UNIVERSITAT DE BARCELONA

Computational and biophysical approaches for the discovery of Tribbles pseudokinases modulators

Juan Salamanca Viloria

UNIVERSITAT DE BARCELONA

UNIVERSITAT DE BARCELONA

FACULTAT DE FARMÀCIA I CIÈNCIES DE L'ALIMENTACIÓ
Director de Tesi: Dr. Xavier Barril Alonso

Computational and biophysical approaches for the discovery of Tribbles pseudokinases modulators

Juan Salamanca Viloria
2022

UNIVERSITAT DE BARCELONA

FACULTAT DE FARMÀCIA I CIÈNCES DE L'ALIMENTACIÓ

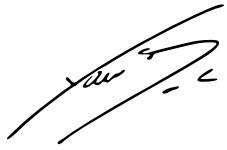
PROGRAMA DE DOCTORAT EN BIOMEDICINA

COMPUTATIONAL AND BIOPHYSICAL APPROACHES FOR THE DISCOVERY OF TRIBBLES PSEUDOKINASES MODULATORS

Memòria presentada per Juan Salamanca Viloria per optar al títol de doctor per la Universitat de Barcelona

Xavier Barril Alonso
Director de tesi i tutor

Juan Salamanca Viloria
Doctorand



Juan Salamanca Viloria
2022

A mis padres y hermanos,

No bird soars too high,
if he soars with his own wings.
-William Blake

Acknowledgements

Quiero dar las gracias a todas las personas que me han ayudado, ya que sin ellos no habría sido posible. Supongo que el doctorado se podría asemejar a un viaje, quizás en mi caso es incluso más literal y después de haber pasado por tantos sitios no me gustaría dejarme a nadie sin dar las gracias.

Primero me gustaría agradecer a Óscar por haberme dado la oportunidad de iniciar este proyecto, siento que no haya podido estar hasta el final, pero sin su ayuda inicial nada de esto habría sido posible. Por supuesto también a Esther y Emiliana que decidieron apostar por mí y también a todo el mundo en Intelligent Pharma que me ayudaron a adaptarme rápido y a sentirme cómodo al llegar a Barcelona.

Part of this thesis (most of it) would not be even possible without the collaboration of Peter Mace. I know my stay in New Zealand was short, but it was one of the most fruitful stays in my career. I want to thank you for allowing me to go to your lab and for all the help and support. Of course, all the people in your lab, especially Sam for all the patience and wisdom around Tribbles.

Por supuesto esta tesis no sería posible sin Xavi, gracias de corazón por haberme aceptado en tu grupo y por haberme ayudado durante todo este tiempo. Creo que he tenido muchísima suerte de poder tenerte como director, tanto en lo profesional como en lo personal. Gracias de verdad por haberme ayudado y por todo lo que me has enseñado durante el camino. También quiero agradecer a todo el grupo pasado y presente por su infinita paciencia frente a mis dudas en el laboratorio, muchas gracias de verdad a Andrea, Miriam y Salvo. También a Carles, Dylan, Maciej, Marina, Moira, Serena y por supuesto a Álvaro y Guillermo por todas las horas de conversaciones sobre ciencia, por haberme ayudado a adaptarme rápido y por crear un ambiente tan bueno.

Quiero también agradecer al consorcio TRAIN por todo el apoyo que me dieron en momentos complicados y en especial me gustaría agradecer a Miguel por todas las risas, cervezas y conversaciones que hemos tenido durante estos años, y por supuesto por su colaboración e interés en trabajar juntos, gracias de verdad. Por supuesto también agradecer a Laura por su infinita paciencia cuando estuve en Sheffield y por todas las risas.

I would like to thanks all the funding agencies that have supported me

through my PhD (Marie Skłodowska-Curie and Fulbright) and allowed me to train in the best possible conditions, thanks to all the PIs and partners.

Por úlitmo y no por ello menos importante quiero agradecer a todos mis amigos y familia que me han aguantado y soportado durante todo este tiempo. A mis padres y hermanos ya que sin ellos no sería la persona que soy hoy en día, gracias de verdad por todo lo que me habéis dado.

Y por supuesto a María, por haber aguantado a alguien como yo todo este tiempo y haberme apoyado tanto sin pedir nada a cambio, gracias de corazón.

