

POLITECNICO DI TORINO
Master's Degree in Biomedical Engineering

Master's Degree Thesis

Computer-aided drug design of Ca²⁺ agonists to
enhance the effect of anti-CD20 monoclonal
antibodies for CD20-positive non-Hodgkin's
lymphoma and chronic lymphocytic leukemia



Supervisor

Prof. DERIU MARCO AGOSTINO

Candidate

ANGELA FERRARA

Co-Supervisor

Prof. CHARLES DUMONTET
Prof. PRETO JORDANE
PALLANTE LORENZO

ACADEMIC YEAR 2019-2020

TABLE OF CONTENTS

Abstract
Estratto

1. INTRODUCTION
2. BIOLOGICAL BACKGROUND
 - 2.1 Non-Hodgkin's lymphoma (NHL) and Chronic lymphocytic leukemia (CLL)
 - 2.2 Anti-CD20 monoclonal antibodies (mAbs)
 - 2.3 Ca²⁺ channels expression in lymphomas
 - 2.3.1 L-type voltage-gated Ca²⁺ channels
 - 2.3.2 Agonist and Antagonist of LTCC
3. MATERIALS AND METHODS
 - 3.1 Computational Modelling of Biomolecular Systems
 - 3.2 Molecular Mechanics
 - 3.2.1 *Potential Energy Function*
 - 3.2.2 *Treatment of Bond and Non-Bond interactions*
 - 3.2.3 *Periodic Boundary Conditions*
 - 3.2.4 *Potential Energy Minimization*
 - 3.3 Molecular Dynamics
 - 3.3.1 *Statistical Ensemble*
 - 3.3.2 *Molecular Dynamics implementation scheme*
 - 3.3.3 *Software package*
 - 3.4 Methods of Analysis
 - 3.4.1 *Cluster Analysis*
 - 3.4.2 *Distance analyses*
 - 3.5 Docking and Virtual screening
 - 3.5.1 *Ligand and receptor*
 - 3.5.2 *MaxMin-Picker algorithm*
 - 3.5.3 *Search Algorithms*
 - 3.5.4 *Binding Energy*
 - 3.5.5 *Software package*
4. STUDY OF NEW COMPOUNDS WITH CA²⁺ AGONIST ACTIVITY
 - Abstract
 - 4.1 Introduction
 - 4.2 Materials and Methods
 - 4.2.1 *Structure preparation of channel*
 - 4.2.2 *Preparation of ligands*
 - 4.2.3 *MD Production*
 - 4.2.4 *Docking for Virtual Screening*
 - 4.2.4.1 Preparation of Binding site
 - 4.2.4.2 Zinc ID
 - 4.2.4.3 Molecular Docking
 - 4.3 Results
 - 4.3.1 *Analysis of ligand interaction*

- 4.3.2 Analysis of conformational changes*
 - 4.3.3 Docking*
 - 4.4 Discussion

5. CONCLUSIONS AND FUTURE PERSPECTIVE

6. SUPPORTING INFORMATION

7. ACKNOWLEDGMENTS

8. REFERENCES

ABSTRACT

Non-Hodgkin's lymphomas (NHL) are clonal lymphoproliferative diseases that originate from B lymphocytes. NHL also comprises the Chronic Lymphatic Leukemia (CLL), a type of blood cancer. Patients affected by these pathologies possess high amounts of differential antigens CD20. In this context, anti-CD20 monoclonal antibodies (mAbs) increase the influx of Ca^{2+} and lead to cell apoptosis after interacting with CD20. Anti-CD20 mAbs bind to CD20 transmembrane antigen, they were found on pre-B lymphocytes and mature B lymphocytes. The CD20 receptor is involved in the regulation of Ca^{2+} influx in B lymphocytes and the cells exposure to anti-CD20 antibodies causes a strong induction of cell death, suggesting a correlation between the and Ca^{2+} influx and the cytotoxic effect of the antibodies.

Therefore, it is worth interesting to characterize small molecules able to keep the calcium channels in an open form and to develop new molecules that can enhance the influx of Ca^{2+} , such as agonists of L-type calcium channels (LTCC). The main hypothesis addressed in this study is improve the actions of anti-CD20 monoclonal antibodies in inducing apoptosis by increasing the flow of calcium inside the cell.

After studying the interactions of the dihydropyridine (DHP) agonist of the calcium channel BAY-K 8644, the aim of the thesis work is to find a drug with the same characteristics in terms of structure and binding sites involved in the protein-binding interactions. All compounds that will be inside the binding pocket of the reference agonist will be discriminated and those with greater interactions in terms of higher binding energy will be chosen.

With these considerations, the present project will combine computer-aided-drug design (CADD), experimental results of previous studies obtained by the Anticancer Antibodies team on calcium channel blockers and the antitumor activity of anti-CD20 monoclonal antibodies and Molecular Dynamics (MD) simulations for the analysis and research of new compounds to support the action of anti-CD20 mAbs, selectively targeting the calcium ions channel and consequently increasing the calcium flow.

The outcome of this work will show the conformational effects on the calcium channel after the agonist/antagonist binding and will perform an extensive virtual screening among a huge number of compounds, exploring the mode of binding of investigated chemicals and highlighting differences in terms of binding sites and main residues involved.

ESTRATTO

I linfomi non Hodgkin (NHL) sono malattie linfoproliferative clonali che hanno origine dai linfociti B. La NHL comprende anche la leucemia linfatica cronica (CLL), un tipo di cancro del sangue. I pazienti affetti da queste patologie possiedono elevate quantità di antigeni differenziali CD20. In questo contesto, gli anticorpi monoclonali anti-CD20 (mAbs) aumentano l'afflusso di Ca^{2+} e portano all'apoptosi cellulare dopo aver interagito con CD20. Gli mAb anti-CD20 si legano all'antigene transmembrana CD20, sono stati trovati sui linfociti pre-B e sui linfociti B maturi. Il recettore CD20 è coinvolto nella regolazione dell'afflusso di Ca^{2+} nei linfociti B e l'esposizione delle cellule agli anticorpi anti-CD20 provoca una forte induzione della morte cellulare, suggerendo una correlazione tra l'afflusso di Ca^{2+} e l'effetto citotossico degli anticorpi.

Pertanto, vale la pena caratterizzare piccole molecole in grado di mantenere i canali del calcio in forma aperta e sviluppare nuove molecole in grado di potenziare l'afflusso di Ca^{2+} , come gli agonisti dei canali del calcio di tipo L (LTCC). La principale ipotesi affrontata in questo studio è migliorare le azioni degli anticorpi monoclonali anti-CD20 nell'indurre l'apoptosi aumentando il flusso di calcio all'interno della cellula.

Dopo aver studiato le interazioni dell'agonista dihidropiridinico (DHP) BAY-K8644 del canale del calcio, lo scopo del lavoro di tesi è quello di trovare un farmaco con le stesse caratteristiche in termini di struttura e siti di legame coinvolti nelle interazioni di legame proteico. Verranno discriminati tutti i composti che saranno all'interno della sacca di legame dell'agonista di riferimento e saranno scelti quelli con maggiori interazioni in termini di maggiore energia di legame.

Con queste considerazioni, il presente progetto combinerà computer-aided-drug design (CADD), risultati sperimentali di studi precedenti ottenuti dal team di Anticancer Antibodies sui bloccanti dei canali del calcio e l'attività antitumorale degli anticorpi monoclonali anti-CD20 e Molecular Dynamics (MD) simulazioni per l'analisi e la ricerca di nuovi composti a supporto dell'azione degli mAbs anti-CD20, mirando selettivamente al canale degli ioni calcio e aumentando di conseguenza il flusso del calcio.

Il risultato di questo lavoro mostrerà gli effetti conformazionali sul canale del calcio dopo il legame con l'agonista e con l'antagonista ed eseguirà un ampio screening virtuale tra un enorme numero di composti, esplorando le modalità di legame delle sostanze chimiche studiate ed evidenziando le differenze in termini di siti di legame e principali residui coinvolti.

Il risultato di questo lavoro mostrerà le interazioni che i diversi composti hanno con il canale del calcio. In particolare, verrà utilizzata la struttura del canale del calcio legata ad un agonista e la stessa struttura legata ad un antagonista.

Verranno identificate le differenze in termini di siti di legame e residui coinvolti nel legame e i risultati saranno confrontati con i risultati dei diversi composti analizzati.

CHAPTER I – Introduction

In this chapter we will introduce the present master's thesis project, the starting hypothesis and the steps for its validations. At the end there is a summary about all the chapters treated in this composition.

Cancer research has always been committed to the discovery of new therapies for the treatment of cancer. The deepening of the pathologies involved allows to produce new strategies to be implemented.

Non-Hodgkin's lymphomas (NHL) affect the lymphatic system and it consists in transformation of lymphocytes creating neoplastic cells. The lymphatic system it is composed by lymphatic organs, such as the bone marrow, tonsils, thymus and spleen, and a series of lymph nodes and lymphatic vessels. The large structure of this system means that this type of lymphoma does not have well-defined areas in which it can develop and rarely affects a single organ.

Chronic Lymphatic Leukemia (CLL) is a group of blood cancer and a subtype of NHL. CLL is the consequence of numerous biological and genetic events affecting mature B lymphocytes and is characterized by an initial proliferation of these cells in the bone marrow and peripheral blood, followed by the involvement of the lymphatic structures that increase in size. The main therapeutic agents in the treatment of patients with differential antigens CD20-positive NHL and CCL are monoclonal antibodies (mAbs) [8]. They work through by complement-mediated cytotoxicity that provides the desired anti-cancer effect or they release toxins and radiation to selectively target affected cells. The mAbs are used for the treatment of these diseases, since monoclonal antibodies, such as anti-CD20. Recent results obtained by the Anticancer Antibodies team^[1] reported that inhibition of the calcium flux with Ca²⁺ channel blockers as used to treat hypertension or angina, but the patients receiving CCB and concomitant therapy anti-CD20 for diseases such as NHL or CLL have the worst survival. This show that CCBs negatively impact the outcomes of patients receiving anti-CD20 mAbs.

Anti-CD20 monoclonal antibodies (mAbs) upregulates the expression of early growth response-1 (EGR-1), by increasing Ca²⁺ concentration inside the cells which trigger ERG-1 expression. These data lead to hypothesize that calcium agonists can enhance and support the therapeutic effects of anti-CD20 antibodies.

Regarding calcium channel antagonists, there is a lot of information regarding the binding sites^[20-23]. The starting point for our project will be to identify the preferential state for the correct binding of the compounds.

The first class we will analyze will be that of dihydropyridines (DHPs) which have both an agonistic and antagonistic effect. In general, agonists are hypothesized to increase calcium flux within the channel because they prolong the opening time, a difference of antagonists who prefer the closed inactivated state. Some Cav agonists are known from literatures: FPL64176 ^[24], and CGP 485065, murrayafoline A ^[25] at low concentration.

Computational approaches are used for the analysis of complex allowing to study in detail the molecular mechanisms, moreover the docking is a powerful means to study the position that a compound assumes in the binding with the protein and to evaluate its affinity in terms of binding energy. In this study, Molecular Dynamics (MD) simulation, Ensemble Docking and the calculation of binding free energy will be used to perform a skimming of the compounds chosen for the analysis and then guide the research towards a certain type of compound for the search of new medications.

This manuscript is divided into 6 sections briefly described below.

Chapter 1 is this introductory part.

Chapter 2 describes the biological context of this work. In details, section 2.1 describes the two main diseases that are of interest to our project: Non-Hodgkin's Lymphoma and Chronic Lymphocytic Leukemia. In section 2.2, the agents used for the treatment of such diseases are introduced. Section 2.3 deals with the mechanisms of calcium channels with an overview on the structure and the different types of calcium channels. In this context, section 2.3.1 is a description of L-type calcium channel (LTCC) and in section 2.3.2 agonists and antagonists of this type of channel will be presented.

Chapter 3 describes in a general way the materials and methods used in this work. In section 3.1 a description of Computational Modeling of Biomolecular Systems is made. Molecular mechanics and molecular dynamics are described in sections 3.2 and 3.3 analyzing specifically the theoretical and physical aspects. Section 3.4 shows the analysis criteria followed by the description in section 3.4 of Molecular Docking and Virtual Screening, with a focus on search algorithms, calculation of free energy binding and scoring functions. The software used is also presented and described.

Chapter 4 presents the body of the work with an introduction of the context in which the work takes place and the purpose of the project in chapter 4.1.

Subsequently, in section 4.2 the materials and methods applied are described in detail, with a focus on channel preparation, ligand preparation, MD production and molecular docking.

Particular attention is paid in the description of the preparation of binding sites in the section 4.2.4.1 and in the choice of compounds used subsequently for molecular docking in 4.2.4.2 Section 4.3 shows the analysis of the results obtained in this work and the conclusions.

Chapter 5 is devoted to discussions and future developments.

Chapter 6 includes supporting information for a greater understanding of some aspects of the work and for further information.

CHAPTER II - Biological background

This chapter provides the biological context in which the thesis work was developed. In details, section 2.1 describes the two main diseases that are of interest to our project: Non-Hodgkin's Lymphoma and Chronic Lymphocytic Leukemia. In section 2.2, the agents used for the treatment of such diseases are introduced. Section 2.3 deals with the mechanisms of calcium channels with an overview on the structure and the different types of calcium channels. In this context, a L-type calcium channel (LTCC) will be investigated and agonists and antagonists of this type of channel will be presented.

2.1 Non-Hodgkin's lymphoma and Chronic lymphocytic leukemia.

Non-Hodgkin's lymphoma (NHL) is a group of blood cancers characterized by clonal proliferation of B lymphocytes (80-85% of cases), T lymphocytes (15-20%), or natural killer, NK, cells (rare)^[5]. NHL represents 4-5% of all cancers both in the male and in the female population and it is the ninth leading cause of cancer death in men and sixth in women. NHLs are classified based on the cell type (B, T or NK lymphocytes), on morphological, immunophenotypic, genetic and molecular criteria, to be integrated with the clinical presentation characteristics. First-line chemotherapy treatment in patients is based on the standard RCHOP schedule, i.e., rituximab, cyclophosphamide, doxorubicin 50 mg/m², vincristine, prednisone, combined with rituximab and, if is required, radiotherapy^[6]. The number of cycles depends on the initial presentation and on the response to therapy. Second-line therapy can either consist in (1) high-dose chemotherapy (HDS) followed by stem cell autograft (ASCT) for young patients who have relapsed or are resistant to the first-line protocol or in (2) rituximab therapy combined with traditional chemotherapy regimens for elderly patients who have relapsed, are resistant to the first-line treatment, or are not candidates for ASCT. Novel therapeutic agents like obinutuzumab, brentuximab vedotin, ibrutinib, polatuzumab, CAR-T, are being studied for the treatment of these patient categories.

Chronic Lymphatic Leukemia (CLL) is a group of blood cancer and a subtype of NHL. CLL is the consequence of numerous biological and genetic events affecting mature B lymphocytes and is characterized by an initial proliferation of these cells in the bone marrow and peripheral blood, followed by the involvement of the lymphatic structures that increase in size.

Given the proliferation of neoplastic lymphocytes in the bone marrow, CLL is characterized by a reduced production of normal blood cells, with typical subsequent signs and symptoms: anemia, reduction of red blood cells (resulting in chronic fatigue and easy fatigue), platelet disorders, that is, reduction of platelets (with a consequent greater risk of bleeding) and neutropenia corresponding to a reduction in the population of cells responsible for defense against microorganisms and greater predisposition to develop infections. Autoimmune disorders can also be found. CLL is a disease with an extremely heterogeneous course and it remain stable for years and require only periodic checks, and aggressive forms that evolve rapidly. In most cases, however, it is a slowly progressing disease. In asymptomatic patients, the current strategy is not to treat the patient. In symptomatic ones, current therapies include both chemoimmunotherapy and target therapies and biologics^[7].

2.2 *Anti-CD20 monoclonal antibodies*

The main therapeutic agents in the treatment of patients with differential antigens CD20-positive NHL and CCL are monoclonal antibodies (mAbs) ^[8]. They work through by complement-mediated cytolysis that provides the desired anti-cancer effect or they release toxins and radiation to selectively target affected cells. MAbs used for cancer therapy are antibodies belonging to the Class G of immunoglobulins (IgG), and as the components of that class contain two light chains and two heavy chains. Based on the applications of these antibodies they can be divided in two groups. The type of mAbs interested in this study is belonging to the first group and antibodies are directed towards the antigens expressed by the normal cells, such as CD20 (differential antigens). For several years, chemotherapy alone was the only approach used in the treatment of CLL and NHL ^[9]. The introduction of monoclonal antibodies such as anti-CD20 (rituximab) into clinical practice has allowed evaluating the efficacy of the chemo-immunotherapy association in numerous clinical trials. Anti-CD20 mAbs bind to CD20 transmembrane antigen, a non-glycosylated phosphoprotein, which is found on pre-B lymphocytes and mature B lymphocytes.

Exposure of B-lymphoma cell lines to anti-CD20 monoclonal antibodies (mAbs) upregulates the expression of early growth response-1 (EGR-1), by increasing Ca²⁺ concentration inside the cells which trigger ERG-1 expression.

EGR-1 encodes a zinc finger transcription factor whose expression is highly regulated by various extracellular stimuli, including calcium influx.

Moreover, recent results obtained by the Anticancer Antibodies team indicate the involvement of the CD20 receptor in the regulation of Ca²⁺ influx in B lymphocytes and that compounds that block calcium channels have been shown to impair the action of anti-CD20 antibodies [10-11].

EGR-1 is a key protein in the modulation of apoptosis by CD20 signaling. After exposure to anti-CD20 antibodies, there is a strong overexpression of EGR-1 at the transcriptional and protein levels. At the same time, exposure to anti-CD20 antibodies causes a strong induction of cell death, suggesting a correlation between the induction of EGR-1 and the cytotoxic effect of the antibodies.

With these data, the main hypothesis addressed in this study is improve the actions of anti-CD20 monoclonal antibodies in inducing apoptosis, in particular, the goal is to find new compounds that have agonists activity to validate the hypothesis.

2.3 Voltage-gated calcium channels

Calcium ion channels are present in all cells and perform many fundamental functions for the body, widespread in all cell types and regulate many fundamental physiological functions [12].

Two main classes of channels mediate the regulation of Ca²⁺ homeostasis voltage-gated Ca²⁺ channels and voltage-free calcium channels.

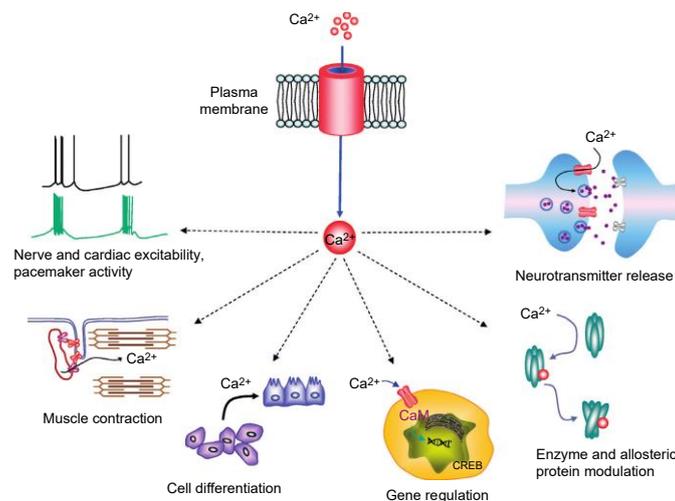


FIGURE 1 | Schematic representation of the influence of the opening of calcium channels on different types of cells. Adapted from Encyclopedia of Neuroscience, volume 10, pp. 427-441.

2.3.1 *L-type voltage-gated Ca²⁺ channels*

Voltage-dependent channel is a membrane protein that regulates the transport of calcium ions across the membrane, the membrane potential controls the opening and closing of the channel ^[13].

The membrane in question has a resting potential of -75mV, once the stimulus is applied there is a depolarization (the membrane potential ranges from -80mV to +30mV). Subsequently, there is a repolarization where the membrane potential goes from +30mV up to values more negative than those of the resting potential (hyperpolarization).

The voltage-dependent channels have at least one "gate" for opening and closing. The voltage variation is able to determine a conformational variation of these channels.

Voltage-gated Ca²⁺ (Cav) channels by membrane depolarization allow Ca²⁺ to enter excitable cells ^[14].

Calcium currents can be distinguished according to the membrane potential: the low voltage-activated calcium channels (type T) transiently open in response to small changes from resting potentials; high voltage-activated calcium which have a higher depolarization for activation.

Previous studies of Doumondet ^[1] have shown that L-type calcium channel was involved in the action of anti-CD20 antibodies.

Interestingly, Cav1.1 gene (CACNA1S) is the L-type Cav1.1 calcium channel gene and it exhibits overexpression in tumor cells compared to normal fabrics, in particular it is overexpressed in acute myeloid leukemia ^[15].

In human, Cav channels are proteins, composed of a pore-forming α_1 subunit, a transmembrane complex of 2 and subunits, an intracellular regulatory subunit, and in some cases a transmembrane subunit. The Cav1.1, like other Cav channels, is composed by a main subunit α_1 and other auxiliary subunits $\alpha_2\delta$, β , γ (Figure 2). The α_1 subunit of the calcium channel (170-240 kDa) is formed by 4 motif domains (I - IV), each consisting of 6 segments of membrane α -helices (S1 to S6) connected by P-loop between segments S5 and S6 (Figure 3). The extracellular domain α_2 and membrane-anchored δ subunits derive from the proteolytic cleavage of a single gene produced and are structurally bound by a disulfide bond.

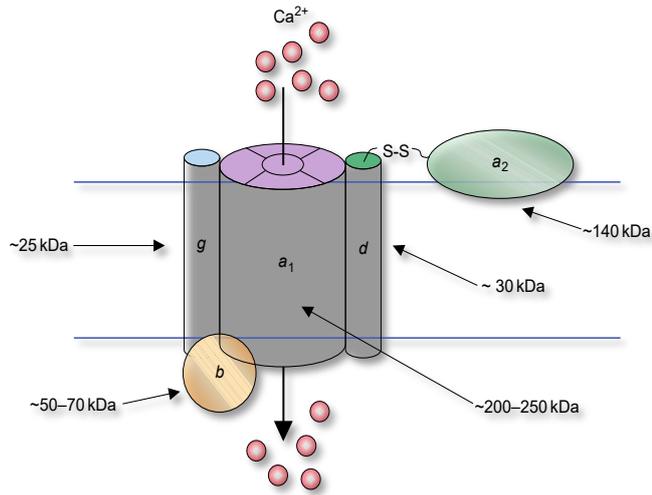


FIGURE 2| Molecular organization of the voltage gated Ca^{2+} channel subunits in the plasma membrane. [13]

β -subunits are cytoplasmic proteins capable of controlling the passage of calcium through the channels and are connected to the Cav subunits.

The $\alpha_2\delta$ subunit is present in the extracellular zone and binds to the membrane through the transmembrane segment formed by the δ subunit. The co-expression of $\alpha_2\delta$ subunits with the other calcium channel subunits affects current density, kinetics, and voltage-current relationships.

The γ subunits headed four transmembrane propellers (TM1 to TM4) with terminals N and C on the cytosolic side^[16-17].

