POLITENICO OF TURIN

Faculty of Engineering

Course of Biomedical Engineering

Master Thesis

# Drug discovery to inhibits processes lead by $\alpha 5\beta 1$-fibronectin interaction

Computational modeling of fibronectin and a search for small moleculeinhibitors of its interactions with integrins


**Advisor:**
prof. Tuszynski Jacek Adam

**CoAdvisor:**
prof. Deriu Marco Agostino

**Candidate:**
Domiziano Doria

**Supervisor**
Prof. Maral Aminpour

July 2022

# Acknowledgment

I would like to acknowledge my professor Jack Tuszynski, for his endless support and trust, my dear supervisor Maral Aminpour for her advices, tools, teachings, to be always next to me when I needed her, plus my coadivisor prof. Deriu. Another big thanks to the entire Toronto researcher's group of CSTS Healthcare Organization for their technical support, related to any kind of analysis and all their help to perform invitro experiments.

# Abstract

Tissue integrity and hence human health are both influenced by cell attachment to the extracellular matrix. Integrins are heterodimer cell surface receptors made up of two non-covalently coupled alpha and beta subunits that primarily govern cell motility, adhesion, differentiation, migration, and proliferation by interacting with cell-cell adhesion and cell-extracellular matrix. Integrin α5β1, also known as the 'fibronectin receptor', is a heterodimer composed of α5 and β1 subunits that has emerged as a key mediator in several human carcinomas. Several forms of human malignancies, including cell proliferation, angiogenesis, and tumor spread, are intimately associated with this kind of integrin. The purpose of this work is the search for small inhibitor molecules capable of downregulating the interactions between α5β1 and fibronectin. The aim has been to target specific sites used for the protein-protein binding between fibronectin and integrin. In order to efficiently find the proper compounds, a docking procedure has been performed in order to screen and rank different molecules. Then, applying molecular dynamic simulations and methods for evaluation of chemical drug-likeness, our research has resulted in a final ranking which we can use to build a 'pharmacophore model' ready to be used for constructing a database of novel molecules or for repurposing old ones. This drug discovery method will also give a better knowledge about the required interactions inhibitor molecules should engage to correctly bind into the α5β1 sites. The outcome of this drug development process will be to cause a decrease in the spread of cancer in different physiological environments.

# Contents

# Introduction

This introduction section will give an overview about different chapters and the timeline used to perform different steps:

## Chapter 1

In this chapter we'll have a general view about the biological environment in which we are going to work, including structure interactions and molecular bonds we are interested in. Among this general line we'll also analyze the portions of the proteins we want to aim during next steps.

## Chapter 2

In this section we'll have a look into the database used for the first screening, the molecule's properties, and the way to build a new set of compounds for novel studies.

## Chapter 3

In this chapter we analyze integrins different from our main target which could potential help enhance the object of our work, we'll call them 'secondary targets'. A review of their molecular pathway will be discussed to clarify the way we want to treat them during analysis.

## Chapter 4

In this paragraph using consensus docking we make a ranking of all our molecules binding to our main target. Then, through a simple equation we rearrange the rank also considering their docking score on secondary targets.

## Chapter 5

Once we decide which molecules fits the best way for our purpose, we consider their properties in terms of biological impact on human body: to validate the final ranking in this section we show results of their Chemical absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties and molecular dynamic simulations.

## Chapter 6

In order to expand the research to novel molecules we use the best inhibitors we got to generate a pharmacophore model which included the main features and interactions we want to hold.

## Chapter 7

This work will end with the creation of a molecule's model (pharmacophore) that will be used to build a larger database for expanding the research to molecules not yet approved by FDA.

# Chapter 1

# 1 Biological environment

## 1.1 Identify molecular pathway for cancer spreading

The role of fibronectin (FN) in tumor genesis and malignant progression has been always taken as main character due to its highly presence in different cancers processes. On the other side of this molecular pathway, we find different types of integrins that contribute to progression of different oncological processes in many cancers, dependent from the specific integrin the FN is interacting with. FN play an important role in the pathobiology of cancer. However, its role in tumor genesis and malignant progression has been highly controversial: the first aspect, reported FN expression in tumor cells plays a tumor suppressive role to prevent tumor transformation, on the other hand, abundant evidence reveals that FN provokes late stages of cancer metastasis, that's the reason why there is an high risk to simply target FN for controlling cancer[1]. We'll focus our inhibition strategies on the other side of the bond, the integrin one. Cell adhesion to the extracellular matrix has important roles in tissue integrity and human health. During years have been found 18 α subunits and 8 β subunits and so far, 24 distinct types of αβ integrin heterodimers. Integrin α5β1, also known as the fibronectin receptor, has emerged as an essential mediator in many human carcinomas. Integrin α5β1 alteration is closely linked to the progression of several types of human cancers, including cell proliferation, angiogenesis, tumor metastasis, and cancerogenesis[2]. The ability of cancer cells to invade locally and further form distant metastasis is partly determined by integrin-mediated attachment to ECM. Integrin α5β1 function as a critical regulator for tumor cell migration and invasion by affecting cytoskeleton rearrangement, cell adhesion, and the production of matrix metalloproteinase (MMP). Some studies have demonstrated that integrin α5β1 enhanced keratinocyte adhesion to fibronectin and promoted invasion and metastasis via activating various signaling pathways[3].

## 1.2 Protein portion focus

Extracellular matrix molecules fibrinogen, fibronectin, and fibrillin-1 could be recognized and bound by integrin α5β1. CD97, CD87 and CD154, transmembrane proteins contain RGD peptide, have shown to interact with integrin α5β1 and induce cell adhesion, intracellular signaling, and angiogenesis.[2] These consideration lead our focus to the head portion of α5β1 region, where we looked for the RGD binding spot and any possible sites used by FN for its bond.

### 1.2.1 Material and methods

In this section we used a complex between FN and α5β1 from Protein Data Bank (PDB), code: 7NWL[4]. Using the software molecular operating environment (MOE) the molecular chaperon has been removed and after adjusting protonation state according to a physiological pH of 7 and minimizing the potential energy, we proceed to analyze the structure Fig 1.

Fig 1. *Fibronectin (chano) on the top in complex with alpha5 (orange) betha1 (pink)*

Using this crystallography structure, we manage to get a full view on the bond between the two structures and performing a Contact analysis between the proteins we obtain a complete vision about the sites of our interest.

### 1.2.2 Sites of interest

Previous analysis give us the possibility to confirm and isolate what we'll call from now on, the 'main site' of interest. This particular site have an high affinity for RGD sequence present on FN and potentially would be our main spot for inhibition Fig 2.

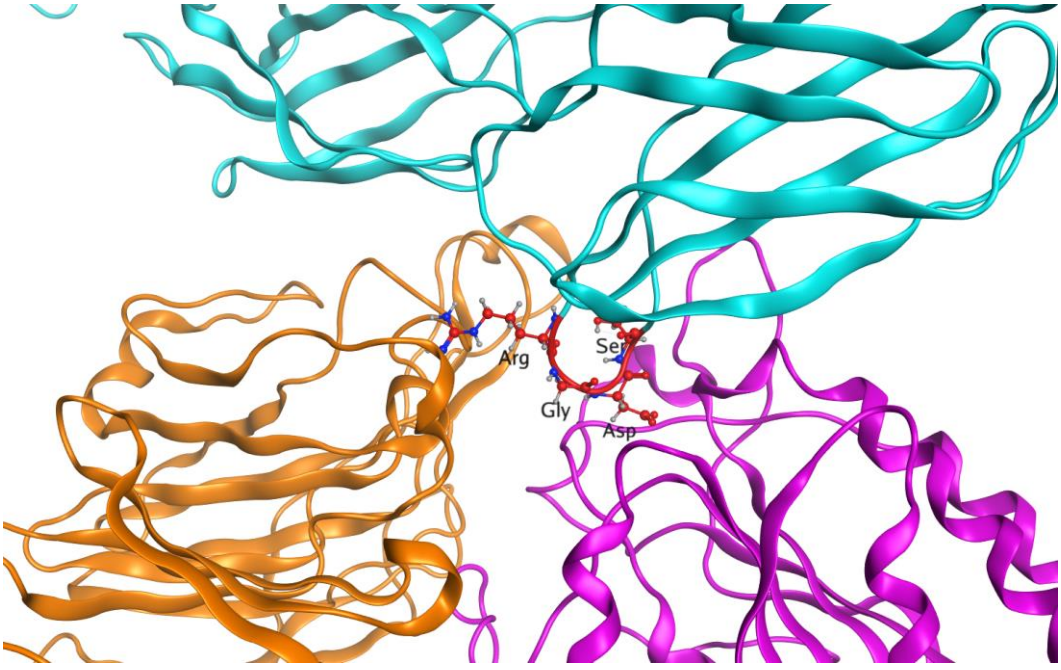



Fig2. *Focus on the spot of attach of FN (chano) into α5β1 (orange, pink), highlight in red the R (Arg), G (Gly), D (Asp) sequence.*

Even if we hold our main site, also region of interest has be defined to compute inhibition analysis in minor spots still involved in FN- α5β1 bond Fig3.

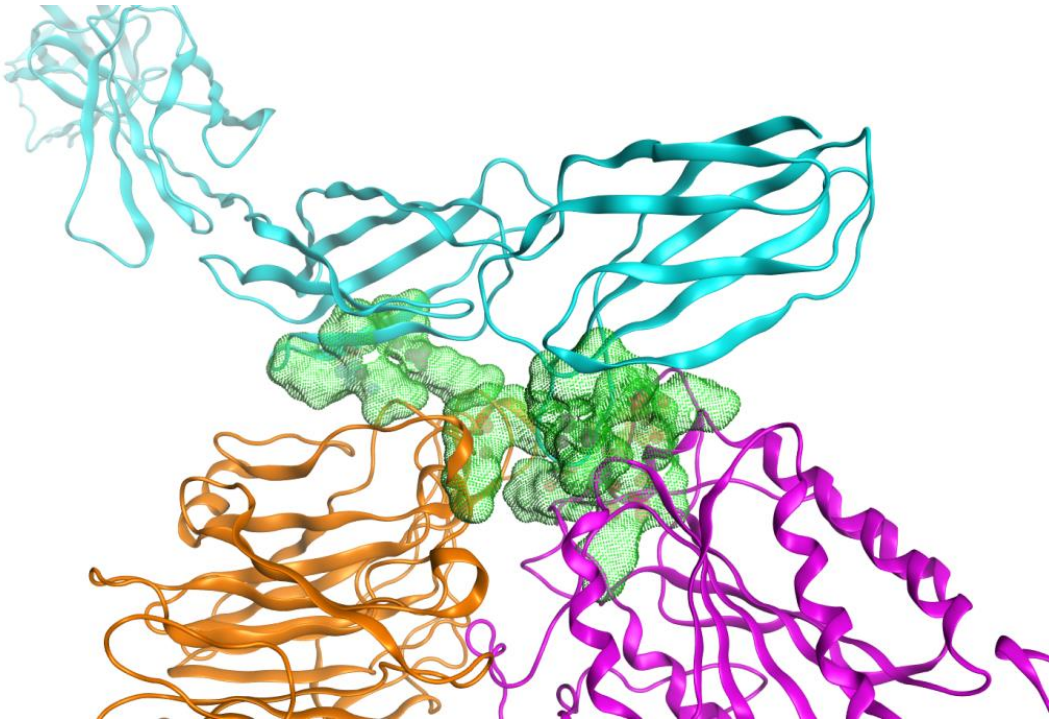


Fig3. *FN (chano) into α5β1 (orange, pink), highlight in green the areas involved in the bond bettwen the 2 structures.*

Once the region of interest has been defined, the MOE routine 'Site finder' was used to find the residues that will define different sites among all the α5β1 integrin. This procedure gives us the possibility to generate potential docking sites ranked accordingly to their Protein ligand-binding (PLB) index, a measure of that particular spot to receive a ligand and let him binding to the structure.
The algorithm has been used among all the integrin and then the sites useful to inhibit the bond between FN and α5β1 have been chosen depending on their position and their PLB index. Final sites are showed in Table 1. , we have particular interest for Site 5 (Fig 4.) that is located in RGB binding position mainly used by FN for the attach.

| Integrin | Site | Size | PLB |
|---|---|---|---|
| α5β1 | 5 | 45 | 0.76 |
| α5β1 | 9 | 36 | 0.27 |
| α5β1 | 14 | 15 | 0.12 |
| α5β1 | 21 | 57 | -0.01 |
| α5β1 | 27 | 10 | -0.27 |

Tab 1. *This table shows the sites selected for the following protein-ligand docking.*

Fig 5. *Region of Site 5 location, high affinity spot for RGD sequence of FN*

# Chapter 2

# 2 Molecular database

## 2.1 Main molecular features

For the initial screening we decide to take advantage of existing database present in oncological major and already validate by Food and Drug Administration (FDA). In particular development therapeutic program (DTP) maintains a repository of synthetic compounds and pure natural products that are available to investigators for non-clinical research purposes. The Repository collection is a uniquely diverse set of more than 200,000 compounds that have been either submitted to DTP for biological evaluation. Inside their collection we looked for a set of molecules suitable for drug delivery and having known anticancer properties.

## 2.2 Database for virtual screening

The set we selected is called Approved Oncology Drugs (AOD) X and come from DTP of Division of Cancer Treatment and Diagnosis (DCTC), Tab 2.

### 2.2.1 Material

| DRUG NAME | MW | LIG N° |
|---|---|---|
| Methotrexate | 454,44 | 1 |
| Busulfan | 246,3 | 2 |
| Thioguanine | 167,19 | 3 |
| Mercaptopurine | 152,18 | 4 |
| Mechlorethamine hydrochloride | 192,52 | 5 |
| Allopurinol | 136,11 | 6 |
| Dactinomycin | 1255,43 | 7 |
| Chlorambucil | 304,22 | 8 |
| Thiotepa | 189,22 | 9 |
| Melphalan hydrochloride | 341,66 | 10 |
| Triethylenemelamine | 204,23 | 11 |
| Altretamine | 210,28 | 12 |
| Aminolevulinic acid hydrochloride | 167,59 | 13 |
| Fluorouracil | 130,08 | 14 |
| Plicamycin | 1085,16 | 15 |
| Pipobroman | 356,06 | 16 |
| Cyclophosphamide | 261,09 | 17 |
| Mitomycin | 334,33 | 18 |
| Floxuridine | 246,19 | 19 |
| Hydroxyurea | 76,05 | 20 |
| Uracil mustard | 252,1 | 21 |
| Mitotane | 320,04 | 22 |
| Dacarbazine | 182,18 | 23 |
| Methoxsalen | 216,19 | 24 |

| | | |
|---|---|---|
| Vinblastine sulfate | 909,06 | 25 |
| Cytarabine hydrochloride | 279,7 | 26 |
| Thalidomide | 258,23 | 27 |
| Vincristine sulfate | 923,04 | 28 |
| Megestrol acetate | 384,51 | 29 |
| Trifluridine | 296,2 | 30 |
| Procarbazine hydrochloride | 257,76 | 31 |
| Lomustine | 233,7 | 32 |
| Daunorubicin hydrochloride | 563,98 | 33 |
| Streptozocin | 265,22 | 34 |
| Arsenic trioxide | 197,84 | 35 |
| Azacitidine | 244,21 | 36 |
| Cladribine | 285,69 | 37 |
| Ifosfamide | 261,09 | 38 |
| Cisplatin | 300,06 | 39 |
| Tretinoin | 300,44 | 40 |
| Teniposide | 656,66 | 41 |
| Doxorubicin hydrochloride | 579,99 | 42 |
| Bleomycin sulfate | 1512,61 | 43 |
| Paclitaxel | 853,92 | 44 |
| Decitabine | 228,21 | 45 |
| Bendamustine hydrochloride | 394,73 | 46 |
| Etoposide | 588,56 | 47 |
| Dexrazoxane | 268,27 | 48 |
| Tamoxifen citrate | 563,65 | 49 |
| Pentostatin | 268,27 | 50 |
| Sirolimus | 914,18 | 51 |
| Carboplatin | 371,25 | 52 |
| Valrubicin | 723,64 | 53 |
| Idarubicin hydrochloride | 533,96 | 54 |
| Epirubicin hydrochloride | 579,99 | 55 |
| Oxaliplatin | 397,29 | 56 |
| Mitoxantrone | 444,49 | 57 |
| Amifostine | 214,22 | 58 |
| Fludarabine phosphate | 365,21 | 59 |
| Temozolomide | 194,15 | 60 |
| Imiquimod | 240,31 | 61 |
| Carmustine | 214,05 | 62 |
| Clofarabine | 303,68 | 63 |
| Vinorelbine tartrate | 1079 | 64 |
| Topotecan hydrochloride | 457,91 | 65 |
| Gemcitabine hydrochloride | 299,65 | 66 |
| Irinotecan hydrochloride | 623,15 | 67 |
| Docetaxel | 807,89 | 68 |
| Temsirolimus | 1030,29 | 69 |
| Vorinostat | 264,32 | 70 |
| Estramustine phosphate sodium | 564,35 | 71 |
| Capecitabine | 359,35 | 72 |
| Exemestane | 296,4 | 73 |

| | | |
|---|---|---|
| Gefitinib | 446,9 | 74 |
| Erlotinib hydrochloride | 429,9 | 75 |
| Fulvestrant | 606,75 | 76 |
| Anastrozole | 293,37 | 77 |
| Letrozole | 285,3 | 78 |
| Celecoxib | 381,37 | 79 |
| Zoledronic acid | 272,09 | 80 |
| Dasatinib | 488,01 | 81 |
| Everolimus | 958,24 | 82 |
| Pazopanib hydrochloride | 473,98 | 83 |
| Selumetinib | 457,68 | 84 |
| Imatinib | 493,61 | 85 |
| Lapatinib | 581,06 | 86 |
| Nilotinib | 529,51 | 87 |
| Sorafenib | 464,82 | 88 |
| Lenalidomide | 259,26 | 89 |
| Ixabepilone | 506,7 | 90 |
| Raloxifene | 473,59 | 91 |
| Abiraterone | 349,51 | 92 |
| Sunitinib | 398,47 | 93 |
| Afatinib | 485,94 | 94 |
| Olaparib | 434,46 | 95 |
| Romidepsin | 540,69 | 96 |
| Pralatrexate | 477,48 | 97 |
| Niraparib hydrochloride | 356,85 | 98 |
| Pemetrexed, Disodium salt, Heptahydrate | 471,38 | 99 |
| Enzalutamide | 464,42 | 100 |
| Lenvatinib | 426,86 | 101 |
| Nelarabine | 297,27 | 102 |
| Vismodegib | 421,3 | 103 |
| Rucaparib phosphate | 421,36 | 104 |
| Crizotinib | 450,34 | 105 |
| Bortezomib | 384,24 | 106 |
| Neratinib | 557,05 | 107 |
| Axitinib | 386,47 | 108 |
| Trametinib | 615,4 | 109 |
| Palbociclib | 447,54 | 110 |
| Carfilzomib | 719,92 | 111 |
| Omacetaxine mepesuccinate | 545,63 | 112 |
| Ixazomib citrate | 517,13 | 113 |
| Ponatinib | 532,55 | 114 |
| Belinostat | 318,35 | 115 |
| Idelalisib | 415,42 | 116 |
| Vandetanib | 475,36 | 117 |
| Cabozantinib | 501,51 | 118 |
| Panobinostat | 349,43 | 119 |
| Erismodegib | 485,49 | 120 |
| Plerixafor | 502,79 | 121 |
| Vemurafenib | 489,91 | 122 |

