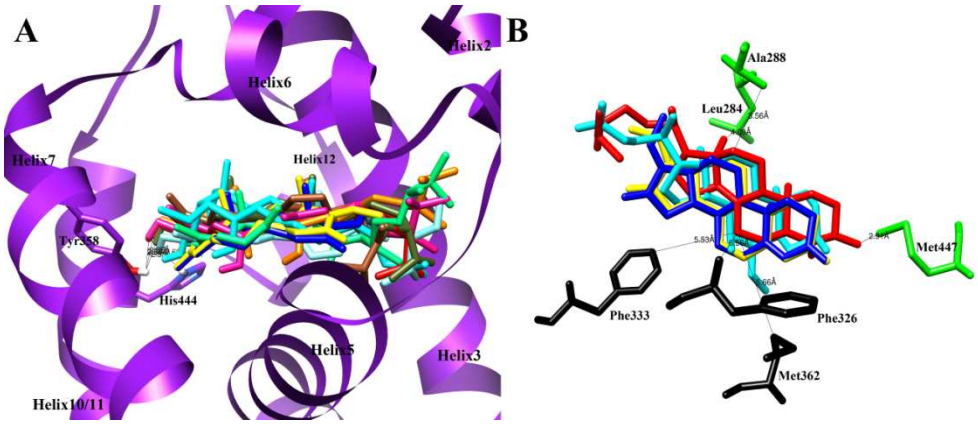


Figure 6. 30 (A) Superimposition of 6-ECDCA (**192**, sky blue), and Z (yellow)/E (blue) gugglusterone (**196**) with **214** (orange), **217** (pink), **218** (dark green), **219** (light blue), **199** (red), **220** (brown), and **223** (emerald) in the binding pocket of FXR (pdb code:1OSV). The FXR molecule is depicted by purple ribbon and the crucial amino acids by sticks (by atom type: C, purple; O, red; N, dark blue, H, white). (B) Amino acids interacting with 6-ECDCA (sky blue) and *Z* (yellow)/*E* (blue) gugglusterone are depicted in black; amino acids interacting with **199** (red) are depicted in green.

Notably, as shown in Figure 6.30B for conicasterol (**199**), the trans junction between A/B rings of this family, combined with diverse unsaturations positioning, causes a different spatial arrangement with respect the well known ligands missing some hydrophobic interactions with the amino acids of the FXR binding site (Phe326, Phe333, Met362). Also the molecule side chain, as reported in our previous study,[294] influences the activity on FXR; in fact, the methyl group at position 24 of conicasterol (**199**) (red), conicasterol K (**218**) and swinhosterol B (**223**) is in close contact with Ile332 on the surface

receptor with respect to theonellasterol J (**219**), bearing an ethyl group with different configuration at the same position (Figure 6.31A).

On the other hand, the major rigidity of the side chain of dehydroconicasterol (**220**) in comparison with the parent conicasterol (**199**) (red) does not allow further hydrophobic interactions with Met262 (Coil 2) and His291 (Helix 3). Moreover, the different positioning of the unsaturations of conicasterol K (**218**) and theonellasterol J (**219**) with respect to conicasterol (**199**) (red) and dehydroconicasterol (**220**) causes a further loss of hydrophobic interactions with the amino acids Leu284, Ile349, and Ile354 (Figure 6.31A). On the other hand, the unusual open nucleus of swinhosterol B (**223**) maintains the same hydrophobic interactions with respect to **199** with the exception of Leu284 and Met262 (Figure 6.31A).

Considering the other nucleus substitutions, the inverted positions of CO and OH groups at C-7 and C-15 of conicasterols G (**214**) and J (**217**) cause a different pattern of hydrogen bonds; in fact conicasterol G (**214**) is in close contact only with Ser329, while the α-OH at C-7 of conicasterols J (**217**) forms two hydrogen bonds with FXR binding pocket, as H-bond acceptor with OH of Ser329 and as H-bond donor with the OH group of Tyr366. Moreover its CO at C-15, as well as the carbonyl group at C-14 of 11, establishes a further weak interaction with Ser329 (Figure 6.31B).

In summary, the different substitutions on the steroid skeleton cause a great discrepancy in the pharmacological activity among this class of sterols; in particular only swinhosterol B (**223**), conicasterol (**199**) and conicasterol J (**217**) exhibit antagonist activity.

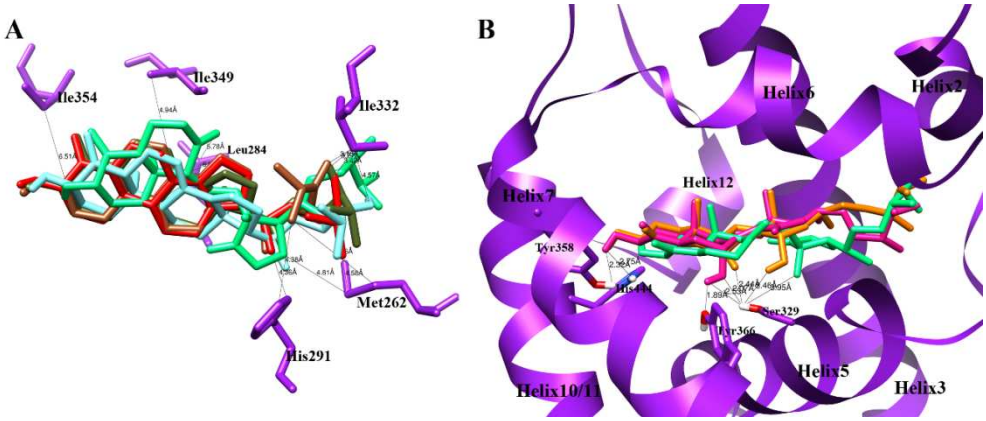


Figure 6. 31 (A) Superimposition between **218** (dark green), **219** (light blue), **199** (red), **220** (brown), and **223** (emerald) in the FXR binding site. (B) Three dimensional model of the different hydrogen bond pattern of **217** (pink), **214** (orange) and **223** (emerald) with FXR. In both figures the crucial amino acids of FXR receptor are depicted by purple sticks.

The lack, in fact, of simultaneous and efficient hydrophobic additional interactions of **218**, **219**, and **220** with respect to **199** and **223**, and different pattern of hydrogen bond of **214** with respect to **217** and **223**, suggest that these contacts are critical for the competition with 6-ECDCA in occupying FXR binding site, but they are not sufficient for the receptor activation.

Moreover, on the basis of the above biological results, among the whole library (Figures 6.24-25), swinhosterol B (**223**) was also disclosed as unique new potential therapeutic agent with double opposite activity on FXR and PXR. For this reason, it was also tried to rationalize, through molecular modeling studies, the bases of its PXR agonist behavior. As reported above, the Autodock4.2 software[56] was utilized for the docking experiments using the human nuclear pregnane-X-receptor reported by Watkins et al.[468] as model receptor, which was successfully used in previous works.[294,474]

PXR[470,471,472] presents a large ligand binding cavity[468] allowing the accommodation of different kind of molecules. As already reported by

Watkins et al.,[470] the possible binding modes are characterized by the different hydrogen bonds and Van der Waals patterns established between a small molecule and the receptor. Also the reported results, in fact, have pointed out three different docking poses for the marine swinhosterol B (**223**). As shown in Figure 6.32, the OH at C-3 forms hydrogen bonds with CO of Gln285, OH of Ser247, and with CO of His407 for the docking poses A, B, C, respectively.

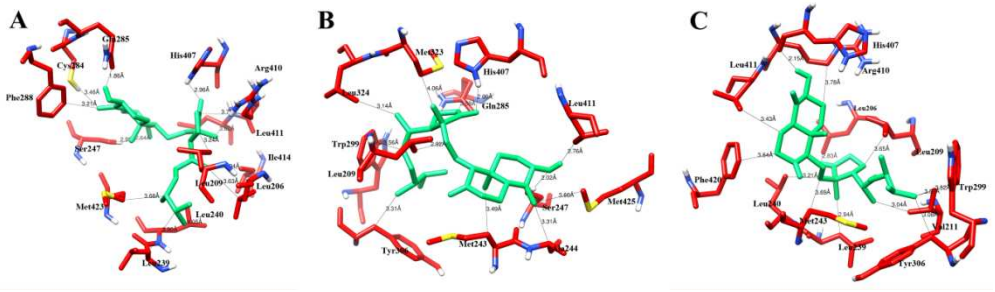


Figure 6. 32 Three dimensional models of the possible docking poses (A-C) of swinhosterol B (**223**) with hPXR.

The CO at C-8 establishes further hydrogen bonds with OH of Ser 247 for pose A, while the CO at C-14 interacts with NH of His407 for poses A and B.

Besides the different pattern of hydrogen bonds, all the three poses establish different Van der Waal interaction with PXR large ligand binding pocket (Figure 6.32) formed by hydrophobic (Cys284, Leu206, Leu209, Leu239, Leu240, Leu411, Met243, Met423, Met425, Phe288, Phe420, Trp299, Val211), polar, and charged (Arg410, Tyr306) amino acids. In conclusion, we identify in the swinhosterol B (**223**) ability to respect these well-known key interactions the rationalization of its agonist activity on PXR.

In summary, swinhosterol B (**11**) was shown to act as potent PXR agonist endowed with antagonistic activity for FXR. The ability to modulate immune

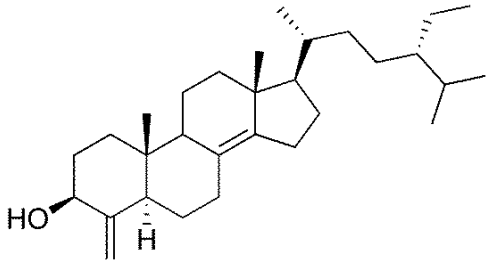
function strongly supports the exploitation of this compound in rodent model of liver inflammation and cholestasis.

6.4.1. Computational Details.

Molecular docking calculations were performed by Autodock 4.2 software[56] on 4 x AMD Opteron SixCore 2.4Ghz. A grid box size of 94 x 96 x 68 for chain A of FXR (pdb code:1OSV),[462] and 90 x 106 x 92 for PXR (pdb code: 1M13)[468] was used with spacing of 0.375 Å between the grid points, and centered at 20.689 (x), 39.478 (y), 10.921 (z) between the SCH_3 of Met262 and the OH group of Thr267 for FXR, and at 14.282 (x), 74.983 (y), 0.974 (z) between the ring of His407 and the side chain of Leu209 for PXR, covering the active site of both the receptors. To achieve a representative conformational space during the docking studies and for taking into account the variable number of active torsions, 10 calculations consisting of 256 runs were performed, obtaining 2560 structures for each ligand. The Lamarckian genetic algorithm (LGA) was employed for docking experiments, choosing an initial population of 600 randomly placed individuals. The maximum number of energy evaluations and of generations was set up to 5×10^6 and to 6×10^6 respectively. For all the docked structures, all bonds were treated as active torsional bonds except the bonds in cycles, which are considered fixed together with the receptors. Results differing by less than 3.5 Å in positional root-mean-square deviation (RMSD) were clustered together and represented by the result with the most favorable free energy of binding. Illustrations of the 3D models were generated using the Python[165] and Chimera[174] software.

6.5 Discovery of theonellasterol a marine sponge sterol as a highly selective FXR antagonist that protects against liver injury in cholestasis

In this paragraph, it will be described is a highly selective FXR antagonist 4-methylenesteroid isolated from the *Theonella swinhoei* sponge (Figure 6.33).[359]



Theonellasterol **198**

Figure 6. 33 Chemical structure of theonellasterol isolated from *Theonella swinhoei.*

The initial processing of the *Theonella swinhoei* was conducted according to procedures described previously.[294] The identity of theonellasterol was secured by comparison of its NMR and MS spectrum with those previously reported.[479]

By transactivation and microarray analyses carried out in HepG2 cells, a human hepatocyte cell line, it was found that theonellasterol is a selective FXR antagonist, devoid of any agonistic or antagonistic activity on a number of human nuclear receptors including the vitamin D receptor, PPARs, PXR, LXR and progesterone, estrogen, glucorticoid and thyroid receptors, among others. Exposure of HepG2 cells to theonellasterol antagonizes the effect of natural and synthetic FXR agonist on a number of FXR target genes, including SHP, OSTα, BSEP and MRP4. Using this agent a proof-of-concept study was have

carried out to investigate whether FXR antagonism rescues mice from liver injury induced by the ligation of the common bile duct, a model of cholestasis. Results from this experiment demonstrates that theonellasterol rescues mice from liver injury caused by bile duct ligation as measure by assessing serum alanine aminostrasferase levels and extent of liver necrosis at histopathology. Analysis of genes involved in bile acid uptake and excretion by hepatocytes reveals that theonellasterol increases the liver expression of MRP4. Administering bile duct ligated mice with a FXR agonist failed to rescue from liver injury and profoundly downregulated the expression of MRP4. Present results demonstrate that FXR antagonism effectively regulates expression of MRP-4 in the liver and is a feasible strategy to target obstructive cholestasis.

As already reported by Soissonet al.,[485] the flexible nature of FXR side chains suggests that its ligand binding domain (LBD) may have considerable ability to accommodate differently shaped ligands changing in response to the binding with them. For these reason, in order to rationalize the binding mode of theonellasterol on FXR receptor, molecular docking calculations (by Autodock 4.2 software)[56] were performed on several FXR structures co-crystallized with different compounds[462,486,487,488] with the aim to predict the position of the LBD in complex with marine sterol **198**. As reported, the activation of the FXR by the sterols molecules is, among the others, regulated by the interaction between the OH at C-3 of steroid skeleton and the amino acids of the catalytic triad (namely Tyr in Helix 7, His in Helix 10/11, Trp Helix12).[462] In all the three dimensional models (Figure 6.34A) the theonellasterol is able to interact with the catalytic triad, and in particular it forms a hydrogen bond with Tyr358 and 365 in Helix 7 for the pdbs 1OSV[462] and 1OSH[486] respectively, and with His447 for the pdbs 3DCT,[487] 3BEJ,[485] 3RUU,[488] and His444 for 1OSV[462] (Helix 10/11).

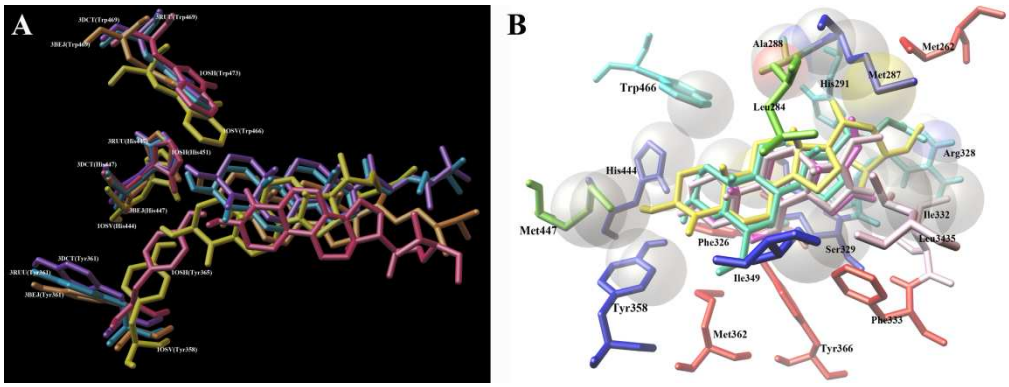


Figure 6. 34 (A) Superimposition of the different docking poses of theonellasterol (**198**) in the rat FXR (**198** and 1OSV[462] yellow), and human FXRs (**198** and 3BEJ[485] orange; **198** and 1OSH[486] pink; **198** and 3DCT[487] purple; **198** and 3RUU[488] light blue). (B) Superimposition of theonellasterol (yellow) with 6-ECDCA (sky blue), and Z (pink)/ E (light pink) gugglusterone in the binding pocket of FXR (1OSV). Amino acids interacting with theonellasterol (yellow) are depicted in green, amino acids interacting with 6-ECDCA and theonellasterol are depicted in sky blue, amino acids interacting with Z/E gugglusterone and theonellasterol are depicted light pink, amino acids interacting with 6-ECDCA and Z/E gugglusterone in red, and amino acids interacting with all molecules are depicted in blue.

On the other hand, only in the three dimensional models with 3DCT,[487] 3RUU,[488] and 1OSV[462] the marine sterol establishes hydrophobic interactions with Helix 12 and in particular with the Trp469 (3DCT,[487] 3RUU,[488]) or Trp466 (1OSV[462]). Considering also the others interactions with the LBD, the complex with the 1OSV[462] was chosen for the molecular docking analysis because **198** and 6-ECDCA (**192**, Figure 6.4) show similar chemical features with respect to the other molecules co-crystallized with the FXR considered structures.

On this basis, in addition to the two hydrogen bonds (Figure 6.34B) between *theonellasterol* with Tyr358 (Helix 7) and His444 (Helix 10/11)

reported above, the *trans* junction between A/B rings and its peculiar unsaturation between C-8 and C-14 cause a different spatial arrangement (Figure 6.34B) with respect to the semi-synthetic agonist 6-ECDCA[462] (**192**, Figure 6.4) and the natural antagonist guggulsterone[445c,d] (**196**, Figure 6.4) isomers not allowing the hydrophobic contact with Met362, Phe326, Phe333 and Tyr366. On the other hand, the **198** steroid skeleton interacts with Leu345 and Trp466 in the same manner of guggulsterone and 6-ECDCA respectively, and with Ala288, Leu284, and Met447; while its alkylic chains is in close contacts with Arg328 and His291, and with Ile332 (See Figure 6.34B) as reported for 6-ECDCA and guggulsterone respectively.

Furthermore, **198** maintains the same hydrophobic interactions of 6-ECDCA and guggulsterone with His444, Ile349, Met287, Met325, Ser329, Tyr358 in the ligand binding site. In conclusion, even if the sterol **1** shows a more simple skeleton with respect to the 4-methylene steroids previously reported in the paragraoh 6.2,[294] these docking results suggest that the different pattern of hydrophobic interactions established with FXR may be efficient and critical for the competition of *theonellasterol* with 6-ECDCA in occupying the FXR binding site.[370,294,462,445c,d,489]

6.5.2. Computational Details

The molecular docking calculations were performed by Autodock 4.2 software[56] on 4 x AMD Opteron SixCore 2.4Ghz, using a grid box size of 94 x 96 x 68, with spacing of 0.375 Å between the grid points, and centered at 20.689 (x), 39.478 (y), 10.921 (z), covering the active site of the FXRs.[462,485,486,487,488] To achieve a representative conformational space during the docking studies and for taking into account the variable number of active torsions, 10 calculations consisting of 256 runs were performed, obtaining

2560 structures for the sterol **198**. The Lamarckian genetic algorithm (LGA) was employed for docking calculations, choosing an initial population of 600 randomly placed individuals. The maximum number of energy evaluations and of generations was set up to 5×10^6 and to 6×10^6 respectively. Results differing by less than 3.5 Å in positional root-mean-square deviation (RMSD) were clustered together and represented by the most favorable free energy of binding. Illustrations of the 3D models were generated the Python software.[165]

-CHAPTER 7-

The Bile Acid Receptor GPBAR-1 (TGR5) Agonists

7.1 The Bile Acid Receptor GPBAR-1 (TGR5) Modulates Integrity of Intestinal Barrier and Immune Response to Experimental Colitis

In this chapert, it will be described the rationalization of the binding mode of two known anti-inflammatory drugs taurolithocholic (**224**, Scheme 7.1) acid and ciprofloxacin (**225**, Scheme 7.1) with the Bile Acid Receptor GPBAR-1 (TGR5), acting as agonist of this nuclear receptor.[296]

Bile acids play an essential role in integrating multiple homeostatic functions in the liver and gastrointestinal tract. In recent years these end-product of cholesterol metabolism have been shown to signal through activation of variety of nuclear and cell surface receptors.[490] Activation of Farnesoid-x-receptor (FXR), pregnane-x-receptor (PXR), and constitutive androstane receptor (CAR), along with the vitamin D receptor (VDR), by primary bile acids chenodeoxycholic acid (CDCA, **191** in Figure 6.4) elicits a series of genomic effects that have been deemed essential for regulation of lipid, cholesterol and bile acid homeostasis, local immune response and insulin signaling in intestinal and liver tissues.[490,491] Knocking down the expression of FXR, the main bile acid receptor, results in a multilevel dysregulation of glucose, lipid, cholesterol and protein metabolism, highlighting the essential role of this receptor in maintaining homeostasis in entero-hepatic tissues.[490,491] In addition, bile acids exert non-genomic effects.[490,491] These non-genomic effects have been ascribed to the activation of a cell surface receptor named TGR5 or M-BAR, a member of the rhodopsin-like superfamily of G protein coupled receptor (GPCR), recently christened as a bile acid-activated GPCR (GP-BAR1).[492,493] GP-BAR1 is restricted to a limited number of tissues, with

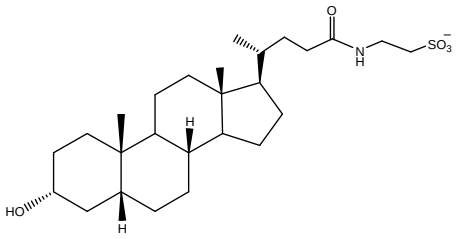
the highest expression detected in brown adipose tissue, spleen, macrophages/monocytes, gallbladder and intestine.[492,493,494]

In the small and large intestine, GP-BAR1 has been detected in the enteric ganglia of the myenteric and submucosal plexus, in the muscularis externa and in the mucosa, in enterocytes of the crypts and villi, while in the cecum and colon the receptor is expressed, thought at lower, in muscle layers and mucosa.[495] In target cells, GP-BAR1 activation by secondary bile acids, lithocolic acid (LCA) and tauro-LCA (TLCA), increases the intracellular concentrations of cyclic adenosine monophosphate (cAMP) and causes the receptor internalization.[490,491,492,493] In intestinal endocrine L-cells that are higly enriched in receptor expression, GP-BAR1 activation by bile acids and dietary agents stimulates the secretion of glucagon-like peptide (GLP)-1, an insulinotropic hormone that regulates insulin and glucagon secretion along with gastrointestinal motility and appetite.[490,491,492,493,496] In addition to its intestinal localization, GP-BAR1 has been detected in peripheral blood derived macrophages and liver macrophages where it exerts an immune-modulatory activity.[491,493] This activity is inhibitory in nature and manifests itself by attenuation of macrophage's effector functions including reduction of phagocytic activity as well and generation of lipopolysaccharide (LPS)-stimulated cytokines (TNF-a, IL-1a, IL-1b, IL-6, and IL-89).[491,497]

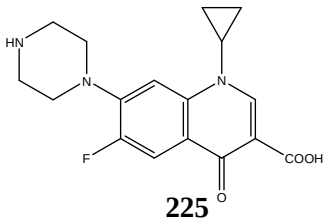
Despite its role in integrating intestinal homeostasis and glucose metabolism is well defined, it is not known whether GP-BAR1 participates into local regulation of intestinal inflammation and whether its ablation would manifest by an exaggerated inflammatory response to intestinal antigens. Because the expression of GPBAR1 is highly restricted to the intestine and identification of a regulatory role would be of interest to ground intestine-specific anti-inflammatory therapies, it was investigated whether GPBAR1

plays a functional role in regulating intestinal homeostasis and inflammation-driven immune response.

As for the majority of G-protein-coupled receptors (GPCRs), the structure of TGR5 has not yet been determined experimentally, and so a homology modeling study was performed in order to obtain the three dimensional structure of the receptor. In particular, basing on the good recent results obtained by Hov et al.[498] the human adenosine A2a receptor was used as template for our modeling studies. After this preliminary step, the predicted TGR5 structure (Figure 7.1) was used to analyze, by means of molecular docking calculations, the interactions of the agonist taurolitholic acid (**224,** Scheme 7.1), and ciprofloxacin (**225,** Scheme 7.1) with the model of the G-protein-coupled receptor in order to generate a structure-activity relationship and obtain information on their binding mode at atomic level.



224



225

Scheme 7. 1 Chemical structure of t-LCA (**224**), and ciprofloxacin (**225**).

As reported in the previous work the binding site of the agonist is located at the N-terminal-extracellular portion,[499] and so the reported docking calculations (Autodock4.2 software)[56] was focused around this protein portion.

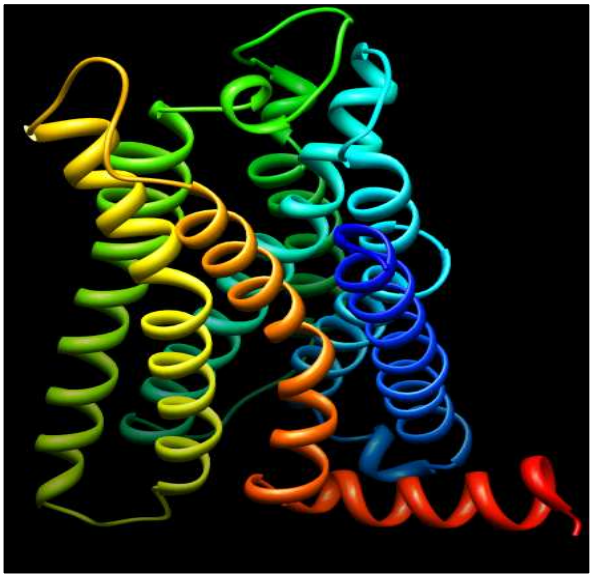


Figure 7. 1 Predicted three dimensional model of the TGR5 receptor.

The TLCA derivative presents an EC_{50} of ~300 nM; as previously reported by other research groups, the hydroxyl group at C-3 is involved in hydrogen bond interactions with the receptor and removal of this group causes decrease in potency toward TGR5 activation.[499,500,501] As shown in Figure 7.2, the BA accommodates on the TGR5 receptor surface, and it is involved in van der Waals interactions with Trp65, Asn66, Ser68, Arg69, Asn144, Cys145, Ser146, Leu231, Leu235. The OH at C-3 interacts with hydroxyl groups of Tyr209 and Ser147, the NH of the taurine interacts with CO of Gln67, also in agreement with the hydrogen bond model proposed by Tiwari et al.,[499] and the sulfate group protrudes toward the solvent.

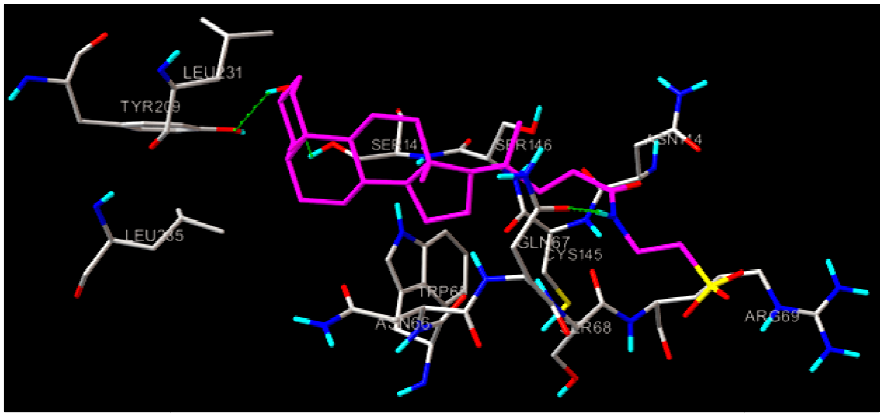


Figure 7. 2 Three dimensional model between TLCA (purple) and calculated structure of TGR5 receptor.

For what concerns the ciprofloxacin interactions, its CO and COOH groups are involved in hydrogen bond interaction with Gln67 and Ser68 respectively, and also its piperazine group form hydrogen bond OH of Tyr209. Moreover the molecule establishes hydrophobic interactions with the cavity pocket formed by Ser146, Ser147, Gln148, Leu231, Leu213, Trp65, Asn66.

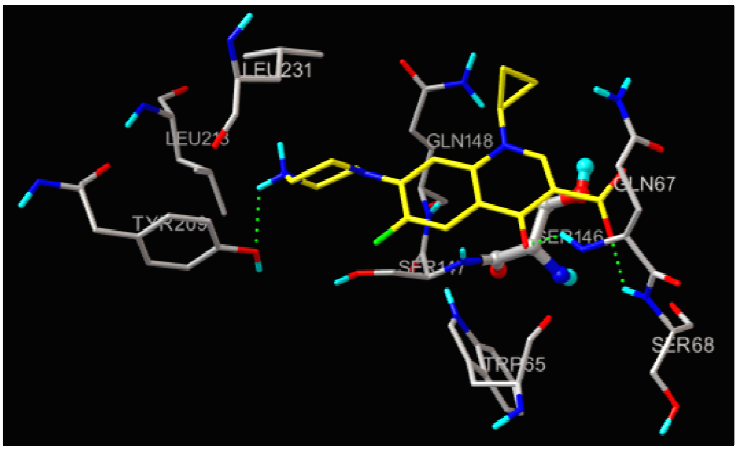


Figure 7. 3 Three dimensional model between ciprofloxacin (yellow) and calculated structure of TGR5 receptor.

All these findings, regarding the binding pose of ciprofloxacin in the same N-terminal region of TLCA, are compatible with its observed agonist activity on TGR5.

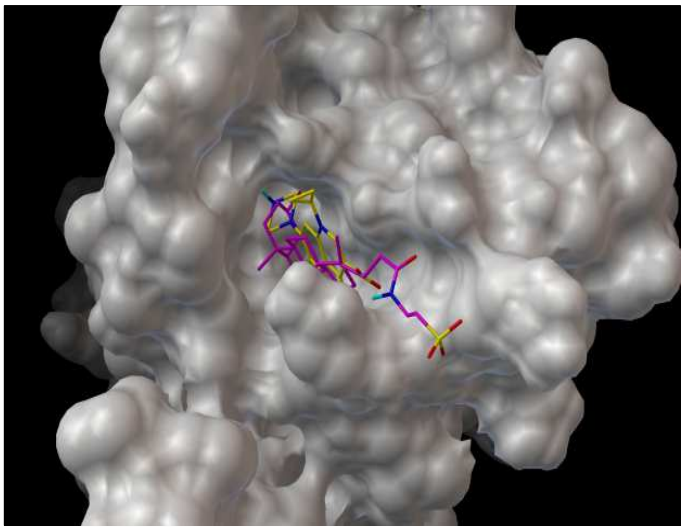


Figure 7. 4 Superimposition of TLCA (purple) and ciprofloxacin (yellow) on and calculated structure of TGR5 receptor.the TGR5 surface.

Results of docking studies were confirmed by in vitro studies in GLUTag cells, a cell line highly enriched in GP-BAR1.[502] Exposure to ciprofloxacin (Figure 7.5A) resulted in a concentrationdependent increase of ($[cAMP]_i$) with an EC_{50} of <8 mM (n =3). Ciprofloxacin, 10 mM, was as effective as 10 mM TLCA (n =3).

In the search for ligand that could be exploited therapeutically as GP-BAR1 ligand, we have identified ciprofloxacin as a GP BAR1 agonist. Ciprofloxacin, is widely used in the treatment of infections due to Gram negative bacteria in Crohn's disease. In addition, ciprofloxacin has been shown to increase ($[cAMP]_i$) in monocytes and macrophages, and by this mean to exert a

counterregulatory effect on cytokine production triggered by LPS.[503] So far the molecular mechanisms mediating these effects were left unknown.

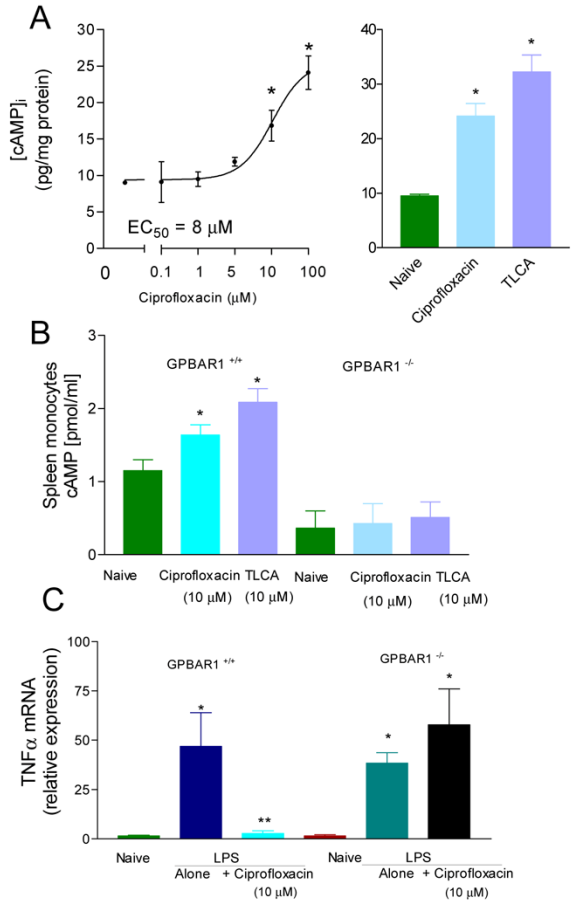


Figure 7. 5 Panel A. Ciprofloxacin and TLCA increases ($[cAMP]_i$) in GLUTag cells. N= 4; P,0.05. Panel B. Ciprofloxacin and taurolithocholic acid , 10 mM, caused a 2–3 fold increase in ($[cAMP]_i$) in spleen-derived monocytes isolated from GP-BAR1 wild type mice (n = 4–5; P,0.05), but not in cells isolated from GP-BAR12/2 mice. Panel C. Ciprofloxacin, 10 mM, inhibits LPS-induced TNFa release in GP-BAR1+/+ monocytes but not in cells isolated from GPBAR1 2/2 mice (n = 6; P,0.05 *versus* naive).

By *in silico* screening, docking calculation and in vitro experiments have shown that ciprofloxacin functions as GPBAR1 agonist. Indeed, not only ciprofloxacin entertains meaningful interaction with key aminoacids in the binding site of GPBAR1, as demonstrated by docking experiments, but it triggers changes in [cAMP]i in GLUTag cells, an L-like cell line generated from an entero-endocrine tumor and highly enriched in GPBAR1, and in spleen-derived monocytes.[490,491,492,493,502] In summary, GP-BAR1 is involved in regulating intestinal homeostasis and that its absence manifests by an increased intestinal permeability and enhanced susceptibility to develop colitis in response to barrier braking agents. Expression of GP-BAR1, in fact, increases in response to inflammation in rodent models of colitis and in inflamed tissues obtained from Crohn' disease patients. Moreover, here it was discovered that ciprofloxacin, a widely used antibiotic, is a GP-BAR1 agonist and that activation of GP-BAR1 with this agent or oleanolic acid, a natural GP-BAR1 ligand, attenuates colon inflammation in rodent models of colitis.