However, α_1 subunit represents the central part governing of the mechanism of the channel, it can also be considered as a separate entity from the rest, i.e., the other parts are not essential to make the channel functional.

L-type calcium channels (LTCCs) can dwell in a resting state (bottom) and in an activated state (top), the pore can be in an open or closed state.

The voltage domains (VDs) are at rest in a bottom position and are activated in response to a depolarization. (Figure 4). Under these conditions the pore opens, calcium enters through the channel and there is a hyperpolarization that closes the channel^[18-19].

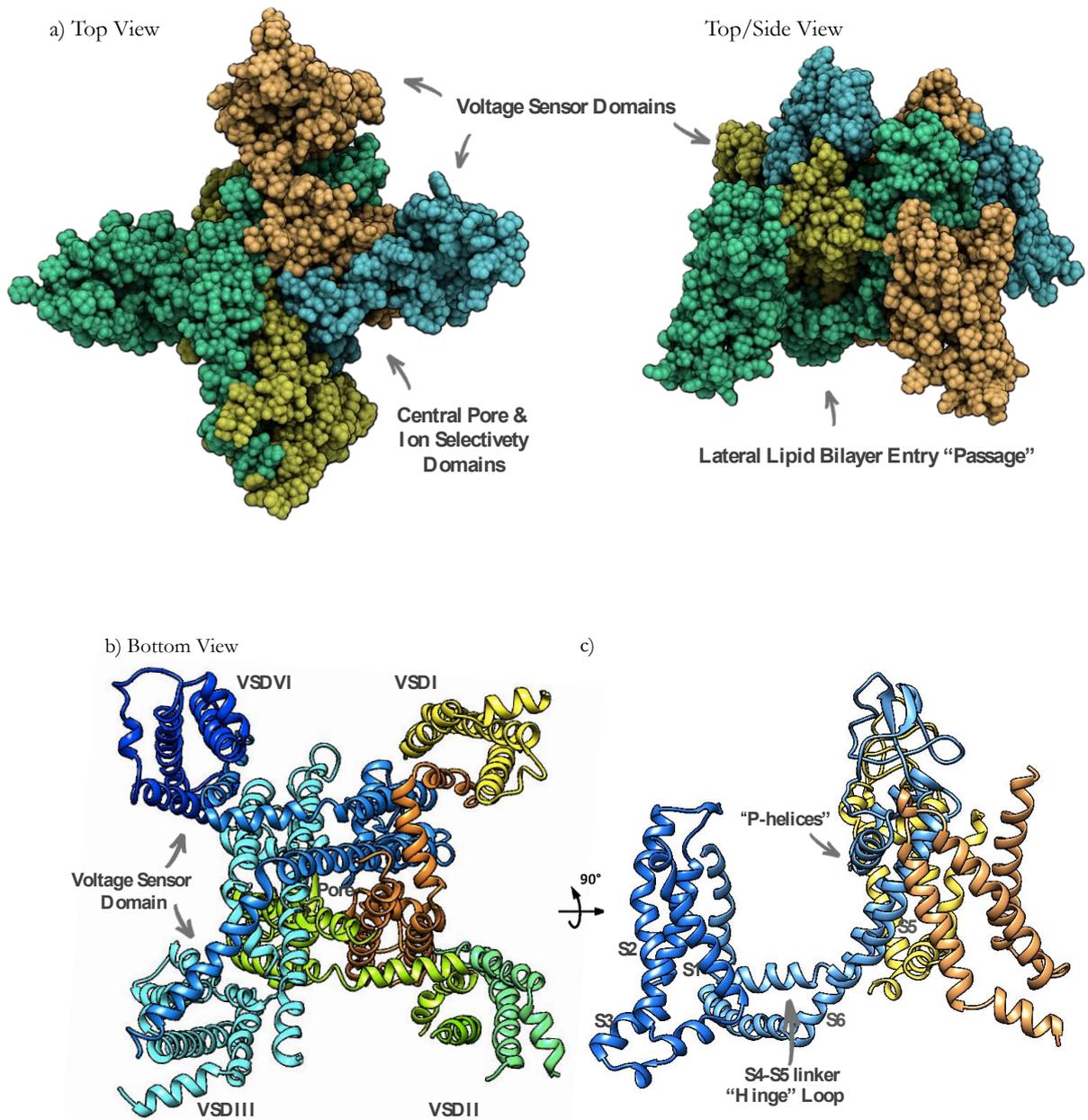


FIGURE 3 | Structure of calcium channels. (a) 3d illustration of the alpha1 subunit, viewed from above and from the side, respectively. the pore of the central channel is formed by a homo-tetrameric association. the voltage-sensitive domains surround the central pore. (b) bottom view of the channel (flat 3d representation) where the domains are highlighted (from vsdi to vsdiv). the helices are represented as cylinders. domains i to iv are colored green, yellow, blue and pink, respectively. d) view from the canal side. s1 to s3 (green) and s4 (blue) represent the voltage sensor domain. s4 is connected to the s5 via the s4 - s5 linker (red). the transmembrane regions are shown in beige, the loops in blue. (c) representation of the surface of the structure (b), the colors represent the heteroatoms (images rendered using qtemol and vmd, the analyzed structure file is 6jp8).

Specifically, the channel states are: R (closed at rest state) where the VDs block the pore which remains closed, A (closed activated state) where the VDs change their position passing to the Up state but the pore remains closed, O (open state activated): the pore is open and the VS are in the up position, D (open state disabled) where the VDs have changed their conformation, but the pore is still open.

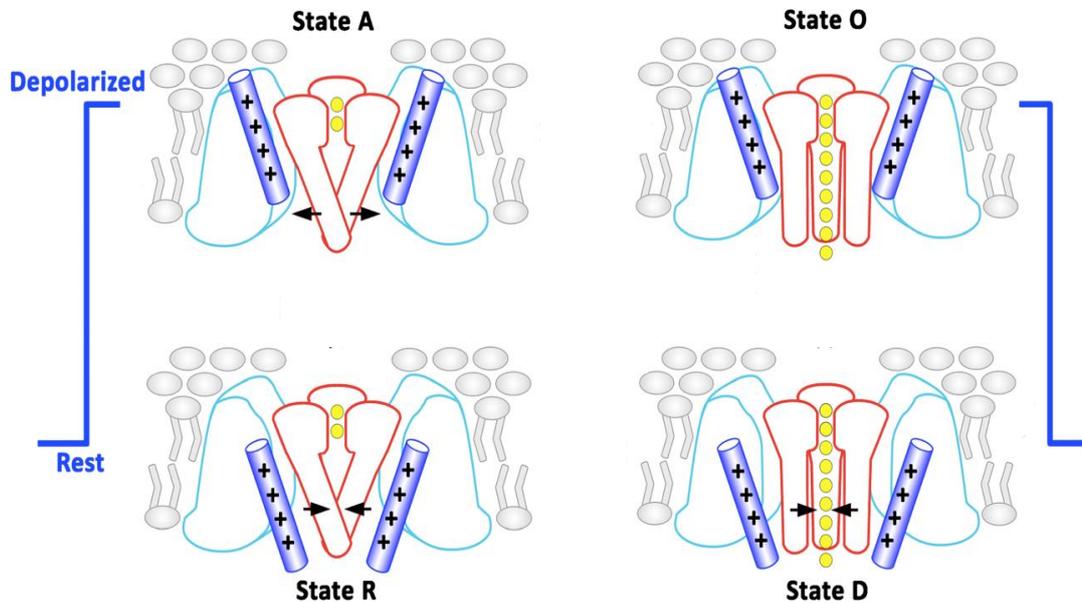


FIGURE 4| Schematic of the 4 states of calcium channels. Channel activation is influenced by two functionally separate processes: a voltage detection mechanism (++++) and the conducting pore.[16] For more information see figure S1 in SI.

2.3.2 Agonists and Antagonists of LTCC

Calcium channel blockers (CCBs) are molecular compounds that act directly on calcium influx which is why they are commonly used for the treatment of cardiovascular disease.

The use in patients with hypertension, cardiovascular decompensation, cardiac arrhythmias, patients with left ventricular diastolic dysfunction was effective.

Although calcium blockers are available with different chemicals, they share the common function of blocking the transmembrane flow of calcium ions through L-type and T-type channels.

Based on the chemical structure, calcium channel antagonists can be classified into dihydropyridines (DHPs), phenylalkylamines, benzothiazepines [20].

The drug class of dihydropyridines (DHP) are L-type calcium channel ligands and they are potent and selective blockers for Cav1.2 and Cav1.1 [21-23].

DHPs can operate as agonists or antagonists based on the chemical structure. Indeed, these compounds can change their function and go from agonist to antagonist (or vice versa) on a site-specific mutation of the calcium channel or due to a change in experimental conditions (Figure 5).

The non-dihydropyridine class includes phenylalkylamines and benzothiazepines which have a purely antagonistic function.

Recent results obtained by the Anticancer Antibodies team reported that inhibition of the calcium flux with Ca²⁺ channel blockers as used to treat hypertension or angina, but the patients receiving CCB and concomitant therapy anti-CD20 for diseases such as NHL or CLL have the worst survival. This shows that CCBs negatively impact the outcomes of patients receiving anti-CD20 mAbs.

There is some information of pore-binding Cav channel agonists compared with antagonists. This will be our starting point. Some Cav agonists are known from literatures: FPL64176 ^[24], and CGP 485065, murrayafoline A ^[25] at low concentration, and dihydropyridines (DHP) like (S)-BayK 8644 that can have a blocker or activator activity depending on the enantiomer³. Both the close or open state of the channel were used in previous in-silico studies.

However, it seems that LTCC ligands bind preferentially to the open inactivated state present at depolarized membrane potentials.

DHPs can have both an agonist and an antagonist function for this reason the chemical structure of these compounds has been investigated to understand which groups are discriminating to assume one function rather than the other.

In figure 6 all groups of the DHP chemical structure are presented. Structures of 1,4-DHPs were introduced in ^[26] as flatten-boat structures with the dihydropyridine group as the stern, an aromatic moiety at the bowsprit and various substituents at the port and starboard sides.

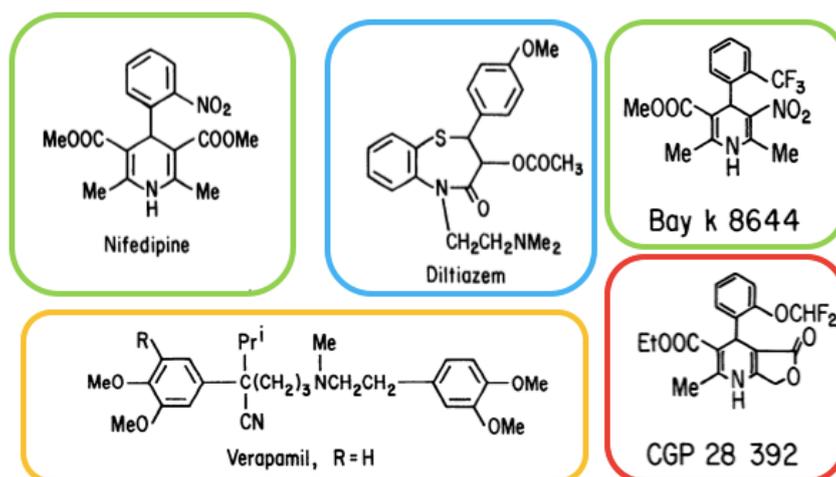


FIGURE 5 | Structural formulae chemical structure of three major classes of antagonists and agonists: blue benzothiazepine, green dihydropyridines antagonist (Nifedipine) and agonist (Bay-k8644), yellow phenylalkylamine, red non-dihydropyridines agonist.

The activity of the compound is due to the port-side group and the agonist or antagonist function depends on the fact that by binding to the target they can exhibit allosteric properties. The main feature of DHP is presented by the bowsprit group and the presence of the nitroxide functional group.

Experimental data reveal that the agonistic or antagonistic action is mainly determined by the nature of the port group in the ortho position of the DHP ring with respect to the bowsprit. In compounds with antagonist function there are two COOMe hydrophobic groups while the agonist structure has a hydrophilic substitute group, in the case analyzed the Trifluoromethyl group.

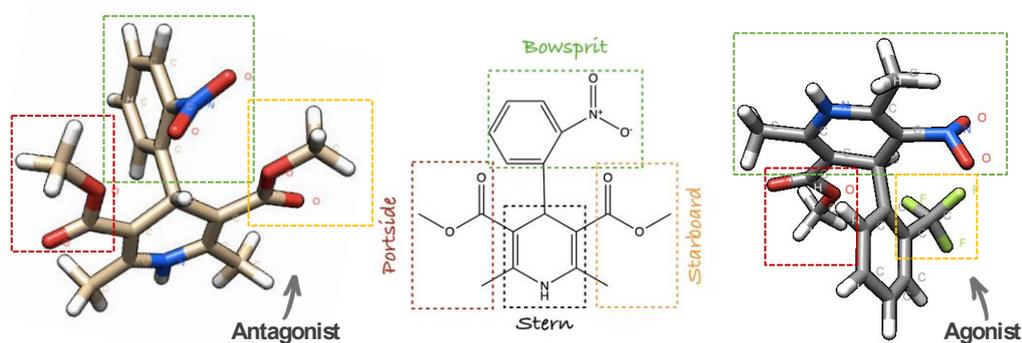


FIGURE 6 | 3D representation of a calcium antagonist (beige) and an agonist (gray). the discriminating functional groups are boxed according to the legend of the central 2D chemical structure.

The binding sites of the agonist and antagonist compounds appear to be the same and therefore their binding cannot be used to determine the function of the compounds.

In this context, the present project will combine computer-aided design (CADD), molecular modeling techniques capable of representing molecular mechanisms with atomistic resolution, experimental results to explore the druggability of L-type calcium channels in B-cell lymphoma.

CHAPTER III – Materials and methods

This chapter provides a general description of the materials and methods used for the master's thesis work. In detail, a brief introduction to Computational Modeling of Biomolecular Systems will be made (section 3.1), Molecular mechanics and molecular dynamics are described in sections 3.2 and 3.3 analyzing specifically the theoretical and physical aspects. Section 3.4 shows the analysis criteria followed by the description in section 3.4 of Molecular Docking and Virtual Screening, with a focus on search algorithms, calculation of free energy binding and scoring functions.

3.1 Computational Modeling of Biomolecular Systems

... An intelligence which could, at any moment, comprehend all the forces by which nature is animated and respective positions of the beings of which it is composed, and moreover, if this intelligence were far-reaching enough to subject these data to analysis, it would encompass in that formula both the movements of the largest bodies in the universe and those of the lightest atom: to it nothing would be uncertain, and the future, as well as the past would be present to its eyes. The human mind offers us, faint sketch of this intelligence [27].

Computer simulation used in different fields of application turns out to be a very powerful tool as it is possible to study specific aspects in detail acting as a bridge between the microscopic and macroscopic world of the laboratory. At the molecular level, Molecular Mechanics (MM) allows to use microscopic inputs in terms of masses of atoms making up a system and the same interactions between the various atoms.

The purpose of molecular dynamics (MD) is therefore to analyze and understand the macroscopic properties of even very complex systems because they generally consist of a large number of molecules [28-29]. To do this, the type of approach used is to solve classical and quantum physics equations. In particular, the forces on each atom can be calculated for each step and updated from time to time to analyze the velocity and position of these atoms. The trajectories of the simulation allow to obtain the atomic configuration of the system in a defined time range.

It is possible to investigate many properties in terms of protein folding and unfolding, the attachment of compounds to proteins and any conformation changes as this type of biochemical mechanisms occur on nanosecond or microsecond time scales.

MM and MD will be discussed in detail in the next sections.

3.2 Molecular Mechanics

Molecular mechanics uses Newton's equations for the modeling of complex systems such as proteins, macromolecules, nucleic acids. Atoms are considered as spherical particles with radius and charge. Bonds between atoms, on the other hand, are treated as springs and the stiffness value (k) clearly depends on the type of bonding atoms. This modeling allows to identify the stretching, torsional modification and bending of the bonds.

With molecular mechanics the energy associated with different molecular conformations can be measured, the estimate of this energy as a function of the atoms including the structure, is given by the set of equations that make up the molecular force field (FF).

3.2.1 Potential energy function

To estimate the molecular energy in its state of rest, the nuclear coordinates must be known. The energetic surface or surface tension thus defined is given by the variation of any variable of the system. For a molecular system composed of N atoms, the surface energy is given by two components, that is component of the bound and unbound terms.

$$V = V_{bonded} + V_{unbonded}$$

EQUATION 1 | Potential energy function.

These components are in turn composed in this way:

$$V_{Bonded} = V_{Bonds} + V_{Angles} + V_{Dihedrals}$$

$$V_{unbonded} = V_{van\ der\ Waals} + V_{Electrostatic}$$

EQUATION 2 | Equation of potential energy function's components.

According to the characteristics of the system and the simulations to be carried out, the components can be modeled in different ways.

3.2.2 Treatment of bond and non-bond interactions

First, it is important to define the potential energy function as:

$$\begin{aligned}
 V(r_N) = & \sum_{\text{Bonds}} \frac{1}{2} k_l [l - l_0]^2 + \sum_{\text{Angles}} \frac{1}{2} k_\theta [\theta - \theta_0]^2 \\
 & + \sum_{\text{Dihedrals}} k_\varphi [1 + \cos(n\varphi + \delta)] \\
 & + \sum_{i=1}^N \sum_{j=i+1}^N 4\varepsilon_{i,j} \left[\left(\frac{\sigma_{i,j}^{12}}{r_{i,j}} \right) - \left(\frac{\sigma_{i,j}^6}{r_{i,j}} \right) \right] + \frac{Q_i Q_j}{4\pi\varepsilon_0 \varepsilon_r r_{i,j}}
 \end{aligned}$$

EQUATION 3 | Explanation of the components of the term bond of potential energy.

This is composed of bond components, angles, dihedral and non-bond positions.

The first component concerning the bonded components is modeled as a harmonic potential and therefore increases as the bond length increases. The stiffness k is none other than the stiffness of the bond between the atoms under consideration.

The second component instead describes the angles that are created by the rotation between the atoms that have a covalent bond. The third component concerns the interactions of the dihedral bond and includes, unlike the first bond where a stiffness value k was present, a torsional potential for the description of the rotation.

In this case $k\varphi$ is defined as the energy barrier of deformation of the angle, δ is the phase, n is the multiplicity. The last component is precisely the one concerning non-bond interactions.

The latter component is modeled with Van der Waals forces and electrostatic interactions.

The Van der Waals potential is the sum of the attractive and repulsive forces within a molecule and between two entities with neutral charge. These forces are generated by the phenomena described by the Pauli exclusion principle and by the phenomena defined as London dispersion. Clearly Van der Waals forces have a weaker bond strength than covalently bonded atoms, yet they are fundamental to the description of molecular properties. In particular, these describe the solubility of the compounds within polar and non-polar solutions, stericity and molecular conformation.

For the definition of this potential, the Lennard-Jones equation is used, which describes the Van Der Waal forces of repulsion and attraction based on the distance between atoms:

$$V_{Lennard-Jones} = 4\varepsilon_{i,j} \left[\left(\frac{\sigma_{i,j}^{12}}{r_{i,j}} \right) - \left(\frac{\sigma_{i,j}^6}{r_{i,j}} \right) \right]$$

EQUATION 4 | Lennard-Jones equation.

The repulsion forces are described by the first component while the attractive ones by the second. $\sigma_{i,j}$ describes the distance limit after which the Van der Waals interactions are zero, $\varepsilon_{i,j}$ describes the depth, i.e. the energy minimum.

instead, for the description of the electrostatic interactions, Coulomb's law is used:

$$V_E = \frac{Q_i Q_j}{4\pi \varepsilon_0 \varepsilon_r r_{i,j}}$$

EQUATION 5 | Coulomb's law for electrostatic interactions.

In this equation ε_0 is the permittivity in the vacuum while ε_r is the permittivity in the medium.

This type of voltage contribution is inversely proportional to the distance between the atoms involved and is therefore defined as long range. To reduce computational costs, the energy term can be calculated only if the distance between them reaches a certain value called the cut-off radius. Another method is the Ewald sum where from real space we pass to Fourier space.

3.2.3 Periodic Boundary Condition

Setting boundary effects within the simulations is very important because they have the ability to influence the behaviour of the entire system under analysis.

These conditions (PBC) can be applied to a system of infinite dimensions allowing to reduce the edge effects.

To remove these on-board effects, the system being analysed is usually placed in a box containing solvents (water or other types) and the atoms are surrounded by a copy of them.

Clearly these conditions alter the result of the simulations, but these errors are in any case negligible if compared to the errors deriving from the presence of an artificial boundary in contact with the vacuum.

The main problem is that the particles inside the box, if periodic boundary conditions are present, then they interact with each other and with the box adjacent to infinity. To avoid this, various methods can be applied such as the Lennard-Jones distance limit, Particle Mesh Ewalds^[32], Multipole Cells^[33] and Reaction Fields^[34].

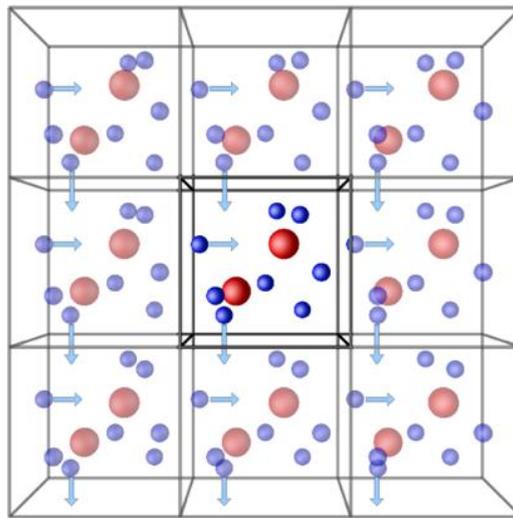


FIGURE 7 | Periodic boundary conditions. An example of a box with a system of particles inside that is replicated identical to itself.

3.5.6 Potential Energy Minimization

The potential energy function is defined on the basis of the dimensional characteristics of the system, in fact for a system composed of a number N of atoms it is necessary to define $3N$ Cartesian coordinates and $3N-6$ are the internal coordinates involved in the system, in particular bonds, angles and dihedrals .

When a system is in the state of equilibrium, we speak of stability and the potential energy function reaches a minimum defined Potential Energy Surface (PES), any alteration of the system will lead to a greater value of energy. The minima can be local (there can be many) and global (for each system there is only one global minimum).

When we talk about energy minimization, we mean to find the minimum energy point.

There are several methods for energy minimization, which can be divided into two macro classes: derivative and non-derivative. The derivative methods are divided into first and second order. In the first order through the direction of the gradient it is possible to identify the position of the minimum, an example is Steepest Descent [35], Conjugate Gradient [36]. For second order methods, based on the coordinates it is possible to understand where the energy changes direction. To do this, it is necessary to calculate the inverse Hessian matrix of each second derivative, which is why these methods are considered to be of high computational cost, examples are Newton-Raphson [37], L-BFGS [38].

