UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES CHIMIQUE (222)

Ingénierie des fonctions moléculaires, Institut de Chimie, UMR 7177

THÈSE présentée par :

Katia GALENTINO

soutenue le : **12 Octobre 2022**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Chimie / Chimie théorique et informatique

Rational design of allosteric modulators in biomolecular motors

Insights into myosins'stability: from molecular modelling to vHTS

THÈSE dirigée par :

M. CECCHINI Marco	Dr, Maitre de Conférence, CNRS - Université de Strasbourg

RAPPORTEURS :

Mme. COURNIA Zoe	Dr, Associate Professor – Biomedical Research Foundation of Academy of Athens (BRFAA)
Mme. FORNILI Arianna	Dr, Senior Lecturer – Queen Mary University of London

---

AUTRES MEMBRES DU JURY :

Mme. HOUDUSSE-JUILLÉ Anne	Dr, Directeur de Recherche, CNRS – Institut Curie
M. STOTE Roland	Dr, Directeur de Recherche, CNRS – Université de Strasbourg
M. HORVATH Dragos	Dr, Directeur de Recherche, CNRS – Université de Strasbourg

RATIONAL DESIGN OF ALLOSTERIC MODULATORS IN BIOMOLECULAR MOTORS

Insights into Myosins' stability: from molecular modelling to vHTS

Katia Galentino

Supervisor: Dr. Marco Cecchini
Co-encadrant: Dr. Anne Houdusse

Laboratoire d'Ingénierie des Fonctions Moléculaires (IFM Lab)
UMR7177

École doctorale des Sciences Chimiques (ED222)

Strasbourg

Université
de Strasbourg

Ad maiora semper

Dedicated to my parents, Italo & Maria Luisa

ABSTRACT

Myosins are biomolecular motors that coordinate life and motion in eukaryotic cells. These molecular machines are able to exploit chemical energy derived from ATP-hydrolysis to generate mechanical work and walk across actin filament. This movement is tuned by actin, in a cyclic manner, describing the so-called acto-myosin cycle. Understanding the chemo-mechanical transduction of these motors has been a crucial topic since decades, but many features remained unknown. Moreover, dysfunction of myosins are associated with severe human disorders, ranging from myopathies to cancers. The lack of safe treatments makes these proteins fundamental therapeutic targets. In this project of thesis, we have studied the dynamics of a peculiar conformational state of the acto-myosin cycle, known as pre-powerstroke (PPS), both in cardiac myosin and in myosinVI. While exploring the stability of this state in myosinVI, we captured a spontaneous transition towards a different intermediate state of the force generation cycle. Comparison with the dynamics of cardiac myosin denoted a remarkable plasticity of the structure. Simultaneously, we investigated these two proteins from a pharmacological point of view, by applying a wide range of computational approaches to comprehend how to design cardiac myosin's allosteric activators, as well as to perform a virtual high throughput screening (vHTS) on myosinVI. Finally, we describe the design of selective inhibitors for two different protein kinases (MST2 and S6) performed in collaboration with the group of Dr. Mobian and Dr. Gaiddon. This collaboration culminated in the selection of peculiar compounds generated by coordination chemistry, that have been experimentally tested.

ACKNOWLEDGMENTS

I would like to thank everybody who took part in this complicated journey. First, I would like to thank my parents, Italo e Maria Luisa, who always supported me. It has not been easy. We were distant, during a global pandemic and a series of unfortunate events that challenged us. Nevertheless, you were always there, ready to participate in my adventure, giving me wise (and funny) advice. I love you more than you can imagine.

I want to thank my sister Sara, my twin, my Demi. We shared all the emotions of these past years, being always there, one for the other. This will never change.

A thought also to my big brother Andrea, who is always excited by listening about scientific evolution. Thank you for your enthusiasm and the energy you show us every day.

I would like to thank all my friends, from Italy to France! You were all there in this important moment of my life. I think about my "historical" friends from Andria, my best friend Cessa (Alessandra for the others), and all the people I met in Strasbourg. I thank all of you for your love, your support, and your craziness.

I thank Gilberto and Marion, my colleagues-friends. Gilberto was my first friend in the IFM Lab and since then we build a friendship that I hope will last forever...Even if you think that to speak Italian you just need to add a double consonant and a "-ino" at the end of each word... Froggolino...

Marioncina, you became my sister of PhD, as you define us! Thank you for all the games of Apex! Thank you for the aperitif we took together as soon as we were unconfined as well as the ones we organized online. Thank you for being the first person to make me discover the Picon. I will never forget all our moments together.

I want to thank Florian, for the advice, the few races we did together, and the help in my research. Also, thank you for the pictures of your wonderful daughter you shared with me during my writing process. It really helped me in smiling once more.

A thank to Alessio for our conversations and for the funny faces you give us every day. I am sure you will be a great group leader! Ti voglio bene.

I would like to thank Luchino for the months spent together in the lab. I was your "supervisor", but actually, you never needed me. Thank you for the countless laughs and the affection you demonstrated to me. I am so proud of you and I will always support your bull****s!

Mariia thank you for teaching me some Python stuff and for being patient with me, in our lab. Hang on and never give up! You will do great!

I want to thank Marco for welcoming me into his lab and allowing me to meet all these amazing people. Thank you for the freedom you gave me in my research and for the sympathy you have shown in some particular moments.

A big thanks to my co-encadrant Anne Houdusse. You are the Queen of Myosins. Thank you for the very interesting discussions we had with you and your group and for showing me how interesting is Myosins' world. I learnt a lot from you!

I want to thank Pierre Mobian and Matthieu Scarpi-Luttenauer for the enthusiasm shared during our collaboration and for challenging me with the huge guys you designed. It has been really cool to work with you!

Thanks also to Mme Paola Sager and Mme Kostmann, for the unconditional help they give to all of us students from beginning to end. You are like two guardian angels.

I would like to thank Ginette and Jean-Marie for becoming my second family here in France. Thank you for welcoming me into your house and for sharing with me this journey, from beginning to end. You make everyone feel at home and this is a gift!

I want to thank the rock of my life, the most important person for me, mon chéri Adrien. Adrien, you helped me in the most difficult moments and you even survived all the confinements with me! I think you are a hero or a masochist. Together we lived so many emotions that we cannot even explain with words. I was able to reach this damned finish line mostly thanks to you. You have always supported me, as no one ever has. You guided me and you have always believed in me and my stupid ideas. I love you immensely and this will never change.

RÉSUMÉ DÉTAILLÉ

INTRODUCTION

Myosines: structure

Les myosines sont des protéines cytoplasmiques, isolées pour la première fois en 1864 par le physiologiste Wilhelm Kühne (Kühne, 2011). Elles représentent une classe de moteurs biomoléculaires classés comme des protéines moléculaires associées à l'actine (Howard, 1997), puisqu'elles se déplacent sur le filament d'actine en convertissant l'énergie chimique dérivée de l'hydrolyse de l'ATP en travail mécanique (Rdowicz, 2002). La superfamille des myosines compte 35 classes, basées sur l'analyse de la séquence de la partie globulaire de la structure appelée domaine moteur (Odronitz and Kollmar, 2007). Les myosines sont divisées en deux catégories principales : conventionnelles et non conventionnelles (Batters and Veigel, 2016; Hartman and Spudich, 2012; Thompson and Langford, 2002). Les myosines dimériques sont classées en deux catégories : les myosines processives et les myosines non processives, en fonction de leur capacité à empêcher le détachement simultané des deux têtes de la piste, sous l'effet d'une charge (Howard, 1997; Lee Sweeney and Houdusse, 2004; O'Connell et al., 2007). Les myosines processives ont une vitesse lente de libération de l'ADP du site actif, lorsqu'elles sont liées à l'actine, et ont des pas plus grands sur le filament (Batters and Veigel, 2016). Cela affecte le ratio cyclique (Batters and Veigel, 2016; Lee Sweeney and Houdusse, 2004), c'est-à-dire le temps passé par la tête du moteur à peupler l'état étroitement lié à l'actine (état de génération de force) par rapport au temps du cycle cinétique global (Batters and Veigel, 2016; Lee Sweeney and Houdusse, 2004). Les myosines non processives ont un ratio cyclique faible en l'absence de charge (Lee Sweeney and Houdusse, 2004), tandis que les myosines processives ont un ratio cyclique plus élevé pour permettre leur mouvement processif (Rock et al., 2000). Chaque monomère de myosine présente une chaîne lourde (MHC), contenant une partie globulaire, et deux chaînes légères (MLC), nommées respectivement Essential et Regulatory light chain (ELC et RLC) (AL-Khayat, 2013). La partie globulaire est connue sous le nom de *chef de myosine* ou sous-fragment 1 (S1), ou *domaine moteur* ou encore tête de myosine. Il s'agit du domaine le plus conservé de la protéine (Foth et al., 2006; Howard, 1997), car c'est là que résident les éléments structurels et fonctionnels responsables du mécanisme de génération de la force. La tête de la myosine est reliée à la longue queue C-terminale par une région centrale en forme de col. Au niveau de ce col se trouvent les MLC, membres des protéines calmoduline (CaM) et CaM-like (Heissler and Sellers, 2014), importantes pour l'intégrité structurelle de l'enzyme (England and Loughna, 2013; Robert-Paganin et al., 2018; Syamaladevi et al., 2012) (voir Figure 0.1).

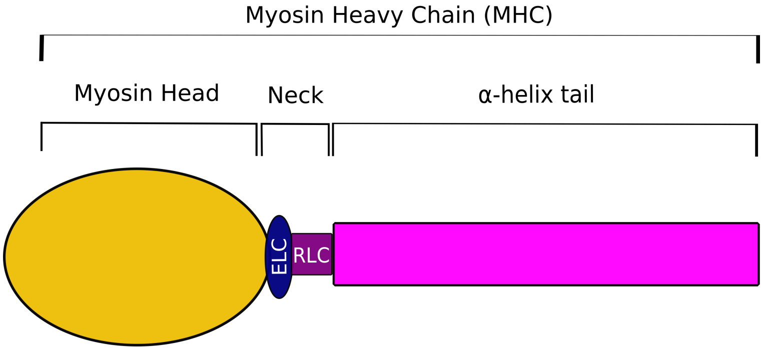


Figure 0.1: Esquisse de la chaîne lourde de la myosine (MHC).

Dans le HMC se trouvent le site de liaison de l'ATP et l'interface de liaison à l'actine (England and Loughna, 2013; Robert-Paganin et al., 2018; Syamaladevi et al., 2012), entre deux sous-domaines appelés Upper50 (U50) et Lower50 (L50), en raison de leur dimension de 50 kDa (Figure 0.2) (Syamaladevi et al., 2012; Wulf et al., 2016). Les sous-domaines U50 et L50 forment une fente qui est ouverte lorsque la myosine est dans l'état de faible affinité pour l'actine et, au contraire, est fermée lorsque la myosine est étroitement liée à l'actine (Wulf et al., 2016). La molécule d'ATP se fixe dans le site actif de la myosine, situé au cœur du domaine moteur, entre les sous-domaines U50 et N-terminal (Syamaladevi et al., 2012).

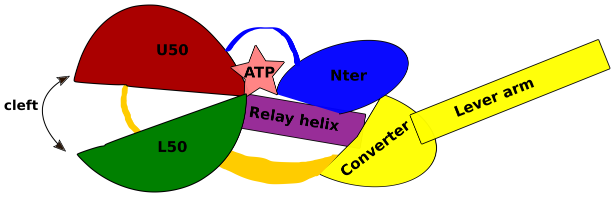


Figure 0.2: Plan des principaux sous-domaines de la tête de la myosine. Upper50 (U50) en rouge et Lower50 (L50) en vert décrivent l'interface de liaison à l'actine (fente indiquée par une flèche noire). Entre l'U50 et le sous-domaine N-terminal (Nter) se trouve le site actif où se fixe l'ATP (l'étoile rose). Les sous-domaines du convertisseur et du bras de levier sont représentés en jaune. Le convertisseur est relié au reste du domaine moteur par l'hélice Relay (partie de L50). La rotation du convertisseur est fondamentale pour amplifier les changements structurels dans le domaine moteur pendant le cycle de génération de force afin d'accomplir la rotation finale du bras de levier.

Le site actif est composé de trois boucles principales appelées commutateur I (swI), boucle P et commutateur II (swII) (Syamaladevi et al., 2012), qui piègent le nucléotide et médient l'hydrolyse de l'ATP (Geeves and Holmes, 1999; Mesentean et al., 2007). En particulier, swII se ferme sur le phosphate γ et hydrolyse l'ATP en ADP $\cdot P_i$ (Onishi et al., 2002). D'autres connecteurs importants dans le domaine moteur sont l'hélice de relais (RH) et l'hélice SH1 (Cecchini et al., 2008; Llinas et al., 2015). Ces deux éléments relient le sous-domaine du convertisseur au reste de l'architecture (Cecchini et al., 2008). Le convertisseur est un élément flexible du domaine moteur et sa rotation est un mouvement clé qui affectera directement l'orientation du bras de levier pendant le cycle (Cecchini et al., 2008).

Au centre du domaine moteur se trouvent sept feuillets β et leurs boucles associées. Les trois derniers brins forment ce qu'on appelle le transducteur (Figure 0.3) et jouent un rôle critique dans le contrôle des réarrangements dans le domaine moteur (Cecchini et al., 2008). Le transducteur subit une déformation au fur et à mesure de la progression du cycle de la myosine et on pense qu'il est le cœur de la transduction chimio-mécanique, régulant la fermeture de la fente de liaison à l'actine (Coureux et al., 2004; Robert-Paganin et al., 2020; Wulf et al., 2016).

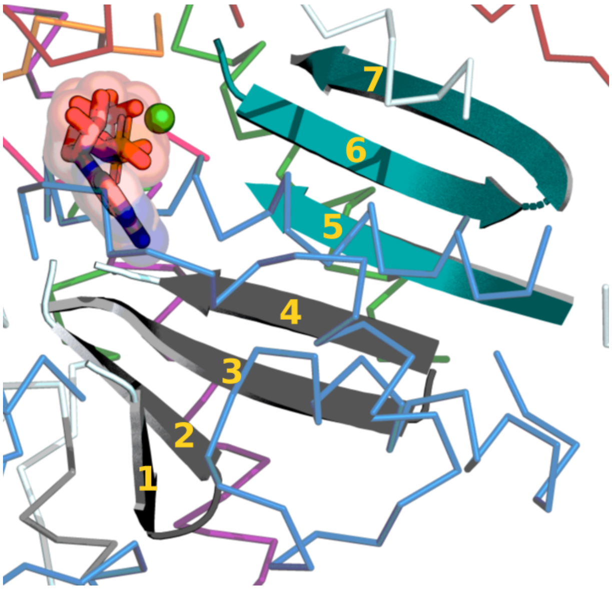


Figure 0.3: Feuillets β centraux. En gris foncé, les quatre premiers feuillets β et en sarcelle les trois derniers qui constituent le transducteur.

Aujourd'hui, grâce à la disponibilité croissante de structures cristallographique aux rayons X à haute résolution du sous-fragment-1 (S1), nous pouvons effectuer des analyses détaillées et faire la lumière sur la transduction chimio-mécanique dans les myosines.

Myosinses: cycle de génération de force

L'interaction des têtes de myosine et de l'actine est un processus cyclique piloté par l'hydrolyse de l'ATP : (Mesentean et al., 2007). Le processus est décrit par le cycle de Lymn-Taylor (Taylor, 1980), une séquence d'interactions entre la myosine, l'actine et l'ATP (voir Figure 0.4). Selon le cycle de Lymn-Taylor, les têtes de myosine se lient fortement à l'actine en l'absence de nucléotide pour ensuite se dissocier du filament par un mécanisme piloté par l'ATP : (Taylor, 1980). La production efficace de travail mécanique est réglée par un système raffiné de communication allostérique dans le domaine moteur (Cecchini et al., 2008; Mesentean et al., 2007). Le principal réarrangement structurel qui se produit au cours du cycle de l'actomyosine est le changement d'orientation du convertisseur et du bras de levier. Le cycle de l'actomyosine est divisé en deux phases ou oscillations, nommées respectivement Power-stroke et Recovery-stroke. Le power-stroke est la phase du cycle impliquée dans le mécanisme de génération de force, déclenchée par la fixation de la F-actine (Houdusse and Sweeney, 2016). L'étape clé du mécanisme de génération de force est la libération de P_i (P_i release step) (Llinas et al., 2015) suivie d'un réarrangement majeur de la position du bras de levier (Houdusse and Sweeney, 2016). Le mouvement du bras de levier pousse la myosine vers l'avant le long de la F-actine (Shiroguchi et al., 2011). Une fois que la P_i s'échappe du site actif, le domaine moteur est étroitement lié à la F-actine, dans l'état dit ADP (Houdusse and Sweeney, 2016) (voir Figure 0.4). L'état suivant, libéré des nucléotides et présentant la plus grande affinité pour la F-actine (De La Cruz et al., 1999), est connu sous le nom d'état de rigueur (R) (Wulf et al., 2016). Une molécule d'ATP induit la dissociation de la F-actine, et le moteur passe dans un état connu sous le nom d'état de pré-récupération ou de post-rigueur (PR) (Houdusse and Sweeney, 2016; Kühner and Fischer, 2011; Mesentean et al., 2007). Le power-stroke se termine et le recovery-stroke commence avec la tête de myosine dans une nouvelle position sur le filament d'actine. Le recovery-stroke est le mouvement fondamental qui réoriente le bras de levier dans une position amorcée, appelée position de pre-power-stroke (PPS), lorsque la tête est détachée de l'actine et que l'ATP est lié dans le site actif (Shiroguchi et al., 2011) (voir Figure 0.4). PPS et PR existent en équilibre l'un avec l'autre et PPS est stabilisé par $\cdot ADP \cdot P_i$, produits de l'hydrolyse de l'ATP (Cecchini et al., 2008; Yu et al., 2007). En présence d'actine il y a la transition de l'état PPS à l'état P_iR (Llinas et al., 2015; Wulf et al., 2016) et un nouveau power-stroke peut commencer.

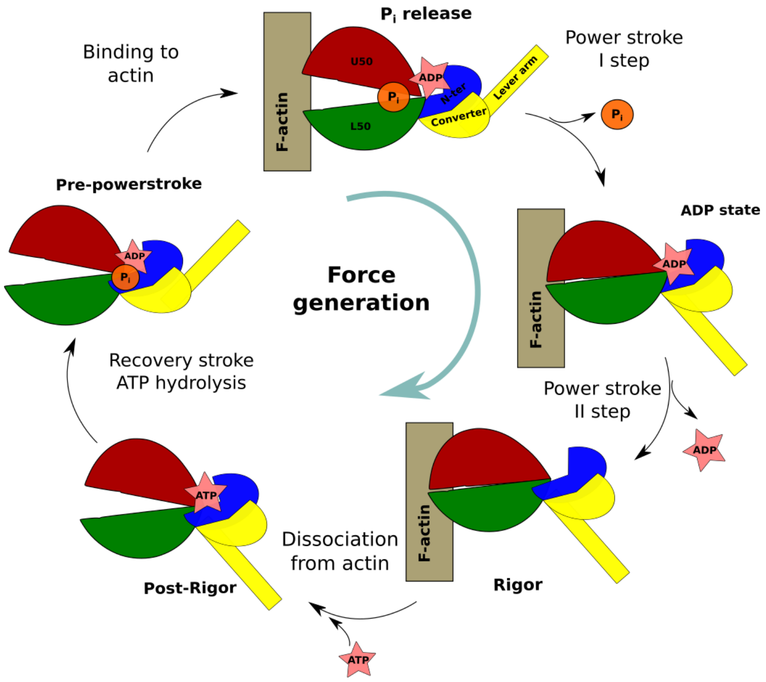


Figure 0.4: Le cycle ATP-ase de la myosine. La figure est inspirée de celle réalisée par Wulf et ses collègues dans (Wulf et al., 2016).

L'allostérie et la conception de médicaments allostériques

Les protéines allostériques, telles que les myosines, sont des molécules dans lesquelles de petits changements structurels distaux sont amplifiés en changements conformationnels et fonctionnels. Ce mécanisme est est connu sous le nom d'allostérie et a été étudié pendant plus de 60 ans (Bohr et al., 1904; Koshland et al., 1966; Monod et al., 1965). Aujourd'hui, de nombreuses approches informatiques, allant de la bioinformatique aux simulations de dynamique moléculaire avec des techniques de échantillonnage accélérée, nous permettent de dévoiler des perspectives mécanistiques possibles dans les processus allostériques et la coévolution (Bolhuis et al., 2002; Clarke et al., 2016; Cui and Karplus, 2008; Greener and Sternberg, 2018; Hilser et al., 2012; Livesay et al., 2012). Puisque les protéines allostériques remplissent des fonctions vitales, de la transduction de signaux, au mouvement en passant par la régulation des enzymes et l'expression des gènes (Tsai and Nussinov, 2014) et que leur dérégulation est associée à des pathologies humaines graves (Huang et al., 2010; Munro et al., 2009; Taly et al., 2009; Yang et al., 2021), elles deviennent des cibles thérapeutiques fondamentales (Greener and Sternberg, 2018; Huang et al., 2019). Le défi consiste à concevoir des modulateurs allostériques, de petites molécules capables de se lier à un site spécifique, appelé "site allostérique", éloigné du site orthostérique (Huang et al., 2010), afin d'améliorer une fonction spécifique. Ces molécules garantiraient une toxicité moindre et une grande sélectivité, en contournant les problèmes de résistance aux médicaments, causés par les mutations du site orthostérique (Lu et al., 2019; Wenthur et al., 2014). Pour réaliser efficacement une campagne de découverte de nouveaux modulateurs, on peut s'appuyer sur des prédictions in silico, qui accélèrent le processus (Gorgulla et al., 2020; Lionta et al., 2014). L'une des approches informatiques les plus utilisées pour cribler des chimiothèques ultralarges de molécules chimiques est appelée criblage virtuel à haut débit (vHTS) (Patrick Walters and Wang, 2019). Une approche largement utilisée dans le vHTS est la découverte de médicaments basée sur la structure (SBDD) (Lionta et al., 2014). La SBDD repose sur l'étude des interactions entre la cible biologique d'intérêt et les ligands afin de comprendre la base moléculaire des interactions que l'on veut étudier (Lionta et al., 2014). L'objectif principal du vHTS est de prédire les affinités de liaison des ligands examinés pour la cible d'intérêt (Patrick Walters et al., 1998) et cela peut être fait avec différents algorithmes (Cournia et al., 2017; Li et al., 2021).

La prédiction précise des affinités de liaison dans le vHTS reste un défi (Li et al., 2021). Figure 0.5 un aperçu du processus SBDD, depuis les composés filtrés extraits de la chimiothèque jusqu'à la prédiction finale du ou des résultats intéressants qui pourraient devenir des candidats médicaments.

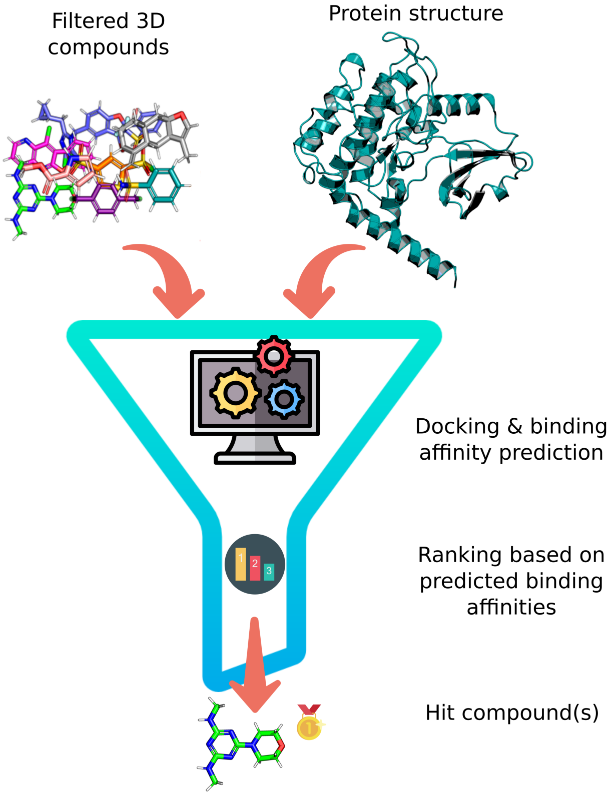


Figure 0.5: Schéma de la SBDD. La procédure prévoit différentes étapes alimentées par des méthodes de calcul. En entrée, on trouve une série de structures chimiques 3D de composés préalablement préparés et filtrés (voir le texte) et une structure de la protéine cible. Les ligands seront dockés afin de générer des complexes protéine-ligand. Les affinités de liaison sont prédites pour chaque mode de liaison retenu. Les ligands sont ensuite triés en fonction de ces prédictions. Enfin, certaines molécules sont sélectionnées pour l'étape suivante de la découverte de médicaments (tests *in vitro* et optimisation ultérieure du candidat).

Dans cette thèse, nous allons nous concentrer à la fois sur les protéines allostériques, en parlant de la myosine cardiaque et de la myosine VI, et sur la SBDD que nous avons réalisée avec notre programme interne, ChemFlow, que j'ai contribué à développer.

Pour les deux myosines, nous avons utilisé des simulations de dynamique moléculaire (MD) pour étudier la dynamique de la protéine en solution dans un état spécifique du cycle de génération de force : l'état PPS.

De plus, nous avons utilisé certaines approches computationnelles, allant du docking moléculaire, du rescoring de l'énergie libre avec des méthodes (MM/PB(GB)SA) aux simulations de dynamique moléculaire à accélération aléatoire (τRAMD) pour comprendre comment définir les activateurs des myosines cardiaques. Nous avons également réalisé une campagne vHTS complète sur la myosine VI afin de trouver quelques molécules qui pourraient être des inhibiteurs dans le PPS.

La myosine cardiaque β dans l'état PPS : analyse structurale et modulation allostérique

La myosine cardiaque β est l'isoforme lente du muscle cardiaque principalement exprimée dans le ventricule du cœur (Walklate et al., 2021a) et impliquée dans des dysfonctionnements graves, tels que l'insuffisance cardiaque (HF), qui constituent l'une des principales causes de décès dans la société actuelle (Afzal, 2021; Gaziano et al., 2010). Compte tenu de l'impact croissant des maladies cardiovasculaires (CVD), qui représentaient 32 % des décès dans le monde en 2019 (Benjamin et al., 2019) avec une tendance qui ne semble pas s'améliorer, il est essentiel d'étudier des traitements innovants et sûrs. De nouvelles voies thérapeutiques pourraient impliquer des modulateurs qui ciblent directement la myosine cardiaque pour réguler à la hausse ou à la baisse son activité (Hwang and Sykes, 2015), sans déclencher de mécanismes de signalisation de second messager qui aggravent l'état des patients. Cette considération a renforcé notre décision d'étudier d'un point de vue structurel et pharmacologique l'isoforme β de la myosine cardiaque, dont la contraction est altérée dans l'HF. Notre analyse a commencé par le choix de la structure cristallographique du β-CM bovin en PPS (Planelles-Herrero et al., 2017). En effet, notre objectif est de concevoir des activateurs non compétitifs de la myosine cardiaque, capables de cibler directement le moteur, sans altérer l'homéostasie calcique intracellulaire (Liu et al., 2015; Morgan et al., 2010). Nous pensons que les activateurs de la myosine cardiaque accélèreraient la transition vers l'état PPS (Liu et al., 2015; Winkelmann et al., 2015), l'état de début du power-stroke, en déplaçant l'équilibre conformationnel vers le PPS. De cette manière, il y aura plus de têtes prêtes à lier l'actine et à augmenter la contractilité cardiaque qui est altérée dans l'HF (Liu et al., 2015; Planelles-Herrero et al., 2017). L'équilibre d'hydrolyse sera donc :

$$\mathrm{M + ATP} \rightleftharpoons \mathrm{M + ADP + P_i} \tag{0.1}$$

où M représente la myosine et P_i le phosphate inorganique, généré par l'hydrolyse de l'ATP.

Nous sommes partis de la structure cristallographique la plus récente et de la plus haute résolution (résolution de 2,45 Å) résolue à ce jour pour le fragment S1 de la myosine cardiaque dans le PPS (Planelles-Herrero et al., 2017). Le domaine moteur est résolu en complexe avec ADP-VO4-MG2+ dans le site actif et l'activateur allostérique Omecamtiv Mecarbil (OM) (Morgan et al., 2010) dans sa poche, à l'interface entre le sous-domaine N-terminal, le transducteur, l'hélice de relais (RH) et le convertisseur (Planelles-Herrero et al., 2017) (Figure 0.2). Le fait que la structure soit la tête de myosine. Dans l'état PPS avec un activateur allostérique co-cristallisé dans sa poche, rend cette structure initiale encore plus attrayante pour nos objectifs pharmacologiques.

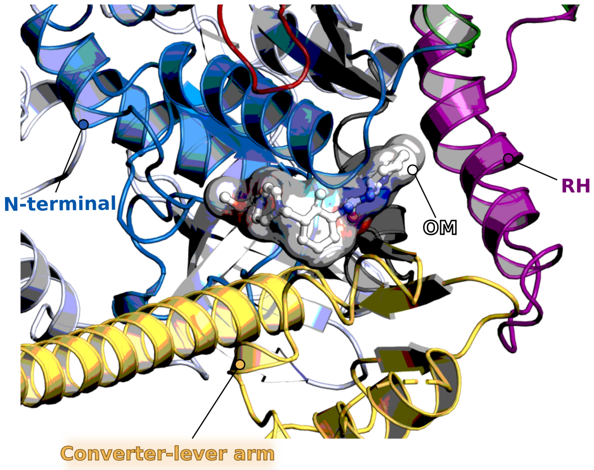


Figure 0.6: Zoom sur l'espace occupé par OM dans le β-CM bovin (code PDB : 5N69.pdb). OM est représentée en bâtonnets et entourée de sa surface, colorée selon la polarité des atomes. En cartoon, les sous-domaines autour du modulateur : RH en violet ; N-terminal en bleu et convertisseur en jaune.

Modélisation de β-CM & stabilité du PPS

Dans un premier temps, nous avons modélisé la partie manquante de la structure cristallographique afin d'obtenir une structure initiale complète. Notre première analyse visait à étudier la stabilité de l'état PPS dans la myosine cardiaque avec et sans le modulateur allostérique OM. Nos analyses ont montré que notre modèle initial en absence d'OM était très instable. Ceci a conduit à l'amélioration du modèle, en changeant l'état de protonation de certains résidus. En particulier, nous avons choisi un état de protonation non standard pour les résidus de U50, swII et le sous-domaine N-terminal (voir Figure 0.7).

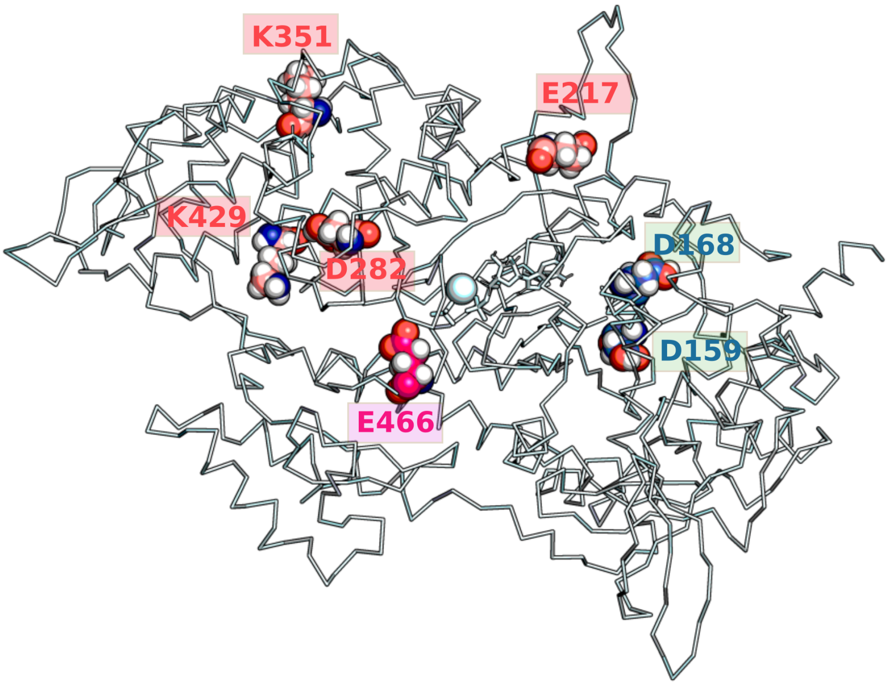


Figure 0.7: Dans les sphères, les sept résidus auxquels nous avons attribué un état de protonation non standard. En rouge saumon les résidus situés dans l'U50 ; en magenta un résidu dans la boucle swII et en bleu deux résidus du sous-domaine N-terminal.

Après cette modification du modèle, la structure était stable aussi bien avec que sans OM.

Conception d'activateurs allostériques de la myosine cardiaque

Nous avons utilisé la structure PPS liée à la OM la plus représentative comme coordonnées initiales pour notre investigation, extraite d'une simulation MD convergé de notre dernier modèle. Nous avons choisi comme jeu de données une série d'analogues d'OM conçus par Morgan et ses collègues, pour lesquels les activités biochimiques sont quantifiées en AC40 (Morgan et al., 2010). Nous avons utilisé ChemFlow pour docker ces ligands dans β-CM et pour effectuer des simulations MD, à partir desquelles nous avons prédit les affinités de liaison en utilisant MMGBSA. Le premier protocole utilisé n'a pas donné les résultats escomptés, car les molécules étaient mal classées si on les comparait à leur activité expérimentale. Nous avons alors amélioré le protocole en changeant de programme de docking et en générant un meilleur mode de liaison pour toutes les molécules. Nous avons obtenu le classement souhaité des molécules (voir Figure 0.8). Lorsque les affinités de liaison prédites ont été comparées aux activités expérimentales, nous n'avons remarqué aucune corrélation. Nous avons calculé le temps de résidence de chaque composé avec des simulations τ-RAMD, puis nous avons à nouveau calculé la corrélation avec les activités expérimentales. Une fois de plus, il n'y avait pas de corrélation, mais lorsque nous avons comparé les deux méthodologies, elles ont partagé une corrélation intéressante de 0,58.

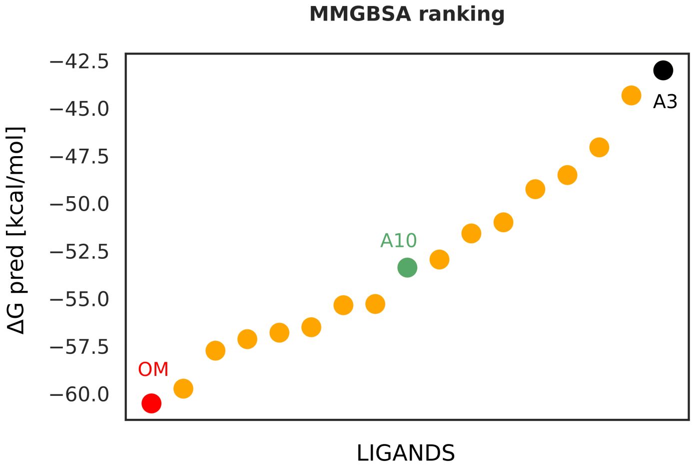


Figure 0.8: Classement des composés de Morgan en fonction des scores MMGBSA. Dans le point rouge OM (ΔG prédit = -60,5 kcal/mol) ; dans le point noir A3 (ΔG prédit = -43 kcal/mol) et dans le point vert A10 (ΔG prédit = -53,4 kcal/mol). L'OM est classée comme meilleur ligand selon les scores du MMGBSA.

Transition spontanée de la Myosine VI du PPS à un autre état intermédiaire

La myosine VI est une myosine non conventionnelle et c'est la seule qui se déplace vers l'extrémité négative du filament d'actine (Pylypenko et al., 2011). De nombreuses caractéristiques structurelles rendent cette protéine unique (Park et al., 2006; Pylypenko et al., 2011; Rock et al., 2000) et nous avons décidé d'explorer la stabilité de l'état PPS qui est un état de transition du cycle de production de la force (Blanc et al., 2018). Nous avons commencé par modéliser la structure porcine de la PPS (code PDB : 2v26), résolue en complexe avec $ADP \cdot VO_4^3-$ (Ménétrey et al., 2007). Nous avons effectué des simulations MD et nous avons remarqué une grande instabilité du domaine moteur, en particulier une remarquable souplesse du sous-domaine convertisseur (Figure 0.9).

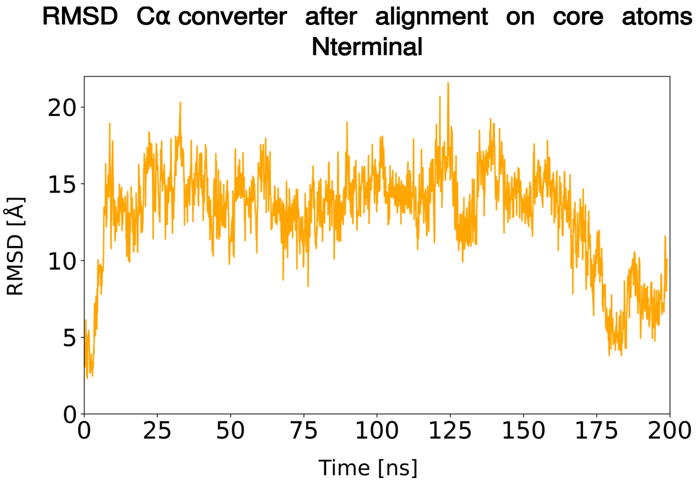


Figure 0.9: Série temporelle sur 200 ns de la RMSD des atomes $C\alpha$ du convertisseur après alignement sur les atomes centraux du sous-domaine Nterminal de la structure initiale. La structure initiale a une RMSD de 0,1 Å par rapport aux coordonnées cristallographique de l'état PPS (2v26.pdb).

Le convertisseur semble tourner et se translater vers une position inférieure, à mi-chemin entre PR et PPS, ce qui suggère que la structure pourrait explorer l'état précédent du cycle de l'actomyosine, appelé état de prétransition (PTS) (Blanc et al., 2018). Après avoir suivi une série d'observables géométriques sur la trajectoire, nous avons défini une séquence plausible d'événements décrivant cette transition spontanée que nous avons capturée avec une simulation MD classique (voir Figure 0.10).

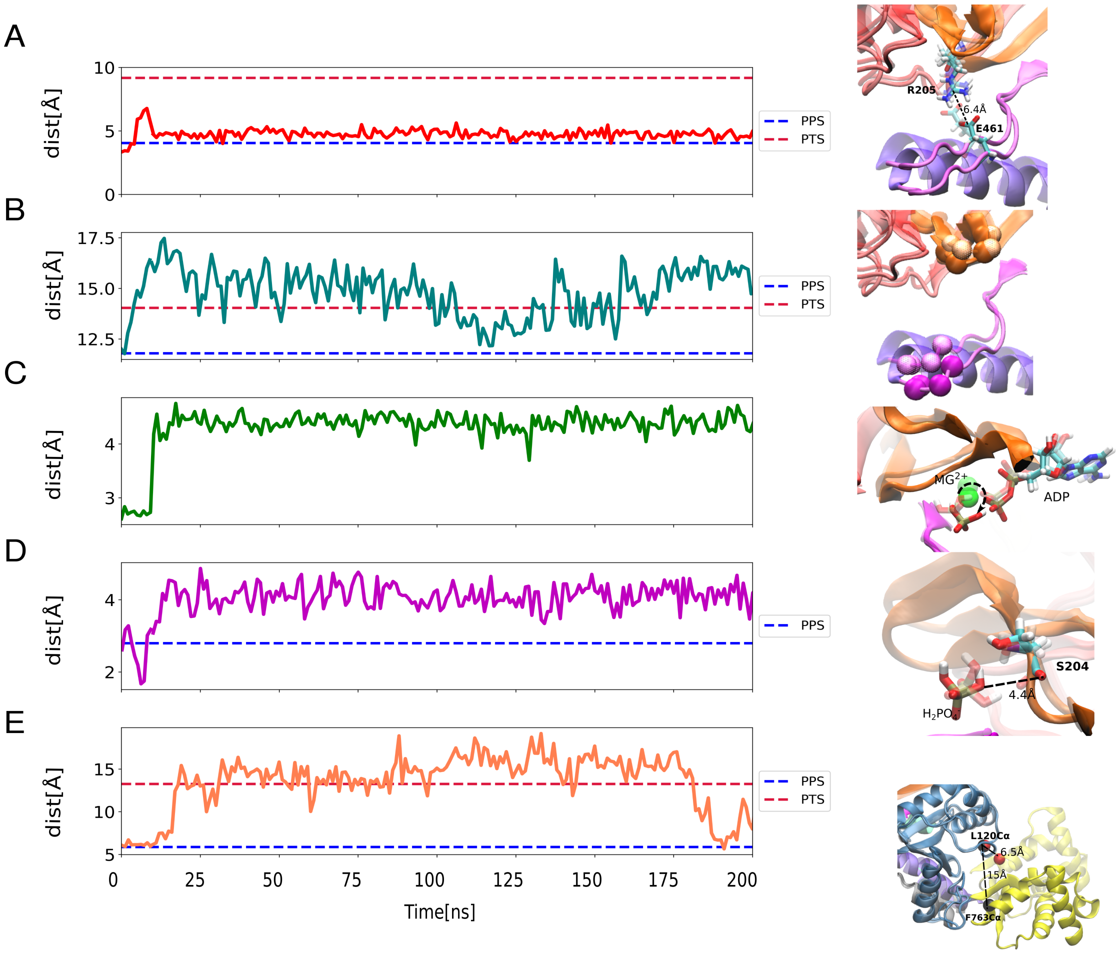


Figure 0.10: Séquence possible d'événements menant à la transition vers un état de type PTS, capturée dans 200 ns MD.

Voici également un lien pour la vidéo qui montre la séquence des réarrangements dans la myosineVI : `https://zenodo.org/record/6907258#.YuAP8exBxH4`.

Pour explorer plus en détail la stabilité du PPS dans la myosine VI, nous avons exécuté 10 répliques de 200 ns avec des simulations MD tout-atome et nous avons suivi les mêmes observables. Dans trois des dix répliques, nous observons la même souplesse du convertisseur et dans deux d'entre elles, dans lesquelles le convertisseur se déplace vers une position plus basse, nous avons enregistré la même séquence d'événements pour la transition que celle définie pour la première simulation. De plus, après avoir suivi la déformation du transducteur, qui joue un rôle central dans la transition conformationnelle de la myosine, nous avons observé que les brins β s'éloignent d'une configuration PPS et préfèrent une configuration PR ou P_iR. Cela suggère que le moteur pourrait se diriger vers un état à plus faible énergie libre, sans libération inutile de produits d'hydrolyse du site actif.

vHTS sur Myosin VI : inhibiteurs allostériques

La myosine VI joue un rôle dans les processus d'endocytose, ainsi que dans la préservation de la morphologie du Golgi et de l'intégrité des stéréocils de l'oreille interne (Jung et al., 2006). Dans les cellules de mammifères, il contrôle également l'adhérence à la jonction des cellules épithéliales (Sweeney et al., 2007) et il a été démontré qu'elle est fortement exprimée dans le noyau, modulant l'activité transcriptionnelle de l'ARN polymérase II (Vreugde et al., 2006). Chez l'homme, la surexpression de la Myosine VI est observée dans le cancer du sein et de la prostate de grade moyen ainsi que dans le cancer de l'ovaire et pour ce dernier, il a également été trouvé une corrélation avec la malignité de la maladie (Dunn et al., 2006; Hari-Gupta et al., 2022; Makowska et al., 2015; Vreugde et al., 2006). En raison du taux d'incidence élevé de ces cancers dans le monde, qui va continuer à s'aggraver (Lukasiewicz et al., 2021; Rahib et al., 2021) et du manque de traitements pour ces pathologies, la MyosineVI représente une cible de choix dans la thérapie du cancer. Étant donné les avantages qui peuvent dériver du développement de nouveaux inhibiteurs allostériques sélectifs pour cette protéine, nous avons réalisé une campagne vHTS sur MyosinVI, en utilisant une approche basée sur la structure (SBDD) avec notre software interne ChemFlow. Nous avons criblé environ 67 000 molécules et, à la fin, nous avons classé par ordre de priorité 8 composés qui ont été achetés pour les tests in vitro. Des essais d'ATPase sont en cours au laboratoire Sweeney, et des expériences cristallographiques au laboratoire Houdusse, afin d'obtenir un co-cristal de Myosin VI avec un bon liant.

Protéine kinase : inhibiteurs potentiels à coordination métallique

Les protéines kinases sont des enzymes capables de médier le transfert réversible d'un groupe phosphate à d'autres protéines (Attwood et al., 2021). Les protéines kinases sont une cible médicamenteuse potentielle (Parsons and Parsons, 2004) pour de petites molécules qui pourraient agir comme des médicaments anticancéreux (Scarpi-luttenauer, 2022; Scarpi-Luttenauer et al., 2022), puisque leur dérégulation provoque une croissance cellulaire anormale dans certaines tumeurs (Downward, 2003; Scarpi-Luttenauer et al., 2022). Depuis le début des années 2000, l'accent a été mis sur les complexes de coordination pour concevoir de nouveaux inhibiteurs de protéine kinase, en imitant la structure de l'inhibiteur fort nommé Staurosporine, qui est trop toxique et non sélectif pour trouver une utilité comme traitement possible (Attwood et al., 2021; Gani and Engh, 2010; Tamaoki et al., 1986).

Conclusions générales

Nous avons exploré la transuction chimio-mécanique dans deux myosines, en étudiant l'état PPS. Nous avons confirmé une hypothèse antérieure de flexibilité caractéristique du convertisseur dans la MyosineVI dans cet état (Ménétrey et al., 2005) et grâce à des simulations MD nous avons capturé une transition spontanée du domaine moteur vers un état intermédiaire. La myosine cardiaque semble plus stable que la MyosinVI et cela ouvre la perspective d'une étude future visant à comprendre si la grande flexibilité observée dans la MyosinVI est caractéristique de toutes les myosines processives, par rapport aux myosines non processives, comme la myosine cardiaque. Nous avons également étudié ces protéines d'un point de vue pharmacologique. Nous avons développé un pipeline capable de discriminer les activateurs puissants de la myosine cardiaque, même s'il n'est pas encore clair s'il peut exister une corrélation entre l'affinité et l'activité. En ce qui concerne la MyosinVI, nous attendons les résultats de nos prédictions d'inhibiteurs plausibles du moteur. En utilisant notre logiciel interne develoe3d pour SBDD, nous avons collaboré avec le groupe de Mobian pour discriminer de bons inhibiteurs de protéine kinase prometteurs qui ont été développé par Scarpi et Mobian avec une chimie de coordination.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1

MYOSINS

Back in 1864 the physiologist Wilhelm Kühne first isolated a protein from the muscle and called it "myosin" (Kühne, 2011). Since then, several studies have been conducted to explore the mechanism of action of these cytoplasmic proteins, classified as actin-based molecular proteins (Howard, 1997). This classification is derived from their capability of moving across actin filament, by converting chemical energy, derived from ATP hydrolysis, into mechanical work (Rdowicz, 2002). The myosin superfamily counts 35 classes, based on the sequence analysis of the globular portion of the structure, known as motor domain (Odronitz and Kollmar, 2007). The classification in several classes underlines how each class has evolved to be specialized in a unique and specific cellular function (Hoklotubbe, 2018; Odronitz and Kollmar, 2007). Myosins are defined as "conventional" or "unconventional", according to which class they belong to (Brawley and Rock, 2009). In particular, the isoforms involved in muscular contraction (myosins 2) are the conventional ones and were the first ones to be discovered (Hartman and Spudich, 2012). All the non-muscular myosins are denominated unconventional and they carry on many different functions in cells, such as cell crawling, exo- and endocytosis, and cell division and they could even be involved in nuclear transcription, chromatin remodeling, and signal transduction in the inner ear, just to mention a few (Batters and Veigel, 2016; Hartman and Spudich, 2012; Woolner and Bement, 2009a). Myosins that exist as a dimer in the cells are further classified into processive and non-processive motors, respectively, according to their capability of preventing simultaneous detachment from actin filament of both myosin heads, under loading (Howard, 1997; Lee Sweeney and Houdusse, 2004; O'Connell et al., 2007). Thanks to this characteristic, when one head of a processive motor detaches from the actin, upon ATP-binding, the other head remains strongly bound to the track until the detached head hydrolyses the ATP and rebinds to a more distant site on actin (Lee Sweeney and Houdusse, 2004). Conventional myosins fall into the classification of non-processive motors (Greenberg et al., 2017) and this mechanochemical behavior was discovered by the pioneering study of a three-bead assay conducted in 1994 by Spudich and his team on skeletal myosin (Finer et al., 1994). In this study, two beads were attached near the ends of a taut filament of actin and held in an optical trap (Finer et al., 1994). This assembling was then lowered on a surface with a third bead coated with skeletal myosin heads and attached to a microscope (Finer et al., 1994). Once myosin has bound F-actin, the beads were displaced and a single step movement along the filament could have been observed before detachment (Finer et al., 1994; Greenberg et al., 2017). The processivity of unconventional myosins, such as Myosin V (Rief et al., 2000), VI (Altman et al., 2004) and X (Takagi et al., 2014), has been demonstrated in the same manner. Processive myosins share a slow ADP release from the active site, when bound to actin, and are capable of larger steps on the filament (Batters and Veigel, 2016). The ratio of ADP release from the active site affects the duty ratio of myosins (Batters and Veigel, 2016; Lee Sweeney and Houdusse, 2004). The duty ratio indicates the time spent by the motor head populating the state tightly bound to actin (force generation state), compared to the overall kinetic cycle time (Batters and Veigel, 2016; Lee Sweeney and Houdusse, 2004). Non-processive myosins have low duty ratio in absence of load (Lee Sweeney and Houdusse, 2004), whereas processive myosins have a higher duty ratio, to tune their processive movement (Rock et al., 2000). The directionality of myosins on actin is strictly connected to the functions they carry in the cells (Cramer, 2000). For instance, plus-end-directed myosins are involved in the trafficking of vesicles towards the plasma membrane, as well as the transportation of cargoes towards the cell surface (Cramer, 2000). On the other hand, minus-end-directed myosins are involved in endocytosis and movement of cargo towards the interior of the cell (Cramer, 2000; O'Connell et al., 2007). Within the characterized processive myosins, Myosin VI was thought to be the only one able to move towards the minus end of the polarized filament of actin, both in monomeric and dimeric form (Schliwa, 1999; Wells et al., 1999). Instead, the same reverse movement is shared by another processive myosin, named Myosin IXb (Myo9b) (Inoue et al., 2002). This finding disputes the idea that the reverse directionality of Myosin VI is regulated by a unique sequence located at the level of the neck region of its heavy chain (Park et al., 2006; Wells et al., 1999). It opens the debate about the possibility that there may be other minus-end-directed myosins, for which directionality is not driven by the presence of a unique insert between the motor domain and the neck region, but rather by elements in the core of the motor domain (Homma et al., 2001). The exact definition of the structural elements involved in the determination of reverse movement is not yet clarified and requires further investigations (Homma et al., 2001; Inoue et al., 2002).

1.1 MYOSINS' ARCHITECTURE: A SCHEMATIC VIEW

Despite their ample classification, all myosins share the same structural organization. The dimers look symmetric and each monomer is composed of one heavy chain (MHC) and two light chains (MLC), named Essential and Regulatory light chain (ELC and RLC), respectively (AL-Khayat, 2013). The MHC weights around 223 kDa (Jin et al.,

2017) and presents a globular head and a long tail of variable length (Syamaladevi et al., 2012) (Figure 1.1). The globular portion is known as Subfragment-1 (S1) and includes an MHC folded with two MLCs (England and Loughna, 2013). In S1 there is the *motor domain*, also known as myosin head, which is the most conserved domain of the protein (Foth et al., 2006; Howard, 1997). Myosin head is connected to the long C-terminal tail with a central neck region, that constitutes the regulatory domain (England and Loughna, 2013). At the level of this neck region, indeed, are wrapped the MLCs, members of the calmodulin (CaM) and CaM-like proteins (England and Loughna, 2013; Heissler and Sellers, 2014).

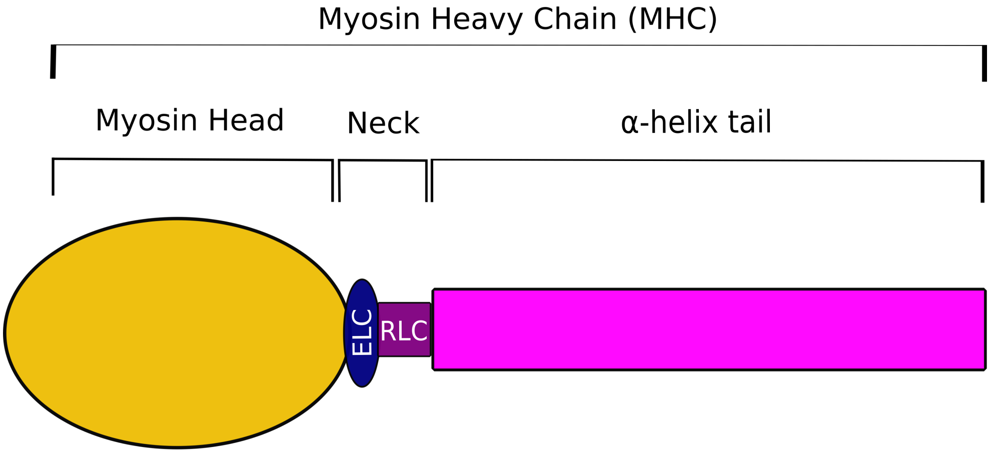


Figure 1.1: Sketch of the Myosin Heavy Chain (MHC).

In the MHC reside the ATP binding sites and the actin-binding interface (England and Loughna, 2013; Syamaladevi et al., 2012). In particular, the actin-binding region is found between two subdomains named, respectively, Upper50 (U50) and Lower50 (L50), due to their dimension of 50 kDa (Figure 1.2) (Syamaladevi et al., 2012; Wulf et al., 2016). U50 and L50 form a cleft that is open when the myosin is in the weak affinity state for actin and, instead, is closed when myosin is tightly bound to actin (Wulf et al., 2016). The change in the affinity for actin and, accordingly, the movement of myosin on its track, is powered by actin itself and by the cyclic binding of ATP, the substrate of myosin (Llinas et al., 2015; Wulf et al., 2016). Actin serves as ATP-ase activator, promoting the release of ATP hydrolysis products (Llinas et al., 2015; Wulf et al., 2016). The molecule of ATP binds in the active site of myosin, located in the core of the motor domain, between the U50 and N-terminal subdomains (Syamaladevi et al., 2012).

1.2 A DETAILED GAZE ON THE MOTOR DOMAIN

The motor domain is the most conserved portion in all myosins since here reside the structural and functional elements, responsible for

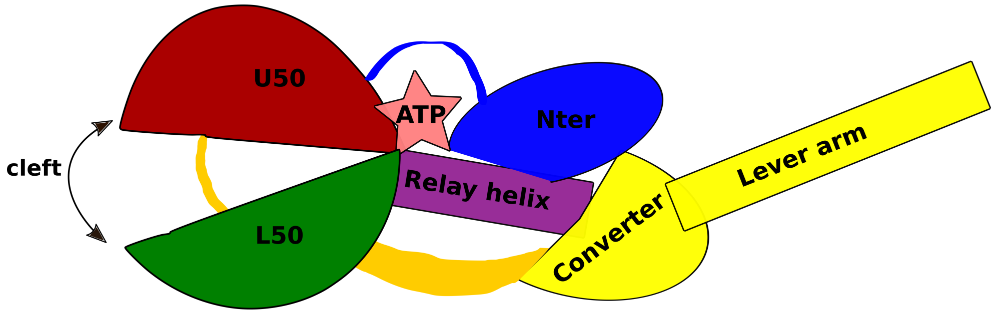


Figure 1.2: Blueprint of the main subdomains of the myosin head. Between Upper50 (U50), in red, and Lower50 (L50), in green, there is the actin-binding interface (cleft). Between the U50 and the N-terminal subdomain (Nter), instead, there is the active site where ATP binds (pink star). Converter and lever arm subdomains are depicted in yellow. The converter is linked to the rest of the motor domain with the Relay helix (part of L50). The rotation of the converter is fundamental for amplifying structural changes in the motor domain during the force generation cycle to accomplish the final rotation of the lever arm.

force generation mechanism (Howard, 1997).

Nowadays, thanks to the ever-growing availability of high-resolution X-ray structures of the subfragment-1 (S1) we can perform detailed analyses and shed light on the chemo-mechanical transduction in myosins. In Figure 1.3 a cartoon representation of the motor domain in bovine cardiac myosin (PDB code: 5N69) (Planelles-Herrero et al., 2017). U50 and L50 subdomains are linked by several loops and describe a wide cleft whose opening changes according to the step of the cycle in which the molecular motor is (Llinas et al., 2015; Wulf et al., 2016) (see Section 1.3). The nucleotide binding pocket (active site) is located at the interface between U50, N-terminal and L50 subdomains (Cecchini et al., 2008).

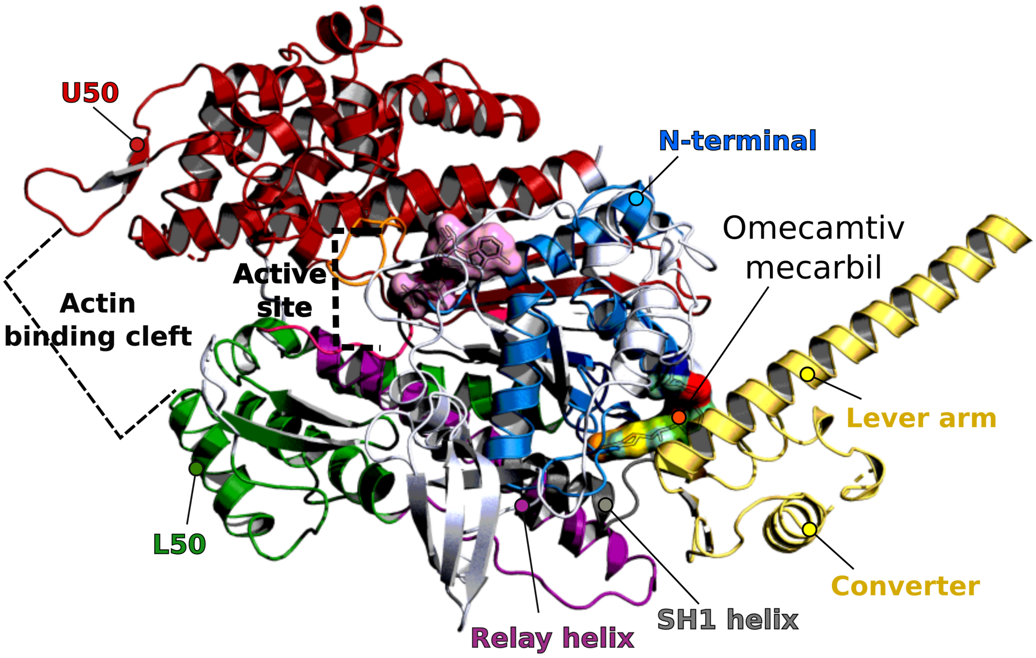


Figure 1.3: Cartoon representation of the motor domain in cardiac myosin (PDB code: 5N69 (Planelles-Herrero et al., 2017)). The main four subdomains are represented as follows: U50 in red; L50 in green; N-terminal in blue, the converter and the lever arm in yellow. Between the U50 and the N-terminal subdomain (Nter), there is the active site in which ADP-Vanadate (VO4) is depicted as a pink surface. In purple the Relay helix (RH), which is part of the L50 subdomain, and in grey SH1 helix that connects the converter subdomain to the N-terminal one. In rainbow, the allosteric activator Omecamtiv mecarbil (Morgan et al., 2010) bound at the interface between RH, Converter, and N-terminal subdomains (Planelles-Herrero et al., 2017).

The active site is composed of three main loops named switch I (swI), P-loop, and switch II (swII) (Syamaladevi et al., 2012) (see Figure 1.4), which trap the nucleotide and mediate the ATP-hydrolysis (Geeves and Holmes, 1999; Mesentean et al., 2007). In particular, swII closes on the γ phosphate and hydrolyzes the ATP in ADP $\cdot$ P_i (Onishi et al., 2002). Other important connectors in the motor domain are the Relay helix (RH) and the SH1 helix (Cecchini et al., 2008; Llinas et al., 2015). These two elements are linking the converter subdomain to the rest of the architecture (Figure 1.3) (Cecchini et al., 2008). The converter is a flexible element of the motor domain and its rotation is a key motion that will directly affect the orientation of the lever arm during the cycle (Cecchini et al., 2008).

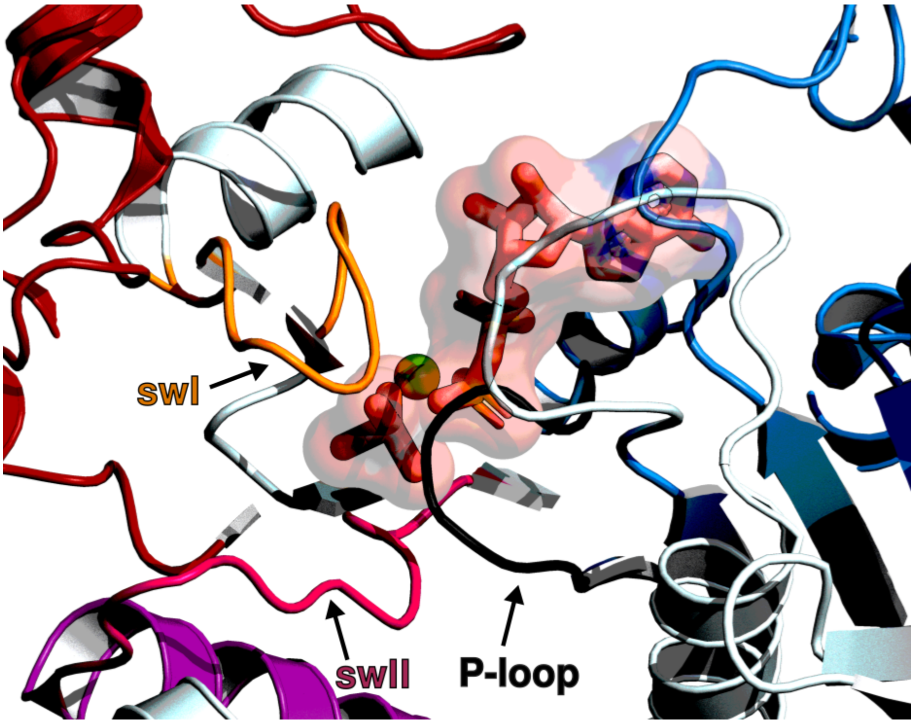


Figure 1.4: Zoom on the active site. In pink sticks the ADP-VO4 and their surfaces. The green sphere represents the Magnesium ion (MG) and the arrows point to the three connectors that facilitate the rearrangements of the subdomains: switch I (swI), P-loop, and switch II (swII).

Thanks to the rotation of the converter subdomain, small structural changes at the level of the active site are communicated and amplified, through a refined mechanism, that culminates in the rotation of the lever arm (Cecchini et al., 2008; Mesentean et al., 2007; Park et al., 2006). The lever arm is a long α-helix portion of variable length, attached to the converter subdomain (Figure 1.2). It consists of different calmodulin (CaM) and calmodulin-like (CaM-like) binding sites named IQ motifs (Ménétrey et al., 2012; Park et al., 2006; Warshaw, 2004). If we explore the core part of the motor domain, we find a region constituted of seven β-sheets and their associated loops. The last three strands shape the so-called transducer (Figure 1.5) and play a critical role in controlling the rearrangements in the motor domain (Cecchini et al., 2008). In particular, the transducer undergoes a distortion as the myosin cycle progresses and it is thought to be the heart of the chemomechanical transduction, regulating the closure of the actin-binding cleft (Coureux et al., 2004; Robert-Paganin et al., 2020; Wulf et al., 2016).

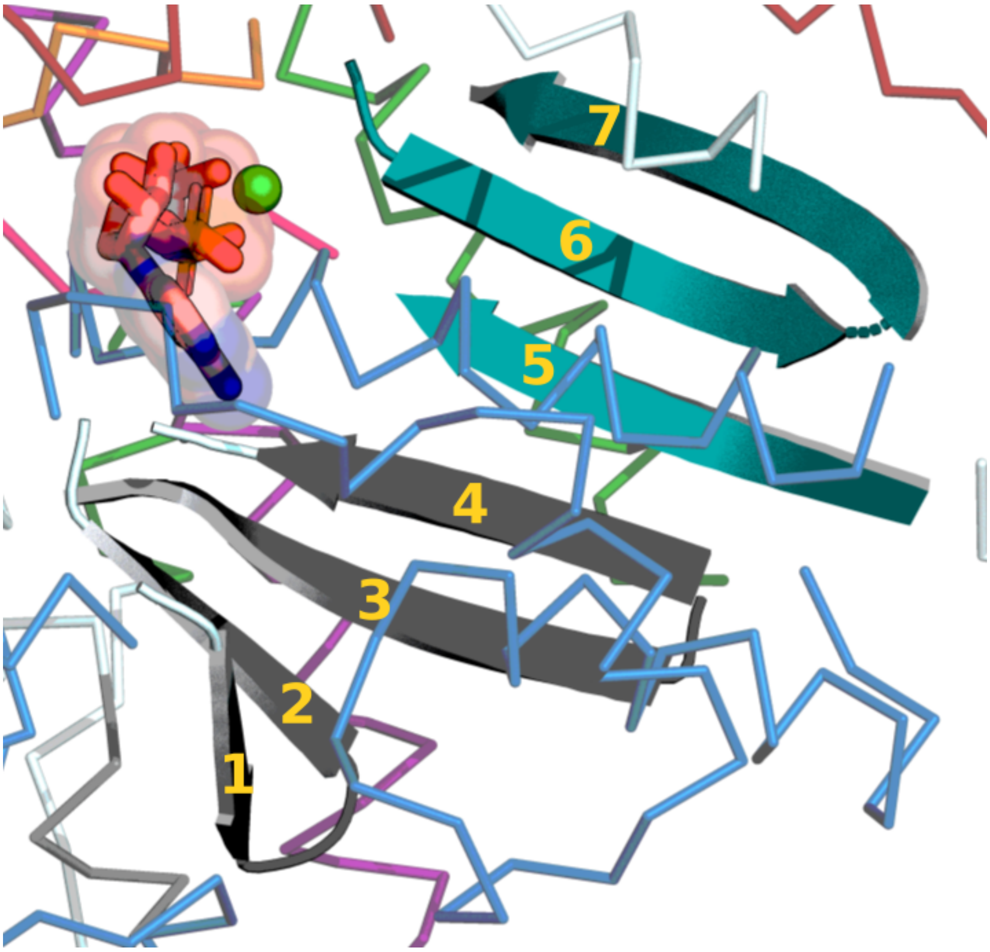


Figure 1.5: Central β-sheets. In dark grey, the first four β sheets, and in teal the last three constitute the transducer.

1.3 FORCE GENERATION CYCLE IN MYOSIN

How does myosin move on the actin filament? According to the most widely held theory, known as *swinging lever-arm hypothesis* and elaborated by Holmes in 1997, myosin can convert ATP-hydrolysis into mechanical work due to the swing of the lever arm (Holmes, 1997). This hypothesis is still considered the dominant principle of the process of myosin movement. It has been derived from the study of a structure (1MYS.pdb) of the S1 fragment, with the hydrolysis products trapped in the active site (Holmes, 1997). This structure had a different position of the tail (the lever arm) concerning previous existing X-ray structures (Holmes, 1997). Holmes realized that the tail is the movable portion of the myosin head, able to rotate around a fulcrum close to the SH1 helix (Holmes, 1997). The rotation of the lever arm, towards a cocked position, is driven by the binding of ATP and its subsequent hydrolysis in $ADP \cdot P_i$ (Holmes, 1997). The following release of hydrolysis products, tuned by actin, is coupled to structural rearrangements in the active site that are then amplified and revealed in the large swinging of the lever arm (Holmes, 1997; Wulf et al., 2016). The final rotation of the lever arm occurs to optimize strain-dependent ADP release (Sweeney and Houdusse, 2010).

1.3.1 Power-stroke & Recovery-stroke

The interaction of myosin heads and actin is a cyclic process driven by ATP hydrolysis (Mesentean et al., 2007). The process is described by the Lymn-Taylor cycle (Taylor, 1980), a sequence of interactions between myosin, actin, and ATP (see Figure 1.6). According to the Lymn-Taylor cycle, myosin heads strongly bind to actin in absence of nucleotide to then dissociate from the track with an ATP-driven mechanism (Taylor, 1980). The chemomechanical cycle encompasses the coupling of small changes at the level of the active site, with large structural changes in the actin-binding interface and converter-lever arm subdomain (Mesentean et al., 2007; Wulf et al., 2016) (see Figure 1.6). The efficient production of mechanical work is tuned by a refined system of allosteric communication in the motor domain (Cecchini et al., 2008; Mesentean et al., 2007). The major structural rearrangement that occurs during the actomyosin cycle is the change in the orientation of the converter and lever arm (Mesentean et al., 2007). The rotation of the lever arm has different amplitudes for various myosin isoforms, to control their step-size (Köhler et al., 2003). For example, in class II myosins, the rotation is of around 60^o (Mesentean et al., 2007), in Myosin I of around 90^o (Köhler et al., 2003), and Myosin VI of 180^o (Reifenberger et al., 2009). The actomyosin cycle is divided into two phases or swings, named, respectively, Power-stroke and Recovery-stroke.

1.3.1.1 *Power-stroke*

The power stroke is the phase of the cycle implied in the force generation mechanism, triggered by F-actin binding (Houdusse and Sweeney, 2016). The consensus states that the power stroke begins when actin catalyzes and controls the sequential release of P_i and $ADP - MG^{2+}$ from the active site (Houdusse and Sweeney, 2016; Llinas et al., 2015) (see Figure 1.6). The movement and the force are generated as a consequence of this release, coupled to conformational changes in the motor domain (Houdusse and Sweeney, 2016; Llinas et al., 2015). The key step for the force generation mechanism is the release of P_i (P_i release step) (Llinas et al., 2015), followed by a major rearrangement in the position of the lever arm (Houdusse and Sweeney, 2016). The motion of the lever arm prompts myosin forward along F-actin (Shiroguchi et al., 2011). The exact sequence of events, from the P_i released to the movement of the lever arm, remains controversial (Houdusse and Sweeney, 2016; Moretto et al., 2022). Is it the release of inorganic phosphate that triggers the power stroke or the main structural changes that occur during the force generation that gates the release of P_i? Very recently Moretto *et al* proposed a "multistep P_i release" hypothesis (Moretto et al., 2022). According to their study, P_i may bind various sites, outside the active site, in the release tunnel (namely back door) (Llinas et al., 2015; Moretto et al., 2022). Llinas *et al* proposed this tunnel as a possible site in Myosin VI where the inorganic phosphate could move before escaping (Llinas et al., 2015). Moretto and co-workers found the same evidence in myosin II, where plausible multiple binding sites for P_i are proposed (Moretto et al., 2022). Their findings would corroborate the hypothesis that P_i release precedes and causes the power-stroke (Moretto et al., 2022). Once the P_i escapes from the active site, the motor domain is tightly bound to F-actin, in the so-called ADP-state (Houdusse and Sweeney, 2016) (see Figure 1.6). In this step, the cleft closes to form a strong actin-myosin interface (Wulf et al., 2016). The closure of the cleft is mediated by an intense distortion of the transducer, whose β-strands twist (Wulf et al., 2016). The twisting of the β-strands triggers some movements of the U50 and N-terminal subdomain, tuning the release of the nucleotide (Wulf et al., 2016). The duration of the ADP-state varies among different myosins (Wulf et al., 2016). Myosins with high duty ratio exhibit longer ADP-state, respect myosins with lower duty ratio (Wulf et al., 2016). The following nucleotide-free state, with the highest affinity for F-actin (De La Cruz et al., 1999), is known as Rigor state (R) (Wulf et al., 2016) [1]. Myosin's head remains tightly bound to actin until a molecule of ATP induces the opening of the cleft and the detachment of the head from the track (Houdusse and Sweeney, 2016; Kühner and Fischer, 2011). Once the head, upon ATP binding, dissociates from F-actin, it populates a state is known as pre-recovery state or Post-rigor (PR) (Houdusse and Sweeney, 2016; Kühner and Fischer, 2011; Mesentean et al., 2007). The power-stroke ends and the recovery-stroke starts, with the myosin head in a new position along the actin filament.

1.3.1.2 *Recovery-stroke*

The recovery-stroke is the fundamental motion that re-orients the lever arm in a primed position, known as the pre-power-stroke (PPS) position when the head is detached from actin and ATP is bound in the active site (Shiroguchi et al., 2011) (see Figure 1.6). The lever arm occupies a primed position in all myosins in PPS, while the exact orientation of the converter varies among the various isoforms (Wulf et al., 2016). During the recovery stroke, structural changes in the active site promote the hydrolysis of ATP in $ADP \cdot P_i$ (Fischer et al., 2005). The hydrolysis products are trapped in the active site and the

---

[1] Small curiosity: from this state it derives the name of *rigor mortis* state, describing the stiffening in a dead body. In particular, depletion of ATP in the cells causes myosin to remain tightly bound to actin, keeping the muscles in a contracted state (Kühner and Fischer, 2011)

motor domain populates the PPS state (Houdusse and Sweeney, 2016). PPS and PR exist in equilibrium with each other and PPS is stabilized by $\cdot ADP \cdot P_i$ (Cecchini et al., 2008; Yu et al., 2007). PPS is not a force generation state of the chemomechanical cycle, since the actin-binding cleft is open and the motor unbound from actin (Houdusse, 2020; Wulf et al., 2016). In presence of actin, the stereo-specificity of the actin-binding interface changes, triggering the transition from PPS to P_iR state (Llinas et al., 2015; Wulf et al., 2016). In this stage, the cleft is still open and minor re-orientations of the converter-lever arm are observed (Llinas et al., 2015). Switch II opens the escape route for the phosphate (Llinas et al., 2015) and hydrophobic residues of the helix-loop-helix (the actin-binding region) in the L50 subdomain re-orient, driven by actin, to favor stereo-specific interactions with the filament (Llinas et al., 2015). The motor domain is in P_iR and now another power-stroke swing can begin.

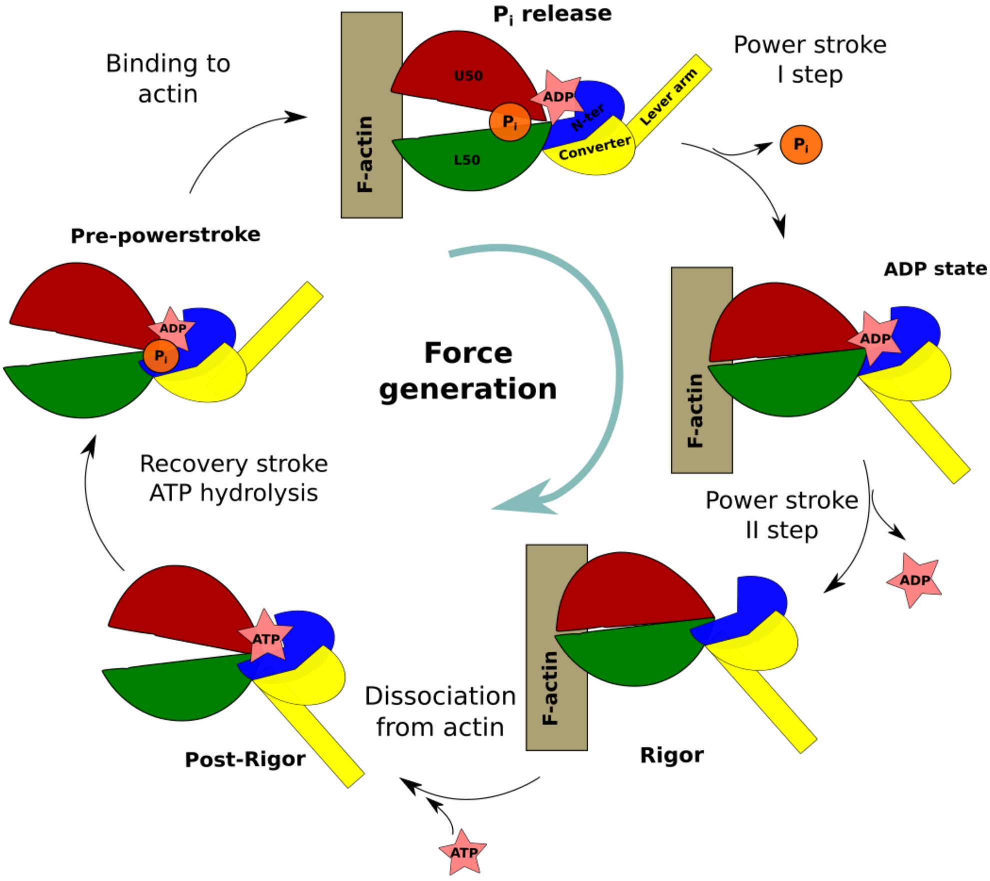


Figure 1.6: The ATP-ase cycle in myosin. The figure is inspired by the one produced by Wulf and co-workers in (Wulf et al., 2016).

2

ALLOSTERISM

What is "protein allostery"? The term "allosteric" was used for the first time in 1961 by Jacques Monod and François Jacob (MONOD and JACOB, 1961) to describe the mechanism of action of a molecule that was not "a steric analog of the substrate" (Liu and Nussinov, 2016). This assertion describes a molecule that can modulate the activity of a protein after binding to a different site of regulation, to one of the substrates. This binding leads to some consequential conformational changes in the protein, that propagate and transmit a specific signal. Even if the term "allosterism" first appeared in 1961, the phenomenon had already been addressed in the early twentieth century by Christian Bohr, when he was intent on explaining the modulation by carbon dioxide on the binding of oxygen to the hemoglobin (Bohr et al., 1904). Since then, the investigation of the phenomenon of enzyme regulation at distant sites, for the active site, as well as the description of the conformational changes in allosteric proteins, have continued unabated. Different theories have been developed over the years, notably the Monod-Wyman-Changeux (MWC) model (Monod et al., 1965) as well as the Koshland-Nemethy-Filmer (KNF) model (Koshland et al., 1966). These two models are considered the classical mechanisms to explain conformational changes in allosteric proteins (Cui and Karplus, 2008). Both have been developed by studying the quaternary protein hemoglobin (Hb) (Koshland et al., 1966; Monod et al., 1965). According to the MWC model, also known as the "concerted" model, the Hb exists in two interconvertible states: Tense (T) and Relax (R), respectively deoxy-Hb and oxy-Hb (Monod et al., 1965). The two conformational states are two fully-symmetric quaternary structures in equilibrium with each other and with different affinities for the substrate (S) (see Figure 2.1) (Monod et al., 1965). We assume that T (represented as squares in Figure 2.1) is the state with low affinity for S and R (depicted as circles in Figure 2.1) the one with high affinity for S. According to this symmetric model, there are no intermediate states between T and R, meaning that the four subunits exist only either in T or R conformation and not in a melange of the two.

The homotetrameric protein can bind from zero to 4 molecules of the substrate and, when the substrate binds, it shifts the equilibrium towards the R state, favorable to the binding and for which the ligand shows higher affinity. In the unbound state the protein exists predominantly in the T conformation, as well as when just one molecule of S is bound. Things start changing when more molecules of S are binding

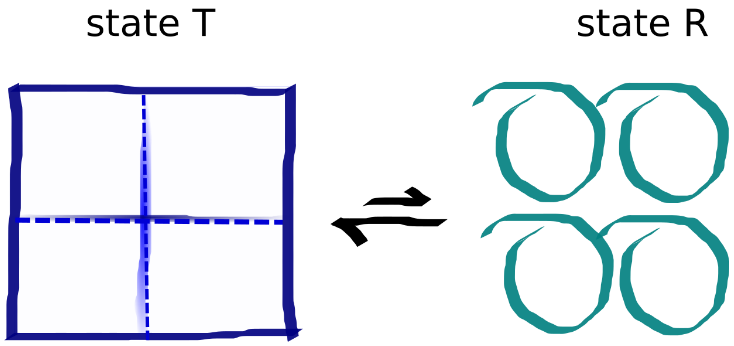


Figure 2.1: State T and R of hemoglobin, at equilibrium, when no substrate is bound

to the enzyme. When three or four molecules of S bind to the protein, the equilibrium is shifted towards the R state.

Since the system exists in these two states, T and R, we can define the dissociation constants for each of them, as K_T and K_R, respectively. In particular, we can define:

$$K_T = \frac{[T] \cdot [S]}{[TS]} \tag{2.1}$$

$$K_R = \frac{[R] \cdot [S]}{[RS]}, \tag{2.2}$$

with [T], [R], and [S] being the concentrations of the unbound species of the system and [TS] and [RS] the concentrations of the bound state both for T and R subunits.

To better understand how the cooperative model works, we can use some equations developed by Monod, Wyman, and Changeux (MONOD and JACOB, 1961; Monod et al., 1965). When the two states are at equilibrium, $T \rightleftharpoons R$ and the equilibrium constant of the system (K) is

$$K = \frac{[T]}{[R]} \tag{2.3}$$

The expression Equation 2.3 represents the fraction of the proteins in the T state over the R state, in the absence of S. The ligand starts binding to the proteins and, as the concentration of ligand molecules increases, the equilibrium constant is modified by a factor "**c**", proportional to the affinity of the substrate for the R state :

$$c = \frac{K_R}{K_T} \tag{2.4}$$

The MWC model can be represented by a sigmoidal curve (Monod et al., 1965), because changes in ligand concentration lead to a significant increase of protein in R conformation (see Figure 2.2), indicating a transition from the low-affinity state (T) to the high-affinity state (R).

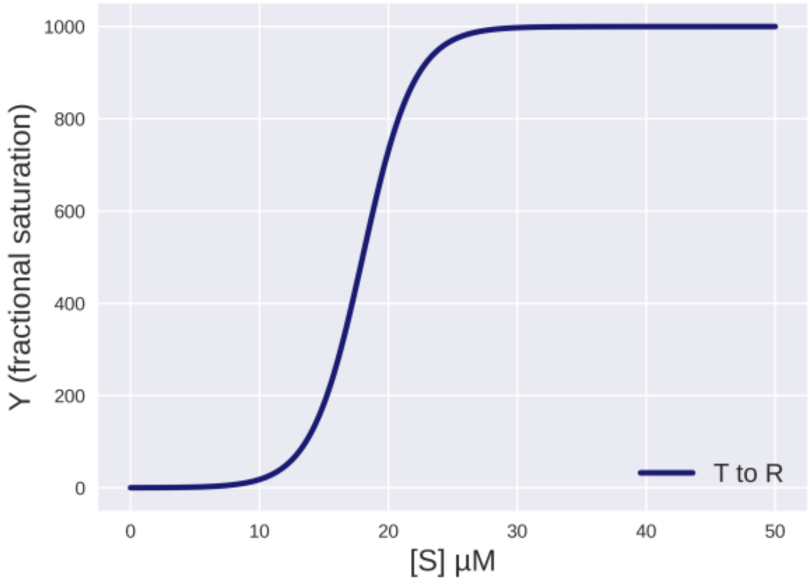


Figure 2.2: Sigmoidal curve of adaptive response in MWC model

Contrary to the "concerted" model of MWC, in which the conformational changes in the whole enzyme occur all at the same moment, in the second most widely recognized model, named KNF, they proposed a sequential mechanism (Koshland et al., 1966). According to this mechanism, the subunits of the protein depend on each other. Each binding event of the substrate to a subunit will correspond to a modification in the adjacent binding site, resulting in a change of affinity of this next subunit (Koshland et al., 1966). The conformational change in one subunit, upon ligand binding, seems to thermodynamically favor the switch to the high-affinity state, in the adjacent binding site. The main differences between the MWC and KNF models reside in the conception of the independence of the subunits of a multimeric protein, or their interdependence during their conformational change upon binding of the substrate and in the concept of the pre-existence at equilibrium, even in the absence of ligand, of both T and R form. We can summarize this last concept with the description of the KNF model as induce-fit theory, according to which the active form of the protein is reached only after binding of the ligand, and through an asymmetric process (Koshland et al., 1966). In Figure 2.3 a schematic representation of the KNF mechanism.

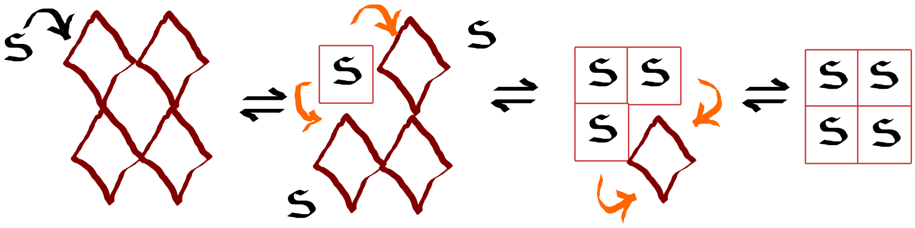


Figure 2.3: Schematic representation of the KNF model. The orange arrows indicate the conformational change induced in the adjacent binding site after binding of the substrate (S).

The MWC model is a probabilistic model, unlike the KNF which is more of a mechanistic model. Despite the mathematical accuracy of these two models, they appear to be too limiting in describing a complex mechanism such as allosteric regulation, if taken individually (Cui and Karplus, 2008). Furthermore, these two models describe the sequence of possible conformational changes in the protein following the binding of its substrate, but they do not outline how an allosteric modulator acts on the protein, by binding an allosteric site (Cui and Karplus, 2008). To cope with this, other theories have been developed in more recent years thanks to the advent of structural, thermodynamic, and kinetic studies (Cui and Karplus, 2008; Gunasekaran et al., 2004).

