Centre Armand Frappier

SEARCHING FOR REPURPOSED INHIBITORS OF RICIN THROUGH MOLECULAR MODELING TECHNIQUES

Par

Tanos Celmar Costa França

Mémoire présenté(e) pour l'obtention du grade de

Maître en Sciences (M.Sc.) en

Sciences Expérimentales de la Santé (toxico. & pharmaco. expérim.)

Jury d'évaluation

Président du jury et   Geraldine Delbes
Examinatrice interne   CAF - INRS

Examinateur externe   Rafael Josef Najmanovich
                                       Université de Montreal

Directeur de recherche   Steven R. LaPlante
                                       CAF-INRS

ACKNOWLEGMENTS

RÉSUMÉ

L'huile de ricin est un produit important pour l'industrie ainsi que pour l'économie de plusieurs pays. Par contre il est également le source de la toxine ricine. Un composé incurable utilisable comme arme chimique. Afin d'identifier des composés-antidotes réutilisable contre la ricine nous avons construit, à l'aide de différentes banques de données, une bibliothèque de 82 composés sélectionnés par criblage virtuel (VS) et par des études d'amarrage. Ces composés sont de potentiels liants du site catalytique de la chaine *Alpha* de la ricine (RTA) ainsi que de sa poche de liaison secondaire. Ici nous rapportons des études de modélisation moléculaire supplémentaires sur un groupe de 15 composés issus de cette bibliothèque. Des cycles d'amarrage flexible suivis de simulations de dynamique moléculaire (DM) ont permis de tracer les empreintes digitales de ces composés à l'intérieur du RTA et ainsi mettre à jour la liste des résidus les plus importants pour la liaison du ligand. Enfin, d'autres calculs de MM-PBSA (Molecular Mechanics Poisson–Boltzmann Surface Area) ont permis de classer ces composés, d'élucider leur comportement dynamique à l'intérieur du RTA, et de sélectionner ceux qui devraient maintenir les interactions à l'intérieur des poches catalytiques et secondaires du RTA.

Mots-clés : Ricin, Repurposing, Drug discovery, Computational chemistry,

ABSTRACT

Castor oil is an important product for the industry as well as the economy of several countries. However, it is also the source of ricin toxin, a chemical weapon with no known antidote. In order to identify repurposed antidotes against ricin, we have built, using different databases, a library of 82 compounds selected by virtual screening (VS) followed by docking studies. These compounds are potential binders of both the catalytic and secondary pockets of the ricin chain A (RTA). Here we report additional molecular modeling studies on a group of 15 compounds selected from this library. Steps of flexible docking followed by molecular dynamics (MD) simulations enabled to elucidate the fingerprints of these compounds inside RTA besides updating the list of the most important residues for the ligand binding. Finally, additional MM-PBSA (Molecular Mechanics Poisson–Boltzmann Surface Area) calculations allowed ranking these compounds, as well as elucidating their dynamic behavior inside the RTA, and pointing to those capable of maintaining interactions in both pockets of RTA.

Keywords: Ricin, Repurposing, Drug discovery, Computational chemistry,

SOMMAIRE RÉCAPITULATIF

La ricine est une toxine dimérique classée comme protéine inactivant les ribosomes de type 2 (RIP). Cette famille de protéines présentes dans les bactéries et les plantes agissent comme des ARNr N-glycosylases (EC 3.2.2.22) au niveau des cellules eucaryotes, inhibant la synthèse des protéines. On pense que les RIP jouent un rôle de défense contre les agents pathogènes et les insectes, et peuvent être trouvés dans la plupart des plantes consommées par l'homme. Cependant, seuls quelques-uns d'entre eux sont toxiques. La sous-unité A de la ricine (appelée RTA) est capable de libérer une base adénine spécifique du squelette de l'ARNr 28S, provoquant l'inactivation des sous-unités ribosomiques 60S et conduisant à la mort cellulaire.

La ricine fait partie des RIP les plus toxiques et est également l'une des substances les plus mortelles connues (Janik et al., 2019). On le trouve dans les graines de l'arbuste tropical connu sous le nom de ricin (*Ricinus communis*). Ces graines sont également une source d'huile de ricine, une matière première largement utilisée dans le monde pour la production industrielle de lubrifiants, de carburants verts et de médicaments (Patel et al., 2016).

Le résidu organique obtenu après extraction de l'huile de ricin, appelé le tourteau de ricine, est riche en minéraux et autres nutriments, étant pour cette raison utilisé comme complément alimentaire pour le bétail, après élimination de la ricine résiduelle (Doan, 2004). Cependant, l'élimination complète de la ricine n'est pas assurée et, par conséquent, les intoxications des travailleurs de l'industrie de l'huile de ricine et des bovins nourris avec des tourteaux sont assez fréquentes. Cela entraîne des problèmes de santé publique et des pertes économiques dans des pays comme l'Inde, la Chine et le Brésil, les plus grands producteurs mondiaux d'huile de ricine (Doan, 2004 ; Sousa et al., 2019).

Du fait de l'inexistence d'antidotes ou de vaccins, combinée à sa toxicité élevée, sa facilité d'obtention, sa stabilité chimique et sa solubilité dans l'eau, cela a encouragé l'utilisation abusive de la ricine comme agent de guerre chimique (Audi et al., 2005 ; Janik et al., 2019 ; Knight, 1979 ; Pita et Romero, 2014). En conséquence, cette toxine a été répertoriée comme arme chimique dans la Convention sur les armes chimiques (CAC) (https://www.opcw.org/chemical-weapons-convention) de l'Organisation pour l'interdiction des armes chimiques (OIAC) (https://www.opcw.org/).

Le mode de liaison de la RTA à l'ARNr a été révélé par Ho et collaborateurs (Ho et al., 2009) qui ont montré qu'en plus du site catalytique, il existe également un site secondaire destiné

à accueillir une base guanine de la séquence cible ribosomique invariante GAGA (Ho et al., 2009). Les principaux résidus du site catalytique sont: Val81, Gly121, Glu177 et Arg180, tandis que les principaux résidus du secondaire sont Asp75, Asn78, Asp96 et Asp100.

Les sites catalytique et secondaire de la RTA sont donc les cibles ultimes pour la conception d'antidotes contre la ricine. Les ligands capables de se lier aux deux sites en même temps, appelés liants doubles, fonctionneront certainement comme des inhibiteurs efficaces de la RTA. Cependant, la plupart des inhibiteurs potentiels rapportés dans la littérature jusqu'à présent fonctionnent comme des liants simples. De plus, aucun n'a encore atteint des valeurs d'IC_{50} dans la gamme du nM. Cela suggère qu'il y a encore suffisamment de place pour la conception/découverte de médicaments d'inhibiteurs nouveaux et plus puissants ciblant en même temps les sites catalytiques et secondaires de la RTA, qui pourraient certainement atteindre une inhibition dans la gamme nM.

Dans des études antérieures (Botelho et al., 2020a; Botelho et al., 2020c), nous avons créé une bibliothèque de 82 composés, qui ont été sélectionnés par criblage virtuel (VS) et des études d'amarrage à partir de différentes banques de données. Parmi ces composés, 6 sont de potentiels liants doubles, et ils pourraient être réutilisés comme antidotes contre cette toxine. Maintenant nous rapportons de nouvelles études de modélisation moléculaire effectuées sur un groupe de 15 molécules. Ces dernières, représentent mieux les caractéristiques structurelles de la bibliothèque d'origine et devraient également se porter comme liants doubles, et déclenchant une inhibition plus forte. Des cycles d'amarrage flexible suivis de simulations de dynamique moléculaire (DM) ont permis de tracer les empreintes digitales de ces composés à l'intérieur de la RTA et de mettre à jour la liste des résidus les plus prometteurs pour la liaison du ligand. De plus, d'autres calculs MM-PBSA (*Molecular Mechanics Poisson–Boltzmann Surface Area*) ont permis de classer ces composés, en plus d'élucider leur comportement dynamique à l'intérieur de la RTA, et de sélectionner ceux qui devraient qualifier comme les liants doubles.

Notre étude a permis d'affiner le protocole utilisé auparavant (Botelho et al., 2020a; Botelho et al., 2020c) pour l'investigation théorique des médicaments réutilisés comme antidotes potentiels contre la ricine. Bien qu'aucun changement significatif n'ait été observé en utilisant le nouveau protocole d'amarrage, l'extension du temps de simulation par DM pour 500 ns a montré une étape fondamentale pour corroborer les résultats d'amarrage et filtrer les ligands qui sont vraiment capables d'effectuer des interactions stables dans les deux poches de la RTA. Ce raffinement a permis d'orienter avec plus de confiance quatre composés vers d'autres tests *in vitro* : CID 135977982, CID 136132835 et la naldémidine, en tant que liants potentiels de la RTA,

et le nilotinib, en tant que liant unique de la poche secondaire avec le potentiel de bloquer l'entrée du site catalytique de la RTA. Nous pensons que ces quatre composés présenteront des valeurs IC_{50} dans la gamme nM après avoir été évalués expérimentalement.

Nos résultats ont également corroboré le complexe RTA/C2X rapportée par Ho et collaborateurs (Ho et al., 2009) en tant que modèle cohérent de la liaison de la RTA à la boucle GAGA de l'ARNr 28S. Ce complexe est une ressource puissante pour la conception de médicaments contre RTA. En ce sens, la découverte aussi des liaisons H cohérentes avec les résidus Arg213 et Arg258 de la RTA, observés pour la plupart des ligands dans les simulations d'amarrage et de DM, ouvre des nouvelles opportunités pour la conception de nouveaux antidotes contre la ricine.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

7CP	7-carboxy pterin
BWAT	Biological Weapons Anti-Terrorism
CWC	Chemical Weapons Convention
DBV	Database Viewer
DMSO	Dimethylsulphoxide
DM	Dynamique Moleculaire
ER	Endoplasmic Reticulum
ERAPD	ER-Associated Protein Degradation
FDA	Food and Drugs Administration
GOLD	Genetic Optimization for Ligand Docking
ICM	Internal Coordinate Mechanics
LBVS	Ligand-Based Virtual Screening
MD	molecular dynamics
MM-PBSA	Molecular Mechanics Poisson–Boltzmann Surface Area
MOE	Molecular Operating Environment
MVD	Molegro Virtual Docker
OPCW	Organization for Prohibition of Chemical Weapons
PDB	Protein Data Bank

PDI   Protein Disulphide Isomerase

PTA   Pteroic Acid

RBVS   Receptor-Based Virtual Screening

RIP   Ribossome-Inactivating Protein

RMSD   Root-Mean-Square Deviation

RNA   Ribonucleic Acid

rRNA   Ribossomic RNA

RTA   Ricin Chain A

RTB   Ricin Chain B

USAMRIID   US Army Medical Research Institute of Infectious Diseases

VdW   Van der Walls

VS   virtual screening

WWII   World War II

1   INTRODUCTION

Ricin is a toxin classified as type 2 ribossome-inactivating protein (RIP). This family of proteins present in bacteria and plants, act as rRNA N-glycosylases (EC 3.2.2.22) in eukaryotic cells, inhibiting the protein synthesis. They are capable of releasing a specific adenine base from the backbone of the 28S rRNA, causing inactivation of the 60S ribosomal subunits. RIPs are believed to play a defense role against pathogens and insects and can be found in most of the plants consumed by humans. However only few of them are toxic (ENDO *et al.*, 1988; Funatsu *et al.*, 1991; May *et al.*, 1989; Zhou *et al.*, 1994).

All RIPs are structurally related and share a similar mechanism of action involving two conserved catalytic residues, a glutamic acid and an arginine. They are classified in three different types according to their protein domains (Figure 1.1), and also have shown potential use as anticancer agents and against HIV-1 infection (ENDO *et al.*, 1988; Funatsu *et al.*, 1991; May *et al.*, 1989; Zhou *et al.*, 1994). Type I RIPs are composed only of an A domain which is the one with glucosidase activity, while Type II RIPs are composed of an A and a B domain connected by a disulfide bond. It's the B domain which enables the entrance of those proteins into the cell, making this type more cytotoxic; Type III subdivides into two subtypes: the (AC) which bears a C-terminal domain with unknown functionality, and the (AD) which contains a site for inactivation.

Ricin is amongst the most toxic RIPs and is also one of the most lethal substances known (Janik *et al.*, 2019). It is found in the seeds of the tropical shrub known as Castor bean (*Ricinus communis*) (see Figure 1.1). These seeds are also source of the castor oil, a raw material largely used worldwide for the industrial production of lubricants, green fuel and drugs (Patel *et al.*, 2016). The world's largest producers of ricin oil are India, China and Brazil (Doan, 2004; Sousa *et al.*, 2019).
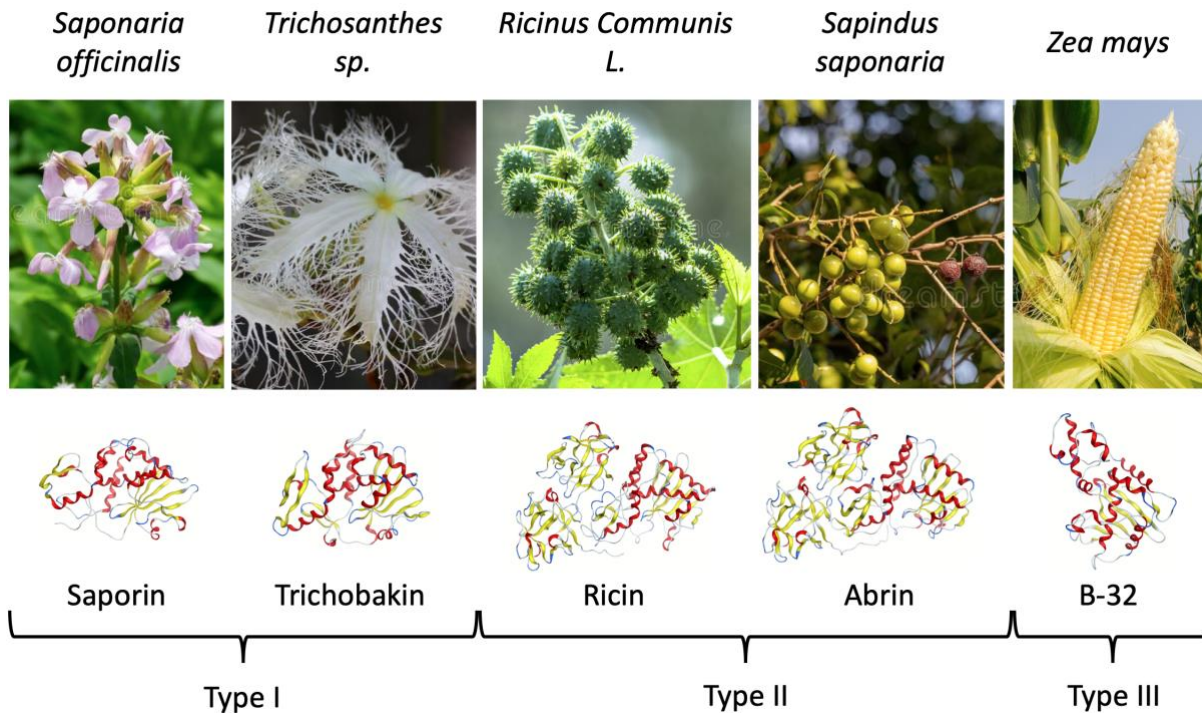
Figure 1.1. Examples of RIPs and their plants of origin. Photos obtained from https://www.dreamstime.com/. 3D structures drawn with MOE® from the Protein Data Bank (PDB) (Berman *et al.*, 2002) IDs: 1QI7, 6Z1Y, 2AAI, 1ABR and 2PQI, respectively.

The organic residue obtained after castor oil extraction, named oil cake, or press cake, is rich in minerals and other nutrients, being for this reason used as food supplement for the cattle, after the removal of the residual ricin (which can be up to 5%) through *salting out* (Doan, 2004). However, the complete elimination of ricin is not warranted and, in consequence, sporadic intoxications of workers in the ricin oil industry and cattle fed with oil cake might happen (Worbs *et al.*, 2011). There are no reports in the literature on basal human exposition due to the consumption of industrial products derived from ricin, fact that reflects an efficient removal of the toxin from these products.

The most common cases of accidental poisoning with ricin happen with children attracted by the appearance of the seeds followed by adults that confounded the ricin seeds with nuts (Worbs *et al.*, 2011). However, no significant number of cases is registered anually (Worbs *et al.*, 2011).

The symptoms of ricin intoxication occur between 3 and 20 h after ingestion or injection, with severity depending on the amount of toxin incorporated. They are characterized by abdominal pain, emesis, diarrhea with or without blood, muscular pain, cramps in the limbs, circulatory collapse, dyspnea and dehydration. Increase in the levels of white blood cells, blood

urea nitrogen (BUN), aspartate aminotransferase (AST) and alanine aminotransferase (ALT) also happen and indicate dysfunction of liver and kidneys. Fatal cases show hemorrhagic necrosis in intestines and heart and edema in lungs (Worbs *et al.*, 2011).

The inexistence of antidotes, combined with its high toxicity, facility of obtention, chemical stability, and water solubility, have encouraged the misuse of ricin as a chemical warfare agent, which represent the higher risk posed by this toxin today (Audi *et al.*, 2005; Janik *et al.*, 2019; Knight, 1979; Pita & Romero, 2014). In consequence this toxin was listed as a chemical weapon in the Chemical Weapons Convention (CWC) (https://www.opcw.org/chemical-weapons-convention) of the Organization for Prohibition of Chemical Weapons (OPCW) (https://www.opcw.org/).

1.1   The misuse of ricin

Ricin was first considered as potential chemical warfare agent at the beginning of the 20th century when, in 1903, the US army described studies using ricin enriched ammunitions (Augerson, 2000). This toxin was also investigated to be used in both World War I and II by the United States, Canada, the European powers and Japan. Many studies were performed aiming weaponizing ricin to be used by aerosolization to target the lungs, and also in ammunitions, targeting the blood stream. However, none of those studies succeeded due probably to the deactivation of ricin by the thermal effects of the munitions. After WWII the Soviet Union and Iraq still searched the weaponization of ricin but also with no success (Audi *et al.*, 2005).

Despite never being successfully weaponized for mass destruction ricin has been largely used in assassination attempts and terrorist actions throughout the world in the last decades (Jansen *et al.*, 2014). The most famous case happened in 1978 when the Bulgarian journalist exiled in London Georgi Markov was injected in the leg with a tiny hollowed sphere stuffed with 500 µg of ricin and covered with wax, while he waited for a bus on the Waterloo Bridge. It was found out that the hypodermic needle used to inject him was disguised in the tip of an umbrella handled by a Bulgarian secret service agent. In the same day Markov started feeling strong pain in the injection site, fever and nausea, being hospitalized the next day due to the worsening of his condition and dying of heart failure 4 days after (Janik *et al.*, 2019; Knight, 1979; Pita & Romero, 2014). A similar attack happened 10 days before Markov's; this time in Paris, against Vladimir Kostov, another Bulgarian exile. However, he survived, probably due to its heavy clothing which avoided the complete penetration of the needle (Eitzen Jr & Takafuji, 1997).

Besides injection, other routes of intoxication with ricin have been explored in assassination attempts. Dermal exposure to a mix of ricin, DMSO and an aloe vera gel meant to be applied on doorknobs was in the plans of an antigovernment group sentenced in 1995 in the USA under the Biological Weapons Anti-Terrorism (BWAT). Also, inhalation was in the minds of the groups that mailed a letter containing ricin powder to the white house or the pentagon in 2003, (Audi *et al.*, 2005; Musshoff & Madea, 2009). Action which was repeated in 2013, 2018 and 2020 (https://www.npr.org/sections/thetwo-way/2014/05/20/314219596/tupelo-man-who-sent-ricin-letters-to-obama-gets-25-year-sentence; https://www.npr.org/2018/10/04/652954240/fbi-makes-ricin-case-arrest-after-man-sends-castor-seeds-to-pentagon, https://www.nytimes.com/2020/09/20/us/politics/ricin-white-house-trump.html). The terrorist group Al Qaeda was suspected of being behind some of those letters.

1.2   Mechanism of action of ricin

The RTB domain of ricin facilitates its entrance into eukaryotic cells through the binding to superficial galactosyl moieties or mannose receptors (Figure 1.2). According to the literature (Sphyris *et al.*, 1995) RTB enables the binding of 10^6-10^8 ricin molecules per cell surface. After internalization ricin ends up in the lumen of the Endoplasmic Reticulum (ER) via retrograde transport (Spooner *et al.*, 2006). It resists to acid degradation in endosomes or lysosomes due to its stability under a wide range of pH. Once inside the ER RTB is separated from RTA due to the action of the protein disulphide isomerase (PDI) (Bellisola *et al.*, 2004; SPOONER *et al.*, 2004). This exposes the RTA active site, triggering its rRNA N-glycosylase activity which consists on the depurination of the adenine A-4324 found in the GAGA loop of rRNA 28S. This happens at a rate of 1500 ribosomes per minute per RTA unit (Endo & Tsurugi, 1988) and quickly stops the protein synthesis, leading to cell death (Lord *et al.*, 1994; Olsnes, 2004; Olsnes *et al.*, 1975; Olson *et al.*, 2004). The tagging by ubiquitination and eventual elimination of RTA through ER-Associated Protein Degradation (ERAPD) is considered unlikely to happen with RTA because it brings few lysine residues needed for the ubiquitin binding (Deeks *et al.*, 2002).
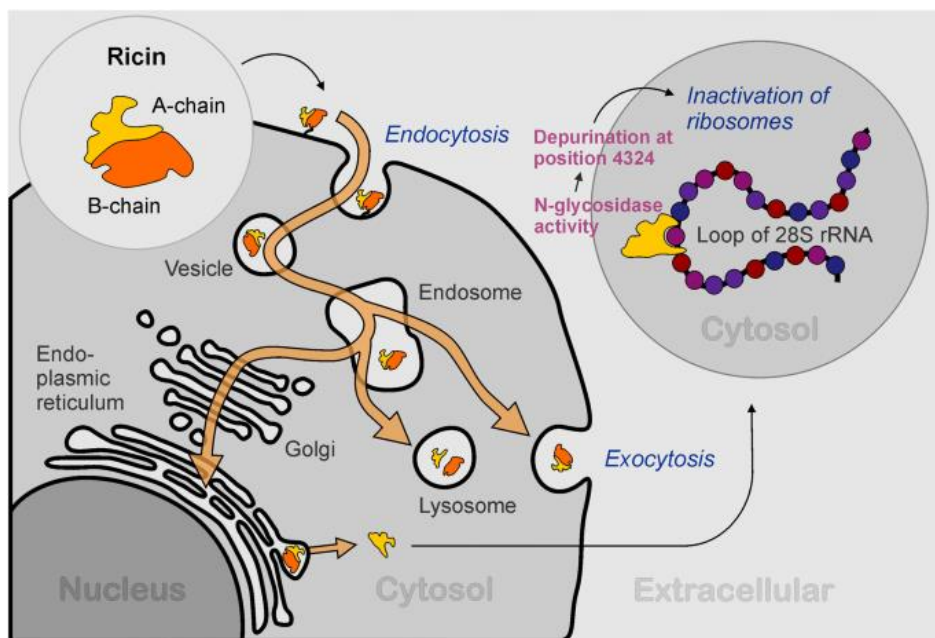
Figure 1.2. Mechanism of action of ricin [copied from (Franke *et al.*, 2019)].

The binding mode of RTA to rRNA was revealed by Ho and co-workers (Ho *et al.*, 2009) who published the crystallographic structure PDB ID: 3HIO which brings RTA complexed with the cyclic tetranucleotide inhibitor C2X (Figure 1.3), designed to mimic the recognition loop of 28S rRNA (Figure 1.3). As one can see in Figure 1.3, besides the active site there is also a secondary site which is meant to accommodate a guanine base from the invariant GAGA ribosomal target sequence (Ho *et al.*, 2009). The main residues responsible for the binding and removal of the adenine nucleotide are Val81, Gly121, Glu177 and Arg180. These residues are found in a pocket of the RTA binding site named the catalytic site. At the right side of the catalytic site there is another pocket were the guanine neighbor of the recognition loop of 28S rRNA binds as shown in Figure 1.3. This pocket was named the secondary site and is mainly composed by the residues Asp75, Asn78, Asp96 and Asp100 (See Figure 1.3).

The catalytic and secondary sites of RTA are, therefore, the ultimate targets for the design of antidotes against ricin. Ligands capable of reproducing the binding mode of C2X, qualifying as dual binders, certainly will work as efficient RTA inhibitors. However, most of the potential inhibitors reported in the literature so far work as single binders. Also, none has achieved IC_{50} values in the nM range yet. This suggests that there is still enough room for the drug design/discovery of new and more potent inhibitors targeting at the same time the catalytic and secondary sites of RTA, that might certainly achieve inhibition at the nM range.
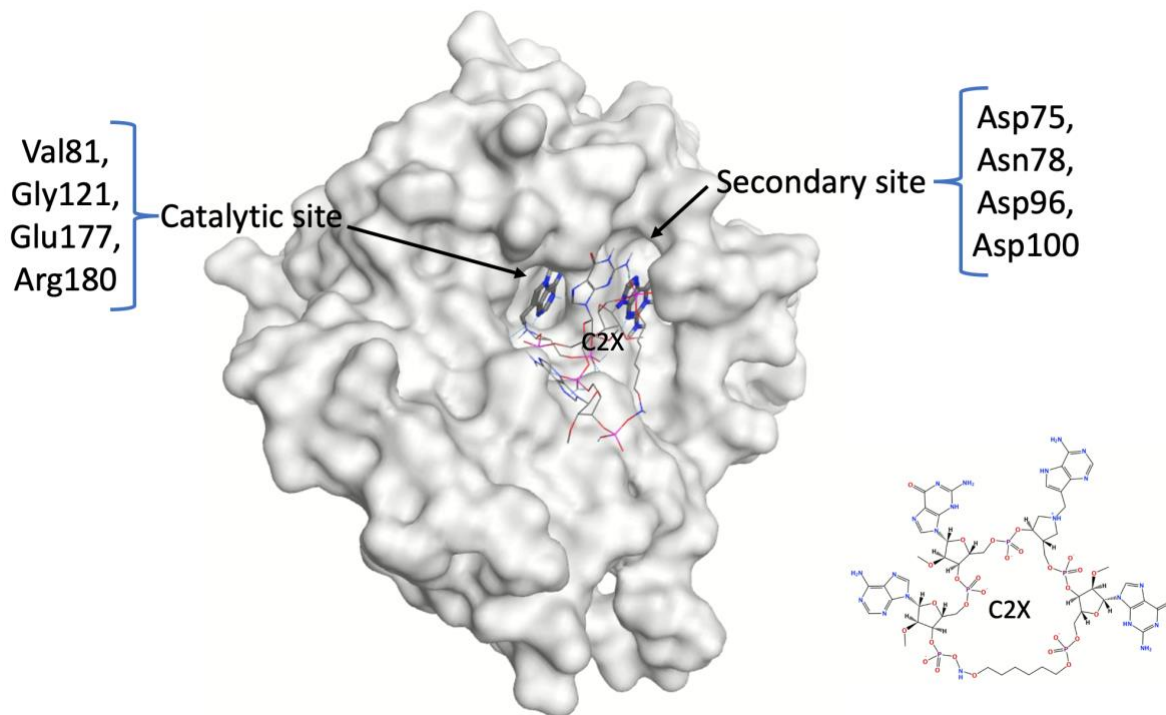
Figure 1.3. Binding mode of C2X inside RTA. For clarity most of C2X 3D structure is shown in wire representation. The C2X 2D structure is shown in the right corner. Figure constructed with MOE® and Power Point®.