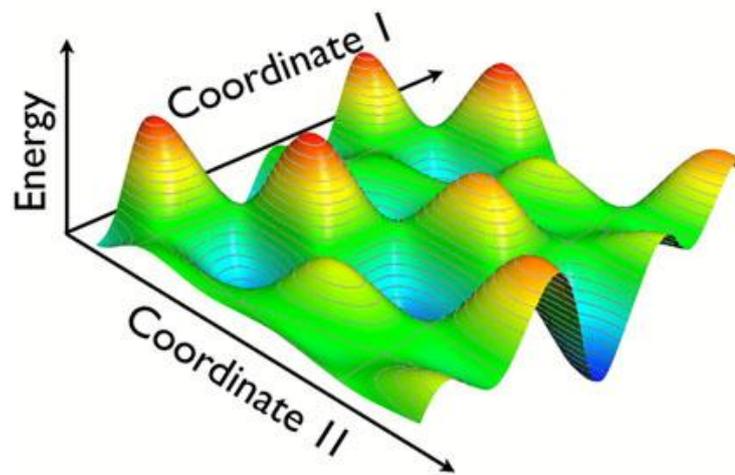


FIGURE 8 | Potential Energy Surface (PES). In the graph there are a lot of local minimum points and a global minimum point.

3.3 Molecular Dynamics

For the analysis of the dynamics of complex systems such as proteins, nucleic acids etc., Molecular Dynamics (MD) is used. This technique, by solving the Newtonian equation, allows to analyse average properties of the system under analysis.

Remembering the dynamic equation:

$$m_i \frac{\partial^2 r_i}{\partial t^2} = F_i$$

EQUATION 6 | Dynamic Newtonian equation.

in which m_i is the total mass of the system, r the position and F_i the sum of the forces applied to the system.

The force acting can also be expressed using the potential energy as follows:

$$F_i = - \frac{\partial U(r_i, \dots, r_N)}{\partial r_i}$$

EQUATION 7 | Equation of Force expressed with potential energy.

3.3.1 Statistical Ensemble

A chemical system can be investigated by means of static balance properties and non-equilibrium properties. The phase space is the representation of all the states that the system can assume in terms of position and momentum. This implies that for the definition of a point in the space of states, the 3 coordinates which represent the position and the three coordinates which represent the momentum are necessary. A point in the state space is defined as the point that represents the current state in which the system is present, all other points are the possible states that the system could assume over time. Ensembles containing points can be defined on the basis of thermodynamic characteristics. Ensembles can be defined as follows:

- The Micro-Canonical Ensemble (NVE): constant number of particles of the system (N), volume (V) and energy (E). It can be defined as an isolated system.
- The Canonical Ensemble (NVT): system particle number (N), volume (V) and constant temperature (T). It can be defined as a closed system.
- The Grand Canonical Ensemble (μ VT): constant chemical potential (μ), volume (V) and temperature (T). It can be called an open system as there is an exchange of matter.
- The isobaric-isothermal set (NPT): number of particles in the system (N), constant pressure (P) and temperature (T).

With molecular dynamics it is possible to sample the entire phase space and calculate the aggregate averages of the various properties of the overall system.

Mediating the properties means representing them as A function of position and moments:

$$\langle A \rangle_{Ensemble} = \iint dp^N dr^N A(p^N, r^N) \rho(p^N, r^N)$$

EQUATION 8 | Function A expressed by moments and position.

In the equation r is the position of the various atoms in the system, p is the momentum and $\rho(p^N, r^N)$ is the probability density function. If we wanted to represent the probability density within the ensemble (NVT), this can be expressed with the Boltzmann distribution:

$$\rho(p^N, r^N) = \frac{1}{Q} \exp[-H(p^N, r^N)/k_b T]$$

EQUATION 9 | Probability density expressed by Boltzmann distribution.

where k_b is the Boltzmann factor, H the Hamiltonian and T the temperature. Regarding the term Q, this is defined as the partition function, it can also be written as a function of the Hamiltonian H as follows:

$$Q = \iint dp^N dr^N \exp[-H(p^N, r^N)/k_b T]$$

EQUATION 10 | Partition function expressed by Hamiltonian H.

This partition function acts as a bridge between microstate and macrostate, in fact it can be defined as the sum of the Boltzmann factor on all the microstates of the system under analysis. Clearly, from the point of view of the solution to this equation, it is impossible to calculate why an ergodic hypothesis has been introduced. This hypothesis is based on the concept that the temporal average for each property can be replaced with the average of the set of the same property over a fairly long period of time. The formula for calculating this average is as follows:

$$\langle A \rangle_{Time} = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_{t=0}^{\tau} A(p^N(t), r^N(t)) dt \cong \frac{1}{M} \sum_{t=1}^M A(p^N, r^N)$$

EQUATION 11 | Average function of Equation 8

where t represents the length of the simulation in terms of time, M is the number of steps of the simulation and $A(p^N, r^N)$ is the instantaneous value of property A .

3.3.2 Molecular Dynamics implementation scheme

The main concept on which molecular dynamics is based is that of solving Newton's equations for the analysis of a complex system. Newton's law is reported for the description of acceleration:

$$a = -\frac{1}{m} \frac{dV}{dr}$$

EQUATION 12 | Acceleration equation expressed by Newton's law.

For functions for which there is no analytical solution, numerical integration is used by means of some methods: the Verlet^[39] algorithm, the Leap-Frog^[40] algorithm, Velocity Verlet^[41] and many others. The state space must be sampled, it is important to choose a correct time step (usually fs).

The flowchart (Figure 9) describes the steps to follow in molecular dynamics:

- The initial structure is provided as input and the initial positions are also given by notes according to the PDB structure used (an example is to download the structure from the Protein Data Bank).
- The velocity values, on the other hand, are assigned arbitrarily by a Maxwell-Boltzmann distribution at a given temperature. With these inputs the first calculation made is the potential energy.
- The potential energy is then derived to calculate the forces acting on the system. At each step a new set of coordinates and speed is generated.
- Finally, the simulation output trajectory is extracted from which the other properties of the analysed system can be calculated.

Using the MD exit trajectory, macroscopic thermodynamic properties (e.g., energy, temperature, pressure) can be calculated as time averages.

3.3.3 Software package

For molecular dynamics, different AMBER^[42], CHARMM^[43], GROMACS^[44], etc. are available. For this master's thesis the only software used was AMBER.

The acronym AMBER stands for Assisted Model Building and Energy Refinement and is software for MD simulations on complex systems.

- Leap: defines the coordinates, the different parameters and topology of the system used at the start of the simulation for setting the parameters.
- Anteroom: Used to automate Amber Force Field compatible parameterization.
- Sander: is the core of the development of simulations.
- Pbsa and mmpbsa: calculate energy.

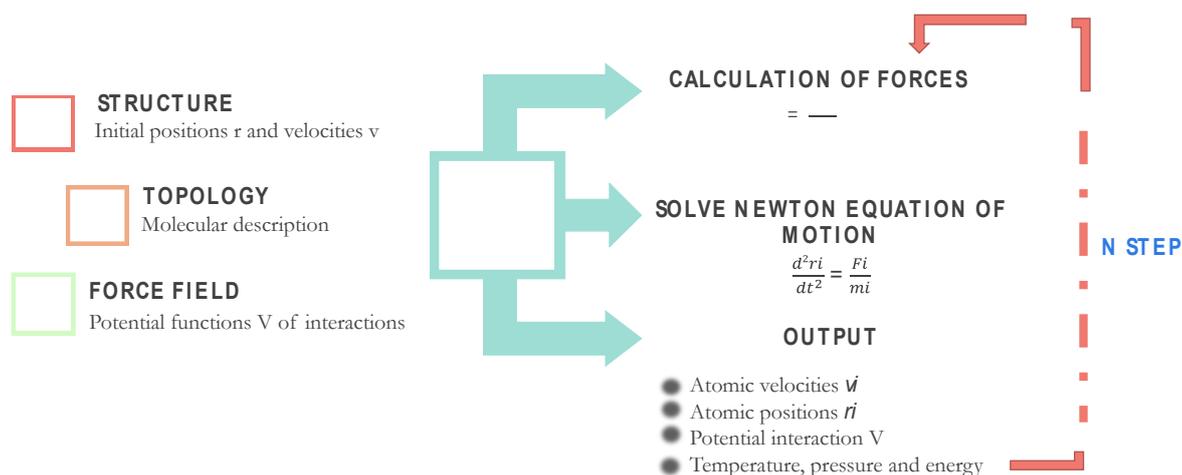


FIGURE 9 | Flowchart of Molecular Dynamic (MD) steps for the correct simulation setup.

3.4 Methods of analysis

After the molecular dynamic simulations, numerous analyses are carried out.

In this work AMBER was used and different analyses were carried out by means of the final trajectory of the simulations.

3.4.1 Cluster analysis

The molecular dynamics exit trajectory contains a large number of frames which are used for the analysis of the system properties.

Grouping methods are used and in particular it was used the kmeans clustering algorithm that minimizes the variances within the cluster (Euclidean distances squared).

This method divides, according to the characteristics set, the data or conformations into different macro-groups by extracting the reference centroid for each group (cluster).

Each point within the group is similar to the other points contained in the same cluster and has differences compared to the other points belonging to different clusters. The software used in this work is Amber and CPPTRAJ was used as a tool.

3.4.2 Distance analyses

Distance analyzes are performed to calculate differences between atoms or molecules. In this context, this analysis was used to evaluate the conformational differences that the protein presented after the simulation of molecular dynamics. For these analyzes groups or single atoms are chosen and the distance between them is calculated from the remaining. Amber CPPTRAJ tool was used.

3.5 Molecular Docking and Virtual Screening

Molecular Docking is a computational method used for the discovery of new drugs and to model the interaction between proteins and new compounds analyzed.

This is allowed thanks to the prediction of the binding site of the ligand to the protein and also to the prediction of the conformation that the ligand assumes.

Thus, docking allows not only to evaluate the poses of the various compounds with the protein but also to estimate a binding affinity by calculating the receptor-binding energy.

Virtual screening instead has the purpose of skimming the analyzed molecules by classifying them as active and inactive.

3.5.1 Ligand and receptor

For Docking analysis, the main structures are ligands and receptors. The receptors are in most cases proteins, or 3D structures with atomistic resolution.

There is a database containing many ready-made crystallographic structures that can be downloaded for free (Protein Data Bank ^[45]). Clearly better is the resolution,

better will be the results, usually less than 2.5 angstrom. Not all structures are present within the databank, the structure is very often present, but the organism of interest is different. Bacteria and non-human calcium channels may be present. For this reason, homology modeling is very often used, that is the modeling of homology ^[46], this allows to obtain a new model starting from a structure already present and already validated but going to compare the amino acid sequences. Before docking, the structure is usually prepared by not only balancing it but also by carrying out molecular dynamics to evaluate the flexibility of the chains and the binding site. The most difficult step from the point of view of the investigation is to understand which is the binding site between the ligand and the receptor called binding pocket. As for the ligands, they too are already present in various databanks, in particular PubChem^[47]. Clearly, before docking the structure must be optimized, if the information is in 2D the 3D structure must be created, controlling the atoms, the charge and the protonation. Not necessarily all the structures are present in the databank, they can be designed using software such as MOE ^[48] or Avogadro^[49], clearly going to optimize after the structure.

3.5.2 *MaxMinPicker algorithm*

The MaxMinPicker algorithm is an algorithm that allows you to skim the data and divide it into large groups. this allows you to work even with large data sets.

The steps of the algorithm are as follows:

1. For each molecule belonging to the dataset to be analyzed, whether they are seed or not, create a descriptor.
2. If there are no starting molecules (seeds), it randomly chooses molecules to be analyzed.
3. Of the molecules identified for the analysis, look for the molecule that has the maximum value for its minimum distance from the molecules of the analysis set.
4. It records all the distances between the molecules and when it finds the largest distance, it transfers the molecule to the collection set.

The figure 10 schematically describes the steps of the algorithm. The table should be read from left to right and for each column. You take the first column and read all the way to the bottom. From the analysis of the first column, the first maximum value is obtained.

For each other column proceed in the same way but comparing the maximum value found with the previous one.

If the new value is in the range of the previous one, this indicates that the processed molecule is similar to that of the previous column. This is done for each column. Clearly, if at the end of a column the minimum value is greater than the minimum value of the following columns, then the molecule representing that column automatically becomes the reference molecule as it has a greater distance from the origin. it should be noted that each distance is calculated once and recorded, there are no repetitions.

		Candidate Pool									
		0	1	2	3	4	5	6	7	8	9
Picks	0	3	1	4	1	5	9	2	6	5	3
	1	5	8	9	7	9	3	2	3	8	4
	2	6	2	6	4	3	3	8	3	2	7
	3	9	5	0	2	8	8	4	1	9	7
< Bounds		3	1	0	1	3	3	2	3	2	3
Maximum		3									

FIGURE 10 | Scheme for understanding how the min-max picker algorithm works.

3.5.3 Search Algorithms

During molecular docking the aim is to find the best pose, this implies that many conformations of the system must be explored. the best pose is the pose that shows the highest possible bond energy in absolute value.

There are two possible ways from: the first is to carry out a local search, the is the exploration of a global minimum and therefore the space is explored more widely. Depending on the type of program used to perform the Docking analysis, it is possible to implement different methods that use matching algorithms, systematic search or stochastic methods.

The first methods combine the structure of the compound to be analysed and the chemical complementarity of the site in which the bond is expected.

The compounds used in this type of docking must be known and the coordinates that identify the location of the ligand can be calculated with can be determined with mean square deviation (RMSD) by taking the lowest compared to the centers of the characteristics.

Systematic search methods, on the other hand, create a sampling of the conformation space in well-established intervals. this allows to investigate all possible degrees of freedom. These methods can be classified into exhaustive research and fragmentation algorithms. The exhaustive research aims to investigate all the possible conformations that the system can assume by creating rotations between the bonds.

The number of conformations is calculated according to the following equation:

$$N_{conform} = \prod_{i=1}^N \prod_{j=1}^{n=inc} \left(\frac{360}{\theta_{ij}} \right)$$

EQUATION 13 | Number of conformations in a system.

where N is the number of bonds that allow rotation and θ_{ij} is the amplitude of the incremental rotation angle j for the bond i.

The last method, namely that of fragmentation, involves the fragmentation of the ligand and the sampling of the conformational space by increasing one fragment at a time within the binding site. the fragments are covalently connected which allows to decrease the conformation energy.

The advantage of stochastic methods is certainly the rapid achievement of an almost optimal solution by randomly varying the degrees of freedom of the analysed compound. Clearly, they are not always suitable in fact they are mostly used for large molecules where the degrees of freedom are high in number.

Examples of algorithms that used these methods are genetic algorithms (GA) ^[50], simulated annealing (SA) ^[51] and Monte Carlo ^[52] (MC). This type of algorithms are genetic algorithms and reproduce the process of evolution. In this process, functions called fitness are used which decide which configuration, defined as the representation of the individual, is able to survive and therefore automatically generate the next configuration. The surviving generation will then be used later for the iteration optimization steps.

The idea is to use these algorithms for docking analysis, to do this you need to make assumptions that are:

1. Each ligand configuration representing the individual has a chromosome and its associated genes. these identify the state variables.
2. The fitness function no longer represents the surviving individuals but the best configuration in terms of state space.
3. The condition that allows the conclusion of the research can be to reach convergence or the achievement of a maximum term of iterations, or the achievement of a chosen fitness value.

The Monte Carlo method performs random changes in the configuration of the compound involved in docking. The interesting thing is that the method takes into account the already sampled states and this allows for no repetitions.

Usually this is used simultaneously with simulated annealing methods: these also allow you to create random changes, but they are reviewed and based on a criterion defined Metropolis Criterion, it is decided whether these changes are accepted or not.

The criterion is structured as follows:

1. It starts from a configuration with a starting energy E_1 .
2. A rendering change is generated and is accepted only if, with this change, the new energy is less than the starting one or with a probability equal to $P = \exp \{- [(E_2 - E_1) / k_b T]\}$ where in the present equation T is the temperature and k_b is the Boltzmann constant.

The temperature does not remain constant on the ground for the duration of the simulation but is reduced in such a way as to be able to lower the probability and therefore accept conformations even at high energy.

The use of this criterion is functional to overcome the problem of accepting the local minimum, and therefore the acceptance of a conformation with higher and therefore worse binding energy.

3.4.3 *Binding Energy*

The main analysis carried out in molecular docking and virtual screening is certainly the analysis of the binding energy value and the conformation (poses) that the ligand has assumed.

The equation describing the binding energy is as follows:

$$\Delta G^0 = \Delta H^0 - T\Delta S^0 = -RT \ln K_i$$

EQUATION 14 | Binding energy equation.

The energy is evaluated under standard conditions, this is indicated by the presence of the zero coefficient, furthermore there is the enthalpy H, entropy S and the temperature T, as well as the constant of the gases R. Clearly, the presence of S and h indicates that a correct evaluation of the binding energy is due to the evaluation of both entropy and enthalpy.

The enthalpy term represents the electrostatic forces, of Van der Walls (VdW), the hydrogen bonds, polar ^[53].

The entropic term instead represents the flexibility of the compound. Molecular Mechanics Generalized Born / (Poisson Boltzmann) Surface Area model (MMGB / (PB) SA) ^[54] are models used to evaluate the energy score due to docking by finding a compromise between accuracy and analysis time.

This method evaluates the differences between two states of the same molecule in the bound and unbound state which is why it can only be used for a post-processing of some algorithms such as molecular dynamics. All the frames that have been generated by the dynamics are in fact used for the calculation of the energy between the two states (bound and unbound). during the simulations we will also have solvent-solvent interactions that will certainly show very slight fluctuations.

This may mask the binding energy by thinking that no engagement has occurred. to avoid this, the energy calculation is done according to a thermodynamic cycle.

According to this cycle, the binding energy can be evaluated using the following equations:

$$\Delta G_{bind,solv}^0 = \Delta G_{bind,vacuum}^0 + \Delta G_{solv,complex}^0 - \Delta G_{solv,ligand}^0 - \Delta G_{solv,receptor}^0$$

EQUATION 15 | Equation of binding Energy.

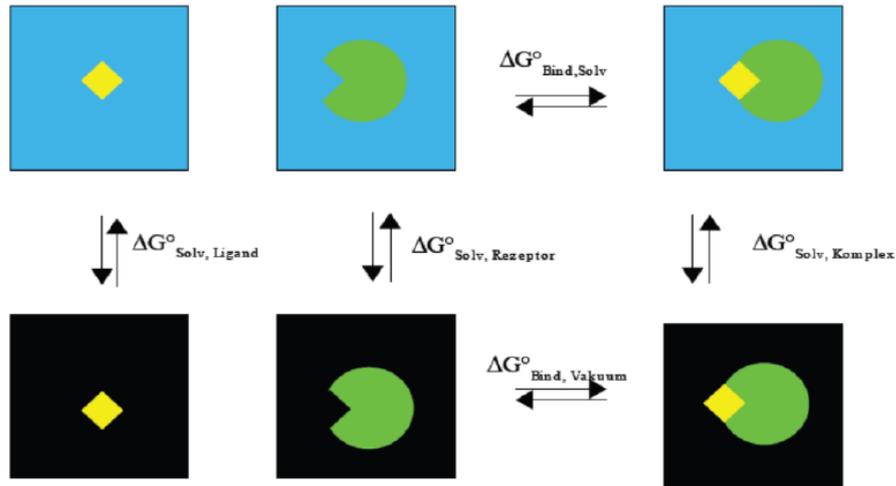


FIGURE 11 | Thermodynamic cycle for the calculation of the binding energy used in the MMGB / PBSA method.

Through this thermodynamic cycle, the final result of the binding energy is the composition of two terms: a contribution of solvation and a contribution in the gas phase. Free energy can be expressed according to this equation:

$$\Delta G_{vacuum} = \langle E_{MM} \rangle - T\Delta S$$

EQUATION 16 | Free Energy equation.

where E_{MM} represents molecular mechanics and takes into account the Van der Waals components, the electrostatic energy and bond, the angular, torsional component. The second is an entropic term and therefore represents the internal entropy of the system. Usually this second term, having a high computational weight, is neglected^[55] or is calculated with low quality approximations. The free energy of solvation includes two terms that describe: the first the electrostatic / polar interactions, the second the non-polar ones. Polar interactions can be described by the solution of the Poisson-Boltzmann equation (PB) or also by the generalized Born equation (GB).

Using the latter to represent the electrostatic contribution, the following equation can be used:

$$\Delta G_{polar} = -\frac{1}{2} \sum_{i,j} \frac{q_i q_j}{f(r_{ij} a_{ij})} \left(1 - \frac{1}{\epsilon}\right)$$

EQUATION 17 | Polar contribution of free Energy expressed by electrostatic contribution.

In the equation $\epsilon^{[56]}$ is the dielectric constant in the solvent while the function f depends on two factors, namely interparticle distances and Born radii.

As regards the non-polar contribution, this is described by the following equation, which indicates that this contribution is proportional to the area accessible to the solvent:

$$\Delta G_{non-polar} = \gamma SASA(x)$$

EQUATION 18 | Non Polar contribution of free Energy expressed by SASA.

3.4.4 *Software package*

There are numerous software that can be used for molecular docking. The most used are AutoDock ^[57], AutoDock Vina ^[58], DOCK ^[59], LigandScout ^[60]. In this master's thesis, the Vina software was used for the first round and then AutoDock. The choice was dictated by the fact that Vina is much faster than Autodock and this initially allowed for a faster analysis.

CHAPTER IV – Computer-aided drug design of Ca²⁺ agonists to enhance the effect of anti-CD20 monoclonal antibodies for CD20-positive non-Hodgkin's lymphoma and chronic lymphocytic leukemia.

In this chapter the study of molecular dynamics and molecular docking on calcium channel agonists will be described. In particular, it will be researched new compounds with similar characteristics in terms of binding energies, binding sites and drug specificity.

Abstract

Non-Hodgkin's lymphomas (NHL) are clonal lymphoproliferative diseases that originate from B lymphocytes. NHL also comprises the Chronic Lymphatic Leukemia (CLL), a type of blood cancer. Patients affected by these pathologies possess high amounts of differential antigens CD20. In this context, anti-CD20 monoclonal antibodies (mAbs) increase the influx of Ca²⁺ and lead to cell apoptosis after interacting with CD20. Anti-CD20 mAbs bind to CD20 transmembrane antigen, they were found on pre-B lymphocytes and mature B lymphocytes. The CD20 receptor is involved in the regulation of Ca²⁺ influx in B lymphocytes and the cells exposure to anti-CD20 antibodies causes a strong induction of cell death, suggesting a correlation between the and Ca²⁺ influx and the cytotoxic effect of the antibodies.

Therefore, it is worth interesting to characterize small molecules able to keep the calcium channels in an open form and to develop new molecules that can enhance the influx of Ca²⁺, such as agonists of L-type calcium channels (LTCC). The main hypothesis addressed in this study is improve the actions of anti-CD20 monoclonal antibodies in inducing apoptosis by increasing the flow of calcium inside the cell.

After studying the interactions of the dihydropyridine (DHP) agonist of the calcium channel BAYK 8644, the aim of the thesis work is to find a drug with the same characteristics in terms of structure and binding sites involved in the protein-binding interactions. All compounds that will be inside the binding pocket of the reference agonist will be discriminated and those with greater interactions in terms of higher binding energy will be chosen.