2.0.1 Hint to a few of novel computational methods & application in protein allostery

Nowadays, we can merge the theory behind the MWC and KNF model to fine-tune a more complete and unified perspective on allosterism, making also use of state-of-the-art techniques, such as computational and bioinformatics approaches (Cui and Karplus, 2008). Thanks to these advances, we dispose of a broader knowledge of protein sequences data and availability of computational resources, that were unimaginable at the time of MWV and KNF and that facilitated the development of new theories such as the Ensemble Allosteric Model (EAM) (Hilser et al., 2012). The EAM tries to combine the ensemble nature of the allosteric process and the structural information that can describe it, by taking into account both the thermodynamic of the protein and the propagation of energy changes in the structure itself (Hilser et al., 2012). According to the EAM, allostery is an ensemble of protein conformations with their intrinsic stability and framework of interactions among the different substructures (Hilser et al., 2012). One refers to structure-based allosterism to indicate a shift in the occupancy of conformations due to a perturbation, at any site of the protein, caused by the binding of a ligand, whether orthosteric or allosteric, as well as by a change in the protonation or interaction with another protein (Greener and Sternberg, 2018; Wodak et al., 2019). The perturbation propagates in the structure of the protein in a wave-like manner, affecting the interactions of the various domains (Greener and Sternberg, 2018; Hilser et al., 2012). The effects are evident both in changes in binding affinities and in the free energy landscape of the conformational equilibria of the different structures (Hilser et al., 2012; Wodak et al., 2019). A method that allows one to move along the energy landscape, through different trajectories, to define the shortest transition path of the system, was developed by Bolhuis and co-workers and is known as transition path algorithm (TPS) (Bolhuis et al., 2002; Dellago et al., 1998). Alongside this method, there is the cheaper, in terms of computational cost, but also more approximate one, named self-consistent path sampling (SCPS) (Orioli et al., 2017). Other sampling methods employ non-equilibrium molecular dynamics simulations to derive a time-dependent representation of the transition (Stock and Hamm, 2018). Bahar *et al* proposed elastic network models (ENMs) to sample conformational ensemble near the native minimum energy state, generated by thermal fluctuations of the system (Bahar, 2010; Wodak et al., 2019). These are only a few of the modern computational approaches used to tackle structure-based allosterism (Greener and Sternberg, 2018; Wodak et al., 2019). Another branch of computer science that proves extremely useful is bioinformatics, coupled with next-generation sequencing technique (Clarke et al., 2016; Wodak et al., 2019). Bioinformatics approaches allow us to analyze the pattern of conserved sequence, which could unveil possible mechanistic insights into allosteric processes and co-evolution, as well as allosteric hotspots (Clarke et al., 2016; Dima, 2006; Livesay et al., 2012; Wodak et al., 2019).

2.1 ALLOSTERIC DRUG DESIGN

Why is allosterism important? Allosterism is important because it is strictly connected to the concept of life. Allosteric proteins fulfill vital functions, from signal transduction to motion, to enzyme regulation, to gene expression (Tsai and Nussinov, 2014), to mention a few. The importance of these proteins goes *pari passu* with their complexity. Allosteric proteins are fundamental therapeutic targets, whose deregulation is associated with severe human disorders, such as Alzheimer's disease, diabetes, myopathies, and inflammation (Huang et al., 2010; Munro et al., 2009; Taly et al., 2009; Yang et al., 2021). It, therefore, becomes critical to design modulators for allosteric protein (Greener and Sternberg, 2018; Huang et al., 2019). As previously mentioned, allosteric modulators are small molecules able to bind to a specific site, named "allosteric site", distant from the orthosteric one (Huang et al., 2010). This event of binding leads to conformational changes through the protein, which propagate and introduce structural changes in the orthosteric site (Huang et al., 2010). The etymology of the adjectives *allosteric* and *orthosteric* also denotes the ability of the molecule to specifically bind a different or the same binding site as the endogenous or exogenous elements. From the Greek, indeed, “αλλος” (allos) means "another", and “ορθος” (orthós) stands for "correct".

Thence, the main difference between orthosteric and allosteric ligands is that the allosteric modulators are non-competitive binders, unlike the orthosteric ones, which are competitive binders since they bind the same site as the endogenous (or exogenous) molecules. This characteristic makes allosteric modulators appealing in the drug discovery field because they would cooperate (positively or negatively) with orthosteric ligands, ensuring less toxicity and high selectivity, given that the allosteric binding site is less conserved than the orthosteric one in homologous proteins, because of lower evolutionary pressure (Lu et al., 2019; Wenthur et al., 2014). Moreover, allosteric modulators would by-pass drug resistance caused by mutations in the orthosteric site; would sharpen the modulation of a specific protein conformation and bind "undruggable" targets (Chatzigoulas and Cournia, 2021; Lu et al., 2019). Nevertheless, allosteric drug discovery faces some disadvantages, indeed the mechanism of regulation and the interactions in allosteric modulation are often not completely understood (Greener and Sternberg, 2018; Huang et al., 2010). Additional challenges in allosteric drug discovery are the recognition of plausible allosteric pockets, which, as mentioned before, are not evolutionary conserved, and the identification of binders with acceptable binding affinity and activity (Chatzigoulas and Cournia, 2021; Lu et al., 2019). To cope with these problems, one can combine computational predictions and experimental validations. Computational approaches range from allosteric site prediction, with specific tools, like AlloPred (Greener and Sternberg, 2015) or AlloDriver (Song et al., 2019) and databases, such as AlloSteric Database (ASD) (Huang et al., 2011) and ASBench (Huang et al., 2015), to enhance sampling algorithms in molecular dynamic (MD) simulations to analyze hidden pockets in protein structures (Chatzigoulas and Cournia, 2021). The exploitation of computational methods is made possible due to the vast amount of available structural data, obtained by X-ray crystallography, NMR spectroscopy, and cryo-EM of protein-allosteric modulator complexes (Lu et al., 2019). If on one hand, computational approaches accelerate the drug discovery process, allowing the analysis of big data, on the other hand, experimental validation is of paramount relevance. In light of these observations, it is of great importance to pursue the development of novel techniques and to refine existing ones, so that the field of allosteric drug design can progress both efficiently and rationally.

2.2 STRUCTURE-BASED VIRTUAL SCREENING (SBVS)

Computational, or *in silico*, predictions expedite the drug discovery process, facing the dilemma of medicinal chemists in selecting the compounds to test or to synthesize, given the incredible amount of compounds that populate large chemical libraries (Gorgulla et al., 2020; Lionta et al., 2014; Patrick Walters and Wang, 2019). The process that allows computational chemists to search for drug-like molecules is known as Virtual Screening (VS) (Patrick Walters et al., 1998). VS already exists for about 20 years, but is a technique in continuous evolution (Patrick Walters and Wang, 2019). Suffice it to know that in the last years they have been able to virtually screen very large databases of compounds (millions to billions of chemicals), speaking now of virtual High Throughput Screening (vHTS) (Patrick Walters and Wang, 2019). vHTS became a decisive step in drug discovery to identify initial hits from ultra-large chemical libraries, such as ZINC or Enamine databases (Bender et al., 2021; Gentile et al., 2022). ZINC database comprises more than 1 billion molecules (Irwin et al., 2020), whereas the Enamine REAL database (`https://enamine.net/compound-collections/real-compounds/real-database`) around 4.5 billion make-on-demand compounds. Furthermore, nowadays one has at one's disposal computational resources that make vHTS processes faster: Central Processing Unit (CPU) and Graphical Processing Unit (GPU) clusters or clouds (Perez-Sanchez and Wenzel, 2010; Vamathevan et al., 2019). These computational resources have marked the breakthrough of this branch of science. Thanks to vHTS one can identify the most promising hit in a relatively short amount of time, reducing the number of compounds to synthesize or acquire for the following *in vitro* testing (Bender et al., 2021; Lionta et al., 2014; Neves et al., 2018; Patrick Walters et al., 1998). Time and money saved during a drug discovery campaign are non-negligible aspects. Hence, the ever-increasing importance and exploitation of vHTS in the early stage of drug discovery. A wide approach used in vHTS is structure-based drug discovery (SBDD) (Lionta et al., 2014). SBDD relies on the study of interactions between the biological target of interest and the putative ligands, to understand the molecular basis of the dysfunction one wants to investigate (Lionta et al., 2014). By analyzing the 3D structure of the protein of interest and the interactions involved in protein-ligand binding, it becomes possible to conduct rational SBDD (Lionta et al., 2014). But what does vHTS aim at? The main purpose of vHTS is to predict, *in silico*, the binding affinity of small molecules for a target protein and, then, rank the screened compounds according to this interaction value (Patrick Walters et al., 1998). In the end, the best ones can be selected for the *in vitro* testing. The process starts with the choice of a therapeutic target and the identification of possible binding pockets (Lionta et al., 2014). The 3D information of the proteins is often retrieved from experimental data (X-ray crystallography or NMR), otherwise, if this information is not available, from homology modeling (Patrick Walters et al., 1998) or MD simulations, by selecting a representative structure of the various configurations sampled (Lionta et al., 2014). Already at this step of SBDD, some setbacks can occur: the selection of the initial structure; it is modeling, and its druggability (Lionta et al., 2014). Moreover, if the protein of interest is an allosteric target, the identification of druggable pockets can be tricky i.e due to the presence of cryptic pockets not detectable in experimental structures (Chatzigoulas and Cournia, 2021). Taking into consideration the flexibility of the protein becomes, therefore, paramount (Chatzigoulas and Cournia, 2021). Once the structure of the macromolecule has been analyzed and selected, the second fundamental initial choice is to select the chemical library to be screened (Lionta et al., 2014). This choice is strictly related to the knowledge of the binding pocket one wants to target, as well as to the chemical space one wants to explore.

2.2.1 Common step to many vHTS pipelines

Nowadays, there are several diverse chemical libraries, classified as commercial or academic, as well as databases of known bio-active molecules or previously approved drugs (Sisquellas and Cecchini, 2021). The selected compounds should be prepared for the vHTS, by assigning proper protonation state, generating possible tautomers, as well as determining 3D coordinates (Lionta et al., 2014; Sisquellas and Cecchini, 2021). Before proceeding with the docking of the prepared molecules, one should filter out compounds according to their absorption, distribution, metabolism, elimination/excretion (ADME) properties, as well as toxicity (Norinder and Bergström, 2006). Other physico-chemical properties, that should be considered to filter molecules, are the ones that satisfied Lipinski's rule of five (Lipinski, 2000). Moreover, to reduce the rate of false positives that can be prioritized is possible to filter out Pan-assay interference compounds (PAINS) (Baell and Holloway, 2010). All these steps can be automatized by user-friendly tools, such as PrepFlow (Sisquellas and Cecchini, 2021), for the initial preparation of compounds, and Knime (Berthold et al., 2007), for the filtering. The final 3D structures are, then, ready to be used as input, along with the structure of the protein, for the *in silico* prediction of their most probable configuration in the binding site with molecular docking (described in Section 3.1).

2.2.2 SBVS: *in silico* predictions

After generating virtual complexes with molecular docking, the following step is the prediction of protein-ligand binding affinity (Lionta et al., 2014). In the last years several algorithms for predicting protein-ligand binding affinity have been developed, ranging from the more approximate empirical scoring functions of molecular docking to the most accurate rigorous methods, such as alchemical free energy perturbation (FEP) (Cournia et al., 2017; Li et al., 2021). The computational costs of a vHTS procedure increase simultaneously with the accuracy of the prediction one wants to get (Li et al., 2021). It takes seconds or minutes for the docking of a compound, rising to several hours for each FEP calculation of docked poses that have been selected (Patrick Walters and Wang, 2019). Limited computational resources and setbacks in the application of more accurate methods challenge their implementation in vHTS pipelines of ultra-large chemical databases (Li et al., 2021; Patrick Walters and Wang, 2019). On the other hand, a significant limitation of faster and more approximate methods, such as molecular docking, is the prioritization of many false positive hit compounds in the early stage of drug discovery (Fischer et al., 2021; Lyu et al., 2019). Accurate prediction of binding affinities in vHTS remains challenging (Li et al., 2021; Patrick Walters and Wang, 2019). In Figure 2.4 an outline of the SBDD process from the filtered compounds, extracted from the chemical library to screen, to the final prediction of hit(s) of interest that could become drug candidates. The application of innovative methodologies that make use of artificial intelligence (AI), in particular of its sub-fields machine learning (ML) (Walters and Barzilay, 2021) and deep learning (DL) (Mater and Coote, 2019), seems to be the most promising avenue in vHTS, to deal with big data and to improve the efficiency of the process (Gentile et al., 2022; Gromski et al., 2019; Patrick Walters and Wang, 2019; Vamathevan et al., 2019). ML builds models, from a fraction of the available data in the chemical space, to make predictions of binding affinities or physicochemical properties of interest (Gromski et al., 2019). DL is a subset of machine learning, which uses deep neural network (Gawehn et al., 2016) and has the advantage to be able to work both with fixed, expertly crafted descriptors and with task-specific representations (Yang et al., 2019a,b). Because of the efficiency of ML, DL, and other chemoinformatics tools, it is possible to explore a larger portion of the existing chemical space, than what SBDD can do. Moreover, DL gave a second youth to *de novo* drug design, which takes advantage of the nearly infinite (10^{60} (Reymond et al., 2010) to 10^{100} (Lipinski and Hopkins, 2004) possible molecules) chemical space, to design and generate novel chemical entities, that fall in the volume of interest only (Popova et al., 2018; Schneider and Fechner, 2005).

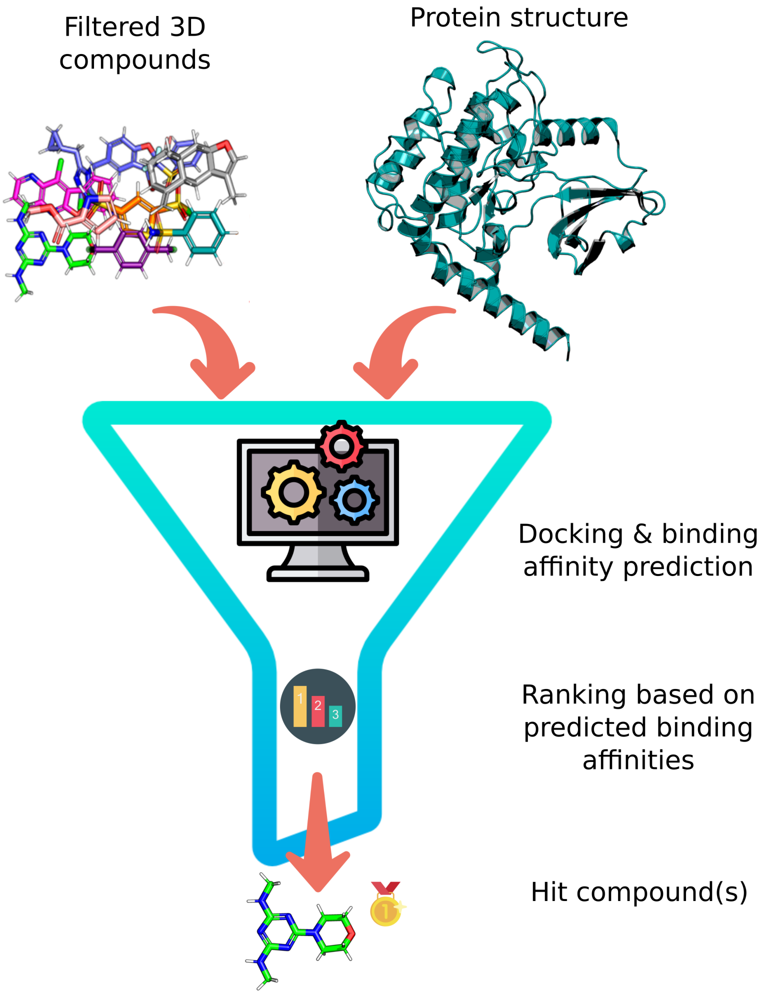


Figure 2.4: SBDD pipeline. The procedure envisages different steps, powered by computational methods. As input, there is a series of 3D chemical structures of compounds, previously prepared and filtered (see the text), and a 3D structure of the target protein. The ligands will be docked to generate protein-ligand complexes. Binding affinities are predicted for each retained binding mode. The ligands are, then, sorted according to these predictions. In the end, some molecules are selected as hit compounds for the next steps of drug discovery (*in vitro* testing and further lead optimization).

The success of SBVS is of primary importance in advancing the field of pharmacology and drug discovery, providing possible treatments for diseases hitherto considered incurable. SBDD has undoubtedly revolutionized the field of drug discovery, but it still faces some limitations, such as the lack of initial 3D structures, now mitigated by new tools, such as Alpha Fold (Jumper et al., 2021; Varadi et al., 2022); the crucial choices that must be taken during the preparation steps, as well as the impossibility of generalizing a protocol that remains strictly system-dependent (Lionta et al., 2014).

2.3 BINDING REACTION OF SMALL MOLECULES

So far we have spoken about the most recent advances in the field of drug discovery, but let's take a step back to better understand the concept behind all this. The general idea is that two molecules, a small molecule named ligand and a macromolecule named protein, can establish attractive and repulsive interactions, generating a stable system (Bissantz et al., 2010). These interactions can be irreversible or reversible and, in the case of small molecules, we often refer to reversible, or non-covalent, interactions, although covalent ligands play an important role in pharmacology, as well (Bissantz et al., 2010; Du et al., 2016; Singh et al., 2011). The binding of the two molecules is characterized by high specificity between the two species (Figure 2.5 illustrates the elements taking part in a reaction of binding). Henceforth we will focus on non-covalent protein-ligand interactions, although the same expression stands for protein-protein, protein-nucleic acids, protein-substrates, etc.

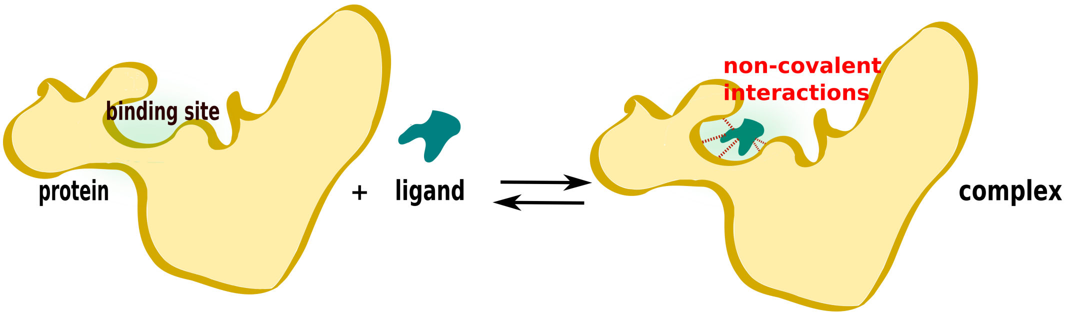


Figure 2.5: Reaction of reversible binding between a protein (in yellow) and a small molecule (in teal). The ligand binds the binding site of the protein establishing non-covalent interactions, depicted with red dashed lines.

2.3.1 Kinetic constants and Free energy of binding

In this section, we will briefly describe the process of protein-ligand association, exploring the kinetics of binding, because the drug-target lifetime is fundamental so that the drug exerts an effect.

Let's assume that we have a protein (P) and a ligand (L) in the solution. The ligand binds the protein and the reaction of binding is:

$$P + L \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} PL \tag{2.5}$$

where k_{on} and k_{off} represent, respectively, the rate constant of reversible association and dissociation of a protein-ligand complex, expressed in $M^{-1} \cdot s^{-1}$ and s^{-1} (Corzo and Santamaria, 2006; Du et al., 2016). At equilibrium, the velocity of the forward reaction is equivalent to the backward one (Du et al., 2016), and the binding and unbinding reactions are balanced according to the relation:

$$k_{on}[P][L] = k_{off}[PL] \tag{2.6}$$

The binding affinity can be expressed as K_d, equilibrium dissociation constant:

$$K_d = \frac{[P][L]}{[PL]} = \frac{k_{off}}{k_{on}} \tag{2.7}$$

In Equation 2.6 and Equation 2.7 [P], [L] and [PL] are, respectively, the concentration of the protein and ligand in the unbound state and the concentration of the protein-ligand complex. The dissociation constant indicates the ligand concentration at which half of the available protein binding sites are occupied (Pan et al., 2013; Zheng et al., 2018).

Another important measurement of binding kinetic is the residence time (or t_R), which describes the average time of binding of the ligand to its target (Lu and Tonge, 2010; Pan et al., 2013).

t_R is expressed as:

$$t_R = \frac{1}{k_{off}} \tag{2.8}$$

and, unlike the equilibrium dissociation constant, it does not depend on the concentration of the drug (Pan et al., 2013). The inverse of K_d, named constant of binding (K_b) (Du et al., 2016) is defined as follows:

$$K_b = \frac{1}{K_d} \tag{2.9}$$

This means that a ligand with a fast binding rate and low dissociation rate has a high binding affinity for the protein (Du et al., 2016). K_b (in M^{-1}) is directly correlated to the change of the standard binding free energy between the bound and unbound state of the species in the reaction, with the Gibbs relation:

$$\Delta G^o = -RTlnK_b \tag{2.10}$$

In Equation 2.10 R is the universal gas constant (1.987 $\cdot 10^{-3}$ kcal $\cdot$K^{-1} $\cdot$mol^{-1}); T is the temperature expressed in Kelvin and ΔG^o is the free energy change at standard conditions, viz. pressure of $1atm$, the temperature of $298K$ and concentrations of the reactants 1 M. During a binding reaction, that does not necessarily occur at standard state conditions, the change in Gibbs free energy (ΔG) is:

$$\Delta G = \Delta G^{o} + RTlnQ \tag{2.11}$$

where Q is the reaction quotient (Du et al., 2016). If Q = K_b, then the system is at equilibrium and ΔG is zero. If, instead, Q < K_b, then the reaction is defined as a spontaneous process (Du et al., 2016). Vice versa the process is non-spontaneous in case Q > K_b.

As for all spontaneous processes, for the protein-ligand binding reaction to occur, the Gibbs free energy must be negative (Du et al., 2016) and, since the free energy is a state function, it only depends on the initial and final state of the reaction, no matter the path connecting these two states. This means that:

$$\Delta G = G_B - G_U \tag{2.12}$$

where G_B and G_U stand, respectively, for the free energy of the Bound and Unbound state, endpoints of the reaction of binding.

Figure 2.6 is a schematic simplification of the change in the binding free energy, during a reversible binding reaction between a ligand and a protein. In the figure, the intermediate state (transition state) exists between the unbound and bound ones. The transition state represents the configuration state with the highest free energy. Consequently, it is the most unstable one. This intermediate state affects the rate at which the reaction happens (Pan et al., 2013) and it is also called the "activation state".

In a binding reaction because the reactants must be able to overcome the energetic barrier of the activation state to form the product of the reaction (in this case the protein-ligand complex). The smaller the energy of activation, the faster will be the reaction.

We can, therefore, assert that thermodynamics and kinetics of binding go hand in hand: thermodynamics describes the spontaneity of the binding process and its more general properties, while kinetics gives us an analysis of the rate at which the reaction proceeds.

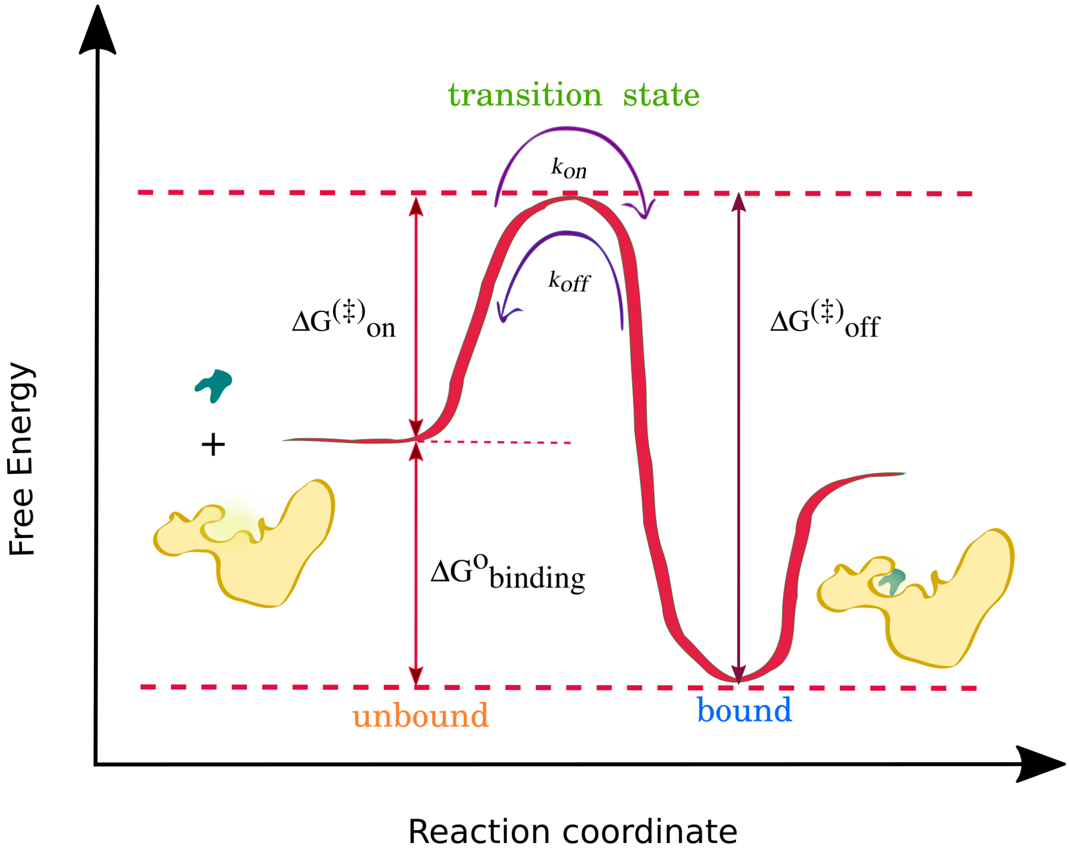


Figure 2.6: Simplified free energy profile of a ligand (in teal) binding and unbinding its protein target (in yellow). On the left is the minimum free energy of the unbound state of the drug and the protein. On the right is the minimum free energy of the bound state. The free energy of binding results from a difference between the two minima (bound and unbound) and is reported as $\Delta G^{o}_{binding}$. $\Delta G^{\ddagger}_{on}$ and $\Delta G^{\ddagger}_{off}$ are the free energy of association and dissociation and are related to the k_{on} and k_{off} rate constants.

3

METHODS

In this chapter, the reader will read about the basic theory behind the main methodologies used in the projects of this thesis.

3.1 MOLECULAR DOCKING

As mentioned in the previous chapter in Section 2.2, one can address the drug discovery problem computationally to optimize predictions of possible hits. In the early 1980s, it has been developed a tool able to predict molecular interactions between two elements, in particular between a macromolecule and a ligand (Kuntz et al., 1982), named Molecular Docking (Kitchen et al., 2004). Molecular docking oversimplifies the complex mechanism of interaction between two molecules, by reducing the number of degrees of freedom to be considered in the process and by defining them as fixed (Kuntz et al., 1982). The method requires the knowledge of the 3D structure of the elements taking part in the binding reaction and, in the last four decades, the number of available experimental high-resolution structures of macromolecules has grown in a conspicuous way (Westbrook et al., 2003). More recently, in 2021, Jumper and co-workers developed the first computational approach able to predict the accurate 3D structure of proteins, using artificial intelligence (AI) (Jumper et al., 2021). This deep learning system is named Alpha-Fold and it further predicts the number of three-dimensional structures available for the study of inter-molecular interactions (Jumper et al., 2021; Varadi et al., 2022). The ever-growing number of 3D available structures is of extreme help for molecular docking since it predicts how a small molecule (here referred to as ligand) or a protein can bind a macromolecule, such as a protein, DNA, or RNA, starting from their 3D information. Molecular docking tries to mimic *in silico* the inter-molecular interactions, returning plausible three-dimensional (3D) structures of the final complex. From now on we will talk about docking between ligand and protein, just as a matter of simplicity. The aims of molecular docking are mostly two. First, finding the best orientation of the ligand in the binding pocket of the targeted macromolecule, called binding mode, and, second, quantifying how well the ligand binds its target. The method requires a specific search algorithm able to generate multiple binding modes of the ligand and a scoring function that evaluates these binding poses and computes an interaction energy score for each of them, known as docking score (Sousa et al., 2006; Tripathi and VA, 2018). Several docking algorithms have been developed, with varying degrees of approximation, ranging from those that treat both the ligand and the protein as rigid bodies to those that, instead, consider the molecules' flexibility (Sousa et al., 2006). Nowadays, one can also choose different combinations and decide to treat only one of the two molecules as flexible. According to the level of approximation chosen, the number of degrees of freedom that the search algorithm has to consider varies, and with it also the speed of the prediction (Leach, 2001; Sousa et al., 2006; Tripathi and VA, 2018). The most approximate algorithms consider both the protein and the ligand as rigid bodies, taking into account just six rotational and translational degrees of freedom of the ligand only, without accounting for the conformational ones (Leach, 2001; Sanchez, 2013). Once the search algorithm generates the various binding poses, these must be evaluated and ranked according to their docking score (Sousa et al., 2006). The docking score is assigned by a mathematical function to which we refer as a scoring function (Bentham Science Publisher, 2006; Sousa et al., 2006).

3.1.1 Scoring functions in molecular docking

The role of the scoring function is to distinguish true active from inactive compounds (Kitchen et al., 2004), as well as the correct binding mode from all the others, despite the various approximations and simplifications from which they are affected (Kitchen et al., 2004; Sousa et al., 2006). The scoring functions can be classified as force-field-based, empirical, and knowledge-based (Bentham Science Publisher, 2006). The force-field (FF) based scoring functions estimate the interaction energy between the ligand and the receptor and the internal energy of the ligand, summing up non-bonded interactions, computed as Van der Waals and electrostatic terms (Huang et al., 2010). A few examples of FF-based docking scoring functions are AutoDock (Morris et al., 1991); DOCK (Meng et al., 1992) and GOLD (Jones et al., 1997). Empirical scoring functions divide the interaction energy into several uncorrelated terms that are weighted with specific coefficients and then summed up (Huang et al., 2010). Among empirical scoring functions are FlexX (Rarey et al., 1996); MedusaScore (Yin et al., 2008) and Glide (Friesner et al., 2004). The knowledge-based scoring function tries to reproduce crystal structures by analyzing interatomic contact frequencies between the ligand and the protein, to, then, assign a greater frequency to the more favorable interactions (Huang et al., 2010; Kitchen et al., 2004; Sanchez, 2013). Some knowledge-based scoring functions are ITScore (Huang and Zou, 2006); DrugScore (Gohlke et al., 2000) and BLEEP (Mitchell et al., 1999).

3.1.1.1 *AutoDock Vina*

One of the most widely used docking programs is AutoDock Vina (VINA)(Alhossary et al., 2015; Eberhardt et al., 2021), which employs an empirical scoring function (Trott and Olson, 2009). The scoring function of VINA takes inspiration from the X-core function (Quiroga and Villarreal, 2016; Trott and Olson, 2009; Wang et al., 2002). To predict the binding energy (termed c in Equation 3.1.1.1), it sums the interaction function for each pair of atoms able to move with respect to each other. In other words:

$$c = \sum_{i<j} f_{t_i t_j}(d_{ij}) \tag{3.1}$$

with i and j being the atoms in the pair; t their type and $f_{t_i t_j}$ the interaction functions at the interatomic distance d_{ij}. In Equation 3.1.1.1 c is, therefore, the sum on inter- and intra-molecular contributions (Trott and Olson, 2009):

$$c = c_{inter} + c_{intra} \tag{3.2}$$

Being an empirical scoring function, VINA assigns weights to each energetic term by linear regression (Trott and Olson, 2009). In particular, steric interactions are taken into account using three different terms and relative weights ($w1$ to $w3$ in Equation 3.3) for attractive and repulsive functions; hydrophobic interactions are considered if both the atoms are hydrophobic (fourth term in Equation 3.3) and if the pair of atoms involves a hydrogen bond acceptor and donator, then the last term is considered with relative weight ($w5$ in Equation 3.3) (Quiroga and Villarreal, 2016; Trott and Olson, 2009).

$$c = \sum_{i<j} f_{t_i t_j}(d_{ij}) = \begin{cases} w1 * Gauss_1(d_{ij}) \\ w2 * Gauss_2(d_{ij}) \\ w3 * Repulsion(d_{ij}) \\ w4 * Hydrophobic(d_{ij}) \\ w5 * HBonds(d_{ij}) \end{cases} \tag{3.3}$$

In Equation 3.3 the steric terms $Gauss_1$ and $Gauss_2$ are defined as follows:

$$\begin{aligned} Gauss_1(d_{ij}) &= e^{-(d/0.5Å)^2} \\ \mathrm{Gauss}_2(d_{ij}) &= e^{-((d-3Å)/2Å)^2} \end{aligned} \tag{3.4}$$

The repulsion term is:

$$Repulsion(d_{ij}) = \begin{cases} d^2, & \text{for} \quad d < 0 \\ 0, & \text{for} \quad d > 0 \end{cases} \tag{3.5}$$

The hydrophobic term corresponds to :

$$Hydrophobic(d_{ij}) = \begin{cases} 1 \quad \text{for} \quad d \leq p_1 \\ p_2 - d \quad \text{for} \quad p_1 > d < p_2 \\ 0 \quad \text{for} \quad d \geq p_2 \end{cases} \tag{3.6}$$

Finally, the HBonds term refers to :

$$HBonds(d_{ij}) = \begin{cases} 1 \quad \text{for} \quad d \leq h_1 \\ \frac{d}{-h_1} \quad \text{for} \quad h_1 < d < 0 \\ 0 \quad \text{for} \quad d \geq 0 \end{cases} \tag{3.7}$$

The search algorithm of VINA includes different iteration levels. At first an iterated local Monte Carlo (MC) search method generates a random conformation of the ligand (Tang et al., 2022). This conformation is then perturbed in its position, orientation, and torsion, and the binding affinity is evaluated (Tang et al., 2022). Next, the initial conformation is optimized with a Broyden–Fletcher–Goldfarb–Shanno (BFGS) method (Fletcher, 2000), a quasi-Newton method, that takes into account the gradient of the scoring function (derivatives of the scoring function with respect to the position, orientation and torsion of the ligand). The purpose of this extra optimization is to predict a better binding mode with a lower score. In the end, the difference between the docking score of the new proposed conformation and the initial one is evaluated to conclude whether the optimized pose can be accepted or not (Tang et al., 2022; Trott and Olson, 2009).

3.1.1.2 *PLANTS (Protein-Ligand ANT System)*

Another docking program is PLANTS (Protein-Ligand ANT System) (Korb et al., 2007, 2009). PLANTS has been developed with different empirical scoring functions to the one of AutoDock Vina: CHEMPLP and PLP (Korb et al., 2007; Verkhivker et al., 2000). PLP stands for *Piecewise linear potential* and it is used to model the steric complementarity in protein-ligand interactions (Korb et al., 2009; Verkhivker et al., 2000). Furthermore, CHEMPLP includes terms of the GOLD Chemscore implementation (Verdonk et al., 2003) to account for hydrogen bonding and metal binding (Korb et al., 2009). The scoring functions are expressed as follows:

$$f_{PLANTS_{PLP}} = f_{plp} + f_{clash} + f_{tors} + c_{site} \tag{3.8}$$

and

$$f_{PLANTS_{CHEMPLP}} = f_{plp} + f_{hb} + f_{hb-ch} + f_{hb-CHO} + +f_{met} + f_{met-coord} + \\ + f_{met-ch} + f_{met-coord-ch} + +f_{clash} + f_{tors} + csite \quad (3.9)$$

Equation 3.8 and Equation 3.9 were taken from the paper by Korb *et al* (Korb et al., 2009). In the equations Equation 3.8 and Equation 3.9 the meaning of the common terms are as follows: f_{plp} stands for piecewise linear potential function for repulsive and attractive interactions (Korb et al., 2009); f_{clash} represents an empirical heavy-atom potential calculated for each pose to avoid clashes (Korb et al., 2009); f_{tors} is the torsional potential of all rotatable bonds in the ligand, excluding terminal hydrogen-bond groups (Korb et al., 2009); c_{site} is a quadratic potential, calculated as the difference between the initial coordinate of the ligand, if outside of the sphere defined by the search algorithm, and the binding site radius (Korb et al., 2009). In Equation 3.9 additional terms appear. $f_{hb}, f_{hb-ch} + f_{hb-CHO}$ are terms used to consider three different types of hydrogen bonds (H-bonds), respectively: all types of H-bonds between atoms that are both uncharged or of which only one is charged; H-bonds between pair of atoms which are both charged or H-bonds between an oxygen-acceptor group and CH-donor one (Korb et al., 2009). The terms $f_{met}, f_{met-coord}, +f_{met-ch} and f_{met-coord-ch}$ are picewise linear functions that takes into account metal interactions (Korb et al., 2009). Unlike VINA, PLANTS uses a so-called *ants colony optimization* (ACO) algorithm to search for the global minimum of a docking run. The ACO algorithm is inspired by the natural behavior of ants, which can find the shortest path between their nest and the source of food. While foraging, the ants release pheromones along the path. Since on the shortest path the amount of hormones that has been released is the highest, the insects can choose this most convenient track with higher probability (Korb et al., 2007). In PLANTS the pheromones are numbers associated with each possible random docking solution. Here we have mentioned just two, free for academics, and docking programs, which we will also encounter again in further sections of the manuscript. Nowadays, there are countless docking programs and a challenge for the user is also knowing which of them is the most appropriate for his or her system. The selection is dictated by several factors, such as whether or not one can afford to purchase the possible required license; the level of accuracy and speed one wants to achieve; the target of the experiment itself (to decide what kind of scoring function will work best), etc. A reasonable choice requires running a benchmark on the system of interest, as well as testing several docking programs in a consensus manner, to pick the most suitable one for the system under investigation. Molecular docking remains one of the cheapest methods in structure-based drug design, with performances tending to improve with the usage of single- and multi-GPU acceleration for large-scale virtual screening (Fan et al., 2021; Tang et al., 2022).

3.2 MOLECULAR DYNAMICS (MD) SIMULATIONS

How can we emulate the real motion of biological molecules at the atomic level? It is possible to reproduce *in silico* the actual behavior of biomolecules, whether they are proteins or small molecules or nucleic acids, thanks to Molecular Dynamics (MD) simulations (Karplus and Petsko, 1990). MD is an approach able to predict the state of the system, meaning its atomic position, based on the current state, by integrating Newton's laws of motion (Karplus and Petsko, 1990). This integration generates an ensemble of successive configurations of the system, named trajectory (Leach, 2001). A trajectory defines both the position and the velocities of the particle in the system, by solving Newton's second law of motion (Leach, 2001):

$$F_i = m_i \cdot a_i \tag{3.10}$$

where the force applied on the particle (F_i) is directly proportional to the mass (m_i) and the acceleration (a_i) of the particle itself.

From Equation 3.2, since a_i can be expressed as derivative of the velocity (v_i) with respect to the time (t):

$$a_i = \frac{dv_i}{dt} \tag{3.11}$$

and, in turn, velocity can be expressed as the derivative of the atomic position (s_i) with respect to time

$$v_i = \frac{ds_i}{dt} \tag{3.12}$$

it becomes:

$$F_i = m_i \cdot a_i = m_i \cdot \frac{d(\frac{ds_i}{dt})}{dt} = \frac{d^2 s_i(t)}{dt^2} \tag{3.13}$$

Equation 3.13 describes the motion of a particle i with mass m_i along the dimension $s_i(t) = (x_i(t), y_i(t), z_i(t))$ under the force F_i.

The first MD simulation of a biomolecule was performed around 25 years ago (McCammon, 1998). McCammon and co-workers simulated the Bovine Pancreatic Trypsin Inhibitor (BPTI) in a vacuum, for only 9.2 ps (McCammon, 1998). Nowadays, we can simulate systems for a time longer than a millisecond (Lindorff-Larsen et al., 2011) thanks to High-Performance Computing (HPC) resources. More recently, the group of Amaro simulated the entire SARS-CoV-2 viral envelope (Casalino et al., 2021) and Thornburg and co-workers simulated a 3D-whole cell system, using a minimal bacterial cell (cell with a genome that does not include genes which are non-essential in laboratory) (Thornburg et al., 2022).

3.2.1 Molecular Dynamics and Monte Carlo methods

In MD successive configurations of the system are connected in time, whereas in Monte Carlo simulations the current configuration depends from the one immediately preceding (Leach, 2001). MD has a kinetic contribution to the total energy, while in Monte Carlo approach the total energy is directly derived from the potential energy function (Leach, 2001). The two approaches can sample different ensembles: microcanonical; canonical, isothermal-isobaric and grand canonical. These ensemble describe the conditions under which the simulation is performed. In canonical, microcanonical (or constant) and isothermal-isobaric ensembles the number of particles of the system (N) is constant, while in grand canonical ensemble the number of particles can vary (Leach, 2001). MD is usually performed under microcanonical ensemble, with constant N, volume (V) and energy (E) of the system. Monte Carlo simulations, instead, sample in the canonical ensemble, which means at constant N, V and Temperature (T) (Leach, 2001).

3.2.2 Potential Energy in MD

Since MD deals with configurations of a system, it depends on the potential energy of the system itself (U) (Karplus and Petsko, 1990). To simulate in detail the behavior of the system, one should use quantum mechanics, but it becomes unfeasible in terms of computational time to compute the electronic motions of all the atoms (Leach, 2001). In quantum-MD is, therefore, introduced an approximation named Born-Oppenheimer, which separates the motions of electrons from the nuclear ones, assuming that the electrons can adapt to the change in the nuclear position, being the nuclei much heavier than electrons (Leach, 2001). In MD atoms are, instead, defined as spheres. It follows that the potential energy of a system in MD is calculated as a function of the nuclear position only, by using functional forms named *Force fields*.

3.2.2.1 *Force-fields in molecular mechanics*

A force field (ff) contains four main terms to describe intra- and inter-molecular forces acting in the system. The first three usually describe the energy of bonds, angles, and torsion, denoted as bonded interactions, and the last one refers to the non-bonded or long-range interactions. The total potential energy of the system becomes:

$$U = U_{bonded} + U_{non-bonded} \tag{3.14}$$

If we decompose the terms into their elements, then Equation 3.14 becomes:

$$U = U_{bonds} + U_{angles} + U_{torsions} + U_{electrostatics} + U_{vanderWaals} \tag{3.15}$$

There are several force fields described in the literature, whose complexity may change according to the nature of the systems to treat, but, since only two force fields were explored in this thesis project, we will describe the terms of only these two. The first one is Assisted Model Building with Energy Refinement (AMBER) ff, whose functional form is (Case et al., 2018):

$$\begin{gathered} U(r^N) = \sum_{bonds} k_b(l - l_0)^2 + \\ + \sum_{angles} k_\theta(\theta - \theta_0)^2 + \\ + \sum_{dihedrals} V_n[1 + cos(n\Phi - \gamma)] + \\ + \sum_{i=1}^{N-1} \sum_{i=i+1}^{N} f_{ij} \left\{ \epsilon_{ij} \left[\left(\frac{rij^0}{r_{ij}} \right)^{12} - 2 \left(\frac{rij^0}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right\} \end{gathered} \tag{3.16}$$

b is the bond length between two atoms that are covalently bonded; l_0, θ_0 are equilibrium reference values for the length and the angle, respectively; k_b, k_θ and V_n are force constant specific for each atom type; n is the multiplicity and γ is the phase angle for the torsional angle parameter.

In the non-bonded term, it figures the non-covalent interactions between inter- or intra-molecular atom pairs, decomposed in van der Waals and electrostatic contributions. The van der Waals term (the first one of the fourth term in Equation 3.16) accounts for the *Lennard-Jones 12-6 function*, characterized by an attractive part that varies of r_{ij}^{-6} and a repulsive one that varies of r_{ij}^{-12}.

The final contribution in Equation 3.16 is the electrostatic potential between two atoms, defined as the sum of interactions between pairs of point charges, as determined in Coulomb's law. The second force field explored in this thesis is Chemistry at Harvard Molecular Mechanics

(CHARMM) (Best et al., 2012; Vanommeslaeghe et al., 2012). The formalism of CHARMM is:

$$\begin{aligned} U = & \sum_{bonds} k_b (l - l_0)^2 + \\ & + \sum_{angles} k_\theta (\theta - \theta_0)^2 + \\ & + \sum_{dihedrals} k_\phi [1 + cos(n\Phi - \gamma)] + \\ & + \sum_{impropers} k_\omega (\omega - \omega_0)^2 + \\ & + \sum_{Urey-Bradley} k_u (u - u_0)^2 + \\ & + \sum_{non-bonded} \epsilon \left[\left(\frac{R_{min_{ij}}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{min_{ij}}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{\epsilon r_{ij}} \end{aligned} \tag{3.17}$$

We can notice that there are some similarities between the formalism of AMBER and CHARMM. Nevertheless, in CHARMM is made explicit the improper dihedral angle for non-consecutive atoms (ω) and the Urey-Bradley term assess the distance u between two atoms separated by a third one. The k_b , k_θ, k_ω and k_u are force constants specific for each atom type.

3.2.3 Numerical integrator algorithms

In classical MD simulations, Newton's law of motion is numerically integrated over a finite time-step (Δt) to follow the time evolution of the system. This is made possible by numerical integrator algorithms (Leach, 2001). The purpose of these algorithms is to determine the position $s_i(t + \Delta t)$ at time $t + \Delta t$, knowing position s, velocity v, and acceleration a at time t. Among them, the Verlet algorithm is one of the most stable and widely implemented in MD codes (Verlet, 1968). It is similar to the *leap-frog* integration, with the exception that velocity and position are calculated at the same time value. The Verlet algorithm uses the position and acceleration at time t and the position of the previous step (t- Δt) to compute the new position at time t+ Δt. The leap-frog method is based on the Taylor series expansion and it is accurate until the fourth power of Δt (Verlet, 1968). If we consider the expansion to the second order of Δt, for each particle i we have forward in time:

$$s_i(t + \Delta t) = s_i(t) + \Delta v_i(t) + \frac{1}{2} \Delta t^2 a_i(t) \tag{3.18}$$

and backward in time:

$$s_i(t - \Delta t) = s_i(t) - \Delta v_i(t) + \frac{1}{2} \Delta t^2 a_i(t) \tag{3.19}$$

Adding together Equation 3.18 and Equation 3.19 it becomes:

$$s_i(t + \Delta t) = 2s_i(t) - s_i(t - \Delta t) + \Delta t^2 a_i(t) \tag{3.20}$$

This algorithm allows us to calculate the new (forward) position $s_i(t + \Delta t)$, by knowing the actual ($s_i(t)$) and the previous ones ($s_i(t - \Delta t)$).

3.2.3.1 *Periodic Boundary Conditions (PBC)*

Important elements in MD simulations are the Periodic Boundary Conditions (PBC). These are a collection of boundaries chosen to approximate an infinite system and the forces it could experience in a real bulk fluid (Leach, 2001). PBC allows to simulate of a fairly small number of particles to calculate the thermodynamic properties of a system, within unit cells (Leach, 2001). Each cell, referred to as a periodic box, is surrounded by others in all directions. This ensures that if a particle leaves the box during the simulation, another particle would replace it from the other side (Leach, 2001). In this way, the number of particles of the system remains constant (Leach, 2001).

3.3 MMPB(GB)SA FREE ENERGY RE-SCORING

Starting from classic MD simulations one can compute several properties of the system, such as free energy of binding or kinetic quantities, which can be employed in pipelines of Computer-aided drug design (CADD). Let's focus for the moment on the free energy of binding. As mentioned in Section 2.3, Equation 2.12, the free energy of binding of a drug to its target can be estimated as the free energy difference between the end point states of the reaction of binding (the protein-ligand complex and the unbound states):

$$\Delta G_b^o = G_{complex} - G_{protein} - G_{ligand} \tag{3.21}$$

This is the theory behind two of the most popular end-point methods used in drug discovery campaigns to calculate the relative free energies of binding: molecular mechanics Poisson Boltzmann surface area (MM/PBSA) and molecular mechanics generalized Born surface area (MM/GBSA) (Genheden and Ryde, 2015; Montalvo-Acosta and Cecchini, 2016). These two end-point methods were first developed by Kollman et., al (Srinivasan et al., 1998) and represent a compromise between efficiency and accuracy (Montalvo-Acosta and Cecchini, 2016). MM/PB(GB)SA method can be used to compute the average energy of binding, over a trajectory extracted from MD simulation. Therefore, Equation 3.21 becomes:

$$\Delta G_b^o =< G_{complex} > - < G_{protein} > - < G_{ligand} > \tag{3.22}$$

Since this approach computes the solvation free energy (Rastelli et al., 2010) on the sampling of the final states of the system, it achieves higher accuracy with respect to the molecular docking (Montalvo-Acosta and Cecchini, 2016), but at the same time, it does not become as computationally expensive as the rigorous methods, such as absolute binding free energy (ABFE) (Gapsys et al., 2021), for which the simulation of the entire reaction path is needed (Montalvo-Acosta and Cecchini, 2016). The free energy averages of Equation 3.22 are computed starting by three different MD simulations, respectively for the complex, the protein, and the ligand in solution, and this approach is known as three average MM/PB(GB)SA (3A-MM/PB(GB)SA) (Wang et al., 2019). A way to reduce the computational cost of the calculation is to run one single trajectory of the complex, from which to extract the contributions of the complex, the free protein, and the ligand unbound (Wang et al., 2019). This second approach is named one-average MM/PB(GB)SA (1A-MM/PB(GB)SA) and is the most widely used protocol (Wang et al., 2019). In the latter approach, the major approximation comes from the assumption that both protein and ligand do not undergo conformational changes upon binding.
According to the second law of thermodynamics:

$$\Delta G_b^o = \Delta H - T\Delta S \tag{3.23}$$

with H being the enthalpy of the system and S its entropy. T is the temperature expressed in Kelvin.

The enthalpic contribution can be rewritten as follows:

$$\Delta H = \Delta E_{MM} + \Delta G_{solv} \tag{3.24}$$

being E_{MM} the molecular mechanical energy and ΔG_{solv} the solvation free energy. These terms can be further decomposed in:

$$\Delta E_{MM} = \Delta E_{internal} + \Delta E_{vdW} + \Delta E_{elec} \tag{3.25}$$

and

$$\Delta G_{solv} = \Delta G_{pol} + \Delta G_{non-pol} \tag{3.26}$$

In Equation 3.25, $\Delta E_{internal}$ represents the intra-molecular energy, in other words, the energy associated with bonds, angles, and torsions in the system (Hou et al., 2011). ΔE_{vdW} and ΔE_{elec}, instead, stand for van der Waals and electrostatic interactions, respectively (Hou et al., 2011). The solvation free energy in Equation 3.26 is decomposed in polar and non-polar contributions (Hou et al., 2011). The polar term is obtained either by solving the Poisson–Boltzmann equation (in MM/PBSA) or by considering the Generalized Born (MM/GBSA) model (Hou et al., 2011; Onufriev et al., 2002; Rastelli et al., 2010). The non-polar term is predicted from a linear relation to the Solvent Accessible Surface Area (SASA) (Gilson et al., 1997; Wang et al., 2006). Essentially, the polar solvation term takes into account the electrostatic interactions between the solute and the solvent, whereas the non-polar free energy is associated with the attractive and repulsive solute-solvent interactions, as well as with the cost of creating the cavity within the solvent for the solute (Genheden and Ryde, 2015; Izairi and Kamberaj, 2017). The main difference between MM/PBSA and MM/GBSA is that MM/PBSA solves the differential Poisson-Boltzmann (PB) equation to compute the electrostatic contribution to the solvation free energy, whereas MM/GBSA applies the Generalized Born (GB) approximation to the exact PB equation (Onufriev et al., 2002; Wojciechowski and Lesyng, 2004). Since the differential equation of PB is rather computationally expensive to solve, the GB approximation is preferred, which, while maintaining a certain level of accuracy, reduces the computational time required (Wang et al., 2019). Despite this difference, both methods consider the solvent as a continuum (Montalvo-Acosta and Cecchini, 2016), meaning that these methods are implicit solvation approaches. In continuum solvation models the solvent is treated as a homogeneous mean, with high dielectric constant (ε), with respect to the solute that is embedded in a continuum with lower ε (Limongelli, 2020). By treating the solvent as a continuum and considering just the end states of the binding reaction, the calculation speeds up.

Since end-point approaches (such as MM/PB(GB)SA) are both more accurate than molecular docking and less computationally challenging than rigorous methods, they are gradually playing a larger role in drug discovery protocols to predict the binding free energies of small molecules.

3.4 ENTROPY CORRECTION (QHMB)

As shown in Equation 3.23, the ΔG_b^o is decomposed in a term that takes into account the conformational entropy change upon binding ($T\Delta S$). This term is often calculated by Normal Mode Analysis (NMA) (Case, 1994) or Quasi Harmonic Analysis (QHA) (Karplus and Kushick, 1981; Polyansky et al., 2012), that, respectively, under- or over-estimate the entropy upon binding. Moreover, the introduction of the entropic correction often worsens the results (Pereira and Cecchini, 2021), therefore it is often neglected in the calculation of the free energy of binding. Nevertheless, recently, a novel approach to tackle this problem has been developed (Pereira and Cecchini, 2021). The new method is named Quasi-Harmonic Multi-Basin (QHMB) and it seems to improve the accuracy of the calculations, overcoming the limitations of both NMA and QHA approaches (Pereira and Cecchini, 2021). The approach consists of clustering, from a converged MD simulation, the possible conformers of the ligand to then weight them with a certain equilibrium probability (Pereira and Cecchini, 2021). Each cluster center is, then, used to evaluate the local fluctuations (per-basin entropy) with standard QHA approach (Pereira and Cecchini, 2021). The ensemble of these various per-basin entropies constitutes the multi-basin scheme of the method (Pereira and Cecchini, 2021). Finally, each quasi-harmonic entropy and equilibrium probability are combined to determine the intramolecular entropy of a molecule with multiple conformers (Pereira and Cecchini, 2021). QHMB is the method that we have chosen to predict entropy correction in one of the projects that will be discussed later in this manuscript.

4

AUTOMATION OF VHTS

We took a little dive into the field of computer-assisted drug discovery (CADD) with some of the theory behind this virtual reality, but now the time has come to look for a tool that practically carries on the virtual screening campaign. Nowadays is paramount to dispose of tools that fully automatize the vHTS campaign, especially since the trend is to screen billions of chemically diverse compounds (Gorgulla et al., 2020; Lyu et al., 2019). The success of the initial steps of SBDD, by prioritizing highly specific and affine hit molecules for the target of interest, would save time and cost for the lead optimization process in early drug discovery stages (Gorgulla et al., 2022). In 2020 Gorgulla *et al* developed *VirtualFlow* (VF), an open-source structure-based platform to prepare and screen *in silico* ultra-large diverse libraries of compounds (Gorgulla et al., 2020). VF was designed to be user-friendly and to optimally exploit any type of computational resources, from local computer clusters to cloud computing platforms (Gorgulla et al., 2020). The authors used VF to screen a virtual library of more than one billion commercially available compounds, targeting the Kelch-like ECH-associated protein 1 (KEAP1)(Gorgulla et al., 2020). In particular, they targeted the interface of the protein-protein interaction nuclear factor erythroid-derived 2-related factor 2 (NRF2) and (KEAP1) (Gorgulla et al., 2020). Their strategy consisted of a multi-stage docking, performing different runs of docking with various implemented programs (Gorgulla et al., 2020). Finally, after ranking the ligands according to their docking score and considering drug-likeness and availability factors, 590 hits compounds were selected for experimental validation (Gorgulla et al., 2020). The results show that VF prioritized potential binders of KEAP1, at the interaction interface with NRF2 (Gorgulla et al., 2020). Another tool that can be exploited for vHTS is *ChemFlow* (*Gomes et al., ChemFlow – From 2D chemical libraries to protein-ligand binding free energies*, accepted). This is one of our in-house programs, which allows us to automatically run a vHTS introducing MD simulations and free energy rescoring with end point methods. It will be more thoroughly described in Chapter 5.

4.1 RANDOM ACCELERATION MOLECULAR DYNAMICS SIMULATION (τRAMD)

Resuming the discussion of the use of MD simulations in CADD, this time we focus on the prediction of the kinetic properties of the system. Why kinetic properties should be used in drug discovery? The kinetics of association and dissociation of a drug may reveal crucial information about the mechanism of action of the compound and its *in vivo* efficacy (Cecchini and Changeux, 2022; Lu and Tonge, 2010; Schuetz et al., 2017), This is particularly important since many approved drugs show non-equilibrium binding characteristics (Schuetz et al., 2017). This evidence emphasizes the potential role of computing drug-target residence time for predicting *in vivo* efficacy (Schuetz et al., 2017), since a drug results active when bound to the target (Lu and Tonge, 2010). The concept of a drug-target lifetime as a crucial parameter for pharmacological properties of drug candidates was introduced in 2006 by Copeland and colleagues (Copeland et al., 2006) and since then has been embraced by the scientific community. The correlation of drug residence time with *in vivo* efficacy has been documented for many targets, ranging from various G-protein-coupled receptors (GPCRs) (Guo et al., 2012; Tee et al., 2007; Watson et al., 2005) to protein kinase (Puttini et al., 2008; Schuetz et al., 2017). Nevertheless, the role of the kinetic of binding should be evaluated case by case, because of the toxicity that long binding time can cause for several targets (Schuetz et al., 2017). Unquestionably, the thermodynamic stability of the drug-target complex is the driving force of the reaction of binding, but the kinetic selectivity of the compounds gives information about the rate of association (k_{on}) and dissociation (k_{off}) of the complex. This last concept could be imported into drug-discovery protocols, to prioritize molecules with a slow off-rate. This is because:

$$K_d = \frac{k_{off}}{k_{on}} \tag{4.1}$$

where K_d is the equilibrium dissociation constant and k_{on} and k_{off}, respectively, are the association and dissociation rate of the complex. Molecules with higher affinity for a specific target should have higher k_{off} for it and, consequently, higher residence time. The residence time will henceforth be indicated by the letter τ and it is defined as:

$$\tau = \frac{1}{k_{off}} \tag{4.2}$$

τ is a better indicator of *in vivo* efficacy of a drug (Copeland, 2016; Copeland et al., 2006; Schuetz et al., 2017), meaning the duration of action of a drug, rather than its binding affinity (Copeland et al., 2006). τ depends only on the average time spent by the drug bound to its target and not on its concentration (Copeland et al., 2006; Pan et al., 2013). This is also evident from its metric unit, which is seconds.

A question now arises: how can we compute the residence time *in silico*? And here we go again to the usefulness of all-atom MD simulations, which are exploited in various computational approaches to explore unbinding pathways of small molecules and estimate relative dissociation rates. These methods range from metadynamics-based approach (Capelli et al., 2020; Tiwary et al., 2015), to weighted ensemble simulations (Dixon et al., 2018; Lotz and Dickson, 2018), to random expulsion MD (Lüdemann et al., 2000). The latter was then renamed as Random Acceleration Molecular Dynamics (RAMD) to generalize its applicability extent (Kokh et al., 2020) and was implemented in various MD packages, such as AMBER8; GROMACS, CHARMM and NAMD (Kokh et al., 2020). Nevertheless, the need to improve the performance of the approach led the group of Wade, in 2018, to make a fundamental modification of the method (Kokh et al., 2018, 2020). They decided to impose a random Force as the input parameter, rather than an acceleration (Kokh et al., 2018). This novel approach was named τRAMD, standing for τ Random Acceleration Molecular Dynamics approach to compute the relative residence time of drug-like molecules (Kokh et al., 2018). In 2020, Kokh co-workers developed a workflow able to compute this relative residence time, starting from classic MD simulations (Kokh et al., 2020). The code has been implemented in Gromacs2020 (`https://github.com/HITS-MCM/gromacs-ramd`), assuring high performances (Kokh et al., 2020). The goal of the calculation is to record many unbinding events and relative paths from various initial configurations, in order to compute average relative τ (Kokh et al., 2020). The procedure (see the scheme in Figure 4.1) consists in running a minimization and long equilibrations to then perform 4 equilibrations, starting with random velocities. For each equilibration, 15 different productions will then be performed (Kokh et al., 2020). These productions will enhance the sampling of ligand unbinding. During the productions, a random Force is applied to the center of mass (COM) of the ligand (Kokh et al., 2020). Based on a certain distance cut-off, the simulations will automatically stop as soon as the ligand reaches this distance and it is, thus, considered unbound from its target pocket (Kokh et al., 2020).

This innovative and easy-to-use tool, although very computationally expensive for large sets of compounds, opens the door to new implementations in drug discovery campaigns. An example of the scientific community's commitment in this direction has been demonstrated by the creation in 2012 of a consortium, called Kinetics for Drug Discovery (K4DD) (`https://cordis.europa.eu/project/id/115366`). This was the first of possible future projects going towards these implementations in the drug-discovery field.

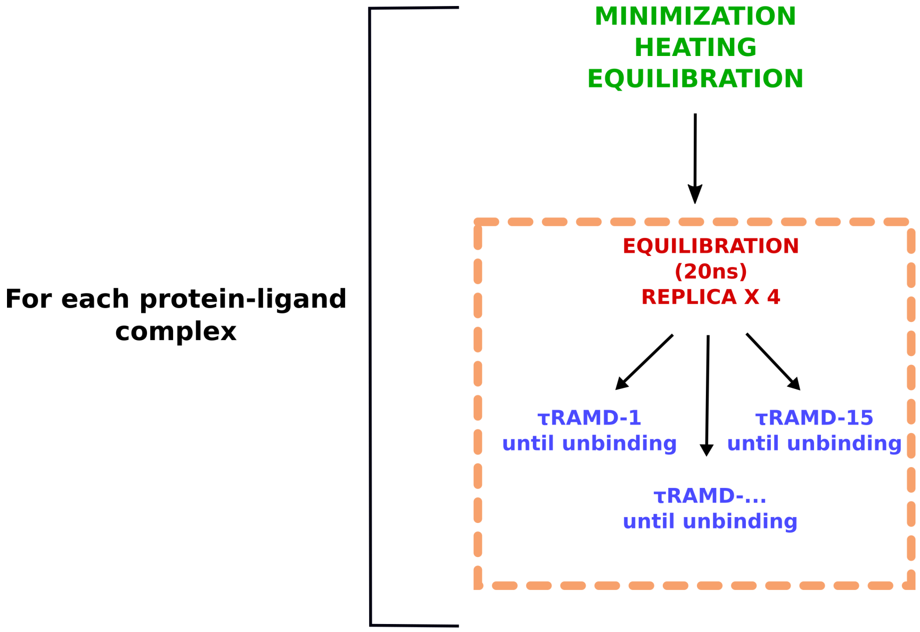


Figure 4.1: Scheme of the τRAMD protocol. For each protein-ligand complex, one has to run initial minimization, heating, and equilibration with conventional MD simulations. The equilibrated system will then be the input of at least four new longer equilibrations (i.e 20 ns). Then, for each equilibration, at least 15 different productions will be performed. In these productions, the PULL code implemented in Gromacs2020 is used and the MD simulations continue until the unbinding of the ligand from its target pocket. The unbinding is driven by a random Force applied on the center of mass of the small molecule. The MD automatically stops once the ligand unbinds and the average residence time is computed over the 15 replicas, thanks to an analysis package named *tauRAMD-v2.py*, provided by the developers of the τRAMD code. In the dashed box the part of the procedure that is the most expensive from a computational point of view.

5

CHEMFLOW

The need for an automated tool that allows us to perform vHTS arose in our laboratory. In particular, we wanted to be able to predict the relative binding free energy (ΔG_b) of compounds in a more accurate way than just docking. We, therefore, wrote a program named Chem*Flow*, designed to predict ΔG_b for a set of ligands in a fully automated way, still balancing accuracy and efficiency. Chem*Flow* is a user-friendly interface between the various programs that are implemented. The pipeline envisages an initial step of docking, that can be performed with different docking programs and respective scoring function(s), followed by partial charges assignment for the docked ligands and the last step of free energy rescoring, performed with end-point methods approach (MM/PB(GB)SA), over a trajectory produced by MD simulation (see Figure 5.1).

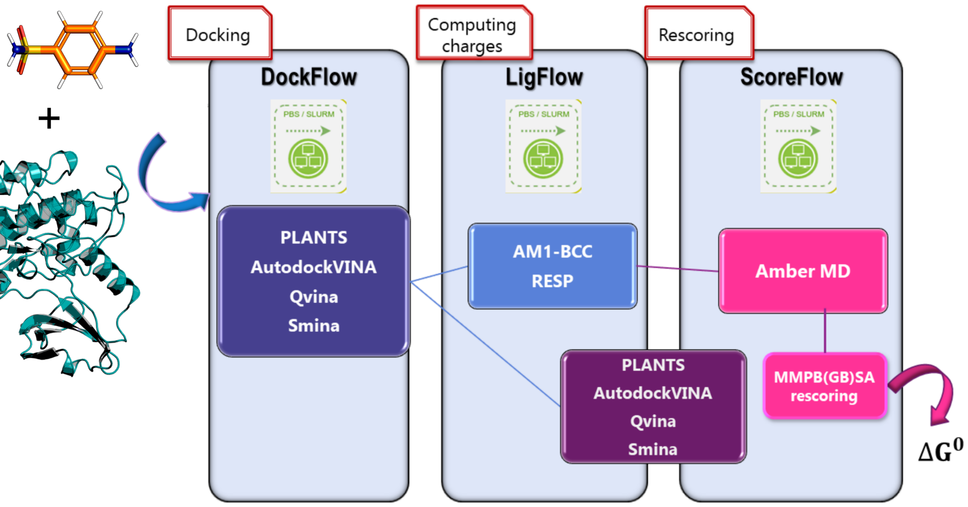


Figure 5.1: Scheme of Chem*Flow* protocol. The user needs to provide as input the 3D structure of the targeted protein and the library of compounds (enclosed in a representative single structure in the figure) to be screened. In the boxes the three modules of *ChemFlow*

The software is organized in three different workflows named respectively Dock*Flow*, Lig*Flow* and Score*Flow*. Dock*Flow* performs molecular docking from initial 3D structures of the target protein of interest and of the chemical library to screen. Docking can be performed running one of the following programs: PLANTS (Korb et al., 2007), AutoDock VINA (Trott and Olson, 2009), QVINA (Alhossary et al., 2015) or SMINA (Ryan Koes et al., 2013). After the docking, the user can select the number of poses to keep for each ligand for the next step of charge attribution, carried out by Lig*Flow*, using AmberTools (Case et al., 2021). The partial charges computed by default are the semi-empirical with bond charge correction (AM1-BCC) (Jakalian et al., 2002), but the user can also choose the restrained electrostatic potential (RESP) (Bayly et al., 1993) ones. The performance of Lig*Flow* depends on the size of the molecules and the type of charges chosen by the user since the quantum calculation of RESP charges is more computationally costly than the AM1-BCC ones. Once this step is completed, the user can continue with the last stage and run Score*Flow*. This last module allows running MD simulations in implicit or explicit solvent, with some default parameters that the user can easily modify according to his/her needs. MD simulation is followed by free energy rescoring with MM/PB(GB)SA approaches. The user can again choose which approach to use in a very smooth way. Also, in this case, the analysis is carried out with AmberTools (Case et al., 2021). Furthermore, Chem*Flow* gives the possibility to run a simple rescoring of the docked poses using a different docking scoring function (consensus docking (Houston and Walkinshaw, 2013; Ren et al., 2018)). Either way, it is possible to analyze the ranking of the compounds according to the classification reported in the file provided at the end of the post-processing of the results at each step. An additional feature, which makes Chem*Flow* a very useful tool for vHTS campaigns, is the ability to launch all the calculations with High-performance computing (HPC), taking advantage of the improved performance of both CPU and GPU clusters. Further developments include the implementation of other docking programs as well as of the Gromacs package for MD simulations (Abraham et al., 2015; Lindahl et al., 2001; Markidis and Laure, 2015). If the reader is moved by curiosity, the software and the relative documentation are available on GitHub at the following link: `https://github.com/IFMlab/ChemFlow`. I personally contributed to the implementation of Qvina and Smina in the module of Dock*Flow*, as well as to the coding of Score*Flow* and Lig*Flow*. Moreover, I wrote both the documentation and the tutorial on the GitHub page.

6

β CARDIAC MYOSIN: STRUCTURAL ANALYSES & DYNAMICS

6.1 CARDIAC MYOSIN'S ISOFORMS & HEART FAILURE

Among all conventional myosins II, there is cardiac myosin, responsible for heart contraction (Barrick and Greenberg, 2021). Cardiac myosin was first identified in the 1980s by Jeacocke and England during phosphorylation experiments on rats' hearts (Jeacocke and England, 1980), but the study of the excitation and contraction mechanism in the heart began almost a century earlier in 1883 with the discovery made by Ringer about the implication of Ca^{2+} in heart contraction (Ringer, 1883). Several studies have followed thereafter, culminating in the explanation of how calcium release is finely regulated in the sarcomere and coupled to muscle contraction (Ebashi, 1963; WEBER and WINICUR, 1961), through a mechanism of filament sliding (HUXLEY, 1961, 1953). Two distinct sarcomeric genes (MyHC-α, MyHC-β) encode for Myosin Heavy Chain (MHC) of cardiac myosin in the mammalian heart (Mahdavi et al., 1984; Weiss et al., 1999). The MHC isoforms are respectively named α and β (Kooij et al., 2014) and they show high sequence identity (93.2%) (Weiss et al., 1999). These isoforms are differentially expressed in the cardiac muscle (Komuro et al., 1987; Kooij et al., 2014) during the development (Lyons et al., 1990; Mahdavi et al., 1984) and exhibit different speeds of contraction (Sartore et al., 1978). The faster α isoform (Miyata et al., 2000), mostly expressed in the atrium (Walklate et al., 2021b), shows up to three times higher ATPase activity (Herron and McDonald, 2002), with respect to the slower β one (Miyata et al., 2000), expressed in the ventriculum (Walklate et al., 2021b). We turned our attention to β cardiac myosin (β-CM), since we were interested in studying the modulation of the protein in a pathological condition, known as heart failure (HF) (Harris, 1983), during which the β isoform is over-expressed with respect to the α one (Miyata et al., 2000). HF is generally defined as the incapability of the heart to pump enough blood to the tissues (Harris, 1983; Mudd and Kass, 2008) and this dysfunction are related to a shift of expression between isoforms, from the faster α to the slower β (Hamdani et al., 2008). As already mentioned, the β isoform is mostly expressed in the ventriculum (Komuro et al., 1987; Kooij et al., 2014), responsible for systolic contraction of the heart muscle (Walklate et al., 2021b). Malfunction of the left ventricle (LVD) is associated with an increased risk of systolic heart failure (Cleland et al., 2005; McMurray et al., 2013), a leading cause of death in today's society, especially in the most developed or in developing countries (Afzal, 2021; Gaziano et al., 2010). Systolic HF is a pathology characterized by a broad clinical picture associated with high morbidity and mortality for both women and men (Bloom et al., 2017; Heidenreich et al., 2022; Murphy et al., 2020). HF can be classified into three different categories according to the ejection fraction (EF): HS with preserved, borderline, or reduced EF (Kapoor et al., 2016; Mentzer and Hsich, 2019). Even if HF with reduced ejection fraction (HF-rEF) is the most common case (in more than 50 % of the population in the USA (Bloom et al., 2017)), primarily in men (Colvin et al., 2015), until today there are no medicament able to directly improve cardiac contractility without side effects caused by the activation of compensatory mechanisms (Bloom et al., 2017). Patients with systolic HF-rEF manifest a decrease in cardiac contractility (Adams et al., 2005), hypoxia or hypercapnia (CO_2 retention in the blood) (Giannoni et al., 2009), EF equal to or less than 40% (Mentzer and Hsich, 2019), as well as dyspnea, fatigue and congestion (Murphy et al., 2020), just to mention a few symptoms. As a compensatory response, the body increases the sympathetic tone of the muscle (Malik et al., 2011) and activates neurohormonal pathways (Giannoni et al., 2009; Malik et al., 2011), which results in a higher concentration of plasma norepinephrine and brain natriuretic peptide (BNP) (Giannoni et al., 2009). These reactions worsen cardiac dysfunction over time (Malik et al., 2011) that exacerbates urgent hospitalization and premature death (McMurray, 2010). Current pharmacologic therapies consist of the administration of diuretics to cope with dyspnea and fluid retention (McMurray, 2010), as well as Angiotensin-Converting–Enzyme (ACE) inhibitors and *beta*-adrenergic blockers to reduce the ventricular size and increase EF, even if moderately (Dickstein et al., 2008; Hunt et al., 2009; McMurray, 2010). These therapies aim at impeding neurohormonal activation, but an alternative strategy relies on the administration of positive inotropic drugs, such as β-agonists or phosphodiesterase inhibitors to increase cardiac contractility (Malik et al., 2011). All these drugs act by impelling second-messenger signaling pathways with an effect on myocardial consumption of oxygen and on myocytes' calcium concentration level to enhance cardiac contractility (Kass and Solaro, 2006). The enhanced cardiac contractility results in prolonged systolic ejection that can cause severe arrhythmias and hypotension, resulting in increased HF disease-related mortality rates (Packer, 1993). Given the increasing impact of cardiovascular diseases (CVD), which accounted for 32% of global deaths in 2019 (Benjamin et al., 2019) with a trend that does not appear to be improving, it is essential to study innovative and safe treatments. Novel therapeutic avenues could involve modulators that directly target cardiac myosin to up or down-regulate its activity (Hwang and Sykes, 2015), without triggering second-messenger signaling mechanisms that worsen patients' conditions. This consideration strengthened our decision to study, both from a structural and pharmacological perspective, the β isoform of cardiac myosin, whose contraction results impaired in HF.

6.2 OMECAMTIV MECARBIL (OM)

6.2.1 OM: mode of action

Omecamtiv mecarbil (OM) is a small molecule designed in 2010 by Morgan and co-workers with HTS of a small synthetic library (Morgan et al., 2010). OM acts as allosteric activator of β-CM, increasing the fractional shortening (FS) by 123 % at 2 μM (Morgan et al., 2010). The efficacy of OM is expressed in terms of AC_{40}, corresponding to "compound concentration resulting in a 40% increase in the cardiac sarcomere ATPase activity at the calcium concentrations that produced 25-50% of maxi- mum calcium-dependent activation." (Morgan et al., 2010) Following kinetic studies, conducted both by Malik *et al* and Liu *et al*, demonstrated that OM modulates the kinetics of β-CM by accelerating the transition of the head from weakly to strongly bound to actin, without modifying the ADP release rate (Liu et al., 2015; Malik et al., 2011). By contrast, OM accelerates the P_i-release in presence of actin by 4-fold (Liu et al., 2015), reducing the energy barrier for the transition from PPS to phosphate release (P_iR) state (Malik et al., 2011). Isothermal titration calorimetry (ITC) measurements show that OM binds β-CM in a nucleotide-dependent manner and the highest affinity ($K_d = 0.29\mu$M) is recorded when there is $MG^{2+} - ADP - VO_4^{3-}$ in the active site (Planelles-Herrero et al., 2017), where VO_4^{3-} stands for Vanadate, the analog of inorganic phosphate (Gresser and Tracey, 1990). A lower affinity ($K_d = 1.8\mu$M) is measured when $ADP \cdot BeF_x$, the analog of ATP, is bound in the active site, corresponding to the PR state whereas no binding is sensed in absence of nucleotide (R or PR states) (Planelles-Herrero et al., 2017). These data suggest that OM binds preferentially the PPS state (Planelles-Herrero et al., 2017). Small-angle X-ray scattering (SAXS) studies also corroborated the previous results, showing that the state populated in presence of OM is, indeed, PPS (Planelles-Herrero et al., 2017).

6.2.2 OM: binding in PPS β-CM

From a structural point of view, OM bridges N-terminal, converter, and RH, as demonstrated by the X-ray structure of bovine β-CM in PPS, solved in presence of OM, by Planelles-Herrero and co-workers in 2017 (Planelles-Herrero et al., 2017) (see Figure 6.1).

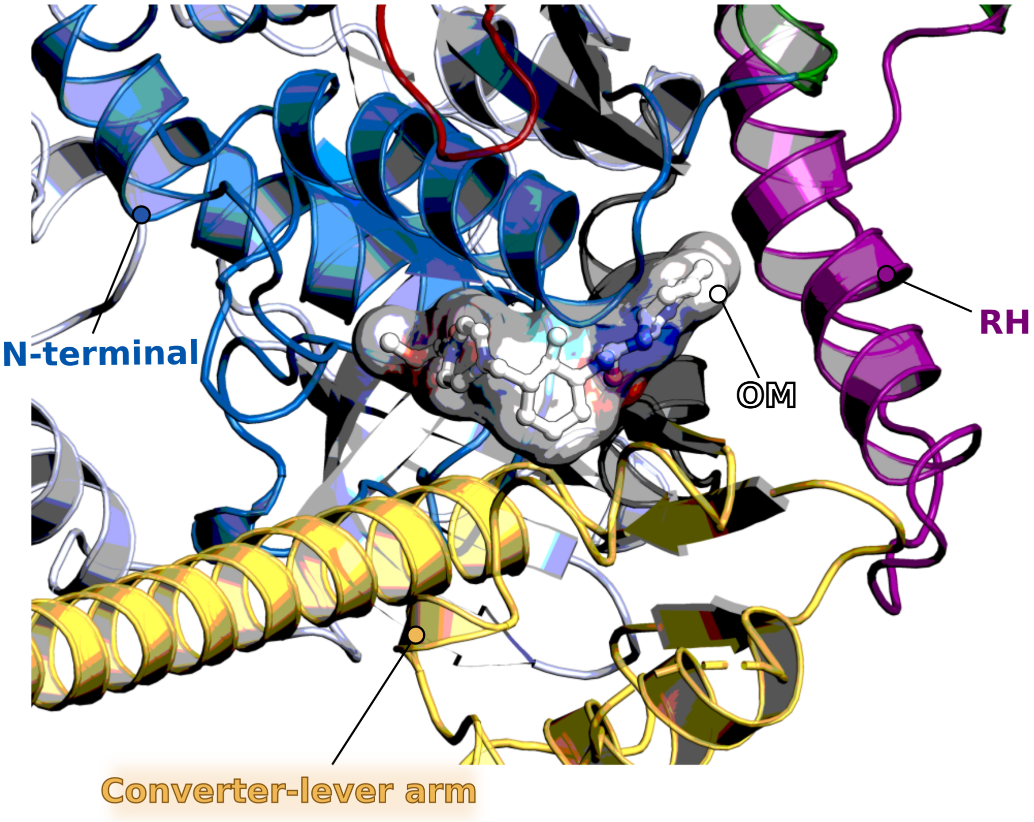


Figure 6.1: Zoom on the space occupied by OM in bovine β-CM (PDB code: 5N69.pdb). OM is represented in sticks and surrounded by its surface, colored according to the polarity of the atoms. In cartoon the subdomains around the modulator: are RH in purple; N-terminal in blue and the converter in yellow.

The interactions (Figure 6.2) that OM establishes in its buried pocket are mostly hydrophobic (Planelles-Herrero et al., 2017). Nevertheless, the carbamoyl-amino linker of OM forms polar contacts both with the converter and the N-terminal subdomains (Planelles-Herrero et al., 2017). Polar contacts are also found between the converter and the methyl-pyridinyl ring of OM (Planelles-Herrero et al., 2017).

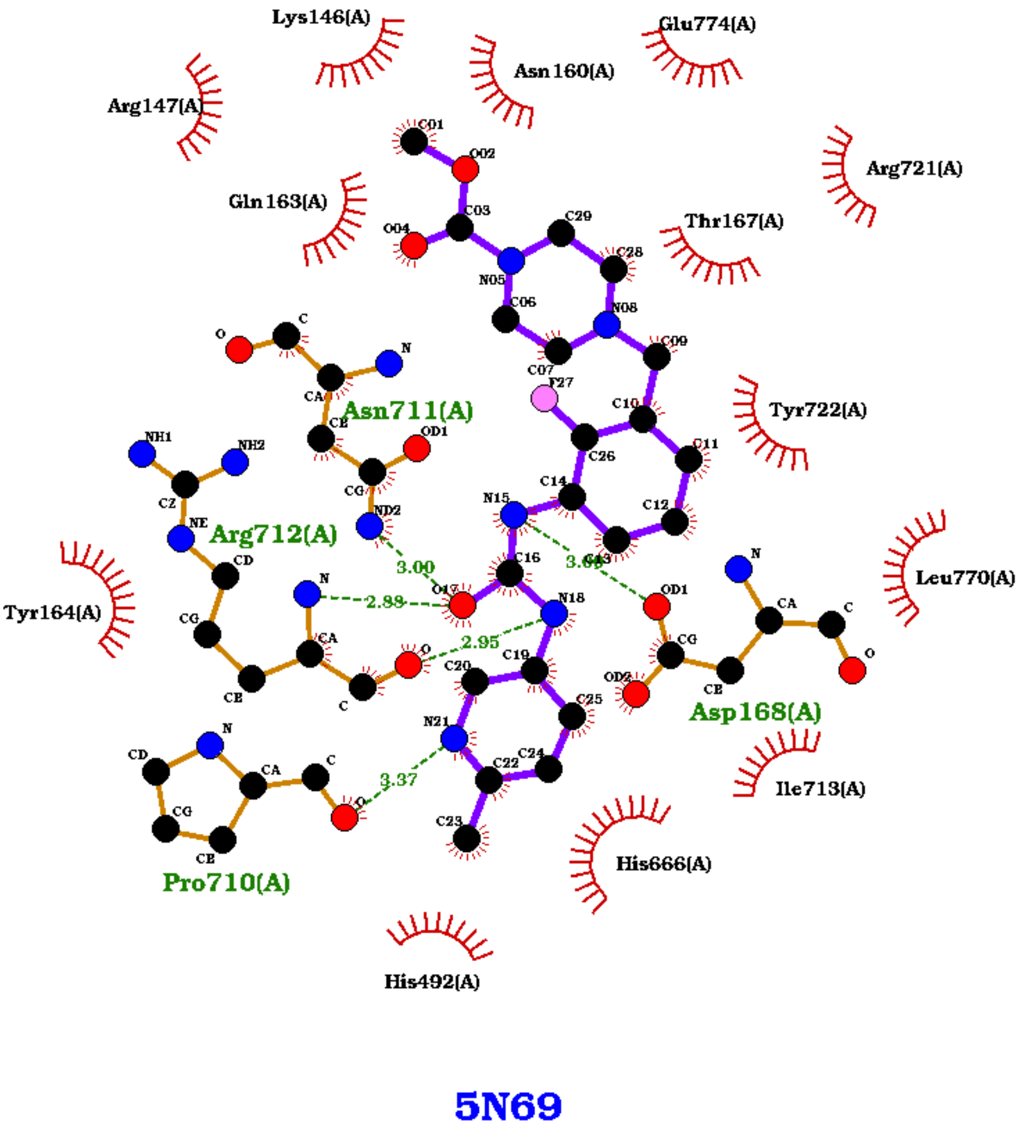


Figure 6.2: 2D representation of the interactions that OM establishes in its pocket in bovine β-CM (5N69.pdb). The analysis has been run with LigPlus (Laskowski and Swindells, 2011) using the chain A of the X-ray structure.

In the next sections, we will address the first part of our research, aimed at creating a stable *in silico* model of β-CM, both in the presence and absence of OM, from which to build all further analyses.

6.3 STRUCTURE OF β-CARDIAC MYOSIN IN PPS

Our study commences with selecting the structure of bovine β-CM in the pre-powerstroke (PPS) state, starting from the coordinates of 5N69.pdb (Planelles-Herrero et al., 2017). One might wonder why we chose to start from this specific state of the force cycle. The reason lies behind the purpose of studying the dynamics of the PPS state to design cardiac allosteric modulators, particularly non-competitive activators, able to directly target the motor, without altering intracellular calcium homeostasis (Liu et al., 2015; Morgan et al., 2010). This will avoid adverse effects caused by activation of second-messengers signaling cascades (Malik et al., 2011). Our starting point is the hypothesis that cardiac myosin co-exists in different states, in equilibrium with each other, and an allosteric activator may accelerate the transition toward the PPS state (Liu et al., 2015; Winkelmann et al., 2015), notoriously the powerstroke onset state (Houdusse and Sweeney, 2016; Sun et al., 2022; Wulf et al., 2016). By shifting the equilibrium of the hydrolysis step towards PPS, there will be more myosin heads ready to bind actin, start the powerstroke and, consequently, increase cardiac contractility (Liu et al., 2015; Planelles-Herrero et al., 2017).

In this situation the hydrolysis equilibrium will be:

$$\mathrm{M} + \mathrm{ATP} \rightharpoonup \mathrm{M} + \mathrm{ADP} + \mathrm{P_i} \tag{6.1}$$

where M stands for Myosin and P_i for inorganic phosphate, generated upon ATP hydrolysis.

Further clarification to be made concerns the choice of the X-ray structure from which we started: 5N69.pdb (Planelles-Herrero et al., 2017). This is the crystallographic structure with the highest resolution (resolution of 2.45 Å) solved so far for the S1 fragment of cardiac myosin in PPS (Planelles-Herrero et al., 2017). The motor domain is solved in complex with ADP-VO_4-MG^{2+} in the active site and the allosteric activator Omecamtiv Mecarbil (OM) (Morgan et al., 2010) in its pocket, at the interface between N-terminal, transducer, Relay helix (RH) and converter subdomain (Planelles-Herrero et al., 2017). ADP-VO_4 is the analog of ADP-P_i (inorganic phosphate) (Gresser and Tracey, 1990) in terms of distances between the beta and the gamma phosphate of ATP, after ATP-hydrolysis. The group of Houdusse has also crystallized β-CM in PPS in absence of an allosteric modulator (PDB code: 5N6A) (Planelles-Herrero et al., 2017) and, after comparison of the structures with and without OM, it appears that the drug stabilizes a more primed position of the converter-lever arm domain, by establishing specific interactions with the elements responsible of the swinging of the lever arm, in its buried hydrophobic pocket, formed only in this state (Planelles-Herrero et al., 2017). Most likely this allosteric pocket is formed by induced fit (Planelles-Herrero et al., 2017). The specificity of OM for β-CM is ensured by a network of interactions with some residues highly conserved in both human and bovine cardiac myosin and not in other myosins II isoforms (Planelles-Herrero et al., 2017). This makes OM allosteric binding site and allosteric pockets in general, very interesting sites to be exploited in drug discovery.