| | | |
|---|---|---|
| Cabazitaxel | 835,94 | 123 |
| Ibrutinib | 440,5 | 124 |
| Ruxolitinib | 306,36 | 125 |
| Regorafenib | 482,8 | 126 |
| Alectinib | 482,62 | 127 |
| Binimetinib | 441,22 | 128 |
| Dabrafenib mesylate | 615,65 | 129 |
| ARRY-380 | 480,52 | 130 |
| Bosutinib | 530,45 | 131 |
| Dacomitinib | 469,94 | 132 |
| Alpelisib | 441,47 | 133 |
| Venetoclax | 868,45 | 134 |
| Talazoparib | 380,35 | 135 |
| Fedratinib | 524,67 | 136 |
| Cobimetinib | 531,32 | 137 |
| Abemaciclib | 506,59 | 138 |
| Apalutamide | 477,42 | 139 |
| Duvelisib | 416,87 | 140 |
| Entrectinib | 560,63 | 141 |
| Pomalidomide | 273,25 | 142 |
| Glasdegib | 374,43 | 143 |
| Ceritinib | 558,14 | 144 |
| Tazemetostat | 572,73 | 145 |
| Capmatinib | 412,41 | 146 |
| Encorafenib | 540,01 | 147 |
| Ribociclib | 434,54 | 148 |
| Osimertinib | 499,61 | 149 |
| Lorlatinib | 406,41 | 150 |
| Selinexor | 443,3 | 151 |
| Erdafitinib | 446,54 | 152 |
| Larotrectinib | 428,43 | 153 |
| Brigatinib | 584,1 | 154 |
| Gilteritinib | 552,71 | 155 |
| Enasidenib | 473,35 | 156 |
| Uridine triacetate | 370,32 | 157 |
| Ivosidenib | 582,96 | 158 |
| Pexidartinib | 417,81 | 159 |
| Acalabrutinib | 465,41 | 160 |
| Avapritinib | 498,55 | 161 |
| Copanlisib tris-HCl | 589,9 | 162 |
| Pemigatinib | 487,49 | 163 |
| Selpercatinib | 525,6 | 164 |
| Zanubrutinib | 471,55 | 165 |
| Darolutamide | 398,84 | 166 |

Tab 2. *This table shows the compounds set used for virtual screening, molecule's name and molecular weight.*

# Chapter 3

# 3 Secondary targets

## 3.1 Different integrins involved in cancer spreading

During our investigation we start considering the role of other integrins and how our ligand's kinetics could be influenced by them. In particular some of them composed by alpha + beta units could show an high affinity for the compounds we are going to test; that's the reason why we include other kind of integrins in our simulations. Basing on crystallography structures available and useful features for our study we selected 3 more integrins. For every new considered proteins, we performed the 'Site finder' algorithm using MOE software. Detect candidate protein-ligand and protein-protein binding sites has been found using a fast geometric algorithm based on Edelsbrunner's Alpha Shapes. Each site on a macromolecular structure is ranked according to its PLB (propensity for ligand binding) index.

## 3.2 Alpha1beta1

The integrin α1β1 is known to plays an important role in fibroblast proliferation. In the small intestine, the single α1 subunit is present in the crypt proliferative compartment and absent in the villus. Recently has been shown that the integrin α1 protein and transcript (*ITGA1*) are highly present in colorectal cancers (CRC). The study[5] finally postulated that the integrin α1β1 has a pro-tumoral contribution to CRC and all analysis suggest how α1β1 is involved in tumor cell proliferation, survival and migration.

### 3.2.1 Materials and methods

The crystallography structure of the protein has been gathered from PDB , code: 1PT6[6]. Once it has been adjusted in terms of protonation state and energy minimization through MOE software, the 'Site Finder' routine has been performed on the entire integrin to generate a list a suitable sites for our compounds, then the best sites has been selected considering their position on the protein and their PLB, Tab 3.

| Integrin | Site | Size | PLB |
|---|---|---|---|
| α1β1 | 1 | 19 | 2.43 |
| α1β1 | 2 | 19 | 1.08 |
| α1β1 | 3 | 13 | 0.28 |

Tab 3. *List on sites compute on α1β1*

## 3.3 Alpha2beta1

The α2β1 integrin is a major collagen receptor that is widely expressed and known to promote cell migration and control tissue homeostasis. Growing evidence suggests that it can be a key pathway in cancer. Recent studies have shown that α2β1 integrin is a regulator of cancer metastasis either by promoting or inhibiting the dissemination

process of cancer cells. The α2β1 integrin signaling can also enhance tumor angiogenesis. In addition, α2β1 integrin has been associated with cancer stem cells.[7]

### 3.3.1 Materials and methods

The crystallography structure of the protein has been gathered from PDB , code: 1AOX[8]. Protonation state and energy minimization through MOE software has been performed, then the 'Site Finder' routine has been used on the entire integrin to generate a list a suitable sites for our compounds, then the best sites has been selected considering their position on the protein and their PLB, Tab 4.

| Integrin | Site | Size | PLB |
|---|---|---|---|
| α2β1 | 1 | 45 | 2.19 |
| α2β1 | 2 | 35 | 1.97 |
| α2β1 | 3 | 14 | 1.54 |
| α2β1 | 4 | 14 | 1.40 |
| α2β1 | 5 | 22 | 1.06 |
| α2β1 | 6 | 18 | 0.61 |

Tab 4. *List on sites compute on α2β1*

## 3.4 AlphaVbeta3

Among all integrin αvβ3 seems to have a high affinity for RGD sequence as well as our main target α5β1 [9]. Integrin αvβ3 is highly expressed on activated endothelial cells, new-born vessels as well as some tumor cells, but is not present in resting endothelial cells and most normal organ systems, making it a suitable target for anti-angiogenic therapy. These two features in combination make this integrin an important secondary target to take in consideration [10].

### 3.4.1 Materials and methods

The crystallography structure comes from PDB , code: 1L5G[11]. Protonation state and energy minimization through MOE software has been performed, then the 'Site Finder' routine has been used on the entire integrin to generate a list a suitable sites for our compounds, then the best sites has been selected considering their position on the protein and their PLB, Tab 5.

| Integrin | Site | Size | PLB |
|---|---|---|---|
| αVβ3 | 1 | 740 | 7.63 |
| αVβ3 | 2 | 200 | 2.92 |
| αVβ3 | 3 | 92 | 0.89 |
| αVβ3 | 4 | 94 | 0.81 |
| αVβ3 | 5 | 66 | 0.72 |
| αVβ3 | 6 | 87 | 0.72 |
| αVβ3 | 7 | 89 | 0.7 |

Tab 5. *List on sites compute on αVβ3*

# Chapter 4

# 4 Database filtering criteria

## 4.1 Docking procedure

Once we define our target proteins and the Database for virtual screening, we proceed to define docking strategy. In order to get the most realist prediction possible we used consensus docking strategy instead of single docking one.
The consensus docking approaches demonstrated to improve the quality of docking and virtual screening results compared to the single docking methods. From a qualitative point of view, the improvement in pose prediction accuracy was obtained by prioritizing ligand binding poses produced by a high number of docking methods, whereas with regards to virtual screening studies, high hit rates were obtained by prioritizing the compounds showing a high level of pose consensus.
Consensus scoring is a method whereby the binding affinities of compounds for a particular target are predicted by using more than one scoring algorithm. Several different studies have found that this is superior in accuracy to using a single scoring algorithm alone. For example, Kukol investigated various consensus algorithms and found that, for a given docking pose, a simple combination of AutoDock and Vina scores gave the most consistent performance that showed early enrichment of known ligands for all receptor targets investigated.[12] Chang et al. performed extensive comparisons of Autodock and Vina docking results using the DUD decoy database[13] and, like Kukol, found that taking both of their scores into account improved the overall binding affinity predictions.[14] Cheng et al. compared a number of different standalone scoring algorithms and showed that a combination of them almost always out performed even the best (although the precise nature of the combinations were found to vary depending on the target).[15]
In our case we used Vina, Autodock and Dock6 for consensus docking performing all simulation on Graham cluster owned by Compute Canada association.
Algorithm used to compare data coming from different docking programs, to perform consensus, come from DockBox: a python wrapper library designed to facilitate the use of standard docking programs either alone or in combination.
All results have been ranked using Vina rescoring method.
Finally, we used a code to extract the best poses from every sites on every protein, doing that we find out the compound that best fit in every single site associated with his specify affinity value.

### 4.1.1 Results

Once docking procedure has been performed, we obtained different affinity values for every sites among all protein: we hold for our future analysis the lower energy state among all sites for every integrin, except for the α5β1, in that case we maintain the affinity for the RGD specific site Tab 6.

| LIG | Alpha5beta1 | Alpha1beta1 | Alpha2Beta1 | AlphaVbeta3 |
|---|---|---|---|---|
| 1 | -6,515 | -6,6 | -6,374 | -7,836 |
| 2 | -5,543 | -5,11 | -4,771 | -6,22 |
| 3 | -5,254 | -5,48 | -4,457 | -4,576 |
| 4 | -3 | -3,86 | -2,5 | -4,542 |

| | | | | |
|---|---|---|---|---|
| 5 | -3,4 | -4,391 | -4,235 | -4,43 |
| 6 | -5,72 | -4,04 | -4,185 | -4,535 |
| 7 | -8,249 | -4 | -4,38 | -11,159 |
| 8 | -5,496 | -5,3 | -5,342 | -5,6 |
| 9 | -3,111 | -4,122 | -4,062 | -4,969 |
| 10 | -6,2 | -5,42 | -4,985 | -6,096 |
| 11 | -5,03 | -4,733 | -4,745 | -5,283 |
| 12 | -9,81 | -4,647 | -4,92 | -3,7 |
| 13 | -3,9 | -5,77 | -5,29 | -4,731 |
| 14 | -4,365 | -4,34 | -3,4 | -3,941 |
| 15 | -11,8 | -15,65 | -9,523 | -8 |
| 16 | -5,14 | -5,222 | -5,404 | -5,481 |
| 17 | -5,823 | -4,5 | -3,16 | -4,2 |
| 18 | -5,451 | -6,1 | -5,162 | -6,743 |
| 19 | -6,444 | -6,15 | -5,096 | -5,337 |
| 20 | -3,88 | -3,78 | -3,533 | -4,227 |
| 21 | -5,189 | -5,077 | -4,765 | -4,721 |
| 22 | -5,6 | -5,83 | -4,311 | -5,311 |
| 23 | -4,523 | -4 | -4,639 | -5,043 |
| 24 | -4,451 | -5,32 | -4,12 | -5,104 |
| 25 | -6,869 | -7,91 | -6,233 | -7,354 |
| 26 | -6,263 | -4,59 | -4,785 | -5,59 |
| 27 | -4,927 | -5,72 | -4,76 | -5,59 |
| 28 | -7,415 | -7,5 | -6,356 | -8,37 |
| 29 | -6,445 | -5,912 | -4,99 | -6,849 |
| 30 | -5,341 | -5,59 | -5,09 | -5,816 |
| 31 | -4,41 | -5,77 | -5,288 | -5,906 |
| 32 | -4,44 | -5,28 | -4,884 | -5,564 |
| 33 | -5,567 | -7,2 | -5,884 | -7,065 |
| 34 | -4,939 | -5,7 | -5,463 | -5,775 |
| 36 | -3,05 | -6,61 | -4,937 | -5,555 |
| 37 | -5,8 | -5,076 | -4,95 | -6,087 |
| 38 | -4,678 | -3,5 | -4,667 | -5,329 |
| 40 | -6,252 | -5,4 | -5,72 | -7,01 |
| 41 | -7,891 | -7,235 | -6,44 | -8,698 |
| 42 | -6,408 | -6,7 | -6,274 | -7,481 |
| 44 | -6,855 | -8,1 | -6,512 | -8,772 |
| 45 | -4,17 | -6,44 | -4,66 | -5,923 |
| 46 | -3,52 | -5,37 | -5,7 | -6,499 |
| 47 | -9,2 | -7,25 | -7,571 | -8,766 |
| 48 | -5,278 | -4,8 | -7,3 | -5,827 |
| 49 | -5,4 | -6,3 | -8 | -6,791 |
| 50 | -4,1 | -5,9 | -4,3 | -5,645 |
| 51 | -8,014 | -7,7 | -7,968 | -11,767 |
| 53 | -7,077 | -6,97 | -7,21 | -8,791 |
| 54 | -6,547 | -9,16 | -6,744 | -6,92 |
| 55 | -6,502 | -6,813 | -7,156 | -7,904 |
| 57 | -6,947 | -5,3 | -6,115 | -7,545 |
| 58 | -5,894 | -6,6 | -4,853 | -5,442 |
| 59 | -4,17 | -6,6 | -5,504 | -6,639 |
| 60 | -4,5 | -5,34 | -4,58 | -5,148 |

| | | | | |
|---|---|---|---|---|
| 61 | -4,92 | -4,83 | -4,636 | -5,373 |
| 62 | -5,29 | -4,99 | -4,509 | -5,2 |
| 63 | -5,2 | -6 | -7,74 | -5,516 |
| 64 | -5,66 | -6,859 | -6,76 | -7,53 |
| 65 | -9,04 | -6,5 | -5,5 | -7,178 |
| 66 | -4,4 | -5,41 | -5,92 | -5,007 |
| 67 | -8,6 | -8,2 | -7,14 | -8,98 |
| 68 | -6,722 | -7,229 | -9,2 | -9,34 |
| 69 | -2,93 | -8,3 | -7,848 | -8,495 |
| 70 | -6,338 | -6,03 | -5,185 | -6,22 |
| 71 | -8,84 | -6,85 | -6,91 | -7,567 |
| 72 | -6,843 | -5,8 | -5,584 | -6,869 |
| 73 | -4,593 | -6,65 | -5,134 | -5,202 |
| 74 | -8,7 | -5,9 | -5,814 | -8,198 |
| 75 | -6,101 | -5,3 | -5,808 | -7,722 |
| 76 | -6,775 | -5,2 | -9,2 | -8,873 |
| 77 | -5,499 | -5,6 | -5,396 | -6,337 |
| 78 | -5,38 | -5,828 | -4,935 | -5,564 |
| 79 | -5,125 | -6,21 | -5,86 | -6,694 |
| 80 | -4,477 | -4,706 | -4,824 | -5,501 |
| 81 | -6,665 | -6,7 | -5,3 | -7,729 |
| 82 | -7,7 | -7,718 | -10,6 | -3,9 |
| 83 | -7,01 | -9,22 | -5,84 | -8,369 |
| 84 | -5,841 | -6 | -5,459 | -6,801 |
| 85 | -6,979 | -7,273 | -6,57 | -12,62 |
| 86 | -9,5 | -9,67 | -6,904 | -8,422 |
| 87 | -8,33 | -7,9 | -9,4 | -8,577 |
| 88 | -4,98 | -8,02 | -6,132 | -7,589 |
| 89 | -4,9 | -6,5 | -5,138 | -5,254 |
| 90 | -5,749 | -6,402 | -6,192 | -7,366 |
| 91 | -6,014 | -7,84 | -5,899 | -6,765 |
| 92 | -5,406 | -5,395 | -5,247 | -6,928 |
| 93 | -7,63 | -7,3 | -5,752 | -6,286 |
| 94 | -6,995 | -7,66 | -6,062 | -7,904 |
| 95 | -6,023 | -7,65 | -6,357 | -7,637 |
| 96 | -6,439 | -5,685 | -6,085 | -7,367 |
| 97 | -6,962 | -7,1 | -6,241 | -7,593 |
| 98 | -5,144 | -7,38 | -5,148 | -5,739 |
| 99 | -6,875 | -6,3 | -6,42 | -7,698 |
| 100 | -5,838 | -6,8 | -6,159 | -7,06 |
| 101 | -9,31 | -6,635 | -6,247 | -7,543 |
| 102 | -7,67 | -6,02 | -5,238 | -5,991 |
| 103 | -7,13 | -8,45 | -5,54 | -6,866 |
| 104 | -7,69 | -5,904 | -5,264 | -5,841 |
| 105 | -7,7 | -6,388 | -5,774 | -7,672 |
| 107 | -9,4 | -6,3 | -2,91 | -8,399 |
| 108 | -8,15 | -6,9 | -6,027 | -6,916 |
| 109 | -7,08 | -8,46 | -6,515 | -7,151 |
| 110 | -6,002 | -7,08 | -6,422 | -7,643 |
| 111 | -7,722 | -7,2 | -7,208 | -8,76 |
| 112 | -5,11 | -5,8 | -5,929 | -10,5 |