With these considerations, the present project will combine computer-aided-drug design (CADD), experimental results of previous studies obtained by the Anticancer Antibodies team on calcium channel blockers and the antitumor activity of anti-CD20 monoclonal antibodies and Molecular Dynamics (MD) simulations for the analysis and research of new compounds to

support the action of anti-CD20 mAbs, selectively targeting the calcium ions channel and consequently increasing the calcium flow.

The outcome of this work will show the conformational effects on the calcium channel after the agonist/antagonist binding and will perform an extensive virtual screening among a huge number of compounds, exploring the mode of binding of investigated chemicals and highlighting differences in terms of binding sites and main residues involved.

4.1 Introduction

Cancer research has always been committed to the discovery of new therapies for the treatment of cancer. The deepening of the pathologies involved allows to produce new strategies to be implemented.

Non-Hodgkin's lymphomas (NHL) affect the lymphatic system and it consists in transformation of lymphocytes creating neoplastic cells. The lymphatic system it is composed by lymphatic organs, such as the bone marrow, tonsils, thymus and spleen, and a series of lymph nodes and lymphatic vessels. The large structure of this system means that this type of lymphoma does not have well-defined areas in which it can develop and rarely affects a single organ.

Chronic Lymphatic Leukemia (CLL) is a group of blood cancer and a subtype of NHL. CLL is the consequence of numerous biological and genetic events affecting mature B lymphocytes and is characterized by an initial proliferation of these cells in the bone marrow and peripheral blood, followed by the involvement of the lymphatic structures that increase in size. The main therapeutic agents in the treatment of patients with differential antigens CD20-positive NHL and CCL are monoclonal antibodies (mAbs) [8]. They work through by complement-mediated cytolysis that provides the desired anti-cancer effect or they release toxins and radiation to selectively target affected cells. The mAbs are used for the treatment of these diseases, since monoclonal antibodies, such as anti-CD20.

Anti-CD20 mAbs bind to CD20 transmembrane antigen, a non-glycosylated phosphoprotein, which is found on pre-B lymphocytes and mature B lymphocytes.

Exposure of B-lymphoma cell lines to anti-CD20 monoclonal antibodies (mAbs) upregulates the expression of early growth response-1 (EGR-1), by increasing Ca^{2+} concentration inside the cells which trigger ERG-1 expression. EGR-1 encodes a zinc finger transcription factor whose expression is highly regulated by various extracellular stimuli, including calcium influx.

Moreover, recent results^[1] indicate the involvement of the CD20 receptor in the regulation of Ca²⁺ influx in B lymphocytes and that compounds that block calcium channels have been shown to impair the action of anti-CD20 antibodies^[10-11]. Exposure to anti-CD20 antibodies causes a strong induction of cell death, suggesting a correlation between the induction of EGR-1 and the cytotoxic effect of the antibodies. Recent results obtained by the Anticancer Antibodies team^[1] reported that inhibition of the calcium flux with Ca²⁺ channel blockers as used to treat hypertension or angina, but the patients receiving CCB and concomitant therapy anti-CD20 for diseases such as NHL or CLL have the worst survival. This show that CCBs negatively impact the outcomes of patients receiving anti-CD20 mAbs.

These data lead to hypothesize that calcium agonists can enhance and support the therapeutic effects of anti-CD20 antibodies.

Regarding calcium channel antagonists, there is a lot of information regarding the binding sites^[20-23]. **The starting point for our project will be to identify the preferential state for the correct binding of the compounds.** The first class we will analyze will be that of dihydropyridines (DHPs) which have both an agonistic and antagonistic effect. In general, agonists are hypothesized to increase calcium flux within the channel because they prolong the opening time, a difference of antagonists who prefer the closed inactivated state. Some Cav agonists are known from literatures: FPL64176^[24], and CGP 485065, murrayafoline A^[25] at low concentration.

Computational approaches are used for the analysis of complex allowing to study in detail the molecular mechanisms, moreover the docking is a powerful means to study the position that a compound assumes in the binding with the protein and to evaluate its affinity in terms of binding energy. In this study, Molecular Dynamics (MD) simulation, Ensemble Docking and the calculation of binding free energy will be used to perform a skimming of the compounds chosen for the analysis and then guide the research towards a certain type of compound for the search of new medications.

4.2 *Materials and Methods*

4.2.1 *Structure preparation of calcium channel embedded in the lipid bilayer*

The most recent structures of the calcium channel have been downloaded from the RCSB database: 6JP8 Rabbit Cav1.1-Bay K8644 Complex (2.70 Å), 6JP5 Rabbit Cav1.1- Nifedipine Complex (2.90 Å), 6JPA Rabbit Cav1.1-Verapamil Complex (2.60 Å), 6JPB Rabbit Cav1.1-Diltiazem Complex (2.90 Å)^[13]. The initial crystal

structures contained γ -1, β -1, α -1S, α -2 / δ -1 subunits and agonist and antagonist ligands. Considering that our interest is focused on the center of the channel and therefore on the α -1 subunit, the other subunits were not considering in the final model (Figure 12).

Structures were prepared by aligning the calcium channel to its corresponding Orientation of Protein in Membranes (OPM) structure^[62] in order to have the pore oriented along the z-axis and the center of mass of the protein at the center of the structures. The channel was embedded into a DOPE (1,2-Dioleoyl-sn-glycero-3-phosphoethanolamine) lipid bilayer using the membrane generator CHARMM-GUI webserver¹¹. A pore was generated using a cylindrical radius of 12 Å to allow the insertion of water inside the pore.

4.2.2 *Structure preparation of ligands*

Ligands for the initial study of the binding sites (Nifedipine, Verapamil, Bay K8644, Diltiazem) were extracted from the respective crystal structures (PDB ID: 6JP5,6JPA, 6JP8,6JPB). By means of the MOE 2018 software all the extracted ligands were protonated with a Ph 7.40, using the PropKa algorithm.

4.2.3 *MD simulation protocol*

All MD simulations were performed using the Amber16 package^[64]. The AMBER-ff99SB-ILDN force-field to define protein topology, while ligands (Bay-K8644) and Nifedipine) were defined using Antechamber program of the AMBER16 package, with general amber force field.

Using Amber's Leap module, an octahedral box was used to place the system inside and filled with TIP3P waters requiring a distance of at least 15 Å between any peptide atom and the edges of the box.

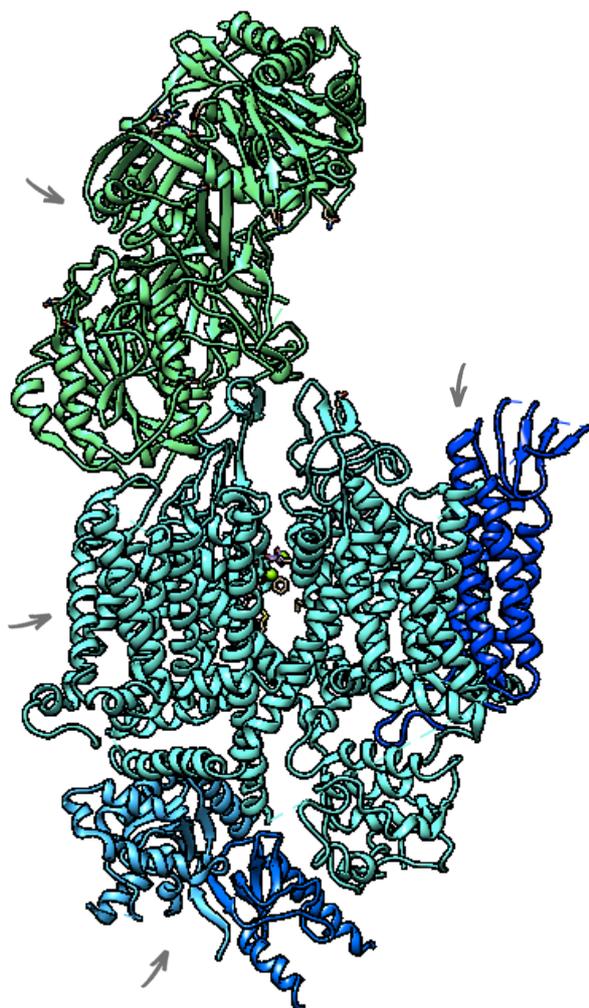


FIGURE 12| Initial Crystal structure (RCSB ID 6JP5) with all calcium channel subunits used for the creation of the model.

To neutralize the system while maintaining a concentration of 0.01 M, Na^+ and Cl^- ions were added in a manner consistent with the experimental conditions used for a measurements of NMR chemical shifts^[65].

After solvation of the structure, equilibrium was performed with Amber's pmemd tool, first minimizing the system (peptide and solvent) during 5000 steps using the steepest descent method followed by 5000 steps of conjugate gradient.

Next, simulations were performed for each structure in NVT (Canonical ensemble) simulations for 500 ps using a speed rescaling thermostat with the tau constant at 2 ps and a reference temperature at 298 K and restraining heavy atoms of the peptide. NPT (Isothermal-Isobaric) simulations was done at $P = 1$ bar and $T = 298\text{K}$ for a total time of 500 ps and was used the Berendsen barostat to maintain the pressure constant. These latest NPT simulations were run without restraints in order to support the equilibration process.

Finally, a MD simulation was done for 50 ns without any restrain with a time step of 2 fs and coordinates saved at every 2 ps.

The software MOE 2018 and Chimera^[66] were used for the visualization of the systems, while for the quantitative analyzes, i.e. Root-Mean-Square Deviation (RMSD) Root-Mean-Square Fluctuation (RMSF) and Cluster, CPPTRAJ was used.

4.2.4 Virtual Screening

4.2.4.1 ZINC-ID compounds

Compounds to perform the virtual screening were chosen and downloaded from the ZINC15 database^[67]. The most physiological 3D structures were selected, choosing a lead-like predefined subset with a lipophilicity (LogP) in the range -1-3.5 and a molecular weight in the range 300-350 Dalton. 2.8 M of compounds were selected. Next, to further reduce the number of compounds, the min-max picker algorithm was applied.

The MinMaxPicker algorithm is an algorithm that allows you to skim the data and divide them into groups. This allows you to work even with large data sets.

The steps of the algorithm are as follows:

5. For each molecule belonging to the dataset to be analyzed create a structural interatomic distance-based descriptor.
6. The algorithm randomly chooses one molecule to be analyzed.
7. The chosen molecule is compared to all the other molecules in the dataset by a structural similarity measure, the similarity index, based on descriptors defined in step 1.
8. Minimum similarity indices between the chosen molecule and all the other chemicals are computed, and the compound characterized by the maximum value is selected for the subsequent step.
9. The process restarts from step 3 until all the chemicals are analyzed.
10. When the algorithm stops, ligands are divided into separate groups based on their similarity index.

Applying the MinMaxPicker algorithm, the most similar group to BAY-K8644 is selected for further analysis, leading to a total of 96.912 compounds.

4.2.4.2 *Molecular docking setup*

Docking was performed on only two structures: 6JP8 (agonist-bound calcium channel) and 6JP5 (antagonist-bound calcium channel).

A cluster analysis was conducted from the final trajectory of the MD simulations, keeping the centroid of each obtaining cluster and resulting in 10 representative structures as starting target for the molecular docking. Two rounds of docking were performed for the two structures, the first using VINA, the second using AUTODOCK. Using the AUTODOCK software^[68], the box containing only the domains interested in binding was built to detect the correct coordinates of center of mass (Figure 13). The size of the cubic search box for the docking was set to 25.0x 25.0x25.0 nm. Particular attention was paid to excluding the pore of the channel since was not of interest.

Among the ten ligand poses, only the best one in terms of binding energy was selected. To verify if the ligands were more specific for agonist or antagonist binding sites, a specificity coefficient was calculated. This calculation was done by evaluating the VINA scores obtained at the end of the docking simulations. These values represented affinity of the ligand in term of binding energies.

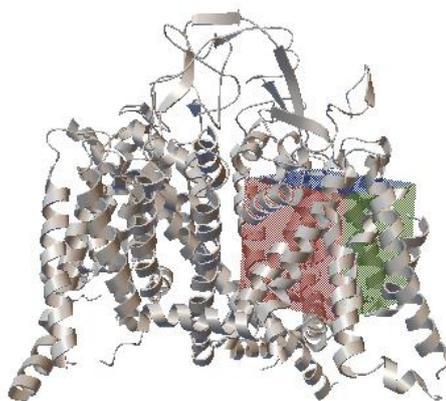


FIGURE 13 | Graphical representation of the box construction for docking analysis.

For each ligand the coefficient was evaluated as follows:

$$SPECIFICITY = \frac{vina_{score}Agonist}{vina_{score}Antagonist}$$

From this equation the ligands that had a value less than or equal to 1 are less specific for calcium channel bound to agonist (6JP8), whereas values greater than 1 indicate compounds with higher specificity. The ligands with a specificity value higher than 1.2 and a binding energy value (vina score) greater than or equal to -9.3 Kcal/mol were chosen.

4.3 Results

4.3.1 Analysis of ligand interaction

The positions of the three groups of calcium channel antagonists and the agonist were visually analyzed after downloading the crystal structures from the Protein Data Bank-PDB (RCSB) and superimposed on the 6JPA structure downloaded from the Orientation of Protein in Membranes (OPM) database.

Bay K8644 (DHP agonist) and nifedipine (DHP antagonist) were found to be located in the same site and not in the pore, whereas Diltiazem (benzothiazepine) and Verapamil (phenylalkylamine) occupy the part of the pore (Figure 14).

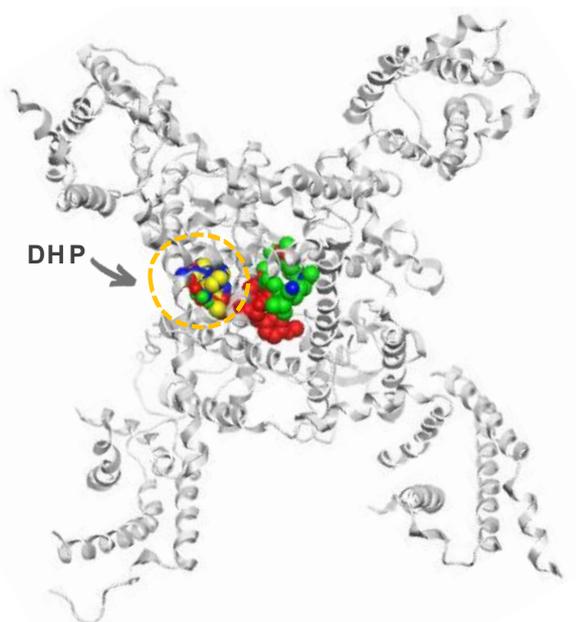


FIGURE 14| Calcium channel structure containing all ligands, obtained by superimposing all structures after MD (Red Diltiazem, yellow nifedipine, verapamil green, Bay blue K8644).

We focused on the binding sites of DHP, characterizing the binding features. It is worth mentioning that residues involved in binding-protein interaction are mainly present on domains III and IV and the segments involved are S5III, P1III, S6III and S6IV. This is consistent with the data from previous work^[16].

The different interactions in terms of electrostatic and hydrophobic interactions were analyzed. DHP antagonists showed hydrogen bonds between polar residues on S5III and P1III and the amine N1 and the C3 ester of the phenyl ring.

In particular, the amine N1 is bound in a hydrophilic way and therefore with a hydrogen bond to the residues S804 while C3 is bound to the residues Thr728 and Gln732 on S5III forming hydrogen bond with O₂ groups (Figure 15). Hydrophobic interactions were also studied, finding that the nitrophenyl ring has interactions with the Val725 residue on S5III and Met850 and Phe853, on S6III.

The dihydropyridine ring creates bonds with Phe801, on P1III, and Tyr1102 and Met1101, on S6IV with methyl groups. While presenting the same residue pocket as the antagonists, a targeted analysis of the ligand interactions in terms of electrostatic and hydrophobic interactions was also carried out for the antagonists. The structures were also superimposed to analyze the conformational differences of the residues involved in the interactions.

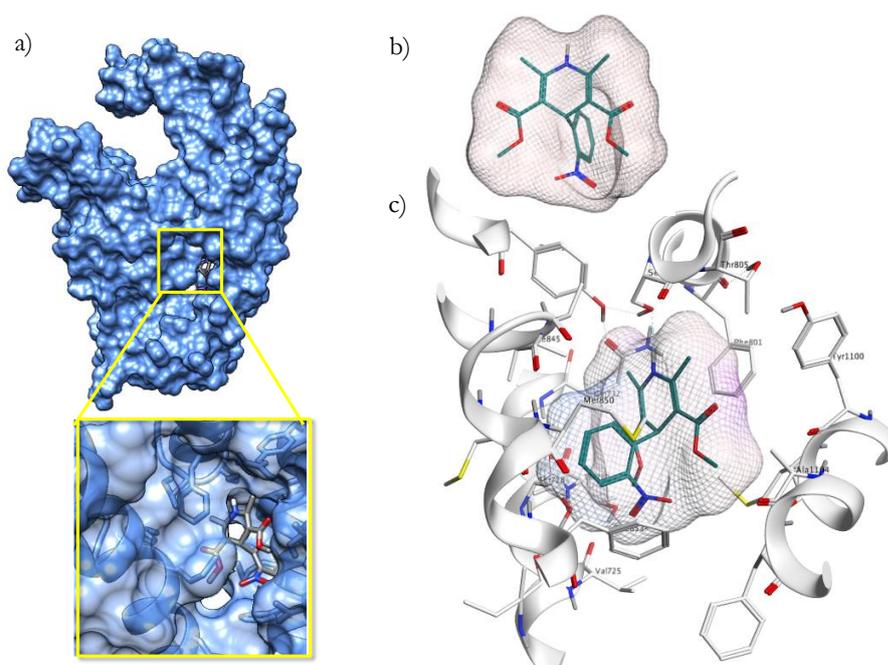


FIGURE 15| Specific calcium channel interactions and Nifedipine. a) the refilled calcium channel with its surface and a set view is done on the ligand present inside. b) EM map for the nifedipine made with MOE where the surface was made with a pink mesh. c) set view of the ligand and the binding residues.

The substantial difference lies in the dihydropyridine ring which creates bonds with Phe801, on P1III, and Tyr1101 and Met1101, on S6IV with methyl groups. The dihydropyridine ring of nifedipine and Bay K 8644 are completely superimposed while the phenyl ring is rotated towards the hydrophobic residues on S6IV (Figure 16a).

In particular, Phe1370 it appears to exist in two conformers in both structures. As regards the conformation change of the domains, it can be noted that the residue Phe1105 in the calcium channel linked with nifedipine antagonist is in the "up" position, while as regards the calcium channel linked to the agonist (Bay K8644) the same residue is in "down" position.

An analysis of the electrostatic and hydrophobic interactions of the ligands (agonist and antagonist) with the calcium channel was performed (Figure 17). For more information see the supporting information chapter.

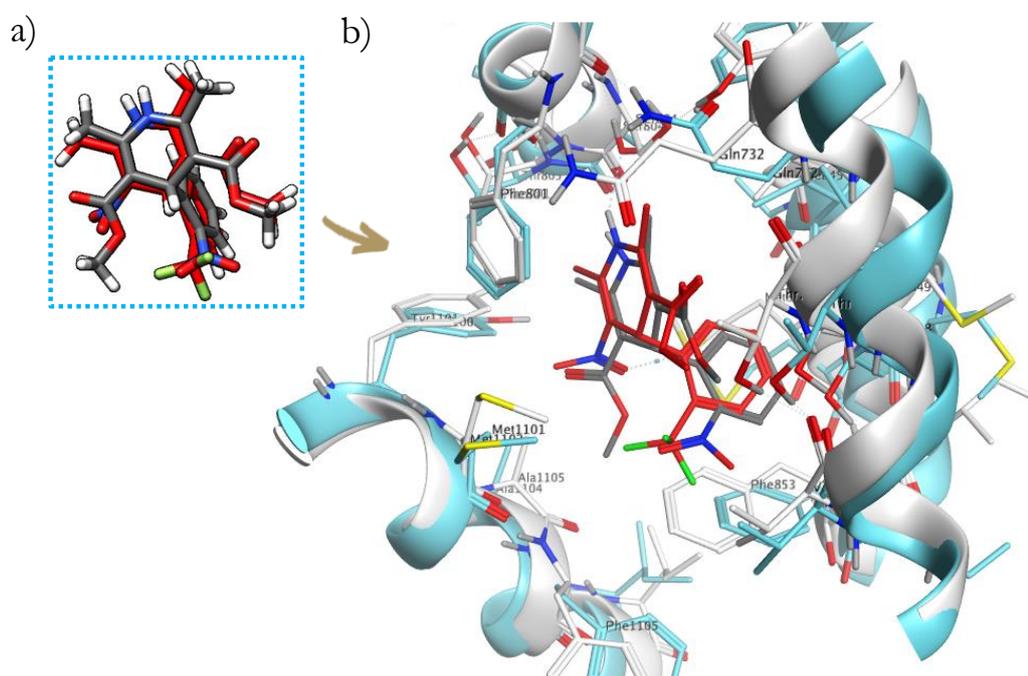


FIGURE 16 | Representations of the binding differences between calcium channel- Bay K8644 and calcium channel- Nifedipine. a) superposition of Bay K8644 red and gray Nifedipine ligands. b) analysis of the residues involved in the interactions with the ligands, overlap of the structures (white calcium channel- Bay K8644, blue calcium channel and Nifedipine).

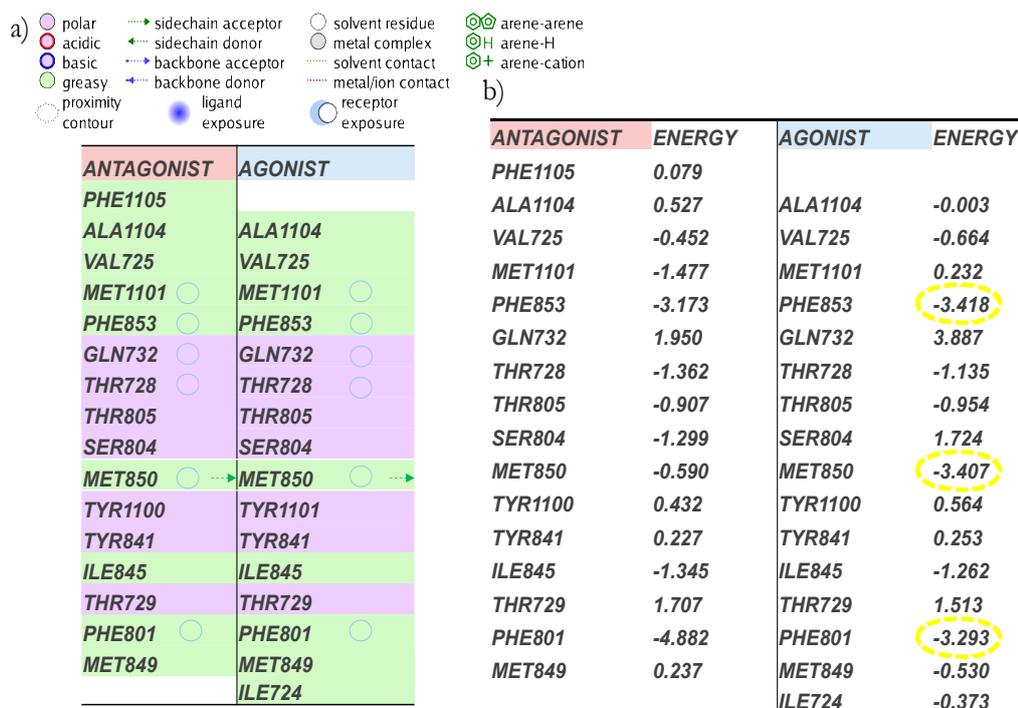


FIGURE 17 | Analysis of residues involved in the binding interaction: a) Electrostatic interaction, b) Hydrophobic interaction. Residues with higher binding energy are circled in yellow.