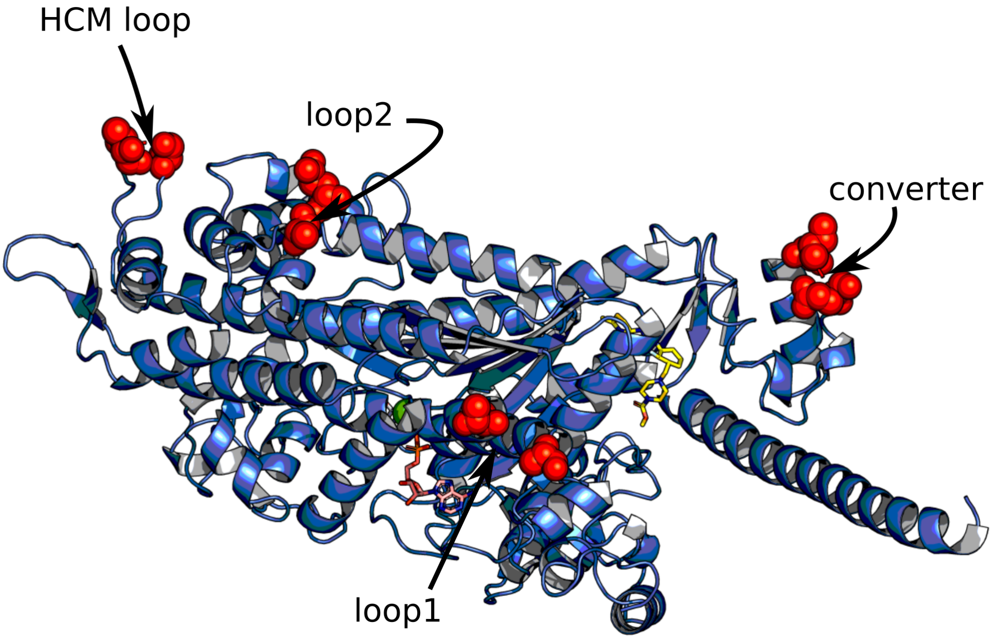
6.3.1 Molecular modelling and Dynamics of bovine β-CM in PPS

6.3.1.1 *Modelling of β-CM & System preparation*

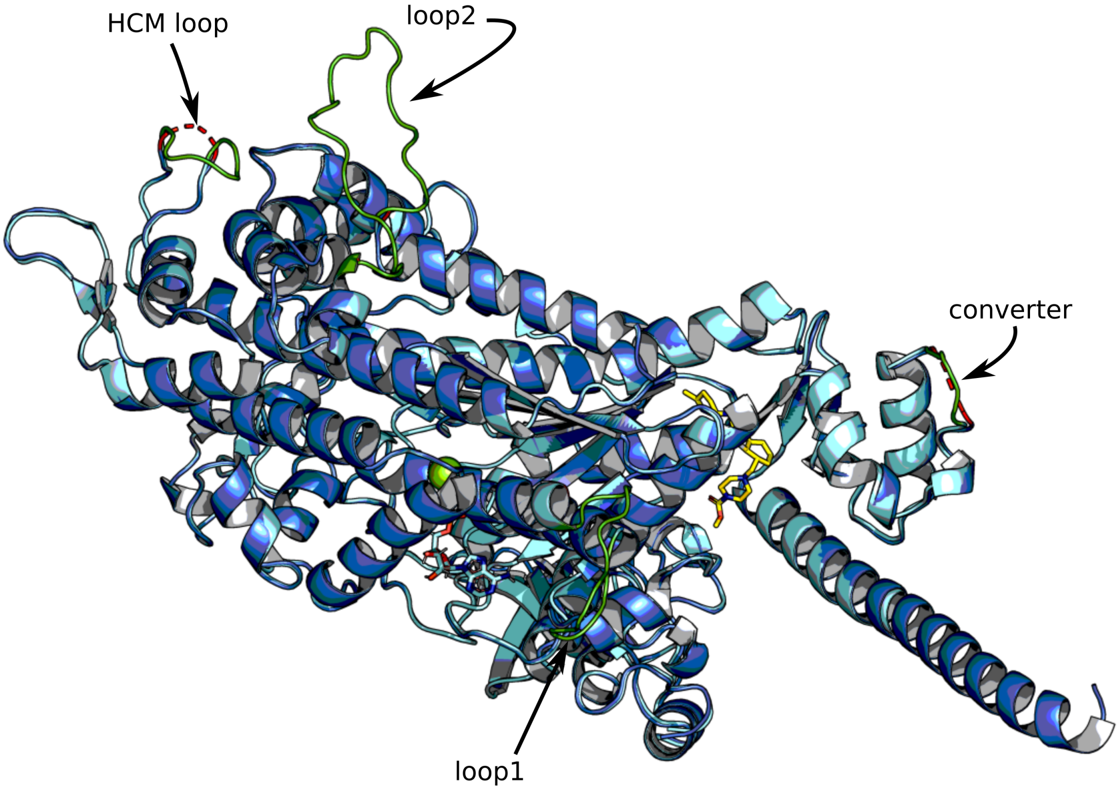
The missing loops in bovine β-CM's structure (5N69.pdb) (residues: 201 to 214; 405 to 410; 623 to 644) were modeled using MODELLER (Webb and Sali, 2016) and the missing residues at the level of the converter subdomain (residues 733 to 735) were modeled starting from the template 4PA0.pdb (Winkelmann et al., 2015), PR state of human β-CM motor domain, always with MODELLER. Bovine and human β-CM share 96% of sequence identity (data obtained with HHPRED (Söding et al., 2005)). The best model was selected according to the lowest Discrete Optimized Protein Energy (DOPE) score. The structure was then uploaded on the MolProbity web server (`http://molprobity.manchester.ac.uk/`) to add missing hydrogens, optimize H-bond networks and detect the more reasonable rotameric states of the side chains of Asn, Gln and His. The standard protonation state was assigned to all the titratable residues and for the histidine, the most probable protonation state at neutral pH was predicted using Karlsberg (Rabenstein, 2000). The missing portions of the X-ray structure are depicted in Figure 6.3 sub-figure (a) and the final model is in sub-figure (b). The inorganic phosphate was modeled in its form $H_2PO_4^-$ (Cecchini et al., 2010) and its topology and parameter files generated from AmberTools (Case et al., 2018) using General AMBER Force-Field2 (GAFF2) (Case et al., 2018; Wang et al., 2004), as well as topology and parameter files for ADP, MG^{2+} and OM. The system was ionized and neutralized with Sodium and Chloride using Tleap (Case et al., 2018). The final size of the system was 193362 atoms, including OM and the nucleotide ($ADP - P_i - MG^{2+}$) in the active site. The same initial structure was used to prepare a system without OM, following the same procedure aforementioned.

6.3.1.2 *Molecular Dynamics (MD) simulations and Visualization tool*

Amber18 engine (Case et al., 2018) with Amberff19SB for the protein (Ponder and Case, 2003; Tian et al., 2020) and GAFF2 (Case et al., 2018; Wang et al., 2004) for the ligands were used to perform MD simulation of the system both with and without OM bound. Four successive minimizations, of 5000 steps each, were run to progressively remove restraints on the system. Then, 2ns of gradual heating in NVT was performed to reach a final temperature of 300^oC, followed by 2.5 ns equilibration in NPT with Berendsen thermostat and barostat (Berendsen et al., 1984). Finally, a production of $375ns$ with Langevin thermostat (Case et al., 2018) and Monte Carlo barostat (Åqvist et al., 2004) was performed for the system with OM. The same procedure was followed to run $330ns$ MD of the system without OM. To visualize



(a) Cartoon representation of 5N69.pdb. In red spheres the position of the missing portion in the X-ray structure. In yellow sticks OM and pink $ADP - VO_4^{3-}$. The green sphere represents $MG^{2}+$



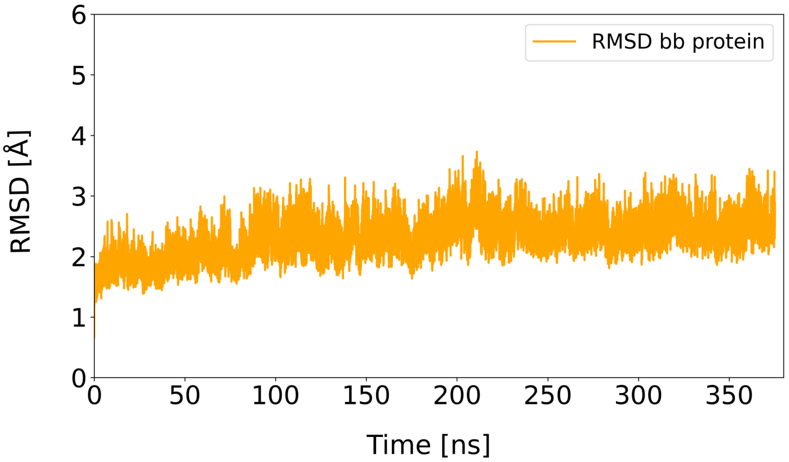
(b) Cartoon representation of the model of 5N69.pdb. In green the missing portion was added to the model.

Figure 6.3: In sub-figure (a) the X-ray structure of β-CM in PPS (5N69.pdb). Loop1 and loop2 are hypervariable loops that cannot be solved by X-ray crystallography due to their disorder (Krenz et al., 2003). Loop2 connects U50 and L50 subdomains and controls the actin-activated ATPase activity, with electrostatic interactions with actin's binding site (Murphy and Spudich, 2000). Loop1 is at the level of the nucleotide binding site, connecting the 50 kDa subdomains to N-terminal (Krenz et al., 2003; Murphy and Spudich, 2000). It influences motility and ADP release rate (Murphy and Spudich, 2000). Hypertrophic cardiomyopathy (HCM) loop is on the surface of U50 and contributes to the strong interactions with actin both in presence of ADP or in the nucleotide-free state (Houdusse and Sweeney, 2001). Sub-figure (b) is a cartoon representation of the model obtained starting from 5N69.pdb.

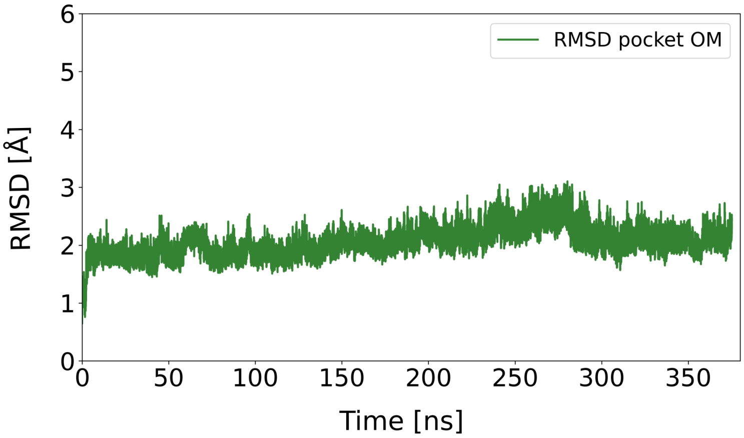
and analyze the trajectory, we used VMD (Humphrey et al., 1996) with some in-house *tcl* scripts.

6.3.2 Exploring the stability of PPS in β-CM with and without OM

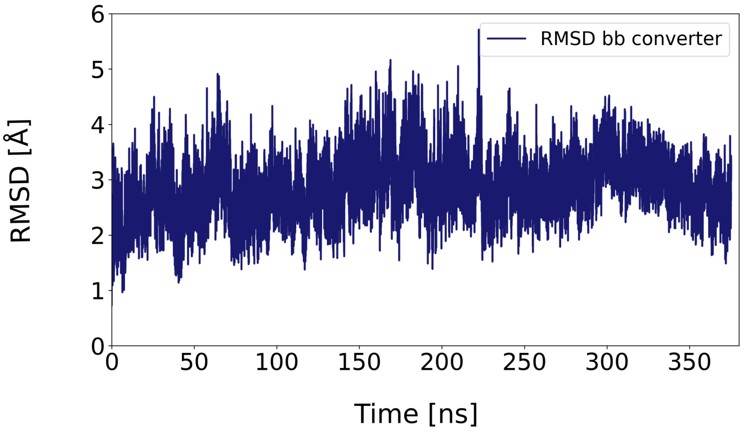
At first, we analyzed the stability of the system with OM bound. As shown in Figure 6.4, sub-figure (a) and (b) the protein is stable for the whole trajectory. The average RMSD of the backbone is equal to 2.3 ± 0.3 Å with respect to the coordinates of the X-ray structure of reference. Likewise, OM binding pocket remains stable with an average RMSD of 2 Å. To complete the analysis of the stability of the system with OM, we investigated the stability of the converter subdomain, being a highly flexible element in myosin and a force-generating element of the motor domain (Ménétrey et al., 2008). In Figure 6.4, sub-figure (c), is plotted the RMSD of the backbone atoms of the converter. We can notice that it fluctuates around an average value of 2.8 Å. There is also a moment in the trajectory during which it fluctuates the most, deviating from the initial coordinate of more than 5 Å.



(a) Timeseries of the RMSD of the backbone of the protein over 375 ns MD simulation, after alignment on the core atoms of the N-terminal subdomain of the reference structure (5N69.pdb).



(b) Timeseries of the RMSD of the side chains of the residues of OM binding pocket over 375 ns MD simulation, after alignment on the core atoms of the N-terminal subdomain of the reference structure (5N69.pdb).



(c) Timeseries of the RMSD of the side chains of the backbone of the converter over 375 ns MD simulation, after alignment on the core atoms of the N-terminal subdomain of the reference structure (5N69.pdb).

Figure 6.4: In sub-figure (a) the time-series of the RMSD of the backbone (bb) of the protein. The average RMSD value is 2.3 ± 0.3 Å. In sub-figure (b) the time series of the RMSD of the side chains of OM binding pocket. The average value of this RMSD is 2 ± 0.2 Å. In the last sub-figure (c) the RMSD of the bb of the converter subdomain. The average RMSD over 375 ns id 2.8 ± 0.5 Å.

6.3.3 Structural observables in PPS β-CM

To monitor more in detail the stability of the PPS structure we used some collective variables (geometrical observables), defined by Blanc for MyosinVI (Blanc, 2018) and adapted them for cardiac myosin. The monitoring of these observable is important to discriminate the structure of myosin in the various conformational states of the actomyosin cycle. Structural rearrangements are descriptive of different functional states of the motor when it undergoes transitions between states (from power-stroke to recovery-stroke). Geometric observables are not easy to define, but they help in understanding the dynamics of the main subdomains during conformational transitions. The set of structural observables designed by Blanc is descriptive of sub-transitions during the recovery stroke in MyosinVI (Blanc, 2018). The recovery stroke is the motion of the motor domain from the end of the rigor (R) to the pre-power stroke (PPS) state, during which the affinity of myosin for the track changes and the head unbinds from it (Fischer et al., 2005). The following are the observables considered to describe the end state of recovery stroke in cardiac myosin.

Observable	5N69.pdb [Å]	average in 375ns MD [Å]
c1	16.6	18.8 ± 1.6
c2	13.1	14.6 ± 1
c3	30.5	31.1 ± 1.2
c4	16.2	16.6 ± 0.6

Table 6.1: Collective variables in βCM: opening of the actin-binding cleft and inner cleft. The values reported are four distances (c1,c2,c3,c4) between Cα of residues of U50 and L50 measured over 375 ns MD. The values are compared with the ones calculated in bovine β-CM with OM bound (5N69.pdb) Here is the definition of the four distances (the numbering of the residues is the one of 5N69.pdb, chain A):

- c1: I274-F470;
- c2: K234-E466;
- c3: L366-M539;
- c4: V420-K596.

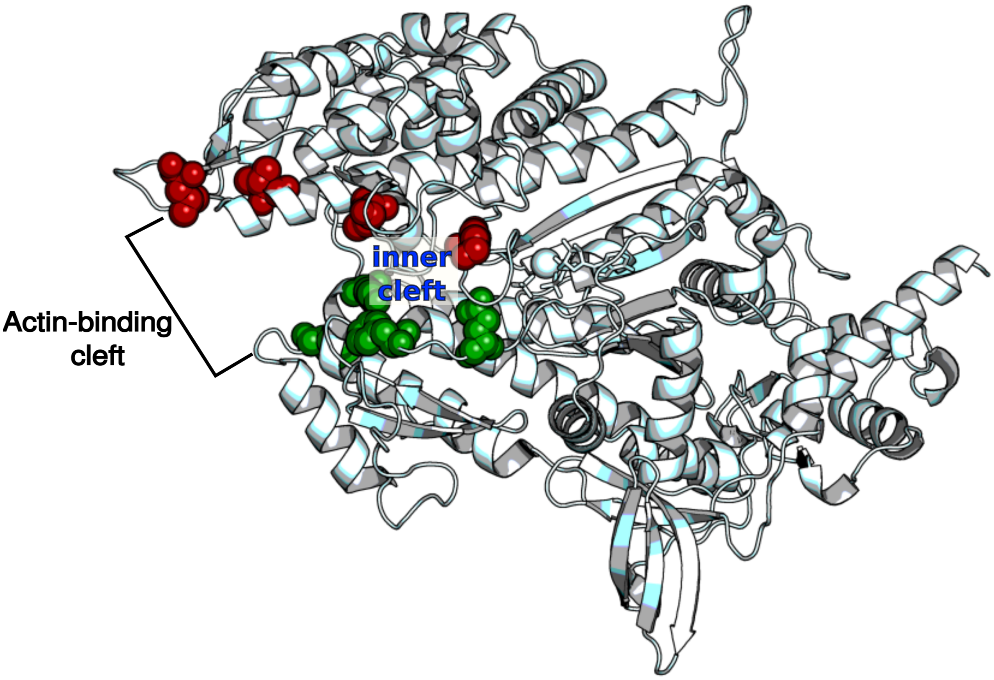


Figure 6.5: Cartoon representation of the model of β-CM used in the MD simulation. In spheres are represented the residues considered for the description of the opening of the clefts. In red the residues in the U50 subdomain and in green the ones in the L50 subdomain.

Observable	5N69.pdb [Å]	average in 375ns MD [Å]
d1	4.3	5.7 ± 1
d2	7.2	8.3 ± 1.8
dγ	3.6	4.9 ± 1

Table 6.2: Collective variables in βCM: interactions between switch loops (swI and swII) and between swII and $H_2PO_4^-$. The values reported are three distances (d1,d2,dγ), representing respectively d1 and d2 two salt bridges between swI and swII (d1: R243CA-E466CD and d2: R237CZ-E466CD) and dγ the interaction between G464 and the γ phosphate of ATP, being now $H_2PO_4^-$, upon hydrolysis. dγ is defined as the distance between G461N and O1$H_2PO_4^-$. The numbering of the residues is the one of 5N69.pdb, chain A.

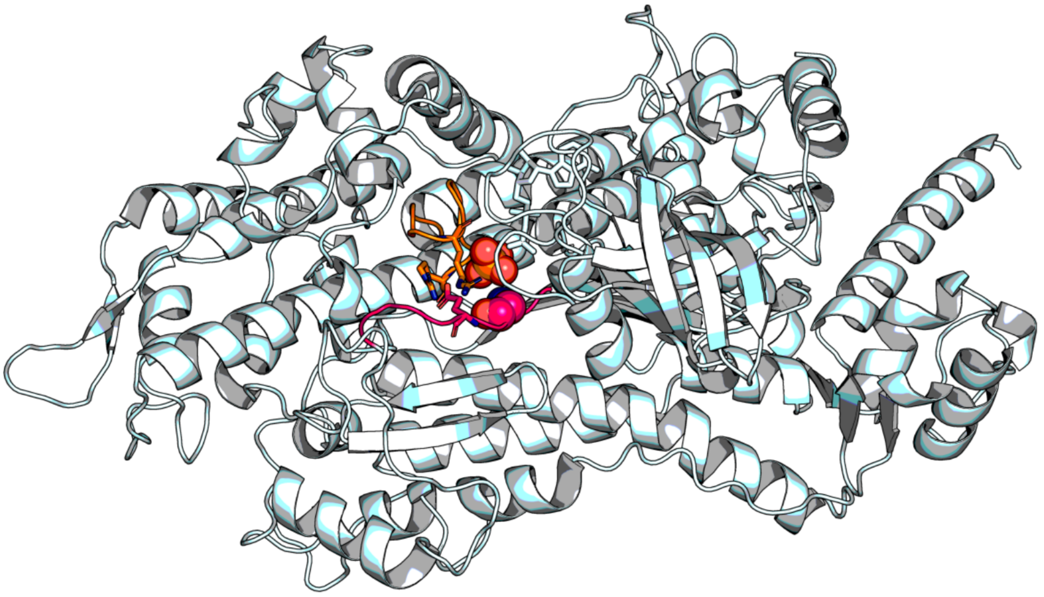


Figure 6.6: Cartoon representation of the model of β-CM used in the MD simulation. In sticks are represented the residues considered for the description of the interactions between the switch loops (swI and swII). In orange the residues of the swI and magenta the ones of swII. In spheres G464 (residue belonging to swII) and $H_2PO_4^-$.

Observable	5N69.pdb [Å]	average in 375ns MD [Å]
b1	2.8	2.8 ± 0.1
b2	3.0	3.1 ± 0.2
b3	2.8	2.9 ± 0.1
b4	2.9	2.9 ± 0.1
b5	2.8	2.8 ± 0.1

Table 6.3: Collective variables in βCM: interactions between the three central β-strands of the transducer: $\beta 4$; $\beta 5$ and $\beta 6$. The distances reported are H-bonds formed between the three strands. The H-bonds between $\beta 4$ and $\beta 5$ from b1 to b3 are defined as follows:

- b1: I174N ($\beta 4$)- GLY458O ($\beta 5$);
- b2: GLN172N ($\beta 4$) -F456O ($\beta 5$);
- b3: GLN172O ($\beta 4$)-GLY458N($\beta 5$).

The H-bonds between $\beta 5$ and $\beta 6$ are named b4 and b5 and are defined as follows:

- b4: I457O($\beta 5$)-I250N($\beta 6$);
- b5: V459N($\beta 5$)-I248O($\beta 6$).

The numbering of the residues is the one of 5N69.pdb, chain A.

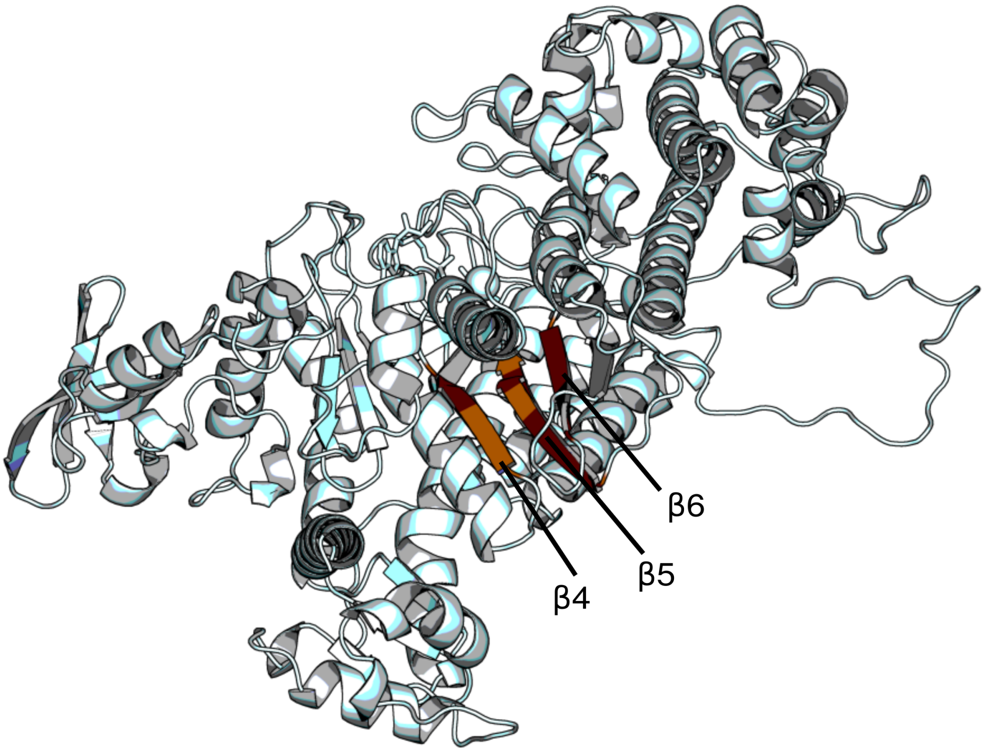


Figure 6.7: Cartoon representation of the model of β-CM used in the MD simulation. In brown are colored the three central β-sheets of the transducer for which we computed the distances reported in table Table 6.3

The data collected for the three classes of observables show that the cleft is partially open (c distances in Table 6.1), typical of PPS, since it is a weakly actin-bound state (Franz et al., 2020; Kühner and Fischer, 2011). In particular, the partially open cleft closes at the beginning of the power stroke, when myosin strongly binds actin (Holmes, 1997; Kühner and Fischer, 2011). The cleft in our simulation is representative of a PPS state and it seems to be a little more open, as found in the first three distances described in Table 6.1, which assume slightly larger values than the ones in the reference structure.

Moreover, swII is closed as evinced by the distance d1, whose average value is close to the reference one (see Table 6.2). This distance describes the formation of a salt bridge between the two switch loops (swI and swII), as defined in Table 6.2. The other two distances, chosen to define the closure of swII (d2 and dγ), variate more from the respective reference value (see Table 6.2). Probably d2 is not a precise observable for this simulation of cardiac myosin, since this salt bridge is not formed in the reference structure chosen as starting coordinates. The distance between G464 (swII) and $H_2PO_4^-$, described by dγ, is the H-bond that mediates the hydrolysis of ATP in ADP-P_i (Sasaki et al., 2003; Smith and Rayment, 1996). This distance in our trajectory seems unstable (Table 6.2 dγ). The distance increases since the beginning of the simulation and does not stabilize back to the reference value. This is due to a rotation of the inorganic phosphate in the active site, which occurs after the loss of H-bond with ILE462, a residue of $\beta 5$ (see Figure 6.8).

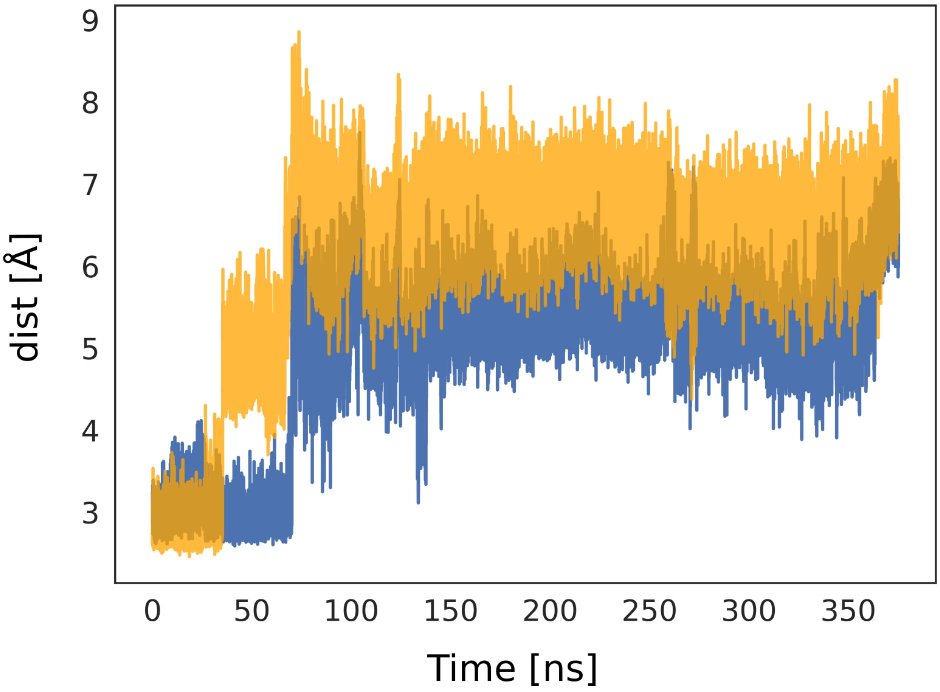


Figure 6.8: Timeseries of the evolution of the H-bonds formed by $H_2PO_4^-$ with swII (G464) and β5 (ILE462). The H-bond with G464N is depicted in light blue and the one with I462O is in light orange.

Based on the analysis of these structural observables, we can assert that the simulation of our PPS model with OM is a faithful representation of the PPS state of cardiac myosin throughout the whole trajectory.

6.3.4 RMSD analyses of β-CM in absence of OM

The RMSD analysis of the backbone of the protein showed completely different behavior in the system without OM (see Figure 6.9). Looking at the simulation, the system was unstable due to a huge movement of the converter (see Figure 6.10).

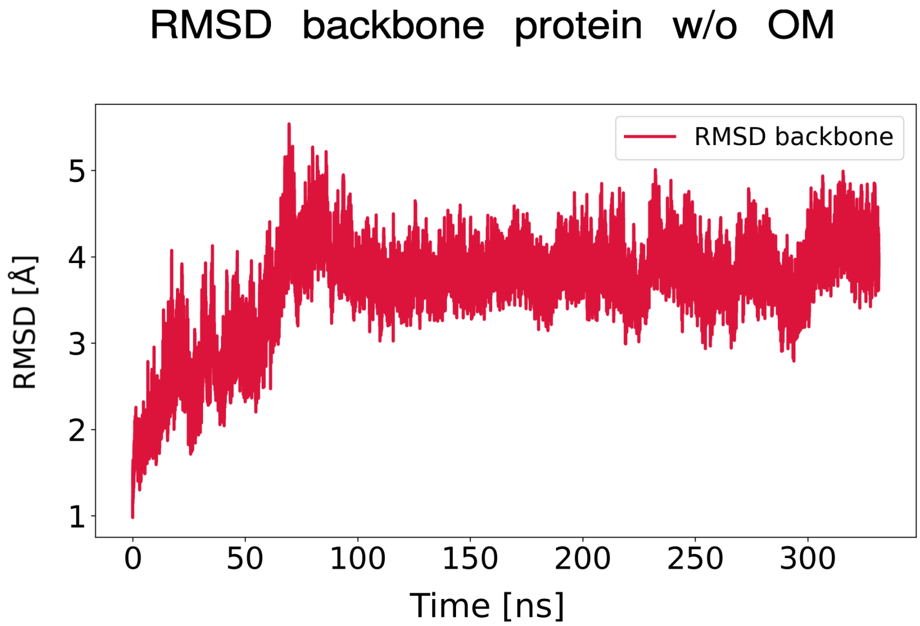


Figure 6.9: Timeseries of the RMSD of the backbone of the protein during 330 ns MD, to the X-ray reference structure. The average value of this RMSD is 3.7 ± 0.6 Å.

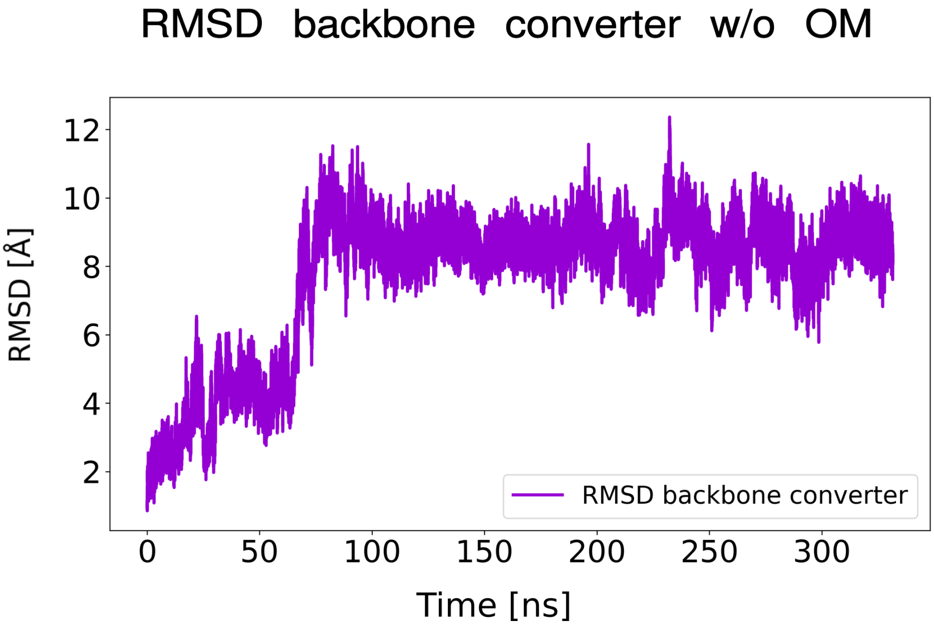


Figure 6.10: Timeseries of the RMSD of the backbone of the converter during 330 ns MD, to the X-ray reference structure. The average value of this RMSD is 7.7 ± 2 Å. The time-series shows that the position of the converter deviates considerably compared to the reference structure, reaching values higher than 10 Å.

We run two replicas of the same system without OM, using different initial velocities for each, to check whether it was a recurrent behavior of the protein or not. Also, in this case, we observed the same motion of the converter, whose position varies greatly from the reference one (see Figure 6.11).

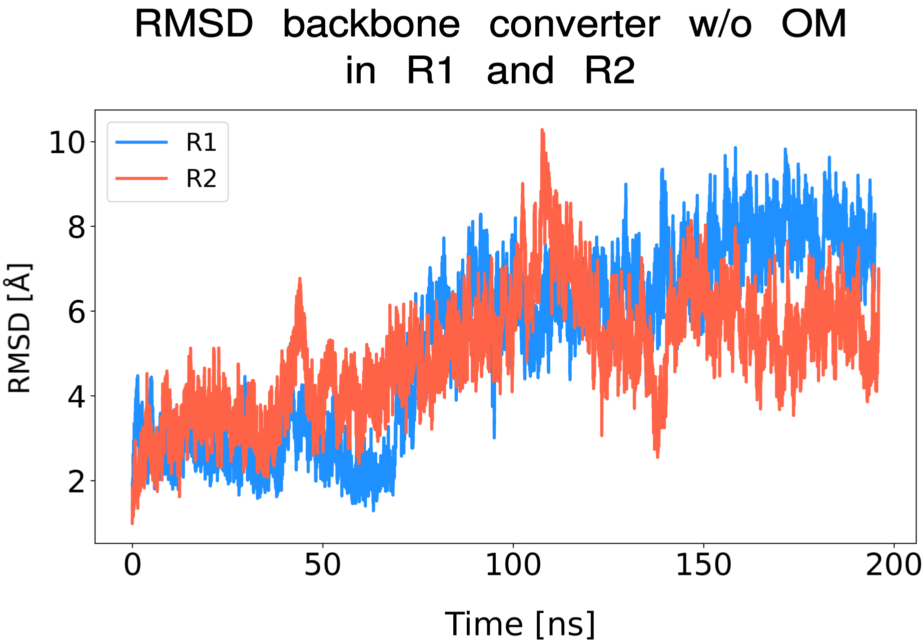


Figure 6.11: Timeseries of the RMSD of the backbone of the converter during 200 ns MD in two replicas R1 and R2, respectively in blue and coral. The average value of the RMSD is 5.3 ± 2 Å for R1 and 5 ± 1.4 Å in R2. The time-series show that the position of the converter deviates compared to the X-ray reference structure in both replicas.

We hypothesized that this behavior could relate to the forcefield chosen for the MD. We therefore, re-built the same system (β-CM with $ADP - MG^{2+}$ and $H_2PO_4^-$) with CHARMM (Chemistry at HARvard Molecular Mechanics) (Brooks et al., 2009), using CHARMM General Force Field (CGenff (`https://cgenff.umaryland.edu/`) (Vanommeslaeghe et al., 2012), version 2.2.0, to generate the parameters for the ligands. As shown in Figure 6.12 and Figure 6.13, also in this case the converter was deviating remarkably from the reference structure.

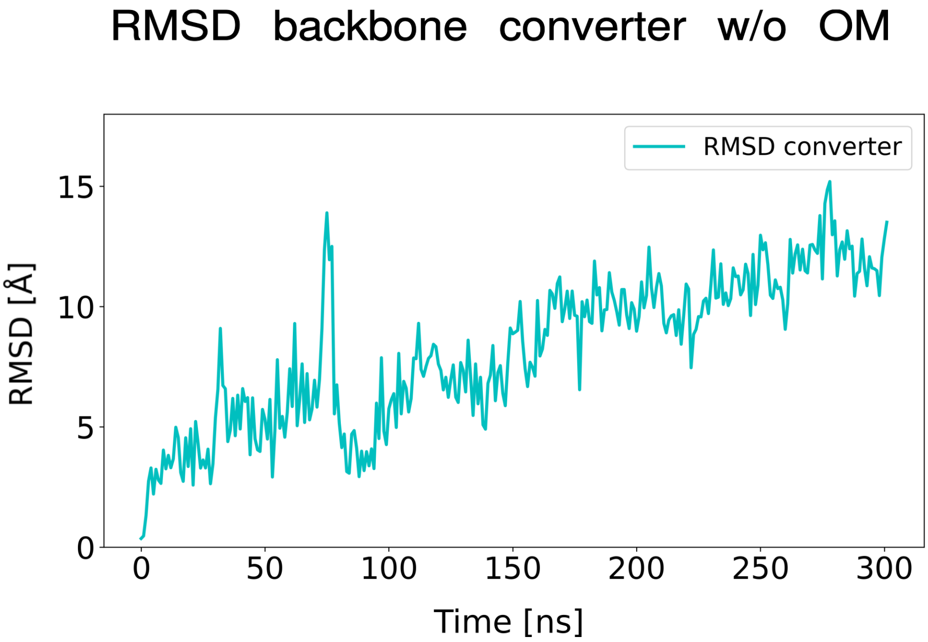


Figure 6.12: Timeseries of the RMSD of the backbone of the converter during 300 ns MD run with CHARMMff. The average value of the RMSD is 8.2 ± 3 Å. The time-series shows that the position of the converter deviates from the X-ray reference structure even using a different force field.

A brief video of the movement of the converter in this latest MD of β-CM can be found at the following link: `https://zenodo.org/record/6957242#.YupKHmFBxH4`.

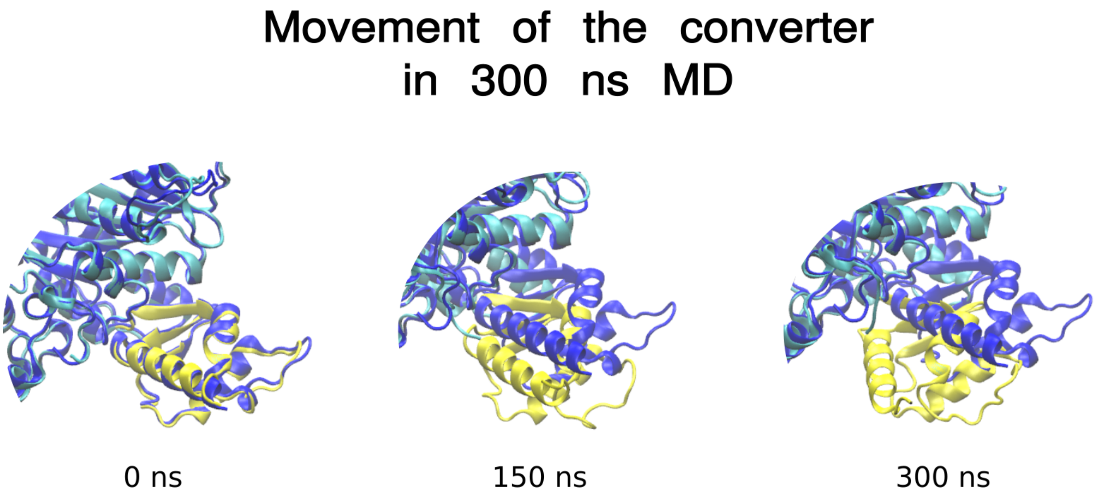


Figure 6.13: Zoom on the converter subdomain (in yellow) in three distinct frames of the MD of 300 ns. In blue is depicted the reference structure and in cyan the protein during the MD.

The huge repositioning of the converter seems to explore an intermediate position between the PR and PPS states. To verify this, we selected a frame of the latest trajectory in which the converter undergoes this rearrangement and we compared the coordinates of the converter with the ones of PPS and PR (Figure 6.14). For the PR reference structure, we chose the bovine X-ray structure (PSB code: 6FSA) (Planelles-Herrero et al., 2017), solved in complex with ADP and MG^{2+} and for PPS the one used until now (5N69.pdb). Figure 6.14 shows that the converter, in the currently used model, could explore an intermediate position, between PPS and PR. This motion can be driven by thermal fluctuations that favor the flexibility of the converter and possibly of other force-generating elements, such as Src homology 1 helix (SH1) and Relay Helix (RH).

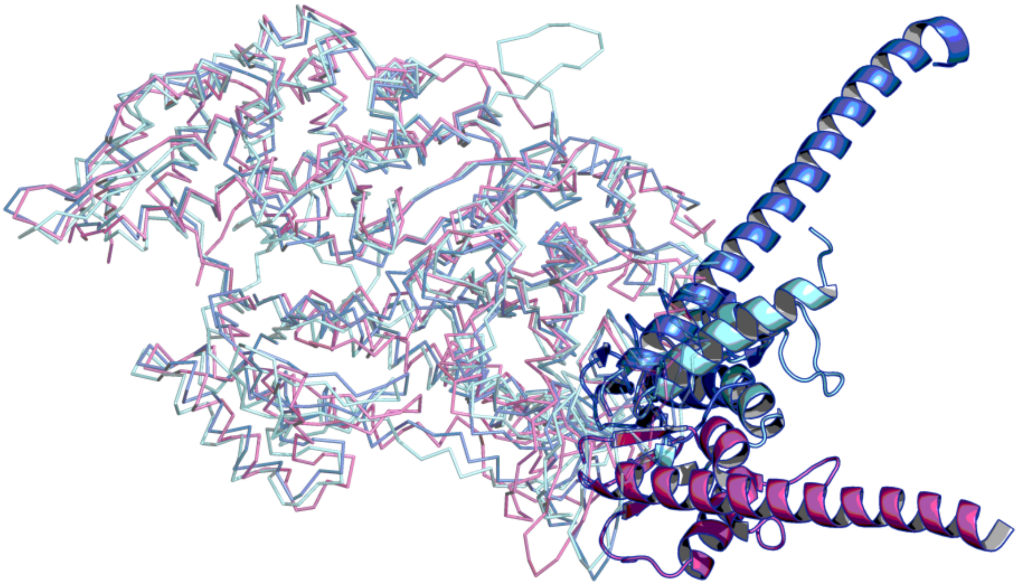


Figure 6.14: Visual comparison of the position of the converter in PPS (5N69.pdb), PR (6FSA.pdb), and a random frame of one of our trajectory in which it deviates from the PPS position. The PPS reference structure (5N69.odb) is depicted in blue, the frame of our MD is in cyan and the PR reference structure (6FSA.pdb) is in magenta. The converter-lever arm subdomain is represented with cartoons and the rest of the motor head with ribbons.

6.3.5 Improvement in the model of β-CM

So far the system without OM looks unstable, regardless of the force field chosen. What might then be the reason for this movement of the converter? Can it be due to the modeling of the protein? We, therefore, analyzed the protonation state of the titratable residues as a possible cause of instability of the system without OM, since it could play an important role in the dynamics of the protein.

6.3.5.1 *System preparation and MD simulation*

In the initial model, we chose the standard protonation state for all titratable residues. In this case, the most probable protonation state of these at neutral pH was predicted using Karlsberg (Rabenstein, 2000). According to this prediction, we assigned a non-standard protonation state to seven residues: ASP159 (N-terminal); ASP168 (N-terminal); ASP282 (U50); GLU217 (U50); GLU466 (swII); LYS351 (U50) and LYS429 (U50).

In Figure 6.15 are highlighted the seven residues that were ionized in a non-standard state.

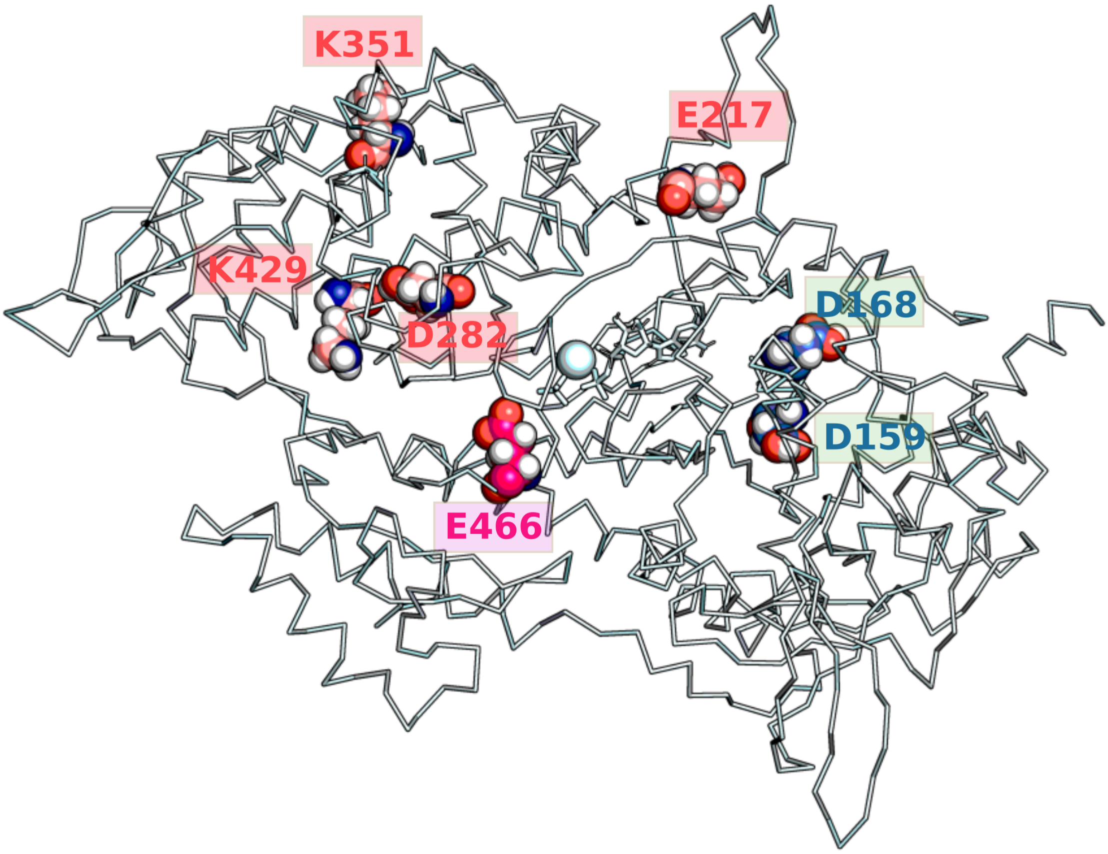


Figure 6.15: In spheres the seven residues to which we assigned a non-standard protonation state. In salmon red the residues located in the U50; in magenta a residue in the swII loop and in blue two residues of the N-terminal subdomain.

After re-building the complex with CHARMM, as described in the previous section, we used the Solution Builder module of CHARMM-GUI (Jo et al., 2008) to solvate and ionize the system. The final system without OM was composed of 193231 atoms, of which 180252 of the TIP3 water model, 170 Sodium ions, and 176 Chloride ions. The system's size with OM was 193362 atoms, of which 180330 of TIP3 water model, 170 Sodium ions, and 176 Chloride ions. We simulated both the system with and without OM for 300 ns with Gromacs2020 (Abraham et al., 2015) and CHARMM36mff (Huang et al., 2016). The protocol of these simulations envisages a minimization until reaching a $F < 1000.0kJ/mol/nm$, an equilibration in NVT with Nose-Hoover thermostat (Evans and Holian, 1985) for $125ps$ followed by $100ps$ equilibration in NPT with Parrinello-Rahman barostat (Parrinello and Rahman, 1981). Finally, a production with leap-frog integrator algorithm with velocity-rescale thermostat (Bussi et al., 2007) and Parrinello-Rahman barostat (Parrinello and Rahman, 1981).

6.3.5.2 *RMSD analysis of the new system with different protonation states of titratable residues*

We computed the RMSD of the backbone atoms of the converter for both systems (with and without OM). As shown in Figure 6.16 the converter is always stable if OM is bound, but it still fluctuates if the modulator is not present.

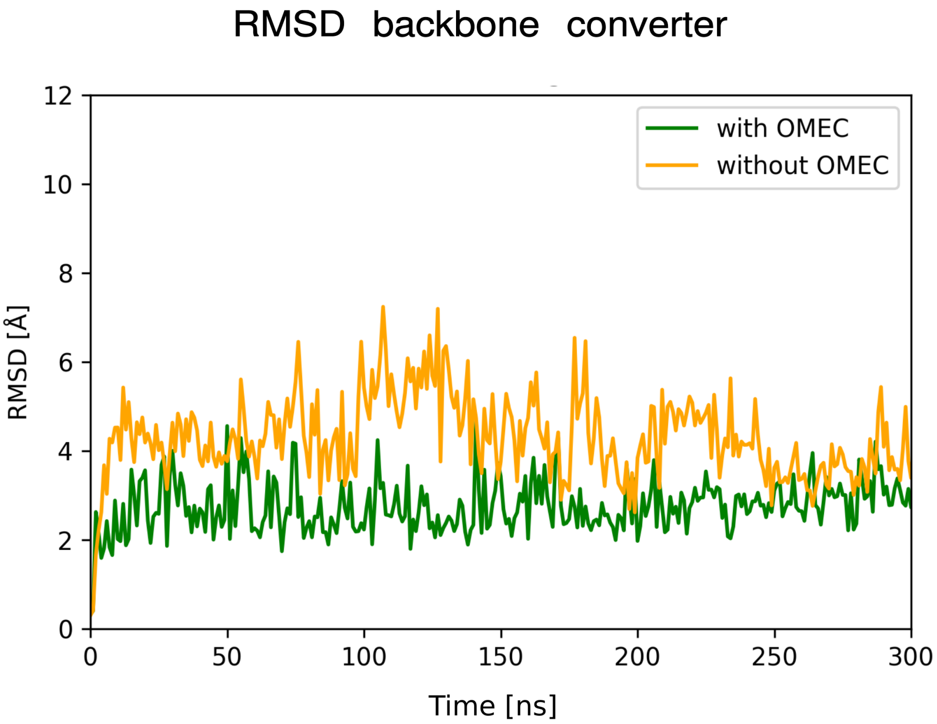


Figure 6.16: Timeseries of the RMSD of the backbone of the converter, after alignment on the N-terminal subdomain of the reference structure. The two time-series represent 300 ns MD simulations run with CHARMMff both with (green) and without OM (orange) bound. The average value of the RMSD in the system with OM is 2.7 ± 0.6 Å. For the system without OM the average value is 4.3 ± 0.5 Å. The converter seems stable in both systems.

This fluctuation concerns the priming of the lever arm. In the simulation without OM, the lever arm is more primed, compared to the system with the allosteric modulator bound (see `https://zenodo.org/record/6958265#.YuqBY2FBxH4`, a video in which in green is colored the converter-lever arm domain of the system with OM and in yellow the one without OM).

6.4 CONCLUSIONS & FUTURE PERSPECTIVES

Having a stable model of Cardiac myosin in PPS is of paramount importance for us since we want to design allosteric modulators of the motor in this state of the actomyosin cycle. Our analysis of the protonation state of titratable residues shows that the choices made at beginning of the modeling may result in very different stability of the system. This is not negligible, considering the amount of both computational and human time one may spend to find a plausible solution to the problem. In the end, our model appears stable both in the presence and absence of Omecamtiv mecarbil. The relevance of the protonation state of titratable residues for the stability of the system in PPS could be further explored, by investigating the surrounding of said residues. This could result in a better understanding of the cause of the motion observed for the converter subdomain, as well as of the possible physiological relevance of the structure in which this motion occurs. It might have intrinsic stability between PR and PPS states. This could be in agreement with what was hypothesized by Fornili and co-workers for their simulations of β-CM in PPS, without any modulator (Hashem et al., 2020). In particular, they have shown how, in absence of OM, the converter-lever arm domain samples a different state (Hashem et al., 2020). It is not clear if the motor domain moves toward a following power-stroke state or to a previous structure of the recovery state (Hashem et al., 2020). The question of the existence of a putative intermediate state between PR and PPS remains open. Something similar has been recently discovered, by our group and the group of Houdusse, in MyosinVI, for which a putative intermediate state of the recovery stroke, linking PR and PPS has been described and solved with X-ray crystallography (Blanc et al., 2018). Another interesting movement observed in the simulations is the different priming of the lever arm, depending on whether OM is bound or not. It opens new questions about the mechanism mediated by OM to stabilize the PPS state. Planelles-Herrero and co-workers proposed an induced fit mechanism for the binding of OM (Planelles-Herrero et al., 2017), however, in our "apo" simulations we observed a more primed lever arm with respect to the structure they crystallized in absence of OM (5N6A.pdb). The latter was obtained from the structure with OM bound, after in situ proteolysis (Planelles-Herrero et al., 2017). This, probably, results in a retention of some memory of the presence of the drug bound. This memory is, instead, lost in our MD of 300 ns, so that the lever arm can adopt a different, more primed position. Most likely, the re-priming captured during our MD simulation is a more realistic fluctuation of the lever arm in "apo" PPS β-CM, since it is not biased by the presence of a modulator. This suggests that OM could bind with an induced fit mechanism, by entering the cavity that is formed in a certain moment and blocking the converter-lever arm domain in a less primed position. The remarkable pliancy of the converter in PPS could be a common feature among myosin classes (Ménétrey et al., 2005).

Our goal was to model a stable structure of β-CM in PPS starting from the most recent X-ray structure (5N69.pdb), both in the presence and absence of OM. The motor domain was already fully stable in the simulation with OM, regardless of the protonation state of titratable residues. The change in the stability of the motor domain becomes considerable in the system without the allosteric modulator, especially for the dynamics of the converter. The non-standard protonation state of some titratable residues improved the stability of the converter, keeping it in a primed position, characteristic of the PPS state. Future analyses may involve the creation of different models of β-CM without modulator and by sequentially protonating the seven residues identified in our study. The analysis of the chemical surrounding of these residues may be useful to understand which protonation state is crucial to maintain the stability of the converter in the system without any modulator. Once the model is stable, it can be used i.e. for a virtual screening campaign to find good binders of cardiac myosin in PPS. This can, for example, be done with our protocol implemented in Chem*Flow* (earlier described in Chapter 5). In the next chapter, we will present our pharmacological studies conducted using the stable model of β-CM to answer some important dilemmas.

7

β CARDIAC MYOSIN: ALLOSTERIC MODULATION

7.1 CURRENT ALLOSTERIC MODULATORS OF CARDIAC MYOSIN

Cardiomyopathies represent one of the leading causes of premature death in the world (Benjamin et al., 2019; Clippinger et al., 2019). Among them, Dilated cardiomyopathy (DCM) and hypertrophic cardiomyopathy (HCM) are two of the most common progressive genetic diseases (Day et al., 2022; Jefferies and Towbin, 2010; Rohde et al., 2018) that lead to HFrEF and arrhythmias (Day et al., 2022; Grillo et al., 2021). DCM and HCM are characterized by opposite pathophysiological indicators: dilated left ventricular chamber and impaired contractility in DCM (Jefferies and Towbin, 2010); hypertrophic non-dilated left ventricle and excessive ejection fraction (EF) in HCM (Day et al., 2022; Maron and Maron, 2013). Current treatments aim at alleviating symptoms, without effectively eradicating the source of these diseases ((Chuang et al., 2021)). Groundbreaking discoveries to treat these disorders are allosteric modulators, small molecules able to modulate the impaired force producing of the motor, directly targeting the myosin isoform. Allosteric modulators act on the motor's mechanochemistry without affecting the intracellular calcium concentration (Day et al., 2022; Houdusse, 2020), resulting in less invasive and more efficient therapies (Masri and Olivotto, 2022). In recent years both allosteric activators and inhibitors of β-CM have been designed (Kawas et al., 2017; Morgan M. Rogers-Carter1 2 and Christianson1, 2017; Morgan et al., 2010; Olivotto et al., 2020; Rosano and Ponikowski, 2019).

7.1.0.1 *Mavacamten*

A promising small molecule that acts as an allosteric inhibitor of β-CM is Mavacamten (Mava) (Figure 7.1) (Morgan M. Rogers-Carter1 2 and Christianson1, 2017) also known as MYK-461 (Kawas et al., 2017).

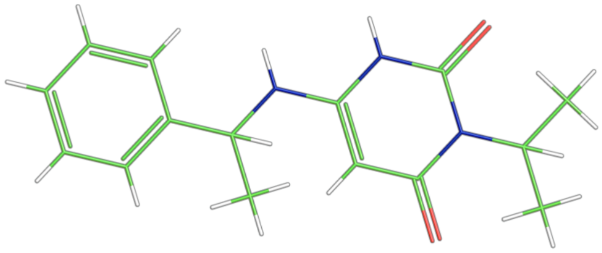


Figure 7.1: 2D representation of Mava (6-(1S)-1-phenylethylamine-3-propane-2-yl-1H-pyrimidine-2,4-dione)

Since HCM is characterized by an excessive sarcomeric power output (Morgan M. Rogers-Carter1 2 and Christianson1, 2017), defined as a product of the force generated by myosin's heads and the velocity of movement along actin (Morgan M. Rogers-Carter1 2 and Christianson1, 2017; Spudich, 2014), the strategy used to design Mava envisages the decrease in force-generation (Morgan M. Rogers-Carter1 2 and Christianson1, 2017). Mava inhibits β-CM contractility in HCM, by modifying the kinetics of actin-activated phosphate release (Rohde et al., 2018), without slowing the release of ADP (Kawas et al., 2017; Morgan M. Rogers-Carter1 2 and Christianson1, 2017). Therefore Mava reduces, in a dose-dependent manner, the phosphate release and the duty ratio of cardiac myosin (Morgan M. Rogers-Carter1 2 and Christianson1, 2017). Consequently, the sarcomere contraction decreases (Kawas et al., 2017; Morgan M. Rogers-Carter1 2 and Christianson1, 2017). Mava stabilizes a "super-relaxed" (SRX) state (Robert-Paganin et al., 2018; Rohde et al., 2018), viz. a state in which myosin heads are inactive, thus not interacting with actin (Robert-Paganin et al., 2018). By stabilizing this state, Mava reduces the number of active myosin heads, able to undergo the powerstroke (Robert-Paganin et al., 2018; Rohde et al., 2018). Mava was tested in Phase 3 of a clinical trial on 251 patients with HCM (Helms et al., 2022). This study reports that Mava improves exercise capacity, left ventricular outflow tract (LVOT) obstruction, and health condition (Helms et al., 2022). Moreover, it reduces left ventricular hypertrophy and left atrial volumes (Helms et al., 2022). Mava is currently waiting for FDA approval (Helms et al., 2022).

7.1.0.2 *Aficamten*

Another myosin inhibitor Aficamten (CK-274) (Figure 7.2) (Chuang et al., 2021) is in phase 3 of clinical trial (Masri and Olivotto, 2022). The design of an additional small molecule to inhibit β-CM rises from the need to overcome some clinical limitations of Mava. More precisely, Chuang and co-workers first aimed at reducing the human half-life ($t_{1/2}$) of 7-9 days of Mava (Grillo et al., 2019) to a value that could allow a daily administration of the drug (Chuang et al., 2021). Besides, they want a compound with a clear pharmacokinetic (PK)/pharmacodynamic (PD) relationship (Chuang et al., 2021) and minimal cytochrome (CYP) P450-induced drug-drug interactions (DDIs) (Chuang et al., 2021).

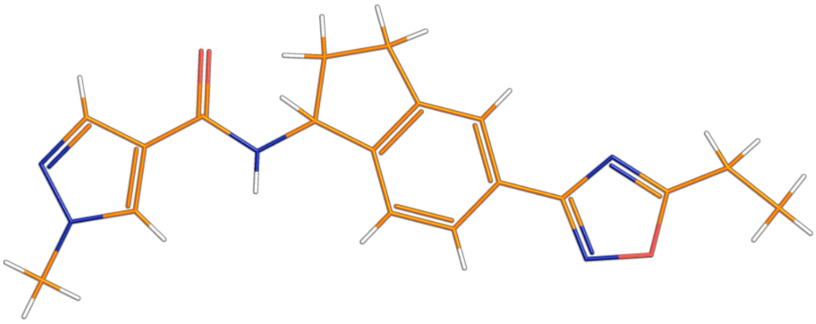


Figure 7.2: 2D representation of Aficamten (CK-274) (N-(1R)-5-(5-ethyl-1,2,4-oxadiazol-3-yl)-2,3-dihydro-1H-inden-1-yl-1-methylpyrazole-4-carboxamide)

Chuang and co-workers conducted a high-throughput screening (HTS) campaign on bovine cardiac muscle measuring ATP hydrolysis rate (Chuang et al., 2021). After various stages of optimization of a selective indoline compound (Chuang et al., 2021), Chuang *et al* selected Aficamten as a next-generation cardiac inhibitor. Aficamten satisfies the three improvement criteria set at the beginning of their study. It shows shorter $t_{1/2}$ with respect to Mava, reaching a steady state within two weeks from administration (Chuang et al., 2021; Masri and Olivotto, 2022). It has a wide therapeutic window (Chuang et al., 2021) and in phase II of the clinical trial, it exhibits a favorable safety profile even at high dosage (Masri and Olivotto, 2022). Outcomes on phase 3 of a clinical trial are expected soon (Masri and Olivotto, 2022).

7.1.0.3 *Omecamtiv Mecarbil*

In 2010 Morgan and co-workers conducted HTS, designed and optimized the first allosteric activator of β-CM, named Omecamtiv Mecarbil (OM) (Morgan et al., 2010) (Figure 7.3). OM is also known as CK1827452/AMG423 (Morgan et al., 2010). OM enhances cardiac contractility in HF disorder, by accelerating the release of inorganic phosphate from the active site, without changes in sarcomeric calcium concentration (Day et al., 2022; Helms et al., 2022; Morgan et al., 2010).

The data reported by Morgan *et al* provide activity values for all the series of compounds they analyzed, expressed in terms of AC_{40} (Morgan et al., 2010). AC_{40} is the concentration of each compound resulting in a 40% increase of ATPase activity of the sarcomere, at a calcium concentration that yields 25-50% of the maximum calcium-dependent activation (Morgan et al., 2010).

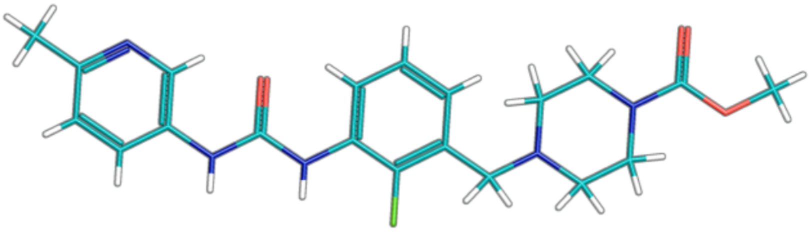


Figure 7.3: 2D representation of OM (methyl-4-(2-fluoro-3-(3-(6-Methylpyridin-3-yl)ureido)benzyl)piperazine-1-carboxylate)

OM has the best AC_{40} value in the series of compounds, with a value of 0.58 μM (Morgan et al., 2010). Isothermal titration calorimetry (ITC) experiments measured an affinity of OM for β-CM of 0.29 μM (Planelles-Herrero et al., 2017). In 2021 OM underwent Phase 3 of a clinical trial in GALACTIC-HF (Global Approach to Lowering Adverse Cardiac Outcomes Through Improving Contractility in Heart Failure) (Teerlink et al., 2021). This approach compares OM with placebo on more than eight thousand patients with HFrEF and EF <= 35% (Teerlink et al., 2021). The outcome of this clinical trial shows that patients with more severe HF benefit most from OM treatment, with a progressive risk reduction for the primary endpoint of HF events (Teerlink et al., 2021). However, it does not reduce cardiovascular death or improve total symptoms (Helms et al., 2022). Given these rather humble results, it remains uncertain whether OM will be approved as a drug for the treatment of HF (Helms et al., 2022).

7.1.0.4 *Danicamtiv*

More recently a novel potential allosteric activator of β-CM has been designed: Danicamtiv (Voors et al., 2020), also known as MYK-491 (Figure 7.4) (Fernandes et al., 2019; Grillo et al., 2021). Like OM, also Danicamtiv accelerates the release of inorganic phosphate from myosin, increasing cross-bridge formation of cardiac muscle (Grillo et al., 2021). Pre-clinical tests from a 2a trial show a direct dose-dependent effect of Danicamtiv on atrial contractility, as well as improved left ventricular systolic function (Grillo et al., 2021; Voors et al., 2020). In particular, it has a half-maximal active concentration (AC_{50}) of 6 μM for ventricular myofibrils and of 3 μM for atrial ones (Voors et al., 2020). In general, Danicamtiv increases the ATPase rate of cardiac myosin at 3 μM (Voors et al., 2020).

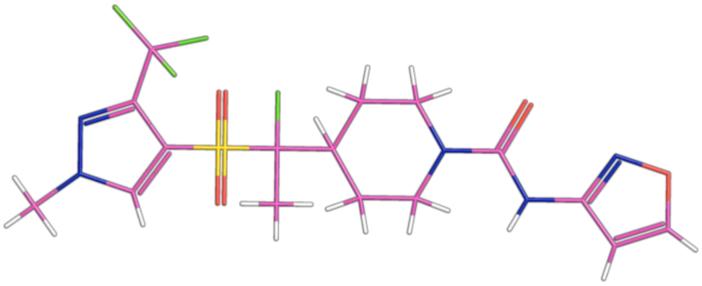


Figure 7.4: 2D representation of Danicamtiv (4-(1R)-1-3-difluoromethyl-1-methylpyrazol-4-yl-sulfonyl-1-fluoroethyl-N-1,2-oxazol-3-yl-piperidine-1-carboxamide)

Danicamtiv is selective for cardiac S1 myosin (Voors et al., 2020) and it is currently in phase 2 of clinical investigation.

7.2 *In silico* STUDIES OF β-CM'S ALLOSTERIC ACTIVATOR OM

As previously discussed, OM represents the first attempt at drug design of selective allosteric activators of β-CM. Although its mechanochemical action on cardiac myosin is clear (accelerating the kinetics of P_i release from the motor), its effect on the stability and dynamics of the protein has not been clarified yet. We performed atomistic studies, with all-atom MD simulations (as presented in Chapter 6) of β-CM in PPS state both in complex with OM and not. Our work is not the first computational study of the dynamics of β-CM in complex with the modulator OM. Notoriously, the group of Fornili has conducted pioneering *in silico* studies about the effects of OM on the dynamics of the protein as well as its mode of action (Ha et al., 2020; Hashem et al., 2017).

7.2.0.1 *OM and β-CM in Near-Rigor state*

In 2017, Fornili and co-workers investigated the mechanochemical role of OM from the X-ray structure of human β-CM, in the near-rigor state of the actomyosin cycle (Hashem et al., 2017; Winkelmann et al., 2015). In this study (Hashem et al., 2017), they conducted MD simulations in the micro-second scale of their β-CM model both in the presence and absence of OM (Hashem et al., 2017; Winkelmann et al., 2015). The system without OM will henceforth be referred to as Apo. Their analyses show that OM reduces the quasi-rigid rotation of the converter-lever arm subdomain, compared to the Apo system, lessening the overall flexibility of the protein (Hashem et al., 2017). This effect is mediated by the contacts that OM establishes with the subdomains with which it interacts in its binding pocket (N-terminal, L50, RH, SH1, and converter-lever arm), but also by the stabilization of pre-existing interactions among the domains (Hashem et al., 2017; Winkelmann et al., 2015). Another interesting observation concerns the interaction between OM and two regions (Loop1 and G-helix that is part of the U50 subdomain) containing two residues (I201 and A223) which, if mutated into Threonine, are associated with dilated cardiomyopathy (DCM) phenotype (Al-Numair et al., 2016; Hashem et al., 2017). Fornili *et al* monitored the interaction network between the drug and this region, calculating a connection score (Hashem et al., 2017). The favorable connection score measured, suggests that OM interacts with this region, stabilizing the structure (Hashem et al., 2017). Therefore, the drug could be beneficial for the patient affected by DCM, with the above-mentioned mutations (I201T and A223T), because it would counteract the effect of the mutations (Hashem et al., 2017). Even binding to the near-rigor state, OM enhances cardiac contractility, by possibly modifying the dynamics of G-helix, involved in the closure and opening of the actin-binding cleft, respectively in weakly- or strongly-actin bound state (Hashem et al., 2017; Kühner and Fischer, 2011). Further studies may involve specific cardiac myosin mutants with OM bound, to verify if the drug could reverse the effect of these mutations on the dynamics and stability of the protein (Hashem et al., 2017).

7.2.0.2 *OM and β-CM in PPS state*

The previous work of Fornili *et al* was published simultaneously with the release of the most recent structure of β-CM, always in complex with OM, but in another state of the actomyosin cycle. This structure is the bovine cardiac myosin in PPS state (PDB code: 5N69), which we have also used in our study (Planelles-Herrero et al., 2017). The group of Fornili decided, therefore, to investigate the dynamics of β-CM in presence of OM, in PPS (Hashem et al., 2020). Analyses of the micro-second scale MD simulations of the system in presence of OM, show that OM remains tightly bound keeping a conformation close to the one in the initial structure (Hashem et al., 2020). Moreover, the different subdomains are more coupled in presence of OM, with respect to the Apo simulations (Hashem et al., 2020), as already observed in the simulations of the near-rigor state (Hashem et al., 2017), thanks to stable contacts that OM establishes with them (Hashem et al., 2020). OM stabilizes the overall structure reducing the flexibility of the protein (Hashem et al., 2020). The Apo simulations of the PPS state show higher flexibility of the motor domain, in particular in the converter-lever arm domain (Hashem et al., 2020), as similarly observed for the Apo trajectories of near-rigor state (Hashem et al., 2017). Among the long-distance interactions that OM couples, there are interactions between swII and nearby regions, closing possible routes for P_i release (Hashem et al., 2020). This evidence is in agreement with some experiments that reported an OM-mediated decrease in P_i release, in absence of actin (Liu et al., 2015; Malik et al., 2011). However, these observations do not exclude a different escape path of P_i from a tunnel, that involves the opening of P-loop (backdoor II), in the transition to a state with high affinity for actin (Cecchini et al., 2010; Hashem et al., 2020). In this case, OM could favor the opening of this last route and mediate the P_i release in presence of actin (Hashem et al., 2020; Liu et al., 2015).

Besides the stability of β-CM in PPS, Fornili and co-workers explored the interactions between the protein and OM, as well as with a set of its analogs (Manickam et al., 2018, 2017), to determine their effects on the motor function (Hashem et al., 2020). Molecular docking of OM both in PPS and near-rigor structure hints that OM binds with higher affinity β-CM in PPS, rather than in near-rigor state (Hashem et al., 2020). This result is in agreement with experimental binding affinities measured with ITC for the motor in complex with different nucleotide forms (Planelles-Herrero et al., 2017). Fornili *et al* also docked both in PPS and near-rigor state structures some of the analogs of OM for which activity values are known (Hashem et al., 2020; Manickam et al., 2018, 2017). Then, they evaluated the relative predicted binding free energies. Also, in this case, the molecules have a higher affinity for the PPS state than for the near-rigor state, demonstrating their selectivity for PPS (Hashem et al., 2020). By computing the difference between the predicted affinity for a molecule in two different conformational states of the target protein, one could get the efficacy of the molecule and then compare it with experimental activity values (Cecchini and Changeux, 2022). The group of Fornili did so, designing various linear regression models that were then fit with experimental data of the analogs of OM (Hashem et al., 2020). The comparison between predicted and experimental activities of these series of compounds, OM included, shows that the most important contributions in the binding are H-bonds patterns and steric interactions (Hashem et al., 2020). Moreover, the compounds that belong to the class of the more active ones, tend to couple different subdomains, just like OM (Hashem et al., 2020). Finally, they assume that the model derived using this dataset is probably too restrictive to be considered a general predictive model for β-CM activators (Hashem et al., 2020).

7.3 DESIGN OF CARDIAC MYOSIN'S ALLOSTERIC ACTIVATORS

In this section, we will describe how we tackled the arduous question of how to design β-CM allosteric activators.
Our approach was different with respect to the one used by Fornili *et al* (Hashem et al., 2020). We did not use a linear regression method to predict activities. We asked ourselves if the binding affinity was discriminant for a molecule to be a cardiac activator, since we suppose that activators of β-CM shift the conformational equilibrium towards the PPS state, by stabilizing it. We will present the analyses run over the model of β-CM in PPS with both a standard protonation state of ionizable residues and the latest one, in which we considered the non-standard state of said residues.

7.3.0.1 *Initial structure of β-CM*

The first step of our study involves choosing the structure of the protein to be used. We have previously described the analyses of the dynamics of β-CM in PPS, in presence of OM proposed both by us (Chapter 6) and the group of Fornili (Hashem et al., 2020) and we have seen that there are some common conclusions. In particular, we both observed that OM stabilizes the PPS structure, especially by limiting the motion of the converter-lever arm domain that shows high pliancy in absence of the modulator. The hypothesis that activators of β-CM stabilize the PPS state has been demonstrated by the molecular dynamics studies with OM, analyzed by us and by Fornili and co-workers (Hashem et al., 2020). Moreover, the allosteric binding pocket of OM, either is formed with a possible induced fit mechanism (Planelles-Herrero et al., 2017), or it does not seem opened in the short timescale Apo simulations, due to the movement of the lever arm that explores a more primed position (see Chapter 6 and the video `https://zenodo.org/record/6958265#.YuqBY2FBxH4`). Based on this evidence, we decided to use a representative OM-bound PPS structure as initial coordinates for our investigation.
The most representative structure of the protein, corresponding to the most populated cluster center, was extracted from the MD simulation with OM bound and used as initial coordinates, both in the system with the standard state of titratable residues and in the one with the non-standard state of said residues. Since the trajectory of the system with a standard protonation state of ionizable residues was produced by Amber, we performed a cluster analysis with AmberTools, imposing an RMSD of 2 Å to the residues of the binding site of OM, excluding the hydrogens. The MD with the different protonation state of titratable residues was run with Gromacs, thus the centroid was chosen after a cluster analysis performed with the "gromos" cluster method (Daura et al., 1999), defining an RMSD cut off of 1.5 Å with respect to the heavy atoms of OM binding site.

7.3.0.2 *Initial dataset*

We had to select a dataset to start with, to define a protocol for designing allosteric activators of cardiac myosin. We chose the series of analogs of OM, described by Morgan *et al*, for which biochemical activities are quantified as AC_{40} (Morgan et al., 2010). OM and its analogs, which we will call A1, A2... A14 and A22, A23, and A24 from now on are the seventeen compounds extracted from the work of Morgan and co-workers (Morgan's compounds) (Morgan et al., 2010). MarvinSketch (Marvin version 17.12, ChemAxon (`https://chemaxon.com`)) was used to draw these molecules. The final 3D structures were prepared (protonation ad neutral pH, searching for tautomers and stereoisomers) with our in-house software *PrepFlow* (Sisquellas and Cecchini, 2021). Various conformers were then generated using RDKit (Landrum and Others, 2016). The molecules show a range of activity values whose extremes are OM and A3, respectively the best and the worst activator in the series (Morgan et al., 2010).

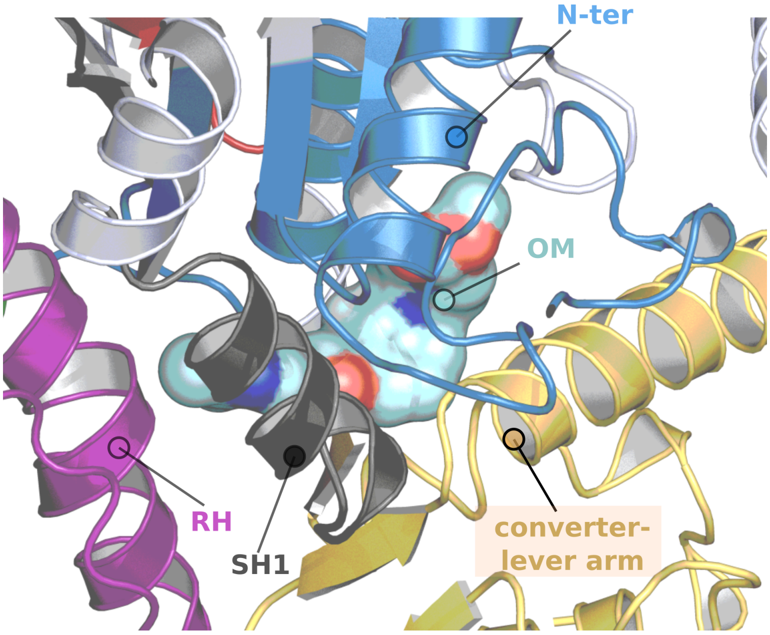
LIGAND	AC_{40} **[μM]**
A1	1.4
A2	2.5
A3	32
A4	1.8
A5	0.7
A6	1.4
A7	1.1
A8	2.5
A9	7.3
A10	2.5
A11	3.6
A12	0.6
A13	0.7
A14	0.8
A22	1.7
A23	4.5
A24	0.58

Table 7.1: Experimental activity values for the congeneric series of ligands measured as AC_{40} (Morgan et al., 2010). A24 corresponds to OM and it has the best activity value among all the molecules.

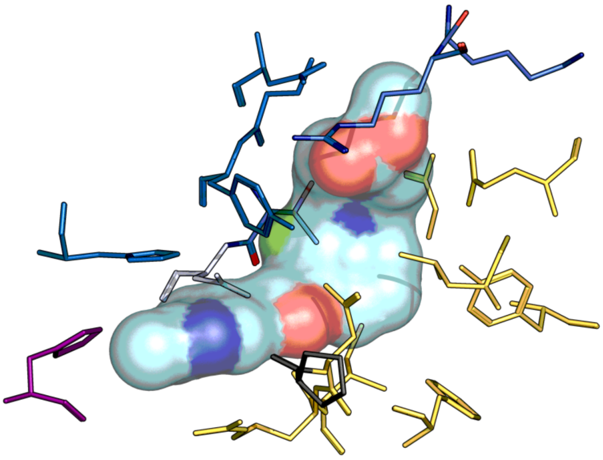
7.3.1 Chem*Flow* on β-CM PPS

7.3.1.1 *DockFlow on β-CM PPS*

To predict binding affinities of the series of compounds under investigation, we used Chem*Flow*. At first, we performed a re-docking of OM in its binding pocket using the bovine cardiac X-ray structure (5N69.pdb) (Planelles-Herrero et al., 2017). In Figure 7.5 a representation of the buried pocket of OM in bovine β-CM (5N69.pdb). The center of mass of OM was computed from the X+ray structure and used as the center for defining the sphere of the searching space in PLANTS, for the first experiment of re-docking performed with Dock*Flow*. The radius of the sphere was set to 15 Å and the default values were kept for the other parameters (Korb et al., 2007).



(a) Cartoon representation of bovine cardiac myosin (5N69.pdb), zooming on OM bound in its buried pocket.



(b) Representation of the binding surface of OM and in lines the residues within 5 Å of OM.

Figure 7.5: In sub-figure (a) a cartoon representation of the subdomains with whom OM interacts. In sub-figure (b) represented in lines the residues within 5 Å of OM, keeping the color code of the subdomain as in sub-figure (a). In both images, OM is depicted as its surface in the binding pocket.

OM was properly re-docked, as shown in Figure 7.6, thus we decided to use PLANTS to dock all the molecules in the representative structure of the protein chosen as starting coordinates.

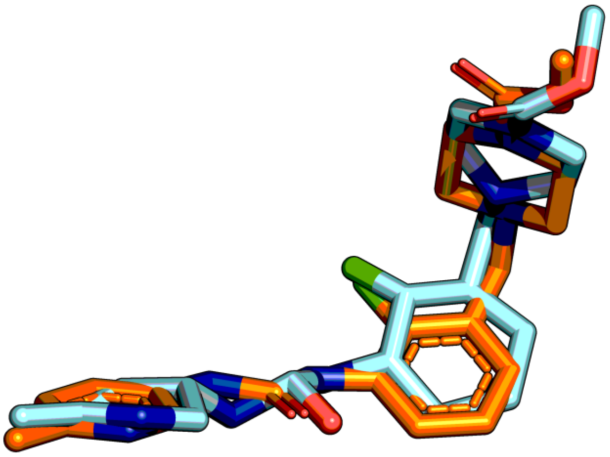


Figure 7.6: Result of the re-docking performed with PLANTS (Korb et al., 2007). The searching space was defined as a sphere with the center corresponding to the center of mass of OM and the radius was set to 15 Å. In orange sticks the re-docked molecule and cyan OM in bovine β-CM. The docking score for this pose is -99.5 and the RMSD to OM in the crystal structure is 1.8 Å.

Since the dataset is a series of structurally similar molecules (all analogs of OM), we hypothesised that they could bind with a similar orientation to the co-crystallized OM (Boström et al., 2006; Cournia et al., 2020). Moreover, since our goal is to predict relative binding free energies with a more accurate method than docking, such as end-point free energy calculations (MM/PB(GB)SA) (Montalvo-Acosta and Cecchini, 2016; Srinivasan et al., 1998), initial coordinates of the systems, obtained by docking, must be optimal. The initial results were not satisfactory, since there were some molecules for which PLANTS(ChemPLP) was not able to find a binding mode close to the one of OM. In Figure 7.7 an example of the best poses obtained with PLANTS for Morgan's compounds.

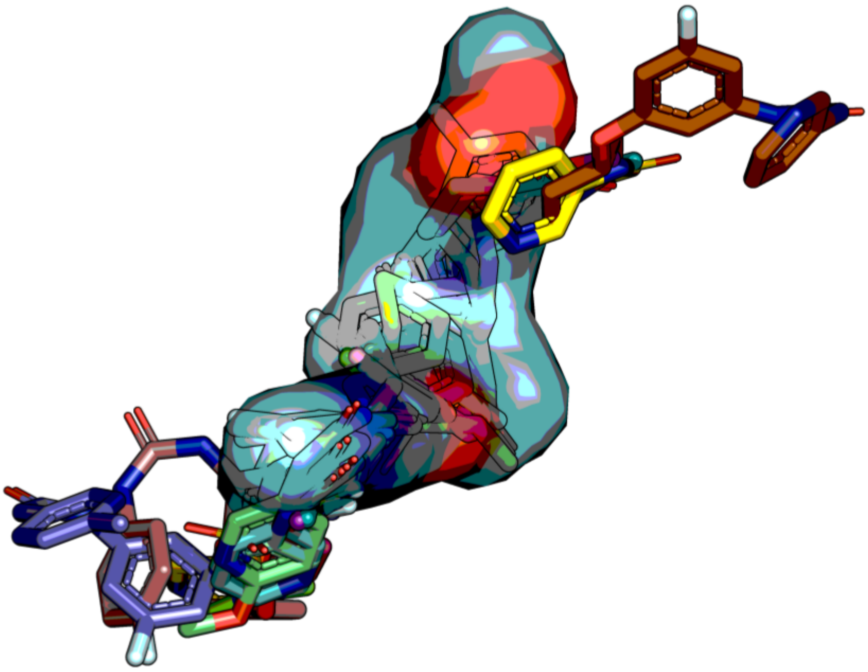


Figure 7.7: Best docking poses (first ones) of the analogs of OM. Not all the molecules are correctly docked in the OM binding pocket. Some are more exposed to the solvent losing key interactions. In sticks the analogs of OM and cyan, color-coded according to the polarity of the atoms, the surface of OM in bovine cardiac myosin (5N69.pdb).

As we can see, some molecules are shifted from the OM binding pocket occupying a space more exposed to the solvent, losing some possible key interactions. We, therefore, tried several combinations of the various parameters that can be set in PLANTS, such as the flexibility of some chemical groups or side chains, the radius of the sphere, the number of ants, or the evaporation rate factor (Korb et al., 2007, 2006), to find the more suitable set-up for our molecules. After several attempts to optimize the parameters, we chose the ones that gave the best results: scoring function ChemPLP; search speed 1; number of ants 20; evaporation factor 0.25; iteration scaling factor 1 (Korb et al., 2006). After visual inspection, we selected the first pose of the conformers that best overlaps OM in the pocket. (see Figure 7.8 as example for A3).

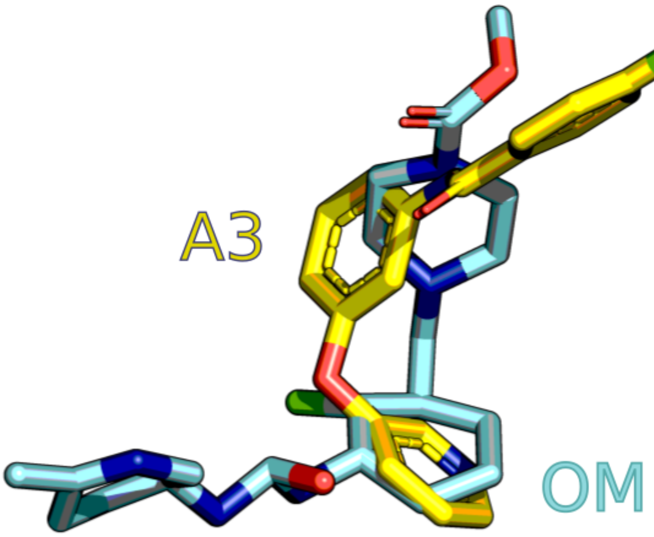


Figure 7.8: Best docking pose obtained with PLANTS for A3 (in yellow) with the new set-up. The docking score is -90.7. The molecules overlap the shape of OM (in cyan) and better fit the pocket.