| | | | | |
|---|---|---|---|---|
| 114 | -10,6 | -7,6 | -7,6 | -8,466 |
| 115 | -6,6 | -6,82 | -5,8 | -6,648 |
| 116 | -5,615 | -7,62 | -5,979 | -6,147 |
| 117 | -5,58 | -6,65 | -6,235 | -7,806 |
| 118 | -6,42 | -8,17 | -6,655 | -8,12 |
| 119 | -7,025 | -7,54 | -5,715 | -5,507 |
| 120 | -9,6 | -7,8 | -6,442 | -8,674 |
| 121 | -10,55 | -7,5 | -6,025 | -8,076 |
| 122 | -7,285 | -7,18 | -6,248 | -9,6 |
| 123 | -7,943 | -7,3 | -5,761 | -7,801 |
| 124 | -9,25 | -6,58 | -5,43 | -7,864 |
| 125 | -6,071 | -6,182 | -4,44 | -6,28 |
| 126 | -6,444 | -8,01 | -6,1 | -7,452 |
| 127 | -6,352 | -6,655 | -6,4 | -7,13 |
| 128 | -6,329 | -7,63 | -5,4 | -5,947 |
| 129 | -6,269 | -9,92 | -5,65 | -7,024 |
| 130 | -6,963 | -8,2 | -9,4 | -7,991 |
| 131 | -6,595 | -6,96 | -4,43 | -8,7 |
| 132 | -6,721 | -6,5 | -6,4 | -8,131 |
| 133 | -6,395 | -6,071 | -2,72 | -6,838 |
| 134 | -8,766 | -10,64 | -7,502 | -11,098 |
| 135 | -7,26 | -5,22 | -3,87 | -5,15 |
| 136 | -6,25 | -7,46 | -7,2 | -10,55 |
| 137 | -5,3 | -6,214 | -6,8 | -7,6 |
| 138 | -7,24 | -7,08 | -6,4 | -7,756 |
| 139 | -5,64 | -8 | -4,96 | -6,464 |
| 140 | -6,41 | -7,67 | -5,03 | -6,663 |
| 141 | -11,52 | -7,6 | -8 | -8,345 |
| 142 | -7,75 | -5,93 | -4,8 | -5,3 |
| 143 | -10,02 | -6,3 | -6,97 | -8,07 |
| 144 | -6,622 | -6,3 | -6,5 | -8,697 |
| 145 | -7,034 | -8,52 | -7,14 | -9,045 |
| 146 | -5,886 | -8,2 | -5,944 | -7,145 |
| 147 | -9,5 | -8,98 | -6,513 | -8,217 |
| 148 | -11,04 | -6,558 | -5,578 | -7,9 |
| 149 | -7,22 | -7,36 | -6,263 | -8,465 |
| 150 | -5,407 | -5,494 | -4,8 | -7,156 |
| 151 | -8,9 | -6,05 | -5,826 | -7,614 |
| 152 | -6,203 | -6,1 | -6,836 | -8,1 |
| 153 | -5,987 | -6,26 | -4,78 | -7,37 |
| 154 | -7,278 | -8,49 | -6,695 | -8,58 |
| 155 | -7,351 | -6,7 | -6,737 | -8,202 |
| 156 | -7,13 | -6,557 | -6,142 | -7,708 |
| 157 | -5,605 | -5,822 | -6,017 | -6,255 |
| 158 | -13,49 | -7,64 | -6,057 | -7,058 |
| 159 | -6,88 | -6,41 | -6,407 | -6,504 |
| 160 | -5,75 | -6,31 | -6,38 | -7,094 |
| 161 | -6,293 | -9,03 | -5,735 | -8,431 |
| 162 | -9,7 | -6,89 | -6,139 | -8,84 |
| 163 | -7,089 | -8,4 | -6,701 | -9,121 |
| 164 | -8,6 | -7,566 | -6,727 | -10,8 |

| | | | | |
|---|---|---|---|---|
| 165 | -7,052 | -7,03 | -6,329 | -7,288 |
| 166 | -8,5 | -6,08 | -6,295 | -9 |

Tab 6. *Affinity values associated to their ligands corresponding to the site with the best affinity among all protein. (kcal/mol measure unit)*

## 4.2 Optimization algorithm

Having all the docking values we proceed to process these results, comparing the different energy states to build a final ranking. We made some consideration to build a valid equation to rank compounds: first, our main target (α5β1) should our main objective, so we included a mandatory feature for final filtering: affinity value relative to α5β1 need to be higher than all the others; second, even if our secondary targets have a positive impact on cancer spreading if they receive some quantity of inhibition molecules, we don't want them to have the same impact of our main target. These are the reasons why our final rank has been build using an equation that associate to secondary targets a different impact on final result:

$$Impact = |\alpha5\beta1| + \frac{|(\alpha1\beta1 + \alpha2\beta1 + \alpha V\beta3)|}{3}$$

This simple formula decreases the influence of secondary target leading to the final ranking Tab 7.

| LIG | Alpha5beta1 | Alpha1beta1 | Alpha2Beta1 | AlphaVbeta3 | Impact | Name |
|---|---|---|---|---|---|---|
| 158 | -13,49 | -7,64 | -6,057 | -7,058 | 20,4083 | Ivosedinib |
| 141 | -11,52 | -7,6 | -8 | -8,345 | 19,5017 | Entrectinib |
| 114 | -10,6 | -7,6 | -7,6 | -8,466 | 18,4887 | Ponatinib |
| 121 | -10,55 | -7,5 | -6,025 | -8,076 | 17,7503 | Plerixafor |
| 148 | -11,04 | -6,558 | -5,578 | -7,9 | 17,7187 | Ribociclib |
| 147 | -9,5 | -8,98 | -6,513 | -8,217 | 17,4033 | Encorafenib |
| 120 | -9,6 | -7,8 | -6,442 | -8,674 | 17,2387 | Erismodegib |
| 143 | -10,02 | -6,3 | -6,97 | -8,07 | 17,1333 | Glasdegib |
| 47 | -9,2 | -7,25 | -7,571 | -8,766 | 17,0623 | Etoposide |
| 162 | -9,7 | -6,89 | -6,139 | -8,84 | 16,9897 | Copanlisib tris-HCl |
| 101 | -9,31 | -6,635 | -6,247 | -7,543 | 16,1183 | Lenvatinib |
| 71 | -8,84 | -6,85 | -6,91 | -7,567 | 15,949 | Estramustine phospate sodium |
| 124 | -9,25 | -6,58 | -5,43 | -7,864 | 15,8747 | Ibrutinib |

Tab 7. *Top ranking of compounds based of their final impact on cancer spreading.*

# Chapter 5

# 5 Validate best molecules

## 5.1 ADMET properties

Before proceeding with laboratory test which will physical test some of the molecules, we performed some analysis to guarantee their biological safety features and provide a solid model about their solubility and biological diffusion among tissues. It is often reported that the failure to meet requisite ADMET criteria are a common cause for the high attrition rates of drug candidates[16]. Early ADMET profiling is indeed desirable so as to mitigate the risk of attrition. Various medium and high-throughput in vitro ADMET screens have therefore been developed, that have contributed to the available experimental data. These are nonetheless quite expensive especially when thousands of compounds are involved. Furthermore, reducing animal testing has now become a priority.

With the aim of facilitating rapid and inexpensive means of ADMET profiling, various in silico tools have been developed[17]. Other efforts have made use of ADMET predictions to evaluate drug-likeness of a compound [18]. While some of the models are available as part of commercial software packages based on proprietary datasets, there has been a significant push for open source software and web services[19,20,21,22,23,24].

### 5.1.1 Swiss ADME

To be effective as a drug, a potential molecule must reach its target in the body in sufficient concentration and stay there in a bioactive form long enough for the expected biologic events to occur. Drug development involves assessment of absorption, distribution, metabolism and excretion (ADME) increasingly earlier in the discovery process, at a stage when considered compounds are numerous but access to the physical samples is limited. In that context, computer models constitute valid alternatives to experiments. Here, we present the new SwissADME web tool that gives free access to a pool of fast yet robust predictive models for physicochemical properties, pharmacokinetics, drug-likeness and medicinal chemistry friendliness, among which in-house proficient methods such as the BOILED-Egg, iLOGP and Bioavailability Radar. Easy efficient input and interpretation are ensured thanks to a user-friendly interface through the login-free website http://www.swissadme.ch. Specialists, but also nonexpert in cheminformatics or computational chemistry can predict rapidly key parameters for a collection of molecules to support their drug discovery endeavours.[25] The predictions are based on a combination of fragmental methods (for solubility), as well as machine-learning based binary classification methods for other ADMET properties (cytochrome-P450 inhibitor, P-glycoprotein substrate)[25].

### 5.1.2 Swiss ADME Results

In this section we explore the results of the best ten compounds, coming from docking analysis. First, we represent all molecules inside a chart to easy understand how they move among different biological barriers Fig 6.
This representation use two physicochemical descriptors, the WLOGP on y axis and Topological Polar Surface Area (TPSA) on x one, the first one represent the lipophilicity and the second the apparent polarity of different molecules. Using these features we are able to allocate the drugs depending on their ability to pass different barriers: the molecules located in the white region are predicted to be passively absorbed by the gastrointestinal tract; the one inside the orange portion passively absorbed by Blood-Brain Barrier.

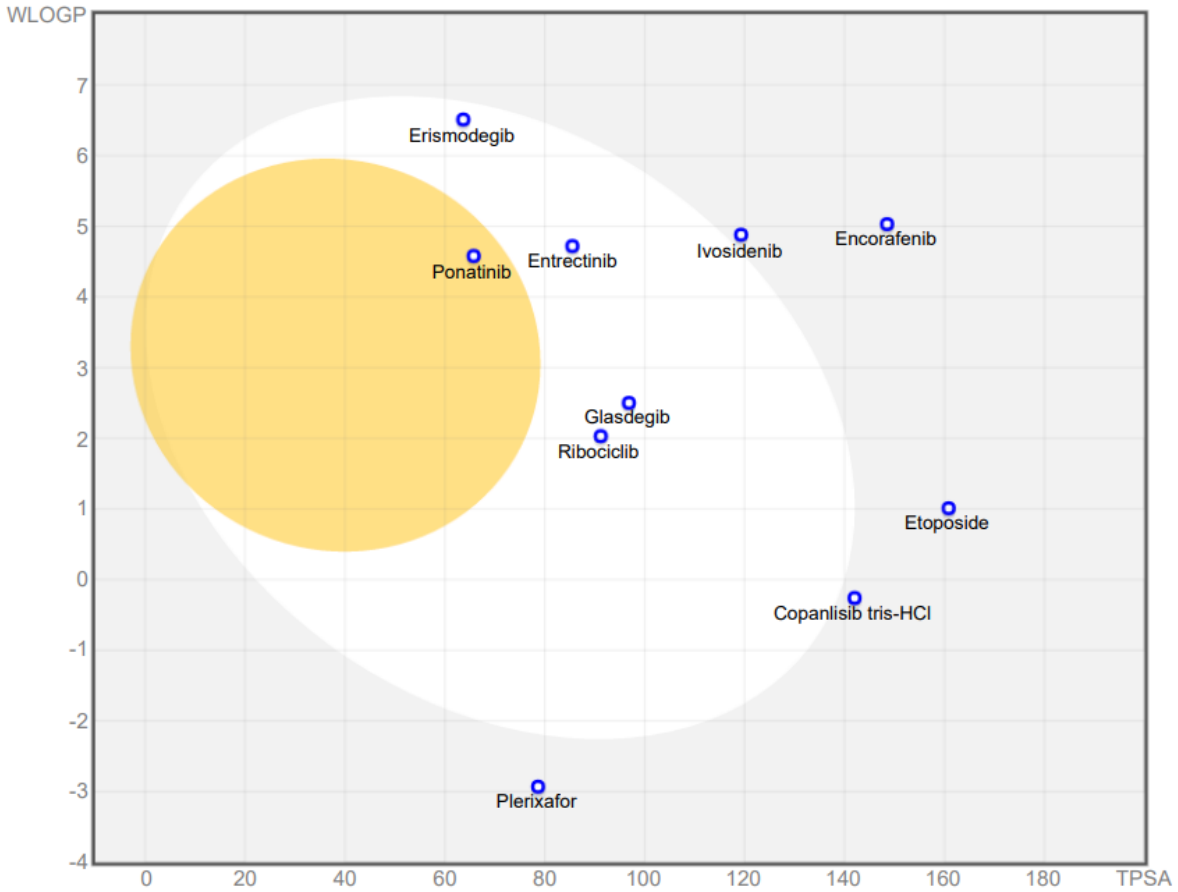



Fig.6 *Boiled-egg representation*

After that general overview, we proceed to evaluate specific features for every single molecule. Fig. 7 to Fig. 16

That include:

Physicochemical properties, simple molecular and physicochemical descriptors like molecular weight (MW), molecular refractivity (MR), count of specific atom types and polar surface area (PSA). The PSA is calculated using the fragmental technique called topological polar surface area (TPSA), considering sulfur and phosphorus as polar atoms.
Lipholiphcity section, the partition coefficient between n-octanol and water (log Po/w) is the classical descriptor for Lipophilicity. It has a dedicated section in SwissADME due to the critical importance of this physicochemical property for pharmacokinetics drug discovery. Many computational methods for

log Po/w estimation were developed with diverse performance on various chemical sets. Common practice is to use multiple predictors either to select the most accurate methods for a given chemical series or to generate consensus estimation. The models behind the predictors should be as diverse as possible to increase the prediction accuracy through consensus log Po/w.

Water solubility, having a soluble molecule greatly facilitates many drug development activities, primarily the ease of handling and formulation. Moreover, for discovery projects targeting oral administration, solubility is one major property influencing absorption. As well, a drug meant for parenteral usage has to be highly soluble in water to deliver a sufficient quantity of active ingredient in the small volume of such pharmaceutical dosage. All predicted values are the decimal logarithm of the molar solubility in water (log S). SwissADME also provides solubility in mol/l and mg/ml along with qualitative solubility classes.

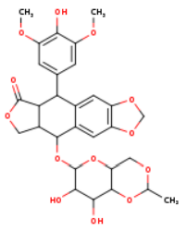
Pharpacocinetics, evaluate individual ADME behaviors of the molecule under investigation.

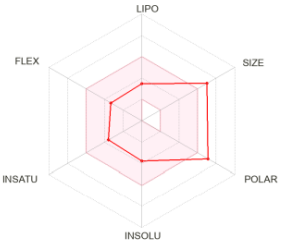
One model is a multiple linear regression, which aims at predicting the skin permeability coefficient (Kp). The more negative the log Kp (with Kp in cm/s), the less skin permeant is the molecule. Other binary classification models are included, which focus on the propensity for a given small molecule to be substrate or inhibitor of proteins governing important pharmacokinetic behaviors.

Druglikness, assesses qualitatively the chance for a molecule to become an oral drug with respect to bioavailability. Drug-likeness was established from structural or physicochemical inspections of development compounds advanced enough to be considered oral drug-candidates. This notion is routinely employed to perform filtering of chemical libraries to exclude molecules with properties most probably incompatible with an acceptable pharmacokinetics profile. This SwissADME section gives access to five different rule-based filters, with diverse ranges of properties inside of which the molecule is defined as drug-like.

Medicinal chemistry, the purpose of this section is to support medicinal chemists in their daily drug discovery endeavors.[25]

## Etoposide



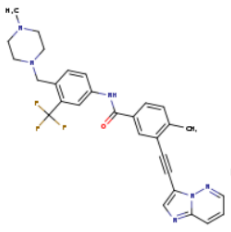


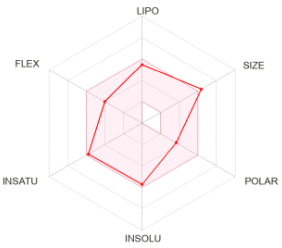

SMILES COc1cc(cc(c1O)OC)C1C2C(=O)OCC2C(c2c1cc1OCOc1c2)OC1OC2COC(OC2C(C1O)O)C

| Physicochemical Properties | |
|---|---|
| Formula | C29H32O13 |
| Molecular weight | 588.56 g/mol |
| Num. heavy atoms | 42 |
| Num. arom. heavy atoms | 12 |
| Fraction Csp3 | 0.55 |
| Num. rotatable bonds | 5 |
| Num. H-bond acceptors | 13 |
| Num. H-bond donors | 3 |
| Molar Refractivity | 139.11 |
| TPSA | 160.83 Å² |
| **Lipophilicity** | |
| Log $P_{o/w}$ (iLOGP) | 3.31 |
| Log $P_{o/w}$ (XLOGP3) | 0.60 |
| Log $P_{o/w}$ (WLOGP) | 1.01 |
| Log $P_{o/w}$ (MLOGP) | -0.14 |
| Log $P_{o/w}$ (SILICOS-IT) | 0.95 |
| Consensus Log $P_{o/w}$ | 1.15 |

| Water Solubility | |
|---|---|
| Log *S* (ESOL) | -3.75 |
| Solubility | 1.05e-01 mg/ml ; 1.78e-04 mol/l |
| Class | Soluble |
| Log *S* (Ali) | -3.55 |
| Solubility | 1.65e-01 mg/ml ; 2.81e-04 mol/l |
| Class | Soluble |
| Log *S* (SILICOS-IT) | -3.18 |
| Solubility | 3.85e-01 mg/ml ; 6.55e-04 mol/l |
| Class | Soluble |
| **Pharmacokinetics** | |
| GI absorption | Low |
| BBB permeant | No |
| P-gp substrate | Yes |
| CYP1A2 inhibitor | No |
| CYP2C19 inhibitor | No |
| CYP2C9 inhibitor | No |
| CYP2D6 inhibitor | Yes |
| CYP3A4 inhibitor | No |
| Log $K_p$ (skin permeation) | -9.46 cm/s |
| **Druglikeness** | |
| Lipinski | No; 2 violations: MW>500, NorO>10 |
| Ghose | No; 3 violations: MW>480, MR>130, #atoms>70 |
| Veber | No; 1 violation: TPSA>140 |
| Egan | No; 1 violation: TPSA>131.6 |
| Muegge | No; 2 violations: TPSA>150, H-acc>10 |
| Bioavailability Score | 0.17 |
| **Medicinal Chemistry** | |
| PAINS | 0 alert |
| Brenk | 0 alert |
| Leadlikeness | No; 1 violation: MW>350 |
| Synthetic accessibility | 6.27 |