1.3   The search for vaccines against ricin

Due to the rapid internalization of ricin into the body, it's believed that the best strategy to deal with this toxin is the preventive vaccination in order to create antibodies that will enter in action as soon as ricin is present. On this line the main approach that have been used so far in the development of vaccines to prevent ricin intoxication involves recombinant RTA mutants. Currently the most promising candidates are the vaccines RiVax from the University of Texas Southwestern Medical Center (Smallshaw *et al.*, 2002; Smallshaw & Vitetta, 2011) and RVEc, developed by the Army Medical Research Institute of Infectious Diseases (USAMRIID) (Carra *et al.*, 2007; McLain *et al.*, 2012).

1.3.1   RiVax

RiVax was formulated by the company Soligenix and it's currently at phase I of clinical trials. It consists on the mutant V76M, Y80A of RTA alum-adjuvanted and thermostabilized through lyophilization. The two mutations don't change the 3D structure of RTA (PDB ID: 3SRP)

compared to wild type RTA while are capable of eliminating its enzymatic activity as well as the ability of inducing vascular leak syndrome in humans (Smallshaw *et al.*, 2002; Smallshaw *et al.*, 2003). It can be stored for long periods of time and have been shown to be capable of triggering an immune response against ricin, two weeks after second vaccination (6 weeks after the first vaccination), in macaques for cases of inhalation, ingestion or injection, preventing RTA from entering the cells.

1.3.2   RVec

RVec is a truncated recombinant mutant of RTA where the hydrophobic loop in N-terminal region between residues 33-44 was removed as well as the C-terminal hydrophobic extremity after residue Ser198 (see PDB ID: 5SV3) (Carra *et al.*, 2007; McLain *et al.*, 2011; McLain *et al.*, 2012; Porter *et al.*, 2011). RTA 1-33/44-198 showed thermal stability to denaturation, better water solubility, and low toxicity. When essayed with mice it triggered immune response, protecting them against aerosol exposure to ricin four weeks after exposure (Porter *et al.*, 2011).

1.4   The search for antidotes against ricin

The strategies used in the drug discovery against RTA have been mostly based on structure-based design, fragment identification, and virtual screening. Those studies have identified many potential binders to the active site of RTA, with a predominance of the pterin derivatives among them (Jasheway *et al.*, 2011; Monzingo & Robertus, 1992; Pruet *et al.*, 2011; Pruet *et al.*, 2012; Saito *et al.*, 2013; Wiget *et al.*, 2013). This is not a surprising result if we consider the structural similarities between the pterin and adenine scaffolds. A brief overview of the most relevant ricin inhibitors reported so far in the literature according to the approaches mentioned above is given in the next sub-sections.

1.4.1   Pterin derivatives

As mentioned above, most of the potential RTA inhibitors reported in literature are pterin derivatives. The rationale behind it comes from the search for subtract analogues as competitive inhibitors (Monzingo & Robertus, 1992). Yan and co-workers (Yan *et al.*, 1997) first reported the crystallographic structures of RTA complexed with neopterin (PDB ID 1BR5) and pteroic acid (PTA) (PDB ID 1BR6) at its catalytic pocket and determined the apparent K_i's as > 2mM and 0.6mM respectively. This result pointed to the PTA as a potential competitive RTA inhibitor and motivated the search for pterin derivatives as efficient antidotes against intoxication with RTA.

Some years latter Pruet (Pruet *et al.*, 2011) reported that 7-substitued pterins might be more efficient binders than the 6-substitued pterins, like the PTA. They synthesized a series of derivatives and showed that 7-carboxy pterin (7CP) can achieve an RTA inhibition of 200 μM. Further developments on this research enabled the discovery of new pterin derivatives with activities in the range of 6 - 115 μM (Pruet *et al.*, 2012; Saito *et al.*, 2013; Wiget *et al.*, 2013) with the best results obtained so far being observed for NNPC, NNPT, and the triazol derivative shown in Figure 1.4. The 3D structure of RTA complexed with NNCP under the PDB ID 4HUO (Figure 1.5) reveals the binding mode of those compounds and confirms that the secondary pocket of RTA has not been explored yet in the drug discovery against ricin.

N-{[1-(phenylthio-methyl)-1H-1,2,3-triazol-4-ill]methyl}-pterin-7-carboxamide

IC_{50} = 15 μM

N-[N-(pterin-7-il)carbonilglycil]-L-phenylalanine (NNCP)

IC_{50} = 20 μM

N-[N-(pterin-7-il)carbonilglycil]-L-tyrosine (NNCT)

IC_{50} = 6 μM
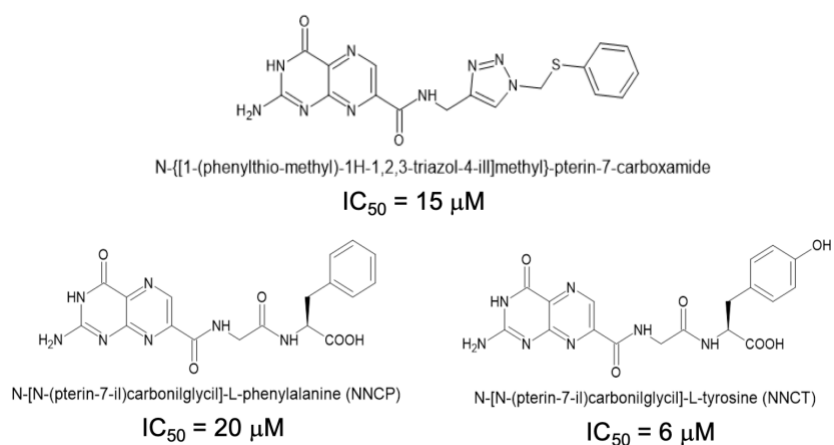
Figure 1.4. Best RTA inhibitors available in the literature.

Val81
Gly121 — Catalytic site
Glu177
Arg180

Asp75
Asn78 — Secondary site
Asp96
Asp100
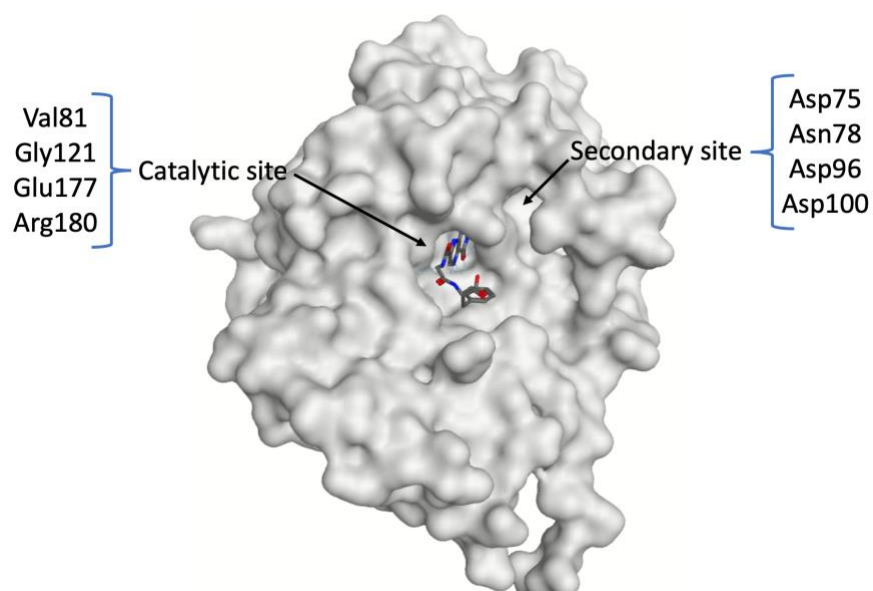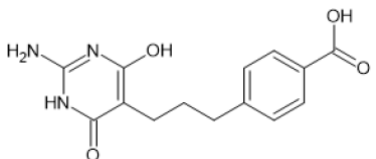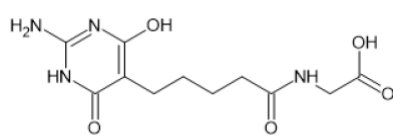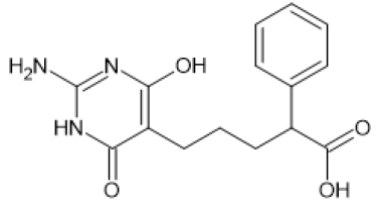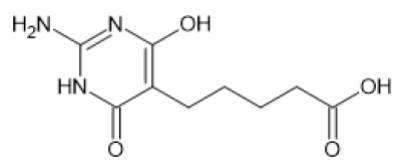
Figure 1.5. RTA complexed with NNCP. Figure constructed with MOE® and Power Point®.

1.4.2   Non-pterin derivatives

Despite showing promising activities towards RTA the pterin derivatives present solubility issues that might impair their effective use as antidotes. In order to address this Bai and co-workers (Bai *et al.*, 2009) evaluated a series of 2-amino-4,6-dihydroxy-pyrimidines 10 times more soluble than the pterin derivatives and showed that they also bind to the catalytic site of RTA and might represent a useful new class of ricin inhibitors. However, only 4 among the 10 compounds synthesized presented some inhibition towards RTA, with the best IC_{50} observed = 0.27 mM (see Table 1.1). The PDB ID 3EJ5, solved by Bai and co-workers (Bai *et al.*, 2009), shows compound 1 in Table 1.1 complexed in the catalytic site of RTA.
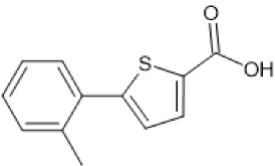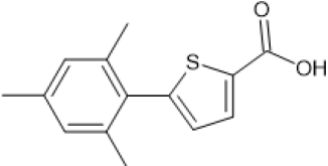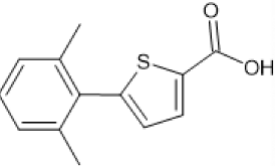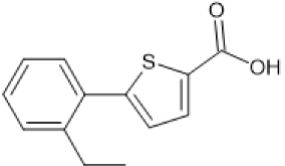
Table 1.1. Compounds reported by Bai and co-workers (Bai et al., 2009) as potential RTA binders.

Compound		IC_{50}	Compound		IC_{50}
1		0.27 mM	3		2mM
2		>1mM	4		> 3mM

More recently Li and co-workers (Li *et al.*, 2021; Li *et al.*, 2020) proposed a different approach for the search for RTA inhibitors. They screened, through plasmon resonance, small molecules capable of binding to RTA and solved their interactions by crystallography. In the first round of experiments they found 5 fragments binding to RTA at the mid-micromolar range but none at the catalytic site, neither at locations that might cause any significant change in the catalytic site geometry. Further improvements using a structure-guided approach afforded four compounds named RU-NT-70, RU-NT-75, RU-NT- 93, and RU-NT-102 (Table 1.2) with efficiency raised by one order of magnitude, which were characterized by X-ray crystallography. It was found that these compounds, despite not binding to the catalytic site, cause enough local conformational

changes capable of inhibiting the depuration of rRNA and cytotoxicity to mammalian cells. According to the authors those are the first fragments capable of disrupting the interaction RTA-ribosome. RTA complexed with the fragments selected by Li and co-workers (Li *et al.*, 2021) can be found in the PDB under the IDs: 7MLN (RU-NT-70), 7MLO (RU-NT-75), 7MLP (RU-NT-93), and 7MLT (RU-NT-102).

Table 1.2. Fragments proposed by Li and co-workers (Li *et al.*, 2021; Li *et al.*, 2020) as capable of disrupting the interaction RTA-ribosome.
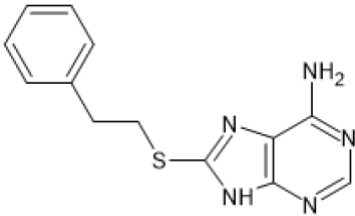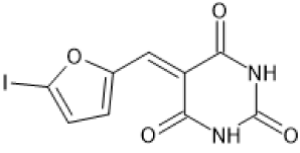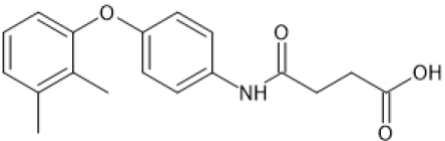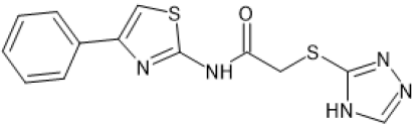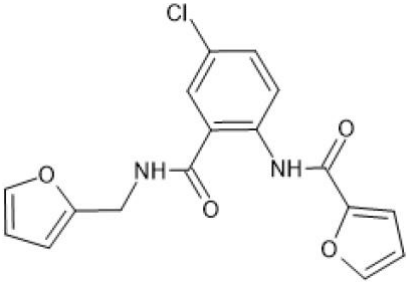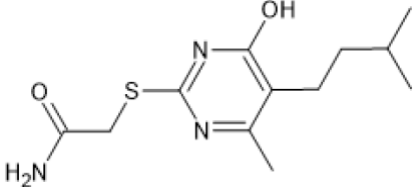
Compound	K_D	Compound	K_D
RU-NT-70	372±34	RU-NT-75	319±9
RU-NT-93	130±6	RU-NT-102	306±22

1.4.3   Virtual screening and repurposing studies

VS studies have also been reported in the literature as strategy in the drug discovery against RTA. Bai and co-workers (Bai *et al.*, 2010) have screened a diversity library of 47,797 compounds from ChemBridge®, using the two VS programs: Internal Coordinate Mechanics (ICM) (Abagyan *et al.*, 1994) and Genetic Optimization for Ligand Docking (GOLD) (Jones *et al.*, 1997) and combined the results obtained to select 306 compounds which were further purchased and evaluated through a kinetic assay against RTA. Six among them showed RTA inhibition ranging from 250 to 1800 μM (see Table 1.3) while 2 (CID 609529 and CID 767227) showed effectiveness to protect cells from ricin in a cell-based assay using Vero cells (5000 cells/well) grown in Dulbecco's minimal essential medium (DMEM) +10% of fetal bovine serum (FBS). CID 609529

showed strong ricin inhibition at 25 μM and little toxicity at $IC_{50} > 30$ μM with the corresponding $EC_{50} = 10$ μM. CID 767227 in the other hand showed strong toxicity at concentration > 60 μM in which it showed a modest ricin inhibition.

Table 1.3. Compounds selected by VS (Bai *et al.*, 2010).

Compound	K_D (μM)	Compound	K_D (μM)
CID 785308	1800	CID 609529	180
CID 766985	250	CID 767227	550
CID 849809	250	CID 933869	500

Mishra an co-workers (Mishra & Prasad, 2011) performed a ligand-based virtual screening (LBVS) in the ZINC databank (https://zinc.docking.org/) using PTA as reference compound and downloaded 34 compounds that were further docked inside RTA using Molegro Virtual Docker (MVD)® (Thomsen & Christensen, 2006). ZINC (https://zinc.docking.org/) is a free database of commercially-available compounds for VS containing over 230 million purchasable compounds in ready-to-dock, 3D formats, while MVD® (Thomsen & Christensen, 2006) is a docking software very reputed for VS studies.

Based on the results authors pointed to the pterin derivatives ZINC05156321, ZINC05156324 and ZINC08555900 (Figure 1.6) as potential binders to be investigated experimentally against ricin.
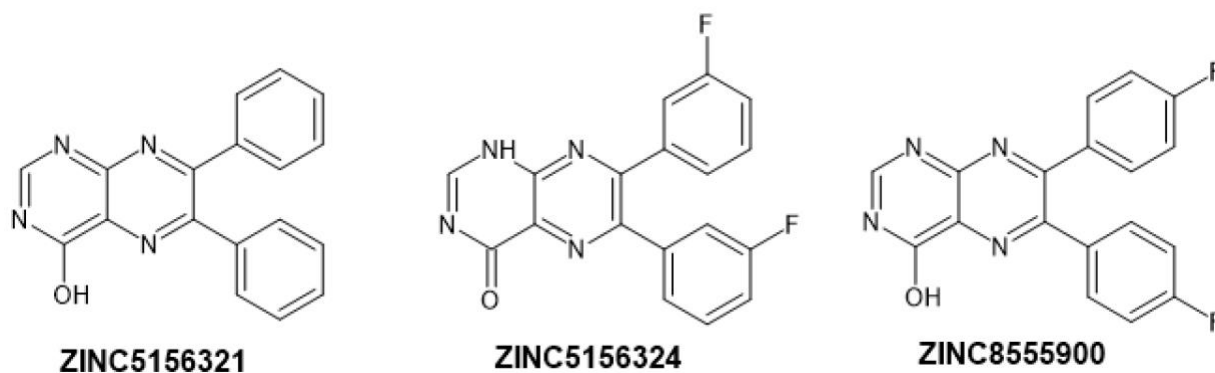
Figure 1.6. Compounds pointed by Mishra an co-workers (Mishra & Prasad, 2011) as potential RTA binders.

Recently a more sophisticated VS study was performed by Botelho and co-workers (Botelho *et al.*, 2020a; Botelho *et al.*, 2020c) where both, the receptor-based virtual screening (RBVS) and LBVS approaches were applied to search for RTA dual binders among FDA-approved drugs available at Cheminfo (https://chemoinfo.ipmc.cnrs.fr/), the approved drugs library available at DrugBank (https://www.drugbank.ca/), and the PubChem database (https://pubchem.ncbi.nlm.nih.gov/). As shown in Figure 1.7, the RBVS and LBVS searches permitted the selection of a library of 6795 compounds. After selection of the best scored ones and elimination of those with undesirable side effects, 180 compounds were submitted to rigid docking using MVD® (Thomsen & Christensen, 2006). According to the docking studies 82 compounds potentially capable of acting as dual binders to RTA were selected. Further MD simulations and MM-PBSA studies of 15 among those 82 compounds, using GROMACS (Abraham *et al.*, 2015), permitted pointing to 6 compounds (Figure 1.8) as potential dual binders to RTA that might be repurposed against ricin intoxication.
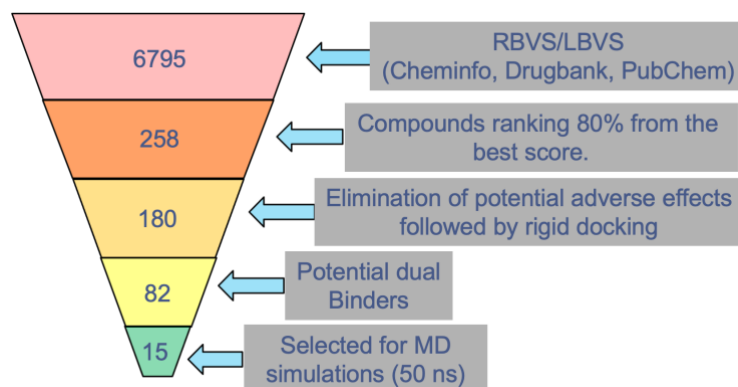
Figure 1.7. Virtual screening approach adopted by Botelho and co-workers (Botelho et al., 2020a; Botelho et al., 2020c).
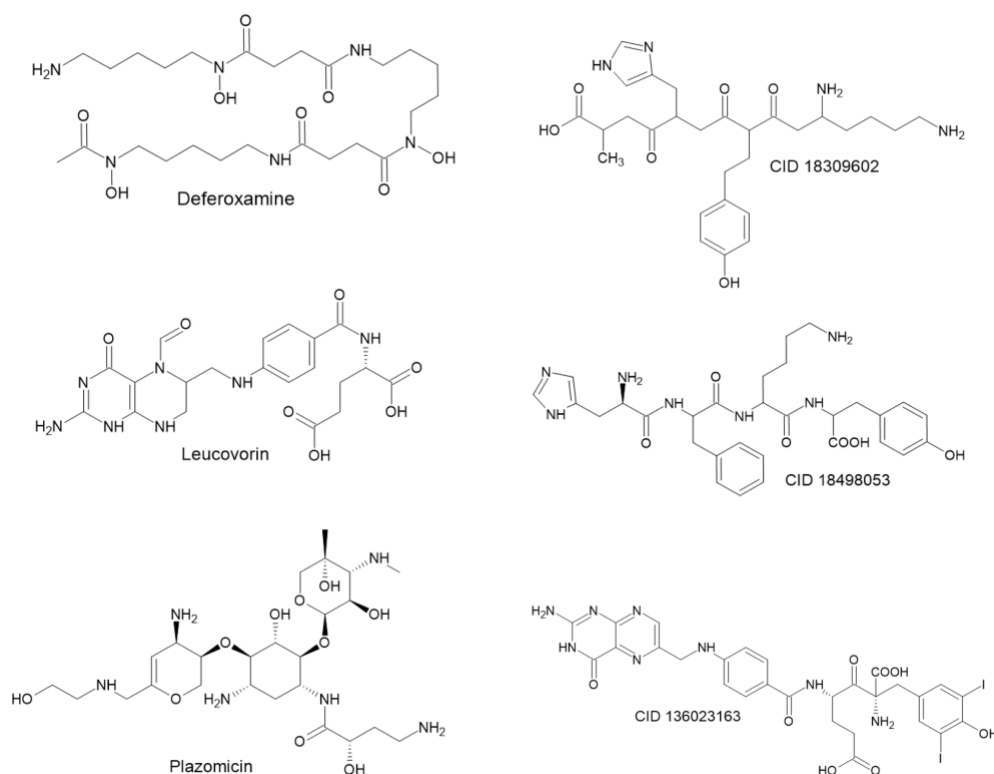
Figure 1.8. First set of compounds pointed by Botelho and co-workers (Botelho et al., 2020a; Botelho et al., 2020c) as potential dual binders to RTA.

1.5   Hypothesis and goals of the project

As commented above there is no effective antidote available yet against intoxication with ricin. Besides, the most promising compounds reported so far in the literature (Saito *et al.*, 2013)

are not capable of reproducing the binding mode of RTA to the GAGA loop of rRNA 28S, as illustrated by the binding of C2X (see Figure 1.3) and, therefore, don't qualify as dual binders to both the catalytic and the secondary sites of RTA. This fact inspired our hypothesis that dual binders to RTA, would render more effective antidotes against ricin which, besides mimicking the binding of RTA to its natural target, will also increase the number of interactions, contributing to a stronger binding and, consequently, lower IC_{50}. In order to quickly check this hypothesis, we have constructed the library of 82 potential dual binders to RTA obtained from databanks of FDA-approved drugs and started investigating the dual binding potential of the compounds in this library through rigid docking and molecular dynamics (MD) simulations, which resulted in the selection of 6 compounds (Botelho *et al.*, 2020a; Botelho *et al.*, 2020c). On the present project we have the main goal of continuing challenging this library through more sophisticated modeling studies on these 6 compounds plus 9 additional ones and also refine the methodology used. Flexible docking (instead of rigid) followed by 10 times longer MD simulations were performed in order to achieve the following specific goals: 1) better map the interactions of those compounds inside RTA; 2) finding the most important residues for the binding, 3) selecting the ones capable of stablishing and keeping interactions in the catalytic and secondary sites, and 4) ranking them according to the potential of inhibiting RTA.

1.6   Computational tools used in this project

Molecular modeling techniques have proved to be an important asset in the drug discovery and gained more and more space in the last decades (Adelusi *et al.*, 2022). Techniques like homology modeling (Franca, 2015), Quantitative Structure Activity Relationships (QSAR) studies (Topliss, 2012), VS (Walters *et al.*, 1998), molecular docking (Fan *et al.*, 2019), molecular dynamics simulations (Karplus & McCammon, 2002), artificial intelligence drug discovery (Paul *et al.*, 2021), among others are today totally incorporated into the research labs in both the private sector and academia. Three of those techniques: the VS, docking and the MD simulation have been explored to achieve our main goal of finding more effective binders to RTA among repurposed drugs. On this section we will discuss a bit more about them.

1.6.1   Virtual Screening (VS)

The VS (see illustration in Figure 1.7) consists on a search in virtual libraries of drugs for compounds with potential of binding to a given molecular target (protein, membrane or nucleic

acid). Currently this can be done following two different approaches: The Ligand-Based Virtual Screening (LBVS) (Sun, 2008; Willett *et al.*, 1998) and the Receptor-Based Virtual Screening (RBVS) (Toledo Warshaviak *et al.*, 2014). LBVS is performed based on the structure of a molecule which already has the properties of interest, like a substrate, cofactor, agonist/antagonist or a similar molecule that is previously known to bind into the target under study. In this case the search is for analogues meant to show more affinity for this target and act as inhibitors. RBVS on the other hand is used when no ligand of the active site or pocket targeted is known. In this case the tridimensional structure of the biological target is used to screen in the virtual databanks for molecules that might bind to it based on the interaction affinity or complementarity to the site of interest.