7.1.1 Computational Details

The results recently published by Hov et al[498] was used as reference for homology modeling studies. In their work, the human adenosine A2a receptor (pdb code:3EML) was used as template for the structure studies. In order to obtain the three dimensional model of the TGR5 receptor we used the alignment reported by Hov et al., see Figure 7.6.

The alignment was used as input for the automated homology modeling program MODELER.[216] The number of generated loops was set to 5 along with high optimization level for models and loops. The generated models of TGR5 showing the lowest energy, lowest restraint violations were selected for further refinement of the main and side chains.

```
hTGR5           SPIPKGALGLSLALASLIITANLLLALGIAWDRRLRSPPAGCFFLSLLLAGLLTGLALPT
Template 3EML   IMGSSVYITVELAIAVLAILGNVLVCWAVWLNSNLQ-NVTNYFVVSLAAADIAVGVLAIP
hTGR5           LPGLWNQS-RRGYWSCLLVYLAPNFSFLSLLANLLLVHGERYMAVLRPLQ-----PPGSI
Template 3EML   FAITISTGFCAACHGCLFIACFVLVLTQSSIFSLLAIAIDRYIAIRIPLRYNGLVTGTRA
hTGR5           RLALLLTWAGPLLFASLPALGWNHWT-------PGANCSSQAIF-PAPYLYLEVYGLLLP
Template 3EML   KGIIAICWVLSFAIGLTPMLGWNNCGQSQGCGEGQVACLFEDVVPMNYMVYFNFFACVLV
hTGR5           AVGAAAFLSVRVLATAHRQL-------------------------------------
Template 3EML   PLLLMLGVYLRIFLAARRQLNIFEMLRIDEGLRLKIYKDTEGYYTIGIGHLLTKSPSLNA
hTGR5           ------------------------------------------------------------
Template 3EML   AKSELDKAIGRNTNGVITKDEAEKLFNQDVDAAVRGILRNAKLKPVYDSLDAVRRAALIN
hTGR5           ------------------------------------------------------------
Template 3EML   MVFQMGETGVAGFTNSLRMLQQKRWDEAAVNLAKSRWYNQTPNRAKRVITTFRTGTWDAY
hTGR5           -----LTWRQARAQAGAMLLFGLCWGPYVATLLLSVLAYEQRPPLGPGTLLSLLSLGSAS
Template 3EML   RSTLQKEVHAAKSLAIIVGLFALCWLPLHIINCFTFFCPD-CSHAPLWLMYLAIVLSHTN
hTGR5           AAAVPVAMGLGDQRYTAPWRAAAQRCLQGL
Template 3EML   SVVNPFIYAYRIREFRQTFRKIIRSHVLRQ
```

Figure 7. 6 The sequence alignment of 3EML and human TGR5.

7.1.1.1 Model refinement

Hydrogen atoms were added to the selected models using Maestro 8.5 software package.[159] In details the charges of side chains were assigned considering their pK_a at physiological pH. The geometry of the model for the target was optimized in three steps:

1) optimization of model with the added hydrogen atoms by Amber force field[504] (steepest descent method, 500 steps and convergence threshold of 0.05 kJ mol^{-1} $Å^{-1}$).

2) optimization of the side chains using the same criteria of step 1.

3) optimization of the whole structure by Polak-Ribiere Conjugate Gradient (PRCG, 9 x 10^7 steps, convergence threshold 0.001 kJ mol^{-1} $Å^{-1}$).

The quality of the obtained model for TGR5 was validated using the software PROCHECK[218] (Figure 7.7).

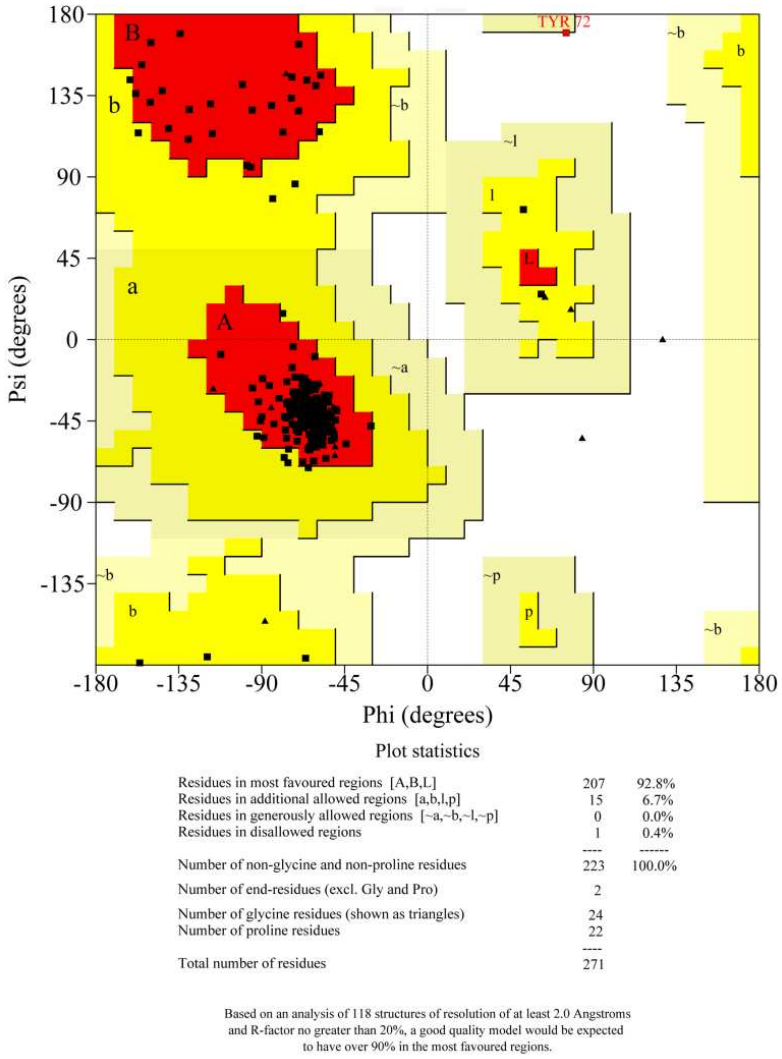


Figure 7. 7 Ramachandran plot of TGR5 calculated by PROCHECK using a hypothetical resolution of 2 Å.

The homology model for TGR5 was used in molecular docking calculations using AutoDock 4.2.[56]

To achieve a representative conformational space during the docking studies and for taking into account the variable number of active torsions, 10 calculations consisting of 256 runs were performed, obtaining 2560 structures

for each ligand. The Lamarckian genetic algorithm was employed for docking calculations, choosing an initial population of 450 randomly placed individuals. The maximum number of energy evaluations and of generations was set up to 5×10^6 and the maximum number of generations to 6×10^6. A mutation rate of 0.02 and a crossover rate of 0.8 were used, and the local search frequency was set up at 0.26. Results differing by less than 2 Å in positional root-mean-square deviation (rmsd) were clustered together and represented by the result with the most favourable free energy of binding. For all the docked structures, all bonds were treated as active torsional bonds except the amide bonds. Illustrations of the 3D models were generated using the Python software.[165]

-CHAPTER 8-

Structural Studies of Natural Products

8.1 DFT/NMR integrated approach: a valid support to the total synthesis of chiral molecules

The entire stereochemical knowledge of a natural product is of fundamental importance in many different fields, spanning from chemical physics to biochemistry. The correct assignment of the configurational pattern in chiral organic compounds, containing more than one stereocenter, is undoubtedly a key step of the structure elucidation process. This process, in fact, is essential in several fields that not only include the total synthesis of the molecules under investigation but also the understanding, at the molecular level, of the biological mechanism of active natural compounds; stereochemical knowledge is also fundamental for structure-activity studies of drug–receptor systems.[505] Basically, there are different approaches to identify the exact structure and/or stereochemistry of organic products. The classic chemical approach, represented by the total synthesis, has played a major role in the structural assignments and revision for a long time, but the additional costs in terms of time and money represent some of its most important weaknesses. The classical chemical approach is in fact often replaced by a series of more rapid methods, such as NMR, circular dichroism (CD), X-ray crystallography, and mass spectrometry (MS). Nevertheless, the correct structural assignment of unknown or known natural products is often very difficult to obtain, especially when a small quantity of the natural compound is available, limiting the full possibility of the currently available spectroscopic methods. For this reason, many examples of structural revision[75] have appeared in literature, where total synthesis is able to show that previous assignments were wrong. This situation is particularly unfortunate in all cases where the total synthesis of a (wrong) proposed

structure does not afford the natural products, and, especially for complicated molecules, this translates in an enormous waste of time and money. The use of quantum mechanical (QM) methods, in fact, was suggested as a rapid, efficient, and economical method for the resolution of stereochemical problems, such as in the example of the structural determination and/or revision of active natural compounds.[81,82,506] Moreover, the enormous enhancement in the performance of computers nowadays along with the user-friendliness of the commercial quantum chemistry software packages should also encourage the synthetic chemists to perform such kind of analysis prior to proceeding to the chemical synthesis, excluding in this way candidate molecules that are not in accordance with the experimental data.

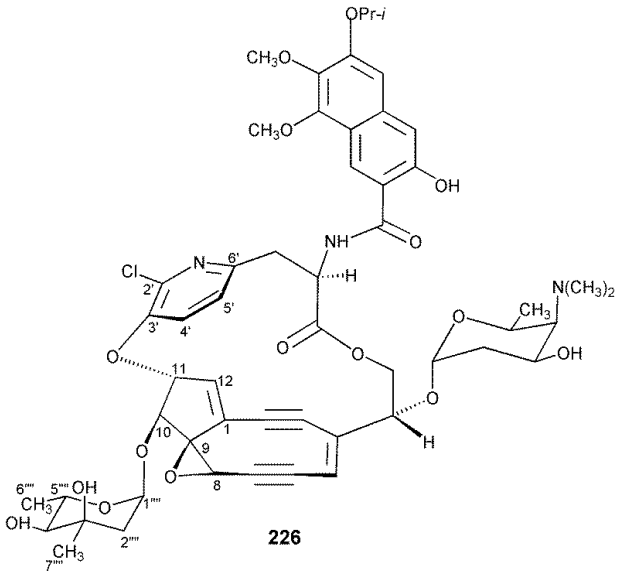
In this paragraph, it was demonstrate that a fast quantum chemical analysis of the stereoisomers of the two natural compounds *kedarcidin chromophore* (**226**–**228**) (from the oldest to the revised structure) and *palau'amine* (**229**) would have avoided the synthesis of the wrong proposed structures by means of density functional theory (DFT) calculation of the coupling constant and GIAO (gauge including atomic orbitals) calculation of ^{13}C chemical shifts (cs).[507] The DFT/NMR integrated approach was applied to these two compounds to show that the reported results could have been sufficient to exclude the proposed structures and could have directly suggested the structure of the correct compounds.

Kedarcidin chromophore is a compound that belongs to the enediyne family of antitumor antibiotics. These compounds are characterized by a unique molecular architecture, an intriguing mode of action, and high potent biological and pharmacological activities. They include, along with kedarcidin chromophore, calicheamicins, esperamicins, dynemicin A, neocarzinostatin chromophore, and C1027 chromophore. All these

compounds possess a common enediyne reactive center along with a chemical functionality that can trigger a cascade of events to form a highly reactive aryl-diradical. These molecules also possess a recognition unit responsible for delivering and establishing their position to the biological target, i.e. DNA. Under diverse conditions, these molecules may undergo a Bergman cyclization,[508] which transforms the enediyne moiety into a very reactive aryl-diradical.[509] The diradical abstract hydrogen atoms from each DNA strand lead to scission of the duplex, such damage cannot be repaired by any of the intracellular DNA repair paths. A first-structure elucidation of kedarcidin chromophore was proposed in 1992 by Leet and coworkers. It is here displayed as **226**[510] (Scheme 8.1).

Subsequently, in 1997, Hirama *et al*. in the course of their synthetic study toward the synthesis of kedarcidin demonstrated that the fragment degraded from the chromophore was not an α-amino acid derivative, but a β-amino ester.[511] They also disputed the absolute configuration of all the stereogenic centers of the molecule and proposed that the structure of kedarcidin chromophore had to be revised as **227**.[511] Very recently, Myers *et al*. synthesized **227** through an unambiguous enantioselective approach and, comparing its spectroscopical data with the data of the natural compounds, discovered some significant inconsistencies that led them to suggest **228** as the structure of kedarcidin chromophore (Scheme 8.2).[512]

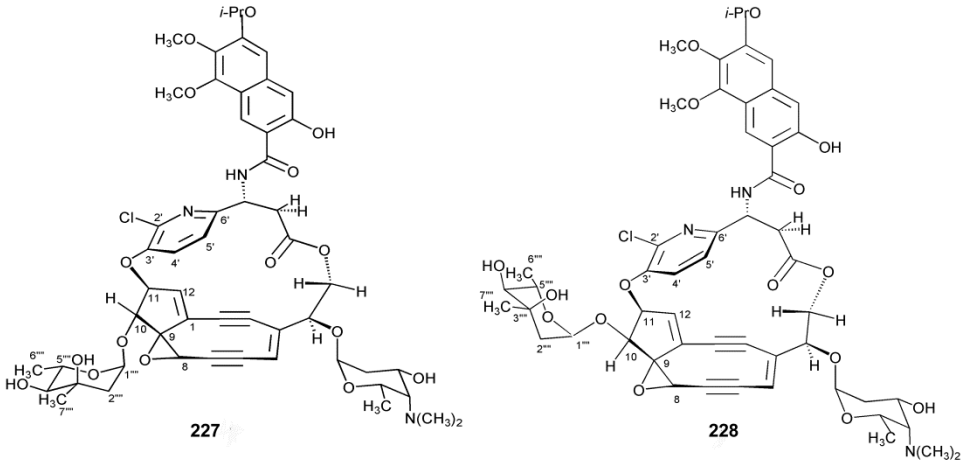
In particular, they observed for **228** a *ca* zero coupling constant between H-10 and H-11 whereas such coupling results were 5.4 Hz in $CDCl_3$ for authentic kedarcidin chromophore. Moreover, they noticed the lack of a nuclear overhauser effect (NOE) between the pyridyl H-4' and H-10 and the presence of an NOE between H-10 and H-12b in the product of borohydrideinduced cycloaromatization of kedarcidin chromophore.



226

Scheme 8. 1 Structure of kedarcidin chromophore as α-amino acid derivative (**226**) proposed by Leet in 1993.

Finally, they showed that synthesized **227** is a mixture of atropisomers in both DMSO-d_6 and C_6D_6, probably due to a slow flipping of the pyridyl ring around two preferential arrangements, one with the chlorine atom pointing toward the central core and the other one resulting from a 180° rotation of the pyridyl ring and with the chlorine atom pointing in the opposite direction with respect to the enedyine moiety. On the other hand, they argued that structure **228**, derived by epimerization of the mycarose-bearing carbon 10, could not exist in the atropisomeric form in which the chlorine atom is oriented toward C-10, in accordance with the single atropisomer spectra displayed by the natural product.[512]

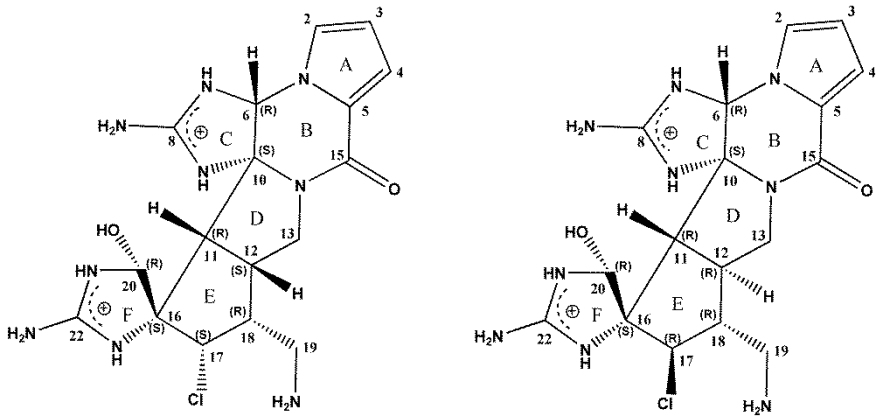


Scheme 8. 2 Previous structure proposed by Hirama in 1997 (**227**) and revised structure (**228**) proposed by Myers in 2007 of kedarcidin chromophore.

Palau'amine is a marine natural product originally isolated in 1993 by Scheuer and coworkers from the sponge *Stylotella aurantium*.[513] The natural product is an oroidin dimer, belonging to the class of pyrrole-imidazole alkoloid family together with the styloguanidines, the konbu'acidins, brominated palau'amines, and the stylissadines A and B.

Palau'amine is a potent immunosuppressive agent that possesses an interesting biological activity; in fact, its anticancer, antifungal, and antibacterial properties are undergoing preclinical studies.[514] More precisely, this marine natural product is a contiguous hexacyclic alkaloid that contains two guanidine units, a pyrrole carboxylic acid and an unbroken chain of eight chiral centers. However, owing to the presence of these chiral centers, the exact stereochemistry of palau'amine has been subject to controversies. The initial structure proposed by Scheuer *et al*. **229a**[513] (Scheme 8.3) had the relative configuration assigned as 6*R**, 10*S**, 11 *R**, 12 *S**, 16 *S**, 17 *S**, 18 *R**,20 *R**, and this resulted in a *cis*-fusion between D and E rings, and the

chlorine atom on the α-face of the molecule. Very recently, the isolation and structural elucidation of two additional palau'amine congeners, the tetrabromostyloguanidine[515] (also carteramine A[516]) and the konbu'acidin B,[517] prompted the revision of the initial assignment of the relative configuration of palau'amine. These three different research groups with others, in fact, proposed the revision of its stereochemistry using NMR spectroscopy, computational methods,[515,518,519] correlation with related compounds,[516,517,520] and reisolation and characterization of palau'amine itself.[521] The proposed revised structure **229b** (Scheme 8.3) has the relative configuration assigned as 6 *R**, 10 *S**, 11 *R**, 12 *R**, 16 *S**, 17 *R**, 18 *R**, 20 *R** and thus there is a *trans* junction between D and E rings and a chlorine atom on the β-face of the molecule.



Scheme 8. 3 Original **229a** and revised **229b** structure of palau'amine.

The main difference between the two possible models of the palau'amine is the junction from E–D rings and the relative configuration of carbon 17. In the first assignment of Scheuer and coworkers, the fusion was assigned as *cis*, based on the *J*-coupling constant (14.1 Hz) between H-11 and H-12 so that

'even relatively large for a *cis*-fusion, was considered comparable to other values of *J* constant observed in similar rigid, spiroannulated five-members rings.[513] Together with the reassignment of the junction as *trans*, the relative configuration of carbon 17 was corrected by NMR spectroscopy[517] and molecular modeling of other congeners.[518] Nevertheless, despitemany creative and insightful attempts in the absence of X-ray crystallographic data, the confirmation of the correct structure would eventually rely on the total synthesis of the natural product; however, unfortunately, the total synthesis of palau'amine has not yet been achieved.

In light of the above considerations, here, the use of QM methods was highlighted, in particular, the DFT calculation of the *J* values and the GIAO calculation of ^{13}C cs in the structure revision.

In the case of kedarcidin chromophore, in fact, to give further support to the configurational arrangement proposed by Myers *et al*.[512] and to investigate the overall conformational features and the atropisomerism associated with the natural and the synthetic compounds, a conformational analysis of **227** and **228** at the QM level was underteken, followed by a DFT calculation of significant *J* values for the two main atropisomers of **227** and **228**, whereas, in the case of palau'amine, in order to obtain its exact relative configuration, the calculation of significant *J* values was followed by a GIAO calculation of ^{13}C cs.

In order to obtain support on the configurational assignment of C-10 of kedarcidin, the $^{3}J_{H-H}$-coupling value for H-10 and H-11 was primarily considered. First, a conformational search on both the possible atropisomers for **227** and **228,** depicted as **227a** and **227b** and as **228a** and **228b** in Figure 8.1, was performed by means of molecular dynamics at different temperatures (400, 600, and 800 K) and with a Monte Carlo conformational search using

the MM3[522] force field included in the MacroModel software package.[159] Even though **227** and **228** are extremely flexible molecules, the chromophore core is rather rigid; therefore, only the lowest energy atropisomers were considered for the two diastereomers (**227a** and **227b** and **228a** and **228b** in Figure 8.1). On the geometries thus obtained, QM optimization of the energies and the geometries were performed *in vacuo* at the DFT level using the mPW1PW91 functional and the 6-31G(d) basis set,[84] and single-point calculations using a polar continuum model(PCM) were performed with the 6-31G(d,p)[84] basis for the simulation of the DMSO solvent.

The calculations for **227** are in accordance with the presence of two atropisomers that are very close in energy; in fact, the predicted differences are 1.8 and 1.3 kJ/mol accounting for a 1.8 : 1 and 1.3 : 1 ratio *in vacuo* and DMSO respectively. These results are in good qualitative accordance with the reported 2 : 1 ratio in DMSO-d_6. Interestingly, it is structure **227a** (Figure 8.1), which presents the chlorine atom directed away from the central core, to be the lowest in energy based on the results of the calculations, in accordance with what was proposed by Myers and coworkers.[512]

Also, the evidence of the presence of a single atropisomers for **3** is well reproduced by the results of the DFT calculation *in vacuo* and DMSO, which indicate differences of 5.4 and 2.5 kJ/mol respectively. As expected, it is atropisomer **228a**, presenting the chlorine atom away from C-10, which was predicted to have the lowest energy, in accordance with what was proposed by Myers *et al*.[512]

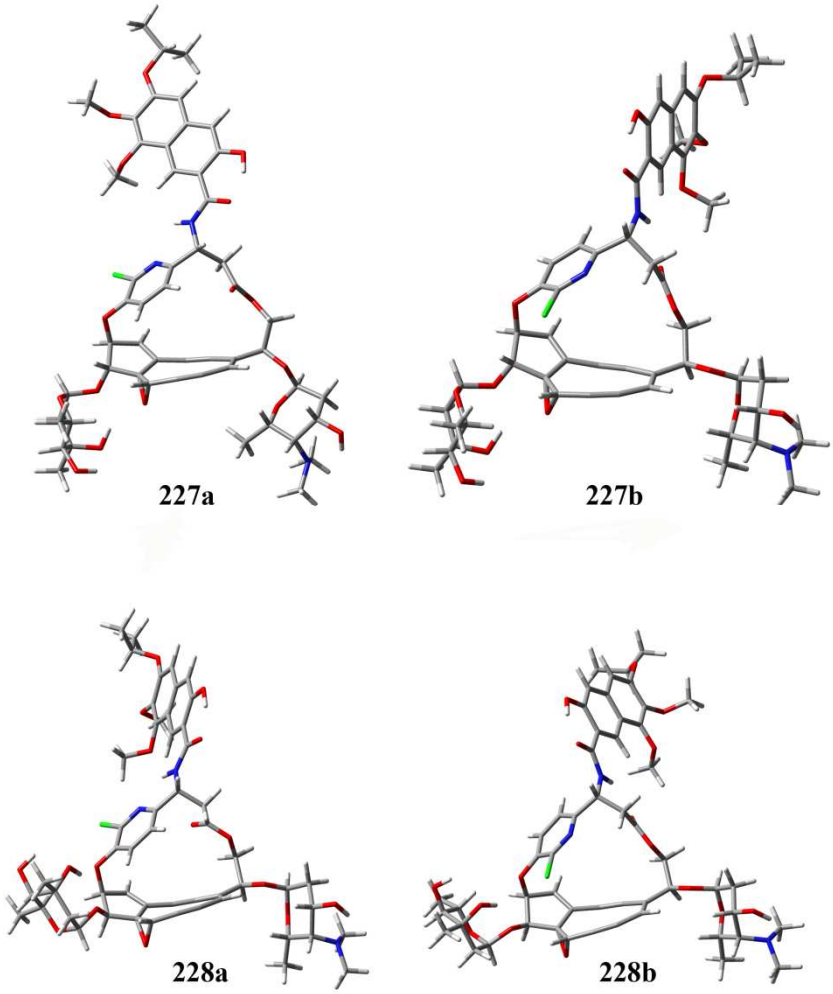


Figure 8. 1 Three-dimensionalmodels of possible kedarcidin chromophore atropisomers for **227** and **228**.

Apart from the above observations regarding the conformational behavior of kedarcidin chromophore, which is in accordance with what was observed in the NMR spectra of the natural product and of **227**, the results of the DFT calculation of the kedarcidin core fragment was examined (Computation Details) *J* values between H-10 and H-11 for **227a, 227b, 228a**, and **228b**

using the mPW1PW91 functional and the 6-311++G(2df, 2p) basis set.[84] In accordance with the experimental data collected by Myers, such values are predicted to be 0.1 Hz for **227a** and 0.1 Hz for **227b**. The values predicted for **228a** and **228b** are of 6.1 and 4.7 Hz respectively, and, compared to the experimental value of 5.4 Hz reported in $CDCl_3$, they strongly support that the kedarcidin chromophore structure be revised as **228,** confirming the hypothesis proposed by Myers *et al*.,[512] in particular, suggesting a 1 : 1 ratio between **228a** and **228b**. On the other hand, the theoretical energy difference of 2.5 kJ/mol in DMSO (5.4 kJ/mol *in vacuo*) between **228a** and **228b** would indicate a preponderance of conformation **3a**. Such results, even if not in perfect accordance with the results of the *J* analysis, are in good agreement with what was observed by Myers *et al*.,[512] suggesting that kedarcidin chromophore could have in its most stable conformation only one predominant atropisomer (**228a**) characterized by the chlorine atom of the pyridyl ring oriented away from the core portion of the molecule (Figure 8.1). Such structural attributes, together with other features of this conformation may be useful in better understanding the mechanism of action of kedarcidin chromophore at the molecular level.

The same approach was used for the palau'amine. In fact, the $^3J_{H-H}$-coupling value for H-11 and H-12 was primarily considered to obtain support or denial of the revised structure. Conformational search on both the possible diastereoisomers **229a** and **229b** depicted in Scheme 8.3 was performed by molecular dynamics at different temperatures (450, 700, and 800 K), and by the Monte Carlo conformational search using the MMFFs[158] force field included in the MacroModel software package.[159] Minimum energy conformer for each model were subjected to full geometry and energy optimization by QM methods at DFT level *in vacuo* using them

PW1PW91functional and the 6-31G(d) basis set.[84] On the basis of the obtained geometries (Figure 8.2), the calculation, *in vacuo*, of the *J*-coupling constants was executed always using the same functional and the 6-311++G(2df, 2p) basis set.

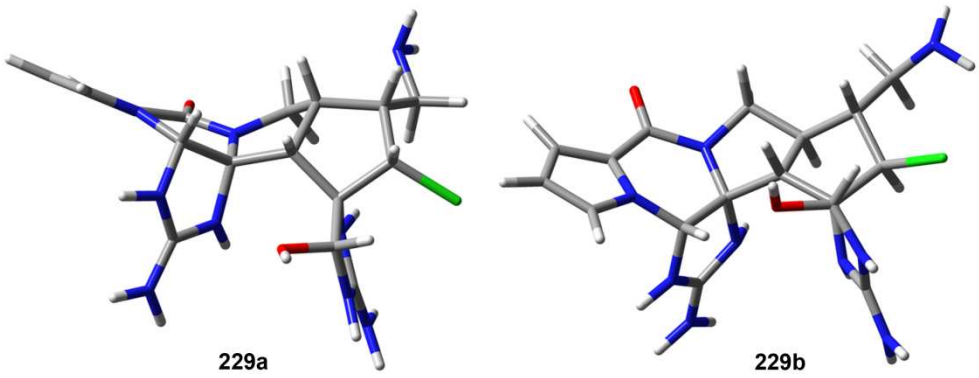


Figure 8. 2 Three-dimensionalmodels of the two diastereoisomers (**229a** and **229b**) of palau'amine.

In the reported calculations, the coupling constants $^{3}J_{H–H}$ of the D and E rings were only considered, which are diagnostic in the correct assignment of natural product configuration, and, in particular, the calculated values with the experimental *J*s recorded in DMSO (Table 8.1) were compared. As shown in Table 8.1, the value of 13.2 Hz, corresponding to the value of $^{3}J_{H11–12}$ of compound **229b**, displays the best agreement with the experimental data (14.4 Hz) with respect to compound **229a** that has a calculated *J* value of 8.8 Hz. This result is aligned and is consistent with the recently proposed revised structure, indicating that the junction between D and E rings is trans.

Moreover, the value of 9.6 Hz, related to the value of the coupling constant of $^{3}J_{H18–17}$ of compound **229b**, is in good agreement with experimental value of 9.0 Hz, suggesting that the relative configuration of carbon 17 may be revised as *R**.

Table 8. 1 Comparison between experimental (DMSO)[518] and calculated (*in vacuo*) (stereoisomers **229a** and **229b**) $^{3}J_{H\text{-}H}$ values in Hz of rings D and E.

	Calculated		**Exp DMSO**
	229a	**229b**	
$^{3}J_{H11\text{-}12}$	8.8	13.2	14.4
$^{3}J_{H12\text{-}18}$	9.0	9.6	9.0
$^{3}J_{H18\text{-}17}$	10.3	9.6	9.0
$^{3}J_{H12\text{-}13\alpha}$	8.5	9.0	10.2
$^{3}J_{H12\text{-}13\beta}$	9.9	7.0	7.2
[a]MAE	**2.3**	**0.7**	
[a]Mean Average Error $=\sum[\lvert J_{exp}\text{-}J_{calc}\rvert]/n$			

In summary, analyzing all the other calculated *J*-coupling constants of the two possible diastereoisomers **229a** and **229b**, we noted, as also outlined by the mean average error (MAE = $\sum[|J_{exp}-J_{calc}|]/n$ summation through *n* of the absolute values of the differences of the corresponding experimental and calculated *J*-coupling constants) in Table 8.1, that model **229b** fits better with the experimental NMR data reported in literature,[513,514,517] and therefore with the proposed revised structure.

Moreover, to give further support to the revised configurational arrangement of palau'amine, we performed single-point GIAO calculations using the mPW1PW91 functional and the 6-31G(d,p) basis set.[84] The analysis of model compounds was carried out taking into consideration the calculated ^{13}C values in DMSO (Table 8.2).

It is noteworthy that a high accuracy in reproducing the experimental chemical shifts is provided by this level of theory, which has proved successful in the configurational analysis of several natural compounds.[523,524,525,526]

Table 8. 2 Comparison between experimental ^{13}C chemical shifts in DMSO[518] with calculated ^{13}C chemical shifts in DMSO of diastereoisomers **229a** and **229b**, and their $^{a}|\Delta\delta|$ values.

		229a		**229b**	
carbon	$\boldsymbol{\delta_{Exp}}$	$\boldsymbol{\delta_{calc}}$	$^{a}\|\Delta\delta\|$, ppm	$\boldsymbol{\delta_{calc}}$	$^{a}\|\Delta\delta\|$, ppm
6	67.5	73.0	5.5	67.1	0.4
10	79.4	84.7	5.3	84.0	4.6
11	55.3	54.8	0.5	54.7	0.6
12	40.6	42.9	2.3	39.1	1.5
13	44.8	51.0	6.2	46.2	1.4
16	70.3	77.5	7.2	73.1	2.8
17	73.7	69.9	3.8	76.4	2.7
18	47.0	41.1	5.9	51.5	4.5
19	40.2	43.5	3.3	44.7	4.5
20	82.2	85.6	3.4	83.7	1.5
^{b}MAE ^{13}C			**4.3**		**2.5**

$^{a}|\Delta\delta| = |\delta_{Exp} - \delta_{calc}|$, absolute differences for experimental *versus* calculated ^{13}C NMR chemical shifts.

^{b}Mean Average Error $= \sum[|\delta_{exp} - \delta_{calc}|]/n$.

Because such accuracy is seldom observed on sp^2 carbon atoms, they have not been reported and considered in the configurational assignment in the present paper and in preceding contributions reported in the literature.[525] In particular, for what concerns ^{13}C calculated results, we have considered the $\Delta\delta$ parameter (differences in experimental *vs* calculated ^{13}C NMR chemical shifts) and the MAE parameter (MAE = $\sum[|\delta_{exp} - \delta_{calc}|]/n$, summation through *n* of the absolute values of the differences in the correspondingexperimental and calculated ^{13}C chemical shifts); such parameters have been successfully used in the characterization of unknown stereostructures by us and by other research groups.[76] The analysis of Table 8.2 suggests that the structure of palau'amine is **229b**. In fact, taking into consideration the calculated

chemical shifts, there is a very large difference, as shown in Table 8.2, between the $|\Delta\delta|$ ^{13}C cs values of the two diastereoisomers. It was observeed, in fact, that the larger differences in $|\Delta\delta|$ ^{13}C of **229a** and **229b** not only regard C-12 of the junction (2.3 *vs* 1.5 respectively) but, as shown in Figure 8.3, also C-6, C-13, C-16, C-20 (5.5 *vs* 0.4, 6.2 *vs* 1.4, 7.2 *vs* 2.8, and 3.4 *vs* 1.5 respectively), while the values of absolute differences for experimental *versus* calculated ^{13}C NMR chemical shifts are comparable for the other carbon atoms.

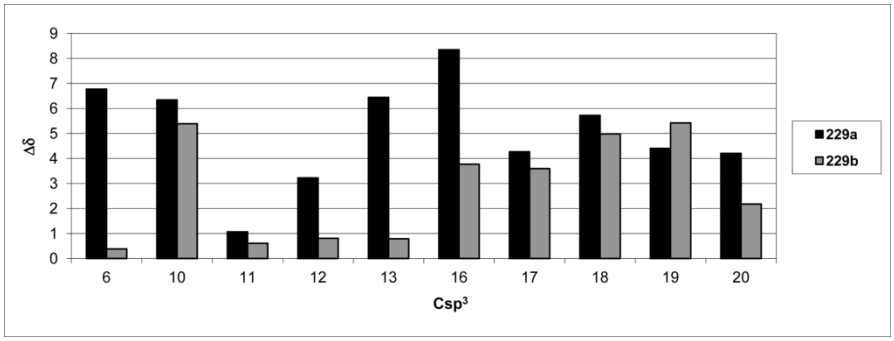


Figure 8. 3 $|\Delta\delta|$ values (parts per million) of the experimental (DMSO) *versus* theoretical carbon chemical shifts (DMSO) for sp^3 Cs of diastereoisomers **229a** and **229b**.

