MINING PUBCHEM AND PROTEIN DATABASES FOR UNKNOWN TARGETS OF POLLUTANTS AND THERAPEUTIC MOLECULES

Ph.D. thesis by

María Paulina Cabarcas Montalvo

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy in the subject of

Biomedical Sciences

Advisor

Prof. Jesús Olivero Verbel. Ph.D.

University of Cartagena

Institute for Immunological Research

Ph.D. Program in Biomedical Sciences

Cartagena, Colombia

2016

Contents

List of tables

List of figures

RESUMEN

Las uniones proteína-ligando corresponden a un proceso dinámico, que permite que los residuos que se encuentran en los bolsillos puedan interactuar con una variedad de ligandos de forma, tamaño y composición diferente a los ligandos naturales. Dado que la teoría convencional "una molécula, un blanco, una enfermedad" es limitado y poco común en la naturaleza, la estrategia de análisis de la interacción entre redes de ligandos y proteínas diana surge como un método sistemático para sondear los mecanismos moleculares en eventos complejos, que son difíciles de predecir a través de los enfoques *in vitro* o *in vivo*.

Los estudios de cribado virtual se han convertido en una parte importante de la investigación en toxicología y farmacología dirigida al descubrimiento de nuevos blancos proteicos y fármacos, respectivamente. Una de las estrategias utilizadas consiste en evaluar las diferentes vías metabólicas involucradas y seleccionar el objetivo biológico potencial. Para esto, existen dos enfoques fundamentales: el primero basado en ligando, cuya finalidad consiste en la identificación de aquellas estructuras que tienen más probabilidades de unirse a una molécula diana; y el segundo basado en la estructura 3D de proteínas, de tal forma que sea posible seleccionar compuestos candidatos que puedan interactuar favorablemente con los residuos que hacen parte del sitio activo de la proteína, o sitios alostéricos.

En el presente documento se exploran ambas aproximaciones: Las dioxinas son un grupo de moléculas altamente tóxicas, que ejercen su toxicidad a través de la activación del receptor de hidrocarburos aromáticos (Ahr). Su agonista más importante corresponde a 2,3,7,8-tetraclorodibenzo-p-dioxina (TCDD). Aunque la mayoría de los efectos relacionados con la exposición a TCDD se han relacionado con la activación de Ahr, el uso de herramientas de cribado virtual permitió la identificación de nuevas dianas farmacológicas. Los resultados mostraron la presencia de sitios con alta afinidad por TCDD en diversas proteínas, tales como metalopeptidasas 8 y 3, oxidoescualeno ciclasa, y mieloperoxidasa. Algunas de estas proteínas son bien conocidas por su papel bioquímico en algunos efectos patológicos de la exposición a dioxinas, incluyendo endometriosis, diabetes, inflamación y daño en el

hígado. Estos resultados sugieren que TCDD también podría interactuar con dianas celulares en vías Ahr-independientes.

En el campo de la farmacología, hasta el momento, no se tiene un tratamiento eficaz contra la enfermedad del dengue. El complejo proteasa NS2B/NS3 del virus dengue (DENV) corresponde a un blanco para el diseño de antivirales específicos debido a su importancia durante la replicación viral y su alto grado de conservación entre los serotipos. La información estructural del complejo proteasa NS2B/NS3 se empleó para encontrar moléculas pequeñas capaces de inhibir la actividad del complejo enzimático. Dicha actividad inhibitoria se evaluó mediante ensayos *in vitro*, utilizando un sustrato fluorescente y el complejo NS2B/NS3 obtenido por técnicas de ADN recombinante. Además, se utilizaron células HepG2 infectadas con virus dengue serotipo 2, para probar la actividad contra la replicación del virus del dengue. Un total de 210,903 moléculas pequeñas almacenadas en PubChem fueron acopladas *in silico* a la estructura NS2B/NS3 (PDB: 2FOM) para encontrar moléculas capaces de inhibir este complejo proteico. Cinco de los mejores 500 compuestos (valores de afinidad entre -11,6 y -13,5 kcal/mol) fueron utilizados para evaluar la actividad inhibidora de proteasa en ensayos con la proteína recombinante y ensayos antivirales. Los productos químicos $C_{35}H_{27}NO_9$ (CID54681617), $C_{30}H_{25}NO_5$ (CID54692801) y $C_{34}H_{23}NO_7S_2$ (CID54715399) se mostraron como inhibidores potentes de NS2B/NS3, con valores de IC_{50} (μM) y porcentajes de reducción de títulos virales de 19.9, 79.9%; 17.5, 69.8%; y el 9.1, el 73.9%, respectivamente. Métodos multivariantes aplicados a los descriptores moleculares mostraron dos compuestos estructuralmente diferentes de otros inhibidores de DENV. Este descubrimiento abre nuevas posibilidades para la obtención de candidatos a fármacos contra el virus del dengue.

ABSTRACT

Protein-ligand binding corresponds to a dynamic process, which allows that residues in the binding pocket can interact with a variety of ligands with shapes, sizes and composition different from those of their natural ligands. Given that the conventional theory "a molecule, a target, a disease" is limited and unusual in nature, the analysis of interactions between networks of ligands and target proteins arises as a systematic method for probing molecular mechanisms in complex events, which are difficult to predict through *in vitro* or *in vivo* approaches.

Virtual screening studies have become an important part of toxicology and pharmacology research aimed at discovering of new protein targets and drugs, respectively. One of the strategies used consists in evaluating the different metabolic pathways involved, selecting the potential biological target. For this, there are two fundamental approaches: the first one is based on ligand, whose purpose is to identify those protein structures that are more likely to bind to a target molecule; and the second based on the 3D structure of proteins, selecting candidate compounds that may interact favorably with residues that are part of the active site of the protein, or allosteric sites.

In this document both approaches are explored: dioxins are a group of highly toxic molecules that exert their toxicity through the activation of the aryl hydrocarbon receptor (AhR). The most important agonist of the AhR corresponds to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). Although most of the effects related to TCDD exposure have been linked to the activation of AhR, the use of bioinformatic approaches allowed the identification of new targets for TCDD. The results showed the presence of binding sites with high affinity for TCDD in diverse proteins, such as metallopeptidases 8 and 3, oxidosqualene cyclase, and myeloperoxidase. Some of these proteins are well known for their biochemical role in some pathological effects of dioxin exposure, including endometriosis, diabetes, inflammation and liver damage. These results suggest that TCDD could also be interacting with cellular targets through AhR-independent pathways.

In the field of pharmacology, so far, there is no effective treatment against dengue disease. The dengue virus (DENV) NS2B/NS3 protease

complex is a target for designing of specific antivirals due to its importance in viral replication and its high degree of conservation. NS2B/NS3 protease complex structural information was employed to find small molecules that are capable of inhibiting the activity of the enzyme complex. This inhibitory activity was probed with *in vitro* assays using a fluorescent substrate and the complex NS2B/NS3 obtained by recombinant DNA techniques, for testing the activity against dengue virus replication. In addition, HepG2 cells infected with dengue virus serotype 2 were used to test the activity against dengue virus replication. A total of 210,903 small molecules from PubChem were docked *in silico* to the NS2B/NS3 structure (PDB: 2FOM) to find molecules that were capable of inhibiting this protein complex. Five of the best 500 leading compounds, according to their affinity values (-11.6 and -13.5 kcal/mol), were used to test the inhibitory protease activity on the recombinant protein and antiviral assays. Chemicals $C_{35}H_{27}NO_9$ (CID54681617), $C_{30}H_{25}NO_5$ (CID54692801) y $C_{34}H_{23}NO_7S_2$ (CID54715399) were strong inhibitors of NS2B/NS3, with IC_{50} values (μM) and percentages of viral titer reductions of 19.9, 79.9%; 17.5, 69.8%; and 9.1, 73.9 %, respectively. Multivariate methods applied to the molecular descriptors showed two compounds structurally different from other DENV inhibitors. This discovery opens new possibilities for obtaining drug candidates against dengue virus.

CHAPTER 1.
Thesis Overview

Virtual screening methods are widely used in the field of toxicology and pharmacology to elucidate new mechanisms of action of environmental pollutants and molecules with antiviral activity. The main purpose of this project was the scanning of alternative pathways of TCDD toxicity and the identification of dengue virus antiproteases through data mining. In this chapter, the rationale, hypothesis, pertinence, objectives and structure of this thesis will be discussed. The introduction to protein-ligand binding is presented in Chapter 2. This describes the use of the three-dimensional structure of a protein and functional characteristics derived from it, to predict potential interactions between biological molecules, as well as a general description of some of the computational methods used for the study of protein-ligand couplings. In Chapter 3 is showed the use of virtual screening to detect new posible targets for TCDD, discussing the possible responses induced by aryl hydrocarbon receptor-independent pathways. Chapter 4 presents the search for molecules that interfere with the proteolytic activity of the DENV NS2B/NS3 protease by using bioinformatics tools and verifying the results obtained from these methods through *in vitro* assays. Finally, the main contributions of this thesis project are presented in Chapter 6, entitled as conclusions and final remarks. Published articles are presented in the annexes.

1.1. Rationale

For several decades, there has been an interest to increase the knowledge on the dynamics and physiological activity of endogenous and exogenous compounds on cell signaling. A widely used approach in this field consists of simulating interactions between molecules and their targets of interest, both available in data bases, in order to understand their mode of action. The advantage of using this strategy lies on saving time and money by identifying candidates, both, putative ligands and targets, more efficiently compared to high-throughput screening (HTS) procedures. Computational-based coupling methods allow to performe protein-ligand interaction simulations from two initiatives: from the ligand, for which binding potencies are known; and from the 3D structure of the protein of interest, in particular information generated by experimental techniques such as crystallography (X-ray methods) or nuclear magnetic resonance (NMR).

In this document, both strategies were used to address two health problems with major consequences in human populations: the physiological dynamics of dioxins in organisms, and the search for inhibitors of the dengue virus life cycle (Figure 1.1). In the first case, it is well known that most TCDD effects are mediated through the AhR, altering a variety of biochemicals processes such as chloracne (Feeley, 1995), thymus atrophy (Schuurman *et al.*, 1991), liver damage (Czepiel *et al.*, 2010), epidermal changes (Puhvel *et al.*, 1989), immunotoxicity (Chastain and Pazdernik, 1985), birth defects (Mastroiacovo *et al.*, 1988), fertility reduction (Bruner-Tran and Osteen, 2011), endometriosis (Cummings *et al.*, 1999) and cancer (Boffetta *et al.*, 2011). However, there is a growing number of studies suggesting that the toxicity of TCDD is not exclusively dependent on its binding to AhR (EPA, 2010), as in the case of some effects after *in utero* exposure to TCDD in AhR-null mice (Peters *et al.*, 1999); dose-dependent reduction of liver phosphoenolpyruvate carboxykinase activity, probably not mediated by AhR (Stahl *et al.*, 1993); or pericardial edema induction in *Oryzias latip*es fish embryos, without induction of CYP1A, which is AhR-dependent (Dong *et al.*, 2010). These processes lead to multiple secondary effects by this organochlorine chemical, although the triggering targets are unknown. Thus, it is necessary to use structural data from protein databases in order to extract useful information to build mechanistic escenarios for the new actions of TCDD.

On the other hand, despite of multiple efforts to combat dengue disease, which is considered one of the mosquito-borne disease with the most rapid growth in the world, there are still many challenges to control this disease; even when the pharmaceutical company Sanofi Pasteur completed dengue vaccine Phase III evaluation, showing a 60.8% reduction (50.3%, 42.3%, 74.0%, and 77.7% protection against serotypes 1, 2, 3, and 4, respectively) (Wilder-Smith and Yoon, 2016) after three doses for Latin American trial (Pasteur, 2014). Additionally, the mechanism of action of the vaccine is still not well understood, so the risks of vaccination are still uncertain (Vannice *et al.*, 2016). Furthermore, dengue virus develops a process known as antibody-dependent enhancement (ADE), wherein non-neutralising antibodies facilitate virus entry into host cells, produced severe forms of the disease after infection with a different serotype (Flipse *et al.*, 2013), being able to exacerbate the reaction to subsequent infections. Given these disadvantages, the search

for inhibitors against viral proteins emerges as an useful alternative. Inhibitors have been developed against dengue viral proteins of envelope, capsid, NS4B, protease, helicase, methyltransferase, and polymerase (Lim *et al.*, 2013); however, NS3 and NS5 constitute important drug targets for antivirals since they are well-characterized, conserved for all serotypes, and with multifunctional activities in the virus life cycle, such as polyprotein processing and genome replication (Oliveira *et al.*, 2014). Despite the existence of good inhibitors for dengue virus protease with activity at very low concentrations (de Sousa *et al.*, 2015; Wu *et al.*, 2015; Yang *et al.*, 2011), the disease continues to challenge their efficacy or clinical usefulness, making the discovery of new alternatives a necessary achievement.

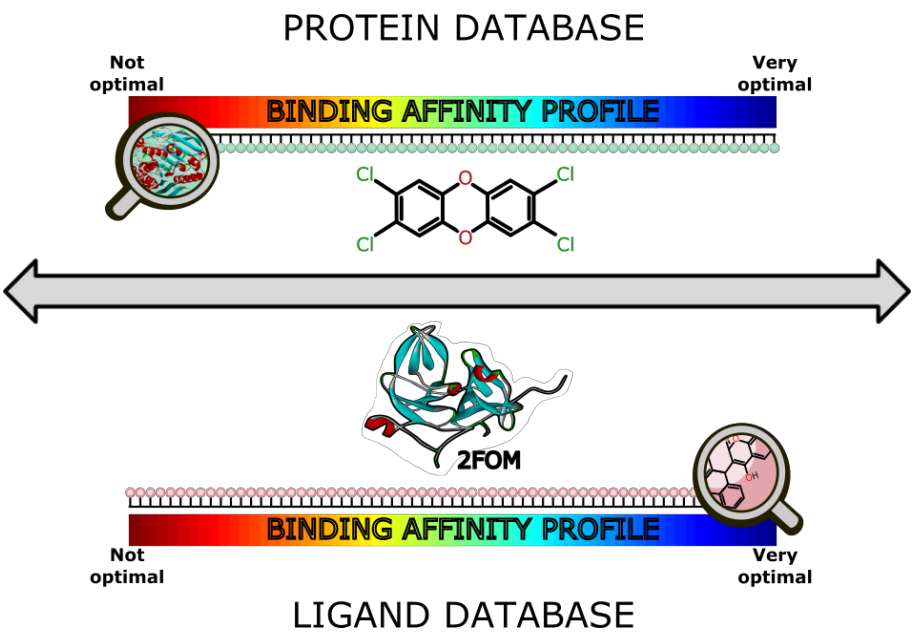


Figure 1.1. Data mining based on ligand and macromolecular structures as a source of targets of environmental pollutants and therapeutic agents in drug design.

Based on these premises, this thesis intends to provide insights about interference of protein function induced by ligand binding in two major public health problems, such as exposure to dioxins and dengue disease. The main contributions obtained were the identification of potential new

protein targets for TCDD in AhR-independent pathways, as well as small molecules with inhibitory capacity for protease activity of dengue virus that could be used in the treatment of dengue. Virtual screening methods allowed predicting the possibilities for different ligand conformations and the binding affinity of the protein and the ligand during coupling. Besides, particularly for the dengue study, *in vitro* assays, including fluorimetric enzyme activity assay and cell-based assays, permitted to validate the protease viral inhibitory effect.

1.2. Hypotheses

- Protein-ligand interactions present in the structures of Protein databases reveal new targets for TCDD, aiming to explain its non-AhR dependent secondary effects.
- PubChem is a source of new scaffolds for potent therapeutic agents acting as dengue virus protease inhibitors.

1.3. Pertinence

The use of traditional methods for the elucidation of the physiological activity of external agents, as well as the process of drug discovery and development, are time and money consuming. Advances developed so far in computational techniques have permitted the implementation of *in silico* methods, which are useful in the recognition of such processes. As a result, in recent years there has been an increasing utilization of these protocols to understand the various phenomena linked to molecular recognition that may occur between macromolecule(s) and ligand candidate(s). Given the advantages provided by these techniques, this thesis describes their application in the investigation of two public health problems: dioxin poisoning and dengue disease.

Dioxins are compounds with potent hormone blocking activity and ubiquitous global distribution. Therefore, we all have a basal level of concentration in our body, called body burden. Due to the aforementioned and its high half-life in the organisms, the interest to identify a possible interaction with other proteins besides the AhR arises, to give an explanation of the extent of toxic effects.

Dengue disease has been identified as a global public health problem of the latter part of the 20th century and the early 21st century. Due to its role in post-transcriptional processing of the viral polyprotein, the role of

NS2B/NS3 protease in dengue virus life cycle and replication mechanism is evident. Besides, it is sufficiently conserved within the four serotypes to allow the design of active compounds against all dengue viral strains and related flavivirus. Therefore, it is expected that inhibition of this enzyme may interfere with viral replication in human host cells.

1.4. Objectives

- To discover new putative targets for TCDD that could explain some AhR-independent effects.
- To find new molecules capable of inhibiting dengue virus proteases.
- To assess the experimental activity of discovered dengue virus protease inhibitors.

1.5. References

- Boffetta, P., K.A. Mundt, H.O. Adami, P. Cole, and J.S. Mandel. 2011. TCDD and cancer: a critical review of epidemiologic studies. Critical reviews in toxicology. 41:622-636.
- Bruner-Tran, K.L., and K.G. Osteen. 2011. Developmental exposure to TCDD reduces fertility and negatively affects pregnancy outcomes across multiple generations. Reprod Toxicol. 31:344-350.
- Chastain, J.E., Jr., and T.L. Pazdernik. 1985. 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD)-induced immunotoxicity. International journal of immunopharmacology. 7:849-856.
- Cummings, A.M., J.M. Hedge, and L.S. Birnbaum. 1999. Effect of prenatal exposure to TCDD on the promotion of endometriotic lesion growth by TCDD in adult female rats and mice. Toxicological sciences : an official journal of the Society of Toxicology. 52:45-49.
- Czepiel, J., G. Biesiada, M. Gajda, W. Szczepanski, K. Szypula, Z. Dabrowski, and T. Mach. 2010. The effect of TCDD dioxin on the rat liver in biochemical and histological assessment. Folia biologica. 58:85-90.
- de Sousa, L.R.F., H. Wu, L. Nebo, J.B. Fernandes, M.F.d.G.F. da Silva, W. Kiefer, M. Kanitz, J. Bodem, W.E. Diederich, T. Schirmeister, and P.C. Vieira. 2015. Flavonoids as noncompetitive inhibitors of Dengue virus NS2B-NS3 protease: Inhibition kinetics and docking studies. *Bioorganic & medicinal chemistry*. 23:466-470.
- Dong, W., F. Matsumura, and S.W. Kullman. 2010. TCDD induced pericardial edema and relative COX-2 expression in medaka (*Oryzias*

Latipes) embryos. Toxicological sciences : an official journal of the Society of Toxicology. 118:213-223.

- EPA, U.S.E.P.A. 2010. Recommended Toxicity Equivalence Factors (TEFs) for Human Health Risk Assessments of 2,3,7,8-Tetrachlorodibenzo-p-dioxin and Dioxin-Like Compounds Washington, DC.
- Flipse, J., J. Wilschut, and J.M. Smit. 2013. Molecular mechanisms involved in antibody-dependent enhancement of dengue virus infection in humans. *Traffic*. 14:25-35.
- Lim, S.P., Q.Y. Wang, C.G. Noble, Y.L. Chen, H. Dong, B. Zou, F. Yokokawa, S. Nilar, P. Smith, D. Beer, J. Lescar, and P.Y. Shi. 2013. Ten years of dengue drug discovery: progress and prospects. *Antiviral research*. 100:500-519.
- Mastroiacovo, P., A. Spagnolo, E. Marni, L. Meazza, R. Bertollini, G. Segni, and C. Borgna-Pignatti. 1988. Birth defects in the Seveso area after TCDD contamination. Jama. 259:1668-1672.
- Oliveira, A.S.D., M.L.D. Silva, A.F. Oliveira, C.C.D. Silva, R.R. Teixeira, and S.O. De Paula. 2014. NS3 and NS5 proteins: important targets for anti-dengue drug design. *Journal of the Brazilian Chemical Society*. 25:1759-1769.
- Pasteur, S. 2014. Sanofi Pasteur's Dengue Vaccine Candidate Successfully Completes Final Landmark Phase III Clinical Efficacy Study in Latin America. Press Release, Lyon, France.
- Puhvel, S.M., M. Sakamoto, and R.M. Reisner. 1989. Effect of TCDD on the density of Langerhans cells in murine skin. Toxicol Appl Pharmacol. 99:72-80.
- Schuurman, H.J., E.J. DE WAAL, H. Van Loveren, and J.G. Vos. 1991. The toxicity of dioxin to the thymus. In The Thymus in Immunotoxicology. M.A.R. M. D. Kendall, editor. Harwood Academic Publishers, ChurSwitzerland.
- Stahl, B.U., D.G. Beer, L.W. Weber, and K. Rozman. 1993. Reduction of hepatic phosphoenolpyruvate carboxykinase (PEPCK) activity by 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) is due to decreased mRNA levels. Toxicology. 79:81-95.
- Vannice, K.S., A. Durbin, and J. Hombach. 2016. Status of vaccine research and development of vaccines for dengue. Vaccine. In Press, Corrected Proof.

- Wilder-Smith, A., and I.K. Yoon. 2016. Edging closer towards the goal of a dengue vaccine. Expert review of vaccines. 15:433-435.
- Wu, H., S. Bock, M. Snitko, T. Berger, T. Weidner, S. Holloway, M. Kanitz, W.E. Diederich, H. Steuber, C. Walter, D. Hofmann, B. Weissbrich, R. Spannaus, E.G. Acosta, R. Bartenschlager, B. Engels, T. Schirmeister, and J. Bodem. 2015. Novel dengue virus NS2B/NS3 protease inhibitors. *Antimicrob Agents Chemother*. 59:1100-1109.
- Yang, C.C., Y.C. Hsieh, S.J. Lee, S.H. Wu, C.L. Liao, C.H. Tsao, Y.S. Chao, J.H. Chern, C.P. Wu, and A. Yueh. 2011. Novel dengue virus-specific NS2B/NS3 protease inhibitor, BP2109, discovered by a high-throughput screening assay. Antimicrob Agents Chemother. 55:229-238.

CHAPTER 2.
Introduction to protein-ligand binding

2.1. Introduction

Proteins are important molecular components of biological systems, responsible for structural support, storage, movement and defense against foreign substances, among other cellular functions (Saada and Pearson, 2011). This process depends on protein residues and the interactions that their side chains (group R) can form. Any molecule that can bind to a protein is known as ligand, which may include other proteins or peptides, nucleic acids, membrane components, small compounds, and solvent molecules. Understanding the three-dimensional structure of a protein and functional characteristics derived from it, is a significant finding for different scientific fields, such as toxicology and pharmacology. In toxicology, it allows to identify the main effect generated by a toxic agent, in order to detect the affected organs and prevent irreversible consequences or debilitating diseases. In pharmacology, especially in drug design, it can strengthen the interactions between drugs and proteins, ensuring its effectiveness. One of the most important characteristics in the process of drug development is the identification of the active site, which can be recognized through the study of features in the structural surface properties (Wasserman *et al.*, 2012).

Although low molecular weight compounds may interact with macromolecules by means of covalent interactions, a broad spectrum of molecules interact through non-covalent protein-ligand interactions to modulate cell behavior (Böhm and Klebe, 1996). In chemistry and biology, molecular recognition refers to the process in which two or more biological molecules form a complex through complementary non-covalent bonding, such as, hydrogen bond, metal coordination, van der Waals forces, and π−π, hydrophobic, or electrostatic interactions (Cleaves, 2011). These kinds of interactions may occur specifically to one or few molecules, or may be promiscuous with many partners; nevertheless, for molecular recognition interactions, the specificity is the first defining characteristic of the process, and it must exist a threshold affinity between molecules involved. In the same way, for this interaction to occur, it is necessary that molecules implicated are near one another and at concentrations that promote interaction, which corresponds to the second prerequisite for molecular recognition (Schreiber and Keating, 2011).

The binding between two molecules comprises changes in the structure and dynamics of the involved parts, which are related to their enthalpic and entropic components. As in any spontaneous processes, the search for a more stable state implies the decrease of the free energy of the system to a minimum, because when a system exists in two alternate states, spontaneity will occur moving toward lower energy. Changes in structure can be subtle or on a large scale, including alterations in folding, conformation, size, shape, spatial distribution, charge distribution and hydration and protonation, likewise, changes in the accessibility of the surface, polarity, and intra- and intermolecular entropy factors (Chowdhry and Harding, 2001). As a result, the free binding energy (G) corresponds to the difference between protein (P) and ligand (L) energy when they are in complex and separated (Kessel, 2010), in other words,

$$\Delta G_{bind} = G_{PL} - (G_P + G_L)$$

Therefore, since the final state of the system has a lower free energy (G_{PL}) compared to the initial state, when the ligand and protein are separated, the change of standard free energy accompanying the process (ΔG) will be negative.

ΔG could also be described as the sum between the enthalpy and entropy of a system (Bronowska, 2011). The following equation is used to calculate it:

$$\Delta G = \Delta H - T\Delta S$$

ΔG is the free energy (kcal/mol), ΔH is the enthalpy of the system (kcal/mol), indicating changes in the free energy from non-covalent bonding, ΔS its entropy (cal/mol*K) and T (K), temperature.

In this case, the enthalpy reflects the internal energy change of the system after the formation of the protein-ligand complex, that is, those energy changes produced by forming or breaking of covalent bonds, Van der Waals interactions, atom movement induced by heat, as well as the loss of non-covalent bonds between protein-solvent and ligand-solvent, called desolvation. The enthalpy changes are due to modifications in the type of non-covalent interactions (Kessel, 2010). Entropy, meanwhile, allows measuring the dynamics of the overall system; so it can be

understood as a measure of disorder or freedom of movement for a given atom. Variations in system entropy reflect the lost motion caused by changes in degrees of freedom of translation and rotation (Du *et al.*, 2016). So, ΔG determines the stability and binding affinity present in a protein-ligand complex. Consequently, in this process is observed a favorable enthalpy change and an unfavorable entropy decrease (Kessel, 2010).

It is important to note that both the ligand and the binding sites of the protein are completely solvated prior to complex formation; their polar groups form hydrogen bonds with the solvent. Therefore, it is clear that the process of ligand binding leads to an important reorganization of solvent on both surfaces, which contributes significantly to the free energy of binding. This results in the loss of part of the solvent layer surrounding both surfaces, involving numerous hydrogen bond breaks formed with water molecules of the first solvating layer (Ramos *et al.*, 2001).

Of these water molecules, some are mobile and easily removed, and others are strongly bound to the protein or ligand. These solvent molecules can play an important role in its biological activity, so it can be considered part of its structure. Its mode of action may be direct, participating in specific interactions, or indirect, influencing the dynamic behavior of the environment. However, the effect of their presence could not be easily predicted, because, in the case of proteins, depends on the characteristics of the binding site. Under those circumstances, among other reasons, it is possible to obtain substantial errors in ΔG values predictions; likewise, in experimental tests with inhibitors (Bronowska, 2011).

Equally important is the flexibility of the molecules involved in the binding process, since the ligands often have several rotational degrees of freedom in them, resulting in a large number of low energy conformations in their free state, which are lost when joining with the acceptor given the steric constraints of the binding site. In addition, this flexibility allows the ligand to fit in different protein targets, and thus, be promiscuous (Haupt *et al.*, 2013). On the other hand, proteins are also flexible and their free state conformation may differ significantly from the complexed state, or joined to the ligand, with a consequent loss in some

degrees of internal flexibility. Thus, it allows residues to fit the shape of the ligand, optimizing geometric complementarities. As expected, both effects consequently bring an entropic cost associated with joining processes (Böhm and Klebe, 1996; Ramos *et al.*, 2001).