7.3.1.2 *LigFlow and ScoreFlow on β-CM PPS*

We continued our analysis using Lig*Flow* and Score*Flow*. These are the other two modules of Chem*Flow*, developed to compute charges for the ligands and then run MD simulations for free-energy rescoring with end-point methods, respectively (see Chapter 5). For our dataset, we computed RESP charges and then calculated relative binding free energies. For the latter calculation, we used the MMGBSA approach over a trajectory of 10 ns for each system.

7.3.1.3 *Results: in silico predictions VS experimental data*

Finally, to understand if with this protocol we were able to properly classify Morgan's compounds according to their biochemical activities, we ranked them for predicted ΔG of binding and compared the ranking with the activity values. In Figure 7.9 the plot shows that we were not able to rank OM as the best binder (lowest value of predicted ΔG) and A3 as the worst one, in line with the experimental data (Morgan et al., 2010). Moreover, there was no correlation between the predicted binding affinities and the activity values (see Figure 7.10).

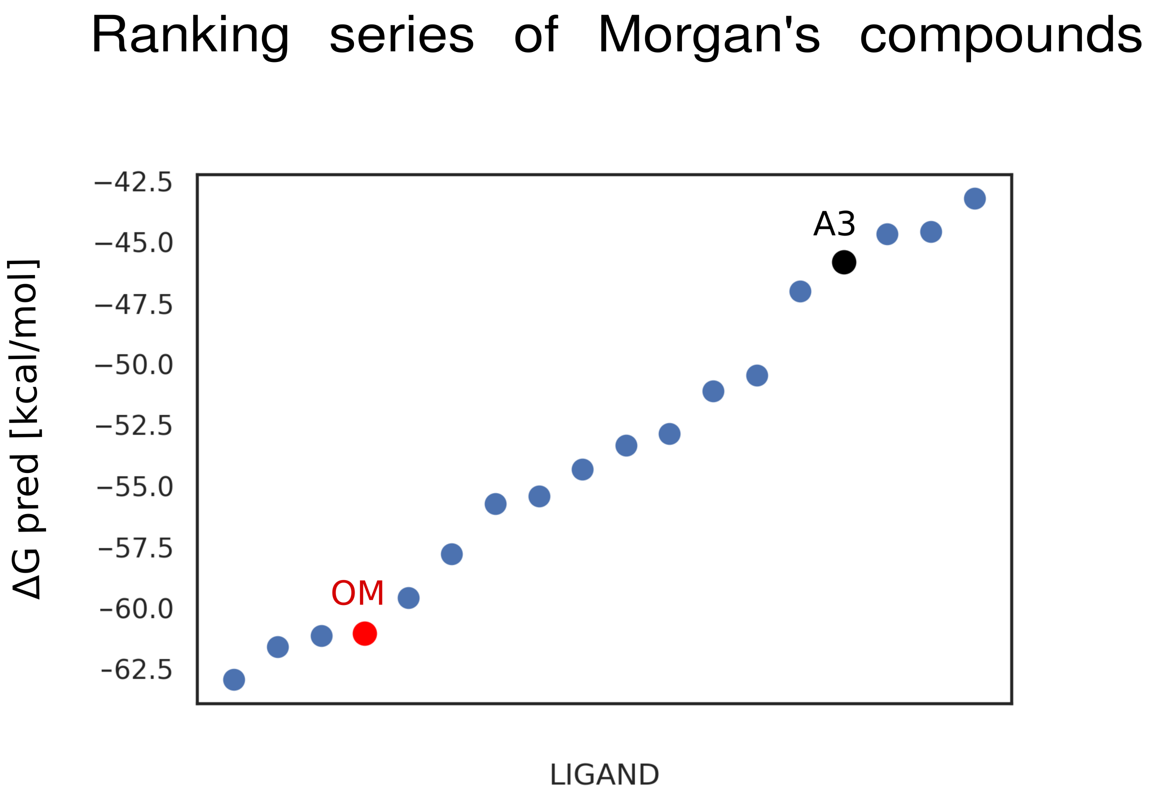


Figure 7.9: Ranking of the series of the compounds under investigation, according to the predicted free energy of binding (ΔG pred). The values are the results of the MMGBSA calculation performed with GB model 2 (Case et al., 2021) and internal dielectric constant (ϵ) equal to 4, over a trajectory of 10 ns per each system.

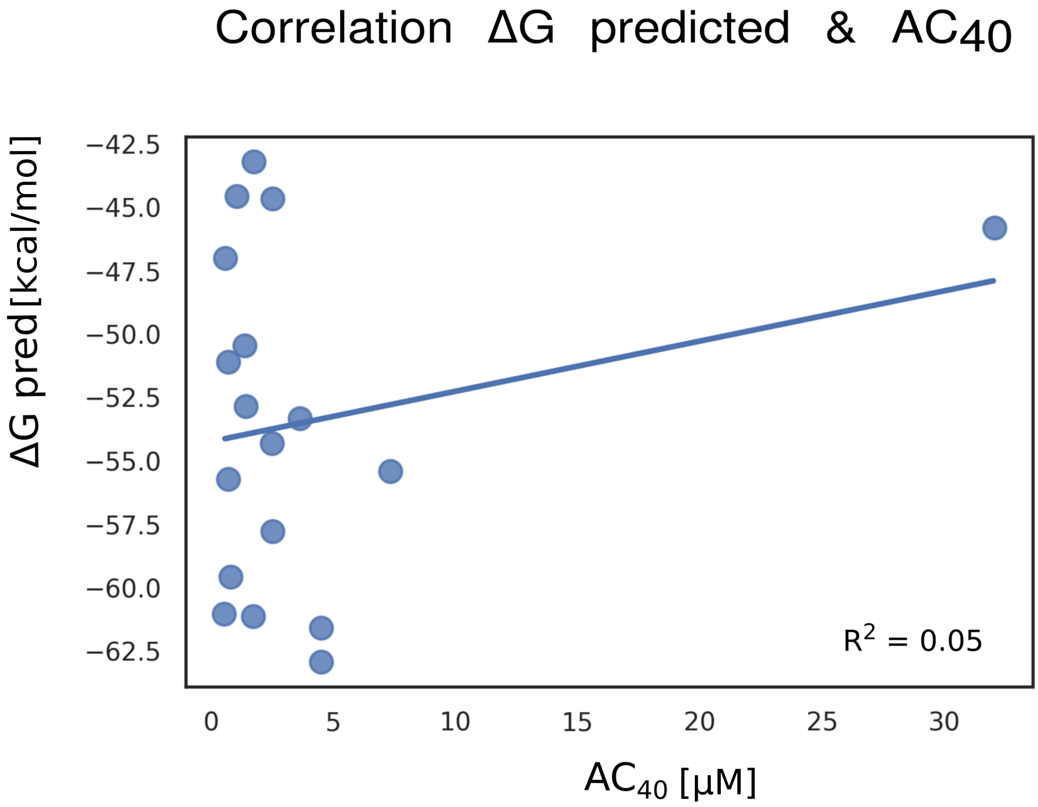


Figure 7.10: Correlation plot between the predicted ΔG obtained with MMGBSA calculation and AC_{40} values (Morgan et al., 2010). The R^2 is 0.05, meaning no correlation between the two quantities.

7.3.2 Improvements in the protocol

The negative results obtained so far were a reflection of a sub-optimal protocol for the systems we were investigating. Therefore, we decided to make substantial changes in the docking protocol and then relaunch all calculations with Chem*Flow*.

7.3.3 Chem*Flow* 2.0 on β-CM PPS

7.3.3.1 *DockFlow 2.0 on β-CM PPS*

First of all, as the initial structure for the docking, we selected the cluster center of the MD simulation run for the model of β-CM with the different protonation state of titratable residues (see Chapter 6). Then, we change the docking program and we used AutoDock VINA (VINA) (Trott and Olson, 2009), already implemented in Chem*Flow*. The search space was defined as a 15 x 15 x 15 Å grid box, centered on OM, and 10 poses were generated for each compound. The results were astonishing even for those molecules for which PLANTS had failed to find a correct binding mode. In Figure 7.11 is reported the example of the new docking pose of A3. The molecule occupies the binding pocket with an orientation similar to OM in the X-ray of bovine β-CM. In Figure 7.12 we can see how all the Morgan's compounds are docked overlapping the shape of OM in the pocket.

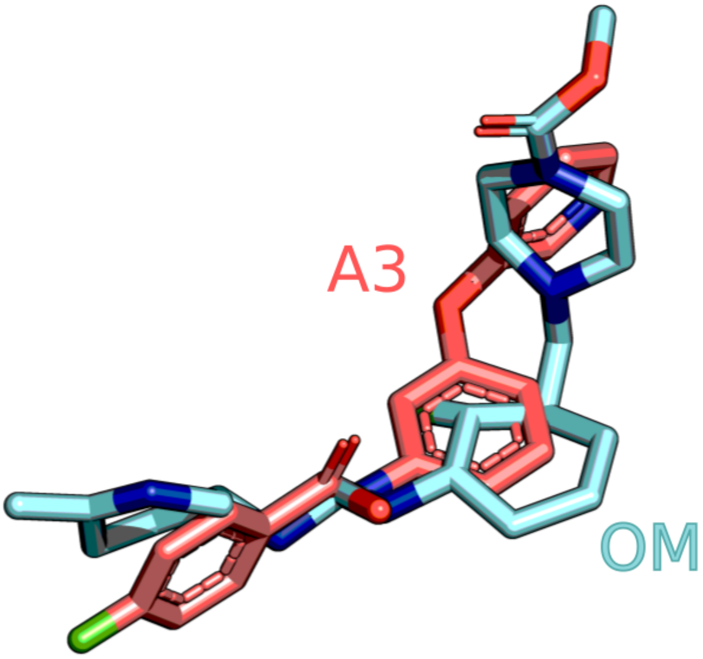


Figure 7.11: Best docking pose of A3 (in coral) obtained with VINA (Trott and Olson, 2009) compared with OM (in cyan) in the X-ray structure of bovine cardiac myosin (5N69.pdb).

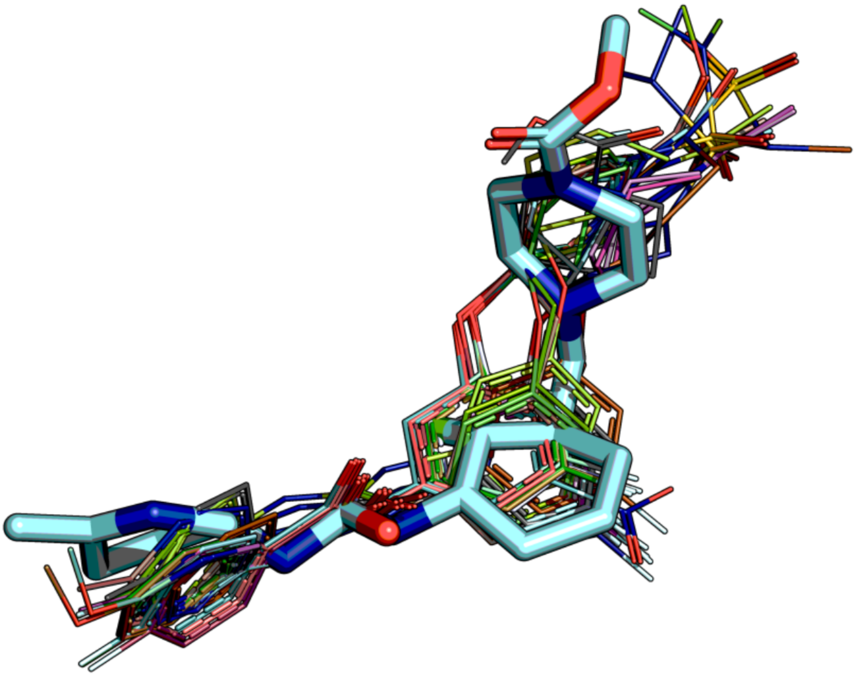


Figure 7.12: Best docking pose of all Morgan's compounds obtained with VINA (in colored lines), compared with OM (in cyan) in the X-ray structure of bovine cardiac myosin (5N69.pdb).

Given the much better docking results, we looked at the ranking of the compounds based on the scoring of VINA. The results in Figure 7.13 show that the ranking based on the docking scores is not enough accurate to classify OM as the best binder. A10, indeed, results the best binder with a docking score of -10.8 kcal/mol. By contrast, VINA can rank A3 as the worst binder (black dot in Figure 7.13), in agreement with the activity data provided by Morgan and co-workers (Morgan et al., 2010).

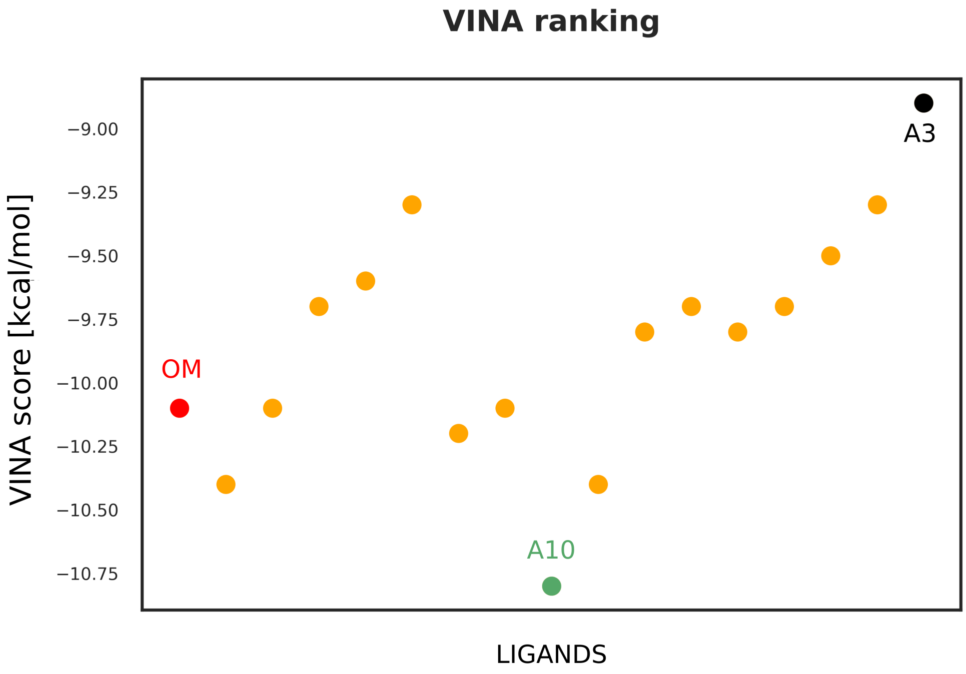


Figure 7.13: Ranking of Morgan's compounds based on VINA scores. In red dot OM (VINA score = -10.1 kcal/mol); in black dot A3 (VINA score = -8.9 kcal/mol) and in green dot A10 (VINA score = -10.8 kcal/mol). A10 is ranked as the best binder according to VINA scores.

7.3.3.2 *LigFlow and ScoreFlow 2.0 on β-CM PPS*

Since predictions based on docking are still not accurate enough to properly rank Morgan's compounds, we investigated the ranking after a more accurate level of prediction. We run 10 ns MD of all the complexes generated by VINA and we computed relative binding free energy with MMGBSA, using the same protocol as in Section 7.3.1.

7.3.3.3 *Results: in silico predictions VS experimental data 2.0*

Morgan's compounds were correctly ranked, being OM (A24) the best one in terms of predicted binding affinity and A3 the worst one (see Figure 7.14). A10, instead, is predicted as a medium affinity binder, in agreement with the AC_{40} data that classify it as such (see Table 7.1).

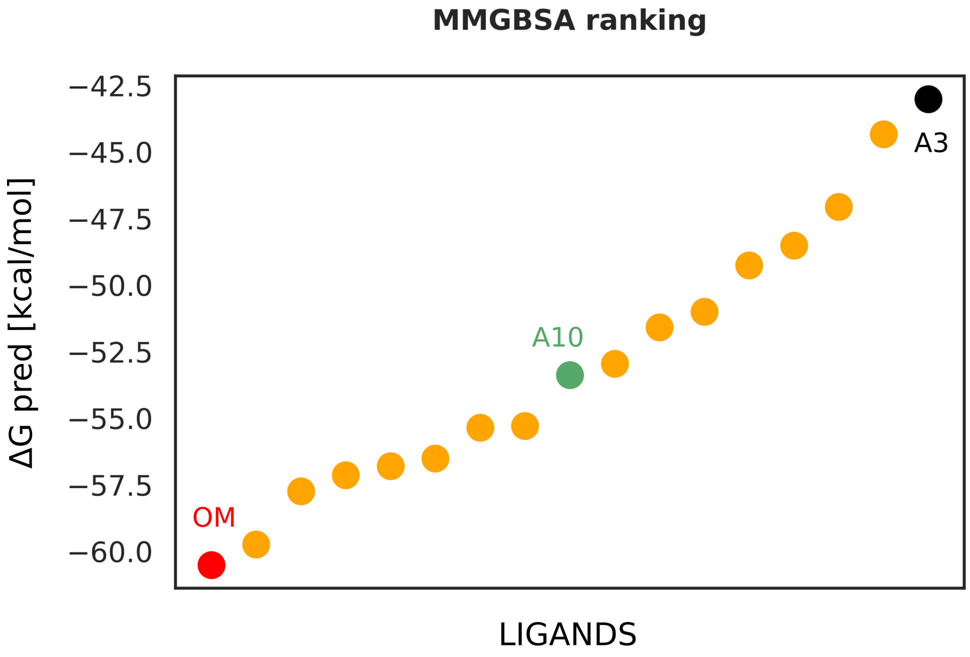


Figure 7.14: Ranking of Morgan's compounds based on MMGBSA scores. In red dot OM (ΔG predicted = -60.5 kcal/mol); in black dot A3 (ΔG predicted = -43 kcal/mol) and in green dot A10 (ΔG predicted = -53.4 kcal/mol). OM is ranked as the best binder according to MMGBSA scores.

We tested the same protocol adding other four molecules to our dataset: one inhibitor (F33), two activators (D33 and C29), and a molecule that is a good binder but shows no significant activity (A28). The structures of these molecules are confidential. Moreover, we introduced the known inhibitor Mava (Morgan M. Rogers-Carter1 2 and Christianson1, 2017). Once again, our protocol can classify activators (Morgan's compounds; C29 and D33) from inhibitors (Mava and F33) or inactive compounds (A28) (see Figure 7.15).

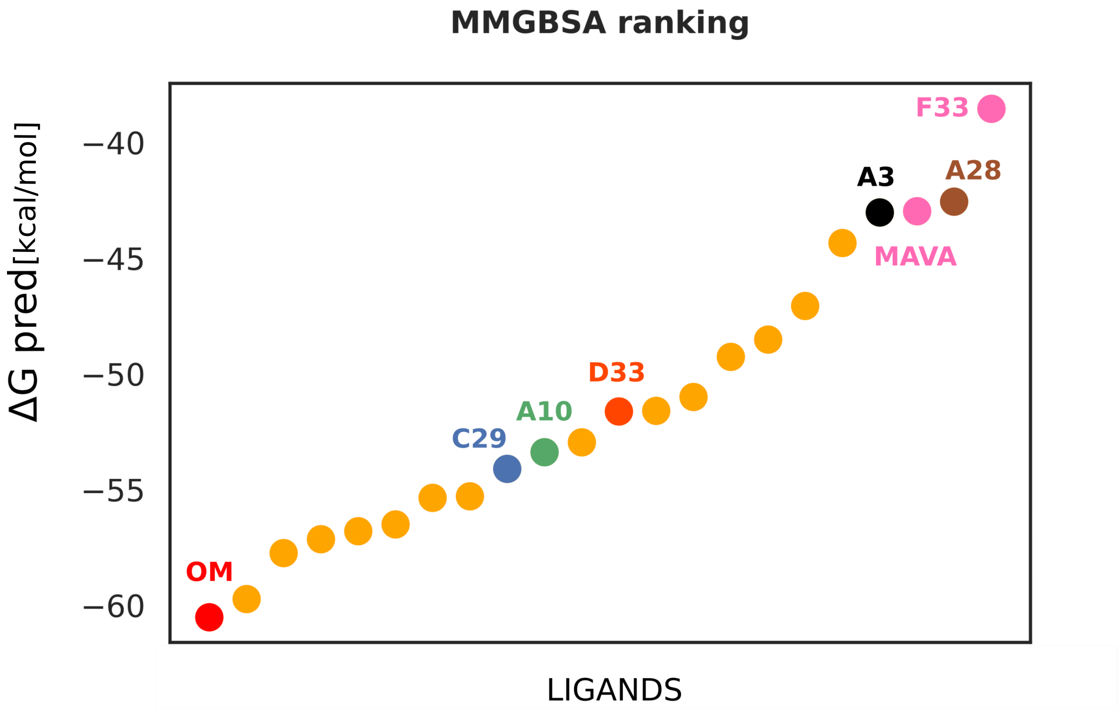


Figure 7.15: Ranking based on MMGBSA scores of Morgan's compounds; Mava and four other molecules with known efficacy. In red OM (ΔG predicted = -60.5 kcal/mol); in black A3 (ΔG predicted = -43 kcal/mol); in green A10 (ΔG predicted = -53.4 kcal/mol); in pink the two inhibitors Mava (ΔG predicted = -43 kcal/mol) and F33 (ΔG predicted = -38.5 kcal/mol); in brown the inactive molecule A28 (ΔG predicted = -42.5 kcal/mol); in dark orange one activator D33 (ΔG predicted = -54 kcal/mol) and in blue the other activator C29 (ΔG predicted = -51.6 kcal/mol). OM is ranked as the best binder whereas A3, the known inhibitors, and the inactive molecules as the worst binders, according to MMGBSA scores.

From the values of predicted affinity, computed with MMGBSA, one could define a boundary around -46 kcal/mol for hit prediction in a vHTS campaign to identify plausible activators of cardiac myosin.

7.3.4 What makes a small molecule a cardiac myosin activator?

So far, we set a protocol capable of qualitatively discriminating the Morgan's compounds, by classifying the best activators from the worst/inactive/inhibitors ones (Morgan et al., 2010). We, therefore, explored the hypothesis of a possible correlation between the predicted relative affinities and the biochemical activity values (AC_{40}). After computing this relationship no correlation emerged ($R^2 = 0.23$) (see Figure 7.16).

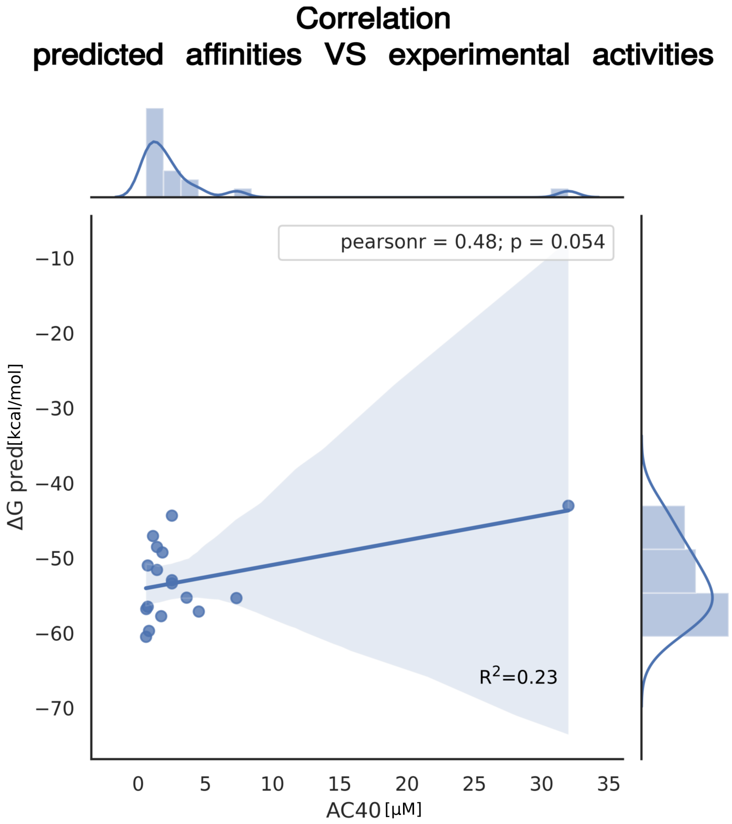


Figure 7.16: Correlation plot between predicted binding affinities (ΔG pred) and biochemical activity values (AC_{40}) for the Morgan's compounds. The data show no correlation ($R^2 = 0.23$).

7.3.4.1 *τ-RAMD*

Afterward, we applied an orthogonal approach: the estimation of the residence time of Morgan's compounds (k_{off}). To compute the residence time of the molecules in our dataset, we used the method developed by the group of Wade (Kokh et al., 2020) named τ-Random Acceleration Molecular Dynamics (τ-RAMD) simulations. As described in Section 4.1, the procedure allows the computation of the drug residence time in the binding site, by exploring ligand dissociation mechanisms (Kokh et al., 2020). As output, we get the values of t_R that equals the reciprocal of the dissociation rate as in Equation 7.3.4.1:

$$t_R = \frac{1}{k_{off}} \tag{7.1}$$

t_R is a predictor of the efficacy of a drug, viz. the maximum response achievable from a small molecule (Holford and Sheiner, 1982; Pan et al., 2013) and it has been demonstrated that it highly correlates with functional efficacy (Pan et al., 2013). For the compounds under analysis, we defined a constant pulling Force of $4972 kJ \cdot mol^{-}1 \cdot nm^{-}1$

and a center of mass (COM) displacement threshold of 7 nm. This means that, when the COM of the ligand reaches a distance of 7 nm relative to the COM of the protein, it is considered dissociated. In Figure 7.17 an example of dissociation of one Morgan's compound during a τ-RAMD simulation.

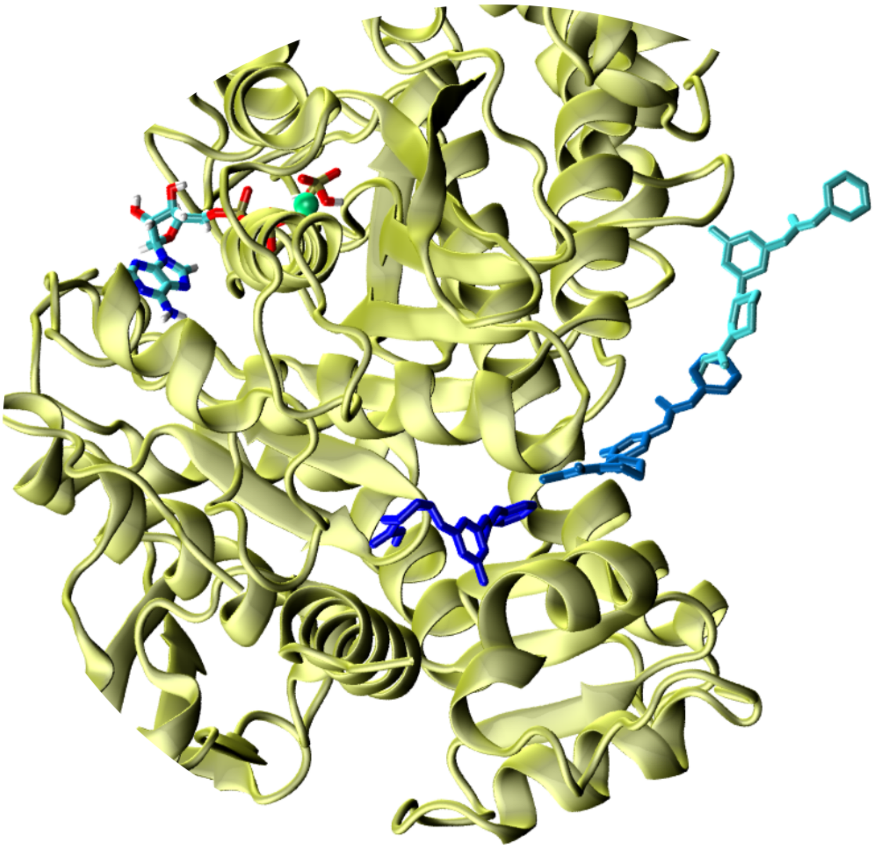


Figure 7.17: Example of unbinding of one Morgan's compound (A11) during a trajectory of τ-RAMD. The ligand is colored with a gradation of blue ranging from deeper blue (bound) to cyan (dissociated), depending on the bond status.

We computed the correlation between predicted residence time and predicted binding affinity, previously calculated with MMGBSA. Pearson's coefficient was 0.58 (see Figure 7.18).

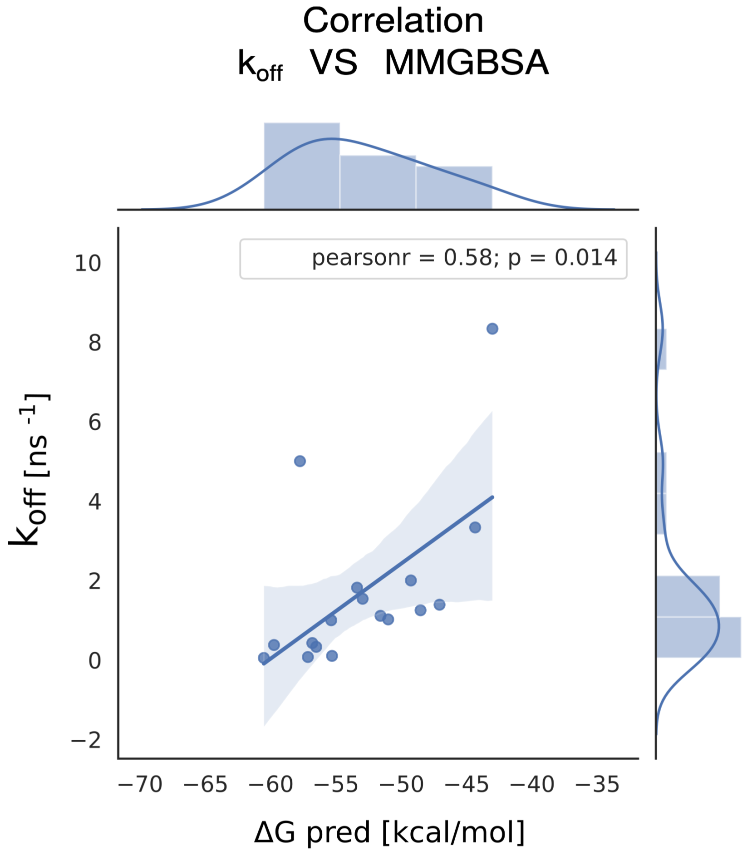


Figure 7.18: Correlation plot between predicted binding affinities (ΔG pred) and predicted residence time (k_{off}) for the Morgan's compounds. The data show an interesting correlation ($R^2 = 0.33$).

To improve the protocol, we computed the free energy of Morgan's compounds with MMPBSA (see Table 7.2). We then recomputed the correlation between predicted k_{off} and predicted ΔG, as well as the correlation between predicted k_{off} and AC_{40}.

LIGAND	$t_R [ns]$	std	$k_{off} [ns^{-1}]$	ΔG $[kcal/mol]$
A1	0.8	0.56	1.2	-17.06
A2	0.3	0.07	3.3	-15.7
A3	0.1	0.04	8.3	-12.76
A4	0.5	0.14	2	-19.2
A5	0.98	0.32	1	-18.9
A6	0.9	0.22	1.1	-19.8
A7	0.72	0.29	1.38	-16.23
A8	0.65	0.39	1.53	-17.2
A9	1	0.35	1	-17.15
A10	0.55	0.18	1.82	-14.2
A11	9.8	6.6	0.1	-22.27
A12	2.35	1.61	0.42	-19.7
A13	3.02	1.54	0.33	-19.56
A14	2.65	0.72	0.38	-17.92
A22	0.2	0.07	5	-11.21
A23	13.38	7.44	0.07	-24.87
A24	19.8	0.84	0.05	-22.46

Table 7.2: Residence time (t_R), respectively k_{off} and binding affinities (ΔG) predicted for the Morgan's compounds. The values of t_R are the average ones of fifteen trajectories of unbinding. In the second column, std stands for the standard deviation of the average value of t_R. ΔG have been computed with MMPBSA using internal ϵ equal to 4 and default values for the other parameters (Case et al., 2018). Highlighted in red are the values of A3 and A24 (OM), used as boundaries for the calculation, being respectively the worst and the best compounds of the series, in terms of biochemical activities (Morgan et al., 2010). The calculations were run on the HPC supercomputers Mesocentre (`https://hpc.pages.unistra.fr/`) and ROMEO Super Computer Center (`https://romeo.univ-reims.fr/pages/aboutUs`).

7.3.4.2 *Results: in silico predictions VS experimental data*

The results (see Figure 7.19) show an interesting correlation between predicted affinities and residence time.

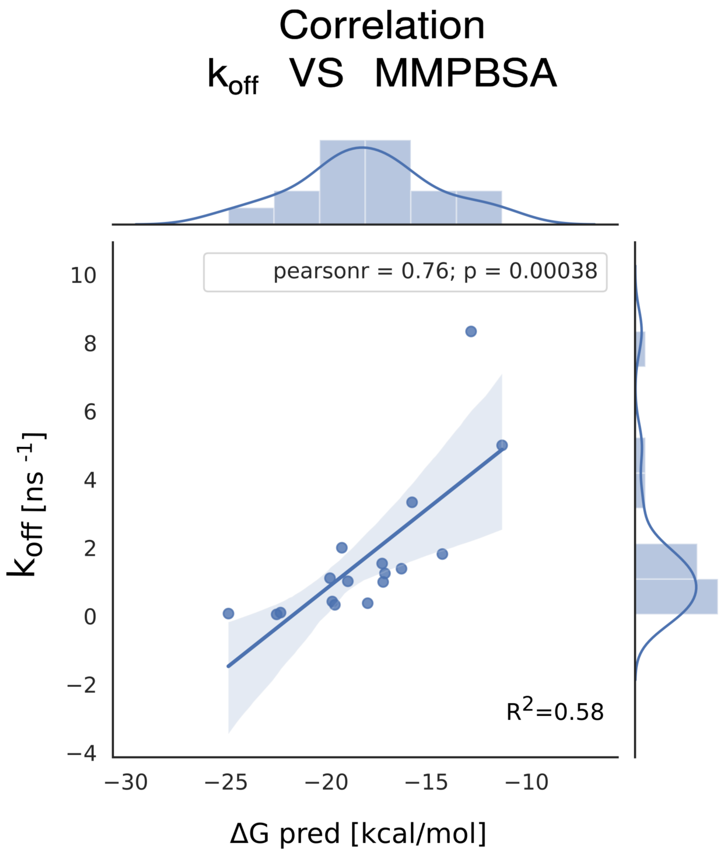


Figure 7.19: Correlation plot between predicted binding affinities (ΔG pred) and predicted residence time (k_{off}) for the Morgan's compounds. The data show an interesting correlation ($R^2 = 0.58$).

In Figure 7.20 it seems that it also exists a correlation between predicted residence times and biochemical activities (Pearson's coefficient = 0.78). Actually, this high correlation is affected by A3, the Morgan's compound with the highest AC_{40} value ($AC_{40} = 32\mu$M). Once removed from the dataset, the correlation drops down to 0, as shown in Figure 7.21.

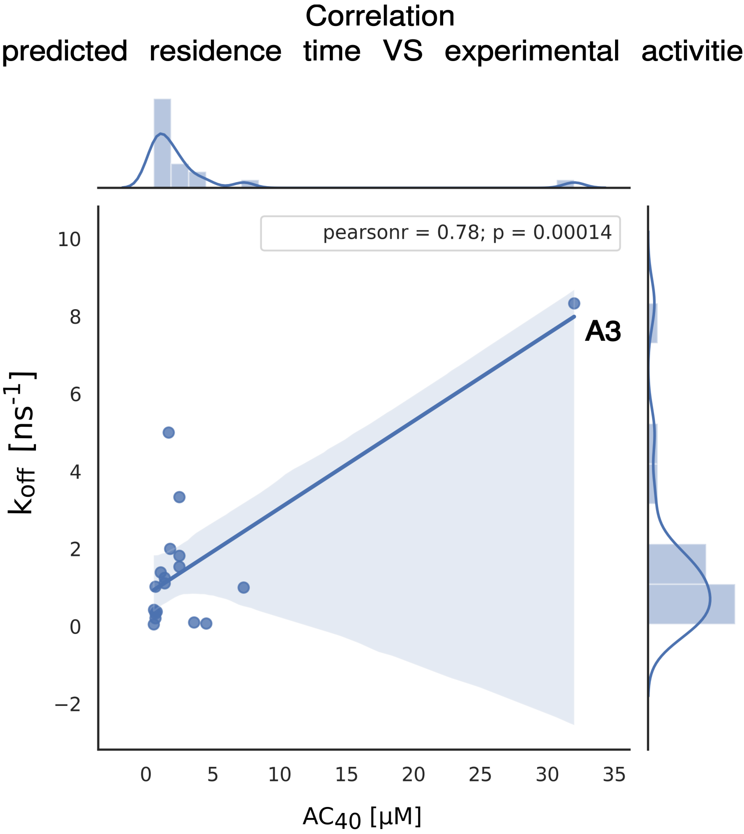


Figure 7.20: Correlation plot between predicted residence time (k_{off}) and biochemical activities (AC_{40}) for the Morgan's compounds. The data share a Pearson's coefficient of 0.78.

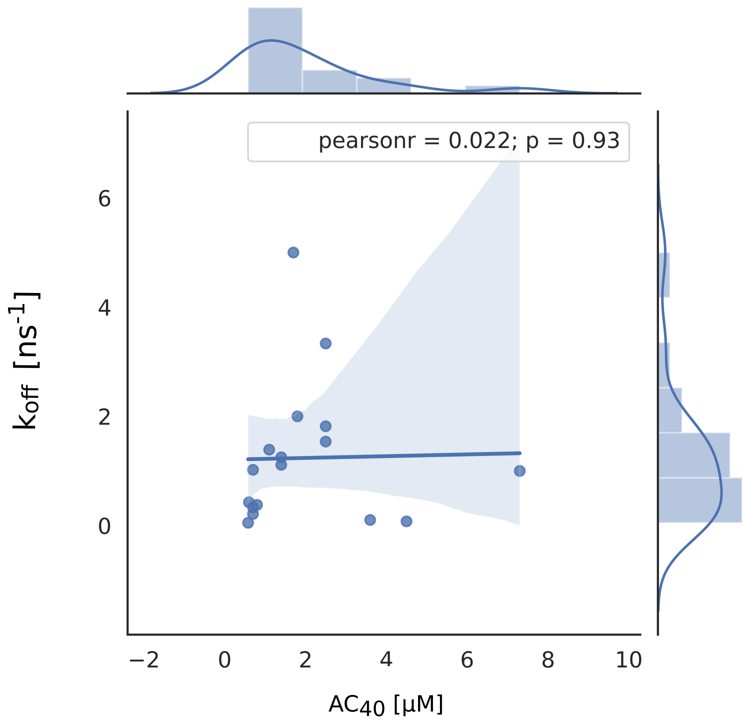


Figure 7.21: Correlation plot between predicted residence time (k_{off}) and biochemical activities (AC_{40}) for the Morgan's compounds, excluding A3. The data share a Pearson's coefficient of 0.

7.4 CONCLUSIONS & FUTURE PERSPECTIVES

What makes a small molecule a cardiac myosin activator? The question remains unanswered, for now. Our (computationally expensive) analyses demonstrated that both predicted affinities and residence time calculations do not quantitatively correlate with the experimental activity data. Nevertheless, our results are in qualitative agreement, i.e activators vs inhibitors, with available data. This finding could be used in a virtual screening campaign to prioritize potent activators of cardiac myosin in PPS. A possible approach that could raise the quantitative outcome of our investigation is a rigorous method, such as Free Energy Perturbation (FEP) calculations for OM and its analogs. This would allow confirmation or disproof of the possible correlation between activity and affinity in PPS. The FEP calculation is, anyway, very challenging for the big and flexible system as cardiac myosin, for which our preliminary absolute binding free energy (ABFE) for OM, has shown serious convergence issues. Alternatively, this problem could be overcome by computing relative binding free energy (RBFE) between the most structurally similar compounds in the series of Morgan's molecules. A computationally cheaper solution could be to compute $\Delta\Delta G$ with MM/PB(GB)SA for the Morgan's compounds between PPS and another state, in equilibrium with it, such as near rigor. In this case, we would predict the efficacy of the molecules, rather than their potency (Cecchini and Changeux, 2022) and then, to fix the missing term for the correlation with experimental activity values, we could fit the data with a linear regression method, as recently proposed by Fornili and co-workers (Hashem et al., 2020). We could combine both approaches by using linear regression based on MM/PB(GB)SA scores computed in both states (PR and PPS). However, the two orthogonal methods we explored correlate with each other, underlying that the calculations are both individually meaningful. The lack of correlation of both methods with experimental activities may suggest that data computed in a single state (here PPS) are not enough to be compared with experimental activities. Similar hints have been by Fornili and co-workers (Hashem et al., 2020), therefore future combination of both approaches might be successful at improving the correlation between our predictions and experimental data.

8

MYOSIN VI: STRUCTURAL FEATURES & DYNAMICS

Myosin VI (MyoVI) was discovered in 1992 by Kellerman and Miller after an analysis of complementary DNA (cDNA) clones encoding for cytoskeletal proteins involved in *Drosophila melanogaster*'s embryonic development (Kellerman and Miller, 1992). Since then, it was shown that it is involved in numerous life-sustaining functions (Buss et al., 2004) and it is expressed in distinct species, ranging from worm (i.e. *C.elegans*) (Kelleher et al., 2000) to mammals (Avraham et al., 1995; Hasson and Mooseker, 1994).

8.1 STRUCTURAL FEATURES & BACKWARD MOTION OF MYOSIN VI

The peculiarity of MyoVI is that it moves toward the *pointed* (minus) end of actin filament (Pylypenko et al., 2011), unlike all the other myosins (Schliwa, 1999; Wells et al., 1999). This peculiar behavior possibly accounts for a structural adaptation for the different roles of MyoVI in the cell, such as transcription in the nucleus (Vreugde et al., 2006), organization of actin filaments (Woolner and Bement, 2009b), intracellular transport (Buss et al., 2004) and mechanical function of stereocilia in the inner ear (Avraham et al., 1995). A hypothesis raised to explain the reverse directionality of MyoVI from a structural point of view, proposes the role of a unique 53-aminoacids insert (insert 2) at the C-terminus of the converter subdomain (Ménétrey et al., 2005) (Ovchinnikov et al., 2011) as responsible of this feature (Park et al., 2006). Homma *et al* in 2001 created chimeras of the motor domain of MyoV combined with the converter and lever arm of MyoVI and vice-versa chimeric MyoVI whose motor domain was coupled to the lever arm of MyoV (Homma et al., 2001). These chimeras were then tested with *in vitro* motility assay with dual fluorescence-labeled F-actin(Homma et al., 2001). Based on the results of this experiment, the core part of the motor domain is modulating the reverse directionality of MyoVI and possibly other myosins could share this feature (Homma et al., 2001). In 2006 Sweeney and co-workers recreated some chimeras of MyoVI by replacing the lever arm with the one of Myosin V either with or without the insert 2 (Park et al., 2006). Sweeney *et al* convincingly argued that a chimeric myosin V with a myosin VI converter would lack crucial interactions between insert 2 and lever arm, resulting in structural instability (Park et al., 2006). This incompatibility in the structure might be the cause of the negative results concerning the implication of insert 2 in minus-end movement obtained by Homma *et al* (Park et al., 2006). The results by Sweeney and co-workers demonstrated that insert 2 is, instead, an optimized element fundamental for reverse directionality specific for MyoVI and not for any other class of Myosins (Park et al., 2006). MyoVI exists as a monomer in cells, able to dimerize to accomplish its functions (Mukherjea et al., 2014). The dimerization process is not merely electrostatic as proposed in 2010 by Kim *et al* (Kim et al., 2010), but it is regulated by cargo adaptor proteins as postulated in 2009 by Phichith (Phichith et al., 2009) and demonstrated more recently by other groups (Hu et al., 2019; Mukherjea et al., 2014; Shang et al., 2017). The step-size over the actin filament changes accordingly to the monomeric or dimeric form and in the latter it shows high variability (Mukherjea et al., 2014; Park et al., 2006; Rock et al., 2000). The monomeric form of MyoVI is thought to work in a non-processive way with a step-size of $18nm$ (Lister et al., 2004). By contrast, the dimer walk processively on actin with a hand-over-hand mechanism (Park et al., 2006) and a step size of $30 - 36 \pm 12nm$ (Buss et al., 2004; Mukherjea et al., 2014; Sun et al., 2007). In both cases, the lever arm attains a rotation of 180^o (Bryant et al., 2007; Reifenberger et al., 2009), This large step-size of dimers of MyoVI is unexpected, considering the short lever arm in which resides only one IQ motif (conventional calmodulin binding site) (Ménétrey et al., 2005) (Park et al., 2006), compared to the lever arm of another processive myosin, Myosin V, that comprises six IQ motifs (Rief et al., 2000). Despite this major difference, both dimers of MyoVI and Myosin V exhibit similar $36nm$ step-sizes (Kodera et al., 2010; Park et al., 2006; Warshaw, 2004). Various hypotheses have been raised to explain this large step-size of MyoVI: one asserts that it could be explained by the unique conformation of the converter in the PPS state by Menetréy *et al* in 2007 (Ménétrey et al., 2007) and already proposed by Sweeney, Park *et al* in 2006 (Park et al., 2006); a different one proposes the pliancy of the region between converter and lever arm as key factor (Sun et al., 2007) and yet another one suggests a possible uncoupling of the lever arm of the lead head as responsible of the large step-size (Ménétrey et al., 2012). The step size in MyoVI is not only surprisingly large, but it also shows high variability. Rock and co-workers discussed how MyoVI dimers can have a step size that varies from 30 nm with a difference of ± 12 nm (Rock et al., 2001). It was further demonstrated by Park and co-workers with chimeric myosins that the variable step size does not depend on the presence of insert 2 (Park et al., 2007). Later, it was proposed by Ménétrey *et al* that the converter of the roar-head may undergo a transition while bound to actin, switching from PPS to R (Ménétrey et al., 2007). In 2011 Ovchinnikov *et al* complemented these studies with string calculations (Maragliano et al., 2006; Ovchinnikov et al., 2011). By computing the minimum free energy pathway (MFEP), they demonstrated that the isomerization of the converter from the canonical folding, known as

R-fold, to a conformation named P-fold characteristic of PPS and P_i release state (Llinas et al., 2015), is responsible for the variable step-size in MyoVI (Ovchinnikov et al., 2011). The monomeric form of MyoVI is thought to work in a non-processive way with a step-size of $18nm$ (Lister et al., 2004). Underneath, in Figure 8.1, we show a blueprint of the backward motion of a monomeric head of MyoVI along a filament of F-actin, with a description of the steps of the force cycle.

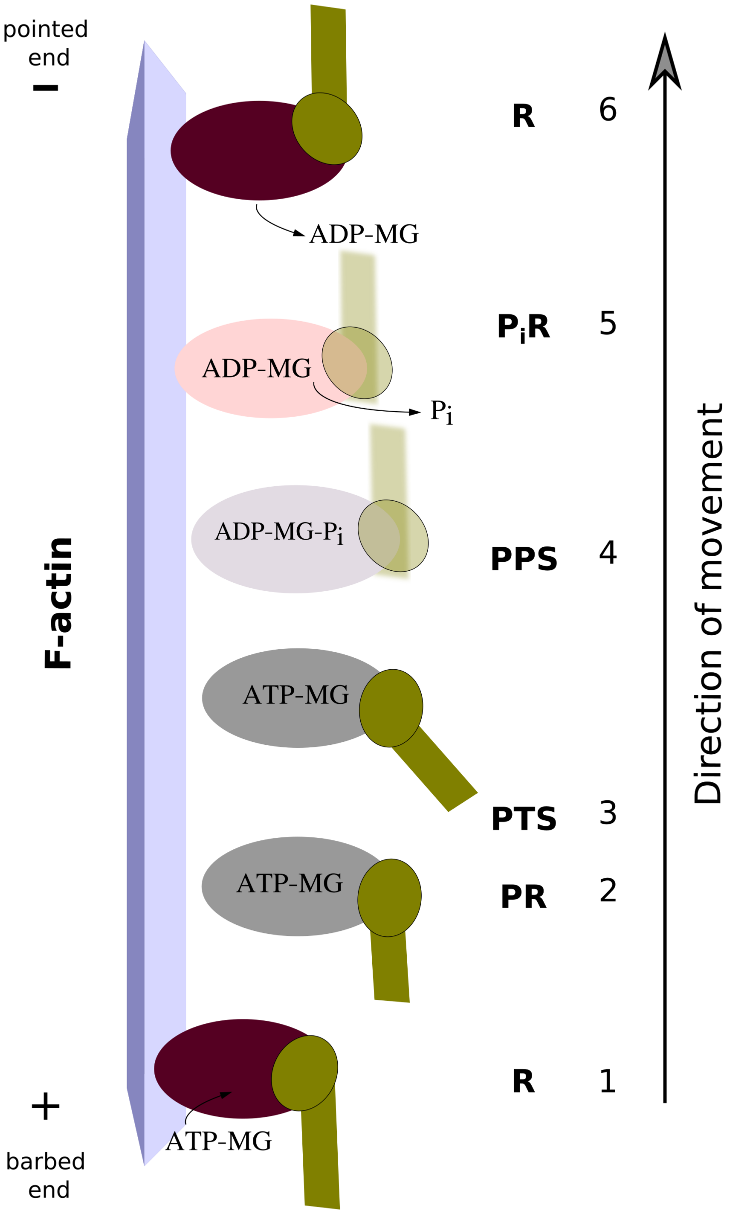


Figure 8.1: Schematic representation of the movement of one head of MyosinVI along an actin filament. The converter-lever arm domain is depicted in green. The movement can be described in six main steps:
1. binding of a molecule of ATP in the nucleotide-free active site of a myosin head, bound in a high-affinity state for actin, named Rigor state (R);
2. detachment of myosin head from F-actin in the ATP-bound state known as Post-Rigor. The affinity for actin decreases at this stage;
3. intermediate state known as the pre-transition state (PTS) (Blanc et al., 2018) with ATP-bound in the active site and a different orientation of the converter-lever arm;
4. conformational change at the level of the lever arm domain accompanied by the hydrolysis of ATP. In this state (Pre-Power stroke state (PPS)) the motor domain traps the hydrolysis products in the active site;
5. following conformational change that drives the release of the inorganic phosphate from the active site (P_iR state) (Llinas et al., 2015). The affinity for actin is higher;
6. strong binding of myosin to actin that triggers the powerstroke, a key step for force generation. Subsequent conformational changes in the head trigger the release of ADP from the active site.
The motor domain is back to the R state, with myosin temporarily bound to actin in a different position than the one occupied at the beginning, ready to re-bind ATP and close the cycle. The color of the motor domain varies from dark magenta (highest affinity for actin) to grey (high-affinity state for ATP e ADP-P_i) to light pink (P_iR-state) to dark magenta again according to the change in affinity respectively for actin and ATP during the cycle.

One more feature that makes MyoVI unique is the presence of another insert, called insert 1, positioned in the U50 subdomain near the nucleotide-binding site (Pylypenko et al., 2011) (Ménétrey et al., 2005) (see Figure 8.2). This 26-aminoacids insert regulates the rate of ATP binding (Sweeney et al., 2007) (Pylypenko et al., 2011), producing strain-dependent slowing of ATP-binding (Sweeney and Houdusse, 2010). The nucleotide-binding kinetics is the gating mechanism for the activity of the two myosin's heads in the dimeric form of MyoVI (Sweeney et al., 2007), i.e. the communication between the myosin's heads to achieve coordinated hand-over-hand motion along actin filament (Yildiz et al., 2021). The kinetics of ATP binding is controlled by the protrusion of a loop interposed in insert 1 (depicted in cyan in Figure 8.2) that interacts with swI (Sweeney et al., 2007). In particular, a Leucine of this loop (L310) could impact ATP-induced actin dissociation, modulating the kinetic transition from strong to weak actin-binding (Pylypenko et al., 2011). This model of gating mediated by control of ATP-binding in the front head has been proposed in the early 2000s and then demonstrated in 2011 by Sweeney and co-workers (De la Cruz et al., 2001; Pylypenko et al., 2011; Sweeney et al., 2007). According to this theory, the processive movement and high duty ratio of MyoVI (meaning the time spent by the motor domain strongly bound to actin during the cycle (De la Cruz et al., 2001; O'Connell et al., 2007)) are regulated by ATP binding (Sweeney et al., 2007), rather than ADP release as in another processive myosin, such as Myosin V (O'Connell et al., 2007; Veigel et al., 2005) (Pylypenko et al., 2011). The ATP binding rate in MyoVI is 50 to 100 times slower than for other myosins (De la Cruz et al., 2001). A different model proposed for the gating mechanism in MyoVI posits ADP release from the front head as the rate-limiting step in the MyoVI cycle when both heads are bound to actin (Dunn et al., 2010; Oguchi et al., 2008). ADP release is accelerated under forward loads and slowed down by backward loads (De la Cruz et al., 2001). In both models (rate limiting step found in ATP binding or ADP dissociation) the motor domain remains strongly bound to F-actin because it populates either the nucleotide-free or ADP state which shows high affinity for the track. Furthermore, the gating mechanism is likely to improve the efficiency of the motor, avoiding wasting ATP (Elting et al., 2011) and ensuring processivity (Sweeney et al., 2007).

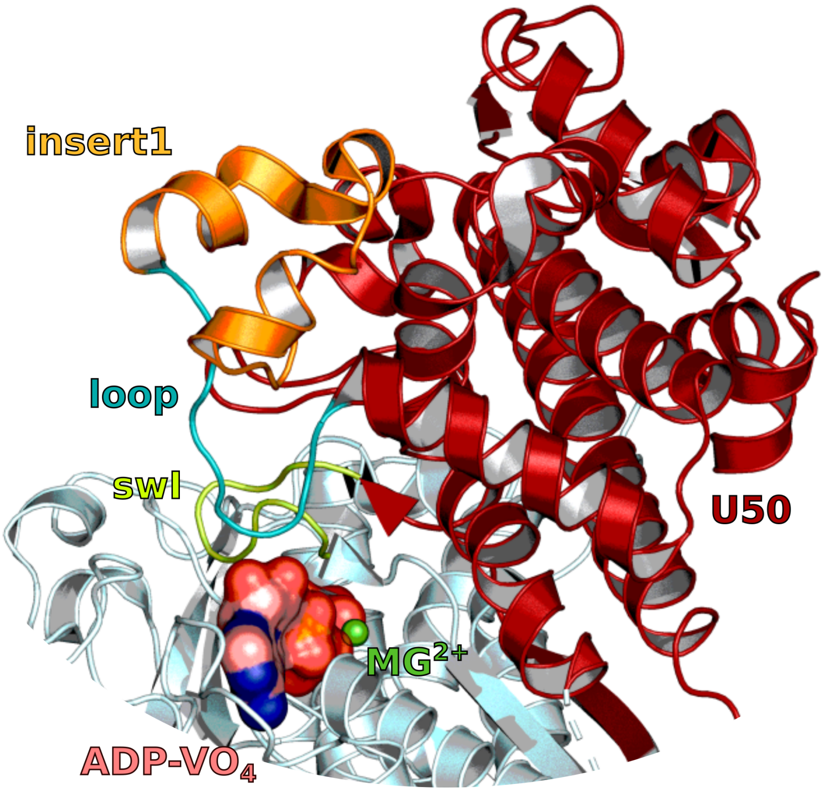


Figure 8.2: Cartoon portrayal of 2V26.pdb (Ménétrey et al., 2007), zooming on the insert1 (residues C278 to A303). In red U50 subdomain, in orange insert 1, in cyan the loop interposed in insert 1 (residues G304-D313) and lime swI. In the pink surface ADP-VO_4^{3-} and in the green sphere the magnesium ion in the active site.

8.1.1 A deep look into power- and recovery-stroke of MyoVI

In all myosins, the mechanical motion of the converter coupled to the forward swing of the lever arm, when bound to actin, is known as Powerstroke (Holmes and Geeves, 2000; Ménétrey et al., 2008). The powerstroke starts when the myosin head is in a pre-powerstroke (PPS) state, with the ATP-hydrolysis products trapped in the active site, able to bind to F-actin (Lee Sweeney and Houdusse, 2004). The binding to the track is linked to the release of P_i from the active site, a process catalyze by actin (Llinas et al., 2015). The following release of MG^{2+} and ADP from the active site is coupled to a progressive change in binding affinity for the track (Llinas et al., 2015; Ménétrey et al., 2008; Swanson et al., 2004). Actin drives the sequential release of ATP-hydrolysis products from myosin's head (Llinas et al., 2015). The transition to the nucleotide-free state strongly bound to actin is coupled to conformational changes of active site elements (switch loops) in the motor domain (Llinas et al., 2015). These changes allow the escape of P_i from the active site and subsequent movement of the lever arm (Llinas et al., 2015). The power stroke ends when myosin is in this nucleotide-free state, the so-called rigor state (R), tightly bound to the thin filament (Lee Sweeney and Houdusse, 2004; Ménétrey et al., 2008). At this stage, ATP can rebind to myosin that detaches from the actin filament and populates the post-rigor (PR) state (Lee Sweeney and Houdusse, 2004). Major rearrangements during this transition involve the distortion of the β-strands of the transducer (Coureux et al., 2004) that are twisted in R state and untwisted in PR and PPS state (Cecchini et al., 2008). The distortion of the transducer promotes rearrangements of P-loop and swI that adapt for the binding of ATP (Ménétrey et al., 2008). These changes induce the opening of the actin-binding cleft and then the detachment from actin filament (Ménétrey et al., 2008). Minor rearrangements affect, instead, the lever arm during the ATP-driven unbinding from actin, to optimize the power stroke (Ménétrey et al., 2008). Some structural elements of the motor domain take part in this sequence of transition from the state at a high affinity for actin and low affinity for ATP to a state at a high affinity for ATP and low affinity for actin (Ménétrey et al., 2008). In particular, actin weakens the affinity for the nucleotide by promoting the closure of the actin-binding cleft, between U50 and L50, while the loops in the active site (especially swI and P-loop) mediate the release of the ATP-hydrolysis products (Coureux et al., 2004). The conformational transitions characterizing the power stroke (viz the transition from PPS to R state) envisages the release of P_i and thereupon of ADP, coupled to the movement of the lever arm (Lee Sweeney and Houdusse, 2004). In dimeric MyoVI the unbinding of the hydrolysis products from the active site is inhibited in the lead head until the rear head binds to actin, becoming the new lead head (Ovchinnikov et al., 2011). This ensures the hand-over-hand motion of the motor (Park et al., 2006). Overall the transition from PPS to R state is defined as Powerstroke and the second part of the cycle, from R to PPS is named recovery stroke (Ménétrey et al., 2008). The understanding of these processes in MyoVI presented additional levels of complexity from trying to infer a description of the cycle that would be compatible with the reverse motion along the actin filament (Ménétrey et al., 2008). In 2008 Ménétrey and co-workers demonstrated that the converter maintains a rigor-like conformation also in the PR state, preserving the interactions with the N-terminal subdomain. (Ménétrey et al., 2008). Moreover, the unique insert2, responsible for the reverse directionality in MyoVI (Park et al., 2007), would play a role in placing the lever arm only 8 Å apart with respect to its position in the R state, limiting its movement during the ATP-induced unbinding from actin (Ménétrey et al., 2008). Further elucidation on the recovery stroke in MyoVI has recently been provided by the group of Houdusse, in collaboration with our (Blanc et al., 2018). In 2018, indeed, Houdusse *et al* solved a structure representing a putative intermediate state in the recovery stroke of MyoVI, linking PR and PPS, called pre-transition state (PTS) (Blanc et al., 2018). The structure is solved in complex with an analog of ATP, namely ADP $\cdot$ $B_E F_x$ (Blanc et al., 2018). Blanc and Cecchini, then, performed unbiased MD simulations and adaptive biasing force (ABF) free energy calculations to better characterize this structure (Blanc et al., 2018). Their results show two potential mechanisms of the recovery stroke in MyoVI, which both culminate with the late closure of swII (Blanc et al., 2018). PTS is also characterized by the remarkable flexibility of the converter, which explores an ensemble of configurations between the two ultimate states (Blanc et al., 2018). This behavior is driven by thermal fluctuations and it is in agreement with the intermediate nature of PTS since the converter could fluctuate between different positions before finding the one compatible with the post-transition state (PPS) (Blanc et al., 2018). Unveiling the sequence of conformational changes in the recovery stroke would allow us to understand the chemomechanical transduction in myosins since the recovery stroke occurs at the end of the force generation process. Given the importance of understanding the critical features of the chemomechanical transduction in myosin and the opportunity provided by the solved structure of this newly characterized stage of the cycle, we decided to investigate the recovery stroke in MyoVI.

8.2 SPONTANEOUS TRANSITION FROM PPS TO ANOTHER INTERMEDIATE STATE

Our investigation began with modeling and performing an all-atom MD simulation of the PPS state of MyoVI. PPS is the transition state of the force-producing cycle: end state of the recovery stroke and initial step of the powerstroke (Blanc et al., 2018; Coureux et al., 2004; Ménétrey et al., 2008). We aimed at characterizing the conformational stability and dynamics of MyoVI in this state by defining a series of geometrical observables (collective variables) characteristic of the motor domain in this stage of the force generation cycle.

8.2.0.1 *System preparation*

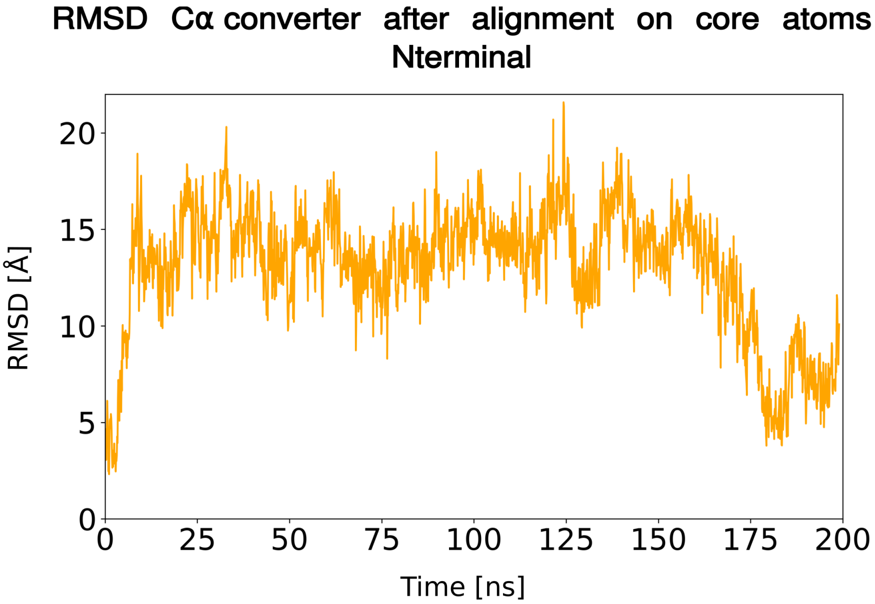
The motor domain of MyoVI in its various conformational states (R-PR-PTS-PPS-P_iR) has been crystallized in absence of actin, producing X-ray structures at high-resolution (Blanc et al., 2018; Llinas et al., 2015; Ménétrey et al., 2005, 2008, 2007). For the PPS state, we used the structure corresponding to the Protein Data Bank (PDB) entry 2V26 (Ménétrey et al., 2007) This is a structure of the S1-fragment of MyoVI solved in complex with $ADP.VO_4^{3-}$, where VO_4^{3-} stands for Vanadate, analogue of inorganic phosphate (Gresser and Tracey, 1990). We replaced the VO_4^{3-} in the active site with inorganic Phosphate in its form $H_2PO_4^-$. The missing portions in the PBD structure (residues: 1 to 4; 34 to 37; 174 to 179; 396 to 406; 622 to 637) were modeled using MODELLER (Webb and Sali, 2016) and the best model was chosen according to the lowest Discrete Optimized Protein Energy (DOPE) score. The structure was then uploaded on the MolProbity webserver (`http://molprobity.manchester.ac.uk/`) to add missing hydrogens, optimize H-bond networks and detect the more reasonable rotameric states of the side chains of Asn, Gln and His. The topology and parameter files for the ligands (ADP, $H_2PO_4^-$ and MG^{2+}) were obtained from CHARMM General Force Field (CGenff) (`https://cgenff.umaryland.edu/`) (Vanommeslaeghe et al., 2012) and the entire system was modeled with CHARMM (Chemistry at HARvard Molecular Mechanics) (Brooks et al., 2009). The most probable protonation state of titratable residues at neutral pH was predicted by Karlsberg (Rabenstein, 2000). The Solution Builder module of CHARMM-GUI (Jo et al., 2008) was used to solvate and ionize the system in 11 Å edge water box (final box size was 13 nm per side) in which 205 Sodium ions and 202 Chloride ions were added to neutralize.

8.2.0.2 *Molecular Dynamics (MD) simulations and Visualization tool*

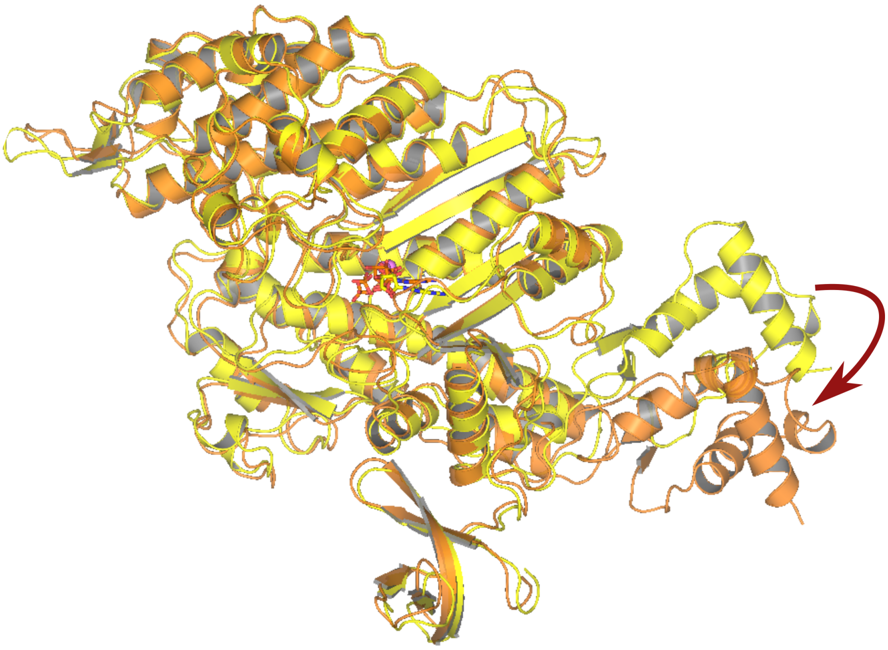
Once the input files for GROMACS for the MD simulation were generated from CHARMM-GUI, a steepest-descent minimization until reaching a $F < 1000.0 kJ/mol/nm$, by using GROMACS 2020.3 (Abraham et al., 2015; Lindahl et al., 2001; Markidis and Laure, 2015) (`https://doi.org/10.5281/zenodo.3923644`). The minimized system was then equilibrated in NVT for $125 ps$ and NPT for $100 ps$ before the production of $200 ns$ run with the leap-frog algorithm. No restraints were applied to the structure. The MD simulation was performed on the GPU partition of the HPC of the University of Strasbourg Mesocentre (`https://hpc.pages.unistra.fr/`). The reference structures chosen for the different states of MyoVI have been selected from the PDB database: 2BKH for the Rigor (R) state (Ménétrey et al., 2005); 4PFO for the P_i release (P_iR) state (Llinas et al., 2015); 2VAS for the post rigor (PR) state (Ménétrey et al., 2008); 5O2L for the Pre-Transition State (PTS) (Blanc et al., 2018) and 2V26 for the PPS state (Ménétrey et al., 2007). To visualize and analyze the trajectory we used VMD with some in-house *tcl* scripts.

8.2.1 Structural & dynamic characterization of spontaneous transition to a PTS-like state

In this section, we will deepen the analysis of the MD simulation of our model of MyoVI in PPS. The first step of our analysis consisted of a visual inspection of the trajectory that revealed a remarkable displacement of the converter for most of the simulation (see Figure 8.3).



(a) RMSD of the *Cα* atoms of the converter in 200ns MD.



(b) Cartoon representation of the X-ray structure of *Sus-scrofa* myosin VI in PPS (PDB: 2v26.pdb) (in yellow) and of a frame of the trajectory in which the converter deviates from its initial coordinates (in orange).

Figure 8.3: In sub-figure (a) the time-series over 200 ns of the RMSD of the *Cα* of the converter after alignment over the core atoms of the N-terminal subdomain of the initial frame. The initial structure has an RMSD of 0.1 Å with respect to the X-ray coordinates of the PPS state (2v26.pdb). Sub-figure (b) is a cartoon representation of the movement of the converter in the MD simulation, as compared to 2v26.pdb (in yellow).

The swing of the converter toward a lower position, halfway between PR and PPS (see Figure 8.4), suggests that the structure could explore the previous state of the actomyosin cycle: the PTS state (Blanc et al., 2018).

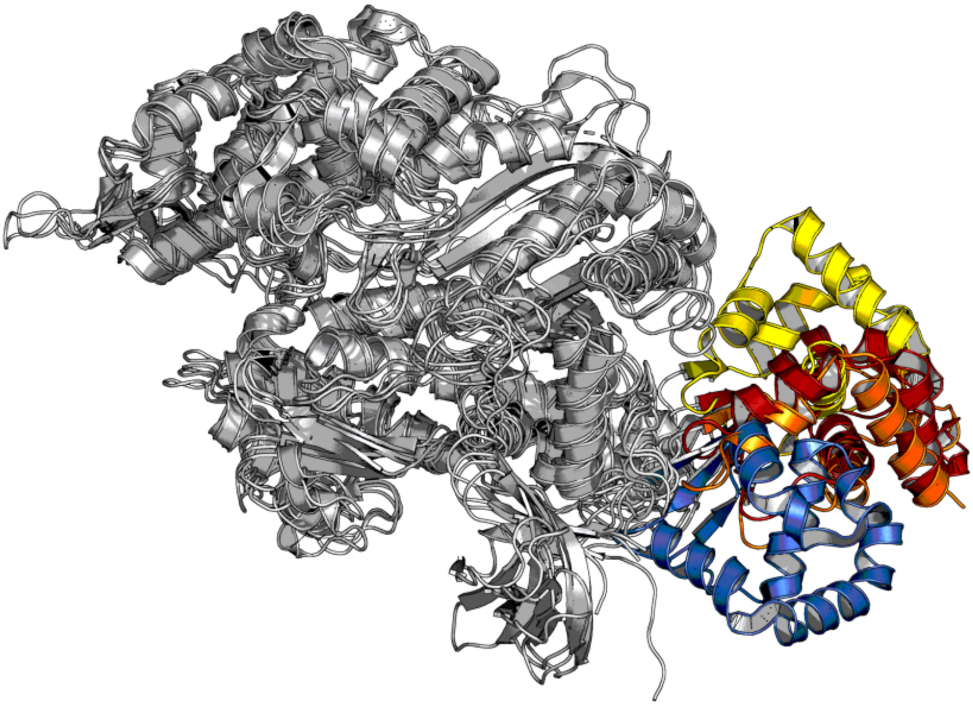


Figure 8.4: Cartoon representation of the position of the converter in three X-ray structures: PPS (yellow); PTS (red) and PR (blue). In orange is a frame of the trajectory in which the converter deviates from its initial coordinates.

8.2.1.1 *Monitoring of geometrical observables: active site*

The PTS (PDB code: 5O2L.pdb) has been characterized as an intermediate structure of the recovery stroke between the post-rigor (PR) and PPS state, exhibiting peculiar geometrical differences if compared to PPS and PR, respectively (Blanc et al., 2018). We, therefore, decided to describe the conformational dynamics of the PPS state of MyoVI by looking at some collective variables descriptive of this specific stage of the cycle. At first, we monitored the stability of the critical salt bridge between swII (E461CD) and swI (R205CZ), which is pivotal for the hydrolysis of ATP in PPS (Onishi et al., 2006) and is solely formed in PPS state (Blanc et al., 2018; Onishi et al., 2006). From our data, it seems that the critical salt bridge between E461CD-R205CZ breaks at the beginning of the MD. This loss can make the structure unstable and even if the aforementioned salt bridge tends to reform later in the trajectory (see time-series in Figure 8.5), the stability of the PPS state is compromised, as inferred from further beneath analyses in the text.

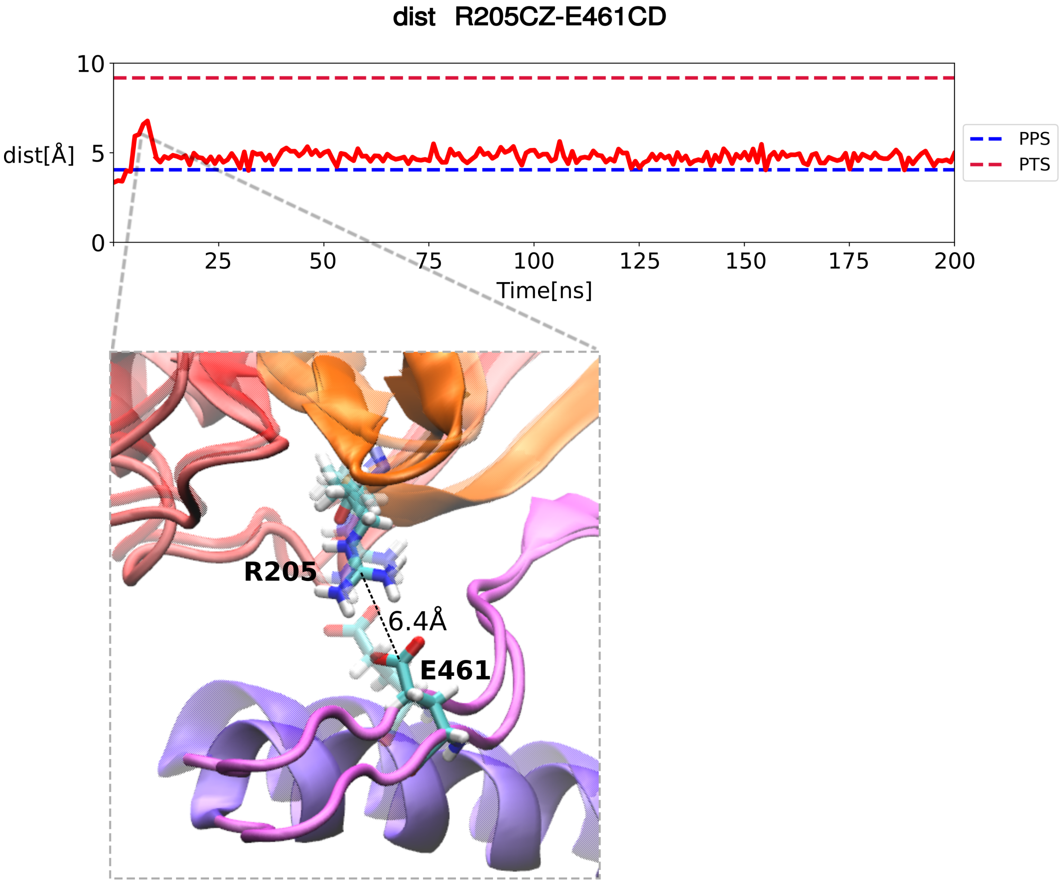


Figure 8.5: Timeseries along the 200 ns MD of the primary salt bridge measured as the distance between R205CZ and E461CD (top plot). In dashed lines the reference values of the same distance in X-ray structures of PTS (red) and PPS (blue). The reference value in PTS (5O2L.pdb) is 9.2 Å e in PPS (2V26.pdb) is 4 Å. In the bottom square a zoom of a frame of the MD in which this interaction is lost. The two residues (R205 and E461) are depicted in sticks. In magenta is the swII; in purple the RH; in orange the swI and in red U50 subdomain. In transparent colors the initial position of the different elements of the motor domain. In opaque color the same portions at the frame when the primary salt-bridge is destabilized.

Another observable that can discriminate PPS from PTS is the distance between switchI (swI) and switchII (swII), (here named as "opening"), seeing that the closing of swII is required to switch on the ATPase activity in PPS (Ménétrey et al., 2008). For this observable, we defined the distance of the center of mass of the Cα of some residues of swI (residues 197 to 200) and swII (residues 463 to 466). The instability of the system emerges also from this opening of the switch loops (Figure 8.6), probably related to the previous loss of the critical salt bridge. A movement at the level of the N-terminal portion of the Relay helix

(RH), directly connected to the swII loop (residues 468 to 483) (see Figure 8.6 bottom square), resembles a step of the seesaw motion of RH, initially proposed for Myosin II by Fisher and co-workers(Fischer et al., 2005), that is incomplete in PTS (Blanc et al., 2018).

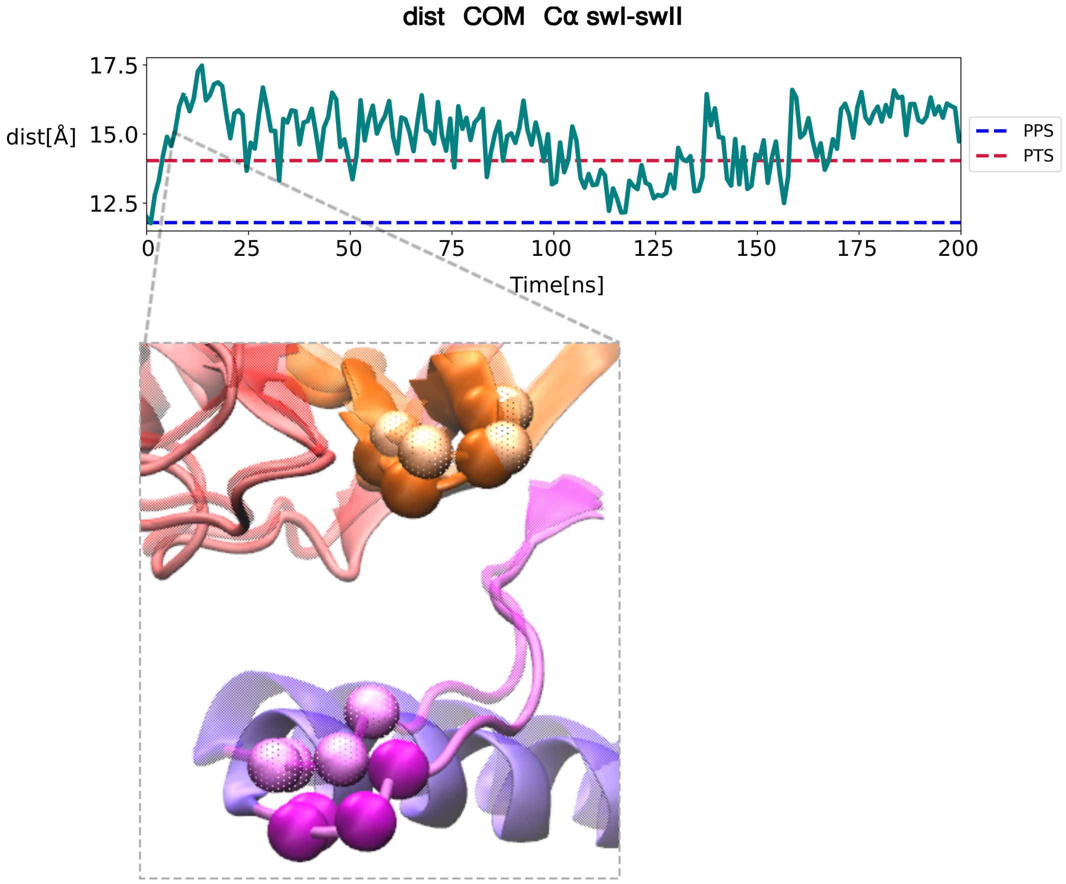


Figure 8.6: Timeseries along the 200 ns MD of the opening of swI and swII (top graph), computed as the distance of the COM of the Cα atoms of residues 197 to 200 belonging to swI (orange spheres) and residues 463 to 466 of swII (magenta spheres). In dashed lines the reference values of the same distance in X-ray structures of PTS (red) and PPS (blue). The reference value in PTS (5O2L.pdb) is 14 Å e in PPS (2V26.pdb) is 11.8 Å. In the bottom figure the cartoon representation of the frame in which the switches diverge from their initial position. The color codes of the elements are the same as in Figure 8.5.

8.2.1.2 *Monitoring of geometrical observables: β-sheets interactions*

Additional instability could be driven by the breaking of fundamental H-bonds connecting the U50 and N-terminal subdomains through three of the seven central β-sheets (transducer) (Coureux et al., 2004). In particular, β4, β5 and β6 strands are interacting through H-bonds to keep the active site in a closed state in PPS as a way of trapping the products of ATP-hydrolysis (Coureux et al., 2004). Looking at a pattern of H-bonds between β5-β4 (Figure 8.8) and β5-β6 (Figure 8.7) we observed that the destabilization and the subsequent breaking of the H-bonds between D456 and K208 (Figure 8.7) occur simultaneously with the opening of swII from swI, as shown in the two time-series in Figure 8.6 and Figure 8.7.

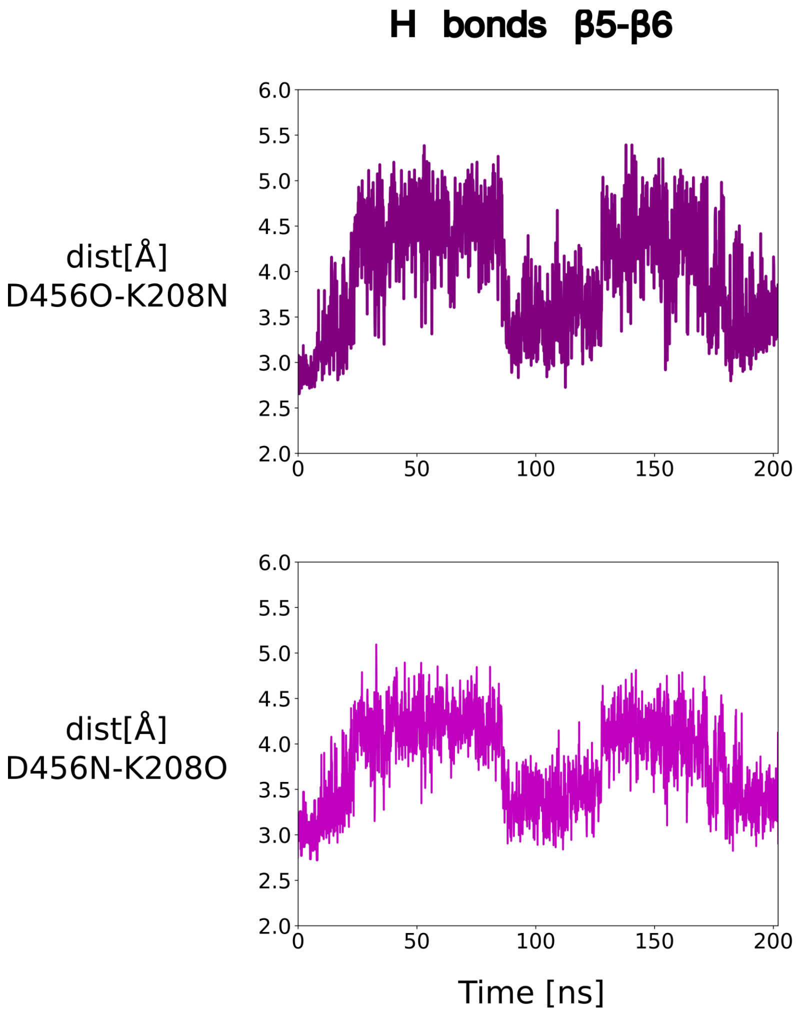


Figure 8.7: Timeseries of H-bonds pattern among $\beta5$-$\beta6$ strands of the transducer, crucial for the stability of U50 subdomain. The panel shows the evolution of the H-bond between D456 and K208, respectively residue of the $\beta5$ and $\beta6$ strand.

These bonds interest the $\beta5$ sheet (with residue D456) directly connected to swII which ends in the N-terminal portion of the RH and the $\beta6$ sheet (with residue K208), directly connected to swI. The breaking of these H-bonds could therefore go along with the opening of the two switch loops, as a further consequence of the loss of the critical salt bridge, an event that seems to trigger everything. The increased distance between swI and swII is typical of the PTS state and it never stabilizes back to a PPS-like value for the whole trajectory (see Figure 8.6). Despite the loss of these interactions, the H-bonds between the β5-β4 strands are upheld during the whole trajectory (Figure 8.8), preserving the interface between the N-terminal and U50 subdomains.

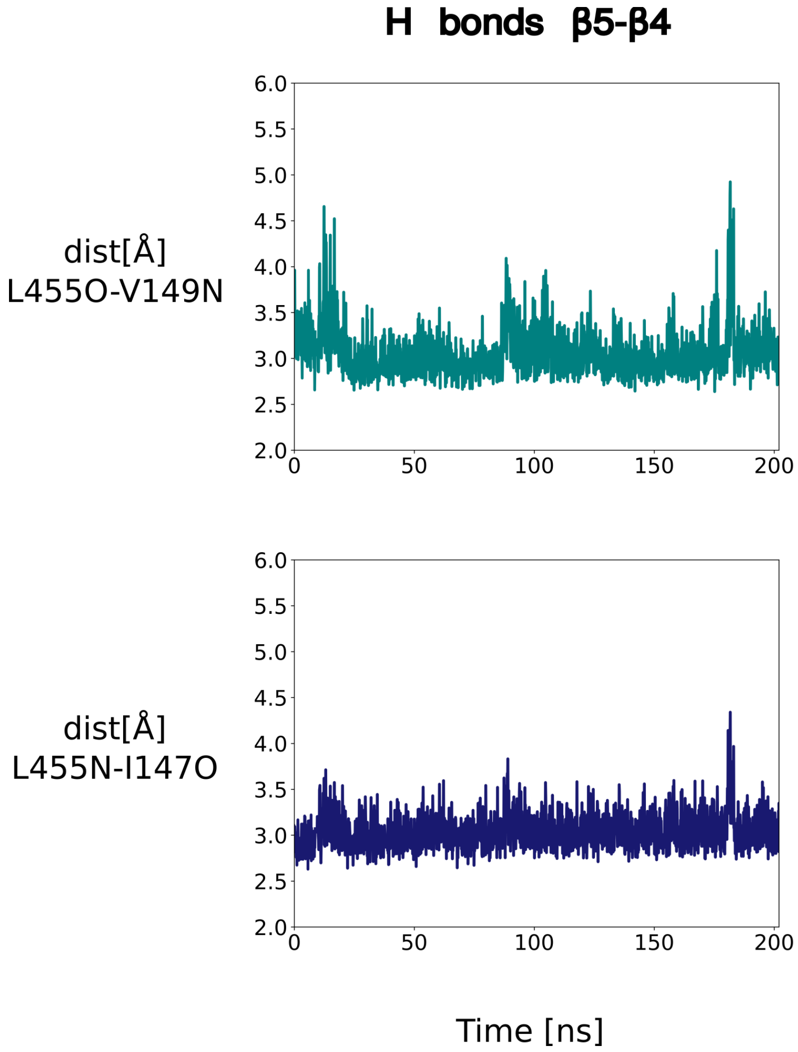


Figure 8.8: Timeseries of H-bonds pattern among β5-β4 strands of the transducer, important for the stability of U50 subdomain. The panel shows the progress of the H-bond between L455O-V149N and L455N-I147O. L455 is a residue of β5 and the other two are residues of β4 strand.