Abstract

Tribbles (TRIB) homologues proteins are a family of three pseudokinases with an important role in controlling immunity, metabolism, and cancer through protein-protein interactions (PPIs). TRIB pseudokinases have the same overall tertiary structure as the eukaryotic protein kinase domain but lack some of the main motifs needed for its catalytic function, such as the DFG motif and a modified glycine-rich loop. This is why this family of proteins lacks catalytic function and has been shown to function as scaffolds. TRIB can interact with kinases (e.g., AKT), E3 ligases (e.g., COP1, SIAH1), and even transcription factors (e.g., C/EBPα). Nevertheless, little is known about how TRIB can interact and modulate other protein functions. For example, TRIB have been seen to act as tumor-suppressors or oncogenes in colorectal cancer in a context-dependent manner. This highlights the need of studying TRIB function in a way that provides temporal and spatial control, such as a chemical probe. A small molecule targeting TRIB will potentially modulate TRIB activity and therefore could help to better understand their role. Moreover, the compound could serve as a potential starting point for drug discovery.

Therefore, this PhD thesis aims to find small molecules that could regulate TRIB function and help to better understand this specific family of pseudokinases. To date, only TRIB2 has been targeted with a covalent compound probing an effect of destabilization and subsequent degradation. In contrast, TRIB1 is unable to interact with ATP but could potentially bind small molecules. In this thesis we have applied a combination of computational and biophysical techniques to study TRIB1 ability to bind small molecules. Thanks to these findings, we demonstrate that TRIB1 can bind small molecules in a similar way to canonical kinases.

At the same time, we wanted to look in more detail some of the PPIs reported for TRIB. As described before, TRIB can interact with kinases, such as AKT/PKB which is a critical regulator of cell survival and proliferation. TRIB3 was described to regulate AKT phosphorylation in the liver, although the mechanism is still poorly understood. Another case under study has been TRIB3 and SIAH1, an E3 ligase linked to ubiquitylation and degradation of TRIB3. The interaction of SIAH1 with TRIB3 has been already described, however, the specific mechanism of how these two proteins interact remains unknown. In this thesis, we have been interested on elucidate these PPIs since they are very relevant for understanding how TRIB functions.

Contents

List of Figures

List of Tables

Abbreviations

AD	Activation Domain
ADMET	Absorption Distribution Metabolism Excretion and Toxicity
AGC	Protein kinase A, G, and C
AP-MS	Affinity Purification combined with MS
ATP	Adenosine Triphosphate
CADD	Computer-Aided Drug Design
CaMK	Calcium/ Calmodulin-dependent Kinases
CASP	Critical Assessment of Protein Structure Prediction
CK1	Casein Kinases
CMGC	Cyclin-dependent Kinases
Co-IP	Co-immunoprecipitation
CryoEM	Cryo-Electron Microscopy
DSF	Differential Scanning Fluorimetry
EMA	European Medicines Agency
FDA	Food and Drug Administration
FP	Fluorescent Polarization
FRET	Förster Resonance Energy Transfer
IDP	Intrinsically Disordered Protein
IDR	Intrinsically Disordered Regions
IM-MS	Ion mobility Mass Spectrometry

ITC	Isothermal Titration Calorimetry
MD	Molecular Dynamics
NMR	Nuclear Magnetic Resonance
PCA	Protein Complementation Assay
PDB	Protein Data Bank
POI	Protein of Interest
PPIs	Protein-Protein Interactions
RGC	Receptor Guanylate Cyclase
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
SBD	Substrate Binding Domain
SBDD	Structure-Based Drug Discovery
SPR	Surface Plasmon Resonance
STE	Homologs of Sterile 7
TBM	Template-Based Modeling
TK	Tyrosine Kinases
TKL	Tyrosine Kinase-like
TRIB	Tribbles
Y2H	Yeast Two-Hybrid

1 Introduction

Research and curiosity are essential parts of human evolution. We cannot consider one without the other. Throughout history, questions have arisen from our ignorance and curiosity allowing us to build up new fields of knowledge. For example, Biology or Chemistry have grown over centuries thanks to making the right questions and finding the correct answers. From an etymological point of view, biology comes from the Greek word "bios" which means life, and "logos" meaning study, and altogether is defined as "the science of life and living things". On the other hand, chemistry has different versions of its origin. It seems that chemistry evolved from alchemy, the Arabic "al-kimia", which is a modification of the ancient Greek "chemeia" that denotes the art of transmutation [1]. Chemistry can be defined as the study of the composition of materials and the changes they undergo.

The combination of both fields gave rise to a new branch of knowledge, biochemistry, the chemistry of life. Trying to understand and depict in a meaningful way all the processes carried out in a living organism is not an easy task. Like any other problem, we need to break down its components to successfully address it.