## Ponatinib





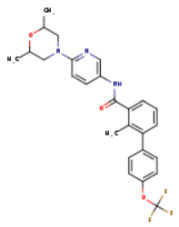

SMILES CN1CCN(CC1)Cc1ccc(cc1C(F)(F)F)NC(=O)c1ccc(c(c1)C#Cc1cnc2n1nccc2)C

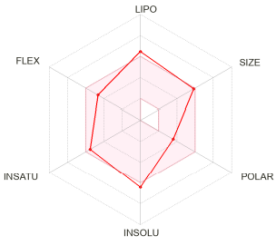
| Physicochemical Properties | |
|---|---|
| Formula | C29H27F3N6O |
| Molecular weight | 532.56 g/mol |
| Num. heavy atoms | 39 |
| Num. arom. heavy atoms | 21 |
| Fraction Csp3 | 0.28 |
| Num. rotatable bonds | 6 |
| Num. H-bond acceptors | 8 |
| Num. H-bond donors | 1 |
| Molar Refractivity | 150.10 |
| TPSA | 65.77 Å² |
| **Lipophilicity** | |
| Log $P_{o/w}$ (iLOGP) | 4.42 |
| Log $P_{o/w}$ (XLOGP3) | 4.11 |
| Log $P_{o/w}$ (WLOGP) | 4.58 |
| Log $P_{o/w}$ (MLOGP) | 3.90 |
| Log $P_{o/w}$ (SILICOS-IT) | 4.49 |
| Consensus Log $P_{o/w}$ | 4.30 |

| Water Solubility | |
|---|---|
| Log *S* (ESOL) | -5.73 |
| Solubility | 9.83e-04 mg/ml ; 1.85e-06 mol/l |
| Class | Moderately soluble |
| Log *S* (Ali) | -5.20 |
| Solubility | 3.38e-03 mg/ml ; 6.35e-06 mol/l |
| Class | Moderately soluble |
| Log *S* (SILICOS-IT) | -8.46 |
| Solubility | 1.84e-06 mg/ml ; 3.45e-09 mol/l |
| Class | Poorly soluble |
| **Pharmacokinetics** | |
| GI absorption | High |
| BBB permeant | Yes |
| P-gp substrate | Yes |
| CYP1A2 inhibitor | No |
| CYP2C19 inhibitor | Yes |
| CYP2C9 inhibitor | Yes |
| CYP2D6 inhibitor | Yes |
| CYP3A4 inhibitor | No |
| Log $K_p$ (skin permeation) | -6.63 cm/s |
| **Druglikeness** | |
| Lipinski | Yes; 1 violation: MW>500 |
| Ghose | No; 2 violations: MW>480, MR>130 |
| Veber | Yes |
| Egan | Yes |
| Muegge | Yes |
| Bioavailability Score | 0.55 |
| **Medicinal Chemistry** | |
| PAINS | 0 alert |
| Brenk | 1 alert: triple_bond |
| Leadlikeness | No; 2 violations: MW>350, XLOGP3>3.5 |
| Synthetic accessibility | 3.97 |

Fig. 7-8

## Erismodegib



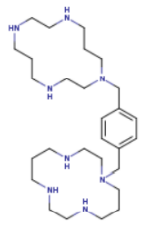


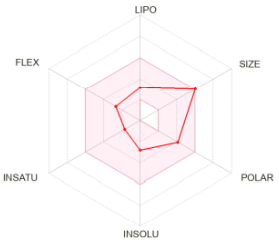

SMILES CC1OC(C)CN(C1)c1ccc(cn1)NC(=O)c1cccc(c1C)c1ccc(cc1)OC(F)(F)F

| Physicochemical Properties | |
|---|---|
| Formula | C26H26F3N3O3 |
| Molecular weight | 485.50 g/mol |
| Num. heavy atoms | 35 |
| Num. arom. heavy atoms | 18 |
| Fraction Csp3 | 0.31 |
| Num. rotatable bonds | 7 |
| Num. H-bond acceptors | 7 |
| Num. H-bond donors | 1 |
| Molar Refractivity | 130.88 |
| TPSA | 63.69 Å² |
| **Lipophilicity** | |
| Log $P_{o/w}$ (iLOGP) | 4.02 |
| Log $P_{o/w}$ (XLOGP3) | 5.81 |
| Log $P_{o/w}$ (WLOGP) | 6.51 |
| Log $P_{o/w}$ (MLOGP) | 3.25 |
| Log $P_{o/w}$ (SILICOS-IT) | 4.88 |
| Consensus Log $P_{o/w}$ | 4.89 |

| Water Solubility | |
|---|---|
| Log *S* (ESOL) | -6.43 |
| Solubility | 1.81e-04 mg/ml ; 3.72e-07 mol/l |
| Class | Poorly soluble |
| Log *S* (Ali) | -6.92 |
| Solubility | 5.87e-05 mg/ml ; 1.21e-07 mol/l |
| Class | Poorly soluble |
| Log *S* (SILICOS-IT) | -8.41 |
| Solubility | 1.89e-06 mg/ml ; 3.89e-09 mol/l |
| Class | Poorly soluble |
| **Pharmacokinetics** | |
| GI absorption | High |
| BBB permeant | No |
| P-gp substrate | Yes |
| CYP1A2 inhibitor | No |
| CYP2C19 inhibitor | Yes |
| CYP2C9 inhibitor | Yes |
| CYP2D6 inhibitor | Yes |
| CYP3A4 inhibitor | Yes |
| Log $K_p$ (skin permeation) | -5.14 cm/s |
| **Druglikeness** | |
| Lipinski | Yes; 0 violation |
| Ghose | No; 3 violations: MW>480, WLOGP>5.6, MR>130 |
| Veber | Yes |
| Egan | No; 1 violation: WLOGP>5.88 |
| Muegge | No; 1 violation: XLOGP3>5 |
| Bioavailability Score | 0.55 |
| **Medicinal Chemistry** | |
| PAINS | 0 alert |
| Brenk | 0 alert |
| Leadlikeness | No; 2 violations: MW>350, XLOGP3>3.5 |
| Synthetic accessibility | 4.16 |

## Plerixafor





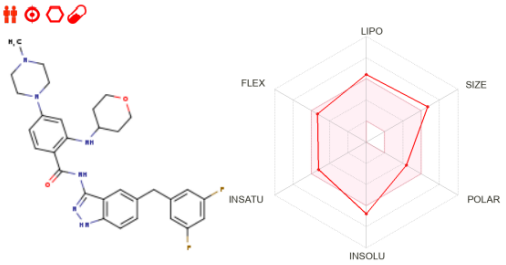

SMILES C1CNCCN(CCCNCCNC1)Cc1ccc(cc1)CN1CCNCCCNCCNCCC1

| Physicochemical Properties | |
|---|---|
| Formula | C28H54N8 |
| Molecular weight | 502.78 g/mol |
| Num. heavy atoms | 36 |
| Num. arom. heavy atoms | 6 |
| Fraction Csp3 | 0.79 |
| Num. rotatable bonds | 4 |
| Num. H-bond acceptors | 8 |
| Num. H-bond donors | 6 |
| Molar Refractivity | 182.21 |
| TPSA | 78.66 Å² |
| **Lipophilicity** | |
| Log $P_{o/w}$ (iLOGP) | 4.17 |
| Log $P_{o/w}$ (XLOGP3) | -0.01 |
| Log $P_{o/w}$ (WLOGP) | -2.93 |
| Log $P_{o/w}$ (MLOGP) | 0.44 |
| Log $P_{o/w}$ (SILICOS-IT) | 1.71 |
| Consensus Log $P_{o/w}$ | 0.68 |

| Water Solubility | |
|---|---|
| Log *S* (ESOL) | -2.81 |
| Solubility | 7.78e-01 mg/ml ; 1.55e-03 mol/l |
| Class | Soluble |
| Log *S* (Ali) | -1.19 |
| Solubility | 3.23e+01 mg/ml ; 6.42e-02 mol/l |
| Class | Very soluble |
| Log *S* (SILICOS-IT) | -7.57 |
| Solubility | 1.34e-05 mg/ml ; 2.67e-08 mol/l |
| Class | Poorly soluble |
| **Pharmacokinetics** | |
| GI absorption | Low |
| BBB permeant | No |
| P-gp substrate | Yes |
| CYP1A2 inhibitor | No |
| CYP2C19 inhibitor | No |
| CYP2C9 inhibitor | No |
| CYP2D6 inhibitor | No |
| CYP3A4 inhibitor | No |
| Log $K_p$ (skin permeation) | -9.37 cm/s |
| **Druglikeness** | |
| Lipinski | No; 2 violations: MW>500, NHorOH>5 |
| Ghose | No; 4 violations: MW>480, WLOGP<-0.4, MR>130, #atoms>70 |
| Veber | Yes |
| Egan | Yes |
| Muegge | No; 1 violation: H-don>5 |
| Bioavailability Score | 0.17 |
| **Medicinal Chemistry** | |
| PAINS | 0 alert |
| Brenk | 0 alert |
| Leadlikeness | No; 1 violation: MW>350 |
| Synthetic accessibility | 4.31 |

Fig. 9-10

## Entrectinib




SMILES CN1CCN(CC1)c1ccc(c(c1)NC1CCOCC1)C(=O)Nc1n[nH]c2c1cc(cc2)Cc1cc(F)cc(c1)F

| Physicochemical Properties | |
|---|---|
| Formula | C31H34F2N6O2 |
| Molecular weight | 560.64 g/mol |
| Num. heavy atoms | 41 |
| Num. arom. heavy atoms | 21 |
| Fraction Csp3 | 0.35 |
| Num. rotatable bonds | 8 |
| Num. H-bond acceptors | 6 |
| Num. H-bond donors | 3 |
| Molar Refractivity | 163.44 |
| TPSA | 85.52 Å² |

| Lipophilicity | |
|---|---|
| Log $P_{o/w}$ (iLOGP) | 3.75 |
| Log $P_{o/w}$ (XLOGP3) | 5.72 |
| Log $P_{o/w}$ (WLOGP) | 4.72 |
| Log $P_{o/w}$ (MLOGP) | 4.00 |
| Log $P_{o/w}$ (SILICOS-IT) | 5.12 |
| Consensus Log $P_{o/w}$ | 4.66 |

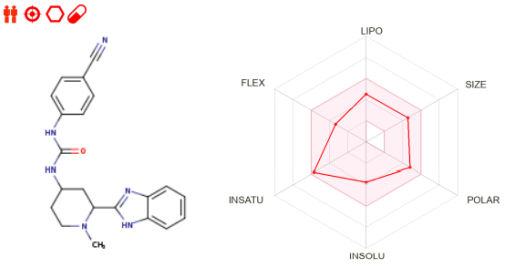
| Water Solubility | |
|---|---|
| Log *S* (ESOL) | -6.77 |
| Solubility | 9.51e-05 mg/ml ; 1.70e-07 mol/l |
| Class | Poorly soluble |
| Log *S* (Ali) | -7.28 |
| Solubility | 2.92e-05 mg/ml ; 5.21e-08 mol/l |
| Class | Poorly soluble |
| Log *S* (SILICOS-IT) | -9.78 |
| Solubility | 9.39e-08 mg/ml ; 1.67e-10 mol/l |
| Class | Poorly soluble |

| Pharmacokinetics | |
|---|---|
| GI absorption | High |
| BBB permeant | No |
| P-gp substrate | Yes |
| CYP1A2 inhibitor | No |
| CYP2C19 inhibitor | Yes |
| CYP2C9 inhibitor | No |
| CYP2D6 inhibitor | Yes |
| CYP3A4 inhibitor | No |
| Log $K_p$ (skin permeation) | -5.66 cm/s |

| Druglikeness | |
|---|---|
| Lipinski | Yes; 1 violation: MW>500 |
| Ghose | No; 3 violations: MW>480, MR>130, #atoms>70 |
| Veber | Yes |
| Egan | Yes |
| Muegge | No; 1 violation: XLOGP3>5 |
| Bioavailability Score | 0.55 |

| Medicinal Chemistry | |
|---|---|
| PAINS | 0 alert |
| Brenk | 0 alert |
| Leadlikeness | No; 3 violations: MW>350, Rotors>7, XLOGP3>3.5 |
| Synthetic accessibility | 4.06 |

## Glasdegib




SMILES N#Cc1ccc(cc1)NC(=O)NC1CCN(C(C1)c1nc2c([nH]1)cccc2)C

| Physicochemical Properties | |
|---|---|
| Formula | C21H22N6O |
| Molecular weight | 374.44 g/mol |
| Num. heavy atoms | 28 |
| Num. arom. heavy atoms | 15 |
| Fraction Csp3 | 0.29 |
| Num. rotatable bonds | 5 |
| Num. H-bond acceptors | 4 |
| Num. H-bond donors | 3 |
| Molar Refractivity | 111.53 |
| TPSA | 96.84 Å² |

| Lipophilicity | |
|---|---|
| Log $P_{o/w}$ (iLOGP) | 2.41 |
| Log $P_{o/w}$ (XLOGP3) | 2.37 |
| Log $P_{o/w}$ (WLOGP) | 2.50 |
| Log $P_{o/w}$ (MLOGP) | 1.64 |
| Log $P_{o/w}$ (SILICOS-IT) | 2.07 |
| Consensus Log $P_{o/w}$ | 2.20 |

| Water Solubility | |
|---|---|
| Log *S* (ESOL) | -3.72 |
| Solubility | 7.12e-02 mg/ml ; 1.90e-04 mol/l |
| Class | Soluble |
| Log *S* (Ali) | -4.04 |
| Solubility | 3.38e-02 mg/ml ; 9.03e-05 mol/l |
| Class | Moderately soluble |
| Log *S* (SILICOS-IT) | -6.16 |
| Solubility | 2.59e-04 mg/ml ; 6.92e-07 mol/l |
| Class | Poorly soluble |

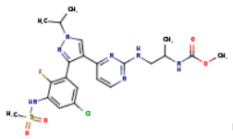
| Pharmacokinetics | |
|---|---|
| GI absorption | High |
| BBB permeant | No |
| P-gp substrate | Yes |
| CYP1A2 inhibitor | No |
| CYP2C19 inhibitor | Yes |
| CYP2C9 inhibitor | No |
| CYP2D6 inhibitor | Yes |
| CYP3A4 inhibitor | Yes |
| Log $K_p$ (skin permeation) | -6.90 cm/s |

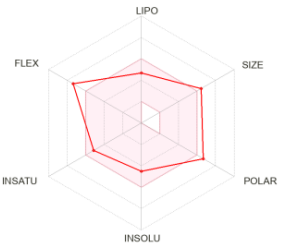
| Druglikeness | |
|---|---|
| Lipinski | Yes; 0 violation |
| Ghose | Yes |
| Veber | Yes |
| Egan | Yes |
| Muegge | Yes |
| Bioavailability Score | 0.55 |

| Medicinal Chemistry | |
|---|---|
| PAINS | 0 alert |
| Brenk | 0 alert |
| Leadlikeness | No; 1 violation: MW>350 |
| Synthetic accessibility | 3.63 |

Fig. 11-12

## Encorafenib






SMILES COC(=O)NC(CNc1nccc(n1)c1cn(nc1c1cc(Cl)cc(c1F)NS(=O)(=O)C)C(C)C)C

| Physicochemical Properties | |
|---|---|
| Formula | C22H27ClFN7O4S |
| Molecular weight | 540.01 g/mol |
| Num. heavy atoms | 36 |
| Num. arom. heavy atoms | 17 |
| Fraction Csp3 | 0.36 |
| Num. rotatable bonds | 11 |
| Num. H-bond acceptors | 8 |
| Num. H-bond donors | 3 |
| Molar Refractivity | 135.81 |
| TPSA | 148.51 Å² |

| Lipophilicity | |
|---|---|
| Log $P_{o/w}$ (iLOGP) | 2.53 |
| Log $P_{o/w}$ (XLOGP3) | 2.70 |
| Log $P_{o/w}$ (WLOGP) | 5.03 |
| Log $P_{o/w}$ (MLOGP) | 1.27 |
| Log $P_{o/w}$ (SILICOS-IT) | 2.10 |
| Consensus Log $P_{o/w}$ | 2.73 |

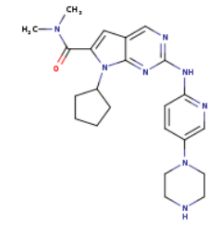
| Water Solubility | |
|---|---|
| Log *S* (ESOL) | -4.51 |
| Solubility | 1.66e-02 mg/ml ; 3.07e-05 mol/l |
| Class | Moderately soluble |
| Log *S* (Ali) | -5.47 |
| Solubility | 1.82e-03 mg/ml ; 3.38e-06 mol/l |
| Class | Moderately soluble |
| Log *S* (SILICOS-IT) | -7.58 |
| Solubility | 1.43e-05 mg/ml ; 2.65e-08 mol/l |
| Class | Poorly soluble |

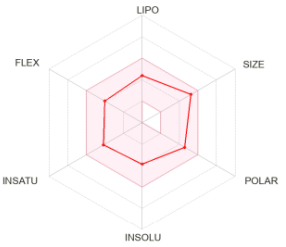
| Pharmacokinetics | |
|---|---|
| GI absorption | Low |
| BBB permeant | No |
| P-gp substrate | Yes |
| CYP1A2 inhibitor | No |
| CYP2C19 inhibitor | Yes |
| CYP2C9 inhibitor | No |
| CYP2D6 inhibitor | No |
| CYP3A4 inhibitor | Yes |
| Log $K_p$ (skin permeation) | -7.68 cm/s |

| Druglikeness | |
|---|---|
| Lipinski | No; 2 violations: MW>500, NorO>10 |
| Ghose | No; 2 violations: MW>480, MR>130 |
| Veber | No; 2 violations: Rotors>10, TPSA>140 |
| Egan | No; 1 violation: TPSA>131.6 |
| Muegge | Yes |
| Bioavailability Score | 0.17 |

| Medicinal Chemistry | |
|---|---|
| PAINS | 0 alert |
| Brenk | 0 alert |
| Leadlikeness | No; 2 violations: MW>350, Rotors>7 |
| Synthetic accessibility | 4.35 |

## Ribociclib






SMILES CN(C(=O)c1cc2c(n1C1CCCC1)nc(nc2)Nc1ccc(cn1)N1CCNCC1)C

| Physicochemical Properties | |
|---|---|
| Formula | C23H30N8O |
| Molecular weight | 434.54 g/mol |
| Num. heavy atoms | 32 |
| Num. arom. heavy atoms | 15 |
| Fraction Csp3 | 0.48 |
| Num. rotatable bonds | 6 |
| Num. H-bond acceptors | 5 |
| Num. H-bond donors | 2 |
| Molar Refractivity | 132.50 |
| TPSA | 91.21 Å² |

| Lipophilicity | |
|---|---|
| Log $P_{o/w}$ (iLOGP) | 3.50 |
| Log $P_{o/w}$ (XLOGP3) | 2.19 |
| Log $P_{o/w}$ (WLOGP) | 2.03 |
| Log $P_{o/w}$ (MLOGP) | 1.38 |
| Log $P_{o/w}$ (SILICOS-IT) | 1.09 |
| Consensus Log $P_{o/w}$ | 2.04 |

| Water Solubility | |
|---|---|
| Log *S* (ESOL) | -3.86 |
| Solubility | 5.93e-02 mg/ml ; 1.37e-04 mol/l |
| Class | Soluble |
| Log *S* (Ali) | -3.74 |
| Solubility | 7.92e-02 mg/ml ; 1.82e-04 mol/l |
| Class | Soluble |
| Log *S* (SILICOS-IT) | -5.51 |
| Solubility | 1.33e-03 mg/ml ; 3.06e-06 mol/l |
| Class | Moderately soluble |