2.2. Binding models between proteins and ligands

Molecular diffusion allows collisions between proteins and ligands, which promote their union. This process is driven by entropy, and can result from kinetic energy of the solute molecules or thermal energy. Before the collision, the water molecules are surrounding the surface of the solute, allowing a negative enthalpy change due to hydrogen bonds and van der Waals interactions between protein and solvent, as well as a decrease in entropy for the loss of degrees of freedom of the water molecules. After the collision, non-covalent bonds are lost, interrupting the water molecules network on the surface (Du *et al.*, 2016).

Protein binding sites correspond to pockets, grooves or cavities, generally hydrophobic, where residues can form chemical bonds with ligands (Henrich *et al*., 2010). Generally, not all residues on the surface are involved in these interactions, but are confined to defined areas, as the active sites, formed by a few amino acids (Nisius *et al*., 2012). Three models have been described to explain such binding mechanisms: lock and key, induced fit and conformational selection model. The first model was proposed in the 19th century by Emil Fischer. This theory states that both molecules are rigids and their surfaces have a geometric perfect matching. However, evidence suggests that molecules containing no complementary surfaces can be coupled as a result of conformational changes occurring in the binding site of the protein. The second one, induced fit model, was proposed by Daniel Koshland in 1959, and assumed that the binding site is flexible in the protein during binding to ligand. In this model only the ligand binding pocket is considered flexible, and the rest of the protein is assumed rigid. In the conformational model selection, the native state of a protein is assumed as a set of substates that coexist in equilibrium even in the absence of ligand (Bronowska, 2011). Given that the three models have been experimentally observed, the three mechanisms can be found simultaneously or sequentially (Du *et al.*, 2016).

In the lock and key model, after the initial collision, the complete displacement of water molecules surrounding the surfaces is induced, producing an increase in entropy. Because of this, it is said that the lock and key model is a process dominated by entropy. Since there is no perfect complementarity in the induced fit model, it is necessary that multiple collisions take place so that proper coupling is given. For new interactions to occur, it is necessary that the initial interactions present in the binding site are interrupted. Breaks of the original interactions generate positive change in enthalpy and entropy negative variation. Due to this, it is said that the induced fit model corresponds to a process dominated by enthalpy. Finally, in the conformational selection model, it is difficult to differentiate whether enthalpy or entropy contribute more to reducing the free energy of the system. The population distribution of the different protein conformational substates allow the ligand to bind with that substate whose shape facilitates better mating between them. Thus, it can not induce conformational change, and similar to lock-key model, dominated by entropy of the solvent. However, the conformational flexibility of the protein allows conformational changes to facilitate strong non-covalent intermolecular interactions, similar to that described in the induced fit model and as such, this process is dominated by a decrease in enthalpy (Du *et al.*, 2016). A general description of these protein-ligand binding models is presented in Figure 2.1.

Nevertheless, complementarity between ligands and protein binding sites is due not only to the geometry of both, but also for their electrostatic characteristics (Henrich *et al.*, 2010). That is, there is not any specific protein structural motif responsible for binding to the ligands (Kessel, 2010).

2.3. Protein-ligand interactions

The protein binding sites are modules that interact with other macromolecules, and small ligands. These interactions are responsible of the protein complex formation and regulation of biological pathways. Usually, the nature of the binding sites is partly hydrophobic, with a larger number of these kinds of residues exposed on the surface and large but variable extensions of apolar surface area (Halperin *et al.*, 2003). Likewise, although the amino acid composition in the binding sites differs between proteins, amino acids such as arginine, histidine, tryptophan and

tyrosine are most common in binding pockets than in other regions of them (Cohen, 1996).

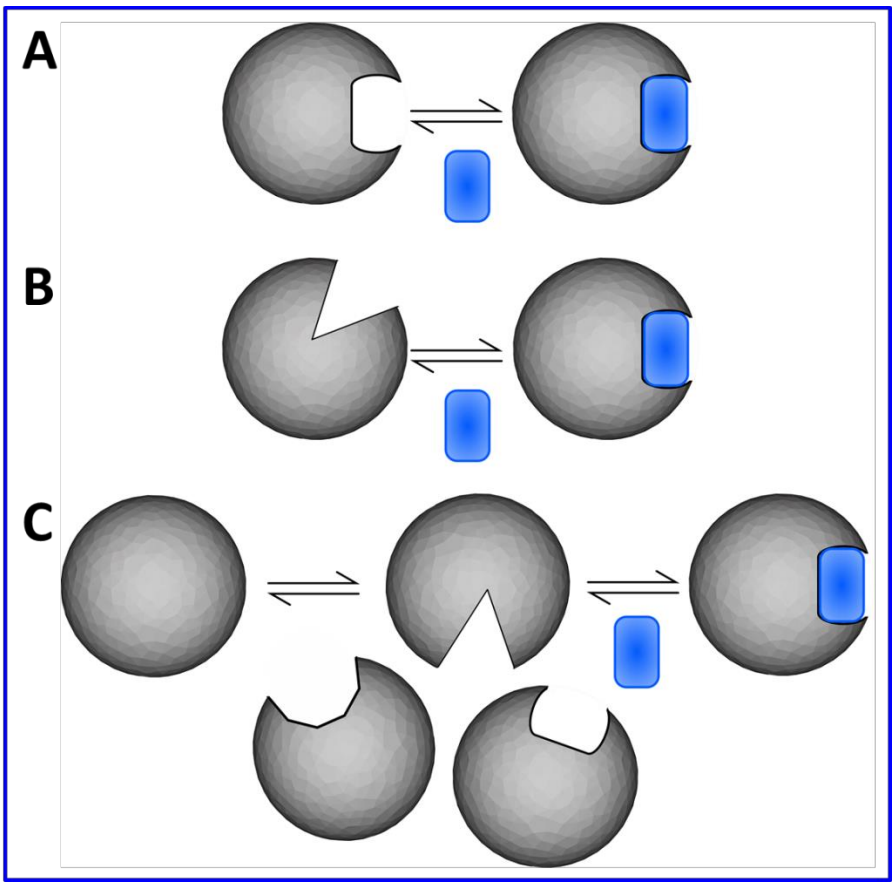


Figure 2.1. Protein-ligand binding models schematization. A Lock-and-key. B. Induced fit. C. Conformational selection. Modified from (Du *et al.*, 2016).

The structural stability of binding sites varies from very low to very high, showing a dual character. This difference is due to the type of interactions that can form residues found in them. In stable regions, their residues allow certain types of interactions corresponding to well-defined stereochemical arrangements, while flexible regions permit binding of different ligands, because they admit that proteins change

their shape for a better coupling with ligands of different chemical structures. Therefore, for binding to occur, it is necessary to ensure shape complementarity as well as the physicochemical characteristics (Nisius *et al.*, 2012). Figure 2.2, displays the residues in glycerol binding pockets of the crystal structure of NS2B/NS3 (PDB ID: 2FOM) protein complex.

The size and shape of the binding sites vary greatly between proteins, enabling them to interact with molecules of different structures. Some forms found are deep and large clefts, as in enzymes, or flat and unstructured as in the regions of protein-protein interactions Furthermore, they may have more than one active site, large proteins specially (Halperin *et al.*, 2003). For example, the catalytic domain of neutrophil collagenase (MMP-8) is a hydrophobic pocket, of intermediate size (Gupta, 2012), consisting of two regions: a groove for the catalytic ion zinc and subsites S3-S1 (158-163), and S1'-S3' (214-220) (Figure 2.3). Strong interactions involve chelation of Zn, hydrogen bonding and van der Waals interactions (Kontogiorgis *et al.*, 2005). For all dengue virus serotypes, the protease has a flat pocket, although with sufficient plasticity in their sub-pockets (S1 - S4). Some possible interactions are hydrogen bonds with Gly133 and Gly153 residues, and in less conserved positions, Ser135, Gly151 and Thr134; furthermore, the presence of aromatic residues such as His51 and Tyr161 allow the formation of π interactions (de Almeida *et al.*, 2013).

As previously stated, the main interactions between proteins and ligands correspond to non-covalent interactions such as electrostatic interactions and van der Waals forces. Although Van der Waals force is very weak in relation to other forces, the large number of such interactions occurring in large protein molecules makes them an important part for the protein-ligand bindings (Pace *et al*., 1996). From electrostatic interactions, the major protein-ligand contributions are usually given by numerous hydrogen bonds (Figure 2.4), which confer stability and selectivity between protein and ligand (Williams and Daviter, 2013). In addition to hydrogen bonds, oppositely charged functional groups are paired in the protein-ligand complex forming salt bridges (Figure 2.5.A).

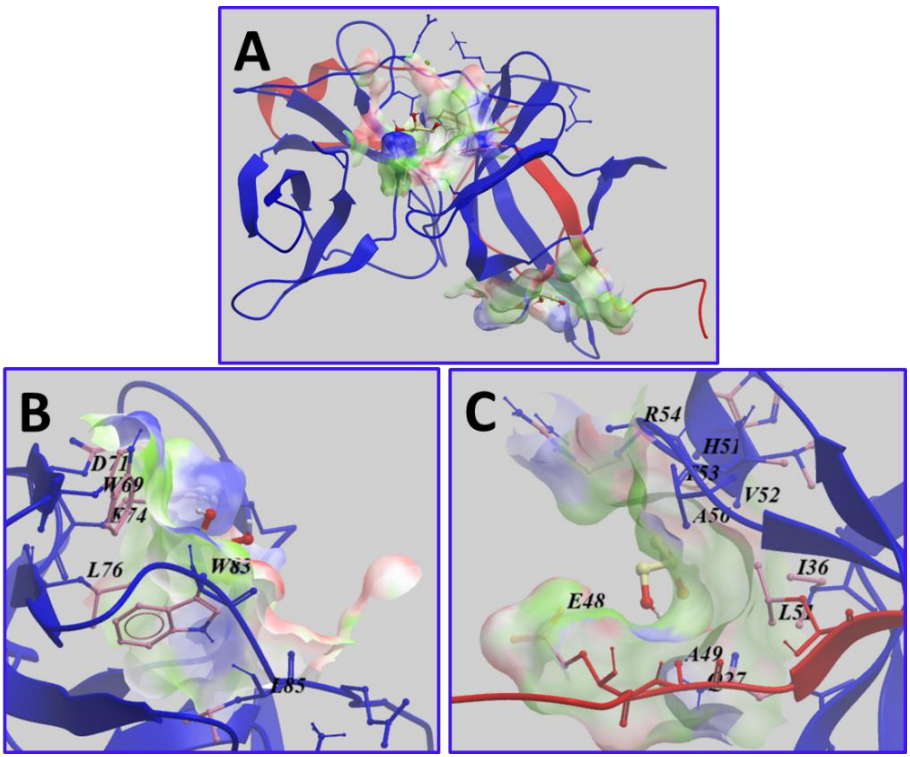


Figure 2.2. Tridimensional view of glycerol binding pockets in the crystal structure of NS2B/NS3 (PDB ID: 2FOM) protein complex. A. Full view of dengue protease complex. B-C. Enlarged view of pocket 1 and 2, respectively. The segment of NS2B is represented in red ribbon, NS3 segment in blue ribbon, and glycerol binding pockets in smooth transparent mesh, colored by binding property: white, neutral surface; green, hydrophobic surface; red, hydrogen bonding acceptor potential; blue, hydrogen bond donor potential. Labeled residues in black corresponds to interacting ones.

Lipophilic groups of the ligand could be located in the hydrophobic pockets of the protein formed by nonpolar side chains of amino acids. The presence of these residues, such as aliphatic and aromatic amino acids, exclude water molecules and other polar groups, attracting nonpolar ligand groups (Figure 2.5.B). Most of the time, aromatic residues in binding sites form aryl-aryl interactions (Figure 2.5.C).

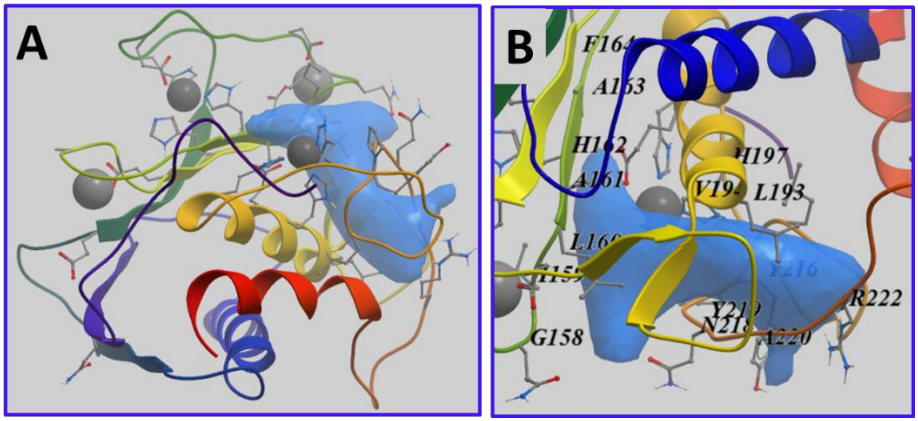


Figure 2.3. Tridimensional view of catalytic domain of neutrophil collagenase (MMP-8), showing the location of residues forming subsites S3-S1, and S1'-S3'.

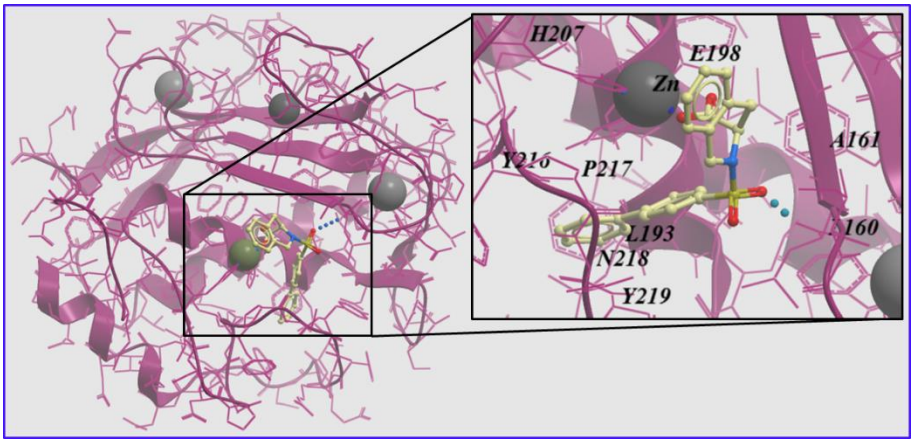


Figure 2.4. Neutrophil Collagenase MMP-8 (PDB: 1I76), showing hydrogen bonds with the crystal structure ligand (2-(biphenyl-4-sulfonyl)-1,2,3,4-tetrahydro-isoquinoline-3-carboxylic acid). Residues surrounding the ligand are labeled in black.

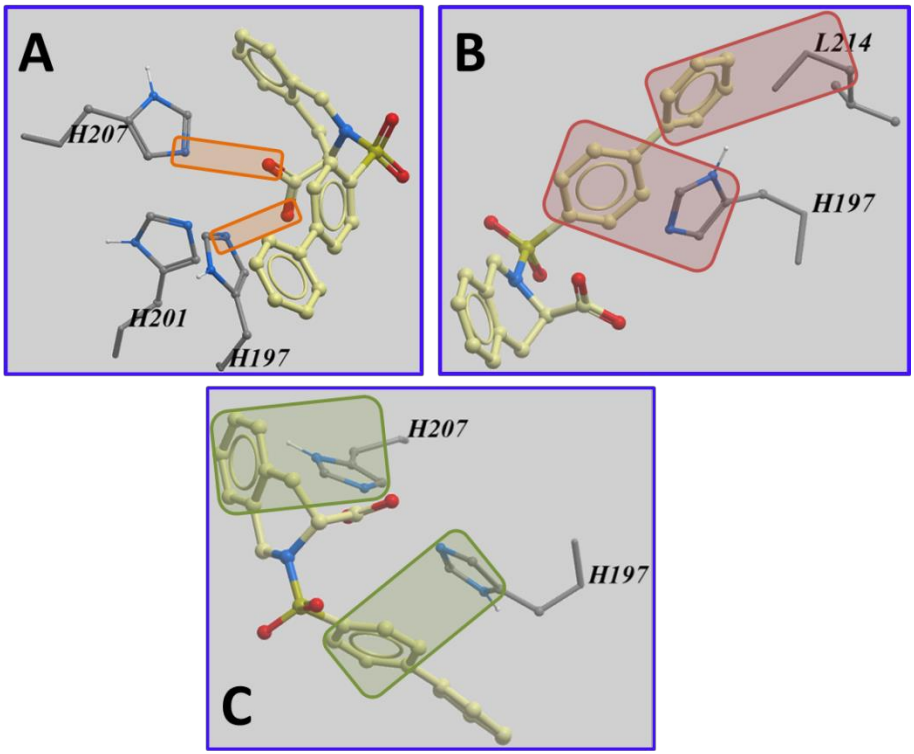


Figure 2.5. Types of protein-ligand interactions between Neutrophil Collagenase MMP-8 (PDB: 1I76), and 2-(biphenyl-4-sulfonyl)-1,2,3,4-tetrahydro-isoquinoline-3-carboxylic acid. A. Salt bridges. B. Hydrophobic interactions. C. Aryl-aryl interactions.

2.4. Computational methods used for the study of protein-ligand couplings

The functional characterization of proteins is a challenge for biochemical, medical and computer science. Computer simulation techniques to model biomolecular bindings are a promising tool to extend our understanding of biological systems. Historically, the search for compounds with biological activity in design drugs, have been performed using high throughput screening made (HTS), however, despite advances in its automation techniques, this approach is still very laborious, time-consuming and expensive (Cavasotto and Phatak, 2009). Therefore, virtual screenings of databases are an increasingly important factor in the search for new active compounds (Lengauer *et al.*, 2004). The knowledge

of protein structure gives an idea of the function of the protein (Lengauer *et al.*, 2004), as well as their contributions to changes in enthalpy, entropy and free energy, being often the starting point for drug design. Thus, the function of a protein can be predicted by computational methods, which can extract even more information than the experimental methods (Pierri *et al.*, 2010). In this way, the process drastically reduces the number of compounds to be synthesized, which is necessary for the optimization of a potential drug (Breitenlechner *et al.*, 2005).

Since protein-ligand docking approaches are one of the most extensively used methods to predict biomolecule interactions, such as macromolecules with small molecules (Forli *et al.*, 2016), it has become an important tool in research. Consequently, the number of softwares available is increasing (Sousa *et al.*, 2013). Such approaches allow accurately predict the binding affinity between a protein and its ligand, identify the most favorable binding pose between them, and therefore the type of interactions involved (Wang and Lin, 2013).

Protein-ligand docking studies can be divided into two strategies: The first one is the technique based on the ligand. It consists on identifying protein similarities to the properties of the ligand of interest. For this, a pharmacophore is used, which is built from functional groups that are necessary to participate in bindings with macromolecules (Hawkins, 2006). A pharmacophore is defined as the abstract representation of the steric and electronic features in the ligand structure, that are necessary to allow binding with proteins, either activating them or inhibiting them (Koes and Camacho, 2012). Among the properties that are considered in the construction of a pharmacophore are hydrogen bond acceptors and donors groups, ionizable groups, hydrophobic regions and aromatic rings, among others (Wieder *et al.*, 2016).The second approach is based on the structure of a target protein, whose normal function is to be altered, blocked in many cases, by binding the pharmacologically active molecule to it. Once the protein is identified, the next step is the selection of a suitable drug to adhere to the protein, it must be easy to synthesize, bio-accessible and have no adverse or toxic effects (Kang *et al.*, 2008). However, it is not necessary to have information on the protein. Indirect information on the protein could be can be taken from reference proteins with known activity to create the

structural model of the protein of interest using a criterion of molecular similarity activity (Lengauer *et al.*, 2004).

Methods of molecular dynamics simulation provide information related to changes in the conformations of proteins and other macromolecules of biological interest, in complex or separated. It is extremely useful to know the change of conformational entropy upon binding; as well as kinetics and thermodynamics information of processes. There are two main advantages of protein-ligand docking simulation: to rapidly predict the best binding pose of the complex, and to score the binding affinity (Grant *et al.*, 2010).

In general, docking methods help to elucidate each one of the ligand orientations inside of a protein pocket, estimating the corresponding binding affinity, classify it according to its value, and identify the more favorable binding mode. Theoretically, the search for the best pose involves analysis of all possible configurations of the protein and ligand separately, and then, all possible conformations of the ligand within each of the possible states of the protein. However, these analyses are very inefficient because of the thoroughness required to consider each of the possibilities. On the other hand, they do not take into account a large number of these degrees of freedom, producing unreliable results. Because of this, the search for the best conformation methods have evolved between rigid, flexible-ligand, and flexible ligand-flexible protein (Du *et al.*, 2016).

The critical essential component of protein-ligand docking protocols is the free binding energy calculation, a mathematical approach that measures the strength of non-covalent interactions between protein and ligand. Although for the accuracy, many physical interactions should be considered, involving each one of them would increase the computational complexity, such that, to facilitate the calculation, certain simplifications are made, which reduces the accuracy of the results (Du *et al.*, 2016).

2.4.1. Identification of protein binding sites

Predicting active sites in proteins is important for understanding the biology of their functions (Kalidas and Chandra, 2008). For the detection of these regions, different methods such as geometric, energetic, and structure-sequence comparison are used. The geometric identification

consists of measuring the volume of the cavity by placing different figures. For this, all atoms must be localized and classified as part of the protein or the solvent. The distances allowed between atom and atom depend on the method, grids, spheres or triangles can be placed between each pair of atoms, with an allowed distance of 3 Å; 1 and 4 Å (radii); or 8 Å, respectively (Henrich *et al.*, 2010). Generally, the region that brings together the largest cluster of geometric figures corresponds to the binding site (Brooijmans, 2009). For an energetic approach, no binding pockets are identified as such, but regions in which the protein is highly likely to interact with different ligands (Henrich *et al.*, 2010). Figure 2.6, shows the predicted binding pockets for NS2B/NS3 (PDB ID: 2FOM) protein complex, the identification was determined using a grid potential map of van der Waals interaction of the receptor and construction of equipotential surfaces along the maps (MolSoft, 2000). This method recognized three different pockets with volumes between 109.7 and 395.6 $Å^3$, hydrophobicity ranged 0.34 and 0.55, and buriedness higher than 0.5 for all of them.

Binding site similarity is another method used in identifying binding pockets. It is based on the search for similar regions in terms of shape and physicochemical properties on different proteins. Two proteins are considered different if they show a percentage of sequence identity lower than 95%. The detection of such binding sites between different proteins can be done in two ways. In the first one, the binding sites of the protein to assess are matched without involving the ligands, and in the second one, the different positions adopted by the ligand bound to each evaluated protein are compared. The root mean squared deviation (RMSD) values <3 Å are considered corresponding to similar binding sites (Haupt *et al.*, 2013). This value corresponds to the average distance between atoms of overlapping proteins.

Some of the softwares widely used to predict protein binding sites are: MED-SUMO, distributed by MEDIT, detects similar surfaces between macromolecules from 3D databases (Doppelt-Azeroual *et al.*, 2009); FINDSITE provided by Skolnick, identifies ligand binding sites between evolutionarily related protein (Srinivasan *et al.*, 2014); LIGSITE recognizes pockets on protein surfaces, using the calculation of depth and superficialness of protein regions, as well as probes of different sizes

(Kawabata, 2010); ICM-PocketFinder, provided by Molsoft, calculates the probability of drug binding (MolSoft, 2000).

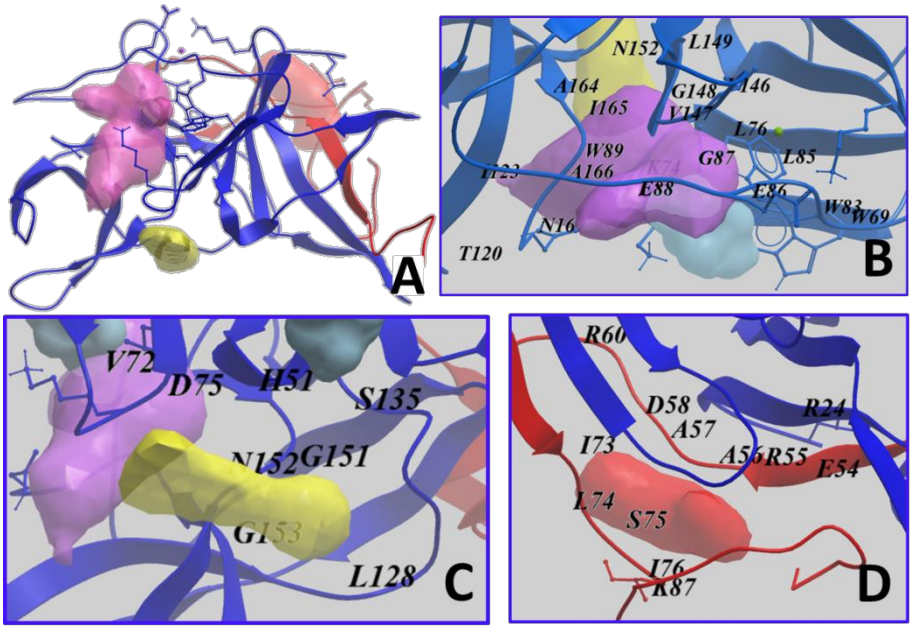


Figure 2.6. Tridimensional view of predicted binding pockets for NS2B/NS3 (PDB ID: 2FOM) protein complex. (A) Full view of dengue protease complex. (B) Enlarged view of pocket 1, (C) pocket 2, and (D) pocket 3. Contact residues of each pocket are shown. The segment of NS2B is represented in red ribbon and NS3 segment in blue ribbon.

2.5. References

- Böhm, H.-J. and G. Klebe (1996). "What Can We Learn from Molecular Recognition in Protein–Ligand Complexes for the Design of New Drugs?" Angewandte Chemie International Edition in English 35(22): 2588-2614.
- Breitenlechner, C. B., D. Bossemeyer, *et al*. (2005). "Crystallography for protein kinase drug design: PKA and SRC case studies." Biochim Biophys Acta 1754(1-2): 38-49.
- Bronowska, A. K. (2011). Thermodynamics of Ligand-Protein Interactions: Implications for Molecular Design. Thermodynamics -

Interaction Studies - Solids, Liquids and Gases. J. C. Moreno-Pirajan, InTech.

- Brooijmans, N. 2009. Docking methods, ligand design, and validating data sets in the structural genomics era. In Structural Bioinformatics. J. Gu and P.E. Bourne, editors. John Wiley & Sons, Inc., Hoboken, New Jersey.
- Cavasotto, C. N. and S. S. Phatak (2009). "Homology modeling in drug discovery: current trends and applications." Drug Discov Today 14(13-14): 676-683.
- Chowdhry, B. Z. and S. E. Harding (2001). Protein-ligand interactions and their analysis. Protein-ligand interactions: Hydrodynamics and calorimetry. B. Z. C. S.E. Harding, Oxford University Press.
- Cleaves, H. J. (2011). Molecular Recognition. Encyclopedia of Astrobiology. M. Gargaud, R. Amils, J. C. Quintanilla*et al*. Berlin, Heidelberg, Springer Berlin Heidelberg: 1079-1080.
- Cohen, N. C. (1996). Guidebook on Molecular Modeling in Drug Design. Boston, Academic Press.
- de Almeida, H., I. M. Bastos, *et al*. (2013). "New binding site conformations of the dengue virus NS3 protease accessed by molecular dynamics simulation." PLoS One 8(8): e72402.
- Doppelt-Azeroual, O., F. Moriaud, F. Delfaud, and A.G. de Brevern. 2009. Analysis of HSP90-related folds with MED-SuMo classification approach. Drug design, development and therapy. 3:59-72.
- Du, X., Y. Li, *et al*. (2016). "Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods." Int J Mol Sci 17(2).
- Forli, S., R. Huey, M.E. Pique, M.F. Sanner, D.S. Goodsell, and A.J. Olson. 2016. Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nature protocols*. 11:905-919.
- Grant, B. J., A. A. Gorfe, *et al*. (2010). "Large conformational changes in proteins: signaling and other functions." Curr Opin Struct Biol 20(2): 142-147.
- Gupta, S. P. (2012). "Matrix metalloproteinase inhibitors: specificity of binding and structure-activity relationships." EXS 103: v-vi.
- Halperin, I., H. Wolfson, *et al*. (2003). "SiteLight: binding-site prediction using phage display libraries." Protein Sci. 12(7): 1344-1359.
- Haupt, V. J., S. Daminelli, *et al*. (2013). "Drug Promiscuity in PDB: Protein Binding Site Similarity Is Key." PLoS One 8(6): e65894.