4.3.2 Analysis of conformational changes

The structures with agonists and antagonists were compared after the MD simulations to evaluate the conformational differences due to the interactions with the ligands.

The starting systems were represented in figure 18.

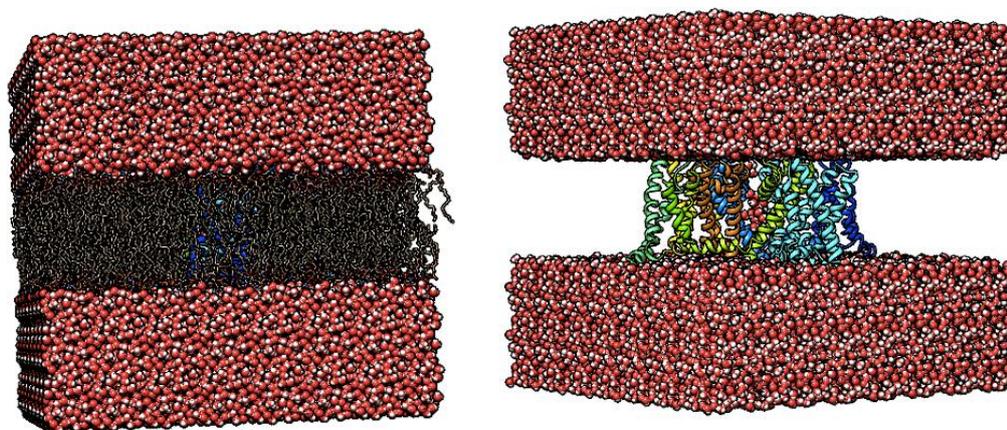
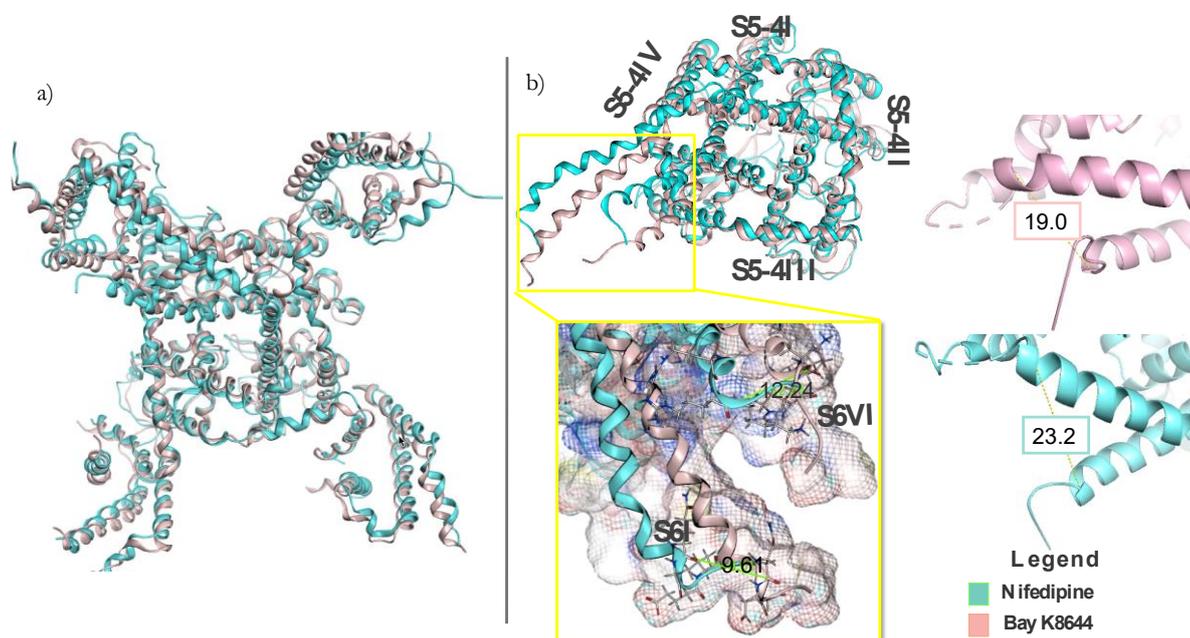


FIGURE 18 | Structure of calcium channel inserted in the membrane. Right a view of the calcium channel with the addition of water above and below the channel and inside the pore.

Human calcium channel structures are not present in the RCSB Protein Data Bank-PDB.

With the swiss-model software^[69], the similarity of the sequences of human voltage-dependent L-type calcium channel alpha-1S subunits were analyzed with the crystal structures already present, using the human FASTA sequences from the UNIPROT database^[61] Q13698 (CAC1S_HUMAN). 92.6% similarity was found with rabbit sequences, so there is a high sequence conservation with the human being and the hypothesis is that the channel mechanism is the same (Figure S1). For this reason in this project we used rabbit structures.

Structural displacements of the S4 and S5 segment domains and of the “Hinge” Loop were visualized (Figure 19). The S6I segment, instead of being in a more closed position in the presence of the antagonist, opens in the presence of the calcium channel agonist. Particular attention is paid to the S6I and S6IV segments. It can be seen that in the presence of the agonist these segments approach each other thus decreasing their distance. Details on distance measurements can be found in the supporting information chapter.



ffIGURE 19| Analysis of the conformation changes of the calcium channel structure with ligand interaction. a) superimposition of the calcium channel structure with bay k 8644 (pink) and the calcium channel with nifedipine (blue) after md simulation. b) focus on s6i and s6vi segment and distance measurement between segments.

A measurement of the distance between the aforementioned segments was carried out and evaluated over time. There is a clear decrease of the inter-domain distance in the presence of the agonist with respect to the calcium channel with the antagonist (figure 20).

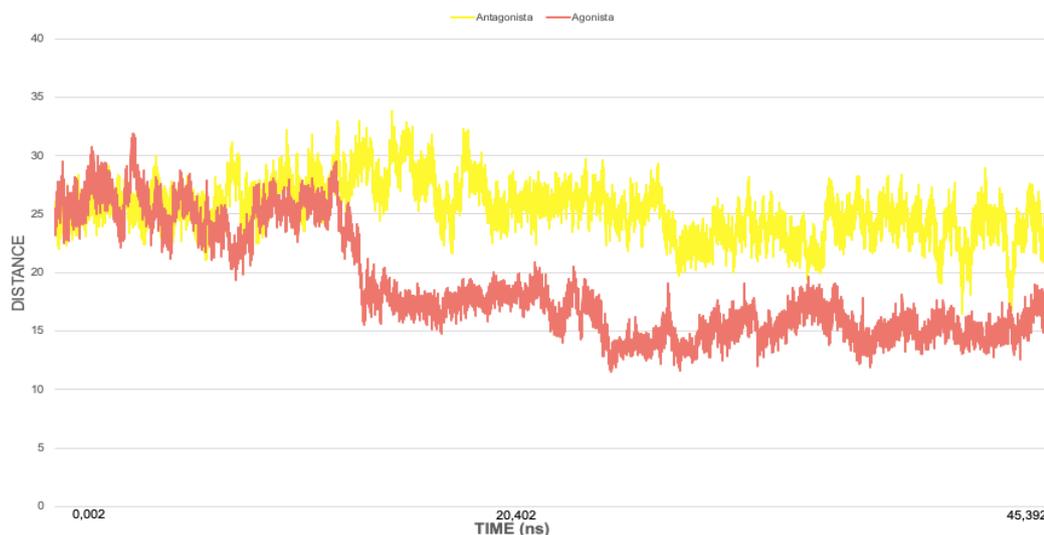


FIGURE 20 | Graph of the trend of the distance between segments s6I and s6IV as a function of the frames.

An analysis of the fluctuations of the atoms of the structure over time was also conducted and no remarkable differences were highlighted, suggesting a similar flexibility of the structures after the ligand binding (see supporting information).

Secondary structures were then investigated. In particular, only the secondary structures of the segments that are part of the ligand-receptor interaction were evaluated. In figure 21, the binding sites of the agonists and antagonists were analyzed to evaluate a rearrangement of the structures.

This was done to understand how the protein is structured in the presence of the agonist and antagonist. In the analysis it can be seen how some residues seem to destabilize in the presence of the antagonist.

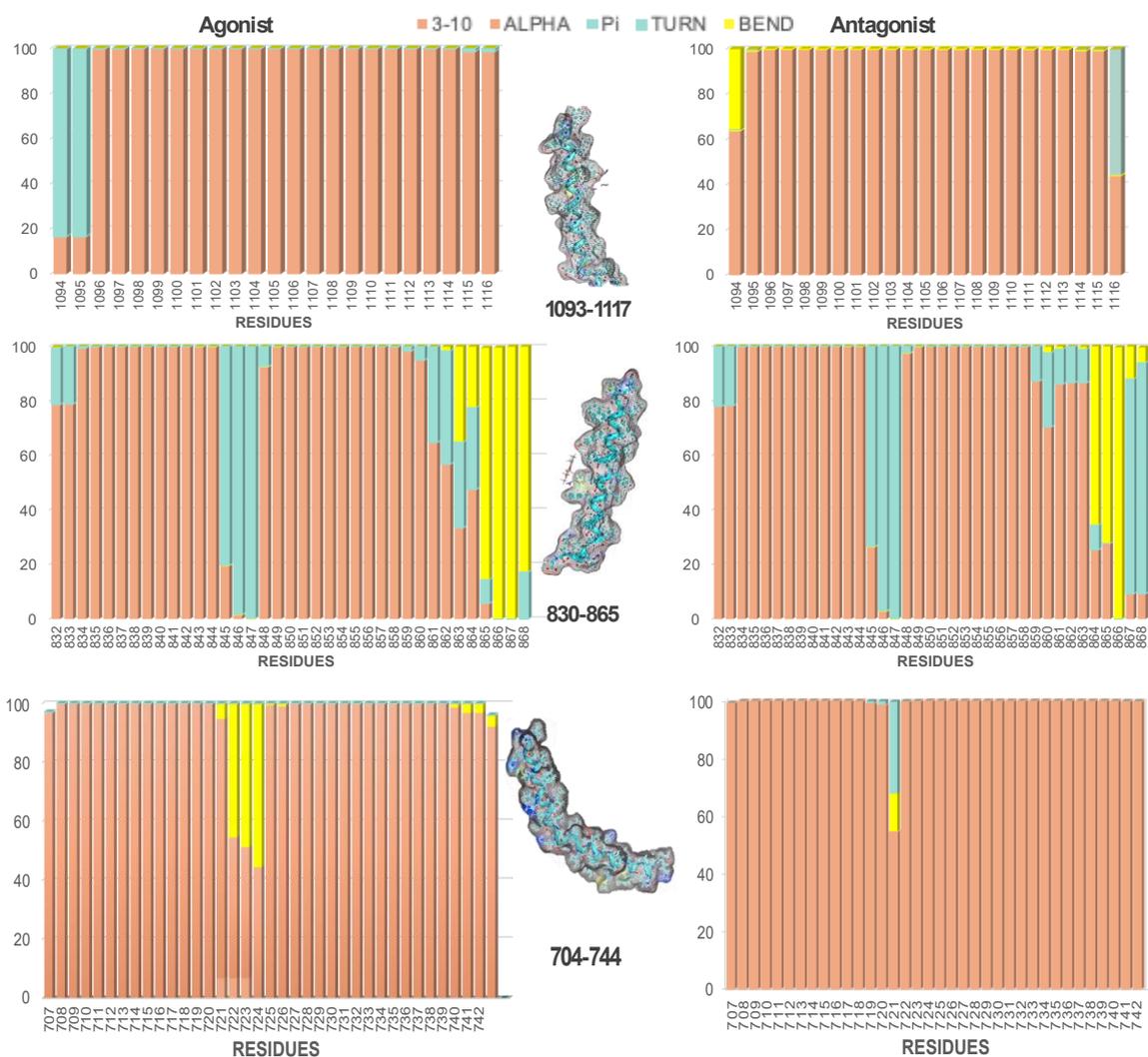


FIGURE 21 | Comparison of the secondary structures of the segments involved in the ligand-receptor binding. a-c) Comparison of the indicated residues, d) highlighting of residues that show changes in secondary structures.

4.3.3 Virtual Screening

All 96,912 compounds were evaluated to analyze the binding affinity with the imposed docking site. Affinities were assessed for the 6jp5 and 6jp8 objectives in terms of energy. Specifically, from the molecular docking simulations, values (vina scores) were obtained that refer to the binding energies.

The range of values is included from -11 to -4.2 kcal * mol⁻¹. Clearly, the ligands that had very low values were discarded, and it was decided to continue the analyzes only with the 10,000 compounds that had a better binding-target coupling in terms of binding energy.

The first 45 compounds with the highest vina scores have been shown in the figure 22 for both 6jp5 and 6jp8 targets. The control value has also been included in the graphs, that is the highest energy value obtained from docking the BAY K8644 agonist.

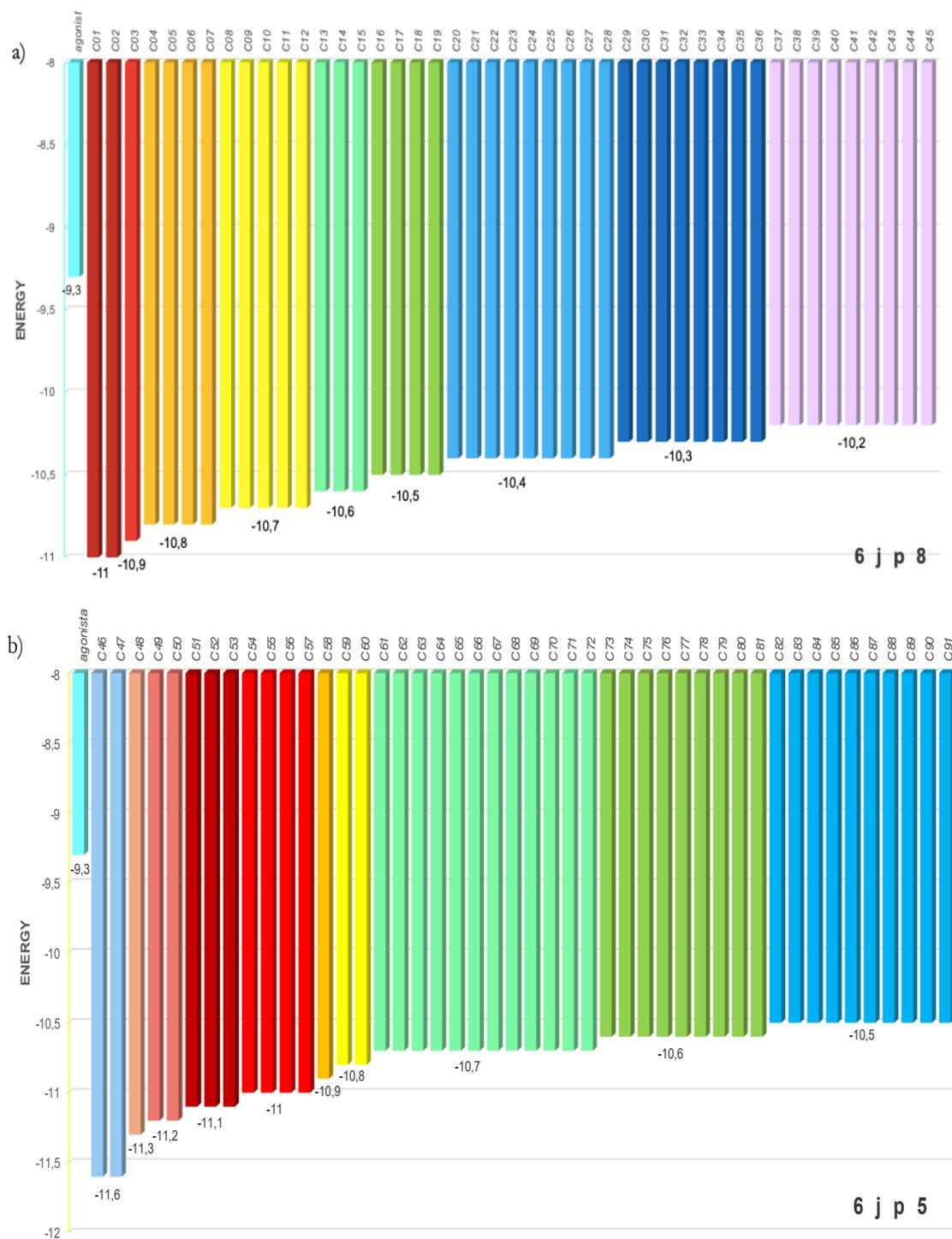


FIGURE 22 | Graph of the first 45 ligands with higher Vina scores: a) ligands that have docked with 6jp8 b) ligands that have docked with 6jp5 structure. Ligands with the same vina score are represented with the same color.

To understand how the compounds behaved in the different cases, the specificity value was calculated for each compound dividing the binding energy value of the agonist by the binding energy value of the antagonist.

A specificity index was therefore evaluated as follows:

$$SPECIFICITY = \frac{vina_{score}Agonist}{vina_{score}Antagonist}$$

All values less than or equal to unit were discarded as they presented low specificity for agonist targets. It is clearly possible that ligands with a high specificity index do not have a high binding energy.

For this reason, two thresholds have been imposed.

In particular:

- for the specificity index was chosen as the threshold value 1.2 since, evaluating all the values, there seems to be a high value. Specificity > 1.2
- for the binding energy the value of -9.3 kcal / mol was imposed, which represents the binding energy value obtained at the end of the docking simulations for the starting antagonist. Energy < - 9.3 kJ/mol

Figure 23 represents a scatter plot where the energy of the ligands and the specificity were driven. On the basis of the applied thresholds, the ligands represented with a purple circle were chosen for the subsequent analysis. For more information on ligand specificity see the supporting information chapter.

Selected compounds were superimposed on the starting agonist to the analysis of the poses, evaluating no longer the energy but the structural affinity of the compounds and the residues involved in the ligand-channel interaction of calcium. The electrostatic interactions of the compounds with calcium channels were analyzed. The residues involved in the binding were compared those involved for the starting agonist.

The data have been reported in Table 1 for easier viewing. It can be seen that the residues that had a higher energy in absolute value PHE853, MET850, PHE801 are always present. It is possible to think that these bonds may confer specificity to the compounds with respect to the binding site.

In reality, the residue TRH 728, VAL725, MET1101, THR729 are also always present but a previous analysis has highlighted that the energy of this residue is low due to the binding of the agonist to the antagonist. This led us to focus on the residues circled in yellow.

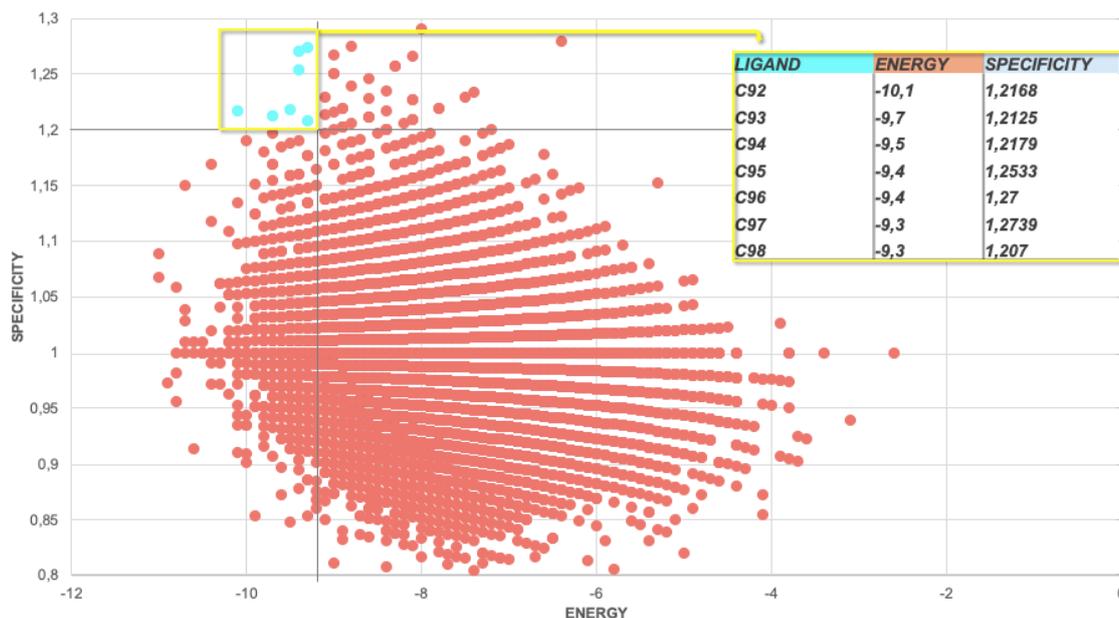
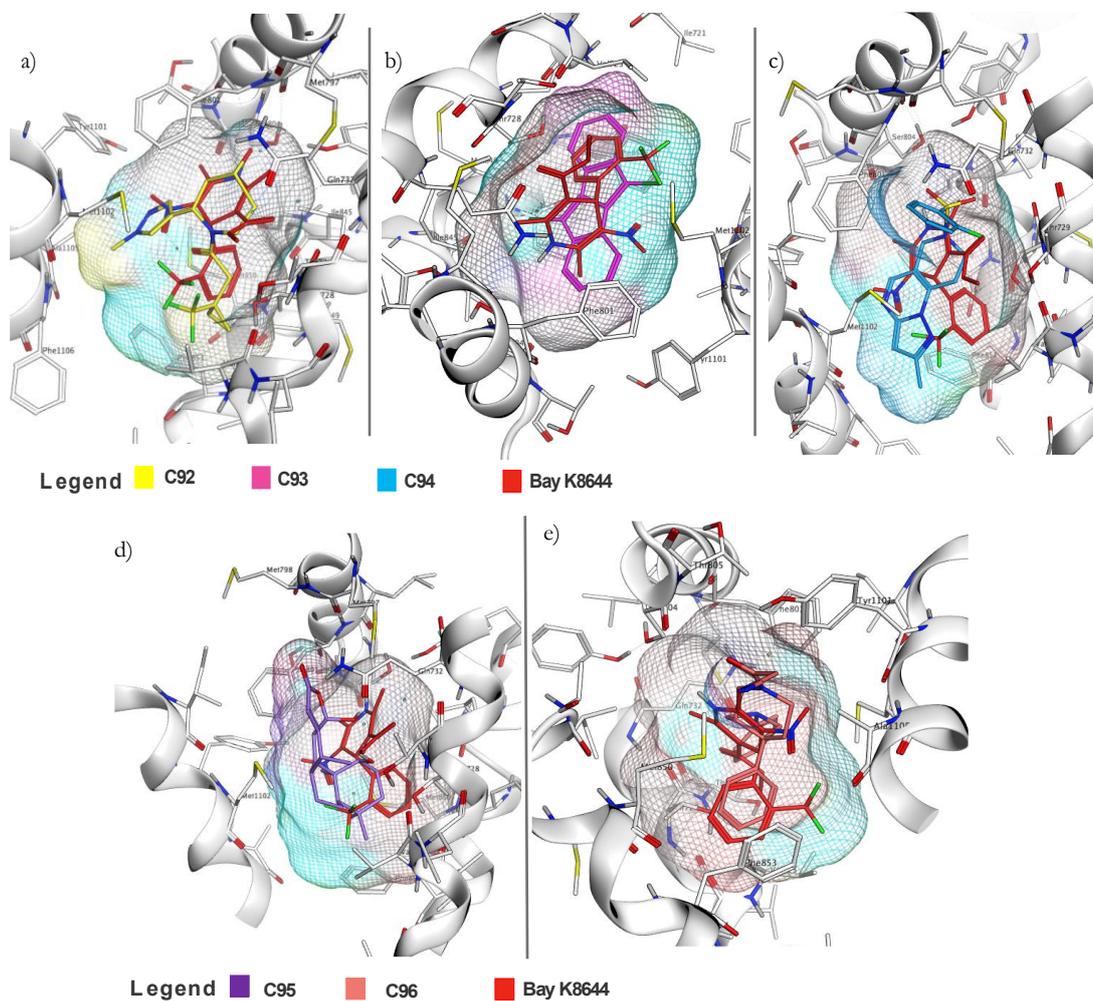


FIGURE 23| Scatter plot of ligand energy as a function of specificity. the purple dots are the ligands chosen for further analysis.