1.6.2   Molecular docking

The main goal of the molecular docking is to predict the binding mode and affinity in systems composed of two or more molecules which 3D structures are known. The most common types of system investigated by docking involve interactions of a macromolecule with another macromolecule (like protein-protein or DNA/RNA-protein) or with a small and usually flexible molecule (like protein-ligand or DNA/RNA-ligand). The docking protein-ligand, illustrated in Figure 1.9, is very useful for the drug design and helps to save time and money in the discovery of new drugs by reducing many experimental steps in the synthesis and activity essays (Fan *et al.*, 2019; Huang & Zou, 2010; Thomsen & Christensen, 2006). In this type of computational study thousands of possible orientations and conformations of the ligand inside the binding pocket of the protein (named poses) are tested and evaluated. The pose showing the lowest energy is considered the most likely binding mode of the ligand inside that molecular target. Different ligands can also be ranked and prioritized with regard to a given target based on their poses optimizing, this way, the search for the most promising ligand. This optimization is performed in two steps named sampling and scoring. The first consists basically on the generation of several possible poses in a region involving the chosen binding pocket of the protein. When this process takes into consideration the flexibility of the protein it's called flexible docking (Jones *et al.*, 1997; Rosenfeld *et al.*, 1995). Otherwise it's a rigid docking (Sauton *et al.*, 2008). The scoring function predicts the strength of the binding protein-ligand for a given orientation/conformation. Those functions determine directly the accuracy of a docking algorithm and can be classified as: 1) Force field functions, which decompose the binding energy on its components (Van der Waals, electrostatics, stretching/torsion of bonds, etc…); 2) Empirical functions, which calculate the biding energy of a complex through a weighted average of different energies; and 3) Punctuation

functions based on knowledge, which use information from experimental protein-ligand structures (Huang & Zou, 2010).

There are several docking software available today for private and academic use. The most known are DOCK (Moustakas *et al.*, 2006), Molegro Virtual Docker (MVD) (Thomsen & Christensen, 2006), AutoDock (Österberg *et al.*, 2002), GOLD (Jones *et al.*, 1997), ICM (Abagyan *et al.*, 1994), Glide (Friesner *et al.*, 2004; Halgren *et al.*, 2004), LigandFit (Venkatachalam *et al.*, 2003), Protein-Ligand ANT System (PLANTS) (Korb *et al.*, 2009), FlexX (Rarey *et al.*, 1996) and MOE (Vilar *et al.*, 2008). This latter was the one used to perform the docking calculations of the present work.
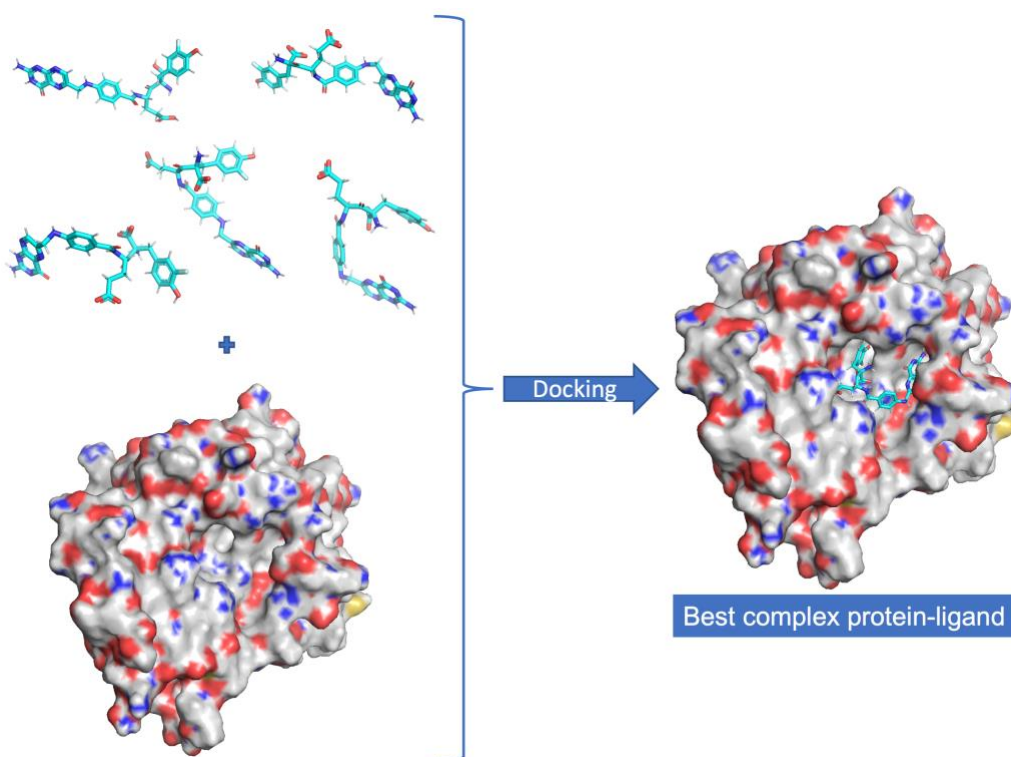
Figure 1.9. Illustration of the protein-ligand docking: Several conformations of the ligand are created and tried into the binding pocket of the protein. The most stable one is selected as the best pose. Figure prepared with the software PyMol® and PowerPoint®.

1.6.3   Molecular dynamics simulations

Despite being a useful tool in the drug design molecular docking is unable to capture the dynamical behavior of a ligand inside its molecular target. This happens because in real systems collisions and conformational changes happen all the time. As the docking approach simulates only static or semiflexible interactions it is not capable of representing exactly what happens in nature. Therefore, in order to push the computational models closer to the real systems, it's

necessary the application of another computational technique known as molecular dynamic (MD) simulation (Hansson *et al.*, 2002). In MD simulations the movement equations are numerically integrated to generate a dynamic trajectory of the system that is useful for the investigation of structural, dynamic and thermodynamic properties of the system from each frame of the trajectory containing a snapshot of the system at that given moment (Braun *et al.*, 2018).

Due to the huge numbers of atoms involved it's impossible to treat biological systems with the more accurate Quantum Mechanics (QM) approach (Liu *et al.*, 2001) which would imply in a prohibitive computational cost. Therefore, approximations have to be applied and the approach for the MD simulations of biological systems have to be performed using Newtonian laws of movement [named here Molecular Mechanics (MM)] to describe the particles of the system. In this approach the forces acting on each atom are estimated considering the bonded interactions (relative to changes in the length, angle and torsions of chemical bonds) and non-bonded interactions [related to the intermolecular interactions, usually electrostatic and Van der Waals (VdW)]. Chemical bonds and angular deformations are approximated to the spring-mass system while the changes in dihedral angles are modeled by sinusoidal functions representing energy oscillations between different conformations. The non-bonding forces originated from VdW interactions are modeled according the Lennard-Jones 6-12 potential (Leach, 2001) while the electrostatic interactions are represented by the Coulomb Law (Braun *et al.*, 2018; Durrant & McCammon, 2011).

The forces acting on each atom in the MM approach are calculated according to a force-field equation like the one shown in Figure 1.9. The three first terms describe the bonded interactions where r and θ represent the bond length and angle, respectively; r_{eq} and θ_{eq} represent the corresponding equilibria length and angles, and K_r and K_θ are the corresponding force constants. In the term of deformations of dihedral angles V_n is the energy barrier for torsion; n is the number of maximum or minima of energy in a complete torsion, φ is the dihedral angle and γ is the de lag angle that can generate a minimal or maximal point at the position $\varphi = \theta$. In the term for the non-bonded interactions A_{ij} and B_{ij} are parameters of the potential 6-12 of Lennard-Jones, R_{ij} represents the interatomic distance, and qi and q_j are the charges of the atoms involved (Leach, 2001).

Bonded interactions · Non-bonded interactions

$$E_{Total} = \sum_{ligands} K_r(r - r_{eq})^2$$

$$+ \sum_{angles} K_\theta(\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2}[1 + \cos(n\varphi - \gamma)] + \sum \left(\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\sum R_{ij}} \right)$$
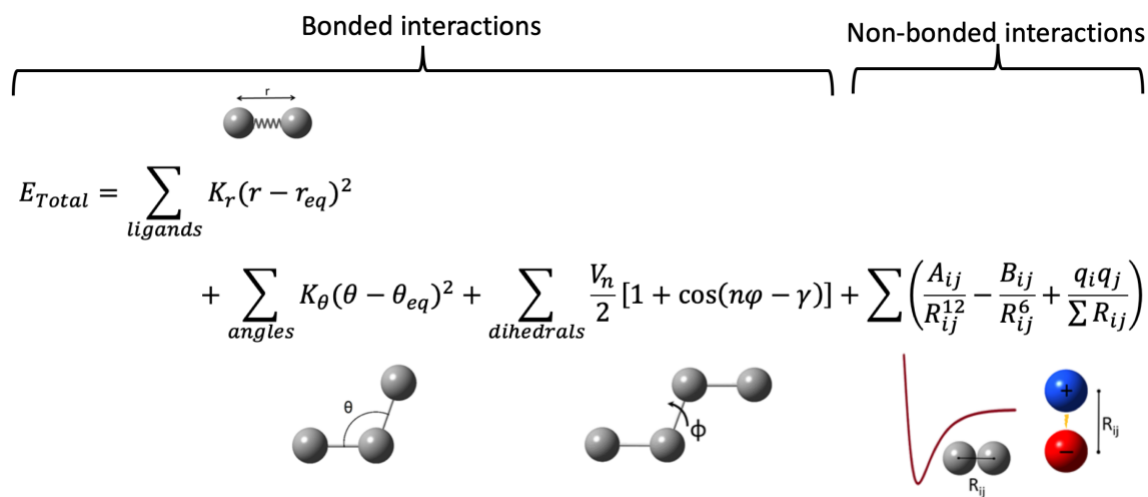
Figure 1.10. Typical energy equation of a force field (Durrant & McCammon, 2011).

Several force-fields have been developed for MD simulations of biological systems. The most common are: Molecular Mechanics Force Field (MMFF) (Halgren, 1996), Groningen Molecular Simulation (GROMOS) (Reif *et al.*, 2012), Chemistry at Harvard Molecular Mechanics (CHARMM) (Vanommeslaeghe *et al.*, 2012), Optimized potentials for Liquid Simulations – All atoms (OPLS/AA) (Jorgensen *et al.*, 1996; Kaminski *et al.*, 2001), and Assisted Model Building with Energy Refinement (AMBER) (Wang *et al.*, 2004).

The force acting on each atom is calculated through the derivative of the potential energy (Figure 1.9) related to the atom coordinate (Leach, 2001). Once the atoms movements can be described by Newton Laws, the acceleration of each atom of mass m_i subjected to a force F_i can be calculated by the equation: $F_i = m_i.a_i$. This will define the new positions of the atoms and, therefore the new configuration of the system (Durrant & McCammon, 2011). The iterative repetition of this process over a given time will reveal the dynamical behavior of the atoms and, consequently, of the whole system.

The first step needed before starting an MD simulation of a biological system is the creation of a simulation box containing the appropriate solvent (usually water) for the molecules under study. After, the appropriate number of ions are added in order to neutralize the charge of the system and reproduce the charge zero of the real system. Then, an energy minimization step is performed in order to assure the appropriate geometries and distances between atoms. After the energy minimization the system is submitted to equilibration steps with the goal of bringing the system to the temperature and density adequate to a real system. Usually this involve two sub-steps: NVT (where Number of particles, Volume and Temperature are kept constant)

(Berendsen *et al.*, 1984; Bussi *et al.*, 2007) and NPT (Number of particles, Pressure and Temperature constants) (Parrinello & Rahman, 2005). The last step of the MD simulation is the production, when position restrictions of ligand solvent and/or protein are removed, and the simulation started for acquisition of dada regarding the system behavior through time and construction of the trajectory.

Several software to run MD simulations are available today. The most known are: AMBER (Wang *et al.*, 2004), GROMOS (Reif *et al.*, 2012), CHARMM (Vanommeslaeghe *et al.*, 2012) (named after their respective force-fields), Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) (Plimpton, 1995), Nanoscale Molecular Dynamics (NAMD) (Phillips *et al.*, 2005; Phillips *et al.*, 2020), and GROMACS (Abraham *et al.*, 2015). These last two where the ones used to develop this work.

1.6.4   Binding energy calculations

The binding energies involving two molecules during a MD simulation of a biological system can be evaluated and quantified through some computational methods. These provide a refinement of the interacting energy values obtained in the docking getting closer to known experimental results (Kumari *et al.*, 2014).

Some of binging energy methods, like the Molecular mechanics – Poisson-Boltzmann Surface Area (MM-PBSA) (Kollman *et al.*, 2000) estimate the binding energies using a set of structures at the initial and final states. This increases the efficiency of the method an make it comparable to the more precise but much more computational costly methods (Kumari *et al.*, 2014). Three terms compose the binding energy calculated according the MM-PBSA method. The first is the potential energy in vacuum including bonding and non-bonding interactions, the second is the term for solvation which considers the polar and apolar solvation energies, and the third is the entropic term associated to the complex in gas phase.

The MM-PBSA method was used to access the binding energy values of the complexes protein-ligand in the present work, through the g-mmpbsa tool (Kumari *et al.*, 2014) from the GROMACS package (Abraham *et al.*, 2015).

2 THEORETICAL SEARCH FOR REPURPOSED DRUGS CAPABLE OF BINDING TO THE TWO POCKETS OF THE SUBUNIT A OF RICIN

Recherche théorique de médicaments reutilisés capables de se lier aux deux poches de la sous-unité A du ricin

Auteurs :

Tanos Celmar Costa Franca[a,b,c], Fernanda Diniz Botelho[b], Michael L. Drummond[d] Steven R. LaPlante[a]

[a]*INRS - Institut Armand-Frappier, 531 boulevard des Prairies, Laval, H7V 1B7, Quebec, Canada.*

[b]*Laboratory of Molecular Modeling Applied to Chemical and Biological Defense, Military Institute of Engineering, Rio de Janeiro, 22290-270, Brazil.*

[c]*Department of Chemistry, Faculty of Science, University of Hradec Kralove, Rokitanskeho 62, 50003, Hradec Kralove, Czech Republic.*

[d]*Chemical Computing Group, Montreal, Quebec H3A 2R7, Canada.*

Contributions of the authors:

This work was conceived, designed, executed and written by Tanos C. C. Franca under the supervision of Prof. Steven R. LaPlante. Fernanda Diniz Botelho helped with the MM-PBSA calculations, and Michael L. Drummond helped in the design of the docking and MD protocols using MOE®, NAMD (Nelson *et al.*, 1996) and Calcul Quebec.

2.1 Abstract

Recently we reported a library of 82 compounds, selected from different databanks through virtual screening (VS) and docking studies, and pointed to 6 among them as potential

repurposed dual binders to both the catalytic site and the secondary binding pockets of the subunit A of ricin (RTA). Here we report additional molecular modeling studies of an extended list of compounds from the original library. Rounds of flexible docking followed by molecular dynamics (MD) simulations and further rounds of MM-PBSA calculations using a more robust protocol, enabled a better investigation of the interactions of these compounds inside RTA, the elucidation of their dynamical behaviors, and updating the list of the most important residues for the ligand binding. Four compounds were pointed as potential repurposed ricin inhibitors that worth being experimentally investigated.

Keywords: Ricin inhibitors; Drug repurposing; Biological warfare agents; Ligand fingerprint; molecular dynamics simulations.

2.2 Introduction

Ricin is a highly toxic plant toxin present in the seeds of *Ricinus communis,* a bush native from the Mediterranean area, Eastern Africa and India, which can also be found in all tropical regions today. Also known as castor oil plant, *R. communis* is the source of the ricin oil, a raw material extensively used in the production of lubricants, green fuel, drugs and cosmetics (Doan, 2004; Sousa *et al.*, 2019). However, the high toxicity of ricin occasionally causes accidental intoxication of workers involved in the oil extraction. Also, the castor oil cake, a byproduct of the ricin oil extraction, is rich in minerals and other nutrients, being, for this reason, used as fertilizer and to feed livestock. As this byproduct is not always totally free of ricin, intoxication or even death of livestock also happen as well (Worbs *et al.*, 2011).

Due to its easiness of obtention, water solubility, and toxicity higher than that of the nerve agents, ricin has also been used as a chemical-biological warfare agent for assassination and terrorism purposes. Many reports of this misuse can be found in the literature (Audi *et al.*, 2005; Janik *et al.*, 2019; Knight, 1979; Pita & Romero, 2014), with the most notorious being the assassination, in London, of the Bulgarian journalist Georgi Markov in September 1978. The murderer used an umbrella adapted with a hypodermic needle at its extremity to inject a tiny sphere stuffed with around 0.4 mg of ricin, and covered with wax, in Markov's leg, causing his death four days after. Literature also reports the tentative of aerolisation or mass production of ricin by terrorist groups in 1989 (Audi *et al.*, 2005) and 2010 (Pita & Romero, 2014), besides several cases of letters contaminated with ricin being delivered to public authorities in the USA in the 2000's and 2010's (Audi *et al.*, 2005; Musshoff & Madea, 2009).

Ricin is a N-glycosidase classified as a type II ribosome-inactivating protein (RIP). This family of proteins found in bacteria and plants, is capable of inhibiting the protein synthesis in eukaryotic cells, playing an important defense role against pathogens and insects (ENDO *et al.*, 1988; Funatsu *et al.*, 1991; May *et al.*, 1989; Zhou *et al.*, 1994). Despite sharing similar mechanisms of action and being structurally related, the RIP types I and III are not as cytotoxic as ricin and the others RIP type II, like abrin. This happens because those RIP lacks the B domain responsible to enable their entrance into the cells (ENDO *et al.*, 1988; Funatsu *et al.*, 1991; May *et al.*, 1989; Zhou *et al.*, 1994).

The mechanism of action of ricin comprises a permanent damage to the rRNA due to the abstraction of the adenine 4324 (A-4324) from the loop GAGA of the rRNA 28S, which is highly conserved in eukaryotic cells. This causes an interruption in the protein's synthesis and lead to cell death (Lord *et al.*, 1994; Olson *et al.*, 2004). The ricin unit responsible for this is known as RTA (or Ricin Toxin A) which unites to RTB (Ricin Unit B) through a disulfide bond to compose the whole ricin structure. Once inside the endoplasmic reticulum of the cell RTA is separated from RTB, due to the action of the enzyme disulfide isomerase, and moves to the cytosol where it will promote the rRNA damage. To date there is no antidote against such action yet and the vaccines developed so far to prevent ricin intoxication are not effective (Gal *et al.*, 2015; Hu *et al.*, 2013; Legler *et al.*, 2011; Olson *et al.*, 2004; Roy *et al.*, 2015).

The search for antidotes against ricin has already afforded many compounds capable of binding to the catalytic site of RTA (comprised of residues Val81, Gly121, Glu177 and Arg180) and work as competitive inhibitors (Pruet *et al.*, 2012; Saito *et al.*, 2013; Wiget *et al.*, 2013). However, no activity below the micromolar range was achieved yet, with the most promising compounds reported so far (Figure 2.1) presenting IC_{50} between 6 and 20 μM (Saito *et al.*, 2013). The publication of the crystallographic structure of NNCP (Figure 2.1) complexed at the catalytic site of RTA in the Protein Data Bank (PDB) (https://www.rcsb.org/) under the ID: 4HUO, elucidated the binding mode [see Figure 5.1(a) in Annex I] of those compounds. Besides it also revealed that they are not capable of also binding to the secondary binding pocket of RTA (comprising residues Asp75, Asn78, Asp96 and Asp100). This pocket had been previously revealed by Ho et al. (2009) (Ho *et al.*, 2009) through the crystallographic structure of RTA complexed with the cyclic tetranucleotide inhibitor C2X [see Figure 5.1(b) in Annex I], which mimics the recognition loop of 28S rRNA [PDB (https://www.rcsb.org/) ID: 3HIO]. This cyclic tetranucleotide mimics the recognition loop of the 28S rRNA, and was designed by Ho et al. (2009) (Ho *et al.*, 2009) with the goal of establishing the catalytic site features contributing to the RTA

catalytic activity. It's also a transition state structure meant to guide the design and synthesis of potent RIP inhibitors. Analysis of 4HUO [Figure 5.1(b)] shows clearly that there is still enough room for the drug design/discovery of new and more potent inhibitors, targeting at the same time the catalytic and secondary binding pockets of RTA.

N-{[1-(phenylthio-methyl)-1H-1,2,3-triazol-4-ill]methyl}-pterin-7-carboxamide
$IC_{50} = 15 \ \mu m$

N-[N-(pterin-7-il)carbonilglycil]-L-phenylalanine (NNCP)
$IC_{50} = 20 \ \mu m$

N-[N-(pterin-7-il)carbonilglycil]-L-tyrosine (NNCT)
$IC_{50} = 6 \ \mu m$
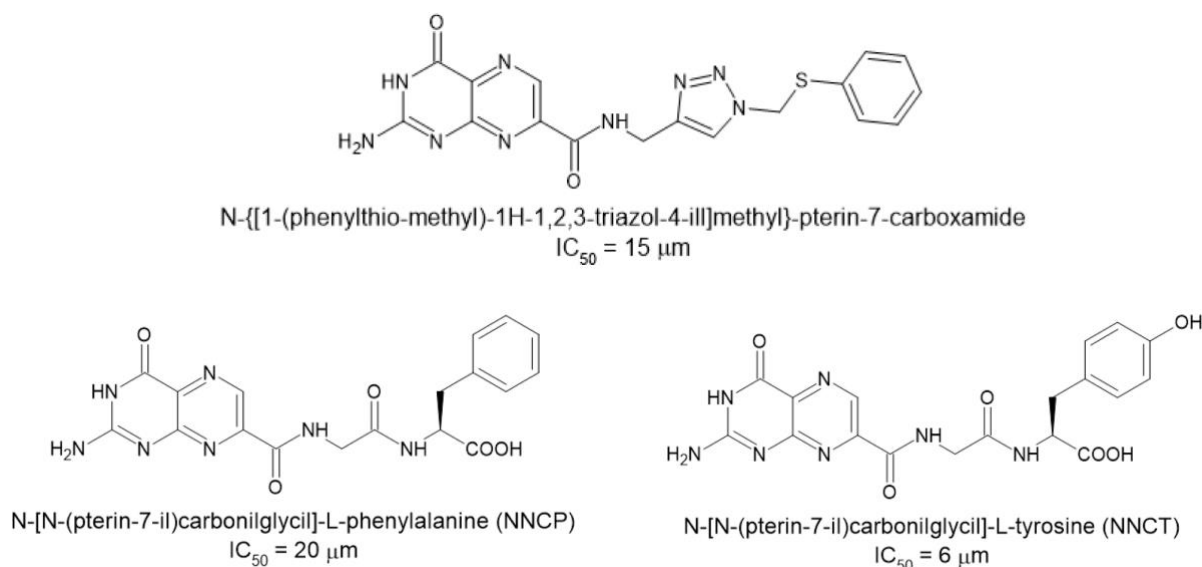
Figure 2.1. Structures of the best competitive inhibitors of RTA currently reported in the literature (Saito et al., 2013).

Recently we proposed the repurposing approach as a promising strategy to follow in the search for more effective antidotes against ricin intoxication capable of binding into both the catalytic and the secondary pockets of RTA. On this line we performed virtual screening (VS) searches in the FDA-approved drugs data set, available at Cheminfo (https://chemoinfo.ipmc.cnrs.fr/), the approved drugs library available at DrugBank (https://www.drugbank.ca/), and the PubChem database (https://pubchem.ncbi.nlm.nih.gov/). This search enabled the selection of 6795 potential binders to RTA which, after additional refinements, afforded a list to 82 compounds. Further molecular dynamics simulations on 20 compounds selected from this list enabled pointing 6 (see Figure 5.2 in Annex I) as potential antidotes against RTA (Botelho *et al.*, 2020b; Botelho *et al.*, 2020d). Here we moved forward on this project through additional theoretical studies on these 6 molecules shown in Figure 5.2 plus 9 more selected compounds (Figure 2.2). Flexible docking of these compounds enabled drawing their fingerprints inside RTA and plotting the most important residues for the ligand binding. Also, longer molecular dynamics simulations of 500 ns of the best poses of these compounds inside

RTA enabled reevaluating the former results and selecting the ones that should keep the interactions inside the two pockets, besides ranking them according to the potential of inhibiting RTA.
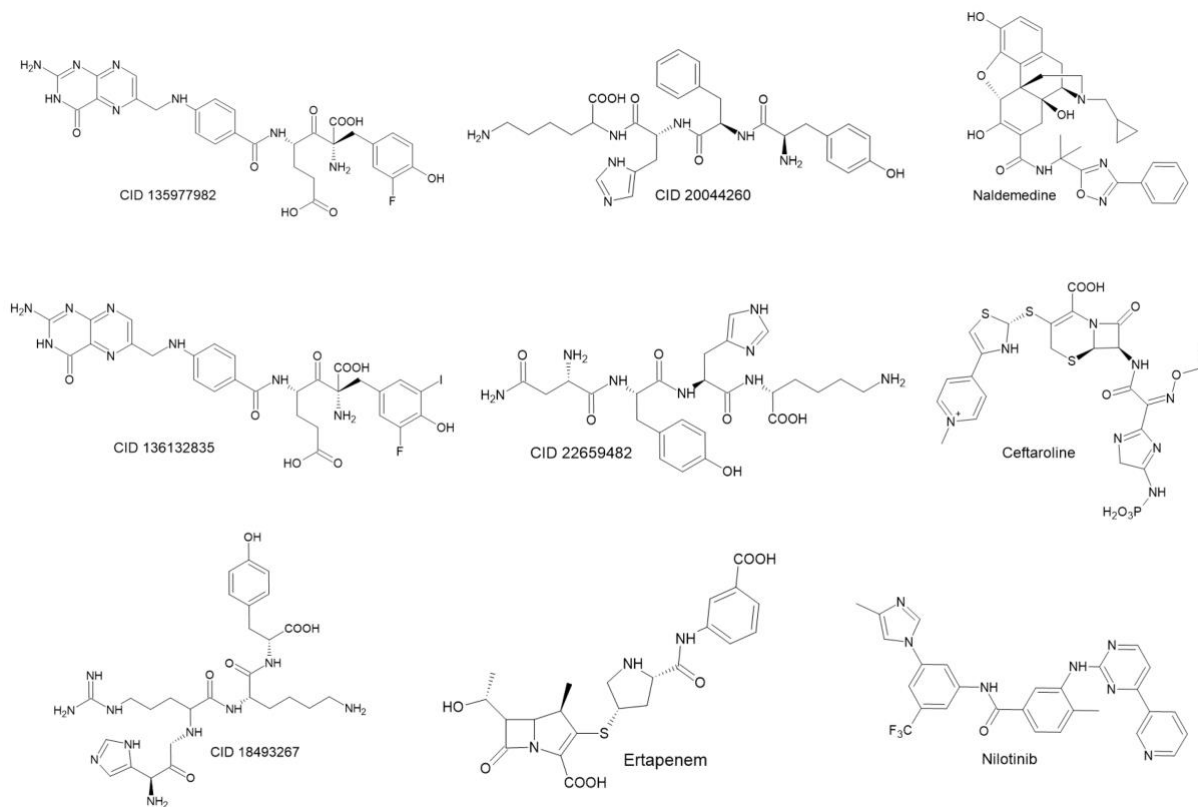
Figure 2.2. 2D structures of the 9 additional compounds selected for the current study.