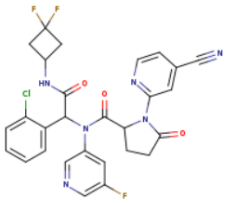
| Pharmacokinetics | |
|---|---|
| GI absorption | High |
| BBB permeant | No |
| P-gp substrate | Yes |
| CYP1A2 inhibitor | No |
| CYP2C19 inhibitor | No |
| CYP2C9 inhibitor | Yes |
| CYP2D6 inhibitor | Yes |
| CYP3A4 inhibitor | Yes |
| Log $K_p$ (skin permeation) | -7.40 cm/s |

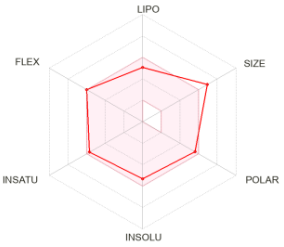
| Druglikeness | |
|---|---|
| Lipinski | Yes; 0 violation |
| Ghose | No; 1 violation: MR>130 |
| Veber | Yes |
| Egan | Yes |
| Muegge | Yes |
| Bioavailability Score | 0.55 |

| Medicinal Chemistry | |
|---|---|
| PAINS | 0 alert |
| Brenk | 0 alert |
| Leadlikeness | No; 1 violation: MW>350 |
| Synthetic accessibility | 3.47 |

Fig. 13-14

## Ivosidenib



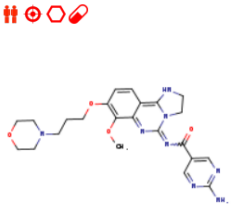


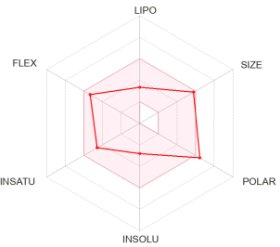

SMILES N#Cc1ccnc(c1)N1C(=O)CCC1C(=O)N(C(c1ccccc1Cl)C(=O)NC1CC(C1)(F)F)c1cncc(c1)F

| Physicochemical Properties | |
|---|---|
| Formula | C28H22ClF3N6O3 |
| Molecular weight | 582.96 g/mol |
| Num. heavy atoms | 41 |
| Num. arom. heavy atoms | 18 |
| Fraction Csp3 | 0.29 |
| Num. rotatable bonds | 9 |
| Num. H-bond acceptors | 9 |
| Num. H-bond donors | 1 |
| Molar Refractivity | 144.84 |
| TPSA | 119.29 Å² |
| **Lipophilicity** | |
| Log $P_{o/w}$ (iLOGP) | 2.76 |
| Log $P_{o/w}$ (XLOGP3) | 3.37 |
| Log $P_{o/w}$ (WLOGP) | 4.88 |
| Log $P_{o/w}$ (MLOGP) | 2.07 |
| Log $P_{o/w}$ (SILICOS-IT) | 4.04 |
| Consensus Log $P_{o/w}$ | 3.42 |

| Water Solubility | |
|---|---|
| Log *S* (ESOL) | -5.31 |
| Solubility | 2.87e-03 mg/ml ; 4.92e-06 mol/l |
| Class | Moderately soluble |
| Log *S* (Ali) | -5.55 |
| Solubility | 1.63e-03 mg/ml ; 2.80e-06 mol/l |
| Class | Moderately soluble |
| Log *S* (SILICOS-IT) | -8.25 |
| Solubility | 3.26e-06 mg/ml ; 5.60e-09 mol/l |
| Class | Poorly soluble |
| **Pharmacokinetics** | |
| GI absorption | Low |
| BBB permeant | No |
| P-gp substrate | Yes |
| CYP1A2 inhibitor | No |
| CYP2C19 inhibitor | Yes |
| CYP2C9 inhibitor | Yes |
| CYP2D6 inhibitor | Yes |
| CYP3A4 inhibitor | Yes |
| Log $K_p$ (skin permeation) | -7.46 cm/s |
| **Druglikeness** | |
| Lipinski | Yes; 1 violation: MW>500 |
| Ghose | No; 2 violations: MW>480, MR>130 |
| Veber | Yes |
| Egan | Yes |
| Muegge | Yes |
| Bioavailability Score | 0.55 |
| **Medicinal Chemistry** | |
| PAINS | 0 alert |
| Brenk | 0 alert |
| Leadlikeness | No; 2 violations: MW>350, Rotors>7 |
| Synthetic accessibility | 4.42 |

## Copanlisib tris-HCl






SMILES COc1c(OCCCN2CCOCC2)ccc2c1nc(=NC(=O)c1cnc(nc1)N)n1c2NCC1

| Physicochemical Properties | |
|---|---|
| Formula | C23H28N8O4 |
| Molecular weight | 480.52 g/mol |
| Num. heavy atoms | 35 |
| Num. arom. heavy atoms | 16 |
| Fraction Csp3 | 0.43 |
| Num. rotatable bonds | 8 |
| Num. H-bond acceptors | 9 |
| Num. H-bond donors | 2 |
| Molar Refractivity | 135.06 |
| TPSA | 142.01 Å² |
| **Lipophilicity** | |
| Log $P_{o/w}$ (iLOGP) | 3.40 |
| Log $P_{o/w}$ (XLOGP3) | 0.33 |
| Log $P_{o/w}$ (WLOGP) | -0.26 |
| Log $P_{o/w}$ (MLOGP) | -0.20 |
| Log $P_{o/w}$ (SILICOS-IT) | 1.56 |
| Consensus Log $P_{o/w}$ | 0.97 |

| Water Solubility | |
|---|---|
| Log *S* (ESOL) | -2.84 |
| Solubility | 6.99e-01 mg/ml ; 1.45e-03 mol/l |
| Class | Soluble |
| Log *S* (Ali) | -2.88 |
| Solubility | 6.40e-01 mg/ml ; 1.33e-03 mol/l |
| Class | Soluble |
| Log *S* (SILICOS-IT) | -5.63 |
| Solubility | 1.12e-03 mg/ml ; 2.34e-06 mol/l |
| Class | Moderately soluble |
| **Pharmacokinetics** | |
| GI absorption | Low |
| BBB permeant | No |
| P-gp substrate | Yes |
| CYP1A2 inhibitor | No |
| CYP2C19 inhibitor | No |
| CYP2C9 inhibitor | Yes |
| CYP2D6 inhibitor | No |
| CYP3A4 inhibitor | Yes |
| Log $K_p$ (skin permeation) | -9.00 cm/s |
| **Druglikeness** | |
| Lipinski | Yes; 1 violation: NorO>10 |
| Ghose | No; 2 violations: MW>480, MR>130 |
| Veber | No; 1 violation: TPSA>140 |
| Egan | No; 1 violation: TPSA>131.6 |
| Muegge | Yes |
| Bioavailability Score | 0.55 |
| **Medicinal Chemistry** | |
| PAINS | 0 alert |
| Brenk | 0 alert |
| Leadlikeness | No; 2 violations: MW>350, Rotors>7 |
| Synthetic accessibility | 3.91 |

Fig. 15-16

### 5.1.3 ADMET Predictor

ADMET Predictor is a machine learning software tool that quickly and accurately predicts over 175 properties, including solubility, logP, pKa, sites of CYP metabolism, and Ames mutagenicity. The latest version integrates market-leading ADMET modeling with compound design, data analysis, SAR, and cheminformatics capabilities to support scientists across computational chemistry, medicinal chemistry, DMPK, and other disciplines.

We will analyze different features relative to the higher ten docking score compounds, using different index to summarize in a proper way their biological risk.

Full ADMET Risk: a score in the 0-22 range indicating the number of potential ADMET problems a compound might have. Exceeds 7 for 10% of a focused World Drug Index (WDI) subset when ALL component risks are included. Absorption contributes up to 8, distribution up to 2, CYP metabolism up to 6, and Toxicity up to 6.

Full ADMET Risk rule codes: Size, RotB=rotatable bonds, HBD=H-bond donors, HBA=H-bond acceptors, ch=charge, Kow=lipophilicity, Peff=permeability, Sw=water solubility, fu=fraction unbound, Vd=volume of distribution, hERG=hERG inhibition, rat=acute rat toxicity, Xr=carcinogenicity in rat, Xm=carcinogenicity in mice, HEPX=hepatotoxicity, MUT=likely Ames positive; 1A2=high clearance by CYP 1A2, etc., CL=high microsomal clearance.

Water solubility (mg/mL) and Native pH calculated at S+Sw solubility.

Absorption Risk: a score in the 0-8 range indicating the number of potential oral absorption problems a compound is likely to have. Exceeds 4 for 9% of a focused WDI subset.

Predicts CYPi, whether or not the compound is human CYP 1A2 substrate (Yes/No), with relative accuracy.

Risk TOX connected with predicted toxicity: a score in the 0-6 range indicating the number of potential toxicity problems a compound might have. Exceeds 2.0 for 9% of a focused WDI subset.

### 5.1.4 ADMET Predictor Results

| Compound | Molecular weight | Full ADMET Risk | Full ADMET Risk rule codes |
|---|---|---|---|
| Etoposide | 588,569 | 6,019 | Size; HBA; ch; hERG-; rat; Xr+; Xm-; MUT |
| Ponatinib | 532,572 | 6,62 | Size; Kow; Sw; fu; Vd; Xr+; Xm; 3A4 |
| Erismodegib | 485,509 | 5,765 | Size; Kow; fu; Xr; 2C9; 3A4; CL |
| Plerixafor | 502,795 | 10,18 | Size; HBD; HBA; ch; Kow-; Peff; Sw-; fu-; Vd+; Xr-; Xm+; MUT; 1A2-; 2C9-; 2D6+; 3A4+; CL+ |
| Entrectinib | 560,651 | 6,914 | Size; ch; Kow; fu; Vd; Xr-; Xm; 3A4; CL |
| Glasdegib | 374,448 | 2 | Xr; 2C9 |
| Encorafenib | 540,019 | 5,514 | Size; RotB; HBA; ch; Xr-; Xm; 2C9 |
| Ribociclib | 434,547 | 2,657 | Size; Vd; Xr-; Xm+; 3A4 |
| Ivosidenib | 582,973 | 4 | Size; RotB; rat; Xr+; 2C9 |
| Copanlisib tris-HCl | 480,53 | 8,753 | Size; HBA; ch; Peff; Vd; hERG-; Xr-; Xm-; MUT; 1A2; CL- |

| Water solubility (mg/mL) | Native pH | Absorption Risk | Predicts CYPi | Risk TOX |
|---|---|---|---|---|
| 0,454 | 6,344 | 3 | No (90%) | 3,019 |
| 0,005 | 8,406 | 2,956 | No (90%) | 1,5 |
| 0,012 | 7,191 | 2 | No (90%) | 0,523 |
| 23,961 | 11,215 | 4,68 | No (80%) | 2 |
| 0,027 | 8,625 | 2,393 | No (97%) | 1,039 |
| 0,019 | 8,439 | 0 | Yes (74%) | 1 |
| 0,194 | 5,716 | 3,5 | No (97%) | 1,014 |
| 2,614 | 10,401 | 0,4 | No (85%) | 1 |
| 0,036 | 7,004 | 1,5 | No (90%) | 1,5 |
| 0,301 | 8,991 | 3,753 | Yes (57%) | 2,5 |

Tab 8. *Properties of best ten compounds coming from ADMET Predictor algorithm*

## 5.2 Molecular dynamics simulations

Molecular dynamics simulations have evolved into a mature technique that can be used effectively to understand macromolecular structure-to-function relationships. These simulations capture the behavior of proteins and other biomolecules in full atomic detail and at very fine temporal resolution. Major improvements in simulation speed, accuracy, and accessibility, together with the proliferation of experimental structural data, have increased the appeal of biomolecular simulation to experimentalists. We used this kind on analysis to guarantee the stability of our compounds inside the specific binding socket (RGD related one) of the main target (α5β1), doing that we will provide a solid time simulation about their movement and behavior through time.

The root-mean-square deviation (RMSD) is a similarity measure widely used in analysis of macromolecular structures and dynamics.

This kind of chart that shows rmsd changes during time track conformational changes of the system. This method measures the deviation of atom positions compared to a reference set of atom positions. RMSD plots can be employed to depict the relative movement of whole proteins or their domains during folding or drug binding. Also, RMSD calculations can be used to compare computed docking configurations to known binding poses to validate the docking protocol

### 5.2.1 Ivosidenib Results

Ivosidenib (lig 158) turn out to have the higher positive impact on our research in terms of docking binding energy and proper ADMET properties features, here we show its dynamics characteristics. Fig. 17

Another important aspect to be consider is the rmsd relative to the protein itself, to be sure that the compound attached doesn't negative influenced the behavior of it, we provide another chart relative to the time evolution of the protein. Fig. 18

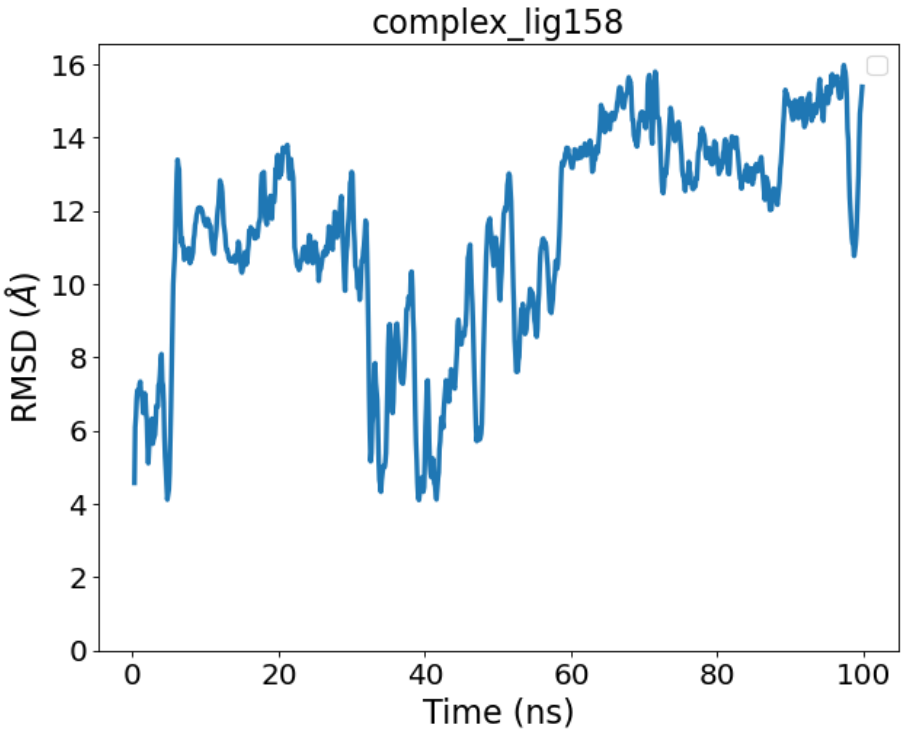



Fig. 17 *RMSD chart relative to Ivosidenib inside the protein's pocket*

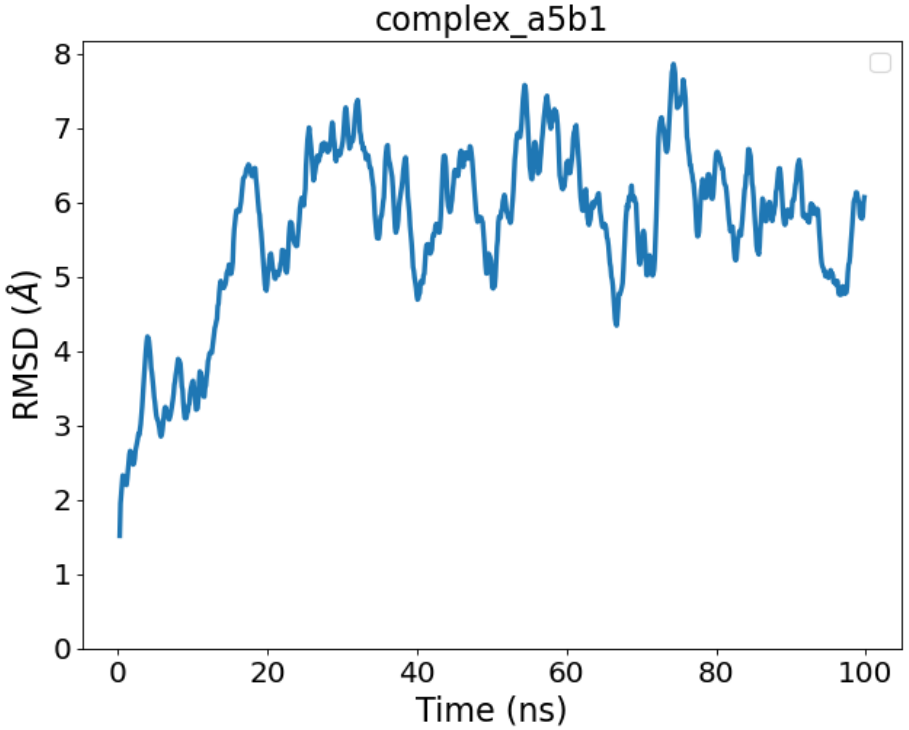



Fig. 18 *RMSD chart relative to the protein with Ivosidenib as ligand*

### 5.2.2 Entrectinib Results

Entrectinib (lig 141) turn out to have an high positive impact on our research in terms of docking binding energy and ADMET properties features, here we show its dynamics characteristics. Fig. 19
The rmsd relative to the protein itself is presented in Fig. 20.