AGONIST	C92	C93	C94	C95	C96	C97	C98
MET797							
ALA1104							
VAL725							
MET1101							
PHE853							
GLN732							
THR728							
THR805							
SER804							
MET850							
TYR1100							
TYR841							
ILE845							
THR729							
PHE801							
MET849							
ILE724							
LEU800							
PHE1105							
MET798							
ILE718							

TABLE 1 | Comparison between the binding residues of the starting agonist and of the compounds analyzed. The colors are representative of the legend on the right of the table.

Compared to Bay-K8644, these drugs, despite being in the same region, have different structures and the aromatic rings are oriented differently. Compounds were superimposed with the starting ligand, representing the starting agonist ligand in red (Figure 24). Clear structural analogies with the starting group Dihydropyridine were presented by these ligands. It is possible to see how all present the characteristic aromatic ring of the dihydropyridine class.



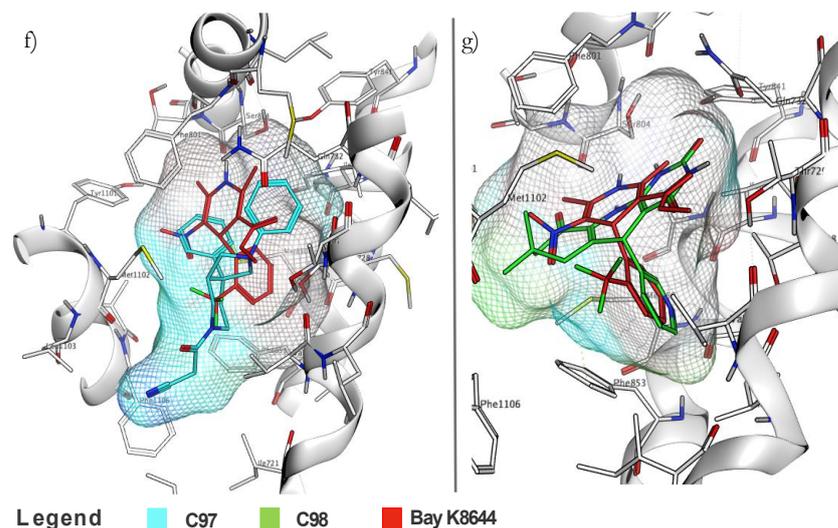


FIGURE 24 | Comparison of a pose for the chosen ligands, superimposed on the pose of the starting agonist.

This ring in some compounds such as C92, C93, C96, C98 is almost completely superimposed on that of BAY-K8644. In addition, the first compound analyzed C92 also has the same starboard group that characterizes BAY-K8644.

4.4 Discussion

Calcium channels are attractive targets for many drugs as they are involved in many functions for the cell. In the present study the interest lies in increasing the calcium flux within the cell, so the new compounds will have to be ungrateful to interact with the channel as activators. The influx of calcium could favor the therapeutic effect of some drugs, so these new compounds will be used to support therapy, especially for the treatment of chronic lymphocytic leukemia.

First, we have focused our attention on the binding site of the calcium channel agonists and the different residues involved in the binding of the same class of compounds, the Dihydropyridines. The residues that had a higher binding energy with respect to the antagonist were identified and they are PHE 853, MET 850, PHE801. Furthermore, it was noted the presence of ILE724 which was not present in the residues involved in the binding with the antagonist.

The conformation changes of the structure linked to the agonist and antagonist were evaluated. It has been seen how the structure with the agonist presents an approach of the

residues S6I and S6IV, assuming there is an opening of the pore of the channel. Figures 19 and 20 show a change in the protein when bound to ligands, in particular the distance between the S6I and S6IV residues (segment forming the pore) decreases over time during the simulation with the agonist. In addition, the secondary structures of the segments involved in the ligand binding of the calcium channel were investigated. Figure 21 shows the secondary structures of the residues involved in binding with the ligands: they were graphed in percentage and slight differences between residues of calcium channel with agonist and antagonist are noted. In particular the residues 862-864. Residues 719-724 show an increase in the secondary bend structure and residue 1116 of alpha.

This shows that the channel with agonist is better structured while in the presence of the antagonist there is a destabilization of these secondary structures.

Subsequently, new compounds were analyzed and the binding residues were compared with those of the agonist. We started from a dataset of 96,912 compounds.

After the molecular docking, the binding energies of the compounds were analyzed and it was decided to proceed with further analyzes only with 10,000 compounds. Figure 22 shows the first 45 compounds that had binding energies, both with docking with the targets extracted from the 6jp8 structure (calcium channel complex and agonist) and for the targets extracted from the 6jp5 structure (calcium channel complex and antagonist).

An analysis was carried out on the specificity of the compounds where we tried to understand which is more specific for the calcium channel linked to the agonist or the calcium channel linked to the antagonist.

The specificity indices were calculated and all compounds with values less than or equal to unity were eliminated as they were evaluated as non-specific.

In addition, a specificity threshold was defined based on the values obtained and it was decided to take only compounds with an index greater than 1.2. Moreover, a second threshold was set on the binding energy, all the compounds with a similarity index more than or equal 1.2 but with energies greater than -9.3 were taken. This value was chosen after a docking with the calcium channel with the starting agonist. In the scatter plot in figure 23 they are highlighted the compounds chosen according to the thresholds imposed for evaluating the poses. The selected compounds C92, C93, C94, C95, c96, c97, c98 were superimposed on the agonist and the residues involved in the bond were evaluated. The presence of all the residues leads us to think that that compound may have the same characteristics as the agonist and the same effects. Other residues in the presence of the

binding with the agonist had lower energy values than the same residues in the presence of the antagonist, this leads us to think that they are discriminating for agonist compounds. The differences between the pose of the compound under analysis and the starting agonist are shown in figure 24, where it is possible to notice that almost all the compounds have a strong similarity from the structural point of view with the agonist. This is an interesting result since the structural affinity is not found instead in compounds with a low energy such as the ligand analyzed in figure S13. On the other hand, by analyzing the residues involved in the binding with the new compounds, it can be seen that not all the residues that bind the agonist with the calcium channel are present. However, the three residues that have been identified are always present because they had a higher binding energy than the others and they are C92,C93,C94,C8.

The compounds that have a higher energy are also the most affinities from the point of view of structural similarity with the starting agonist. The aromatic rings are found with the same orientation and are almost superimposable with the Bay-K8644. In addition, C92 also has the same starboard group that characterizes BAY-K8644.

This result is important because we know that starboard group of DHP is discriminating for a compound with already agonistic activities and we hope that the new compound therefore has not only the same structure but also the same functions.

The aim of this master's thesis was to find a drug with the same characteristics as calcium channel agonists, such as Bay K8644. The four compounds identified seem interesting: they have the highest energies on the calcium channel and a high specificity compared to the other compounds.

It would be interesting to perform more analyzes on these compounds.

CHAPTER V – Conclusions and future perspective

In this master's thesis work, two disease were considered: Non-Hodgkin's lymphomas (NHL) and Chronic Lymphatic Leukemia (CLL). Patients affected by these pathologies possess high amounts of differential antigen CD20 differential antigens. In this context, anti-CD20 monoclonal antibodies (mAbs) attacking CD20 increase the influx of Ca^{2+} and lead to cell apoptosis.

The main hypothesis addressed was improve the actions of anti-CD20 monoclonal antibodies in inducing apoptosis by increasing the flow of calcium inside the cell.

After studying the interactions of the dihydropyridine (DHP) agonist of the calcium channel BAY-K 8644, the aim of the thesis work was to find a drug with the same characteristics in terms of structure and binding sites involved in the protein-binding interactions. All compounds that were inside the binding pocket of the reference agonist were discriminated and those with greater interactions in terms of higher binding energy will be chosen.

With these considerations, the present project combined computer-aided-drug design (CADD), experimental results of previous studies obtained by the Anticancer Antibodies team on calcium channel blockers and the antitumor activity of anti-CD20 monoclonal antibodies and Molecular Dynamics (MD) simulations for the analysis and research of new compounds to support the action of anti-CD20 mAbs, selectively targeting the calcium ions channel and consequently increasing the calcium flow.

The first analysis performed was to identify the binding sites of the dihydropyridine agonists and antagonists within the calcium channel. After this, we moved on to the study of the interactions of the ligands involved in the calcium-binding channel link. The differences between agonists and antagonists in terms of residues forming the binding pocket and related binding energies were evaluated. The conformation differences of the calcium channel bound to the agonist ligand and the antagonist were analyzed and particular attention was paid to the secondary structures of the segments involved in the biding pocket. Once these data were obtained, we moved on to the molecular docking analysis where 96.912 compounds with physiological properties were extracted from a ZINC15 database and subsequently added to the channel for the Docking analysis.

The compounds were analyzed in terms of binding energy and the energy values was compared with starting agonist energy. Two filters have been inserted in order to choose the best compounds.

In particular, the specificity indices of compounds were calculated in term of difference between binding energies obtained by the docking with the calcium channel bound to the agonist and binding energies obtained by the docking with the calcium channel bound to the antagonist. This specificity analysis was performed to understand which compound was more specific for the channel linked to the agonist or antagonist. 8 compounds were then selected for structural analysis by superimposing the compounds with the starting agonist and by evaluating the orientation of the functional groups. The residues involved in binding to the calcium channel were also analyzed.

These 8 compounds have the structural similarities with the starting.

Future studies could test the effectiveness of the compound in increasing the calcium flux inside the cell by carrying out a molecular dynamic and subsequently analyzing different characteristics of the protein. From the analysis of fluctuations conducted on the atoms no major differences were noted in this study when the calcium channel was bound to the agonist and when it was bound to the antagonist. The dynamics of the channels could be analyzed in more detail with the new identified compounds, it could be evaluated how the pore behaves in the presence of these, if important fluctuations are present.

Furthermore, one could study the conformation change of the channel with the new bound compound and in vivo study could be conducted.

The existing LTCC calcium channels are of 4 types Cav1.1, Cav1.2, Cav1.3 and 1.4 and each channel is involved in different functions in different excitable cells. Transcripts for all L-type channel isoforms were detected in lymphocytes. CaV1.1 is co-expressed in neurons that produce gamma-aminobutyric acid (GABA), CaV1.2 and CaV1.3 exhibit a highly overlapping expression pattern in many tissues and particularly in cardiac and neuronal cells [70]. CaV1.2 and CaV1.3 are found in neurons and in the sinoatrial node acting as a pacemaker as they are involved in excitation-contraction coupling. CaV1.3 are mainly present in pancreatic cells and kidneys while Cav1.4 in cells of the retina. It would be interesting to use the compounds found to evaluate the different expressions of calcium channels. The interactions of the compounds found with the other calcium channels could be studied. Moreover, the same project could be carried out using a non-dihydropyridine, such as FPL64176, CGP 48506 and murrayafoline A.

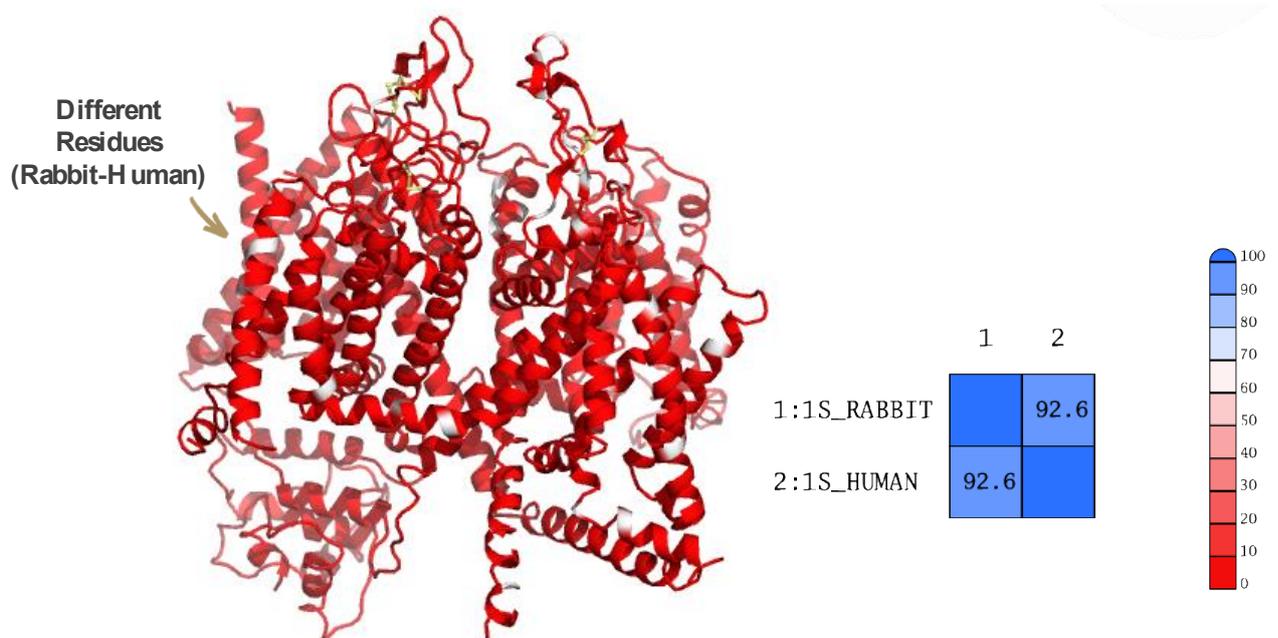
A further future development could be to create a homology model of the human calcium channel for analysis.

A biological evaluation of the calcium channel linked to the compounds found and in combination with direct anti-CD20 antibodies could be performed both in vitro and in vivo. Preliminary toxicity experiments will address questions relating to the therapeutic index.

CHAPTER VI – Supporting information

This chapter contains supporting information for a better understanding some topics covered within the thesis work and further analysis. There is a table at the end of the chapter that contains all the software used for image processing and analysis.

For a better understanding of the similarity of rabbit to human structures, the alignment of the alpha 1 subunits of the human and rabbit calcium channel is reported. a model was also built which highlights the different residues.



SUPPORTING FIGURE 1 | The figure represents a model of the rabbit alpha 1 subunit calcium channel. The white portions within the model represent residues that are different in the human calcium channel sequence. On the right in the figure is the graph of the similarity between the structures of the calcium channel of the rabbit and the calcium channel of the human. The sequences were downloaded from the UNIPROT database and then compared with the MOE software. The Uniprot IDs for sequences are: rCav1.1: P07293; hCav1.1: Q13698.

HUMAN 1-1874

HUMAN/1-1874 1 MEESSPQDEGLRKRKOPKPKVPELPPRPDRAFLCTLENELRKAACISIVEWKPFETIILLT
RABBIT/1-1874 1 MEESSPQDEGLRKRKOPKPKVPELPPRPDRAFLCTLENELRKAACISIVEWKPFETIILLT

HUMAN 1-1874^Q TT

HUMAN/1-1874 61 IFANCVALAVYLPMPEDDNNSLNLGLEKLEYFFLIVFSTEAMKIIAYGFLFHODAYLRS
RABBIT/1-1874 61 IFANCVALAVYLPMPEDDNNSLNLGLEKLEYFFLIVFSTEAMKIIAYGFLFHODAYLRS

HUMAN 1-1874^Q

HUMAN/1-1874 121 GWNVLDFFIVFLEVFTVILEQVNVIOSHTAPMSSKAGLDVKALRAFVLRPLRLVSGVP
RABBIT/1-1874 121 GWNVLDFFIVFLEVFTVILEQVNVIOSHTAPMSSKAGLDVKALRAFVLRPLRLVSGVP

HUMAN 1-1874^Q

HUMAN/1-1874 181 SIQVVLNSIFRAMLPLFHAIIVLFMVIIYAITGLELFGKRMHKTCTYFHTDIVATVENE
RABBIT/1-1874 181 SIQVVLNSIFRAMLPLFHAIIVLFMVIIYAITGLELFGKRMHKTCTYFHTDIVATVENE

HUMAN 1-1874^Q

HUMAN/1-1874 241 EFSFCARTGSGRRCTINGSECRGGWPGPNHGITHFDNFGFSMLTVYQCTIMEGWTDVLYW
RABBIT/1-1874 241 EFSFCARTGSGRRCTINGSECRGGWPGPNHGITHFDNFGFSMLTVYQCTIMEGWTDVLYW

HUMAN 1-1874^Q

HUMAN/1-1874 301 VNDATGNWFWIYFVTLILLGSEFFILNLVVLGVLGSEFTKEREKAKSRGTFOKLRKQQLD
RABBIT/1-1874 301 VNDATGNWFWIYFVTLILLGSEFFILNLVVLGVLGSEFTKEREKAKSRGTFOKLRKQQLD

HUMAN 1-1874^Q

HUMAN/1-1874 361 EDLRGYMSWITQGEVMDVEDFREGLKSLDEGGSDTESYEIAENKTIQFIRHWROWNRV
RABBIT/1-1874 361 EDLRGYMSWITQGEVMDVEDFREGLKSLDEGGSDTESYEIAENKTIQFIRHWROWNRV

HUMAN 1-1874^Q TT

HUMAN/1-1874 421 FRWCHDIIKSKVFWYVWVLIIVALTLSIASEHNNQPLWTRHODIANRVILSIFTEML
RABBIT/1-1874 421 FRWCHDIIKSKVFWYVWVLIIVALTLSIASEHNNQPLWTRHODIANRVILSIFTEML

HUMAN 1-1874^Q

HUMAN/1-1874 481 MKMYGLGLQYFMSIFNRFDCFVVCSGILEILLVESGAMTPLGISVLCRIRLLRIFKTK
RABBIT/1-1874 481 MKMYGLGLQYFMSIFNRFDCFVVCSGILEILLVESGAMTPLGISVLCRIRLLRIFKTK

HUMAN 1-1874^Q

HUMAN/1-1874 541 YWTSLSNVLVSLNSIRSIASLLLLLFLLFIVIFALLGMQLFGGRYDFEDTEVRRSNFDNF
RABBIT/1-1874 541 YWTSLSNVLVSLNSIRSIASLLLLLFLLFIVIFALLGMQLFGGRYDFEDTEVRRSNFDNF

HUMAN 1-1874^Q

HUMAN/1-1874 601 POALISVFOVLTGEDWTSMYNGLMAYGGPSYPGMVLCYFIFLFCGNVILLNVFLATA
RABBIT/1-1874 601 POALISVFOVLTGEDWTSMYNGLMAYGGPSYPGMVLCYFIFLFCGNVILLNVFLATA

HUMAN 1-1874^Q

HUMAN/1-1874 661 VDNLAEAESLTSQAKAEEKRRRMSKSLPDRSEEEKSTMAKKLEOKPKGEGIPTTAKI
RABBIT/1-1874 661 VDNLAEAESLTSQAKAEEKRRRMSKSLPDRSEEEKSTMAKKLEOKPKGEGIPTTAKI

HUMAN 1-1874^Q

HUMAN/1-1874 721 KIDFESNVNEVKDEYPSADFPGDDEEDEPEITLSPRRPLAELQLKRENAVPIPEASSFF
RABBIT/1-1874 721 KIDFESNVNEVKDEYPSADFPGDDEEDEPEITLSPRRPLAELQLKRENAVPIPEASSFF

HUMAN 1-1874^Q

HUMAN/1-1874 781 IFSFTNKIVRVLCHRIVNATWFTNFILLFILLSSAALAAEDPIRADSMRNOILKHFIDIGFT
RABBIT/1-1874 781 IFSFTNKIVRVLCHRIVNATWFTNFILLFILLSSAALAAEDPIRAESVNRNOILGYFDIAFT

HUMAN 1-1874^Q

HUMAN/1-1874 841 SVFTVEIVLKMTTYGAFLHKGSFCRNYFNLDDLVAVAVSLISMGLESSAISVVKILRVLR
RABBIT/1-1874 841 SVFTVEIVLKMTTYGAFLHKGSFCRNYFNLDDLVAVAVSLISMGLESSAISVVKILRVLR

HUMAN 1-1874^Q

HUMAN/1-1874 901 VLRPLRAINRAKGLKHVVQCMFVAISTIGNIVLVTTLLQFMFACIGVQLFRKGFPSQDL
RABBIT/1-1874 901 VLRPLRAINRAKGLKHVVQCMFVAISTIGNIVLVTTLLQFMFACIGVQLFRKGFPSQDL

HUMAN 1-1874^Q

HUMAN/1-1874 961 SKMTEEECRGYYYVYKDDPDMQTELRLRHWVHSDFFDNVLSAMMSLFTVSTFEWQPLL
RABBIT/1-1874 961 SKMTEEECRGYYYVYKDDPDMQTELRLRHWVHSDFFDNVLSAMMSLFTVSTFEWQPLL

HUMAN 1-1874^Q

HUMAN/1-1874 1021 YKAIDSNABEDVGGPHYNNRVEMATFFIYIILIAFFMMNIFVGFVIVTFOEQGETEYKNC
RABBIT/1-1874 1021 YKAIDSNABEDMGGPHYNNRVEMATFFIYIILIAFFMMNIFVGFVIVTFOEQGETEYKNC

HUMAN 1-1874^Q

HUMAN/1-1874 1081 LDKNORQCVOYALKARPLRCYIPKNPYQYQVWYVVTSSYFEYLMFALIMLNTICLGMQHY
RABBIT/1-1874 1081 LDKNORQCVOYALKARPLRCYIPKNPYQYQVWYVVTSSYFEYLMFALIMLNTICLGMQHY