Finally, through the evaluation of the MAEs for **229a** and **229b** models (4.3 *vs* 2.5 respectively), the GIAO calculated ^{13}C chemical shifts indicated that compound **229b** fits better with the experimental data, confirming the revised structure of palau'amine recently proposed with a trans junction between the rings D and E, and the chlorine atom on the *β*-face of the molecule.

8.1.1 Computational details

Molecular mechanics/dynamics calculations on each of the compounds under examination, compounds **227** and **228** for kedarcidin chromophore and

compounds **229a** and **229b** for palau'amine, were performed at different temperatures (between 400 and 800 K/10 ns) on on quad-core Intel® Xeon® 3.4 GHz using the MM3[522] force field in the first case and the MMFFs[158] in the second case within the MacroModel software package.[159] All the structures so obtained (numbering 100) for each structure were minimized using the Polak-Ribier conjugate gradient algorithm (PRCG, 1000 steps, maximum derivative less than 0.05 kcal/mol). This led to the selection of the lowest energy minimum atropisomers of **227** and **228**, namely **227a** and **227b** and **228a** and **228b** (Figure 8.1) for kedarcidin chromophore, and of the lowest energy minimum conformers for palau'amine. Such geometries were in accordance, in both cases, with the results of a parallel conformational search performed with the Monte Carlo multiple minimum (MCMM) method of the MacroModel software package (MM3 force field for kedarcidin and MMFFs for palau'amine, 50 000 steps). The empirical geometries either of the minimum energy conformers **227a, 227b, 228a**, and **228b** (kedarcidin chromophore) or of the minimum energy stereoisomers **229a** and **229b** (palau'amine) were optimized at the DFT mPW1PW91 levelusing the 6-31G(d) basis set[84] (Gaussian 03 software package).[160] *J*-coupling calculations were performed using the mPW1PW91 functional and the 6-311++G(2df, 2p) basis set for a significant fragment of kedarcidin core (C-1, C-8, C-9, C-10, C-11, C-12, O-9, O-10, O-11) and for palau'amine; in all the cases, the previously optimized geometry at the mPW1PW91/6-31G(d) level[84] was considered as input. For kedarcidin chromophore, single-point calculations using the IEF-PCM solvent continuum model, as implemented in Gaussian (DMSO),[527] were performed for estimating the relative energy differences between **227a** and **227b**, and **228a** and **228b**.

For palau'amine, moreover, single-point ^{13}C cs calculations, carried out using inputs of the mPW1PW91/6-31G(d)[84] optimized structures, were performed employing the same functional combined with the 6-31G(d,p)[84] either *in vacuo* or using the IEF-PCM solvent continuum model, as implemented in Gaussian (DMSO and D_2O solvent).[527] The calculated values of chemical shifts of palau'amine were referred to the theoretical tetramethylsilane ^{13}C cs value (previously optimized at DFT level), computed at the same level of theory, except for carbon 17, which was referred to the chlorocyclohexane.

8.2 Quantitative NMR-derived interproton distances combined with quantum mechanical calculations of ^{13}C chemical shifts in the stereochemical determination of conicasterol F, a nuclear receptor ligand from *Theonella swinhoei.*

Historically, natural products have represented a principal source of therapeutic agents, and so complete knowledge of their constitution and three dimensional structure is necessary to design and characterize new molecular platforms with potential utility in therapy. In this context, NMR spectroscopy together with modern computational techniques represents an efficient and effective approach to stereochemical determination in solution. It is absolutely critical that the structural information obtained from NMR spectroscopy is as accurate and reliable as possible, because when it fails to provide an unequivocal answer, the traditional fallback position is the time-intensive and expensive solution of total synthesis. In NMR spectroscopy, the Nuclear Overhauser Effect (NOE) is undoubtedly a powerful tool for establishing stereochemical and conformational details of chemical structures. However, because many factors may perturb NOE intensities, including spin diffusion, additional cross-relaxation pathways, selective polarisation transfer, variation in τ_c between spins, accuracy of signal integration and conformational flexibility,[96] the analysis of NOE data are generally qualitative (NOE/no NOE) or semi-quantitative (strong/medium/weak) when applied to stereochemical determination of small molecules such as natural products.[96] On the other hand, recent improvements in NMR hardware (non-quadrature detection, improved RF generation and digital receivers etc), NOE experimental methods, (*e.g.* 1D-DPFGSE[99,100,101] sequences, zero-quantum suppression)[102]

and data analysis, such as the PANIC method of Macura[108,109] have made quantitative measurements of NOEs more reliable. On this basis, it was recently reported[103,104,105] that many of the 'perturbing' factors outlined above do not contribute substantially to NOE experiments conducted on small molecules, and indeed the NOE can be applied quantitatively with surprisingly high levels of accuracy of interproton distances in both rigid and flexible organic test molecules. Specifically, the accurate NOE-distance methodology compares the relative NOE intensities (and hence relative build-up rates) for pairs of spins in transient NOESY (or ROESY) experiments when the molecule of interest is in the fast tumbling regime and the measurements are made within the Initial Rate Approximation limits.[103] The relative intensity values are obtained from the standardization of each NOE peak intensity *versus* the irradiated peak in the same selective inversion experiment, because in this way any perturbation, which proportionately affects all spins in a given experiment, is minimized. In fact, Macura[109] et al. and others[110] have highlighted that employing relative, rather than absolute, intensities of NOEs (the so-called 'PANIC' method) from within a single experiment corrects for other forms of relaxation in 1D- or 2D-NOESY experiments and effectively extends the period during which the Initial Rate Approximation holds for the relative values of NOE enhancements.[528]

Also recently, the QM/NMR[81,82] has demonstrated success[76,529] as a tool for differentiating diastereoisomers, to understand the mechanism of action of a determinate molecule,[505] as support to the total synthesis,[507] or to determine the structure of unusual natural substances,[530] being able to identify the correct configuration in an efficient and accurate fashion. In the QM/NMR approach, DFT values of NMR parameters, *i.e.* chemical shifts (^{13}C and ^{1}H) and scalar coupling constants ($^{3}J_{\text{H-H}}$ and $^{2,3}J_{\text{C-H}}$), are compared with experimental values

to assign constitution and configuration. In fact, in many cases if the system is treated at a sufficient level of theory, the correct molecular arrangement will predictably be the isomer characterized by the best match between calculated and measured spectroscopic properties, where the best fit is identified using statistic parameters such as mean absolute error (MAE), Total Absolute Deviation (TAD), or others.[531] In particular, the chemical shift values are especially diagnostic in cases where there are multiple quaternary centres – which strongly limit stereochemical analysis by traditional coupling constant and/or NOE analysis.

On this basis, the aim was to combine these two different validated approaches (accurate NOE-distance analysis and QM/NMR analysis) to obtain a new efficient and robust method for the stereostructural determination of the organic compounds (Figure 8.4).

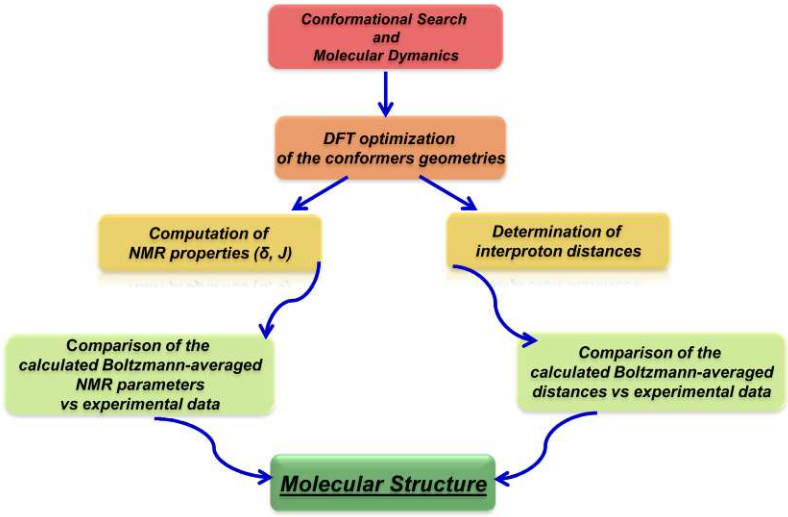


Figure 8. 4 General protocol for combining quantitative interproton distances by NOE/ROE with quantum chemical calculations of NMR parameters in stereochemical determination.

The combination of the two methods allows the stereochemistry to be determined by two experimentally independent methods, providing greater confidence in the final structural assignment. This is especially important in cases where one technique or the other is not completely unequivocal in distinguishing all structural candidates. Here, it was described the results of the first application of the combined method to unknown structures, and in particular the stereostructure assignments of conicasterol F (**230**) and theonellasterol I (**231**) (Figure 8.5), two new 4-methylene polyhydroxylated steroids isolated from the marine sponge *Theonella swinhoei*.[532] These compounds are especially challenging for stereochemical determination (*vide infra*) due to the 5 contiguous, quaternary stereogenic centres from C8-C9-C10-C13-C14.

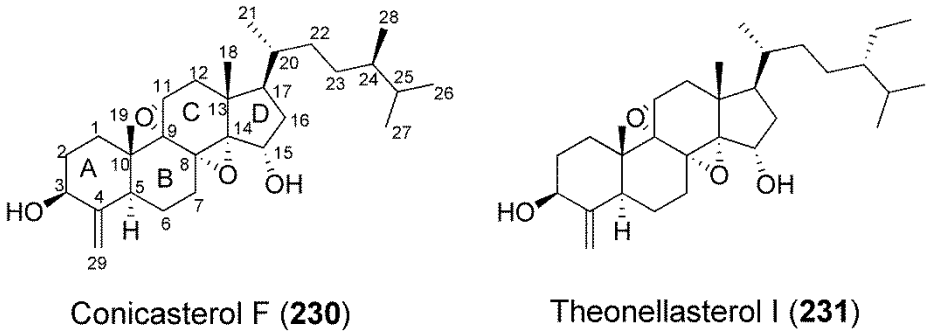


Figure 8. 5 Conicasterol F (**1**) and theonellasterol I (**2**) from the marine sponge *Theonella swinhoei*.

Marine sponges of the genus Theonella have attracted great interest from the scientific community for the impressive variety of bioactive secondary metabolites with unusual structures and powerful biological activity. In particular, the recent exploration of a specimen of *Theonella swinhoei* collected at the Solomon islands led to the isolation of new cyclic peptides

perthamides C-F,[458,459] and solomonamides,[460] endowed with anti-inflammatory activity. From the same sponge two sulfated sterols, solomonsterols A and B and new theonellasterols[294,533] and conicasterols,[533] have shown to be potent ligands of human nuclear receptors with pharmacological potential in the treatment of immune-driven inflammatory bowel diseases.[473]

The molecular formula of conicasterol F (**230**) was established as $C_{29}H_{46}O_4$ from HR ESIMS (*m/z* 465.3532, calcd. 465.3556 $[M+Li]^+$), and ^{13}C NMR data (Table 8.8). The seven degrees of unsaturation implied by the formula were ascribed to six rings and one C=C bond (δ_C 151.6 and 104.5). The ^1H NMR spectrum showed signals for six Me groups [(δ_H 0.66 (s), 0.73 (s), 0.78 (d), 0.86 (d), 0.91 (d) and 0.84 (d)], two olefinic H-atoms (δ_H 4.64 and 5.30) and one O-bearing CH group (δ_H 3.67), consistent with a 3-hydroxy-4-methylene-24-methyl steroidal system. The HMBC correlations from CH_2-29 to C3, C4 and C5 and the allylic ^1H-^1H COSY correlations of CH_2-29 with H-3 and H-5 confirmed the presence of exocyclic CH_2 group at C4. In addition to the signals arising from the 4-methylen-3β-hydroxy ring A, the ^{13}C NMR spectrum of **1** (Table 8.8) showed the presence of one oxygen-bearing methine carbon (δ_C 65.0) and three oxygen-bearing quaternary carbons (δ_C 79.8, 61.5, 60.8). The HMBC correlation between the angular methyl Me-18 protons and the carbon at δ_C 79.8 indicated that C-14 was an oxygen-bearing carbon, whereas the position of the secondary alcoholic function at C15 (δ_C 65.0, δ_H 4.18) was deduced by ^1H-^1H COSY spin system from H17 (δ_H 1.61) to H15 (δ_H 4.18). The HMBC correlations of H-7 at δ_H 1.86 with C-14, C-8 (δ_C 60.8) and C-9 (δ_C 61.5) suggests an epoxide ring between C14 and C-8 and the presence of an oxygen bearing carbon at C9 position which is further substantiated by an HMBC crosspeak Me-19/C-9 (δ_C 61.5).

Table 8. 3 ^{1}H and ^{13}C NMR data (500 and 700 MHz, C_6D_6) for conicasterol F (**230**).

Conicasterol F (**1**)[a]		
Position	δ_H	δ_C
1α	1.34 ovl	30.3
1β	0.88 m	
2α	1.74 ovl	32.2
2β	1.27 m	
3	3.67 dd (4.6, 11.4)	72.6
4	-	151.6
5α	1.99 br d (11.6)	45.3
6α	1.38 m	22.6
6β	1.45 m	
7α	1.86 ovl	27.3
7β	2.26 ddd (1.9, 4.3, 14.3)	
8	-	60.8
9	-	61.5
10	-	39.5
11β	2.51 dd (2.1, 7.3)	50.1
12α	1.92 d (14.0)	35.7
12β	1.67 dd (7.3, 14.0)	
13	-	44.9
14	-	79.8
15β	4.18 dd (1.4, 8.8)	65.0
16α	2.07 m	40.9
16β	1.77 ovl	
17	1.61 m	50.6
18	0.66 s	14.8
19	0.73 s	18.5
20	1.21 m	35.0

21	0.78 d (6.7)	18.7
22	1.24 ovl 1.47 ovl	33.4
23	1.23 ovl 1.34 ovl	30.9
24	1.25 ovl	39.3
25	1.54 m	32.7
26	0.86 d (6.7)	18.6
27	0.91 d (6.7)	20.6
28	0.84 d (6.5)	15.8
29a	4.64 br s	104.5
29b	5.30 br s	
[a]Coupling constants are in parentheses and given in Hertz. ^{1}H and ^{13}C assignments were made based on COSY, HSQC, HMBC and ROESY experiments as described in the main text.		

The additional epoxy functionality at C9/C11 of ring C is suggested in the ^{1}H NMR spectrum by a resonance at δ_H 2.51 (dd, *J*= 7.3, 2.1 Hz) that is correlated in the HSQC spectrum with C-11 at δ_C 50.1 and which showed COSY crosspeaks exclusively with the protons at δ_H 1.67 and 1.92 assigned to H_2-12. These data thus also account for the additional degree of unsaturation established on the basis of mass data.

The stereochemical assignment of the majority of the steroid skeleton is straightforward. The absolute steroidal configuration as depicted in Figure 8.14 was assumed on biogenetic grounds. The coupling constants between H-3 [δ_H 3.67 (dd, *J*=4.6, 11.4 Hz)] and H_2-2, and the ROESY correlations H-3/H-5α, indicated that H-3 was axial and therefore the OH-3 was β-oriented. The axial disposition of H-11 [δ_H 2.51 (dd *J* =2.1, 7.3 Hz)] and, consequently, the

α-orientation of the C9/C11 epoxide ring, was evident from the $^{3}J_{HH}$ vicinal coupling to H_2-12 and ROESY correlations to axial protons Me-19, H-8 and Me-18. The α-orientation of 15-OH was similarly assigned on the basis of a ROESY correlation between H-15 and Me-18. The configuration at C-24 on the flexible side chain was determined was determined to be the same of conicasterol by comparison of ^{1}H and ^{13}C chemical shifts.[533]

Theonellasterol I (**231**) was isolated as an optically active pale yellow oil and a molecular formula of $C_{30}H_{48}O_4$, was established by high resolution mass spectrometry. The ^{1}H and ^{13}C NMR spectra of **231** were very similar to those of conicasterol F (Table 8.4), with the data for the steroidal core being essentially identical ($\Delta\delta$ <0.5 ppm).

Table 8. 4 ^{1}H and ^{13}C NMR data (500 and 700 MHz, C_6D_6) for theonellasterol I (**231**).

Position	δ_H	δ_C	HMBC
1α	1.33 ovl	30.1	-
1β	0.89 m		
2α	1.74 ovl	32.0	-
2β	1.28 m		
3	3.65 dd (4.7, 11.3)	72.5	C4
4	-	151.1	-
5α	1.99 br d (11.6)	45.2	C30
6	1.38 m 1.44 m	22.9	-

7α	1.87 ovl	27.3	C6, C8, C14
7β	2.26 ddd (1.8, 4.3, 14.5)		
8	-	60.8	-
9	-	61.6	-
10	-	39.7	-
11β	2.50 dd (2.1, 7.3)	49.8	C12
12α	1.93 d (14.1)	35.5	C9, C13, C14, C17, C18
12β	1.66 dd (7.3, 14.1)		
13	-	45.0	-
14	-	79.9	-
15β	4.18 d (8.7)	65.0	-
16α	2.08 m	40.7	C13, C14, C15, C17
16β	1.78 ov		
17	1.61 m	50.4	-
18	0.66 s	14.5	C12, C13, C14, C17
19	0.72 s	18.4	C1, C5, C9, C10
20	1.20 m	35.0	-
21	0.80 d (6.5)	18.5	C17, C20, C22

22	0.93 m 1.34 ovl	33.3	-
23	1.08 m 1.36 ovl	26.7	-
24	0.99 m	46.2	-
25	1.72 ovl	29.1	-
26	0.89 d (7.1)	19.1	C24, C25, C27
27	0.91 d (7.1)	19.7	C24, C25, C26
28	1.18 m 1.38 m	23.3	-
29	0.93 t (7.4)	12.4	C24, C28
30	4.64 br s 5.31 br s	104.5	C3, C4, C5
[a]Coupling costants are in parentheses and given in hertz. ^{1}H and ^{13}C assignments aided by COSY, TOCSY, ROESY, HSQC and HMBC experiments.			

The only difference between these two molecules lies in the steroidal side chain, with an ethyl group replacing the C24 methyl in **230**. On the basis of chemical shift similarities, both the constitution and stereochemistry of the tetracyclic nucleus of theonellasterol I (**231**) was assumed to be the same as conicasterol F (**230**). On the other hand, the configuration at C24 were determined by comparison of ^{13}C-NMR data with the epimeric steroidal side chain.[534,535]

8.2.1 Assignment of the relative configuration of 8-14 epoxy ring in conicasterol F

The presence of four contiguous quaternary centres in ring C of **230** severely hampers the stereochemical assignment of the relative configuration of the 8-14 epoxy ring, with the lack of protons at C8 and C14 making conventional $^{3}J_{HH}$ and NOE analysis impossible. Therefore this molecule was selected as a suitable case study to test the proposed combination of quantitative ROE-derived interproton distances with QM calculation of NMR parameters (Figure 8.4). The relative rigidity, the molecular size, and the solubility in non-viscous benzene, should ensure appropriate conditions for accurate interproton measurements by ROE.[103] On the other hand, because the only difference between the two possible diasteroisomers of the compound **230** is the stereochemistry of C8 and 14 (Figure 8.15), substantial chemical shift differences are expected for the carbon atoms close to the junction between C-D rings, suggesting the potential for stereostructure validation by GIAO calculation of ^{13}C NMR chemical shifts.

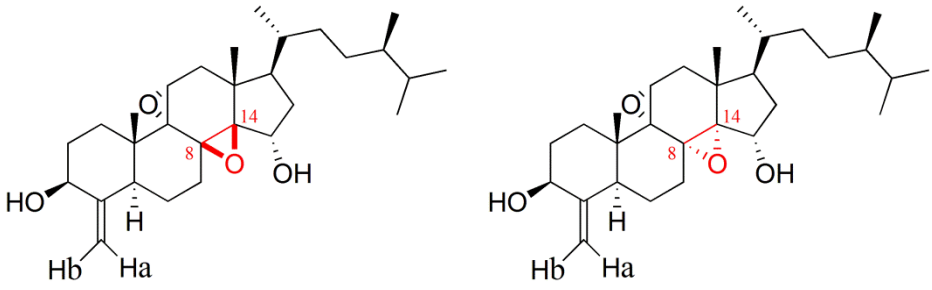


Figure 8. 6 Molecular structure of the two possible diasteroisomers (**230a** and **230b**) of conicasterol F.

Following our protocol (Figure 8.4), molecular dynamics and Monte Carlo conformational search calculations (see Computational Details) were performed on both possible stereoisomers **230a** and **230b** (Figure 8.6), using the MMFF[158] force field (MacroModel software package).[159] In each case, only a single conformer of the steroid ring system was found. The resulting geometries for **230a** and **230b** were optimized (Figure 8.7) at the DFT level using the MPW1PW91 functional and 6-31G (d)[84] basis set (Gaussian 09 Software Package).[536] Calculated interproton distances were obtained directly from the DFT-optimised geometries without further corrections.

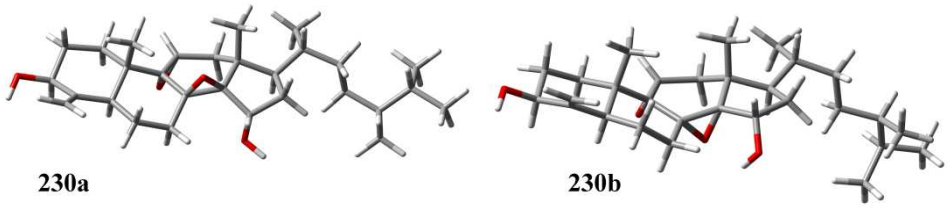


Figure 8. 7 Optimized geometries of diastereoisomers **1a** and **1b** of conicasterol F.

Quantitative experimental interproton distances for **230** were obtained from 1D-ROESY spectra.[103] Figure 8.8 shows an example of the data obtained, with the selective 1D-ROESY spectrum of H29a, with clear ROE enhancements for H29b, H6α, H6β and Me19 of **230**.

Selective 1D-ROESY experiments were performed only on the key protons[537] around the junction C-D (See Figure 8.5), namely proton H7β, H11β, H15β, Me18 (See Figures 8.8-12)

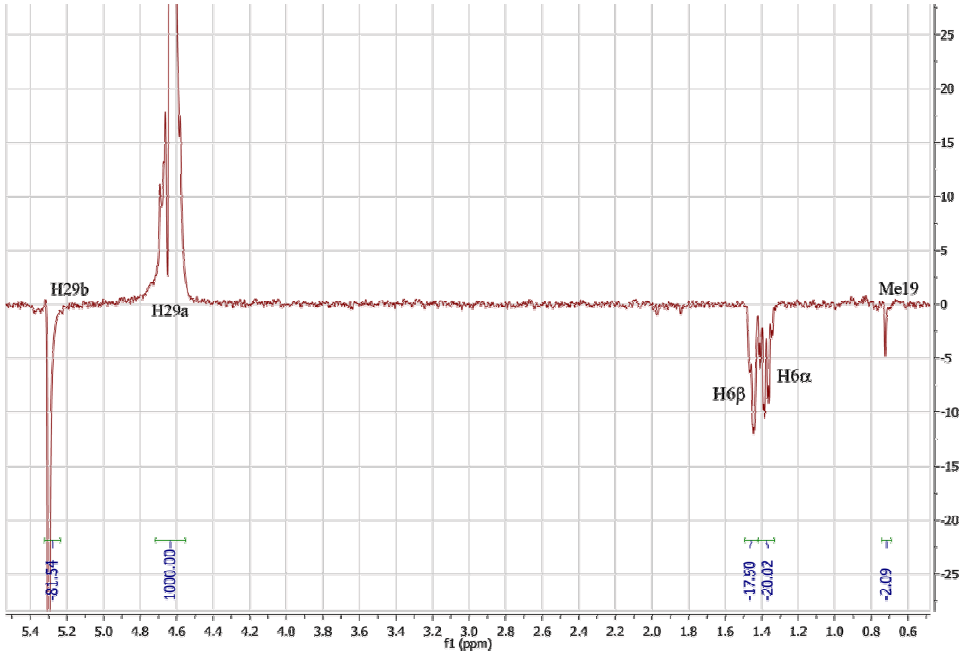
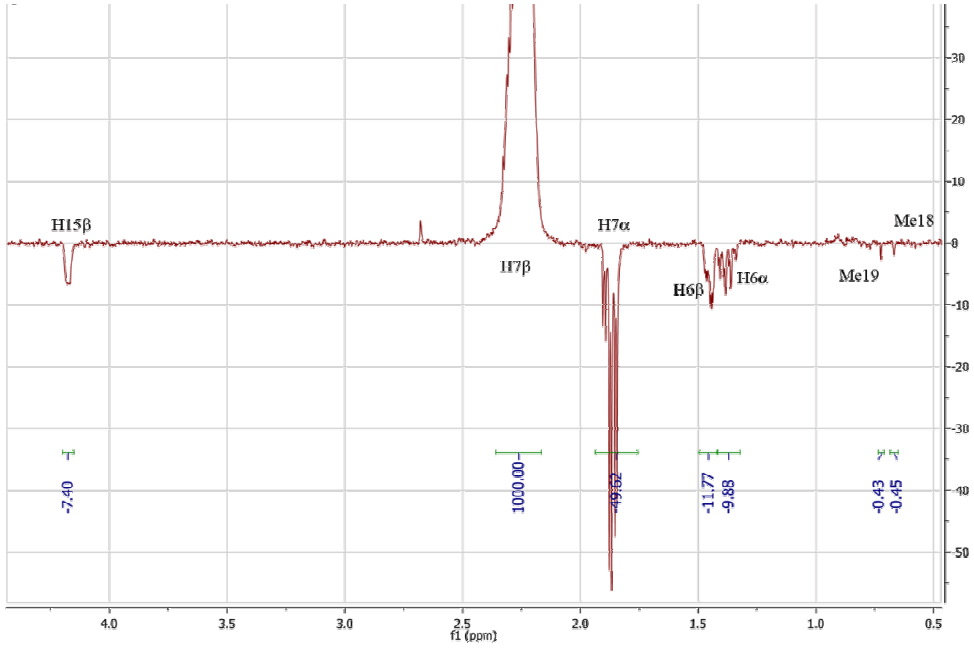
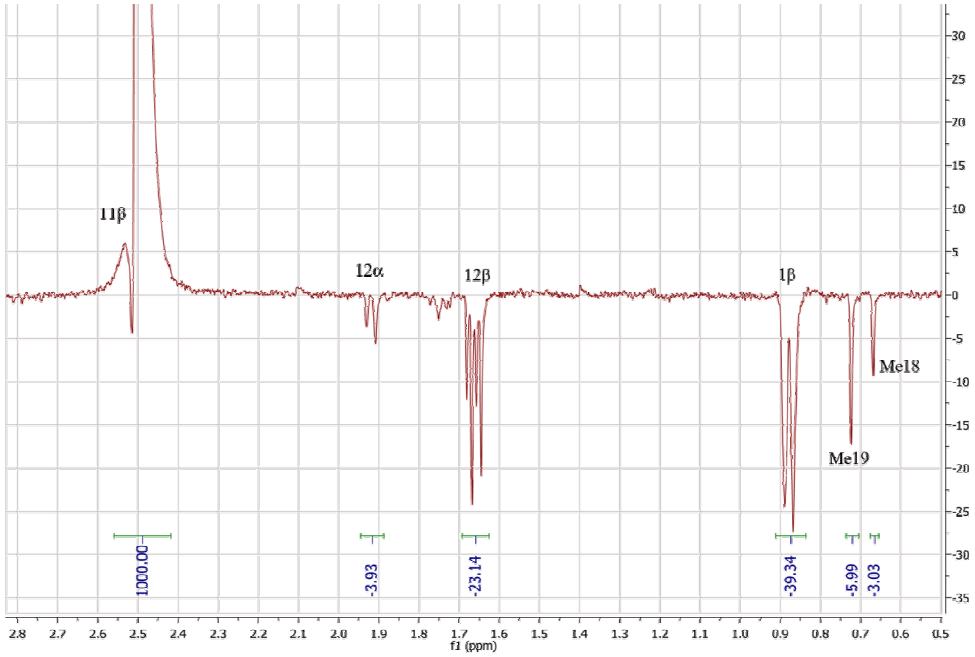


Figure 8. 8. Selective 1D-ROESY spectrum (600 MHz, C_6D_6) of H29a of conicasterol F.



ɔnicasterol F.



of conicasterol F.

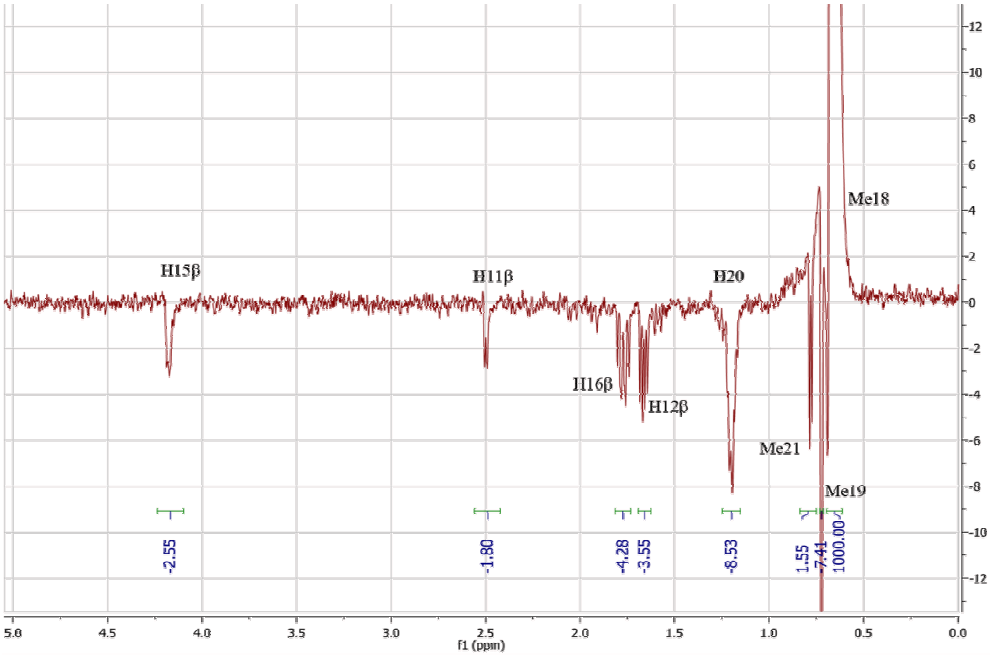


Figure 8. 11 Selective 1D-ROESY spectrum (600 MHz, C_6D_6) of Me18 of conicasterol F.

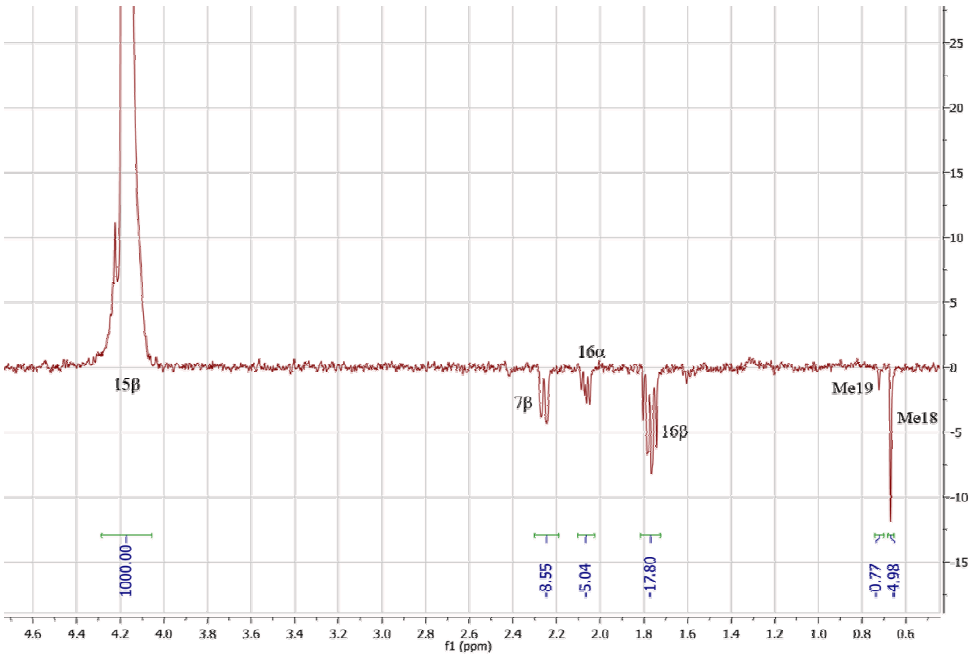


Figure 8. 12 Selective 1D-ROESY spectrum (600 MHz, C_6D_6) of H15β of conicasterol F.

The 1D-ROESY spectra were calibrated in the PANIC fashion[109,110] by setting the integral value for the irradiated peak in each case arbitrarily to 1000[103] and thus standardizing the absolute values of the ROE intensities to this in every 1D-ROESY spectrum. Interproton distances were then calculated using Equation 1.22

$$\frac{\eta_{IS}}{\eta_{ref}} = \frac{{r_{IS}}^{-6}}{{r_{ref}}^{-6}}$$

Equation 1. 22

where η_{IS} is the intensity of the NOE for a given proton pair I and S, r_{IS} is the corresponding interproton distance, η_{ref} and r_{ref} are reference values for a single chosen NOE for which the interproton distance is assumed based on

geometric constraint. The methylene protons H29a-H29b was chosen as the reference ROE (η_{ref} = 81.54) for the 1D-ROESY dataset with the corresponding intermethylene distance of 1.85 Å as the reference distance (r_{ref}).

Applying equation 1.22 in turn to each ROE resonance (η_{IS}) observed in the 1D-ROESY spectra, allows the calculation of the corresponding interproton distance (r_{IS}). Methyl groups were treated by using r^{-3} averaging as described in reference 96, as this gives slightly improved fits over r^{-6} averaging on molecules of the size of conicasterol. A subset of these values are shown in Table 8.5 which were identified as useful for stereochemical structure elucidation *i.e.* those where DFT-calculated interproton distances in **230a** and **230b** differed by >0.03Å (~1%) from each other.

On the basis of the data in Table 8.5, it is clear that diastereomer **230b** shows the best fit with the estimated experimental data (MAE 3.0%, standard deviation (STD) 2.6%), while diastereomer **230a** is clearly less satisfactory (MAE 7.8%, STD 5.9%). Critically, the MAE and STD values for **230b** fall within the expected range of errors previously reported[103] for this accurate NOE/ROE-distance method (MAE and STD both <4%), while those for **230a** fall well outside this range.