1.1 Proteins: essential elements of life

One of these essential elements is the protein, it is not a coincidence that protein is the Greek word prōteios, which means "the first quality" [2].

Proteins are composed of amino acids connected by a covalent peptide bond. An amino acid, as the name implies, is an amino group and a carboxylic group with a side chain that confers specificity between the 20 possible combinations encoded in the human genome. The peptide bond is formed between the carbonyl group of one amino acid and the amide group of another. Several amino acids connected give rise to the peptides, where they differ from proteins on their smaller size, usually between 2 and 50 residues. Although there is no restricted definition of the actual size of peptides [3].

The composition of amino acids determines how the protein will fold and therefore its 3D structure. This is due to the type of interactions that can occur between the amino acids which leads to the overall folding of the protein. The types of interactions within a protein can be classified as covalent or non-covalent. Apart from the peptide bond, there is another covalent interaction that can occur between cysteine residues, disulfide bridges. The reduced sulfhydryl or thiol group (SH) of two cysteine residues react with one another to make a disulfide bond. Of course, this type of interaction is limited to the orientation and proximity of the cysteine side chains. Moreover, intracellular proteins are usually in a reducing environment which prevents the formation of disulfide bridges, sulfhydryl groups are favored over disulfide bridges. In contrast, extracellular compartments are less exposed to the reducing environment and therefore more prone to this interaction.

Non-covalent interactions are the remaining forces to stabilize protein folding. These types of interactions depend on the electrostatic attraction between opposite charges. The most known are the van der Waals interaction, hydrogen bond, and salt bridge, sorted from lower to higher free energy (van der Waals 0.5-1 Kcal/mol, hydrogen bond 1-10 Kcal/mol, and salt bridge 4 Kcal/mol) [4]. Hydrogen bonds cover a broad range from very strong, having a covalent character, to very weak, having energies slightly above van der Waals interactions [5, 6]. These forces are weaker in comparison to covalent interaction where the free energy is much higher (e.g., covalent bond 80 kcal/mol). Even though they contributed far less to the enthalpy contributed by a single covalent bond, they appear in a large proportion, so their contribution is still very important [6]. Moreover, thanks to not being so restrictive, it enables proteins to have structural flexibility and therefore conformational changes important for their function.

1.1.1 Protein Structure

From a structural point of view, proteins can be divided into several hierarchy levels (Figure 1.1). The first level of protein structure hierarchy is the ordered sequence of amino acids linked by peptide bonds, referred to as the primary structure. The primary sequence tends to form three-dimensional structural motifs such as α helix, β sheets, or loops. These motifs constitute the secondary structure of the protein. The tertiary structure is the placement of the secondary structures with each other to form the whole 3D structure. The quaternary structures correspond to the arrangement of two or more polypeptide

chains to form a complex. Usually, the tertiary structure is the biological unit, which corresponds with the functional form of the protein. Even though there are cases where the protein needs to form a complex to carry out its function. Proteins can be still functional without a defined structure, known as intrinsically disordered proteins (IDPs). In this case, proteins are unable to fold into a well-defined 3D structure. Instead, they are dynamically disordered and change within a range of conformations. This lack of structure allows proteins to interact and function within a wide range of processes [7]. There are some cases where the full protein will be disordered whereas in others just a domain or region, referred to as intrinsically disordered regions (IDRs) [8]. An important state about proteins is that structure is more conserved than sequence. This was exposed in the late 80s by Chotia and Lesk, where they showed the relationship between the root mean square deviation (RMSD) of the atoms and sequence identity. Later it was proven how the structure is three to ten times more conserved than sequence [9, 10]. How proteins fold is still a paradigm, known as the protein folding problem or Levinthal's paradox. Levinthal stated that a protein cannot find its correct native state by a random search since this will take a huge amount of time. Surprisingly, proteins do not sample all possible configurations since they fold in few seconds, that is why the paradox [11].

Nevertheless, over the last decades, several methods have been used to solve the 3D structures of proteins. Thanks to experimental techniques such as X-ray crystallography, nuclear magnetic resonance (NMR), and cryo-electron microscopy (CryoEM) we can see the 3D structure of proteins. Since 1971, all the structures solved worldwide have been deposit in the Protein Data Bank (PDB).