- Hawkins, P. S., G. (2006). Ligand-Based Design Workflow, OpenEye Scientifi c Software.
- Henrich, S., O. M. Salo-Ahen, *et al*. (2010). "Computational approaches to identifying and characterizing protein binding sites for ligand design." J Mol Recognit 23(2): 209-219.
- Kalidas, Y. and N. Chandra (2008). "PocketDepth: a new depth based algorithm for identification of ligand binding sites in proteins." J Struct Biol 161(1): 31-42.
- Kang, L., B. G. Chung, *et al*. (2008). "Microfluidics for drug discovery and development: from target selection to product lifecycle management." Drug Discov Today 13(1-2): 1-13.
- Kawabata, T. 2010. Detection of multiscale pockets on protein surfaces using mathematical morphology.
- Kessel, A. B.-t. N. (2010). Protein-Ligand Interactions: Structure, Function, and Motion. Introduction to Proteins, CRC Press: 515-591.
- Koes, D.R., and C.J. Camacho. 2012. ZINCPharmer: pharmacophore search of the ZINC database. Nucleic acids research. 40:W409-414.
- Kontogiorgis, C. A., P. Papaioannou, *et al*. (2005). "Matrix metalloproteinase inhibitors: a review on pharmacophore mapping and (Q)SARs results." Curr Med Chem 12(3): 339-355.
- Lengauer, T., C. Lemmen, *et al*. (2004). "Novel technologies for virtual screening." Drug Discov Today 9(1): 27-34.
- MolSoft. 2000. ICM 2.8 Program Manual. MolSoft LLC, San Diego, CA.
- Nisius, B., F. Sha, *et al*. (2012). "Structure-based computational analysis of protein binding sites for function and druggability prediction." J Biotechnol 159(3): 123-134.
- Pace, C. N., B. A. Shirley, *et al*. (1996). "Forces contributing to the conformational stability of proteins." FASEB J 10(1): 75-83.
- Pierri, C. L., G. Parisi, *et al*. (2010). "Computational approaches for protein function prediction: a combined strategy from multiple sequence alignment to molecular docking-based virtual screening." Biochim Biophys Acta 1804(9): 1695-1712.
- Ramos, M. J., A. Melo, *et al*. (2001). Chapter 13 – Modeling enzyme – ligand interactions. Theoretical and Computational Chemistry. A. E. Leif, Elsevier. Volume 9: 539-595.
- Saada, I. and J. K. Pearson (2011). "A theoretical study of the structure and electron density of the peptide bond." Computational and Theoretical Chemistry 969(1–3): 76-82.

- Schreiber, G. and A. E. Keating (2011). "Protein binding specificity versus promiscuity." Curr Opin Struct Biol 21(1): 50-61.
- Sousa, S.F., A.J. Ribeiro, J.T. Coimbra, R.P. Neves, S.A. Martins, N.S. Moorthy, P.A. Fernandes, and M.J. Ramos. 2013. Protein-ligand docking in the new millennium--a retrospective of 10 years in the field. Current medicinal chemistry. 20:2296-2314.
- Srinivasan, B., H. Zhou, J. Kubanek, and J. Skolnick. 2014. Experimental validation of FINDSITEcomb virtual ligand screening results for eight proteins yields novel nanomolar and micromolar binders. J. Cheminformatics. 6:16.
- Wang, J.C., and J.H. Lin. 2013. Scoring functions for prediction of protein-ligand interactions. *Current pharmaceutical design*. 19:2174-2182.
- Wasserman, S. R., J. W. Koss, *et al*. (2012). "Rapid-access, high-throughput synchrotron crystallography for drug discovery." Trends Pharmacol Sci 33(5): 261-267.
- Williams, M.A., and T. Daviter. 2013. Protein-ligand interactions [electronic resource] : methods and applications. Humana Press, New York.
- Wieder, M., U. Perricone, T. Seidel, S. Boresch, and T. Langer. 2016. Comparing pharmacophore models derived from crystal structures and from molecular dynamics simulations. Monatshefte fur chemie. 147:553-563.

CHAPTER 3.
New molecular targets for TCDD

3.1. Introduction

Polychlorinated aromatic hydrocarbons are ubiquitous environmental pollutants that affect human health, resulting in a diversity of systemic and tissue-specific pathological changes. The term dioxin includes about hundreds of chemicals that are highly persistent in the environment, such as certain polychlorinated dibenzofurans (PCDF), polychlorinated biphenyls (PCBs), polychlorinated diphenyl ethers and polychlorinated naphthalenes, among others; of which 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) (Figure 3.1) is the most toxic. Dioxins are formed naturally during volcanic eruptions and forest fires, and as an intermediate product of fossil fuels, pesticides manufacturing, metal recycling, waste incineration, polyvinylchloride plastic production, and chlorine bleaching in paper production, among others (Hutz, 1999). It has also been found in automobile emissions, cigarette smoke and coal burning (Navratil and Minarik, 2011).

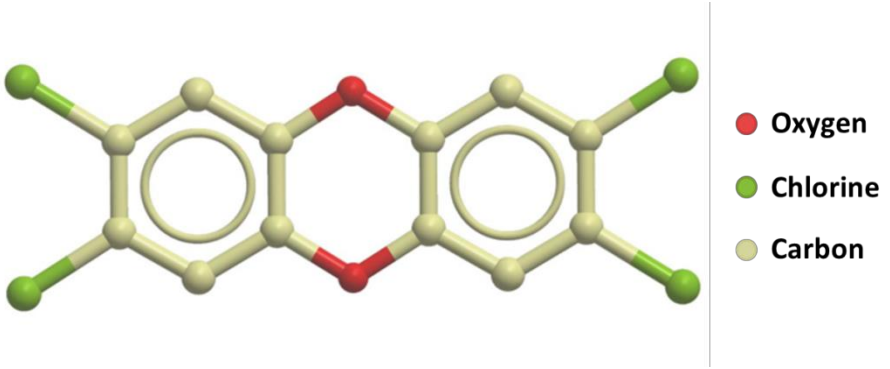


Figure 3.1. Tridimensional structure of 2,3,7,8-Tetrachlorodibenzo-p-dioxin (CID: 15625).

Dioxins are fat soluble and bioaccumulate in most animals, with an average half-life between 7 and 11 years in humans, remaining and biomagnifying along the food chain. From the environment, TCDD is taken through ingestion or respiration, however, the main source of exposure is food, such as meat, dairy products, fish and shellfish. Once in the organism, it acts as a hormone, so it is able to interfere in different signaling pathways (Figure 3.2), including celular differentiation, regulation and development (Schecter and Gasiewicz, 2003).

Figure 3.2. Data mining to alteration of gene expression and type of cells related to TCDD toxicity.

Some of its effects on organisms are shown in Figure 3.3. TCDD exposure results in weightloss, damage to various organs such as liver, lung, thymus, pancreas, adrenal glands, and central nervous system (Pohjanvirta and Tuomisto, 1994), carcinogenicity in humans (Chan *et al*., 2004; Knerr and Schrenk, 2006), reproductive toxicity (Jin *et al*., 2008a,b; Arima *et al*., 2009; Porpora *et al*., 2009), immunosuppression (Gogal and Holladay, 2008; Smialowicz *et al*., 2008), neurological dysfunction (Urban *et al*., 2007; Akahoshi *et al*., 2009), hepatotoxicity (Boutros *et al*., 2008; Nukaya *et al*., 2009), and teratogenicity in laboratory animals (Bryant *et al*., 2001; Thackaberry *et al*., 2005), skeletal deformities, kidney defects and a possible association with sarcomas, lymphomas and carcinomas (ATSDR, 1998). The proper disposal of this type of compounds requires to be treated as a hazardous waste, and be incinerated at temperatures higher than 850 °C (WHO, 2014).

3.1.1. Physiological dynamic of TCDD

TCDD have a relatively extraordinary potency of action due to its planar shape, high stability and accumulation in fatty tissues. It is well known that once inside the body (Figure 3.4), TCDD toxicity effects are mediated by a ligand-dependent transcription factor, the aryl hydrocarbon receptor (AhR)(Burcham, 2014), which displays a high binding affinity for it (Mimura and Fujii-Kuriyama, 2003). Studies using AhR-null mutant mice have confirmed the critical role that AhR plays on the toxic effects of TCDD (Fernandez-Salguero *et al*., 1996; Mimura *et al*., 1997).

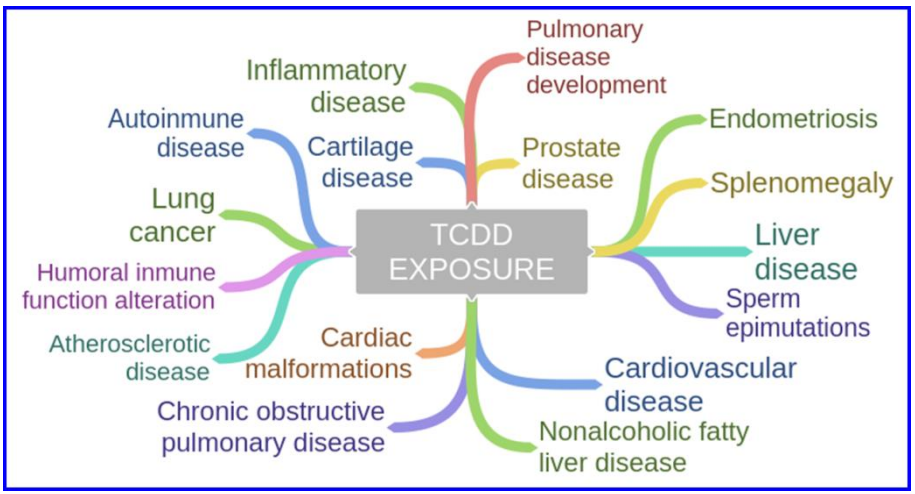


Figure 3.3. Adverse health effects of TCDD.

AhR is located in the cytoplasm, associated with a heat shock protein 90 (HSP90) dimer and the co-chaperone protein X-associated protein 2 (XAP2 or AIP1). Once binding with dioxin, the nuclear localization signal (NLS) in AhR is exposed, allowing the complex to pass into the nucleus of the cell, after being recognized by alpha karyopherin α (Schecter and Gasiewicz, 2003). There, HSP90-XAP2 complex is released, and AhR is heterodimerizated with AhRNT (AhR nuclear translocator), creating a ligand–receptor–translocator complex (Swanson and Bradfield, 1993). When activated by TCDD, the expression of several target genes regulated by the receptor is altered (Sato *et al*., 2008), inducing cytochrome P450 enzymes (CYP450) transcription, by binding to dioxin-responsive elements (DRE) in its upstream region (enhancers). These enzymes metabolize and activate carcinogens (Murray *et al.*, 2014).

Some of CYP450 enzymes induced are CYPIA1, CYPIA2, and CYPIB1. The main function of these enzymes is to oxidize xenobiotics, which are generally lipophilic. Thus, hydrophilic groups could be added to increase their water solubility, increasing their excretion by the kidneys, and in consequence, reducing their potential toxic effects (Shimada, 2006). However, in the case of dioxins, halogens in its structure prevent oxidation by induced CYP450 enzymes, decreasing the removal rate. Because of this, the half-life time in the body, and therefore the effects

produced, increase (Schecter and Gasiewicz, 2003). According to microarray studies, AhR participates in the regulation of various genes involved in various biochemical pathways, such as cholesterol synthesis, energy metabolism, lipid and xenobiotic, cell cycle, cell migration and immune function (Murray *et al*., 2014).

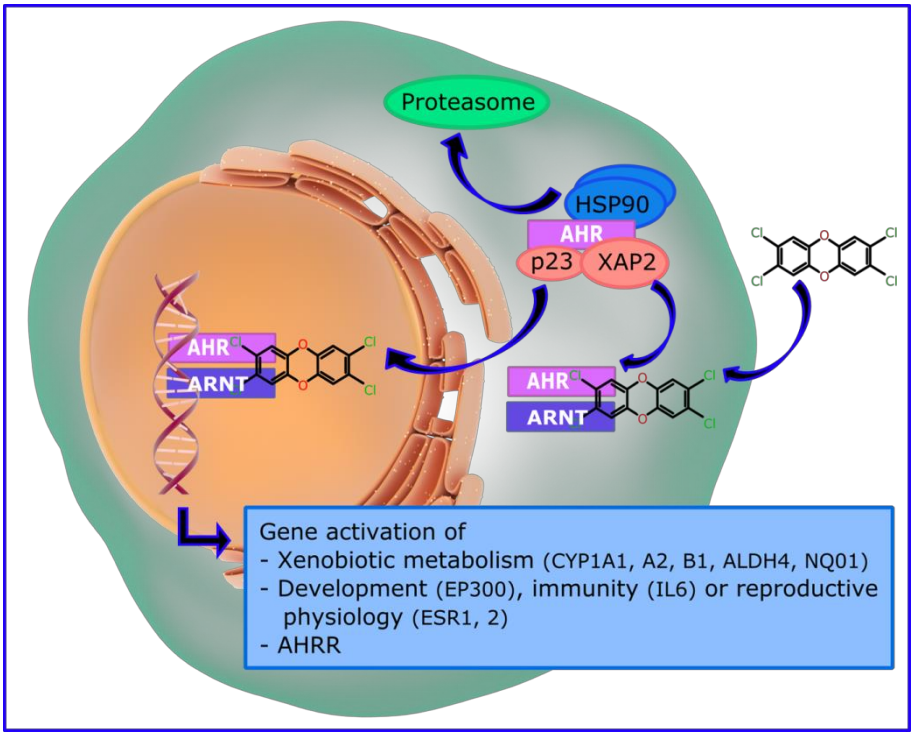


Figure 3.4. AhR activation and signaling after TCDD exposure. Modified from Bersten *et al.* (2013).

Alternative pathways have been identified. Some studies reveal that the TCDD–AhR complex activates cytosolic protein kinase in nuclear-free subcellular homogenates (Enan and Matsumura, 1995). *In vitro* studies have raised the possibility that TCDD could affect cellular pathways in AhR-independent pathways. For instance, it has been shown that TCDD is able to induce Ca^{2+}/Calmodulin signals that regulate apoptosis in EL-4 cells (Kobayashi *et al*., 2009), activate the mitogen activated protein kinase (MAPK) pathway in RAW 264.7 murine macrophages (Park *et al*., 2005), and induce immunotoxic effects in EL-4 cells through a

mechanism mediated by insulin-like growth factor-binding protein-6 (IGFBP- 6) (Park *et al*., 2003), among other AhR-independent effects. Although these apparently receptor-independent processes are not clear yet, it is possible that some of the toxicologic and pathologic effects were caused by targets different from AhR.

TCDD, as well as any ligand, alters the functions of target proteins acting as agonist or antagonist. Testing individual molecules to find specific protein targets is a time-consuming and expensive process. An approach that could help the discovery of new targets for toxic compounds is virtual screening. Screening for ligand conformations can be performed using a ligand-based approach (Lyne, 2002; Stahura and Bajorath, 2005), wich is established on the assumption that structurally similar compounds are likely to exhibit similar biological activities (Stockwell, 2000). Its design includes pharmacophore searches (Khedkar *et al*., 2007) using the structure of a ligand or series of ligands that are active against the target to determine ligand–protein interactions. This technique categorizes ligands that bind to the protein of interest and allows making predictions about activation or inhibition of the protein (Lewis *et al*., 2009). Ligand-based screening methods have been conventionally used in the early stages of a variety of drug discovery projects to mine chemical databases with the aim of identifying new hit compounds or optimizing leads (Agrawal *et al*., 2007; Kumar *et al*., 2009; Pérez-Nueno *et al*., 2009). Target Fishing Dock (TarFisDock) is a web server that docks small molecules with protein structures in the potential drug target database (PDTD), in an effort to discover new drug targets (Li *et al.*, 2006). It works performing reverse molecular docking. This process allows docking a particular compound into known protein binding pockets found in the potential drug target database (PDTD). This server has been used to predict binding sites for multiple target proteins, such as vitamin E and 4H-tamoxifen, with predictions that were nearly 50% correct, after experimental verification, indicating the relative reliability of this server tool (Li *et al*., 2006).

In this report, we used virtual screening to detect new posible targets for TCDD, and discuss the possible implications in human health.

3. 2. Methods

A three-step approach has been used to find new protein targets for TCDD. This includes TCDD optimization, virtual screening of new targets and validation. First, the TCDD structure has been optimized using DFT at the B3LYP/6-31G level. Calculations were carried out with Gaussian 03 package program (Frisch *et al*., 2003). The resultant geometry was translated to Mol2 format with Open Babel (Guha *et al*., 2006), and the Gasteiger partial charge were calculated. The optimized structure was submitted to TarFisDock (Li *et al*., 2006) to detect proteins with tri-dimensional structures having theoretical binding sites to TCDD. The search started using the ‘‘targets in all categories” option, and ligand docking was performed on all proteins (1207 proteins available) present in the potential drug target database (PDTD). The output consists of the top 2%, 5% or 10% best hits, ranked by an energy score, providing binding conformations and a table with related target information. The 25 proteins with the best binding score for TCDD, were scanned for pocket verification using the algorithm SiteID using the SYBYL8.1 package.

In order to calculate theoretical affinities, those proteins for which a binding pocket had been detected were docked with TCDD using AutoDock Vina 1.0 program. The geometrical orientation of TCDD on the protein generated by TarFisDock and that given by AutoDock Vina 1.0 were compared calculating the RMSD, using the RMSD Calculator, tools from the SoongSil University Bioinformatics eBioLab. Interactions between TCDD and proteins were checked with LigandScout 2.0 (Wolber and Langer, 2005). This software extracts and interprets ligands and their macromolecular environment from a PDB file, previously prepared in SYBYL8.1, showing a 2D image of the interactions.

As the main target for TCDD is the AhR, a mammal (*Mus musculus*) 3D model of this protein (PM0074603), available at the publicly accessible Protein Model Database (Castrignano *et al*., 2006), was used to perform docking with TCDD using AutoDock Vina 1.0.

3. 3. Results

The top 10% TarFisDock output identified 106 protein targets with binding affinities to TCDD over -28.41 kcal/mol. After testing each one with AutoDock Vina program, top 2% of the target list were selected

according to the docking score (Table 3.1), and those with a binding affinity over -7.0 kcal/mol were chosen for the analyses. RMSDs values calculated for the ligand position in the protein as given by TarFisDock and AutoDock vina revealed that TCDD acquires the same position using both tools.

Based on the docking score, we found that TCDD could interact in an AhR-independent way with different enzymes such as hydrolases, isomerases, oxidoreductases, oxidases and other receptors such as the nuclear orphan receptor Lxr-beta (LXRb). As it can be seen from Table 3.1, proteins with the greater binding affinity were neutrophil collagenase (MMP8) (Figure 3.5), stromelysin-1 MMP3 (Figure 3.6), oxidosqualene cyclase (Figure 3.7), and myeloperoxidase (Figure 3.8). For these proteins, the binding sites for TCDD are embedded in a hydrophobic region, with interactions of both the chlorine atoms and the aromatics rings.

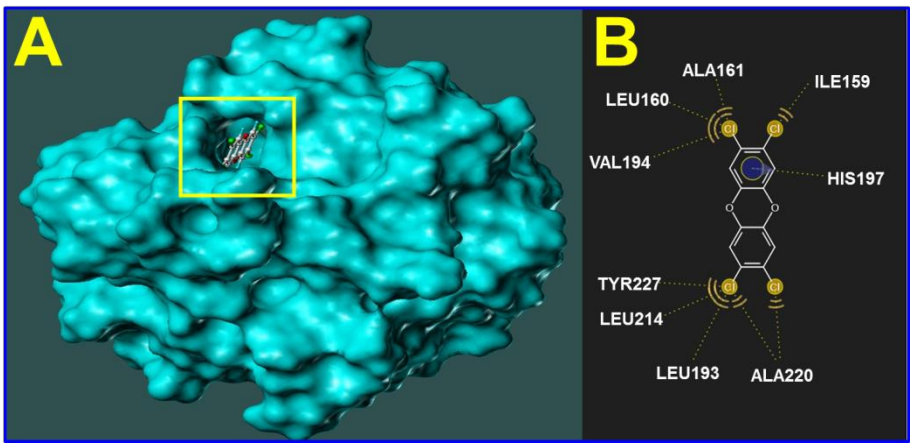


Figure 3.5. Tridimensional view of TCDD docking with neutrophil collagenase, MMP8. (A) 3D structure of the TCDD-protein complex. The yellow bow shows TCDD in the binding site. (B) Interactions between aminoacids in the protein and TCDD.

The AutoDock Vina-generated docking of TCDD with the available theoretical PM0074603 AhR model in Protein Model Database (Figure 3.9), showed that as observed for hypothetical TCDD targets, the binding site is also hydrophobic with prevalence of aromatic and aliphatic

residues. However, the calculated affinity (-7.5 kcal/mol) is not as good as the ones observed for other TCDD targets predicted by TarFisDock.

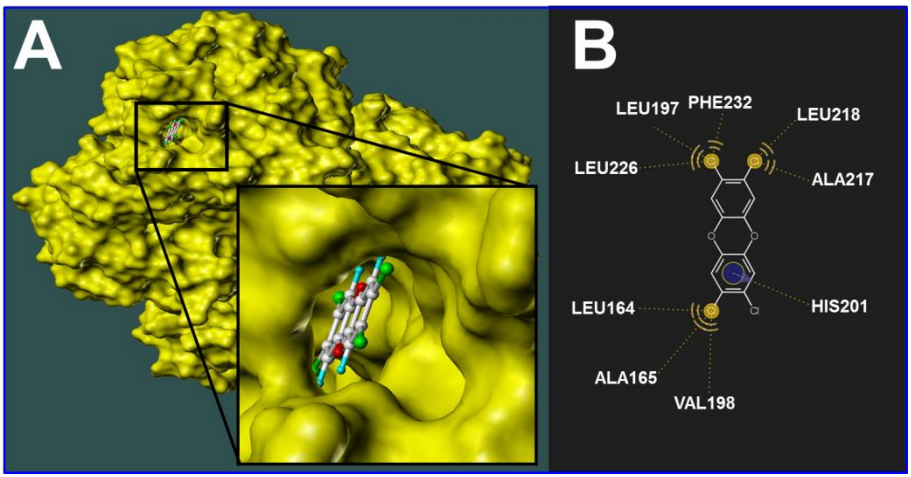


Figure 3.6. Tridimensional view of TCDD docking with stromelysin-1, MMP3. (A) 3D structure of the TCDD-protein complex, showing binding TCDD in the site cavity. (B) Interactions between aminoacids in the protein and TCDD.

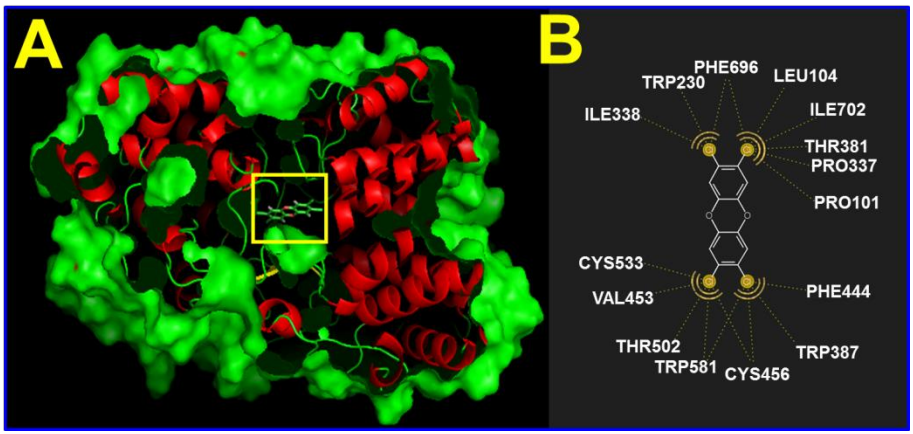


Figure 3.7. Tridimensional view of TCDD docking with oxidosqualene cyclase. (A) 3D structure of the TCDD-protein complex, showing binding TCDD in the site cavity. (B) Interactions between aminoacids in the protein and TCDD.

Table 3.1. Proteins with theoretical binding sites for TCDD.

PDB ID	Protein Name	Gene Name	Energy score (Kcalmol^{-1})		RMSD (Å)
			TarFisDock	AutoDockVina	
1I76	Neutrophil Collagenase, MMP8	MMP8	-35,90	-10,4	0.741
1QIA	Stromelysin-1, MMP3	MMP3 Synonym: STMY1	-33.84	-9.8	0.493
1W6K	Oxidosqualene cyclase (Lanosterol synthase).	LSS Synonym: OSC	-33.52	-9.8	0.414
1D2V	Myeloperoxidase	MPO	-34.11	-8.9	0.688
2JFE	Cytosolic beta-Glucosidase	GBA3 Synonym: CBG, CBGL1	-32.59	-8.7	0.251
1K4W	Nuclear Orphan Receptor Lxr-beta	NR1H2 Synonym: LXRB, NER, UNR	-28.52	-8.4	0.685
1XOS	cAMP-specific 3',5'-cyclic phosphodiesterase 4B	PDE4B Synonym: DPDE4	-31,78	-8.4	0.174
1S2A	Prostaglandin D2 11-ketoreductase	AKR1C3 Synonyms: DDH1, HSD17B5, KIAA0119, PGFS	-34,38	-8.1	0.683
1B41	Acetylcholinesterase	ACHE	-30.5	-8.2	0.358
4QUP	Monooxygenase	NA	-30.06	-8.1	0.255

PDB ID	Protein Name	Gene Name	Energy score (Kcalmol^{-1})		RMSD (Å)
			TarFisDock	AutoDockVina	
2NZL	Hydroxyacid oxidase 1	HAO2 Synonyms: HAOX2	-34.91	-8.0	0.326
1PY5	TGF-beta receptor type I	TGFBR1	-28.69	-7.9	3.061

NA. Not available.

3. 4. Discussion

The virtual screening of protein targets for TCDD has shown that several of them have the potential to bind this toxic compound. It is considered that a ligand is promiscuous when it is capable of forming chemical interactions with more than one protein target. Although the physiological and toxicological relevance of this finding is still unknown, it brings the opportunity to rethink new mechanisms by which TCDD could be interfering with many biochemical systems in the body.

Even though in this article were taken into account only those proteins with the best values of affinity for the ligand of interest, promiscuous ligands with low binding affinities, may also show high efficiency (Haupt *et al.*, 2013). The binding energy between a protein and a ligand varies from -22 and -2.5 kcal/mol. Evidence suggests that the biological purpose and strength of the bond (affinity value) between protein and ligand could be related. Complexes formed by interactions with weak binding force generally correspond to proteins involved in signaling networks, facilitating the partner change. Reversible and strong bindings, as is the case of interactions between enzymes and cofactors, show values between -5.5 and -9.5 kcal/mol. Finally, other proteins have lower affinity values indicating better binding strength, generally are related with metals, small anions and ligands which form covalent bonds (Kessel, 2010).