2.3   Methodology

2.3.1   Protein preparation

The receptor used to perform our studies was the three-dimensional structure of RTA in complex with the cyclic tetranucleotide inhibitor C2X, which mimic the sarcin-ricin recognition loop of the 28S rRNA, available in the PDB (https://www.rcsb.org/) under the code: 3HIO (Ho *et al.*, 2009). This structure was downloaded and optimized using the default configuration of the quickprep tool of the MOE package (https://www.chemcomp.com/Products.htm), in order to remove gaps, optimize bond lengths and angles, calculate charges, and properly protonate residues and ligand according to the physiologic pH. Crystallographic water molecules were also removed, and only ligand and RTA were kept in the structure for the theoretical studies.

2.3.2 Ligand preparation

The 3D structures of the ligands studied here (Figures 2.2 and 5.2), were constructed using the builder tool of the MOE® package (https://www.chemcomp.com/Products.htm) and added to a MOE® databank named as "ligands.mdb". After, each entry of the databank was "washed", using the compute/molecule/wash tool of the MOE® package (https://www.chemcomp.com/Products.htm), in order to optimize bond lengths, angles and charges, and also to afford the dominant species of each ligand under physiologic pH.

2.3.3 Docking studies

The entries of the *ligands.mdb* file described above were submitted to rounds of docking calculations on the active site of the crystallographic structure of RTA (PDB ID: 3HIO), using the *dock* module of the MOE package (https://www.chemcomp.com/Products.htm). The docking grid was defined to include all residues present in the whole binding pocket of C2X and the docking placement method used was Triangle Matcher with induced fit refinement and the *generate fingerprints* options active. First, the 100 best ranked poses (lowest energies) of each ligand were collected using the London dG scoring function and, after, those poses were re-scored using the GBVI/WSA dG function with the 20 best results being collected for further analysis of the fingerprints of each ligand inside RTA. This docking protocol was validated by re-docking of the crystallographic structure of C2X inside 3HIO where the five best ranked poses were collected. The docking energy of NNCP inside RTA was obtained after energy minimization of the crystallographic structure 4HUO. Regarding NNCT, once its reported in the literature that it binds the same way as NNCP (Saito *et al.*, 2013), the docking energy was calculated through energy minimization of 4HUO after changing the structure of NNCP to NNCT by adding an –OH group in the *para* position of the NNCP phenyl ring.

2.3.4 MD simulation

The *dynamics* tool of the MOE package (https://www.chemcomp.com/Products.htm) was used to run the MD simulations. The Compute/Simulations/Dynamics path was used to prepare each system according to the parameters of the NAMD (Nelson *et al.*, 1996) software and using the forcefield AMBER10:EHT (Case *et al.*, 2008), with cut off of 10 for electrostatic and 8,10 for VdW interactions. Each complex was centered in a cubic box containing around 9,000 water molecules and neutralized with NaCl ions. The rounds of simulations involved first 10 ps of an energy minimization step followed by 100 ps of NPT and 200 ps of NVT. After, production steps of 500

ns of free MD simulation were performed. The MD simulation results were analyzed using the md_analysis tool and the database viewer (DBV) menu of MOE (https://www.chemcomp.com/Products.htm). The trajectory files of each system, used for analysis of the MD simulations, were generated after superposition of the initial and simulated frames of protein and ligand.

2.3.5   MM-PBSA calculations

Hybrid sequential QM/MM–MD methods are amongst the most accurate ways to estimate the free energy associated to the protein-ligand binding (Lipparini & Mennucci, 2021). However, as many configurations are generated in large MD simulations, the number of QM calculations required is too high, since in every step an energy evaluation of the system is needed. Therefore, a great computational effort is necessary to carry out this kind of simulation. Aiming, then, to reduce the number of QM calculations without loss of the relevant information from the simulation, new methods based on the statistical inefficiency and wavelet analysis for selecting MD conformations had been reported in the literature (Gonçalves *et al.*, 2017). However, the computational cost involved is still high. One alternative to reduce this computational cost with an acceptable impact on the accuracy of the calculations is the use of the MM-PBSA approach (Kumari *et al.*, 2014), which has become one of the most used method for estimation of free energy in biological systems (Homeyer & Gohlke, 2012).

Despite not explicitly calculating the entropy effect, the MM-PBSA method considers bonded and nonbonded interactions (vacuum potential energy), as well as polar and nonpolar terms (free energy of solvation) (Kumari *et al.*, 2014). The Poisson-Boltzmann equation is used to estimate the solvation energy term (usually with dielectric constant set to 1), while the surface area (SASA) method (Kumari *et al.*, 2014) is used to calculate the nonpolar solvation energy term.

Here the MM-PBSA calculations were performed as before (Botelho *et al.*, 2020a; Botelho *et al.*, 2020c) using the g_mmpbsa tool, compatible with the GROMACS 2019.4 software (Kumari *et al.*, 2014). For this it was necessary first to run 100 ns of MD simulation of the best poses of the ligands inside RTA obtained in the docking studies, in order to afford the trajectory frames needed to feed the g_mmpbsa tool. Those MD simulations were performed following the same protocol used before (Botelho *et al.*, 2020a; Botelho *et al.*, 2020c) and described in detail in the supplementary material. The g_mmpbsa tool was used to predict the binding free energy

of the ligands that showed capable of keeping interactions in both pockets of RTA during the 500 ns of MD simulations with NAMD (Nelson *et al.*, 1996).

2.4   Results and discussion

As mentioned in the introduction we studied in this work the 6 molecules pointed before as potential antidotes against ricin (see Figure 5.2 in Annex I) (Botelho *et al.*, 2020a; Botelho *et al.*, 2020c) plus 9 additional compounds (Figure 2.2) originally selected by VS, and pointed by docking studies as potential dual binders to RTA, but not investigated through MD simulations. By extending the study to these 9 compounds we advanced one more step towards the ultimate goal of investigating through MD simulations the whole library of 82 compounds former selected by VS (Botelho *et al.*, 2020a; Botelho *et al.*, 2020c). These 9 compounds were selected based on a visual inspection to ensure a structural diversity that represent the most the whole library.

2.4.1   Docking study

The total energy values of the five poses returned from the re-docking of C2X over its crystallographic structure (3HIO) ranged from -15.23 to -15.62 kcal.mol^{-1}, while the root-mean-square deviation (RMSD) values ranged from 1.66 to 2.01 Å. Since a RMSD < 2.0 Å is considered valid according to the literature recommendation (Kontoyianni *et al.*, 2004), the poses obtained during our re-docking studies were good enough to validate the docking protocol used. The superposition with the lowest RMSD observed (1.66 Å) is shown in Figure 5.3 in Annex I.

It's well known that an appropriate theoretical strategy for selecting docking poses is crucial for the subsequent steps of MD simulations in order to avoid conformational changes in the time scale of few nanoseconds, which could compromise the MD simulation viability. For this reason, distinct selecting poses approaches have been reported in literature (da Cunha *et al.*, 2008; Farahani *et al.*, 2022; Franca *et al.*, 2005; Franca *et al.*, 2006; Guimaraes *et al.*, 2011; Santos *et al.*, 2022). In the current work, as we are looking for potential dual binders to RTA, we selected the best ranked pose (lowest energy) of each ligand showing interactions with at least one residue of both the catalytic and secondary pockets of RTA. The only exception was C2X which the pose selected was the one with lowest RMSD from the re-docking, shown in Figure 5.3. The docking results obtained for the selected poses are summarized in Table 2.1.

As shown in Table 2.1 no ligand presented more negative energy value compared to C2X. This was already expected once this compound mimics the sarcin-ricin recognition loop of the 28S rRNA. However, the best poses of most ligands presented total energy more negative than NNCP and NNCT. This suggests a higher affinity for the RTA active site than the reference ligands and, therefore, a higher potential to act as more effective RTA inhibitors. Regarding the 9 additional compounds included in this work, all ranked in the same range of values (between -

7.44 and -9.74 kcal.mol^{-1}) as the 6 compounds repurposed before(Botelho *et al.*, 2020a; Botelho *et al.*, 2020c), with the most promising among all these molecules being CID 135977982 and CID 136132835 ranking respectively, -9.74 and -9.72 kcal.mol^{-1}. This means that they might also be good options of repurposed ligands against RTA.

Results in Table 2.1 also show that amongst the residues not belonging to the catalytic or the secondary pockets Arg213 and Arg258 showed up in interactions with most of the ligands. These two residues are located in the border of the active site of RTA which includes the catalytic or the secondary pockets, but in opposition to them, as illustrated in Figure 5.4, which shows the best pose of CID 135977982 inside RTA. As far as we know Arg213 and Arg258 have not been reported yet in the literature as important residues for the binding of inhibitors to RTA.

Table 2.1. Interactions of the best poses of each ligand after induced fit docking inside RTA. Residues of the catalytic site are shown in red while the ones of the secondary pocket are shown in blue.

Ligand	Interacting residues (interaction type)	Energy (kcal mol^{-1})
C2X	Asp75(DD), Asn78(DDAA), Val81(ddaa), Gly121(dd), Asp124(DD), Glu177(II), Arg180(AA), Asn209(a), Arg213(AAaII), Thr216(AA), Arg258(AAII)	-15.31
NNCP	Tyr80(R), Val81(dd), Gly121(dd), Asn122(a), Tyr123(aa), Arg180(AA), Gly212(a), Arg258(AAII)	-8.44
NNCT	Tyr80(R), Val81(dd), Gly121(dd), Asn122(a), Tyr123(aa), Arg180(AA), Gly212(a), Arg258(AAII)	-8.39
CID 135977982	Pro95(RR), Asp96(DD), Arg180(AI), Arg213(AAaaII), Arg258(AA)	-9.74
CID 136132835	Asp75(DD), Glu177(DDII), Arg180(AAII), Glu208(dd), Arg258(AAII)	-9.72
CID 136023163	Asp75(DD), Asn78(AA), Gly121(d), Glu208(d), Gly212(a), Arg213(AAII), Arg258(AAII)	-9.40
CID 20044260	Asn78(A), Asp96(DD), Asn122(A), Glu177(DDII), Glu208(dd), Arg258(AAII)	-9.35
Ceftaroline	Asp75(I), Asn78(AA), Asp96(I), Asp100(I), Arg180(II), Gly212(a), Arg213(I), Arg258(AAII)	-9.19

Ligand	Interacting residues (interaction type)	Energy (kcal mol^{-1})
Deferoxamine	Asp96(DII), Asp100(DDII), Glu177(DD), Arg180(AA), Arg213(AA)	-9.16
CID 18498053	Asn78(A), Glu177(D), Glu208(dd), Arg258(AAII)	-9.01
Naldemedine	Asn78(AA), Arg180(A), Asn209(D)	-9.00
CID 18493267	Asn78(A), Pro95(RR), Gly121(dd), Arg258(AAII)	-8.94
CID 22659482	Asp96(DD), Asn122(AA), Glu177(DDII), Glu208(d), Arg258(AAII)	-8.79
CID 18309602	Asp96(DDIIRR), Asp100(DDII), Glu177(DDII), Arg180(AAII),	-8.70
Nilotinib	Asp75(DD), Asn78(R), Gly121(a)	-8.32
Plazomicin	Asn78(AA), Asp96(DI), Asp100(II), Tyr123(R), Glu177(DDII), Glu208(d)	-8.23
Leucovorin	Asp96(D), Gly121(a), Asn122(AA), Arg213(AAII), Arg258(AAII)	-7.53
Ertapenem	Asp96(DR), Glu177(DD), Arg258(AAII)	-7.44

*D = donor to side chain; A = Acceptor from sidechain; I = Ionic attraction; d = donor to backbone; a = acceptor from backbone; R = Arene attraction.

2.4.2 Molecular Dynamics simulation

The plots of total energy for the 18 complexes studied by MD simulation (see Figures 5.5 - 5.10 in Annex I) show that all systems stabilized since the beginning of the simulation time, with an average energy ranging between -0.9 x 10^{-6} and -1.3 x 10^{-6} kcal.mol^{-1}. Also, the RMSD plots calculated after superposition of initial and simulated structures for RTA and ligands (Figure 2.3), show that the fluctuations never passed 6.0 and 2.5 Å for the ligands and protein, respectively, except for plazomicin that was not capable of stabilizing inside RTA and left after 50 ns of MD simulation (see also movie: Plazomicin_RTA.mpg in the supplementary information). For the other 17 systems the total energy and RMSD results point to stability over the simulated time and accommodation of the ligands inside RTA.

Table 2.2 lists all H-bonds formed during the MD simulations with their respective percentages of occupancy while the plots for the ligands capable of dual biding to RTA during more than 10% of the simulated time are shown in Figures 5.11 - 5.14 in Annex I. In those Figures the graphs in red lines represent the number of H-bonds formed with residues of the catalytic pocket of RTA, while the graphs in blue lines represent H-bonds formed with residues of the secondary pocket. As can be seen most of the interactions pointed in the docking studies (Table 2.1) were confirmed by the MD simulations, with some additional interactions showing up during the MD simulations. It can also be seen in Table 2.2 that both Arg213 and Arg258 show up h-bonding to C2X with occupancies > 10%. Arg258 also shows up h-bonding to many other ligands with high percentages of occupancy. This corroborates the docking findings and suggest that these two residues might be key for the binding into RTA.

Analysis of Table 2.2 and Figure 5.11 show that C2X was capable of maintaining H-bonds with the residues of both RTA binding sites during the whole simulated time, showing occupancies over 10% with at least three residues from each site. Also, the RMSD plot of C2X (see its red line in Figure 2.3) show stabilization around 3.5 Å, what means a good accommodation inside RTA. This is well illustrated in the movie: C2X_RTA.mpg in supplementary information and by the frames shown in Figure 2.4 (where the external loop of C2X was omitted for better clarity). As can be seen, the cytosine and guanosine portions of C2X, found in the catalytic and secondary sites of RTA in the crystallographic structure (PDB code 3HIO), stay inside the pockets during the whole simulation time. Therefore, our theoretical results confirmed that C2X is capable of binding and stabilizing inside RTA, mimicking the recognition loop of 28S rRNA. This corroborates the results reported by Ho et al. (2009) (Ho *et al.*, 2009) and the mechanism of action reported in the literature for RTA (Ho *et al.*, 2009).

Only CID 20044260, CID 22659428, Naldemedine, CID 135977982, and CID 136132835 kept H-bonds with residues of both sites of RTA during the whole simulated time (see Figures 5.11-5.14). Analysis of the dynamical behavior of those compounds (see the respective movies supplied as supplementary information) revealed that only Naldemidine, CID 135977982, and CID 136132835 showed a good accommodation inside RTA as can be also observed in the superposition of frames collected during the MD simulations shown in Figure 2.4. This is also reflected in Figure 2.3 which show very little RMSD fluctuations for these compounds. CID 20044260 and CID 22659428, besides showing larger RMSD fluctuations (Figure 2.3), also didn't accommodate well inside RTA as can be seen in the superposition of frames collected during the MD simulations shown in Figure 5.15 and the respective movies in the supplementary information.
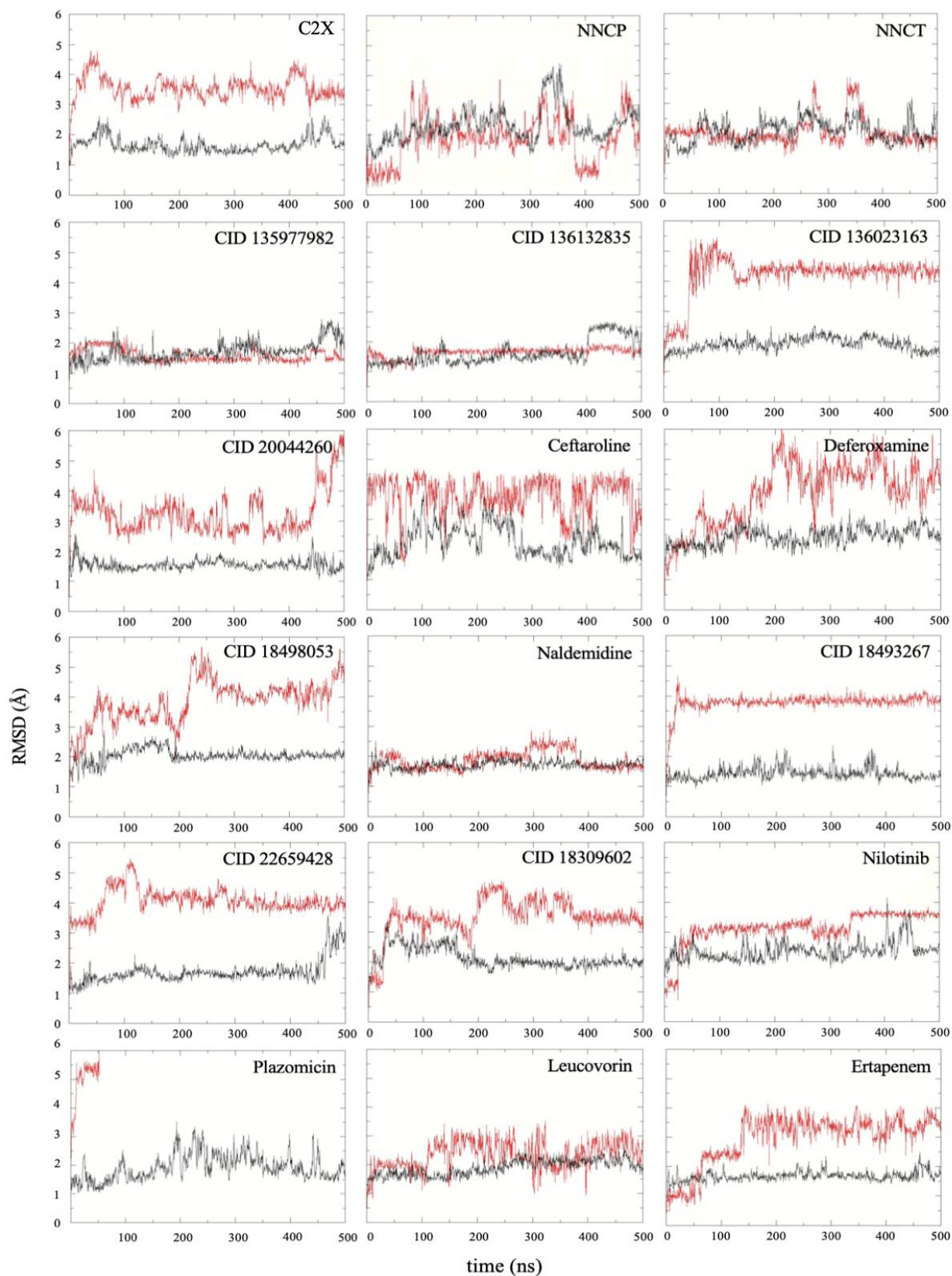
Figure 2.3. RMSD plots for the 18 systems during the MD simulations. Black lines = RTA, Red lines = ligands.

31

The large RMSD fluctuations observed in some systems are due to unstable interactions between protein and ligand which make both to move more. Larger fluctuation means worse behavior and, therefore, not good accommodation of the ligand inside RTA. They reflect changes in positions of the ligands during the MD simulations and the corresponding effect in the RTA behavior. These movements can be observed in the movies of the MD simulations supplied as supplementary information.

Table 2.2. H-bonds formed during the MD simulations. Residues of the catalytic site are shown in red while residues of the secondary sites are shown in blue.

Ligand	Interacting residues (interaction type)	
	Acceptor	Donor
C2X	Asp75(26.60%), Asn78(42.90%), Val81(4.00%), Asp100(0.20%), Gly120(0.20%), Gly121(68.30%), Asn122(0.10%), Thr123(0,10%), Asp124(1.00%), Glu177(0.20%), Glu208(2.80%), Asp209(1.60%), Arg213(0.70%), Thr216(0.10%), Ser228(2.20%), Pro229(3.00%), Gln231(0.40%),	Arg48(7.40%), Asn78(42.80%), Val81(56.70%), Asp96(62.80%), Arg180(36.30%), Asn122(22.70%), Thr123(0,30%), Asp124(1.10%), Arg125(15.90%), Asn209(52.70%), Arg213(24.50%), Thr216(6.20%), Ser228(0.40%), Gln231(0.10%), Arg258(51.60%),
NNCP	Val81(41.40%), Gly121(41.30%), Asn122(6.80%), Glu208(0.10%),	Asn78(0.30%), Tyr80(0.70%), Val81(41.60%), Asn122(6.6%), Tyr123(15.00%), Arg125(0.20%), Ser176(2.20%), Arg180(58.90%), Asn209(21.50%), Arg213(0.50%)
NNCT	Tyr80(0.10%), Val81(40.00%), Asp96(0.20%), Gly121(36.90%), Asn122(7.30%), Asp124(0.20%), Glu177(13.10%), Glu208(0.20%),	Tyr80(0.50%), Val81(39.60%), Asn122(5.5%), Tyr123(4.10%), Arg125(0.20%), Ser176(2.90%), Arg180(53.10%), Asn209(5.70%), Trp211(0.10%), Arg213(0.50%)
CID 135977982	Val81(41.60%), His94(0.30%), Asp96(23.30%), Asp100(0.20%), Gly121(41.40%), Glu177(0.70%), Arg180(0.90%)	Arg48(109.70%), Asn78(33.30%), Tyr80(7.70%), Val81(45.20%), Asp96(0.40%), Gly121(0.70%), Asn122(11.70%), Tyr123(6.60%), Ser176(1.00%), Arg180(45.50%), Arg213(0.10%), Arg258(80.10%)
CID 136132835	Val81(50.10%), Gly121(45.00%), Asn78(50.90%), His94(17.50%), Asp96(10.20%), Arg180(1.60%),	Arg48(102.10%), Asn78(8.10%), Tyr80(1.10%), Val81(31.30%), Asn97(2.00%), Gly121(34.60%),

Ligand	Interacting residues (interaction type)	
		Tyr123(4.60%), Ser176(2.30%), Arg180(57.00%), Arg258(114.10%)
CID 136023163	Asp96(17.10%), Glu208(3.50%), Glu220(0.80%), Tyr257(47.30%), Arg258(0.20%), Asn209(1.20%)	Asn47(2.40%), Asn78(1.00%), Tyr80(0.10%), Asn97(9.00%), Asn122(2.50%), Asp124(0.10%), Arg134(0.20%), Arg180(0.20%), Asn209(0.10%), Arg213(3.60%), Thr216(29.20%), Tyr257(0.10%), Arg258(120.20%)
CID 20044260	Thr77(0.20%), His94(0.30%), Asp96(3.10%), Gly120(1.00%), Gly121(6.60%), Asn122(2.60%), Glu177(72.40%), Asn209(0.10%)	Arg48(32.10%), Asn78(33.90%), Asp96(1.00%), Asn97(1.20%), Asn122(3.10%), Tyr123(27.00%), Asp124(0.40%), Arg125(0.30%), Arg180(2.00%), Arg258(15.10%)
Ceftaroline	Asp96(0.20%), Ala260(0.10%), Pro261(0.10%), Gln266(0.20%)	Asn47(0.70%), Arg48(36.40%), Asn78(0.40%), Asp96(0.20%), Asn97(0.40%), Arg213(31.90%), Thr216(4.90%), Arg258(22.40%), Cys259(5.80%), Ala260(0.10%), Ser265(0.10%), Gln266(0.30%),
Deferoxamine	Asn78(0.30%), Tyr80(3.00%), Val81(4.90%), His94(25.50%), Pro95(0.30%), Asp96(10.00%), Asp100(0.80%), Ala118(0.40%), Gly120(3.80%), Gly121(1.30%), Asn122(0.20%), Tyr123(0.10%), Asp124(4.30%), Glu127(18.90%), Glu135(7.20%), Ile205(0.50%), Thr206(0.10%), Glu208(8.90%), Asn209(0.70%), Pro229(0.30%),	Arg48(0.10%), Tyr80(3.10%), Asn78(12.30%), Asp96(6.20%), His94(0.30%), Asn97(3.10%), Gln98(0.10%), Asn122(3.90%), Gly120(0.10%), Gly121(1.90%), Tyr123(15.10%), Asp124(0.90%), Arg125(1.50%), Arg134(0.20%), Arg180(18.80%), Thr206(0.50%), Asn209(3.90%), Arg213(1.50%), Gln231(0.90%), Gln233(0.20%), Arg258(0.40%),
CID 18498053	Asn47(1.40%), Thr77(0.10%), Asn78(0.10%), Tyr80(0.30%), Pro95(0.30%), Asp96(32.70%), Gly121(0.20%), Asn122(0.50%), Asp124(1.90%), Glu208(0.10%), Asn209(0.60%), Thr216(0.20%), Tyr257(5.50%), Arg258(0.30%)	Asn47(0.50%), Arg48(5.60%), Asn78(21.10%), Tyr80(0.20%), Asn97(1.00%), Gln98(0.10%), Gly120(0.10%), Gly121(0.60%), Asn122(1.10%), Arg125(0.40%), Arg180(4.00%), Asn209(2.70%), Gly212(0.10%), Arg213(2.60%), Arg258(36.70%)
Naldemedine	Asn78(70.90%), Glu208(66.40%), Asn209(0.10%)	Arg48(0.20%), Asn78(0.30%), Asn122(3.60%), Arg180(26.10%),
CID 18493267	Asp75(15.80%), Asn78(2.20%), Tyr80(0.20%), Phe93(7.60%), His94(42.90%), Pro95(4.50%),	Arg48(28.60%), Asn78(62.10%), Val81(0.10%), Asp96(1.40%), Asn122(0.60%), Arg125(0.30%),