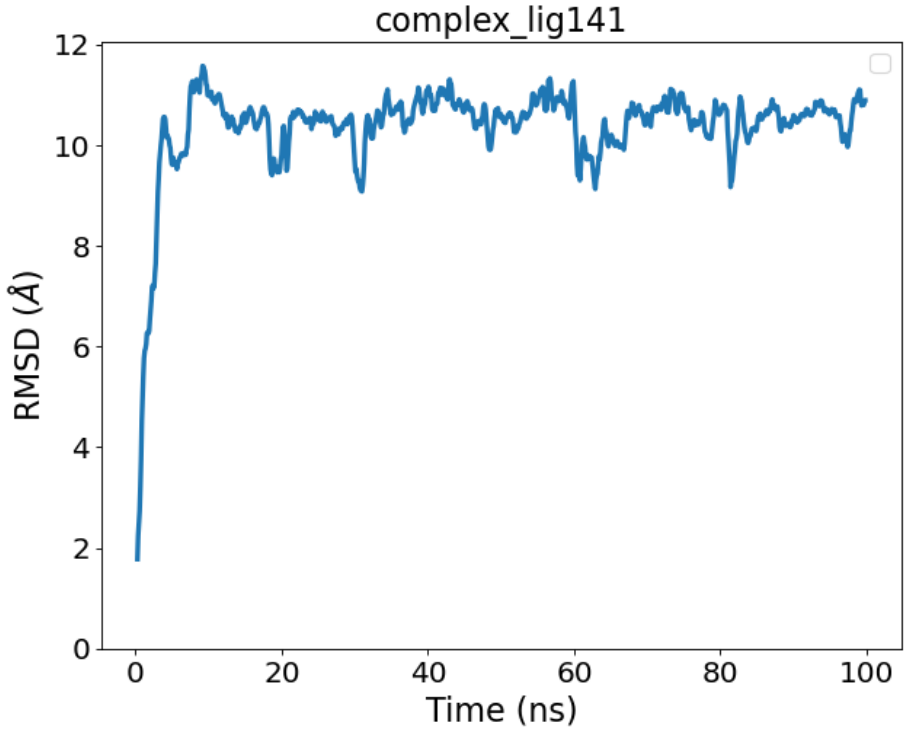



Fig. 19 *RMSD chart relative to Entrectinib inside the protein's pocket*

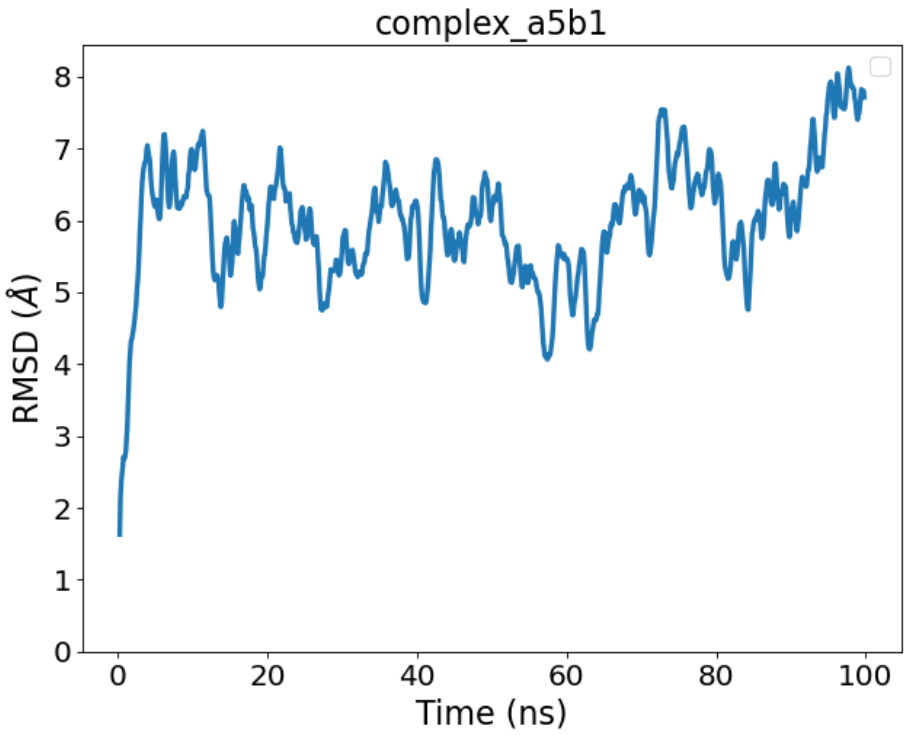



Fig. 20 *RMSD chart relative to the protein with Entrectinib as ligand*

### 5.2.3 Ponatinib Results

Ponatinib (lig 114) turn out to have an high positive impact on our research in terms of docking binding energy and ADMET properties features, here we show its dynamics characteristics. Fig. 21
The rmsd relative to the protein itself is presented in Fig. 22.

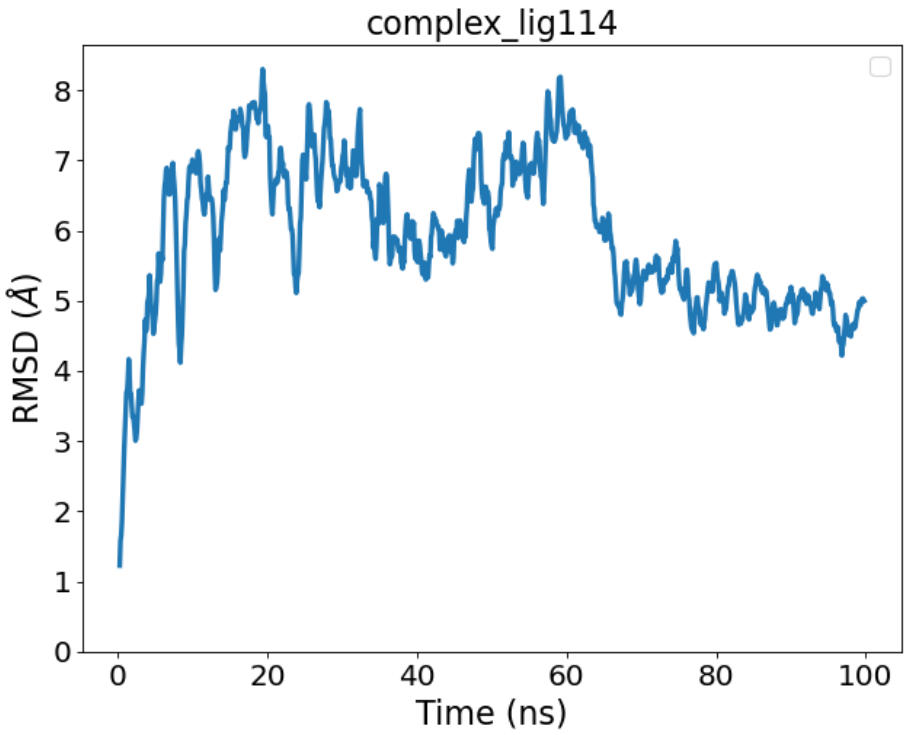



Fig. 21 *RMSD chart relative to Ponatinib inside the protein's pocket*

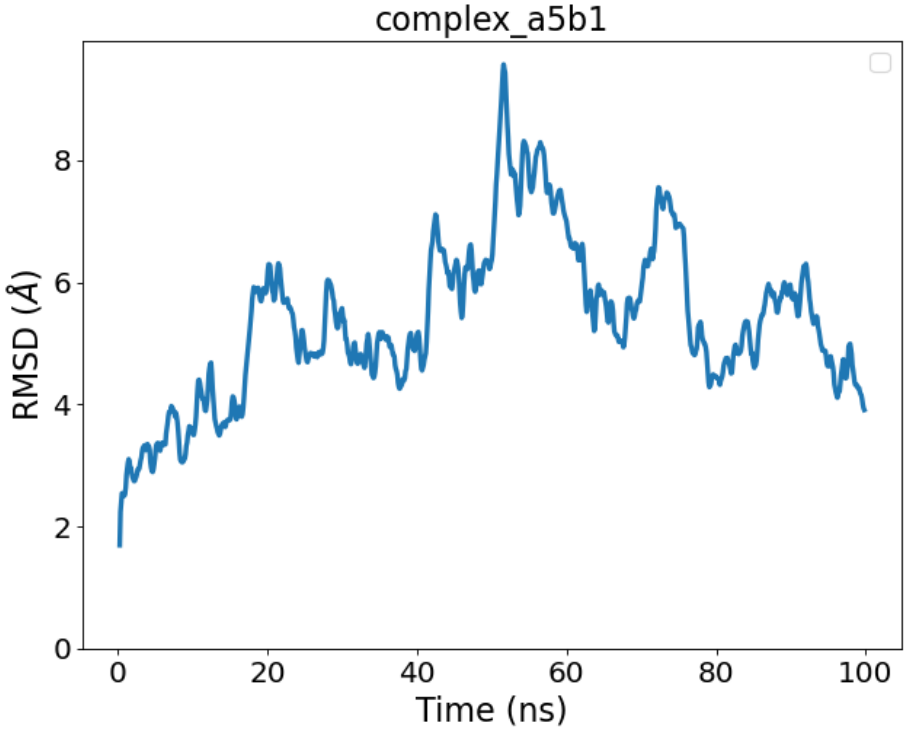



Fig. 22 *RMSD chart relative to the protein with Ponatinib as ligand*

### 5.2.4 Plerixafor Results

Plerixafor (lig 121) turn out to have an high positive impact on our research in terms of docking binding energy and ADMET properties features, here we show its dynamics characteristics. Fig. 23
The rmsd relative to the protein itself is presented in Fig. 24.

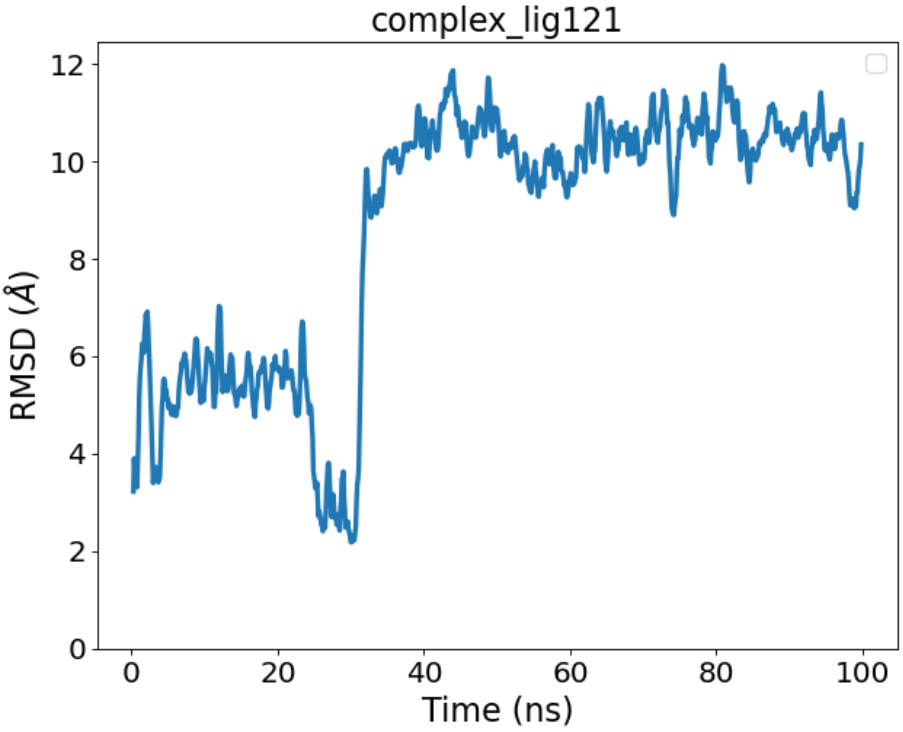



Fig. 23 *RMSD chart relative to Plerixafor inside the protein's pocket*

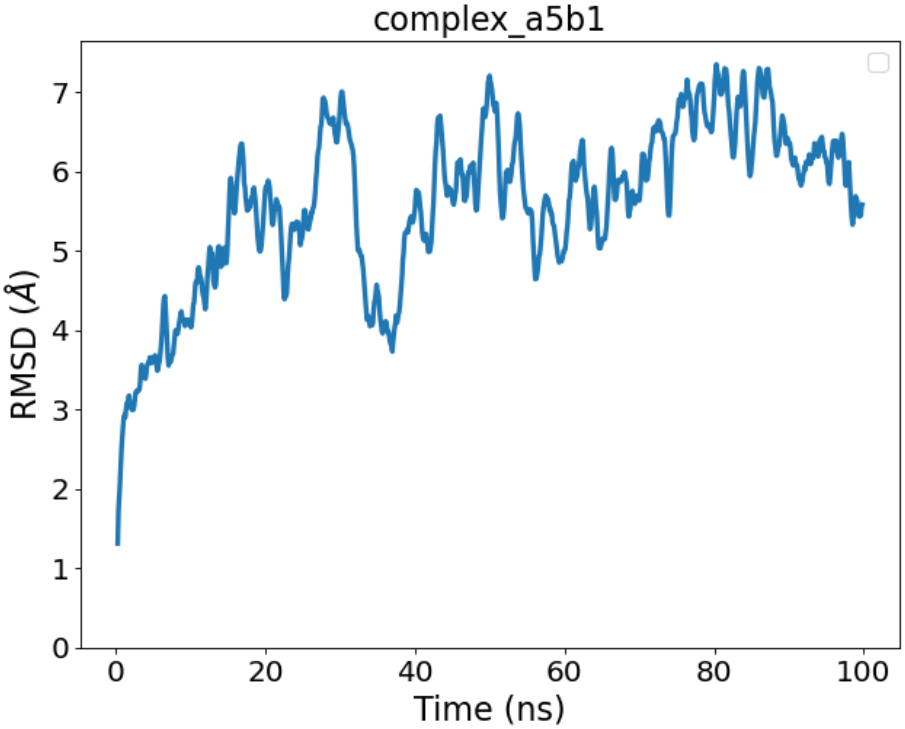



Fig. 24 *RMSD chart relative to the protein with Plerixafor as ligand*

# Chapter 6

# 6 Interactions analysis

## 6.1 Contacts and interactions

### 6.1.1 Materials and methods

Considering the complexes between the main integrin α5b1 and the 4 best ligands, all of the interactions involved in the binding were analyzed using the MOE software and particularly its “Contacts” analysis function, as this tool provides information about the bond type and the interacting residues involved.
The interaction between the integrin and ligands was evaluated using six types of contacts: hydrogen bonds (H) relative to hydrogen bond contacts; metal, relative to metal interactions which are bonded, or are close enough to be within bonded distance; ionic, for ionic bonds; arene, for arene interactions, these include π:π, π-H, and π:cation contacts; covalent, relative to covalent bonds; and distance (D) interactions relative to van der Waals distance interactions.
As starting point to perform our analysis we take the best poses, on for each ligand, in the main binding site relative to α5b1 integrin; the different structures are shown in figure 25-26-27-28.

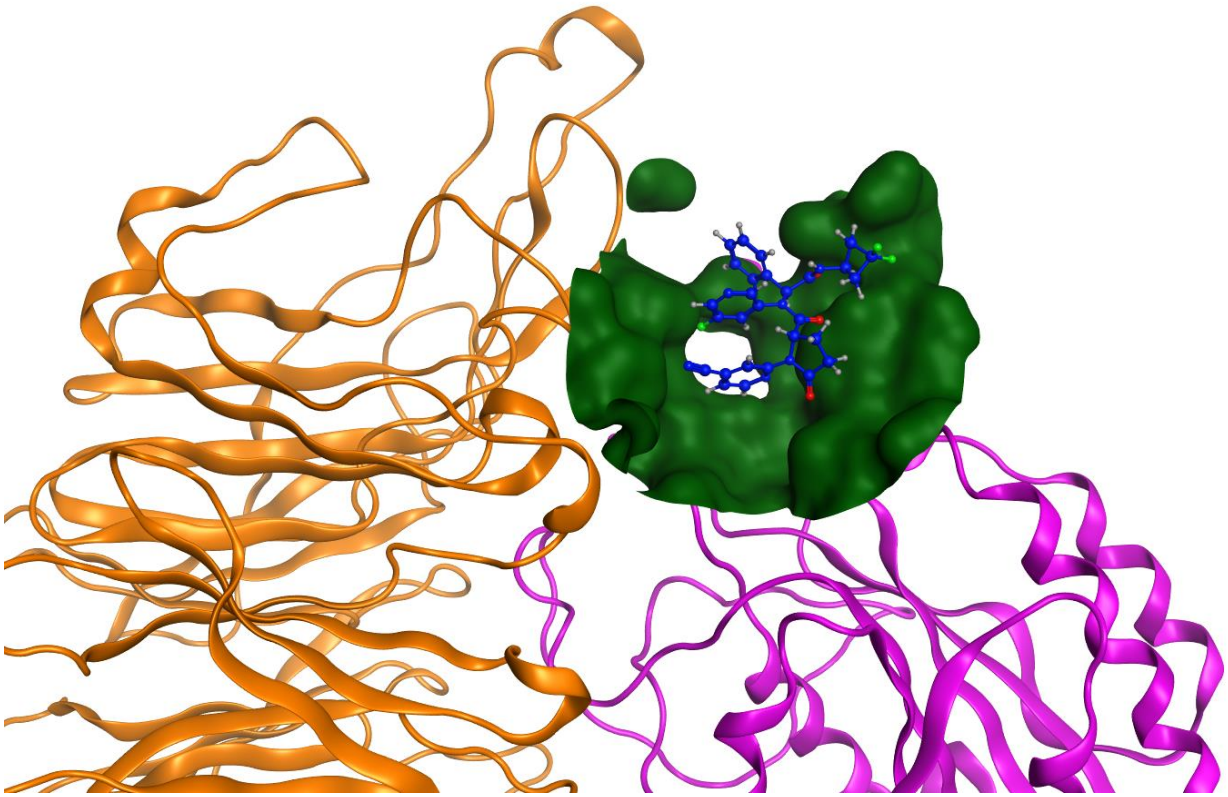


Fig 25. *Shown the best pose of Ivosidenib (lig 158) inside the RGD binding socket of α5b1, α5 portion in orange, b1 portion violet color, the ligand blue and the interaction area between the molecule and the integrin green color.*