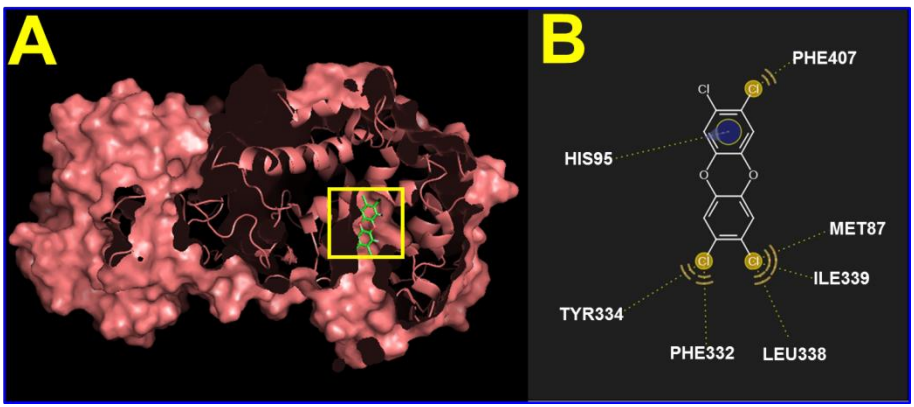


Figure 3.8. Tridimensional view of the TCDD docking with myeloperoxidase. (A) 3D structure of the TCDD-protein complex, showing

binding TCDD in the site cavity. (B) Interactions between aminoacids in the protein and TCDD.

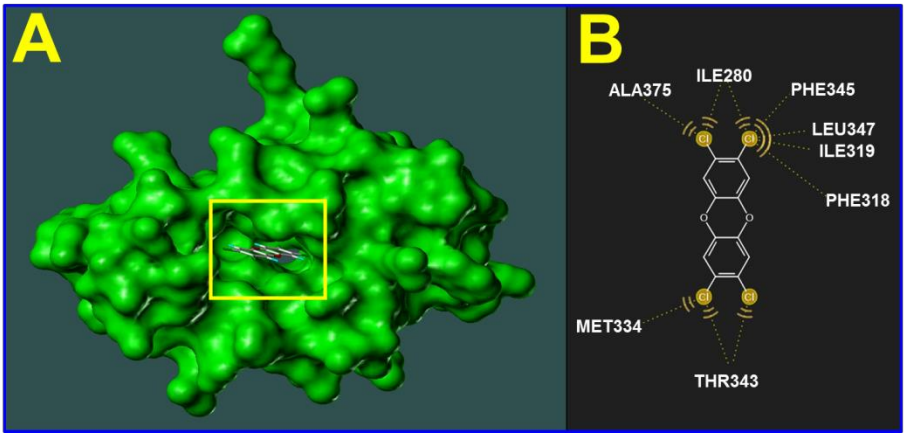


Figure 3.9. Tridimensional view of the docking of TCDD with AhR. (A) 3D structure of the protein showing binding site cavity. (B) Interactions between aminoacids in the protein and TCDD.

The highest binding score for TCDD was obtained for metalloproteinase 8 (MMP8), known as neutrophil collagenase. The binding pocket to TCDD is formed by residues between positions 280 and 361, which correspond to an allosteric site on the protein; according to the associated entries in PDB with MMP8 (such as 1A85, 1ZP5 and 3DPF) the active site is formed by residues between positions 171 and 250 (Kufareva *et al.*, 2012), as shown in Figure 2.3. This protein is a member of the interstitial collagenase family, capable of degrading fibrillar type I, II, and III collagens. Its expression and activity has been associated with chronic inflammatory and fibrotic diseases such as cystic fibrosis, rheumatoid arthritis, periodontal disease, and chronic skin wounds (Matsuki *et al*., 1996; Nwomeh *et al*., 1999; Allport *et al*., 2002; Ratjen *et al*., 2002).

Interestingly, rheumatoid arthritis (Kobayashi *et al*., 2008), and periodontal diseases (Yoshizawa *et al*., 2005) have been observed as a consequence of TCDD exposure. Similarly, TCDD can activate macrophage and neutrophil accumulation, resulting in neutrophilia, inducing an inflammatory response in certain tissues (Kerkvliet, 1995).

It has been demonstrated that normal melanocytes and melanoma cells exposed to dioxin have increased expression and activity of metallopeptidases, such as MMP1, MMP2 and MMP9 (Villano *et al*., 2006). Particularly in the case of MMP8, mice exposed for 6 to 12 hours to TCDD show an upregulated expression for this gene. So far, there is no evidence that this response is mediated by AhR, however, some authors suggest that transcriptional regulation occurs by via June or Fos (Hillegass *et al.*, 2006)

Stromelysin-1 (MMP3), another metallopeptidase, also had a high theoretical docking affinity with TCDD. The TCDD binding pocket coincides with the active site identified from the PDB entries associated with this protein by using Pocketome software (Kufareva *et al*., 2012). This protein can promote the proteolysis of basal membrane collagen and induce synthesis of other matrix metalloproteinases (MMPs) such as MMP1 and MMP9 (Brinckerhoff *et al*., 2000; Van Themsche *et al*., 2004). The proper function of the normal human endometrium relies on well organized cell–cell interactions regulated locally by cytokines and growth factors under the direction of steroid hormones. In ectopic endometrium, MMP3 may participate in the process of invasion and tissue remodelling that is hypothesized to occur in the pathogenesis of endometriosis (Cox *et al*., 2001). Experimental evidence suggests that TCDD alters the endometrial response to estrogen, facilitating the onset of endometriosis, and contributing to its development by increasing the invasion of endometrial stromal cells in patients with endometriosis (Yu *et al*., 2008). It also affects the pathophysiology of endometriosis by modulation of immune and endocrine function (Pauwels *et al*., 2001). Although metalloproteinases are reported to be involved in this process, it is clear that at least two of the proteins belonging to this family are potential targets for TCDD, and this could in part, explain, some of the toxic effects of dioxin on endometriosis. However, the specific mechanism of dioxin-mediated toxicity in the pathogenesis of endometriosis remains unclear (Pauwels *et al*., 2001).

Oxidosqualene cyclase catalyzes the cyclization of squalene 2,3-epoxide to form the tetracyclic lanosterol, the precursor to all sterols (Kramer, 2006). Interestingly, the TCDD binding site is the same for squalene, both interacting with Phe 696. Several studies have shown that TCDD alters cholesterol biosynthesis (Sato *et al*., 2008). Although this

process has been linked to the activation of the AhR (Tanos *et al.*, 2012), it is quite intriguing to determine whether TCDD binding to this enzyme can also induce some changes on colesterol homeostasis.

Myeloperoxidase (MPO) is a lysosomal enzyme that is found in white blood cells and neutrophils, and its primary function is to mediate a defense response against invading pathogens (Shetty *et al*., 2006). MPO uses hydrogen peroxide and chloride to produce hypochlorous acid (HOCl), a potent oxidant that reacts with proteins, DNA, and lipids to cause cellular injury. However, MPO has also been shown to metabolize chlorophenols such as 2,4,5-tri-, 2,3,4,6-tetra and pentachlorophenol, into dibenzo-p-dioxins and dibenzofurans (PCDD/F) (Wittsiepe *et al*., 2000). This metabolic pathway not only could increase the body burden of TCDD, but also could explain the binding affinity of this compound by MPO.

Virtual screening also showed that transforming growth factor beta type I (TGFb1) can also be a potential target for TCDD. This protein belongs to a superfamily of paracrine-acting peptides, known to elicit a variety of biological activities in many cell types, including effects on cell proliferation, cell differentiation, cell adhesion, cell migration, and regulation of extracellular matrix components. Gene expression studies have showed that TCDD down regulated TGFb1 (Sahlberg *et al*., 2007). In mice, TCDD induces oxidative stress on testes, related to an increase of TGFb1 activity, triggering TGF-beta1-Smad pathway (Jin *et al*., 2008a).

Another protein found to be a possible target for TCDD is the nuclear orphan receptor Lxr-beta (LXRb), a ligand-activated transcription factor that belongs to the nuclear receptor superfamily (Apfel *et al*., 1994). It is an isoform of the liver X receptors (LXRs), classified as orphan nuclear receptors because their natural ligands are unknown. LXRb is expressed in almost all the organism tissues (Cortés *et al*., 2005).

There are other interesting proteins that also were predicted to have binding sites for TCDD: cAMP cyclic phosphodiesterase 4 (PDE4) has four isoforms (A–D) (Lugnier, 2006). PDE4B is predominantly expressed in a variety of inflammatory and structural lung cells (Torphy, 1998), and it has been shown to modulate inflammatory responses (Jin *et al*., 2005). Prostaglandin D2 11-ketoreductase (AKR1C3) is a member of the aldo–

keto reductase (AKR) gene superfamily that can act on a variety of substrates including hormones, prostaglandins and PAHs (Palackal *et al*., 2002). It has been related to carcinogenesis due to its implication with PAH activation (Penning and Drury, 2007). Cytosolic beta-glucosidase (hCBG), is present in the liver, kidney, intestine and spleen of humans but primarily in enterocytes (Daniels *et al*., 1981). This enzyme hydrolyses many common dietary xenobiotics, including glycosides of phytoestrogens, flavonoids, simple phenolics and cyanogens (Berrin *et al*., 2002).

Finally, acetylcholinesterase (AChE) is the enzyme involved in the deactivation of acetylcholine at nerve endings, preventing continuous nerve firings. It is vital for normal functioning of sensory and neuromuscular systems (Murphy, 1986). Although it is a common molecular target for organophosphate and carbammate pesticides, its binding capacity to TCDD is surprising weak.

As it was observed for proteins TCDD targets predicted by TarFisDock, the docking of TCDD with an available mouse model of AhR, the known receptor for this compound, also showed that the protein binding site for the chlorinated xenobiotic is predominantly hydrophobic, as published by Pandini *et al*. (2009), for which several interacting residues are the same as those depicted in our working model (Figure 3.9).

Finally, the 2D views of the interaction patterns for TCDD and TarFisDock-detected targets show that as suggested by Kortagere *et al*. (2008), halogens form mainly hydrophobic interactions with amino acids such as leucine, phenylalanine and hydrogen bonds with serine or threonine.

In short, although it is safe to state that this virtual screening approach provides insight about hypothetical targets for TCDD, it is also worth emphasizing that currently, the use of these tools might have limitations, especially when considering environmental samples where activity depends on mixtures of TCDD-like compounds. Future developments must include prediction of the toxicity induced by dioxins and related compounds, by obtaining parameters such as the toxic equivalence factors (TEF) (Van den Berg *et al*., 1998; 2006), directly from

the extent of individual ligand– AhR binding interactions, detected for each one of the components in a sample.

3. 5. References

- Agrawal, H., Kumar, A., Chandra-Bal, N., Imran-Siddiqia, M., Arora, A., 2007. Ligand based virtual screening and biological evaluation of inhibitors of chorismate mutase (Rv1885c) from Mycobacterium tuberculosis H37Rv. Bioorg. Med. Chem Lett. 17, 3053–3058.
- Akahoshi, E., Yoshimura, S., Uruno, S., Ishihara-Sugano, M., 2009. Effect of dioxins on regulation of tyrosine hydroxylase gene expression by aryl hydrocarbon receptor: a neurotoxicology study. Environ. Health 8, 24.
- Allport, J.R., Lim, Y.C., Shipley, J.M., Senior, R.M., Shapiro, S.D., Matsuyoshi, N., Vestweber, D., Luscinskas, F.W., 2002. Neutrophils from MMP-9- or neutrophil elastase-deficient mice show no defect in transendothelial migration under flow *in vitro*. J. Leukocyte Biol. 71, 821–828.
- Apfel, R., Benbrook, D., Lernhardt, E., Ortiz, M.A., Salbert, G., Pfahl, M., 1994. A novel orphan receptor specific for a subset of thyroid hormone-responsive elements and its interaction with the retinoid/thyroid hormone receptor subfamily. Mol. Cell. Biol. 14, 7025–7035.
- Arima, A., Kato, H., Ooshima, Y., Tateishi, T., Inoue, A., Muneoka, A., Ihara, T., Kamimura, S., Fukusato, T., Kubota, S., Sumida, H., Yasuda, M., 2009. *In utero* and lactational exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) induces a reduction in epididymal and ejaculated sperm number in rhesus monkeys. Reprod. Toxicol. 28, 495–502.
- ATSDR. 1998. Toxicological Profile for Chlorinated Dibenzo-p-Dioxins. U.S.D.o.H.a.H.S. Public Health Service, editor. Agency for Toxic Substances and Disease Registry, Atlanta, GA.
- Berrin, J.G., McLauchlan, W.R., Needs, P., Williamson, G., Puigserver, A., Kroon, P.A., Juge, N., 2002. Functional expression of human liver cytosolic b-glucosidase in Pichia pastoris: insights into its role in the metabolism of dietary glucosides. Eur. J. Biochem. 269, 249–258.
- Bersten, D.C., A.E. Sullivan, D.J. Peet, and M.L. Whitelaw. 2013. bHLH-PAS proteins in cancer. Nature reviews. Cancer. 13:827-841.

- Boutros, P.C., Yan, R., Moffat, I.D., Pohjanvirta, R., Okey, A.B., 2008. Transcriptomic responses to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in liver: comparison of rat and mouse. BMC Genom. 9, 419.
- Brinckerhoff, C.E., Rutter, J.L., Benbow, U., 2000. Interstitial collagenases as markers of tumor progression. Clin. Cancer Res. 6, 4823–4830.
- Bryant, P.L., Schmid, J.E., Fenton, S.E., Buckalew, A.R., Abbott, B.D., 2001. Teratogenicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin (tcdd) in mice lacking the expression of egf and/or tgf-a. Toxicol. Sci. 62, 103–114.
- Burcham, P.C. 2014. Toxicodynamics: How Chemicals Harm Cells. Springer, London.
- Castrignano, T., De Meo, P.D., Cozzetto, D., Talamo, I.G., Tramontano, A., 2006. The PMDB protein model database. Nucleic Acids Res. 34, 306–309.
- Chan, C.Y.Y., Kim, P.M., Winn, L.M., 2004. TCDD affects DNA double strand-break repair. Toxicol. Sci. 81, 133–138.
- Cortés, V., Quezadaa, N., Rigott, A., Maiz, A., 2005. Nuevos receptores nucleares heterodiméricos: reguladores metabólicos con impacto en fisiopatología y su proyección terapéutica en dislipidemias y diabetes mellitus. Rev. Méd. Chile 133, 1483–1492.
- Cox, K.E., Piva, M., Sharpe-Timms, K.L., 2001. Differential regulation of matrix metalloproteinase-3 gene expression in endometriotic lesions compared with endometrium. Biol. Reprod. 65, 1297–1303.
- Daniels, L.B., Coyle, P.J., Chiao, Y.B., Glew, R.H., Labow, R.S., 1981. Purification and characterization of a cytosolic broad specificity b-glucosidase from human liver. J. Biol. Chem. 256, 13004–13013.
- Enan, E., Matsumura, F., 1995. Evidence for a second pathway i n the action mechanism of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). Significance of Ahreceptor mediated activation of protein kinase under cell-free conditions. Biochem. Pharmacol. 49, 249–261.
- Fernandez-Salguero, P.M., Hilbert, D.M., Rudikoff, S., Ward, J.M., Gonzalez, F.J., 1996. Aryl-hydrocarbon receptor-deficient mice are resistant to 2,3,7,8- tetrachlorodibenzo-p-dioxin-induced toxicity. Toxicol. Appl. Pharmacol. 140, 173–179.
- 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., Pople, J.A., 2003. Gaussian 03. Revision B.02. Gaussian, Inc., Pittsburgh PA.

- Gogal Jr., R.M., Holladay, S.D., 2008. Perinatal TCDD exposure and the adult onset of autoimmune disease. J. Immunotoxicol. 5, 413–418.
- Guha, R., Howard, M.T., Hutchison, G.R., Murray-Rust, P., Rzepa, H., Steinbeck, C., Wegner, J., Willighagen, E.L., 2006. The blue obelisk-interoperability in chemical informatics. J. Chem. Inf. Model. 46, 991–998.
- Haupt, V.J., S. Daminelli, and M. Schroeder. 2013. Drug Promiscuity in PDB: Protein Binding Site Similarity Is Key. PloS one. 8:e65894.
- Hillegass, J.M., K.A. Murphy, C.M. Villano, and L.A. White. 2006. The impact of aryl hydrocarbon receptor signaling on matrix metabolism: implications for development and disease. Biological chemistry. 387:1159-1173.
- Hutz, R.J., 1999. Reproductive endocrine disruption by environmental xenobiotics that modulate the estrogen-signaling pathway, particularly tetrachlorodibenzop-dioxin (TCDD). J. Reprod. Develop. 45, 1–12.
- Jin, S.L.C., Lan, L., Zoudilova, M., Conti, M., 2005. Specific role of phosphodiesterase 4b in lipopolysaccharide-induced signaling in mouse macrophages. J. Immunol. 175, 1523–1531.
- Jin, M.H., Hong, C.H., Lee, H.Y., Kang, H.J., Han, S.W., 2008a. Enhanced TGF-beta1 is involved in 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) induced oxidative stress in C57BL/6 mouse testis. Toxicol. Lett. 178, 202–209.

- Jin, M.H., Ko, H.K., Hong, C.H., Han, S.W., 2008b. *In utero* exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin affects the development of reproductive system in mouse. Yonsei Med. J. 49, 843–850.
- Kerkvliet, N.I., 1995. Immunological effects of chlorinated dibenzo-p-dioxins. Environ. Health Perspect. 103, 47–53.
- Kessel, A.B.-t.N. 2010. Protein-Ligand Interactions: Structure, Function, and Motion. In Introduction to Proteins. CRC Press. 515-591.
- Khedkar, S.A., Malde, A.K., Coutinho, E.C., Srivastava, S., 2007. Pharmacophore modeling in drug discovery and development: an overview. Med. Chem. 3, 187-197.
- Knerr, S., Schrenk, D., 2006. Carcinogenicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin in experimental models. Mol. Nutr. Food Res. 50, 897–907.
- Kobayashi, D., Ahmed, S., Ishida, M., Kasai, S., Kikuchi, H., 2009. Calcium/calmodulin signaling elicits release of cytochrome c during 2,3,7,8-tetrachlorodibenzo-pdioxin-induced apoptosis in the human lymphoblastic T-cell line. L-MAT Toxicol. 258, 25–32.
- Kobayashi, S., Okamoto, H., Iwamoto, T., Toyama, Y., Tomatsu, T., Yamanaka, H., Momohara, S., 2008. A role for the aryl hydrocarbon receptor and the dioxin TCDD in rheumatoid arthritis. Rheumatology 47, 1317–1322.
- Kortagere, S., Ekins, S., We, W.J., 2008. Halogenated ligands and their interactions with amino acids: Implications for structure–activity and structure–toxicity relationships. J. Mol. Graph. Model. 27, 170–177.
- Kramer, M.A., 2006. Focus on Cholesterol Research. New York, Nova Science Publishers, Inc.
- Kufareva, I., A.V. Ilatovskiy, and R. Abagyan. 2012. Pocketome: an encyclopedia of small-molecule binding sites in 4D. Nucleic acids research. 40:D535-540.
- Kumar, A., Imran-Siddiqi, M., Miertus, S., 2009. New molecular scaffolds for the design of Mycobacterium tuberculosis type II dehydroquinase inhibitors identified using ligand and receptor based virtual screening. J. Mol. Model. 16, 693–712.
- Lewis, S.N., Bassaganya-Riera, J., Bevan, D.R., 2009. Virtual screening as a technique for PPAR modulator discovery. PPAR Res. 2010, 861238.
- Li, H., Gao, Z., Kang, L., Zhang, H., Yang, K., Yu, K., Luo, X., Zhu, W., Chen, K., Shen, J., Wang, X., Jiang, H., 2006. TarFisDock: a web server for

identifying drug targets with docking approach. Nucleic Acids Res. 34, W219–W224.

- Lugnier, C., 2006. Cyclic nucleotide phosphodiesterase (PDE) superfamily: a new target for the development of specific therapeutic agents. Pharmacol. Therapeut. 109, 366–398.
- Lyne, P.D., 2002. Structure-based virtual screening: an overview. Drug Discov. Today 7, 1047–1055.
- Matsuki, H., Fujimoto, N., Iwata, K., Knauper, V., Okada, Y., Hayakawa, T., 1996. A one-step sandwich enzyme immunoassay for human matrix metalloproteinase 8 (neutrophil collagenase) using monoclonal antibodies. Clin. Chim. Acta 244, 129–143.
- Mimura, J., Fujii-Kuriyama, Y., 2003. Functional role of AhR in the expression of toxic effects by TCDD. Biochim. Biophys. Acta 1619, 263–268.
- Mimura, J., Yamashita, K., Nakamura, K., Morita, M., Takagi, T.N., Nakao, K., Ema, M.,Sogawa, K., Yasuda, M., Katsuki, M., Fujii-Kuriyama, Y., 1997. Loss of teratogenic response to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in mice lacking the Ah (dioxin) receptor. Genes Cells 2, 645–654.
- Murphy, S.D., 1986. Toxic effects of pesticides in Casarett and Doull's toxicology. In: Klaassen, Curtis.D., Amdur, Mary.O., John, Doull (Eds.), The Basic Science of Poisons, third ed. Macmillan Publishing Company, New York.
- Navratil, T., and L. Minarik. 2011. Trace Elements and Contaminants. In Earth System - History and Natural Variability. . Vol. 8. V. Cilek, Smith, R. H. , editor. EOLSS Publishers, Oxford.
- Nukaya, M., Moran, S., Bradfield, C.A., 2009. The role of the dioxin-responsive element cluster between the Cyp1a1 and Cyp1a2 loci in aryl hydrocarbon receptor biology. Proc. Natl. Acad. Sci. USA 106, 4923–4928.
- Nwomeh, B.C., Liang, H.X., Cohen, I.K., Yager, D.R., 1999. MMP-8 is the predominant collagenase in healing wounds and nonhealing ulcers. J. Surg. Res. 81, 189–195.
- Palackal, N.T., Lee, S.H., Harvey, R.G., Blair, I.A., Penning, T.M., 2002. Activation of polycyclic aromatic hydrocarbon trans-dihydrodiol proximate carcinogens by human aldo–keto reductase (AKR1C) enzymes and their functional overexpression in human lung carcinoma (A549) cells. J. Biol. Chem. 277, 24799–24808.

- Pandini, A., Soshilov, A.A., Song, Y., Zhao, J., Bonati, L., Denison, M.S., 2009. Detection of the TCDD binding-fingerprint within the Ah receptor ligand binding domain by structurally driven mutagenesis and functional analysis. Biochemistry 48, 5972–5983.
- Park, S.J., Yoon, W.K., Kim, H.J., Son, H.Y., Cho, S.W., Jeong, K.S., Kim, T.H., Kim, S.H., Kim, S.R., Ryu, S.Y., 2005. 2,3,7,8-Tetrachlorodibenzo-p-dioxin activates ERK and p38 mitogen-activated protein kinases in RAW 264.7 cells. Anticancer Res. 25, 2831–2836.
- Park, J.H., Hahn, E.J., Kong, J.H., Cho, H.J., Yoon, C.S., Cheong, S.W., Oh, G.S., Youn, H.J., 2003. TCDD-induced apoptosis in EL-4 cells deficient of the aryl hydrocarbon receptor and down-regulation of IGFBP-6 prevented the apoptotic cell death. Toxicol. Lett. 145, 55–68.
- Pauwels, A., Schepens, P.J.,D´ Hooghe, T., Delbeke, L., Dhont, M., Brouwer, A., Weyler, J., 2001. The risk of endometriosis and exposure to dioxins and polychlorinated biphenyls: a case-control study of infertile women. Human Reprod. 16, 2050– 2055.
- Penning, T.M., Drury, J.E., 2007. Human aldo–keto reductases: function, gene regulation, and single nucleotide polymorphisms. Arch. Biochem. Biophys. 464, 241–250.
- Pérez-Nueno, V.I., Pettersson, S., Ritchie, D.W., Borrell, J.I., Teixidó, J., 2009. Discovery of novel HIV entry inhibitors for the CXCR4 receptor by prospective virtual screening. J. Chem. Inf. Model. 49, 810–823.
- Peters, J.M., M.G. Narotsky, G. Elizondo, P.M. Fernandez-Salguero, F.J. Gonzalez, and B.D. Abbott. 1999. Amelioration of TCDD-induced teratogenesis in aryl hydrocarbon receptor (AhR)-null mice. Toxicological sciences: an official journal of the Society of Toxicology. 47:86-92.
- Pohjanvirta, R., Tuomisto, J., 1994. Short-term toxicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin in laboratory animals: effects, mechanisms, and animal models. Pharmacol. Rev. 46, 483–549.
- Porpora, M.G., Medda, E., Abballe, A., Bolli, S., De Angelis, I., di Domenico, A., Ferro, A., Ingelido, A.M., Maggi, A., Panici, P.B., De Felip, E., 2009. Endometriosis and organochlorinated environmental pollutants: a case-control study on Italian women of reproductive age. Environ. Health Perspect. 117, 1070–1075.
- Ratjen, F., Hartog, C.M., Paul, K., Wermelt, J., Braun, J., 2002. Matrix metalloproteases in BAL fluid of patients with cystic fibrosis and their modulation by treatment with dornase. Thorax 57, 930–934.

- Sahlberg, C., Peltonen, E., Lukinmaa, P.L., Alaluusua, S., 2007. Dioxin alters gene expression in mouse embryonic tooth explants. J. Dent. Res. 86, 600–605.
- Sato, S., Shirakawa, H., Tomita, S., Ohsaki, Y., Haketa, K., Tooi, O., Santo, N., Tohkin, M., Furukawa, Y., Gonzalez, F.J., Komai, M., 2008. Low-dose dioxins alter gene expression related to cholesterol biosynthesis, lipogenesis, and glucose metabolism through the aryl hydrocarbon receptor-mediated pathway in mouse liver. Toxicol. Appl. Pharm. 229, 10–19.
- Schecter, A., and T.A. Gasiewicz. 2003. Dioxins and Health. John Wiley & Sons, Inc., Hoboken, New Jersey.
- Shetty, K., Paliyath, G., Pometto, A., Levin, R.E., 2006. Food Biotechnology. CRC Press, Taylor & Francis Group, Cleveland.
- Shimada, T. 2006. Xenobiotic-metabolizing enzymes involved in activation and detoxification of carcinogenic polycyclic aromatic hydrocarbons. Drug metabolism and pharmacokinetics. 21:257-276.
- Smialowicz, R.J., DeVito, M.J., Williams, W.C., Birnbaum, L.S., 2008. Relative potency based on hepatic enzyme induction predicts immunosuppressive effects of a mixture of PCDDS/PCDFS and PCBS. Toxicol. Appl. Pharm. 227, 477–484.
- Stahura, F.L., Bajorath, J., 2005. New methodologies for ligand-based virtual screening. Curr. Pharm. Design. 11, 1189–1202.
- Stockwell, B.R., 2000. Chemical genetics: ligand-based discovery of gene function. Nat. Rev. Genet. 1, 116–125.
- Swanson, H.I., Bradfield, C.A., 1993. The AH-receptor: genetics, structure and function. Pharmacogenetics 3, 213–230.
- Tanos, R., R.D. Patel, I.A. Murray, P.B. Smith, and G.H. Perdew. 2012. Ah receptor regulates the cholesterol biosynthetic pathway in a dioxin response element-independent manner. Hepatology (Baltimore, Md.). 55:1994-2004.
- Thackaberry, E.A., Jiang, Z., Johnson, C.D., Ramos, K.S., Walker, M.K., 2005. Toxicogenomic profile of 2,3,7,8-tetrachlorodibenzo-p-dioxin in the murine fetal heart: modulation of cell cycle and extracellular matrix genes. Toxicol. Sci. 88, 231–241.
- Torphy, T.J., 1998. Phosphodiesterase isozymes: molecular targets for novel antiasthma agents. Am. J. Resp. Crit. Care 157, 351–370.
- Urban, P., Pelclová, D., Lukás, E., Kupka, K., Preiss, J., Fenclová, Z., Smerhovsky´ , Z., 2007. Neurological and neurophysiological

examinations on workers with chronic poisoning by 2,3,7,8-TCDD: follow-up 35 years after exposure. Eur. J. Neurol. 14, 213–218.