As might be expected, the largest contributions to the MAE for **1a** (underlined in the Table 8.10) relate to the protons close to the C-D junction (Figure 8.14), e.g. H7β-H15β, and Me18-H15β, where the conformation of ring C is influenced by the relative stereochemistry of the epoxide rings.

However, it must be emphasised that in both of these cases (H7β-H15β, and Me18-H15β), a qualitative ROE/NOE analysis would not have differentiated between the diastereomers, as the ROE peaks would have been present, but with very weak intensities for either diastereomer.

Table 8. 5 Interproton distances determined by ROE for H7β, H11β, H15β, Me18 and H29a-b of conicasterol F (**230**) in d_6-benzene and comparison with DFT calculated value for **1a** and **1b**. Value in bold was used to calibrate the ROEs.

		Exp.	**230a**		**230b**	
Proton		R_{ROE}(Å)	R_{calc}(Å)	ABS % error[a]	R_{calc}(Å)	ABS % error[a]
H29a	**H29**b	**1.85**	**1.85**	**-**	**1.85**	-
H7β	H15β	2.73	2.45	11.4%	2.71	0.5%
H11β	H12α	3.07	2.87	6.9%	2.92	5.1%
H11β	Me19	3.43	3.83	10.4%	3.20	7.4%
H11β	H1β	2.09	2.26	7.4%	2.12	1.7%
H15β	H16α	2.94	2.69	9.3%	2.87	2.5%
H15β	H16β	2.38	2.48	4.0%	2.32	2.9%
Me18	H15β	3.42	4.37	21.7%	3.18	7.5%
Me18	H11β	3.67	3.71	1.2%	3.54	3.7%
Me18	H12β	3.12	2.89	8.1%	2.93	6.3%
Me18	H16β	3.02	3.09	2.2%	3.01	0.4%
Me18	H20	2.70	2.61	3.1%	2.74	1.6%
Me18	Me19	3.31	4.00	17.2%	3.28	1.0%
H29a	H6α	2.34	2.44	4.2%	2.35	0.5%
H29a	H6β	2.39	2.33	2.7%	2.38	0.5%
[b]MAE				7.8%		3.0%
STD				5.9%		2.6%

[a]|% error|= $|r_{calc}-r_{ROE}|/r_{calc}$, absolute differences for calculated *versus* ROE derived distances/calculated distances

[b]MAE=∑[%error|]/n.

2D-ROESY spectra were also obtained for **230** (See Figures 8.13-18) and confirmed the trend of the 1D-ROESY data.

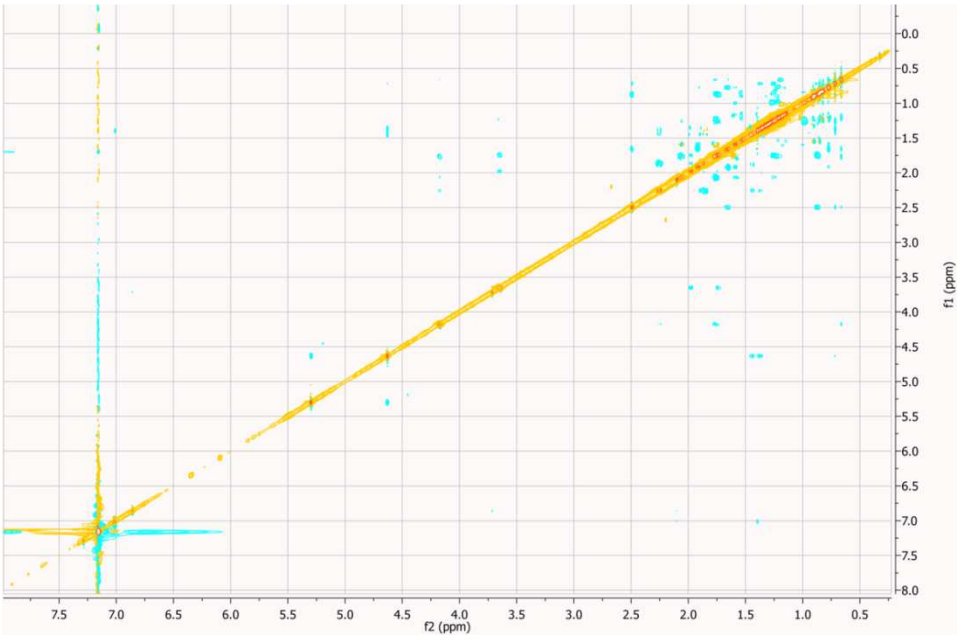


Figure 8. 13 2D-ROESY spectrum (600 MHz, C_6D_6) of conicasterol F.

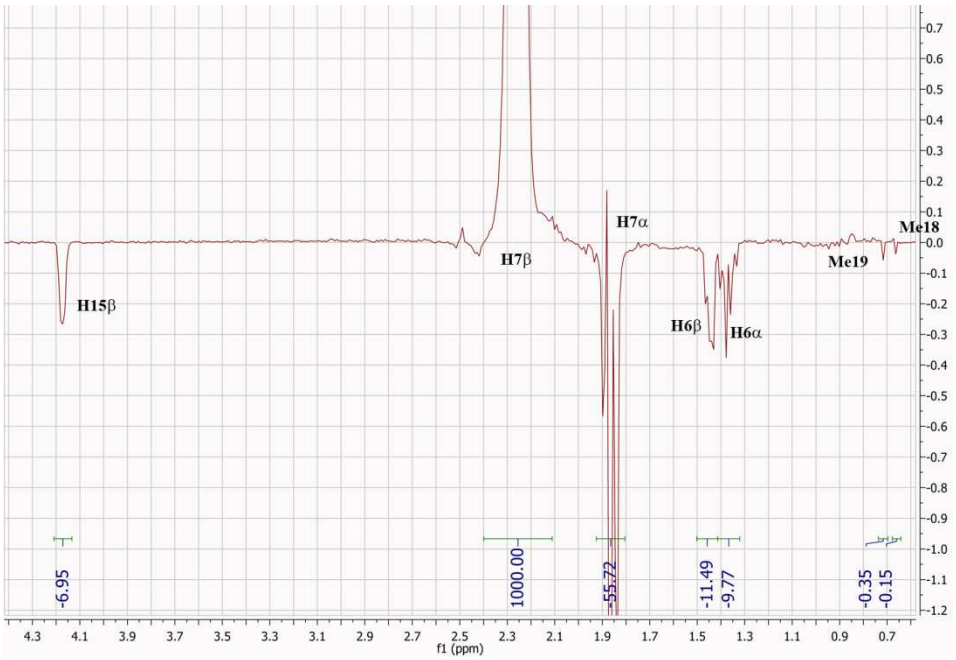


Figure 8. 14 2D-ROESY F2-slice at F1 chemical shift (600 MHz, C_6D_6) of H7β of conicasterol F.

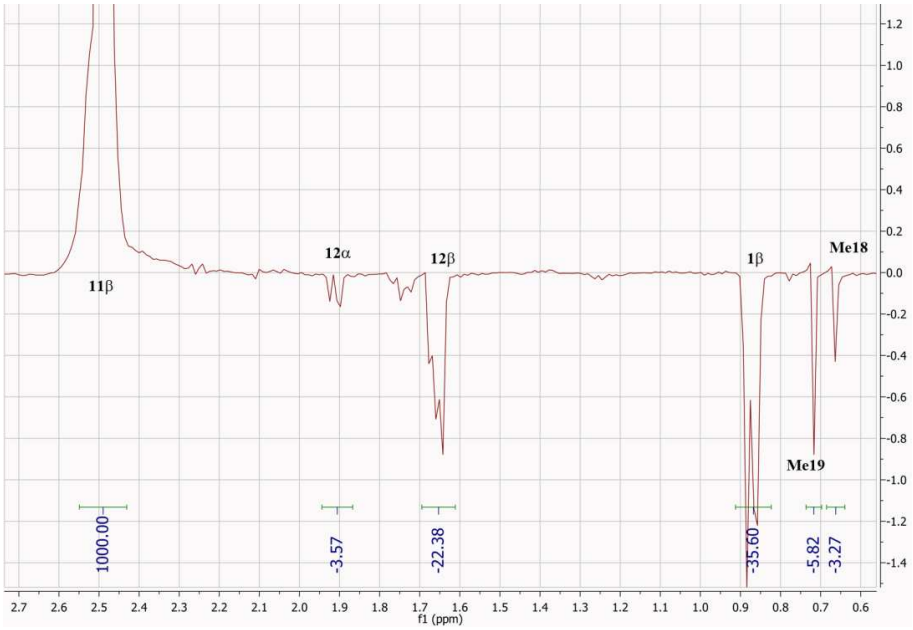


Figure 8. 15 2D-ROESY F2-slice at F1 chemical shift (600 MHz, C_6D_6) of H11β of conicasterol F.

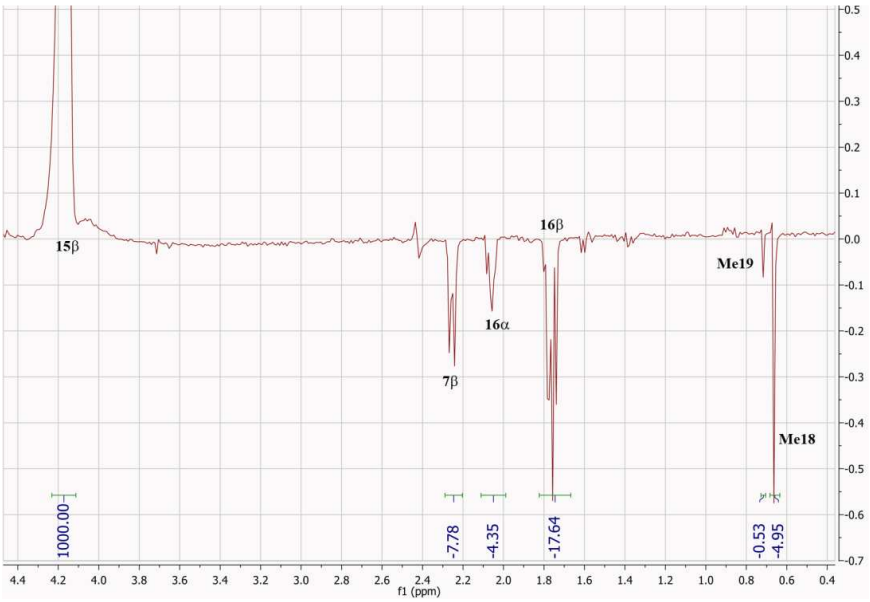


Figure 8. 16 2D-ROESY F2-slice at F1 chemical shift (600 MHz, C_6D_6) of H15β of conicasterol F.

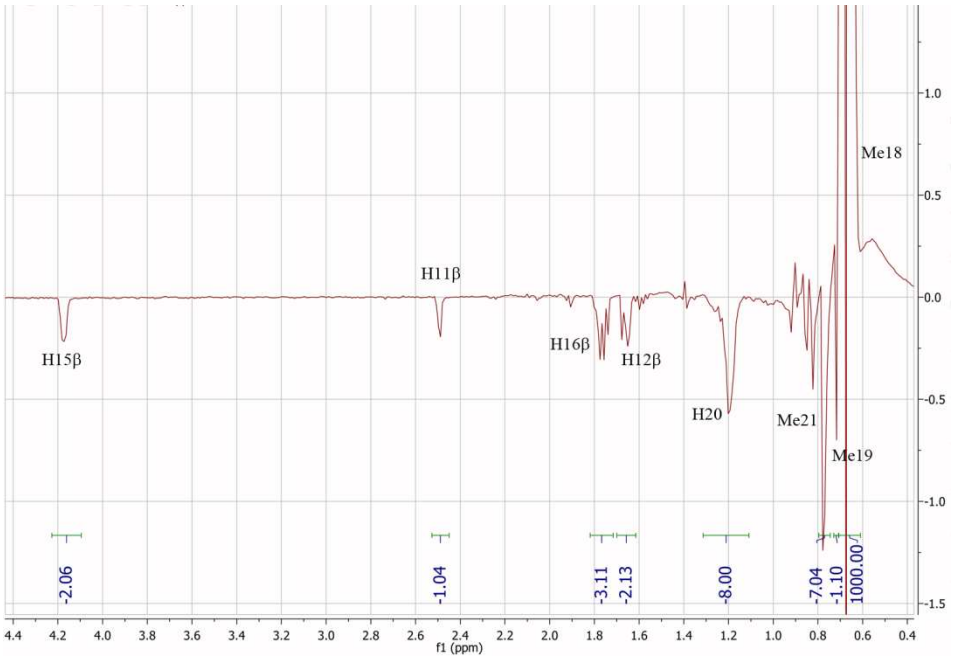


Figure 8. 17 2D-ROESY F2-slice at F1 chemical shift (600 MHz, C_6D_6) of Me18 of conicasterol F.

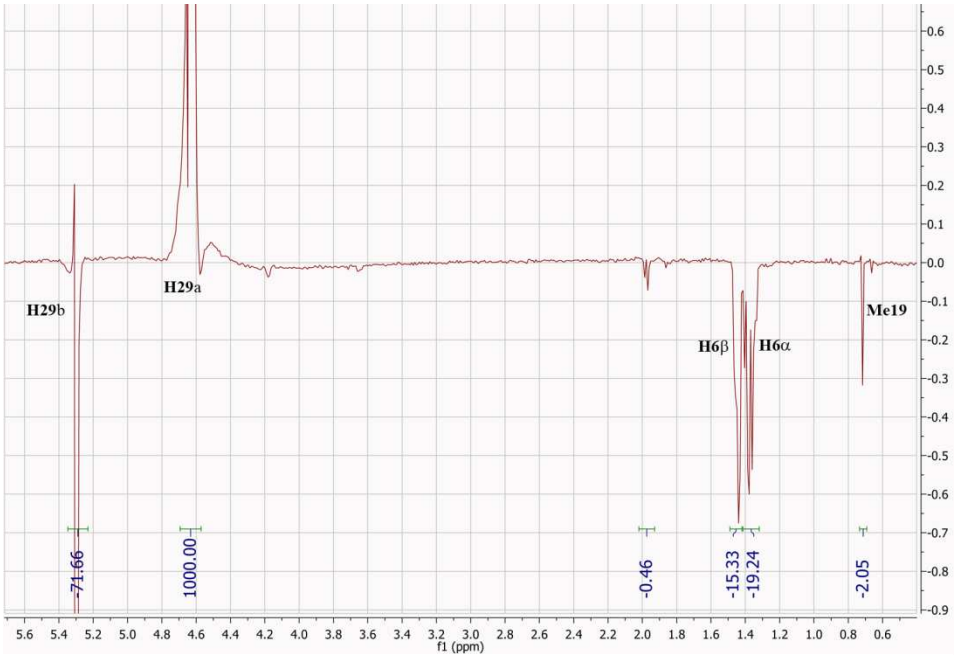


Figure 8. 18 2D-ROESY F2-slice at F1 chemical shift (600 MHz, C_6D_6) of H29a of conicasterol E (**1**).

Informative plots of computed *vs* ROE-determined interproton distances for each diastereomer (Figure 8.19) clearly show that **1b** provides a far superior fit for both 1D- and 2D- data.

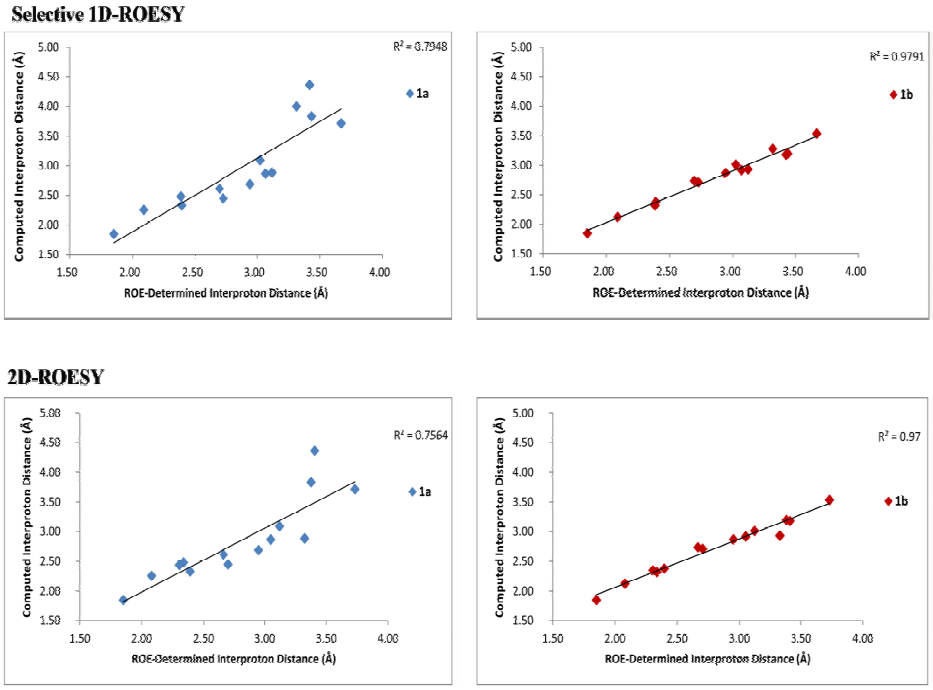


Figure 8. 19 DFT calculated interproton distances *vs* experimental from 1D-ROESY experiments (top view) and from 2D-ROESY experiments (bottom view).

However there is in fact a very slight deterioration in the quality of the overall fit from the 2D-ROESY data, for the 'correct' stereoisomers **1b** (MAE 3.9%, STD 3.5%) (Table 8.6) which is in line with our previous observations.

These encouraging results are confirmed by QM-NMR structural analysis. In particular, we performed single-point GIAO calculations on the optimized

230a and **230b** geometries (see above and Figure 8.7) using the MPW1PW91 functional and the 6-31G (d,p) basis set[84] (Gaussian 09 Software Package).[536]

Table 8. 6 Interproton distances determined by 2D-ROESY for H7β, H11β, H15β, Me18 and H29a of Conicasterol F (**230**) in C_6D_6 and comparison with DFT calculated value for both the diastereoisomers. Value in bold were used to calibrate the ROEs.

		Exp.	**230a**		**230b**	
Proton		R_{ROE}(Å)	R_{calc}(Å)	ABS % error[a]	R_{calc}(Å)	ABS % error[a]
H29a	**H29**b	**1.85**	**1.85**	**-**	**1.85**	-
H7β	H15β	2.70	2.45	10.4%	2.71	0.4%
H11β	H12α	3.05	2.87	6.3%	2.92	4.5%
H11β	Me19	3.38	3.83	11.9%	3.20	5.6%
H11β	H1β	2.08	2.26	7.8%	2.12	2.1%
H15β	H16α	2.95	2.69	9.6%	2.87	2.8%
H15β	H16β	2.34	2.48	5.9%	2.32	0.8%
Me18	H15β	3.41	4.37	22.0%	3.18	7.1%
Me18	H11β	3.73	3.71	0.5%	3.54	5.5%
Me18	H12β	3.32	2.89	15.2%	2.93	13.3%
Me18	H16β	3.12	3.09	1.0%	3.01	3.6%
Me18	H20	2.67	2.61	2.0%	2.74	2.7%
H29a	H6α	2.39	2.44	5.6%	2.35	0.7%
H29a	H6β	2.30	2.33	2.9%	2.38	1.9%
[b]MAE				7.7%		3.9%
STD				6.2%		3.5%

[a]|% error|= $|r_{calc}-r_{ROE}|/r_{calc}$, absolute differences for calculated *versus* ROE derived distances/calculated distances

[b]MAE=∑[%error|]/n.

For the ^{13}C chemical shift analysis, the attention was docused on the key carbon atoms near the C ring because this is the region likely to be influenced by the stereochemistry of the epoxy ring.

Moreover, the data evaluation was performed considering the Δδ parameter (differences in experimental *vs* calculated ^{13}C NMR chemical shifts) and the MAE parameter (MAE $=\sum[|\delta_{exp}-\delta_{calc}|]/n$, summation through n absolute values of the differences in the corresponding experimental and calculated ^{13}C chemical shifts); which have been successfully used in the characterization of unknown stereostructures by us and by other research groups.[76,529,531] Table 8.7 lists selected experimental and calculated chemical shifts for carbons near the C ring in **230a** and **230b**.

Once again, the chemical shift data strongly suggests that diastereomer **230b** is the correct structure for **230**, with a very large difference between the average |Δδ| values of the MAE values for **230a** and **230b** (3.7 *vs* 0.8 respectively).

Critically, the |Δδ| MAE for **230b** in Table 8.7 falls within the MAE calculated for the whole molecule (1.4 ppm, Table 8.8) and is comparable to the and is comparable to the expected error range (≤2.0 ppm) for QM/NMR calculations of this sort, while the |Δδ| MAE in Table 8.7 for **230a** falls outside these bounds.

Table 8. 7 Comparison between experimental ^{13}C chemical shifts in benzene with calculated ^{13}C chemical shifts in *vacuo* of diastereoisomers **1a** and **1b** and their $^{a}|\Delta\delta|$ values.

		230a		**230b**	
Carbon	δ_{exp}	δ_{calc}	$^{a}\|\Delta\delta\|$,ppm	δ_{calc}	$^{a}\|\Delta\delta\|$,ppm
7	27.3	31.4	4.1	28.1	0.8
8	60.8	62.8	2.0	61.3	0.5
9	61.5	62.9	1.4	61.6	0.1
14	79.8	72.8	7.0	79.1	0.7
13	44.9	42.6	2.3	46.1	1.2
12	35.7	41.5	5.8	36.4	0.7
11	50.1	54.1	4.0	49.6	0.5
18	14.8	19.1	4.3	15.9	1.1
15	65.0	75.3	10.3	64.3	0.7
16	40.9	40.2	0.7	41.4	0.5
17	50.6	50.5	0.1	48.1	2.5
20	35.0	33.2	1.8	35.0	0.0
$^{b}MAE\ ^{13}C$			**3.7**		**0.8**
$^{a}\|\Delta\delta\|= \|\delta_{exp}-\delta_{calc}\|$, absolute differences for experimental *versus* calculated ^{13}C NMR chemical shifts.					
$^{b}MAE= \sum[\|\delta_{exp}-\delta_{calc}\|]/n.$					

Table 8. 8 Comparison between experimental ^{13}C chemical shifts in C_6D_6 and calculated ^{13}C chemical shifts in *vacuo* of diastereoisomers **230a** and **230b** and their $^{a}|\Delta\delta|$ values.

		1a		**1b**	
Carbon	δ_{exp}	δ_{calc}	$^{a}\|\Delta\delta\|$,ppm	δ_{calc}	$^{a}\|\Delta\delta\|$,ppm
1	30.3	29.5	0.8	30.8	0.5
2	32.2	30.9	1.2	30.8	1.3
3	72.6	72.6	0.0	72.5	0.1
4	151.6	150.8	0.8	150.8	0.8
5	45.3	46.6	1.3	46.0	0.7

6	22.6	24.5	1.9	23.7	1.1
7	27.3	31.4	4.1	28.1	0.8
8	60.8	62.8	2.0	61.3	0.5
9	61.5	62.9	1.4	61.6	0.1
10	39.5	43.4	3.9	41.8	2.3
11	50.1	54.1	4.0	49.6	0.5
12	35.7	41.5	5.8	36.4	0.7
13	44.9	42.6	2.3	46.1	1.2
14	79.8	72.8	7.0	79.1	0.7
15	65.0	75.3	10.3	64.3	0.7
16	40.9	40.2	0.7	41.4	0.5
17	50.6	50.5	0.1	48.1	2.5
18	14.8	19.1	4.3	15.9	1.1
19	18.5	16.7	1.8	20.1	1.6
20	35.0	33.2	1.8	35.0	0.0
21	18.7	20.4	1.7	20.2	1.5
22	33.4	33.5	0.1	35.2	1.8
23	30.9	27.6	3.3	29.9	1.0
24	39.3	39.7	0.4	40.2	0.9
25	32.7	28.2	4.5	35.7	3.0
26	18.6	14.1	4.5	22.9	4.3
27	20.6	22.7	2.1	15.9	4.7
28	15.8	15.8	0.0	12.4	3.4
29	104.5	102.1	2.4	103.0	1.5
[b]*MAE* ^{13}C			**2.6**		**1.4**

[a]$|\Delta\delta| = |\delta_{exp} - \delta_{calc}|$, absolute differences for experimental *versus* calculated ^{13}C NMR chemical shifts.

[b]MAE= $\sum[|\delta_{exp} - \delta_{calc}|]/n$.

The graphical representation of the errors in calculated chemical shifts (Figure 8.20) clearly shows that **230b** is the best match with the experimental

data for every carbon with the one exception of C-17 where both diastereomers give calculated values which are within reasonable error limits.

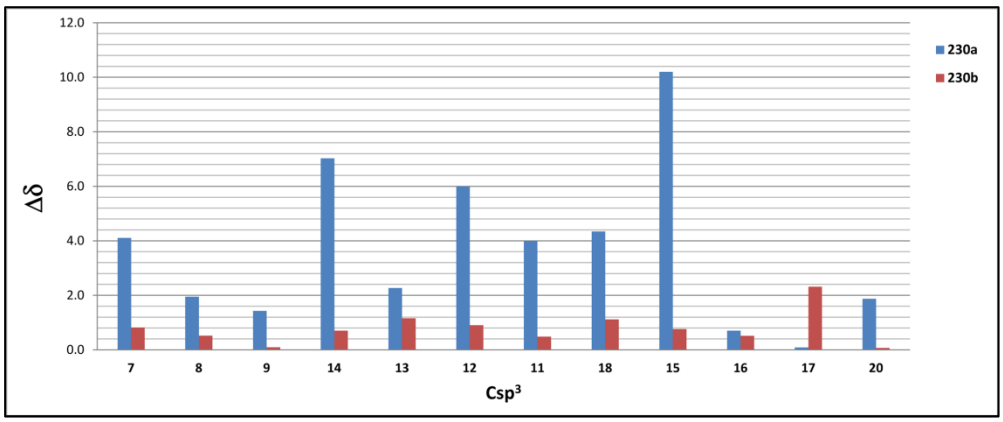


Figure 8. 20 |Δδ| values (parts per million) of the experimental (benzene) *versus* theoretical carbon chemical shifts (*vacuo*) for Csp^3 of diastereoisomers **1a** and **1b**.

In summary, the evaluation of the GIAO calculated ^{13}C chemical shifts MAEs for **230a** and **230b** geometries (3.7 *vs* 0.8 ppm respectively) confirm the stereochemical assignment made on the basis of the accurate ROE-distance analysis. Both methods independently identify a *trans* C/D ring junction, allowing the assignment of conicasterol F as **230b** depicted in Figure 8.6. Moreover, here it was demonstrated the proposed method as a powerful tool in stereochemical structure determination, and its application is strongly recommend prior to, or indeed as a complete alternative to, total synthesis when traditional NMR data analysis is not sufficient to distinguish stereochemical alternative, such as in the case of conicasterol F (**230**).

Further, conicasterol F (**230**) and theonellasterol I (**231**) are ligands of two well-known nuclear receptors, PXR and FXR, Farnesoid-X-receptor (FXR) and the pregnane-X-receptor (PXR), two nuclear receptors both functioning as bile acid activated receptors, have emerged as the main receptors involved in

regulating bile acid synthesis, detoxification and excretion in the liver and gastro-intestinal tract.[366,367,538,539]

From the pharmacology stand point, a dual ligand, such as conicasterol F (**230**), holds potential in the treatment of liver disorders characterized by cholestasis and/or impaired metabolism of xenobiotics and, because both FXR and PXR exert anti-inflammatory effects in the intestine, in the treatment of inflammatory bowel diseases. In conclusion this discovery reaffirms the utility of examining natural product libraries for identifying novel receptor ligands potentially useful in the treatment of liver-related immune disorders.

8.2.1 Computational details

Molecular dynamics and Monte Carlo calculations were both performed on conicasterol F (**230**) on 4 x AMD Opteron SixCore 2.4Ghz. The dynamics calculations of **230** were performed at two different temperatures (600 and 750 K for 8 ns using 1.5 fs as time-step) using the MMFF[158] force field (MacroModel software package).[159] During both calculations, a standard constant temperature velocity-Verlet algorithm was used to integrate the equations of motions.[540] All the obtained structures (numbering =150, selected at regular intervals throughout the simulation) from Molecular Dynamics calculations for each isomer were minimized using the Polak-Ribier conjugate gradient algorithm (PRCG, 100000 steps, convergence threshold 0.005 kJ mol^{-1} $Å^{-1}$), leading to the selection of the lowest energy minimum conformer for both the diastereoisomers. The distribution of resulting geometries were in accordance with the results of a parallel conformational search performed with the Monte Carlo multiple minimum (MCMM) method (MMFFs,[158] 10000 steps, numbering 150, stored on a similarity and an energy criterion), where the variables used for the calculations included all the possible rotable

torsions. The empirical geometries from the conformer search for diastereoisomers **230a** and **230b** were optimized in vacuo at the DFT MPW1PW91 level using the 6-31G(d)[84] basis set (Gaussian 09 software package),[536] and then the optimized structures are used as inputs for the single-point ^{13}C chemical shift calculations performed in vacuo employing the same functional combined with the 6-31G(d,p)[84] basis set. The calculated values of chemical shifts of conicasterol F (**230**) were referred to the theoretical tetramethylsilane ^{13}C chemical shift value (previously optimized at DFT level), computed at the same level of theory.

-CONCLUSIONS-

A critical issue that determines the specificity of pharmacological response is related to the ability of a cellular receptor to recognize certain molecules. Relatively weak interactions with the active molecule are critical for the interaction with a receptor, and these are all dependent on its chemical structure and configuration. Another fundamental issue is the chirality of the ligand; thus, in most cases the receptor can bind, more or less specifically, only one of the two enantiomers according to the three-point model. As a consequence, in order to design and characterize new molecular platforms useful for any therapy it is necessary to know exactly the conformation and the configuration of the ligand.

In this context, the combination between NMR spectroscopy[78b,541] and modern computational techniques[542] (molecular docking, molecular dynamics and conformational search) certainly represents one of the most effective approaches, which were applied in order to design, rationalize, and perform structural studies on new potential antitumor and/or antiinflammatory molecules.

The results obtained described above can be outlined in three main areas of activity:

a) <u>Support in the design of original scaffolds for the generation of libraries potentially utilizable in therapy.</u> A molecular docking technique was exclusively used to conduct this rational design taking into account the analysis of ligand-target interactions and the synthetic possibilities. This kind of approach was successfully applied leading to the identification of new potential inhibitors for HDAC enzymes -both ciclic (mono and bis amides,[156] paragraph 2.2; conformationally locked calixarenes, paragraph 2.4), and linear (hydroxamic tertiary amines, paragraph 2.3),[167] and selective[191] (paragraph

2.6)- and for microsomal prostaglandin E_2 synthase (mPGES)-1 enzyme (two series of triazole-based compounds; paragraphs 4.2 and 4.3).[269,297]

In the paragraph 2.2, it was described the design collection of mono and bis amides (**18**-**24**) that have showed a satisfying level of antiproliferative activity, and from whose a new series of hydroxamic tertiary amines were designed (paragraph 2.3). In particular, a good accordance between molecular modeling predictions and biological results was found, and in fact, all the synthesized compounds (**25**-**34**) displayed a considerable HDAC inhibition activity. Especially, as predicted by docking calculations, compound **30** showed the highest inhibitory activity in the nanomolar range (IC_{50} 0,07 μM **30** vs 0,022 μM TSA),while **28** has showed a calculated and experimental inactivity.

Moreover, in the paragraph 2.4, alkyl- and arylamidocalix[4]arene derivatives **35**-**45** have been designed and theoretically evaluated by docking studies as potential histone deacetylase inhibitors (HDACi). On the basis of the trimodal distribution of the calculated inhibition constants (K_d), five alkyl- or arylamido derivatives (**37**, **41**, **42**, **43**, and **45**) were synthesized and tested. A qualitative accordance between the experimental results and the theoretical predictions was obtained, confirming that appropriately substituted arylamidocalix[4]arenes are active HDACi.

On the other hand in the paragraph 2.6 the structural elements responsible of the selective binding towards a specific isoform of HDAC have been investigated. It should be highlighted that class I and II proteins present a considerable sequence similarity in the catalytic site. In order to verify the theoretical findings, selective inhibitors able to discriminate between HDAC1 and 2 are designed, and then synthesized and tested by biological assays on all considered HDAC isoforms. The compound **110** exclusively binds HDAC2,

whereas the **111** inhibits HDAC8, suggesting that the structural modification of the appendage of the metal binder can lead to isoform selectivity in agreement with theoretical analysis. The **112** has showed similar inhibitory activity on HDAC3 and 8, confirming the theoretical prediction that structural modification of the linker alters the interaction of the cap group with HDAC2. Even though the small molecules showed a modest potency, the experimental data confirm theoretical observations, opening a new avenue for a targeted rational design of selective inhibitors towards the different HDAC isoforms.

Alongside the results reported above, in this thesis the design of two series of mPGES-1 inhibitors was described. In the pagraph 4.2, *in silico* screening to rapidly direct the synthesis, based on the copper-catalyzed 3 + 2 Huisgen's reaction (click chemistry) of potential mPGES-1 inhibitors was described.

26 (**140**-**165**) new triazole-based compounds were designed in accordance with the pocket binding requirements of human mPGES-1. Docking results, in agreement with ligand efficiency values, suggested the synthesis of **15** compounds that at least in theory were shown to be more efficient in inhibiting mPGES-1. Biological evaluation of these selected compounds has disclosed three new potential anti-inflammatory drugs: (I) compound **143** displaying selectivity for mPGES-1 with an IC_{50} value of 3.2 μM, (II) compound **159** that dually inhibits 5-lipoxygenase and mPGES-1, and (III) compound **146** apparently acting as 5-lipoxygenase-activating protein inhibitor (IC_{50} = 0.4 μM).

Moreover, in the paragraph 4.3 the design and synthesis, and biological evaluation of a second series (**167**-**181**) of mPGES-1 inhibitors based on a triazole scaffold are described. The reported studies allowed us to draw a SAR profile and to optimize this series with the identification of compounds **175**, **176** and **179**-**180** which displayed potent mPGES-1 inhibition in a cell-free

assay. In addition, compounds **170**, **175**, **177** and **179-181** also blocked 5-LO activity in cell-free and cell-based test systems, emerging as very promising candidates for the development of safer and more effective anti-inflammatory drugs. In particular taken together, compound **179** turned out to be the most potent dual mPGES-1/5-LO inhibitor out of these series with 4- and 7-fold lower IC_{50} values versus the parental lead compound **143** (IC_{50} = 0.68 μM).

b) <u>Rationalization of the biological activity of compounds by the study of the drug-receptor interactions.</u> In this thesis, the molecular docking was used to rationalize the binding modes of several compounds with known biological activities. In particular, the reported results are related to Ugi products derivatives of CHAP 1[161a] (HDAC inhibitors, paragraph 2.5), new and potent inhibitor of NMPRTAse analogs of FK866 and CHS 828[221] (chapter 3), marine natural products as inhibitors of $hsPLA_2$ (BLQ[291] and CLDA,[292] chapter 5), 4-methylen sterols extracted from *Theonella swinhoei* as ligands of FXR and PXR (chapter 6),[294,357,358,359] and moreover of known compounds as taurolitholic acid and ciprofloxacin (chapter 7),[296] agonists of TGR5.