In 2018 Blanc observed structural instability of the PPS active site, when simulating the system with ATP ((Blanc, 2018), Chapter10). In his dynamics, the pattern of H-bonds between β5-β4 and β5-β6 seems to be mutually exclusive in controlling the closure of swII. If H-bonds are established between β5-β4, then swII is open, otherwise, if H-bonds are formed between β5-β6, swII is closed. Once swII is closed, the swII-ATP H-bond, needed for ATP-hydrolysis, can be formed. The instability of PPS is described by Blanc with a sequence of four main events: 1) β5-β4 bonds are not formed at the beginning of the simulation, but formed upon active site opening; 2) β5-β6 bonds break;

3) the critical salt-bridge remains formed for the whole simulation time and 4) the swII-ATP hydrogen bond breaks due to a re-opening of swII. Considering both Blanc's results and ours, it seems that the transducer could control the opening state of the active site (i.e swII), through different pathways by alternating interactions of $\beta 4$, $\beta 5$ and $\beta 6$ ((Blanc, 2018) Chapter10).

8.2.1.3 *Monitoring of geometrical observables: coordination* $H_2PO_4^- \cdot MG^{2+}$

The previous rearrangements at the level of the active site (Section 8.2.1.1), such as the breaking of the critical salt bridge and the movements of swI and swII, could lead to the loss of coordination between $H_2PO_4^-$ and MG, as shown in Figure 8.9. This could be a consequence of a rotation of the inorganic phosphate following the movement of the swI with which it was establishing an H-bond (O1-S204O) that then breaks (Figure 8.10).

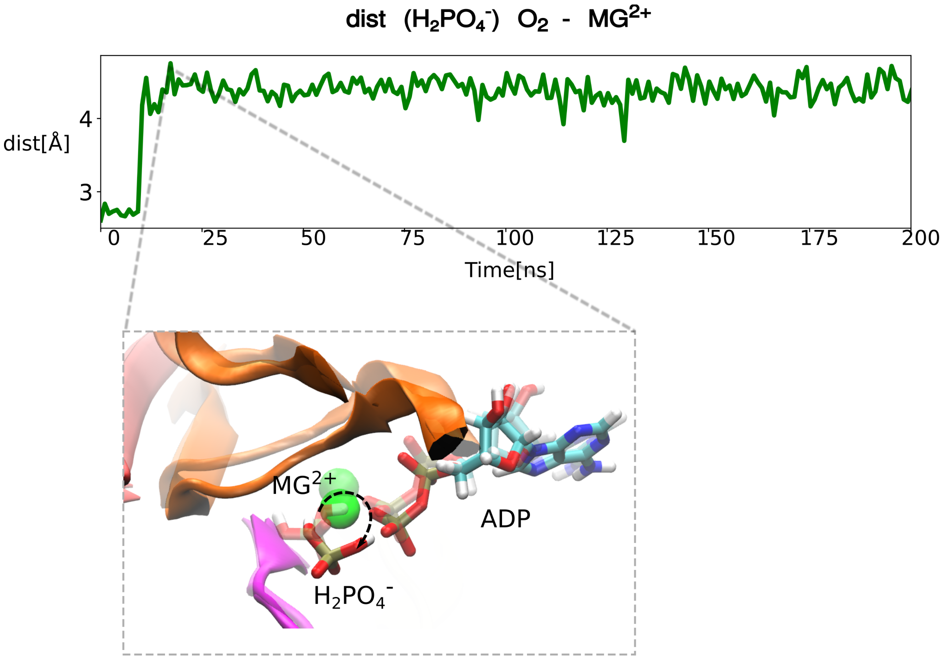


Figure 8.9: Timeseries along the 200 ns MD of the coordination between $H_2PO_4^-$ and MG^{2+} (plot on the top). This observable has been monitored as the distance between the O_2 of $H_2PO_4^-$ and MG^{2+}. At the bottom, a zoom of a frame after the rotation of $H_2PO_4^-$ (dotted arch) causes an increase in this distance to 5.2 Å, resulting in a loss of coordination with the co-factor. In transparent colors the different elements before the rearrangement and in opaque colors the same elements of the frame chosen to show the observable under investigation.

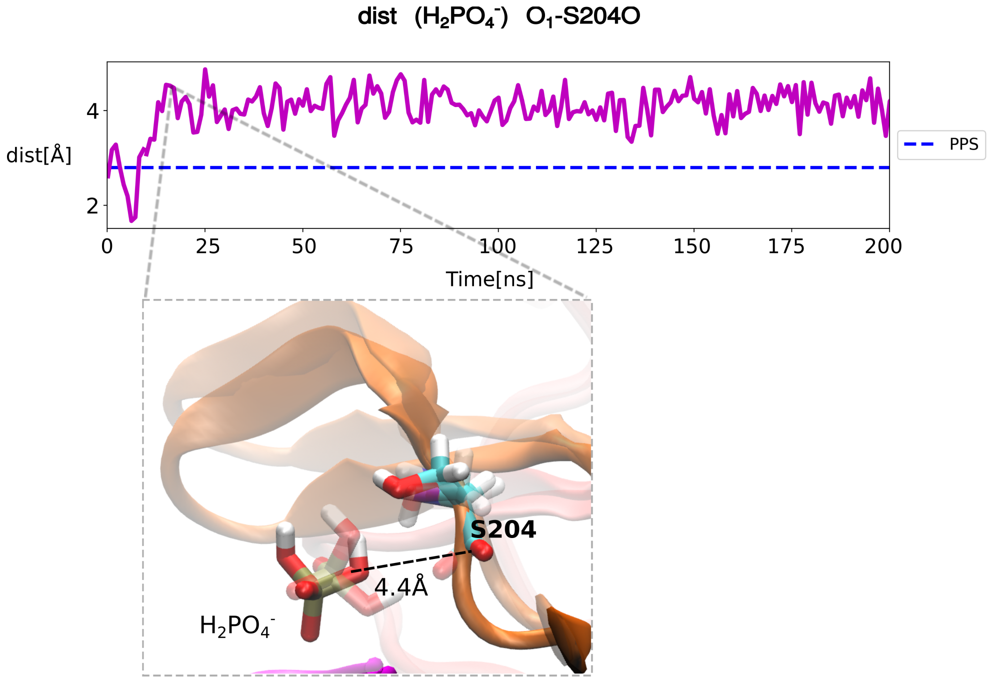


Figure 8.10: Timeseries along the 200 ns MD of the H-bond formed between O1 of $H_2PO_4^-$ and O of S204 (residue of swI in U50) in the plot on the top. This H-bond results formed in the PPS X-ray structure 2V26.pdb (dashed blue line in the plot corresponding to a value of 2.9 Å). On the bottom a zoom of a frame of the trajectory in which the distance between the two atoms considered increases to a value of 4.4 Å, resulting in the breaking of the H-bond. In transparent are depicted the swI, $H_2PO_4^-$ and S204 at the beginning of the MD, and in opaque colors the same elements at the frame chosen to show the breaking of the H-bond under analysis.

The rotation of the inorganic phosphate that provokes a loss of coordination with MG^{2+} has recently been addressed by Mugnai and co-workers (Mugnai and Thirumalai, 2021a), as a key step for subsequent movement of the converter, highlighting once again a possible communication path between nucleotide binding pocket and force-generating element of the motor.

8.2.1.4 *Movement of the ADP in the active site*

Focusing on the rotation of $H_2PO_4^-$ and its effects, we monitored another contact, defined by Mugnai and co-workers (Mugnai and Thirumalai, 2021a), which could be indicative of a movement of ADP in the active site. This contact is represented by the distance between the nitrogen of ADP (N6) and the Cα of N98 (residue of N-terminal subdomain). In Figure 8.11 we can appreciate how this contact between the nucleotide and the residue of the N-terminal is lost for around 100 ns. The ADP flips in the active site, yet without unbinding, as a consequence of the rotation of the inorganic phosphate, which generates a "stress" that weakens the interaction with the motor (Mugnai and Thirumalai, 2021a).

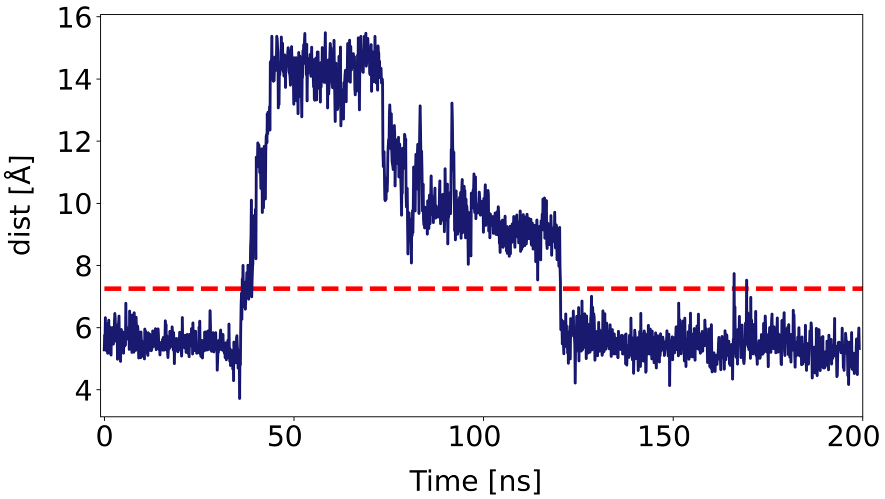


Figure 8.11: Timeseries of the distance between ADP(N6) and N98, residue of the extremity of the second β-strand. The average value is 7.9 ± 3.3 Å. The red dashed line is the threshold distance (7.25 Å) defined by Mugnai *et al* in (Mugnai and Thirumalai, 2021a)

This result would be another evidence more in agreement with a possible progression towards P_iR state, rather than a reversal process to PTS.

8.2.1.5 *Monitoring of geometrical observables: uncoupling of the converter*

The loss of coordination between phosphate and magnesium seems to worsen also in our simulation the structural instability of the motor, which can be transmitted along the motor domain, through the RH, reaching the converter which then uncouples from the N-terminal subdomain. This can be seen in Figure 8.12 in which is plotted the distance between the Cα of F763 and L120, residues of the converter and N-terminal subdomain, respectively. This distance stabilizes at PTS value (13.3 Å) with an average value of 13.6 ± 2.5 Å, far from the reference value of the PPS structure of 5.9 Å (Figure 8.12).

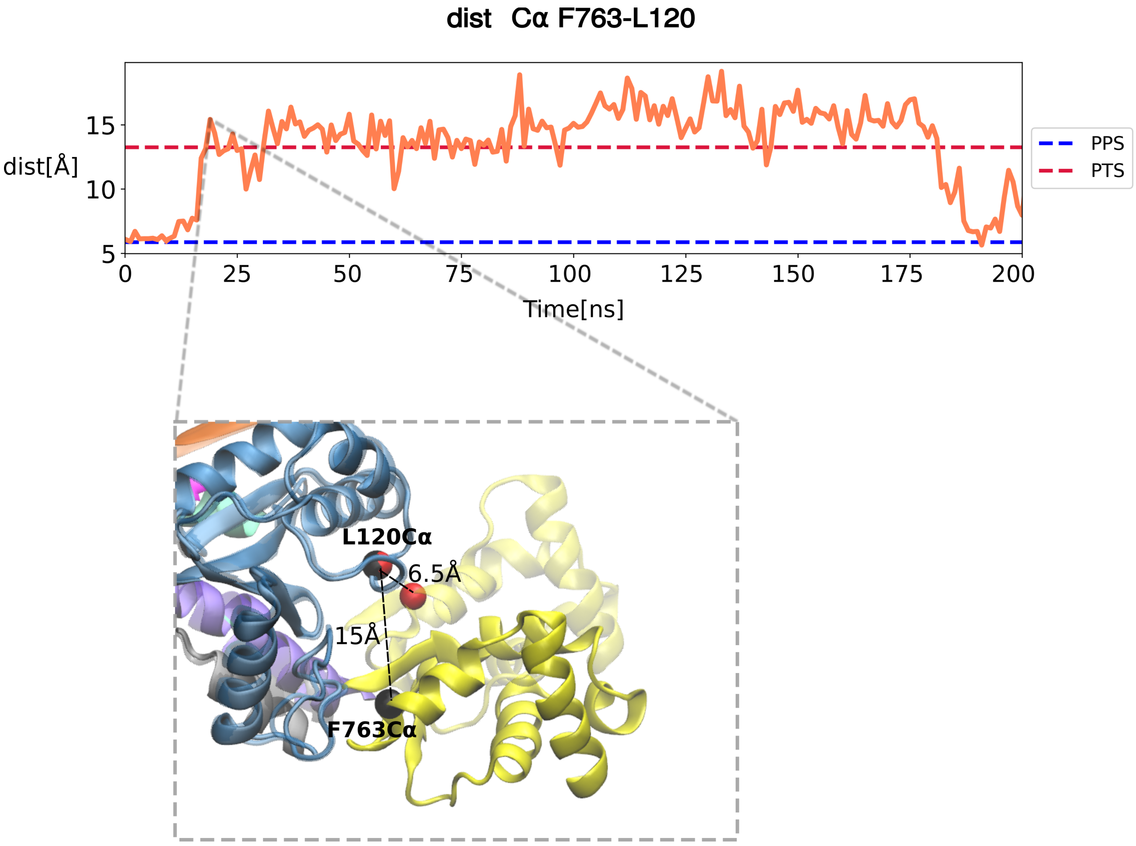


Figure 8.12: Timeseries along the 200 ns MD of the coupling between the converter and N-terminal subdomains (plot at the top). To define this observable, we chose to monitor the distance between the Cα of F763 and L120, a residue of the converter and N-terminal subdomain, respectively. The distance between these two atoms varies considerably in the two states PPS and PTS (red and blue dashed lines in the top panel). In the square at the bottom, a zoom of this distance at the frame is extracted from the trajectory when the converter undergoes its rotation and moves away from the initial position (in transparent yellow) in respect of the N-terminal subdomain (in sky-blue). As a consequence of this movement, the distance between the two selected atoms increases at values closer to the PTS reference one (13.3 Å measured from the X-ray structure 5O2L.pdb), rather than PPS one (5.9 Å measured from the X-ray structure 2V26.pdb). In red spheres are represented the Cα of F763 and L120 at the initial frame of the MD and in black spheres the same atoms at 20 ns. The distance between these atoms increases from 6.5 Å to 15 Å.

What drives the uncoupling of the converter from the N-terminal subdomain? This spontaneous and reversible transition, captured by MD simulation, is likely governed by statistical coupling between subdomains (HUXLEY, 1957). This means that changes in the different collective variables are not strongly correlated to each other and the motions are caused by thermal fluctuations and the intrinsic flexibility of the protein. This makes the system fluctuate between different energy minima and the protein is, thus, able to explore an ensemble of configurations. It could be possible that, thanks to thermal fluctuations, SH1-helix drives the initial rotation of the converter. In particular, from the very beginning of the simulation, this helix fluctuates toward a more tilted orientation as for the PPS reference structure (see figure Figure 8.13). This tilting is maybe linked to a rotation of M701 towards RH (see Figure 8.14 sub-figure (a)).

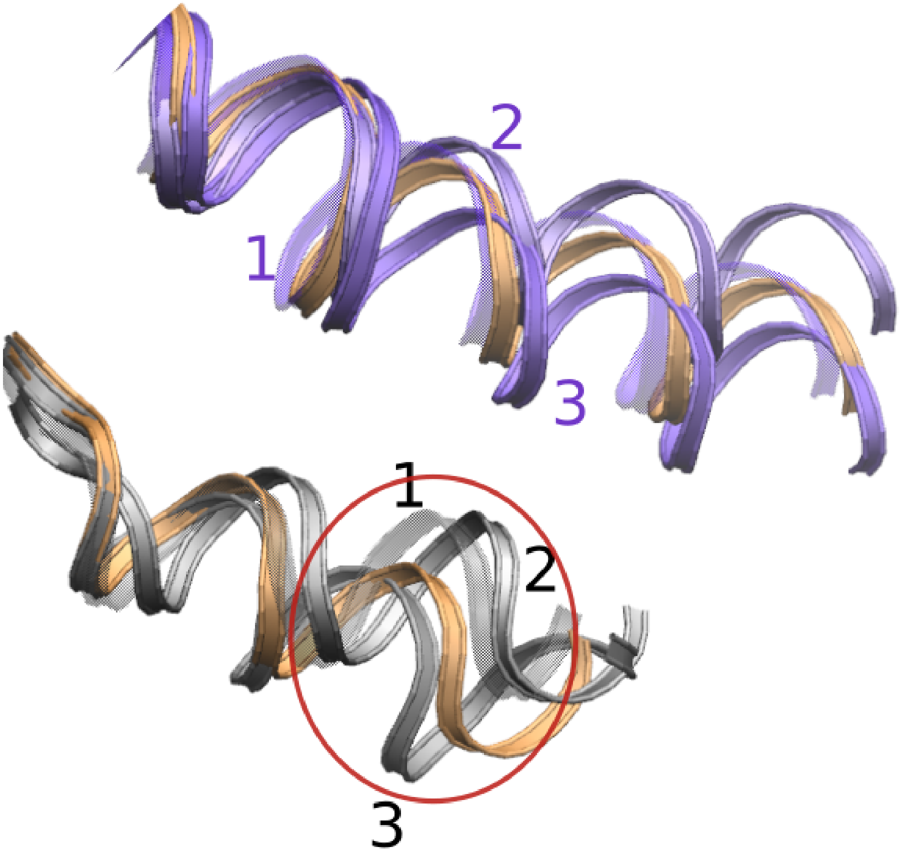
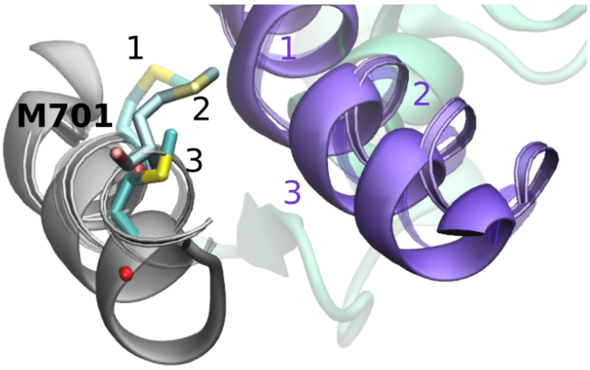
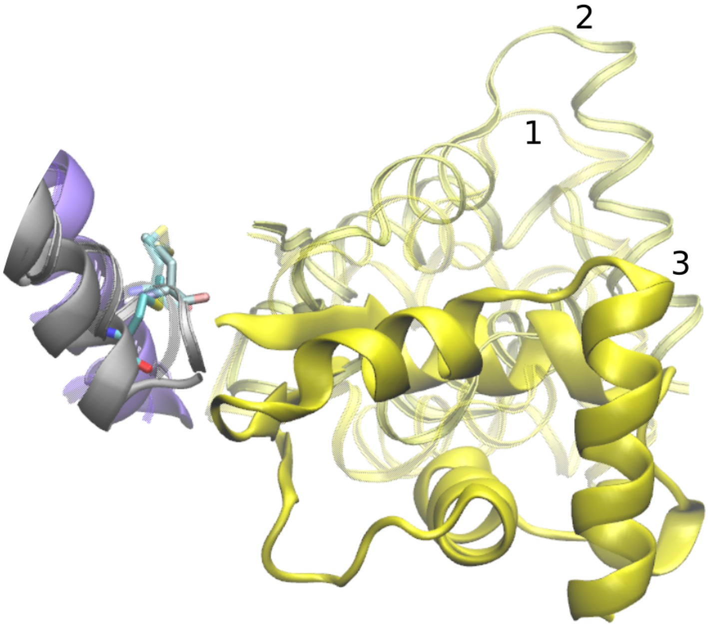


Figure 8.13: Zoom on the SH1 (in grey ribbons) and C-terminal portion of RH (in violet ribbons) of three different frames extracted from the MD and of the PPS reference structure (in orange ribbons). In transparent is depicted the first frame of the trajectory, in light grey the second one and in dark grey the third one. As well as for SH1 also for RH is represented in transparent, light, and dark violet the C-terminal portion at the three frames chosen. In red is circled the tilted portion of SH1 described in the text.



(a) Zoom on M701 (in cyan sticks), a residue of SH1 helix (in grey), that changes orientation during the MD and on the RH (in violet), which moves concurrently. The elements in the first and the second frames are illustrated as ribbons and in the third frame chosen in cartoon.



(b) Zoom on the converter fluctuating between three positions. In ribbons is shown the converter in the first and second frames chosen. With cartoon is depicted the converter in the last frame selected for the analysis.

Figure 8.14: Sub-figure (a) is a zoom on the M701, residue of SH1 helix. The figure depicts three different frames extracted from the trajectory, in which we observe a change in the orientation of this residue that could be involved in the uncoupling of the converter from the N-terminal subdomain. The numbers indicate the sequence of this change, progressively from the beginning of the MD (number 1) to the frame in which we observe the uncoupling of the converter. In sub-figure (b) zoom on the converter in the same frames as in sub-figure (a). Number 1 represents the position of the converter at the beginning of the MD. The converter is next exploring an upper position (number 2) and then it stabilizes to a lower position at the stage of uncoupling from the N-terminal subdomain (number 3).

The tilting of the terminal part of SH1, directly connected to the converter, is coupled to a movement of the converter which fluctuates between three main positions (see Figure 8.14 sub-figure (b)). From an energetic point of view, it seems that in our simulation the most favorable position among the ones explored by the converter is the lower one (number 3 in Figure 8.14 sub-figure (b)) since it shifts here for almost the whole simulation time (as shown in the time-series Figure 8.3 sub-figure (a)).

The interdependence between rearrangements at the level of the active site, RH, and SH1 has already been described as linked to the re-orientation of the converter (Koppole et al., 2006; Ménétrey et al., 2005, 2007). In particular, the higher flexibility of swII upon ATP-hydrolysis (Koppole et al., 2006) could play an important role in the motion of RH, SH1, and converter, being this loop directly connected to the N-terminal portion of RH, in turn, connected to SH1 and converter.

The increased pliancy in the motion of the converter is linked also to the loss of some interactions at the converter/N-terminal interface, which triggers the uncoupling. The correlation between the uncoupling of the converter from the N-terminal subdomain and the simultaneous loss of a hydrophobic interaction at the converter/N-terminal interface is depicted in Figure 8.15. The initial rotation of RH-SH1 and converter with respect to the N-terminal subdomain correlates with the breaking up of the hydrophobic interaction established between L120 and F763 (Pearson's coefficient equal to 0.81).

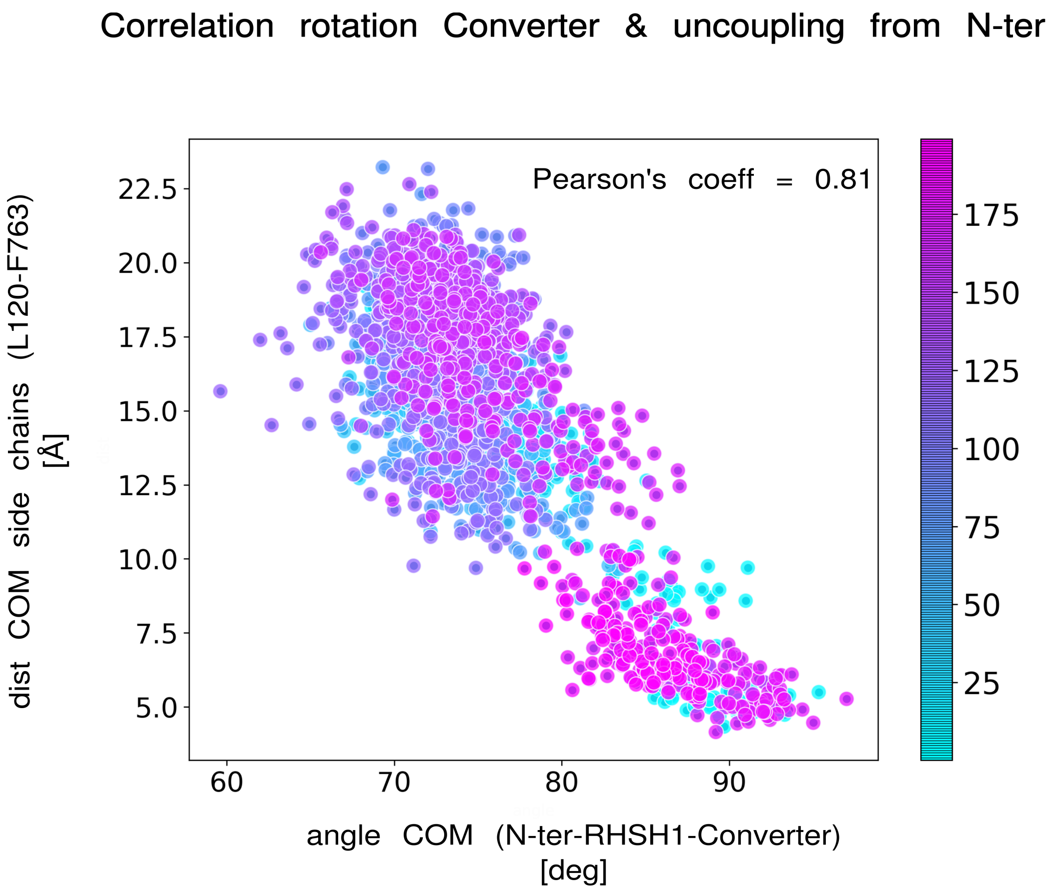
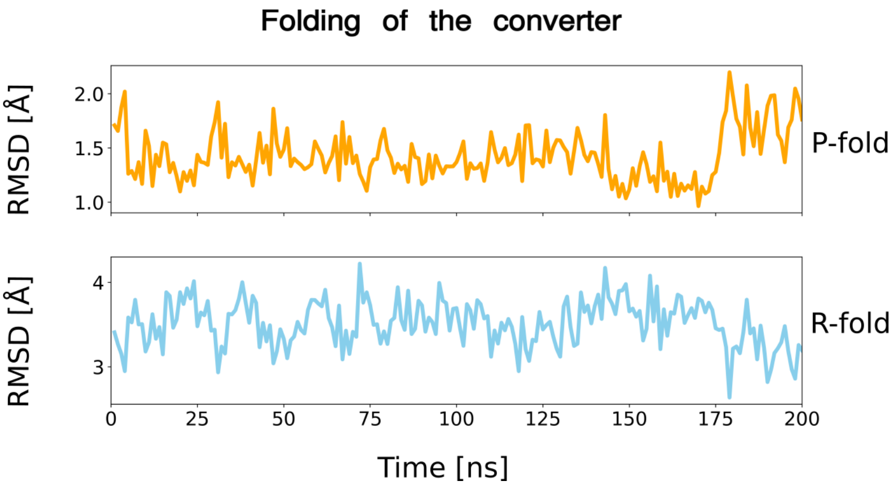
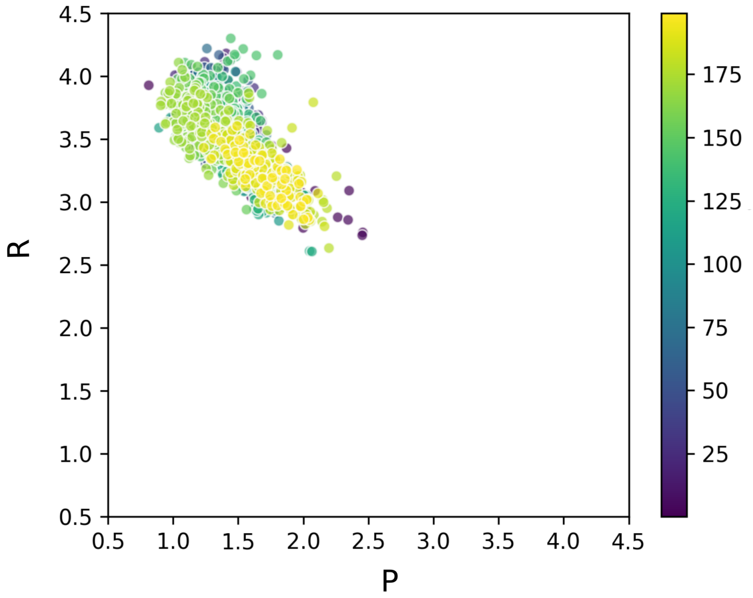


Figure 8.15: Correlation between the initial rotation of the converter and the uncoupling from the N-terminal subdomain along the MD. The dots in the plot are colored according to the color map on the right side, which indicates the trajectory of 200 ns. On the x-axis is plotted the angle formed between the center of mass of $C\alpha$ atoms of some residues of the N-terminal subdomain (residues 128 to 140), of RH-SH1, here considered merged into a single element, (residues 489 to 500 and residues 700 to 705) and of the converter (residues 750 to 770). On the y-axis the distance of the center of mass of the side chains of L120 (residue of N-terminal) and F763 (residues of the converter) which establish an initial hydrophobic interaction that then breaks simultaneously with the uncoupling.

Nevertheless, the converter still maintains a fold typical of the PPS state, named P-fold (Ovchinnikov et al., 2011) (Ménétrey et al., 2012) and different from the canonical R-fold conformation, described both for R (Ménétrey et al., 2005) and PR structures (Ménétrey et al., 2008) (see Figure 8.16). We analyzed the fold of the converter by computing its RMSD after superimposition on the same domain of PPS and PTS, respectively (Figure 8.16 top and bottom plot respectively). The average value of this measurement converges to the P-fold configuration with an estimation of 1.5 Å.



(a) RMSD of the converter vs PPS reference structure (2V26.pdb) in the panel at the top and RMSD of the converter vs PTS reference structure (5O2L.pdb) in the panel at the bottom.



(b) Scatter plot of the RMSD of the backbone of the converter with respect to PPS reference structure (2V26.pdb) and PTS reference structure (5O2L.pdb).

Figure 8.16: In sub-figure (a) the RMSD time-series of the backbone atoms of the converter computed with respect to PPS and PTS reference structure. In sub-figure (b) a scatter plot of the same RMSD. As we can see the converter keeps a P-fold conformation, with an average RMSD over 200 ns MD of 1.4 ± 0.2 Å. The average RMSD with respect to the R-fold conformation is 3.5 ± 0.3 Å.

8.2.1.6 *Sequence of events*

A summary of a plausible sequence of events of the spontaneous transition of the motor domain of MyoVI in PPS toward a PTS-like conformation is presented in Figure 8.17 and in the video beneath.

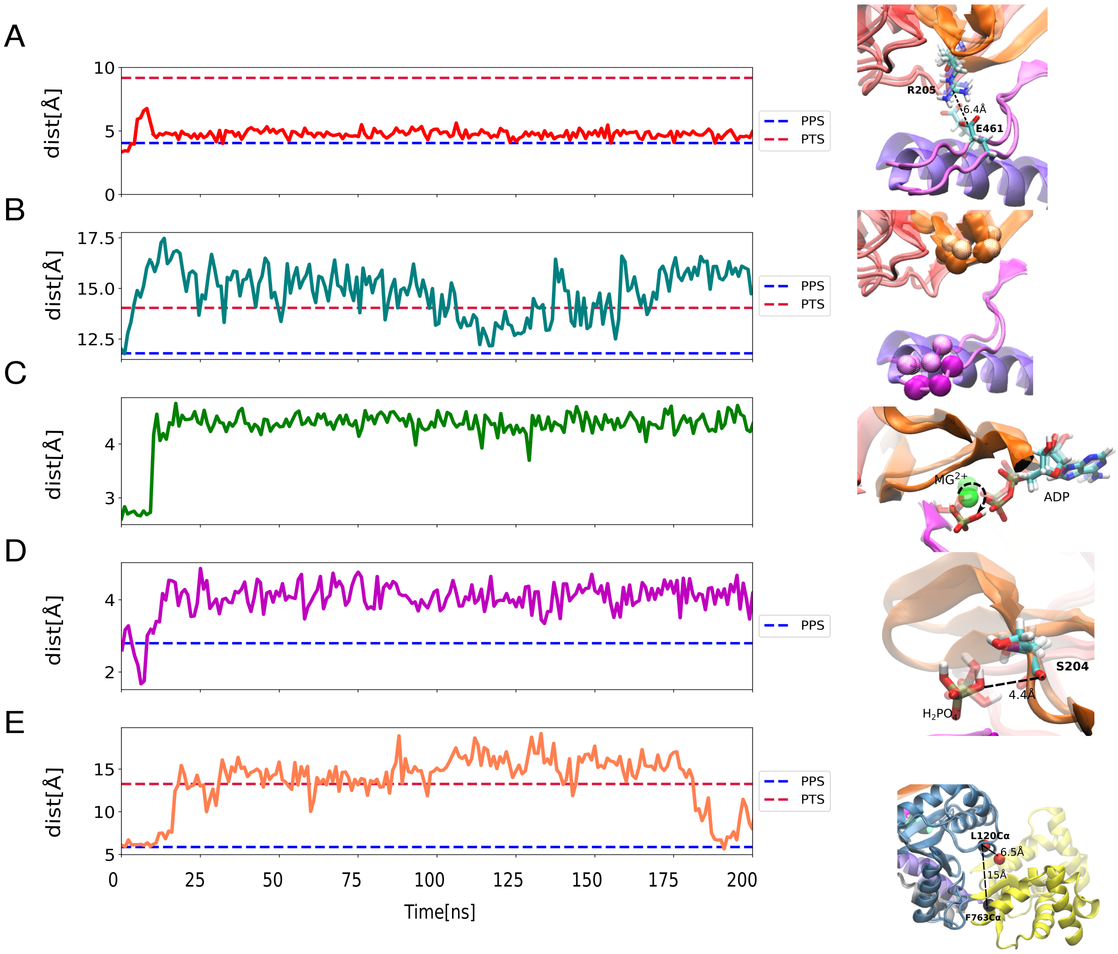


Figure 8.17: Possible sequence of events that leads to the transition towards a PTS-like state, captured in 200 ns MD.

Link for the video that shows the sequence of rearrangements presented in the text: `https://zenodo.org/record/6907258#.YuAP8exBxH4`.

8.2.1.7 *Additional analysis: twisting of the transducer*

The twisting state of the central β-sheet, or transducer, is involved in the functional transitions of the motor domain (Cecchini et al., 2008; Coureux et al., 2004; Yang et al., 2007). To investigate whether the spontaneous transition we capture is accompanied by large rearrangements of the transducer, we analyze its twisting angle during the simulation. We represented the sheets as two vectors, considering the Cα of two extreme residues of the first β-sheet (β1) (residue 86 and 91) and of the seventh (last strand of the transducer) β-sheet (β7) (residue 222 and 228). We then calculated the scalar product between these vectors. The results of this analysis run over the trajectory of 200ns described in Figure 8.18 reveal that the transducer twists towards a post-rigor like (PR) or between PR and Pi-release (P_iR) conformation (see Figure 8.18).

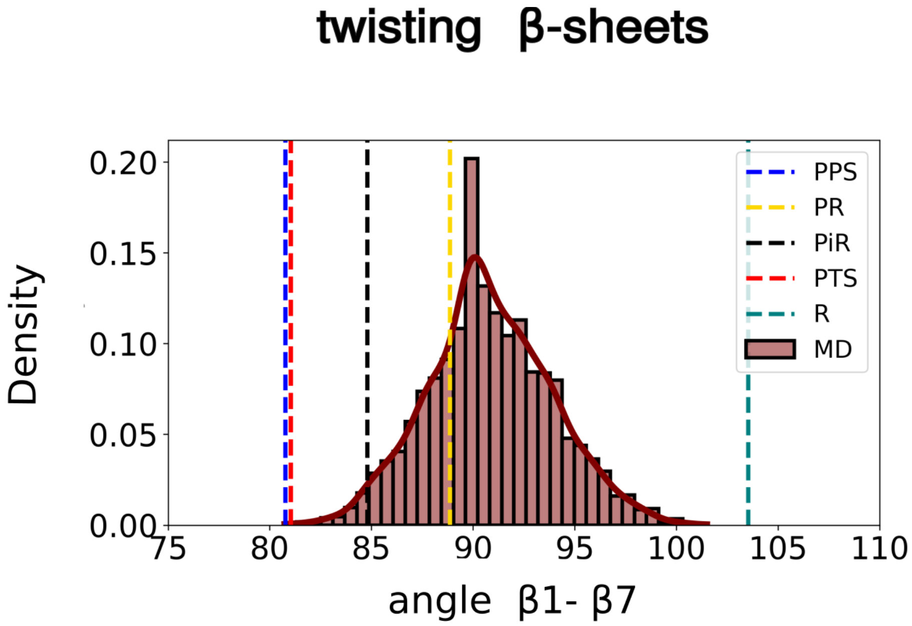


Figure 8.18: Distribution of the twisting of the β-sheets computed for the 200 ns MD.

We checked whether the difference in transducer twist still holds in simulations of the various states, by computing the same angle in the MD of PTS and PR states, respectively. For PTS we run 10 replicas of 100 ns each, starting from the structure modeled by Blanc (Blanc, 2018). The protocol for the MD was the same as the one used for the PPS simulation (Section 8.2.0.1). In Figure 8.19 we can see that the twisting of the transducer has a bimodal distribution, ranging from PTS to PR.

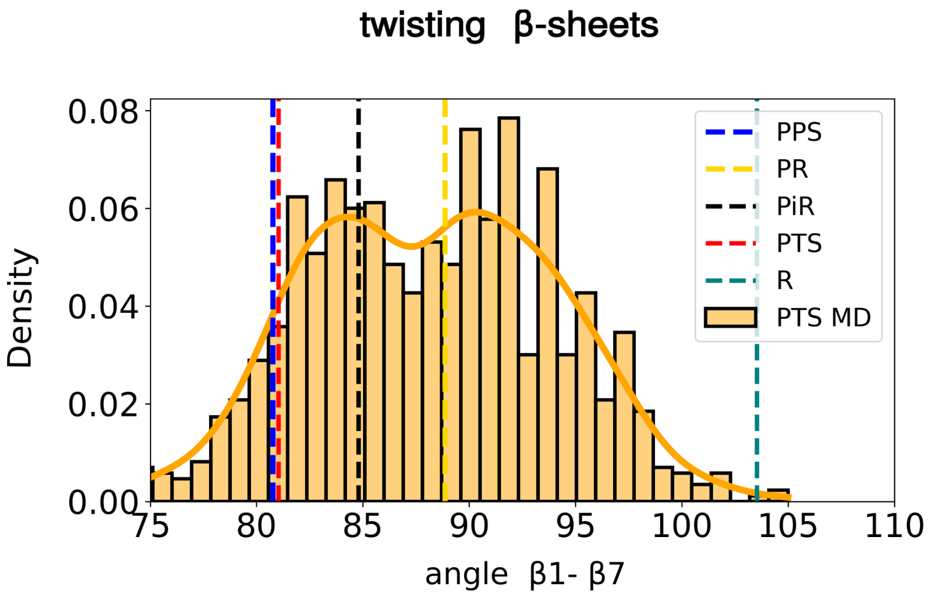


Figure 8.19: Distribution of the twisting of the β-sheets computed for 1μs MD of PTS state.

We then computed the same twisting angle for a simulation of the PR state run by Blanc. In Figure 8.20 we can see that the twisting of the transducer picks on the PR value of reference.

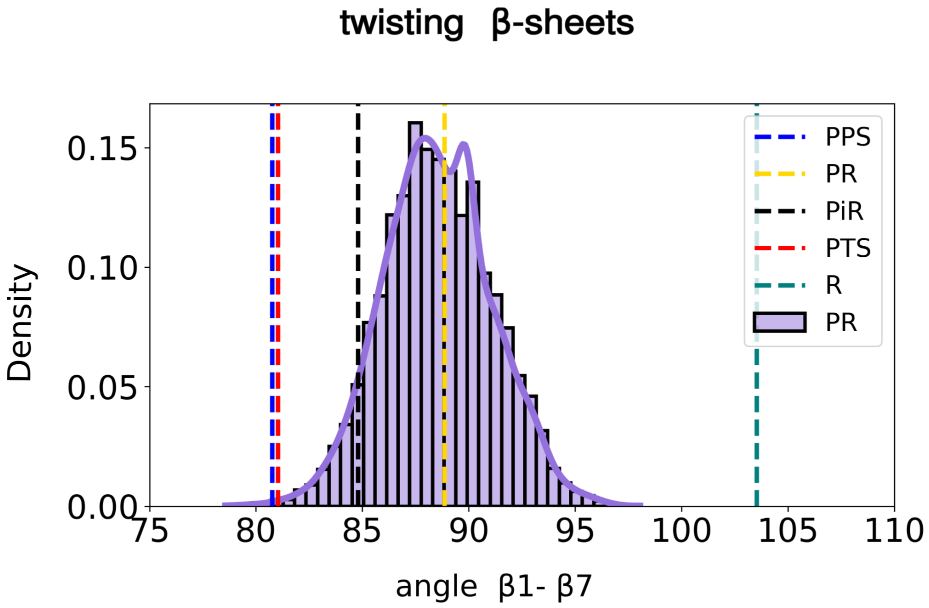


Figure 8.20: Distribution of the twisting of the β-sheets computed for 100 ns MD of PR state.

8.3 CONCLUSIONS & FUTURE PERSPECTIVES

Based on our results, we propose that our classical MD simulation captured a reversible transition from PPS to what we defined as being a PTS-like state, given the melange of PPS and PTS features. This transition is most evident in the unique dynamic of the converter and its uncoupling from the N-terminal subdomain. This major displacement is driven by entropy and thermal fluctuations that characterize myosin's intrinsic flexibility. In particular, the rearrangements observed in the nucleotide-binding site are amplified and affect the behavior of the converter as well as the rotation of the lever arm (Ménétrey et al., 2005; Mesentean et al., 2007; Mugnai and Thirumalai, 2021a; Warshaw, 2004). Moreover, the initial rotation of the inorganic phosphate might play a major role in communicating structural instability to the force-generating elements, namely RH, SH1, and converter (Lawson et al., 2004; Mugnai and Thirumalai, 2021a), but despite this, the phosphate is still trapped in the active site, probably thanks to other interactions formed with active-site binding loops (swI, swII, and P-loop) (Mugnai and Thirumalai, 2021a). Besides, the partially open swII and the position of the connector elements RH and SH1 resemble the PTS state rather than PPS, but the folding of the converter is still descriptive of the PPS state (P-fold). It has already been demonstrated that the converter of MyoVI in PPS has a unique orientation and could plausibly bind actin after uncoupling from the motor domain (Ménétrey et al., 2005). The combination of these observations leads us to think that we simulated a possible pathway that the motor domain could take in a monomeric form, starting from a PPS configuration and going to a pre-transition-like state. Besides, the motion that we captured is connected with the distortion of the transducer that communicates with the active site (Coureux et al., 2004) and may be involved in energy storage (Cecchini et al., 2008). The twisting of the β-sheets towards a PR or P_iR conformation supports this hypothesis. This distortion can trigger additional rearrangements, leading to a conformational transition from a high free-energy state (PPS) to lower free-energy states (PR or P_iR) (Karatzaferi et al., 2004) (Yang et al., 2007), without an unnecessary release of hydrolysis products from the active site. Moreover, as we already mentioned, the converter keeps a P-fold conformation for much of the trajectory, compatible with PPS and P_iR states. These results could be interpreted as a progression in a non-force generating pathway, rather than a backward transition. Conversely, they can suggest a behavior in agreement with the hypothesis proposed by Ménétrey and co-workers of reversible uncoupling of the converter in MyoVI for a normal progression of a working stroke (Ménétrey et al., 2005). The structural transition (from PPS to PTS or from PPS to P_iR) is not entirely achieved, because we analyzed a trajectory of few ns, whereas complete conformational transitions occur after a timescale of milliseconds (Mesentean et al., 2007; Mugnai and Thirumalai, 2021b; Trivedia et al., 2015). Nevertheless, the reason behind this transition is not entirely clear. Unless caused by actual functional reorganizations of the motor domain, it could be caused by the protonation state chosen for the titratable residues as well as by the force-field selected (CHARMM36mff (Huang et al., 2016)) for the simulation or, yet, the parameters used for the ligands in the active site.

In this work, we aimed at proposing a possible pathway of spontaneous fluctuations that the motor domain may undergo, with a detailed sequence of events. What the actual captured state remains unclear because of the latest evidence of a possible forward transition to the P_iR state. Future studies may shed light on the chemomechanical transduction in MyoVI, through additional investigations of the various conformational states of the actomyosin cycle. Interesting analyses may involve the determination of free energy barriers between all states, applying some of the most widely used methods of enhanced sampling that relies on the use of collective variables (Branduardi and Faraldo-Gómez, 2013; Fukunishi et al., 2002; Maragliano et al., 2006).

8.4 EXPLORING THE STABILITY OF PPS

We investigated the stability of PPS in MyoVI to understand the possible causes of the peculiar dynamic observed in Section 8.2. It might depend on the protonation state chosen for the titratable residues of the PPS model since this might have a crucial effect on the dynamics of the protein itself. Furthermore, we questioned whether the behavior observed for MyoVI was unique to this class of myosins or was common to other classes. We, therefore, compare the stability of a processive myosin such as MyoVI with that of a non-processive motor, such as cardiac myosin.

8.4.1 Exploring the stability of PPS in Myosin VI and Cardiac myosin

The behavior of MyoVI in the PPS state was monitored over a pseudo-trajectory of $2\mu s$, obtained from the concatenation of ten different replicas of $200ns$, run starting from a new model of the same X-ray structure of MyoVI in PPS previously used, solved in complex with $ADP.VO_4^{3-}$ (PDB code: 2v26.pdb). For each simulation, we used different initial coordinates. The analysis envisages also a comparison with a different myosin: β cardiac myosin (β-CM) (Planelles-Herrero et al., 2017). Notoriously cardiac myosin belongs to class II of myosin superfamily (Barrick and Greenberg, 2021) and, unlike MyoVI, it is a non-processive myosin (Greenberg et al., 2017) (Altman et al., 2004).

8.4.1.1 *System preparation*

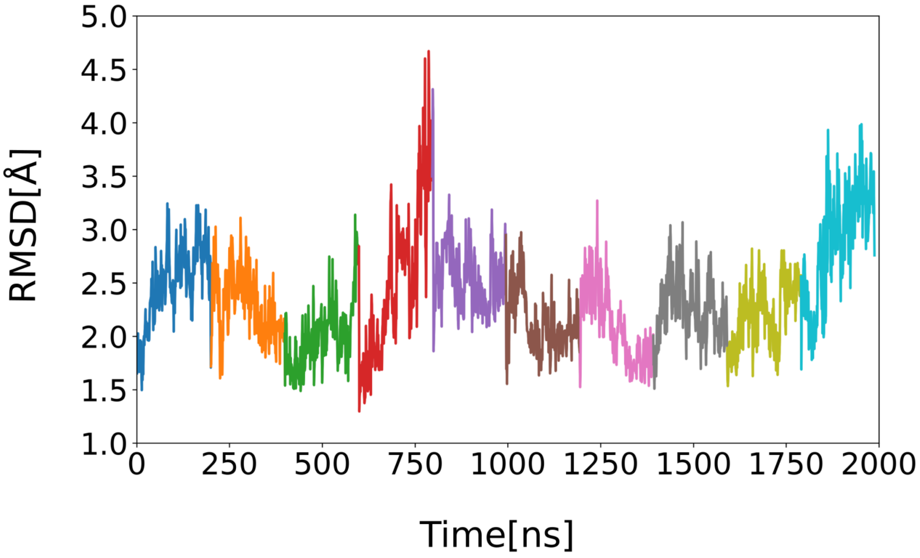
The new model of 2v26.pdb was modeled in the same way as described in Section 8.2.0.1. The parameters used for the ligands (ADP and $H_2PO_4^-$) were obtained by the more recent version of CGenff, than the one used in the previous experiment (Section 8.2.0.1) and the final system (MyoVI in complex with ADP-MG^{2+}-H_2PO4^-) was modeled with CHARMM. For this experiment we have chosen the non-standard protonation state for the following residues: ASP137; ASP312; ASP450; GLU170; GLU681; LYS39; LYS208; LYS562 and LYS740. The Solution Builder module of CHARMM-GUI (Jo et al., 2008) was used to solvate and ionize the system in 11 Å edge water box in which 171 Sodium ions and 174 Chloride ions were added to neutralize. The same procedure was followed to prepare the S1 fragment of Cardiac myosin (β-CM), starting in this case from the PPS structure 5N69.pdb (Planelles-Herrero et al., 2017), solved in complex with $ADP - MG^{2+} \cdot VO_4^{3-}$ in the active site and the allosteric modulator Omecamtiv Mecarbil (OM) (Morgan et al., 2010; Planelles-Herrero et al., 2017). For β-CM we used our model described in Section 6.3.5. The inorganic phosphate was modeled with CHARMM in its form $H_2PO_4^-$ and its topology and parameter files were obtained with the more recent version of CGenff. OM was removed from its binding pocket, to simulate the motor domain only with the ligands ($ADP - MG^{2+} \cdot H_2PO_4^-$) in the active site, to avoid further allosteric effect.

8.4.1.2 *Molecular Dynamics (MD) simulations and Visualization tool*

The protocol used for the MD simulations was the same as in Section 8.2.0.2 and was applied both for MyoVI and for β-CM. The equilibration and production steps were repeated ten times for the different replicas of the two systems. All the MD simulations were run on the GPU partition of the supercomputer Jean Zay (`http://www.idris.fr/eng/jean-zay/jean-zay-presentation-eng.html`). The different replicas will be referred to from now on by the letter "R" both in the text and in the figures.

8.4.1.3 *Structural & dynamic characterization of MyoVI and β-CM in PPS state*

We firstly analyzed the stability of the motor domain in the simulations of MyoVI, by computing the RMSD of the core atoms with respect to the initial minimized structure, ignoring the loops (Figure 8.21). The motor domain results overall stable, with an average deviation from the minimized structure of 2.3 Å. The density plot of this RMSDs (Figure 8.21, panel at the bottom) shows that four replicas diverge from the average stability (R1, R4, R5, and R10), displaying a higher RMSD.



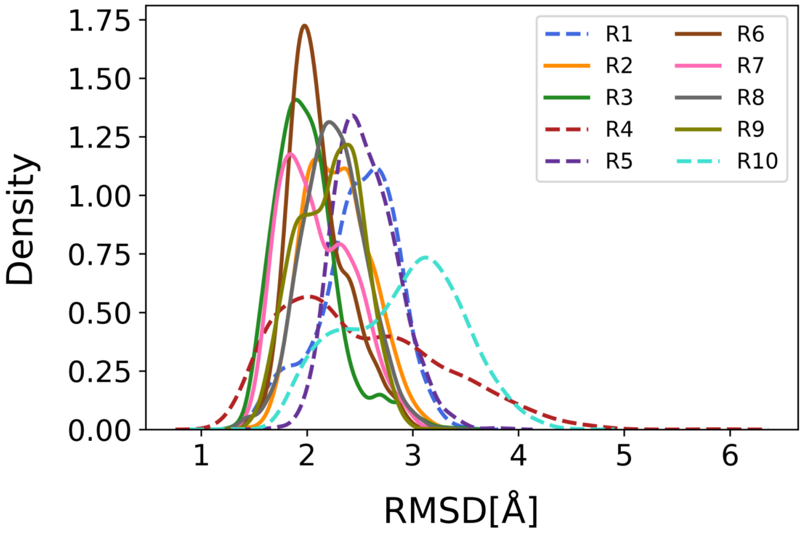


Figure 8.21: Timeseries of the RMSD of the Cα of the core atoms motor domain in ten replicas of MyoVI (top plot). Distribution plot of the same RMSDs on the bottom. In dashed lines the curves of the replicas in which the motor domain shows higher deviation from the minimized structure. The average value of the RMSD over the pseudo-trajectory is 2.3 ± 0.48 Å.

After analyzing the trajectories, we noticed once again remarkable flexibility in the converter. This is shown in Figure 8.22, in which RMSD analysis shows that the converter is flexible, with some replicas sampling deviations up to 15 Å (case in R4).

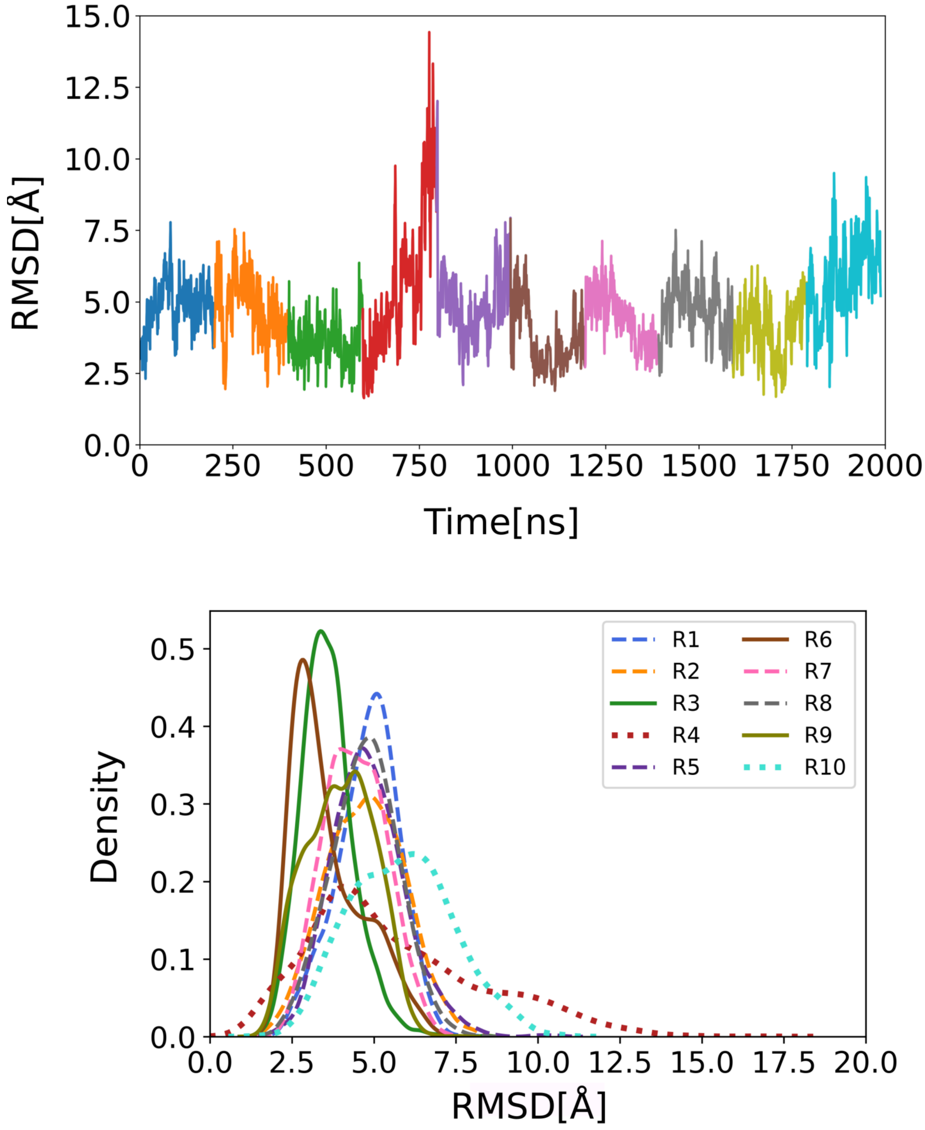
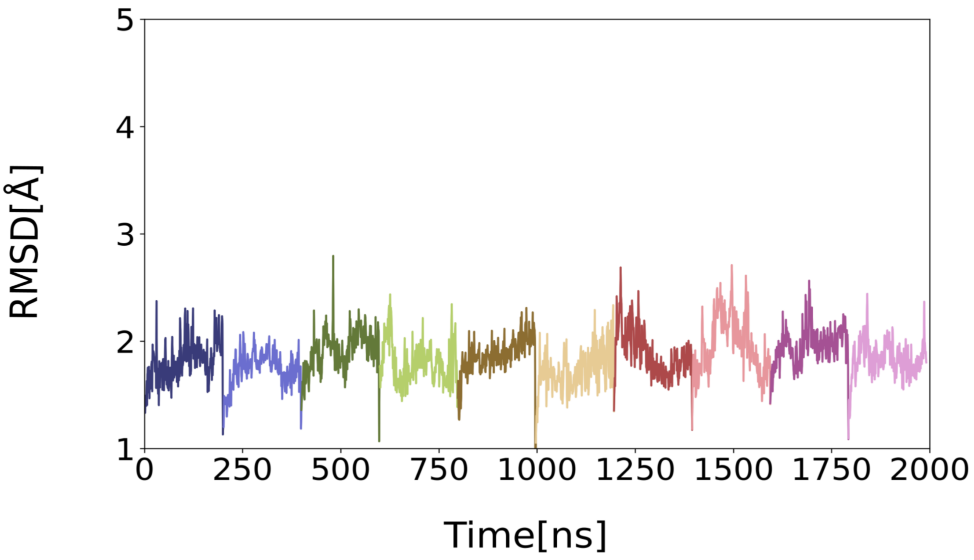


Figure 8.22: RMSD of the Cα of the core atoms of the converter subdomain after alignment to the N-terminal domain (top) and density plots of this RMSD for each replica (bottom). The average RMSD value over the pseudo-trajectory is 4.6 ± 1.51 Å.

By investigating in more details the deviation of the converter, in the ten replicas, it samples three distinct configurations, as highlighted in the bottom panel in Figure 8.22, by the curves pictured with continuous, dashed and dotted lines.

This data is quite interesting if compared to the behavior of the motor domain of β-CM (Planelles-Herrero et al., 2017), over a pseudo-trajectory of $2\mu s$, analyzed after concatenation of ten trajectories of 200 ns. Throughout the whole trajectory, the motor domain of β-CM is stable in its core atoms, with an average RMSD of 1.8 ± 0.22Å (see Figure 8.23), with respect to the minimized structure.



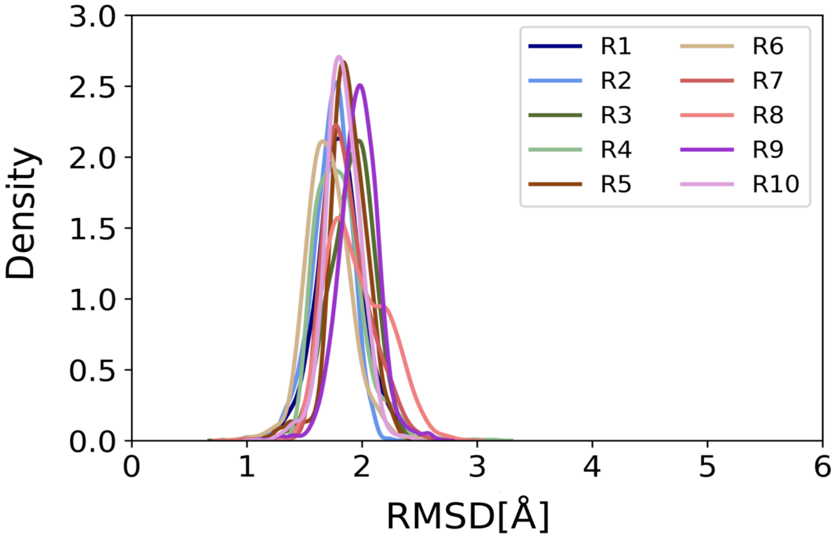


Figure 8.23: Timeseries of the RMSD of the Cα of the core atoms motor domain in ten replicas of β-CM (top plot). Distribution plot of the same RMSDs on the bottom. The average value of the RMSD over the pseudo-trajectory is 1.8 ± 0.22 Å. Also in this case the RMSD of the converter was computed after alignment to the N-terminal domain of the minimized structure of β-CM.

The analysis of the RMSD of the Cα of the converter, after alignment to the N-terminal domain (Figure 8.24), shows a constant fluctuation around 3.5 Å with exception of R3, whose average RMSD is 3.9 Å. These data suggest that the converter in β-CM samples a unique configuration in PPS, compared to the PPS of MyoVI, in which instead it undergoes again a peculiar dynamic.

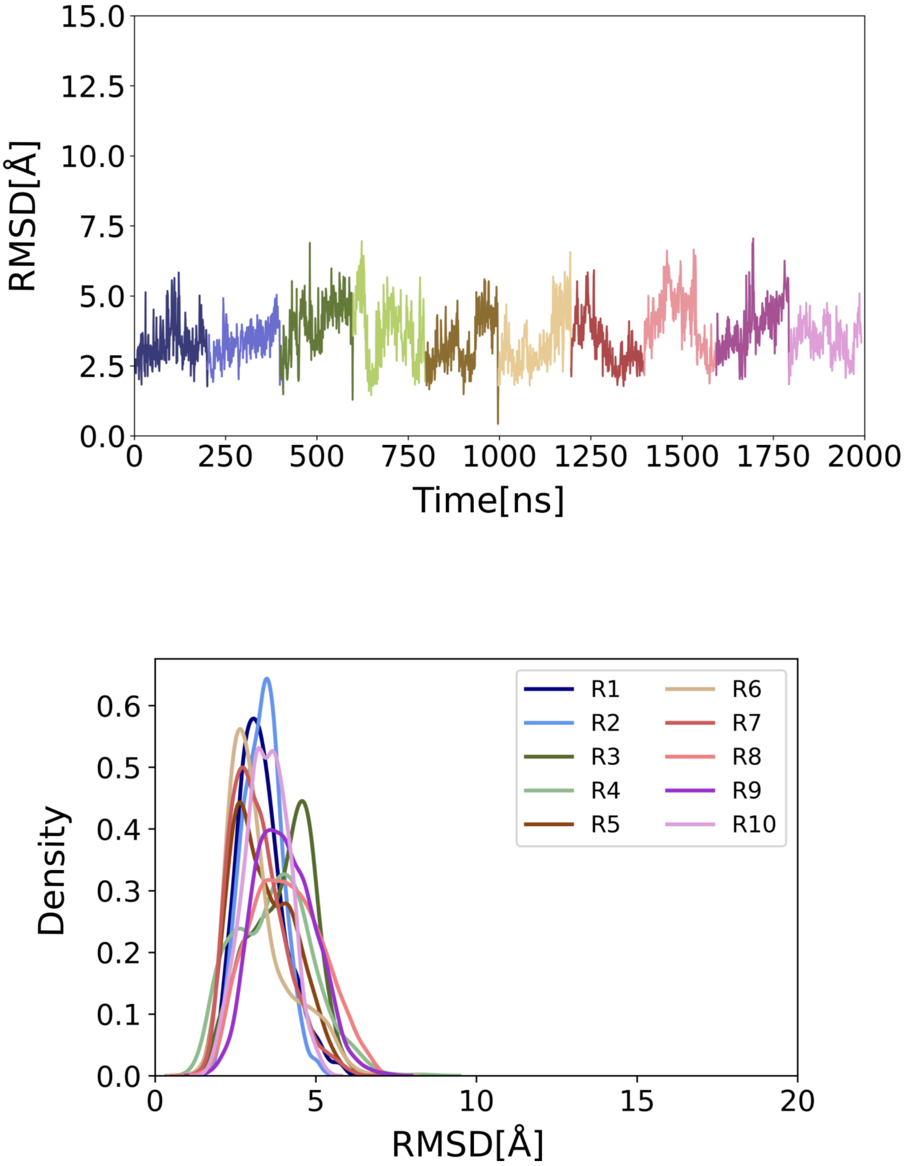


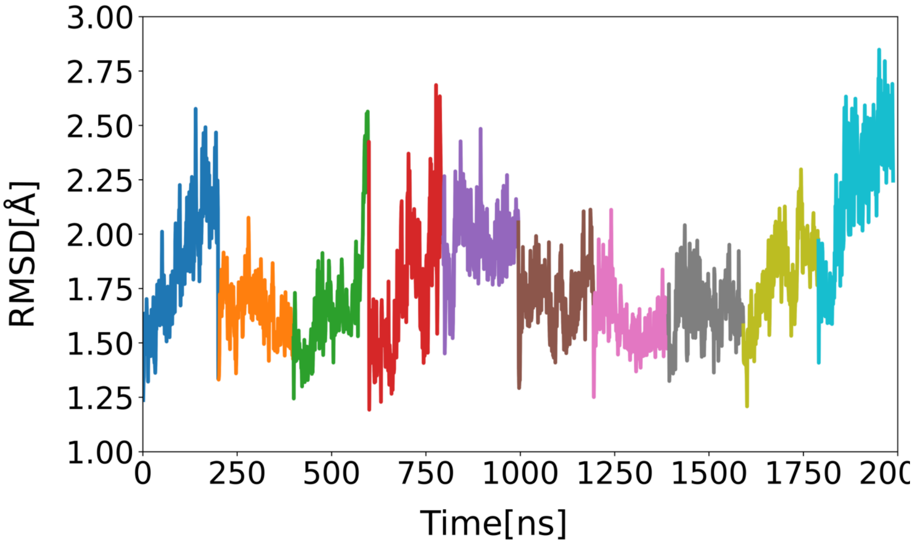
Figure 8.24: RMSD of the Cα of the core atoms of the converter subdomain after alignment to the N-terminal domain (top) and density plots of this RMSD for each replica (bottom). The average RMSD value over the pseudo-trajectory is 3.5 ± 0.9 Å.

RMSD analysis shows that the converter is flexible in MyoVI PPS, with some replicas sampling deviations up to 15 Å. This suggests that the large fluctuations of the converter are robust to the change in the protonation state added in the model. By contrast, the same analysis shows that the converter of cardiac myosin is more stable in PPS. The increased dynamics of the converter thus appears to be a feature of myosin VI.

8.4.2 Further characterization of MyoVI in PPS state

8.4.2.1 *Movements of the subdomains*

Aside from the intrinsic pliancy of the converter, which reflects what was previously observed in Section 8.2.1, we further investigated the reason for the variability in the RMSD of the core atoms of the motor domain in MyoVI. We computed the RMSD of all the subdomains except the converter: Upper50 (U50); Lower50 (L50) and N-terminal (Nter). Figure 8.25 shows that even after excluding the converter from the RMSD calculation, the same four replicas (R1, R4, R5, and R10) deviate more from the minimized structure with respect to the other ones (Figure 8.25, density plot on the bottom).



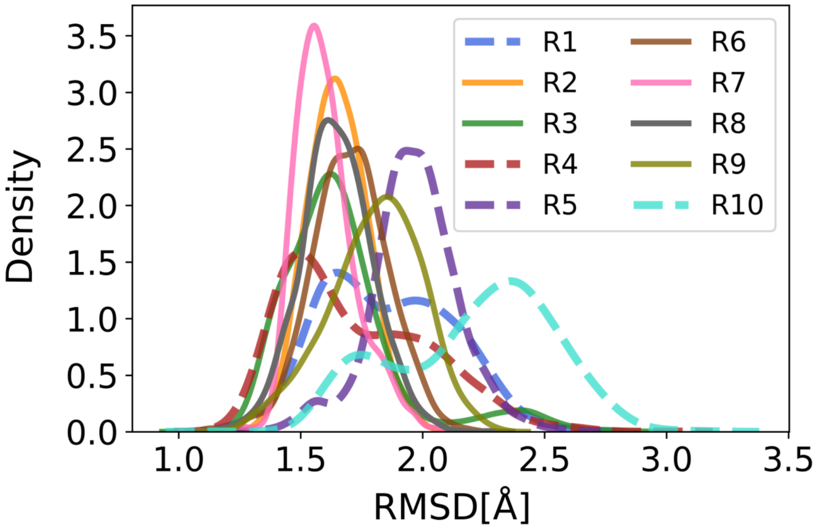


Figure 8.25: RMSD of the core atoms of the motor domain without the converter subdomain. On the top is the time-series over $2\mu s$ and on the right is the distribution of the RMSD per replica. In dashed lines the replicas whose RMSD deviates the most from the reference structure (minimized structure of the model of 2v26.pdb).

To understand the reason behind this divergence we analyzed the behavior of the different subdomains in relation to each other (plots reported in Section A.2) and we found that U50 moves with respect to the N-terminal subdomain in all the replicas whose RMSD deviates the most from the reference structure, often with a bimodal distribution (Figure 8.26).

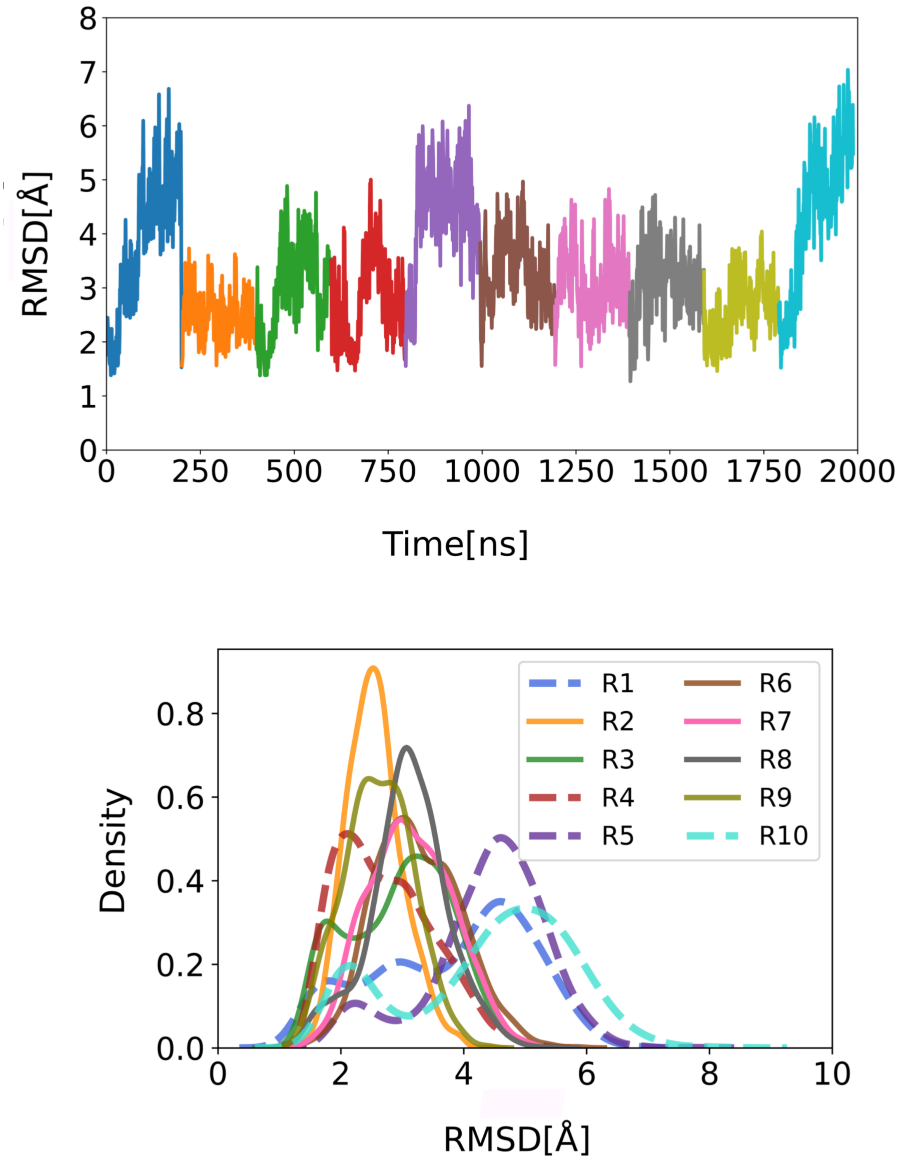
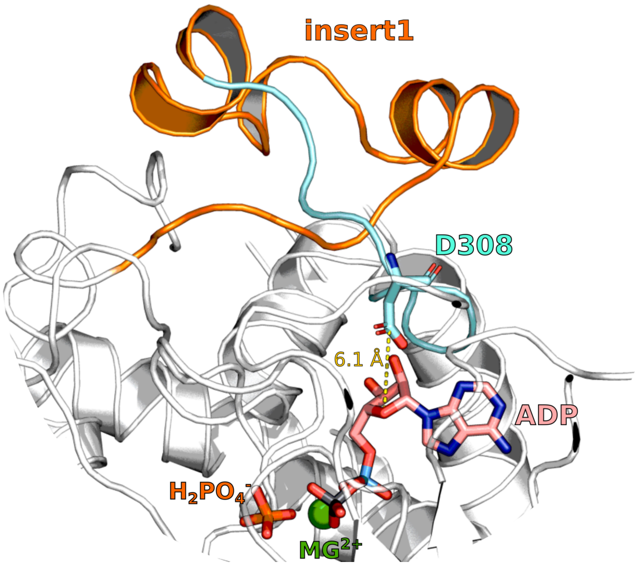


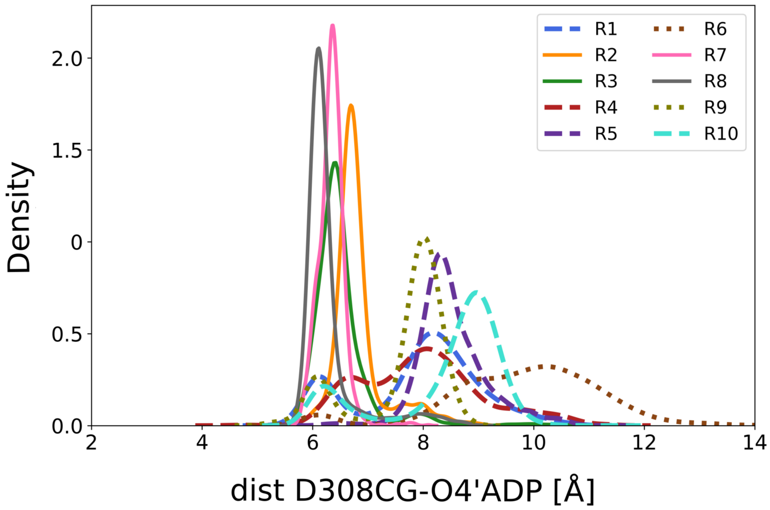
Figure 8.26: RMSD of the core atoms of the U50 subdomain over $2\mu s$, after alignment on the $C\alpha$ of the N-terminal subdomain. On the top the time-series over the whole trajectory and on the bottom the density plot of the same RMSD for each replica. In dashed lines the replicas whose RMSD deviates the most from the reference structure (minimized structure of the model of 2v26.pdb).

We tried to identify a possible "pinning point", which, once broken, could mediate the rotation of U50. By inspecting the trajectories we identified a plausible "pinning point" in a residue (ASP308) belonging to a loop of the U50, in which is inserted the unique insert1 of MyoVI (Pylypenko et al., 2011) (see Figure 8.27 sub-figure (a)). The side chain of this residue seems to interact with the hydroxyl groups of the ADP ribose sugar and, once the interaction breaks, the subdomain is destabilized. After monitoring the distance between the side chain of

ASP308 and the ribose of ADP in the different replicas, we noticed that in the replicas in which the U50 rotates with respect to the N-terminal (Figure 8.26) the above-mentioned distance diverges the most with respect to the reference value (see Figure 8.27 sub-figure (b)). In particular R1, R4, R5, and R10 (represented in dashed lines in Figure 8.27 sub-figure (b)) are the four replicas in which we observed both a deviation of the converter (Figure 8.22) and a rotation of the U50 (Figure 8.26) and the loss of the interaction between D308 and ADP (Figure 8.27 sub-figure (b)). This interaction is also broken in R6 and R9 (Figure 8.27 sub-figure (b)), replicas in which we observed neither full destabilization of the U50 (Figure 8.26) nor different configurations for the converter (Figure 8.22). This probably means that the interaction D308-ADP is not sufficient to destabilize U50 and the converter.



(a) Cartoon representation of the interaction between D308 and ADP in the structure used as reference. In orange the unique insert1 (residues from C278-A303) and cyan the interposing loop (residues G304-D312). In sticks the residue ASP308 and the nucleotide in the active site ($ADP \cdot MG^{2+} \cdot H_2PO_4^-$). In yellow the distance between CGD308-O4′ADP.



(b) Density plot of the distance expressed in Å between $ASP308$ and ADP O4′ in the different replicas of MyoVI

Figure 8.27: ASP308 is a residue of the loop belonging to U50 subdomain in which is interposed the insert1 (Pylypenko et al., 2011). This loop is close to the active site and the interaction of ASP308 with the hydroxyl groups of the ribose in the ADP structure could stabilize the U50, preventing its rotation. The distance considered is the one between ASP308CG and O4′ of ADP (the oxygen of the ribose). This distance is representative of the interaction of the side chain of ASP308 with the two hydroxyl groups of the ribose. The distance in the reference structure (minimized model) is 6.1 Å (see sub-figure (a)). In sub-figure (b) is plotted the density of this distance in the various replicas of PPS MyoVI. In dashed/dotted lines are depicted the replicas (R1, R4, R5, R6, R9, and R10) in which the distance ASP308-O4′ diverges significantly from the reference value.

8.4.2.2 *Hydrogen bonds pattern between β-strands*

Investigating in detail the possible cause of this instability, we observed that the movement of the U50 could be determined by the breaking of the same pattern of fundamental H-bonds connecting the U50 and Nter subdomains analyzed in Section 8.2.1.2. As previously explained (in Section 8.2.1.2), β4, β5, and β6 strands are interacting through H-bonds to keep the active site in a closed configuration in PPS to trap the products of ATP-hydrolysis (Coureux et al., 2004). As shown in Figure 8.28 and Figure 8.29, the H-bonds that break in the simulations in which the RMSD of U50 has a bimodal distribution or a higher value to the N-terminal subdomain (R1, R3, R4, R5, R7, R10) are the ones between L455 (β5)-V149 (β4) and between D456 (β5)-K208 (β6).

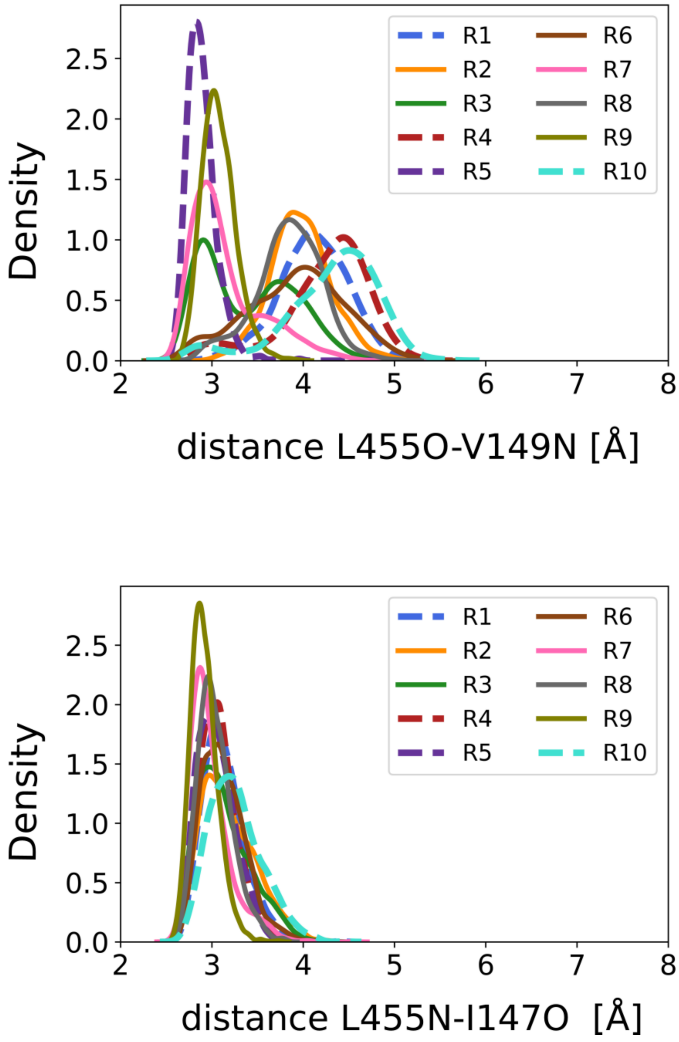


Figure 8.28: H-bonds pattern between β5-β4. L455 is a residue of β5 and V149 and I147 are residues of β4.

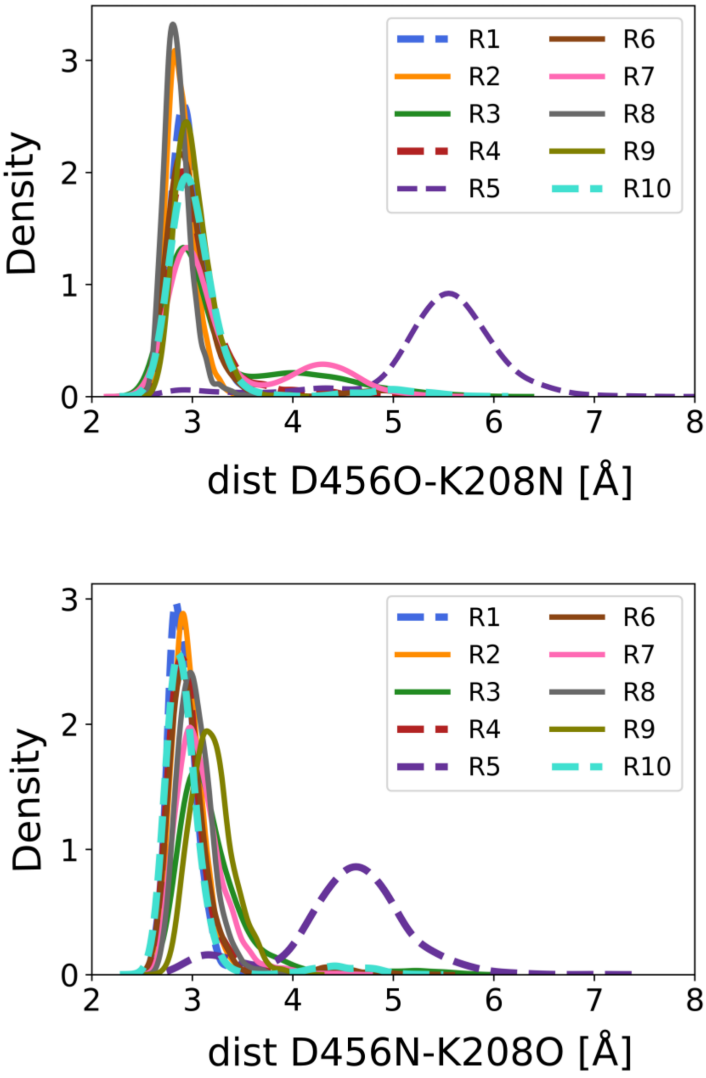


Figure 8.29: H-bonds pattern between $\beta5$-$\beta6$. D456 and K208 are respectively residues of the $\beta5$ and $\beta6$ strands.

Following the loss of these interactions, U50 could uncouple from the N-terminal subdomain, propagating this instability to some key elements of the motor domain such as the switch II (swII) and the Relay helix (RH) with which the U50 relates to the $\beta5$-strand. The RH is part of the L50 subdomain (Ménétrey et al., 2007), and it is connected to the converter with its C-terminal portion. Motion at the level of the U50, RH, and L50 as a result of the break of the above-mentioned hydrogen bonds could explain the higher deviation in the RMSD value of the motor domain previously described (Figure 8.25).

8.4.2.3 *Monitoring geometrical observables over the pseudo-trajectory*

To pursue our characterization of the PPS state in MyoVI, we monitored the stability of the same collective variables analyzed in Section 8.2.1 over the pseudo-trajectory of 2μs. We used the same reference structure as before for the different states: 2BKH for the Rigor (R) state (Ménétrey et al., 2005); 4PFO for the P_i release (P_iR) state (Llinas et al., 2015); 2VAS for the post rigor (PR) state (Ménétrey et al., 2008); 5O2L for the Pre-Transition State (PTS) (Blanc et al., 2018) and 2V26 for the PPS state (Ménétrey et al., 2007). The critical salt bridge between swII (E461) and swI (R205), e is formed with a probability of 23% with a distance between the CD of E461 and the CZ of R205 consistent with that in the PPS X-ray structure (2v26.pdb) of 4 Å. This salt bridge is not stable at the reference distance along the entire trajectory, due to a rotation of the side chain of E461, that breaks this interaction. The distance defined for the second geometrical observable, named "opening of swI", peaks at the reference values of the PPS state (11.8 Å) with respect to the PTS one (14 Å). As mentioned above, myosin in PPS traps in the active site the inorganic phosphate ($H_2PO_4^-$), a product of the ATP-hydrolysis, unlike the PTS state in which the hydrolysis did not occur yet and we still find the ATP in the active site (Blanc et al., 2018). To monitor the trapping of $H_2PO_4^-$ in the active site we inspected its coordination with the magnesium ion (MG). As shown in Panel C of Figure 8.30 the distance is picked at 1.8 Å, equal to the reference value of 1.8 Å (2v26.pdb). This observation suggests that the coordination $H_2PO_4^-$-MG is stable in all PPS replicas despite the breaking of the critical salt bridge. Last but not least, we followed the movement of the converter subdomain, to determine if the three distinct conformations previously described (Figure 8.22) are related to the uncoupling from the N-terminal domain. The average value of the distance defined for this observable over the 2 μs is 6.5 ± 1.4 Å, far from the uncoupled PTS value of 13.2 Å and, instead, close to the reference value of PPS of 5.9 Å (Figure 8.30, Panel D).