X-ray crystallization is one of the most used methods for protein structure determination. The technique consists of several steps from purifying the protein, obtaining a crystal and scattering X-ray by the electrons thrown into the crystal. It is important to know that one of the limitations of this technique is the amount of time needed without a high success rate. Protein purification can be highly challenging due to the complexity, size, and stability of the protein under study. Moreover, protein crystallization is a phase separation phenomenon where several factors affect crystallization. For example, the purity and concentration of the protein, pH, and temperature. Making this a very iterative and heuristic step. Finally, once the crystal is obtained, we will face the next problem, the accuracy of the model. To assess the accuracy and

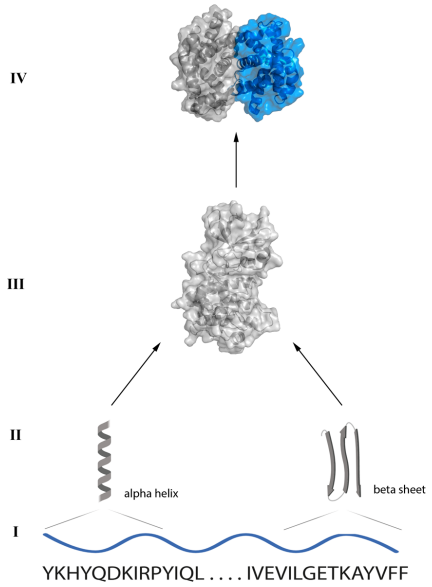


Figure 1.1. Hierarchy classification of protein structure. The roman numbers correspond to each level from first (peptide sequence) through fourth (quaternary system).

quality of the crystal structure two parameters need to be measured, resolution, R value and R-free. The resolution refers to the quality of the data that has been collected. It measures how well the density map calculated fits with the expected atomic positions. High resolution structures have a resolution of 1 Å. Usually, a resolution lower than 2.5 Åis considered good since most of the atoms are seen in the density map and are not inferred. R values measure the quality of the atomic model obtained from the crystallographic data. It is a measure of the diffraction pattern and the predicted intensities from the model. R values lower than 0.2 are considered reliable. Of course, there is a bias by using the same data to optimize the model. To avoid this, it was introduced the R free where a subset of the experimental observations (10%) is removed, and the refinement is performed using the remaining data (90%). The R free is calculated based on how the model predicts the 10% subset not used in the refinement. Apart from the technical limitations, we should also consider the limits of the technique. A protein is a dynamic entity where crystallization will

force and pack the protein in a static conformation. Therefore, assay conditions as solvents, pH, and temperature will conditionate the results and we need to be careful to misinterpret of electron density maps [12, 13].

Nuclear Magnetic Resonance (NMR) spectroscopy can determine the structure of a protein in solution. It is necessary to purify the protein to place in a strong magnetic field and then probed with radio waves. Resonances are analyzed to determine how close are the atomic nuclei from the protein atoms. This information is used to build the model of the protein. One of the limitations of this technique is the size of the protein, so it is used for small-medium proteins (less than 30 kDa) [14]. On the other hand, the advantage of this technique is that allows us to study the flexibility of proteins. That is because the protein is in solution, unlike crystallization. Therefore, NMR structures are an ensemble of protein structures where the flexible regions have fewer restraints in contrast with the more rigid ones.

Another technique used to determine the 3D structure of large macromolecules is electron microscopy, specifically cryo-electron microscopy. This technique is based on transmission electron microscopy where an electron beam transmits through the protein to form an image. The protein is purified and cryo-cooling in a thin layer of vitrified water. Then, 2D projection images are obtained to decipher the 3D structure. A major advantage of this technique is that the proteins do not need to be fixed in contrast with X-ray crystallization. Cryo-EM is used to solve the structure of macromolecular complexes such as ribosome or membrane proteins [15]. The major disadvantage is the very low signal to noise ratio which compromises obtaining high resolution structures. Despite this and thanks to technological advances, cryo-EM is emerging as a technique to obtain the structure of large proteins complexes [16].

In addition to experimental methods, computational techniques are used to predict the protein structure. There are two general approaches to predict the structure of a protein: template-based modeling or template-free modeling. As the name implies, the methods depend on whether an already solved structure is used as a template [17].

The regular steps in standard template-based modeling include:

- Identification of a suitable structural template. It consists of the selection of one or several templates from protein structures deposited in the PDB.

Several tools have been developed to find suitable templates (e.g BLAST, HMMER). Ideally those programs will find the template with the highest similarity with the target. Template structures are mainly chosen on the target-template sequence identity, but other criteria can be considered. For example, the experimental accuracy or conservation of active site or ligand of interest. The choice of the template is a matter of sequence homology, biological function, and structure resolution.