In the paragraph 2.5 a full rationalization of binding mode of novel Ugi products (**36-77**) containing a zinc-chelating moiety is presented. **74** shows improved inhibitory potencies compared to SAHA, demonstrating that hindered lipophilic residues grafted on R-aminoacylamides scaffold with $NHR_1COCHR_2NR_3$ as *cap group* can favour in the interaction with the enzyme. These findings are also confirmed by calculated and experimental **73** inactivity.

In the chapter 3, from a biological screening of a small library of triazole-based compounds analogs (**113-134**), of FX866 and CSH828, the most

representative compounds were analyzed and a full rationalization of their binding mode was reported. Moreover, via molecular docking, the excellent potential of click chemistry for rapidly generating structure-activity relationships and for drug screening was reinforced.

In the paragraph 5.2 and 5.3 the molecular basis of the human group IIA secretory phospholipase A_2 inactivation by bolinaquinone (BLQ, **182**), and cladocoran A (CLD A, **183**) has been investigated for the comprehension of their relevant antiinflammatory properties, through the combination of spectroscopic techniques, biosensors analysis, mass spectrometry (MS) and molecular docking. The reported results suggest a mechanism of competitive inhibition guided by a non-covalent molecular recognition event, disclosing the key role of the BLQ hydroxyl-quinone moiety and of CLD A γ-hydroxybutenolide ring in the chelation of the catalytic Ca^{2+} ion inside the enzyme active site.

Moreover, CLD A is able to react selectively with Ser82, although this covalent event seems to play a secondary role in terms of enzyme inhibition.

The understanding of the $sPLA_2$-IIA inactivation mechanism by BLQ and CLDA could be useful for the development of a new chemical class of PLA_2 inhibitors, able to specifically target the enzyme active site.

In the paragraph 6.2 the putative binding modes to nuclear receptors (NRs) of 10 polyhydroxylated steroids, theonellasterols B-H (**200**-**206**) and conicasterols B-D (**207**-**209**) extracted from marine sponge *Theonella swinhoei*, were described. Pharmacological and structure-activity relationship analysis have demonstrated that these natural polyhydroxylated steroids are potent ligands of human nuclear pregnane receptor (PXR) and modulator of farnesoid-X-receptor (FXR). In addition, the molecular characterization of theonellasterol G allowed the identification of the first FXR modulator and

PXR ligand so far identified, highlighting its pharmacological potential in the treatment of liver disorders.

In the pargraph 6.3 the isolation and pharmacological characterization of conicasterol E (**213**) (*Theonella swinhoei*) in comparison to CDCA, a natural FXR ligand, 6-ECDCA, a synthetic FXR agonist, and rifaximin, a potent PXR agonist, were decribed, demostrating its FXR modulator endowed with PXR agonistic activity. The relative positioning in the ligand binding domain of FXR, explored through docking calculations, demonstrated a different spatial arrangement for conicasterol E and pointed to the presence of simultaneous and efficient interactions with the receptor that might support its biological activity.

In the paragraph 6.4 the isolation and the structural elucidation of a family of polyhydroxylated steroids (**241**-**223**) from the marine sponge *Theonella swinhoei* were reported. Decodification of interactions of this family with nuclear receptors has showed that these steroids are potent agonists of human pregnane-X-receptor (PXR) and antagonists of human farnesoid-X-receptor (FXR) with the putative binding mode to nuclear receptors (NRs) obtained through docking experiments.

Moreover in the paragraph 6.5, the discovery of theonellasterol (**198**) (*Theonella swinhoei*) as a highly selective FXR antagonist that protects against liver injury in cholestasis was reported. In particular, detailed description of the putative binding mode in several FXR crystal structures was reported in order to have a more precise description of the flexible Helix 12 in the ligand binding site.

Finally, in the chapter 7, ongoing the lack of the crystal structure of TGR5, the human adenosine A2a receptor was used as template for homology modeling procedure in order to obtain a predicted TGR5 three dimensional

structure. The so obtained structure was used as receptor model for the docking calculations to rationalize the binding modes of TGR5 agonist: the taurolitholic acid (TLCA, **224**) and ciprofloxacin (**225**).

c) <u>Determination of relative configuration of natural products</u> This kind of approach was successfully used in order to resolve stereostructural assignments of four natural products.[507, 532]

In the paragraph 8.1, the analysis of kedarcidin chromophore (**228**) and palau'amine (**229**) configuration through quantum chemical calculation of *J*s and chemical shifts was suggested as a fast and convenient approach utilizable prior to proceeding to the total synthesis of complex natural compounds in order to avoid loss of time and resources employed in the total synthesis of wrong diastereoisomers.

On the other hand, in the paragraph 8.2 the first application of combined accurate ROE-distance analysis with DFT calculations of NMR chemical shifts was reported to achieve the relative configuration assignment of a marine natural products, conicasterol F (**230**) and its and its 24-ethyl derivative, theonellasterol I (**231**), new polyhydroxylated steroids isolated from the marine sponge *Theonella swinhoei*. In this paragraph, it was demonstrated the substantial advantages of this combined approach as a tool for structural studies of natural products, providing a powerful alternative to, or information to underpin, total synthesis when more classical NMR data analysis fails to provide unequivocal results. Moreover, their pharmacological evaluation as human nuclear receptor modulators were reported.

References

1 Balkwill, F.; Mantovani, A. *Lancet*, **2001**, *357*, 539–545.

2 Dvorak, H. F. *N. Engl. J. Med.* **1986**, *135*, 1650–1659.

3 Dranoff, G. *Curr. Opin. Immunol.* **2002**, *14*, 161–164.

4 Pardoll, D. M. *Nature Rev. Immunol.* **2002**, *2*, 227–238.

5 Philip, M.; Rowley, D.A.; Schreiber, H. *Semin. Cancer Biol.* **2004**, *14*, 433–9.

6 a) Coussens, L.M.; Werb, Z. *Nature* **2002**, *420*, 860–867; b) Okada, F. Redox Rep. **2002**, *7*, 357–368; c) Yang, C. R.; Hsieh, S. L.; Ho, F. M.; Lin, W. W. *J. Immunol.* **2005**, *174*, 1647–1656; d) Nathan, C. *Nature* **2002**, *420*, 846–852.

7 Maiuri, M. C.; Tajana, G.; Iuvone, T.; De Stefano, D.; Mele, G.; Ribecco, M.T.; Cinelli, M. P.; Romano, M. F.; Turco, M. C.; Carnuccio, R. *Am. J. Pathol.* **2004**, *165*, 115–26.

8 Levy, B. D.; Clish, C. B.; Schmidt, B.; Gronert, K.; Serhan, C. N. *Nat. Immunol.* **2001**, *2*, 612–619.

9 Hodge-Dufour, J.; Marino, M. W.; Horton, M. R.; Horton, M. R. Jungbluth, A.; Burdick, M. D.; Strieter, R. M.; Noble, P. W.; Hunter, C. A.; Puré, E. *Proc. Natl. Acad. Sci. U S A* **1998**, *95*, 13806–13811.

10 a) Savill, J.; Wyllie, A. H.; Henson, J. E.; Walport, M. J.; Henson, P. M.; Haslett, C. *J. Clin. Invest.* **1989**, *83*, 865–875; b) Savill, J.; Fadok, V. A. *Nature* **2000**, *407*, 784–788; c) Savill, J.; Dransfield, I.; Gregory, C.; Haslett, C. *Nat. Rev. Immunol.* **2002**, *2*, 965–75.

11 a) Fadok, V. A.; Bratton, D. L.; Konowal, A.; Freed, P. W.; Westcott, J. Y.; Henson, P. M. *J. Clin. Invest.* **1998**, *101*, 890–898; b) McDonald,

P. P.; Fadok, V. A.; Bratton, D.; Henson, P. M. *J. Immunol.* **1999**, *163*, 6164–6172; c) Huynh, M.- L. N.; Fadok, V. A.; Henson, P. M. *J. Clin. Invest.* **2002**, *109*, 41–50.

12 Macarthur, M.; Hold, G. L.; El-Omar, E. M. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2004**, *286*, G515–G520.

13 Berenblum, I. *Cancer Res.* **1941**, *1*, 44–48.

14 Trosko, J. E. *Mol. Carcinog.* **2001**, *30*, 131–137.

15 Hanahan, D.; Weinberg, R. A. *Cell* **2000**, *100*, 57–70.

16 Aggarwal, B. B. *Cancer Cell* **2004**, *6*, 203–208.

17 Mantovani, A. *Nature*, **2005** *435*, 752–753.

18 Coussens, L. M.; Werb, Z. *Nature* **2002**, *420*, 860–867.

19 a) Martey, C. A.; Pollock, S. J.; Turner, C. K.; O'Reilly, K. M.; Baglole, C. J.; Phipps, R. P.; Sime, P. J. *Am. J. Physiol. Lung Cell Mol. Physiol.* **2004**, *287*, L981–991; b) Peek, Jr R. M.; Crabtree, J. E. *J. Pathol.* **2006**, 208, 233–248; c) Castle, P. E.; Hillier, S. L., Rabe, L. K.; Hildesheim, A.; Herrero, R.; Bratti, M. C., Sherman M. E., Burk R. D., Rodriguez A. C., Alfaro M., Hutchinson M. L., Morales J., Schiffman M. *Cancer Epidemiol. Biomarkers Prev.* **2001**, *10*, 1021–1027; d) Di Bisceglie, A. M. *Hepatology* **1997**, *26*, 34S–38S; e) Kanoh, K.; Shimura, T.; Tsutsumi, S.; Suzuki, H.; Kashiwabara, K.; Nakajima. T.; Kuwano, H. *Cancer Lett*. **2001**, *169*, 7–14; f) Offersen, B. V.; Knap, M. M.; Marcussen, N.; Horsman, M. R.; Hamilton-Dutoit, S.; Overgaard, *Cancer* **2002**, *87*, 1422–1430; g) Garcea, G.; Dennison, A. R.; Steward, W. P.; Berry, D. P. *Pancreatology* **2005**, *5*, 514–529; h) Murphy, S. J.; Anderson, L. A.; Johnston, B. T.; Fitzpatrick, D. A.; Watson, P. R.; Monaghan, P.; Murray, L. J. *World J. Gastroenterol.*

2005, *11*, 7290–7295; i) Mossman, B. T.; Kamp, D. W.; Weitzman, S. A. *Cancer Invest.* **1996**, *14*, 466–480; l) Okano, M.; Gross, T. G. *Pediatr. Hematol. Oncol.* **2001**, *18*, 427–442; m) Vagefi, P. A.; Longo, W. E. *Clin. Colorectal Cancer* **2005**, *4*, 313–319; m) Berwick, M.; Armstrong, B. K.; Ben-Porat, L.; Fine, J.; Kricker, A.; Eberle, C.; Barnhill, R. *J. Natl. Cancer Inst.* **2005**, *97*, 195–199; n) Nelson, W. G.; De Marzo, A. M.; DeWeese, T. L.; Isaacs, W. B. *J. Urol.* **2004**, *172*, S6–11.

20 a) Balkwill, F.; Charles, K. A.; Mantovani, A. *Cancer Cell* **2005**, 7, 211–217; b) Coussens, L. M.; Werb, Z. *Nature* **2002**, *420*, 860–867; c) Karin, M. *Nature*, **2006**, *441*, 431–436.

21 a) Koehne, C. H.; Dubois, R. N. *Semin. Oncol.* **2004**, *31*, 12–21; b) Flossmann, E.; Rothwell, P. M. *Lancet* **2007**, *369*, 1603–1613; c) Chan, A. T.; Ogino, S.; Fuchs, C. S. *N. Engl. J. Med.* **2007**, *356*, 2131–2142.

22 Mantovani, A., Allavena, P., Sica, A.; Balkwill, F. *Nature* **2008**, *454*, 436–444.

23 Hanahan, D.; Weinberg, R. A. *Cell* **2000**, *100*, 57–70.

24 Kim, S.; Takahashi, H.; Lin, W. W.; Descargues, P.; Grivennikov, S.; Kim, Y.; Luo, J. L.; Karin, M. *Nature* **2009**, *457*, 102–106.

25 Maeda, H.; Akaike, H. *Biochemistry* **1998**, *63*, 854–65.

26 Fulton, A. M.; Loveless, S. E.; Heppner, G. H. *Cancer Res*. **1984**, *44*, 4308–4311.

27 Pollard, J. W. *Nat. Rev. Cancer* **2004**, *4*, 71–78.

28 Hudson, J. D.; Shoaibi, M.A.; Maestro, R.; Carnero, A.; Hannon, G. J.; Beach, D. H. *J. Exp. Med.* **1999**, *190*, 1375–82.

29 Petrenko, O.; Moll, U. M. *Mol. Cell*. **2005**, *17*, 225–36.

30 Chu, F. F.; Esworthy, R. S.; Chu, P. G.; Longmate, J. A.; Huycke, M. M.; Wilczynski, S.; Doroshow, J. H.*Cancer Res.* **2004**, *64*, 962–968.

31 Lin, E. Y.; Pollard, J. W. *Br. J. Cancer* **2004**, *90*, 2053–2058.

32 Coussens, L. M.; Werb, Z. *J. Exp. Med.* **2001**, *193*, F23–26.

33 a) Clark, W. H.; Elder, D. E.; Guerry, D. IV, Braitman, L. E.; Trock, B. J.; Schultz, D.; Synnestvedt, M.; Halpern, A. C. *J. Natl. Cancer Inst.* **1989**, *81*, 1893–1904; b) Clemente, C. G.; Mihm, M. C.; Bufalino, R.; Zurrida, S.; Collini, P.; Cascinelli, N. *Cancer* **1996**, *77*, 1303–1310; c) Naito, Y.; Saito, K.; Shiiba, K.; Ohuchi, A.; Saigenji, K.; Nagura, H.; Ohtani, H. *Cancer Res.* **1998**, *58*, 3491–3494; d) Nakano, O.; Sato, M.; Naito, Y,; Suzuki, K.; Orikasa, S.; Aizawa, M.; Suzuki, Y.; Shintaku, I.; Nagura, H.; Ohtani, H. *Cancer Res.* **2001**, *61*, 5132–5136; e) Zhang, L.; Conejo-Garcia, J. R.; Katsaros, D.; Gimotty, P. A.; Massobrio, M.; Regnani, G.; Makrigiannakis, A.; Gray, H.; Schlienger, K.; Liebman, M. N.; Rubin, S. C.; Coukos, G. *N. Eng. J. Med.* **2003**, *348*, 203–213;

f) Dunn, G.; Bruce, A.; Ikeda, H.; Old, L.; Schreiber, R. *Nat. Immunol.* **2002**, *3*, 991–998; g) Brigati, C, Noonan, D. M.; Albini, A.; Benelli, R. *Clin. Exp. Metastasis* **2002**, *19*, 247–258; h) Tsung, K.; Dolan, J. P.; Tsung, Y. L.; Norton, J. A. *Cancer Res*. **2002**, *62*, 5069–5075; i) Mihm, M.; Clemente, C.; Cascinelli, N. *Lab. Invest.* **1996**, 74, 43–47.

34 Smyth, M. J.; Cretney, E.; Kershaw, M. H.; Hayakawa, Y. *Immunol. Rev.* **2004**, *202*, 275–293.

35 Khong, H.T.; Restifo, N. P. *Nat. Immunol.* **2002**, *3*, 999–1005.

36 Karin, M. *Nature* **2006**, *441*, 431–436.

37 Yu, H.; Kortylewski, M.; Pardoll, D. *Nature Rev. Immunol.* **2007**, *7*, 41–51.

38 a) Voronov, E.; Shouval, D. S.; Krelin, Y.; Cagnano, E.; Benharroch, D.; Iwakura, Y.; Dinarello, C. A.; Apte, R. N. *Proc. Natl Acad. Sci. USA* **2003**, *100*, 2645–2650; b) Grivennikov, S.; Karin, M. *Cancer Cell* **2008**, *13*, 7–9; c) Szlosarek, P. W.; Balkwill, F. R. *Lancet Oncol.* **2003**, *4*, 565–573; d) Langowski, J. L.; Zhang, X.; Wu, L.; Mattson, J. D.; Chen, T.; Smith, K.; Basham, B.; McClanahan, T.; Kastelein, R. A.; Oft, M. *Nature* **2006**, *442*, 461–465.

39 Aggarwal, B. B. *Nat. Rev. Immunol.* **2003**, *3*, 745–56.

40 Balkwill, F. *Cytokine Growth Factor Rev.* **2002**, 13, 135–41.

41 Courtois, G.; Gilmore, T. D. *Oncogene* **2006**, *25*, 6831–6843.

42 a) Carbia-Nagashima, A.; Gerez, J.; Perez-Castro, C.; Paez-Pereda, M.; Silberstein, S.; Stalla, G. K.; Holsboer, F.; Arzt, E. *Cell* **2007**, *131*, 309–323; b) Mizukami, Y.; Jo, W. S.; Duerr, E. M.; Gala, M.; Li, J.; Zhang, X.; Zimmer, M. A.; Iliopoulos, O.; Zukerberg, L. R.; Kohgo, Y.; Lynch, M. P.; Rueda, B. R.; Chung, D. C. *Nature Med.* **2005**, *11*, 992–997; c) Rius, J.; Guma, M.; Schachtrup, C.; Akassoglou, K.; Zinkernagel, A. S.; Nizet, V.; Johnson, R. S.; Haddad, G. G.; Karin, M. *Nature* **2008**, *453*, 807–811.

43 a) Greten, F. R.; Eckmann, L.; Greten, T. F.; Park, J. M.; Li, Z. W.; Egan, L. J.; Kagnoff, M. F.; Karin, M. *Cell* **2004**, *118*, 285–296; b) Pikarsky, E.; Porat, R. M.; Stein, I.; Abramovitch, R.; Amit, S.; Kasem, S.; Gutkovich-Pyest, E.; Urieli-Shoval, S.; Galun, E.; Ben-Neriah, Y. *Nature* **2004**, *431*, 461–466.

44 a) Garlanda, C.; Riva, F.; Veliz, T.; Polentarutti, N.; Pasqualini, F.; Radaelli, E.; Sironi, M.; Nebuloni, M.; Zorini, E. O.; Scanziani, E.; Mantovani, A. *Cancer Res.* **2007**, *67*, 6017–6021; b) Xiao, H.; Gulen, M. F.; Qin, J.; Yao, J.; Bulek, K.; Kish, D.; Altuntas, C. Z.; Wald, D.; Ma, C.; Zhou, H.; Tuohy, V. K.; Fairchild, R. L.; de la Motte, C.; Cua, D.; Vallance, B. A.; Li, X. *Immunity* **2007**, *26*, 461–475.

45 Biswas, S. K.; Gangi, L.; Paul, S.; Schioppa, T.; Saccani, A.; Sironi, M.; Bottazzi, B.; Doni, A.; Vincenzo, B.; Pasqualini, F.; Vago, L.; Nebuloni, M.; Mantovani, A.; Sica, A. *Blood* **2006**, *107*, 2112–2122.

46 Saccani, A.; Schioppa, T.; Porta, C.; Biswas, S. K.; Nebuloni, M.; Vago, L.; Bottazzi, B.; Colombo, M. P.; Mantovani, A.; Sica A. *Cancer Res.* **2006**, *66*, 11432–1140.

47 Halili, M. A.; Andrews, M. R.; Sweet M. J.; Fairlie, D. P. *Curr. Top. Med. Chem.* **2009**, *9*, 309-319.

48 Gallí, M.; Van Gool, F.; Rongvaux, A.; Andris, F.; Leo, O. *Cancer Res* **2010**, *70*, 8-11.

49 a) Nakanishi, M.; Gokhale, V.; Meuillet, E. J., Rosenberg, D. W. *Biochimie* **2010**, *92*, 660-664; b) Rådmark, O.; Samuelsson, B. *J. Intern. Med.* **2010**, *268*, 5–14.

50 a) Mirtti, T.; Laine, V. J.; Hiekkanen, H.; Hurme, S.; Rowe, O.; Nevalainen, T. J.; Kallajoki, M.; Alanen, K. *APMIS* **2009**, *117*, 151–161; b) Sved, P.; Scott, K. F.; McLeod, D.; King, N. J.; Singh, J.; Tsatralis, T.; Nikolov, B.; Boulas, J.; Nallan, L.; Gelb, M. H.; Sajinovic, M.; Graham, G. G.; Russell, P. J.; Dong, Q. *Cancer Res.* **2004**, *64*, 6934–6940; c) Dong, Q.; Patel, M.; Scott, K.F.; Graham, G. G.; Russell, P. J.; Sved, P. *Cancer Lett.* **2006**, *240*, 9–16; d) Attiga, F.

A.; Fernandez, P. M.; Weeraratna, A. T.; Manyak, M. J.; Patierno, S. R. *Cancer Res.* **2000**, *60*, 4629–4637.

51 Thomas, C.; Pellicciari, R.; Pruzanski, M.; Auwerx, J.; Schoonjans, K.; *Nat. Rev. Drug Discov.* **2008**, *7*, 678-93.

52 Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. *J. Mol. Biol.* **1982**, *161*, 269–288.

53 Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J. Mol. Biol.* **1996**, *261*, 470–489.

54 Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, *267*, 727–748.

55 Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comp. Chem.* **1998**, *19*, 1639–1662.

56 Huey, R.; Morris, G. M.; Olson, A. J.; Goodsell, D. S. *J. Comput. Chem.* **2007**, 28, 1145–1152.

57 Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. *J. Comput. Chem.* **2009,** *30*, 2785-2791.

58 Trott, O.; Olson, A. J. *J. Comput. Chem.* **2010**, *31*, 455–461.

59 Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J. Med. Chem.* **2004**, *47*, 1739–1749.

60 a) Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. *Proteins* **2002**, *47*, 409-443; b) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. *Nat. Rev. Drug Discovery* **2004**, *3*, 935-949.

61 a) Verkhivker, G. M.; Bouzida, D.; Gehlhaar, D. K.; Rejto, P. A.; Arthurs, S.; Colson, A. B.; Freer, S. T.; Larson, V.; Luty, B. A.;

Marrone, T.; Rose, P. W. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 731-751; b) Watson, J. D.; Laskowski, R. A.; Thornton, J. M. *Curr. Opin. Struct. Biol.* **2005**, *15*, 275-284; c) Arakaki, A. K.; Zhang, Y.; Skolnick, J. *Bioinformatics* **2004**, *20*, 1087-1096; d) Park, H.; Lee, J.; Lee, S. 195 *Proteins* **2006**, *65*, 549-554, e) Hetenyi, C.; Van Der Spoel, D. *Protein Sci.* **2002**, *11*, 1729-1737.

62 Leach, A. R. *Molecular Modelling: Principles and Applications*, Addison Wesley Longman Limited, Harlow, **1996**

63 Brooijmans, N.; Kuntz, I. D. *Annu. Rev. Biophys. Biolmol. Struct.* **2003**, *32*, 335–373.

64 Di Nola, A.; Berendsen, H. J. C.; Roccatano, D. *Proteins* **1994**, *19*, 174–182.

65 Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *Proteins* **2006**, *65*, 15-26.

66 Park, H.; Lee, J.; Lee, S. *Proteins* **2006**, *65*, 549-554.

67 W. E. Hart, T. E. Kammeyer, and R. K. Belew, In *Foundations of Genetic Algorithms III*, D. Whitley and M. Vose, Eds., Morgan Kauffman, San Francisco, CA, 1994.

68 J. B. Lamarck, *Zoological Philosophy*, Macmillan, London, 1914.

69 Goodford, P. J. *J. Med. Chem.* **1985**, *28*, 849-857.

70 a) Sharp, K.; Fine, R.; Honig, B. *Science* **1987**, *236*, 1460-1463; b) Allison, S. A.; Bacquet, R. J.; McCammon, J. *Biopolymers* **1988**, *27*, 251-269.

71 Weiner, S. J., Kollman, P. A., Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230-252.

72 Gilson, M. K.; Zhou, H.-X.; *Annu. Re. Biophys. Biomole Struct.* **2007**, *36*, 21-42

73 a) Gilson, M.; Given, J.; Bush, B.; McCammon, J. *Biophys J* **1997**, *72*, 1047-1069; b) Chang, C.-E. A.; Chen,W.; Gilson, M. K. *Proc Nat Acad Sci USA* **2007**, *104*, 1534-1539.

74 Seco, J. M.; Quiñoà, E.; Riguera, R. *Chem. Rev. (Washington, DC, U. S.)* **2004**, *104*, 17-117.

75 Nicolaou, K. C.; Snyder, S. A. *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 1012-1044.

76 Bifulco, G.; Dambruoso, P.; Gomez-Paloma, L.; Riccio, R. *Chem. Rev.* **2007**, *107*, 3744-3779.

77 Karplus, M. *J. Chem. Phys.* **1959**, *30*, 11-15.

78 Overhauser, A. W. *Phys. Rev.* **1953**, *92*, 411-415.

79 Matsumori, N.; Kaneno, D.; Murata, M.; Nakamura, H.; Tachibana, K. *J. Org. Chem.* **1999**, *64*, 866–876.

80 a) Kobayashi, Y.; Lee, J.; Tezuka, K.; Kishi, Y. *Org. Lett.* **1999**, *1*, 2177–2180; b) Lee, J.; Kobayashi, Y.; Tezuka, K.; Kishi, Y. *Org. Lett.* **1999**, *1*, 2181–2184.

81 Barone, G.; Gomez-Paloma, L.; Duca, D.; Silvestri, A.; Riccio, R.; Bifulco, G. *Chem. Eur. J.* **2002**, *8*, 3233-3329.

82 Barone, G.; Duca, D.; Silvestri, A.; Gomez-Paloma, L.; Riccio, R.; Bifulco, G. *Chem. Eur. J.* **2002**, *8*, 3240-3245.

83 a) Cheeseman, J. R.; Trucks, G. W.; Keith, T. A.; Frisch, M. J. *J. Chem. Phys.* **1996**, *104*, 5497-5509.

84 Cimino, P.; Duca, D.; Gomez-Paloma, L.; Riccio, R.; Bifulco, G. *Magn. Reson. Chem.* **2004**, *42*, S26-S33.

85 a) Elyashberg, M.; Blinov, K.; Williams, A. *Magn. Reson. Chem.* **2009**, *47*, 371–389; b) Elyashberg, M. E.; Blinov, K. A.; Williams, A. J.

Magn. Reson. Chem. **2009**, *47*, 333–341; c) Elyashberg, M.; Blinov, K.; Molodtsov, S.; Smurnyy, Y.; Williams, A. J.; Churanova, T. *J. Chem Inf.* **2009**, *3*, 1-26.

86 a) Elyashberg, M. E.; Williams, A. J.; Martin, G. E. *Prog. NMR Spectrosc.* **2008**, *53*, 1-104; b) Blinov, K. A.; Smurnyy, Y. D.; Churanova, T. S.; Elyashberg, M. E.; Williams, A. J. *Chemom. Intell. Lab. Syst.* **2009**, *97*, 91-97; c) Blinov, K. A.; Smurnyy, Y. D.; Elyashberg, M. E.; Churanova, T. S.; Kvasha, M.; Steinbeck, C.; Lefebvre, B. A.; Williams, A. J. *J. Chem. Inf. Model.* **2008**, *48*, 550-555; d) Smurnyy, Y. D.; Blinov, K. A.; Churanova, T. S.; Elyashberg, M. E.; Williams, A. J. *J. Chem. Inf. Model.* **2008**, *48*, 128-134.

87 a) Ditchfield, R. J. *Chem. Phys.* **1972**, *56*, 5688–5691; b) Wolinski, K.; Hinton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251–8260.

88 a) van Gunsteren, W. F.; Berendsen, H. J. C. *Angew. Chem., Int. Ed.* **1990**, *29*, 992-1023; b) Höltje, H. D.; Sippl, W.; Folkers, G. *Molecular Modeling Basic Principles and Applications* (Eds.: R. Mannhold, H. Kubinyi, H. Timmerman), Wiley-VCH, Weinhein, New York, **2003**.

89 Chang, G.; Guida, W. C.; Still, W. C. *J. Am. Chem. Soc.* **1989**, *111*, 4379–4386.

90 Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.

91 Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209-220.

92 P. Hohenberg, W. Kohn, *Phys. Rev. B* **1964**, *136*, 864-871; b) W. Kohn, L. J. Sham, *Phys. Re. A* **1965**, *140*, 1133-1138.

93 Bifulco, G.; Bassarello, C.; Riccio, R.; Gomez-Paloma, L. *Org. Lett.* **2004**, *6*, 1025–1028.

94 a) Thiele, C. M.; Schmidts, V.; Bottcher, B.; Louzao, I.; Berger, R.; Maliniak, A.; Stevensson, B. *Angew. Chem. Int. Ed* **2009**, *48*, 6708-6712; b) Beraud, S.; Bersch, B.; Brutscher, B.; Gans, P.; Barras, F.; Blackledge, M. *J. Am. Chem. Soc.* **2002**, *124*, 13709-13715.

95 a) Liu, M.; Farrant, R. D.; Lindon, J. C. *Magn. Reson. Chem.* **1992**, *30*, 173-176; b) Chazin, W. J.; Colebrok, L. D.; Edward, J. T. *Can. J. Chem.* **1983**, *61*, 1749-1755; c) Liu, M.; Lindon, J. C. *Concepts Magn. Reson.* **1996**, *8*, 161-173; d) Colebrook, L. D.; Hall, L. D. *Can. J. Chem.* **1980**, *58*, 2016-2023; e) Hall, L. D.; Hill, H. D. W. *J. Am. Chem. Soc.* **1976**, *98*, 1269-1270.

96 Neuhaus, D.; Williamson, M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, John Wiley & Sons Inc, 2nd revised edition, **2000**.

97 Andersen, N. H.; Eaton, H. L.; Lai, X. *Magn. Reson. Chem.* **1989**, *27*, 515-528.

98 Keller, C. E.; Carper, W. R. *Magn. Reson. Chem.* **1993**, *31*, 566-572.

99 Stonehouse, J.; Adell, P.; Keeler, J.; Shaka A. J. *J. Am. Chem. Soc.* **1994**, *116*, 6037-6038.

100 Stott, K.; Keeler, J.; Van Q. N.; Shaka, A. J. *J. Magn. Reson.* **1997**, *125*, 302–324.

101 Stott, K.; Stonehouse, J.; Keeler, J.; Hwang, T. L.; Shaka, A. J. *J. Am. Chem. Soc.* **1995**, 117, 4199-4200.

102 Thrippleton, M. J.; Keeler, J. *Angew. Chem., Int. Ed.* **2003**, 42, 3938-3941.

103 Butts, C. P.; Jones, C. R.; Towers, E. C.; Flynn, J. L.; Appleby, L.; Barron, N. *Org. Biomol. Chem.* **2011**, *9*, 177–184.

104 Butts, C. P.; Jones, C. R.; Harvey, J. N. *Chem. Commun.* **2011**, *47*, 1193–1195.

105 Butts, C. P.; Jones, C. R.; Harvey, J. N. *Beilstein J. Org. Chem.* **2011**, *7*, 145–150.

106 Keller, C. E.; Carper, W. R. *Magn. Reson. Chem.* **1993**, *31*, 566-572.

107 Solomon, I. *Phys. Rev.* **1955**, *99*, 559-565.

108 Macura, S.; Ernst, R. R. *Mol. Phys.* **1980**, *41*, 95-117.

109 Macura, S.; Farmer II B. T.; Brown, L. R. *J. Magn. Reson.* **1986**, *70*, 493–499.

110 Hu, H.; Krishnamurthy, K. *J. Magn. Reson.* **2006**, *182*, 173–177.

111 James, T. L., Keepers, J.W. *J. Magn. Reson.* **1984**, *57*, 404-426.

112 Wang, B.; Dossey, A. T.; Walse, S. S.; Edison, A. S.; Merz, K. M. *J. Nat. Prod.* **2009**, *72*, 709-713.

113 Andersen, N. H.; Eaton, H. L.; Lai, X. *Magn. Reson. Chem.* **1989**, *27*, 515-528.

114 Forgo, P.; Kover, K. E.; Hohmann, J. *Monatsh. Chem.* **2002**, *133*, 1249-1261.

115 Messerschmidt, M.; Scheins, S.; Luger, P. *Acta Crystallogr., Sect. B: Struct. Sci* **2005**, *61*, 115-121.

116 Bagno, A.; Rastrelli, F.; Saielli, G. *Chem. Eur. J.* **2006**, *12*, 5514-5525.

117 Strahl, B. D.; Allis, C.D. *Nature* **2000**, *403*, 41–45.

118 Lund, A. H.; van Lohuizen, M. *Genes Dev.* **2004**, *18*, 2315-2335.

119 Kornberg, R. D.; Klug, A. *Sci Am.* **1981**, *244*, 52-64

120 Wu, J.; Grunstein, M. *Trends Biochem. Sci.* **2000**, *25*, 619-623.

121 Gregoretti, I.; Lee, Y.-M.; Goodson, H. V. *J. Mol. Biol.* **2004**, *338*, 17-31.

122 Finnin, M. S.; Donigian, J. R.; Cohen, A.; Richon, V. M.; Rifkind, R. A.; Marks, P. A.; Pavletich, N. P. *Nature*, **1999**, *401*, 188-193.

123 Grozinger, C. M.; Schreiber, S. L. *Chem. Biol.* **2002**, *9*, 3-16.

124 a) Davie, J. R.; Spencer, V. A. *J. Cell. Biochem*. **1999**, *75*, 141–148; b) Langley, B.; Gensert, J. M.; Beal, M. F.; Ratan, R. R. *Curr. Drug Targets CNS Neurol. Disord.* **2005**, *4*, 41–50; c) Bhaumik, S. R.; Smith, E.; Shilatifard, A. *Nature Struct. Mol. Biol.* **2007**, *14*, 1008–1016.