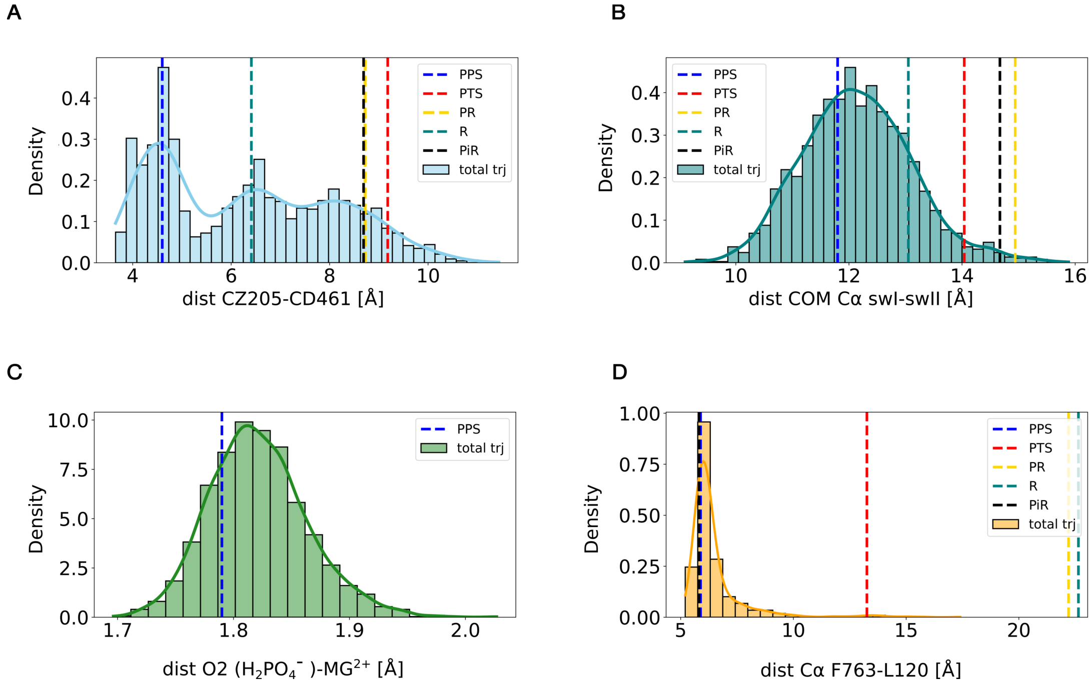


Figure 8.30: Distribution plots of the different observables defined to describe the PPS state of MyoVI in the ten replicas of 200 ns each. Panel **A**: critical salt-bridge between E461CD-R205CZ (Onishi et al., 2006). Panel **B**: distance of the center of mass of the Cα of some residues of swI (residues 197 to 200) and swII (residues 463 to 466). Panel **C**: coordination $H_2PO_4^-$ with MG^{2+}, defined as distance among O2 of $H_2PO_4^-$ and MG^{2+}. Panel **D**: coupling of the converter to the N-terminal subdomain, typical of PPS state.

Nevertheless, we can appreciate small peaks in Figure 8.30, Panel D at values close to 10 and 15 Å. These peaks correspond to three of four replicas in which the converter appeared unstable in Figure 8.22, notably R4, R5, and R10. To verify whether the uncoupling of the converter, which is typical of the PTS state, also occurred in these three replicas, we followed the evolution of the observables defined in Section 8.2.1.6.

8.4.2.4 *Monitoring geometrical observables for three replicas*

We first elongated these three simulations to 300 ns for R5 and R10 and 400 ns for R4 and then we analyzed the sequence of events. Below, the three panels (one for each of the three replications under examination) in which we depicted the time-series of the five observables monitored so far, following the order of the pathway proposed for the previous transition (in Section 8.2.1, Figure 8.17).

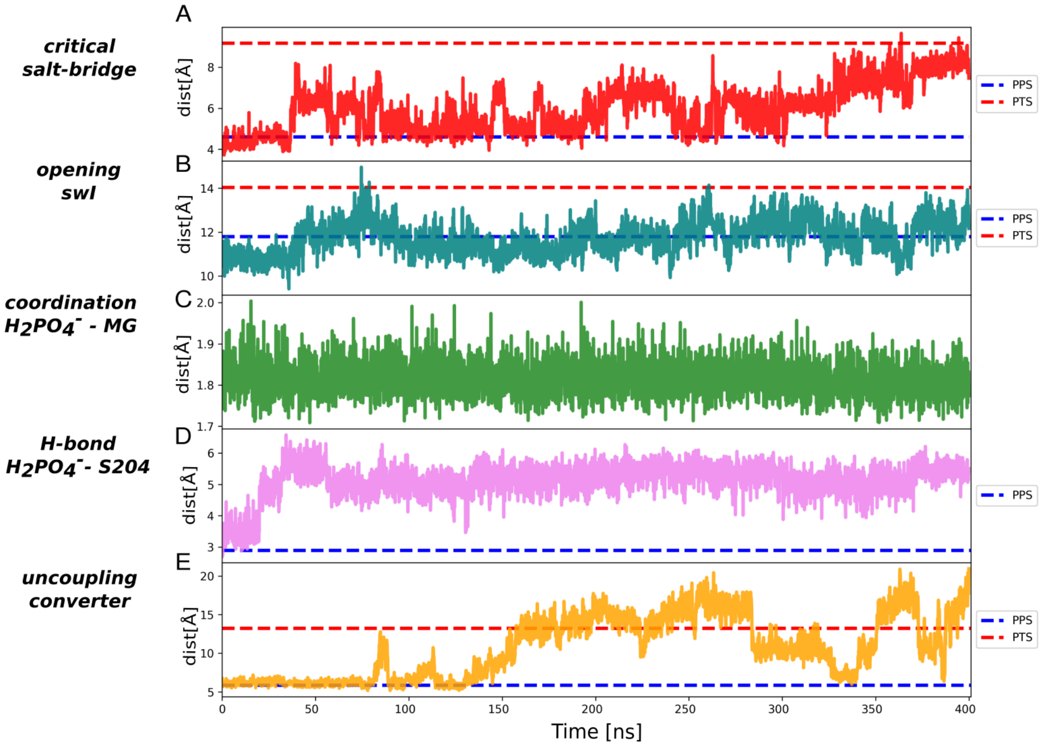


Figure 8.31: Panel representing the possible sequence of events that lead to the uncoupling of the converter from the N-terminal subdomain in R4 along 400 ns MD. Panel A: time-series of the critical salt bridge (CDE461-CZR205). Panel B: time-series of the "opening" of the switches computed as the distance between the center of mass of some residues of swI (residues 197 to 200) and swII (residues 463 to 466). Panel C: coordination between $H_2PO_4^-$ and MG^{2+}, monitored as the distance between O2 of $H_2PO_4^-$ and the ion. Panel D: time-series of the H-bond between O1 of $H_2PO_4^-$ and O of S204 (residue of swI in U50). Panel E: time-series of the uncoupling of the converter from the N-terminal subdomain, monitored as the distance between $C\alpha$ of L120 and F763, residues of N-terminal and converter, respectively.

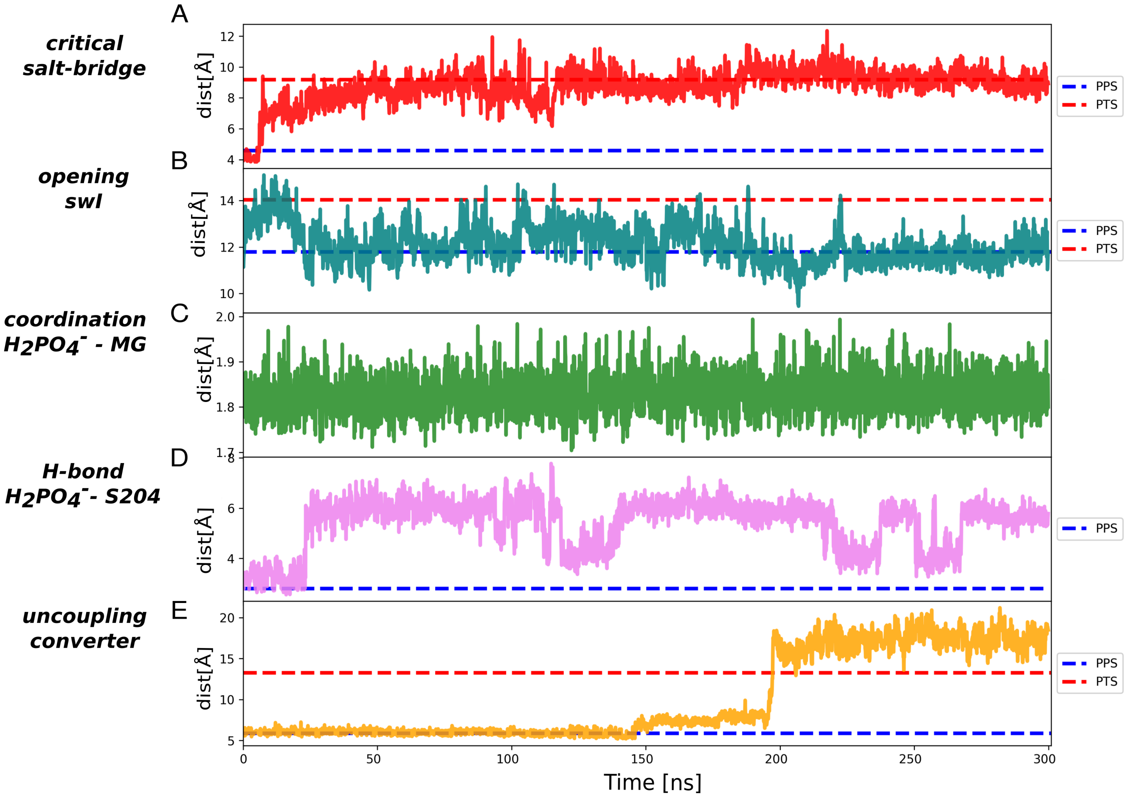


Figure 8.32: Panel representing the possible sequence of events that lead to the uncoupling of the converter from the N-terminal subdomain in R5 over 300 ns MD.Panel A: time-series of the critical salt bridge (CDE461-CZR205). Panel B: time-series of the "opening" of the switches computed as the distance between the center of mass of some residues of swI (residues 197 to 200) and swII (residues 463 to 466). Panel C: coordination between $H_2PO_4^-$ and MG^{2+}, monitored as the distance between O2 of $H_2PO_4^-$ and the ion. Panel D: time-series of the H-bond between O1 of $H_2PO_4^-$ and O of S204 (residue of swI in U50). Panel E: time-series of the uncoupling of the converter from the N-terminal subdomain, monitored as the distance between $C\alpha$ of L120 and F763, residues of N-terminal and converter, respectively.

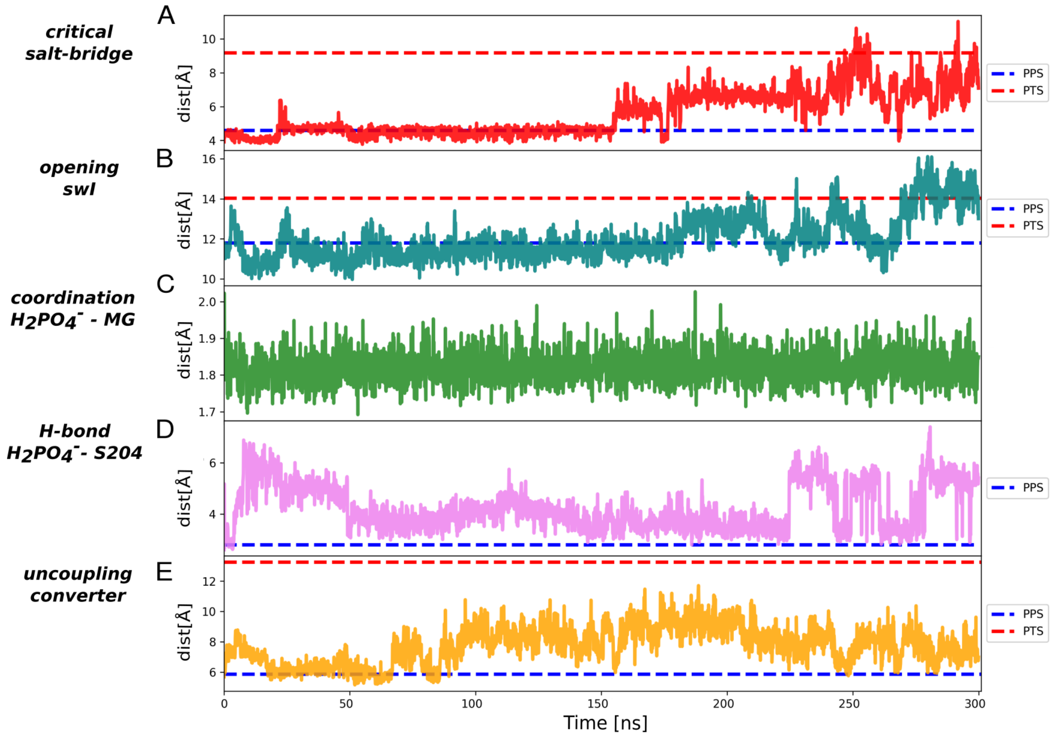


Figure 8.33: Panel representing the possible sequence of events that lead to the uncoupling of the converter from the N-terminal subdomain in R10 over 300 ns MD.Panel A: time-series of the critical salt bridge (CDE461-CZR205). Panel B: time-series of the "opening" of the switches computed as the distance between the center of mass of some residues of swI (residues 197 to 200) and swII (residues 463 to 466). Panel C: coordination between $H_2PO_4^-$ and MG^{2+}, monitored as the distance between O2 of $H_2PO_4^-$ and the ion. Panel D: time-series of the H-bond between O1 of $H_2PO_4^-$ and O of S204 (residue of swI in U50). Panel E: time-series of the uncoupling of the converter from the N-terminal subdomain, monitored as the distance between $C\alpha$ of L120 and F763, residues of N-terminal and converter, respectively.

By inspecting the time-series of Figure 8.31, Figure 8.32 we can appreciate how both in R4 and R5 the breaking of the critical salt bridge (Panel A of the two panels) is the initial event in the sequence of rearrangements that leads to the uncoupling of the converter (following the time-series from A to E). In contrast to these two replicas, in R10, the only one of the three in which the uncoupling of the converter is not reached (Panel C in Figure 8.33), the primary salt bridge results stable for more than half of the simulation time (Figure 8.33, Panel A). In conjunction with the rupture of the critical salt bridge in R4 and R5, the opening of the swI (Panel B in Figure 8.31 and Figure 8.32). The coordination between $H_2PO_4^-$ and MG^{2+} is preserved in all three replicas (Panel C of the three panels) at a constant value of 1.8 Å. Conversely, the H-bond between $H_2PO_4^-$ and S204 (residue of swI) is seldom formed in any of the three replicas (Panel D in Figure 8.31,

Figure 8.32 and Figure 8.33). Lastly, the uncoupling of the converter from the N-terminal subdomain occurs both in R4 and R5, but not in R10 (Panel E in Figure 8.31, Figure 8.32 and Figure 8.33). Since we saw in the previous analysis (Section 8.2.1 Figure 8.15) that the uncoupling of the converter correlates with the loss of hydrophobic interaction with the N-terminal subdomain, we computed this correlation also in R4 and R5, replicas in which the converter uncouples from the N-terminal subdomain.

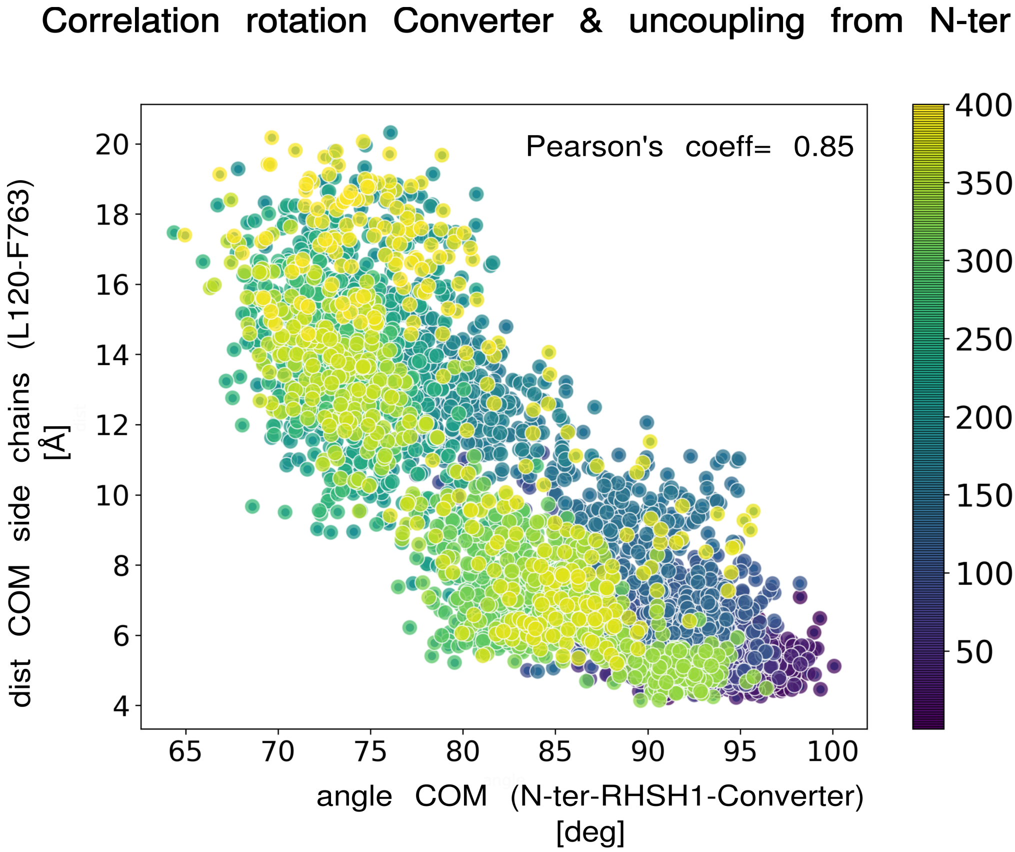


Figure 8.34: Correlation between the initial rotation of the converter and the uncoupling from the N-terminal subdomain along the MD. The dots in the plot are colored according to the color map on the right side, which indicates the trajectory of 200 ns. On the x-axis is plotted the angle formed between the center of mass of Cα atoms of some residues of the N-terminal subdomain (residues 128 to 140), of RH-SH1, here considered merged into a single element, (residues 489 to 500 and residues 700 to 705) and of the converter (residues 750 to 770). On the y-axis the distance of the center of mass of the side chains of L120 (residue of N-terminal) and F763 (residues of the converter) which establish an initial hydrophobic interaction that then breaks simultaneously with the uncoupling.

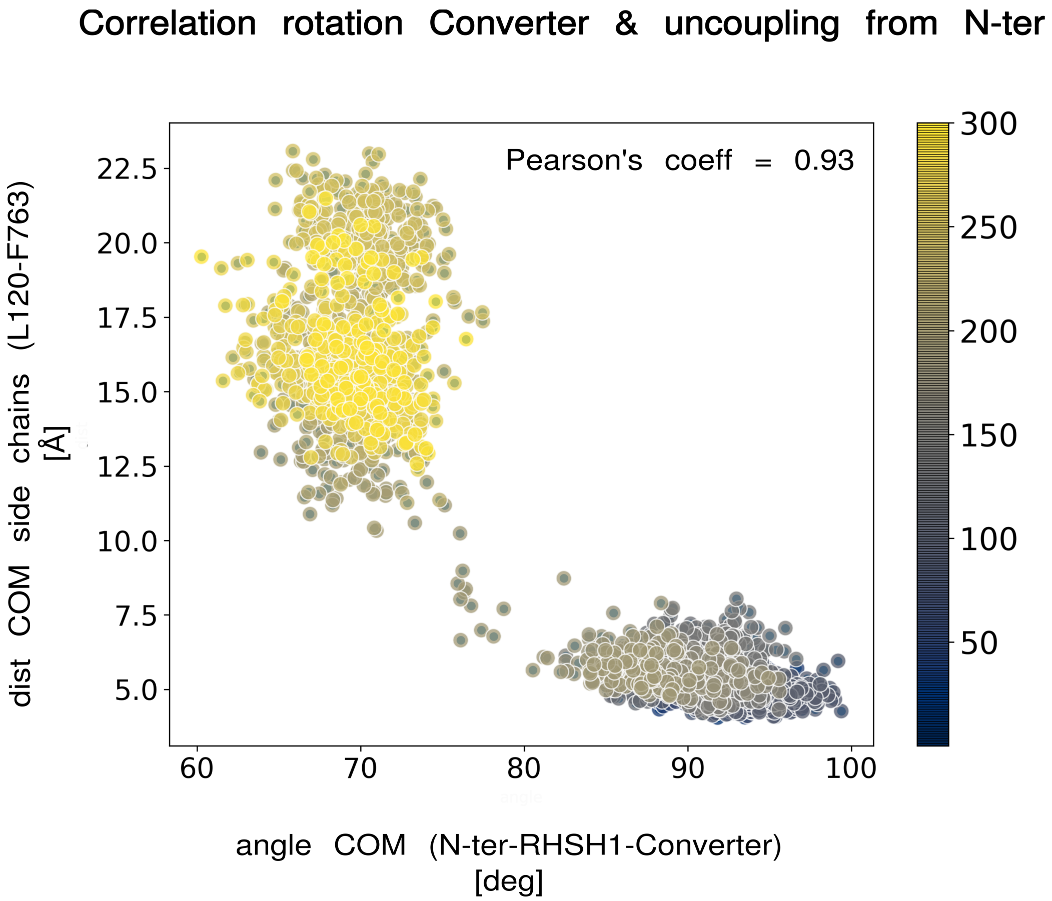


Figure 8.35: Correlation between the initial rotation of the converter and the uncoupling from the N-terminal subdomain along the MD. The dots in the plot are colored according to the color map on the right side, which indicates the trajectory of 200 ns. On the x-axis is plotted the angle formed between the center of mass of $C\alpha$ atoms of some residues of the N-terminal subdomain (residues 128 to 140), of RH-SH1, here considered merged into a single element, (residues 489 to 500 and residues 700 to 705) and of the converter (residues 750 to 770). On the y-axis the distance of the center of mass of the side chains of L120 (residue of N-terminal) and F763 (residues of the converter) which establish an initial hydrophobic interaction that then breaks simultaneously with the uncoupling.

From both Figure 8.34 and Figure 8.35 we can surmise that also in these trajectories the uncoupling of the converter from N-terminal subdomain strongly correlates with an initial rotation of the converter, triggered by thermal fluctuations of RH and SH1. Also, in this case, the uncoupling of the converter is highlighted by the loss of a hydrophobic interaction at the interface between the two subdomains.

As already mentioned, the transducer plays an important role in the transition of the motor domain towards the different states of the actomyosin cycle (Coureux et al., 2004; Ménétrey et al., 2008; Yang et al., 2007), thus we run the same analysis as before (Section 8.2.1, Figure 8.18) to study the twisting of the β-sheets of the transducer (Llinas et al., 2015). The results of this analysis, both for the trajectory of $200ns$ of the initial model of MyoVI in PPS with standard protonation state of some titratable residues and the concatenated ten trajectories of 200*ns* described so far, reveal that the transducer twists towards a post-rigor like (PR) or amidst PR and Pi-release (P_iR) conformation (see Figure 8.36) in both cases.

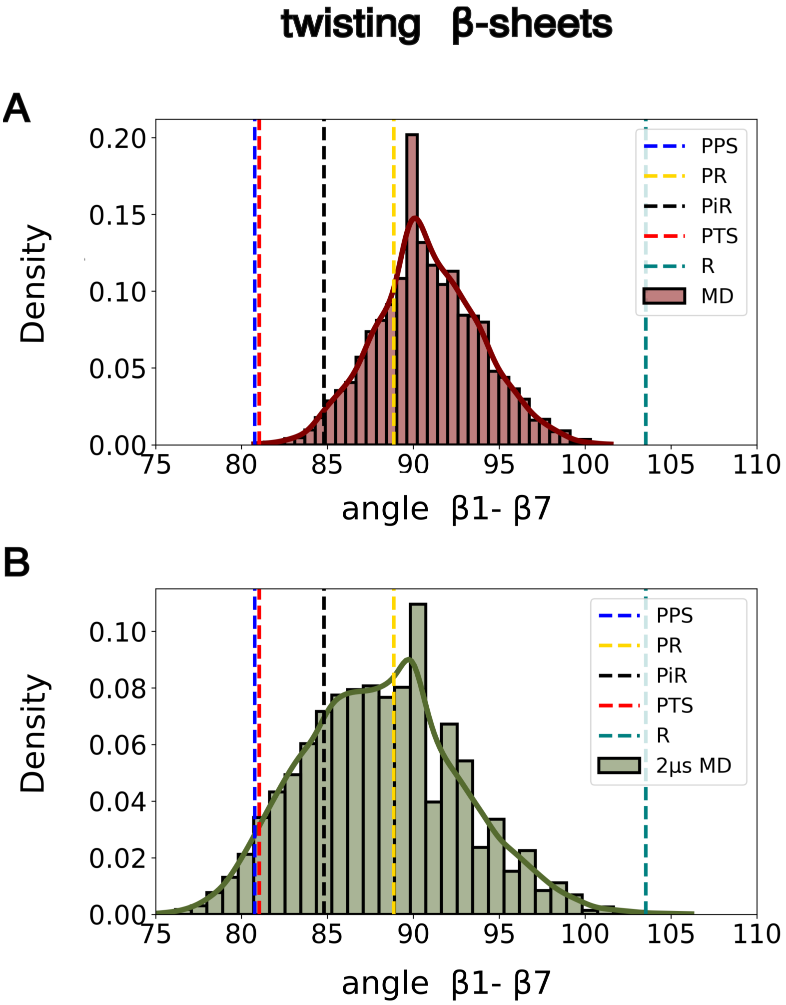


Figure 8.36: Distribution of the twisting of the β-sheets. Panel A: distribution of the angle computed for the 200 ns MD of the model of MyoVI in PPS with a standard protonation state for the titratable residues. Panel B: distribution over the 2 μs trajectory of the initial PPS model. The twisting was computed over the catted trajectory from the ten replicas.

Additional analyses of the MD simulation described in this chapter and the definition of the core atoms of all the subdomains are collected in Appendix A.

8.5 CONCLUSIONS & FUTURE PERSPECTIVES

Overall, we could conclude that our 2 μs cumulated simulation time represents on average a PPS state of MyoVI and, although there are fluctuations in some of the observables we chose (see the case of the primary salt bridge) the analysis is consistent with the flexible nature of the PPS state (Fisher et al., 1995; Gourinath et al., 2003; Houdusse et al., 2000). Perhaps the fluctuations are related to an intrinsic conformational variability distinctive for MyoVI and maybe for processive myosins if compared with non-processive ones (like we did with cardiac myosin in this chapter). Moreover, we can assert that the possible sequence of events that leads to a transition toward a PTS-like state, proposed in the previous Section 8.2.1, is reproducible both in R4 and R5, two of the ten replicas in which the breaking of the critical salt bridge triggers the final uncoupling of the converter. A very interesting counter result in this analysis, if compared with the previous one, is the preservation of the coordination between $H_2PO_4^-$ and MG^{2+}. We previously supposed that the loss of this coordination would have contributed to communicating further instability from the nucleotide binding site to the force-generating element (RH, SH1, and converter), as also hypothesized by Mugnai *et al* (Mugnai and Thirumalai, 2021a). Our results for R4 and R5, instead, suggest that the uncoupling of the converter, driven by thermal fluctuation of RH, SH1, and switch loops would occur even in presence of coordination between inorganic phosphate and magnesium. Thus, we infer that the different parameters used for the phosphate in these simulations could play an important role in the stability of the interactions established by the ligands trapped in the nucleotide binding pocket. Another finding is the independence of the protonation state chosen for some titratable residues in inducing major rearrangements, such as the uncoupling of the converter from the N-terminal subdomain, in the transition to a different intermediate state (PTS-like). Regarding the distortion of the transducer, as previously speculated, the twisting of the β-sheets towards a PR or P_iR conformation in the vast majority of our simulation could be the attempt of the motor to switch, through distortion of the transducer, toward lower free-energy states (PR or P_iR) (Karatzaferi et al., 2004; Yang et al., 2007), without an unnecessary release of hydrolysis products from the active site. A more practical explanation might lie in the choice of $C\alpha$ atoms considered to describe the vectors representing the transducer. With a different selection, the final distribution may change. A potential future study could be the comparison of the behavior of various processive myosins (i.e myosin I, V, and X), versus non-processive ones (different myosins being class II myosins) to delineate specific and/or common properties in the two groups. Furthermore, as already mentioned, many factors could play an important role in the stability of the system, such as the forcefield used, the parametrization of the ligands, as well as the initial model selected. Our study could open a new understanding of the stability of the transition state in MyoVI, paving the way for further analyses.

9

MYOSIN VI: CLINICAL RELEVANCE & RATIONAL DRUG DISCOVERY

9.1 CLINICAL RELEVANCE OF MYOSIN VI

Myosin VI plays a role in endocytotic processes, as well as in the preservation of Golgi's morphology and integrity of stereocilia of the inner ear (Jung et al., 2006). In mammalian cells it also controls the adherence at the junction of epithelial cells (Sweeney et al., 2007) and it has been demonstrated that it is highly expressed in the nucleus, modulating the RNA polymerase II transcriptional activity (Vreugde et al., 2006). In humans, overexpression of Myosin VI is observed in breast and medium-grade prostate cancer as well as in ovarian cancer and for the latter, it has also been found to correlate with the malignancy of the disease (Dunn et al., 2006; Hari-Gupta et al., 2022; Makowska et al., 2015; Vreugde et al., 2006). A significantly reduced expression of Myosin VI has been detected in deaf mice, in which the integrity of the inner ear hair was not preserved, due to a possible lack of cytoplasmic components that leads to disruption of membrane recycling (Avraham et al., 1995). More recently some mutations at the level of the motor domain of Myosin VI (C442Y, H246R, E216V, and D179Y) have been associated with the deregulation of endocytosis processes, resulting in disruption of stereocilia structural integrity (Ahmed et al., 2003; Melchionda et al., 2001; Pylypenko et al., 2015).

Due to the ample range of functions in which it is involved, Myosin VI represents a very important therapeutic target (Preller et al., 2011), particularly to treat human cancers (Buss and Kendrick-Jones, 2008; Naydenov et al., 2021). According to World Health Organization (WHO), breast and prostate cancers have the highest incidence rate worldwide (58.5% and 36%, respectively.
See `https://gco.iarc.fr/today/online-analysis-multi-bars`). Beyond that, recent studies (Lukasiewicz et al., 2021; Rahib et al., 2021) predict that both the incidence and mortality of breast cancer will continue to worsen. It, therefore, becomes pivotal to develop selective inhibitors, able to target Myosin VI.

9.2 ALLOSTERIC INHIBITION OF MYOSIN VI

Given the social impact linked to diseases related to the overactivity of Myosin VI and the benefits that can be derived from the development of novel selective inhibitors, we conducted a vHTS campaign on Myosin VI (MyoVI). The small molecules eligible as promising lead compounds should be allosteric inhibitors, viz. non-competitive inhibitors. Why is this important? It is critical to target myosins on their allosteric pockets rather than on their active site because potent competitive inhibitors would prevent the binding of ATP and its hydrolysis, leading to a permanent rigor state of myosin (strongly bound to actin) (Bond et al., 2013). Competitive inhibitors, structurally similar to ATP, would not be selective for myosins and could target other ATP-ases (Bond et al., 2013). Moreover, myosin's active site is highly conserved among different classes (Smith and Rayment, 1996), so choosing competitive (or orthosteric) ligands would induce a loss of selectivity for the class of myosin to be targeted. Allosteric modulators, instead, would be more selective by binding pockets that do not arise from evolutionary selection. For these reasons, our study was based on the hunt for possible allosteric inhibitors of MyoVI.

9.3 VHTS ON MYOSIN VI

9.3.1 What protein structure to use?

Our purpose is to search for (and perhaps find) allosteric inhibitor(s) of MyoVI, as we would like to suppress the overactivity unveiled in human pathological conditions, such as cancer. For this to be achieved, we need to block the force generation cycle of MyoVI, therefore we decided to target the PPS sate (Ménétrey et al., 2005). This is the state of the myosin cycle that precedes the conformation tightly bound to actin, able to initiate the powerstroke. As initial structure we selected 2v26.pdb (Ménétrey et al., 2005), PPS state of MyoVI with ADP-VO_4 and MG^{2+} trapped in the active site (Figure 9.1). This is the MyoVI S1 expressed in *Sus Scrofa* (`https://www.uniprot.org/uniprot/Q29122`) that shares 96.7% of sequence identity with the human MyoVI's motor domain (data obtained running Basic Local Alignment Search Tool (BLAST) (Madden, 2002) in UniProt. (See this link to the web page with the result).

9.3.1.1 *Protein preparation*

The missing portions in the protein structure (residues: 1 to 4; 34 to 37; 174 to 179; 396 to 406; 622 to 637) were modeled using MODELLER (Webb and Sali, 2016) and the best model was chosen according to the lowest Discrete Optimized Protein Energy (DOPE) score. The structure was then uploaded on the MolProbity webserver (`http://molprobity.manchester.ac.uk/`) to add missing hydrogens, optimize H-bond networks and detect the more reasonable rotameric states of the side chains of Asn, Gln and His. The topology and parameter files for the ligand (ADP, P_i and MG^{2+}) were obtained from CGenff (`https://cgenff.umaryland.edu/`) (Vanommeslaeghe et al., 2012) and the entire system was modeled with CHARMM

Figure 9.1: Cartoon representation of 2v26.pdb. The active site is depicted with a green sphere the MG^{2+} and in cyan sticks the ADP-VO_4. In orange the structure of the motor domain of MyoVI in the PPS (also known as post-transition) state.

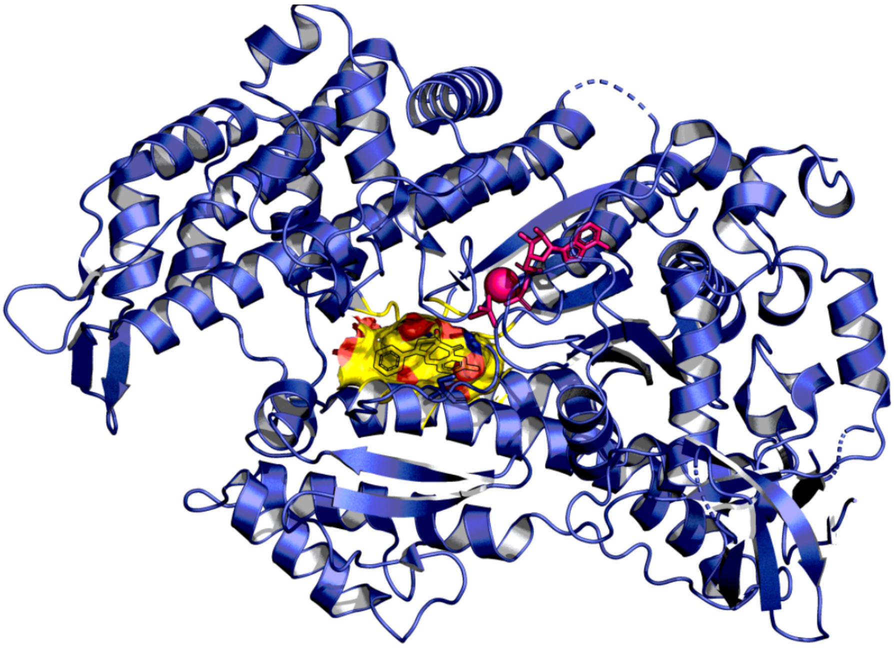
(Chemistry at HARvard Molecular Mechanics) (Brooks et al., 2009). The most probable protonation state of titratable residues at neutral pH was predicted by Karlsberg (Rabenstein, 2000). CHARMM-GUI (Jo et al., 2008) was used to solvate and ionize the system that was then minimized and equilibrated with Gromacs2021 (Abraham et al., 2015; Lindahl et al., 2001; Markidis and Laure, 2015) (`https://zenodo.org/record/4457591#.YrHG43hBxH4`).

9.3.2 What pocket to target?

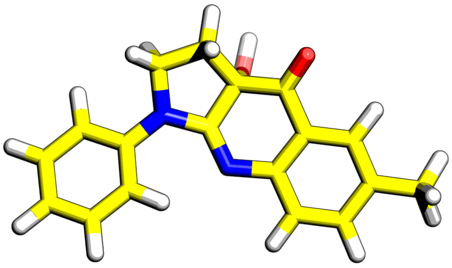
To the best of our knowledge, there are no allosteric inhibitors of MyoVI employed for the treatment of human (breast, prostate, and ovarian) cancers. In 2012 *Hessler et al.,* found a selective inhibitor of human MyoVI, named 2,4,6-triiodophenol (TIP) and classified as a non-competitive inhibitor (Heissler et al., 2012). Actin-activated ATP-ase assays shown that TIP, also known as Bobel-24 or AM24, inhibits MyoVI dependent secretion processes with an IC_{50} of 1.6 ±

0.6 μM (Heissler et al., 2012). Moreover, with the same measurements, Heissler *et al.,* demonstrated that TIP has a "biphasic behavior" in MyoVI (Heissler et al., 2012). This result suggested that TIP may bind to two different binding sites, contributing to 37% or 63% of inhibition of actin-activated ATP-ase activity (Heissler et al., 2012). This is the only currently documented inhibitor of MyoVI, but there has never been conclusive evidence regarding the effective binding site targeted by this molecule. One of the plausible binding sites proposed by Heissler and co-workers (Heissler et al., 2012) is a hydrophobic pocket at the apex of the 50kDa cleft, known as Blebbistatin pocket, from the name of a non-muscle myosin II inhibitor (Straight et al., 2003). Blebbistatin is a small molecule ((S)-blebbistatin (–)-1-phenyl-1,2,3,4-tetrahydro-4-hydroxypyrrolo[2,3-b]-7-methylquinolin-4-one) that has been co-crystallized in *Dictyostelium discoideum* (Dicty) non-muscle myosin II S1 (Allingham et al., 2005) (Figure 9.2) and whose name derives from its ability to arrest cell blebbing (Straight et al., 2003).

Blebbistatin stabilizes the transition state of non-muscle myosin II (PPS state) by blocking the release of inorganic phosphate from the active site, thereby preventing the initiation of the force generation cycle (Allingham et al., 2005). By binding at the apex of the 50kDa cleft, Blebbistatin impedes its closure and the cleft remains partially open. This incomplete closure of the cleft gives long life to the PPS state of the motor domain, delaying the transition to the Rigor state when the cleft is completely closed and myosin is strongly bound to actin. Blebbistatin establishes two H-bonds in the binding site, respectively with the amine of SER456 and the one of GLY240, otherwise, hydrophobic interactions prevail (see Figure 9.3 and Figure 9.4).



(a) 1yv3.pdb



(b) 3D structure of Blebbistatin

Figure 9.2: In (a) the X-ray structure of Dicty Myosin II (PDB code: 1yv3.pdb). In magenta, the ADP-VO4 and MG^{2+} are trapped in the active site. In yellow Blebbistatin and its binding pocket's surface. In (b) the 3D structure of Blebbistatin.

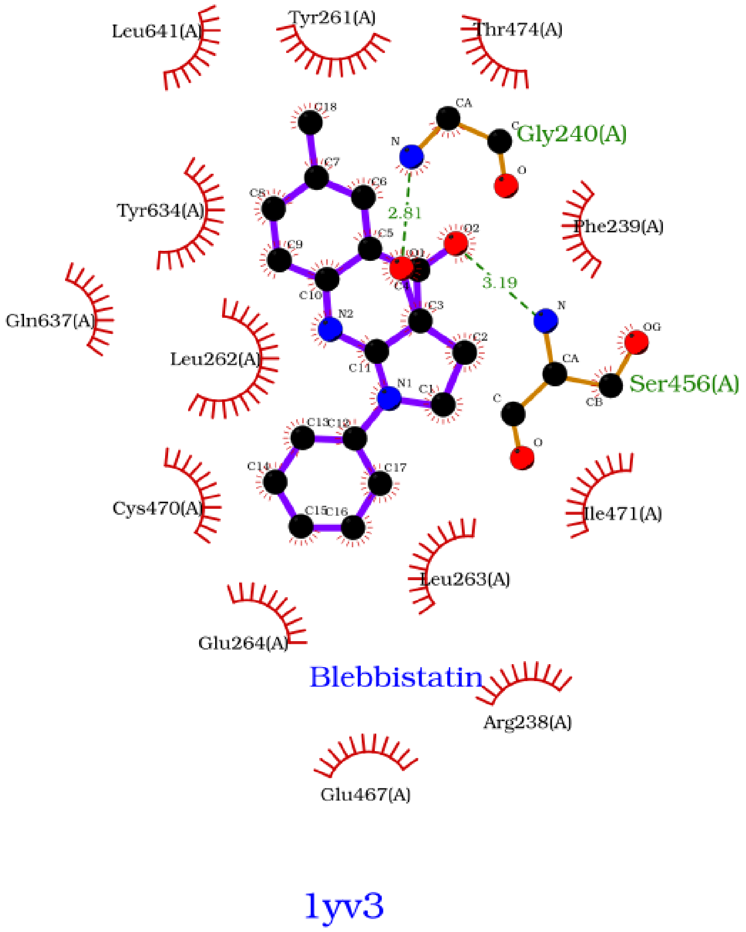


Figure 9.3: 2D scheme of Blebbistatin interactions in the binding pocket. The image has been produced by using LigPlus (Laskowski and Swindells, 2011). The red rays represent the hydrophobic interactions, instead, the dashed green lines indicate the hydrogen bonds. In blue, the PDB code 1yv3 and the name of the small molecule are depicted in violet in the middle (Blebbistatin). "(A)" stands for the name of the chain.

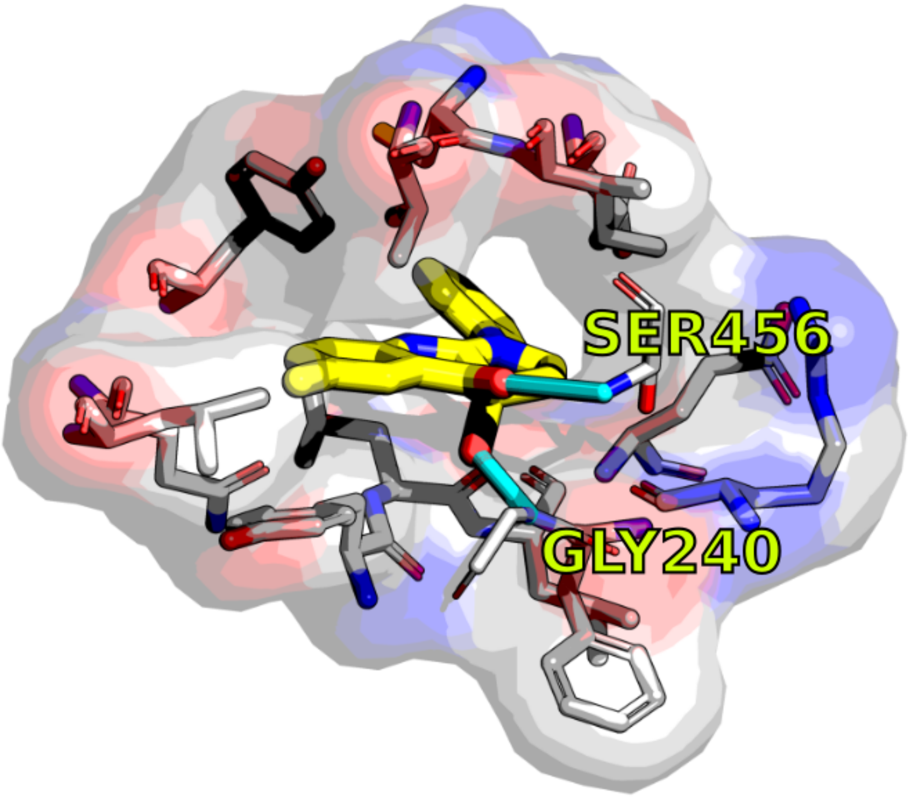


Figure 9.4: 3D representation of Blebbistatin in the binding pocket. In cyan are depicted the two hydrogen bonds formed with S456 and G240 and Blebbistatin is in yellow sticks. The surface of the binding site is colored coded from red to blue based on the polarity of the atoms.

9.3.2.1 *Blebbistatin pocket in MyoVI*

Blebbistatin pocket is a druggable pocket in non-muscle myosin II, as demonstrated by the co-crystallization of the small molecule in the motor domain (Allingham et al., 2005). Furthermore, since there is evidence that this pocket also exists in MyoVI in the PPS state, as shown by Heissler and co-workers (Heissler et al., 2012) and that the motor domains of *Dicty* myosin II and *Sus scrofa* MyoVI share 39% of sequence identity (data computed by using HHPRED (Söding et al., 2005)), we have chosen this as a pocket to target. Moreover, being a pocket at the interface between Relay Helix, L50, and U50 subdomains, which are highly conserved portions of the motor domain, we expected high sequence identity of the binding site between MyoVI and *Dicty* myosin II. Considering the residues within 5 Å from Blebbistatin, only three differ (Figure 9.5): S456-A458; T474-C476; Y634-F647 (on the left the residues name and number in *Dicty* myosin II and on the right the corresponding ones in *Sus scrofa* MyoVI). After aligning *Sus scrofa*

MyoVI (2v26.pdb) over *Dicty* myosin II (1yv3.pdb), the RMSD of the residues in the Blebbistatin pocket was 1.96 Å, considering also the side chains. All these data together strengthen our decision to target this specific pocket in MyoVI.

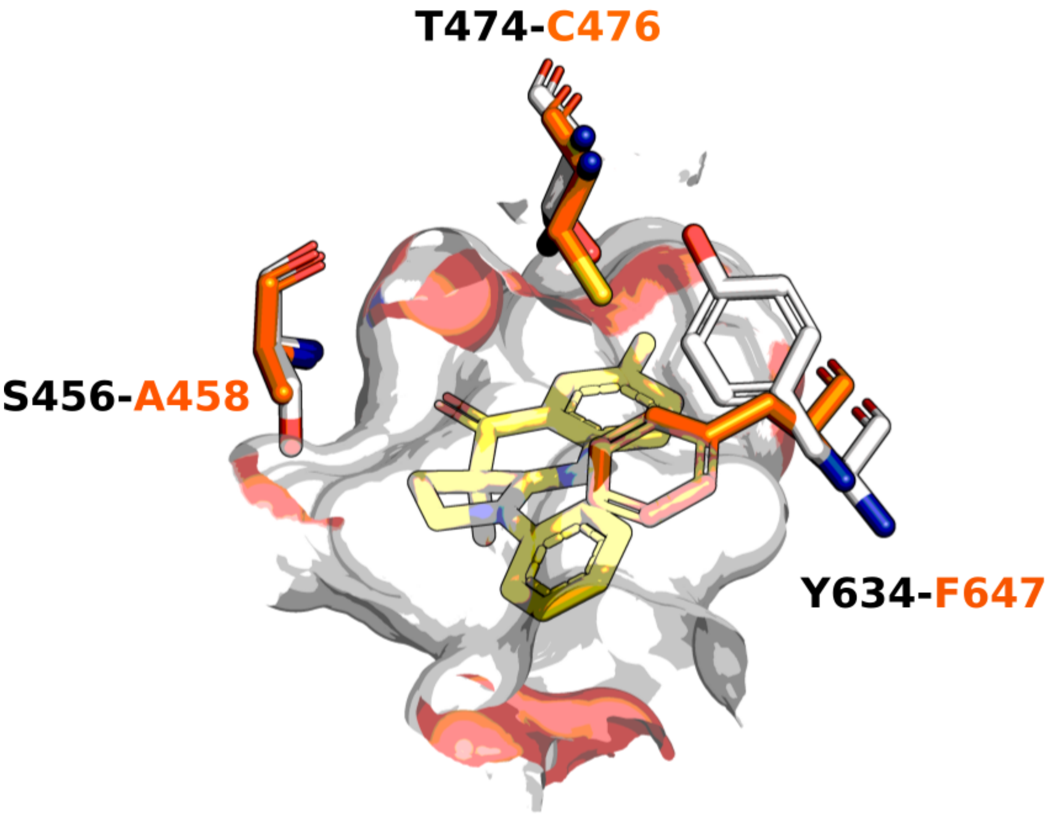


Figure 9.5: Zoom on the Blebbistatin pocket in *Sus scrofa* MyoVI (2v26.pdb), after alignment on *Dicty* myosin II(1yv3.pdb). The RMSD between the two structures is 1.04 Å. In sticks the three residues that differ in MyoVI at the level of this binding site. In orange sticks, the residues in MyoVI and white sticks are the ones in *Dicty* myosin II. F647 assumes a different rotameric state with respect to Y634 in *Dicty* myosin II. This difference creates a new space in the pocket that can be explored. In white the surface of the binding pocket in *Dicty* myosin II and Blebbistatin is in yellow sticks.

9.4 VHTS PROTOCOL

Once the structure (model of 2v26.pdb minimized with CHARMM36mff (Huang et al., 2016)) and the pocket (Blebbistatin binding site) have been selected, all that remains is to select the library of compounds to be screened and the strategy to be adopted.

9.4.1 Chemical library to screen

Regarding the library of compounds to screen, we chose Maybridge, a library containing more than 53K lead-like molecules in stock for efficient order and testing (Arnaud, 2006). These molecules have been previously prepared by using *PrepFlow* (Sisquellas and Cecchini, 2021), to obtain 3D structures of the most probable conformers at neutral pH. After this step of ligand preparation, around 69K structures were generated. Since these 3D structures have been taken from the common virtual archive of compounds existing in our laboratory, we proceeded with a filtering step, to filter out molecules with a molecular weight larger than 500 Da and with more than 10 rotatable bonds. This selection has been performed through Knime (Berthold et al., 2007). The decision to keep only molecules with less than ten rotatable bonds was dictated by a limitation in the use of QHMB (Pereira and Cecchini, 2021) for the estimation of the entropy correction in the final step of prediction of free energy of binding. In addition, a number of ten or fewer rotatable bonds also matches the Veber rules for drug-likeness (Veber et al., 2002). The final number of 3D structures to screen was 66885.

9.4.2 vHTS Strategy

To perform molecular docking we have used our *in house* workflow *DockFlow* and for the free energy rescoring with MMGBSA method a fast procedure involving Gromacs and AmberTools, that will be implemented in the upcoming version of *ChemFlow*. In particular, the docking was run with AutoDock VINA on the supercomputer Jean Zay, by defining as searching space a box with the following coordinates for the center (all dimensions are in Å): x = 28.120; y = -6.105 and z = 4.002. The size of the box was set to 15 Å in the three dimensions. The screening was parallelized, running 50 docking per each slurm job, on a single CPU-node on IDRIS (Institut du développement et des ressources en informatique scientifique : `http://www.idris.fr/`). The docking results were then post-processed to keep only the best pose for each compound (namely the one ranked first by the VINA scoring function). The compounds were ranked according to their docking score and the first 10K were further analyzed. A minimization of maximum 50000 steps was performed with Gromcas2020.3 to then proceed with a conversion of the structure and topology files (*.gro* and *.top* extension) to Amber format (*.prmtop* extension), carried out with Gromber (`https://parmed.github.io/ParmEd/html/parmed.html`). An initial single point MMGBSA (with GB model 7 and internal dielectric constant 4 (Case et al., 2021)) rescoring was performed over the ten thousand minimized structures with AmberTools2021 and the compounds were newly ranked, according to their MMGBSA score. The first thousand complexes were selected from this novel classification and MD simulations of 10 ns, after a short equilibration of 100 ps, were performed with Gromacs2020.3. A second conversion and free energy rescoring with the MMGBSA method was run over these trajectories and the best 80 systems were chosen for the subsequent configurational entropy correction, performed with QHMB (Pereira and Cecchini, 2021) always on the supercomputer Jean Zay. The QHMB approach includes an MD simulation of at least 300 ns of the ligand free in solution and a simulation of at least 50 ns for the protein-ligand complex. We performed the MD simulations of the unbound states on ROMEO Super Computer Center (`https://romeo.univ-reims.fr/pages/aboutUs`) using Gromacs2020.4, whereas for the bound state we elongated the previous MD of 10 ns, for the chosen 80 systems. Unfortunately, for ten compounds QHMB failed for the unbound state, probably due to the high degree of flexibility of these molecules, such that the sampling was not sufficient to populate the clusters which are parsed by the algorithm. We applied the same workflow both for Blebbistatin and TIP. For TIP we did not perform configurational entropy correction, because the molecule does not have any rotatable bond. Moreover, looking at the MD simulations of 50 ns of the bound state, we observed that TIP is not stably bound in the pocket. The molecule is too small and explores different possible cavities. It could be considered as a fragment for the identification of other possible cavities. However, we will not include TIP in our results.

In Figure 9.6 a summary of the strategy adopted in this study.

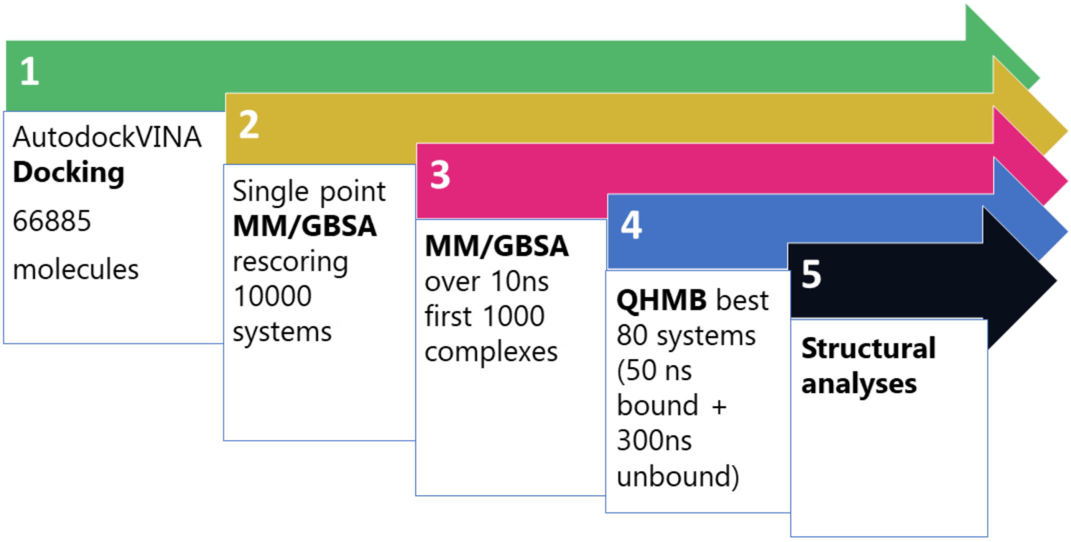


Figure 9.6: Main steps of the vHTS strategy adopted for our case study. The average size of the systems was about 130 thousand atoms.

9.4.3 Structural Analyses & final predictions

In the end, we proceeded with some structural analyses.
One might wonder why we need to conduct structural analyses and go beyond the values of predicted free energy of binding. We previously discussed in Chapter 8 the flexible behavior of MyoVI in PPS and in Chapter 6 the impact of an allosteric modulator on the stability of cardiac myosin in PPS. By analogy, we believe in the importance of an allosteric modulator to be able to stabilize the PPS structure of MyoVI to preserve the functionality of the motor.

9.4.3.1 *RMSD binding pocket*

To limit possible induced fit of the side chains of the targeted pocket, which comes with a free energy cost of the protein that is not accounted for in our calculations, we defined as first observable the RMSD of the binding pocket. To perform these structural analyses we clustered the trajectories of 50 ns of the protein-ligand complexes and we extracted the most populated cluster centers. We then calculated the RMSD of the residues describing the binding pocket, including the side chains, to the initial structure used for the vHTS (the minimized model of 2v26.pdb). We performed these structural analyses for all selected complexes and additionally for Blebbistatin. Even if Blebbistatin could not fit the pocket in MyoVI as it does in *Dicty* myosin II because of severe clashes with F647 (see Figure 9.5), we have forcibly introduced it in our campaign, to exploit it as a negative control, since it experimentally shows lack of activity in MyoVI (personal communication). Our *in silico* results are in agreement with the experimental ones since even with our approach Blebbistatin was predicted as one of the worst binders for MyoVI in the targeted pocket (positive docking score). Nevertheless, we computed also for Blebbistatin a final value of binding free energy, including configurational entropy correction and we obtained a prediction of $-43.97 kcal/mol$. This value will later be used for a final selection of hit compounds. We arbitrarily established a cut-off of 3 Å for the RMSD of the pocket. Ranking the molecules according to this cut-off, a total of 53 compounds bound to the pocket with an RMSD lower than 3 Å. After visual inspection of the cluster centers of the systems, we noticed that the rotameric states of some residues were different from the reference structure. This would have changed the space explored by the molecules in the pocket.

9.4.3.2 *$\chi 1$ dihedral angle F647*

F647 is one of the three residues that differ in the binding pocket to *Dicty* myosin II (see Section 9.3.2). In particular, it assumes a completely different rotameric state in MyoVI, if compared to the Y634 in non-muscle myosin II (see i.e. Figure 9.5). This orientation opens a

different space in the pocket that could be explored, giving the possibility of finding interesting new scaffolds selective for MyoVI. We, therefore, monitored a new observable in the various cluster centers, Blebbistatin included: the dihedral angle $\chi1$ of F647, defined as the angle described by the atoms N-CA-CB-CG.
We compared the value of the angle of the cluster centers with the X-ray structure of porcine MyoVI (2v26.pdb).

As mentioned before and as shown in Figure 9.7, several compounds (53 to be exact) stabilize the pocket keeping its average RMSD value below 3 Å along the 50 ns trajectory. However, very few molecules are keeping both the RMSD of the pocket underneath 3 Å and the $\chi1$ angle value of F647 close to the reference value of the X-ray structure of 272.3° (Figure 9.7).

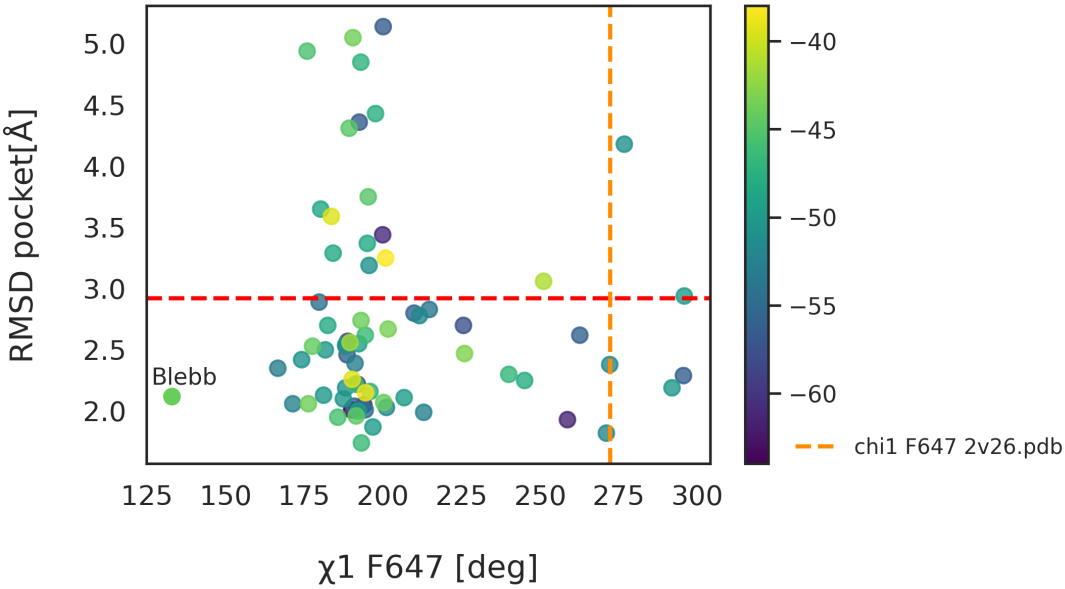


Figure 9.7: Results of the structural analyses conducted on the data of the vHTS on MyoVI in PPS. The dots represent the compounds extracted from the final ranking of MMGBSA, for which the entropic correction was calculated by applying QHMB (Pereira and Cecchini, 2021). Their color is assigned according to the color map shown on the right of the plot, per their final ΔG of binding. On the x-axis is plotted the value of the dihedral angle $\chi1$ of F647 (N-CA-CB-CG) for each system, computed by using an in-house script of Wordom (Seeber et al., 2007). The y-axis reports the RMSD value of the binding pocket, including the side chains of residues within 5 Å of Blebbistatin. The label "Blebb" refers to Blebbistatin. The red dotted lines are the arbitrary cut-offs selected for the RMSD of the targeted pocket. In orange dotted line the $\chi1$ value of F647 in the X-ray structure of porcine MyoVI (272.3°).

In Figure 9.7 we also noticed that in several cluster centers the value of this dihedral angle was around 175 and 200°. Monitoring the distribution of the value of this angle in the pseudo-trajectory of 2μs of MyoVI in PPS, discussed in Chapter 8, we observed a higher frequency in the same range of values (see Figure 9.8). In particular, we can define a couple of ranges of $\chi 1$ values that exist with higher frequency. The first is between 175 and 200° and it is populated in 57% of the total frames of the trajectory. The second one is between 260 and 290° and it is populated for 25% of the whole trajectory.

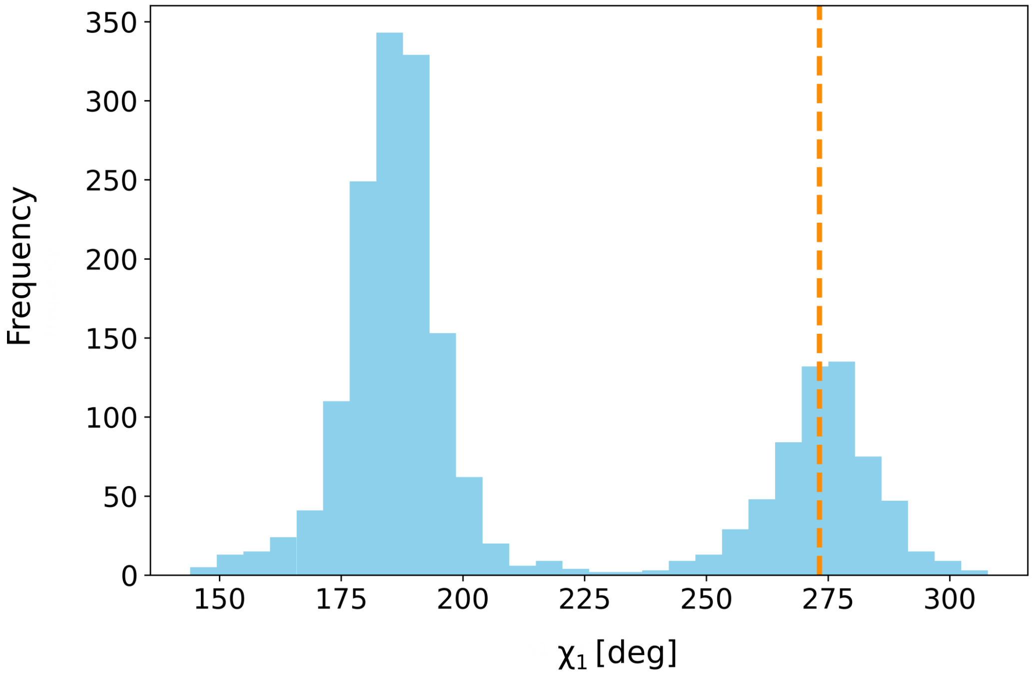


Figure 9.8: Distribution of $\chi 1$ angle in the pseudo-trajectory of 2μs of PPS MyoVI discussed in Chapter 8. In orange dashed line the value of the angle in the X-ray structure of reference. The frequency of the $\chi 1$ angle value between 175 and 200° is 57.6 %. Instead, the frequency of a value between 272 and 273° is 1.2%.

This suggests that we cannot use the reference value of the $\chi 1$ of F647 as a criterion for compounds to be prioritized in our vHTS, because it would be too restrictive for the final selection.

9.4.4 Solubility of molecules

For the final selection of molecules, we determined their solubility. The solubility of the molecules to opt for experimental testing is a fundamental parameter so that compounds that would otherwise precipitate in solution are not tested. We used DataWarrior (Sander et al., 2015) for *in silico* prediction of aqueous solubility of the seventy compounds, expressed in terms of LogS (10-based logarithm of the solubility measured in mol/liter). We then plotted the predicted ΔG of binding and the LogS values for the 70 compounds, including

Blebbistatin (Figure 9.9). We set two cut-offs, one for the free-energy of binding and one for the LogS, respectively. We chose $-43.97 kcal/mol$ as free-energy energy cutoff because this was the predicted ΔG for Blebbistatin. As mentioned before, we know that Blebbistatin is not a good binder for MyoVI, so we aim at keeping molecules with a better free-energy prediction (lower ΔG). As concerns the LogS cut-off, we fixed it at -4, since 62% of the FDA approved drugs available on the market share a calculated LogS greater than -4 (data computed with Datawarrior over the 3545 approved drugs extracted from ChEMBL database (Gaulton et al., 2017)). Looking at the plot in Figure 9.9, we see that around 30 compounds meet the selection criteria (blue dots on the bottom right). It should not be surprising that many other molecules have better predicted ΔG at lower LogS values than the cut-off since we are targeting a predominantly hydrophobic pocket.

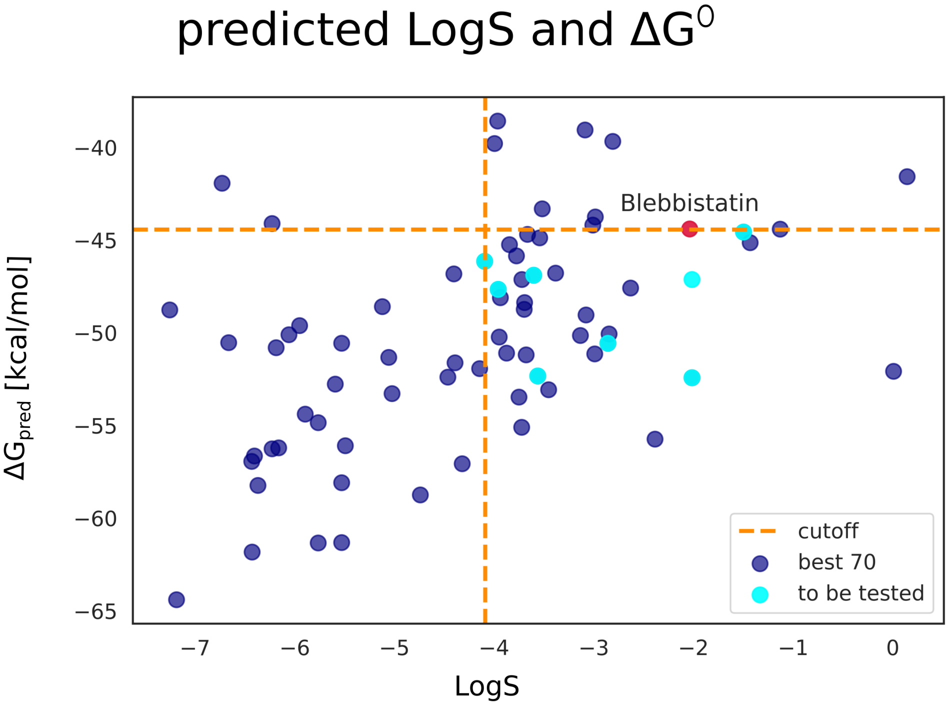


Figure 9.9: Aqueous solubility (LogS) and ΔG predicted for the final 70 compounds (blue dots) and Blebbistatin (red dot). In cyan are highlighted the compounds acquired for experimental testing. The dashed orange lines represent the cut-off established both for the ΔG and for the LogS value for the final selection of compounds.

To summarize, the final criteria for the selection of molecules to be tested are:

1. compounds with a predicted binding affinity (MMGBSA+QHMB) better than Blebbistatin.

2. a LogS greater than -4.
3. RMSD of the targeted pocket lower than 3 Å.

In the end, we extracted 22 molecules that we proposed as possible hits to our collaborator at Institut Curie in Paris (Dr. Anne Houdusse's laboratory). After a visual inspection of the systems, eight compounds were acquired for the following experimental testing.

No scaffold will be shown for confidentiality reasons.

Below in Figure 9.10, depicted in mesh, one of the prioritized molecules in the cavity of interest. We can appreciate how the molecule nicely accommodates in the cavity. All the molecules that were selected for *in vitro* testing occupy the space of this cavity of interest by establishing mostly hydrophobic interactions. Nevertheless, few recurrent H-bonds are found in all the systems with L229 and Y228. Moreover, a few molecules interact with some of the three residues that differ from the Blebbistatin pocket in *Dicty* myosin II (F647, A458, and C476). This could ensure selectivity for MyoVI to *Dicty* myosin II.

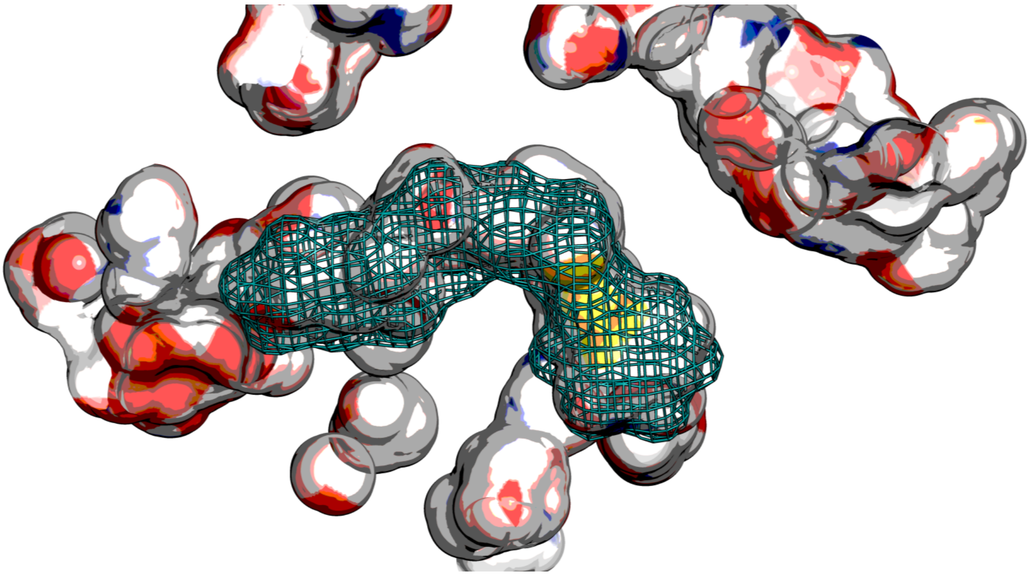


Figure 9.10: Representation of how one of the molecules selected for the *in vitro* testing settles into the cavity of interest (represented in surface whose color changes according to the polarity of the atoms).

The molecules occupy the cavity of interest to stabilize key elements in allosteric transmission in the motor domain (U50, RH, and the central β-strands of the transducer).

9.4.5 Experimental testing: ATP-ase assay

Once purchased from Thermo Fisher Scientific, the eight molecules were re-suspend in 100% DMSO and keep them at -20°C at 50mM to then be tested with ATPase assay. This assay is the experiment optimized for studying the *in vitro* inhibition of small molecules targeting myosin (Cheung et al., 2002; De La Cruz and Michael Ostap, 2009). Via a phosphoryl transfer reaction, the enzyme Pyruvate kinase transfers a phosphate group from phospho-enol-pyruvate (PEP) to ADP, a product of ATP-hydrolysis of myosin. The reaction produces one molecule of ATP and one molecule of pyruvate (De La Cruz and Michael Ostap, 2009). The pyruvate serves as a substrate for the lactate dehydrogenase (LDH), which catalyzes its conversion to lactate (De La Cruz and Michael Ostap, 2009). The reaction proceeds with the consumption of NADH that is oxidated to NAD+ (De La Cruz and Michael Ostap, 2009). Since NADH absorbs light at 340 nm and since one molecule of NADH is consumed in the last reaction per molecule of ATP that is generated in the first one, the assay measures the change in absorbance of NADH at 340 nm as an indirect measurement of ADP obtained by ATP-hydrolysis by myosin (De La Cruz and Michael Ostap, 2009). To measure the inhibitory activity of small molecules, one has to set a baseline of the reaction to measure the basal enzymatic activity of myosin in absence of modulators. Once the reference rate is set, one can monitor the inhibition of myosin mediated by a small molecule by incubating the actomyosin complex with the plausible inhibitor. In case of a positive result (inhibitory activity), the slope of the curve should decrease to the basal rate of ATPase activity. In case no inhibition is registered, the slope of the curve of absorbance reaction of NADH does not change over time.

9.4.5.1 *Experimental testing: Results*

The results of the first ATP-ase assay performed by Lee Sweeney's group, in Philadelphia, for the eight molecules are shown in Figure 9.11. As we can infer rather clearly, unfortunately, none of the 8 molecules showed inhibitory activity at neither a concentration of 50 μM nor at 200 μM.

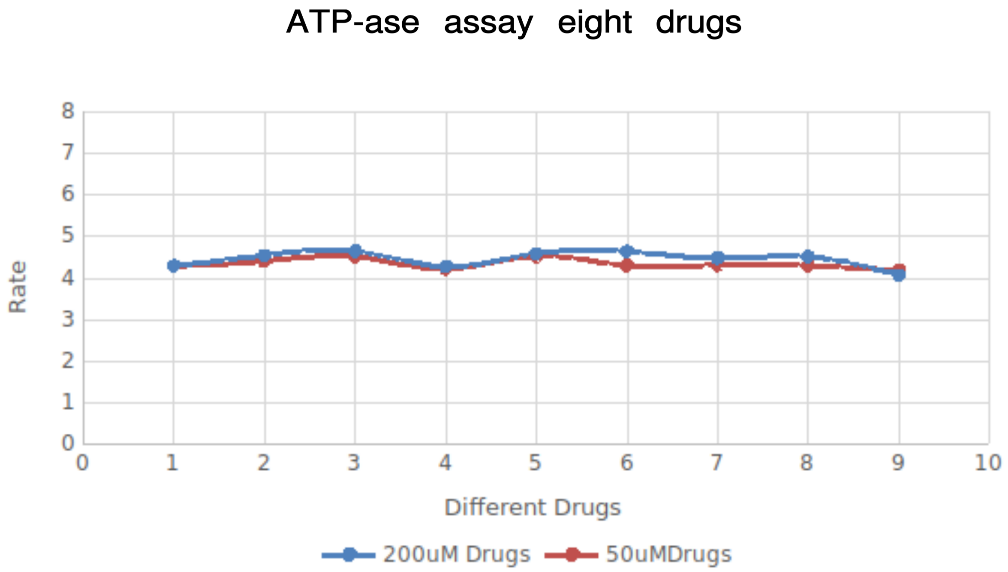


Figure 9.11: ATP-ase assay carried out for the eight molecules selected from vHTS. The assay has been performed at 25^oC with two different concentrations of drugs: 50μM and 200μM. The concentration of actin used for the assay was 40μM. The ATPase rate of reference was $4.3 sec^{-1}$.

9.5 CONCLUSIONS & FUTURE PERSPECTIVES

Our first vHTS campaign on MyoVI ended with no detected active compounds, for the moment. It could be due to misfortune, as often happens in the vHTS campaign, or to the small number of molecules, we could afford to test, among which there was not an active one. It could also be that the molecules were degraded and this interfered with the experiment. Otherwise, it could be that the molecules we prioritized are good binders (based on our prediction of binding affinity) but not good inhibitors of MyoVI. Our collaborators in Paris are, therefore, trying to co-crystallize MyoVI with the drugs selected for the tests, to verify this hypothesis. In a favorable scenario, where the group of Houdusse can co-crystallize MyoVI with at least one of the molecules, we would finally have a reference structure from which to start our predictions. This could open up new avenues for future vHTS campaigns to find hit compounds and inhibitors of MyoVI. Undoubtedly, the best possible scenario would be to have a reference structure of MyoVI co-crystallized with a modulator (inhibitor in our case) with known IC_{50} and K_d. Moreover, the group of Sweeney in Philadelphia is currently conducting other biochemical assays over these eight molecules. For the time being, the results obtained open up many questions, such as which computational approaches should be used in a vHTS campaign since we predict binding affinities and not the activities of the molecules. Additionally, we performed a blind vHTS not having a straightforward definition of the binding site to target, as well as no co-crystallized known binders from which to obtain useful information. Other questions that arise are what can be changed in our protocol to ameliorate the predictions (different docking program(s) and scoring function(s)? Another possible binding pocket to target? Consider a larger and more diverse library from Maybridge? Run more replicas to better sample the binding mode of the molecules? etc). Future developments of screening strategies might be built on this first challenging attempt, which looked, anyway, very promising.

10

PROTEIN KINASES

Since the 1950s, when the regulatory role of reversible phosphorylation of glycogen phosphorylase was discovered, the study of protein phosphorylation in regulating vital functions in cells has assumed great importance (Manning et al., 2002). A class of enzymes that mediate the reversible transfer of a phosphate group to other proteins is known as protein-kinase (Kinases) (Attwood et al., 2021; Manning et al., 2002). Kinases play a key role in regulating cellular metabolism, transcription, growth, and proliferation (Manning et al., 2002). Deregulation in their activity, especially abnormal activation due to point mutation (Attwood et al., 2021), takes part in tumor-cell growth (Downward, 2003; Scarpi-Luttenauer et al., 2022). Protein kinases became a potential drug target (Parsons and Parsons, 2004) for small molecules that could act as anticancer drugs (Scarpi-Luttenauer et al., 2022). In 1977 Omura and co-workers discovered Staurosporine (STU) while screening for microbial alkaloids (ŌMura et al., 1995). The molecule is an indolo-carbazole derivative (ŌMura et al., 1995) (see Figure 10.1) which was later shown to be a great inhibitor of phospholipid/CA^{2+} dependent protein kinases (protein kinase C), with an $IC_{50} = 2.7$ nM (Tamaoki et al., 1986). STU binds the pocket of Kinases establishing a pattern of H-bonds between its lactame fragment and some residues of the enzyme (Gani and Engh, 2010; Scarpi-Luttenauer et al., 2022) Albeit its strong inhibitory activity, STU was not highly selective and therefore too toxic to be used as a drug (Attwood et al., 2021; Gani and Engh, 2010). STU represented the starting point for the development and synthesis of highly selective analogs, with drug-like properties (Attwood et al., 2021).

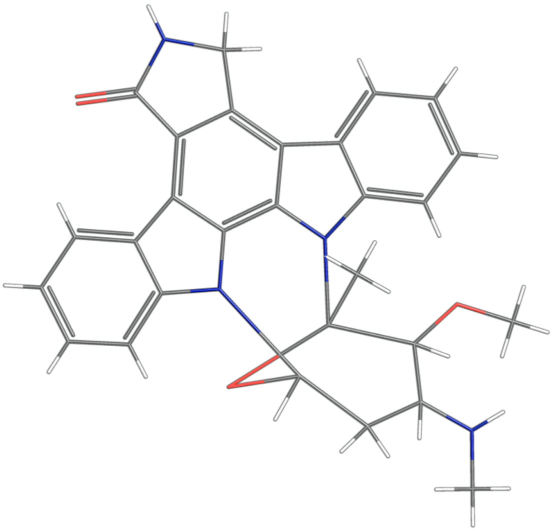


Figure 10.1: 2D representation of Staurosporine ((2S,3R,4R,6R)-3-methoxy-2-methyl-4-(methylamino)-29-oxa-1,7,17-triazaoctacyclo[12.12.2.12,6.07,28.08,13.015,19.020,27.021,26]nonacosa-8,10,12,14,19,21,23,25,27-nonaen-16-one).

10.1 DESIGN OF PROMISING METAL COORDINATED INHIBITORS

10.1.1 Phenanthroline-based ligands and Titanium complexes

Since the early 2000s, there has been a focus on coordination complexes to design novel protein kinase inhibitors, by mimicking the structure of Staurosporine with organometallic molecules (Bencini and Lippolis, 2010; Dörr and Meggers, 2014). In particular, since the coordination chemistry of 1,10-phenanthroline (1,10-phen) derivatives is well documented (Bencini and Lippolis, 2010) and the backbone of these ligands can be modified and optimized with H-bond donor-acceptor groups (Scarpi-Luttenauer et al., 2022) to mimic the interactions of STU in the pocket (Gani and Engh, 2010), Mobian and co-workers proposed to design a series of 1,10-phen derivatives as novel kinases' inhibitors (Scarpi-Luttenauer et al., 2022). Scarpi-Luttenauer and Mobian synthesized a series of 1,10-phen derivatives introducing in some of them a pyridinone moiety to mimic the structure of STU (Scarpi-Luttenauer et al., 2022) (see Figure 10.2).

Moreover, in 2020, Scarpi-Luttenauer *et al* demonstrated that coordination complexes based on a TiO_4N_2 core incorporating 1,10-phen derivatives could generate highly robust molecules, which could be of particular interest for biochemical applications as protein kinase inhibitors (Scarpi-Luttenauer et al., 2020). Due to their expertise with this coordination chemistry and to the fact that Ti(IV) is a good candidate element for the design of therapeutic molecules (Scarpi-Luttenauer et al., 2020), Mobian and co-workers introduced the TiO_4N_2 motif

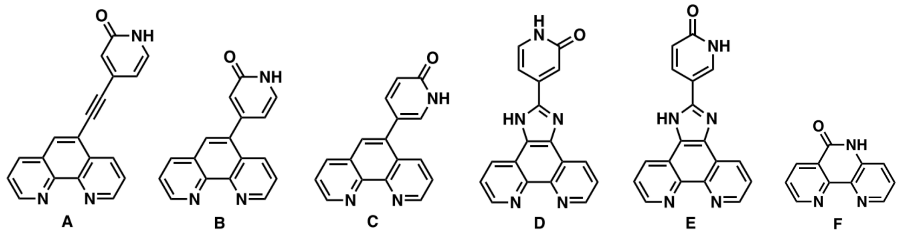
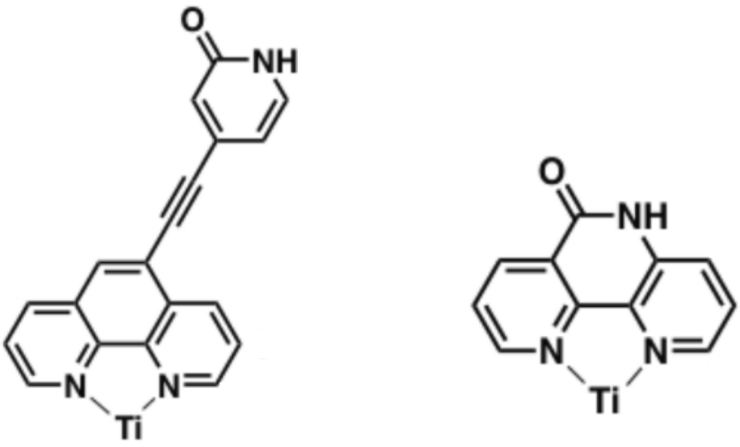


Figure 10.2: Series of 1,10-phenanthroline derivatives (A-E) linked with a pyridinone group through various spacers and the last derivative F without the pyridinone group (Scarpi-Luttenauer et al., 2022). The image has been taken from (Scarpi-Luttenauer et al., 2022) without modifications.

in the series of 1,10-phen derivatives (Scarpi-Luttenauer et al., 2022). The synthesis was successful for only two complexes, named $Ti(1)_2A$ and $Ti(1)_2F$, depicted in Figure 10.3. These two complexes were characterized by NMR spectroscopy, mass spectrometry, and elemental analysis (Scarpi-Luttenauer et al., 2022). Both Titanium complexes and 1,10-phen derivatives were tested for cytotoxicity, by measuring the cell viability in human gastric adenocarcinoma cells (AGS) (Scarpi-Luttenauer et al., 2022). Unfortunately, the Ti(IV) compounds were not soluble in the medium used in the test (DMSO) and their toxicity could not be evaluated. Among the 1,10-derivatives only one compound was not toxic (F) (Scarpi-Luttenauer et al., 2022). However, the cytotoxicity of the other 1,10-phen derivatives is not surprising since it was already detected as a consequence of intercalation between DNA strands for some complexes incorporating a phenanthroline, that could not be entirely repaired by cellular mechanisms (Roy et al., 2008; Scarpi-Luttenauer et al., 2022).

10.1.1.1 *Molecular docking on MST2 kinase*

To understand the possible therapeutical application of the synthesized molecules as anticancer drugs, we performed molecular docking on MST2 kinase (PDB code: 5DH3) (Fan et al., 2016; Scarpi-Luttenauer et al., 2022). MST2 is a kinase that plays a key role in the Hippo signaling pathway (Qin et al., 2013). This pathway is involved in cell proliferation, cell differentiation, and control of apoptosis (Qin et al., 2013; Scarpi-Luttenauer et al., 2022). Deregulation of the Hippo pathway is associated with the occurrence of gastric cancer, for which effective therapies are lacking (Qiao et al., 2018; Scarpi-Luttenauer et al., 2022). At first, we defined the docking protocol by re-docking in MST2 the co-crystallized molecule XMUMP-1 (Fan et al., 2016), after preparing the ligand using PrepFlow (Sisquellas and Cecchini, 2021). We performed the docking with *DockFlow*, choosing PLANTS



$T_i(1)_2A$ $T_i(1)_2F$

Figure 10.3: Complexes $Ti(1)_2A$ and $Ti(1)_2F$ (Scarpi-Luttenauer et al., 2022). The image has been taken from (Scarpi-Luttenauer et al., 2022).

as docking program and ChemPLP as scoring function. As shown in Figure 10.4 the molecule was correctly re-docked in its binding pocket, with an RMSD of 0.5 Å to the reference coordinates.