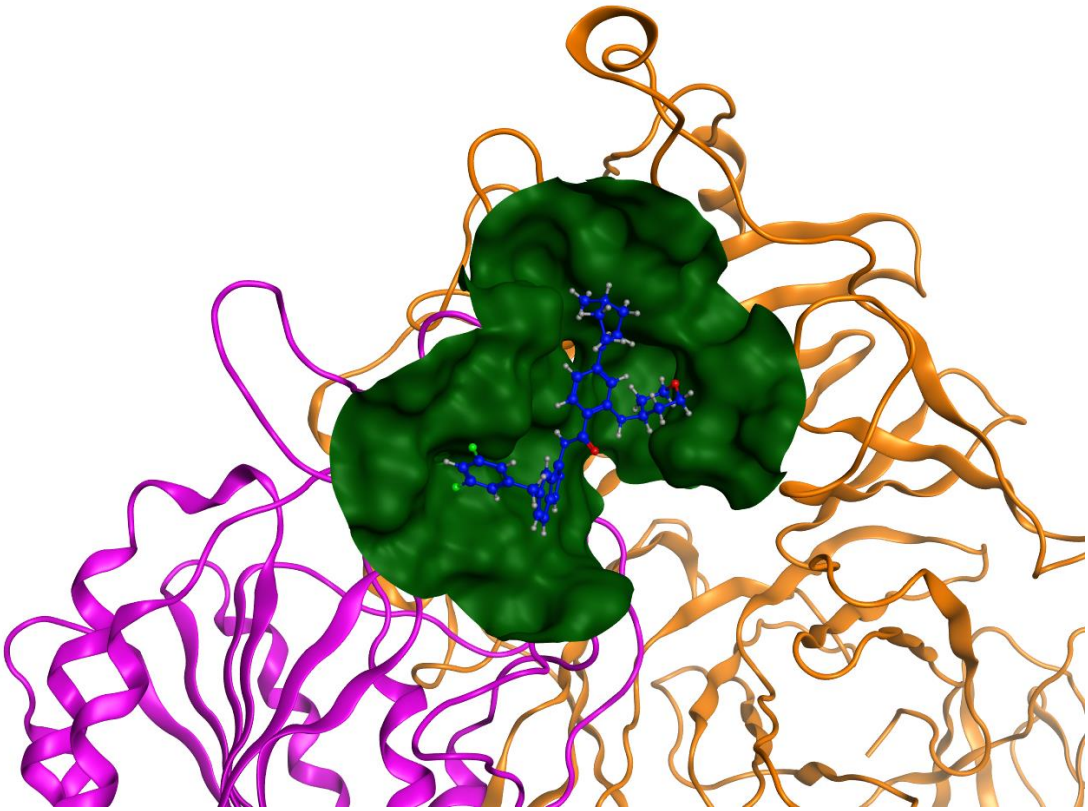


Fig 26. *Shown the best pose of Entrectinib (lig 141) inside the RGD binding socket of α5b1, α5 portion in orange, b1 portion violet color, the ligand blue and the interaction area between the molecule and the integrin green color.*

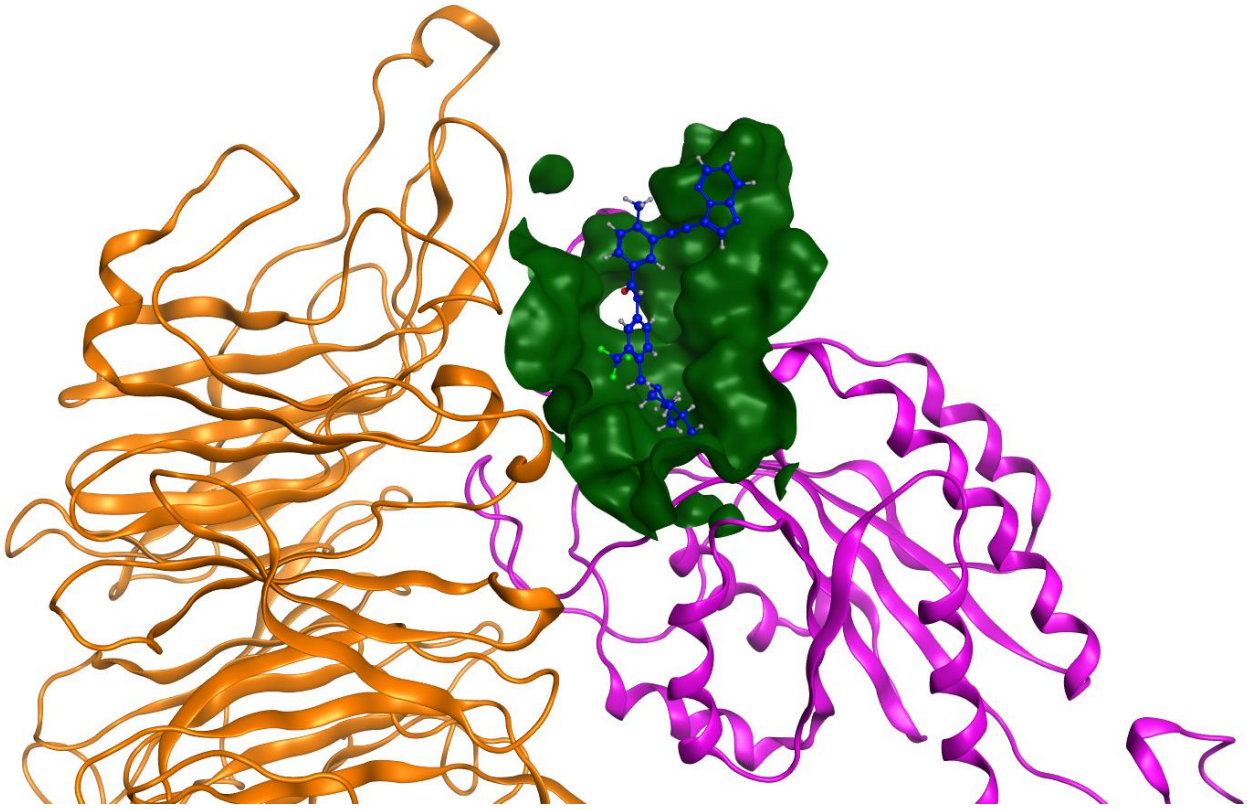


Fig 27. *Shown the best pose of Ponatinib (lig 114) inside the RGD binding socket of α5b1, α5 portion in orange, b1 portion violet color, the ligand blue and the interaction area between the molecule and the integrin green color.*

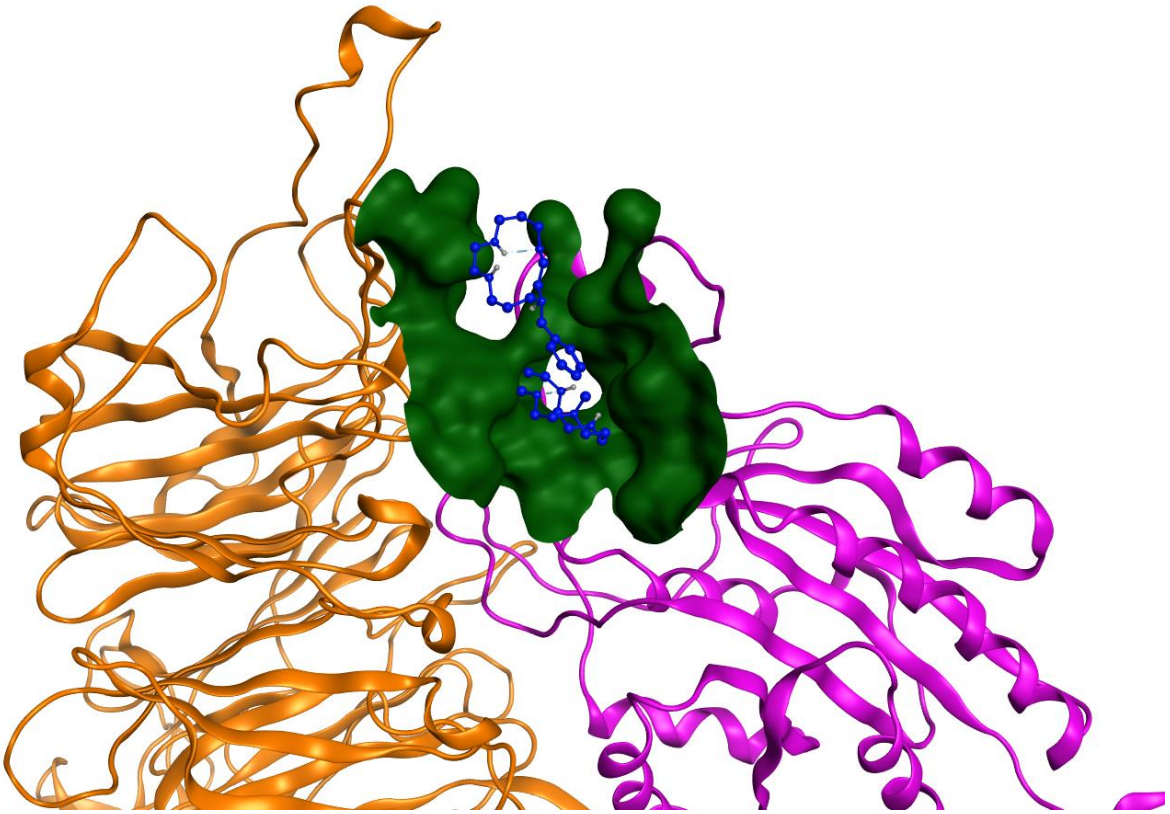


Fig 28. *Shown the best pose of Plerixafor (lig 121) inside the RGD binding socket of α5b1, α5 portion in orange, b1 portion violet color, the ligand blue and the interaction area between the molecule and the integrin green color.*

### 6.1.2 Results

Starting from the complexes previously presented we perform 2 types of analysis: 'Ligand interactions', which show ligand's atom position and contour interactions with the receptor; then 'Contact' analysis which give a better knowledge about the type of bond between the molecule and all the protein. All results are shown in Figure 29, Table 9 (Ivosidenib), Figure 30, Table 10 (Entrectinib), Figure 31, Table 11 (Ponatinib), Figure 32, Table 12 (Plerixafor).

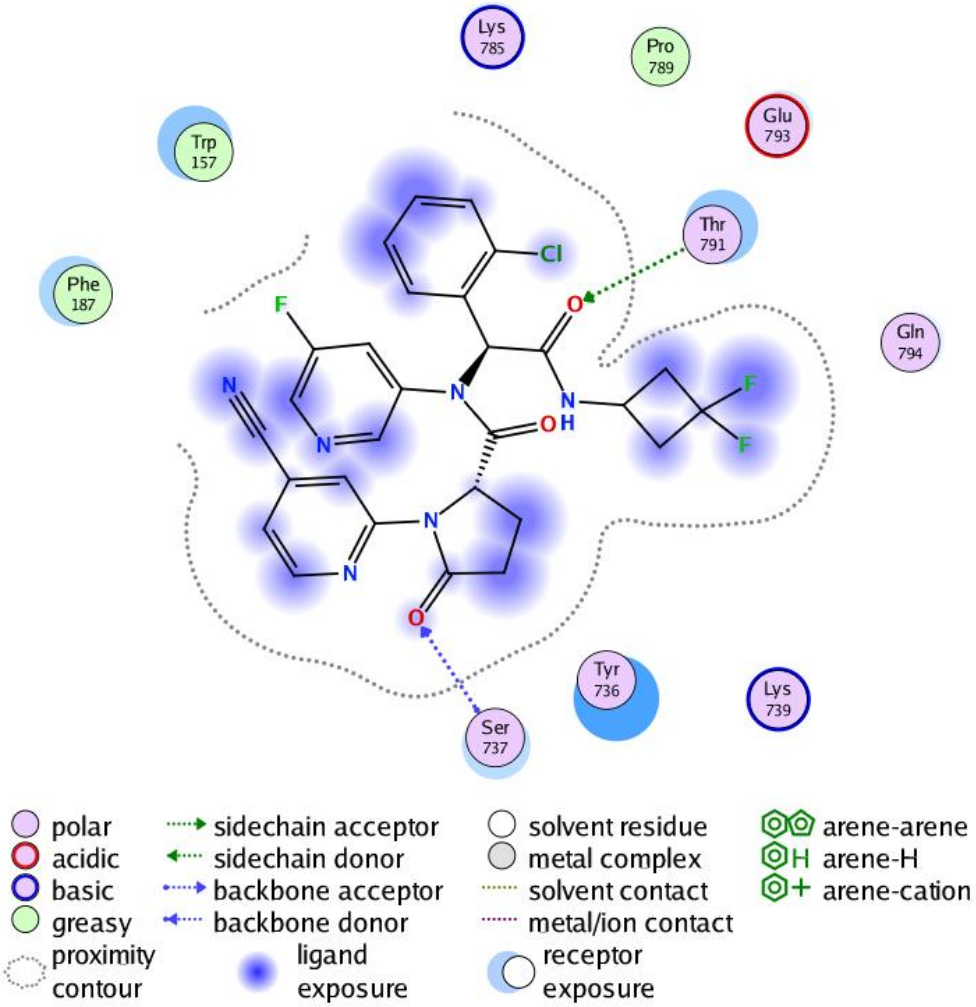



Fig 29. *Shown the interactions between Ivosidenib and the receptor contour, with a legend for the different interactions.*

| Type | ChainA | SetA | ChainB |
|---|---|---|---|
| D | receptor | Tyr736 | Ivosidenib |
| DH | receptor | Ser737 | Ivosidenib |
| D | receptor | Lys739 | Ivosidenib |
| D | receptor | Trp157 | Ivosidenib |
| D | receptor | Lys785 | Ivosidenib |
| D | receptor | Pro789 | Ivosidenib |
| D | receptor | Phe187 | Ivosidenib |
| DH | receptor | Thr791 | Ivosidenib |
| D | receptor | Glu793 | Ivosidenib |
| D | receptor | Gln794 | Ivosidenib |

Table 9. *Shown the contacts between the α5b1 (chain A) and the ligand (chain B), in column SetA a focus on the interaction's residue relative to the receptor.*

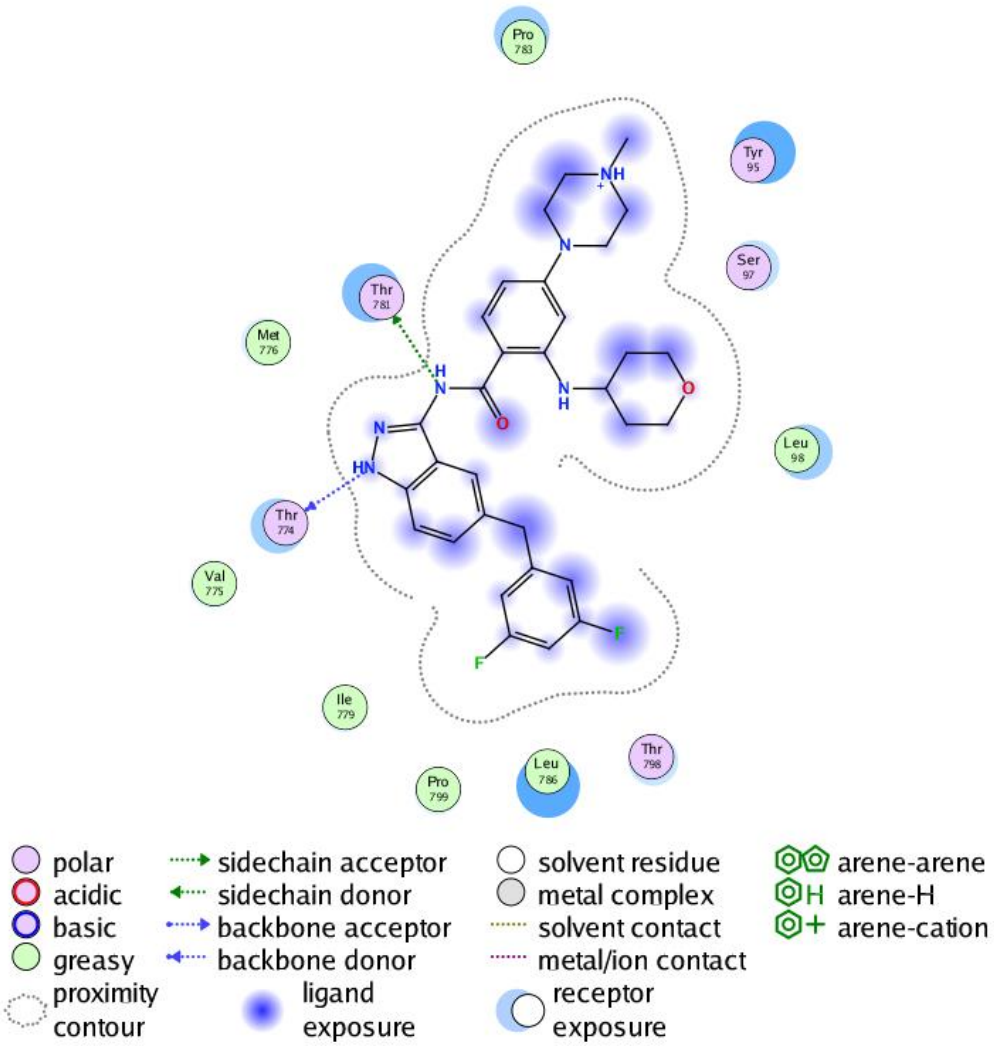



Fig 30. *Shown the interactions between Entrectinib and the receptor contour, with a legend for the different interactions.*

| Type | ChainA | SetA | ChainB |
|---|---|---|---|
| D | receptor | Tyr95 | Entrectinib |
| D | receptor | Ser97 | Entrectinib |
| D | receptor | Leu98 | Entrectinib |
| DH | receptor | Thr774 | Entrectinib |
| D | receptor | Val775 | Entrectinib |
| D | receptor | Met776 | Entrectinib |
| D | receptor | Ile779 | Entrectinib |
| DH | receptor | Thr781 | Entrectinib |
| D | receptor | Pro783 | Entrectinib |
| D | receptor | Leu786 | Entrectinib |
| D | receptor | Thr798 | Entrectinib |
| D | receptor | Pro799 | Entrectinib |

Table 10. *Shown the contacts between the α5b1 (chain A) and the ligand (chain B), in column SetA a focus on the interaction's residue relative to the receptor.*