- Alignment of the target sequence to the template structure. The higher the sequence similarity between the target and the template the better the results will be. As a rule of thumb, lower than 30% of similarity, known as the twilight zone, will not give a very reliable model [18]. Multiple sequence alignment is usually preferred to overcome problems with high variable regions. Moreover, since we are using as a template an already solved structure the alignment can be always improved by including structural information from the template. The first two steps are critical since they could lead to a completely wrong model.

- Modelling of variable regions (mutations, insertions, deletions) present in the alignment. Usually, the alignment between model and template can contain gaps. To improve the quality of the variable regions several refinements can be applied. Loops and side chains can be refined through database search in PDB or rotamer library, and conformational search approaches [19]. This allows us to refine the model to eliminate clashes or forced torsion angles. One of the most common strategies is the energy minimization, where the predicted structures are optimized based on a score function or a force field. Conformational sampling can be done by using different strategies such as molecular dynamics or Monte Carlo simulations.

- Model evaluation. Once the model is obtained will be necessary to evaluate and select the best solution. The quality of the model can be inferred from the sequence similarity between the target and the template. Overall protein folding is usually correct when the similarity is higher than 30%. Other parameters like stereochemical issues (e.g., bonds, dihedral angles, clashes) need to be checked to evaluate if the model satisfies spatial restraints [19].

Template-free modeling is much more challenging and relies on large conformational sampling and the application of an energy function or force field.

This method, also known as ab-initio, has been applied to smaller proteins due to the inner complexity as the number of residues increase. There are also semi ab-initio methods that are based on structural fragments or use statistical information to build the model. Ab-initio methods depend on three factors: an accurate energy function to predict the native state which is the most thermodynamically stable; conformational sampling to identify low-energy states and a model evaluation strategy to select the best models [20].

Several programs have been developed to predict protein structure, from webservers to automatize the process (e.g. I-TASSER, SWISS-MODEL) up to computer programs (e.g. MODELLER, Rosetta). Depending on the user experience there will be a preference for one or another. To assess the accuracy of the different methods a competition was launched in 1994. CASP or Critical Assessment of protein Structure Prediction provides independent analysis on the state of the art in protein structure modeling. This competition is conducted every two years where different groups submit their results without knowing the experimental structures, just the sequence. One of the biggest advanced in protein structure prediction was driven by the application of deep learning techniques to predict inter residue distance [21]. Very recently the long-standing problem of protein folding could have been partially solved thanks to this technique. A company based in London, named DeepMind, obtained high resolution models without a template in the CASP competition by using their method, AlphaFold [22]. Later, it has been published a similar approach, but with less accuracy, using neural networks by the research group of David Baker at the University of Washington [23].

The central idea of AlphaFold is a convolutional neural network that is trained on PDB structures to predict the distance between atoms [22, 24]. Of course, without all the effort previously made to obtain experimentally all the structures of the proteins, this could not be achieved. As they even stated, we do not know if it will work with all the proteins or with even more complex problems as the protein-protein or the protein-ligand interactions. In either case, together with experimentally solved structures, computational methods provide relevant information to understand better the protein structure-function relationship. This is a great example of how the progress in scientific research is made, which leads us to be closer to the main goal of understanding the diseases and therefore being able to cure them.

1.2 Molecular recognition

Molecular recognition in biological systems as proteins is essential for the process and function in the cell. This process is highly dependent of many factors as the surface complementarity, thermodynamics, and physicochemical properties such as hydrophobicity and electrostatic forces. The first time this was reported, it was by Emil Fischer in 1894 by the lock and key model. Fischer proposed that a substrate is like a key which fits specifically into a lock or binding site of the enzyme [25]. This basic concept was further developed by Daniel Koshland to propose the theory of induced fit. The ligand and macromolecule are described as dynamic components, conformational changes of the protein are necessary to optimize their fit [26]. The components differ in their unbound (apo) or bound (holo) state. This theory is applicable to most molecular binding processes since rarely have a fixed structure. An alternative model to the two previous theories is the conformational selection, where a conformational change occurs before the binding of a ligand. In this case the ligand will select and stabilize a conformation for binding. On the contrary, in the induce fit, the conformational change occurs after the ligand binding. An extended version of this model includes the induce fit, where conformational selection is followed by conformational adjustment [27].