- Van den Berg, M., Birnbaum, L.S., Denison, M., De Vito, M., Farland, W., Feeley, M., Fiedler, H., Hakansson, H., Hanberg, A., Haws, L., Rose, M., Safe, S., Schrenk, D., Tohyama, C., Tritscher, A., Tuomisto, J., Tysklind, M., Walker, N., Peterson, R.E., 2006. The 2005 World Health Organization reevaluation of human and Mammalian toxic equivalency factors for dioxins and dioxin-like compounds. Toxicol. Sci. 93, 223–241.
- Van den Berg, M., Birnbaum, L., Bosveld, A.T., Brunström, B., Cook, P., Feeley, M., Giesy, J.P., Hanberg, A., Hasegawa, R., Kennedy, S.W., Kubiak, T., Larsen, J.C., van Leeuwen, F.X., Liem, A.K., Nolt, C., Peterson, R.E., Poellinger, L., Safe, S., Schrenk, D., Tillitt, D., Tysklind, M., Younes, M., Waern, F., Zacharewski, T., 1998. Toxic equivalency factors (TEFs) for PCBs, PCDDs, PCDFs for humans and wildlife. Environ. Health Perspect. 106, 775–792.
- Van Themsche, C., Potworowski, E.F., ST-Pierre, Y., 2004. Stromelysin-1 (MMP-3) is inducible in T lymphoma cells and accelerates the growth of lymphoid tumors *in vivo*. Biochem. Biophys. Res. Commun. 315, 884–891.
- Villano, C.M., Murphy, K.A., Akintobi, A., White, L.A., 2006. 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) induces matrix metalloproteinase (MMP) expression and invasion in A2058 melanoma cells. Toxicol. Appl. Pharm. 210, 212–224.
- WHO. 2014. Fact sheet N°225. Dioxins and their effects on human health. http://www.who.int/mediacentre/factsheets/fs225/en/.
- Wittsiepe, J., Kullmann, Y., Schrey, P., Selenka, F., Wilhelm, M., 2000. Myeloperoxidase-catalyzed formation of PCDD/F from chlorophenols. Chemosphere 40, 963–968.
- Wolber, G., Langer, T., 2005. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. J. Chem. Inf. Model. 45, 160–169.
- Yoshizawa, K., Walker, N.J., Jokinen, M.P., Brix, A.E., Sells, D.M., Marsh, T., Wyde, M.E., Orzech, D., Haseman, J.K., Nyska, A., 2005. Gingival carcinogenicity in female harlan sprague-dawley rats following two-year oral treatment with 2,3,7,8-tetrachlorodibenzo-p-dioxin and dioxin-like compounds. Toxicol. Sci. 83, 64–77.

- Yu, J., Wang, Y., Zhou, W.H., Wang, L., He, Y.Y., Li, D.J., 2008. Combination of estrogen and dioxin is involved in the pathogenesis of endometriosis by promoting chemokine secretion and invasion of endometrial stromal cells. Hum. Reprod. 23, 1614–1626.

CHAPTER 4.
Pubchem fishing as a source of chemical scaffolds for dengue virus protease inhibitors.

4.1. Introduction

Dengue is the fastest spreading viral disease that is transmitted by mosquitoes (*Aedes aegypti*, mainly) worldwide (Shroyer, 1986); with two-fifths of the world population at risk for becoming infected (WHO, 2014). Otherwise, dengue fever is a viral infectious disease transmitted to humans through one of the five serotypes of dengue virus, members of genus *Flavivirus,* which are genetically similar but antigenically distinct (Mustafa *et al.*, 2015). Its clinical manifestations range from asymptomatic to dengue shock syndrome, which may cause death. According to the World Health Organization, approximately 3.6 million people are living in endemic areas prone to developing the disease, and at least 2.4 million cases were reported annually in 2010, 2013 and 2015 (WHO, 2016). However, estimates made by cartographic approaches consider that about 390 million (95% credible interval) dengue infections occur annually (Bhatt *et al*., 2013). Until now, there is no vaccine against infection, for this reason current prevention methods are based on eliminating mosquito breeding sites. Urban growth, globalization and lifestyle changes have contributed to increase populations of *Ae. aegypti* and its geographical distribution (Gubler, 2014).

In tropical and sub-tropical countries, dengue is a major public health problem due to multiple factors, such as reemergence and intense transmission, the regular occurrences (every two or three years) of the disease in the same geographical site, the increase in the frequency of severe dengue outbreaks and co-circulation of the serotypes, causing an increase in the number of new cases. Of those people with severe dengue, relatively 10% of the total cases involved dengue hemorrhagic fever, and about 500,000 need to be hospitalized, and (Usme-Ciro *et al.*, 2008; WHO, 2013). Despite its global spread, currently there is no drug that is available for the disease.

Ecological studies of dengue virus in West Africa and Malaysia habitats have identified transmission cycles in non-human primates, with other species of the genus *Aedes* as vectors. This sylvatic cycle seems to be the ancestral condition: moving from non-human primates to humans and from arboreal species of *Aedes* to *Ae. aegypti* and *Ae. albopictus*, as a result of the increased human activity (Weaver and Barrett, 2004).

The first dengue viruses were isolated in 1943 during World War II, however, only until 1944, samples obtained from soldiers settled in India, New Guinea and Hawaii were assigned as members of antigenically similar viruses, named dengue serotype 1 (DENV-1), with Hawaiian strain as the prototype. Also isolated from New Guinea, an antigenically different strain was obtained, known as serotype 2 (DENV-2). Subsequently, in 1956, were obtained two serotypes in the Philippines, known serotypes 3 and 4 (DENV-3 y DENV-4)(Mahy and Van Regenmortel, 2010). Finally, in 2013, was isolated the fifth serotype in the forests of Sarawak Malaysia (DENV-5) (Mustafa *et al.*, 2015).

4.1.1. Dengue disease

Dengue infections may be asymptomatic (especially in children under 15 years) or generate a series of clinical manifestations, even cause death. Population studies have shown an increase in the severity of the symptoms of dengue fever in patients proportional to age and repeated infections (Thai *et al.*, 2011). DENV-2 and DENV-3 have historically been the most prevalent agents related to disease severity (Balmaseda *et al.*, 2006). Infection with any of the DENV serotypes may affect various human organ systems, and can cause a wide variety of clinical manifestations ranging from undifferentiated febrile illness to fatal febrile illness (WHO, 2009).

After a bite from an infected mosquito there is an incubation period of up to 2 weeks, symptoms develop would vary, but usually following three phases – the initial few days (3 to 7 days), known as febrile phase; a critical phase from about 4-5 days from the onset of fever (39 - 40 °C), followed by a phase of spontaneous recovery. It is associated with the entry of the virus into the bloodstream (viremia), along with some non-specific symptoms such as skin rash, generalized body pain, myalgia (muscle pain), arthralgia and headache; anorexia, nausea and vomiting; petechiae and mucosal bleeding, among other less common signs. A positive tourniquet test and a progressive decrease in the total white blood cell count should alert to a high probability of dengue. The critical stage coincides with the plasma extravasation, often accompanied by hemorrhagic manifestations, organ disorders (hepatitis, encephalitis or myocarditis) and hematological abnormalities, being shock, the most severe manifestation. This latter case occurs when a critical volume of plasma is lost, accompanied by a decrease in temperature below normal

levels; and a pulse pressure ≤ 20 mmHg. As a result, hematocrit (red blood cells proportion of the total blood volume) rises at this stage and platelets that were already descending reach their lowest values. In the recovery stage, usually the patient improvement is evident by absorption of extravascular fluid. Hematocrit stabilizes and white blood cell count increases, followed by the rise in platelet count. In some cases, the patient may deteriorate with some warning signs, such as vomiting, mucosal bleeding, high hematocrit and low platelet count. Cases of dengue with warning signs can be treated with intravenous rehydration and usually recover. Otherwise, dengue infections will deteriorate to severe dengue. It is classified as severe dengue if during the course of the disease, the patient shows any of the following symptoms: loss of plasma causing shock, accumulation of fluid with or without respiratory distress, severe bleeding and severe deterioration of organs (WHO, 2009).

4.1.2. Virion structure

Viruses of the Flaviviridae family have an RNA genome size of approximately 11 kb in a positive sense chain, flanked by untranslated regions (5'-UTR and 3'-UTR). The Flavivirus genome presents a methylated cap type I (m7GpppAmp) at the 5' end but lacks the polyadenylated tail at the 3' end. Over 95% of the RNA genome consists of a single open reading frame (ORF) and encodes a single polyprotein which is processed by viral and host proteases to produce individual proteins (Figure 4.1). Gene products are translated in the following order NH_3-C, prM, E, NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5-COOH, where C is the capsid, prM is the pre-membrane, and E is the envelope protein, while NS corresponds to nonstructural proteins. The genome is encapsulated by the capsid protein. This core-capsid is surrounded by a lipid bilayer containing the membrane (M) protein and envelope (E) protein (Lindenbach and Rice, 2003).

The virion is a spherical particle of 48 to 50 nm in diameter with a core of about 30 nm, surrounded by a lipid envelope. The structural proteins are found on the surface of viral particles. Glycoprotein E, main antigenic determinant on the virus particles, is involved in binding steps and viral entry via membrane fusion. The M protein is a small proteolytic fragment of the precursor protein prM produced during maturation of viral particles in the secretory pathway (Zhang *et al.*, 2003). The

nucleocapsid contains the capsid protein (C) interacting with the genomic RNA (Ma *et al.*, 2004).

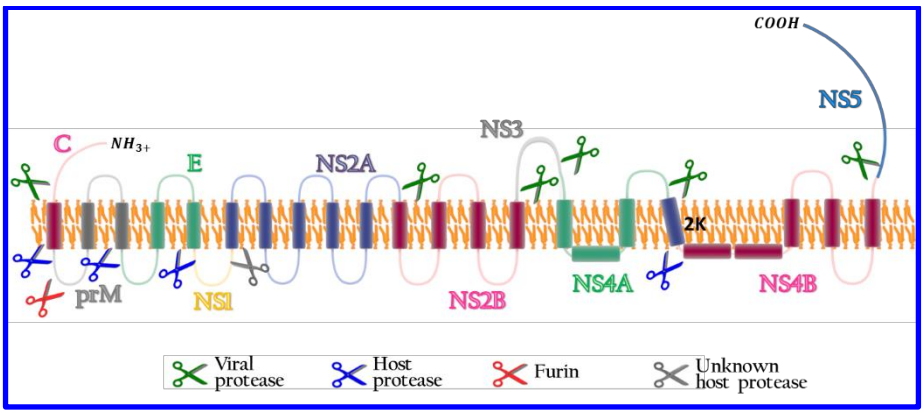


Figure 4.1. Dengue virus polyprotein organization at the endoplasmic reticulum membrane and processing by cellular and viral proteases. Modified from Assenberg *et al.* (2009).

4.1.3. Dengue virus infection

To enter host cells, the virus undergoes endocytosis mediated by binding of the glycoprotein E to cellular receptors (Figure 4.2), such as heparan sulfate in kidney cells, DC-SIGN/L-SIGN in dendritic cells, mannose receptor in macrophages, among others (Hidari and Suzuki, 2011). The low pH in endosomes, allows the rearrangement in the conformation of the glycoprotein, exhibiting the fusion loop and allowing the fusion of the virus envelope with the cell membrane. Once in the cytoplasm, the genome is translated by cell ribosomes. Host proteases and virus are responsible for splitting up each one of the 10 proteins (Figure 4.1). Together with it, the genome is replicated by nonstructural proteins to be packaged by capsid proteins and form new virions in the lumen of the endoplasmic reticulum (ER). When leaving the ER, a membrane is obtained with proteins M (immature, prM) and E, inserted. At this stage, viruses are in an immature form; the maturation occurs just before leaving the cell, when part of the prM protein is cleaved by furin, a cellular protein (Smit *et al.*, 2011).

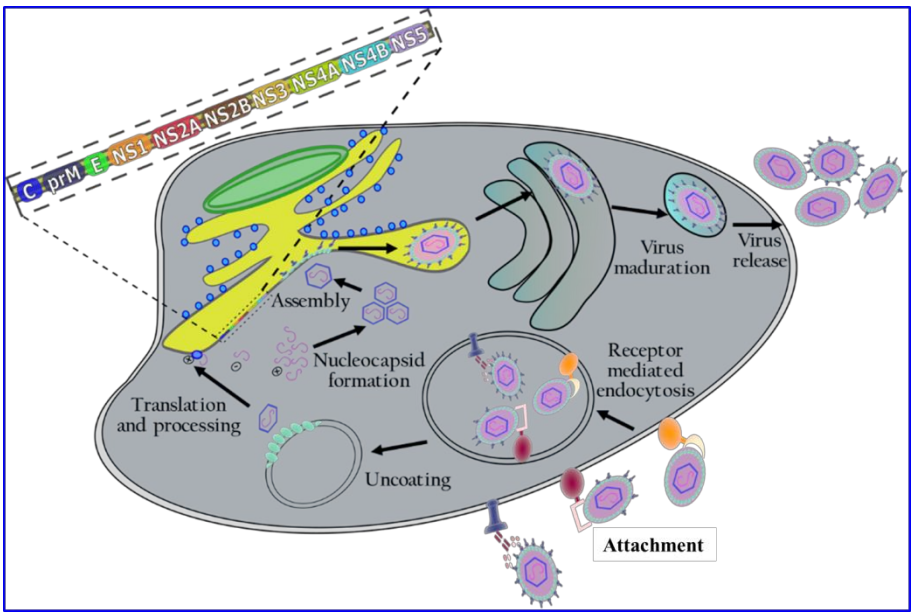


Figure 4.2. Cellular infection by dengue virus. The process begins with the **Attachment** of the virus to host cell receptors; the entry into the cell is mediated via a **Receptor mediated endocytosis**. Virion conformational changes due to vesicular acidification causes the **Uncoating** of viruses and the fusion of the viral and endocytic membrane, releasing the nucleocapsid into the cytoplasm. At the endoplasmic reticulum (ER), the viral RNA is translated into a polyprotein and processed (**Translation and processing**). The viral RNA is transcribed, amplified and packaged by capsid protein for **Nucleocapsid formation**. Subsequently, the **Assembly** of viral components into immature virions occurs in the ER. After its passage through the Golgi apparatus, **Virus maturation** takes place and it is released to extracellular environment (**Virus release**). Modified from Screaton *et al.* (2015).

The functional roles of nonstructural proteins during the viral life cycle are relatively well characterized and discussed below. NS1 is a multifunctional glycoprotein, involved in viral replication, although its role in this process is not known. When secreted, induces host immune responses (Chen *et al.*, 2015). Studies in other flavivirus have shown that NS2A interacts with NS3, NS5 and 3 'UTR (Mackenzie *et al.*, 1998), besides, this protein has shown an important role in downregulating the response of type I interferon to infection, the initial host defense against viruses (Liu *et al.*, 2006). NS2B is an integral membrane protein; a

hydrophilic segment between residues 49-95 plays a central role as an essential cofactor for the NS3 serine protease activity for splitting up the natural polyprotein substrate (Yon *et al.*, 2005). The multifunctional protein NS3 includes a catalytic triad composed of serine in the N-terminal, which in complex with the cofactor NS2B functions as viral protease, required for polyprotein processing; the C-terminal region has helicase activity, essential during replication of the viral genome, plus a nucleoside triphosphatase activity (NTPase) and RNA 5'-triphosphatase (RTPasa) (Matusan *et al.*, 2001), with which is believed to dephosphorylate the 5' end of genomic RNA prior to addition of m7GpppAmp (Wengler and Wengler, 1993). Thereby, NS3 is the main axis in the dengue viral replication process. The C-terminal hydrophobic region of the NS4A protein acts as a signal sequence for translocation of NS4B in the ER lumen (2K). The latter corresponds to an integral membrane protein, which may be involved in replication efficiency in both the mosquito and human hosts (Hanley *et al.*, 2003). Also has been demonstrated its ability to block the signal transduction cascade by IFN-α / β, interfering with induced STAT1 phosphorylation (Munoz-Jordan *et al.*, 2005). Finally, the NS5, also a multifunctional protein, is involved in RNA replication. In the N-terminal has cap-processing activity and at the C-terminal, RNA-dependent RNA polymerase activity (RdRp). Its N-terminal region shares homology with S-adenosyl methionine (SAM) dependent transferases and it has been shown that it transfers methyl groups from SAM for both methylation events (N-7 guanine and ribose 2'-O) required for cap formation at the 5' end of RNA (Ray *et al.*, 2006). Besides, it contains two NLS, importin β binding domain and a protein-protein interacting site for NS3 (Brooks *et al.*, 2002). Because it has been demonstrated that NS5 induces the secretion of IL-8 in the late infection, apparently nuclear translocation of NS5 can modulate transcription of cytokines such as IL-8 to counteract the antiviral effects of innate immunity (Medin *et al.*, 2005). It has also been shown that interactions between proteins NS3 and NS5 improve RTPasa and NTPase activities of NS3 (Yon *et al.*, 2005).

Several authors have focused their efforts looking for inhibitors of viral proteins for treating dengue, blocking their biological roles (Gu *et al.*, 2014; Lou *et al.*, 2014). Even when envelope protein (Schmidt *et al.*, 2012), capsid (Byrd *et al.*, 2013a), protease (Raut *et al.*, 2015), helicase (Byrd *et al.*, 2013b), methyltransferase (Luzhkov *et al.*, 2013), and NS5

RNA-dependent RNA polymerase (Sampath and Padmanabhan, 2009) have been suggested to be therapeutic targets for specific antiviral searches against dengue, NS3 and NS5 have some advantageous features as being the most conserved in all the serotypes of dengue virus (Khan *et al.*, 2008), presenting multifunctional activity and also corresponding to the best characterized DENV nonstructural proteins. Given the importance of NS3 and NS5 proteins functions, both represent an important target for the development of specific antiviral inhibitors (Oliveira *et al.*, 2014).

NS2B/NS3 protease has been one of the most widely used targets, as shown by several authors (Ezgimen *et al.*, 2012; Kwo *et al.*, 2010; Yang *et al.*, 2014). It is clear that flavivirus proteases, including NS2B/NS3, are essential for viral replication and infectivity. Proof thereof is the fact that dengue virus infectivity decreases by 80% when infected cells were treated with peptide inhibitors against dengue protease (Rothan *et al.*, 2012). Moreover, this strategy has proved effective as treatment of other diseases, as in the case of protease inhibitors in HIV infection (Hughes *et al.*, 2011). Similarly, two hepatitis C virus protease inhibitors, Simeprevir (Forns *et al.*, 2014) and Sofosbuvir (Keating and Vaidya, 2014), which were recently approved for therapeutic use, have become the standard of care. Thus, the use of compounds that target the NS2B/NS3 protease becomes a prominent strategy for dengue virus therapy.

The use of bioinformatics tools, molecular modeling programs and high performance computing has been leading the process of designing and *in silico* searching for therapeutically useful molecules. This same approach has been applied to the discovery or development of compounds that could target several diseases (de Rezende *et al.*, 2013; Keri *et al.*, 2013). One of the strategies at the forefront of drug discovery is the virtual screening of databases (Reddy *et al.*, 2007), which can identify powerful new protein ligands (Ananthan *et al.*, 2009; Tuccinardi *et al.*, 2009). Using the structural information on the dengue virus protease and molecules published in PubChem (http://pubchem.ncbi.nlm.nih.gov), virtual screening was performed to search for low molecular weight molecules that could inhibit this protease and that could be potentially employed in the treatment of dengue infection, verifying their ability to inhibit the NS2B/NS3 protease activity of the dengue virus through *in vitro* assays. The discovery of

molecules that interfere with the proteolytic activity of the DENV NS2B/NS3 protease could become a potential alternative for the treatment and control of the disease.

For this purpose, after determining the best leading compounds, the complex NS2B/NS3 was obtained using recombinant DNA techniques. The inhibitory capacity against the proteolytic activity of the isolated complex was determined using the synthetic peptide Boc-Gly-Arg-Arg-AMC as a fluorescent substrate. Finally, the inhibitory effect of these compounds was investigated *in vitro,* using HepG2 cells that were infected with dengue virus serotype 2.

4.2. Materials and Methods

The methodology used in this study comprised the steps detailed below:

4.2.1. *In silico* studies

The dengue virus NS2B/NS3 protease crystal structure was obtained from the Protein Data Bank (PDB: 2FOM). This crystallographic structure has been used recently by several authors for *in silico* molecular docking approaches (Lai *et al.*, 2013; Noble and Shi, 2012; Pambudi *et al.*, 2013). The 3D structure of the protein was prepared using the SYBYL 8.1 software package (Tripos, 2008) by removing all the water molecules and substructures that were present. Furthermore, the side chain amide groups were fixed for each amino acid. The resulting structure was minimized using the Powell method of molecular mechanics, with the combined force fields Kollman united/Kollman All Atom, applying Amber charges and a gradient convergence criterion of 0.005 kcal/mol. The final structure was saved in PDB format. The program MGLTools 1.5.0 (Sanner *et al.*, 1996) was utilized to convert the PDB file to the PDBQT format, adding polar hydrogens for all of the docking calculations.

Three-dimensional structures of ligands that were uploaded into the NCBI PubChem Database from March/2005 to December/2011, were downloaded from this database (http://pubchem.ncbi.nlm.nih.gov/) in SDF format. Structures were optimized with the SYBYL 8.1.1 package (Tripos, 2008) with molecular mechanics methods, employing a Tripos force field, Gasteiger charges, gradient convergence of 0.01 kcal/mol and a maximum number of optimization iterations set to 1000. Optimized

structures were saved in MOL2 format, and converted to PDBQT files with MGLTools 1.5.0 (Sanner, 1999) for molecular docking procedures.

The AutoDock Vina 1.1 program (Trott and Olson, 2010) was employed for all of the fittings that were required for the virtual screening. The protein was maintained rigid throughout the docking process, while the ligands were allowed to be flexible. The molecular parameters used were the following: a grid spacing of 1.000 Å; box size dimensions of 22 × 30 × 32 (x, y, z); coordinates of x = 9.790 Å, y = 10.355 Å, and z = 11.150 Å at the center of the box; number of modes = 20; energy range = 1.5; and exhaustiveness = 20. The substrate-based inhibitor binding pocket shown in the crystallographic structure (PDB: 2FOM) was chosen for placing the grid box to cover the entire enzyme binding site and accommodate ligands to move freely. The interactions between the molecules and dengue virus protease were checked with LigandScout 2.0 (Wolber and Langer, 2005). This software extracts and interprets ligands and their macromolecular environment from a PDB file, displaying tridimensional and 2D images of the interactions.

Docking validation with biological data.

The docking protocol was validated using a test set of 40 molecules, including active and inactive reported inhibitors against dengue virus, as well as reported inhibitors for proteases from other viruses, for comparative purposes (Maldonado-Rojas *et al.*, 2015). For dengue inhibitors, molecules with IC_{50} values less than 15 μM were chosen. For some molecules, the three-dimensional structures were downloaded from the NCBI PubChem Database, and their processing was similar to that described above. Such is the case of the compounds CID_3010818, CID_53487990, CID_10324367, CID_35370, CID_5362440, CID_60825, CID_441243, CID_54682461, CID_44246257, CID_4101471, CID_60194816, and CID_23350827. For remaining ligands, 3D structure was drawn and optimized using DFT at the B3LYP/6-31G level. Calculations were carried out with Gaussian 09 package program (Frisch *et al.*, 2009). SYBYL 8.1.1 package was used to convert files to MOL2 format and MGLTools 1.5.0 to PDBQT format.

Molecules with the best binding affinities were compared to other chemicals that were published as active and inactive inhibitors of dengue protease. This comparison was based on their molecular characteristics using hierarchical agglomerative clustering. Cluster analysis was

performed using the Tanimoto coefficient method with a measure of similarity and an average linkage (Backman *et al.*, 2011).

The web-based software tool FAFDrugs3 (Lagorce *et al.*, 2011), which is hosted on the public domain of The Ressource Parisienne en Bioinformatique Structurale (http://fafdrugs2.mti.univ-paris-diderot.fr/references.html), was used to filter for Pan Assay Interference Compounds (PAINS). The canonical Simplified Molecular Input Line Entry System (SMILES) strings were used as Input data, which were obtained from the PubChem database.

4.2.2. *In vitro* studies

Protein expression

A soluble His-tagged recombinant DENV-2 NS2B/NS3 protease was constructed according to the sequence published by Yang *et al*. (2011). In this study, the vector pET28c was used for the inducible expression of the hexahistidine-tagged recombinant protein. The DENV-2 NS2B/NS3 protease sequence was inserted between the NdeI and XhoI restriction sites with a flexible linker (Gly4) between NS2B and NS3. The vector was then transformed into the *Escherichia coli* strain BL21 star (DE3) for protein expression. Cultures of this *E. coli* strain were grown in 5 L of LB medium that contained 30 μg/mL kanamycin at 37 °C until the absorbance at 600 nm (A_{600}) reached 0.6. The cells were induced for expression by the addition of isopropyl-β-D-thiogalactopyranose to a final concentration of 0.4 mM and were incubated for additional 4 h at 37°C. The cells were then harvested by centrifugation (8900 rpm, 20 min, 4 °C), and the pellets were stored at −80 °C until used. For protein purification, the cells were thawed and re-suspended in BugBuster® Protein Extraction Reagent. The re-suspended cells were stirred at 300 rpm for 20 minutes at room temperature. Then, they were centrifuged 30 min at 5,000 g and 4 °C. The supernatant was purified immediately by passage through a Ni^{+2} column and pre-equilibrated with 50 mM Na_2HPO_4 and 300 mM NaCl, pH 7.0. The column was extensively washed with buffer that contained 10 mM and 150 mM imidazole, and protein was then eluted from the column in buffer that contained 300 mM imidazole. Finally, the elution buffer was changed for storage buffer (Tris 125 mM, pH 7.5, NaCl 0,6 M, glycerol 25%). The elution fractions were analyzed by 12% SDS-PAGE. Samples of pre- and post-induced cells as well as soluble

and insoluble fractions following lysis were collected and also analyzed by 12% SDS-PAGE.

Fluorimetric enzyme activity assay

The bioassay used here is a modification of that reported by Tomlinson and Watowich (2011). The molecules used were those commercially available, with the best binding affinities obtained with the *in silico* studies: $C_{29}H_{25}N_3O_4S$ (PubChem ID: 54710332, STK841825, Purity > 90%), $C_{25}H_{14}O_6$ (PubChem ID: 54692153, STK523898, Purity > 90%), $C_{35}H_{27}NO_9$ (PubChem ID: 54681617, STK530661, Purity > 90%), $C_{30}H_{25}NO_5$ (PubChem ID: 54692801, STK332077, Purity > 90%), and $C_{34}H_{23}NO_7S_2$ (PubChem ID: 54715399, STK529754, Purity 95%), as well as aprotinin, a competitive serine protease inhibitor. The assay was performed in 96-well plates, with a final volume of 100 µL. The NS2B/NS3 dengue virus protease (0.1 µM) was pre-incubated in storage buffer (200 mM Tris, pH 9.5, 20% glycerol) at room temperature for 30 min. Small molecules were tested at 1, 2 and 10 µM. The reaction was initiated by the addition of the substrate Boc-Gly-Arg-Arg-AMC at 100 µM. The reaction progress was monitored continuously for 60 minutes. Fluorescence readings were recorded at 355 nm excitation and 460 nm emission. Triplicate measurements were conducted for each data point. The data are reported as the means ± S.E. The 50% inhibitory concentration (IC_{50}) values were estimated graphically from the plots of percent inhibition versus inhibitor concentration (Fabrick *et al.*, 2002). IC_{50} was defined as the inhibitor concentration at which 50% of the control protease activity was inhibited. All of the protease inhibitor assays were performed in triplicate.