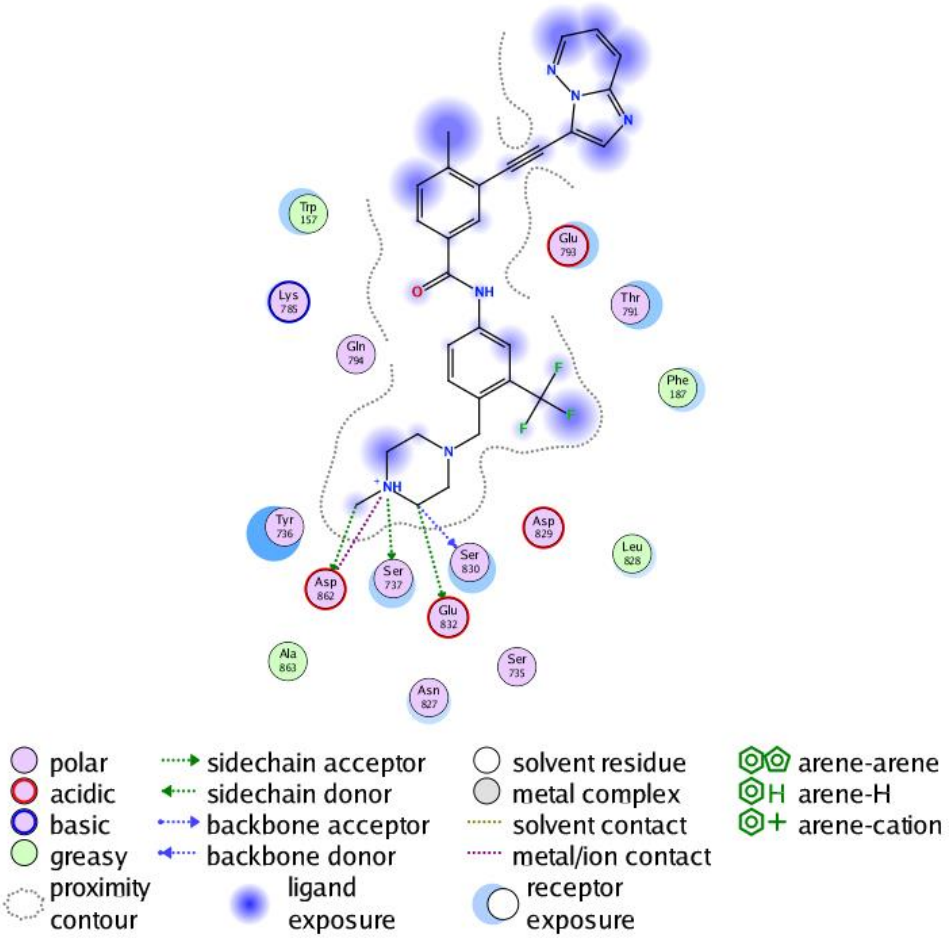



Fig 31. *Shown the interactions between Ponatinib and the receptor contour, with a legend for the different interactions.*

| Type | ChainA | SetA | ChainB |
|---|---|---|---|
| D | receptor | Ser735 | Ponatinib |
| D | receptor | Tyr736 | Ponatinib |
| DH | receptor | Ser737 | Ponatinib |
| D | receptor | Trp157 | Ponatinib |
| D | receptor | Lys785 | Ponatinib |
| D | receptor | Phe187 | Ponatinib |
| D | receptor | Thr791 | Ponatinib |
| D | receptor | Glu793 | Ponatinib |
| D | receptor | Gln794 | Ponatinib |
| D | receptor | Asn827 | Ponatinib |
| D | receptor | Leu828 | Ponatinib |
| D | receptor | Asp829 | Ponatinib |
| DH | receptor | Ser830 | Ponatinib |
| DH | receptor | Glu832 | Ponatinib |
| DIH | receptor | Asp862 | Ponatinib |
| D | receptor | Ala863 | Ponatinib |

Table 11. *Shown the contacts between the α5b1 (chain A) and the ligand (chain B), in column SetA a focus on the interaction's residue relative to the receptor.*

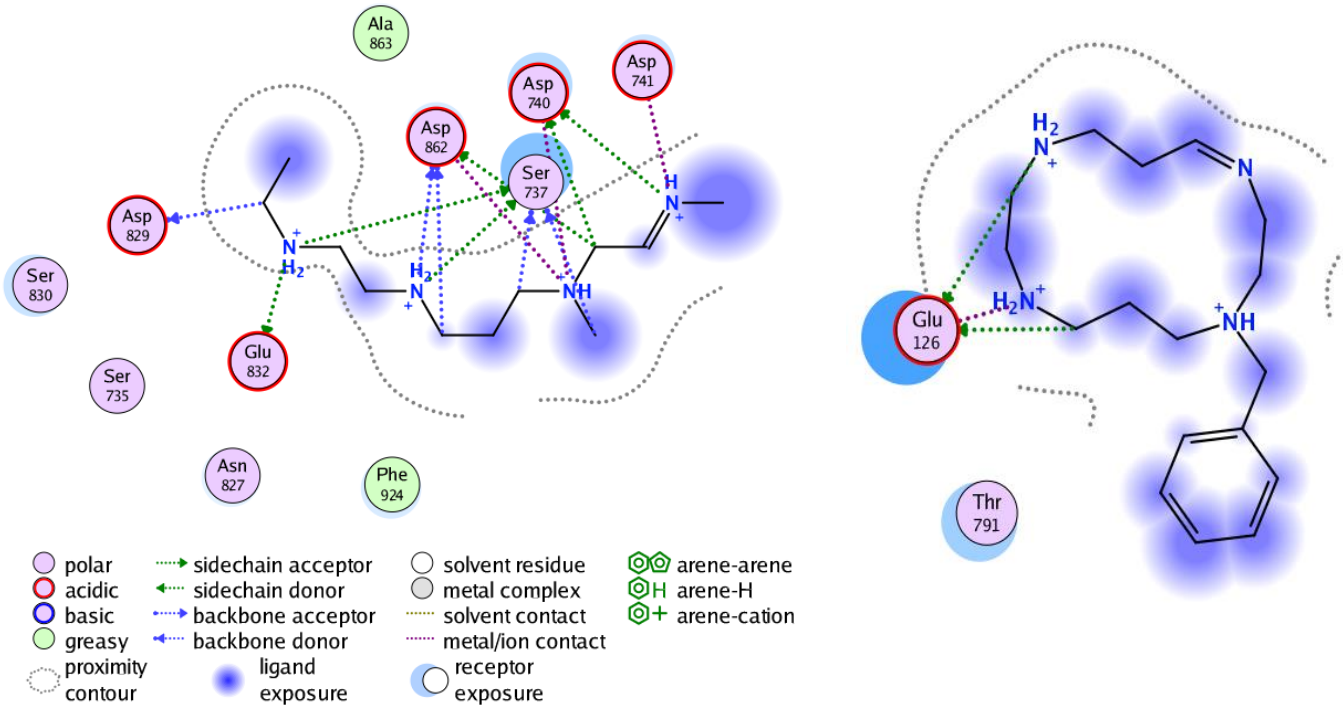



Fig 32. *Shown the interactions between Plerixafor and the receptor contour, with a legend for the different interactions.*

| Type | ChainA | SetA | ChainB |
|---|---|---|---|
| DIH | receptor | Glu126 | Plerixafor |
| D | receptor | Ser735 | Plerixafor |
| DH | receptor | Ser737 | Plerixafor |
| DIH | receptor | Asp740 | Plerixafor |
| DI | receptor | Asp741 | Plerixafor |
| D | receptor | Trp157 | Plerixafor |
| D | receptor | Lys785 | Plerixafor |
| D | receptor | Pro789 | Plerixafor |
| D | receptor | CYX790 | Plerixafor |
| D | receptor | Thr791 | Plerixafor |
| D | receptor | Asn827 | Plerixafor |
| DH | receptor | Asp829 | Plerixafor |
| D | receptor | Ser830 | Plerixafor |
| DIH | receptor | Glu832 | Plerixafor |
| DIH | receptor | Asp862 | Plerixafor |
| D | receptor | Ala863 | Plerixafor |
| D | receptor | Thr921 | Plerixafor |
| D | receptor | Phe924 | Plerixafor |

Table 12. *Shown the contacts between the α5b1 (chain A) and the ligand (chain B), in column SetA a focus on the interaction's residue relative to the receptor.*

# Chapter 7

# 7 Conclusion

## 7.1 Laboratory test

During our work we have relied on Toronto research group to perform a wound healing assay test on the best 4 compounds we got.
Cell migration is a pivotal process involved in physiologic and pathologic events including morphogenesis, immune cell trafficking, inflammation, and cancer metastasis [26].
Melanoma is a highly metastatic cancer that is responsible for 75% of all deaths related to cutaneous tumors. Once melanoma becomes metastatic, the patient has a very poor prognosis with the median survival of 6–12 months [27]. Therapeutic strategies aiming to disrupt the metastatic process are imperatively required for patients with melanoma. Along these lines, a better understanding of melanoma cell migration becomes essential given that cell migration is an event that takes places early during melanoma metastasis formation.
The wound healing assay is an easy, non-expensive, and highly reproducible method to study melanoma cell migration in vitro [28]. Wound healing assays can replicate some features of cell migration that happen in vivo. The foundation of this assay is the fact that the creation of an artificial wound in cells growing in a monolayer initiates cell migration. Cells migrate perpendicularly to the wound edge until cellular contacts are re-established. The wound healing assay is appropriate to study cell–cell and cell–extracellular matrix interactions during cell migration [29]. The conventional wound healing assay requires the formation of a cell monolayer that is scratched in order to create a wound. Cell migration is monitored as images are captured immediately after wound creation as well as at given time points during wound closure. Images are then compared in order to calculate cell migration [30].
Waiting for these analyses we moved on generating a model, coming from the optimal inhibitor molecule according to computational analysis.

## 7.2 Model for future analysis

Final purpose of this work is to construct a model based on one ligand which can be used to screening wide database of compounds or synthesize new type of inhibitors, useful to prevent cancer spreading.
Considering all analyses performed by now, Entrectinib (lig 141) comes out to be a proper choice to build our molecule's model: its second position in docking strength ranking, its good ADMET properties and its highly stability inside the pocket considered in molecular dynamics analysis lead us to that conclusion, this considerations needs to be confirmed by in vitro experiments once results are ready to be analyze.
Pharmacophore, which is defined by the International Union of Pure and Applied Chemistry (IUPAC) as "the ensemble of steric and electronic features that is necessary to ensure the optimal supra-molecular interactions with a specific biological target structure and to trigger (or to block) its biological response" [31], will be help us on that side. According to this definition, the interaction patterns of bioactive molecules with their targets are represented via a three-dimensional (3D) arrangement of abstract features that define interaction types rather than specific functional groups. These

interaction types can, for example, include the formation of hydrogen bonds, charged interactions, metal interactions, or hydrophobic (H) and aromatic (AR) contacts. In its entirety, a pharmacophore model represents one binding mode of ligands with a specific target [32].

The pharmacophore studies were performed using modules of the Molecular Operating Environment (MOE) software package. We used the contacts information coming from previous analyses (Section 6.1.2) to include the proper features in our model (Fig 33): F1,F4,F8,F11 are hydrogen bond donor;F2,F5,F9 are hydrogen bond acceptor;F3,F6,F7,F10 are aromatic center.

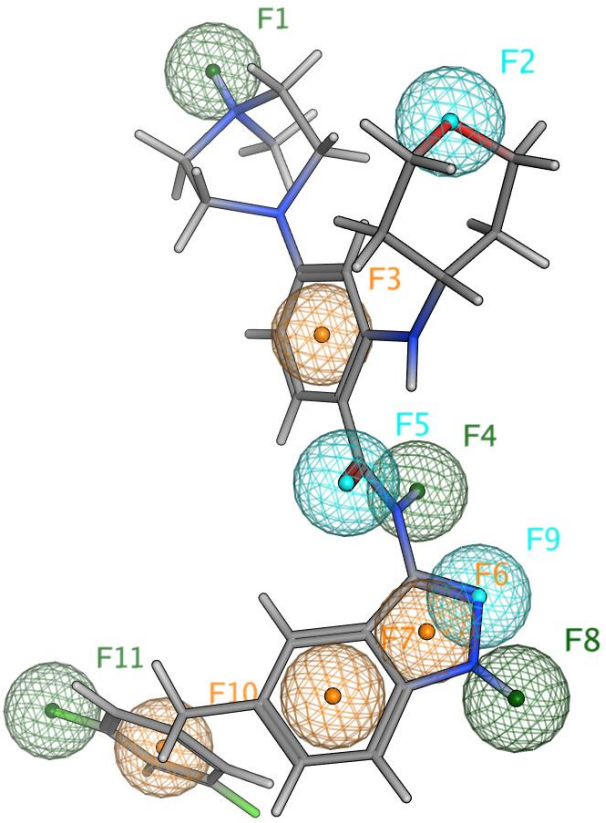



Fig 33. In this figure is shown the *pharmacophore based on Entrectinib molecule with different color depending on the feature they are representing (green: H-B donor, chano: H-B acceptor, orange: aromatic center); the sticks represent the structure of Entrectinib itsef.*

# Bibliography


1. Lin, T.-C. *et al.* Fibronectin in Cancer: Friend or Foe. *Cells* 9, (2020).

2. Hou, J., Yan, D., Liu, Y., Huang, P. & Cui, H. The Roles of Integrin α5β1 in Human Cancer. *OncoTargets Ther.* 13, 13329–13344 (2020).

3. Nieberler, M. *et al.* Exploring the Role of RGD-Recognizing Integrins in Cancer. *Cancers* 9, (2017).

4. Schumacher, S. *et al.* Structural insights into integrin α5β1 opening by fibronectin ligand. *Sci. Adv.* 7, eabe9716 (2021).

5. Boudjadi, S. *et al.* Involvement of the Integrin α1β1 in the Progression of Colorectal Cancer. *Cancers* 9, (2017).

6. Nymalm, Y. *et al.* Jararhagin-derived RKKH peptides induce structural changes in alpha1I domain of human integrin alpha1beta1. *J. Biol. Chem.* 279, 7962–7970 (2004).

7. Naci, D., Vuori, K. & Aoudjit, F. Alpha2beta1 integrin in cancer development and chemoresistance. *Semin. Cancer Biol.* 35, 145–153 (2015).

8. Emsley, J., King, S. L., Bergelson, J. M. & Liddington, R. C. Crystal structure of the I domain from integrin alpha2beta1. *J. Biol. Chem.* 272, 28512–28517 (1997).

9. Liu, Z. *et al.* Integrin (αvβ3) Targeted RGD Peptide Based Probe for Cancer Optical Imaging. *Curr. Protein Pept. Sci.* 17, 570–581 (2016).

10. Liu, Z., Wang, F. & Chen, X. Integrin alpha(v)beta(3)-Targeted Cancer Therapy. *Drug Dev. Res.* 69, 329–339 (2008).

11. Xiong, J.-P. *et al.* Crystal Structure of the Extracellular Segment of Integrin αVβ3 in Complex with an Arg-Gly-Asp Ligand. *Science* 296, 151–155 (2002).

12. Kukol, A. Consensus virtual screening approaches to predict protein ligands. *Eur. J. Med. Chem.* 46, 4661–4664 (2011).

13. Huang, N., Shoichet, B. K. & Irwin, J. J. Benchmarking sets for molecular docking. *J. Med. Chem.* 49, 6789–6801 (2006).

14. Chang, M. W., Ayeni, C., Breuer, S. & Torbett, B. E. Virtual screening for HIV protease inhibitors: a comparison of AutoDock 4 and Vina. *PloS One* 5, e11955 (2010).

15. Cheng, T., Li, X., Li, Y., Liu, Z. & Wang, R. Comparative assessment of scoring functions on a diverse test set. *J. Chem. Inf. Model.* 49, 1079–1093 (2009).

16. Ferreira, L. L. G. & Andricopulo, A. D. ADMET modeling approaches in drug discovery. *Drug Discov. Today* 24, 1157–1165 (2019).

17. Kar, S. & Leszczynski, J. Open access in silico tools to predict the ADMET profiling of drug candidates. *Expert Opin. Drug Discov.* 15, 1473–1487 (2020).

18. Jia, C.-Y., Li, J.-Y., Hao, G.-F. & Yang, G.-F. A drug-likeness toolbox facilitates ADMET study in drug discovery. *Drug Discov. Today* 25, 248–258 (2020).

19. Pires, D. E. V., Blundell, T. L. & Ascher, D. B. pkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures. *J. Med. Chem.* 58, 4066–4072 (2015).

20. Clark, A. M. *et al.* Open Source Bayesian Models. 1. Application to ADME/Tox and Drug Discovery Datasets. *J. Chem. Inf. Model.* 55, 1231–1245 (2015).

21. Lagorce, D., Bouslama, L., Becot, J., Miteva, M. A. & Villoutreix, B. O. FAF-Drugs4: free ADME-tox filtering computations for chemical biology and early stages drug discovery. *Bioinformatics* 33, 3658–3660 (2017).

22. Schyman, P., Liu, R., Desai, V. & Wallqvist, A. vNN Web Server for ADMET Predictions. *Front. Pharmacol.* 8, (2017).

23. Yang, H. *et al.* admetSAR 2.0: web-service for prediction and optimization of chemical ADMET properties. *Bioinformatics* 35, 1067–1069 (2018).

24. Xiong, G. *et al.* ADMETlab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties. *Nucleic Acids Res.* 49, W5–W14 (2021).

25. Daina, A., Michielin, O. & Zoete, V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* 7, 42717 (2017).

26. Pijuan, J. *et al.* In vitro Cell Migration, Invasion, and Adhesion Assays: From Cell Imaging to Data Analysis. *Front. Cell Dev. Biol.* 7, 107 (2019).

27. Schadendorf, D. *et al.* Melanoma. *Nat. Rev. Dis. Primer* 1, 15003 (2015).

28. Marshall, J. Transwell(®) invasion assays. *Methods Mol. Biol. Clifton NJ* 769, 97–110 (2011).

29. Liang, C.-C., Park, A. Y. & Guan, J.-L. In vitro scratch assay: a convenient and inexpensive method for analysis of cell migration in vitro. *Nat. Protoc.* 2, 329–333 (2007).

30. Rodriguez, L. G., Wu, X. & Guan, J.-L. Wound-healing assay. *Methods Mol. Biol. Clifton NJ* 294, 23–29 (2005).

31. Wermuth, C. G., Ganellin, C. R., Lindberg, P. & Mitscher, L. A. Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). *Pure Appl. Chem.* 70, 1129–1143 (1998).

32. Kaserer, T., Beck, K. R., Akram, M., Odermatt, A. & Schuster, D. Pharmacophore Models and Pharmacophore-Based Virtual Screening: Concepts and Applications Exemplified on Hydroxysteroid Dehydrogenases. *Mol. Basel Switz.* 20, 22799–22832 (2015).