The association between the ligand (L) and a protein (P) can be formulated as, for a noncovalent reversible association [28]:

$$P + L \underset{k_{off}}{\overset{k_{on}}{\leftrightarrows}} PL \tag{1.1}$$

Where PL represents the protein-ligand complex, k_{on} and k_{off} are the kinetic rate constant of the association (on) or dissociation (off). The equilibrium binding constant is defined as:

$$K_b = \frac{[PL]}{[P][L]} = \frac{k_{on}}{k_{off}} = \frac{1}{K_d} \tag{1.2}$$

Where the K_d is the dissociation constant. A tight molecular recognition is expected to have a fast association and slow dissociation, leading to a high binding affinity. The stability of a biological complex is determined by the binding free energy (ΔG), which only takes places when it is associated with a negative binding free energy (-ΔG). The binding free energy is expressed as:

$$\Delta G = \Delta H - T\Delta S \tag{1.3}$$

Which is composed of an enthalpic (ΔH) and an entropic ($-T\Delta S$) term. The standard binding free energy is related to the dissociation constant as

$$G = -RTlnK \tag{1.4}$$

Where R is the gas constant, T is the temperature and K the equilibrium constant.

Thermodynamic properties as the binding free energy (ΔG), enthalpy (H), and entropy (S) are directly involved in the ligand-protein complex. The binding free energy (ΔG) determines the stability of the system. The physical meaning of binding enthalpy (ΔH) represents the loss of noncovalent interactions. The ΔH is the results of the formation and breaking of many different bonds, such as hydrogen bonds with water molecules, the subsequent formation of interactions between the ligand and protein, and solvent reorganization near protein surface. Proteins are dynamics entities and therefore flexibility plays an important role. Conformational changes in both ligand and protein contribute to the total energy of the system. The binding entropy (ΔS) is the sum of several contributing effects. The main factor is the displacement of solvent (hydration effect) as well as the reduction of rotational degrees of freedom of ligand and protein. A positive entropy change usually indicates that water molecules haven been released from the complex. Higher binding affinities caused by increased intermolecular interactions will have large negative enthalpic contribution, will be at the expense of increased order leading to a more negative ΔS. Unfavorable contributions include desolvation, strain, and restricted rotation [29, 30].

The balance between these two properties, known as enthalpy-entropy compensation, have been described before although the contrary has been also observed [31]. In any case the nature of these compensatory mechanism is very system dependent and do not obey a single rule [31, 32].

In medicinal chemistry, it has been tried to rationalize the enthalpy-entropy compensation to help increase the binding affinity. The ideal optimization will be the implementation of enthalpic or entropic contributions that lead to minimal penalties.

1.2.1 Intermolecular interactions

Experimental data helps to elucidate the important role that complementarity plays in the binding of a ligand to the protein. Noncovalent intermolecular

interactions are responsible for molecular recognition and subsequent complex formation. A systematic analysis of the atomic protein-ligand interactions in the PDB has been already performed [33]. The authors were able to calculate the frequency of the most common non-covalent interactions observed in protein-ligands. Among the most frequently observed were, ranked from most to least frequent: hydrophobic, hydrogen bonding, pi-stacking, weak hydrogen bonding, salt bridge, amide stacking and cation-pi.

Hydrophobic interactions

According to the IUPAC (international Union of Pure and Applied Chemistry), it is defined as the tendency of hydrocarbons to form intermolecular aggregates in an aqueous medium, and analogous intramolecular interactions [34]. The hydrophobic effect causes a disruption in the structure of bulk water and decreases entropy because of stronger bonding and ordering of water molecules group [32]. Removal of the hydrophobic surface area from water to bind into a hydrophobic pocket, it is estimated to be approximately 0.7 kcal/mol or a 3.5 fold increase in binding constant for a methyl group [32, 35].

Hydrogen bonding

They are considered one of the most important specific interactions in molecular recognition. Hydrogen bonds are electrostatic interactions between two dipoles, hydrogen donor group where the hydrogen is covalently bonded and forms an attractive interaction with another electronegative group, the hydrogen acceptor group. This type of interaction is highly dependent of the geometry and distance of the atoms involved. The angle donor-hydrogen-acceptor is generally above 150^{o} [32]. The distance between the hydrogen bond of the donor and the acceptor atom ranges from 2.6 to 3.1 Å [36]. Hydrogen bonds are one of the main contributors of binding selectivity thanks to their inner flexibility due to the distance and angle constrains. It has been determined that a hydrogen bond can contribute between 0.5 and 4.7 kcal/mol to the binding energy depending on the charges and environment [35]. On the other hand, not satisfying a hydrogen bond donor, a hydroxyl group, has a penalty of approximately 5 kcal/mol because of desolvation [32]. There is also evidence of weak hydrogen bond, where one or both dipoles are of medium-low electronegativity, e.g., carbon atom is acting as hydrogen bond donor. Weak hydrogen bonds have more variable geometry than strong hydrogen bond [37]. Stronger interactions contribute more enthalpically while the weaker interactions could compensate for the desolvation cost.