Ligand	Interacting residues (interaction type)	
	Asp96(0.50%), Asp100(0.30%), Asn122(9.60%), Asp124(7.00%), Glu177(0.10%), Glu208(15.50%), Asn209(2.70%),	Ser176(0.30%), Arg180(1.70%), Asn209(0.10%), Trp211(0.10%), Arg258(85.20%)
CID 22659482	Asp75(0.80%), Thr77(0.10%), Asn78(31.30%), Tyr80(0.30%), Val81(11.50%), His94(0.40%), Asp96(8.70%), Asp100(0.80%), Gly121(3.90%), Asp124(0.90%), Ser176(0.10%), Asp209(1.00%), Glu177(1.30%), Glu208(0.90%), Asn209(0.90%), Asn122(7.70%),	Arg48(12.90%), Asp78(30.00%), Asn97(0.20%), Gly121(0.50%), Asn122(0.50%), Tyr123(21.80%), Arg180(12.10%), Trp211(0.10%), Gly212(3.60%), Arg213(1.50%), Arg258(38.20%)
CID 18309602	Asp75(56.10%), Asp78(1.00%), Asp96(138.00%), Asn97(0.10%), Asp100(46.60%), Asp124(1.80%), Glu177(0.50%), Glu208(25.10%), Asn209(0.40%), Trp211(0.10%),	Arg48(11.60%), Asn78(0.70%), Asn97(5.00%), Asn122(1.00%), Tyr123(1.20%), Arg180(2.00%), Asn209(4.00%), Gly212(0.80%), Arg213(18.20%), Arg258(9.30%)
Nilotinib	His94(0.20%), Asp96(0.10%), Asp100(17.70%)	Arg48(0.20%), Asn78(38.20%), Asn122(0.10%), Tyr123(0.10%), Arg180(0.40%), Arg258(0.20%)
Plazomicin	Asp75(4.00%), Asp78(0.20%), Tyr80(0.40%), Val82(0.60%), Glu99(0.30%), Phe93(0.60%), Asp96(12.60%), Asp100(6.10%), Asn122(0.90%), Tyr123(0.10%), Asp124(0.70%), Glu208(3.10%), Asn209(0.30%), Glu220(0.10%), Arg258(0.20%)	Asn47(0.10%), Asp78(4.70%), Tyr80(1.00%), Phe93(0.10%), Asp96(0.10%), Asn122(0.10%), Arg180(1.40%), Arg258(0.10%)
Leucovorin	Asp75(21.80%), Asp78(22.00%), Asp96(0.10%), Asp100(42.10%), Glu99(0.10%),	Arg48(29.50%), Arg56(0.50%), Asp78(40.00%), Asn97(0.20%), Asn122(13.80%), Arg125(2.60%), Arg180(0.40%), Asn209(0.10%), Arg213(1.10%), Arg258(66.90%)
Ertapenem	Asp96(17.10%), Gly121(0.10%), Glu177(10.30%),	Arg48(18.50%), Asn78(8.10%), Asp96(1.30%), Asn97(1.30%), Tyr80(6.10%), Asn122(0.60%), Tyr123(34.00%), Gly121(0.10%), Asn209(2.50%), Arg213(6.20%), Thr216(0.60%), Arg258(46.10%)
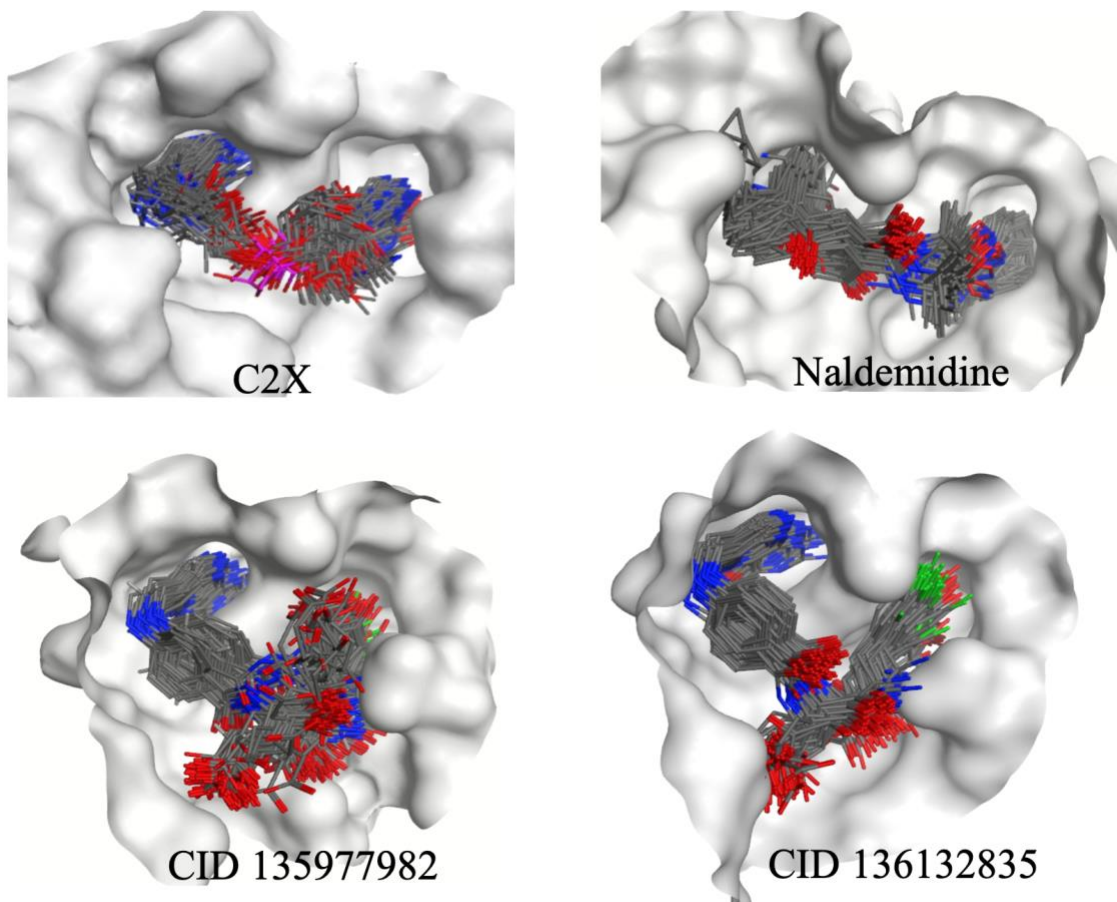
Figure 2.4. Superposition of 100 frames of C2X, Naldemedine, CID 135977982 and CID 136132835 collected during the MD simulations. Hydrogens and the external loop of C2X were omitted for better clarity. The receptor surface is represented in gray. Figure made with MOE®.

Analysis of Figure 5.15 and the movies supplied as supplementary information show that most of the no dual binders didn't stabilize well inside RTA during the MD simulations. Regarding NNCP and NNCT, while the pterinic rings stabilized inside the catalytic site, the other moiety of those molecules didn't find where to anchor and remained unstable the whole simulation. This is a possible reason for the IC_{50} in the range of μM reported for these compounds (Saito *et al.*, 2013). Similar behavior was also observed for CID 136023163, CID 18498053, CID 20044260, Deferoxamine, CID 18309602, Ertapenem, Leucovorin and CID 22659482. Ceftaroline on the other hand moved from both RTA binding sites since the beginning of the simulation and, despite staying close to RTA didn't accommodate well inside it. The only ligands in Figure 5.15 that showed stable dynamic behavior inside RTA were nilotinib and CID 18493267 which, despite not

behaving as dual binders stabilized during the MD simulation, keeping interactions inside the secondary pocket and at the entrance of the catalytic site (see movies: Nilotinib_RTA.mpg and CID18493267_RTA.mpg supplied as supplementary information). This means that these two ligands might also bind strongly to RTA, avoiding the interactions with the loop GAGA of the rRNA 28S.

2.4.3   MM-PBSA calculations

The bars plot in Figure 2.5 show the values of binding energies predicted through MM-PBSA calculations for the reference compounds and the potential dual binders pointed in the MD simulations. As expected, all of them presented negative binding energies, with a good correlation with the docking and MD results. Fact that confirms their affinities for RTA. It's interesting to notice, however, that CID 135977982 and CID 136132835 showed better (more negative) binding energies than C2X. This is probably due to the cyclic part of C2X with didn't stablish stable interactions inside RTA and fluctuated during the MD simulation (see movie C2X_RTA.mpg in the supplementary information). Those fluctuations might be adding unfavorable interactions which contribute to increase the binding energy value.
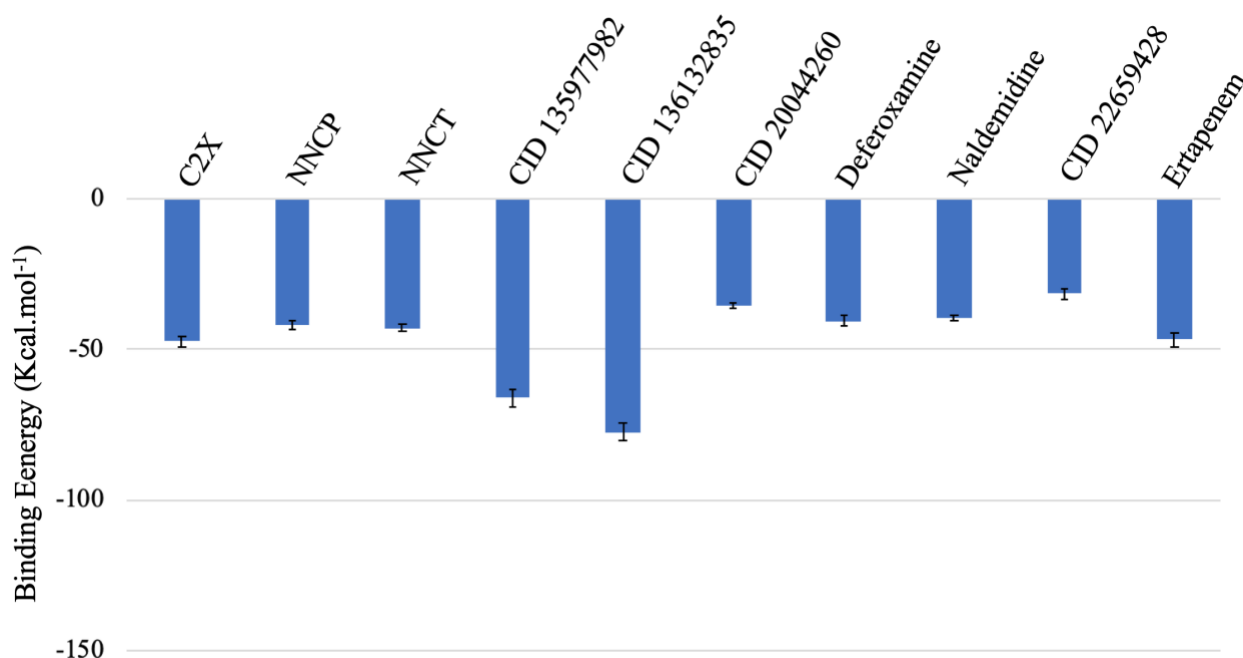
Figure 2.5. MM-PBSA results for RTA complexed with the reference compounds and the potential dual binders.

2.4.4   Comparison of results for the 6 potential dual binders pointed before

The flexible docking studies performed with MOE corroborated the rigid docking results obtained before using Molegro Virtual Docker (MVD) (Botelho *et al.*, 2020a; Botelho *et al.*, 2020c) with all repurposed ligands showing energetically favourable poses (negative values) that can H-bond to at least one residue of each binding site of RTA. Also, the values of docking energy with both methods followed the same tendency, with the only outlier being deferoxamine (see Table 2.3). This might be due to the higher flexibility of this compound that would favor finding a more energetically favorable pose during the flexible docking in comparison to the rigid docking. Despite this, in general, the results suggest that both docking protocols might be suitable for finding appropriate poses for the further MD simulations steps.

Differently from the docking studies not much corroboration was found when comparing the MD protocols applied to the 6 first potential dual binders. The results obtained for the extended MD simulations suggest that the 50 ns of free MD simulation ran before might not be enough to fully characterize the dynamical behavior of these compounds inside RTA. As shown in Table 2.2 none of the 6 compounds was capable of keeping H-bonds with residues of both sites of RTA during the whole simulated time. Deferoxamine was the one with best performance but lost consistency after 200 ns. CID 18498053, CID 18309602 and leucovorin stablished H-bonds mainly with the secondary pocket, while CID 136023163 and plazomicin lost the H-bonds with the pockets during the first third of the simulated time. Plazomicin was not even capable of stabilizing in another position inside RTA and left before 100 ns of simulation, as can be seem in the RMSD plots shown in Figure 2.3 and in the movie Plazomicin_RTA.mpg in the supplementary material.

Table 2.3. Rigid and Flexible docking results for the 6 potential dual binders pointed before (Botelho et al., 2020a; Botelho et al., 2020c).

Ligand	Interacting residues (interaction type)		Energy (kcal mol⁻¹)	
	Rigid docking (Botelho *et al.*, 2020a; Botelho *et al.*, 2020c)	Flexible docking	Rigid docking	Flexible docking
CID136023163	Arg56, Thr77, Asn78, Tyr80, Val81, Asp96, Asp100, Gly121, Arg180, Arg258	Asp75, Asn78, Gly121, Glu208, Gly212, Arg213, Arg258	-203.93	-9.40

Ligand	Interacting residues (interaction type)		Energy (kcal mol^{-1})	
CID18498053	Asn78, Val81, Asp96, Asp100, Asn122, Ser176, Glu177, Arg180, Glu208, Arg258	Asn78, Glu177, Glu208, Arg258	-161.20	-9.01
CID18309602	Asp75, Asp96, Asp100, Tyr123, Asn122, Asp124, Glu177, Arg180, Glu208, Asn209,	Asp96, Asp100, Glu177, Arg180	-152.14	-8.70
Plazomicin	Asn78, Asp96, Asp100, Glu177, Arg180, Asn209	Asn78, Asp96, Asp100, Tyr123, Glu177, Glu208	-120.79	-8.23
Leucovorin	Asp75, Asn78, Asp100, Ser176, Arg180	Asp96, Gly121, Asn122, Arg213, Arg258	-99.35	-7.53
Deferoxamine	Asn78, Asp96, Asp100, Gly121, Asn122, Phe168 Arg180	Asp96, Asp100, Glu177, Arg180, Arg213	-68.57	-9.16

2.4.5   The pterin derivatives

It's important to notice that the compounds CID 135977982, CID 136023163, and CID 136132835 (Figures 2.2 and 5.2) are all pterin derivatives, differing from each other only by the substitution pattern in the phenol ring. CID 136023163 brings two iodine atoms in the ortho positions to the OH group, while CID 136132835 brings one fluorine in one ortho position and an iodine in the other, and CID 135977982 have only one fluorine atom in ortho to the OH group. Despite this single difference the best pose of CID 136023163 was not capable of stabilizing inside RTA forming H-bonds with the two pockets (see Figure 5.15 and the movie CID136023163-RTA.mpg supplied as supplementary information), while the other two analogues displayed the best binding results of our study, showing a binding mode very similar to the reference compound C2X (see Figure 2.4 and the movies CID135977982-RTA.mpg and CID136132835_RTA.mpg supplied as supplementary information). This probably happened because the two large iodine atoms didn't allow the stabilization of the phenol ring of CID 136023163 in neither of the binding pockets of RTA. This resulted in its displacement from the binding site of RTA. Such result adds some valuable Structure Activity Relationship (SAR) information to this project that certainly will

help in the selection/design of the ideal ring size on this part of the molecule to achieve an effective and stable binding.

2.5   Conclusions

Our theoretical study allowed a refinement of the protocol used before (Botelho *et al.*, 2020a; Botelho *et al.*, 2020c) for the theoretical investigation of repurposed drugs selected by VS, as potential antidotes against ricin and also moving forward on the theoretical investigation of the library selected before (Botelho *et al.*, 2020a; Botelho *et al.*, 2020c). Despite no significant changes were observed using the new docking protocol, the extension of the time of MD simulation for 500 ns showed a fundamental step to corroborate the docking results and filter the ligands which are really capable of performing stable interactions in both pockets of RTA. This refinement enabled pointing with more confidence four compounds to further *in vitro* assays: CID 135977982, CID 136132835 and naldemidine, as potential dual binders to RTA, and nilotinib and CID 18493267, as single binders to the secondary pocket of RTA. Despite not interacting directly with residues of the catalytic site, these compounds seem to be capable of interacting with residues at its entrance, blocking the catalytic activity of RTA. This might trigger a stronger inhibition. We believe that these five compounds will present IC_{50} values in the nM range after experimentally evaluated. Our study also corroborated the complex RTA/C2X as a consistent model of the RTA binding to the loop GAGA of the 28S rRNA. This validates the PDB structure 3HIO reported by Ho and co-workers (Ho *et al.*, 2009) for the drug design against RTA. On this sense also the discovery of consistent H-bonds with residues Arg213 and Arg258 observed for most of the ligands in both the docking and MD simulations, opens new opportunities for the drug design of new antidotes against ricin. As far as we know this was not described in the literature before. Still on the context of drug design, the limitation imposed by the two iodine atoms in the phenol ring of the pterin derivatives adds useful information for the design of more effective pterine derivatives as antidotes against ricin.

2.6   Associated content

2.6.1   Supporting information

Binding modes of NNCP and C2X inside RTA (Figure 5.1); 2D structures of 6 compounds pointed before by VS for repurposing (Figure 5.2); Protocol used for MD simulations using GROMACS; Redocking of C2X (Figure 5.3); Best pose of CID 135977982 inside RTA (Figure 5.4); Energy plots after MD simulations (Figures 5.5 - 5.10); H-bonds formed by the dual binders

during the MD simulations (Figures 5.11 - 5.14); Superposition of frames collected during the MD simulations (Figure 5.15); Movies of the 18 MD simulations performed (https://pubs.acs.org/doi/10.1021/acsomega.2c04819).

2.6.2   Disclosure statement

The authors declare the following competing financial interest(s): Michael L. Drummond is employee of Chemical Computing Group ULC, creators of MOE software used to perform most of the modeling work described herein.

2.7   Author information

2.7.1   Corresponding authors

Steven R. LaPlante - INRS - Institut Armand-Frappier, 531 boulevard des Prairies, Laval, H7V 1B7, Quebec, Canada; ORCID: https://orcid.org/0000-0003-2835-5789; Email: Steven.LaPlante@inrs.ca

Tanos C. C. França - INRS - Institut Armand-Frappier, 531 boulevard des Prairies, Laval, H7V 1B7, Quebec, Canada; Laboratory of Molecular Modeling Applied to Chemical and Biological Defense, Military Institute of Engineering, Rio de Janeiro, 22290-270, Brazil; Department of Chemistry, Faculty of Science, University of Hradec Kralove, Rokitanskeho 62, 50003, Hradec Kralove, Czech Republic. ORCID: https://orcid.org/0000-0002-6048-8103; Email: tanosfranca@gmail.com

2.7.2   Authors

Fernanda D. Botelho - Laboratory of Molecular Modeling Applied to Chemical and Biological Defense, Military Institute of Engineering, Rio de Janeiro, 22290-270, Brazil; ORCID: https://orcid.org/0000-0001-6407-5800; Email: fernandinhabotelho@gmail.com

Michael L. Drummond - Chemical Computing Group, Montreal, Quebec H3A 2R7, Canada; ORCID: https://orcid.org/0000-0002-4688-3975; Email: mdrummond@chemcomp.com

2.8   Acknowledgments

3 GENERAL DISCUSSION, CONCLUSIONS, AND PERSPECTIVES

The main purpose of this work was both extending and refining the studies started by Botelho and co-workers (Botelho *et al.*, 2020a; Botelho *et al.*, 2020c). For that we added 9 more compounds to the list of 6 potential dual binders to RTA pointed before and submitted this new group of 15 compounds to a more robust modeling protocol, including flexible docking using MOE®, 10 times longer MD simulations using the NAMD software (Nelson *et al.*, 1996), and further MM-PBSA studies using GROMACS (Abraham *et al.*, 2015). The results obtained corroborated some findings pointed before for the 6 first potential dual binders, but also exposed some limitations of the first protocol. These limitations were properly addressed here, and a new protocol was established, which will allow a more accurate investigation of not only our initial library of 82 compounds, but also any other library of potential RTA binders.

Regarding the changes in the docking protocol we observed a good corroboration between the rigid docking used before (Botelho *et al.*, 2020a; Botelho *et al.*, 2020c) and the flexible docking adopted in this work. The same tendency in terms of the energies obtained for the best poses of the 6 compounds initially investigated (Botelho *et al.*, 2020a; Botelho *et al.*, 2020c) was observed, as well as similar interactions (see Table 3.1). Both protocols, therefore, showed suitable to obtain the initial poses needed for the further MD simulation study of ligands inside RTA. The extended MD simulation protocol, however, revealed that our former MD protocol (Botelho *et al.*, 2020a; Botelho *et al.*, 2020c) is too short to ensure the stability of the ligands inside RTA and/or their capacity of stablishing H-bonds in both the catalytic and the secondary sites. This can be seen in the RMSD plots in Figure 2.3, the overlap of frames in Figure 5.15, and in the plots of H-bonds formed during the MD simulations in Figure 3.1. These results reveal that changes were needed in the MD simulation protocol in order of achieving more accurate results, and suggest that the 6 potential dual binders pointed before (Botelho *et al.*, 2020a; Botelho *et al.*, 2020c) might fail when challenged experimentally as RTA inhibitors.
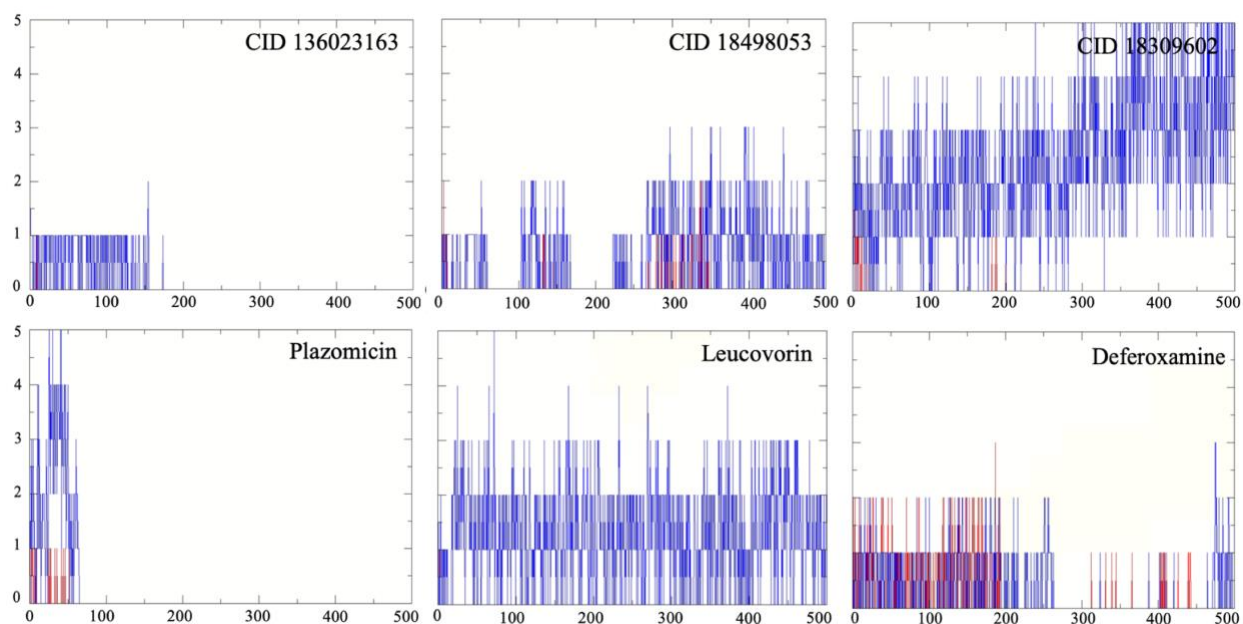
Figure 3.1. H-bonds formed during the MD simulations for the 6 potential dual binders pointed before (Botelho *et al.*, 2020a; Botelho *et al.*, 2020c). H-bonds with residues of the catalytic site are shown in red, while those with residues of the secondary site are shown in blue.

The docking and MD simulation studies of the 15 compounds investigated here allowed mapping the relevant interactions for the dual binding inside RTA and disclosed the potential important role of residues Arg213 and Arg258. This is in line with the specific goals 1) and 2) listed in section 1.5. Also, our new protocol associating flexible docking, longer MD simulations, and MM-PBSA calculations, enabled selecting and ranking the most promising binders to RTA among the set of compounds studied achieving this way the specific goals 3) and 4) listed in section 1.5.

Interestingly, our theoretical study pointed to the two pterine derivatives CID 135977982, and CID 136132835 (Figure 2.2) as the most promising RTA binders. This is aligned with results from the literature (Pruet *et al.*, 2012; Saito *et al.*, 2013; Wiget *et al.*, 2013) showing this family of compounds as the best RTA binders (see section 1.4.1). Our docking results showed that these compounds can adopt stable conformations capable of keeping the pterin part of the molecule interacting at the catalytic site, while the other side of the molecule interact at the secondary site. The dynamic behavior of those compounds accessed through the MD simulations (see movies CID135977982-RTA.mpg and CID136132835_RTA.mpg supplied as supplementary material), corroborated this finding by showing the stabilization of those conformations inside RTA with very low binding energy values, as shown in the plots of MM-PBSA calculations (see Figure 2.5). This encourage us to predict that these compounds will show lower IC_{50} than NNPT (the best RTA

inhibitor currently reported in literature) and reinforce our hypothesis that RTA dual binders will reduce the current values of IC_{50} observed in literature (Pruet *et al.*, 2012; Saito *et al.*, 2013; Wiget *et al.*, 2013). We believe that together with naldemidine and nilotinib, that also showed stable dynamical behavior and low binding energies inside RTA (see movies Naldemidine_RTA.mpg and Nilotinib_RTA.mpg supplied as supplementary material and Figure 2.5), those compounds will qualify as the first antidotes against intoxication with ricin after experimental evaluation. If this scenario comes true, they will quickly be available on the counter, since all are repurposed drugs.

Even if the experimental results are not good enough to qualify these four compounds as effective antidotes against ricin, our study still have a significative impact on the field once it allowed mapping new important interactions for the binding into the catalytic and secondary sites of RTA (see discussion at section 2.4.2), validated the complex RTA/C2X (Ho *et al.*, 2009) for the drug design against ricin, corroborated pterin derivatives as potential antidotes, and provided relevant structure-activity information for the binding to the secondary pocket (see discussion at section 2.4.5).

3.1   Conclusions

In summary the present work enabled disclosing a more accurate protocol for the search of more efficient binders to RTA, besides contributing to reinforce our hypothesis that dual binders can be more effective inhibitors of RTA. We also revealed potential new relevant interactions for the stabilization of ligands inside RTA and validated the C2X-RTA complex as suitable for the design of RTA binders, besides reinforcing the pterin derivatives as promising RTA inhibitors. Finally, four repurposed compounds were pointed as potential antidotes against intoxication with ricin, which worth being experimentally evaluated.