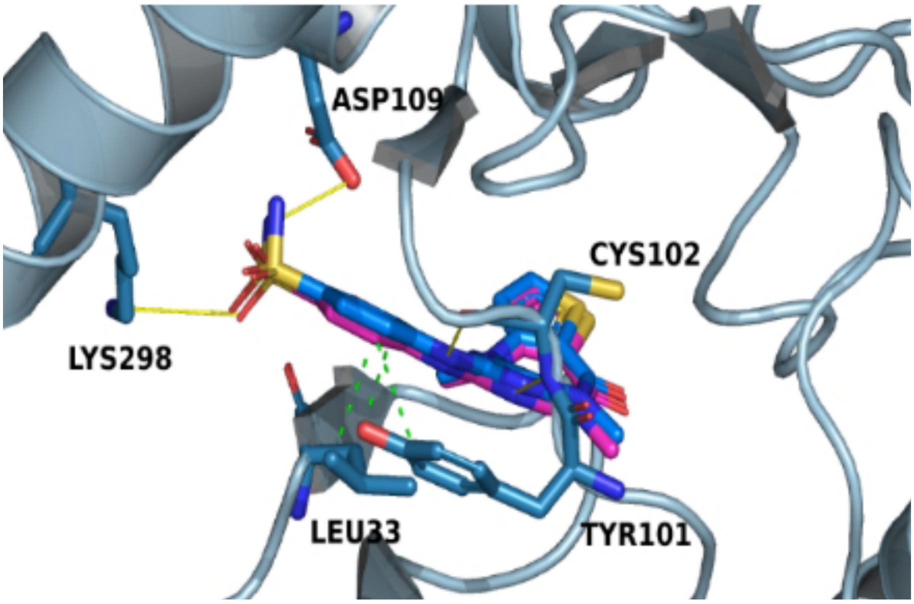


Figure 10.4: Re-docking of XMUMP-1 in MST2 (PDB code: 5DH3) (Fan et al., 2016). In blue sticks is depicted the co-crystallized ligand and in magenta ones the re-docked molecule. In green are highlighted the hydrophobic interactions and in yellow are the hydrogen bonds. The docking score of this binding mode is -76.72. The re-docked molecule establishes both the same H-bonds (with CYS102, LYS298, and ASP109) and hydrophobic interactions (with LEU33 and TYR101) as the reference in the crystal structure (Scarpi-Luttenauer et al., 2022). The analysis of ligand-protein interactions has been produced using PLIP (Adasme et al., 2021).

Once the docking protocol was validated with the re-docking experiment, we docked some more soluble molecules of the series of 1,10-phen derivatives (A, B, C, and F) and the two Ti(IV) complexes with the same procedure. The phenanthroline derivatives properly fit the targeted pocket with a favorable docking score, ranging from -84.91 (compound A) to -64.65 (compound F) (Scarpi-Luttenauer et al., 2022). On the other hand, $Ti(1)_2A$ and $Ti(1)_2F$ turn out to be too bulky for the binding pocket (see Figure 10.5).

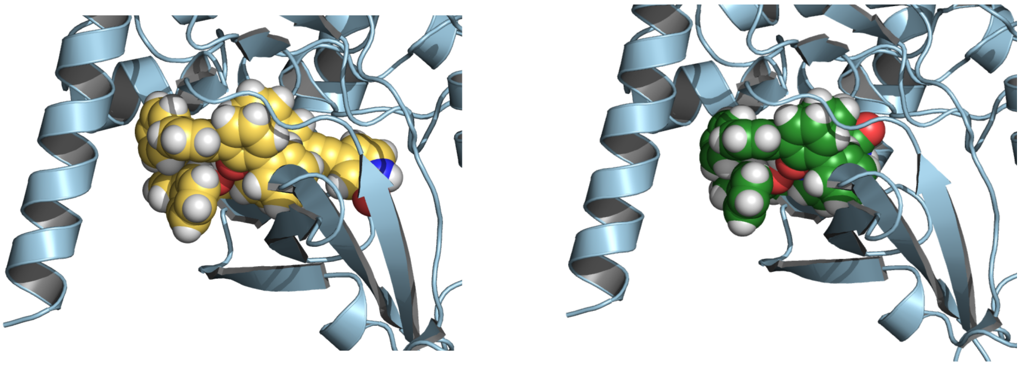


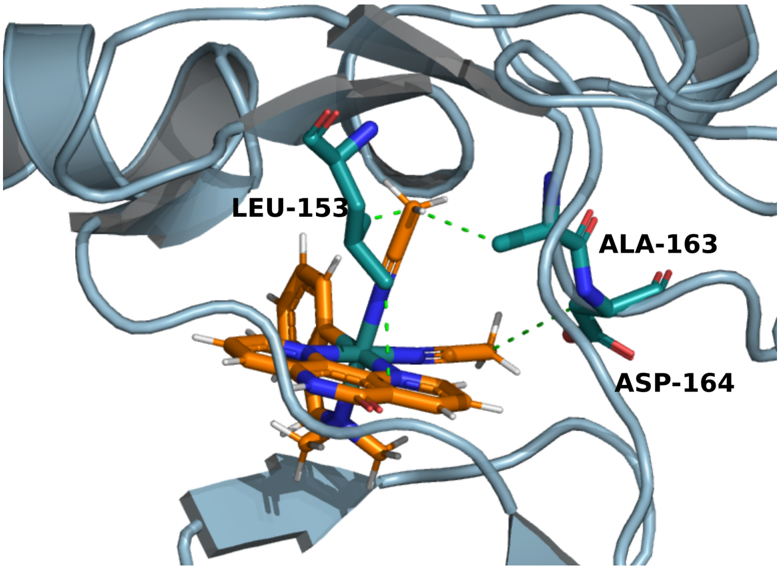
Figure 10.5: Docking results of $Ti(1)_2A$ and $Ti(1)_2F$ compounds. In yellow spheres $Ti(1)_2A$, on the left, and in green spheres $Ti(1)_2F$, on the right. The docking score of the best pose for the two molecules (here reported) is +285.9 and +600.1 for $Ti(1)_2A$ and $Ti(1)_2F$ respectively.

10.1.2 Ruthenium and Platinum complexes

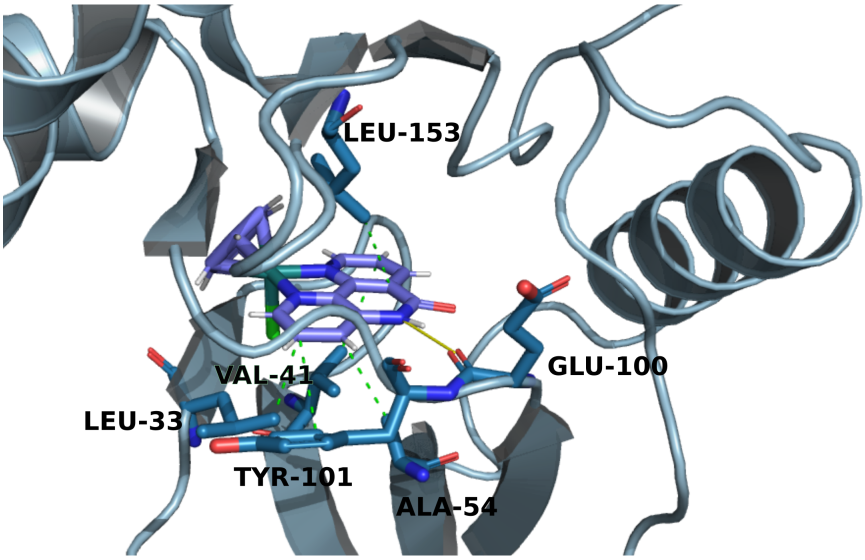
Since molecule F is the only non-toxic one in the series of 1,10-phen derivatives and it can also bind the pocket of MST2 without steric clashes and with a favorable docking score (docking score of -64.65), it can be selected as a promising scaffold for the design of novel coordination complexes able to inhibit MST2 and, consequently, the Hippo signaling pathway (Scarpi-Luttenauer et al., 2022). Intending to develop new generations of less bulky compounds, based on hydrogen bond donor-acceptor phenanthroline ligands (Scarpi-Luttenauer et al., 2022), Mobian and co-workers considered the coordination with other metals, such as Ruthenium (Ru(II)) and Platinum (Pt(II)) (Scarpi-luttenauer, 2022).

10.1.2.1 *Molecular docking on MST2 kinase*

They designed a new series of compounds that we then examined by docking experiments, using the same procedure we defined for the previous series of molecules (ligand prepared with PrepFlow (Sisquellas and Cecchini, 2021) and docked with DockFlow using PLANTS and ChemPLP scoring function). Hereinafter we describe the results of four compounds that are the most promising in the analyzed series and have also been experimentally tested. In Figure 10.6 and Figure 10.7 are reported the docking results of Ru(3), Ru(5) and Pt(3) and Pt(4a), respectively.

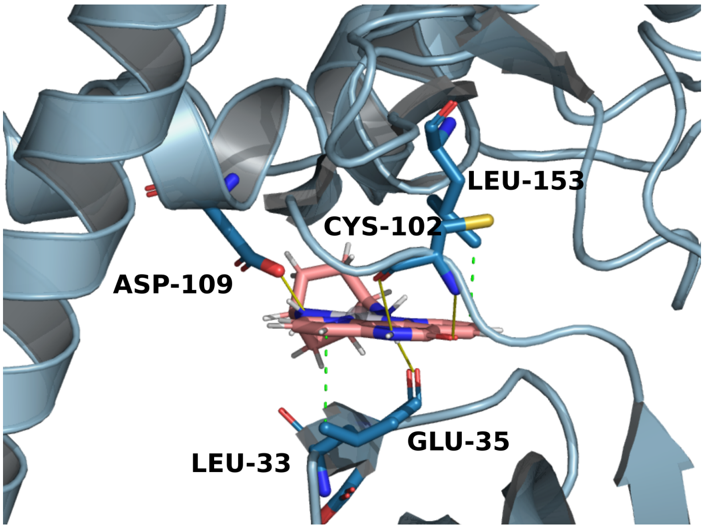


(a) Best binding mode of Ru(3) docked in MST2. The analysis of ligand-protein interactions has been produced using PLIP (Adasme et al., 2021).

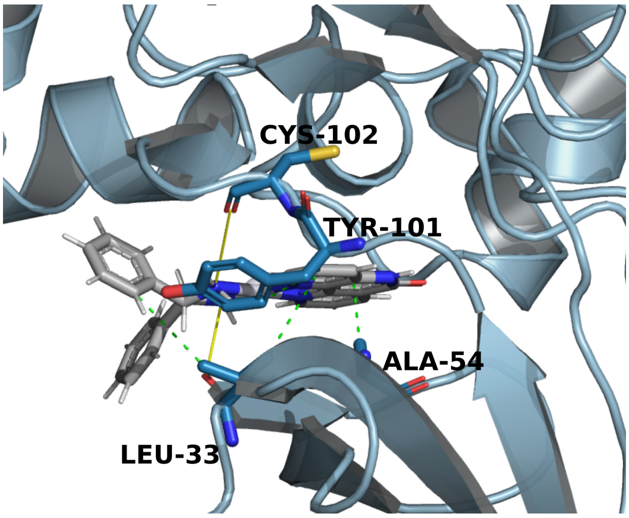


(b) Best binding mode of Ru(5) docked in MST2. The analysis of ligand-protein interactions has been produced using PLIP (Adasme et al., 2021).

Figure 10.6: In sub-figure (a) the best docking pose of Ru(3) in MST2 kinase. The docking score is +13.7. In sub-figure (b) the best docking pose of Ru(5) in MST2 kinase. The docking score is -74.6



(a) Best binding mode of Pt(3) docked in MST2. The analysis of ligand-protein interactions has been produced using PLIP (Adasme et al., 2021).



(b) Best binding mode of Pt(4a) docked in MST2. The analysis of ligand-protein interactions has been produced using PLIP (Adasme et al., 2021).

Figure 10.7: In sub-figure (a) the best docking pose of Pt(3) in MST2 kinase. The docking score is -11.9. In sub-figure (b) the best docking pose of Pt(4a) in MST2 kinase. The docking score is -22.8.

From our analysis, it emerged that Ru(5) interacts stronger with MST2 than all the other molecules of the series. Moreover, Ru(3) shows a positive docking score, due to a high value for the steric contact score and repulsion score of ChemPLP scoring functions. These scores represent the Van der Waals and electrostatic contribution, respectively. Therefore, the binding of Ru3 is unfavorable in this pocket of M2ST, due to steric hindrance of the molecule and repulsive forces.

10.1.2.2 *In vitro testing of inhibition of MST2 kinase*

The inhibitory activity of these four compounds has been tested, monitoring the inhibition of the phosphorylation of YAP (Yes-associated protein), namely co-transcriptional activator of the Hippo signaling pathway (Qin et al., 2013; Scarpi-luttenauer, 2022). AGS cells were incubated with the four complexes at a concentration of 2 and 25 μM. The results in Figure 10.8 show that the proportion of pYAp and YAP is five times smaller than the control, in presence of Ru(3).

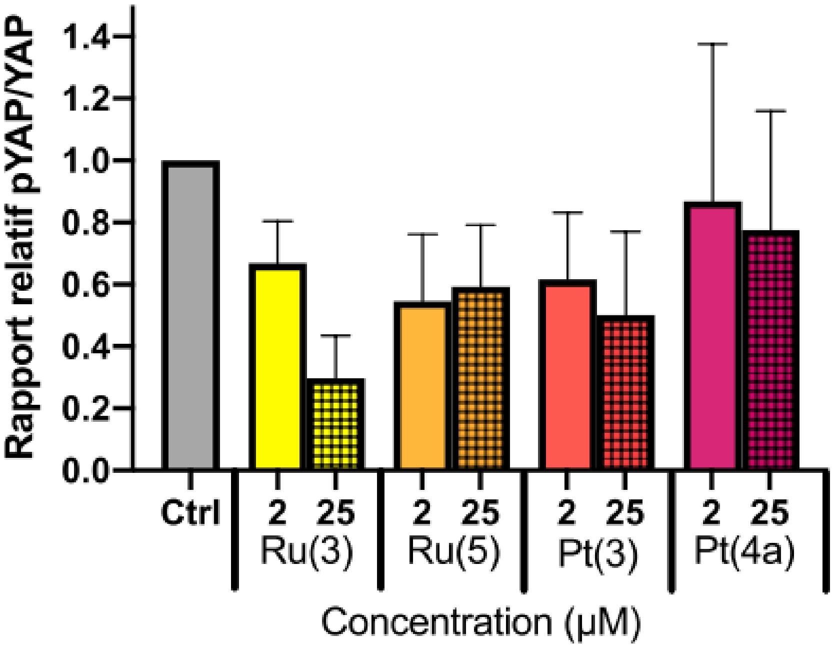


Figure 10.8: Determination of MST2 inhibition for the four complexes under investigation. The data reported are the average values of three different experiments. The image has been taken from (Scarpi-luttenauer, 2022) without modifications. The experiments were conducted by Scarpi and Orvain at Inserm U1113.

10.1.3 Investigation of specificity

To verify the selectivity of these compounds for MST2 to avoid off-target toxicity the same docking experiment and experimental evaluation of inhibitory activity have been performed on a different protein kinase, named S6K1 (Qin et al., 2015). S6K1 is involved in the control of cell size, protein synthesis, and regulation of glucose levels in the cell through a different biological pathway than the Hippo one (Qin et al., 2015; Scarpi-luttenauer, 2022). Since the inhibition of one pathway does not interfere with the other, we can understand if the molecules are selective inhibitors for one kinase rather than the other.

Figure 10.9 shows that Ru(3) does not have the same inhibitory effect on S6K1 and the most remarkable variation in activity results for Pt(3) and Pt(4a). These compounds, indeed, show almost no activity for MST2 but, on the contrary, they strongly inhibit the phosphorylation of S6 (see Figure 10.9).

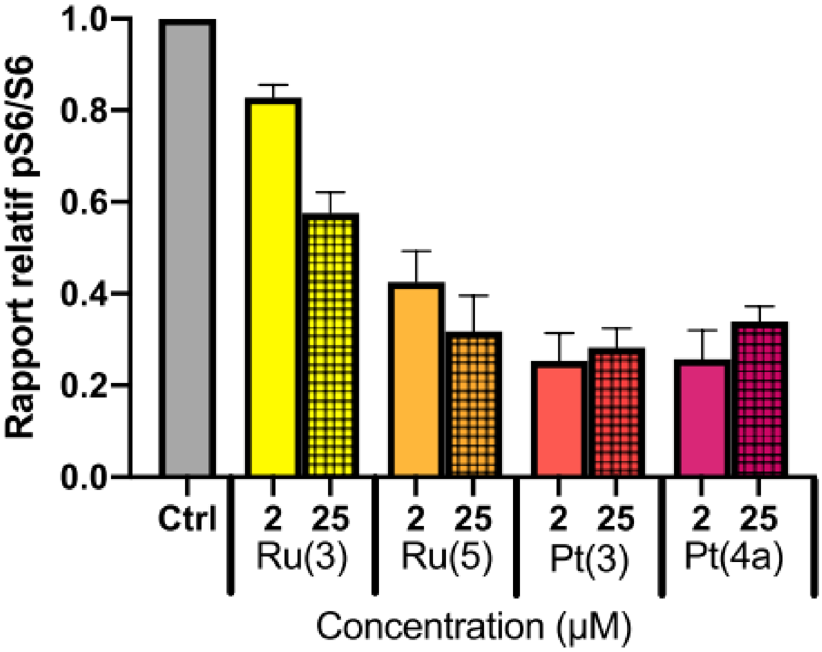


Figure 10.9: Determination of S6K1 inhibition for the four complexes under investigation. The data reported are the average values of three different experiments. The image has been taken from (Scarpi-luttenauer, 2022) without modifications. The experiments were conducted by Scarpi and Orvain at Inserm U1113.

These results (Figure 10.8 and fig:experiment-S6K1) demonstrate that Pt(3) and Pt(4a) show better inhibitory activity of S6K1, associated with the phosphorilation of S6, than both Ru(3) and Ru(5). Vice versa Ruthenium compounds, in particular Ru(3), inhibit more MST2 kinase, thus the phosphorylation of YAP.

10.1.3.1 *Molecular docking on S6 kinase*

To check the propensity of the four molecules to bind MST2 or S6K1, we conducted molecular docking analyses on S6 kinase, using the structure of the human S6K1 solved in complex with a ruthenium-based inhibitor, named FL772 (Qin et al., 2015). At first, we designed the docking protocol by re-docking FL772 in the active site of human S6 kinase with *DockFlow*. We used, again, PLANTS and ChemPLP scoring function. As shown in Figure 10.10 the re-docked compound was overlapping with the co-crystallized molecule with an RMSD of 0.9 Å.

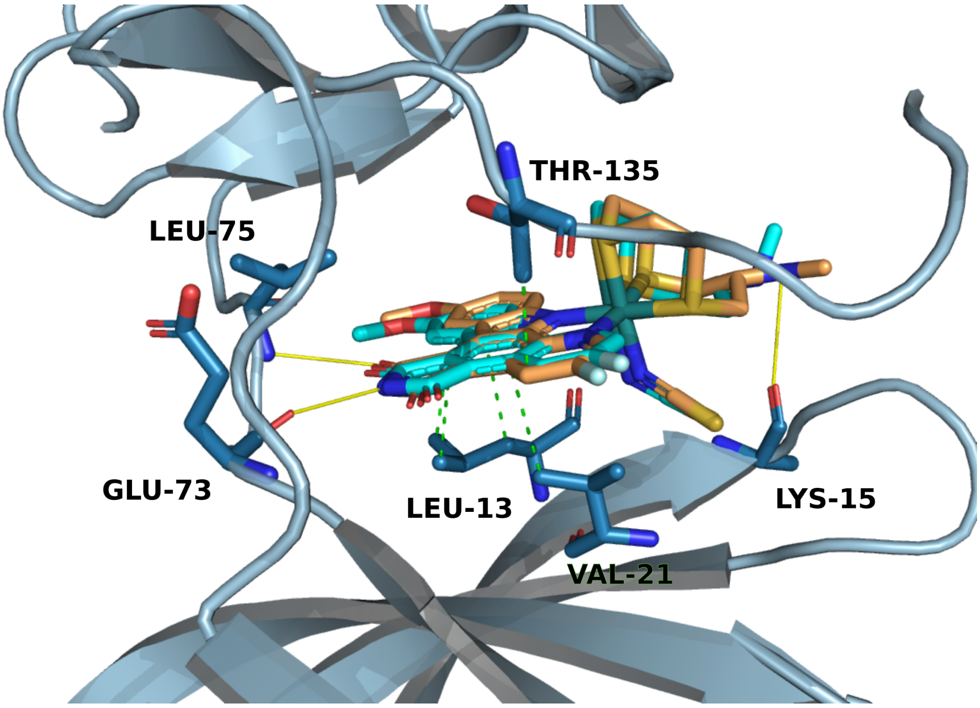
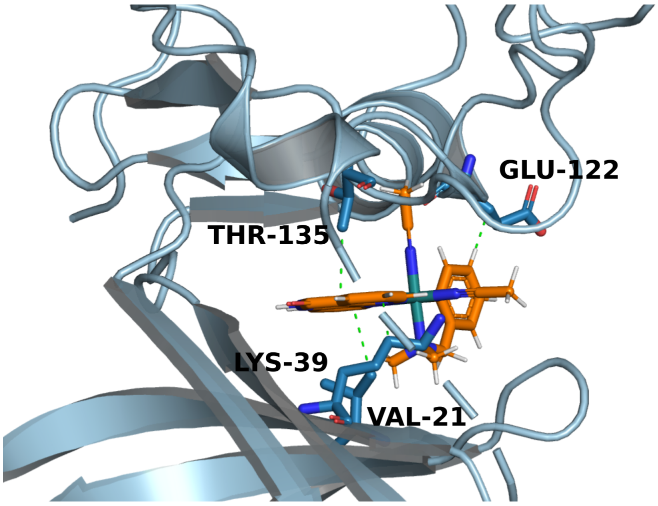
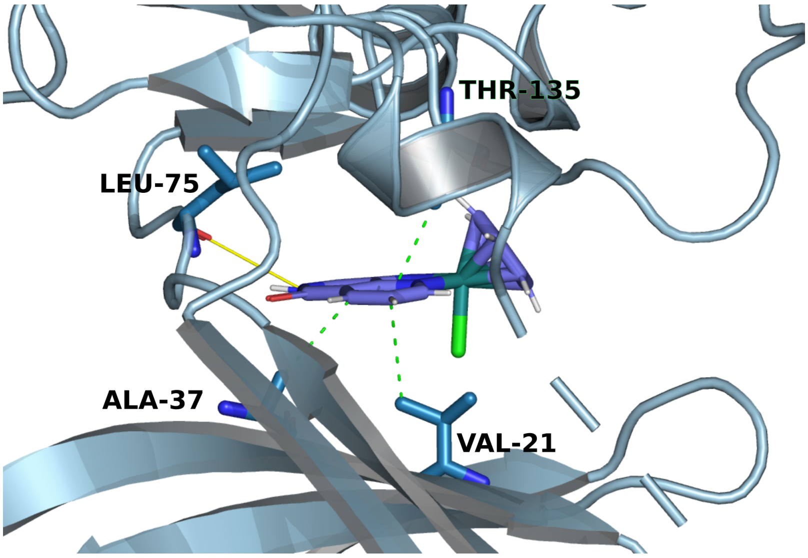


Figure 10.10: Re-docking of FL772 in S6K1 (PDB code: 4RLP). The co-crystallized ligand is depicted in cyan and the re-docked molecule is in orange. In green are highlighted the hydrophobic interactions and in yellow are the hydrogen bonds. The docking score of this binding mode is -97.17. The re-docked molecule establishes the same H-bonds (with LEU75, GLU73, and LYS15) as well as the same hydrophobic contacts (with THR135, LEU13, and VAL21) as the reference structure (FL772) in the X-ray. The analysis of ligand-protein interactions has been produced using PLIP (Adasme et al., 2021).

We, therefore, used the same protocol to dock Ru(II) and Pt(II) compounds under investigation (see Figure 10.11 and Figure 10.12).

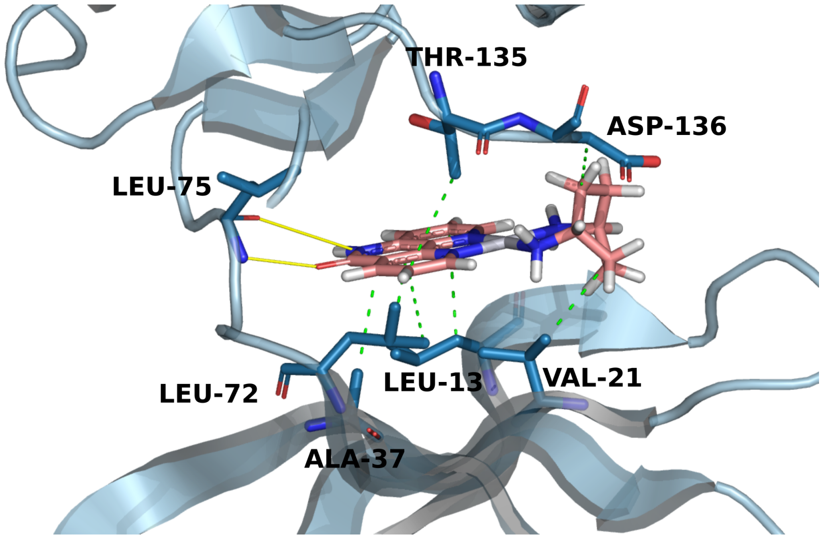


(a) Best binding mode of Ru(3) docked in human S6 kinase. The analysis of ligand-protein interactions has been produced using PLIP (Adasme et al., 2021).

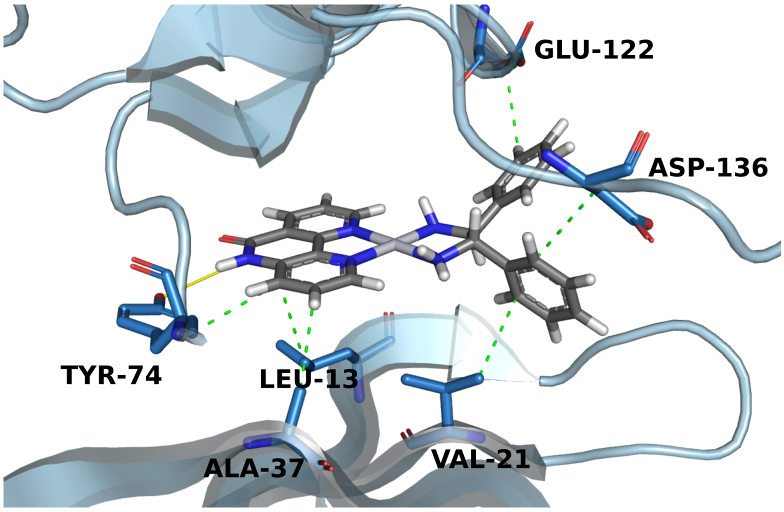


(b) Best binding mode of Ru(5) docked in human S6 kinase. The analysis of ligand-protein interactions has been produced using PLIP (Adasme et al., 2021).

Figure 10.11: In sub-figure (a) the best docking pose of Ru(3) in S6K1. The docking score is -61.3. In sub-figure (b) the best docking pose of Ru(5) in S6K1. The docking score is -68.9.



(a) Best binding mode of Pt(3) docked in human S6 kinase. The analysis of ligand-protein interactions has been produced using PLIP (Adasme et al., 2021).



(b) Best binding mode of Pt(4a) docked in human S6 kinase. The analysis of ligand-protein interactions has been produced using PLIP (Adasme et al., 2021).

Figure 10.12: In sub-figure (a) the best docking pose of Pt(3) in S6K1. The docking score is -79.5. In sub-figure (b) the best docking pose of Pt(4a) in S6K1. The docking score is -91.5.

The docking of the designed inhibitors in the active sites of the two kinases reveals a different propensity of the two classes of compounds to bind S6K1 or MST2 kinase (see Table 10.1). The results in Table 10.1 show that the Platinum compounds prefer binding to S6K1, whereas Ru(5) interacts stronger with MST2. Ru(3) has a positive docking score for MST2 and a negative one for S6K1, showing a preference for binding S6 kinase whose active site is bigger and can better accommodate the molecule. This result is in contrast with the experimental measurement, according to which Ru(3) is a strong inhibitor of MST2 (Figure 10.8).

LIGAND	S6K1	MST2	Δ(S6K1-MST2)
Ru(3)	-61.3	+13.7	-75
Ru(5)	-68.9	-74.6	+5.7
Pt(3)	-79.5	-67.5	-11.9
Pt(4a)	-91.5	-68.6	-22.8

Table 10.1: Docking scores of molecular docking performed for the four molecules both on S6K1 and MST2 kinases. The second and third columns report the docking score of the chosen binding pose of each molecule in the two proteins. In the last column is calculated the delta of the docking scores for S6K1 and MST2 to determine the preference to bind to one kinase or the other.

The docking results are in agreement with the *in vitro* testing, according to which Pt(3) and Pt(4a) show better inhibitory activity in S6K1 than both Ru(3) and Ru(5) (Figure 10.9). Conversely, Ru(5) shows a delta docking score that favors the binding for MST2 kinase, as also shown experimentally (Figure 10.8).

10.2 CONCLUSIONS

In conclusion, we can assert that the *in silico* results and the *in vitro* outcomes agree, showing that the Platinum compounds are putative inhibitors for S6K1 kinase, instead, the Ruthenium compounds are designed as possible inhibitors of MST2 kinase. These findings could represent a novel frontier in protein kinase inhibitor design.

11

CONCLUSIONS & FUTURE PERSPECTIVES

Myosins are a class of biomolecular motors fundamental for a wide range of cellular functions in eukaryotic organisms, such as cell proliferation, cell migration, muscle contraction just to mention a few. Myosins move on actin filament thanks to a refined mechanism of signal amplification that converts chemical energy derived from ATP hydrolysis into mechanical work. Dysfunctions of these molecular motors lead to several human disorders. Deregulation of cardiac myosin, for instance, is involved in both systolic heart failure and hypertrophic cardiomyopathy, depending on a decreased or increased cardiac contractility, respectively (McMurray et al., 2013; Planelles-Herrero et al., 2017; Robert-Paganin et al., 2018). It has also been demonstrated an interesting connection between overexpression of Myosin VI and some human cancers progression, ranging from prostate cancer to ovarian cancer to breast cancer (Dunn et al., 2006; Hari-Gupta et al., 2022; Makowska et al., 2015; Vreugde et al., 2006). Mutations of Myosin VI are connected to deafness due to the role of this myosin in preserving the integrity of stereocilia of the inner ear (Ahmed et al., 2003; Melchionda et al., 2001; Pylypenko et al., 2015). Nowadays, there are drugs able to indirectly improve cardiac contractility through signaling cascades, but mechanism-related side effects limit their application. Moreover, although Myosin VI has been extensively studied, no effective modulators are known. Both heart disease and cancer have a major impact on today's society. The current pharmacological treatments for heart disease rely on drugs that activate second messenger signal cascades, whereas there is a total lack of effective modulators for Myosin VI. It becomes, therefore, essential to find innovative and safe treatments with molecules able to target both cardiac myosin and Myosin VI. To find novel small molecules able to modulate allosteric proteins such as myosins, without interfering with other intracellular processes, one can rely on the rational design of allosteric modulators. These non-competitive modulators have innumerable advantages over other drugs that could target the active site of the enzyme. In particular, allosteric modulators would bind a different binding pocket than the active site with high selectivity for a specific isoform of the protein, since allosteric pockets are less conserved among myosins' classes. To find new small molecules one can run a vHTS campaign using innovative tools like *ChemFlow*, our in-house software presented in Chapter 5, which offers a fully automated workflow for drug discovery with some state-of-the-art techniques in the field of CADD (i.e. molecular docking, free energy re-scoring with end-point methods).

Rational drug design relies also on in-depth studies of both structure and function of the target protein. Nowadays the huge amount of structural data available from X-ray crystallography and NMR experiments allow us to best accomplish these studies. Thanks to the availability of the X-ray structure of both cardiac myosin and Myosin VI we have been able to perform our research. In Chapter 6 and Chapter 8, we have extensively analyzed the behavior of the two myosins in the PPS state of the actomyosin cycle with Molecular Dynamics simulations. We have observed peculiar dynamics of the motor domain in both myosins. Notably, a pliancy of the converter-lever arm domain seems characteristic of the PPS state and that could be modulated by the presence of allosteric modulators binding nearby. This is evident from our analyses on cardiac myosin (Chapter 6), whose PPS structure was solved in presence of the known allosteric activator Omecamtiv Mecarbil (OM) (Planelles-Herrero et al., 2017). Despite the initial huge instability of the PPS structure detected in absence of OM, we succeeded in modeling this state of cardiac myosin so that it was stable both with and without an allosteric modulator. We inferred that our model represents a proper PPS state and can be used for pharmacological studies, as we have done in Chapter 7. For cardiac myosin, we selected this state because we aimed at understanding what makes a small molecule a cardiac myosin's activator and we believed that activators may shift the configurational ensemble of the motor domain towards PPS, being this the onset state for the powerstroke in myosins. We wondered whether the binding affinity (k_d) or dissociation rate constant (k_{off}) of small molecules would correlate with biological activity. In Chapter 7 we described how we predicted both relative binding free energy and residence time of a small dataset of compounds with available experimental activity data (Morgan et al., 2010). We designed a protocol that can be used for future vHTS campaigns on cardiac myosin, aiming at the discovery of novel cardiac activators. Moreover, this strategy can be verified by introducing decoys (false positive molecules) of OM in the screening protocol designed from some analogs of this drug. We also did not find any correlation between biological activity and free energy calculations, nor with predictions of the residence time of binding. These findings left unanswered the question "What makes a small molecule a cardiac myosin activator?" The question may be addressed by performing free energy calculations with a more rigorous approach, such as Free Energy Perturbation (FEP). We have started to develop a FEP protocol to predict absolute binding free energy for OM in PPS and we are still working on it. Another possibility would be to consider a different state of the actomyosin cycle, in equilibrium with PPS, for example, the post-rigor (PR) state. We may apply the same procedure to predict k_d for this state and then compute a $\Delta\Delta G$ of binding between the two states. The $\Delta\Delta G$ of binding would elucidate if the molecules bind preferentially to one state or the other. This approach could be further explored to design a new strategy for the screening of potential small molecules activators of cardiac myosin. Another avenue for predicting the activity of promising activators of cardiac myosin could envisage the calculation of k_{off} of P_i, depending on the nature of the allosteric modulator bound. This is because, as discussed in Chapter 7 the known allosteric modulators either accelerate (if activators) or slow down (if inhibitors) the P_i release. Moreover, using τ-RAMD for P_i release we could predict the various egress routes from the active site (with and without modulators), with a given probability (Kokh et al., 2020). This could allow us to better understand and describe the preferred route of escape of P_i from myosin's head, a concept that still remains to be clarified.

As previously mentioned, we also explored the dynamics of Myosin VI, another relevant target studied in this thesis. In Chapter 8, we wondered if the fluctuations observed in the simulations of Myosin VI reflect an intrinsic conformational variability distinctive for this myosin. A way to shed light on this consideration might be the comparison of various processive myosins (i.e myosin I, V, and X) versus non-processive ones (different myosins being class II myosins) to delineate unusual and/or common properties in the two classes. We investigated in detail the dynamics of Myosin VI in PPS, for some simulations in which the motor domain undergoes a spontaneous transition. We proposed a possible sequence of events (in Chapter 8 Section 8.2) of this transition that results in a huge movement of the converter, concluding that the breaking of the critical salt-bridge (Onishi et al., 2002) could trigger the uncoupling of the converter. We still do not know if the uncoupling of the converter is a possible normal progression of the actomyosin cycle for Myosin VI, which could bind to actin with an uncoupled converter, as proposed by Menetrey *et al* in 2005 (Ménétrey et al., 2005). Moreover, we don't know if the transition we captured is towards a pre-transition state or a P_i release state. A possible future study may involve Potential of Mean Force (PMF) calculations to predict free energy barriers compatible with the unbinding of the inorganic phosphate from the active site. Moreover, we could explore the advantages of enhance sampling techniques (i.e metadynamics; umbrella sampling; rep. exchange; etc) to increase the sampling and overcome energetic barriers between states. More evidence is certainly needed to shed light on the chemo-mechanical transduction in Myosin VI. As discussed in Chapter 9, we conducted also a vHTS on Myosin VI, choosing the PPS state as the target configuration of the protein. We chose the PPS state because this state was already targeted during a screening in 2012 by Heissler (Heissler et al., 2012) who found a molecule named 2,4,6-triiodophenol (TIP) and classified it as a non-competitive inhibitor. Since then, no further evidence has been provided to describe the binding of this molecule.

Given the apparent success of their screening, we pursue the same initial step choosing the structure solved by Menetrey and co-workers in 2007 (Ménétrey et al., 2007). In Chapter 9 we described our vHTS campaign on Myosin VI which, unfortunately, did not yet culminate in the discovery of novel modulators. We performed blind docking since no X-ray structure with a co-crystallized compound for Myosin VI is available. This is, undoubtedly, a factor that increases the difficulty of successful screening. To cope with this initial setback, we could test various docking scoring functions to get a different prioritization of compounds. Furthermore, we may also choose more molecules to test or a different library to be screened. Although the structural dynamics of MyoVI in PPS taught us a lot, the structure-based virtual screening in highly flexible allosteric proteins, such as myosin, remains very challenging. Last but not least, in Chapter 10 we described the design of novel promising kinase inhibitors, by Scarpi-Luttenauer and Mobian. We contributed to this project by performing molecular docking of their metal-coordinated molecules. The results are encouraging for the development of selective inhibitors of the Hippo signaling pathway, using an innovative chemistry that envisages the coordination of transitional metals in the structures.

A

MYOSIN VI

A.1 DEFINITION CORE ATOMS SUBDOMAINS IN MYOSIN VI

Table A.1: Definition core atoms for the various subdomains in MyosinVI

Subdomain	Core atoms
N-terminal	resid 47 to 51 or resid 62 to 65 or resid 70 to 97 or resid 109 to 116 or resid 128 to 151 or resid 157 to 172 or resid 661 to 668 or resid 682 to 689
U50	resid 185 to 198 or resid 200 to 214 or resid 220 to 235 or resid 246 to 263 or resid 285 to 305 or resid 313 to 328 or resid 331 to 352 or resid 372 to 379 or resid 383 to 394 or resid 409 to 442 or resid 450 to 457 or resid 604 to 612 or resid 615 to 622
L50	resid 468 to 499 or resid 511 to 520 or resid 525 to 535 or resid 540 to 550 or resid 576 to 590 or resid 593 to 600 or resid 642 to 660
Converter	resid 707 to 788

A.2 RMSD DISTRIBUTION PLOTS OF THE VARIOUS SUBDOMAINS IN THE TEN REPLICAS

In order to understand which domain was deviating the most in the simulations, we analyzed how the interfaces between sub-domains were changing with a series of RMSD calculations, shown hereafter.

L50 subdomain vs N-terminal

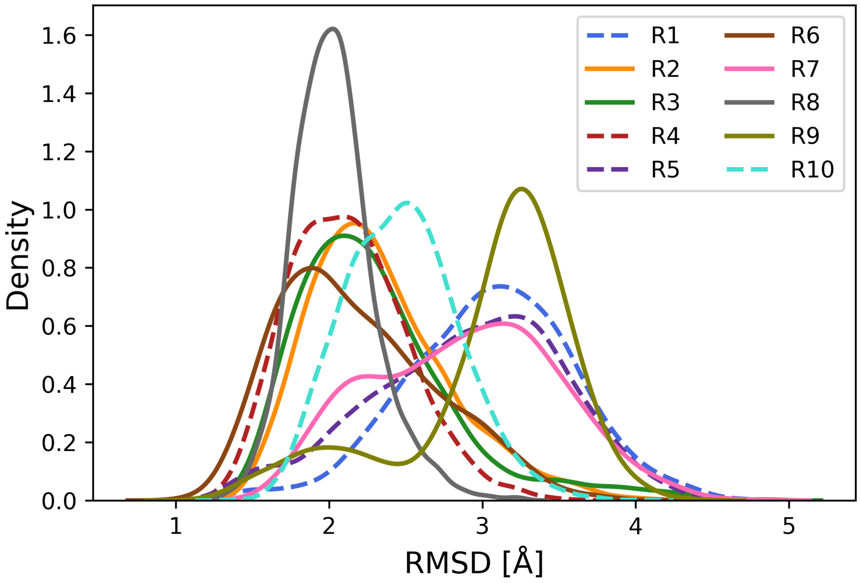


Figure A.1: Distribution plot of the RMSD of the core atoms of the L50 subdomain after alignment on the N-terminal in the ten replicas of MyosinVI.

L50 subdomain vs U50

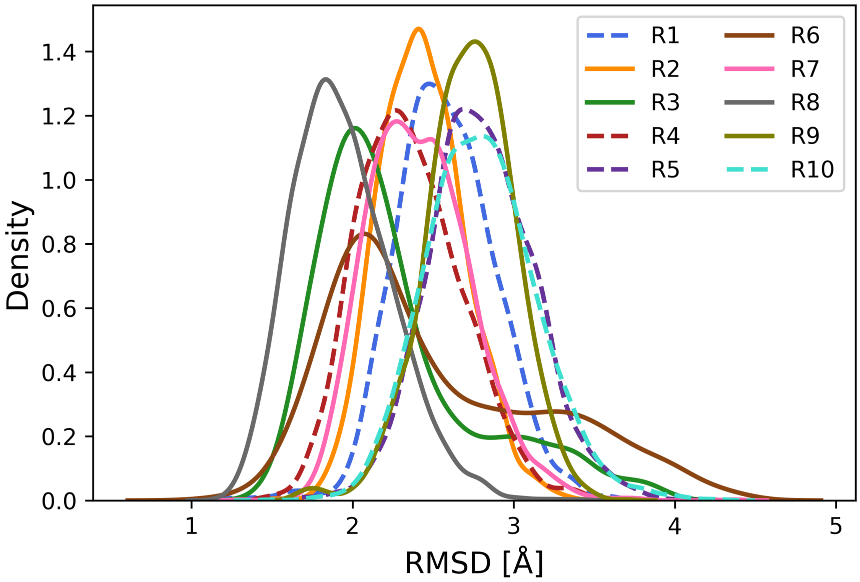


Figure A.2: Distribution plot of the RMSD of the core atoms of the L50 subdomain after alignment on the U50 subdomain in the ten replicas of MyosinVI.

N-terminal subdomain vs L50

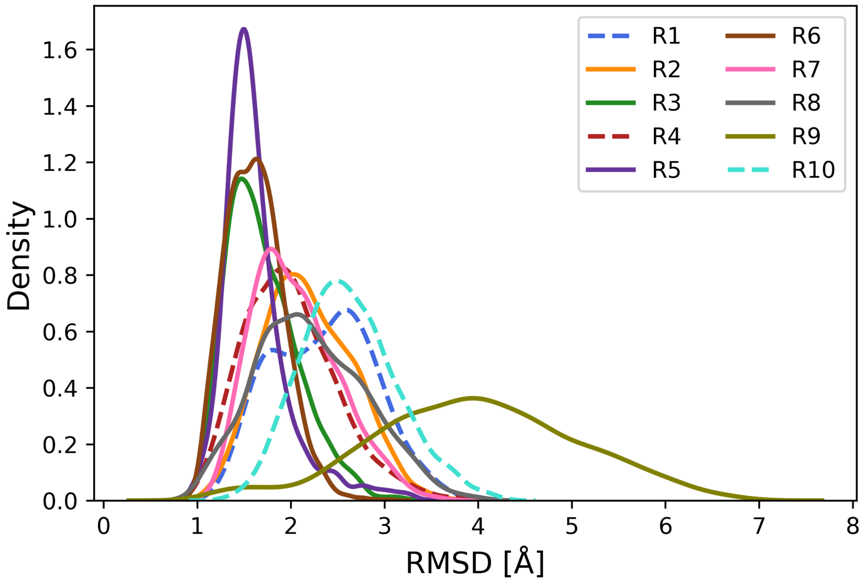


Figure A.3: Distribution plot of the RMSD of the core atoms of the N-terminal subdomain after alignment on the L50 subdomain in the ten replicas of MyosinVI.

N-terminal subdomain vs U50

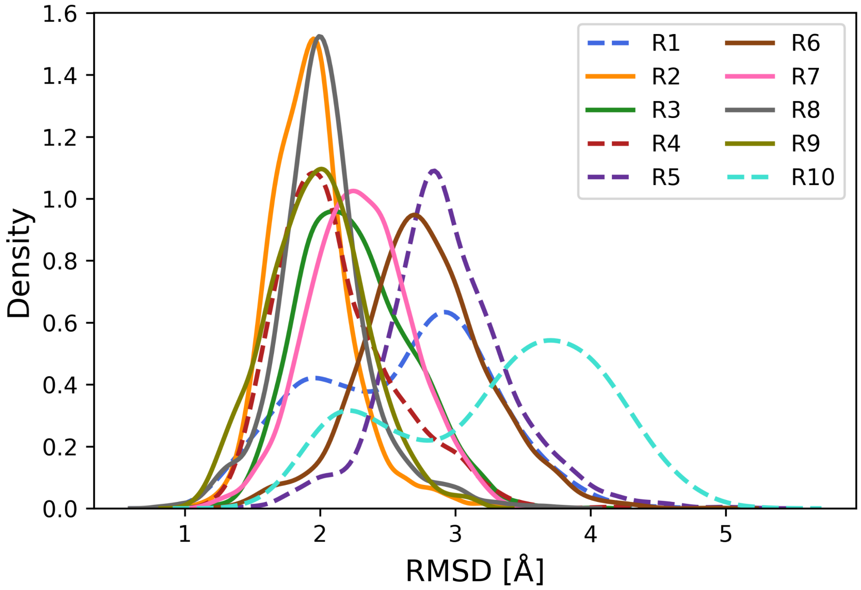


Figure A.4: Distribution plot of the RMSD of the core atoms of the N-terminal subdomain after alignment on the U50 subdomain in the ten replicas of MyosinVI.

U50 subdomain vs L50

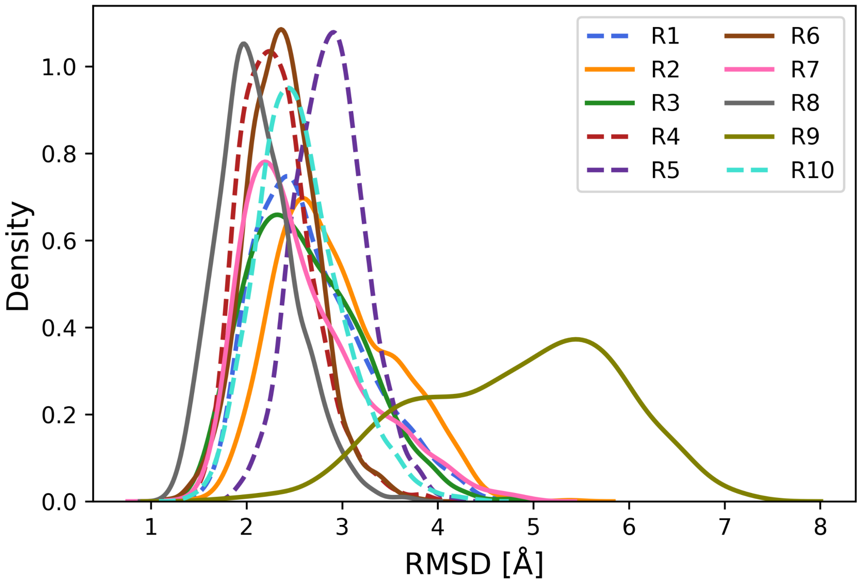


Figure A.5: Distribution plot of the RMSD of the core atoms of the N-terminal subdomain after alignment on the U50 subdomain in the ten replicas of MyosinVI.

U50 subdomain vs N-terminal

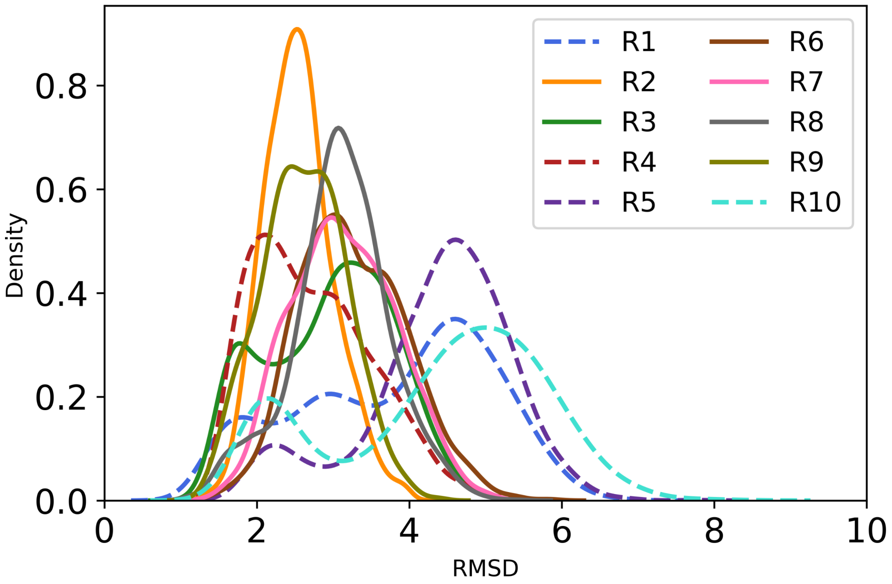


Figure A.6: Distribution plot of the RMSD of the core atoms of the N-terminal subdomain after alignment on the U50 subdomain in the ten replicas of MyosinVI.

A.3 CONTACT ADP - N98 IN THE TEN REPLICAS

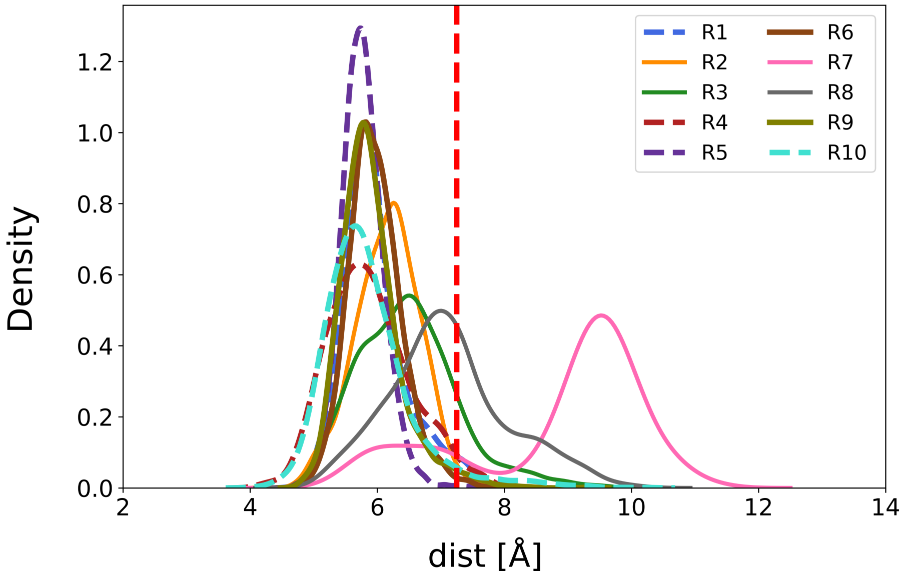


Figure A.7: Distribution plot of the distance between ADP (N6) and N98 (residue of the N-terminal subdomain) in the ten replicas of 200 ns of MyosinVI.

A.4 SUMMARY MD SIMULATIONS PERFORMED

Table A.2: List MD simulations performed

Myosin's state	**MD simulation's length**
PPS with standard protonation state of titratable residues, at pH 7.4	200 ns (CHARMMff)
PPS with non-standard protonation state of titratable residues, at pH 7.4	10 replicas of 200 ns each (CHARMMff)
PTS	10 replicas of 100 ns each (CHARMMff)
PR	100 ns (CHARMMff)

B

CARDIAC MYOSIN

B.1 DEFINITION CORE ATOMS SUBDOMAINS IN β CARDIAC MYOSIN (β-CM)

Table B.1: Definition core atoms for the various subdomains in β-CM

Subdomain	Core atoms
N-terminal	resid 79 to 82 or resid 87to90 or resid 95 to 123 or resid 133 to 140 or resid 151 to 165 or resid 169 to 175 or resid 181 to 199 or resid 663 to 668
U50	resid 208 to 228 or resid 235 to 247 or resid 252 to 261 or resid 265 to 269 or resid 278 to 296 or resid 300 to 304 or resid 319 to 361 or resid 375 to 394 or resi 411 to 442 or resid 601 to 618
L50	resid 404 to 499 or resid 511 to 520 or resid 526 to 534 or resid 539 to 551 or resid 557 to 559 or resid 571 to 583 or resid 587 to 593 or resid 640 to 661
Converter	resid 712 to 785

B.2 SUMMARY MD SIMULATIONS PERFORMED

Table B.2: List MD simulations performed

Myosin's system	**MD simulation's length**
PPS with standard protonation state of titratable residues, at pH 7.4, with OM bound	375 ns (Amberff)
PPS with standard protonation state of titratable residues, at pH 7.4, w/o OM	330 ns and two replicas of 200ns each (Amberff)
PPS with standard protonation state of titratable residues, at pH 7.4, w/o OM	330 ns (CHARMMff)
PPS with non-standard protonation state of titratable residues, at pH 7.4, w/o OM	300 ns (CHARMMff)
PPS with non-standard protonation state of titratable residues, at pH 7.4, with OM	300 ns (CHARMMff)

BIBLIOGRAPHY

Abraham, M. J., T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindah (2015). "Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers." In: *SoftwareX* 1-2, pp. 19–25. DOI: `10.1016/j.softx.2015.06.001`.

Adams, K. F., G. C. Fonarow, C. L. Emerman, T. H. LeJemtel, M. R. Costanzo, W. T. Abraham, R. L. Berkowitz, M. Galvao, and D. P. Horton (2005). "Characteristics and outcomes of patients hospitalized for heart failure in the United States: Rationale, design, and preliminary observations from the first 100,000 cases in the Acute Decompensated Heart Failure National Registry (ADHERE)." In: *American Heart Journal* 149.2, pp. 209–216. DOI: `10.1016/j.ahj.2004.08.005`.

Adasme, M. F., K. L. Linnemann, S. N. Bolz, F. Kaiser, S. Salentin, V. J. Haupt, and M. Schroeder (2021). "PLIP 2021: expanding the scope of the protein-ligand interaction profiler to DNA and RNA." In: *Nucleic Acids Research* 49. DOI: `10.1093/nar/gkab294`.

Afzal, M. (2021). "Recent updates on novel therapeutic targets of cardiovascular diseases." In: *Molecular and Cellular Biochemistry* 476.1, pp. 145–155. DOI: `10.1007/s11010-020-03891-8`.

Ahmed, Z. M., R. J. Morell, S. Riazuddin, A. Gropman, S. Shaukat, M. M. Ahmad, S. A. Mohiddin, L. Fananapazir, R. C. Caruso, T. Husnain, S. N. Khan, S. Riazuddin, A. J. Griffith, T. B. Friedman, and E. R. Wilcox (2003). "Mutations of MYO6 are associated with recessive deafness, DFNB37." In: *American Journal of Human Genetics* 72.5, pp. 1315–1322. DOI: `10.1086/375122`.

AL-Khayat, H. A. (2013). "Three-dimensional structure of the human myosin thick filament: Clinical implications." In: *Global Cardiology Science and Practice* 2013.3, p. 36. DOI: `10.5339/gcsp.2013.36`.

Al-Numair, N. S., L. Lopes, P. Syrris, L. Monserrat, P. Elliott, and A. C. Martin (2016). "The structural effects of mutations can aid in differential phenotype prediction of beta-myosin heavy chain (Myosin-7) missense variants." In: *Bioinformatics* 32.19, pp. 2947–2955. DOI: `10.1093/bioinformatics/btw362`.

Alhossary, A., S. D. Handoko, Y. Mu, and C. K. Kwoh (2015). "Fast, accurate, and reliable molecular docking with QuickVina 2." In: *Bioinformatics* 31.13, pp. 2214–2216. DOI: `10.1093/bioinformatics/btv082`.

Allingham, J. S., R. Smith, and I. Rayment (2005). "The structural basis of blebbistatin inhibition and specificity for myosin II." In: *Nature Structural and Molecular Biology* 12.4, pp. 378–379. DOI: `10.1038/nsmb908`.

Altman, D., H. L. Sweeney, and J. A. Spudich (2004). "The mechanism of myosin VI translocation and its load-induced anchoring." In: *Cell* 116.5, pp. 737–749. DOI: `10.1016/S0092-8674(04)00211-9`.

ŌMura, S., Y. Sasaki, Y. Iwai, and H. Takeshima (1995). *Staurosporine, a Potentially Important Gift from a Microorganism*. Tech. rep. 7, pp. 535–548. DOI: `10.7164/antibiotics.48.535`.

Åqvist, J., P. Wennerström, M. Nervall, S. Bjelic, and B. O. Brandsdal (2004). "Molecular dynamics simulations of water and biomolecules with a Monte Carlo constant pressure algorithm." In: *Chemical Physics Letters* 384.4-6, pp. 288–294. DOI: `10.1016/j.cplett.2003.12.039`.

Arnaud, C. H. (2006). *Thermo fisher scientific*. DOI: `10.1021/cen-v084n048.p011`.

Attwood, M. M., D. Fabbro, A. V. Sokolov, S. Knapp, and H. B. Schiöth (2021). "Trends in kinase drug discovery: targets, indications and inhibitor design." In: *Nature Reviews Drug Discovery* 20.11, pp. 839–861. DOI: `10.1038/s41573-021-00252-y`.

Avraham, K. B., T. Hasson, K. P. Steel, D. M. Kingsley, L. B. Russell, M. S. Mooseker, N. G. Copeland, and N. A. Jenkins (1995). "The mouse snell's waltzer deafness gene encodes an unconventional myosin required for structural integrity of inner ear hair cells." In: *Nature Genetics* 11.4, pp. 369–375. DOI: `10.1038/ng1295-369`.

Baell, J. B. and G. A. Holloway (2010). "New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays." In: *Journal of Medicinal Chemistry* 53.7, pp. 2719–2740. DOI: `10.1021/jm901137j`.

Bahar (2010). "Mechanisms of Membrane Proteins." In: *Chemical Reviews* 110.3, pp. 1463–1497. DOI: `10.1021/cr900095e.Normal`.

Barrick, S. K. and M. J. Greenberg (2021). "Cardiac myosin contraction and mechanotransduction in health and disease." In: *Journal of Biological Chemistry*, p. 101297. DOI: `10.1016/j.jbc.2021.101297`.

Batters, C. and C. Veigel (2016). "Mechanics and Activation of Unconventional Myosins." In: *Traffic* 17.8, pp. 860–871. DOI: `10.1111/tra.12400`.

Bayly, C. I., P. Cieplak, W. D. Cornell, and P. A. Kollman (1993). *A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model*. Tech. rep., pp. 10269–10280.

Bencini, A. and V. Lippolis (2010). "1,10-Phenanthroline: A versatile building block for the construction of ligands for various purposes." In: *Coordination Chemistry Reviews* 254.17-18, pp. 2096–2180. DOI: `10.1016/j.ccr.2010.04.008`.

Bender, B. J., S. Gahbauer, A. Luttens, J. Lyu, C. M. Webb, R. M. Stein, E. A. Fink, T. E. Balius, J. Carlsson, J. J. Irwin, and B. K. Shoichet (2021). *A practical guide to large-scale docking*. DOI: `10.1038/s41596-021-00597-z`.

Benjamin, E. J., P. Muntner, A. Alonso, M. S. Bittencourt, C. W. Callaway, A. P. Carson, A. M. Chamberlain, A. R. Chang, S. Cheng, S. R. Das, F. N. Delling, L. Djousse, M. S. Elkind, J. F. Ferguson, M. Fornage, L. C. Jordan, S. S. Khan, B. M. Kissela, K. L. Knutson, T. W. Kwan, D. T. Lackland, T. T. Lewis, J. H. Lichtman, C. T. Longenecker, M. S. Loop, P. L. Lutsey, S. S. Martin, K. Matsushita, A. E. Moran, M. E. Mussolino, M. O'Flaherty, A. Pandey, A. M. Perak, W. D. Rosamond, G. A. Roth, U. K. Sampson, G. M. Satou, E. B. Schroeder, S. H. Shah, N. L. Spartano, A. Stokes, D. L. Tirschwell, C. W. Tsao, M. P. Turakhia, L. B. VanWagner, J. T. Wilkins, S. S. Wong, and S. S. Virani (2019). *Heart Disease and Stroke Statistics-2019 Update: A Report From the American Heart Association*. Vol. 139. 10, e56–e528. DOI: `10.1161/CIR.0000000000000659`.

Bentham Science Publisher, B. S. P. (2006). "Scoring Functions for Protein-Ligand Docking." In: *Current Protein Peptide Science* 7.5, pp. 407–420. DOI: `10.2174/138920306778559395`.

Berendsen, H. J., J. P. Postma, W. F. Van Gunsteren, A. Dinola, and J. R. Haak (1984). "Molecular dynamics with coupling to an external bath." In: *The Journal of Chemical Physics* 81.8, pp. 3684–3690. DOI: `10.1063/1.448118`.

Berthold, M. R., N. Cebron, F. Dill, T. R. Gabriel, T. Kötter, T. Meinl, P. Ohl, C. Sieb, K. Thiel, and B. Wiswedel (2007). "KNIME: The Konstanz Information Miner." In:

Best, R. B., X. Zhu, J. Shim, P. E. Lopes, J. Mittal, M. Feig, and A. D. MacKerell (2012). "Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain $\chi 1$ and $\chi 2$ Dihedral Angles." In: *Journal of Chemical Theory and Computation* 8.9, pp. 3257–3273. DOI: `10.1021/ct300400x`.

Bissantz, C., B. Kuhn, and M. Stahl (2010). "A medicinal chemist's guide to molecular interactions." In: *Journal of Medicinal Chemistry* 53.14, pp. 5061–5084. DOI: `10.1021/jm100112j`.

Blanc, F. (2018). "Exploration de la transduction chimio- myosine par simulations numériques." In: *Ecole Doctorale ED222*.

Blanc, F., T. Isabet, H. Benisty, H. L. Sweeney, M. Cecchini, and A. Houdusse (2018). "An intermediate along the recovery stroke of myosin VI revealed by X-ray crystallography and molecular dynamics." In: *Proceedings of the National Academy of Sciences of the United States of America* 115.24, pp. 6213–6218. DOI: `10.1073/pnas.1711512115`.

Bloom, M. W., B. Greenberg, T. Jaarsma, J. L. Januzzi, C. S. Lam, A. P. Maggioni, J. N. Trochu, and J. Butler (2017). "Heart failure with reduced ejection fraction." In: *Nature Reviews Disease Primers* 3. DOI: `10.1038/nrdp.2017.58`.

Bohr, C., K. Hasselbalch, and A. Krogh (1904). "Ueber einen in biologischer

Beziehung wichtigen Einfluss, den die Kohlensäurespannung des Blutes auf dessen Sauerstoffbindung übt1." In: *Skandinavisches Archiv Für Physiologie* 16.2, pp. 402–412. DOI: `10.1111/j.1748-1716.1904.tb01382.x`.

Bolhuis, P. G., D. Chandler, C. Dellago, and P. L. Geissler (2002). "Transition path sampling: Throwing ropes over rough mountain passes, in the dark." In: *Annual Review of Physical Chemistry* 53, pp. 291–318. DOI: `10.1146/annurev.physchem.53.082301.113146`.

Bond, L. M., D. A. Tumbarello, J. Kendrick-Jones, and F. Buss (2013). "Small-molecule inhibitors of myosin proteins." In: *Future Medicinal Chemistry* 5.1, pp. 41–52. DOI: `10.4155/fmc.12.185`.

Boström, J., A. Hogner, and S. Schmitt (2006). "Do structurally similar ligands bind in a similar fashion?" In: *Journal of Medicinal Chemistry* 49.23, pp. 6716–6725. DOI: `10.1021/jm060167o`.

Branduardi, D. and J. D. Faraldo-Gómez (2013). "String method for calculation of minimum free-energy paths in cartesian space in freely tumbling systems." In: *Journal of Chemical Theory and Computation* 9.9, pp. 4140–4154. DOI: `10.1021/ct400469w`.

Brawley, C. M. and R. S. Rock (2009). "Unconventional myosin traffic in cells reveals a selective actin cytoskeleton." In: *Proceedings of the National Academy of Sciences of the United States of America* 106.24, pp. 9685–9690. DOI: `10.1073/pnas.0810451106`.

Brooks, B. R., C. L. Brooks, A. D. Mackerell, 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. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus (2009). "CHARMM: The biomolecular simulation program." In: *Journal of Computational Chemistry* 30.10, pp. 1545–1614. DOI: `10.1002/jcc.21287`.

Bryant, Z., D. Altman, and J. A. Spudich (2007). "The power stroke of myosin VI and the basis of reverse directionality." In: *Proceedings of the National Academy of Sciences of the United States of America* 104.3, pp. 772–777. DOI: `10.1073/pnas.0610144104`.

Buss, F. and J. Kendrick-Jones (2008). "How are the cellular functions of myosin VI regulated within the cell?" In: *Biochemical and Biophysical Research Communications* 369.1, pp. 165–175. DOI: `10.1016/j.bbrc.2007.11.150`.

Buss, F., G. Spudich, and J. Kendrick-Jones (2004). "Myosin VI: Cellular functions and motor properties." In: *Annual Review of Cell and Developmental Biology* 20, pp. 649–676. DOI: `10.1146/annurev.cellbio.20.012103.094243`.

Bussi, G., D. Donadio, and M. Parrinello (2007). "Canonical sampling through velocity rescaling." In: *Journal of Chemical Physics* 126.1. DOI: `10.1063/1.2408420`. arXiv: `0803.4060`.

Capelli, R., W. Lyu, V. Bolnykh, S. Meloni, J. M. H. Olsen, U. Rothlisberger, M. Parrinello, and P. Carloni (2020). "Accuracy of Molecular Simulation-Based Predictions of koffValues: A Metadynamics Study." In: *Journal of Physical Chemistry Letters* 11.15, pp. 6373–6381. DOI: `10.1021/acs.jpclett.0c00999`.

Casalino, L., A. C. Dommer, Z. Gaieb, E. P. Barros, T. Sztain, S. H. Ahn, A. Trifan, A. Brace, A. T. Bogetti, A. Clyde, H. Ma, H. Lee, M. Turilli, S. Khalid, L. T. Chong, C. Simmerling, D. J. Hardy, J. D. Maia, J. C. Phillips, T. Kurth, A. C. Stern, L. Huang, J. D. McCalpin, M. Tatineni, T. Gibbs, J. E. Stone, S. Jha, A. Ramanathan, and R. E. Amaro (2021). "AI-driven multiscale simulations illuminate mechanisms of SARS-CoV-2 spike dynamics." In: *International Journal of High Performance Computing Applications* 35.5, pp. 432–451. DOI: `10.1177/10943420211006452`.

Case, D. A. (1994). "Normal mode analysis of protein dynamics." In: *Current Opinion in Structural Biology* 4.2, pp. 285–290. DOI: `10.1016/S0959-440X(94)90321-2`.

Case, D. A., H. M. Aktulga, K. Belfon, I. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. E. Cheatham III, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, et al. (2021). *Amber 2021*. University of California, San Francisco.

Case, D. A., R. C. Walker, T. E. Cheatham, C. Simmerling, A. Roitberg, K. M. Merz, R. Luo, and T. Darden (2018). "Amber 2018." In: *University of California, San Francisco. 2018*, pp. 1–923.

Cecchini, M., A. Houdusse, and M. Karplus (2008). "Allosteric communication in myosin V: From small conformational changes to large directed movements." In: *PLoS Computational Biology* 4.8. DOI: `10.1371/journal.pcbi.1000129`.

Cecchini, M., Y. Alexeev, and M. Karplus (2010). "Pi Release from Myosin: A Simulation Analysis of Possible Pathways." In: *Structure* 18.4, pp. 458–470. DOI: `10.1016/j.str.2010.01.014`.

Cecchini, M. and J. P. Changeux (2022). *Nicotinic receptors: From protein allostery to computational neuropharmacology*. DOI: `10.1016/j.mam.2021.101044`.

Chatzigoulas, A. and Z. Cournia (2021). "Rational design of allosteric modulators: Challenges and successes." In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* June 2020, pp. 1–49. DOI: `10.1002/wcms.1529`.

Cheung, A., J. A. Dantzig, S. Hollingworth, S. M. Baylor, Y. E. Goldman, T. J. Mitchinson, and A. F. Straight (2002). "A small-molecule inhibitor of skeletal muscle myosin II." In: *Nature Cell Biology* 4.1, pp. 83–88. DOI: `10.1038/ncb734`.

Chuang, C., S. Collibee, L. Ashcraft, W. Wang, M. Vander Wal, X. Wang, D. T. Hwee, Y. Wu, J. Wang, E. R. Chin, P. Cremin, J. Zamora, J. Hartman, J. Schaletzky, E. Wehri, L. A. Robertson, F. I. Malik, and B. P. Morgan (2021). "Discovery of Aficamten (CK-274), a Next-Generation Cardiac Myosin Inhibitor for the Treatment of Hypertrophic Cardiomyopathy." In: *Journal of Medicinal Chemistry* 64.19, pp. 14142–14152. DOI: `10.1021/acs.jmedchem.1c01290`.

Clarke, D., A. Sethi, S. Li, S. Kumar, R. W. Chang, J. Chen, and M. Gerstein (2016). "Identifying Allosteric Hotspots with Dynamics: Application to Inter- and Intra-species Conservation." In: *Structure* 24.5, pp. 826–837. DOI: `10.1016/j.str.2016.03.008`.

Cleland, J. G., A. Torabi, and N. K. Khan (2005). "Epidemiology and management of heart failure and left ventricular systolic dysfunction in the aftermath of a myocardial infarction." In: *Heart* 91.SUPPL. 2, pp. 7–13. DOI: `10.1136/hrt.2005.062026`.

Clippinger, S. R., P. E. Cloonan, L. Greenberg, M. Ernst, W. T. Stump, and M. J. Greenberg (2019). "Disrupted mechanobiology links the molecular and cellular phenotypes in familial dilated cardiomyopathy." In: *Proceedings of the National Academy of Sciences of the United States of America* 116.36, pp. 17831–17840. DOI: `10.1073/pnas.1910962116`.

Colvin, M., N. K. Sweitzer, N. M. Albert, R. Krishnamani, M. W. Rich, W. G. Stough, M. N. Walsh, C. A. Westlake Canary, L. A. Allen, M. R. Bonnell, P. E. Carson, M. C. Chan, M. G. Dickinson, D. L. Dries, G. A. Ewald, J. C. Fang, A. F. Hernandez, R. E. Hershberger, S. D. Katz, S. Moore, J. E. Rodgers, J. G. Rogers, A. R. Vest, D. J. Whellan, and M. M. Givertz (2015). "Heart Failure in Non-Caucasians, Women, and Older Adults: A White Paper on Special Populations from the Heart Failure Society of America Guideline Committee." In: *Journal of Cardiac Failure* 21.8, pp. 674–693. DOI: `10.1016/j.cardfail.2015.05.013`.

Copeland, R. A. (2016). "The drug–target residence time model: a 10-year retrospective." In: 15.February, pp. 87–95.

Copeland, R. A., D. L. Pompliano, and T. D. Meek (2006). "Drug-target residence time and its implications for lead optimization." In: *Nature Reviews Drug Discovery* 5.9, pp. 730–739. DOI: `10.1038/nrd2082`.

Corzo, J. and M. Santamaria (2006). "Time, the forgotten dimension of ligand binding teaching." In: *Biochemistry and Molecular Biology Education* 34.6, pp. 413–416. DOI: `10.1002/bmb.2006.494034062678`.

Coureux, P. D., H. L. Sweeney, and A. Houdusse (2004). "Three myosin V structures delineate essential features of chemo-mechanical transduction." In: *EMBO Journal* 23.23, pp. 4527–4537. DOI: `10.1038/sj.emboj.7600458`.

Cournia, Z., B. K. Allen, T. Beuming, D. A. Pearlman, B. K. Radak, and W. Sherman (2020). "Rigorous Free Energy Simulations in Virtual Screening." In: *Cite This: J. Chem. Inf. Model* 2020, pp. 4153–4169. DOI: `10.1021/acs.jcim.0c00116`.

Cournia, Z., B. Allen, and W. Sherman (2017). "Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations." In: *Journal of Chemical Information and Modeling* 57.12, pp. 2911–2937. DOI: `10.1021/acs.jcim.7b00564`.

Cramer, L. P. (2000). *Myosin VI: Roles for a minus end-directed actin motor in cells*. Tech. rep. 6, pp. 121–126. DOI: `10.1083/jcb.150.6.F121`.

Cui, Q. and M. Karplus (2008). "Allostery and cooperativity revisited." In: *Pro-*

tein Science 17.8, pp. 1295–1307. DOI: 10.1110/ps.03259908.

Daura, X., K. Gademann, B. Jaun, D. Seebach, W. F. Van Gunsteren, and A. E. Mark (1999). “Peptide folding: When simulation meets experiment.” In: *Angewandte Chemie - International Edition* 38.1-2, pp. 236–240. DOI: 10.1002/(sici)1521-3773(19990115)38:1/2<236::aid-anie236>3.0.co;2-m.

Day, S. M., J. C. Tardiff, and E. Michael Ostap (2022). “Myosin modulators: Emerging approaches for the treatment of cardiomyopathies and heart failure.” In: *Journal of Clinical Investigation* 132.5. DOI: 10.1172/JCI148557.

De La Cruz, E. M. and E. Michael Ostap (2009). “Chapter 6 Kinetic and Equilibrium Analysis of the Myosin ATPase.” In: *Methods in Enzymology* 455.A, pp. 157–192. DOI: 10.1016/S0076-6879(08)04206-7.

De la Cruz, E. M., E. M. Ostap, and H. L. Sweeney (2001). “Kinetic Mechanism and Regulation of Myosin VI.” In: *Journal of Biological Chemistry* 276.34, pp. 32373–32381. DOI: 10.1074/jbc.M104136200.

De La Cruz, E. M., A. L. Wells, S. S. Rosenfeld, E. M. Ostap, and H. L. Sweeney (1999). “The kinetic mechanism of myosin V.” In: *Proceedings of the National Academy of Sciences of the United States of America* 96.24, pp. 13726–13731. DOI: 10.1073/pnas.96.24.13726.

Dellago, C., P. G. Bolhuis, F. S. Csajka, and D. Chandler (1998). “Transition path sampling and the calculation of rate constants.” In: *Journal of Chemical Physics* 108.5, pp. 1964–1977. DOI: 10.1063/1.475562.

Dickstein, K., A. Cohen-Solal, G. Filippatos, J. J. McMurray, P. Ponikowski, P. A. Poole-Wilson, A. Strömberg, D. J. van Veldhuisen, D. Atar, A. W. Hoes, A. Keren, A. Mebazaa, M. Nieminen, S. G. Priori, K. Swedberg, A. Vahanian, J. Camm, R. De Caterina, V. Dean, K. Dickstein, C. Funck-Brentano, I. Hellemans, S. D. Kristensen, K. McGregor, U. Sechtem, S. Silber, M. Tendera, P. Widimsky, J. L. Zamorano, A. Auricchio, J. Bax, M. Böhm, U. Corrà, P. della Bella, P. M. Elliott, F. Follath, M. Gheorghiade, Y. Hasin, A. Hernborg, T. Jaarsma, M. Komajda, R. Kornowski, M. Piepoli, B. Prendergast, L. Tavazzi, J. L. Vachiery, F. W. Verheugt, and F. Zannad (2008). “ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2008. The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2008 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association of the ESC (HFA) and endorsed by the European Society of Intensive Care Medicine (ESICM).” In: *European Journal of Heart Failure* 10.10, pp. 933–989. DOI: 10.1016/j.ejheart.2008.08.005.

Dima, R. I. (2006). “Determination of network of residues that regulate allostery in protein families using sequence analysis.” In: *Protein Science* 15.2, pp. 258–268. DOI: 10.1110/ps.051767306.

Dixon, T., S. D. Lotz, and A. Dickson (2018). “Predicting ligand binding affinity using on- and off-rates for the SAMPL6 SAMPLing challenge.” In: *Journal of Computer-Aided Molecular Design* 32.10, pp. 1001–1012. DOI: 10.1007/s10822-018-0149-3.

Dörr, M. and E. Meggers (2014). *Metal complexes as structural templates for targeting proteins*. DOI: 10.1016/j.cbpa.2014.01.005.

Downward, J. (2003). *Targeting RAS signalling pathways in cancer therapy*. DOI: 10.1038/nrc969.

Du, X., Y. Li, Y. L. Xia, S. M. Ai, J. Liang, P. Sang, X. L. Ji, and S. Q. Liu (2016). “Insights into protein–ligand interactions: Mechanisms, models, and methods.” In: *International Journal of Molecular Sciences* 17.2. DOI: 10.3390/ijms17020144.

Dunn, A. R., P. Chuan, Z. Bryant, and J. A. Spudich (2010). “Contribution of the myosin VI tail domain to processive stepping and intramolecular tension sensing.” In: *Proceedings of the National Academy of Sciences of the United States of America* 107.17, pp. 7746–7750. DOI: 10.1073/pnas.1002430107.

Dunn, T. A., S. Chen, D. A. Faith, J. L. Hicks, E. A. Platz, Y. Chen, C. M. Ewing, J. Sauvageot, W. B. Isaacs, A. M. De Marzo, and J. Luo (2006). “A novel role of myosin VI in human prostate cancer.” In: *American Journal of Pathology* 169.5, pp. 1843–1854. DOI: 10.2353/ajpath.2006.060316.

Ebashi, S. (1963). “Third component participating in the super precipitation of ’natural actomyosin’.” In: *Nature* 200.4910, p. 1010. DOI: 10.1038/2001010a0.

Eberhardt, J., D. Santos-Martins, A. F. Tillack, and S. Forli (2021). *No Title*. DOI: `10.1021/acs.jcim.1c00203`.

Elting, M. W., Z. Bryant, J. C. Liao, and J. A. Spudich (2011). "Detailed tuning of structure and intramolecular communication are dispensable for processive motion of myosin VI." In: *Biophysical Journal* 100.2, pp. 430–439. DOI: `10.1016/j.bpj.2010.11.045`.

England, J. and S. Loughna (2013). "Heavy and light roles: Myosin in the morphogenesis of the heart." In: *Cellular and Molecular Life Sciences* 70.7, pp. 1221–1239. DOI: `10.1007/s00018-012-1131-1`.

Evans, D. J. and B. L. Holian (1985). "The Nose-Hoover thermostat." In: *The Journal of Chemical Physics* 83.8, pp. 4069–4074. DOI: `10.1063/1.449071`.

Fan, F., Z. He, L. L. Kong, Q. Chen, Q. Yuan, S. Zhang, J. Ye, H. Liu, X. Sun, J. Geng, L. Yuan, L. Hong, C. Xiao, W. Zhang, X. Sun, Y. Li, P. Wang, L. Huang, X. Wu, Z. Ji, Q. Wu, N. S. Xia, N. S. Gray, L. Chen, C. H. Yun, X. Deng, and D. Zhou (2016). "Pharmacological targeting of kinases MST1 and MST2 augments tissue repair and regeneration." In: *Science Translational Medicine* 8.352. DOI: `10.1126/scitranslmed.aaf2304`.

Fan, M., J. Wang, H. Jiang, Y. Feng, M. Mahdavi, K. Madduri, M. T. Kandemir, and N. V. Dokholyan (2021). "GPU-Accelerated Flexible Molecular Docking." In: *Journal of Physical Chemistry B* 125.4, pp. 1049–1060. DOI: `10.1021/acs.jpcb.0c09051`.

Fernandes, S., A. Oikonomopoulos, J.-M. S. K, A.-S. Roozbeh, H. Marcus, S. Marius, G. Qing-Fen, A. R. L, and d. R. C. L (2019). *MYK-491, a Novel Small-Molecule Cardiac Myosin Activator Increases Cardiac Systolic Function and Preserves Mechanical Efficiency: Pre-Clinical in vivo and in vitro Evidence*.

Finer, J. T., R. M. Simmons, and J. A. Spudich (1994). *Single myosin molecule mechanics: Piconewton forces and nanometre steps*. Tech. rep. 6467, pp. 113–119. DOI: `10.1038/368113a0`.

Fischer, A., M. Smies, M. Sellner, and M. A. Lill (2021). "2021, 64, 24892500 Downloaded via UNIV STRASBOURG on September 14." In: *J. Med. Chem* 14, p. 30. DOI: `10.1021/acs.jmedchem.0c02227`.

Fischer, S., B. Windshügel, D. Horak, K. C. Holmes, and J. C. Smith (2005). *Structural mechanism of the recovery stroke in the myosin molecular motor*. Tech. rep. 19, pp. 6873–6878. DOI: `10.1073/pnas.0408784102`.

Fisher, A. J., C. A. Smith, J. Thoden, R. Smith, K. Sutoh, H. M. Holden, I. Rayment, M. Reedy, C. Schutt, Y. Goldman, T. Yanagida, D. Martyn, B. Brenner, E. Reisler, M. Geeves, R. Cooke, K. Holmes, S. Highsmith, and M. Kawai (1995). *Structural studies of myosin:nucleotide complexes: A revised model for the molecular basis of muscle contraction*. Tech. rep. 4 SUPPL.

Fletcher, R. (2000). *Practical Methods of Optimization*. Chichester, West Sussex England: John Wiley Sons, Ltd. DOI: `10.1002/9781118723203`.

Foth, B. J., M. C. Goedecke, and D. Soldati (2006). *New insights into myosin evolution and classification*. Tech. rep. 10, pp. 3681–3686. DOI: `10.1073/pnas.0506307103`.

Franz, P., W. Ewert, M. Preller, and G. Tsiavaliaris (2020). "Unraveling a Force-Generating Allosteric Pathway of Actomyosin Communication Associated with ADP and P i Release." In: *International Journal of Molecular Sciences Article*. DOI: `10.3390/ijms22010104`.

Friesner, R. A., J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, and P. S. Shenkin (2004). "Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy." In: *Journal of Medicinal Chemistry* 47.7, pp. 1739–1749. DOI: `10.1021/jm0306430`.

Fukunishi, H., O. Watanabe, and S. Takada (2002). "On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction." In: *Journal of Chemical Physics* 116.20, pp. 9058–9067. DOI: `10.1063/1.1472510`.

Gani, O. A. and R. A. Engh (2010). "Protein kinase inhibition of clinically important staurosporine analogues." In: *Natural Product Reports* 27.4, pp. 489–498. DOI: `10.1039/b923848b`.

Gapsys, V., A. Yildirim, M. Aldeghi, Y. Khalak, D. van der Spoel, and B. L. de Groot (2021). "Accurate absolute free energies for ligand–protein binding based on non-equilibrium approaches." In: *Communications Chem-*

istry 4.1. DOI: `10.1038/s42004-021-00498-y`.

Gaulton, A., A. Hersey, M. L. Nowotka, A. Patricia Bento, J. Chambers, D. Mendez, P. Mutowo, F. Atkinson, L. J. Bellis, E. Cibrian-Uhalte, M. Davies, N. Dedman, A. Karlsson, M. P. Magariños, J. P. Overington, G. Papadatos, I. Smit, and A. R. Leach (2017). "The ChEMBL database in 2017." In: *Nucleic Acids Research* 45.D1, pp. D945–D954. DOI: `10.1093/nar/gkw1074`.

Gawehn, E., J. A. Hiss, and G. Schneider (2016). "Deep Learning in Drug Discovery." In: *Molecular Informatics* 35.1, pp. 3–14. DOI: `10.1002/minf.201501008`.

Gaziano, T. A., A. Bitton, S. Anand, S. Abrahams-Gessel, and A. Murphy (2010). "Growing Epidemic of Coronary Heart Disease in Low- and Middle-Income Countries." In: *Current Problems in Cardiology* 35.2, pp. 72–115. DOI: `10.1016/j.cpcardiol.2009.10.002`.

Geeves, M. A. and K. C. Holmes (1999). "Structural mechanism of muscle contraction." In: *Annual Review of Biochemistry* 68.1, pp. 687–728. DOI: `10.1146/annurev.biochem.68.1.687`.

Genheden, S. and U. Ryde (2015). "The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities." In: *Expert Opinion on Drug Discovery* 10.5, pp. 449–461. DOI: `10.1517/17460441.2015.1032936`.

Gentile, F., J. C. Yaacoub, J. Gleave, M. Fernandez, A. T. Ton, F. Ban, A. Stern, and A. Cherkasov (2022). *Artificial intelligence–enabled virtual screening of ultra-large chemical libraries with deep docking*. DOI: `10.1038/s41596-021-00659-2`.

Giannoni, A., M. Emdin, F. Bramanti, G. Iudice, D. P. Francis, A. Barsotti, M. Piepoli, and C. Passino (2009). "Combined Increased Chemosensitivity to Hypoxia and Hypercapnia as a Prognosticator in Heart Failure." In: *Journal of the American College of Cardiology* 53.21, pp. 1975–1980. DOI: `10.1016/j.jacc.2009.02.030`.

Gilson, M. K., J. A. Given, B. L. Bush, and J. A. McCammon (1997). "The statistical-thermodynamic basis for computation of binding affinities: A critical review." In: *Biophysical Journal* 72.3, pp. 1047–1069. DOI: `10.1016/S0006-3495(97)78756-3`.

Gohlke, H., M. Hendlich, and G. Klebe (2000). "Knowledge-based scoring function to predict protein-ligand interactions." In: *Journal of Molecular Biology* 295.2, pp. 337–356. DOI: `10.1006/jmbi.1999.3371`.