π-stacking

Type of interaction between aromatic rings thanks to π bonds, it is caused by intermolecular overlapping of p orbitals in π conjugated systems. They can interact through several conformations as: face to face, edge to face (T shape), and offset [38].

Salt bridge

Ionic bonds caused by the attraction of two oppositely charged atoms. Their strength varies between 5-10 kcal/mol and it is very dependent on the environment [29, 39]. Moreover, their contribution is masked by the energetic penalty of desolvating charged groups.

Amide π stacking

Interactions between an amide group and aromatic ring, π-surface of the amide bond stacks against the π-surface of the aromatic ring [40].

Cation π stacking

This type of interaction occurs between a positively charged molecule and a π system (e.g. aromatic ring, alkenes, alkynes) [29].

Halogen Bonds

An attractive interaction that can form between an electrophilic region of a halogen atom with a nucleophilic group [41]. This type of interactions has been widely used to explore steric and electronic effects for lead compounds. Moreover, it is used not only to increase the affinity but also the membrane permeability and metabolic stability of the compound [33].

1.3 Disease: Drug Discovery & Kinases

Proteins are considered essential for life since they are present in almost all cell functions. The biology dogma states that proteins are translated from genes but not all the genes can transcribe proteins. Once the human genome was solved in the early 2000s, one of the milestones in science was achieved. Since then, our knowledge on how genes are regulated and connected between them in a healthy organism, but more importantly in disease context has grown exponentially.

We are beginning to understand the complexity of what is indeed happening thanks to the advance in basic science like molecular biology and other related disciplines. The malfunction of proteins is directly linked to a disease due to their central role in biological function. For example, neurodegenerative diseases such as Alzheimer's or Parkinson's are caused by protein misfolding and aggregation [42]. Other diseases, such as cancer are usually caused by mutations affecting protein's activity or stability by causing an allosteric change [43]. From the 30,000 genes found in the human genome, approximately 20,000 encode for proteins but only about 4% of them have been targeted by drugs. The current number of FDA-approved drugs that are directly related to a human protein according to its mechanism of action is 754 [44]. In 2006 the number of approved targeted drugs acted through 324 protein targets, which increased to 667 by 2017 [45]. This small percentage shows us how far we are still from targeting most of the diseases and more precisely how difficult it is to develop drugs for new targets. Only 20% of the approximately 20,000 protein-coding genes are considered druggable [44]. To target a specific protein, some features need to be taken into consideration. The term druggable is used to define how likely a protein or a protein complex can be targeted by a drug. This definition is very broad and implies several things about the target. Usually, the activity of the protein is well characterized and their function or the absence of it will trigger the disease. Also, the 3D structure of the protein is known, which helps in the design of the small molecule to interact with the protein. To have a crystal structure of a small molecule bound to the target protein will help to rational design and optimize the compound properties and affinity. Although being able to find a small molecule to interact with the protein is not enough to obtain a drug [46, 47].

Indeed, it is just the beginning of a long process to achieve a therapeutic effect. Once a small compound is found to interact with the target (i.e. hits), they need to be optimized to find lead compounds that could be developed into drugs. The next generation of compounds is developed through studying their drug metabolism and pharmacokinetics. This process is usually known as ADMET studies, which is the abbreviation of Absorption, Distribution, Metabolism, Excretion, and Toxicity. If the compound passes this threshold, it will enter clinical trials. This is divided into 4 phases to answer specific research questions. In phase I the purpose is to determine the safety and dosage of the new drug. At this stage, the number of participants is low, around 100 healthy volunteers and it usually takes several months. If the targeted disease is cancer, the drug will be provided to people with the disease.

In phase II the objective is to see the efficacy and side effects involving a few hundred patients. The length of study is very broad from several months up to 2 years. Phase III has the objective to test the efficacy and to demonstrate whether the drug offers a treatment benefit. The number of participants is around a few thousand and the length is from 1 to 4 years. In this stage, most of the safety data is provided due to extended time and a higher number of patients. If the drug is approved by the corresponding agency, it will reach the last stage, phase IV, or post-market safety monitoring. The purpose of this phase is to monitor the possible long-term side effects and to probe the safety and efficacy. The FDA (Food and Drug Administration) in the USA and the EMA (European Medicines Agency) in Europe are examples of agencies in charge of the safety and regulation of medicines, although in Europe the EU has the last word on drug approval [48]. The idea to target a protein is based on being able to control and regulate cell homeostasis. Drugs have been developed against a myriad of diseases with varying levels of success. Looking into ChEMBL, an open-access bioactivity database, we can see which protein targets are more overrepresented and studied than others (Figure 1.2). ChEMBL contains information manually extracted from scientific literature (e.g Medicinal Chemist Journals) and integrated with data on approved drugs and clinical development [49].