Cell and Virus

The human hepatocellular liver carcinoma cell line, HepG-2 (from American Type Culture Collection [ATCC] HB-8065), was grown in DMEM/F-12 medium (GIBCO® Dulbecco's Modified Eagle Medium: Nutrient Mixture F-12) supplemented with 10% (v/v) fetal calf serum and maintained in a humidified 5% (v/v) CO_2 incubator at 37 °C. The DENV-2 New Guinea C strain (NGC) provided by the Center for Disease Control (CDC, Fort Collins, CO) was used in this study. Stock virus was prepared in *Ae. albopictus* (C6/36) cells, and the virus titer was determined by the conventional Reed & Muench method (Reed and Muench, 1938).

Cell proliferation assay

A MTT [3-(4, 5-dimethylthiazol-2-yl)-2, 5-di-phenyl tetrazolium bromide] cell proliferation assay was performed to estimate the potential cytotoxicity of the compounds on HepG-2 cells using the protocol described by Meneses *et al*. (2009). Cells were seeded into 96-well plates at $1x10^4$ cells per well and incubated for 24 h at 37°C; then, the culture medium was removed, and fresh medium with each compound at final concentrations of 200, 100, 50, 25 and 12.5 µM was added. After 72 h of incubation at 37°C, MTT solution (20 µL, 5 mg/mL, Sigma Co.) was added, and the cells were incubated for additional 4 h. The reduction of MTT to a formazan product was quantified by measuring the optical density (OD) at 580 nm using a microplate reader (Thermo Scientific). Data are expressed as the concentration of compound that reduces the OD_{580} of the treated cells to 50% compared to the untreated cells (50% Cytotoxic Concentration, CC50).

Virus yield reduction assay

Viral protein NS1 reduction assays were performed to examine the antiviral activity of the tested compounds. Secreted NS1 in the culture medium of virus-infected cells correlates with viral replication (Ludert *et al.*, 2008). Approximately $1x10^4$ HepG-2 cells per well were seeded in a 96-well plate, and after incubating for 24 h at 37°C, these cells were infected with DENV-2 at a multiplicity of infection (MOI) of 11.0. One hour after virus adsorption, the cells were washed and replenished with medium that contained concentrations of the compound below the Maximum Non-Toxic Concentration (MNTC). Samples of culture medium were collected at 72 h post infection, stored at -20°C, and later tested for the presence of DENV-2/NS1 protein. Assays were performed in duplicate in at least two independent experiments. Virus-infected cells, non-drug treated or treated with ribavirin, were run in parallel as negative and positive controls, respectively. The differences in the NS1 levels between the treated and untreated cells were tested for significance by Student's t-test. The IC_{50} and selectivity index (SI = CC_{50}/CI_{50}) were calculated.

NS1 protein capture assay

The Panbio® Dengue Early kit (PanBio Diagnostic Brisbane, Australia) is a sandwich format microplate enzyme immunoassay for the quantitative detection of dengue NS1 protein. Samples of culture medium

from DENV-2-infected cells treated and untreated with compound or ribavirin were tested according to the instructions of manufacturer. The presence of NS1 was demonstrated by a color development. The optical density (OD) reading was taken at 450 nm, and the amount of NS1 was determined by comparing the OD of the sample with the OD of the cut-off control cell medium. The results are expressed as ''Panbio units: PbU,'' with 11.0 defined as the minimum concentration that was detected (Shrivastava *et al.*, 2011).

Data Analysis

Mean comparisons between groups were performed using the Kruskal-Wallis test. Statistical analyses were performed using GraphPad Prism version 5.01 (GraphPad Software, San Diego, CA). The p-value that was considered to be significant is indicated in each Figure. Error bars are expressed as ± SD.

4.3. Results

This report describes the discovery of compounds that can inhibit NS2B/NS3 dengue virus protease, an essential enzyme for viral replication, which could provide pharmacological benefits for the medical treatment of this disease. It has been found that this enzyme has an essential role in dengue virus replication through the cleavage of viral polyprotein (Geiss *et al.*, 2009). Therefore, disturbing the activity of dengue protease will contribute to blocking the viral replication in host cells. The NS2B/NS3 two-component protease of dengue virus is currently being investigated with the purpose of discovering small molecule inhibitors that are suitable as treatment options for the disease with the same name (Raut *et al.*, 2015).

4.3.1. Screening with the PubChem Database

The use of the dengue virus protease structural information for *in silico* approaches allowed the identification of different compounds with good inhibitory activity for this protein. From the PubChem Database, 210,903 compounds were docked into the DENV protease model (PDB: 2FOM). The scoring of the *in silico* binding affinities ranged from -15.1 to 165.0 kcal/mol. Theoretically, a lower Gibbs free energy of a protein–ligand complex indicates more stability, and in addition, binding requires a negative Gibbs free energy, which is similar to any other spontaneous process (Bronowska, 2011). Of the total number of examined molecules,

113,849 showed equal or better binding affinity values than the average (-7.4 kcal/mol); these are molecules that are less than 2000 Da, and most have more H bond acceptors than donors.

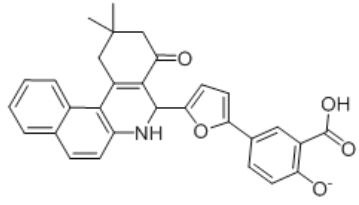
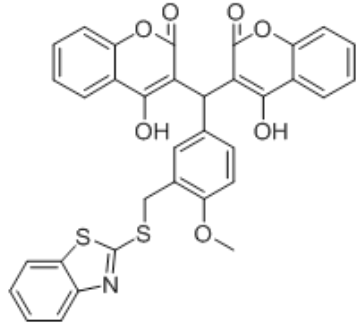
4.3.2. Selecting the inhibitors

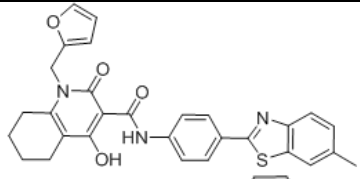
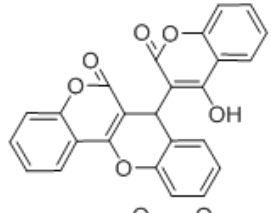
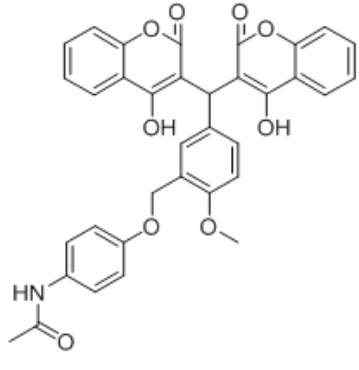
Among the molecules with the best docking scores, five were commercially available, purchased, and tested *in vitro* as dengue protease inhibitors (Table 4.1). All of them have molecular weights between 410.38 and 621.68 Da. For these molecules, the potential drug-like properties were calculated using the OSIRIS property calculator (http://www.organic-chemistry.org/prog/peo). This software suggests that except for $C_{30}H_{25}NO_5$ (CID 54692801), these small molecules do not contain the pharmacophores that are associated with increased risk of adverse mutagenic, tumorigenic or irritant effects. For $C_{30}H_{25}NO_5$ (CID 54692801), naphtalen-1-amine is considered to be a high risk fragment that indicates tumorigenicity and mutagenicity.

According to Bael and Holloway (2010), PAINS appear as frequent hits in high throughput screening (HTS). Some of the molecules with the best docking scores obtained from this type of assay involve false positives as a result of their high protein reactivity. The substructural features identified in promiscuous inhibitors were searched by FAFDRUGS3 (Lagorce *et al.*, 2011) (http://fafdrugs2.mti.univ-paris-diderot.fr/mobyle.html). None of the five hits reported in this report were described as PAINS, which corroborates that these compounds are promising starting points for further exploration in current dengue research.

Protein-ligand complexes for selected compounds are presented in Figure 4.3. Most of the interactions were hydrophobic and hydrogen bonding interactions, van der Waals forces, and a few aromatic ring interactions. Comparing the conformation adopted for these molecules in the binding site of the dengue protease, it is evidenced that certain residues interact with most of these molecules, in particular, Ile65, Trp69, Lys74, Leu76, Thr120, Ile123, Val154, Ala164, Ile165, sharing the same binding site.

Table 4.1. Purchased and tested compounds from the PubChem database with high *in silico* affinity values (kcal/mol) for the DENV-2 protease model (PDB: 2FOM).

PubChem ID	Structure	Molecular formula	Docking affinity scores (kcal/mol)	Molecular weight (g/mol)	log P	H-bond donors	H bond acceptors
54692801		$C_{30}H_{25}NO_5$	-13.5	478.52	5.90	2	6
54715399		$C_{34}H_{23}NO_7S_2$	-11.4	621.68	6.78	2	8

PubChem ID	Structure	Molecular formula	Docking affinity scores (kcal/mol)	Molecular weight (g/mol)	log P	H-bond donors	H bond acceptors
54710332		$C_{29}H_{25}N_3O_4S$	-11.7	511.59	5.18	2	7
54692153		$C_{25}H_{14}O_6$	-11.4	410.38	3.98	1	6
54681617		$C_{35}H_{27}NO_9$	-11.6	605.59	4.63	3	10

Given the high amino acid sequence homology between the NS3 proteases of the Flaviviridae family (Cregar-Hernandez et al., 2011), it is feasible that the inhibitors reported for other viral proteases might have an important inhibitory activity for this enzyme in dengue virus. For this reason, the crystal structure of the dengue NS2B/NS3 protease (PDB: 2FOM) was docked with some inhibitors that were reported for viral proteases such as, Hepatitis C (HCV), West Nile (WNV), Dengue and Human Immunodeficiency Virus (HIV) (Table 4.2). For dengue inhibitors, molecules with IC50 values of less than 15 μM were chosen. The affinity values obtained for these inhibitors were in the range from -10,1 kcal/mol to -5.5 kcal/mol. From these, the best values were those obtained for dengue virus protease inhibitors, as expected. However, the binding affinity values obtained with these molecules were not as good as those calculated for the molecules suggested in this report. The PDB crystal structure 2FOM was also used for docking with some molecules that were reported as not being active inhibitors for dengue protease (Aravapalli et al., 2012; Steuer et al., 2011; Tiew et al., 2012). The molecular structure of inhibitors that were reported for other proteases (HCV, WNV, Dengue and HIV), together with the molecules that were reported as not active inhibitors for dengue protease were docked with the crystal structure of the dengue NS2B/NS3 protease (PDB: 2FOM). The affinity values range varied from -9,0 kcal/mol to -6,2 kcal/mol (Table 4.2), which shows significant differences ($P<0.05$) from the values reported in inhibitors.

Cluster analysis allowed finding similarities between the molecular structures in Table 4.1 and the dengue inhibitors reported and shown in Table 4.2 (Figure 4.4). Four of the five compounds proposed for *in silico* results as dengue protease inhibitor candidates were grouped together, where the last cluster encompasses two of our identified molecules ($C_{34}H_{23}NO_7S_2$, CID 54715399 and $C_{35}H_{27}NO_9$, CID 54681617), which were the ones with the largest structure-based difference when they were compared to some of the reported inhibitory molecules. They both share the structure bis(4-hydroxy-2-oxo-2H-chromen-3-yl)(2-methoxyphenyl)methyl.

4.3.3. *In vitro* assays

The commercially available molecules that had a high theoretical binding affinity to NS2B/NS3 protease were analyzed *in vitro* for their

inhibition capacity (Figure 4.5). The percent of inhibition calculation was obtained by subtracting the experimental value (with inhibitor) from the control (without inhibitor), the resulting value is divided by the control and multiplied by one hundred (Umeadi *et al.*, 2008), by the following equation:

$$Inhibition\ (\%) = \frac{Control - Test}{Control} * 100$$

For all of the tested molecules, the higher the concentration, the higher the percentage of inhibition, this demonstrates a dose-dependent response (Figure 4.5). The IC_{50} values suggested activity against dengue virus protease for four of the tested compounds, with $C_{30}H_{25}NO_5$ (CID: 54692801) and $C_{34}H_{23}NO_7S_2$ (CID: 54715399) showing significant differences ($P<0.05$) compared to the negative control. The lowest activity against dengue virus protease was observed for $C_{29}H_{25}N_3O_4S$ (CID 54710332), which was apparently inactive, with an IC_{50} > 100 µM. The IC_{50} values for the others were between 9.09 µM and 19.87 µM, with $C_{34}H_{23}NO_7S_2$ (CID 54715399) having the highest inhibitory activity.

Evaluation of antiviral activity by using cell-based assays that include the viral inhibitory effect and cell toxicity

To exclude the possibility that the antiviral activities were due to compound-mediated cytotoxicity, uninfected HepG-2 cells were incubated with different concentrations of each compound for 72 h, and the cell viability was measured by the MTT method. Significant cell viability reduction was not observed at concentrations of up to 100.0, 85.7, 19.5, 13.5 and 12.5 µM for $C_{29}H_{25}N_3O_4S$ (CID 54710332), $C_{35}H_{27}NO_9$ (CID 54681617), $C_{30}H_{25}NO_5$ (CID 54692801), $C_{34}H_{23}NO_7S_2$ (CID 54715399) and $C_{25}H_{14}O_6$ (CID 54692153), respectively. Thus, these concentrations were considered to be non-toxic to the cells.

To verify the antiviral activity, the cells were infected with DENV-2 and treated with compounds at non-toxic concentrations. The NS1 viral titer reduction assay was performed in culture medium at 72 h p.i. The level of NS1 in the medium from infected cells that were untreated and treated with ribavirin were 44.5 ± 3.2 PbU and 8.3 ± 2.5 PbU, respectively. Only three compounds were active (Figure 4.6), two of them with a good inhibitory activity against dengue protease, as it was also observed in the

fluorometric enzyme assay. The $C_{30}H_{25}NO_5$, (CID 54692801) compound reduced NS1 by 69.8% (13.3 ± 3.3 PbU) at 19.5 µM and 15.5% (37.6 ± 3.7 PbU) at 9.8 µM; $C_{34}H_{23}NO_7S_2$ (CID 54715399) reduced NS1 by 73.9% (11.6 ± 0.8 PbU) at 13.5 µM with no reduction observed at lower concentrations; and the $C_{35}H_{27}NO_9$ (CID 54681617) compound reduced NS1 by 79.9% (8.9 ± 2.1 PbU) at 85.7 µM and 37.0% (28.2 ± 7.9 PbU) at 49.2 µM. In contrast, the $C_{25}H_{14}O_6$ (CID 54692153) and $C_{29}H_{25}N_3O_4S$ (CID 54710332) compounds did not reduce NS1 production at the highest non-toxic tested concentrations (12.5 µM and 100.0 µM, respectively) under the experimental conditions.

The relevant activity of synthetic compounds is generally expressed by IC_{50} and SI toward the supporting host (Table 4.3). An important parameter for the interpretation of the antiviral of a compound tested by using cell-based assays is the SI, which is calculated by the following formula:

$$SI = {}^{CC_{50}}/_{IC_{50}}$$

The $C_{30}H_{25}NO_5$ (CID 54692801) was the most active, followed by $C_{34}H_{23}NO_7S_2$ (CID 54715399), with both compounds having an IC_{50} value of less than 20 µM and an SI value of more than 3.0. The $C_{35}H_{27}NO_9$ (CID 54681617) compound was active at an IC_{50} of 61.5 ± 4.6 µM with an SI of 2.6. The results suggest that $C_{30}H_{25}NO_5$ (CID 54692801) and $C_{34}H_{23}NO_7S_2$ (CID 54715399) are compounds that exert promising anti-dengue activity through significant inhibition of DENV-2 replication *in vitro* in accordance with the values of IC_{50} and SI.

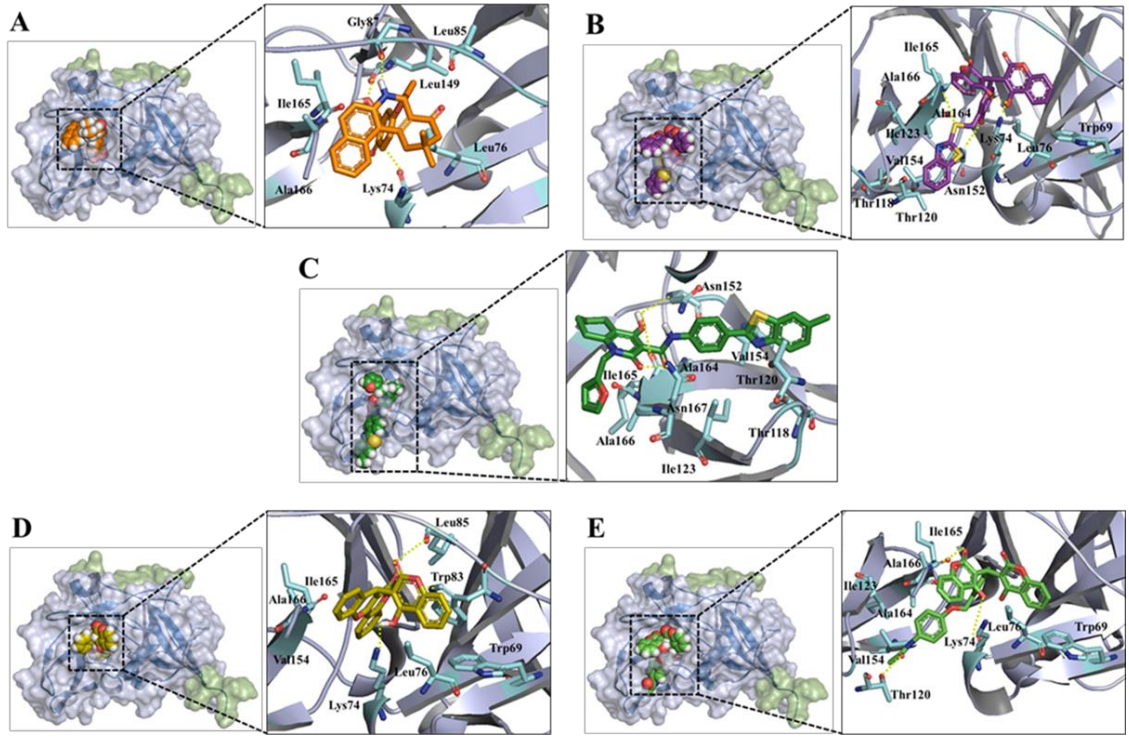


Figure 4.3. Tridimensional views of docking and interacting residues in the binding site cavity of dengue virus protease (PDB: 2FOM) in complex with A. $C_{30}H_{25}NO_5$, CID 54692801. B. $C_{34}H_{23}NO_7S_2$, CID 54715399. C. $C_{29}H_{25}N_3O_4S$, CID 54710332. D. $C_{25}H_{14}O_6$, CID 54692153. E. $C_{35}H_{27}NO_9$, CID 54681617.

Table 4.2. Affinity binding values for viral protease inhibitors and molecules reported as not active dengue virus protease inhibitors, docked with the DENV-2 protease crystal structure (PDB: 2FOM).

Code	Compound	Inhibited virus	Replicates of docking affinity with DENV2 (kcal/mol)			Mean	SD	Reference
			R1	R2	R3			
C1	CID_3010818	HCV[a]	-7,0	-7,0	-7,1	-7,0	0.1	(Gottwein *et al.*, 2011)
C2	CID_53487990	HCV	-6,7	-6,7	-6,7	-6,7	0.3	
C3	CID_10324367	HCV	-7,2	-7,3	-7,1	-7,2	0.1	
C4	CID_35370	HIV[b]	-6,3	-6,3	-6,3	-6,3	0.0	(Gulick *et al.*, 1997)
C5	CID_5362440	HIV	-7,8	-7,9	-8	-7,9	0.2	
C6	CID_60825	HIV	-5,6	-5,6	-5,4	-5,5	0.1	
C7	CID_441243	HIV	-8,1	-8,1	-8	-8,1	0.4	(Stefanidou *et al.*, 2013)
C8	CID_54682461	HIV	-7,1	-7	-8,8	-7,6	0.7	(Turner *et al.*, 1998)
C9	CID_44246257	WNV[c]	-8,2	-8,1	-7,7	-8,0	0.2	(Su *et al.*, 2009)
C10	N1,N4-bis(4-(4-methyl-4,5-dihydro-1H-imidazol-2-yl)phenyl)phthalazine-1,4-diamine	DENV	-9,4	-9,4	-9,3	-9,4	0.1	(Bodenreider *et al.*, 2009)
C11	CID 4101471	DENV	-9,1	-8,6	-9,1	-8,9	0.2	

Code	Compound	Inhibited virus	Replicates of docking affinity with DENV2 (kcal/mol)			Mean	SD	Reference
			R1	R2	R3			
C12	ARDP0006	DENV	-8	-8	-8	-8,0	0.0	(Tomlinson and Watowich, 2011)
C14	Chemdiv K286-0036	DENV	-8,6	-8,6	-8,6	-8,6	0,0	(Nguyen *et al.*, 2013)
C15	CID 60194816	DENV	-8,4	-7,8	-7,9	-8,0	0,3	(Deng *et al.*, 2012)
C16	CID 60194819	DENV	-8,9	-9,4	-8,3	-8,9	0,6	
C17	Compound 7b	DENV	-9,6	-9,7	-8,8	-9,4	0,5	
C18	Compound 7c	DENV	-9,8	-9,8	-9,1	-9,6	0,4	
C19	Compound 7e	DENV	-9,6	-9,5	-8,7	-9,3	0,5	(Liu *et al.*, 2014)
C20	Compound 8a	DENV	-9,4	-9,8	-9	-9,4	0,4	
C21	Compound 8b	DENV	-10,2	-10,2	-9,9	-10,1	0,2	
C22	Compound 8e	DENV	-9,3	-9,8	-9,3	-9,5	0,3	
C23	Compound 7a	NOT ACTIVE DENGUE VIRUS PROTEASE INHIBITORS	-9,2	-9,0	-8,9	-9,0	0,2	(Aravapalli *et al.*, 2012)
C24	Compound 8f		-9,8	-9,9	-10	-9,9	0,1	
C25	Compound 2		-6,3	-6,4	-6,2	-6,3	0,1	(Steuer *et al.*, 2011)
C26	Compound 4a		-6,7	-6,7	-7,6	-7,0	0,5	(Tiew *et al.*, 2012)
C27	Compound 4b		-8,0	-8,7	-8,8	-8,5	0,4	
C28	Compound 4c		-7,7	-8,6	-7,9	-8,1	0,5	

Code	Compound	Inhibited virus	Replicates of docking affinity with DENV2 (kcal/mol)			Mean	SD	Reference
			R1	R2	R3			
C29	Compound 4d		-7,5	-7,5	-7,2	-7,4	0,2	
C30	Compound 4e		-8,6	-8,2	-7,9	-8,2	0,4	
C31	Compound 4f		-8,1	-7,9	-7,8	-7,9	0,2	
C32	Compound 4g		-8,2	-7,9	-7,3	-7,8	0,5	
C33	Compound 4h		-7,8	-7,8	-7,5	-7,7	0,2	
C34	Compound 4i		-7,7	-7,8	-7,5	-7,7	0,2	
C35	Compound 4j		-8,1	-8,4	-8,5	-8,3	0,2	
C36	Compound 4k		-8,8	-7,9	-8,8	-8,5	0,5	
C37	Compound 4m		-8,4	-8,4	-9,0	-8,6	0,3	
C38	Compound 4n		-8,8	-8,3	-8,7	-8,6	0,3	
C39	Compound 4o		-7,3	-7,5	-7,4	-7,4	0,1	
C40	Compound 4p		-7,2	-7,0	-7,0	-7,1	0,1	

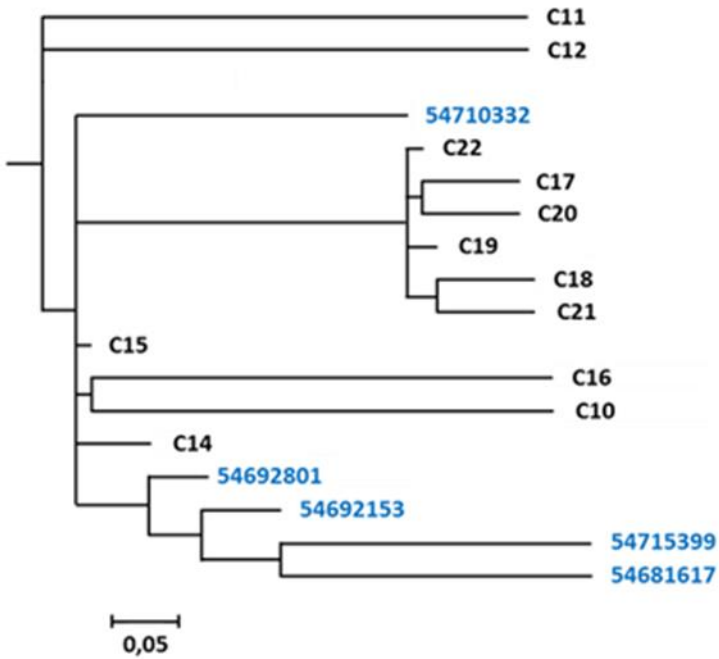


Figure 4.4. Dendrogram that shows the clustering between the structures of the *in vitro* tested molecules and some of the reported dengue virus NS2B/NS3 protease inhibitors. CIDs of the tested molecules are in blue.

Table 4.3. Relevant antiviral activities of active compounds on DENV-2 replication *in vitro*.