125 a) Hubbert, C.; Guardiola, A.; Shao, R.; Kawaguchi, Y.; Ito, A.; Nixon, A.; Yoshida, M.; Wang, X. F.; Yao, T. P. *Nature* **2002**, *417*, 455–458: b) Matsuyama, A.; Shimazu, T.; Sumida, Y.; Saito, A.; Yoshimatsu, Y.; Seigneurin-Berny, D.; Osada, H.; Komatsu, Y.; Nishino, N.; Khochbin, S.; Horinouchi, S.; Yoshida, M. *EMBo J*. **2002**, *21*, 6820–6831.

126 Minucci, S.; Pelicci, P. G. *Nat. Rev. Cancer* **2006**, *6*, 38-51.

127 a) Lin, R. J.; Sternsdorf, T.; Tini, M.; Evans, R. M. *Oncogene* **2001**, *20*, 7204-7215; b) Grignani, F.; De Matteis, S.; Nervi, C.; Tomassoni, L.; Gelmetti, V.; Cioce, M.; Fanelli, M.; Ruthardt, M.; Ferrara, F. F.; Zamir, I.; Seiser, C.; Grignani, F.; Lazar, M. A.; Minucci, S.; Pelicci, P. G. *Nature* **1998**, *391*, 815-818. c) Pandolfi, P. P. *Oncogene* **2001**, *20*, 3116-3127.

128 Lin, H.Y.; Chen, C.S.; Lin, S.P; Weng, J.R.; Chen, C. S.*; Med Res Rev* **2006**, *26*, 397-413.

129 a) Witt, O.; Deubzer, H. E.; Milde, T.; Oehme, I. *Cancer Lett.* **2009**, *277*, 8-21; b) Bolden, J. E.; Peart, M. J.; Johnstone, R. W. *Nat. Rev. Drug Discovery* **2006**, *5*, 769-784.

130 Bertrand, P. *Eur. J. Med. Chem.* **2010**, *45*, 2095-2116.

131 Witt, O.; Deubzer, H. E.; Milde, T.; Oehme, I. *Cancer Lett.* **2009**, *277*, 8–21.

132 Mai, A.; Masse, S.; Rotili, D.; Cerbara, I.; Valente, S.; Pezzi, R.; Simeoni, S.; Ragno, R. *Med. Res. Rev.* **2005**, *25*, 261-309.

133 a) Kuendgen, A.; Schmid, M.; Schlenk, R.; Knipp, S.; Hildebrandt, B.; Steidl, C.; Germing, U.; Haas, R.; Dohner, H.; Gattermann, N. *Cancer* **2006**, *106*, 112-119; b) Bug, G.; Ritter, M.; Wassmann, B.; Schoch, C.; Heinzel, T.; Schwarz, K.; Romanski, A.; Kramer, O. H.; Kampfmann, M.; Hoelzer, D.; Neubauer, A.; Ruthardt M.; Ottmann, O. G. *Cancer* **2005**, *104*, 2717-2725; c) Chavez-Blanco, A.; Segura-Pacheco, B.; Perez-Cardenas, E.; Taja-Chayeb, L.; Cetina, L.; Candelaria, M.; Cantu, D.; Gonzalez-Fierro, A.; Garcia-Lopez, P.; Zambrano, P.; Perez-Plasencia, C.; Cabrera, G.; Trejo-Becerril, C.; Angeles, E.; Duenas-Gonzalez, A. *Mol Cancer* **2005**, *4*, 1-9.

134 a) Gojo, I.; Jiemjit, A.; Trepel, J.B.; Sparreboom, A.; Figg, W.D.; Rollins, S.; Tidwell, M. L.; Greer, J.; Chung, E-J.; Lee, M.-J.; Gore, S. D.; Sausville, E. A.; Zwiebel, J.; Karp, J. E. *Blood* **2007**, *109*, 2781-2790; b) Gore, L.; Rothenberg, M .L.; O'Bryant, C. L.; Schultz, M. K.; Sandler, A. B.; Coffin, D.; McCoy, C.; Schott, A.; Scholz, C.; Eckhardt1, S. G. *Clin Cancer Res* **2008**, *14*, 4517-4525; c) Hauschild, A.; Trefzer, U.; Garbe, C.; Kaehler, K. C.; Ugurel, S.; Kiecker, F.; Eigentler, T.; Krissel, H.; Schott, A.; Schadendorf, D. *Melanoma Res* **2008**, *18*, 274-278.

135 Mai, A.; Massa, S.; Pezzi, R.; Simeoni, S.; Rotili, D.; Nebbioso, A.; Scognamiglio, A.; Altucci, L.; Loidl, P.; Brosch, G. *J. Med. Chem.* **2005**, *48* 3344–3353.

136 a) Cang, S.; Ma, Y.; Liu, D. *J. of Hem. & Oncol.* **2009**, *2*, 1-11; b) Klimek, V. M.; Fircanis, S.; Maslak, P.; Guernah, I.; Baum, M.; Wu, N.; Panageas, K.; Wright, J. J.; Pandolfi, P. P.; Nimer, S. D.*Clin Cancer Res* **2008**, *14*, 826-832; c) Schrump, D. S.; Fischette, M. R.; Nguyen, D. M.; Zhao, M.; Li, X.; Kunst, T. F.; Hancox, A.; Hong, J. A.; Chen, G. A.; Kruchin, E.; Wright, J. J.; Rosing, D. R.; Sparreboom, A.; Figg, W. D.; Steinberg, S. M. *Clin Cancer Res* **2008**, *14*, 188-198; d) Stadler, W. M.; Margolin, K.; Ferber, S.; McCulloch,W.; Thompson, J. A. *Clin Genitourin Cancer* **2006**, *5*, 57-60.

137 Marks, P. A.; Breslow, R. *Nat. Biotechnol.* **2007**, *25*, 84–90.

138 Pankaj, B.; Tehireem, A.; Ian, M. A. *J. Allergy Clin. Immunol.* **2008**, *121*, 580–584.

139 Blanchard, F.; Chipoy, C. *Drug Discov. Today* **2005**, *10*, 197–204.

140 Kazantsev, A. G.; Thompson, L. M. *Nat. Rev. Drug Discov.* **2008**, *7*, 854–868.

141 Mishra, N.; Brown, D. R.; Olorenshaw, I. M.; Kammer, G. M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 2628–2633.

142 Rubenstein, R. C.; Zeitlin, P. L. *Am. J. Respir. Crit. Care. Med.* **1998**, 157, 484–490.

143 Tao, R.; Hancock, W. W. *Hepatobiliary Pancreat Dis. Int.* **2007**, *6*, 348–357.

144 Huang, L. *J. Cell Physiol.* **2006**, *209*, 611-616.

145 Khan, N.; Jeffers, M.; Kumar, S.; Hackett, C.; Boldog, F.; Khramtsov, N.; Qian, X.; Mills, E.; Berghs, S. C.; Carey, N.; Finn, P. W.; Collins, L. S.; Tumber, A.; Ritchie, J. W.; Jensen, P. B.; Lichenstein H. S.; Sehested, M.; *Biochem. J.* **2008**, *409*, 581-589.

146 Bruserud, O.; Stapnes, C.; Ersvaer, E.; Gjertsen, B. T.; Ryningen, A. *Curr. Pharm. Biotechnol.* **2007**, *8*, 388–400.

147 Maulucci, N.; Chini, M. G.; Di Micco, S.; Izzo, I.; Cafaro, E.; Russo, A.; Gallinari, P.; Paolini, C.; Nardi, M. C.; Casapullo, A.; Riccio, R.; Bifulco, G.; De Riccardis, F. *J. Am. Chem. Soc.* **2007**, *129*, 3007-3012.

148 Breneman, C. M.; Wiberg, K. B. *J. Comp. Chem.* **1990**, *11*, 361-373.

149 Gaussian 03 (Revision B.05), Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.;. Tomasi, J; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.;

Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian Inc., 2003, Pittsburgh, PA (USA).

150 Bressi, J. C.; Jennings, A. J.; Skene, R.; Wu, Y.; Melkus, R.; De Jong, R.; O'Connell, S.; Grimshaw, C. E.; Navre, M.; Gangloff, A. R. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 3142–3145.

151 Bottomley, M. J.; Lo Surdo, P.; Di Giovine, P.; Cirillo, A.; Scarpelli, R.; Ferrigno, F.; Jones, P.; Neddermann, P.; De Francesco, R.; Steinkuhler, C.; Gallinari, P.; Carfi, A. *J. Biol. Chem.* **2008**, *283*, 26694-26704.

152 Schuetz, A.; Min, J.; Allali-Hassani, A.; Schapira, M.; Shuen, M.; Loppnau, P.; Mazitschek, R.; Kwiatkowski, N. P.; Lewis, T. A.; Maglathin, R. L.; McLean, T. H.; Bochkarev, A.; Plotnikov, A. N.; Vedadi, M., Arrowsmith, C. H. *J. Biol. Chem.* **2008**, *283*, 11355-11363.

153 Dowling, D. P.; Gantt, S. L.; Gattis, S. G.; Fierke, C. A.; Christianson, D. W. *Biochemistry* **2008**, *47*, 13554-13563.

154 J. R.; Skene, R. J.; Katz, B. A.; Mol, C.; Ho, J. D.; Jennings, A. J.; Luong, C.; Arvai, A.; Buggy, J. J.; Chi, E.; Tang, J.; Sang, B.-C.; Verner, E.; Wynands, R.; Leahy, E. M.; Dougan, D. R.; Snell, G.; Navre, M.; Knuth, M. W.; Swanson, R. V.; McRee, D. E.; Tari, L. W. *Structure* **2004**, *12*, 1325-1334.

155 Gaussian 03, Revision C.02, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.;

Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V., Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian, Inc., Wallingford CT, 2004.

156 Terracciano, S.; Chini, M. G.; Bifulco, G.; D'Amico, E.; Marzocco, S.; Riccio, R.; Bruno, I. *Tetrahedron* **2010**, *66*, 2520-2528.

157 In the calculation, the ammine functionality of structures **2, 4** and **6** were considered protonated at physiological pH.

158 Halgren T. A. *J. Comput. Chem.* **1999**, *20*, 720-729.

159 a) Mohamadi, F.; Richards, N. G.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440-467; b) MacroModel, version 8.5, Schrödinger LLC, New York, NY, **2003**.

160 Gaussian 03, Revision E.01, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.;

Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc., Wallingford CT, 2004..

161 a) Grolla, A. A.; Podesta, V.: Chini, M. G.; Di Micco, S.; Vallario, A.; Genazzani, A. A.; Canonico, P. L.; Bifulco, G.; Tron, G. C.; Sorba, G.; Pirali, G. T. *J. Med. Chem.* **2009**, *52*, 2776-2785; b) Di Micco, S.; Terracciano, S.; Bruno, I.; Rodriquez, M.; Riccio, R.; Taddei, M.; Bifulco, G. *Bioorg. Med. Chem.* **2008**, *16*, 8635-8642; c) Pirali, T.; Faccio, V.; Mossetti, R.; Grolla, A. A.; Di Micco, S.; Bifulco, G.; Genazzani, A. A.; Tron, G. C. *Mol. Diversity* **2010,** *14*, 109–121; d) Wang, D-F.; Wiest, O.; Helquist, P.; Lan-Hargest, H-Y.; Wiech, N. L. *J. Med. Chem.* **2004**, *47*, 3409; e) Park, H.; Lee, S. *J. Comput.-Aided Mol. Des.* **2004**, *18*, 375-388; f) Wang, D-F.; Helquist, P.; Wiech, N. L.; Wiest, O. *J. Med. Chem.* **2005**, *48*, 6936-6947

162 a) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, Jr K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197; b) Weiner, P.K.; Kollman, P. A. *J. Comput. Chem.* **1981**, *2*, 287-303; c) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, Jr S.; Weiner, P. K. *J. Am. Chem. Soc.* **1984**, *106*, 765-784.

163 Still, W.C.; Tempczyk, A., Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127-6129.

164 Stote, R. H.; Karplus, M. *Proteins* **1995**, *23*, 12-31.

165 Sanner, M. F. J. Mol. Graphics Mod. **1999**, *17*, 57- 61.

166 Sanner, M. F.; Olson, A. J.; Spehner, J. C. *Biopolymers* **1996**, *38*, 305-320.

167 Terracciano, S.; Chini, M. G.; Riccio, R.; Bruno, I.; Bifulco, G. *ChemMedChem* **2012**, *7*, 1–9, DOI: 10.1002/cmdc.201100531.

168 Bieliauskas, A. V.; Pflum, M. K.; *Chem. Soc. Rev.* **2008,** *37*, 1402-1413.

169 Chini, M. G.; Terracciano, S.; Riccio, R.; Bifulco, G.; Ciao, R.; Gaeta, C.; Troisi, F.; Neri, P. *Organic Letters* **2010**, *12*, 5382-5385.

170 a) Kuntz, I. D.; Chen, K.; Sharp, K. A.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9997–10002; b) Abad-Zapatero, C.; Metz, J. T. *Drug Discovery Today* **2005**, *10*, 464–469; c) Hetenyi, C.; Maran, U.; Garcia-Sosa, A. T.; Karelson, M. *Bioinformatics* **2007**, *23*, 2678–2685; d) Wells, J. A.; McClendon, C. L. *Nature* **2007**, *450*, 1001–1009.

171 Chan, W. C.; White, Peter D. *Fmoc Solid Phase Peptide Synthesis A Practical Approach* Oxford University. Press, **2000.**

172 a) Aoki, Y., Kobayashi, S. *J. Comb. Chem*. **1999**, *1*, 371-372; b) Abdel-Magid, A. F.; Carson, K. G.; Harris, B. H.; Maryanoff, C. A.; Shah, R. D. *J. Org. Chem*. **1996**, *61*, 3849-3862.

173 a) Sarantakis, D.; Bicksler, J. *Tetrahedron Lett*. **1997**, *38*, 7325-7328; b) Ayesa, S.; Argrypoulos, D.; Maltseva, T.; Sund, C.; Samuelsson, B. *Eur. J. Org. Chem*. **2004**, 2723-2737.

174 Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput. Chem.* **2004**, *25*, 1605-1612.

175 a) Fletcher, S.; Hamilton, A. D. *Curr. Opin. Chem. Biol.* **2005**, *9*, 632-638; b) Peczuh, M. W.; Hamilton, A. D. *Chem. Rev.* **2000**, *100*, 2479-2494.

176 a) Zhou, H., Wang, D. A.; Baldini, L.; Ennis, E.; Jain, R.; Carie, A.; Sebti, S. M.; Hamilton A. *Org. Biomol. Chem.* **2006,** *4*, 2376-2386; b) Park, H. S.; Lin, Q.; Hamilton, A. D. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5105- 5109; c) Blaskovich, M. A.; Lin, Q.; Delarue, F. L.; Sun, J.; Park, H. S.; Coppola, D.; Hamilton, A. D.; Sebti, S. M. *Nat. Biotechnol.* **2000**, *18*, 1065-1070.

177 Analogously, Liu and coworkers have recently shown that *p*-sulfonatocalix[*n*]arenes do not show toxic effect in mices poisoned with viologens derivatives. In addition, the mortality rate of viologen poisoned mice was significantly decreased: Wang, K.; Guo, D. S.; Zhang, H. Q.; Li, D.; Zheng, X. L.; Liu, Y. *J. Med. Chem.* **2009**, 52, 6402-6412.

178 Francese, S.; Cozzolino, A.; Caputo, I.; Esposito, C.; Martino, M.; Gaeta, C.; Troisi, F.; Neri, P. *Tetrahedron Lett.* **2005**, *46*, 1611-1615.

179 Mecca, T.; Consoli, G. M. L.; Geraci, C.; La Spina, R.; Cunsolo, F. *Org. Biomol. Chem.* **2006,** *4*, 3763-3768.

180 Perret, F.; Lazar, A. N.; Coleman, A. W. *Chem. Commun.* **2006**, 2425-2438.

181 Baldini, L; Casnati, A.; Sansone, F.; Ungaro, R. *Chem. Soc. Rev.* **2007,** *36*, 254-266.

182 Horne, W. S.; Olsen, C. A.; Beierle, J. M.; Montero, A; Ghadiri, M. R. *Angew. Chem. Int. Ed.* **2009,** *48*, 4718-4724.

183 Wang, D.-F. *Curr. Top. Med. Chem.* **2009**, *9*, 241-256.

184 Gutsche, C. D. *Calixarenes, An Introduction*; Royal Society of Chemistry: Cambridge, UK, **2008.**

185 a) Troisi, F.; Gaeta, C.; Pierro, T.; Neri, P. *Tetrahedron Lett.* **2009**, *50*, 5113-5115; b) Troisi, F.; Russo, A.; Gaeta, C.; Bifulco, G.; Neri, P. *Tetrahedron Lett.* **2007**, *48*, 7986-7989.

186 Nakao, Y.; Yoshida, S.; Matsunaga, S.; Shindoh, N.; Terada, Y.; K. Nagai, Yamashita, J. K.; Ganesan, A.; van Soest, R. W.; Fusetani, N. *Angew. Chem., Int. Ed.* **2006**, *45*, 7553-7557.

187 a) Furumai, R.; Komatsu, Y.; Nishino, N.; Khochbin, S.; Yoshida, M.; Horinouchi, S. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 87–92; b) Yoshida, M.; Furumai, R.; Nishiyama, M.; Komatsu, Y.; Nishino, N.; Horinouchi, S. *Cancer Chemother. Pharmacol.* **2001**, *48*, 20–26.

188 Komatsu, Y.; Tomizaki, K.; Tsukamoto, M.; Kato, T.; Nishino, N.; Sato, S.; Yamori, T.; Tsuruo, T.; Furumai, R.; Yoshida, M.; Horinouchi, S.; Hayashi, H. *Cancer Res.* **2001**, *61*, 4459–4466.

189 a) Ugi, I.; Meyr, R.; Fetzer, U.; Steinbruckner, C. *Angew. Chem., Int. Ed. Engl.* **1959**, *71*, 386–388; b) Dömling, A.; Ugi, I. *Angew. Chem., Int. Ed. Engl.* **2000**, *39*, 3168–3210.

190 a) Condorelli, F.; Gnemmi, I.; Vallario, A.; Genazzani, A. A.; Canonico, P. L. *Br. J. Pharmacol.* **2008**, *153*, 657–668; b) Pirali, T.; Pagliai, F.; Mercurio, C.; Boggio, R.; Canonico, P. L.; Sorba, G.; Tron, G. C.; Genazzani, A. A. *J. Comb. Chem.* **2008**, *10*, 624–627.

191 Di Micco, S.; Chini, M. G.; Terracciano, S.; Bruno, I.; Riccio, R.; Bifulco, G. *ChemMedChem* **2012**, submitted.

192 a) Galletti, P.; Quintavalla, A.; Ventrici, C.; Giannini, G.; Cabri, W.; Penco, S.; Gallo, G.; Vincenti, S.; Giacomini, D. *ChemMedChem* **2009**, *4*, 1991–2001; b) KrennHrubec, K.; Marshall, B. L.; Hedglin, M.; Verdin, E.; Ulrich, S. M. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 2874–2878; c) Moradei, O. M.; Mallais, T. C.; Frechette, S.; Paquin, I.; Tessier, P. E.; Leit, S. M.; Fournel, M.; Bonfils, C.; Trachy-Bourget, M.-C.; Liu, J.; Yan, T. P.; Lu, A.-H.; Rahil, J.; Wang, J.; Lefebvre, S.; Li, Z.; Vaisburg, A. F.; Besterman, J. M. *J. Med. Chem.* **2007,** *50*, 5543-5546; d) Estiu, G.; Greenberg, E.; Harrison, C. B.; Kwiatkowski, N. P.; Mazitschek, R.; Bradner, J. E.; Wiest, O. *J. Med. Chem.* **2008,** *51*, 2898–2906; e) Schäfer, S.; Saunders, L.; Schlimme, S.; Valkov, V.; Wagner, J. M.; Kratz, F.; Sippl, W.; Verdin, E.; Jung, M. *ChemMedChem* **2009**, *4*, 283–290.

193 a) Itoh, Y.; Suzuki, T.; Kouketsu, A.; Suzuki, N.; Maeda, S.; Yoshida, M.; Nakagawa, H.; Miyata, N. *J. Med. Chem.* **2007**, *50*, 5425-5438; b) Kozikowski, A. P.; Chen, Y.; Gaysin, A.; Chen, B.; D'Annibale, M. A.; Suto, C. M.; Langley, B.C. *J. Med. Chem.* **2007**, *50*, 3054-3061; c)

Mai, A.; Massa, S.; Pezzi, R.; Simeoni, S.; Rotili, D.; Nebbioso, A.; Scognamiglio, A.; Altucci, L.; Loidl, P.; Brosch, G. *J. Med. Chem.* **2005,** *48,* 3344-3353; d) Siliphaivanh, P.; Harrington, P.; Witter, D. J.; Otte, K.; Tempest, P.; Kattar, S.; Kral, A. M.; Fleming, J. C.; Deshmukh, S. V.; Harsch, A.; Secrist, P. J.; Miller, T. A. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 4619–4624.

194 Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Phys.* **2005**, *123*, 161103-1-4.

195 Wang, D. F., Helquist, P., Wiech, N. L.; Wiest, O. *J. Med. Chem.* **2005**, *48*, 6936-6947.

196 a) Methot, J. L.; Hamblett, C. L.; Mampreian, D. M.; Jung, J.; Harsch, A.; Szewczak, A. A.; Dahlberg, W. K.; Middleton, R. E.; Hughes, B.; Fleming, J. C.; Wang, H.; Kral, A. M.; Ozerova, N.; Cruz, J. C.; Haines, B.; Chenard, M.; Kenific, C. M.; Secrist, J. P.; Miller, T. A. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 6104–6109; b) Hamblett, C. L.; Methot, J. L.; Mampreian, D. M.; Sloman, D. L.; Stanton, M. G.; Kral, A. M.; Fleming, J. C.; Cruz, J. C.; Chenard, M.; Ozerova, N.; Hitz, A. M.; Wang, H.; Deshmukh, S. V.; Nazef, N.; Harsch, A.; Hughes, B.; Dahlberg, W. K.; Szewczak, A. A.; Middleton, R. E.; Mosley, R. T.; Secrist, J. P.; Miller, T. A. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 5300–5309.

197 Saito, A.; Yamashita, T.; Mariko, Y.; Nosaka, Y.; Tsuchiya, K.; Ando, T.; Suzuki, T.; Tsuruo, T.; Nakanishi, O. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 4592-4597.

198 Byrd, J. C.; Marcucci, G.; Parthun, M. R.; Xiao, J. J.; Klisovic, R. B.; Moran, M.; Lin, T. S.; Liu, S.; Sklenar, A. R.; Davis, M. E.; Lucas, D.

M.; Fischer, B.; Shank, R.; Tejaswi, S. L.; Binkley, P.; Wright, J.; Chan, K. K.; Grever, M. R. *Blood* **2005**, *105*, 959–967.

199 Singh, S. B.; Zink, D. L.; Liesch, J. M.; Mosley, R. T.; Dombrowski, A. W.; Bills, G. F.; Darkin-Rattray, S. J.; Schmatz, D. M.; Goetz, M. A. *J. Org. Chem.* **2002**, *67*, 815-825.

200 Atadja, P.; Gao, L.; Kwon, P.; Trogani, N.; Walker, H.; Hsu, M.; Yeleswarapu, L.; Chandramouli, N.; Perez, L.; Versace, R.; Wu, A.; Sambucetti, L.; Lassota, P.; Cohen, D.; Bair, K.; Wood, A.; Remiszewski, S. *Cancer Res.* **2004**, *64*, 689–695.

201 a) Methot, J. L.; Hamblett, C. L.; Mampreian, D. M.; Jung, J.; Harsch, A.; Szewczak, A. A.; Dahlberg, W. K.; Middleton, R. E.; Hughes, B.; Fleming, J. C.; Wang, H.; Kral, A. M.; Ozerova, N.; Cruz, J. C.; Haines, B.; Chenard, M.; Kenific, C. M.; Secrist, J. P.; Miller, T. A. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 6104–6109; b) Hamblett, C. L.; Methot, J. L.; Mampreian, D. M.; Sloman, D. L.; Stanton, M. G.; Kral, A. M.; Fleming, J. C.; Cruz, J. C.; Chenard, M.; Ozerova, N.; Hitz, A. M.; Wang, H.; Deshmukh, S. V.; Nazef, N.; Harsch, A.; Hughes, B.; Dahlberg, W. K.; Szewczak, A. A.; Middleton, R. E.; Mosley, R. T.; Secrist, J. P.; Miller, T. A. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 5300–5309.

202 Witter, D. J.; Harrington, P.; Wilson, K. J.; Chenard, M.; Fleming, J. C.; Haines, B.; Kral, A. M.; Secrist J. P.; Miller, T. A. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 726–731.

203 Hu, E.; Dul, E.; Sung, C.-M.; Chen, Z.; Kirkpatrick, R.; Zhang, G.-F.; Johanson, K.; Liu, R.; Lago, A.; Hofmann, G.; Macarron, R.; De Los

Frailes, M.; Perez, P.; Krawiec, J.; Winkler, J.; Jaye, M. *J. Pharmacol. Exp. Ther*. **2003**, *307*, 720–728.

204 Dowling, D. P.; Gantt, S. L.; Gattis, S. G.; Fierke, C. A.; Christianson, D. W. *Biochemistry* **2008**, *47*, 13554-13563.

205 J. R.; Skene, R. J.; Katz, B. A.; Mol, C.; Ho, J. D.; Jennings, A. J.; Luong, C.; Arvai, A.; Buggy, J. J.; Chi, E.; Tang, J.; Sang, B.-C.; Verner, E.; Wynands, R.; Leahy, E. M.; Dougan, D. R.; Snell, G.; Navre, M.; Knuth, M. W.; Swanson, R. V.; McRee, D. E.; Tari, L. W. *Structure* **2004**, *12*, 1325-1334.

206 Balasubramanian, S.; Ramos, J.; Luo, W.; Sirisawad, M.; Verner, E.; Buggy, J. J. *Leukemia* **2008**, *22*, 1026-1034.

207 Mai, A.; Jelicic, K.; Rotili, D.; Di Noia, A.; Alfani, E.; Valente, S.; Altucci, L.; Nebbioso, A.; Massa, S.; Galanello, R.; Brosch, G.; Migliaccio, A. R.; Migliaccio, G. *Mol. Pharmacol.* **2007**, *72*, 1111–1123.

208 Haggarty, S. J.; Koeller, K. M.; Wong, J. C.; Butcher, R. A.; Schreiber, S. L. *Chem. Biol.* **2003**, *10*, 383-396.

209 Suzuki, T.; Kouketsu, A.; Itoh, Y.; Hisakawa, S.; Maeda, S.; Yoshida, M.; Nakagawa, H.; Miyata, N. *J. Med. Chem.* **2006**, *49*, 4809-4812.

210 Butler, K. V.; Kalin, J.; Brochier, C.; Vistoli, G.; Langley, B.; Kozikowski, A. P. A. *J. Am. Chem. Soc.* **2010**, *132*, 10842–10846.

211 Schuetz, A.; Min, J.; Allali-Hassani, A.; Schapira, M.; Shuen, M.; Loppnau, P.; Mazitschek, R.; Kwiatkowski, N. P.; Lewis, T. A.; Maglathin, R. L.; McLean, T. H.; Bochkarev, A.; Plotnikov, A. N.; Vedadi, M., Arrowsmith, C. H. *J. Biol. Chem.* **2008**, *283*, 11355-11363.

212 Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. *J. Mol. Biol.* **1990**, *215*, 403-410.

213 Henikoff, S.; Henikoff, J. G. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 10915–10919.

214 Ayer, D. E. *Trends Cell Biol.* **1999**, *9*, 193-198.

215 Zou, H.; Wu, Y.; Navre, M.; Sang, B. C. *Biochem. Biophys. Res. Commun.* **2006**, *341*, 45-50.

216 a) Fiser, A; Sali, A. *Meth. Enzymol.* **2003**, *374*, 461–91; b) N. Eswar, M. A. Marti-Renom, B. Webb, M. S. Madhusudhan, D. Eramian, M. Shen, U. Pieper, A. Sali. *Current Protocols in Bioinformatics*, John Wiley & Sons, Inc., Supplement **2006**; b) Marti-Renom, M. A.; Stuart, A.; Fiser, A.; Sánchez, R.; Melo, F.; Sali. A. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291-325; c) Sali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*, 779-815; d) Fiser, A.; Do, R. K.; Sali. A. *Protein Sci.* **2000**, *9*, 1753-1773.

217 Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474-6487.

218 Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. App. Cryst.*, **1993,** *26*, 283-291.

219 Hasmann, M.; Schemainda, I. *Cancer Res.* **2003**, *63*, 7436–7442.

220 http://www.clinicaltrials.gov/ct2/results?term=GMX1777; accessed on May 25, 2009.

221 Colombano, G; Travelli, C; Galli, U; Caldarelli, A; Chini, M G; Canonico, P L; Sorba, G; Bifulco, G; Tron, G C; Genazzani, A. *J. Med. Chem.* **2010**, *53*, 616-623.

222 Khan, J. A.; Forouhar, F.; Tao, X.; Tong, L. *Expert Opin. Ther. Targets* **2007**, *11*, 695–705.

223 Denu, J. M. *Curr. Opin. Chem. Biol.* **2005**, *9*, 431–440.

224 Belenky, P.; Bogan, K. L.; Brenner, C. *Trends Biochem. Sci.* **2007**, *32*, 12–19.

225 Berger, F.; Ramírez-Hernández, M. H.; Ziegler, M. *Trends Biochem. Sci.* **2004**, *29*, 111–118.

226 Genazzani, A. A.; Billington, R. A. *Trends Pharmacol. Sci.* **2002**, *23*, 165–167.

227 Malavasi, F.; Deaglio, S.; Funaro, A.; Ferrero, E.; Horenstein, A. L.; Ortolan, E.; Vaisitti, T.; Aydin, S. *Physiol. Rev.* **2008**, *88*, 841–886.

228 Magni, G.; Amici, A.; Emanuelli, M.; Orsomando, G.; Raffaelli, N.; Ruggieri, S. *Cell. Mol. Life Sci.* **2004**, *61*, 19–34.

229 Bieganowski, P.; Brenner, C. *Cell* **2004**, *117*, 495–502.

230 Gallí, M.; Van Gool, F.; Rongvaux, A.; Andris, F.; Leo, Oberdan. *Cancer Res* **2010**, *70*, 1-11.

231 Nahimana, A.; Attinger, A.; Aubry, D.; Greaney, P.; Ireson, C.; Thougaard, A. V.; Tjoernelund, J.; Dawson, K. M.; Dupuis, M.; Duchosal, M A. *Blood* **2009**, *113*, 3276–3286.

232 http://www.clinicaltrials.gov/ct2/results?term=APO866; accessed on May 25, 2009.

233 Schou, C.; Ottosen, E. R.; Petersen, H. J.; Björkling, F.; Latini, S.; Hjarnaa, P. V.; Bramm, E.; Binderup, L. *Bioorg. Med. Chem.* **1997**, *7*, 3095–3100.

234 Olesen, U. H.; Christensen, M. K.; Björkling, F.; Jäättelä, M.; Jensen, P. B.; Sehested, M.; Nielsen, S. J. *Biochem. Biophys. Res. Commun.* **2008**, *367*, 799–804.

235 Galli, U.; Ercolano, E.; Carraro, L.; Blasi Roman, C. R.; Sorba, G.; Canonico, P. L.; Genazzani, A. A.; Tron, G. C.; Billington, R. A. *ChemMedChem* **2008**, *3*, 771–779.

236 Khan, J. A.; Tao, X.; Tong, L. *Nat. Struct. Mol. Biol.* **2006**, *13*, 582–588.

237 a) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem., Int. Ed. Engl.* **2001**, *40*, 2004–2021; b) Tornoee, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057–3064; c) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596–2599.

238 a) Kolb, H. C.; Sharpless, K. B. *Drug Discovery Today* **2003**, *8*, 1128–1137; b) Tron, G. C.; Pirali, T.; Billington, R. A.; Canonico, P. L.; Sorba, G.; Genazzani, A. A. *Med. Res. Rev.* **2008**, *28*, 278–308.

239 Pagliai, F.; Pirali, T.; Del Grosso, E.; Di Brisco, R.; Tron, G. C.; Sorba, G.; Genazzani, A. A. *J. Med. Chem.* **2006**, *49*, 467–470.

240 Pirali, T.; Gatti, S.; Di Brisco, R.; Tacchi, S.; Zaninetti, R.; Brunelli, E.; Massarotti, A.; Sorba, G.; Canonico, P. L.; Moro, L.; Genazzani, A. A.; Tron, G. C.; Billington, R. A. *ChemMedChem* **2007**, *2*, 437–440.

241 a) Pirali, T.; Pagliai, F.; Mercurio, C.; Boggio, R.; Canonico, P. L.; Sorba, G.; Tron, G. C.; Genazzani, A. A. *J. Comb. Chem.* **2008**, *10*, 624–627; b) Brik, A.; Muldoon, J.; Lin, Y.; Elder, J. H.; Goodsell, D. S.; Olson, A. J.; Fokin,V. V.; Sharpless, K. B.; Wong, C. *ChemBioChem* **2003**, *4*, 1246–1248; c) Lee, L. V.; Mitchell, M. L.;

Huang, S.; Fokin, V. V.; Sharpless, K. B.; Wong, C. A. *J. Am. Chem. Soc.* **2003**, *125*, 9588–9589.

242 Beauparlant, P.; Bédard, D.; Bernier, C.; Chan, H.; Gilbert, K.; Goulet, D.; Gratton, M. O.; Lavoie, M.; Roulston, A.; Turcotte, E.; Watson, M. *Anti-Cancer Drugs*, **2009**, *20*, 346-354.

243 Koeberle, A.; Werz, O. *Curr. Med. Chem.* **2009**, *16*, 4274–4296.

244 Funk, C. D. *Science* **2001**, *294*, 1871-1875.

245 Wang, D.; Dubois, R. N. *Gut*. **2006**, *55*, 115-122.

246 Buttgereit, F.; Burmester, G. R.; Simon, L. S. *Am. J. Med*. **2001**, *110*, 13S-19S.

247 McGettigan, P.; Henry, D. *J. Am. Med. Assoc.* **2006**, *296*, 1633-1644.

248 Claveau, D.; Sirinyan, M.; Guay, J.; Gordon, R.; Chan, C. C.; Bureau, Y.; Riendeau, D.; Mancini, J. A. *J. Immunol*. **2003**, *170*, 4738-4744.