3.2   Perspectives

As commented above the main perspective of our work is moving to the experimental evaluation of the compounds pointed by our theoretical study as the most promising binders to RTA. This should follow the standard steps of the drug discovery process (Berdigaliyev & Aljofan, 2020) summarized in Figure 3.2, where (in our case) many steps can be by passed or accelerated. Once our target is already known and the compounds under study are repurposed drugs, the two first steps (target selection and lead discovery) can be considered as already overcame. The work performed here correspond to the *in-silico* screening task of step three (medicinal chemistry) and the following task, chemical synthesis, is not necessary in our case, once the compounds are

commercially available. Therefore, we can tell that step three is also overcame and the investigation should now move towards the *in vitro* studies. On this sense the goal will be first verify if in fact the compounds pointed by the *in-silico* screening are capable of binding to RTA with an IC_{50} value low enough to be qualified as drug candidates and if they are selective to RTA compared to other targets (drug affinity and selectivity). In our case it's not necessary to create a cell disease model because there are no genetic issues involved in the intoxication with ricin. Also, studies of mechanism of action (MOA) are not need because it's already known where and how the antidote candidates should bind into RTA (Pruet *et al.*, 2012; Saito *et al.*, 2013; Wiget *et al.*, 2013). Regarding the task of lead candidate refinement, despite being FDA-approved drugs, some structural modifications on the compounds might be needed in order to improve the potency, oral availability, selectivity, pharmacokinetic properties and safety. This is because they are being evaluated for a different purpose then their original uses.

The *in-vivo* studies needed to evaluate the antidotes should involve animals previously intoxicated with ricin and treated with the potential antidotes. Normally no animal models of disease states, behavioral studies or *ex-vivo* studies are necessary in this case. Finally, after being considered promising in the *in-vitro* and *in-vivo* studies, the antidote candidates should move towards the clinical trials. Here again as we are dealing with FDA-approved drugs, many steps can be by passed or accelerated, like the investigation of pharmacodynamics and pharmacokinetics, human toxicity, and tolerated dosage range. Tasks typical of phases 0 and I. Accordingly, our ricin antidote candidates might likely jump directly to the clinical evaluations of phases II and III where they will be tested with humans before being approved to follow to the market (phase IV).

Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
Target selection	Lead discovery	Medicinal Chemistry	*In-vitro* studies	*In-vivo* studies	Clinical trials and therapeutics
Cellular and genetic targets	Synthesis	Library development	Drug affinity and selectivity	Animal models	Phase 0
Genomics	Combinatorial chemistry	SAR studies	Cell disease models	Behavioural studies	Phase I
Proteomics	Assay development	*In-silico* screening	MOA	Functional imaging	Phase II
Bioinformatics	High-throughput screening	Chemical synthesis	Lead candidate refinement	*Ex-vivo* studies	Phase III
					Phase IV

Figure 3.2. Usual steps in the drug discovery process. Figure created with Power Point®.

4 BIBLIOGRAPHY

Abagyan R, Totrov M,Kuznetsov D (1994) ICM—A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *Journal of computational chemistry* 15(5):488-506.

Abraham MJ, Murtola T,Schulz R (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. (SoftwareX), p 19-25.

Adelusi TI, Oyedele A-QK, Boyenle ID, Ogunlana AT, Adeyemi RO, Ukachi CD, Idris MO, Olaoba OT, Adedotun IO,Kolawole OE (2022) Molecular modeling in drug discovery. *Informatics in Medicine Unlocked* :100880.

Audi J, Belson M, Patel M, Schier J,Osterloch J (2005) Ricin Poisoning - A comprehensive review. *JAMA* 294(18).

Augerson W (2000) *A Review of the Scientific Literature as it Pertains to Gulf War Illnesses: Volume 5: Chemical and Biological Warfare Agents.* RAND Corporation, Santa Monica, CA. https://www.rand.org/pubs/monograph_reports/MR1018z5.html

Bai Y, Monzingo AF,Robertus JD (2009) The X-ray structure of ricin A chain with a novel inhibitor. *Archives of biochemistry and biophysics* 483(1):23-28.

Bai Y, Watt B, Wahome PG, Mantis NJ,Robertus JD (2010) Identification of new classes of ricin toxin inhibitors by virtual screening. *Toxicon* 56(4):526-534.

Bellisola G, Fracasso G, Ippoliti R, Menestrina G, Rosén A, Soldà S, Udali S, Tomazzolli R, Tridente G,Colombatti M (2004) Reductive activation of ricin and ricin A-chain immunotoxins by protein disulfide isomerase and thioredoxin reductase. *Biochemical Pharmacology* 67(9):1721-1731.

Berdigaliyev N & Aljofan M (2020) An overview of drug discovery and development. *Future Medicinal Chemistry* 12(10):939-947.

Berendsen HJC, Postma JPM, Gunsteren WFv, DiNola A,Haak JR (1984) Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* 81(8):3684-3690.

Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S, Fagan P, Marvin J, Padilla D, Ravichandran V, Schneider B, Thanki N, Weissig H, Westbrook JD,Zardecki C (2002) The protein data bank. *Acta Crystallographica Section D: Biological Crystallography* 10.1107/S0907444902003451.

Botelho FD, dos Santos MC, Goncalves AD, Kuca K, Valis M, LaPlante SR, Franca TCC,de Almeida JSFD (2020a) Ligand-Based Virtual Screening, Molecular Docking, Molecular Dynamics, and MM-PBSA Calculations towards the Identification of Potential Novel Ricin Inhibitors. *Toxins* 12(12).

Botelho FD, Dos Santos MC, Gonçalves AdS, Kuca K, Valis M, LaPlante SR, França TCC,de Almeida JSFD (2020b) Ligand-Based Virtual Screening, Molecular Docking, Molecular Dynamics, and MM-PBSA Calculations towards the Identification of Potential Novel Ricin Inhibitors. *Toxins* 12(12).

Botelho FD, Santos MC, Gonçalves AS, França TC, LaPlante SR,de Almeida JS (2020c) Identification of novel potential ricin inhibitors by virtual screening, molecular docking,

molecular dynamics and MM-PBSA calculations: a drug repurposing approach. *Journal of Biomolecular Structure and Dynamics* :1-11.

Botelho FD, Santos MC, Goncalves AS, Franca TCC, LaPlante SR,de Almeida JSFD (2020d) Identification of novel potential ricin inhibitors by virtual screening, molecular docking, molecular dynamics and MM-PBSA calculations: a drug repurposing approach. *Journal of Biomolecular Structure & Dynamics* 10.1080/07391102.2020.1870154.

Braun E, Gilmer J, Mayes H, Mobley D, Monroe J, Prasad S,Zuckerman D (2018) Best Practices for Foundations in Molecular Simulations [Article v1. 0]. Living J Comp Mol Sci. 2018; 1 (1): 5957.).

Bussi G, Donadio D,Parrinello M (2007) Canonical sampling through velocity rescaling. *Journal of Chemical Physics* 126(1).

Carra JH, Wannemacher RW, Tammariello RF, Lindsey CY, Dinterman RE, Schokman RD,Smith LA (2007) Improved formulation of a recombinant ricin A-chain vaccine increases its stability and effective antigenicity. *Vaccine* 25(21):4149-4158.

Case DA, Darden T, Cheatham T, Simmerling CL, Wang J, Duke RE, Luo R, Crowley M, Walker RC,Zhang W (2008) *Amber 10.* University of California, p

da Cunha EFF, Ramalho TC,Reynolds RC (2008) Binding Mode Analysis of 2,4-diamino-5-methyl-5-deaza-6-substituted Pteridines with Mycobacterium tuberculosis and Human Dihydrofolate Reductases. *Journal of Biomolecular Structure and Dynamics* 25(4):377-385.

Da Silva AWS & Vranken WF (2012) ACPYPE-Antechamber python parser interface. *BMC research notes* 5(1):1-8.

Deeks ED, Cook JP, Day PJ, Smith DC, Roberts LM,Lord JM (2002) The Low Lysine Content of Ricin A Chain Reduces the Risk of Proteolytic Degradation after Translocation from the Endoplasmic Reticulum to the Cytosol. *Biochemistry* 41(10):3405-3413.

Doan LG (2004) Ricin: Mechanism of toxicity, clinical manifestations, and vaccine development. A review. *Journal of Toxicology - Clinical Toxicology* 42(2):201-208.

Dodda LS, Cabeza de Vaca I, Tirado-Rives J,Jorgensen WL (2017) LigParGen web server: an automatic OPLS-AA parameter generator for organic ligands. *Nucleic Acids Res* 45(W1):W331-w336.

Durrant JD & McCammon JA (2011) Molecular dynamics simulations and drug discovery. *BMC Biology* 10.1186/1741-7007-9-71.

Eitzen Jr EM & Takafuji ET (1997) Historical overview of biological warfare. *Medical aspects of chemical and biologial warfare*. p 415-423.

Endo Y & Tsurugi K (1988) The RNA N-glycosidase activity of ricin A-chain. The characteristics of the enzymatic activity of ricin A-chain with ribosomes and with rRNA. *The Journal of biological chemistry* 263(18):8735-8739.

ENDO Y, TSURUGI K, YUTSUDO T, TAKEDA Y, OGASAWARA T,IGARASHI K (1988) Site of action of a Vero toxin (VT2) from Escherichia coli O157:H7 and of Shiga toxin on eukaryotic ribosomes. *European Journal of Biochemistry* 171(1-2):45-50.

Fan J, Fu A,Zhang L (2019) Progress in molecular docking. *Quantitative Biology* 7(2):83-89.

Farahani MD, França TCC, Alapour S, Shahout F, Boulon R, Iddir M, Maddalena M, Ayotte Y,LaPlante SR (2022) Jumping from Fragment to Drug via Smart Scaffolds. *ChemMedChem* 17(10):e202200092.

Franca TCC (2015) Homology modeling: an important tool for the drug discovery. *Journal of Biomolecular Structure & Dynamics* 33(8):1780-1793.

Franca TCC, Pascutti PG, Ramalho TC,Figueroa-Villar JD (2005) A three-dimensional structure of Plasmodium falciparum serine hydroxymethyltransferase in complex with glycine and 5-formyl-tetrahydrofolate. Homology modeling and molecular dynamics. *Biophysical Chemistry* 115(1):1-10.

Franca TCC, Wilter A, Ramalho TC, Pascutti PG,Figueroa-Villar JD (2006) Molecular dynamics of the interaction of Plasmodium falciparum and human Serine Hydroxymethyltransferase with 5-formyl-6-hydrofolic acid analogues: Design of new potential antimalarials. *Journal of the Brazilian Chemical Society* 17(7):1383-1392.

Franke H, Scholl R,Aigner A (2019) Ricin and Ricinus communis in pharmacology and toxicology-from ancient use and "Papyrus Ebers" to modern perspectives and "poisonous plant of the year 2018". *Naunyn-Schmiedeberg's archives of pharmacology* 392(10):1181-1208.

Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M,Perry JK (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of Medicinal Chemistry* 47(7):1739-1749.

Funatsu G, Islam MR, Minami Y, Sung-Sil K,Kimura M (1991) Conserved amino acid residues in ribosome-inactivating proteins from plants. *Biochimie* 73(7):1157-1161.

Gal Y, Alcalay R, Sabo T, Noy-Porat T, Epstein E, Kronman C,Mazor O (2015) Rapid assessment of antibody-induced ricin neutralization by employing a novel functional cell-based assay. *Journal of Immunological Methods* 424:136-139.

Gonçalves MA, Santos LS, Prata DM, Peixoto FC, da Cunha EF,Ramalho TC (2017) Optimal wavelet signal compression as an efficient alternative to investigate molecular dynamics simulations: application to thermal and solvent effects of MRI probes. *Theoretical Chemistry Accounts* 136(1):1-13.

Guimaraes AP, Oliveira AA, da Cunha EFF, Ramalho TC,Franca TCC (2011) Design of New Chemotherapeutics Against the Deadly Anthrax Disease. Docking and Molecular Dynamics Studies of Inhibitors Containing Pyrrolidine and Riboamidrazone Rings on Nucleoside Hydrolase from Bacillus anthracis. *Journal of Biomolecular Structure & Dynamics* 28(4):455-469.

Halgren TA (1996) Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of computational chemistry* 17(5-6):490-519.

Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT,Banks JL (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *Journal of Medicinal Chemistry* 47(7):1750-1759.

Hansson T, Oostenbrink C,van Gunsteren W (2002) Molecular dynamics simulations. *Current opinion in structural biology* 12(2):190-196.

Ho M-c, Sturm MB, Almo SC,Schramm VL (2009) Transition state analogues in structures of ricin and saporin ribosome-inactivating proteins. *PNAS* 106(48):20276-20281.

Homeyer N & Gohlke H (2012) Free energy calculations by the Molecular Mechanics Poisson-Boltzmann Surface Area method. *Molecular Informatics* 31(2).

Hu W-g, Yin J, Chau D, Hu CC, Lillico D, Yu J, Negrych LM,Cherwonogrodzky JW (2013) Conformation-Dependent High-Affinity Potent Ricin-Neutralizing Monoclonal Antibodies. *BioMed Research International* 2013.

Huang S-Y & Zou X (2010) Advances and challenges in protein-ligand docking. *International Journal of Molecular Sciences* 11(8):3016-3034.

Jakalian A, Bush BL, Jack DB,Bayly CI (2000) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. *Journal of computational chemistry* 21(2):132-146.

Janik E, Ceremuga M, Saluk-Bijak J,Bijak M (2019) Biological toxins as the potential tools for bioterrorism. *International Journal of Molecular Sciences* 20(5):1181.

Jansen HJ, Breeveld FJ, Stijnis C,Grobusch MP (2014) Biological warfare, bioterrorism, and biocrime. *Clinical Microbiology and Infection* 20(6):488-496.

Jasheway K, Pruet J, Anslyn EV,Robertus JD (2011) Structure-based design of ricin inhibitors. *Toxins* 3(10):1233-1248.

Jones G, Willett P, Glen RC, Leach AR,Taylor R (1997) Development and validation of a genetic algorithm for flexible docking. *Journal of Molecular Biology* 267(3):727-748.

Jorgensen WL, Maxwell DS,Tirado-Rives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society* 118(45).

Kaminski GA, Friesner RA, Tirado-Rives J,Jorgensen WL (2001) Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *The Journal of Physical Chemistry B* 105(28):6474-6487.

Karplus M & McCammon JA (2002) Molecular dynamics simulations of biomolecules. *Nature structural biology* 9(9):646-652.

Knight B (1979) Ricin - a potent homicidal poison. *British Medical Journal* 1(February):350-351.

Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA,Cheatham TE (2000) Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Accounts of Chemical Research* 33(12).

Kontoyianni M, McClellan LM,Sokol GS (2004) Evaluation of Docking Performance: Comparative Data on Docking Algorithms. *Journal of Medicinal Chemistry* 10.1021/jm0302997.

Korb O, Stutzle T,Exner TE (2009) Empirical scoring functions for advanced protein− ligand docking with PLANTS. *Journal of chemical information and modeling* 49(1):84-96.

Kumari R, Kumar R, Consortium OSDD,Lynn A (2014) g _ mmpbsa - A GROMACS tool for MM-PBSA and its optimization for high-throughput binding energy calculations. *J. Chem. Inf. Model.* 54.

Leach AR (2001) *Molecular Modelling: Principles and Applications.* Prentice Hall, Harlow, 2

Legler PM, Brey RN, Smallshaw JE, Vitetta ES,Millard CB (2011) Structure of RiVax: a recombinant ricin vaccine. *Acta Crystallographica Section D: Biological Crystallography* 67(9):826-830.

Li X-P, Harijan RK, Cao B, Kahn JN, Pierce M, Tsymbal AM, Roberge JY, Augeri D,Tumer NE (2021) Synthesis and structural characterization of ricin inhibitors targeting ribosome binding using fragment-based methods and structure-based design. *Journal of Medicinal Chemistry* 64(20):15334-15348.

Li X-P, Harijan RK, Kahn JN, Schramm VL,Tumer NE (2020) Small molecule inhibitors targeting the interaction of ricin toxin A subunit with ribosomes. *ACS infectious diseases* 6(7):1894-1905.

Lipparini F & Mennucci B (2021) Hybrid QM/classical models: Methodological advances and new applications. *Chemical Physics Reviews* 2(4):041303.

Liu H, Elstner M, Kaxiras E, Frauenheim T, Hermans J,Yang W (2001) Quantum mechanics simulation of protein dynamics on long timescale. *Proteins: Structure, Function, and Bioinformatics* 44(4):484-489.

Lord JM, Roberts LM,Robertus JD (1994) Ricin: structure, mode of action, and some current applications. *The FASEB journal* 8(2):201-208.

May MJ, Hartley MR, Roberts LM, Krieg PA, Osborn RW,Lord JM (1989) Ribosome inactivation by ricin A chain: a sensitive method to assess the activity of wild-type and mutant polypeptides. *The EMBO Journal* 8(1):301-308.

McLain DE, Horn TL, Detrisac CJ, Lindsey CY,Smith LA (2011) Progress in biological threat agent vaccine development: a repeat-dose toxicity study of a recombinant ricin toxin A-chain (rRTA) 1-33/44-198 vaccine (RV Ec) in male and female New Zealand white rabbits. *International journal of toxicology* 30(2):143-152.

McLain DE, Lewis BS, Chapman JL, Wannemacher RW, Lindsey CY,Smith LA (2012) Protective effect of two recombinant ricin subunit vaccines in the New Zealand white rabbit subjected to a lethal aerosolized ricin challenge: survival, immunological response, and histopathological findings. *Toxicological Sciences* 126(1):72-83.

Mishra V & Prasad CS (2011) Ligand based virtual screening to find novel inhibitors against plant toxin Ricin by using the ZINC database. *Bioinformation* 7(2):46.

Monzingo AF & Robertus JD (1992) X-ray analysis of substrate analogs in the ricin A-chain active site. *Journal of Molecular Biology* 227(4):1136-1145.

Moustakas DT, Lang PT, Pegg S, Pettersen E, Kuntz ID, Brooijmans N,Rizzo RC (2006) Development and validation of a modular, extensible docking program: DOCK 5. *Journal of computer-aided molecular design* 20(10):601-619.

Musshoff F & Madea B (2009) Ricin poisoning and forensic toxicology. *Drug Testing and Analysis* 1(4):184-191.

Nelson MT, Humphrey W, Gursoy A, Dalke A, Kalé LV, Skeel RD,Schulten K (1996) NAMD: a parallel, object-oriented molecular dynamics program. *The International Journal of Supercomputer Applications and High Performance Computing* 10(4):251-268.

Olsnes S (2004) The history of ricin, abrin and related toxins. *Toxicon* 44(4):361-370.

Olsnes S, Fernandez-Puentes C, Carrasco L,Vazquez D (1975) Ribosome Inactivation by the Toxic Lectins Abrin and Ricin. *European Journal of Biochemistry* 60:281-288.

Olson MA, Carra JH, Roxas-Duncan V, Wannemacher RW, Smith LA,Millard CB (2004) Finding a new vaccine in the ricin protein fold. *Protein Engineering, Design and Selection* 17(4):391-397.

Österberg F, Morris GM, Sanner MF, Olson AJ,Goodsell DS (2002) Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in AutoDock. *Proteins: Structure, Function, and Bioinformatics* 46(1):34-40.

Parrinello M & Rahman A (2005) Polymorphic transitions in single crystals: A new molecular dynamics method Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys* 7182(1981).

Patel VR, Dumancas GG, Viswanath LCK, Maples R,Subong BJJ (2016) Castor oil: properties, uses, and optimization of processing parameters in commercial production. *Lipid insights* 9:LPI. S40233.

Paul D, Sanap G, Shenoy S, Kalyane D, Kalia K,Tekade RK (2021) Artificial intelligence in drug discovery and development. *Drug discovery today* 26(1):80.

Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L,Schulten K (2005) Scalable molecular dynamics with NAMD. *Journal of computational chemistry* 26(16):1781-1802.

Phillips JC, Hardy DJ, Maia JD, Stone JE, Ribeiro JV, Bernardi RC, Buch R, Fiorin G, Hénin J,Jiang W (2020) Scalable molecular dynamics on CPU and GPU architectures with NAMD. *The Journal of Chemical Physics* 153(4):044130.

Pita R & Romero A (2014) Toxins as Weapons: A Historical Review. *Forensic Science Review* 26(2):85-86.

Plimpton S (1995) Fast parallel algorithms for short-range molecular dynamics. *Journal of computational physics* 117(1):1-19.

Porter A, Phillips G, Smith L, Erwin-Cohen R, Tammariello R, Hale M,DaSilva L (2011) Evaluation of a ricin vaccine candidate (RVEc) for human toxicity using an in vitro vascular leak assay. *Toxicon* 58(1):68-75.

Pruet JM, Jasheway KR, Manzano LA, Bai Y, Anslyn EV,Robertus JD (2011) 7-Substituted pterins provide a new direction for ricin A chain inhibitors. *European Journal of Medicinal Chemistry* 46(9):3608-3615.

Pruet JM, Saito R, Manzano LA, Jasheway KR, Wiget PA, Kamat I, Anslyn EV,Robertus JD (2012) Optimized 5-membered heterocycle-linked pterins for the inhibition of Ricin Toxin A. *ACS Medicinal Chemistry Letters* 3(7):588-591.

Rarey M, Kramer B, Lengauer T,Klebe G (1996) A fast flexible docking method using an incremental construction algorithm. *Journal of Molecular Biology* 261(3):470-489.

Reif MM, Hünenberger PH,Oostenbrink C (2012) New interaction parameters for charged amino acid side chains in the GROMOS force field. *Journal of Chemical Theory and Computation* 8(10).

Ribeiro AAST, Horta BAC,De Alencastro RB (2008) MKTOP: A program for automatic construction of molecular topologies. *Journal of the Brazilian Chemical Society* 19(7).

Rosenfeld R, Vajda S,DeLisi C (1995) Flexible docking and design. *Annual Review of Biophysics and Biomolecular Structure* 24(1):677-700.

Roy CJ, Brey RN, Mantis NJ, Mapes K, Pop IV, Pop LM, Ruback S, Killeen SZ, Doyle-Meyers L, Vinet-Oliphant HS, Didier PJ,Vitetta ES (2015) Thermostable ricin vaccine protects rhesus macaques against aerosolized ricin: Epitope-specific neutralizing antibodies correlate with protection. *Proceedings of the National Academy of Sciences* 112(12):201502585.

Saito R, Pruet JM, Manzano LA, Jasheway K, Monzingo AF, Wiget PA, Kamat I, Anslyn EV,Robertus JD (2013) Peptide-conjugated pterins as inhibitors of ricin toxin A. *Journal of Medicinal Chemistry* 56(1):320-329.

Santos MC, Botelho FD, Gonçalves AS, Kuca K, Nepovimova E, Cavalcante SFA, Lima ALS,França TCC (2022) Theoretical assessment of the performances of commercial oximes on the reactivation of acetylcholinesterase inhibited by the nerve agent A-242 (novichok). *Food and Chemical Toxicology* 165:113084.

Sauton N, Lagorce D, Villoutreix BO,Miteva MA (2008) MS-DOCK: accurate multiple conformation generator and rigid docking protocol for multi-step virtual ligand screening. *BMC bioinformatics* 9(1):1-12.

Smallshaw JE, Firan A, Fulmer JR, Ruback SL, Ghetie V,Vitetta ES (2002) A novel recombinant vaccine which protects mice against ricin intoxication. *Vaccine* 20(27-28):3422-3427.

Smallshaw JE, Ghetie V, Rizo J, Fulmer JR, Trahan LL, Ghetie M-A,Vitetta ES (2003) Genetic engineering of an immunotoxin to eliminate pulmonary vascular leak in mice. *Nature biotechnology* 21(4):387-391.

Smallshaw JE & Vitetta ES (2011) Ricin vaccine development. *Ricin and Shiga Toxins* :259-272.

Sousa RB, Lima KSC, Santos CGM, Franca TCC, Nepovimova E, Kuca K, Dornelas MR,Lima ALS (2019) A New Method for Extraction and Analysis of Ricin Samples through MALDI-TOF-MS/MS. *Toxins* 11(4).

Sphyris N, Lord JM, Wales R,Roberts LM (1995) Mutational Analysis of the Ricinus Lectin B-chains: GALACTOSE-BINDING ABILITY OF THE 2γ SUBDOMAIN OF RICINUS COMMUNIS AGGLUTININ B-CHAIN (∗). *Journal of Biological Chemistry* 270(35):20292-20297.

Spooner RA, Smith DC, Easton AJ, Roberts LM,Lord MJ (2006) Retrograde transport pathways utilised by viruses and protein toxins. *Virology Journal* 3(1):1-10.

SPOONER Robert A, WATSON Peter D, MARSDEN Catherine J, SMITH Daniel C, MOORE Katherine AH, COOK Jonathon P, LORD JM,ROBERTS Lynne M (2004) Protein disulphide-isomerase reduces ricin to its A and B chains in the endoplasmic reticulum. *Biochemical Journal* 383(2):285-293.

Sun H (2008) Pharmacophore-based virtual screening. *Current Medicinal Chemistry* 15(10):1018-1024.

Thomsen R & Christensen MH (2006) MolDock: a new technique for high-accuracy molecular docking. *Journal of Medicinal Chemistry* 49(11):3315-3321.

Toledo Warshaviak D, Golan G, Borrelli KW, Zhu K,Kalid O (2014) Structure-Based Virtual Screening Approach for Discovery of Covalently Bound Ligands. *Journal of chemical information and modeling* 54(7):1941-1950.

Topliss J (2012) Quantitative structure-activity relationships of drugs.

Vanommeslaeghe K, Raman EP,MacKerell AD (2012) Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges. *Journal of chemical information and modeling* 52(12).

Venkatachalam CM, Jiang X, Oldfield T,Waldman M (2003) LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *Journal of Molecular Graphics and Modelling* 21(4):289-307.

Vilar S, Cozza G,Moro S (2008) Medicinal chemistry and the molecular operating environment (MOE): application of QSAR and molecular docking to drug discovery. *Current topics in medicinal chemistry* 8(18):1555-1572.

Walters WP, Stahl MT,Murcko MA (1998) Virtual screening—an overview. *Drug discovery today* 3(4):160-178.

Wang J, Wolf RM, Caldwell JW, Kollman PA,Case DA (2004) Development and testing of a general Amber force field. *Journal of computational chemistry* 25(9).