Compound	CC_{50} (μM)	IC_{50} (μM)	SI
$C_{30}H_{25}NO_5$ (CID: 54692801)	58.6 ± 3.0	14.9 ± 2.9	3.9
$C_{34}H_{23}NO_7S_2$ (CID: 54715399)	42.1 ± 1.6	11.8 ± 0.2	3.6
$C_{35}H_{27}NO_9$ (CID: 54681617)	162.4 ± 0.9	61.5 ± 4.6	2.6

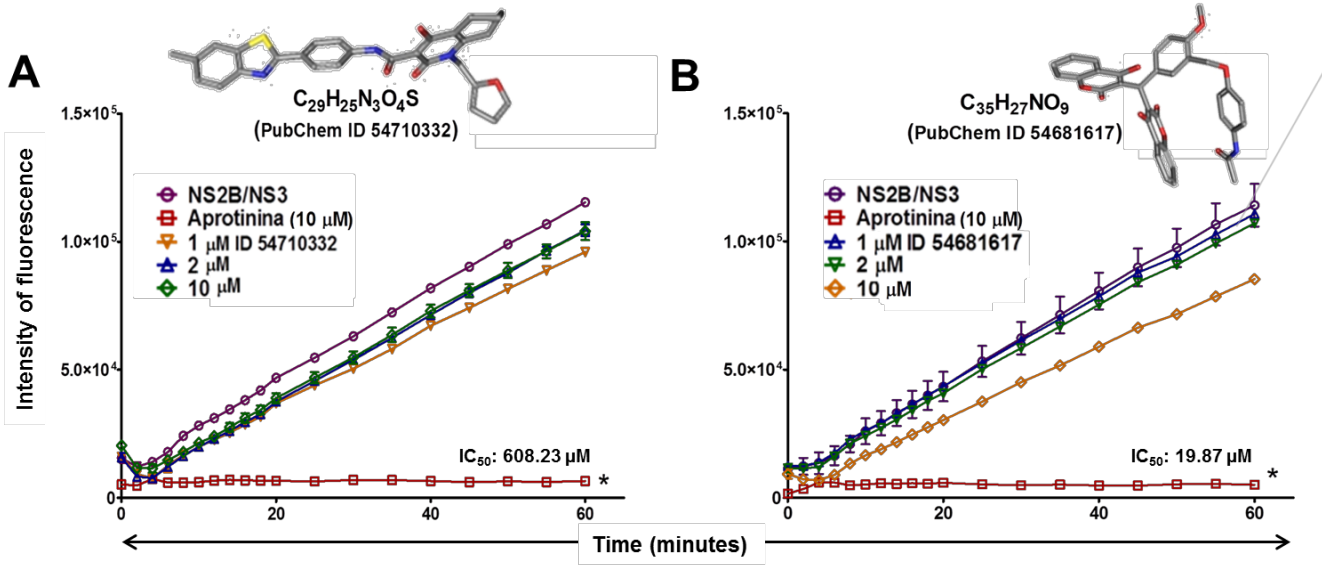
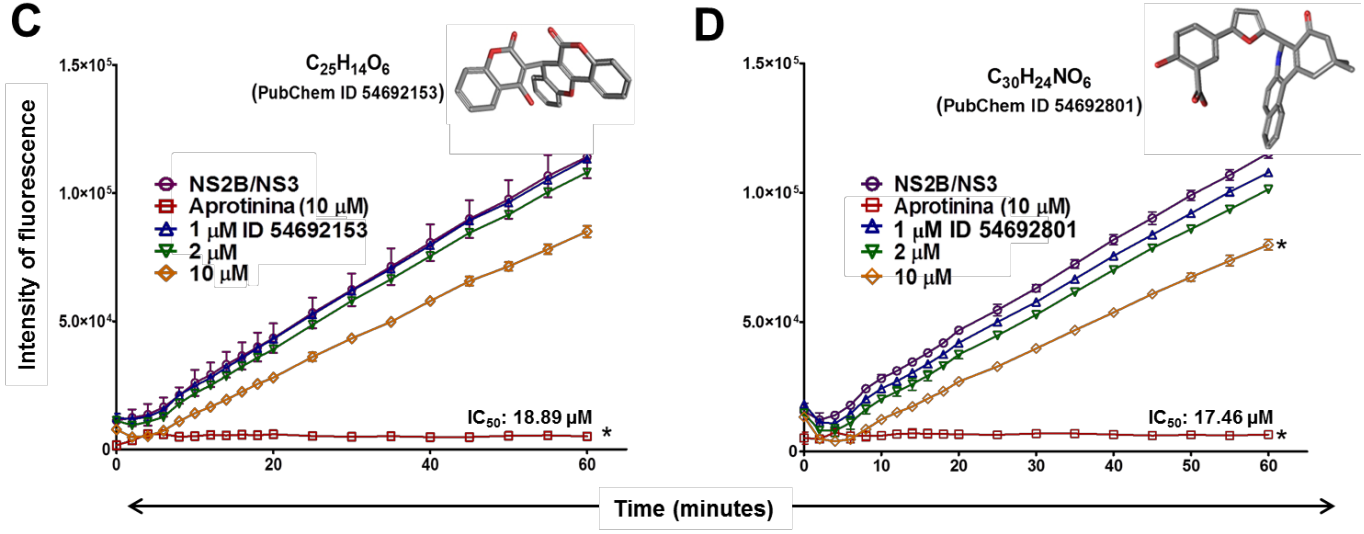
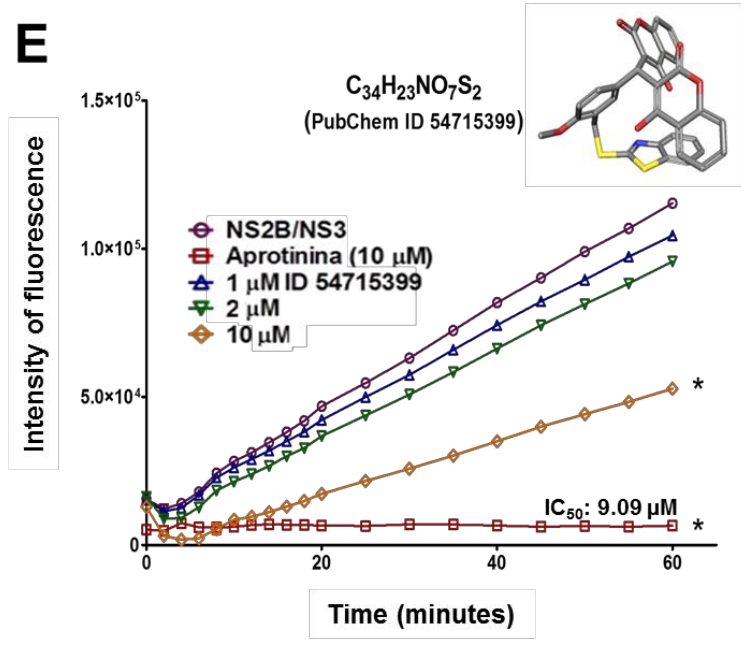


Figure 4.5. Percentage of inhibition of dengue virus protease activity for the *in vitro* tested compounds. A. $C_{29}H_{25}N_3O_4S$ (CID: 54710332). B. $C_{35}H_{27}NO_9$ (CID: 54681617).
* Significantly different ($P<0.05$) when compared to the negative control.



Cont. Figure 4.5. Percentage of inhibition of dengue virus protease activity for the *in vitro* tested compounds. C. $C_{25}H_{14}O_6$ (CID: 54692153). D. $C_{30}H_{25}NO_5$ (CID: 54692801).
* Significantly different ($P<0.05$) when compared to the negative control.



Cont. Figure 4.5. Percentage of inhibition of dengue virus protease activity for the *in vitro* tested compounds.
E. $C_{34}H_{23}NO_7S_2$, (CID 54715399).
* Significantly different ($P<0.05$) when compared to the negative control.

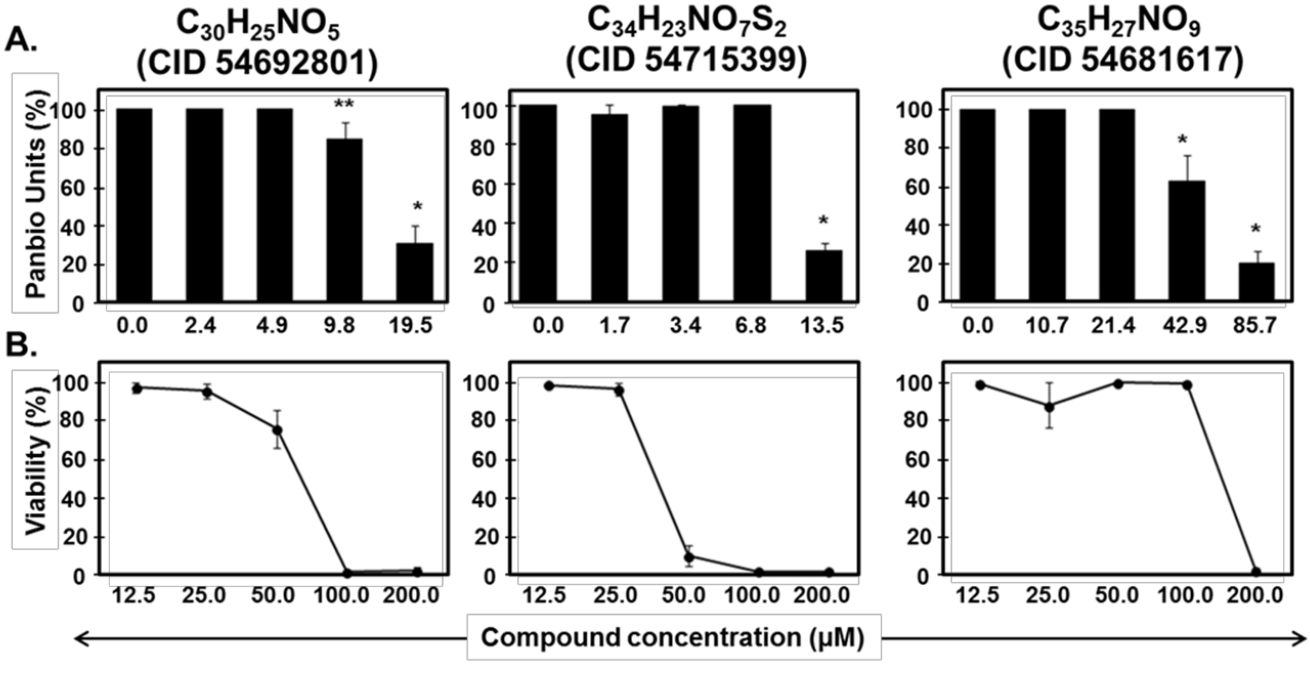


Figure 4.6. Inhibition of DENV-2 and cytotoxicity of the compounds. A. NS1 protein secretion monitoring (%PbU) from virus-infected HepG-2 cells, 72 h after non-toxic concentrations of compounds. The values indicate the percentages of PbU from the compound-treated virus-infected compared with those from untreated infected-cells. *,**. Significantly different $P < 0.01$ and $P < 0.05$, respectively, when compared to non-exposure. B. Compound-mediated cytotoxicity (cell viability percentage) derived from the compound-treated cells compared with that from the untreated cells.

4.4. Discussion

To date, several efforts have been made to find non-peptidic inhibitors for dengue virus protease. However, most of them do not possess the desired properties for drug candidate or the antiviral activity required. Until now, some of the potent known dengue protease inhibitors reported in literature (listed in Table 4.2) include some compounds with specific mode of inhibition for different dengue serotypes like C10 (Bodenreider *et al.*, 2009); others, that interact with conserved residues of the catalytic triad and the active site, decreasing the likelihood of drug-resistant mutations, like CID 4101471 and ARDP0006 (Tomlinson and Watowich, 2011) or Chemdiv K286-0036 (Nguyen *et al.*, 2013); or forming hydrogen bonds with the P4 region, a specificity determinant of dengue virus NS3 protease (Niyomrattanakit *et al.*, 2006), like CID 60194816 and CID 60194819 (Deng *et al.*, 2012) or finally compounds 7b, 7c, 7e, 8a, 8b and 8e with inhibitory rates over 75% and IC_{50} values lower than 5 μM (Liu *et al.*, 2014).

The results obtained in this work showed that the molecules have promising antiviral features against dengue. On the one hand, these molecules exhibited a range of values close to the required standards for the physicochemical parameters according to the rule of 5 (RO5) for drug-like compounds: molecular weight ≤ 500, logarithm of partition coefficient (log P) ≤ 5, H-bond donors ≤ 5, and H bond acceptors ≤ 10 (Lipinski, 2004). Their molecular weights are low, between 410.38 and 621.68, the range of log P is between 3,98 and 6,78, H-bond donors ≤ 3, and H bond acceptors ≤ 10 (Table 4.1). On the other hand, they do not have a risk of causing adverse effects, at least in theory for four of them, as shown by the OSIRIS calculator. Even for $C_{30}H_{25}NO_5$ (CID 54692801), because naphtalen-1-amine, which was the only fragment described as a non-drug-like molecular fragment because of its high toxicity risks of undesired effects, such as tumorigenicity and mutagenicity, could be replaced for drug-like molecular groups.

None of the examined compounds has any report as a dengue antiviral. Of these molecules, $C_{25}H_{14}O_6$ (CID 54692153) was reported by Nolan *et al*. (Nolan *et al.*, 2007) as a human NAD(P)H quinone oxidoreductase-1 (NQO1) inhibitor that was useful for cancer therapeutics. However, no biological information was found for the other four compounds.

As shown for the protein-ligand complexes (Figure 4.3), these molecules do not interact directly with residues of the catalytic triad (His51, Asp75, Ser135), their interactions are at positions close to them and also near to the residues that are reported as being located inside of the S1, S2 and S3 pockets (Noble *et al.*, 2012). Asn152 (identified as an interacting residue for $C_{34}H_{23}NO_7S_2$, CID 54715399 and $C_{29}H_{25}N_3O_4S$, CID 54710332; Figure 4.1) and Ile165 were also reported by Salaemae *et al*. (2010) as determinants of binding with the substrate in the active site of the dengue virus NS3 protease, as it was also noted by other authors (Qamar *et al.*, 2014; Tomlinson and Watowich, 2011; Tomlinson *et al.*, 2009). In addition, some of the interacting residues between dengue virus protease and these molecules, such as Lys 74, Gly 87, Leu 149, Asn 152, Ala 164, and Asn 167, were described by Senthilvel *et al*. (2013)(Senthilvel *et al.*, 2013) as key residues for the binding affinity of the flavonoid quercetin. This compound is one of the seven phenolic molecules that are found in the leaves of *Carica papaya*, which have inhibitory activity on NS2B/NS3 protease.

When comparing the affinity values obtained between the Flaviviridae family protease inhibitors and dengue virus protease, none showed significant inhibitory activity for this enzyme. The low affinity registered for the docking between the dengue protease structure and the protease inhibitors reported for other viruses could be explained by the altered specificity relative to other trypsin-like serine proteases. One of the reasons for this behavior could be related to the flatness of the dengue virus protease active site (Yang *et al.*, 2011). This flatness substantially hampers drug development against this target because conformational changes are required to allow the protein-ligand interaction.

The obtained results when comparing affinity binding values from molecules reported as dengue protease inhibitors (Bodenreider *et al.*, 2009; Deng *et al.*, 2012; Liu *et al.*, 2014; Nguyen *et al.*, 2013; Tomlinson and Watowich, 2011) and for those reported as not active inhibitors, validate the parameters used in the *in silico* experiments, which is supported by the statistically significant differences that are found between these groups of molecules.

Cluster analysis grouped four of the five compounds that were proposed for *in silico* results as dengue protease inhibitor candidates,

with $C_{34}H_{23}NO_7S_2$ (CID 54715399) and $C_{35}H_{27}NO_9$ (CID 54681617) as the compounds that had the largest structure-based differences when they were compared with some of the reported inhibitory molecules. They both share the structure bis(4-hydroxy-2-oxo-2H-chromen-3-yl)(2-methoxyphenyl)methyl. This molecule is an analogue of bis-coumarine, which is a known inhibitor of urease (Khan *et al.*, 2004; Zaheer ul *et al.*, 2008), as well as nucleotide pyrophosphatase phosphodiesterase-1 (Choudhary *et al.*, 2006). As proteases, ureases and pyrophosphatases are hydrolase enzymes. Ureases act on carbon-nitrogen bonds, but in linear amides other than peptide bonds that are involved in the nitrogen cycle (Avellaneda-Torres *et al.*, 2013). Pyrophosphatase hydrolyses and inorganic pyrophosphate is produced in a very large number of biochemical reactions in the cell, such as in DNA or protein biosynthesis (Oksanen, 2009).

A dose-dependent response was shown by all of the tested molecules, which was demonstrated by an increase in the percentage of inhibition as the concentration increases. The inhibitory activity of the compound $C_{34}H_{23}NO_7S_2$ should be noted, which at the highest tested concentration (10 μM) showed an inhibition of more than 50% of the dengue virus protease. These IC_{50} values are better than those reported for other inhibitors (Rothan *et al.*, 2012; Yang *et al.*, 2011), which is evidence that the proposed methodology is useful for the identification of compounds that have dengue virus protease inhibitory activity.

Using concentrations that are non-toxic to the cells of each tested molecule, three compounds showed inhibitory activity. The relevant antiviral activity of the active compounds was defined according to the IC_{50} and SI, as shown in Table 4.3. SI enables to distinguish between the antiviral activity of the compounds evaluated and the possible toxic effects on the cells (Vieira-Barros *et al.*, 2012). With respect to this consideration, the inhibitory activity is related to its ability to induce cell death (FDA, 2006). According to the SI value, it is established if the antiviral activity is not a consequence of the cytotoxic effect on the cells, which indicates the selectivity of the tested compounds on the cell lines assessed (Vieira-Barros *et al.*, 2012). According to Senthil Kumar *et al.* (2014), SI values greater than 3.0 are considered to indicate high selectivity of samples to the cell lines tested. Moreover, very well-known antiviral molecules have SI values lower than 5. For instance, Amantadine

against influenza A(H1N1)pdm09 virus (SI = 4) (Sokolova *et al.*, 2015) and Ribavirin against DENV2 (SI = 4.7) (Tseng *et al.*, 2014). However, in order to avoid false positives, being that for most scientists, promising antiviral compounds should have SI values greater than 10; we consider that the two higher SI values obtained represent a modest anti-dengue protease activity. Additionally, it is also established that the antiviral activity is relevant when the IC_{50} values are below 100 mg/mL (Cos *et al.*, 2006). Taken together, the $C_{30}H_{25}NO_5$ (CID: 54692801) and $C_{34}H_{23}NO_7S_2$ (CID: 54715399) compounds present a moderate anti-DENV-2 activity, which affects the functional activity of the NS2B/NS3 protease complex. $C_{35}H_{27}NO_9$ (CID: 54681617) was active, but it does not appear to be a promising chemical in accordance with the IC_{50} (61.5 μM) and IS (2.6) values. $C_{25}H_{14}O_6$ (CID: 54692153) appears to be very toxic to human hepatocytes since maximum non-toxic concentration obtained was 12.5 μM. In contrast, $C_{29}H_{25}N_3O_4S$ (CID: 54710332) exhibited low cytotoxic effects, although at a high concentration (100 μM), it was inactive.

Studies demonstrate that *in vitro* cytotoxicity is commonly cell-type specific and that cytotoxicity in one cell does not necessarily suggest cytotoxicity in another (Xia *et al.*, 2008). In addition, the inability of compounds to inhibit dengue can depend on the type of cell, and discrepancies could be attributed to differences in the membrane permeability or the molecule uptake into the host cell. In this study, hepatic cells were used because they are target cells for virus infection in dengue patients, which contributes to the reliability of the assay. On the other hand, the primary screening that uses DENV-2 could lead to an inaccurate assessment of the anti-dengue activity due to the possibility that compounds that do not inhibit this serotype could exhibit activity of varying degrees against other infections.

4.5. References

- Ananthan, S., W. Zhang, and J.V. Hobrath. 2009. Recent advances in structure-based virtual screening of G-protein coupled receptors. *The AAPS journal*. 11:178-185.
- Aravapalli, S., H. Lai, T. Teramoto, K.R. Alliston, G.H. Lushington, E.L. Ferguson, R. Padmanabhan, and W.C. Groutas. 2012. Inhibitors of Dengue virus and West Nile virus proteases based on the aminobenzamide scaffold. *Bioorg. Med. Chem.* 20:4140-4148.

- Assenberg, R., E. Mastrangelo, T.S. Walter, A. Verma, M. Milani, R.J. Owens, D.I. Stuart, J.M. Grimes, and E.J. Mancini. 2009. Crystal structure of a novel conformational state of the flavivirus NS3 protein: implications for polyprotein processing and viral replication. *Journal of virology*. 83:12895-12906.
- Avellaneda-Torres, L.M., L.M. Melgarejo, C.E. Narvaez-Cuenca, and J. Sanchez. 2013. Enzymatic activities of potato crop soils subjected to conventional management and grassland soils. *J. Plant Nutr. Soil Sci.* 13:301-312.
- Backman, T.W., Y. Cao, and T. Girke. 2011. ChemMine tools: an online service for analyzing and clustering small molecules. *Nucleic acids research*. 39:W486-491.
- Baell, J.B., and G.A. Holloway. 2010. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *Journal of medicinal chemistry*. 53:2719-2740.
- Balmaseda, A., S.N. Hammond, L. Perez, Y. Tellez, S.I. Saborio, J.C. Mercado, R. Cuadra, J. Rocha, M.A. Perez, S. Silva, C. Rocha, and E. Harris. 2006. Serotype-specific differences in clinical manifestations of dengue. *The American journal of tropical medicine and hygiene*. 74:449-456.
- Bhatt, S., P.W. Gething, O.J. Brady, J.P. Messina, A.W. Farlow, C.L. Moyes, J.M. Drake, J.S. Brownstein, A.G. Hoen, O. Sankoh, M.F. Myers, D.B. George, T. Jaenisch, G.R. Wint, C.P. Simmons, T.W. Scott, J.J. Farrar, and S.I. Hay. 2013. The global distribution and burden of dengue. *Nature*. 496:504-507.
- Bodenreider, C., D. Beer, T.H. Keller, S. Sonntag, D. Wen, L. Yap, Y.H. Yau, S.G. Shochat, D. Huang, T. Zhou, A. Caflisch, X.C. Su, K. Ozawa, G. Otting, S.G. Vasudevan, J. Lescar, and S.P. Lim. 2009. A fluorescence quenching assay to discriminate between specific and nonspecific inhibitors of dengue virus protease. *Analytical biochemistry*. 395:195-204.
- Bronowska, A.K. 2011. Thermodynamics of Ligand-Protein Interactions: Implications for Molecular Design. *In* Thermodynamics - Interaction Studies - Solids, Liquids and Gases. J.C. Moreno-Pirajan, editor. InTech.
- Brooks, A.J., A.V. Johansson M Fau - John, Y. John Av Fau - Xu, D.A. Xu Y Fau - Jans, S.G. Jans Da Fau - Vasudevan, and S.G. Vasudevan. 2002. The

interdomain region of dengue NS5 protein that binds to the viral helicase NS3 contains independently functional importin beta 1 and importin alpha/beta-recognized nuclear localization signals. *J Biol Chem.* 277:36399-36407.

- Byrd, C.M., D. Dai, D.W. Grosenbach, A. Berhanu, K.F. Jones, K.B. Cardwell, C. Schneider, K.A. Wineinger, J.M. Page, C. Harver, E. Stavale, S. Tyavanagimatt, M.A. Stone, R. Bartenschlager, P. Scaturro, D.E. Hruby, and R. Jordan. 2013a. A Novel Inhibitor of Dengue Virus Replication That Targets the Capsid Protein. *Antimicrobial Agents and Chemotherapy*. 57:15-25.
- Byrd, C.M., D.W. Grosenbach, A. Berhanu, D. Dai, K.F. Jones, K.B. Cardwell, C. Schneider, G. Yang, S. Tyavanagimatt, C. Harver, K.A. Wineinger, J. Page, E. Stavale, M.A. Stone, K.P. Fuller, C. Lovejoy, J.M. Leeds, D.E. Hruby, and R. Jordan. 2013b. Novel Benzoxazole Inhibitor of Dengue Virus Replication That Targets the NS3 Helicase. *Antimicrobial Agents and Chemotherapy*. 57:1902-1912.
- Cos, P., A.J. Vlietinck, D.V. Berghe, and L. Maes. 2006. Anti-infective potential of natural products: how to develop a stronger *in vitro* 'proof-of-concept'. *Journal of ethnopharmacology*. 106:290-302.
- Cregar-Hernandez, L., G.S. Jiao, A.T. Johnson, A.T. Lehrer, T.A. Wong, and S.A. Margosiak. 2011. Small molecule pan-dengue and West Nile virus NS3 protease inhibitors. *Antiviral chemistry & chemotherapy*. 21:209-217.
- Chen, J., M.M. Ng, and J.J. Chu. 2015. Activation of TLR2 and TLR6 by Dengue NS1 Protein and Its Implications in the Immunopathogenesis of Dengue Virus Infection. *PLoS pathogens*. 11:e1005053.
- Choudhary, M.I., N. Fatima, K.M. Khan, S. Jalil, S. Iqbal, and R. Atta Ur. 2006. New biscoumarin derivatives-cytotoxicity and enzyme inhibitory activities. *Bioorganic & medicinal chemistry*. 14:8066-8072.
- de Rezende, L.C., F. Fumagalli, M.S. Bortolin, M.G. de Oliveira, M.H. de Paula, V.F. de Andrade-Neto, and S. Emery Fda. 2013. *In vivo* antimalarial activity of novel 2-hydroxy-3-anilino-1,4-naphthoquinones obtained by epoxide ring-opening reaction. *Bioorganic & medicinal chemistry letters*. 23:4583-4586.
- Deng, J., N. Li, H. Liu, Z. Zuo, O.W. Liew, W. Xu, G.K. Chen, X. Tong, W. Tang, and J. Zhu. 2012. Discovery of novel small molecule inhibitors of dengue viral NS2B-NS3 protease using virtual screening and scaffold hopping. *J. Med. Chem.* 55:6278-6293.

- Ezgimen, M., H. Lai, N.H. Mueller, K. Lee, G. Cuny, D.A. Ostrov, and R. Padmanabhan. 2012. Characterization of the 8-hydroxyquinoline scaffold for inhibitors of West Nile virus serine protease. *Antiviral research*. 94:18-24.
- Fabrick, J., C. Behnke, T. Czapla, K. Bala, A.G. Rao, K.J. Kramer, and G.R. Reeck. 2002. Effects of a potato cysteine proteinase inhibitor on midgut proteolytic enzyme activity and growth of the southern corn rootworm, Diabrotica undecimpunctata howardi (Coleoptera: Chrysomelidae). *Insect biochemistry and molecular biology*. 32:405-415.
- FDA 2006. Guidance for Industry Antiviral Product Development — Conducting and Submitting Virology Studies to the Agency. http://www.fda.gov/OHRMS/DOCKETS/98fr/05d-0183-gdl0002-01.pdf (accessed April 2016).
- Forns, X., E. Lawitz, S. Zeuzem, E. Gane, J.P. Bronowicki, P. Andreone, A. Horban, A. Brown, M. Peeters, O. Lenz, S. Ouwerkerk-Mahadevan, J. Scott, G. De La Rosa, R. Kalmeijer, R. Sinha, and M. Beumont-Mauviel. 2014. Simeprevir with peginterferon and ribavirin leads to high rates of SVR in patients with HCV genotype 1 who relapsed after previous therapy: a phase 3 trial. *Gastroenterology*. 146:1669-1679 e1663.
- Frisch, M.J., G.W. Trucks, H.B. Schlegel, and *e. al.* 2009. Gaussian 09. Vol. Gaussian 09. Gaussian, Inc., Wallingford, CT, USA.
- Geiss, B.J., H. Stahla, A.M. Hannah, A.M. Gari, and S.M. Keenan. 2009. Focus on flaviviruses: current and future drug targets. *Future medicinal chemistry*. 1:327-344.
- Gottwein, J.M., T.K. Scheel, T.B. Jensen, L. Ghanem, and J. Bukh. 2011. Differential efficacy of protease inhibitors against HCV genotypes 2a, 3a, 5a, and 6a NS3/4A protease recombinant viruses. *Gastroenterology*. 141:1067-1079.
- Gu, W.G., D.T. Ip, S.J. Liu, J.H. Chan, Y. Wang, X. Zhang, Y.T. Zheng, and D.C. Wan. 2014. 1,4-Bis(5-(naphthalen-1-yl)thiophen-2-yl)naphthalene, a small molecule, functions as a novel anti-HIV-1 inhibitor targeting the interaction between integrase and cellular Lens epithelium-derived growth factor. *Chemico-biological interactions*. 213:21-27.
- Gubler, D.J. 2014. Dengue and Dengue Hemorrhagic Fever, Boston, MA.
- Gulick, R.M., J.W. Mellors, D. Havlir, J.J. Eron, C. Gonzalez, D. McMahon, D.D. Richman, F.T. Valentine, L. Jonas, and A. Meibohm. 1997.

Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy. *N. Engl. J. Med.* 337:734-739.