249 Yoshimatsu, K.; Golijanin, D.; Paty, P. B.; Soslow, R. A.; Jakobsson, P. J.; DeLellis, R. A.; Subbaramaiah, K.; Dannenberg, A. J. *Clin. Cancer Res.* **2001**, *7*, 3971–3976.

250 Yoshimatsu, K.; Altorki, N. K.; Golijanin, D.; Zhang, F.; Jakobsson, P. J.; Dannenberg, A. J.; Subbaramaiah, K. *Clin. Cancer Res.* **2001**, *7*, 2669–2674.

251 Mehrotra, S.; Morimiya, A.; Agarwal, B.; Konger, R.; Badve, S. *J. Pathol.* **2006**, *208*, 356–363.

252 a) Murakami, M.; Naraba, H.; Tanioka, T.; Semmyo, N., Nakatani, Y., Kojima, F., Ikeda, T., Fueki, M., Ueno, A., Oh-Ishi, S. Kudo, I. *J. Biol. Chem.* **2000**, *275*, 32783-32792; b) Lazarus, M.: Kubata, B.K.; Eguchi, N.; Fujitani, Y.; Urade, Y.; Hayaishi, O. *Arch. Biochem. Biophys.* **2002,** *397*, 336-341

253 Matsumoto, H.; Naraba, H.; Murakami, M.; Kudo, I.; Yamaki, K.; Ueno, A.; Oh-ishi, S. *Biochem. Biophys. Res. Commun* **1997**, *230* 110-114.

254 Thoren, S.; Jakobsson, P.J. *Eur. J. Biochem.* **2000**, *267*, 6428-6434.

255 Arai, I.; Hamasaka, Y.; Futaki, N.; Takahashi, S.; Yoshikawa, K.; Higuchi, S.; Otomo, S. *Res. Commun. Chem. Pathol. Pharmacol.* **1993**, *81*, 259-270.

256 Riendeau, D.; Aspiotis, R.; Ethier, D.; Gareau, Y.; Grimm, E. L.; Guay, J.; Guiral, S.; Juteau, H.; Mancini, J. A.; Méthot, N.; Rubin, J.; Friesen, R. W. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3352-3355.

257 Jakobsson, P. J.; Thorén, S.; Morgenstern, R.; Samuelsson, B. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 7220–7225.

258 Friesen, R. W.; Mancini, J. A. *J. Med. Chem.* **2008**, *51*, 4059–4067.

259 Murakami, M.; Naraba, H.; Tanioka, T.; Semmyo, N.; Nakatani, Y.; Kojima, F.; Ikeda, T.; Fueki, M.; Ueno, A.; Oh, S.; Kudo, I. *J. Biol. Chem.* **2000**, *275*, 32783–32792.

260 Jegerschöld, C.; Pawelzik, S. C.; Purhonen, P.; Bhakat, P.; Gheorghe, K. R.; Gyobu, N.; Mitsuoka, K.; Morgenstern, R.; Jakobsson, P. J.; Hebert, H. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 11110–11115.

261 Congreve, M.; Chessari, G.; Tisi, D.; Woodhead A. J. *J. Med. Chem* **2008**, *51*, 3661-3680.

262 Shelke, S. V.; Cutting, B.; Jiang, X.; Koliwer-Brandl, H.; Strasser, D. S., Schwardt, O.; Kelm, S.; Ernst, Beat. *Angew. Chem. Int. Ed.* **2010**, *49*, 5721 –5725.

263 Nestler, H. P. *Curr. Drug Discovery Technol.* **2005**, *2*, 1-12.

264 Sreeman, W.; Mamidyala, K.; Finn, M. G. *Chem. Soc. Rev.* **2010**, *39*, 1252–126.

265 Moses, J. E.; Moorhouse, A. D. *Chem. Soc. Rev.* **2007**, *36*, 1249–1262.

266 Rees, D. C.; Congreve, M.; Murray, C. W.; Carr, R. *Nat. Rev. Drug Discovery* **2004**, *3*, 660-672.

267 Kolb, H. C.; Sharpless, K. B. *DDT* **2003**, *8*, 1128-1137.

268 Sharpless, K. B.; Manetsch, R. *Expert Opinion on Drug Discovery* **2006**, *1*, 525-538.

269 De Simone, R.; Chini, M. G.; Bruno, I.; Riccio, R.; Mueller, D.; Werz, O.; Bifulco, G. *Journal of Medicinal Chemistry* **2011**, *54*, 1565-1575.

270 Pasha, F. A.; Muddassan, M.; Jung, H.; Yang, B. S.; Lee, C.; Soo Oh, J.; Joo Cho, S.; Cho, H. *Bull. Korean Chem.* **2008**, *29*, 647–655.

271 Giroux, A.; Boulet, L.; Brideau, C.; Chau, A.; Claveau, D.; Cote, B.; Ethier, D.; Frenette, R.; Gagnon, M.; Guay, J.; Guiral, S.; Mancini, J.; Martins, E.; Masse, F.; Methot, N.; Riendeau, D.; Rubin, J.; Xu, D.; Yu, H.; Ducharme, Y.; Friesen, R. W. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 5837–5841.

272 Wang, J.; Limburg, D.; Carter, J.; Mbalaviele, G.; Gierse, J.; Vazquez, M. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 1604–1609.

273 San Juan, A. A.; Cho, S. J. *J. Mol. Model.* **2007**, *13*, 601–610.

274 AbdulHameed, M. D.; Hamza, A.; Liu, J.; Huang, X.; Zhan, C. G. *J. Chem. Inf. Model.* **2008**, *48*, 179–185.

275 San Juan, A. A.; Cho, S. J.; Cho, H. *Bull. Korean Chem. Soc.* **2006**, *27*, 1531–1536.

276 Calvin, Y. A. A. *J. Taiwan Inst. Chem.* **2009**, *40*, 155–161.

277 Rorsch, F.; Wobst, I.; Zettl, H.; Schubert-Zsilavecz, M.; Grosch, S.; Geisslinger, G.; Schneider, G.; Proschak, E. *J. Med. Chem.* **2010**, *53*, 911–915.

278 Cote, B.; Boulet, L.; Brideau, C.; Claveau, D.; Ethier, D.; Frenette, R.; Gagnon, M.; Giroux, A.; Guay, J.; Guiral, S.; Mancini, J.; Martins, E.; Masse, F.; Methot, N.; Riendeau, D.; Rubin, J.; Xu, D.; Yu, H.; Ducharme, Y.; Friesen, R. W. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 6816–6820.

279 Hamza, A.; AbdulHameed, M. D.; Zhan, C. G. *J. Phys. Chem. B* **2008**, *112*, 7320–7329.

280 Hamza, A.; Tong, M.; AbdulHameed, M. D. M.; Liu, J.; Goren, A. C.; Tai, H. H.; Zhan, C. G. *J. Phys. Chem.* **2010**, *114*, 5605–5616.

281 Quraishi, O.; Mancini, J. A.; Riendeau, D. *Biochem. Pharmacol.* **2002**, *63*, 1183–1189.

282 Riendeau, D.; Aspiotis, R.; Ethier, D.; Gareau, Y.; Grimm, E. L.; Guay, J.; Guiral, S.; Juteau, H.; Mancini, J. A.; Methot, N.; Rubin, J.; Friesen, R. W. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3352–3355.

283 Holm, P. J.; Bhakat, P.; Jegerschold, C.; Gyobu, N.; Mitsuoka, K.; Fujiyoshi, Y.; Morgenstern, R.; Hebert, H. *J. Mol. Biol.* **2006**, *360*, 934–945.

284 Thoren, S.; Weinander, R.; Saha, S.; Jegerschold, C.; Pettersson, P. L.; Samuelsson, B.; Hebert, H.; Hamberg, M.; Morgenstern, R.; Jakobsson, P. J. *J. Biol. Chem.* **2003**, *278*, 22199–22209.

285 Werz, O. 5-Lipoxygenase: *Curr. Drug Targets Inflamm. Allergy* **2002**, *1*, 23–44.

286 Mendis, C.; Campbell, K.; Das, R.; Yang, D.; Jett, M. *FEBS J.* **2008**, *275*, 3088–3098.

287 Ferguson, A. D.; McKeever, B. M.; Xu, S.; Wisniewski, D.; Miller, D. K.; Yamin, T. T.; Spencer, R. H.; Chu, L.; Ujjainwalla, F.; Cunningham, B. R.; Evans, J. F.; Becker, J. W. *Science* **2007**, *317*, 510–512.

288 Gillmor, S. A.; Villasenor, A.; Fletterick, R.; Sigal, E.; Browner, M. F. *Nat. Struct. Biol.* **1997**, *4*, 1003–1009.

289 Werz, O. *Curr. Drug Targets* **2002**, *1*, 23-44.

290 Charlier, C.; Henichart, J. P.; Durant, F.; Wouters, J. *J. Med. Chem.* **2006**, *49*, 186–195.

291 Monti, M. C.; Chini, M. G.; Margarucci, L.; Tosco, A.; Riccio, R.; Bifulco, G.; Casapullo, A. *J. Mol. Recognit.* **2009**, *22*, 530-537.

292 Monti, M. C.; Chini, M. G.; Margarucci, L.; Riccio, R.; Bifulco, G.; Casapullo, A. *Chembiochem.* **2011**, *12*, 2686-2691.

293 Sepe, V.; Bifulco, G.; Renga, B.; D'Amore, C.; Fiorucci, S.; Zampella, A. *J. Med. Chem.* **2011**, *54*, 1314-1320.

294 De Marino, S.; Ummarino, R.; D'auria, M. V.; Chini, M. G.; Bifulco, G.; Renga, B.; D'Amore, C.; Fiorucci, S.; Debitus, C.; Zampella, A. *J. Med. Chem.* **2011**, *54*, 3065-3075.

295 Festa, C.; De, M. S.; D'auria, M. V.; Bifulco, G.; Renga, B.; Fiorucci, S.; Petek, S.; Zampella, A. *J. Med. Chem.* **2011**, *54*, 401-405.

296 Cipriani, S.; Mencarelli, A.; Chini, M. G.; Distrutti, E.; Renga, B.; Bifulco, G.; Baldelli, F.; Donidi, A.; Fiorucci, S. *PloS One* **2011**, *6*, e25637, 1-11.

297 Chini, M. G.; De Simone, R.; Bruno, I.; Riccio, R.; Dehm, F.; Weinigel, C.; Barz, D.; Werz, O.; Bifulco, G. *J. Med. Chem.* **2011** submitted

298 Huisgen, R. *1,3-Dipolar cycloaddition – introduction, survey, mechanism. In 1,3-Dipolar Cycloaddition Chemistry*; PADWA, A., Ed.; Wiley: New York, **1984**.

299 Miyaura, N.; Yamada, K.; Suzuki, A. *Tetrahedron Letters* **1979,** *20*, 3437-3440.

300 Koeberle, A.; Zettl, H.; Greiner, C.; Wurglics, M.; Schubert-Zsilavecz, M.; Werz, O. *J. Med. Chem.* **2008,** *51*, 8068-8076.

301 Hieke, M.; Greiner, C.; Dittrich, M.; Reisen, F.; Schneider, G.; Schubert-Zsilavecz, M.; Werz, O. *J. Med. Chem.* **2011,** *54*, 4490-4507.

302 Gilbert, N. C.; Bartlett, S. G.; Waight, M. T.; Neau, D. B.; Boeglin, W. E.; Brash, A. R.; Newcomer, M. E. *Science* **2011,** *331*, 217-219.

303 Fang-Yu, L.; Yufeng, J. T. *J. Chem. Inf. Model.* **2011,** *51*, 1703-1715.

304 Li, Y.; Chen, S. H.; Ou, T. M.; Tan, J. H.; Li, D.; Gu, L. Q.; Huang, Z. S. *Bioorg. Med. Chem.* **2011,** *19*, 2074-2083.

305 a) Balsinde, J.; Balboa, M.; Insel, P.; Dennis, E. *Annu. Rev. Pharmacol. Toxicol.* **1999**, *39*, 175–189; b) Berg, O. G.; Yu, B. Z.; Rogers, J.; Jain, M. K. *Biochemistry* **1991**, *30*, 7283–7297.

306 Parente L. *J. Rheumatol.* **2001**, *28*, 2375–2382.

307 Lee, H. Y.; Bahn, S. C.; Shin, J. S.; Hwang, I.; Back, K.; Doelling, J. H.; Ryu, S. B. *Prog. Lipid Res.* **2005**, *44*, 52–67.

308 a) Kuchler, K.; Gmachl, M.; Sippl, M. J.; Kreil, G. *Eur. J. Biochem.* **1989**, *184*, 249–254; b) Ryu, Y.; Oh, Y.; Yoon, J.; Cho, W.; Baek, K. *Biochim. Biophys. Acta* **2003**, *1628*, 206–210.

309 McIntosh, J. M.; Ghomashchi, F.; Gelb, M. H.; Dooley, D. J.; Stoehr, S. J.; Giordani, A. B.; Naisbitt, S. R.; Olivera, B. M. *J. Biol. Chem.* **1995**, *27*, 3518–3526.

310 Kini, R. M. *Toxicon* **2003**, *42*, 827–840.

311 Valentin, E.; Lambeau, G. *Biochim. Biophys. Acta* **2000**, 1488, 59–70.

312 Valentin, E.; Ghomashchi, F.; Gelb, M. H.; Lazdunski, M.; Lambeau, G. *J. Biol. Chem.* **2000**, *275*, 7492–7496.

313 a) Navalainen, T. J.; Aho, H. J.; Eskola, J. U.; Suonpaa, A. K. *Acta Pathol. Microbiol. Immunol. Scand.* **1983**, *91*, 97–102; b) Forster, S.; Ilderton, E.; Norris, J. F. B.; Summerly, R.; Yardley, H. J. *Br. J. Dermatol.* **1985**, *27*, 112–135; c) Seilhamer, J. J.; Pruzanski, W.; Vadas, P.; Plant, S.; Miller, J. A.; Kloss, J.; Johnson, L. K. *J. Biol. Chem.* **1989**, *264*, 5335–5340; d) Hurt-Camejo, E.; Camejo, G.; Peilot, H.; Oorni, K.; Kovanen, P. *Circ. Res.* **2001**, *17*, 298–304.

314 Vadas, P.; Pruzanski, W. *Lab. Invest.* **1986**, *55*, 391–396.

315 Goracci, G.; Ferrini, M.; Nardicchi, V. *J. Mol. Neurobiol.* **2010**, *41*, 274–289.

316 Mirtti, T.; Lain, V. J.; Hiekkanen, H.; Hurme, S.; Rowe, O.; Nevalainen, T. J.; Kallajoki, M.; Alanen, K. *APMIS* **2009**, *117*, 151–161.

317 Potts, B.; Faulkner, D.; Jacobs, R. *J. Nat. Prod.* **1992**, *55*, 1701–1717.

318 Terracciano, S.; Rodriguez, M.; Aquino, M.; Monti, M. C.; Casapullo, A.; Riccio, R.; Gomez-Paloma, L. *Curr. Med. Chem.* **2006**, *13*, 1947–1969.

319 Gomez-Paloma, L.; Monti, M. C.; Terracciano, S.; Casapullo, A.; Riccio, R. *Curr. Org. Chem.* **2005**, *9*, 1419–1427.

320 Dal Piaz, F.; Casapullo, A.; Randazzo, A.; Riccio, R.; Pucci, P.; Marino, G.; Gomez-Paloma, L. *ChemBioChem* **2002**, *3*, 664–671.

321 Monti, M. C.; Casapullo, A.; Riccio, R.; Gomez-Paloma, L. *Bioorg. Med. Chem.* **2004**, *12*, 1467–1474.

322 Monti, M. C.; Casapullo, A.; Riccio, R.; Gomez-Paloma, L. *FEBS Lett.* **2004**, *578*, 269–274.

323 De Silva, E. D.; Scheuer, P. J. *Tetrahedron Lett.* **1981**, *22*, 3147- 3150.

324 Albizati, K. F.; Holman, T.; Faulkner, D. J.; Glaser, K. B.; Jacobs, R. S. *Experientia* **1987**, *43*, 949-950.

325 Kernan, M. R.; Faulkner, D. J.; Jacobs, R. S. *J. Org. Chem.* **1987**, *52*, 3081-3083.

326 Kernan, M. R.; Faulkner, D. J.; Parkanyi, L.; Clardy, J.; De Carvalho, M. S.; Jacobs, R. S. *Experientia* **1989**, *45*, 388-390.

327 a) De Rosa, S.; De Stefano, S.; Zavodnik, N. *J. Org. Chem.* **1988**, *53*, 5020-5023; b) De Rosa, S.; Crispino, A.; De Giulio, A.; Iodice, C.; Pronzato, R.; Zavodnik, N. *J. Nat. Prod.* **1995**, *58*, 1776-1780.

328 Randazzo, A.; Debitus, C.; Minale, L.; Pastor, P. G.; Alcaraz, M. J.; Payà, M.; Gomez-Paloma, L. *J. Nat. Prod.* **1998**, *61*, 571-575.

329 Monti, M. C.; Casapullo, A.; Cavasotto, C. N.; Tosco, A.; Dal Piaz, F.; Ziemys, A.; Margarucci, L.; Riccio, R. *Chem. Eur. J.* **2009**, *15*, 1155–1163.

330 Monti, M. C.; Casapullo, A.; Gomez-Paloma, L. *Oceanis* **2001**, *27*, 499–511.

331 Cimino, G.; De Stefano, S.; Minale, L. *Tetrahedron* **1972**, *28*, 1315–1318.

332 a) Gil, B.; Sanz, M. J.; Terencio, M. C.; De Giulio, A.; De Rosa, S.; Alcaraz, M. J.; Payà M. *Eur. J. Pharmacol.* **1995**, *285*, 281–288; b) Coop, B. R.; Ireland, C.; Barrows, L. *J. Org. Chem.* **1991**, *56*, 4596–4597.

333 De Guzman, F.; Coop, B.; Mayne, C.; Conception, G.; Mangalindan, G.; Barrows, L.; Ireland, C. *J. Org. Chem.* **1998**, *63*, 8042–8044.

334 Soriente, A.; De Rosa, M.; Scettri, A.; Sodano, G.; Terencio, M. C.; Payà, M.; Alcaraz, M. J. *Curr. Med. Chem.* **1999**, *6*, 415–431.

335 Glaser, K. B.; De Carvalho, M.; Jacobs, R.; Kernan, M.; Faulkner, D. J. *Mol. Pharm.* **1989**, *36*, 782–788

336 Jacobson, P.; Marshall, L.; Sung, A.; Jacobs, R. *Biochem. Pharmacol.* **1990**, *39*, 1557–1564.

337 Dal Piaz, F.; Casapullo, A.; Randazzo, A.; Riccio, R.; Pucci, P.; Marino, G.; Gomez-Paloma, L. *ChemBioChem.* **2002**, *3*, 664–671.

338 Lucas, R.; Giannini, C.; D'Auria, M. V.; Payà, M. *J. Pharmacol. Exp. Ther.* **2003**, *304*, 1172–1180.

339 Busserolles, J.; Payà, M.; D'Auria, M. V.; Gomez-Paloma, L.; Alcaraz, M. J. *Biochem. Pharmacol.* **2005**, *69*, 1433–1440.

340 Monti, M. C.; Casapullo, A.; Santomauro, C.; D'Auria, M. V.; Riccio, R.; Gomez-Paloma, L. *ChembioChem* **2006**, *7*, 971–980.

341 Guerrero, M. D.; Aquino, M.; Bruno, I.; Terencio, M. C.; Payà, M.; Riccio, R.; Gomez-Paloma, L. *J. Med. Chem.* **2007**, *50*, 2176–2184.

342 Aquino, M.; Guerrero, M. D.; Bruno, I.; Terencio, M. C.; Payà, M.; Riccio, R. *Bioorg. Med. Chem.* **2008**, *16*, 9056–9064.

343 Mirtti, T.; Lain, V. J.; Hiekkanen, H.; Hurme, S.; Rowe, O.; Nevalainen, T. J.; Kallajoki, M.; Alanen, K. *APMIS* **2009**, *117*, 151–161.

344 a) Szabo, A.; Stolz, L.; Granzow, R. *Curr. Opin. Struct. Biol.* **1995**, *5*, 699–705; b) Myszka, D. G. *Anal. Biochem.* **2004**, *329*, 316–323.

345 Jain, M.; Yu, B.; Rogers, J.; Smith, A.; Booger, E. T.; Ostrander, R. L.; Rheingold, A. *Phytochemistry* **1995**, *39*, 537–547.

346 Schaloske, R. H.; Dennis, E. A. *Biochem. Biophys. Acta* **2006**, *1761*, 1246–1259.

347 Scott, D. L.; White, S. P.; Browning, J. L.; Rosa, J. J.; Gelb, M. H.; Sigler, P. B. *Science* **1991**, *254*, 1007–1010.

348 Schuchmann, M. N.; Bothe, E.; von Sonntag, J.; von Sonntag, C. *J. Chem. Soc., Perkin Trans.* **1998**, *2*, 791–796.

349 Giannini, C.; Debitus, C.; Lucas, R., Ubeda, A.; Payà, M.; Hooper, J.; D'Auria, M. V. *J. Nat. Prod.* **2001**, *64*, 612–615.

350 Scott, D. L.; Otwinowski, Z.; Gelb, M. H.; Sigler, P.B.; *Science* **1990**, *250*, 1563-1566.

351 Miyaoka, H.; Yamanishi, M.; Mitome, H. *Chem. Pharm. Bull.* **2006**, *54*, 268–270.

352 Fontana, A.; Ciavatta, M. L.; Cimino, G. *J. Org. Chem.* **1998**, *63*, 2845–2849.

353 Miyaoka, H.; Yamanishi, M.; Kajiwara, Y.; Yamada, Y. *J. Org. Chem.* **2003**, *68*, 3476–3479.

354 Monti, M. C.; Casapullo, A.; Cavasotto, C. N.; Napolitano, A.; Riccio, R. *ChemBioChem* **2007**, *8*, 1585–1591.

355 Jain, M.; Yu, B.; Rogers, J.; Smith, A.; Booger, E. T.; Ostrander, R. L.; Rheingold, A. *Phytochemistry* **1995**, *39*, 537–547.

356 Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. *J. Comput. Chem.* **2009**, *30*, 2785–2791.

357 Sepe, V.; Ummarino, R.; D'Auria, M. V.; Chini, M. G.; Bifulco, G.; Renga, B.; D'Amore, C.; Debitus, C.; Fiorucci, S.; Zampella, A. *J. Med. Chem.* **2012**, *55*, 84-93.

358 De Marino, S.; Ummarino, R.; D'Auria, M. V.; Chini, M. G.; Bifulco, G.; D'Amore, C.; Renga, B.; Mencarelli, A.; Petek, S.; Fiorucci, S.; Zampella, A. *Steroids* **2012**, *doi:10.1016/j.steroids.2012.01.006*

359 Renga, B.; Mencarelli, A.; D'Amore, C.; Cipriani, S.; D'Auria, M. V.; Sepe, V.; Chini, M. G.; Monti, M. C.; Bifulco, G.; Zampella, A.; Fiorucci, S. *PLoS ONE* **2012**, *7*, e30443, 1-12.

360 a) Blumberg, B.; Evans, R. M. *Genes Dev.* **1998**, *12*, 3149–3155; b) Kliewer, S. A.; Lehmann, J. M.; Willson, T. M. *Science* **1999**, *284*, 757–760; c) Mangelsdorf, D. J.; Evans, R. M. *Cell* **1995**, *83*, 841–850.

361 a) Chawla, A.; Repa, J. J.; Evans, R. M.; Mangelsdorf, D. J. *Science* **2001**, *294*, 1866–1870; b) Shulman, A. I.; Mangelsdorf, D. J. N. *Engl. J. Med.* **2005**, *353*, 604–615.

362 Hofmann, A. F. *Liver* **2002**, *22*, 14-9.

363 Geier, A.; Wagner, M.; Dietrich, C. G.; Trauner, M. *Biochim. Biophys. Acta* **2007**, *1773*, 283-308.

364 Alrefai, W. A.; Gill, R. K. *Pharm. Res.* **2007**, *24*, 1803-1823.

365 Zollner, G.; Marschall, H. U.; Wagner, M.; Trauner, M. *Mol. Pharm.* **2006**, *3*, 231-250.

366 Pellicciari, R.; Costantino, G.; Fiorucci, S. *J. Med. Chem.* **2005**, *48*, 5383-403.

367 Fiorucci, S.; Rizzo, G.; Donini, A.; Distrutti, E.; Santucci, L. *Trends Mol. Med.* **2007**, *13*, 298-09.

368 Sonoda, J.; Xie, W.; Rosenfeld, J. M.; Barwick, J. L.; Guzelian, P. S.; Evans, R. M. *Proc. Natl. Acad. Sci U.S.A.* **2002**, *99*, 13801-13806.

369 Guo, G. L**.;** Lambert, G.; Negishi, M.; Ward, J. M.; Brewer, H. B. J.; Kliewer, S. A.; Gonzalez, F. J.; Sinal, C. J**.** *J. Biol. Chem.* **2003**, *278*, 45062-45071.

370 Renga, B.; Migliorati, M.; Mencarelli, A.; Cipriani, S.; D'Amore, C.; Distrutti, E.; Fiorucci, S. *Biochim. Biophys. Acta* **2011**, *1809***,** 157-165.

371 Stedman, C.; Liddle, C.; Coulter, S.; Sonoda, J.; Alvarez, J. G.; Evans, R. M.; Downes, M. *Proc Natl Acad Sci U.S.A.* **2006**, *103*, 11323-11328.

372 Fiorucci, S.; Cipriani, S.; Mencarelli, A, Baldelli, F.; Bifulco, G.; Zampella, A. *Mini Rev. Med. Chem.* **2011**, *11*, 753-762.

373 Kakizaki, S.; Takizawa, D.; Tojima, H.; Horiguchi, N.; Yamazaki, Y.; Mori, M. *Front. Biosci.* **2011**, *17*, 2988–2905.

374 a) Kliewer, S.; Moore, J. T., Wade, L.; Staudinger, J. L.; Watson, M. A.; Jones, S. A.; McKee, D. D.; Oliver, B. B.; Willson, T. M.; Zetterström, R. H.; Perlmann, T.; Lehmann, J. M. *Cell* **1998,** *92*, 73-82; b) Lehmann J. M.; McKee, D. D.; Watson, M. A.; Willson, T. M.; Moore, J. T.; Kliewer, S. A. *J. Clin. Invest.* **1998**, *102*, 1016-1023; c) Bertilsson, G.; Heidrich, J.; Svensson, K.; Asman, M.; Jendeberg, L.; Sydow-Bäckman, M.; Ohlsson, R.; Postlind, H.; Blomquist, P.; Berkenstam, A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12208-12213;

d) Blumberg, B.; Sabbagh, W. Jr.; Juguilon, H.; Bolado, J. Jr.; van Meter, C. M.; Ong, E. S.; Evans, R. M. *Genes Dev.* **1998**, *12*, 3195-3205.

375 Goodwin, B.; Moore, L. B.; Stoltz, C. M.; McKee, D. D.; Kliewer, S. A. *Mol. Pharmacol.* **2001**, *60*, 427–431.

376 Synold, T. W.; Dussault, I.; Forman, B. M. *Nat Med* **2001**, *7*, 584–590.

377 Staudinger, J.; Liu, Y.; Madan, A.; Habeebu, S.; Klaassen, C. D. *Drug Metab. Dispos.* **2001**, *29*, 1467–1472.

378 Schuetz, E.; Strom, S. *Nat Med* **2001**, *7*, 536–537.

379 Krasowski, M. D.; Yasuda, K.; Hagey, L. R.; Schuetz, E. G. *Mol. Endocrinol.* **2005**, 19, 1720–1739.

380 Sonoda, J.; Chong, L. W.; Downes, M.; Barish, G. D.; Coulter, S.; Liddle, C.; Lee, C. H.; Evans, R. M. *Proc Natl Acad Sci U S A* **2005**, *102*, 2198–2203.

381 El-Sankary, W.; Gibson, G. G.; Ayrton, A.; Plant, N. *Drug. Metab. Dispo.s* **2001**, *29*, 1499–1504.

382 a) Takeshita, A.; Inagaki, K.; Igarashi-Migitaka, J.; Ozawa, Y.; Koibuchi, N. *J. Endocrinol.* **2006**, *190***,** 897–902; b) Takeshita, A.; Koibuchi, N.; Oka, J.; Taguchi, M.; Shishiba, Y.; Ozawa, Y. *Eur. J. Endocrinol.* **2001**, *145*, 513–517.

383 Bachmann, K.; Patel, H.; Batayneh, Z.; Slama, J; White, D.; Posey, J.; Ekins, S.; Gold, D.; Sambucetti, L. *Pharmacol Res* **2004**, *50*, 237–246.

384 Tabb, M. M.; Kholodovych, V.; Grun, F.; Zhou, C.; Welsh, W. J.; Blumberg, B. *Environ Health Perspect* **2004**, *112*, 163–169.

385 Wang, H.; Huang, H.; Li, H.; Teotico, D. G.; Sinz, M.; Baker, S. D.; Staudinger, J.; Kalpana, G.; Redinbo, M. R.; Mani, S. *Clin. Cancer Res.* **2007**, *13*, 2488–2495.

386 Mu, Y.; Stephenson, C. R.; Kendall, C.; Saini, S. P.; Toma, D.; Ren, S.; Cai, H.; Strom, S. C.; Day, B. W.; Wipf, P.; Xie, W. *Mol. Pharmacol.* **2005**, *68*, 403–413.

387 Mani, S.; Huang, H.; Sundarababu, S.; Liu, W.; Kalpana, G., Smith, A. B., Horwitz, S. B. *Clin. Cancer Res.* **2005**, *11*, 6359–6369.

388 Moore, J. T.; Kliewer, S. A. *Toxicology* 2000, *153*, 1–10.

389 Ding, X.; Staudinger, J. L. *J Pharmacol Exp Ther* **2005**, *312*, 849–856.

390 Mu, Y.; Zhang, J.; Zhang, S.; Zhou, H. H.; Toma, D.; Ren, S.; Huang, L.; Yaramus, M.; Baum, A.; Venkataramanan, R.; Xie, W. *J. Pharmacol. Exp. Ther.* **2006**, 316, 1369–1377.

391 a) Rühl, R.; Sczech, R.; Landes, N.; Pfluger, P.; Kluth, D.; Schweigert, F. J. *Eur. J. Nutr.* **2004**, *43*, 336–343; b) Rühl, R. *Biochim. Biophys. Acta* **2005**, *1740*, 162–169.

392 Zhou, C.; Tabb, M. M.; Sadatrafiei, A.; Grun, F.; Blumberg, B. *Drug Metab. Dispos.* **2004**, *32*, 1075–1082.

393 Dussault, I.; Lin, M.; Hollister, K.; Wang, E. H.; Synold, T. W.; Forman, B. M. *J. Biol. Chem.* **2001**, *276*, 33309–33312.

394 Drocourt, L.; Pascussi, J. M.; Assenat, E.; Fabre, J. M.; Maurel, P.; Vilarem, M. J. *Drug Metab. Dispo.s* **2001**, *29*, 1325–1331.

395 Moore, J. T.; Kliewer, S. A. *Toxicology* **2000**, *153*, 1–10.

396 a) Masuyama, H.; Hiramatsu, Y.; Kunitomi, M.; Kudo, T.; MacDonald, P. N. *Mol. Endocrinol.* **2000**, *14*, 421–428; b) Masuyama, H.; Inoshita, H.; Hiramatsu, Y.; Kudo, T. *Endocrinology* **2002**, *143*, 55–61; c)

Takeshita, A.; Koibuchi, N.; Oka, J.; Taguchi, M.; Shishiba, Y.; Ozawa Y *Eur. J. Endocrinol.* **2001**, *145*, 513–517.

397 a) Coumol, X.; Diry, M.; Barouki, R. *Biochem. Pharmacol.* **2002**, *64*, 1513–1519; b) Lemaire, G.; de Sousa, G.; Rahmani, R. *Biochem. Pharmaco.l* **2004**, *68*, 2347–2358.

398 Leesnitzer, L. M.; Parks, D. J.; Bledsoe, R. K.; Cobb, J. E.; Collins, J. L.; Consler, T. G.; Davis, R. G.; Hull-Ryde, E. A.; Lenhard, J. M.; Patel, L.; Plunket, K. D.; Shenk, J. L.; Stimmel, J. B.; Therapontos, C.; Willson, T. M., Blanchard, S. G. *Biochemistry* **2002**, *41*, 6640–6650.

399 Lehmann, J. M.; McKee, D. D.; Watson, M. A.; Wilson, T. M.; Moore, J. T.; Kliewer, S. A. *J. Clin. Inves.t* **1998**, *102*, 1016–1023.

400 Waxman, D. J. *Arch. Biochem. Biophys.* **1999**, *369*, 11–23.

401 Ekins, S.; Erickson, J. A. *Drug Metab. Dispos.* **2002**, *30*, 96–99.

402 Luo, G.; Cunningham, M.; Kim, S.; Burn, T.; Lin, J.; Sinz, M.; Hamilton, G. A.; Rizzo, C.; Jolley, S.; Gilbert, D.; et al. *Drug. Metab. Dispos.* **2002**, *30*, 795–804.

403 Schuster, D.; Laggner, C.; Steindl, T. M.; Palusczak, A.; Hartmann, R. W.; Langer, T. *J. Chem Inf. Model* **2006**, *46*, 1301–1311.

404 Ung, C. Y.; Li, H.; Yap, C. W.; Chen, Y. Z. *Mol. Pharmacol.* **2007**, *71*, 158–168.

405 Xue, Y.; Moore, L. B.; Orans, J.; Peng, L.; Bencharit, S.; Kliewer, S. A.; Redinbo, M. R. *Mol. Endocrinol.* **2007** *21*, 1028–1038.

406 a) Berkhout T. A.; Simon, H. M.; Patel, D. D.; Bentzen, C.; Niesor, E.; Jackson, B.; Suckling, K. E. *J. Biol. Chem.* **1996**, *271*, 14376-14382; b) Berkhout T. A.; Simon, H. M.; Jackson, B.; Yates, J.; Pearce, N.;

Groot, P. H.; Bentzen, C.; Niesor, E.; Kerns, W. D., Suckling, K. E. *Atherosclerosis* **1997**, *133*, 203-212.

407 Kliewer, S. A.; Goodwin, B.; Willson, T. M. *Endocr. Rev.* **2002**, *23***,** 687-702.

408 Mencarelli, A.; Renga, B.; Palladino, G.; D'Amore, C.; Ricci, P.; Distrutti, E.; Barbanti, M.; Baldelli, F.; Fiorucci, S. *Eur. J. Pharmacol.* **2011**, *668*, 317-324.