Wiget PA, Manzano LA, Pruet JM, Gao G, Saito R, Monzingo AF, Jasheway KR, Robertus JD,Anslyn EV (2013) Sulfur incorporation generally improves Ricin inhibition in pterin-appended glycine-phenylalanine dipeptide mimics. *Bioorganic and Medicinal Chemistry Letters* 23(24):6799-6804.

Willett P, Barnard JM,Downs GM (1998) Chemical Similarity Searching. *Journal of Chemical Information and Computer Sciences* 38(6):983-996.

Worbs S, Köhler K, Pauly D, Avondet M-A, Schaer M, Dorner MB,Dorner BG (2011) Ricinus communis intoxications in human and veterinary medicine—a summary of real cases. *Toxins* 3(10):1332-1372.

Yan X, Hollis T, Svinth M, Day P, Monzingo AF, Milne GW,Robertus JD (1997) Structure-based identification of a ricin inhibitor. *Journal of Molecular Biology* 266(5):1043-1049.

Zhou K, Fu Z, Chen M, Lin Y,Pan K (1994) Structure of trichosanthin at 1.88 Å resolution. *Proteins: Structure, Function, and Bioinformatics* 19(1):4-13.

5   ANNEX I

Theoretical investigation of repurposed drugs potentially capable of binding to the catalytic site and the secondary binding pocket of the subunit A of ricin

Tanos C. C. França[a,b,c,*], Fernanda D. Botelho[b], Michael L. Drummond[d],
Steven R. LaPlante[a,*]

[a]*INRS - Institut Armand-Frappier, 531 boulevard des Prairies, Laval, H7V 1B7, Quebec, Canada.*
[b]*Laboratory of Molecular Modeling Applied to Chemical and Biological Defense, Military Institute of Engineering, Rio de Janeiro, 22290-270, Brazil.*
[c]*Department of Chemistry, Faculty of Science, University of Hradec Kralove, Rokitanskeho 62, 50003, Hradec Kralove, Czech Republic.*
[d]*Chemical Computing Group, Montreal, Quebec H3A 2R7, Canada.*

*Corresponding authors: tanos@ime.eb.br ; Steven.LaPlante@inrs.ca

Supplementary Material
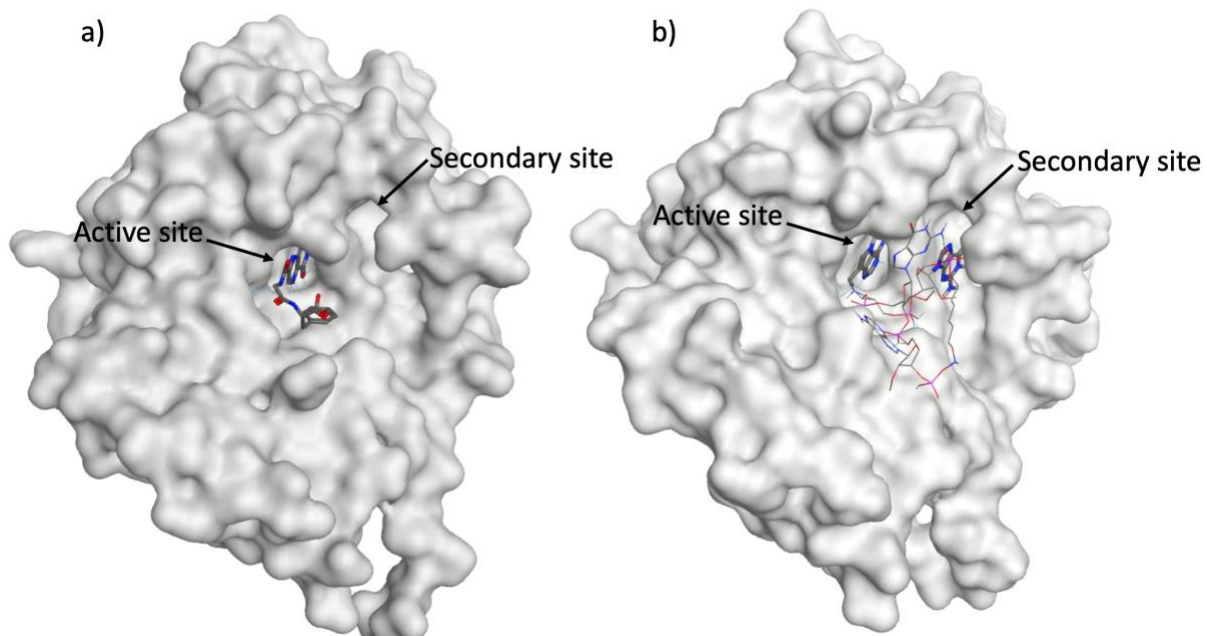
Figure 5.1. Binding modes inside RTA of: a) NNCP (PDB crystal 4HUO) and b) the cyclic tetranucleotide inhibitor reported by Ho et al. (2009) (PDB ID: 3HIO). For clarity most of the tetranucleotide structure is shown in wire representation.
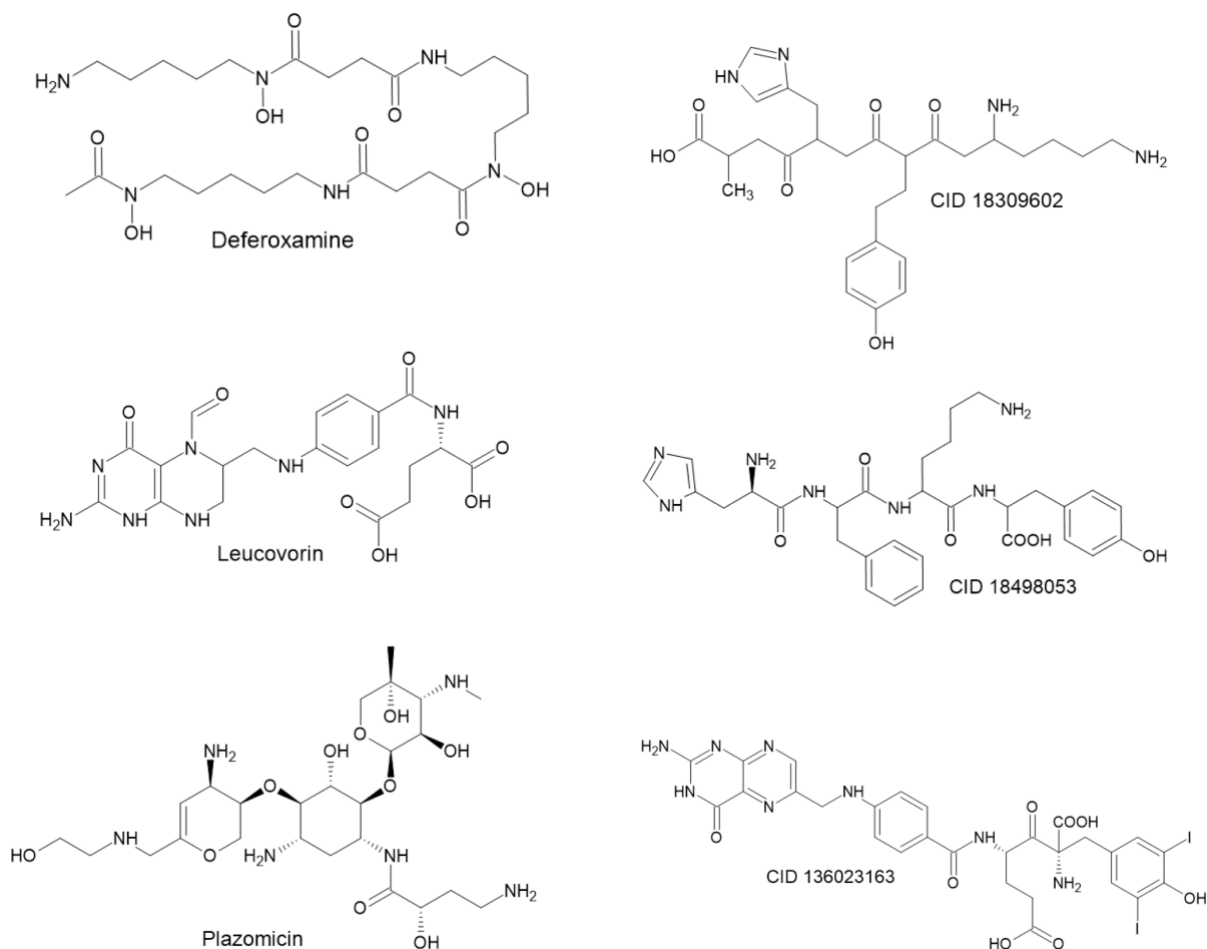
Figure 5.2. Compounds first pointed for repurposing against RTA (Botelho et al., 2020a; Botelho et al., 2020c).

Protocol used for the MD simulations using GROMACS

The ligands were first parameterized in order to be recognizable by the force field OPLS/AA(Jorgensen *et al.*, 1996) used for the MD simulations with the GROMACS 2019.4 program(Abraham *et al.*, 2015). The software ACPYPE(Da Silva & Vranken, 2012) and MKTOP(Ribeiro *et al.*, 2008) (Ribeiro et al., 2008) and the LigParGen online server (http://zarbi.chem.yale.edu/ligpargen)(Dodda *et al.*, 2017) were used to create the topology and coordinate files of the ligands, while the pdb2gmx program of GROMACS 2019.4(Abraham *et al.*, 2015) was used for the protein. After atomic charges calculation using the Restrained Electrostatic Potential (RESP) method(Jakalian *et al.*, 2000) the topology files of ligand and protein were merged to create the protein/ligand complexes needed for the MD simulations, which were carried

out using GROMACS 2019.4(Abraham *et al.*, 2015). For this the complexes were centered in cubic boxes of volume 850 nm^3 under periodic boundary conditions (PBC), filled with around 25,000 Tip4P water molecules(Jorgensen *et al.*, 1996) and neutralized by the addition of counterions. The energy minimization steps involved 100 ps under the steepest descent algorithm with position restraint (PR), followed by plus 100 ps with no restriction with the maximum force set to 100.0 kJ mol^{-1} nm^{-1}. This was followed by two equilibration steps of 100 ps under NPT and NVT ensembles (Bussi *et al.*, 2007; Parrinello & Rahman, 2005), respectively. The production steps consisted of 100 ns of free MD simulation at 310 K and 1 bar using 2 fs as integration time, a cutoff of 1.2nm for short-range (Lennard-Jones and Coulomb) interactions, and the leap-frog integrator algorithm; the coordinates of the complexes were stored every 10ps. In order to reproduce the physiological conditions, the Glu and Asp residues were assigned with negative charges, and the Lys and Arg residues were assigned with positive charges.
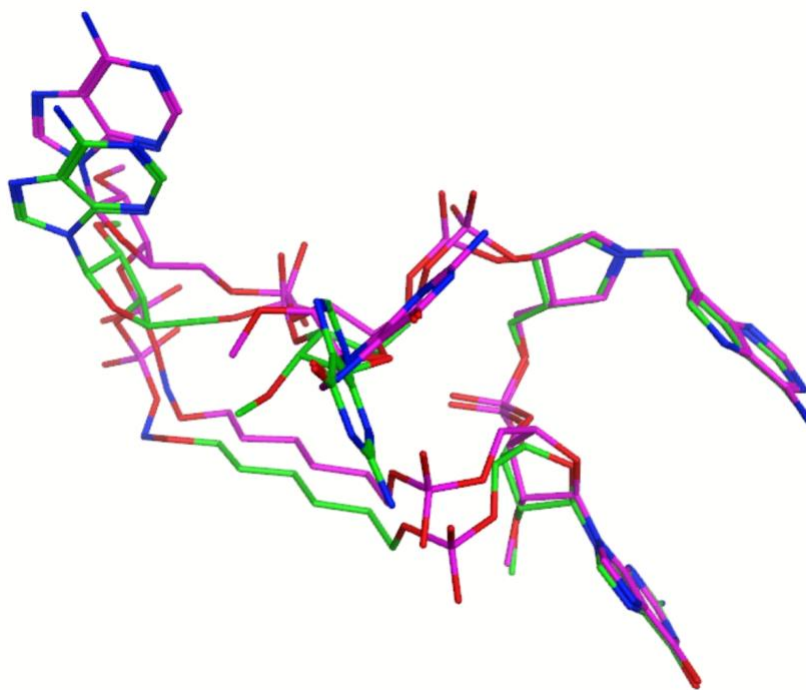
Figure 5.3. Redocking of C2X. Carbons of the crystallographic structure are shown in pink while carbons of the best pose are shown in green.

57
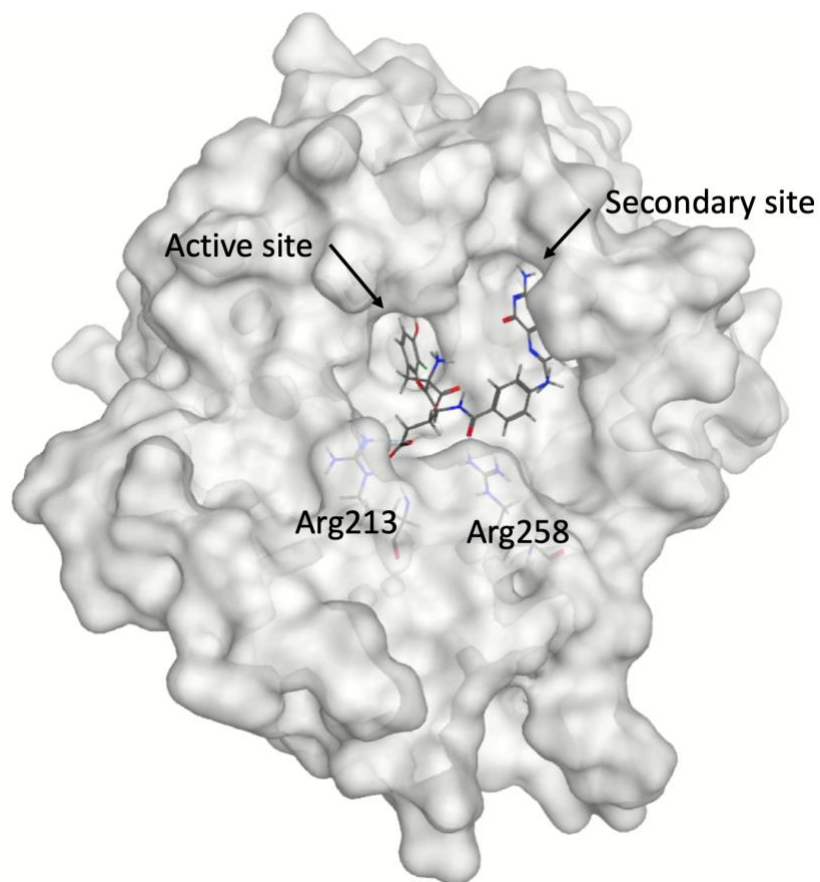
Figure 5.4. Best pose of CID 135977982 inside RTA, illustrating the positions of Arg213 and Arg258 below the RTA surface (in light gray).
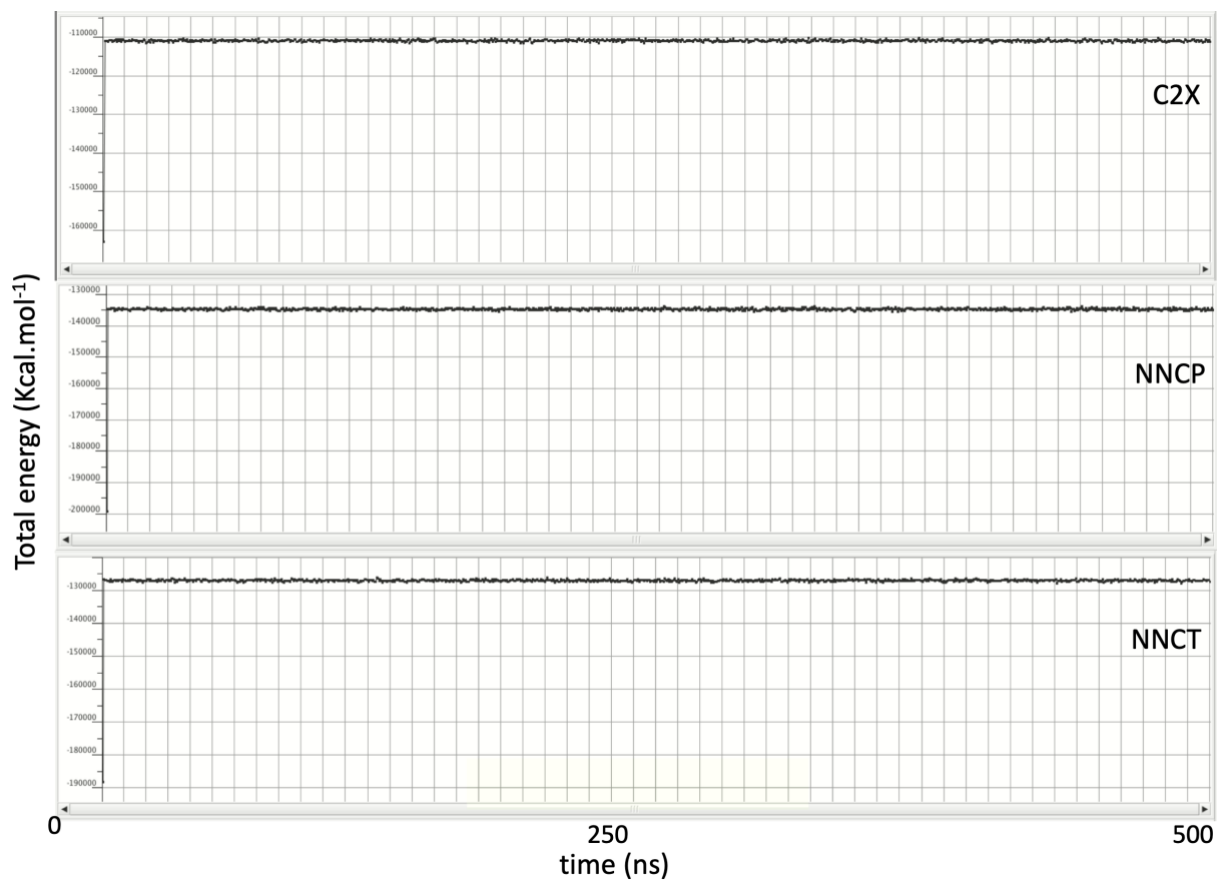
Figure 5.5. Plots of total energy during the MD simulation for C2X, NNCP and NNCT.
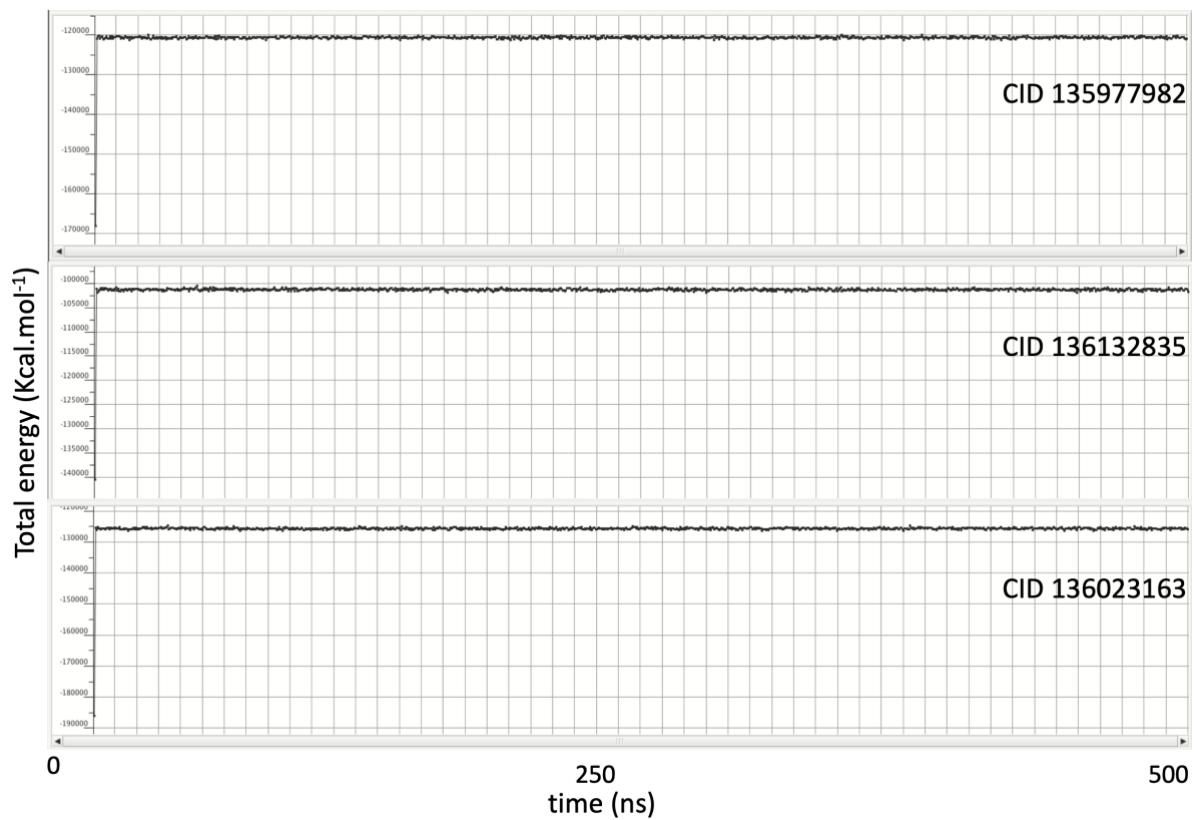
Figure 5.6. Plots of total energy during the MD simulation for CID 135977982, CID 136132835 and CID 136023163.
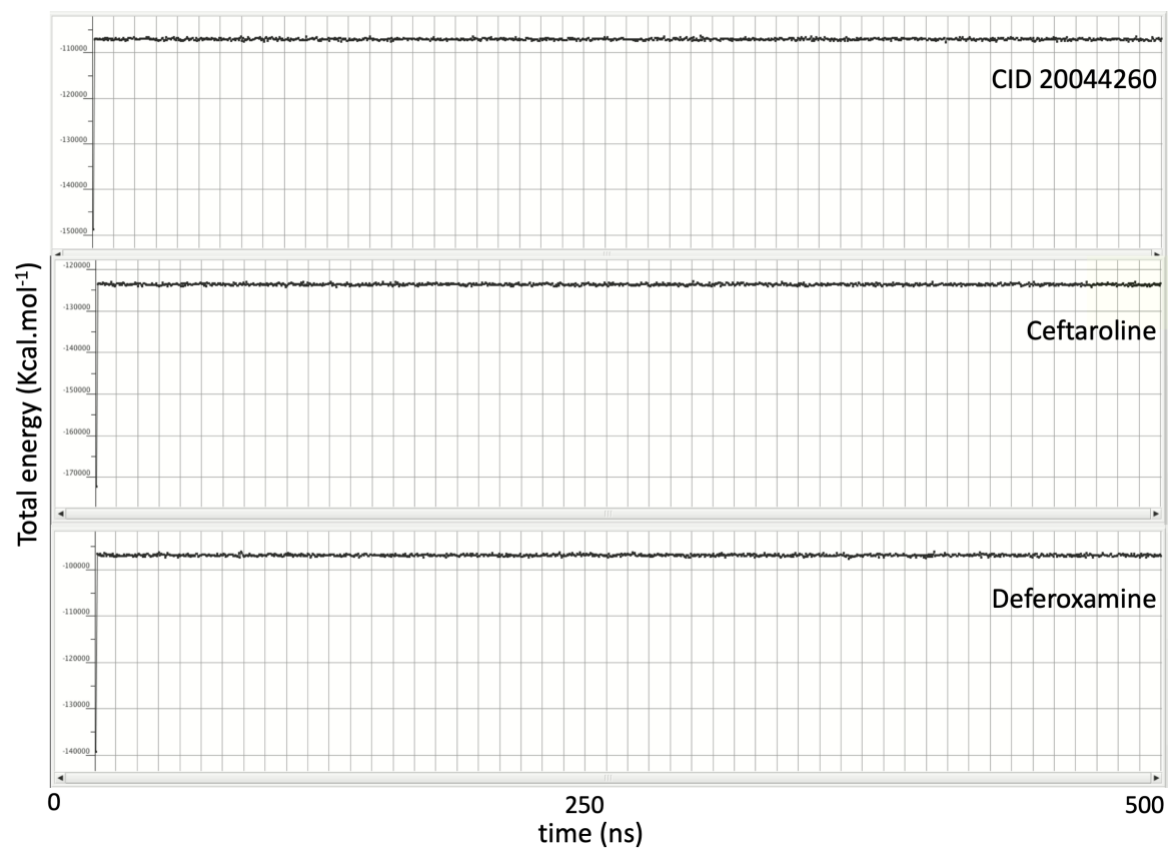
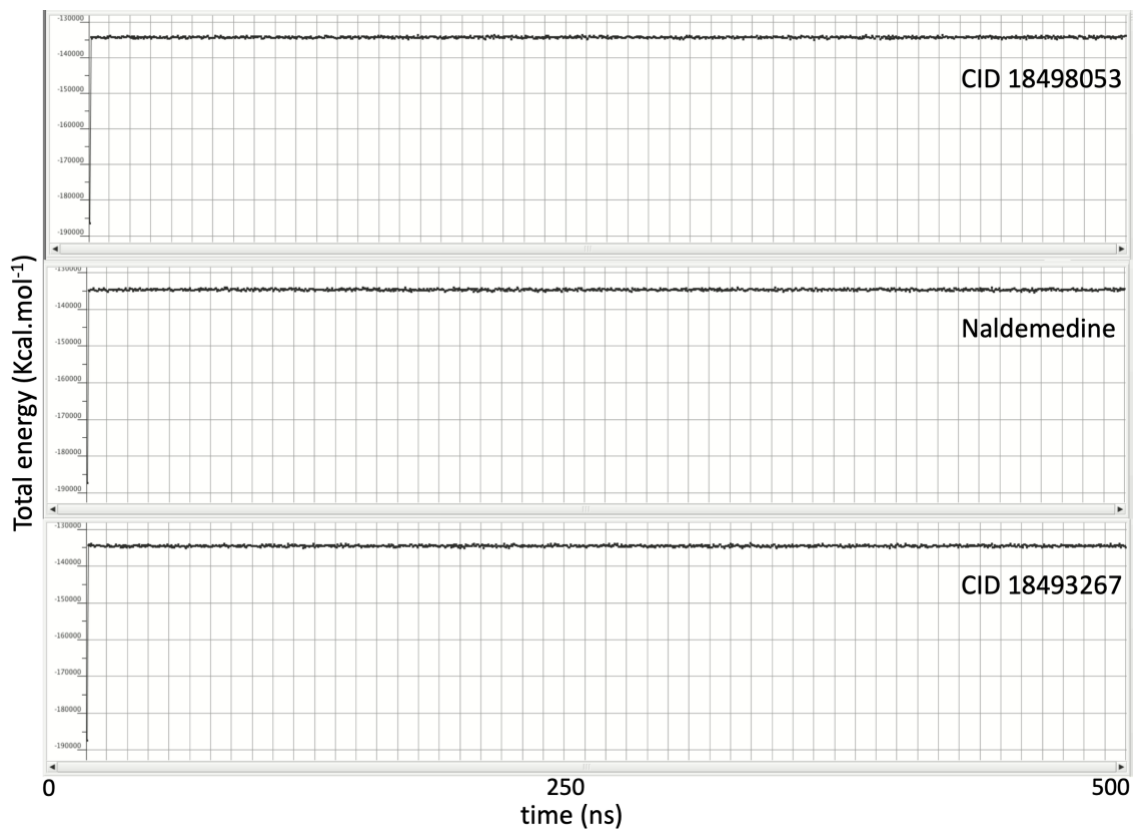**Figure 5.7. Plots of total energy during the MD simulation for CID 20044260 , Ceftaroline and Deferoxamine.**