HUMAN 1-1874^Q

HUMAN/1-1874 1141 NQSPQMNHISDILNVAFITIFTLEMILKLMFAKRGYFSDPWNVDFLIVIGSIIIDVILS
RABBIT/1-1874 1141 HQSEEMNHISDILNVAFITIFTLEMILKLMFAKRGYFSDPWNVDFLIVIGSIIIDVILS

HUMAN 1-1874^Q

HUMAN/1-1874 1201 EIDTFLASSGGIYCLGGCCGNVDDESARISSAFFRLFRVMRLIKLISRAEGVRTLLWTF
RABBIT/1-1874 1201 EIDTFLASSGGIYCLGGCCGNVDDESARISSAFFRLFRVMRLIKLISRAEGVRTLLWTF

HUMAN 1-1874^Q

HUMAN/1-1874 1261 IKSFAQALPYVALLIVMLFFIYAVIGMQMFGKRIALVDGTQINRNNNFQTFPQAVLLLRCA
RABBIT/1-1874 1261 IKSFAQALPYVALLIVMLFFIYAVIGMQMFGKRIALVDGTQINRNNNFQTFPQAVLLLRCA

HUMAN 1-1874^Q

HUMAN/1-1874 1321 TGEAWQEILLACSYGKLCDEPESDYAPGEEYTCGTNFAYYYFTSFYMLCAFLVNLVAVI
RABBIT/1-1874 1321 TGEAWQEILLACSYGKLCDEPESDYAPGEEYTCGTNFAYYYFTSFYMLCAFLVNLVAVI

HUMAN 1-1874^Q

HUMAN/1-1874 1381 MDNFDYITDWSILGPHHLEDFKATWAEYDPEAKGRIKHLDVVTLLRRIQPPLEGFKFCP
RABBIT/1-1874 1381 MDNFDYITDWSILGPHHLEDFKATWAEYDPEAKGRIKHLDVVTLLRRIQPPLEGFKFCP

HUMAN 1-1874^Q

HUMAN/1-1874 1441 HRVACKRVLGMNMPINSDGTVTFNATLFAVLRALKIKTEGNFEQANEELRAIKKIWR
RABBIT/1-1874 1441 HRVACKRVLGMNMPINSDGTVTFNATLFAVLRALKIKTEGNFEQANEELRAIKKIWR

HUMAN 1-1874^Q

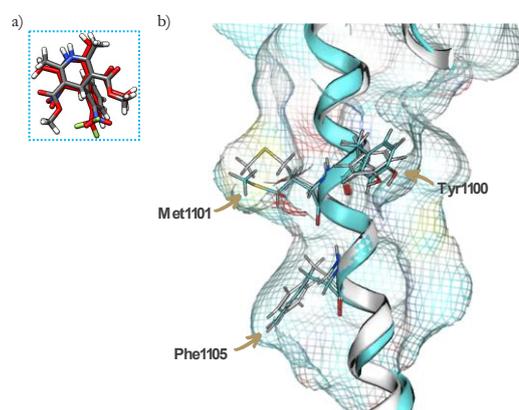
HUMAN/1-1874 1501 TSMKLLDQVIPPIDDEVTGKRYATFLIQEHFRKFMKROEEYYSYRPKKDLVQIQAGLR
RABBIT/1-1874 1501 TSMKLLDQVIPPIDDEVTGKRYATFLIQEHFRKFMKROEEYYSYRPKKDLVQIQAGLR

acc

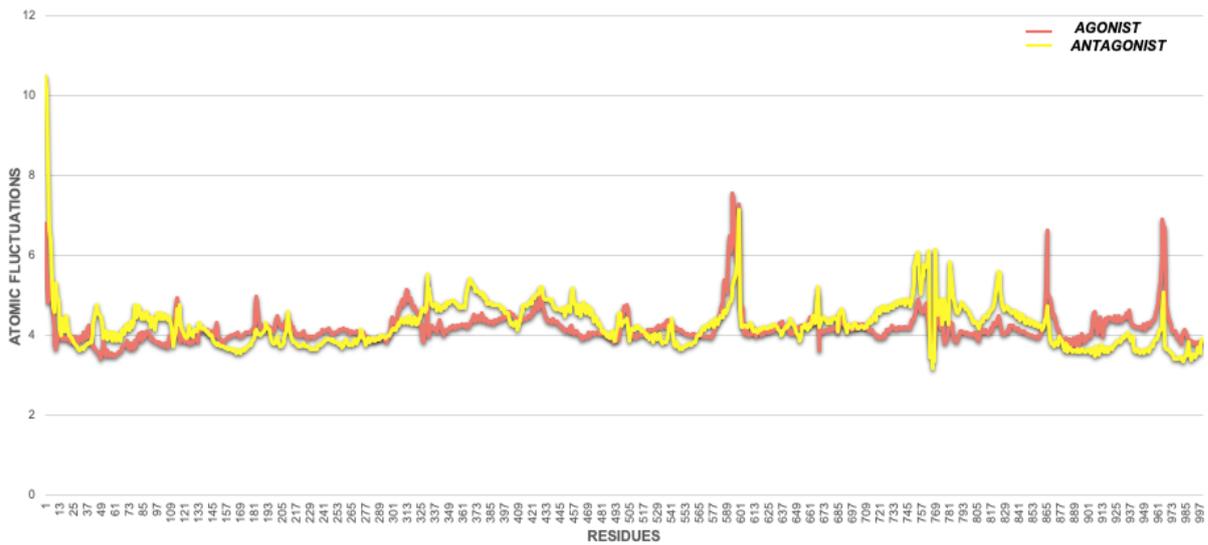
HUMAN 1-1874			
HUMAN/1-1874	1561	T T E E E A A P E T C R T V S G D I A A E E E L E R A M V E A A M E E G I F R R T G G L F G Q V D N F L E R T N S I P P	
RABBIT/1-1874	1561	T T E E E A A P E T R R T I S G D I T A E E E L E R A M V E A A M E E R I F R R T G G L F G Q V D N F L E R T N S I P P	
HUMAN 1-1874^C			
HUMAN/1-1874	1621	V M A N Q R P L Q F A E I E M E E M E S P V F L E D F P Q D P R T N P L A R A N T N N A N A N V A Y G N S N H S N S H V	
RABBIT/1-1874	1621	V M A N Q R P L Q F A E I E M E E M E S P V F L E D F P Q D A R T N P L A R A N T N N A N A N V A Y G N S N H S N N Q M	
HUMAN 1-1874^C			
HUMAN/1-1874	1681	F S S V H Y E R E F F E E T E T P A T R G R A L G Q P C R V L G P H S K P C V E M L K G L L T Q R A M P R G Q A P P A P	
RABBIT/1-1874	1681	F S S V H C E R E F F G E A E T P A A G R G A L S H S H R A L G P H S K P C A G K L N G Q L V Q P G M P I N Q A P P A P	
HUMAN 1-1874^C			
HUMAN/1-1874	1741	C Q C P R V E S S M P E D R K S S T P G S I H E T P H S R S T R E N T S R C S A P A T A L L I Q K A L V R G G I S T L	
RABBIT/1-1874	1741	C Q Q P S T D P P E R G O R R T S L T G S I Q D E A P Q R R S E G S T P R R P A P A T A L L I Q E A L V R G G L D T L	
HUMAN 1-1874^C			
HUMAN/1-1874	1801	A A D A N F I M A T G Q A L A D A C Q M E P E E V E V A T E L L K G R E A P E G M A S S L G S I N L G S S L G S I D Q	
RABBIT/1-1874	1801	A A D A G F V T A T S Q A L A D A C Q M E P E E V E V A T E L L K A R E S V O G M A S V P S I S R R S S L G S I D Q	
HUMAN 1-1874^C			
HUMAN/1-1874	1861	H Q G S Q E T L I P P R L	
RABBIT/1-1874	1861	V Q G S Q E T L I P P R P	
	acc		

SUPPORTING FIGURE 2 | Sequence alignment of the $\alpha 1$ subunit between rabbit Cav1.1 and human Cav channels. Secondary structural elements of the rabbit Cav1.1 are indicated above the sequence alignment and color-coded for the four repeats. Each residue has been highlighted with different colors and where no matches have been found, the residue is white. Furthermore, secondary structures were also reported on the sequences. The Uniprot IDs for alignment of sequences are: rCav1.1: P07293; hCav1.1: Q13698. The sequence alignments were performed using Clustal Omega web server and figures were generated by ESPrnt 3.0 [71]

The analysis of the binding sites of the agonist bayk8466 and antagonist Nifedipine ligands, leads to evaluate a different orientation of some residues. As regards the conformation change of the domains, it can be noted that the residue Phe1105 in the calcium channel linked with nifedipine antagonist is in the "up" position, while as regards the calcium channel linked to the agonist (Bay K8644) the same residue is in "down" position.

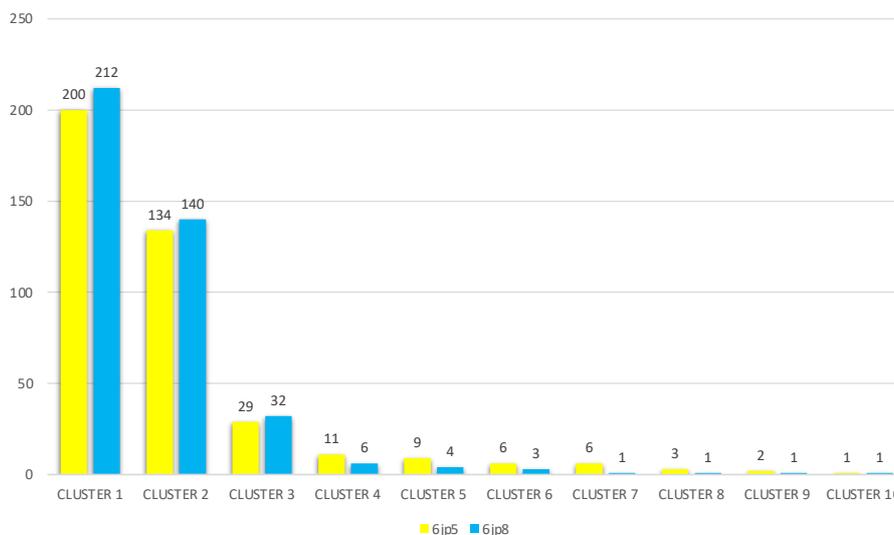


SUPPORTING FIGURE 3 | Overlap of the 6jp8 (calcium channel complex with agonist) and 6jp5 (calcium channel complex with antagonist) structures. Figure a) is a superposition of the two ligands (agonist red BAYK8466, antagonist gray NIFEDIPINE). In figure b) a superimposition of the channels (blue structure with agonist, blue structure with antagonist) was made after the simulation of molecular dynamics and the differences in orientation of some residues were highlighted: MET1101, TYR 1100, Phe 1105.



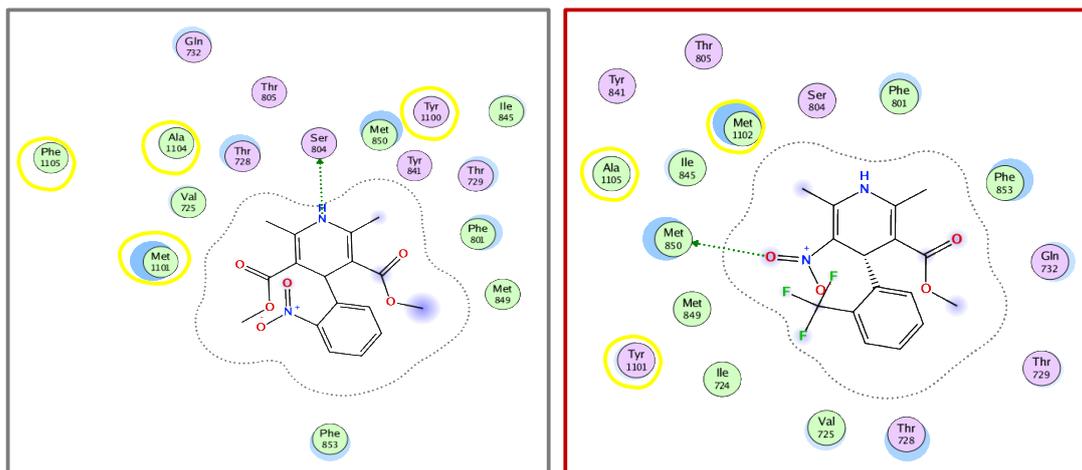
SUPPORTING FIGURE 4 | Analysis of the fluctuations of the calcium channel. The analysis was carried out on the backbone and in particular on the alpha carbons of the calcium channel and agonist structures (pink), calcium channel and antagonist (yellow). The graphs were plotted with Amber's tool gnuplot.

From the simulation of molecular dynamic MD the trajectory was used to conduct an analysis on the clusters. In this way 10 clusters were extracted and for each of these the centroid was extracted. From the following image it is possible to see how the first 5 clusters are the most populated and therefore the most representative ones.



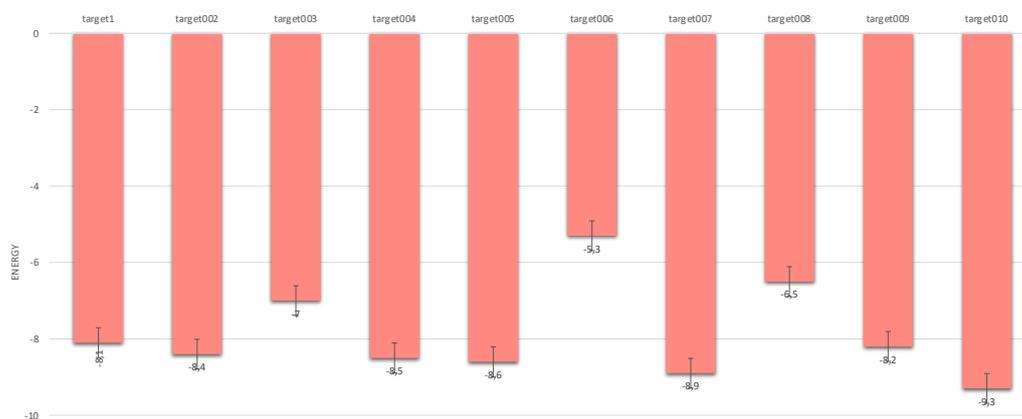
SUPPORTING FIGURE 5 | Cluster analysis obtained from the MD restaurant for the calcium channel structure linked to the antagonist (yellow) and linked to the agonist (blue). the graph shows that the first 5 clusters are the most representative. the histogram was plotted with excel starting from the data format files obtained with Amber's tool cpptraj.

a) Electrostatic Interactions

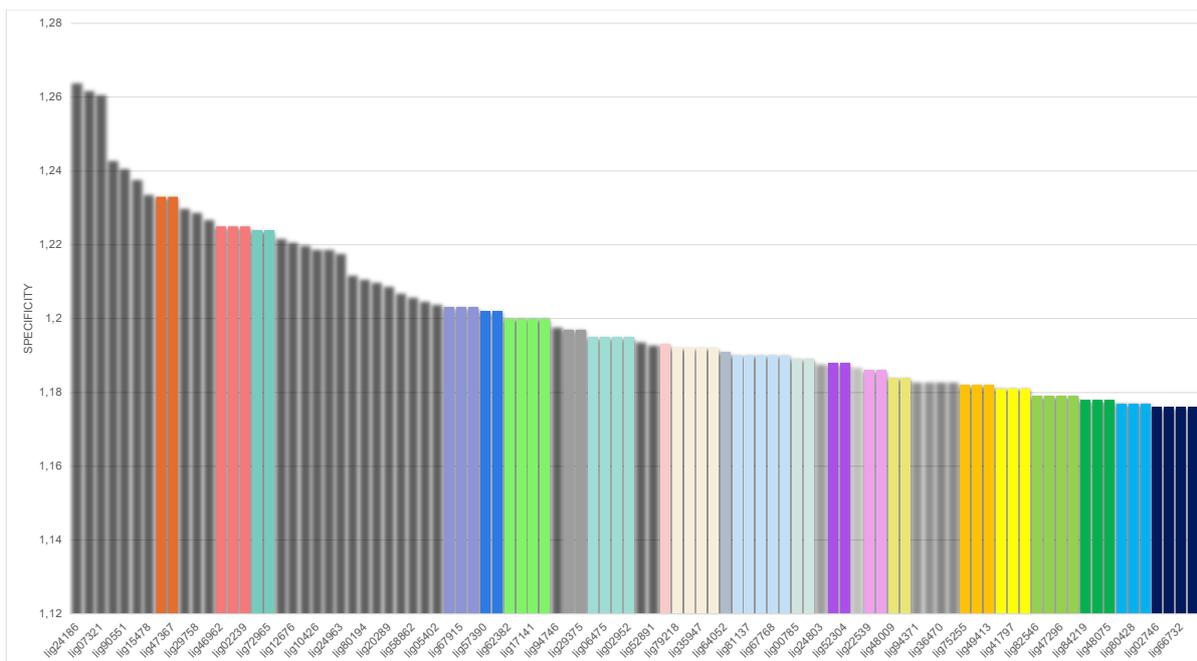


SUPPORTING FIGURE 6 | Analysis of the electrostatic interactions of the residues involved in binding with the ligand. The structure in the red box is the agonist BAYK 8466 while the antagonist NIFEDIPINE is present in the gray box. The circled residues are those that change orientation in the presence of the ligands (see figure S3).

For the molecular docking analysis, a docking was first carried out on the starting ligand BAYK 8466 in order to have the binding energy with the calcium channel to make comparisons with the new compounds.

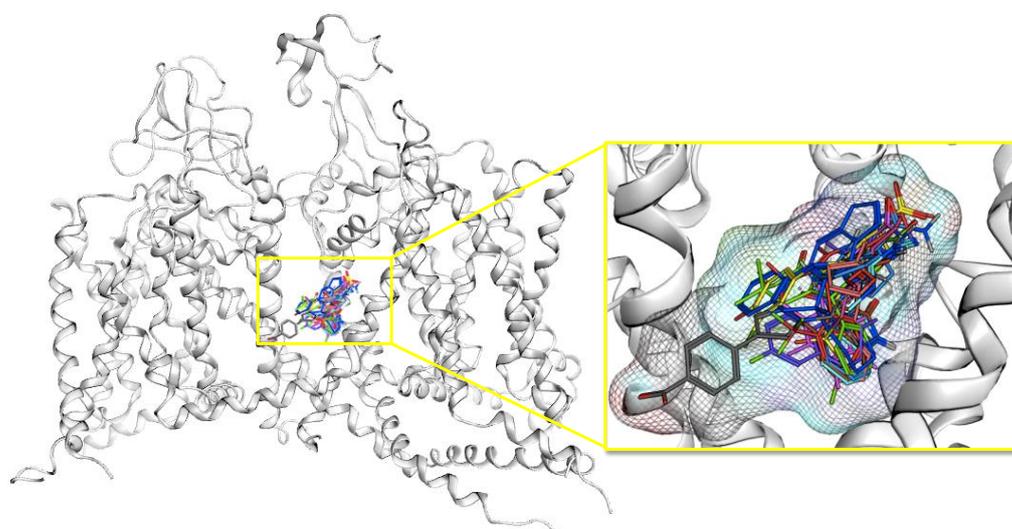


SUPPORTING FIGURE 7 | Histogram of the binding energies of the BAYK8466 agonist with the ten targets extracted from the cluster analysis from the 6jp8 structure after the MD simulation. It can be seen that the highest energy value turns out to be -9.3 kcal / mol.

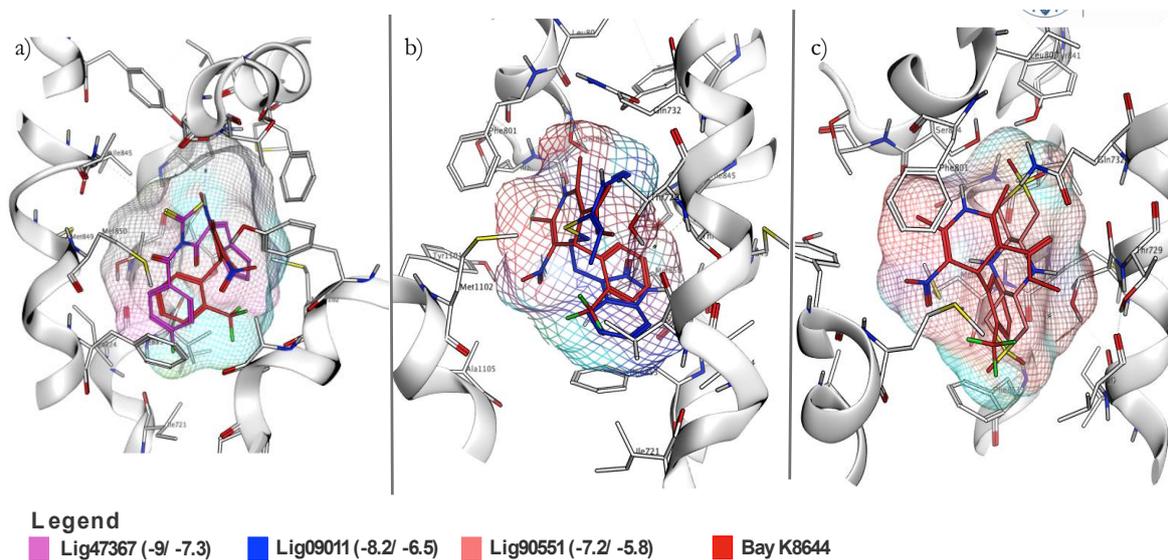


SUPPORTING FIGURE 8 | Histogram for the first 50 compounds with the highest specificity. Ligands with the same specificity value are indicated with the same colors. The plot is made by excel.

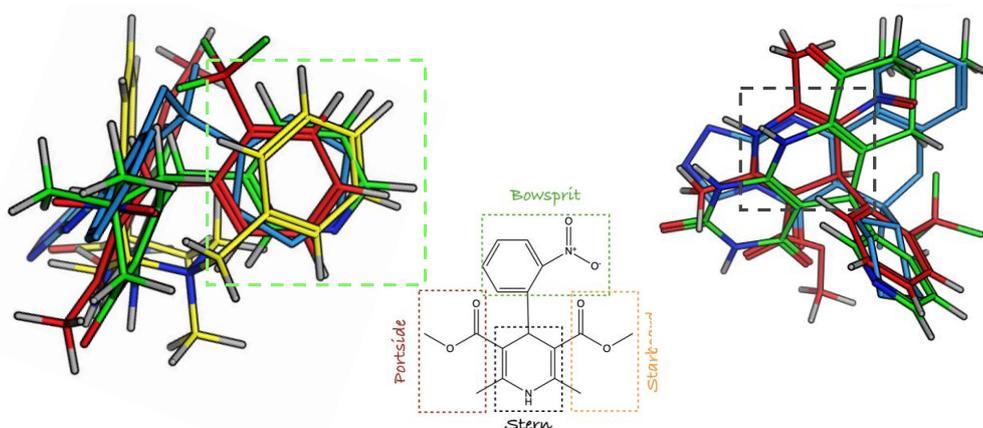
The chosen ligands were then superimposing with the starting one for a structural affinity analysis. Figure S9 shows the superposition of all the ligands to be analyzed to indicate the binding pocket in which they are inserted. The following images will be a comparison of the residues involved in the binding with the compounds that have a high specificity but a lower binding energy than the starting agonist. The binding residues are then compared with the starting agonist.