Gorgulla, C., A. Boeszoermenyi, Z. F. Wang, P. D. Fischer, P. W. Coote, K. M. Padmanabha Das, Y. S. Malets, D. S. Radchenko, Y. S. Moroz, D. A. Scott, K. Fackeldey, M. Hoffmann, I. Iavniuk, G. Wagner, and H. Arthanari (2020). "An open-source drug discovery platform enables ultra-large virtual screens." In: *Nature* 580.7805, pp. 663–668. DOI: `10.1038/s41586-020-2117-z`.

Gorgulla, C., A. Jayaraj, K. Fackeldey, and H. Arthanari (2022). "Emerging frontiers in virtual drug discovery: From quantum mechanical methods to deep learning approaches." In: *Current Opinion in Chemical Biology* 69, p. 102156. DOI: `10.1016/j.cbpa.2022.102156`.

Gourinath, S., D. M. Himmel, J. H. Brown, L. Reshetnikova, A. G. Szent-Györgyi, and C. Cohen (2003). "Crystal Structure of Scallop Myosin S1 in the Pre-Power Stroke State to 2. 6 Å Resolution: Flexibility and Function in the Head." In: *Structure* 11.12, pp. 1621–1627. DOI: `10.1016/j.str.2003.10.013`.

Greenberg, M. J., H. Shuman, and E. M. Ostap (2017). "Measuring the kinetic and mechanical properties of non-processive myosins using optical tweezers." In: *Methods in Molecular Biology* 1486, pp. 483–509. DOI: `10.1007/978-1-4939-6421-5_19`.

Greener, J. G. and M. J. Sternberg (2015). "AlloPred: Prediction of allosteric pockets on proteins using normal mode perturbation analysis." In: *BMC Bioinformatics* 16.1, p. 335. DOI: `10.1186/s12859-015-0771-1`.

Greener, J. G. and M. J. Sternberg (2018). "Structure-based prediction of protein allostery." In: *Current Opinion in Structural Biology* 50, pp. 1–8. DOI: `10.1016/j.sbi.2017.10.002`.

Gresser, M. J. and A. S. Tracey (1990). "Vanadates as Phosphate Analogs in Biochemistry." In: *Vanadium in Biological Systems*, pp. 63–79. DOI: `10.1007/978-94-009-2023-1_4`.

Grillo, M. P., J. C. Erve, R. Dick, J. P. Driscoll, N. Haste, S. Markova, P. Brun, T. J. Carlson, and M. Evanchik (2019). "In vitro and in vivo pharmacokinetic characterization of mavacamten, a first-in-class small molecule allosteric modulator of beta cardiac myosin." In:

Xenobiotica 49.6, pp. 718–733. DOI: `10.1080/00498254.2018.1495856`.

Grillo, M. P., S. Markova, M. Evanchik, M. Trellu, P. Moliner, P. Brun, A. Perreard-Dumaine, P. Vicat, J. P. Driscoll, and T. J. Carlson (2021). "Preclinical in vitro and in vivo pharmacokinetic properties of danicamtiv, a new targeted myosin activator for the treatment of dilated cardiomyopathy." In: *Xenobiotica* 51.2, pp. 222–238. DOI: `10.1080/00498254.2020.1839982`.

Gromski, P. S., A. B. Henson, J. M. Granda, and L. Cronin (2019). "How to explore chemical space using algorithms and automation." In: *Nature Reviews Chemistry* 3.2, pp. 119–128. DOI: `10.1038/s41570-018-0066-y`.

Gunasekaran, K., B. Ma, and R. Nussinov (2004). "Is allostery an intrinsic property of all dynamic proteins?" In: *Proteins: Structure, Function and Genetics* 57.3, pp. 433–443. DOI: `10.1002/prot.20232`.

Guo, D., T. Mulder-Krieger, A. P. IJzerman, and L. H. Heitman (2012). "Functional efficacy of adenosine A 2A receptor agonists is positively correlated to their receptor residence time." In: *British Journal of Pharmacology* 166.6, pp. 1846–1859. DOI: `10.1111/j.1476-5381.2012.01897.x`.

Ha, E. J., C. T. Lwin, and J. D. Durrant (2020). "LigGrep: a tool for filtering docked poses to improve virtual-screening hit rates." In: *Journal of Cheminformatics* 12.1. DOI: `10.1186/s13321-020-00471-2`.

Hamdani, N., V. Kooij, S. Van Dijk, D. Merkus, W. J. Paulus, C. D. Remedios, D. J. Duncker, G. J. Stienen, and J. Van Der Velden (2008). "Sarcomeric dysfunction in heart failure." In: *Cardiovascular Research* 77.4, pp. 649–658. DOI: `10.1093/cvr/cvm079`.

Hari-Gupta, Y., N. Fili, Á. dos Santos, A. W. Cook, R. E. Gough, H. C. Reed, L. Wang, J. Aaron, T. Venit, E. Wait, A. Grosse-Berkenbusch, J. C. M. Gebhardt, P. Percipalle, T. L. Chew, M. Martin-Fernandez, and C. P. Toseland (2022). "Myosin VI regulates the spatial organisation of mammalian transcription initiation." In: *Nature Communications* 13.1, pp. 1–18. DOI: `10.1038/s41467-022-28962-w`.

Harris, P. (1983). *The definition of heart failure*. Tech. rep. 8, p. 600. DOI: `10.1093/oxfordjournals.eurheartj.a061529`.

Hartman, M. A. and J. A. Spudich (2012). *The myosin superfamily at a glance*. DOI: `10.1242/jcs.094300`.

Hashem, S., W. G. Davies, and A. Fornili (2020). "Heart Failure Drug Modifies the Intrinsic Dynamics of the Pre-Power Stroke State of Cardiac Myosin." In: DOI: `10.1021/acs.jcim.0c00953`.

Hashem, S., M. Tiberti, and A. Fornili (2017). "Allosteric modulation of cardiac myosin dynamics by omecamtiv mecarbil." In: *PLoS computational biology* 13.11, e1005826. DOI: `10.1371/journal.pcbi.1005826`.

Hasson, T. and M. S. Mooseker (1994). "Porcine myosin-VI: Characterization of a new mammalian unconventional myosin." In: *Journal of Cell Biology* 127.2, pp. 425–440. DOI: `10.1083/jcb.127.2.425`.

Heidenreich, P. A., B. Bozkurt, D. Aguilar, L. A. Allen, J. J. Byun, M. M. Colvin, A. Deswal, M. H. Drazner, S. M. Dunlay, L. R. Evers, J. C. Fang, S. E. Fedson, G. C. Fonarow, S. S. Hayek, A. F. Hernandez, P. Khazanie, M. M. Kittleson, C. S. Lee, M. S. Link, C. A. Milano, L. C. Nnacheta, A. T. Sandhu, L. W. Stevenson, O. Vardeny, A. R. Vest, and C. W. Yancy (2022). *2022 AHA/ACC/HFSA Guideline for the Management of Heart Failure: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines*. DOI: `10.1161/CIR.0000000000001063`.

Heissler, S. M. and J. R. Sellers (2014). *Myosin light chains: Teaching old dogs new tricks*. DOI: `10.1080/19490992.2015.1054092`.

Heissler, S. M., J. Selvadurai, L. M. Bond, R. Fedorov, J. Kendrick-Jones, F. Buss, and D. J. Manstein (2012). "Kinetic properties and small-molecule inhibition of human myosin-6." In: *FEBS Letters* 586.19, pp. 3208–3214. DOI: `10.1016/j.febslet.2012.07.014`.

Helms, A. S., A. D. Thompson, and S. M. Day (2022). *Translation of New and Emerging Therapies for Genetic Cardiomyopathies*. DOI: `10.1016/j.jacbts.2021.07.012`.

Herron, T. J. and K. S. McDonald (2002). "Small amounts of α-myosin heavy chain isoform expression significantly increase power output of rat cardiac myocyte fragments." In: *Circulation Research* 90.11, pp. 1150–1152. DOI: `10.1161/01.RES.0000022879.57270.11`.

Hilser, V. J., J. O. Wrabl, and H. N. Motlagh (2012). “Structural and Energetic Basis of Allostery.” In: DOI: `10.1146/annurev-biophys-050511-102319`.

Hoklotubbe, S. S. (2018). “Chapter 19.” In: *Betrayal at the Buffalo Ranch* 1486, pp. 111–113. DOI: `10.2307/j.ctt1xhr7d5.22`.

Holford, N. H. and L. B. Sheiner (1982). “Kinetics of pharmacologic response.” In: *Pharmacology and Therapeutics* 16.2, pp. 143–166. DOI: `10.1016/0163-7258(82)90051-1`.

Holmes, K. C. (1997). “The swinging lever-arm hypothesis of muscle contraction.” In: *Current Biology* 7.2, pp. 112–118. DOI: `10.1016/S0960-9822(06)00051-0`.

Holmes, K. C. and M. A. Geeves (2000). *The structural basis of muscle contraction*. Tech. rep. 1396, pp. 419–431. DOI: `10.1098/rstb.2000.0583`.

Homma, K., M. Yoshimura, J. Saito, R. Ikebe, and M. Ikebe (2001). “The core of the motor domain determines the direction of myosin movement.” In: *Nature* 412.6849, pp. 831–834. DOI: `10.1038/35090597`.

Hou, T., J. Wang, Y. Li, and W. Wang (2011). “Assessing the Performance of the MM$_P$*BSAandMM*$_G$*BSAMethods*.1.*TheAccuracyof*”. In: *J. Chem. Inf. Model* 51.1, pp. 69–82.

Houdusse, A., A. G. Szent-Györgyi, and C. Cohen (2000). “Three conformational states of scallop myosin S1.” In: *Proceedings of the National Academy of Sciences of the United States of America* 97.21, pp. 11238–11243. DOI: `10.1073/pnas.200376897`.

Houdusse, A. (2020). “Biological nanomotors, driving forces of life.” In: *Comptes Rendus - Biologies* 343.4, pp. 53–78. DOI: `10.5802/CRBIOL.45`.

Houdusse, A. and H. L. Sweeney (2016). “How Myosin Generates Force on Actin Filaments.” In: *Trends in Biochemical Sciences* 41.12, pp. 989–997. DOI: `10.1016/j.tibs.2016.09.006`.

– (2001). “Myosin motors: Missing structures and hidden springs.” In: *Current Opinion in Structural Biology* 11.2, pp. 182–194. DOI: `10.1016/S0959-440X(00)00188-3`.

Houston, D. R. and M. D. Walkinshaw (2013). “Consensus docking: Improving the reliability of docking in a virtual screening context.” In: *Journal of Chemical Information and Modeling* 53.2, pp. 384–390. DOI: `10.1021/ci300399w`.

Howard, J. (1997). *Molecular motors: Structural adaptations to cellular functions*. Tech. rep. 6651, pp. 561–567. DOI: `10.1038/39247`.

Hu, S., Y. Guo, Y. Wang, Y. Li, T. Fu, Z. Zhou, Y. Wang, J. Liu, and L. Pan (2019). “Structure of Myosin VI/Tom1 complex reveals a cargo recognition mode of Myosin VI for tethering.” In: *Nature Communications* 10.1. DOI: `10.1038/s41467-019-11481-6`.

Huang, J., S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. De Groot, H. Grubmüller, and A. D. MacKerell (2016). “CHARMM36m: An improved force field for folded and intrinsically disordered proteins.” In: *Nature Methods* 14.1, pp. 71–73. DOI: `10.1038/nmeth.4067`.

Huang, R., Z. A. Ripstein, J. L. Rubinstein, and L. E. Kay (2019). “Cooperative subunit dynamics modulate p97 function.” In: *Proceedings of the National Academy of Sciences of the United States of America* 116.1, pp. 158–167. DOI: `10.1073/pnas.1815495116`.

Huang, S. Y., S. Z. Grinter, and X. Zou (2010). “Scoring functions and their evaluation methods for protein-ligand docking: Recent advances and future directions.” In: *Physical Chemistry Chemical Physics* 12.40, pp. 12899–12908. DOI: `10.1039/c0cp00151a`.

Huang, S. Y. and X. Zou (2006). “An iterative knowledge-based scoring function to predict protein-ligand interactions: I. Derivation of interaction potentials.” In: *Journal of Computational Chemistry* 27.15, pp. 1866–1875. DOI: `10.1002/jcc.20504`.

Huang, W., G. Wang, Q. Shen, X. Liu, S. Lu, L. Geng, Z. Huang, and J. Zhang (2015). “ASBench: Benchmarking sets for allosteric discovery.” In: *Bioinformatics* 31.15, pp. 2598–2600. DOI: `10.1093/bioinformatics/btv169`.

Huang, Z., L. Zhu, Y. Cao, G. Wu, X. Liu, Y. Chen, Q. Wang, T. Shi, Y. Zhao, Y. Wang, W. Li, Y. Li, H. Chen, G. ZhangChen, and J. Zhang (2011). “ASD: A comprehensive database of allosteric proteins and modulators.” In: *Nucleic Acids Research* 39.SUPPL. 1. DOI: `10.1093/nar/gkq1022`.

Humphrey, W., A. Dalke, and K. Schulten (1996). “VMD: Visual molecular dynamics.” In: *Journal of Molecular Graphics* 14.1, pp. 33–38. DOI: `10.1016/0263-7855(96)00018-5`.

Hunt, S. A., W. T. Abraham, M. H. Chin, A. M. Feldman, G. S. Francis, T. G. Ganiats, M. Jessup, M. A. Konstam, D. M. Mancini, K. Michl, J. A. Oates, P. S. Rahko, L. W. Stevenson, M. Jesup, D. E. Casey, M. A. Silver, C. W. Yancy, S. C. Smith, A. K. Jacobs, C. E. Buller, M. A. Creager, S. M. Ettinger, H. M. Krumholz, F. G. Kushner, B. W. Lytle, R. A. Nishimura, R. L. Page, and L. G. Tarkington (2009). "2009 focused update incorporated into the ACC/AHA 2005 guidelines for the diagnosis and management of heart failure in adults: A report of the Aamerican College of Cardiology Foundation/American Heart Association Task Force on practice guidelines: Develop." In: *Circulation* 119.14. DOI: `10.1161/CIRCULATIONAHA.109.192065`.

HUXLEY, A. F. (1957). "Muscle structure and theories of contraction." In: *Progress in biophysics and biophysical chemistry* 7, pp. 255–318. DOI: `10.1016/s0096-4174(18)30128-8`.

HUXLEY, H. E. (1961). "The contractile structure of cardiac and skeletal muscle." In: *Circulation* 24, pp. 328–335. DOI: `10.1161/01.cir.24.2.328`.

– (1953). "X-ray analysis and the problem of muscle." In: *Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society (Great Britain)* 141.902, pp. 59–62. DOI: `10.1098/rspb.1953.0017`.

Hwang, P. M. and B. D. Sykes (2015). "Targeting the sarcomere to correct muscle function." In: *Nature Reviews Drug Discovery* 14.5, pp. 313–328. DOI: `10.1038/nrd4554`.

Inoue, A., J. Saito, R. Ikebe, and M. Ikebe (2002). "Myosin IXb is a single-headed minus-end-directed processive motor." In: *Nature Cell Biology* 4.4, pp. 302–306. DOI: `10.1038/ncb774`.

Irwin, J. J., K. G. Tang, J. Young, C. Dandarchuluun, B. R. Wong, M. Khurelbaatar, Y. S. Moroz, J. Mayfield, and R. A. Sayle (2020). "ZINC20 - A Free Ultralarge-Scale Chemical Database for Ligand Discovery." In: *Journal of Chemical Information and Modeling* 60.12, pp. 6065–6073. DOI: `10.1021/acs.jcim.0c00675`.

Izairi, R. and H. Kamberaj (2017). "Comparison Study of Polar and Nonpolar Contributions to Solvation Free Energy." In: DOI: `10.1021/acs.jcim.7b00368`.

Jakalian, A., D. B. Jack, and C. I. Bayly (2002). "Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation." In: *Journal of Computational Chemistry* 23.16, pp. 1623–1641. DOI: `10.1002/jcc.10128`.

Jeacocke, S. A. and P. J. England (1980). *Phosphorylation of a myofibrillar protein of Mr 150 000 in perfused rat heart, and the tentative identification of this as C-protein*. Tech. rep. 1, pp. 129–132. DOI: `10.1016/0014-5793(80)80418-2`.

Jefferies, J. L. and J. A. Towbin (2010). "Dilated cardiomyopathy." In: *The Lancet* 375.9716, pp. 752–762. DOI: `10.1016/S0140-6736(09)62023-7`.

Jin, Y., L. Wei, W. Cai, Z. Lin, Z. Wu, Y. Peng, T. Kohmoto, R. L. Moss, and Y. Ge (2017). "Complete Characterization of Cardiac Myosin Heavy Chain (223 kDa) Enabled by Size-Exclusion Chromatography and Middle-Down Mass Spectrometry." In: *Analytical Chemistry* 89.9, pp. 4922–4930. DOI: `10.1021/acs.analchem.7b00113`.

Jo, S., T. Kim, V. G. Iyer, and W. Im (2008). "CHARMM-GUI: A web-based graphical user interface for CHARMM." In: *Journal of Computational Chemistry* 29.11, pp. 1859–1865. DOI: `10.1002/jcc.20945`.

Jones, G., P. Willett, R. C. Glen, A. R. Leach, and R. Taylor (1997). "Development and validation of a genetic algorithm for flexible docking." In: *Journal of Molecular Biology* 267.3, pp. 727–748. DOI: `10.1006/jmbi.1996.0897`.

Jumper, J., R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis (2021). "Highly accurate protein structure prediction with AlphaFold." In: *Nature* 596.7873, pp. 583–589. DOI: `10.1038/s41586-021-03819-2`.

Jung, E. J., G. Liu, W. Zhou, and X. Chen (2006). "Myosin VI Is a Mediator of the p53-Dependent Cell Survival Pathway." In: *Molecular and Cellular Biology* 26.6, pp. 2175–2186. DOI: `10.1128/mcb.26.6.2175-2186.2006`.

Kapoor, J. R., R. Kapoor, C. Ju, P. A. Heidenreich, Z. J. Eapen, A. F. Hernandez, J. Butler, C. W. Yancy, and G. C. Fonarow (2016). “Precipitating Clinical Factors, Heart Failure Characterization, and Outcomes in Patients Hospitalized With Heart Failure With Reduced, Borderline, and Preserved Ejection Fraction.” In: *JACC: Heart Failure* 4.6, pp. 464–472. doi: `10.1016/j.jchf.2016.02.017`.

Karatzaferi, C., M. K. Chinn, and R. Cooke (2004). “The force exerted by a muscle cross-bridge depends directly on the strength of the actomyosin bond.” In: *Biophysical journal* 87.4, pp. 2532–2544. doi: `10.1529/biophysj.104.039909`.

Karplus, M. and J. N. Kushick (1981). “Method for Estimating the Configurational Entropy of Macromolecules.” In: *Macromolecules* 14.2, pp. 325–332. doi: `10.1021/ma50003a019`.

Karplus, M. and G. A. Petsko (1990). “Molecular dynamics simulations in biology.” In: *Nature* 347.6294, pp. 631–639. doi: `10.1038/347631a0`.

Kass, D. A. and R. J. Solaro (2006). “Mechanisms and use of calcium-sensitizing agents in the failing heart.” In: *Circulation* 113.2, pp. 305–315. doi: `10.1161/CIRCULATIONAHA.105.542407`.

Kawas, R. F., R. L. Anderson, S. R. Bartholomew Ingle, Y. Song, A. S. Sran, and H. M. Rodriguez (2017). “A small-molecule modulator of cardiac myosin acts on multiple stages of the myosin chemomechanical cycle.” In: *Journal of Biological Chemistry* 292.40, pp. 16571–16577. doi: `10.1074/jbc.M117.776815`.

Kelleher, J. F., M. A. Mandell, G. Moulder, K. L. Hill, S. W. L’Hernault, R. Barstead, and M. A. Titus (2000). “Myosin VI is required for asymmetric segregation of cellular components during C. elegans spermatogenesis.” In: *Current Biology* 10.23, pp. 1489–1496. doi: `10.1016/S0960-9822(00)00828-9`.

Kellerman, K. A. and K. G. Miller (1992). *An unconventional myosin heavy chain gene from Drosophila melanogaster*. Tech. rep. 4, pp. 823–834. doi: `10.1083/jcb.119.4.823`.

Kim, H. J., J. Hsin, Y. Liu, P. R. Selvin, and K. Schulten (2010). “Formation of Salt Bridges Mediates Internal Dimerization of Myosin VI Medial Tail Domain.” In: *Structure* 18.11, pp. 1443–1449. doi: `10.1016/j.str.2010.09.011`.

Kitchen, D. B., H. Decornez, J. R. Furr, and J. Bajorath (2004). *Docking and scoring in virtual screening for drug discovery: Methods and applications*. doi: `10.1038/nrd1549`.

Kodera, N., D. Yamamoto, R. Ishikawa, and T. Ando (2010). “Video imaging of walking myosin v by high-speed atomic force microscopy.” In: *Nature* 468.7320, pp. 72–76. doi: `10.1038/nature09450`.

Köhler, D., C. Ruff, E. Meyhöfer, and M. Bähler (2003). “Different degrees of lever arm rotation control myosin step size.” In: *Journal of Cell Biology* 161.2, pp. 237–241. doi: `10.1083/jcb.200212039`.

Kokh, D. B., M. Amaral, J. Bomke, U. Grädler, D. Musil, H. P. Buchstaller, M. K. Dreyer, M. Frech, M. Lowinski, F. Vallee, M. Bianciotto, A. Rak, and R. C. Wade (2018). “Estimation of Drug-Target Residence Times by τ-Random Acceleration Molecular Dynamics Simulations.” In: *Journal of Chemical Theory and Computation* 14.7, pp. 3859–3869. doi: `10.1021/acs.jctc.8b00230`.

Kokh, D. B., B. Doser, S. Richter, F. Ormersbach, X. Cheng, and R. C. Wade (2020). “A workflow for exploring ligand dissociation from a macromolecule: Efficient random acceleration molecular dynamics simulation and interaction fingerprint analysis of ligand trajectories.” In: *Journal of Chemical Physics* 153.12. doi: `10.1063/5.0019088`. arXiv: `2006.11066`.

Komuro, I., K. Nomoto, T. Sugiyama, M. Kurabayashi, F. Takaku, and Y. Yazaki (1987). *Isolation and characterization of myosin heavy chain isozymes of the bovine conduction system*. Tech. rep. 6, pp. 859–865. doi: `10.1161/01.RES.61.6.859`.

Kooij, V., V. Venkatraman, J. A. Kirk, C. Ubaida-Mohien, D. R. Graham, M. J. Faber, and J. E. Van Eyk (2014). “Identification of cardiac myofilament protein isoforms using multiple mass spectrometry based approaches.” In: *Proteomics - Clinical Applications* 8.7-8, pp. 578–589. doi: `10.1002/prca.201400039`.

Koppole, S., J. C. Smith, and S. Fischer (2006). “Simulations of the Myosin II Motor Reveal a Nucleotide-state Sensing Element that Controls the Recovery Stroke.” In: *Journal of Molecular Biology*

361.3, pp. 604–616. DOI: `10.1016/j.jmb.2006.06.022`.

Korb, O., T. Stützle, and T. E. Exner (2007). "An ant colony optimization approach to flexible protein–ligand docking." In: *Swarm Intelligence* 1.2, pp. 115–134. DOI: `10.1007/s11721-007-0006-9`.

– (2009). "Empirical scoring functions for advanced Protein-Ligand docking with PLANTS." In: *Journal of Chemical Information and Modeling* 49.1, pp. 84–96. DOI: `10.1021/ci800298z`.

– (2006). "PLANTS: Application of ant colony optimization to structure-based drug design." In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 4150 LNCS. Springer Verlag, pp. 247–258. DOI: `10.1007/11839088_22`.

Koshland, D. E., J. G. Nemethy, and D. Filmer (1966). "Comparison of Experimental Binding Data and Theoretical Models in Proteins Containing Subunits." In: *Biochemistry* 5.1, pp. 365–385. DOI: `10.1021/bi00865a047`.

Krenz, M., A. Sanbe, F. Bouyer-Dalloz, J. Gulick, R. Klevitsky, T. E. Hewett, H. E. Osinska, J. N. Lorenz, C. Brosseau, A. Federico, N. R. Alpert, D. M. Warshaw, M. B. Perryman, S. M. Helmke, and J. Robbins (2003). "Analysis of myosin heavy chain functionality in the heart." In: *Journal of Biological Chemistry* 278.19, pp. 17466–17474. DOI: `10.1074/jbc.M210804200`.

Kühne, W. (2011). *Untersuchungen über das Protoplasma und die Contractilität / von W. Kühne.* Leipzig: W. Engelmann, pp. 1–204. DOI: `10.5962/bhl.title.11169`.

Kühner, S. and S. Fischer (2011). "Structural mechanism of the ATP-induced dissociation of rigor myosin from actin." In: *Proceedings of the National Academy of Sciences of the United States of America* 108.19, pp. 7793–7798. DOI: `10.1073/pnas.1018420108`.

Kuntz, I. D., J. M. Blaney, S. J. Oatley, R. Langridge, and T. E. Ferrin (1982). "A Geometric Approach to Macromolecule-Ligand Interactions." In:

Landrum, G. and Others (2016). "Rdkit: Open-source cheminformatics software." In:

Laskowski, R. A. and M. B. Swindells (2011). "LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery." In: *Journal of Chemical Information and Modeling* 51.10, pp. 2778–2786. DOI: `10.1021/ci200227u`.

Lawson, J. D., E. Pate, I. Rayment, and R. G. Yount (2004). "Molecular dynamics analysis of structural factors influencing back door Pi release in myosin." In: *Biophysical Journal* 86.6, pp. 3794–3803. DOI: `10.1529/biophysj.103.037390`.

Leach, A. R. (2001). *Molecular modelling _ principles and applications-Prentice Hall (2001).pdf*.

Lee Sweeney, H. and A. Houdusse (2004). "The motor mechanism of myosin V: Insights for muscle contraction." In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 359.1452, pp. 1829–1841. DOI: `10.1098/rstb.2004.1576`.

Li, Z., H. Li, K. Yu, and H.-B. Luo (2021). "Perspective of drug design with high-performance computing." In: *National Science Review* 8.12. DOI: `10.1093/nsr/nwab105`.

Limongelli, V. (2020). "Ligand binding free energy and kinetics calculation in 2020." In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 10.4, pp. 1–32. DOI: `10.1002/wcms.1455`.

Lindahl, E., B. Hess, and D. van der Spoel (2001). "GROMACS 3.0: A package for molecular simulation and trajectory analysis." In: *Journal of Molecular Modeling* 7.8, pp. 306–317. DOI: `10.1007/S008940100045`.

Lindorff-Larsen, K., S. Piana, R. O. Dror, and D. E. Shaw (2011). "How fast-folding proteins fold." In: *Science* 334.6055, pp. 517–520. DOI: `10.1126/science.1208351`.

Lionta, E., G. Spyrou, D. K. Vassilatis, and Z. Cournia (2014). *Send Orders for Reprints to reprints@benthamscience.net Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances*. Tech. rep.

Lipinski, C. A. (2000). "Drug-like properties and the causes of poor solubility and poor permeability." In: *Journal of Pharmacological and Toxicological Methods* 44.1, pp. 235–249. DOI: `10.1016/S1056-8719(00)00107-6`.

Lipinski, C. and A. Hopkins (2004). "for Biology and Medicine." In: 432.December.

Lister, I., S. Schmitz, M. Walker, J. Trinick, F. Buss, C. Veigel, and J. Kendrick-Jones (2004). "A monomeric myosin VI with a large working stroke." In:

EMBO Journal 23.8, pp. 1729–1738. DOI: `10.1038/sj.emboj.7600180`.

Liu, J. and R. Nussinov (2016). “Allostery: An Overview of Its History, Concepts, Methods, and Applications.” In: *PLOS Computational Biology* 12.6, e1004966. DOI: `10.1371/journal.pcbi.1004966`.

Liu, Y., H. D. White, B. Belknap, D. A. Winkelmann, and E. Forgacs (2015). “Omecamtiv Mecarbil Modulates the Kinetic and Motile Properties of Porcine β-Cardiac Myosin.” In: *Biochemistry* 54.10, pp. 1963–1975. DOI: `10.1021/bi5015166`.

Livesay, D. R., K. E. Kreth, and A. A. Fodor (2012). “A critical evaluation of correlated mutation algorithms and coevolution within allosteric mechanisms.” In: *Methods in Molecular Biology* 796, pp. 385–398. DOI: `10.1007/978-1-61779-334-9_21`.

Llinas, P., T. Isabet, L. Song, V. Ropars, B. Zong, H. Benisty, S. Sirigu, C. Morris, C. Kikuti, D. Safer, H. L. Sweeney, and A. Houdusse (2015). “How Actin Initiates the Motor Activity of Myosin.” In: *Developmental Cell* 33.4, pp. 401–412. DOI: `10.1016/j.devcel.2015.03.025`.

Lotz, S. D. and A. Dickson (2018). “Unbiased Molecular Dynamics of 11 min Timescale Drug Unbinding Reveals Transition State Stabilizing Interactions.” In: *Journal of the American Chemical Society* 140.2, pp. 618–628. DOI: `10.1021/jacs.7b08572`.

Lu, H. and P. J. Tonge (2010). “Drug-target residence time: Critical information for lead optimization.” In: *Current Opinion in Chemical Biology* 14.4, pp. 467–474. DOI: `10.1016/j.cbpa.2010.06.176`.

Lu, S., X. He, D. Ni, and J. Zhang (2019). “Allosteric Modulator Discovery: From Serendipity to Structure-Based Design.” In: DOI: `10.1021/acs.jmedchem.8b01749`.

Lüdemann, S. K., V. Lounnas, and R. C. Wade (2000). “How do substrates enter and products exit the buried active site of cytochrome P450cam? 1. Random expulsion molecular dynamics investigation of ligand access channels and mechanisms.” In: *Journal of Molecular Biology* 303.5, pp. 797–811. DOI: `10.1006/jmbi.2000.4154`.

Lukasiewicz, S., M. Czeczelewski, A. Forma, J. Baj, R. Sitarz, and A. Stanisławek (2021). “Breast cancer—epidemiology, risk factors, classification, prognostic markers, and current treatment strategies—An updated review.” In: *Cancers* 13.17, pp. 1–30. DOI: `10.3390/cancers13174287`.

Lyons, G. E., S. Schiaffino, D. Sassoon, P. Barton, and M. Buckingham (1990). *Developmental regulation of myosin gene expression in mouse cardiac muscle*. Tech. rep. 6 I, pp. 2427–2436. DOI: `10.1083/jcb.111.6.2427`.

Lyu, J., S. Wang, T. E. Balius, I. Singh, A. Levit, Y. S. Moroz, M. J. O’Meara, T. Che, E. Algaa, K. Tolmachova, A. A. Tolmachev, B. K. Shoichet, B. L. Roth, and J. J. Irwin (2019). “Ultra-large library docking for discovering new chemotypes.” In: *Nature* 566.7743, pp. 224–229. DOI: `10.1038/s41586-019-0917-9`.

Madden, T. (2002). “NCBI$_{b}lastinformation.pdf$.” In: *The NCBI Handbook*, pp. 1–15.

Mahdavi, V., A. P. Chambers, and B. Nadal-Ginard (1984). *Cardiac ca-and 83-myosin heavy chain genes are organized in tandem (gene organization/gene regulation/tissue-specific expression/DNA sequence/SI mapping analysis)*. Tech. rep., pp. 2626–2630.

Makowska, K. A., R. E. Hughes, K. J. White, C. M. Wells, and M. Peckham (2015). “Specific Myosins Control Actin Organization, Cell Morphology, and Migration in Prostate Cancer Cells.” In: *Cell Reports* 13.10, pp. 2118–2125. DOI: `10.1016/j.celrep.2015.11.012`.

Malik, F. I., J. J. Hartman, K. A. Elias, B. P. Morgan, H. Rodriguez, K. Brejc, R. L. Anderson, S. H. Sueoka, K. H. Lee, J. T. Finer, R. Sakowicz, R. Baliga, D. R. Cox, M. Garard, G. Godinez, R. Kawas, E. Kraynack, D. Lenzi, P. P. Lu, A. Muci, C. Niu, X. Qian, D. W. Pierce, M. Pokrovskii, I. Suehiro, S. Sylvester, T. Tochimoto, C. Valdez, W. Wang, T. Katori, D. A. Kass, Y. T. Shen, S. F. Vatner, and D. J. Morgans (2011). “Cardiac myosin activation: A potential therapeutic approach for systolic heart failure.” In: *Science* 331.6023, pp. 1439–1443. DOI: `10.1126/science.1200113`.

Manickam, M., H. B. Jalani, T. Pillaiyar, P. R. Boggu, N. Sharma, E. Venkateswararao, Y. J. Lee, E. S. Jeon, M. J. Son, S. H. Woo, and S. H. Jung (2018). “Design and synthesis of sulfonamidophenylethylureas as novel cardiac myosin activator.” In: *European Journal of Medicinal Chemistry* 143, pp. 1869–1887. DOI: `10.1016/j.ejmech.2017.10.077`.

Manickam, M., H. B. Jalani, T. Pillaiyar, N. Sharma, P. R. Boggu, E.

Venkateswararao, Y. J. Lee, E. S. Jeon, and S. H. Jung (2017). "Exploration of flexible phenylpropylurea scaffold as novel cardiac myosin activators for the treatment of systolic heart failure." In: *European Journal of Medicinal Chemistry* 134, pp. 379–391. DOI: `10.1016/j.ejmech.2017.04.005`.

Manning, G., D. B. Whyte, R. Martinez, T. Hunter, and S. Sudarsanam (2002). "The protein kinase complement of the human genome." In: *Science* 298.5600, pp. 1912–1934. DOI: `10.1126/science.1075762`.

Maragliano, L., A. Fischer, E. Vanden-Eijnden, and G. Ciccotti (2006). "String method in collective variables: Minimum free energy paths and isocommittor surfaces." In: *Journal of Chemical Physics* 125.2, p. 024106. DOI: `10.1063/1.2212942`.

Markidis, S. and E. Laure (2015). "Solving software challenges for exascale: International Conference on Exascale Applications and Software, EASC 2014 Stockholm, Sweden, April 2–3, 2014 revised selected papers." In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 8759. DOI: `10.1007/978-3-319-15976-8`.

Maron, B. J. and M. S. Maron (2013). "Hypertrophic cardiomyopathy." In: *The Lancet* 381.9862, pp. 242–255. DOI: `10.1016/S0140-6736(12)60397-3`.

Masri, A. and I. Olivotto (2022). "Cardiac Myosin Inhibitors as a Novel Treatment Option for Obstructive Hypertrophic Cardiomyopathy: Addressing the Core of the Matter." In: *Journal of the American Heart Association* 11.9, p. 24656. DOI: `10.1161/JAHA.121.024656`.

Mater, A. C. and M. L. Coote (2019). "Deep Learning in Chemistry." In: *Journal of Chemical Information and Modeling*. DOI: `10.1021/acs.jcim.9b00266`.

McCammon, J. A. (1998). "Theory of biomolecular recognition." In: *Current Opinion in Structural Biology* 8.2, pp. 245–249. DOI: `10.1016/S0959-440X(98)80046-8`.

McMurray, J. J. (2010). "Systolic heart failure." In: *New England Journal of Medicine* 362.3, pp. 228–238. DOI: `10.1056/NEJMcp0909392`.

McMurray, J. J., J. A. Ezekowitz, B. S. Lewis, B. J. Gersh, S. Van Diepen, J. Amerena, J. Bartunek, P. Commerford, B. H. Oh, V. P. Harjola, S. M. Al-Khatib, M. Hanna, J. H. Alexander, R. D. Lopes, D. M. Wojdyla, L. Wallentin, and C. B. Granger (2013). "Left ventricular systolic dysfunction, heart failure, and the risk of stroke and systemic embolism in patients with atrial fibrillation insights from the ARISTOTLE trial." In: *Circulation: Heart Failure* 6.3, pp. 451–460. DOI: `10.1161/CIRCHEARTFAILURE.112.000143`.

Melchionda, S., N. Ahituv, L. Bisceglia, T. Sobe, F. Glaser, R. Rabionet, M. L. Arbones, A. Notarangelo, E. Di Iorio, M. Carella, L. Zelante, X. Estivill, K. B. Avraham, and P. Gasparini (2001). "MYO6, the human homologue of the gene responsible for deafness in Snell's waltzer mice, is mutated in autosomal dominant nonsyndromic hearing loss." In: *American Journal of Human Genetics* 69.3, pp. 635–640. DOI: `10.1086/323156`.

Ménétrey, J., A. Bahloul, A. L. Wells, C. M. Yengo, C. A. Morris, H. L. Sweeney, and A. Houdusse (2005). "The structure of the myosin VI motor reveals the mechanism of directionality reversal." In: *Nature* 435.7043, pp. 779–785. DOI: `10.1038/nature03592`.

Ménétrey, J., T. Isabet, V. Ropars, M. Mukherjea, O. Pylypenko, X. Liu, J. Perez, P. Vachette, H. L. Sweeney, and A. M. Houdusse (2012). *Processive Steps in the Reverse Direction Require Uncoupling of the Lead Head Lever Arm of Myosin VI.* DOI: `10.1016/j.molcel.2012.07.034`.

Ménétrey, J., P. Llinas, J. Cicolari, G. Squires, X. Liu, A. Li, H. L. Sweeney, and A. Houdusse (2008). "The post-rigor structure of myosin VI and implications for the recovery stroke." In: *EMBO Journal* 27.1, pp. 244–252. DOI: `10.1038/sj.emboj.7601937`.

Ménétrey, J., P. Llinas, M. Mukherjea, H. L. Sweeney, and A. Houdusse (2007). "The Structural Basis for the Large Powerstroke of Myosin VI." In: *Cell* 131.2, pp. 300–308. DOI: `10.1016/j.cell.2007.08.027`.

Meng, E. C., B. K. Shoichet, and I. D. Kuntz (1992). "Automated docking with grid-based energy evaluation." In: *Journal of Computational Chemistry* 13.4, pp. 505–524. DOI: `10.1002/jcc.540130412`.

Mentzer, G. and E. M. Hsich (2019). "Heart Failure with Reduced Ejection Fraction in Women: Epidemiology, Outcomes, and Treatment." In: *Heart Failure Clinics* 15.1, pp. 19–27. DOI: `10.1016/j.hfc.2018.08.003`.

Mesentean, S., S. Koppole, J. C. Smith, and S. Fischer (2007). "The Principal Motions Involved in the Coupling Mechanism of the Recovery Stroke of the Myosin Motor." In: *Journal of Molecular Biology* 367.2, pp. 591–602. DOI: `10.1016/j.jmb.2006.12.058`.

Mitchell, J. B., R. A. Laskowski, A. Alex, M. J. Forster, and J. M. Thornton (1999). "BLEEP - Potential of mean force describing protein-ligand interactions: II. Calculation of binding energies and comparison with experimental data." In: *Journal of Computational Chemistry* 20.11, pp. 1177–1185. DOI: `10.1002/(SICI)1096-987X(199908)20:11<1177::AID-JCC8>3.0.CO;2-0`.

Miyata, S., W. Minobe, M. R. Bristow, and L. A. Leinwand (2000). *Myosin heavy chain isoform expression in the failing and nonfailing human heart*. Tech. rep. 4, pp. 386–390. DOI: `10.1161/01.RES.86.4.386`.

MONOD, J. and F. JACOB (1961). "Teleonomic mechanisms in cellular metabolism, growth, and differentiation." In: *Cold Spring Harbor symposia on quantitative biology* 26, pp. 389–401. DOI: `10.1101/sqb.1961.026.01.048`.

Monod, J., J. Wyman, and J. P. Changeux (1965). "On the nature of allosteric transitions: A plausible model." In: *Journal of Molecular Biology* 12.1, pp. 88–118. DOI: `10.1016/S0022-2836(65)80285-6`.

Montalvo-Acosta, J. J. and M. Cecchini (2016). "Computational Approaches to the Chemical Equilibrium Constant in Protein-ligand Binding." In: *Molecular Informatics* 35.11-12, pp. 555–567. DOI: `10.1002/minf.201600052`.

Moretto, L., M. Ušaj, O. Matusovsky, D. E. Rassier, and R. Friedman (2022). "M ultistep orthophosphate release tunes actomyosin energy transduction." In: *Nature Communications* 13.1, p. 4575. DOI: `10.1038/s41467-022-32110-9`.

Morgan M. Rogers-Carter1 2, J. A. V. K. B. G. A. F. P. M. T. M. M. R. and J. P. Christianson1 (2017). " HHS Public Access." In: *Physiology behavior* 176.12, pp. 139–148. DOI: `10.1126/science.aad3456.A`.

Morgan, B. P., A. Muci, P. P. Lu, X. Qian, T. Tochimoto, W. W. Smith, M. Garard, E. Kraynack, S. Collibee, I. Suehiro, A. Tomasi, S. C. Valdez, W. Wang, H. Jiang, J. Hartman, H. M. Rodriguez, R. Kawas, S. Sylvester, K. A. Elias, G. Godinez, K. Lee, R. Anderson, S. Sueoka, D. Xu, Z. Wang, N. Djordjevic, F. I. Malik, and D. J. Morgans (2010). "Discovery of omecamtiv mecarbil the first, selective, small molecule activator of cardiac myosin." In: *ACS Medicinal Chemistry Letters* 1.9, pp. 472–477. DOI: `10.1021/ml100138q`.

Morris, G. M., D. S. Goodsell, M. E. Pique, R. Huey, S. Forli, W. E. Hart, S. Halliday, R. Belew, and A. J. Olson (1991). *User Guide AutoDock Version 4.2 Updated for version 4.2.6 Automated Docking of Flexible Ligands to Flexible Receptors*. Tech. rep.

Mudd, J. O. and D. A. Kass (2008). *Tackling heart failure in the twenty-first century*. DOI: `10.1038/nature06798`.

Mugnai, M. L. and D. Thirumalai (2021a). "Step-Wise Hydration of Magnesium by Four Water Molecules Precedes Phosphate Release in a Myosin Motor." In: *Journal of Physical Chemistry B* 125.4, pp. 1107–1117. DOI: `10.1021/acs.jpcb.0c10004`.

– (2021b). "Step-Wise Hydration of Magnesium by Four Water Molecules Precedes Phosphate Release in a Myosin Motor." In: *Journal of Physical Chemistry B* 125.4, pp. 1107–1117. DOI: `10.1021/acs.jpcb.0c10004`.

Mukherjea, M., M. Y. Ali, C. Kikuti, D. Safer, Z. Yang, H. Sirkia, V. Ropars, A. Houdusse, D. M. Warshaw, and H. L. Sweeney (2014). "Myosin VI Must Dimerize and Deploy Its Unusual Lever Arm in Order to Perform Its Cellular Roles." In: *Cell Reports* 8.5, pp. 1522–1532. DOI: `10.1016/j.celrep.2014.07.041`.

Munro, G., P. K. Ahring, and N. R. Mirza (2009). "Developing analgesics by enhancing spinal inhibition after injury: GABAA receptor subtypes as novel targets." In: *Trends in Pharmacological Sciences* 30.9, pp. 453–459. DOI: `10.1016/j.tips.2009.06.004`.

Murphy, C. T. and J. A. Spudich (2000). *Variable surface loops and myosin activity: Accessories to a motor*. Tech. rep. 2, pp. 139–151. DOI: `10.1023/A:1005610007209`.

Murphy, S. P., N. E. Ibrahim, and J. L. Januzzi (2020). "Heart Failure with Reduced Ejection Fraction: A Review." In: *JAMA - Journal of the American Medical Association* 324.5, pp. 488–504. DOI: `10.1001/jama.2020.10262`.

Naydenov, N. G., S. Lechuga, E. H. Huang, and A. I. Ivanov (2021). “Myosin motors: Novel regulators and therapeutic targets in colorectal cancer.” In: *Cancers* 13.4, pp. 1–24. DOI: `10.3390/cancers13040741`.

Neves, B. J., R. C. Braga, C. C. Melo-Filho, J. T. Moreira-Filho, E. N. Muratov, and C. H. Andrade (2018). *QSAR-based virtual screening: Advances and applications in drug discovery*. DOI: `10.3389/fphar.2018.01275`.

Norinder, U. and C. A. Bergström (2006). “Prediction of ADMET properties.” In: *ChemMedChem* 1.9, pp. 920–937. DOI: `10.1002/cmdc.200600155`.

O’Connell, C. B., M. J. Tyska, and M. S. Mooseker (2007). “Myosin at work: Motor adaptations for a variety of cellular functions.” In: *Biochimica et Biophysica Acta - Molecular Cell Research* 1773.5, pp. 615–630. DOI: `10.1016/j.bbamcr.2006.06.012`.

Odronitz, F. and M. Kollmar (2007). “Drawing the tree of eukaryotic life based on the analysis of 2,269 manually annotated myosins from 328 species.” In: *Genome Biology* 8.9, p. 196. DOI: `10.1186/gb-2007-8-9-r196`.

Oguchi, Y., S. V. Mikhailenko, T. Ohki, A. O. Olivares, E. M. De La Cruz, and S. Ishiwata (2008). “Load-dependent ADP binding to myosins V and VI: Implications for subunit coordination and function.” In: *Proceedings of the National Academy of Sciences of the United States of America* 105.22, pp. 7714–7719. DOI: `10.1073/pnas.0800564105`.

Olivotto, I., A. Oreziak, R. Barriales-Villa, T. P. Abraham, A. Masri, P. Garcia-Pavia, S. Saberi, N. K. Lakdawala, M. T. Wheeler, A. Owens, M. Kubanek, W. Wojakowski, M. K. Jensen, J. Gimeno-Blanes, K. Afshar, J. Myers, S. M. Hegde, S. D. Solomon, A. J. Sehnert, D. Zhang, W. Li, M. Bhattacharya, J. M. Edelberg, C. B. Waldman, S. J. Lester, A. Wang, C. Y. Ho, D. Jacoby, and E.-H. study investigators (2020). “Mavacamten for treatment of symptomatic obstructive hypertrophic cardiomyopathy (EXPLORER-HCM): a randomised, double-blind, placebo-controlled, phase 3 trial.” In: *Lancet (London, England)* 396.10253, pp. 759–769. DOI: `10.1016/S0140-6736(20)31792-X`.

Onishi, H., S. V. Mikhailenko, and M. F. Morales (2006). *Toward understanding actin activation of myosin ATPase: The role of myosin surface loops*. Tech. rep. 16, pp. 6136–6141. DOI: `10.1073/pnas.0601595103`.

Onishi, H., T. Ohki, N. Mochizuki, and M. F. Morales (2002). *Early stages of energy transduction by myosin: Roles of Arg in switch I, of Glu in switch II, and of the salt-bridge between them*. Tech. rep. 24, pp. 15339–15344. DOI: `10.1073/pnas.242604099`.

Onufriev, A., D. A. Case, and D. Bashford (2002). “Effective Born radii in the generalized Born approximation: The importance of being perfect.” In: *Journal of Computational Chemistry* 23.14, pp. 1297–1304. DOI: `10.1002/jcc.10126`.

Orioli, S, S Beccara, and P Faccioli (2017). “Enhanced sampling in molecular dynamics.” In: *The Journal of Chemical Physics* 147, p. 64108. DOI: `10.1063/1.4997197`.

Ovchinnikov, V., M. Cecchini, E. Vanden-Eijnden, and M. Karplus (2011). “A conformational transition in the myosin VI converter contributes to the variable step size.” In: *Biophysical Journal* 101.10, pp. 2436–2444. DOI: `10.1016/j.bpj.2011.09.044`.

Packer, M. (1993). “The development of positive inotropic agents for chronic heart failure: How have we gone astray?” In: *Journal of the American College of Cardiology* 22.4 SUPPL. 1, A119–A126. DOI: `10.1016/0735-1097(93)90474-F`.

Pan, A. C., D. W. Borhani, R. O. Dror, and D. E. Shaw (2013). “Molecular determinants of drug-receptor binding kinetics.” In: *Drug Discovery Today* 18.13-14, pp. 667–673. DOI: `10.1016/j.drudis.2013.02.007`.

Park, H., A. Li, L. Q. Chen, A. Houdusse, P. R. Selvin, and H. L. Sweeney (2007). “The unique insert at the end of the myosin VI motor is the sole determinant of directionality.” In: *Proceedings of the National Academy of Sciences of the United States of America* 104.3, pp. 778–783. DOI: `10.1073/pnas.0610066104`.

Park, H., B. Ramamurthy, M. Travaglia, D. Safer, L. Q. Chen, C. Franzini-Armstrong, P. R. Selvin, and H. L. Sweeney (2006). “Full-length myosin VI dimerizes and moves processively along actin filaments upon monomer clustering.” In: *Molecular Cell* 21.3, pp. 331–336. DOI: `10.1016/j.molcel.2005.12.015`.

Parrinello, M. and A. Rahman (1981). "Polymorphic transitions in single crystals: A new molecular dynamics method." In: *Journal of Applied Physics* 52.12, pp. 7182–7190. DOI: `10.1063/1.328693`.

Parsons, S. J. and J. T. Parsons (2004). *Src family kinases, key regulators of signal transduction*. DOI: `10.1038/sj.onc.1208160`.

Patrick Walters, W., M. T. Stahl, and M. A. Murcko (1998). "Virtual screening - An overview." In: *Drug Discovery Today* 3.4, pp. 160–178. DOI: `10.1016/s1359-6446(97)01163-x`.

Patrick Walters, W. and R. Wang (2019). "New Trends in Virtual Screening." In: *Journal of Chemical Information and Modeling* 59.9, pp. 3603–3604. DOI: `10.1021/acs.jcim.9b00728`.

Pereira, G. P. and M. Cecchini (2021). "Multibasin Quasi-Harmonic Approach for the Calculation of the Configurational Entropy of Small Molecules in Solution." In: *Cite This: J. Chem. Theory Comput* 17, pp. 1133–1142. DOI: `10.1021/acs.jctc.0c00978`.

Perez-Sanchez, H. and W. Wenzel (2010). "Optimization Methods for Virtual Screening on Novel Computational Architectures." In: *Current Computer Aided-Drug Design* 7.1, pp. 44–52. DOI: `10.2174/157340911793743565`.

Phichith, D., M. Travaglia, Z. Yang, X. Liu, A. B. Zong, D. Safer, and H. L. Sweeney (2009). "Cargo binding induces dimerization of myosin VI." In: *Proceedings of the National Academy of Sciences of the United States of America* 106.41, pp. 17320–17324. DOI: `10.1073/pnas.0909748106`.

Planelles-Herrero, V. J., J. J. Hartman, J. Robert-Paganin, F. I. Malik, and A. Houdusse (2017). "Mechanistic and structural basis for activation of cardiac myosin force production by omecamtiv mecarbil." In: *Nature Communications* 8.1, pp. 1–10. DOI: `10.1038/s41467-017-00176-5`.

Polyansky, A. A., R. Zubac, and B. Zagrovic (2012). "Estimation of conformational entropy in protein-ligand interactions: A computational perspective." In: *Methods in Molecular Biology* 819.1, pp. 327–353. DOI: `10.1007/978-1-61779-465-0_21`.

Ponder, J. W. and D. A. Case (2003). *Force fields for protein simulations*. Tech. rep., pp. 27–85. DOI: `10.1016/S0065-3233(03)66002-X`.

Popova, M., O. Isayev, and A. Tropsha (2018). "Deep reinforcement learning for de novo drug design." In: *Science Advances* 4.7. DOI: `10.1126/sciadv.aap7885`. arXiv: `1711.10907`.

Preller, M., K. Chinthalapudi, R. Martin, H. J. Knölker, and D. J. Manstein (2011). "Inhibition of myosin ATPase activity by halogenated pseudilins: A structure-activity study." In: *Journal of Medicinal Chemistry* 54.11, pp. 3675–3685. DOI: `10.1021/jm200259f`.

Puttini, M., S. Redaelli, L. Moretti, S. Brussolo, R. H. Gunby, L. Mologni, E. Marchesi, L. Cleris, A. Donella-Deana, P. Drueckes, E. Sala, V. Lucchini, M. Kubbutat, F. Formelli, A. Zambon, L. Scapozza, and C. Gambacorti-Passerini (2008). "Characterization of compound 584, an Abl kinase inhibitor with lasting effects." In: *Haematologica* 93.5, pp. 653–661. DOI: `10.3324/haematol.12212`.

Pylypenko, O., L. Song, A. Shima, Z. Yang, A. M. Houdusse, and S. H. Lee (2015). "Myosin VI deafness mutation prevents the initiation of processive runs on actin." In: *Proceedings of the National Academy of Sciences of the United States of America* 112.11, E1201–E1209. DOI: `10.1073/pnas.1420989112`.

Pylypenko, O., L. Song, G. Squires, X. Liu, A. B. Zong, A. Houdusse, and H. L. Sweeney (2011). "Role of insert-1 of myosin VI in modulating nucleotide affinity." In: *Journal of Biological Chemistry* 286.13, pp. 11716–11723. DOI: `10.1074/jbc.M110.200626`.

Qiao, Y., T. Li, S. Zheng, and H. Wang (2018). *The Hippo pathway as a drug target in gastric cancer*. DOI: `10.1016/j.canlet.2018.01.062`.

Qin, F., J. Tian, D. Zhou, and L. Chen (2013). *Mst1 and Mst2 kinases: Regulations and diseases*. Tech. rep. 1. DOI: `10.1186/2045-3701-3-31`.

Qin, J., R. Rajaratnam, L. Feng, J. Salami, J. S. Barber-Rotenberg, J. Domsic, P. Reyes-Uribe, H. Liu, W. Dang, S. L. Berger, J. Villanueva, E. Meggers, and R. Marmorstein (2015). "Development of organometallic S6K1 inhibitors." In: *Journal of Medicinal Chemistry* 58.1, pp. 305–314. DOI: `10.1021/jm5011868`.

Quiroga, R. and M. A. Villarreal (2016). "Vinardo: A scoring function based on autodock vina improves scoring, docking, and virtual screening." In: *PLoS ONE* 11.5. DOI: `10.1371/journal.pone.0155183`.

Rabenstein, B. (2000). *Monte Carlo Methods for Simulation of Protein Folding and Titration*. Tech. rep.

Rahib, L., M. R. Wehner, L. M. Matrisian, and K. T. Nead (2021). "Estimated Projection of US Cancer Incidence and Death to 2040." In: *JAMA Network Open* 4.4, pp. 1–14. DOI: `10.1001/jamanetworkopen.2021.4708`.

Rarey, M., B. Kramer, T. Lengauer, and G. Klebe (1996). "A fast flexible docking method using an incremental construction algorithm." In: *Journal of Molecular Biology* 261.3, pp. 470–489. DOI: `10.1006/jmbi.1996.0477`.

Rastelli, G., A. Del Rio, G. Degliesposti, and M. Sgobba (2010). "Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA." In: *Journal of Computational Chemistry* 31.4, pp. 797–810. DOI: `10.1002/jcc.21372`.

Rdowicz, M. J. (2002). *Myosins and pathology: Genetics and biology*. Tech. rep. 4, pp. 789–804. DOI: `10.18388/abp.2002_3739`.

Reifenberger, J. G., E. Toprak, H. J. Kim, D. Safer, H. L. Sweeney, and P. R. Selvin (2009). "Myosin VI undergoes a 180° power stroke implying an uncoupling of the front lever arm." In: *Proceedings of the National Academy of Sciences of the United States of America* 106.43, pp. 18255–18260. DOI: `10.1073/pnas.0900005106`.

Ren, X., Y. S. Shi, Y. Zhang, B. Liu, L. H. Zhang, Y. B. Peng, and R. Zeng (2018). "Novel Consensus Docking Strategy to Improve Ligand Pose Prediction." In: *Journal of Chemical Information and Modeling* 58.8, pp. 1662–1668. DOI: `10.1021/acs.jcim.8b00329`.

Reymond, J. L., R. Van Deursen, L. C. Blum, and L. Ruddigkeit (2010). *Chemical space as a source for new drugs*. DOI: `10.1039/c0md00020e`.

Rief, M., R. S. Rock, A. D. Mehta, M. S. Mooseker, R. E. Cheney, and J. A. Spudich (2000). "Myosin-V stepping kinetics: A molecular model for processivity." In: *Proceedings of the National Academy of Sciences of the United States of America* 97.17, pp. 9482–9486. DOI: `10.1073/pnas.97.17.9482`.

Ringer, S. (1883). "A further Contribution regarding the influence of the different Constituents of the Blood on the Contraction of the Heart." In: *The Journal of Physiology* 4.1, pp. 29–42. DOI: `10.1113/jphysiol.1883.sp000120`.

Robert-Paganin, J., D. Auguin, and A. Houdusse (2018). "Hypertrophic cardiomyopathy disease results from disparate impairments of cardiac myosin function and auto-inhibition." In: *Nature Communications* 9.1. DOI: `10.1038/s41467-018-06191-4`.

Robert-Paganin, J., O. Pylypenko, C. Kikuti, H. L. Sweeney, and A. Houdusse (2020). "Force Generation by Myosin Motors: A Structural Perspective." In: *Chemical Reviews* 120.1, pp. 5–35. DOI: `10.1021/acs.chemrev.9b00264`.

Rock, R. S., S. E. Rice, A. L. Wells, T. J. Purcell, J. A. Spudich, and H. L. Sweeney (2001). *Myosin VI is a processive motor with a large step size*. Tech. rep. 24, pp. 13655–13659. DOI: `10.1073/pnas.191512398`.

Rock, R. S., M. Rief, A. D. Mehta, and J. A. Spudich (2000). "In vitro assays of processive myosin motors." In: *Methods* 22.4, pp. 373–381. DOI: `10.1006/meth.2000.1089`.

Rohde, J. A., O. Roopnarine, D. D. Thomas, and J. M. Muretta (2018). "Mavacamten stabilizes an autoinhibited state of two-headed cardiac myosin." In: *Proceedings of the National Academy of Sciences of the United States of America* 115.32, E7486–E7494. DOI: `10.1073/pnas.1720342115/-/DCSupplemental`.

Rosano, G. and P. Ponikowski (2019). "Myosin Activators." In: *International Cardiovascular Forum Journal* 18, pp. 7–10. DOI: `10.17987/icfj.v18i0.611`.

Roy, S., K. D. Hagen, P. U. Maheswari, M. Lutz, A. L. Spek, J. Reedijk, and G. P. Van Wezel (2008). "Phenanthroline derivatives with improved selectivity as DNA-targeting anticancer or antimicrobial drugs." In: *ChemMedChem* 3.9, pp. 1427–1434. DOI: `10.1002/cmdc.200800097`.

Ryan Koes, D., M. P. Baumgartner, and C. J. Camacho (2013). "Lessons Learned in Empirical Scoring with smina from the CSAR 2011 Benchmarking Exercise." In: *J. Chem. Inf. Model* 22, p. 2. DOI: `10.1021/ci300604z`.

Sanchez, G. (2013). "Las instituciones de ciencia y tecnología en los procesos de aprendizaje de la producción agroalimentaria en Argentina." In: *El sistema argentino de innovación: instituciones, empresas y redes. El desafío de la creación y apropiación de conocimiento.* 26.October 2006, pp. 659–664. DOI: `10.1002/prot.`

Sander, T., J. Freyss, M. Von Korff, and C. Rufener (2015). "DataWarrior: An open-source program for chemistry aware data visualization and analysis." In: *Journal of Chemical Information and Modeling* 55.2, pp. 460–473. DOI: `10.1021/ci500588j`.

Sartore, S., S. Pierobon-bormioli, and S. Schiaffino (1978). "Immunohistochemical evidence for myosin polymorphism in the chicken heart." In: *Nature* 274.5666, pp. 82–83. DOI: `10.1038/274082a0`.

Sasaki, N., R. Ohkura, and K. Sutoh (2003). "Dictyostelium myosin II mutations that uncouple the converter swing and ATP hydrolysis cycle." In: *Biochemistry* 42.1, pp. 90–95. DOI: `10.1021/bi026051l`.

Scarpi-luttenauer, M. (2022). "Complexes de coordination pour l'inhibition de proteine kinases de la voie Hippo." In: *Ecole Doctorale ED222*.

Scarpi-Luttenauer, M., K. Galentino, C. Orvain, M. Cecchini, C. Gaiddon, and P. Mobian (2022). "TiO4N2 complexes formed with 1,10-phenanthroline ligands containing a donor-acceptor hydrogen bond site: Synthesis, cytotoxicity and docking experiments." In: *Inorganica Chimica Acta* 540, p. 121036. DOI: `10.1016/j.ica.2022.121036`.

Scarpi-Luttenauer, M., L. Geminiani, P. Lebrun, N. Kyritsakas, A. Chaumont, M. Henry, and P. Mobian (2020). "Bent 1,10-phenanthroline ligands within octahedral complexes constructed around a TiO4N2core." In: *Inorganic Chemistry* 59.17, pp. 12005–12016. DOI: `10.1021/acs.inorgchem.0c00915`.

Schliwa, M. (1999). "Myosin steps backwards." In: *Nature* 401.6752, pp. 431–432. DOI: `10.1038/46692`.

Schneider, G. and U. Fechner (2005). "Computer-based de novo design of drug-like molecules." In: *Nature Reviews Drug Discovery* 4.8, pp. 649–663. DOI: `10.1038/nrd1799`.

Schuetz, D. A., W. E. A. de Witte, Y. C. Wong, B. Knasmueller, L. Richter, D. B. Kokh, S. K. Sadiq, R. Bosma, I. Nederpelt, L. H. Heitman, E. Segala, M. Amaral, D. Guo, D. Andres, V. Georgi, L. A. Stoddart, S. Hill, R. M. Cooke, C. De Graaf, R. Leurs, M. Frech, R. C. Wade, E. C. M. de Lange, A. P. IJzerman, A. Müller-Fahrnow, and G. F. Ecker (2017). "Kinetics for Drug Discovery: an industry-driven effort to target drug residence time." In: *Drug Discovery Today* 22.6, pp. 896–911. DOI: `10.1016/j.drudis.2017.02.002`.

Seeber, M., M. Cecchini, F. Rao, G. Settanni, and A. Caflisch (2007). "Wordom: A program for efficient analysis of molecular dynamics simulations." In: *Bioinformatics* 23.19, pp. 2625–2627. DOI: `10.1093/bioinformatics/btm378`.

Shang, G., C. A. Brautigam, R. Chen, D. Lu, J. Torres-Vázquez, and X. Zhang (2017). "Structure analyses reveal a regulated oligomerization mechanism of the PlexinD1/GIPC/myosin VI complex." In: *eLife* 6. DOI: `10.7554/eLife.27322`.

Shiroguchi, K., H. F. Chin, D. E. Hannemann, E. Muneyuki, E. M. de la Cruz, and K. Kinosita (2011). "Direct observation of the myosin va recovery stroke that contributes to unidirectional stepping along actin." In: *PLoS Biology* 9.4, p. 1001031. DOI: `10.1371/journal.pbio.1001031`.

Singh, J., R. C. Petter, T. A. Baillie, and A. Whitty (2011). *The resurgence of covalent drugs*. DOI: `10.1038/nrd3410`.

Sisquellas, M. and M. Cecchini (2021). "PrepFlow: A Toolkit for Chemical Library Preparation and Management for Virtual Screening." In: *Molecular Informatics* 40.12, pp. 1–7. DOI: `10.1002/minf.202100139`.

Smith, C. A. and I. Rayment (1996). "Active site comparisons highlight structural similarities between myosin and other P-loop proteins." In: *Biophysical Journal* 70.4, pp. 1590–1602. DOI: `10.1016/S0006-3495(96)79745-X`.

Söding, J., A. Biegert, and A. N. Lupas (2005). "The HHpred interactive server for protein homology detection and structure prediction." In: *Nucleic Acids Research* 33.SUPPL. 2. DOI: `10.1093/nar/gki408`.

Song, K., Q. Li, W. Gao, S. Lu, Q. Shen, X. Liu, Y. Wu, B. Wang, H. Lin, G. Chen, and J. Zhang (2019). "AlloDriver: A method for the identification and analysis of cancer driver targets." In: *Nucleic Acids Research* 47.W1, W315–W321. DOI: `10.1093/nar/gkz350`.

Sousa, S. F., P. A. Fernandes, and M. J. Ramos (2006). "Protein-ligand docking: Current status and future challenges." In: *Proteins: Structure, Function, and Bioinformatics* 65.1, pp. 15–26. DOI: `10.1002/prot.21082`.

Spudich, J. A. (2014). *Hypertrophic and dilated cardiomyopathy: Four decades of basic research on muscle lead to potential therapeutic approaches to these devastating genetic diseases*. DOI: 10.1016/j.bpj.2014.02.011.

Srinivasan, J., T. E. Cheatham, P. Cieplak, P. A. Kollman, and D. A. Case (1998). *Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate-DNA Helices*. Tech. rep.

Stock, G. and P. Hamm (2018). "A nonequilibrium approach to allosteric communication." In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 373.1749. DOI: 10.1098/rstb.2017.0187.

Straight, A. F., A. Cheung, J. Limouze, I. Chen, N. J. Westwood, J. R. Sellers, and T. J. Mitchison (2003). "Dissecting temporal and spatial control of cytokinesis with a myosin II inhibitor." In: *Science* 299.5613, pp. 1743–1747. DOI: 10.1126/science.1081412.

Sun, S., C. Karki, B. Z. Gao, and L. Li (2022). "Molecular mechanisms of cardiac actomyosin transforming from rigor state to post-rigor state." In: *Journal of Chemical Physics* 156.3. DOI: 10.1063/5.0078166.

Sun, Y., H. W. Schroeder, J. F. Beausang, K. Homma, M. Ikebe, and Y. E. Goldman (2007). "Myosin VI Walks "Wiggly" on Actin with Large and Variable Tilting." In: *Molecular Cell* 28.6, pp. 954–964. DOI: 10.1016/j.molcel.2007.10.029.

Swanson, J. M., R. H. Henchman, and J. A. McCammon (2004). "Revisiting Free Energy Calculations: A Theoretical Connection to MM/PBSA and Direct Calculation of the Association Free Energy." In: *Biophysical Journal* 86.1 I, pp. 67–74. DOI: 10.1016/S0006-3495(04)74084-9.

Sweeney, H. L. and A. Houdusse (2010). "Structural and functional insights into the myosin motor mechanism." In: *Annual Review of Biophysics* 39.1, pp. 539–557. DOI: 10.1146/annurev.biophys.050708.133751.

Sweeney, H. L., H. Park, A. B. Zong, Z. Yang, P. R. Selvin, and S. S. Rosenfeld (2007). "How myosin VI coordinates its heads during processive movement." In: *EMBO Journal* 26.11, pp. 2682–2692. DOI: 10.1038/sj.emboj.7601720.

Syamaladevi, D. P., J. A. Spudich, and R. Sowdhamini (2012). "Structural and functional insights on the myosin superfamily." In: *Bioinformatics and Biology Insights* 6, pp. 11–21. DOI: 10.4137/bbi.s8451.

Takagi, Y., R. E. Farrow, N. Billington, A. Nagy, C. Batters, Y. Yang, J. R. Sellers, and J. E. Molloy (2014). "Myosin-10 produces its power-stroke in two phases and moves processively along a single actin filament under low load." In: *Proceedings of the National Academy of Sciences of the United States of America* 111.18. DOI: 10.1073/pnas.1320122111.

Taly, A., P. J. Corringer, D. Guedin, P. Lestage, and J. P. Changeux (2009). "Nicotinic receptors: Allosteric transitions and therapeutic targets in the nervous system." In: *Nature Reviews Drug Discovery* 8.9, pp. 733–750. DOI: 10.1038/nrd2927.

Tamaoki, T., H. Nomoto, I. Takahashi, Y. Kato, M. Morimoto, and F. Tomita (1986). "Staurosporine, a potent inhibitor of phospholipid Ca++dependent protein kinase." In: *Biochemical and Biophysical Research Communications* 135.2, pp. 397–402. DOI: 10.1016/0006-291X(86)90008-2.

Tang, S., R. Chen, M. Lin, Q. Lin, Y. Zhu, J. Ding, H. Hu, M. Ling, and J. Wu (2022). "molecules Accelerating AutoDock Vina with GPUs." In: DOI: 10.3390/molecules27093041.

Taylor, P. (1980). "Molecular Physics: An International Journal at the Interface Between Chemistry and Physics." In: January 2014, pp. 37–41.

Tee, A. K., M. S. Koh, P. G. Gibson, T. J. Lasserson, A. J. Wilson, and L. B. Irving (2007). *Long-acting beta2-agonists versus theophylline for maintenance treatment of asthma*. DOI: 10.1002/14651858.CD001281.pub2.

Teerlink, J. R., R. Diaz, G. M. Felker, J. J. McMurray, M. Metra, S. D. Solomon, K. F. Adams, I. Anand, A. Arias-Mendoza, T. Biering-Sørensen, M. Böhm, D. Bonderman, J. G. Cleland, R. Corbalan, M. G. Crespo-Leiro, U. Dahlström, L. E. Echeverria, J. C. Fang, G. Filippatos, C. Fonseca, E. Goncalvesova, A. R. Goudev, J. G. Howlett, D. E. Lanfear, J. Li, M. Lund, P. Macdonald, V. Mareev, S.-i. Momomura, E. O'Meara, A. Parkhomenko, P. Ponikowski, F. J. Ramires, P. Serpytis, K. Sliwa, J. Spinar, T. M. Suter, J. Tomcsanyi, H. Vandekerckhove, D. Vinereanu, A. A. Voors, M. B. Yilmaz, F. Zannad, L. Sharpsten, J. C. Legg, C. Varin, N. Honarpour, S. A. Abbasi, F. I.

Malik, and C. E. Kurtz (2021). "Cardiac Myosin Activation with Omecamtiv Mecarbil in Systolic Heart Failure." In: *New England Journal of Medicine* 384.2, pp. 105–116. DOI: `10.1056/nejmoa2025797`.

Thompson, R. F. and G. M. Langford (2002). "Myosin superfamily evolutionary history." In: *Anatomical Record* 268.3, pp. 276–289. DOI: `10.1002/ar.10160`.

Thornburg, Z. R., D. M. Bianchi, T. A. Brier, B. R. Gilbert, T. M. Earnest, M. C. Melo, N. Safronova, J. P. Sáenz, A. T. Cook, K. S. Wise, C. A. Hutchison, H. O. Smith, J. I. Glass, and Z. Luthey-Schulten (2022). "Fundamental behaviors emerge from simulations of a living minimal cell." In: *Cell* 185.2, 345–360.e28. DOI: `10.1016/j.cell.2021.12.025`.

Tian, C., K. Kasavajhala, K. A. Belfon, L. Raguette, H. Huang, A. N. Migues, J. Bickel, Y. Wang, J. Pincay, Q. Wu, and C. Simmerling (2020). "Ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution." In: *Journal of Chemical Theory and Computation* 16.1, pp. 528–552. DOI: `10.1021/acs.jctc.9b00591`.

Tiwary, P., V. Limongelli, M. Salvalaglio, and M. Parrinello (2015). "Kinetics of protein-ligand unbinding: Predicting pathways, rates, and rate-limiting steps." In: *Proceedings of the National Academy of Sciences of the United States of America* 112.5, E386–E391. DOI: `10.1073/pnas.1424461112`.

Tripathi, A and B. VA (2018). "Molecular Docking: From Lock and Key to Combination Lock." In: *Journal of Molecular Medicine and Clinical Applications* 2.1. DOI: `10.16966/2575-0305.106`.

Trivedia, D. V., J. M. Murettab, A. M. Swensona, J. P. Davisc, D. D. Thomasb, and C. M. Yengoa (2015). "Direct measurements of the coordination of lever arm swing and the catalytic cycle in myosin V." In: *Proceedings of the National Academy of Sciences of the United States of America* 112.47, pp. 14593–14598. DOI: `10.1073/pnas.1517566112`.

Trott, O. and A. J. Olson (2009). "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading." In: DOI: `10.1002/jcc.21334`.

Tsai, C. J. and R. Nussinov (2014). "A Unified View of "How Allostery Works"." In: *PLoS Computational Biology* 10.2. DOI: `10.1371/journal.pcbi.1003394`.

Vamathevan, J., D. Clark, P. Czodrowski, I. Dunham, E. Ferran, G. Lee, B. Li, A. Madabhushi, P. Shah, M. Spitzer, and S. Zhao (2019). *Applications of machine learning in drug discovery and development*. DOI: `10.1038/s41573-019-0024-5`.

Vanommeslaeghe, K., E. P. Raman, and A. D. MacKerell (2012). "Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges." In: *Journal of Chemical Information and Modeling* 52.12, pp. 3155–3168. DOI: `10.1021/ci3003649`.

Varadi, M., S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Zídek, T. Green, K. Tunyasuvunakool, S. Petersen, J. Jumper, E. Clancy, R. Green, A. Vora, M. Lutfi, M. Figurnov, A. Cowie, N. Hobbs, P. Kohli, G. Kleywegt, E. Birney, D. Hassabis, and S. Velankar (2022). "AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models." In: *Nucleic Acids Research* 50.D1, pp. D439–D444. DOI: `10.1093/nar/gkab1061`.

Veber, D. F., S. R. Johnson, H. Y. Cheng, B. R. Smith, K. W. Ward, and K. D. Kopple (2002). "Molecular properties that influence the oral bioavailability of drug candidates." In: *Journal of Medicinal Chemistry* 45.12, pp. 2615–2623. DOI: `10.1021/jm020017n`.

Veigel, C., S. Schmitz, F. Wang, and J. R. Sellers (2005). "Load-dependent kinetics of myosin-V can explain its high processivity." In: *Nature Cell Biology* 7.9, pp. 861–869. DOI: `10.1038/ncb1287`.

Verdonk, M. L., J. C. Cole, M. J. Hartshorn, C. W. Murray, and R. D. Taylor (2003). "Improved protein-ligand docking using GOLD." In: *Proteins: Structure, Function, and Bioinformatics* 52.4, pp. 609–623. DOI: `10.1002/prot.10465`.

Verkhivker, G. M., D. Bouzida, D. K. Gehlhaar, P. A. Rejto, S. Arthurs, A. B. Colson, S. T. Freer, V. Larson, B. A. Lutty, T. Marrone, and P. W. Rose (2000). *Deciphering common failures in molecular docking of ligand-protein com-*

plexes. Tech. rep. 8, pp. 731–751. DOI: `10.1023/A:1008158231558`.

Verlet, L. (1968). “Computer "experiments" on classical fluids. II. Equilibrium correlation functions.” In: *Physical Review* 165.1, pp. 201–214. DOI: `10.1103/PhysRev.165.201`.

Voors, A. A., J. F. Tamby, J. G. Cleland, M. Koren, L. B. Forgosh, D. Gupta, L. H. Lund, A. Camacho, R. Karra, H. P. Swart, P. Pellicori, F. Wagner, R. E. Hershberger, N. Prasad, R. Anderson, A. Anto, K. Bell, J. M. Edelberg, L. Fang, M. Henze, C. Kelly, G. Kurio, W. Li, K. Wells, C. Yang, S. L. Teichman, C. L. del Rio, and S. D. Solomon (2020). “Effects of danicamtiv, a novel cardiac myosin activator, in heart failure with reduced ejection fraction: experimental data and clinical results from a phase 2a trial.” In: *European Journal of Heart Failure* 22.9, pp. 1649–1658. DOI: `10.1002/ejhf.1933`.

Vreugde, S., C. Ferrai, A. Miluzio, E. Hauben, P. C. Marchisio, M. P. Crippa, M. Bussi, and S. Biffo (2006). “Nuclear Myosin VI Enhances RNA Polymerase II-Dependent Transcription.” In: *Molecular Cell* 23.5, pp. 749–755. DOI: `10.1016/j.molcel.2006.07.005`.

Walklate, J., C. Ferrantini, C. A. Johnson, C. Tesi, C. Poggesi, and M. A. Geeves (2021a). *Alpha and beta myosin isoforms and human atrial and ventricular contraction*. DOI: `10.1007/s00018-021-03971-y`.

– (2021b). *Alpha and beta myosin isoforms and human atrial and ventricular contraction*. DOI: `10.1007/s00018-021-03971-y`.

Walters, W. P. and R. Barzilay (2021). “Critical assessment of AI in drug discovery.” In: *Expert Opinion on Drug Discovery* 16.9, pp. 937–947. DOI: `10.1080/17460441.2021.1915982`.

Wang, E., H. Sun, J. Wang, Z. Wang, H. Liu, J. Z. H. Zhang, and T. Hou (2019). “End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design.” In: DOI: `10.1021/acs.chemrev.9b00055`.

Wang, J., Y. Deng, and B. Roux (2006). *No Title*. DOI: `10.1529/biophysj.106.084301`.

Wang, J., R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case (2004). “Development and testing of a general Amber force field.” In: *Journal of Computational Chemistry* 25.9, pp. 1157–1174. DOI: `10.1002/jcc.20035`.

Wang, R., L. Lai, and S. Wang (2002). *Further development and validation of empirical scoring functions for structure-based binding affinity prediction*. Tech. rep. 1, pp. 11–26. DOI: `10.1023/A:1016357811882`.

Warshaw, D. M. (2004). *Lever arms and necks: A common mechanistic theme across the myosin superfamily*. Tech. rep. 6, pp. 467–474. DOI: `10.1007/s10974-004-1767-z`.

Watson, C., S. Jenkinson, W. Kazmierski, and T. Kenakin (2005). “The CCR5 receptor-based mechanism of action of 873140, a potent allosteric noncompetitive HIV entry inhibitor.” In: *Molecular Pharmacology* 67.4, pp. 1268–1282. DOI: `10.1124/mol.104.008565`.

Webb, B. and A. Sali (2016). “Comparative protein structure modeling using MODELLER.” In: *Current Protocols in Bioinformatics* 2016, pp. 5.6.1–5.6.37. DOI: `10.1002/cpbi.3`.

WEBER, A. and S. WINICUR (1961). “The role of calcium in the superprecipitation of actomyosin.” In: *The Journal of biological chemistry* 236.12, pp. 3198–3202. DOI: `10.1016/s0021-9258(18)93995-5`.

Weiss, A., S. Schiaffino, and L. A. Leinwand (1999). “Comparative sequence analysis of the complete human sarcomeric myosin heavy chain family: Implications for functional diversity.” In: *Journal of Molecular Biology* 290.1, pp. 61–75. DOI: `10.1006/jmbi.1999.2865`.

Wells, A. L., A. W. Lin, L. Q. Chen, D. Safer, S. M. Cain, T. Hasson, B. O. Carragher, R. A. Milligan, and H. L. Sweeney (1999). “Myosin VI is an actin-based motor that moves backwards.” In: *Nature* 401.6752, pp. 505–508. DOI: `10.1038/46835`.

Wenthur, C. J., P. R. Gentry, T. P. Mathews, and C. W. Lindsley (2014). “Drugs for allosteric sites on receptors.” In: *Annual Review of Pharmacology and Toxicology* 54, pp. 165–184. DOI: `10.1146/annurev-pharmtox-010611-134525`.

Westbrook, J., Z. Feng, L. Chen, H. Yang, and H. M. Berman (2003). *The Protein Data Bank and structural genomics*. DOI: `10.1093/nar/gkg068`.

Winkelmann, D. A., E. Forgacs, M. T. Miller, and A. M. Stock (2015). “Structural basis for drug-induced allosteric changes to human β-cardiac myosin motor activity." In: *Nature Communications* 6.1, pp. 1–10. DOI: `10.1038/ncomms8974`.

Wodak, S. J., E. Paci, N. V. Dokholyan, I. N. Berezovsky, A. Horovitz, J. Li, V. J. Hilser, I. Bahar, J. Karanicolas, G. Stock, P. Hamm, R. H. Stote, and J. Eberhardt (2019). "Review Allostery in Its Many Disguises : From Theory to Applications." In: *Structure/Folding and Design*, pp. 1–13. DOI: `10.1016/j.str.2019.01.003`.

Wojciechowski, M. and B. Lesyng (2004). "Generalized Born Model: Analysis, Refinement, and Applications to Proteins." In: DOI: `10.1021/jp046748b`.

Woolner, S. and W. M. Bement (2009a). "Unconventional myosins acting unconventionally." In: *Trends in Cell Biology* 19.6, pp. 245–252. DOI: `10.1016/j.tcb.2009.03.003`.

– (2009b). "Unconventional myosins acting unconventionally." In: *Trends in Cell Biology* 19.6, pp. 245–252. DOI: `10.1016/j.tcb.2009.03.003`.

Wulf, S. F., V. Ropars, S. Fujita-Becker, M. Oster, G. Hofhaus, L. G. Trabuco, O. Pylypenko, H. L. Sweeney, A. M. Houdusse, R. R. Schröder, and R. D. Vale (2016). "Force-producing ADP state of myosin bound to actin." In: *Proceedings of the National Academy of Sciences of the United States of America* 113.13, E1844–E1852. DOI: `10.1073/pnas.1516598113`.

Yang, D., Q. Zhou, V. Labroska, S. Qin, S. Darbalaei, Y. Wu, E. Yuliantie, L. Xie, H. Tao, J. Cheng, Q. Liu, S. Zhao, W. Shui, Y. Jiang, and M. W. Wang (2021). "G protein-coupled receptors: structure- and function-based drug discovery." In: *Signal Transduction and Targeted Therapy* 6.1. DOI: `10.1038/s41392-020-00435-w`.

Yang, K., K. Swanson, W. Jin, C. Coley, P. Eiden, H. Gao, A. Guzman-Perez, T. Hopper, B. Kelley, M. Mathea, A. Palmer, V. Settels, T. Jaakkola, K. Jensen, and R. Barzilay (2019a). "Analyzing Learned Molecular Representations for Property Prediction." In: *Journal of Chemical Information and Modeling* 59.8, pp. 3370–3388. DOI: `10.1021/acs.jcim.9b00237`. arXiv: `1904.01561`.

– (2019b). *Correction to: Analyzing Learned Molecular Representations for Property Prediction (J. Chem. Inf. Model. (2019) 59:8 (3370-3388) DOI: 10.1021/acs.jcim.9b00237)*. DOI: `10.1021/acs.jcim.9b01076`.

Yang, Y., S. Gourinath, M. Kovács, L. Nyitray, R. Reutzel, D. M. Himmel, E. O'Neall-Hennessey, L. Reshetnikova, A. G. Szent-Györgyi, J. H. Brown, and C. Cohen (2007). "Rigor-like Structures from Muscle Myosins Reveal Key Mechanical Elements in the Transduction Pathways of This Allosteric Motor." In: *Structure* 15.5, pp. 553–564. DOI: `10.1016/j.str.2007.03.010`.

Yildiz, A., A. Snoberger, B. Barua, J. L. Atherton, H. Shuman, E. Forgacs, Y. E. Goldman, D. A. Winkelmann, and E. M. Ostap (2021). "Myosin with hypertrophic cardiac mutation R712L has a decreased working stroke which is rescued by omecamtiv mecarbil." In: DOI: `10.7554/eLife.63691`.

Yin, S., L. Biedermannova, J. Vondrasek, and N. V. Dokholyan (2008). "MedusaScore: An accurate force field-based scoring function for virtual drug screening." In: *Journal of Chemical Information and Modeling* 48.8, pp. 1656–1662. DOI: `10.1021/ci8001167`.

Yu, H., L. Ma, Y. Yang, and Q. Cui (2007). "Mechanochemical coupling in the myosin motor domain. I. Insights from equilibrium active-site simulations." In: *PLoS Computational Biology* 3.2. Ed. by F. Tama, pp. 0199–0213. DOI: `10.1371/journal.pcbi.0030021`.

Zheng, M., J. Zhao, C. Cui, Z. Fu, X. Li, X. Liu, X. Ding, X. Tan, F. Li, X. Luo, K. Chen, and H. Jiang (2018). "Computational chemical biology and drug design: Facilitating protein structure, function, and modulation studies." In: *Medicinal Research Reviews* 38.3, pp. 914–950. DOI: `10.1002/med.21483`.

Université de Strasbourg

EDSC
École Doctorale des Sciences Chimiques

Katia GALENTINO

Rational design of allosteric modulators in biomolecular motors

Insights into myosins’stability: from molecular modelling to vHTS

Résumé

Les myosines sont des biomoteurs qui contrôlent la vie et le mouvement des cellules, via la conversion de l'énergie chimique en travail mécanique pour se déplacer sur le filament d'actine, par un mécanisme cyclique. Il est crucial de comprendre la transduction chimiomécanique de ces moteurs, mais plusieurs aspects sont inconnus. Le dysfonctionnement des myosines implique de graves maladies, des myopathies aux cancers. Dans cette thèse, nous avons étudié la dynamique de la myosine cardiaque et de la myosineVI dans un état configurationnel spécifique. Nous avons capturé une transition spontanée vers un état intermédiaire différent du cycle dans la myosineVI et la comparaison avec la dynamique de myosine cardiaque a dénoté une plasticité marquée de myosineVI. Nous avons aussi étudié les deux protéines sur le plan pharmacologique, avec de nombreuses méthodes de calcul pour concevoir des activateurs allostériques de la myosine cardiaque et pour réaliser un criblage virtuel sur la myosineVI.

Myosines ; Dynamique moléculaire ; Allostérie ; Criblage virtuel ; Cinétique de liaison ; Énergie libre de liaison

Résumé en anglais

Myosins are biomolecular motors that coordinate life and motion in cells, by converting chemical energy into mechanical work to walk on actin filament, via a cyclic mechanism . Undertsanding the chemomechanical transduction in these motors has been crucial for decades, but many features remain unknown. Dysfunction of myosins relates to severe human disorders, from myopathies to cancers, becoming relevant targets. In this thesis we studied the dynamics of cardiac myosin and myosinVI in a peculiar configuration state of the actomyosin cycle. We captured a spontaneous transition towards a different intermediate state of the force generation cycle in myosinVI and comparison with the dynamics of cardiac myosin denoted a remarkable plasticity of the structure. We also investigated the two proteins from a pharmacological point of view, with a wide range of computational approaches to design cardiac myosin allosteric activators and to perform a virtual high throughput screening on myosinVI.

Myosins; Molecular Dynamics; Allostery; Virtual Screening; Kinetics of binding; Binding free energy