Figure 5.8. Plots of total energy during the MD simulation for CID 18498053, Naldemedine and CID 18493267.
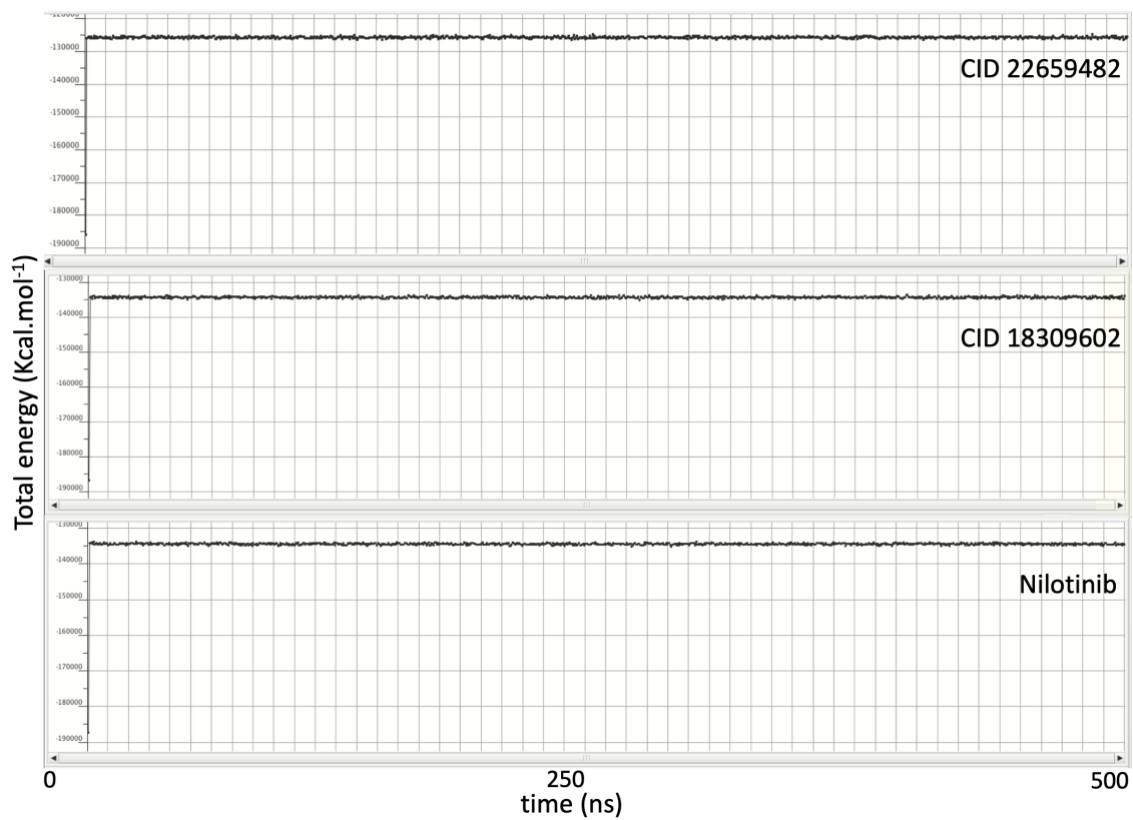
Figure 5.9. Plots of total energy during the MD simulation for CID 22659482, CID 18309602 and Nilotinib.
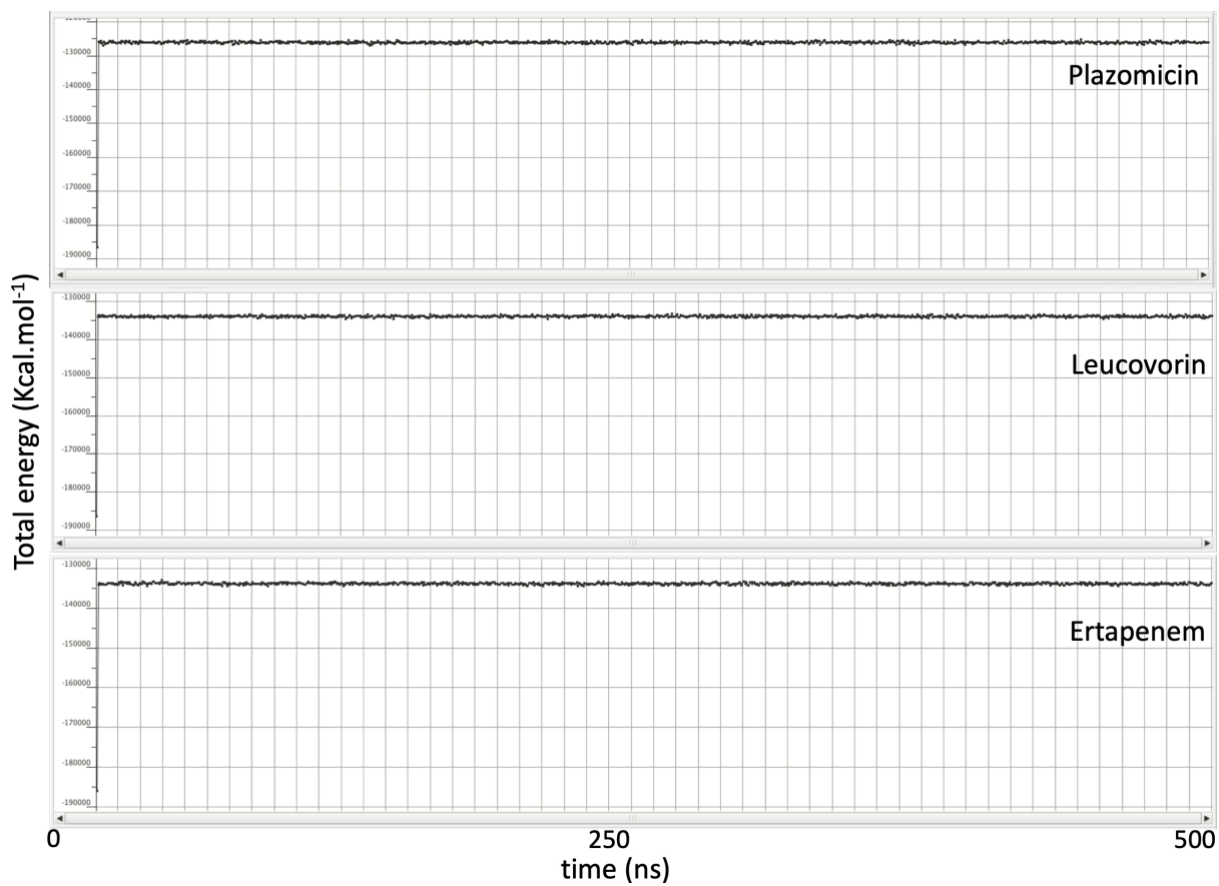
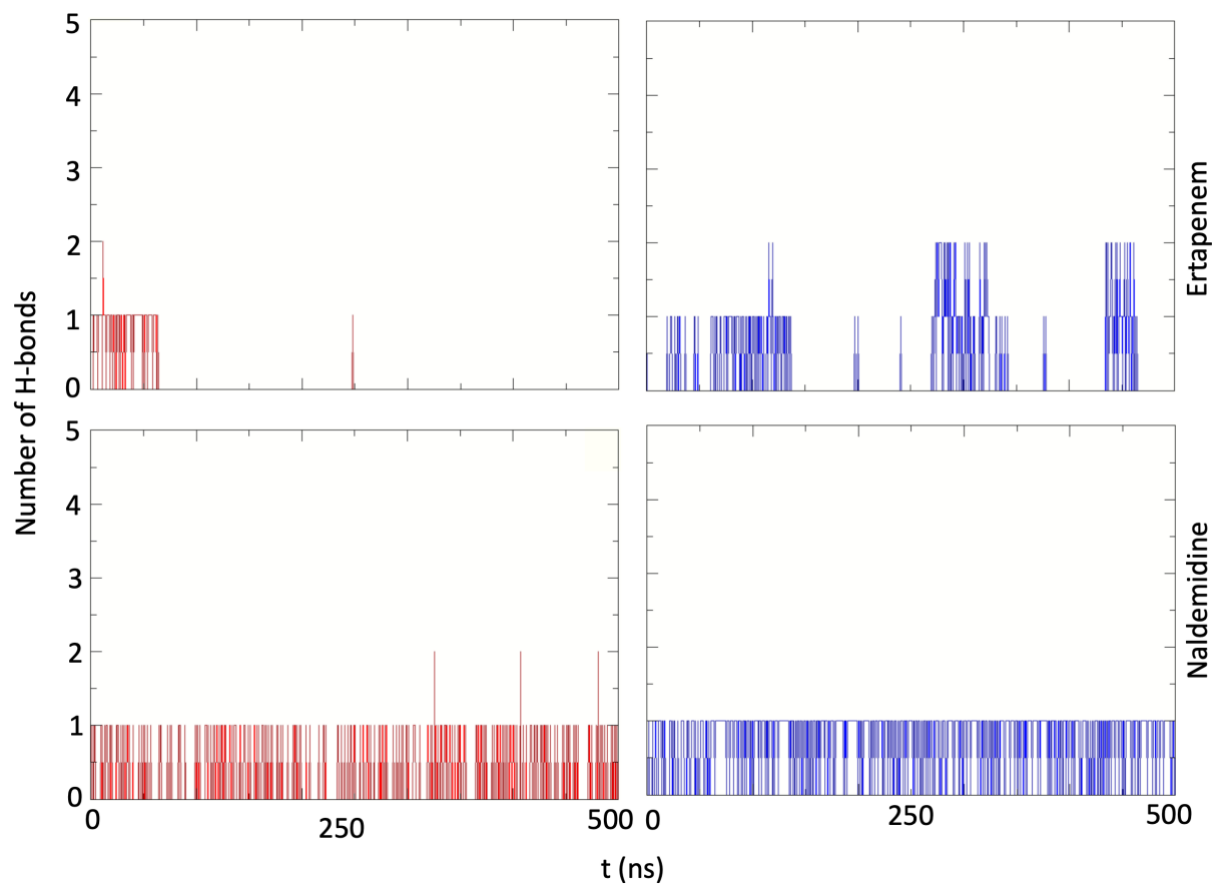**Figure 5.10. Plots of total energy during the MD simulation for Plazomicin, Leucovorin and Ertapenem.**
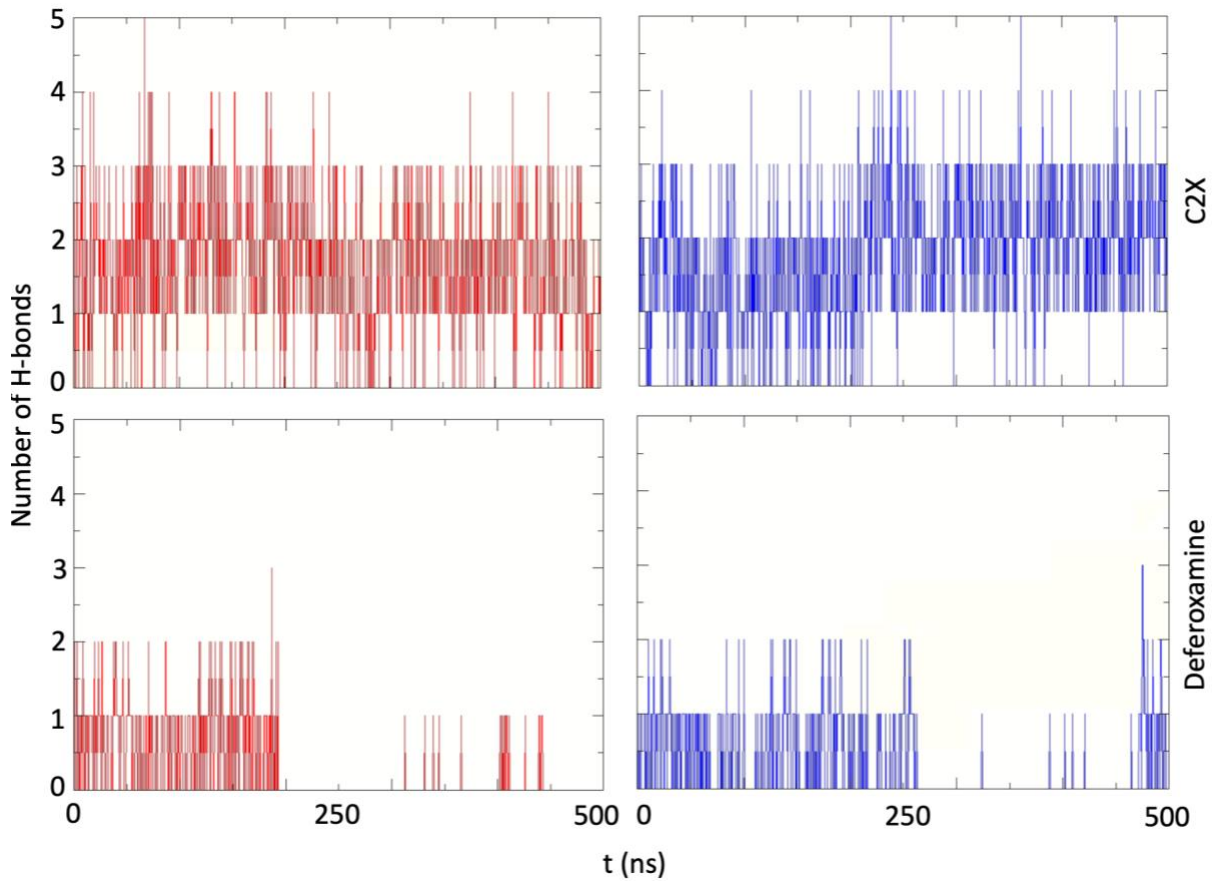
Figure 5.11. Plots of number of H-bonds formed by C2X and deferoxamine during the MD simulations with residues the catalytic (red lines) and the secondary (blue lines) sites of RTA.

Figure 5.12. Plots of number of H-bonds formed by ertapenem and naldemidine during the MD simulations with residues the catalytic (red lines) and the secondary (blue lines) sites of RTA.
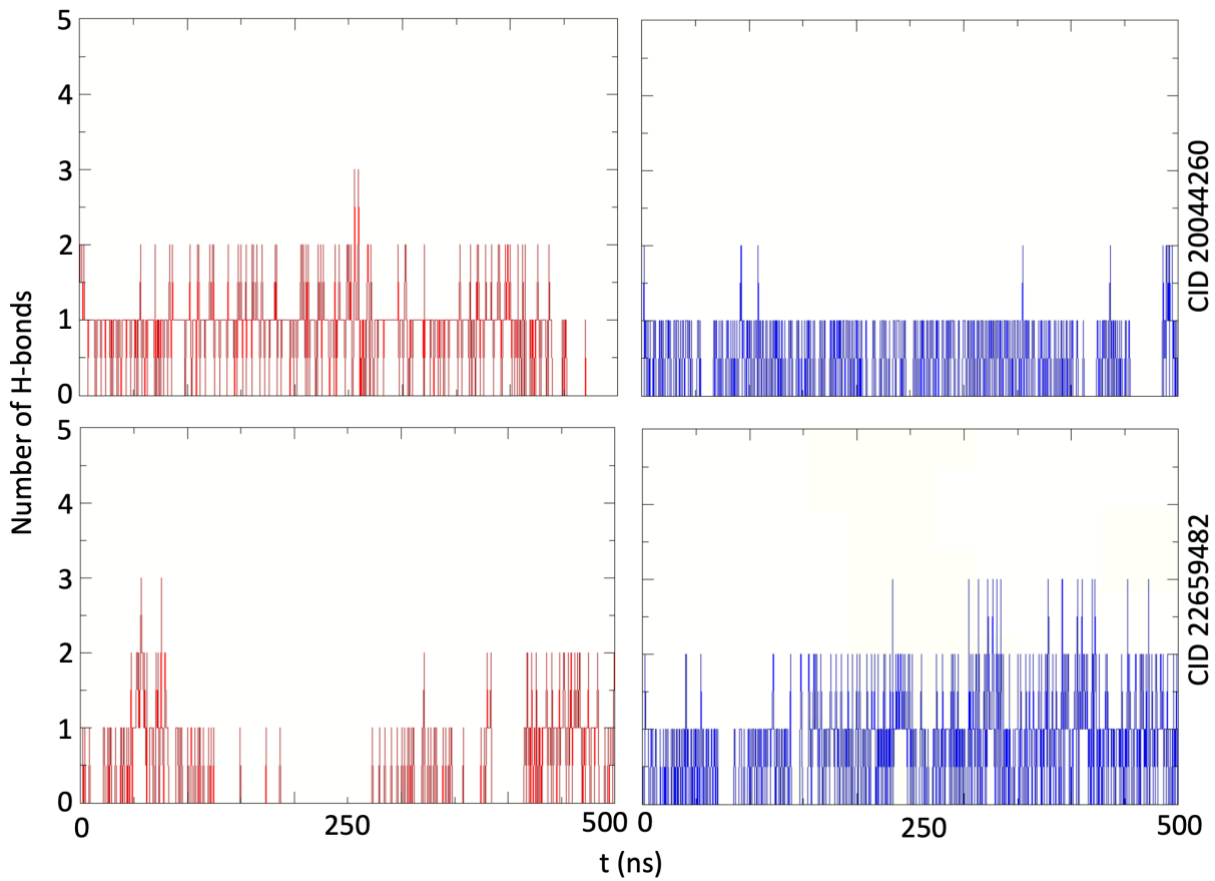
Figure 5.13. Plots of number of H-bonds formed by CID 20044260 and CID 22659482 during the MD simulations with residues the catalytic (red lines) and the secondary (blue lines) sites of RTA.
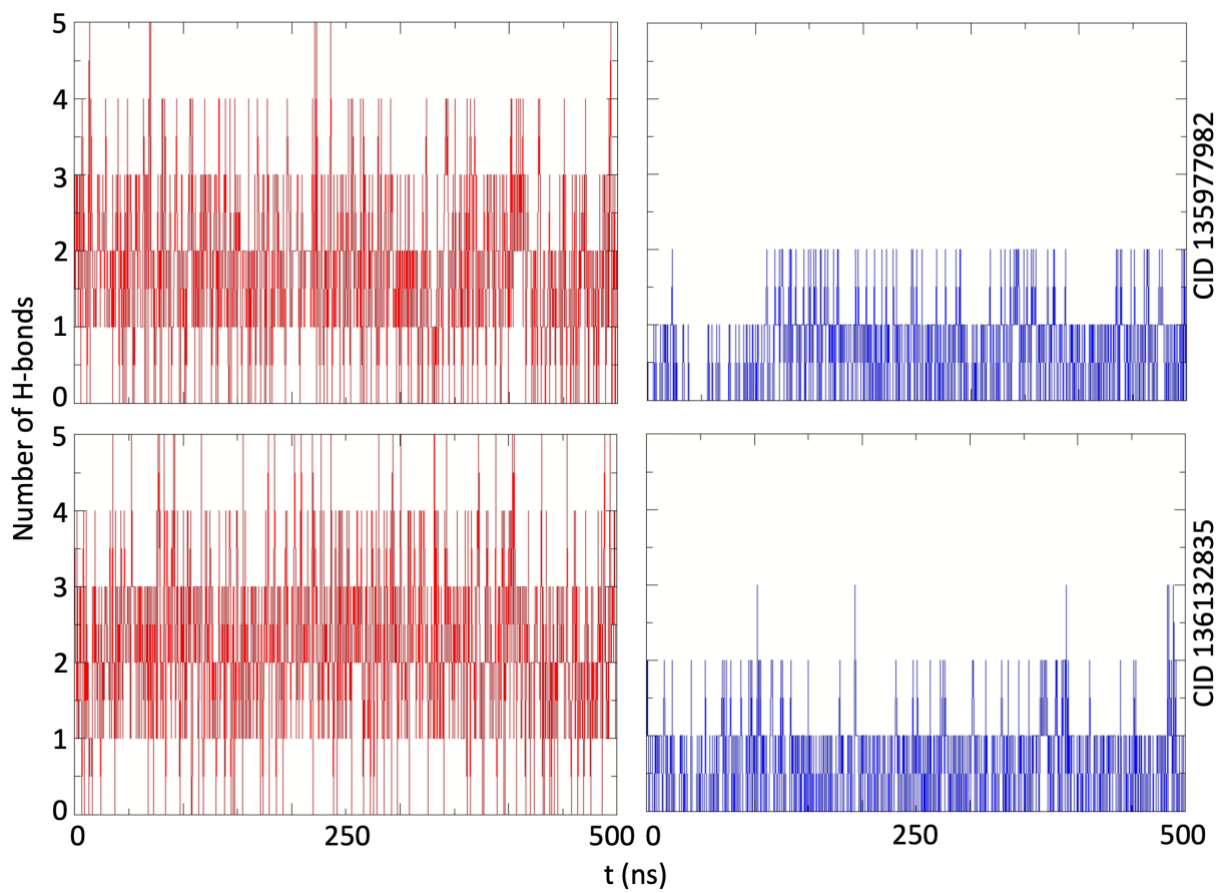
Figure 5.14. Plots of number of H-bonds formed by CID 135977982 and CID 136132835 during the MD simulations with residues the catalytic (red lines) and the secondary (blue lines) sites of RTA.
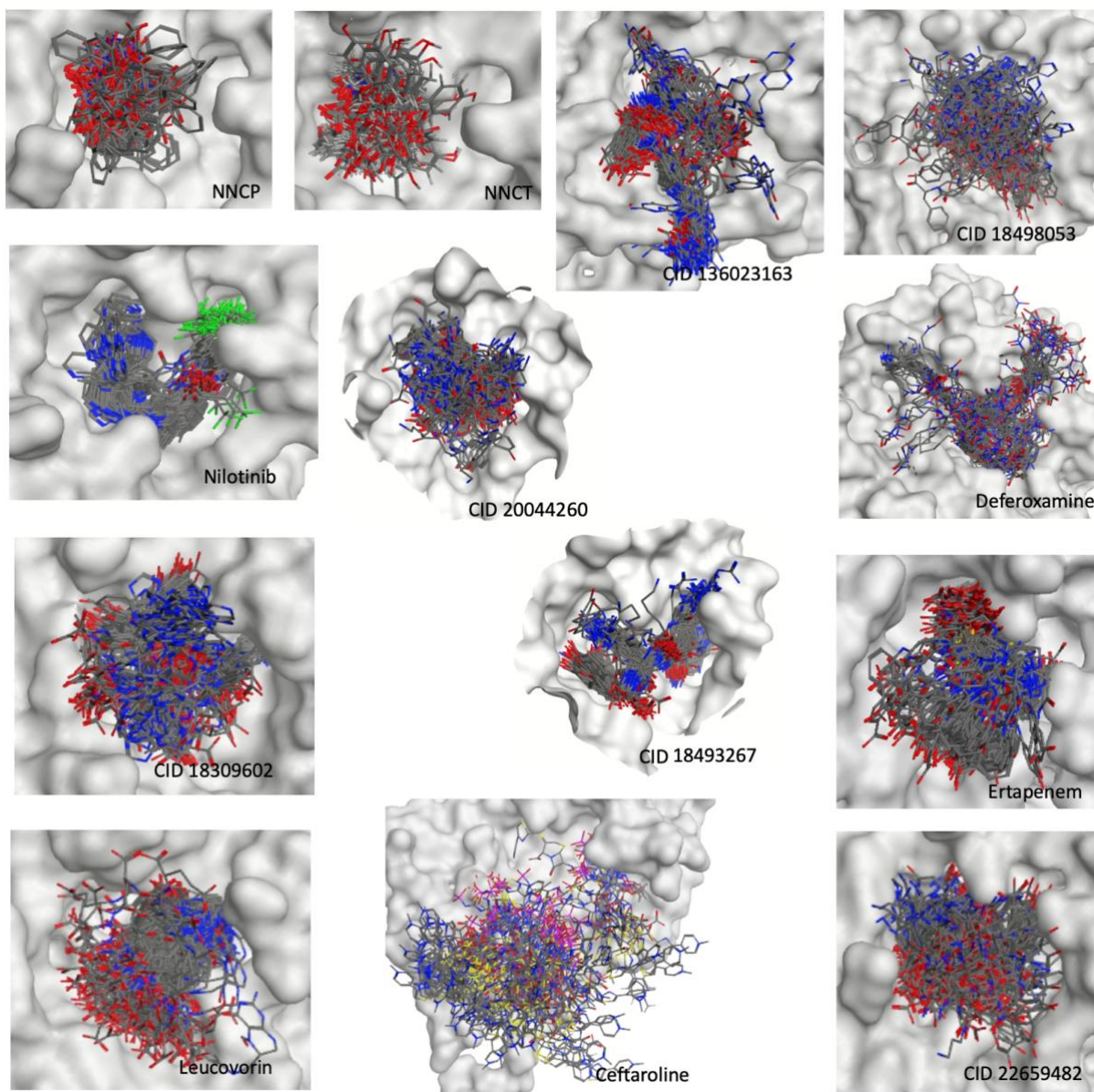
Figure 5.15. Superposition of 100 frames of no dual binders collected during the MD simulations. Hydrogens were omitted for better clarity.