SUPPORTING FIGURE 9 |Representation of the calcium channel with the ligands chosen for the structural affinity analysis with BAY-K8466, the focus is on the identified binding pocket.

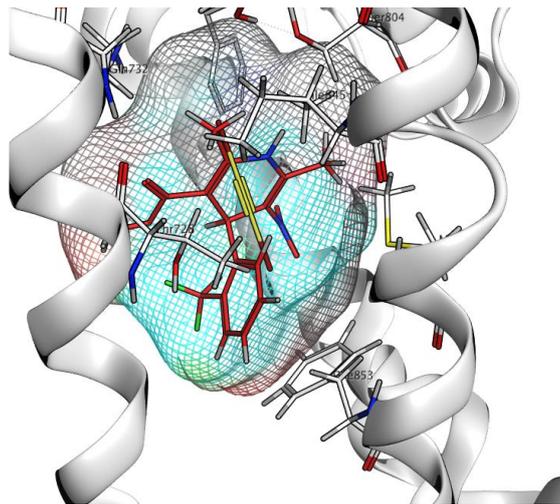


SUPPORTING FIGURE 10 | Comparison of the orientation and of the residues involved in the bond between the starting agonist in red and the ligand47367 (fuchsia), ligand09011 (blue) and ligand90551 (pink). The closer the bond energy is to that of the agonist (-9.3) the more the structure has structural similarity.



SUPPORTING FIGURE 11 | Main structural groups present within the molecules belonging to the dihydropyridine (DHPs) group are highlighted in the central part of the figure. The compounds analyzed are shown on the right and left of the figure and also in this case the presence of the aromatic rings, also present in the DHPs, has been highlighted.

LIGAND	AGONIST
	ALA1104
	VAL725
	MET1101
PHE853	PHE853
GLN732	GLN732
THR728	THR728
	THR805
SER804	SER804
MET850	MET850
	TYR1100
	TYR841
	ILE845
	THR729
	PHE801
	MET849
	ILE724

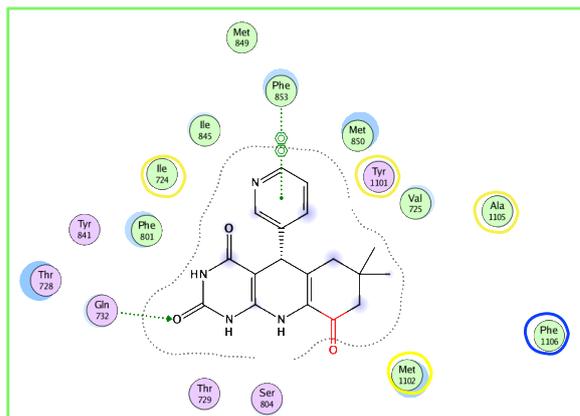


Lig0018

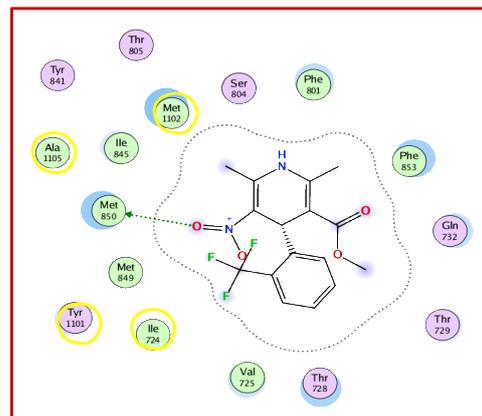
SUPPORTING FIGURE 12 | Comparison of the orientation and of the residues involved in the bond between the starting agonist in red and the ligand0018 that presented a binding energy value of -2.1. On the right of the figure is present a table with the residues involved in the bond for the Ligand 0018 and the Agonist. The low value of binding energy affects not only the structure, which is completely different, there is no presence of aromatic rings, but also the residues involved in the electrostatic interactions.

Legend

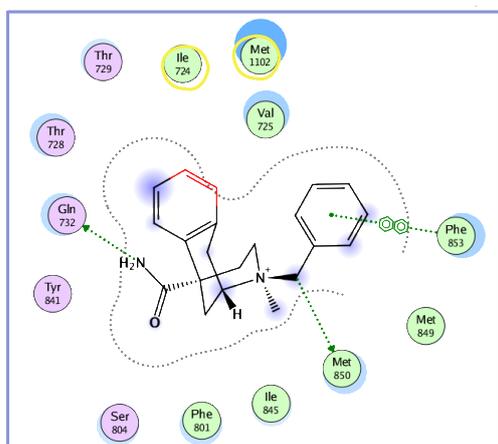
- polar
- acidic
- basic
- greasy
- proximity
- contour
- sidechain acceptor
- ← sidechain donor
- backbone acceptor
- ← backbone donor
- ligand exposure
- solvent residue
- metal complex
- solvent contact
- metal/ion contact
- receptor exposure
- arene-arene
- H arene-H
- + arene-cation



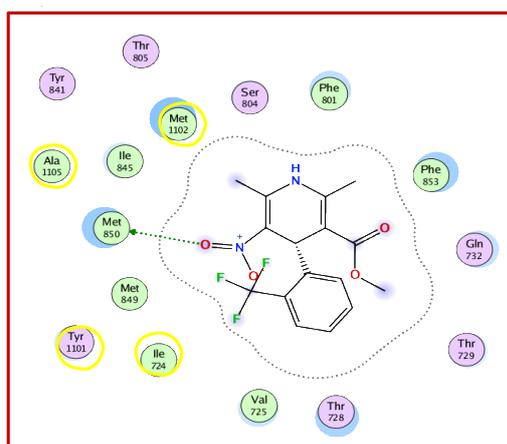
Lig30550 (-11.1/ -9)



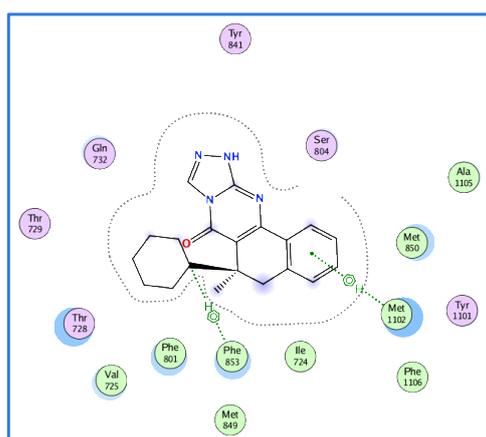
Bay K8644



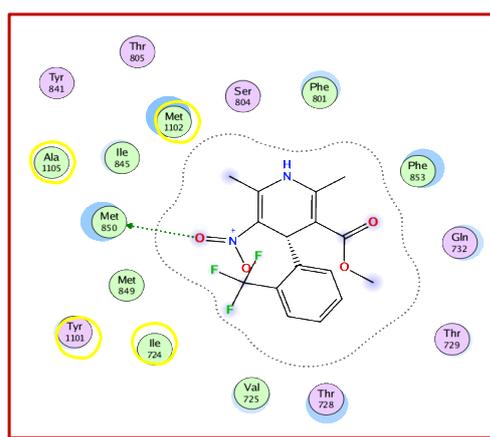
■ Lig24186 (-11/ -8.6)



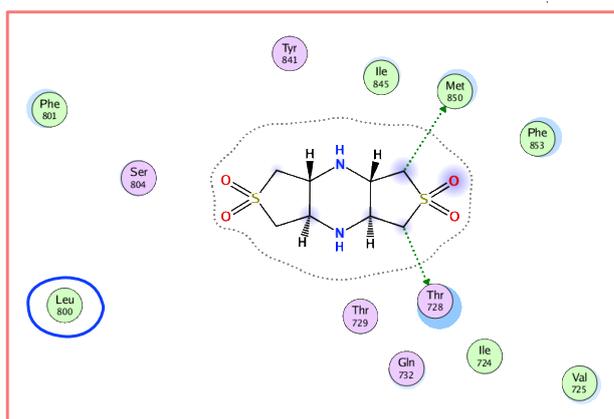
■ Bay K8644



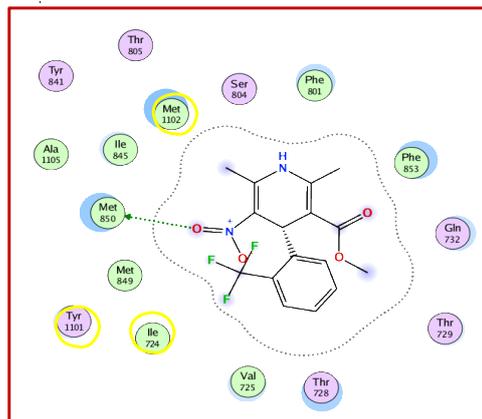
■ Lig53224 (-11)



■ Bay K8644



■ Lig90551 (-7.2/ -5.8)



■ Bay K8644

SUPPORTING FIGURE 13 | Analysis of the electrostatic interactions of the residues involved in binding with the ligand. The structure in the red box is the agonist BAYK 8466 while Ligand90551 (pink), Ligand53224 (blue), Ligand 24186 (purple), Ligand 30550 (green). The circled residues in yellow are those that change orientation in the presence of the ligands (see figure S3)

CHAPTER VII – Acknowledgments

I would like to express my gratitude to my supervisor prof. Marco Agostino Deriu who taught me the basics of molecular modeling. His positivity was a great support to me.

I would like to thank Jordane Preto because he allowed me to become part of the project, for the endless explanations and advices, for the patience with my imperfect English and the ease with which he guided me along this work.

I thank Lorenzo Pallante for supporting me and giving me a great hand in drafting the paper.

I am very grateful to my teammates, Lorenza, Eleonora and Valentina for sharing the good and bad moments of these years and for building a solid and precious friendship.

I thank my longtime friend Federica, a strong shoulder to lean on from the beginning of this journey, even if at a distance.

Finally, I thank my family for always believing in me and for providing me with a safe place to pour out all the anxieties and torments of these years without ever being judged. I would like to tell them that this is also their merit, for their patience and for the silent hand they have extended to me at every moment.

In particular, I would like to thank my life partner Angelo for having been my pillar of strength and for the thousand smiles torn even when everything was difficult.

Finally, I thank my guardian angels for having marked my path, taught me not to give up, having believed in me from afar, perhaps too far away, but love never passes here.

Angela Ferrara

CHAPTER VIII – References

1. Spasevska, I. et al. Calcium channel blockers impair the antitumor activity of anti-CD20 monoclonal antibodies by blocking EGR-1 induction. *Mol. Cancer Ther.* molcanther.0839.2019 (2020) doi:10.1158/1535-7163.mct-19-0839.
2. Lipkind, G. M. & Fozzard, H. A. Molecular modeling of interactions of dihydropyridines and phenylalkylamines with the inner pore of the L-type Ca²⁺ channel. *Mol. Pharmacol.* 63, 499–511 (2003).
3. Grabner, M., Wang, Z., Hering, S., Striessnig, J. & Glossmann, H. Transfer of 1,4-dihydropyridine sensitivity from L-type to class A (BI) calcium channels. *Neuron* 16, 207–218 (1996).
4. Lipkind, G. M. & Fozzard, H. A. Molecular modeling of interactions of dihydropyridines and phenylalkylamines with the inner pore of the L-type Ca²⁺ channel. *Mol. Pharmacol.* 63, 499–511 (2003).
5. Brown, J. R., Weng, A. P. & Freedman, A. S. Hodgkin Disease Associated with T-Cell Non-Hodgkin Lymphomas: Case Reports and Review of the Literature. *Am. J. Clin. Pathol.* 121, 701–708 (2004).
6. Dotan, E., Aggarwal, C. & Smith, M. R. Impact of rituximab (Rituxan) on the treatment of B-cell non-Hodgkin's lymphoma. *P T* 35, 148–157 (2010).
7. Eichhorst, B. et al. Chronic lymphocytic leukaemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* 26, vi50–vi54 (2015).
8. FROM THE EDITORS : Monoclonal Antibodies and Therapies. Group 2004–2004 (2004).
9. Zhang, Q., Chen, G., Liu, X. & Qian, Q. Monoclonal antibodies as therapeutic agents in oncology and antibody gene therapy. *Cell Res.* 17, 89–99 (2007).
10. Dotan, E., Aggarwal, C. & Smith, M. R. Impact of rituximab (Rituxan) on the treatment of B-cell non-Hodgkin's lymphoma. *P T* 35, 148–157 (2010).
11. Grimsley, A., Shah, K. S. & McKibbin, T. Monoclonal antibodies in cancer. *Pharm. Biotechnol. Fundam. Appl.* Fourth Ed. 12, 337–359 (2013).
12. Rosenberg, L. et al. Calcium channel blockers and the risk of cancer. *J. Am. Med. Assoc.* 279, 1000–1004 (1998).
13. Zhao, Y. et al. Molecular Basis for Ligand Modulation of a Mammalian Voltage-Gated Ca²⁺ Channel. *Cell* 177, 1495–1506.e12 (2019).
14. Payandeh, J., Gamal El-Din, T. M., Scheuer, T., Zheng, N. & Catterall, W. A. Crystal structure of a voltage-gated sodium channel in two potentially inactivated states. *Nature* 486, 135–139 (2012).
15. Patel, R., Montagut-Bordas, C. & Dickenson, A. H. Calcium channel modulation as a target in chronic pain control. *Br. J. Pharmacol.* 175, 2173–2184 (2018).
16. Tang, L. et al. Structural basis for Ca²⁺ selectivity of a voltage-gated calcium channel. *Nature* 505, 56–61 (2014).
17. Findeisen, F. & Minor, D. L. Progress in the structural understanding of voltage-gated calcium channel (Ca_v) function and modulation. *Channels* 4, 459–474 (2010).
18. Snutch, T. P. Voltage-Gated Calcium Channels. *Encycl. Neurosci.* 10, 427–441 (2009).
19. Catterall, W. A. Voltage-Gated Calcium Channels. 1–24 (2011).
20. Grabner, M., Wang, Z., Hering, S., Striessnig, J. & Glossmann, H. Transfer of 1,4-dihydropyridine sensitivity from L-type to class A (BI) calcium channels. *Neuron* 16, 207–218 (1996).

21. Domracheva, I. et al. 4-Pyridinio-1,4-Dihydropyridines as Calcium Ion Transport Modulators: Antagonist, Agonist, and Dual Action. *Oxid. Med. Cell. Longev.* 2020, (2020).
22. Xu, L. et al. Binding mechanisms of 1,4-dihydropyridine derivatives to L-type calcium channel Cav1.2: A molecular modeling study. *Mol. Biosyst.* 12, 379–390 (2016).
23. Domracheva, I. et al. 4-Pyridinio-1,4-Dihydropyridines as Calcium Ion Transport Modulators: Antagonist, Agonist, and Dual Action. *Oxid. Med. Cell. Longev.* 2020, (2020).
24. Fan, J. S., Yuan, Y. & Palade, P. FPL-64176 modifies pore properties of L-type Ca²⁺ channels. *Am. J. Physiol. - Cell Physiol.* 280, 565–572 (2001).
25. Saponara, S. et al. Functional, electrophysiological and molecular docking analysis of the modulation of Cav1.2 channels in rat vascular myocytes by murrayafoline A. *Br. J. Pharmacol.* 173, 292–304 (2016).
26. Tikhonov, D. B. & Zhorov, B. S. Structural model for dihydropyridine binding to L-type calcium channels. *J. Biol. Chem.* 284, 19006–19017 (2009).
27. P. S. d. Laplace, *Theorie Analytique des Probabilities*, 3 ed. (Gauthier- Villars, Paris).
28. Sharma, S., Kumar, P. & Chandra, R. Introduction to molecular dynamics. *Mol. Dyn. Simul. Nanocomposites using BIOVIA Mater. Stud. Lammps Gromacs* 23, 1–38 (2019)
29. Hernández, E. R. et al. *Molecular Dynamics: from basic techniques to applications (A Molecular Dynamics Primer)*. 95–123 (2008) doi:10.1063/1.3040265.
30. Hollingsworth, S. A. & Dror, R. O. *Molecular Dynamics Simulation for All*. *Neuron* 99, 1129–1143 (2018).
31. Kumar, H. & Maiti, P. K. *Introduction to Molecular Dynamics Simulation*. (2011). doi:10.1007/978-93-86279-50-7_6.
32. 5. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98, 10089–10092 (1993).
33. 6. Ding, H. Q., Karasawa, N. & Goddard, W. A. Atomic level simulations on a million particles: The cell multipole method for Coulomb and London nonbond interactions. *J. Chem. Phys.* 97, 4309–4315 (1992).
34. 7. Tironi, I. G., Sperb, R., Smith, P. E. & Van Gunsteren, W. F. A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* 102, 5451–5459 (1995).
35. 8. Fletcher, R. & Powell, M. J. D. A Rapidly Convergent Descent Method for Minimization. *Comput. J.* 6, 163–168 (2012).
36. 9. McCarthy, J. F. Block-conjugate-gradient method. *Phys. Rev. D* 40, 2149–2152 (1989).
37. 10. Verbeke, J. & Cools, R. The newton-raphson method. *Int. J. Math. Educ. Sci. Technol.* (1995). doi:10.1080/0020739950260202
38. 11. Head, J. D. & Zerner, M. C. A Broyden-Fletcher-Goldfarb-Shanno optimization

- procedure for molecular geometries. *Chem. Phys. Lett.* (1985). doi:10.1016/0009-2614(85)80574-1
39. Grubmüller, H., Heller, H., Windemuth, A. & Schulten, K. Generalized Verlet Algorithm for Efficient Molecular Dynamics Simulations with Long-range Interactions. *Mol. Simul.* 6, 121–142 (1991).
40. 13. W.F., van G. & H.J.C.A., B. A leap-frog algorithm for stochastic dynamics. *Mol.Simul.* 1, 173–185 (1998).
41. 14. Martys, N. S. & Mountain, R. D. Velocity Verlet algorithm for dissipative-particle-dynamics-based models of suspensions. *Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top.* 59, 3733–3736 (1999).
42. Salomon-Ferrer, R., Case, D. A. & Walker, R. C. An overview of the Amber biomolecular simulation package. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 3, 198–210 (2013).
43. B.R. Brooks, C.L. Brooks, A.D. MacKerell Jr, L. Nilsson, R.J. Petrella, B., Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R., Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V., Ovchinnikov, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaefer, B. Tidor, R. M. V. & H. L. Woodcock, X. Wu, W. Yang, D.M. York, and M. K. CHARMM: The Biomolecular Simulation Program. 30, 1545–1614 (2010).
44. Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* 91, 43–56 (1995).
45. Berman, H. M. et al. The Protein Data Bank Helen. *Nucleic Acids Res.* 28, 235–242 (2000).
46. 21. Krieger, E., Nabuurs, S. B. & Vriend, G. Homology Modeling. in *Structural Bioinformatics* (2005). doi:10.1002/0471721204.ch25
47. 23. Bolton, E. E., Wang, Y., Thiessen, P. A. & Bryant, S. H. Chapter 12 PubChem: Integrated Platform of Small Molecules and Biological Activities. in *Annual Reports in Computational Chemistry* (2008). doi:10.1016/S1574-1400(08)00012-1
48. 1. Roy, U. & Luck, L. A. Molecular modeling of estrogen receptor using molecular operating environment. *Biochem. Mol. Biol. Educ.* 35, 238–243 (2007).
49. 24. Hanwell, M. D. et al. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.* (2012). doi:10.1186/1758-2946-4-17

50. Castillo, O. & Aguilar, L. T. Genetic algorithms. in *Studies in Fuzziness and Soft Computing* (2019). doi:10.1007/978-3-030-03134-3_2
51. 26. Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. P. Optimization by simulated annealing. *Science* (80-.). (1983). doi:10.1126/science.220.4598.671
52. 27. Cervantes, M. The Monte Carlo Method. *Math. Sci. Eng.* (1972). doi:10.1016/S0076-5392(08)61352-1
53. Gohlke, H. & Klebe, G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angewandte Chemie - International Edition* (2002). doi:10.1002/1521-3773(20020802)41:15<2644::AID-ANIE2644>3.0.CO;2-O
54. Genheden, S. & Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* 10, 449–461 (2015).
55. Shirts, M. R., Mobley, D. L. & Brown, S. P. Free-energy calculations in structure-based drug design. in *Drug Design: Structure- and Ligand-Based Approaches* (2010).
56. Clark Still, W., Tempczyk, A., Hawley, R. C. & Hendrickson, T. Semianalytical Treatment of Solvation for Molecular Mechanics and Dynamics. *J. Am. Chem. Soc.* (1990). doi:10.1021/ja00172a038
57. Garrett M. Morris, David S. Goodsell, Michael E. Pique, William “Lindy” Lindstrom, Ruth Huey, Stefano Forli, William E. Hart, Scott Halliday, R. B. and A. J. O. AutoDock Version 4.2. Citeseer (2012).
58. Trott, O. & Olson, A. Autodock vina: improving the speed and accuracy of docking. *J. Comput. Chem.* 31, 455–461 (2010).
59. Ewing, T. J. A., Makino, S., Skillman, A. G. & Kuntz, I. D. DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases. *J. Comput. Aided. Mol. Des.* (2001). doi:10.1023/A:1011115820450
60. Wolber, G. & Langer, T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J. Chem. Inf. Model.* 45, 160–169 (2005).
61. Magrane, M. & Consortium, U. P. UniProt Knowledgebase: A hub of integrated protein data. *Database* 2011, 1–13 (2011).
62. Lomize, M. A., Pogozheva, I. D., Joo, H., Mosberg, H. I. & Lomize, A. L. OPM database and PPM web server: Resources for positioning of proteins in membranes. *Nucleic Acids Res.* 40, 370–376 (2012).
63. Wu, E. L. et al. NIH Public Access. 35, 1997–2004 (2015).

64. Case, D. A. et al. Amber 2016. Univ. California, San Fr. (2016).
65. No, P., No, M. & Tag, M. CHEMISTRY PAPER No . 12 : ORGANIC SPECTROSCOPY MODULE No . 11 : Introduction to NMR CHEMISTRY PAPER No . 12 : ORGANIC SPECTROSCOPY MODULE No . 11 : Introduction to NMR.
66. Goddard, T. D., Huang, C. C. & Ferrin, T. E. Software extensions to UCSF chimera for interactive visualization of large molecular assemblies. *Structure* 13, 473–482 (2005).
67. Sterling, T. & Irwin, J. J. ZINC 15 - Ligand Discovery for Everyone. *J. Chem. Inf. Model.* 55, 2324–2337 (2015).
68. Ravi, L. & Krishnan, K. a Handbook on Protein-Ligand Docking Tool: Autodock4. 4, 1–6 (2016).
69. Goddard, T. D., Huang, C. C. & Ferrin, T. E. Software extensions to UCSF chimera for interactive visualization of large molecular assemblies. *Structure* 13, 473–482 (2005).
70. Feng, T., Kalyaanamoorthy, S. & Barakat, K. L-Type Calcium Channels: Structure and Functions. *Ion Channels Heal. Sick.* (2018) doi:10.5772/intechopen.77305.
71. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* 42, 320–324 (2014).