409 Fiorucci, S.; Cipriani, S.; Mencarelli, A.; Renga, B.; Distrutti, E.; Baldelli, F. *Curr. Mol. Med.* **2010**, *10*, 579–595.

410 Dubrac, S.; Elentner, A.; Ebner, S.; Horejs-Hoeck, J.; Schmuth, M. *J. Immunol.* **2010**, *184*, 2949-2957.

411 Ioannides, C. *Cytochromes P450: Metabolic and Toxicological Aspects*, CRC Press, **1996**.

412 a) Jones, S. A.; Moore. L. B.; Shenk, J. L.; Wisely, G. B.; Hamilton, G. A.; McKee, D. D.; Tomkinson, N. C.; LeCluyse, E. L.; Lambert, M. H.; Willson, T. M.; Kliewer, S. A.; Moore, J. T. *Mol. Endocrinol.* **2000**, *14*, 27-39.

413 Sinz, M.; Kim, S.; Zhu, Z.; Chen, T.; Anthony, M.; Dickinson, K.; Rodrigues, A. D. *Curr Drug Metab* **2006**, *7*, 375–388.

414 Piscitelli, S. C.; Burstein, A. H.; Chaitt, D.; Alfaro, R. M.; Falloon, J. *Lancet* **2000**, *355*, 547-548.

415 Ruschitzka, F.; Meier, P. J.; Turina, M.; Luscher, T. F.; Noll, G. *Lancet* **2000**, *355*, 548-549.

416 Wentworth, J. M.; Agostini, M.; Love, J.; Schwabe, J. W.; Chatterjee, V. K. *J. Endocrinol.* **2000**, *166*, R11-16.

417 Hofmann, A. F. *Arch. Intern. Med.* **1999**, *159*, 2647–2658.

418 Staudinger, J. L.; Goodwin, B.; Jones, S A.; Hawkins-Brown, D.; MacKenzie, K. I.; LaTour, A.; Liu, Y.; Klaassen, C. D.; Brown, K. K.; Reinhard, J.; Willson, T. M.; Koller, B. H.; Kliewer, S. A. *Proc. Natl Acad. Sci. USA* **2001**, *98*, 3369–3374.

419 Staudinger, J.; Liu, Y.; Madan, A.; Habeebu, S.; Klaassen, C. D. *Drug Metab. Dispos.* **2001**, *29*, 1467–1472.

420 Moore, L. B.; Goodwin, B.; Jones, S. A.; Wisely, G. B.; Serabjit-Singh, C. J.; Willson, T. M.; Collins, J. L.; Kliewer, S. A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 7500-7502.

421 Wentworth, J. M.; Agostini, M.; Love, J.; Schwabe, J. W.; Chatterjee, V. K. K. *J. Endocrinol.* **2000**, *166*, R11–R16.

422 Chevallier, A. *The Encyclopedia of Medicinal Plants* (Dorling Kindersley, New York, 1996).

423 a) Fiorucci, S.; Mencarelli, A.; Distrutti, E.; Palladino, G.; Cipriani, S. *Curr. Med. Chem.* **2010**, *17*, 139-159; b) Fiorucci, S.; Cipriani, S.; Baldelli, F.; Mencarelli, A. *Prog. Lipid Res.* **2010**, *49*, 171-185; c) Fiorucci, S.; Mencarelli, A.; Palladino, G.; Cipriani, S. *Trend Pharmacol. Sci.* **2009**, *30*, 570-580; d) Fiorucci., S.; Baldelli, F. *Curr. Opin. Gastroenterol.* **2009**, *25*, 252-259; e) Fiorucci, S.; Rizzo, G.; Donini, A.; Distrutti, E.; Santucci, L. *Trends Mol. Med.* **2007**, *13*, 298-309.

424 a) Eloranta, J. J.; Kullak-Ublick, G. A. *Arch. Biochem. Biophys.* **2005**, *433*, 397–412; b) Pellicciari, R.; Costantino, G.; Fiorucci, S. *J. Med. Chem.* **2005**, *48*, 5383–5403.

425 Laffitte, B. A.; Kast, H. R.; Nguyen, C. M.; Zavacki, A. M.; Moore, D. D.; Edwards, P. A. *J. Biol. Chem.* **2000**, *275*, 10638–10647.

426 Repa, J. J.; Turley, S. D.; Lobaccaro, J. A.; Medina, J.; Li, L.; Lustig, K.; Shan, B.; Heyman, R. A.; Dietschy, J. M.; Mangelsdorf, D. J. *Science* **2000**, *289*, 1524–1529.

427 Yang, Y.; Zhang, M.; Eggertsen, G.; Chiang, J. Y. *Biochim. Biophys. Acta* **2002**, *1583*, 63–73.

428 Ananthanarayanan, M.; Balasubramanian, N.; Makishima, M.; Mangelsdorf, D. J.; Suchy, F. J. *J. Biol. Chem.* **2001**, *276*, 28857–28865.

429 Denson, L. A.; Sturm, E.; Echevarria, W.; Zimmerman, T. L.; Makishima, M.; Mangelsdorf, D. J.; Karpen, S. J. *Gastroenterology* **2001**, *121*, 140–147.

430 Goodwin, B.; Jones, S. A.; Price, R. R.; Watson, M. A.; McKee, D. D.; Moore, L. B.; Galardi, C.; Wilson, J. G.; Lewis, M. C.; Roth, M. E.; Maloney, P. R.; Willson, T. M.; Kliewer, S. A. *Mol. Cell.* **2000**, *6*, 517–526.

431 Fiorucci, S.; Baldelli, F. *Curr. Opin. Gastroenterol.* **2009**, *25*, 252–259.

432 Jonker, J. W.; Liddle, C.; Downes, M. *J Steroid Biochem Mol Biol* **2011**, doi:10.1016/j.jsbmb.2011.06.012

433 Stedman, C.; Liddle, C.; Coulter, S.; Sonoda, J.; Alvarez, J. G.; Evans R. M.; Downes, M. *Proc Natl Acad Sci USA* **2006**, *103*, 11323-11328.

434 Mennone, A.; Soroka, C. J.; Cai, S. Y.; Harry, K.; Adachi, M.; Hagey, L.; Schuetz, J. D.; Boyer, J. L. *Hepatology* **2006**, *43*, 1013-1021.

435 a) Renga, B.; Migliorati, M.; Mencarelli, A.; Cipriani, S.; D'Amore, C.; *Biochim Biophys Acta* **2011**, *3*, 157-165; b) Stedman, C. A.; Liddle, C.; Coulter, S. A.; Sonoda, J.; Alvarez, J. G.; Moore, D. D.;

Evans, R. M.; Downes, M. *Proc Natl Acad Sci USA* **2005**, *102*, 2063–2068.

436 Wang, H.; Chen, J.; Hollister, K.; Sowers, L. C.; Forman, B. M. *Mol. Cell* **1999**, *3*, 543-553.

437 a) Makishima, M.; Okamoto, A. Y.; Repa, J. J.; Tu, H.; Learned, R. M.; Luk, A.; Hull, M. V.; Lustig, K. D.; Mangelsdorf, D. J.; Shan, B. *Science* **1999**, *284*, 1362-1365.

438 Parks, D. J.; Blanchard, S. G.; Bledsoe, R. K.; Chandra, G.; Consler, T. G.; Kliewer, S. A.; Stimmel, J. B.; Willson, T. M.; Zavacki, A. M.; Moore, D. D.; Lehmann, J. M. *Science* **1999**, *284*, 1365-1368.

439 a) Maloney, P. R.; Parks, D.; Haffner, C. D.; Fivush, A. M.; Chandra, G.; Plunket, K. D.; Creech, K. L.; Moore, L. B.; Wilson, J. G.; Lewis, M. C.; Jones, S. A.; Willson, T. M. *J. Med. Chem.* **2000**, *43*, 2971-2974; b) Nicolaou, K. C.; Evans, R. M.; Roecker, A. J.; Hughes, R.; Downes, M.; Pfefferkorn, J. A. *Org. Biol. Chem.* **2003**, *1*, 908-920; c) Deng, G.; Li, W.; Shen, J.; Jiang, H.; Chen, K.; Liu, H. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5497-5502; d) Flatt, B.; Martin, R.; Wang, T. L.; Mahaney, P.; Murphy, B.; Gu, X. H.; Foster, P.; Li, J.; Pircher, P.; Petrowski, M.; Schulman, I.; Westin, S.; Wrobel, J.; Yan, G.; Bischoff, E.; Daige, C.; Mohan, R. *J. Med. Chem.* **2009**, *19*, 904-907; e) Feng, S.; Yang, M.; Zhang, Z.; Wang, Z.; Hong, D.; Richter, H.; Benson, G. M.; Bleicher, K.; Grether, U.; Martin, R. E.; Plancher, J. M.; Kuhn, B.; Rudolph, M. G.; Chen, L. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 2595-2598; f) Abel, U.; Schlu.ter, T.; Schulz, A.; Hambruch, E.; Steeneck, C.; Hornberger, M.; Hoffmann, T.; Perovic.-Ottstadt, S.; Kinzel, O.;

Burnet, M.; Deuschle, U.; Kremoser, C. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 4911-4917.

440 a) Fujino, T.; Une, M.; Imanaka, T.; Inoue, K.; Nishimaki-Mogami, T. *J. Lipid Res.* **2004**, *45*, 132-138; b) Pellicciari, R.; Costantino, G.; Camaioni, E.; Sadeghpour, B. M.; Entrena, A.; Willson, T. M.; Fiorucci, S.; Clerici, C.; Gioiello, A. *J. Med. Chem.* **2004**, *47*, 4559-4569.

441 a) Fujino, T.; Une, M.; Imanaka, T.; Inoue, K.; Nishimaki-Mogami, T. *J. Lipid Res.* **2004**, *45*, 132–138; b) Pellicciari, R.; Fiorucci, S.; Camaioni, E.; Clerici, C.; Costantino, G.; Maloney, P. R.; Morelli, A.; Parks, D. J.; Willson, T. M. *J. Med. Chem.* **2002**, *45*, 3569–3572; c) Nicolaou, K. C.; Evans, R. M.; Roecker, A. J.; Hughes, R.; Downes, M.; Pfefferkorn, J. A. *Org. Biomol. Chem.* **2003**, *1*, 908–920; d) Dussault, I.; Beard, R.; Lin, M.; Hollister, K.; Chen, J.; Xiao, J. H.; Chandraratna, R.; Forman, B. M. *J. Biol. Chem.* **2003**, *278*, 7027–7033; e) Love, J. D.; Gooch, J. T.; Benko, S.; Li, C.; Nagy, L.; Chatterjee, V. K.; Evans, R. M.; Schwabe, J. W. *J. Biol. Chem.* **2002**, *277*, 11385–11391; f) Maloney, P. R.; Parks, D. J.; Haffner, C. D.; Fivush, A. M.; Chandra, G.; Plunket, K. D.; Creech, K. L.; Moore, L. B.; Wilson, J. G.; Lewis, M. C.; Jones, S. A.; Willson, T. M. *J. Med. Chem.* **2000**, *43*, 2971–2974.

442 Pellicciari, R.; Fiorucci, S.; Camaioni, E.; Clerici, C.; Costantino, G.; Maloney, P. R.; Morelli, A.; Parks, D. J.; Willson, T. M. *J. Med. Chem.* **2002**, *45*, 3569-3572.

443 Fiorucci, S; Cipriani, S; Mencarelli, A; Baldelli, F; Bifulco, G; Zampella, A. *Mini-Rev. Med. Chem.* **2011**, *11*, 753-762.

444 a) Li, J. W. H.; Vederas, J. C. *Science* **2009**, *325*, 161–165; b) Koehn, F. E.; Carter, G. T. Nat Rev Drug Discovery **2005**, *4*, 206-220; c) Carlson, E. E. *ACS Chem. Biol.* **2010**, *5*, 639-653.

445 a) Singh, R. B.; Niaz, M. A.; Ghosh, S. *Cardiovasc. Drugs Ther.* **1994**, *4*, 659-664; b) Sinal, C. J.; Gonzalez, F. J. *Trends Endocrinol. Metab.* **2002**, *13*, 275-276; c) Urizar, N. L.; Liverman, A. B.; Dodds, D. T.; Silva, F. V.; Ordentlich, P.; Yan, Y.; Gonzalez, F. J.; Heyman, R. A.; Mangelsdorf, D. J.; Moore, D. D. *Science* **2002**, *296*, 1703−1706; d) Wu, J.; Xia, C.; Meier, J.; Li, S.; Hu, X.; Lala, D. S. *Mol. Endocrinol.* **2002**, *16*, 1590−1597.

446 Nozawa, H. *Biochem. Biophys. Res. Commun.* **2005**, *336*, 754−761.

447 Nam, S. J.; Ko, H.; Shin, M.; Ham, J.; Chin, J.; Kim, Y.; Kim, H.; Shin, K.; Choi, H.; Kang, H. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 5398−5402.

448 Choi, H.; Hwang, H.; Chin, J.; Kim, E.; Lee, J.; Nam, S. J.; Lee, B. C.; Rho, B. J.; Kang, H. *J. Nat. Prod.* **2011**, *74*, 90−94.

449 Nishimura, S.; Arita, Y.; Honda, M.; Iwamoto, K.; Matsuyama, A.; Shirai, A.; Kawasaki, H.; Kakeya, H.; Kobayashi, T.; Matsunaga, S.; Yoshida, M. *Nat Chem Biol* **2010**, *6*, 519-526.

450 Kobayashi, M.; Tanaka, J.; Katori, T.; Kitagawa, I. *Tetrahedron Lett*. **1989**, *30*, 2963–2966.

451 Zhang, H. J.; Yi, Y. H.; Lin, H. W. *Helv. Chim. Acta* **2010**, *93*, 1120–1126.

452 Kho, E.; Imagawa, D. K.; Rohmer, M.; Kashman, Y.; Djerassi, C. Sterols in marine invertebrates. *J. Org. Chem.* **1981**, *46*, 1836–1839.

453 a) Angawi, R. F.; Calcinai, B.; Cerrano, C.; Dien, H. A.; Fattorusso, E.; Scala, F.; Taglialatela-Scafati, O. *J. Nat. Prod.* **2009**, *72*, 2195–2198.

454 Kobayashi, M.; Kawazoe, K.; Katori, T.; Kitagawa, I. *Chem. Pharm. Bull.* **1992**, *40*, 1773–1778.

455 Qureshi, A.; Faulkner, D. J. *J. Nat. Prod.* **2000**, *63*, 841–842.

456 Sugo, Y.; Inouye, Y.; Nakayama, N. *Steroids* **1995**, *60*, 738–742.

457 Coral Reef Initiative in The South Pacific (CRISP).

458 Festa, C.; De Marino, S.; Sepe, V.; Monti, M. C.; Luciano, P.; D'Auria, M. V.; Debitus, C.; Bucci, M.; Vellecco, V.; Zampella, A. *Tetrahedron* **2009**, *65*, 10424–10429.

459 Sepe, V.; D'Auria, M. V; Bifulco, G.; Ummarino, R.; Zampella, A. *Tetrahedron* **2010**, *66*, 7520–7526.

460 Festa, C.; De Marino, S.; Sepe, V.; D'Auria, M. V.; Bifulco, G.; Debitus, C.; Bucci, M.; Vellecco, V.; Zampella, A. *Org. Lett.* **2011**, *13*, 1532–1535.

461 Isolation of conicasterol C (**201**) in very small amounts hampered its pharmacological evaluation.

462 Mi, L. -Z.; Devarakonda, S.; Harp, J. M.; Han, Q.; Pellicciari, R.; Willson, T. M.; Khorasanizadeh, S.; Rastinejad, F. *Mol. Cell* **2003**, *11*, 1093–1100.

463 Fujino, T.; Une, M.; Imanaka, T.; Inoue, K.; Nishimaki-Mogami, T. *J. Lipids Res.* **2004**, *45*, 132–138.

464 Modica, S.; Gadaleta, R. M; Moschetta, A. *Nucl. Recept. Signaling* **2010**, *8*, 1–28.

465 Downes, M.; Verdecia, M. A.; Roecker, A. J.; Hughes, R.; Hogenesch, J. B.; Kast-Woelbern, H. R.; Bowman, M. E.; Ferrer, J. L.; Anisfeld, A.

M.; Edwards, P. A.; Rosenfeld, J. M.; Alvarez, J. G. A.; Noel, J. P.; Nicolaou, K. C; Evans, R. M. *Mol. Cell* **2003**, *11*, 1079–1092.

466 Wang, H.; Chen, J.; Hollister, K.; Sowers, L. C.; Forman, B. M. *Mol. Cell* **1999**, *3*, 543–553.

467 Kuipers, F.; Claudel, T.; Sturm, E.; Staels, B. *Rev. Endocr. Metab. Disord.* **2004**, *5*, 319–326.

468 Watkins, R. E.; Maglich, J. M.; Moore, L. B.; Wisely, G. B.; Noble, S. M.; Davis-Searles, P. R.; Lambert, M. H.; Kliewer, S. A.; Redinbo., M. R. *Biochemistry* **2003**, *42*, 1430–1438.

469 Ekins, S.; Chang, C.; Mani, S.; Krasowski, M. D.; Reschly, E. J.; Iyer, M.; Kholodovych, V.; Ai, N.; Welsh, W. J.; Sinz, M.; Swaan, P. W.; Patel, R.; Bachmann, K. *Mol. Pharmacol.* **2007**, *72*, 592–603.

470 Watkins, R. E.; Wisely, G. B.; Moore, L. B.; Collins, J. L.; Lambert, M. H.; Williams, S. P.; Willson, T. M.; Kliewer, S. A.; Redinbo., M. R. *Science* **2001**, *292*, 2329–2333.

471 Ekins, S.; Kortagere, S.; Iyer, M.; Reschly, E. J.; Lill, M. A.; Redinbo, M. R.; Krasowski, M. D. *PLoS Comput. Biol.* **2009**, *5*, e1000594, 1-12.

472 Xue, Y.; Moore, L. B.; Orans, J.; Peng, L.; Bencharit, S.; Kliewer, S. A.; Redinbo, M. R. *Mol. Endocrinol*. **2007**, *21*, 1028–1038.

473 Sepe, V.; Ummarino, R.; D'Auria, M. V.; Mencarelli, A.; D'Amore, C.; Renga, B.; Zampella, A.; Fiorucci, S. *J. Med. Chem.* **2011**, *54*, 4590−4599.

474 De Marino, S.; Sepe, V.; D'Auria, M. V.; Bifulco, G.; Renga, B.; Petek, S.; Fiorucci, S.; Zampella, A. *Org. Biomol. Chem.* **2011**, *9*, 4856−4862.

475 Wright, J. L. C.; Mclnnes, A. G.; Shimizu, S.; Smith, D. G.; Walter, J. A.; Idler, D.; Khalil, W. *Can. J. Chem.* **1978**, *56*, 1898−1903.

476 Wegerski, C. J.; Hammond, J.; Tenney, K.; Matainaho, T.; Crews, P. *J. Nat. Prod.* **2007**, *70*, 89-94.

477 a) van Soest, R. W. M. *Neth. J. Sea Res.* 1989; *23*, 223-230; b) Hamada T, Matsunaga S, Yano S, Fusetani N. *J. Am. Chem. Soc.* **2005**, *127*, 110- 118; c) Ratnayake, A. S.; Davis, R. A.; Harper, M. K.; Veltri, C. A.; Andjelic, C. D.; Barrows, L. R.; Ireland, C. M. *J. Nat. Prod.* **2005**, *68*, 104-107.

478 De Marino, S.; Festa, C.; D'Auria, M. V.; Cresteil, T.; Debitus, C.; Zampella, A. *Marine drugs* **2011**, *9*, 1133-1141.

479 Kho, E.; Imagawa, D. K.; Rohmer, M.; Kashman, Y.; Djerassi, C. *J. Org. Chem.* **1981**, *46*, 1836−1839.

480 Sugo, Y.; Inouye, Y.; Nakayama, N. *Steroids* **1995**, *60*, 738-742.

481 Angawi, R. F.; Calcinai, B.; Cerrano, C.; Dien, H. A.; Fattorusso, E.; Scala, F.; Taglialatela-Scafati, O. *J. Nat. Prod.* **2009**, *72*, 2195-2198.

482 Umeyama, A.; Shoji, N.; Enoki, M.; Arihara, S. *J. Nat. Prod.* **1997**, *60*, 296-298.

483 Rubinstein, I.; Goad, L. J.; Clague, A. D. H.; Mulheirn, L. J. *Phytochemistry* **1976**, *15*, 195-200.

484 Wright, J. L. C.; Mclnnes, A. G.; Shimizu, S.; Smith, D. G.; Walter, J. A.; Idler, D.; Khalil, W. *Can. J. Chem.* **1978**, *56*, 1898-18903.

485 Soisson, S. M.; Parthasarathy, G.; Adams, A. D.; Sahoo, S.; Sitlani, A.; Parthasarathy, G.; Adams, A. D.; Sahoo, S.; Sitlani, A.; Sparrow, C.; Cui, J.; Becker, J. W. *Proc. Natl. Acad. Sci. U S A* **2008**, *105*, 5337–5342.

486 Downes, M.; Verdecia, M. A.; Roecker, A. J.; Hughes, R.; Hogenesch, J. B.; Kast-Woelbern, H. R.; Bowman, M. E.; Ferrer, J. L.; Anisfeld, A. M.; Edwards, P. A.; Rosenfeld, J. M.; Alvarez, J. G.; Noel, J. P.; Nicolaou, K. C.; Evans, R. M. Mol Cell **2003**, *11*, 1079-1092.

487 Akwabi-Ameyaw, A.; Bass, J. Y.; Caldwell, R. D.; Caravella, J. A.; Chen, L.; Creech, K. L.; Deaton, D. N.; Jones, S. A.; Kaldor, I.; Liu, Y.; Madauss, K. P.; Marr, H. B.; McFadyen, R. B.; Miller, A. B.; Iii, F. N.; Parks, D. J.; Spearing, P. K.; Todd, D.; Williams, S. P.; Wisely, G. B. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4339-4343.

488 Akwabi-Ameyaw, A.; Caravella, J. A.; Chen, L.; Creech, K. L.; Deaton, D. N.; Madauss, K. P.; Marr, H. B.; Miller, A. B.; Navas, F. III; Parks, D. J.; Spearing, P. K.; Todd, D.; Williams, S. P.; Wisely, G. B. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 6154-6160.

489 Fiorucci, S.; Mencarelli, A.; Palladino, G.; Cipriani, S. *Trends Pharmacol. Sci.* **2009**, *30*, 570-580

490. Fiorucci, S.; Cipriani, S.; Baldelli, F.; Mencarelli, A. *Prog. Lipid Res.* **2010**, *49*, 171–185.

491 Fiorucci, S.; Cipriani, S.; Mencarelli, A.; Renga, B.; Distrutti, E., et al. Curr. Mol. Med. **2010**, *10*, 579–595.

492 Maruyama, T.; Miyamoto, Y.; Nakamura, T.; Tamai, Y.; Okada, H.; Sugiyama, E.; Nakamura, T.; Itadani, H.; Tanaka, K. *Biochem. Bophys. Res. Commun.* **2002**, *298*, 714–719.

493 Kawamata, Y.; Fujii, R.; Hosoya, M.; Harada, M.; Yoshida, H.; Miwa, M.; Fukusumi, S.; Habata, Y.; Itoh, T.; Shintani, Y.; Hinuma, S.; Fujisawa, Y.; Fujino, M. *J. Biol. Chem.* **2003**, *278*, 9435–9440.

494 Watanabe, M.; Houten, S. M.; Mataki, C.; Christoffolete, M. A.; Kim, B. W.; Sato, H.; Messaddeq, N.; Harney, J. W.; Ezaki, O.; Kodama, T.; Schoonjans, K.; Bianco, A. C.; Auwerx, J. *Nature* **2006**, *439*, 484–489.

495 Poole, D. P.; Godfrey, C.; Cattaruzza, F.; Cottrell, G. S.; Kirkland, J. G.; Pelayo, J. C.; Bunnett, N. W.; Corvera, C. U. *Neurogastroenterol. Motil.* **2010**, *22*, 814–825.

496 Katsuma, S.; Hirasawa, A.; Tsujimoto, G. *Biochem. Biophis. Res. Comm.* **2005**, *329*, 386–390.

497 Keitel, V.; Donner, M.; Winandy, S.; Kubitz, R.; Häussinger, D. *Biochem. Biophys. Res. Commun.* **2008**, *372*, 78–84.

498 Hov, J. R.; Keitel, V.; Laerdahl, J. K.; Spomer, L.; Ellinghaus, E.; ElSharawy, A.; Melum, E.; Boberg, K. M.; Manke, T.; Balschun, T.; Schramm, C.; Bergquist, A.; Weismuller, T.; Gotthardt, D.; Rust, C.; Henckaerts, L.; Onnie, C. M.; Weersma, R. K.; Sterneck, M.; Teufel, A.; Runz, H.; Stiehl, A.; Ponsioen, C. Y.; Wijmenga, C.; Vatn, M. H.; Stokkers, P. C. F.; Vermeire, S.; Mathew, C. G.; Lie, B. A.; Beuers, U.; Manns, M. P.; Schreiber, S.; Schrumpf, E.; Haussinger, D.; Franke, A.; Karlsen, T. H. *PLoS One* **2010**, *5*, 1-13.

499 Tiwari, A.; Maiti, P. *Drug Discov. Today.* **2009,** *14*, 523-30

500 Sato, H.; Macchiarulo, A.; Thomas, C.; Gioiello, A.; Une, M.; Hofmann, A. F.; Saladin, R.; Schoonjans, K.; Pellicciari, R.; Auwerx, J. *J. Med. Chem.* **2008**, *51*, 1831-1841.

501 Pellicciari, R.; Sato, H.; Gioiello, A.; Costantino, G.; Macchiarulo, A.; Sadeghpour, B. M.; Giorgi, G.; Schoonjans, K.; Auwerx, J. *J. Med. Chem.* **2007**, *50*, 4265-4268

502 Parker, H. E.; Wallis, K.; le Roux, C. W.; Wong, K. Y.; Reimann, F.; Gribble FM. *Br. J. Pharmacol.* **2012,** *165*, 414-23.

503 Mantovani, A.; Sozzani, S.; Locati, M.; Allavena, P.; Sica, A.; *TRENDS Immunology* **2003**, *23*, 549–555.

504 a) Weiner, S. J.; Kollman, P. A.; Case, D.; Singh, U.C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, 106, 765-784; b) Weiner, S. J.; Kollman, P. A.; Nguyen, N. T.; Case D. A. *J. Comput. Chem.***1987**, *7*, 230-252.

505 Di Micco, S.; Boger, D. L.; Riccio, R.; Bifulco, G. *Eur. J. Org. Chem.* **2008**, *14*, 2454-2462.

506 Della Monica, C.; Randazzo, A.; Bifulco, G.; Cimino, P.; Aquino, M.: Izzo, I.; De Riccardis, F.; Gomez-Paloma, L. *Tetrahedron Lett.* **2002**, *43*, 5707-5710.

507 Chini, M. G.; Riccio, R.; Bifulco, G. *Magn. Reson. in Chem.* **2008**, *46*, 962-968.

508 a) Lockart, T. P.; Comita, P. B.; Bergman, R. G. *J. Am. Chem. Soc.* **1981**, *103*, 4082-4090; b) Jones, R. R.; Bergman, R. G. *J. Am. Chem. Soc.* **1972**, *94*, 660-651; c) Bergman, R. G. *Acc. Chem. Res.* **1973**, *6*, 25-31.

509 Nicolaou, K. C.; Dai, W. M. *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 1387-1416.

510 Leet, J. E.; Schroeder, D. R.; Langley, D. R.; Colson, K. L.; Huang, S.; Klohr, S. E.; Lee, M. S.; Colik, J.; Hofstead, S. J.; Doyle, T. W.; Matson, J. A. *J. Am. Chem. Soc.* **1993**, *115*, 8432-8443.

511 Kawata, S.; Ashizawa, S.; Hirama, M. *J. Am. Chem. Soc.* **1997**, *119*, 12012-12013.

512 Ren, F.; Hogan, P. C.; Anderson, A. J.; Myers, A. G. *J. Am. Chem. Soc.* **2007**, *129*, 5381-5383.

513 Kinnel, R. B.; Gehrken, H. P.; Scheuer, P. J. *J. Am. Chem. Soc.* **1993**, *115*, 3376-3377.

514 Kinnel, R. B.; Gehrken, H. P.; Scheuer, P. J.; Gravalos, D. G.; Faircloth, G. T. *Eur. Pat. Appl.* **1994**, EP626383 A1.

515 Grube, A.; M. Köck, *Angew.Chem. Int. Ed. Engl.* **2007**, *46*, 2320-2324.

516 Kobayashi, H.; Kitamura, K.; Nagai, K.; Nakao, Y.; Fusetani, N.; van Soest, R. W.; Matsunaga, M. S. *Tetrahedron Lett.* **2007**, *48*, 2127-2129.

517 Buchanan, M. S.; Carroll, A. R.; Addepalli, R.; Avery, V. M.; Hooper, J. N. A.; Quinn, R. J. *J. Org. Chem.* **2007**, *72*, 2309-2317.

518 Köck, M.; Grube, A.; I. B. Seiple, P. S. Baran, *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 6586.

519 T. A. Cernak, J. L.Gleason, *J. Org. Chem.* **2008**, *73*, 102-110.

520 B. A. Lanman, L. E. Overman, R. Paulini, N. S. White, *J. Am. Chem. Soc.* **2007**, *129*, 12896-12900.

521 M. S. Buchanan, A. R. Carroll, R. J. Quinn, *Tetrahedron Lett.* **2007**, *48*, 4573-4574.

522 a) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. *J.Am.Chem. Soc.* **1989**, *111*, 8551-8566; b) Lii, J. H.; Allinger, N. L. *J. Comput. Chem.* **1991**, *12*, 186-199.

523 Rosselli, S.; Bruno, M.; Maggio, A.; Bellone, G.; Formisano, C.; Mattia, C. A.; Di Micco, S.; Bifulco, G. *Eur. J. Org. Chem.* **2007**, *15*, 2504-2510.

524 Bassarello, C.; Bifulco, G.; P. Montoro, A. Skhirtladze, M. Benidze, E. Kemertelidze, C. Pizza, S. Piacente, *J. Agric. Food Chem.* **2007**, *55*, 6636.

525 E. Manzo, M. Gavagnin, Bifulco, G.; Cimino, P.; Di Micco, S.; Ciavatta, M. L.; Guo, Y. W.; Cimino, G. *Tetrahedron* **2007**, *63*, 9970-9978.

526 Plaza, A.; Piacente, S.; Perrone, A.; Hamed, A.; Pizza, C.; Bifulco, G. *Tetrahedron* **2004**, *60*, 12201-12209.

527 Cancès, M. T.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032-3037.

528 As it was noted previously in reference 103, this approach also offers an alternative test for the fulfillment of the Initial Rate Approximation by using a single 1D-NOESY mixing time (and thus not needing a full NOE build-up curve for every selective irradiation, which is extremely time-consuming) on a given sample i.e. where the Initial Rate Approximation holds true, all of the relative intensities of peaks within the spectra will be constant as a function of mixing time.

529 Di Micco, S.; Chini, M. G.; Riccio, R.; Bifulco, G. *Eur. J. Org. Chem.* **2010**, *8*, 1411-1434.

530 Tähtinen, P.; Saielli, G.; Guella, G.; Mancini, I.; Bagno, A.; *Chem. Eur. J.* **2008**, *14*, 10445-10452.

531 Smith, S. G.; Goodman, J. M. *J. Am. Chem. Soc.* **2010**, *132*, 12946–12959.

532 Chini, M. G.; Jones, C. R.; Zampella, A; D'Auria, MV; Renga, B; Fiorucci, S; Butts CP; and Bifulco, G.*J. Org. Chem.* **2012** dx.doi.org/10.1021/jo2023763.

533 Kho, E.; Imagawa, D. K.; Rohmer, M.; Kashman, Y.; Djerassi, C. *J. Org. Chem.* **1981**, *46*, 1836–1839.

534 Wright, J. L. C.; Mclnnes, A. G.; Shimizu, S.; Smith, D. G.; Walter, J. A.; Idler, D.; Khalil, W. *Can. J. Chem.* **1978**, *56*, 1898-1903.

535 Horibe, I.; Nakai, H.; Sato, T.; Seo, S.; Takeda, K.; Takatsuto, S. *J. Chem. Soc. Perkin 1 Trans. 1*, **1989**, 1957-1967.

536 Gaussian 09, Revision A.02, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2009.

537 The proton 7α and the H2-12 being in a crowed region of the 1H spectra (See Supporting Information) were ignored because the selective irradiation of the correct proton was impossible to perform.

538 Sonoda, J.; Xie, W.; Rosenfeld, J. M.; Barwick, J. L.; Guzelian, P. S.; Evans, R. M. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 13801-13806.

539 Guo, G. L.**;** Lambert, G.; Negishi, M.; Ward, J. M.; Brewer, H. B. J.; Kliewer, S. A.; Gonzalez, F. J.; Sinal, C. J. *J. Biol. Chem.* **2003**, *278*, 45062-45071.

540 Swope,W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. *J. Chem. Phys.* **1982,** *76*, 637-649.

541 Derome, A. E. Modern NMR Techniques for Chemistry Research; Pergamon press: Oxford, **1987**.

542 a) Hehre, W. J.; Radom, L.; Schleyer, P. V. R.; Pople, J. A. Ab initio Molecular Orbital Theory; Wiley: New York, 1986; b) Nàray-Szabó, G.; Surján, P. R.; Ángyán, S. G. In Applied Quantum Chemistry; Reidel: Boston, 1897; c) van Gunsteren, W. F.; Berendsen, H. J. C. *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 992-1023; d) Höltje, H. D.; Folkers, G. Molecular Modeling Basic Principles and Applications; Eds.; Mannhold, R.; Kubinyi H.; Timmerman, H., VCH: Weinhein New York, Vol. 5; e) Allinger, N. L.; Zhu, Z.-q. S.; Chen, K. J. Am. *Chem. Soc.* **1992**, *114*, 6120-6133; f) Allinger, N. L.; Rodriguez, S.; Chen, K. *Theochem* **1992**, *92*, 161-178; g) Weiner, P. K.; Kollman, P. A. *J. Comp. Chem.* **1981**, *2*, 287-303; h) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765-784; i) Kollman, P. A.; Merz, K. M. Acc. *Chem. Res.* **1990**, 23, 246-252.