- Hanley, K.A., L.R. Manlucu, L.E. Gilmore, J.E. Blaney, Jr., C.T. Hanson, B.R. Murphy, and S.S. Whitehead. 2003. A trade-off in replication in mosquito versus mammalian systems conferred by a point mutation in the NS4B protein of dengue virus type 4. *Virology*. 312:222-232.
- Hidari, K.I., and T. Suzuki. 2011. Dengue virus receptor. Tropical medicine and health. 39:37-43.
- Hughes, P.J., E. Cretton-Scott, A. Teague, and T.M. Wensel. 2011. Protease Inhibitors for Patients With HIV-1 Infection: A Comparative Overview. *P & T : a peer-reviewed journal for formulary management*. 36:332-345.
- Keating, G.M., and A. Vaidya. 2014. Sofosbuvir: first global approval. *Drugs*. 74:273-282.
- Keri, R.S., C. Quintanova, S.M. Marques, A.R. Esteves, S.M. Cardoso, and M.A. Santos. 2013. Design, synthesis and neuroprotective evaluation of novel tacrine-benzothiazole hybrids as multi-targeted compounds against Alzheimer's disease. *Bioorganic & medicinal chemistry*. 21:4559-4569.
- Khan, A.M., O. Miotto, E.J. Nascimento, K.N. Srinivasan, A.T. Heiny, G.L. Zhang, E.T. Marques, T.W. Tan, V. Brusic, J. Salmon, and J.T. August. 2008. Conservation and variability of dengue virus proteins: implications for vaccine design. PLoS neglected tropical diseases. 2:e272.
- Khan, K.M., S. Iqbal, M.A. Lodhi, G.M. Maharvi, Z. Ullah, M.I. Choudhary, A.-u. Rahman, and S. Perveen. 2004. Biscoumarin: new class of urease inhibitors; economical synthesis and activity. *Bioorganic & medicinal chemistry*. 12:1963-1968.
- Kwo, P.Y., E.J. Lawitz, J. McCone, E.R. Schiff, J.M. Vierling, D. Pound, M.N. Davis, J.S. Galati, S.C. Gordon, N. Ravendhran, L. Rossaro, F.H. Anderson, I.M. Jacobson, R. Rubin, K. Koury, L.D. Pedicone, C.A. Brass, E. Chaudhri, and J.K. Albrecht. 2010. Efficacy of boceprevir, an NS3 protease inhibitor, in combination with peginterferon alfa-2b and ribavirin in treatment-naive patients with genotype 1 hepatitis C infection (SPRINT-1): an open-label, randomised, multicentre phase 2 trial. *Lancet*. 376:705-716.

- Lagorce, D., J. Maupetit, J. Baell, O. Sperandio, P. Tuffery, M.A. Miteva, H. Galons, and B.O. Villoutreix. 2011. The FAF-Drugs2 server: a multistep engine to prepare electronic chemical compound collections. *Bioinformatics*. 27:2018-2020.
- Lai, H., D. Dou, S. Aravapalli, T. Teramoto, G.H. Lushington, T.M. Mwania, K.R. Alliston, D.M. Eichhorn, R. Padmanabhan, and W.C. Groutas. 2013. Design, synthesis and characterization of novel 1,2-benzisothiazol-3(2H)-one and 1,3,4-oxadiazole hybrid derivatives: potent inhibitors of Dengue and West Nile virus NS2B/NS3 proteases. *Bioorganic & medicinal chemistry*. 21:102-113.
- Lindenbach, B.D., and C.M. Rice. 2003. Molecular biology of flaviviruses. *Advances in virus research*. 59:23-61.
- Lipinski, C.A. 2004. Lead- and drug-like compounds: the rule-of-five revolution. *Drug discovery today. Technologies*. 1:337-341.
- Liu, H., R. Wu, Y. Sun, Y. Ye, J. Chen, X. Luo, and X. Shen. 2014. Identification of novel thiadiazoloacrylamide analogues as inhibitors of dengue-2 virus NS2B/NS3 protease. *Bioorganic & medicinal chemistry*. 22:6344-6352.
- Liu, W.J., X.J. Wang, D.C. Clark, M. Lobigs, R.A. Hall, and A.A. Khromykh. 2006. A single amino acid substitution in the West Nile virus nonstructural protein NS2A disables its ability to inhibit alpha/beta interferon induction and attenuates virus virulence in mice. *Journal of virology*. 80:2396-2404.
- Lou, Z., Y. Sun, and Z. Rao. 2014. Current progress in antiviral strategies. *Trends in pharmacological sciences*. 35:86-102.
- Ludert, J.E., C. Mosso, I. Ceballos-Olvera, I. Ceballos-Olvera, and R.M. del Angel. 2008. Use of a commercial enzyme immunoassay to monitor dengue virus replication in cultured cells. *Virology journal*. 5:51.
- Luzhkov, V., E. Decroly, B. Canard, B. Selisko, and J. Åqvist. 2013. Evaluation of Adamantane Derivatives as Inhibitors of Dengue Virus mRNA Cap Methyltransferase by Docking and Molecular Dynamics Simulations. *Molecular Informatics*. 32:155-164.
- Ma, L., T.D. Jones Ct Fau - Groesch, R.J. Groesch Td Fau - Kuhn, C.B. Kuhn Rj Fau - Post, and C.B. Post. 2004. Solution structure of dengue virus capsid protein reveals another fold. *Proceedings of the National Academy of Sciences of the United States of America*. 101:3414-3419.

- Mackenzie, J.M., M.K. Khromykh Aa Fau - Jones, E.G. Jones Mk Fau - Westaway, and E.G. Westaway. 1998. Subcellular localization and some biochemical properties of the flavivirus Kunjin nonstructural proteins NS2A and NS4A. *Virology*. 245:203-215.
- Mahy, B.W., and M.H. Van Regenmortel. 2010. Desk Encyclopedia of Human and Medical Virology. Academic Press.
- Maldonado-Rojas, W., J. Olivero-Verbel, and Y. Marrero-Ponce. 2015. Computational fishing of new DNA methyltransferase inhibitors from natural products. *Journal of molecular graphics & modelling*. 60:43-54.
- Matusan, A.E., A.D. Pryor Mj Fau - Davidson, P.J. Davidson Ad Fau - Wright, and P.J. Wright. 2001. Mutagenesis of the Dengue virus type 2 NS3 protein within and outside helicase motifs: effects on enzyme activity and virus replication. *J Virol.* . 75:9633-9643.
- Medin, C.L., K.A. Fitzgerald, and A.L. Rothman. 2005. Dengue virus nonstructural protein NS5 induces interleukin-8 transcription and secretion. *Journal of virology*. 79:11053-11061.
- Meneses, R., R.E. Ocazionez, J.R. Martínez, and E.E. Stashenko. 2009. Inhibitory effect of essential oils obtained from plants grown in Colombia on yellow fever virus replication *in vitro*. *Ann. Clin. Microbiol. Antimicrob.* . 8 8.
- Munoz-Jordan, J.L., M. Laurent-Rolle, J. Ashour, L. Martinez-Sobrido, M. Ashok, W.I. Lipkin, and A. Garcia-Sastre. 2005. Inhibition of alpha/beta interferon signaling by the NS4B protein of flaviviruses. *Journal of virology*. 79:8004-8013.
- Mustafa, M.S., V. Rasotgi, S. Jain, and V. Gupta. 2015. Discovery of fifth serotype of dengue virus (DENV-5): A new public health dilemma in dengue control. *Medical journal, Armed Forces India*. 71:67-70.
- Nguyen, T.T., S. Lee, H.K. Wang, H.Y. Chen, Y.T. Wu, S.C. Lin, D.W. Kim, and D. Kim. 2013. *In vitro* evaluation of novel inhibitors against the NS2B-NS3 protease of dengue fever virus type 4. *Molecules*. 18:15600-15612.
- Niyomrattanakit, P., S. Yahorava, I. Mutule, F. Mutulis, R. Petrovska, P. Prusis, G. Katzenmeier, and J.E. Wikberg. 2006. Probing the substrate specificity of the dengue virus type 2 NS3 serine protease by using internally quenched fluorescent peptides. *The Biochemical journal*. 397:203-211.

- Noble, C.G., Y.L. Chen, H. Dong, F. Gu, S.P. Lim, W. Schul, Q.Y. Wang, and P.Y. Shi. 2010. Strategies for development of Dengue virus inhibitors. *Antiviral research*. 85:450-462.
- Noble, C.G., C.C. Seh, A.T. Chao, and P.Y. Shi. 2012. Ligand-bound structures of the dengue virus protease reveal the active conformation. *Journal of virology*. 86:438-446.
- Noble, C.G., and P.Y. Shi. 2012. Structural biology of dengue virus enzymes: towards rational design of therapeutics. *Antiviral research*. 96:115-126.
- Nolan, K.A., H. Zhao, P.F. Faulder, A.D. Frenkel, D.J. Timson, D. Siegel, D. Ross, T.R. Burke, Jr., I.J. Stratford, and R.A. Bryce. 2007. Coumarin-based inhibitors of human NAD(P)H:quinone oxidoreductase-1. Identification, structure-activity, off-target effects and *in vitro* human pancreatic cancer toxicity. *Journal of medicinal chemistry*. 50:6316-6325.
- Oksanen, E. 2009. Enzyme Molecular Choreography - Studies of Soluble InorganicPyrophosphatases. Institute of Biotechnology and Department of Chemistry, Laboratory of Organic Chemistry. Faculty of Science, University of Helsinki, Helsinki.
- Oliveira, A.S.D., M.L.D. Silva, A.F. Oliveira, C.C.D. Silva, R.R. Teixeira, and S.O. De Paula. 2014. NS3 and NS5 proteins: important targets for anti-dengue drug design. *Journal of the Brazilian Chemical Society*. 25:1759-1769.
- Pambudi, S., N. Kawashita, S. Phanthanawiboon, M.D. Omokoko, P. Masrinoul, A. Yamashita, K. Limkittikul, T. Yasunaga, T. Takagi, K. Ikuta, and T. Kurosu. 2013. A small compound targeting the interaction between nonstructural proteins 2B and 3 inhibits dengue virus replication. *Biochem Biophys Res Commun.* 440:393-398.
- Qamar, T.U., A. Mumtaz, U.A. Ashfaq, S. Azhar, T. Fatima, M. Hassan, S.S. Hussain, W. Akram, and S. Idrees. 2014. Computer Aided Screening of Phytochemicals from Garcinia against the Dengue NS2B/NS3 Protease. *Bioinformation*. 10:115-118.
- Raut, R., H. Beesetti, P. Tyagi, I. Khanna, S.K. Jain, V.U. Jeankumar, P. Yogeeswari, D. Sriram, and S. Swaminathan. 2015. A small molecule inhibitor of dengue virus type 2 protease inhibits the replication of all four dengue virus serotypes in cell culture. *Virology journal*. 12:1-7.
- Ray, D., M. Shah A Fau - Tilgner, Y. Tilgner M Fau - Guo, Y. Guo Y Fau - Zhao, H. Zhao Y Fau - Dong, T.S. Dong H Fau - Deas, Y. Deas Ts Fau -

Zhou, H. Zhou Y Fau - Li, P.-Y. Li H Fau - Shi, and P.Y. Shi. 2006. West Nile virus 5'-cap structure is formed by sequential guanine N-7 and ribose 2'-O methylations by nonstructural protein 5. *J Virol.* 80:8362-8370.

- Reddy, A.S., S.P. Pati, P.P. Kumar, H.N. Pradeep, and G.N. Sastry. 2007. Virtual screening in drug discovery -- a computational perspective. *Current protein & peptide science*. 8:329-351.
- Reed, L.J., and H. Muench. 1938. A simple method of estimating fifty percent endpoints. *American journal of epidemiology*. 27:493-497.
- Rothan, H.A., H.C. Han, T.S. Ramasamy, S. Othman, N.A. Rahman, and R. Yusof. 2012. Inhibition of dengue NS2B-NS3 protease and viral replication in Vero cells by recombinant retrocyclin-1. *BMC infectious diseases*. 12:314.
- Salaemae, W., M. Junaid, C. Angsuthanasombat, and G. Katzenmeier. 2010. Structure-guided mutagenesis of active site residues in the dengue virus two-component protease NS2B-NS3. *Journal of biomedical science*. 17:68.
- Sampath, A., and R. Padmanabhan. 2009. Molecular targets for flavivirus drug discovery. *Antiviral research*. 81:6-15.
- Sanner, M.F. 1999. Python: a programming language for software integration and development. *Journal of molecular graphics & modelling*. 17:57-61.
- Sanner, M.F., A.J. Olson, and J.C. Spehner. 1996. Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers*. 38:305-320.
- Screaton, G., J. Mongkolsapaya, S. Yacoub, and C. Roberts. 2015. New insights into the immunopathology and control of dengue virus infection. *Nature Reviews*. 15.
- Schmidt, A.G., K. Lee, P.L. Yang, and S.C. Harrison. 2012. Small-molecule inhibitors of dengue-virus entry. *PLoS pathogens*. 8:e1002627.
- Senthil Kumar, M., K. Selvam, R. Singaravel, and N. Krishnaveni. 2014. Regulation of Operative Biomarkers Production by Treating with Marine Actinomycetes L-Asparaginase in HepG2 Cell Line. *Journal of Applied & Environmental Microbiology*. 2:74-80.
- Senthilvel, P., P. Lavanya, K.M. Kumar, R. Swetha, P. Anitha, S. Bag, S. Sarveswari, V. Vijayakumar, S. Ramaiah, and A. Anbarasu. 2013.

Flavonoid from Carica papaya inhibits NS2B-NS3 protease and prevents Dengue 2 viral assembly. *Bioinformation*. 9:889-895.

- Shrivastava, A., P.K. Dash, N.K. Tripathi, A.K. Sahni, N. Gopalan, and P.V. Lakshmana Rao. 2011. Evaluation of a commercial Dengue NS1 enzyme-linked immunosorbent assay for early diagnosis of dengue infection. *Indian journal of medical microbiology*. 29:51-55.
- Shroyer, D.A. 1986. *Aedes* albopictus and arboviruses: a concise review of the literature. *Journal of the American Mosquito Control Association*. 2:424-428.
- Smit, J.M., B. Moesker, I. Rodenhuis-Zybert, and J. Wilschut. 2011. Flavivirus cell entry and membrane fusion. *Viruses*. 3:160-171.
- Sokolova, A.S., O.I. Yarovaya, A.V. Shernyukov, Y.V. Gatilov, Y.V. Razumova, V.V. Zarubaev, T.S. Tretiak, A.G. Pokrovsky, O.I. Kiselev, and N.F. Salakhutdinov. 2015. Discovery of a new class of antiviral compounds: camphor imine derivatives. *European journal of medicinal chemistry*. 105:263-273.
- Stefanidou, M., I. Ramos, V. Mas-Casullo, J.B. Trepanier, S. Rosenbaum, A. Fernandez-Sesma, and B.C. Herold. 2013. Herpes simplex virus 2 (HSV-2) prevents dendritic cell maturation, induces apoptosis, and triggers release of proinflammatory cytokines: potential links to HSV-HIV synergy. *J. Virol.* 87:1443-1453.
- Steuer, C., C. Gege, W. Fischl, K.H. Heinonen, R. Bartenschlager, and C.D. Klein. 2011. Synthesis and biological evaluation of alpha-ketoamides as inhibitors of the Dengue virus protease with antiviral activity in cell-culture. *Bioorg. Med. Chem. Lett.* 19:4067-4074.
- Su, X.C., K. Ozawa, H. Yagi, S.P. Lim, D. Wen, D. Ekonomiuk, D. Huang, T.H. Keller, S. Sonntag, A. Caflisch, S.G. Vasudevan, and G. Otting. 2009. NMR study of complexes between low molecular mass inhibitors and the West Nile virus NS2B-NS3 protease. *FEBS J*. 276:4244-4255.
- Thai, K.T., H. Nishiura, P.L. Hoang, N.T. Tran, G.T. Phan, H.Q. Le, B.Q. Tran, N.V. Nguyen, and P.J. de Vries. 2011. Age-specificity of clinical dengue during primary and secondary infections. PLoS neglected tropical diseases. 5:e1180.
- Tiew, K.C., D. Dou, T. Teramoto, H. Lai, K.R. Alliston, G.H. Lushington, R. Padmanabhan, and W.C. Groutas. 2012. Inhibition of Dengue virus and West Nile virus proteases by click chemistry-derived benz[d]isothiazol-3(2H)-one derivatives. *Bioorg. Med. Chem.* 20:1213-1221.

- Tomlinson, S.M., and S.J. Watowich. 2011. Anthracene-based inhibitors of dengue virus NS2B-NS3 protease. *Antiviral research*. 89:127-135.
- Tomlinson, T.D., D.E. Huber, C.A. Rieth, and E.J. Davelaar. 2009. An interference account of cue-independent forgetting in the no-think paradigm. *Proceedings of the National Academy of Sciences of the United States of America*. 106:15588-15593.
- Tripos. 2008. SYBYL 8.1. Tripos, 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA.
- Trott, O., and A.J. Olson. 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*. 31:455-461.
- Tseng, C.H., C.K. Lin, Y.L. Chen, C.Y. Hsu, H.N. Wu, C.K. Tseng, and J.C. Lee. 2014. Synthesis, antiproliferative and anti-dengue virus evaluations of 2-aroyl-3-arylquinoline derivatives. *European journal of medicinal chemistry*. 79:66-76.
- Tuccinardi, T., S. Taliani, M. Bellandi, F. Da Settimo, E. Da Pozzo, C. Martini, and A. Martinelli. 2009. A virtual screening study of the 18 kDa translocator protein using pharmacophore models combined with 3D-QSAR studies. *ChemMedChem*. 4:1686-1694.
- Turner, B.J., L. Markson, J. Cocroft, L. Cosler, and W.W. Hauck. 1998. Clinic HIV-focused features and prevention of Pneumocystis carinii pneumonia. *J. Gen. Intern. Med.* 13:16-23.
- Umeadi, C., K. Bentsi-Barnes, F. Kandeel, and I.H. Al-Abdullah. 2008. Endogenous pancreatic protease activity and methods for impeding their function. Transplantation proceedings. 40:355-357.
- Usme-Ciro, J.A., J.A. Mendez, A. Tenorio, G.J. Rey, C. Domingo, and J.C. Gallego-Gomez. 2008. Simultaneous circulation of genotypes I and III of dengue virus 3 in Colombia. *Virology journal*. 5:101.
- Vieira-Barros, A., L. Mendes-Araujo, F.F. de Oliveira, A. Oliveira da Conceiçao, I.C. Simoni, M.J. Bittencourt-Fernandes, and C. Weis-Arns. 2012. *In vitro* evaluation of the antiviral potential of Guettarda angelica against animal herpesviruses. *Acta Sci. Veterinariae*. 40:1068.
- Weaver, S.C., and A.D. Barrett. 2004. Transmission cycles, host range, evolution and emergence of arboviral disease. *Nature reviews. Microbiology*. 2:789-801.

- Wengler, G., and G. Wengler. 1993. The NS 3 nonstructural protein of flaviviruses contains an RNA triphosphatase activity. *Virology*. 197:265-273.
- WHO. 2009. Dengue: guidelines for diagnosis, treatment, prevention and control. World Health Organization. World Health Organization. .
- WHO. 2013. Dengue and Severe Dengue. Fact Sheet No. 117. http://www.who.int/mediacentre/factsheets/fs117/en/ (accessed September 2015).
- WHO. 2014. World Health Day 2014: Small Bite, Big Threat. http://www.who.int/campaigns/world-health-day/2014/en/ (accessed September 2015).
- WHO. 2016. Fact sheet. Dengue and severe dengue. http://www.who.int/mediacentre/factsheets/fs117/en/ (accessed May, 2016).
- Wolber, G., and T. Langer. 2005. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *Journal of chemical information and modeling*. 45:160-169.
- Xia, T., M. Kovochich, M. Liong, L. Madler, B. Gilbert, H. Shi, J.I. Yeh, J.I. Zink, and A.E. Nel. 2008. Comparison of the mechanism of toxicity of zinc oxide and cerium oxide nanoparticles based on dissolution and oxidative stress properties. *ACS nano*. 2:2121-2134.
- Yang, C.C., Y.C. Hsieh, S.J. Lee, S.H. Wu, C.L. Liao, C.H. Tsao, Y.S. Chao, J.H. Chern, C.P. Wu, and A. Yueh. 2011. Novel dengue virus-specific NS2B/NS3 protease inhibitor, BP2109, discovered by a high-throughput screening assay. *Antimicrob Agents Chemother*. 55:229-238.
- Yang, C.C., H.S. Hu, R.H. Wu, S.H. Wu, S.J. Lee, W.T. Jiaang, J.H. Chern, Z.S. Huang, H.N. Wu, C.M. Chang, and A. Yueh. 2014. A novel dengue virus inhibitor, BP13944, discovered by high-throughput screening with dengue virus replicon cells selects for resistance in the viral NS2B/NS3 protease. *Antimicrob Agents Chemother*. 58:110-119.
- Yon, C., T. Teramoto, N. Mueller, J. Phelan, V.K. Ganesh, K.H. Murthy, and R. Padmanabhan. 2005. Modulation of the nucleoside triphosphatase/RNA helicase and 5'-RNA triphosphatase activities of Dengue virus type 2 nonstructural protein 3 (NS3) by interaction with NS5, the RNA-dependent RNA polymerase. *The Journal of biological chemistry*. 280:27412-27419.

- Zaheer ul, H., M.A. Lodhi, S. Ahmad Nawaz, S. Iqbal, K. Mohammed Khan, B.M. Rode, R. Atta ur, and M.I. Choudhary. 2008. 3D-QSAR CoMFA studies on bis-coumarine analogues as urease inhibitors: a strategic design in anti-urease agents. *Bioorganic & medicinal chemistry*. 16:3456-3461.
- Zhang, W., P.R. Chipman, J. Corver, P.R. Johnson, Y. Zhang, S. Mukhopadhyay, T.S. Baker, J.H. Strauss, M.G. Rossmann, and R.J. Kuhn. 2003. Visualization of membrane protein domains by cryo-electron microscopy of dengue virus. *Nature structural biology*. 10:907-912.

CHAPTER 5.
FINAL REMARKS AND CONCLUSIONS

The mining of protein-ligand coupling interactions performed in this thesis allowed the exploration of alternative ways by which TCDD may be affecting cellular targets through AhR-independent pathways, as well as the identification of new antidengue protease structures for therapeutic strategies. Results suggest that several proteins, some linked to effects of TCDD exposure, may, in one way or another, work as direct molecular targets for binding this chlorinated compound. Although different reports have shown AhR-independent activities of TCDD, the cellular targets involved in these cases are still unknown. This thesis presented some data showing that several proteins can, at least theoretically, have suitable binding sites for TCDD, and these direct effects should also be considered when studying its toxicity mechanisms. On the other hand, a similar protein-ligand interaction approach was employed to discover anti-dengue molecules with potential to inhibit the enzymatic activity of NS2B/NS3 dengue serine protease. The moiety bis (4-hydroxy-2-oxo-2H-chromen-3-yl) (2-methoxyphenyl) methyl, could be used as a starting point for developing new drug candidates with advantageous properties as a dengue virus protease inhibitor. Therefore, the data generated in this work could be a useful guide for *in vitro* and *in vivo* screening processes with a view to clarify the role of endogenous and exogenous compounds on cell dynamics and physiology.

Summarizing, the following contributions were made during this work:

- Mine activity based on ligand and macromolecular structures is an effective way to identify targets of environmental pollutants, as a first step to elucidate their biochemical pathways; as well as to work as a source of new scaffolds for therapeutic agents in drug design.

- Several protein targets have the potential to bind TCDD, evidencing the possible existence of AhR-independent mechanisms that affect numerous processes in the body.
- Rheumatoid arthritis, periodontal diseases and endometrosis have been observed as a consequence of TCDD exposure. These illnesses are associated with the expression of metallopeptidases, which showed the highest binding score for TCDD.
- Other proteins likely targets for TCDD are oxidosqualene cyclase, and transforming growth factor beta type I, which have a well known role in some pathological effects of dioxin exposure, including changes on

colesterol homeostasis and effects on cell proliferation, cell differentiation, cell adhesion, cell migration, and regulation of extracellular matrix components.

- PubChem is a notorious source of inhibitors against dengue virus protease.
- Compounds $C_{35}H_{27}NO_9$ (CID 54681617), $C_{30}H_{25}NO_5$ (CID 54692801), and $C_{34}H_{23}NO_7S_2$ (CID 54715399) presented good inhibitory properties for NS2B/NS3 DENV2, in addition to had drug-like characteristics.
- $C_{35}H_{27}NO_9$ (CID 54681617), $C_{30}H_{25}NO_5$ (CID 54692801) have molecular features that have not been previously reported as moieties for the DENV protease.
- $C_{35}H_{27}NO_9$ and $C_{34}H_{23}NO_7S_2$ did not show theoretical risk of causing adverse effects.
- The discovery of molecules that interfere with the proteolytic activity of the DENV NS2B/NS3 protease is a potential alternative for the treatment and control of dengue disease.

ACKNOWLEDGEMENTS

This project would not have been possible if I did not have the support of many people and institutions. Hence, I would like to extend my sincere gratitude to all of them.

First, I would like to express all my deepest appreciation to my advisor, Prof. Dr. Jesús Olivero Verbel, for his guidance, continuous support, patience, motivation, and vast knowledge. During this time, his teachings transcended the academy as my learning from him was not only in the scientific area but also on a personal level.

My sincere thanks also go to Prof. Dr. Irene Wagner-Döbler and all members of Microbial Communication research group in the Helmholtz Centre for Infection Research (Germany) and Prof. Dr. Raquel Ocazionez and the Infectious and Metabolic Diseases research group in the Industrial University of Santander (Colombia) for giving me the opportunity to do the internship in their groups and their wonderful collaboration.

I am also indebted to my friends and our coffee sessions, which helped me to overcome setbacks and stay focused. Likewise, I would like to thank my colleagues and fellows for their encouragement and kind co-operation.

Most importantly, none of this would have been possible without the love and patience of my family: my mom, grandma, godmother and the rest of relatives, for their constant love, wise counsel and for supporting me spiritually throughout my life.

I am also grateful to the University of Cartagena by granting me the opportunity to do advanced studies to the development of my PhD thesis.

Finally, I wish to thank the financial support from Colciencias (Grant 110745921616, 110751929058, and 1107-459-21616), the University of Cartagena (Grant 0342010), and CENIVAM (RC 0572-2012).

ABBREVIATIONS

AhR	Aryl hydrocarbon Receptor
AKR1A1	Aldo-keto reductase family 1, member A1 (aldehyde reductase)
Aldr2	Aldehyde reductase 2
APC	Adenomatous polyposis coli. Syn: protein phosphatase 1
ARNT	Aryl hydrocarbon receptor nuclear translocator
BAX	BCL2-associated X protein
BCL2	B-cell CLL/Lymphoma 2
CASP14	Caspase 14
CCL3	Chemokine (C-C motif) ligand 3
CID	PubChem Compound Identifier
CYP1A1	Cytochrome P450, family1, subfamily A, polypeptide 1
CYP1A2	Cytochrome P450, family1, subfamily A, polypeptide 2
CYP1B1	Cytochrome P450, family1, subfamily B, polypeptide 1
CYP3A4	Cytochrome P450, family3, subfamily A, polypeptide 4
CYP4X1	Cytochrome P450, family4, subfamily X, polypeptide 1
Cyp9b2	Cytochrome P450-9b2
DENV	Dengue Virus

EP300	E1A binding protein p300
ESR1	Estrogen Receptor 1
ESR2	Estrogen Receptor 2
GABPA	GA binding protein transcription factor, alpha subunit 60kDa
GIIIspla2	CG1583 gene product from transcript CG1583-RA
HMOX1	Heme oxygenase (decycling) 1
HPSE	Heparanase
HTS	High Throughput Screening
IC_{50}	Half maximal inhibitory concentration
IL17A	Interleukin 17A
IL22	Interleukin 22
IL6	Interleukin 6
IPTG	Isopropyl 1-thio-β-d-galactopyranoside
IRF6	Interferon regulatory factor 6
MMP2	Matrix metallopeptidase 2
NAT	N-acetyltransferase
NFKB1	Nuclear factor of kappa light polypeptide gene enhancer in B-cells 1
NQO1	NAD(P)H dehydrogenase, quinone 1 PAX3

NR1I2	Nuclear receptor subfamily 1, group I, member 2
PDB	Protein Data Bank
PLA2G12A	Phospholipase A2, group XIIA
PROC	Protein C
PTGS2	Prostaglandin-endoperoxide synthase 2
RELA	V-rel reticuloendotheliosis viral oncogene homolog A (avian)
RMSD	Root Mean Square Deviation
SEC14L5	SEC14-like 5
SMILES	Simplified Molecular Input Line Entry Specification
TCDD	2,3,7,8-tetraclorodibenzo-p-dioxina
TP53	Tumor protein p53
WHO	World Health Organization