The families of proteins that have been mainly targeted are GPCRs (G protein-coupled receptors), ion channels, protein kinases, and nuclear hormone receptors. Indeed, these families of proteins are responsible for the therapeutic effect of most small molecule drugs. The term small molecule is used in drug discovery to refer to the chemical compound able to interact with proteins. The name implies the difference in size since they are small compared with proteins [45, 50]. Apart from small molecules, there are other ways to target proteins, called biologics or biological products. In contrast to chemically synthesized small molecules, biological products are derived from living cells or through biological processes and are more complex in structure. Biologics have high molecular weight and due to their complexity are not entirely characterizable, unlike small molecules which are completely characterizable. As examples of biologics, we can find hormones, vaccines, monoclonal antibodies, recombinant proteins, or even gene and cellular therapies [51].

1.3.1 Protein Kinase Superfamily

Proteins can be classified in several ways according to sequence, structural similarity, and function. Already well characterized proteins are used as a ref-

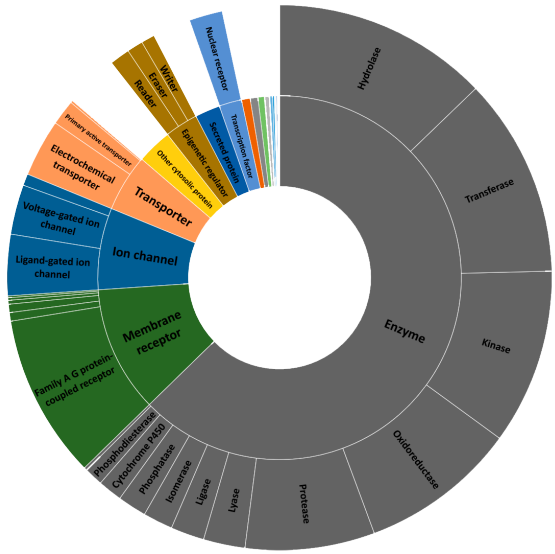


Figure 1.2. Families of proteins present in ChEMBL. The classification is arranged in concentric circles with major families in the center and smaller subfamilies around the edge.

erence to classify in the same family. A protein family is a group of proteins that share a common evolutionary origin and therefore they usually have the same function and similarities in structure and sequence. They can be subdivided into smaller groups or subfamilies which collect smaller groups closely related. Then, all the subfamilies whose structures and functional features suggest a common origin are encompassed in a bigger group called superfamily [52]. From a structural point of view, proteins can be classified into domains that usually determine their overall function. It is important to notice that similar domains can be found in proteins with different functions. Moreover, inside the domains, we can find sequence features (motifs) that can be also used to classify proteins. As an example of a superfamily, kinases are one of the largest superfamilies in humans with 555 members. They have been grouped in the main class of eukaryotic protein kinases (497) and atypical kinases (58) [53]. Inside this superfamily, it has been also described as a subfamily name pseudokinase. This subfamily will be further explained in the following sections.

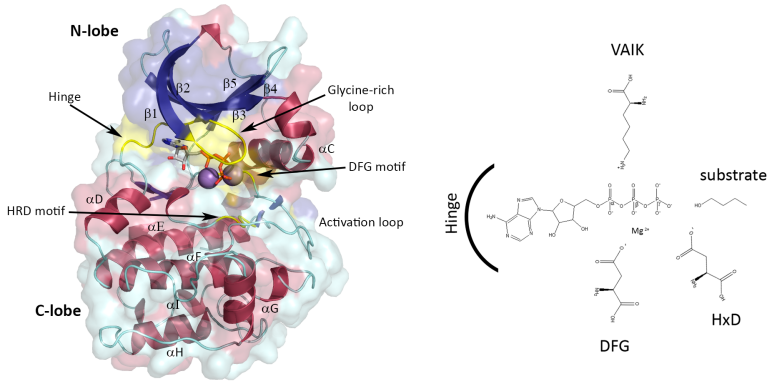


Figure 1.3. Protein kinase and schematic representation of ATP and key residues.Protein kinase structure shown as ribbon diagrams with the surface displayed. β sheets and α helix colored blue and red respectively. Schematic representation of ATP and key residues (motifs) in the canonical kinase pocket. The β3 lysine (K in the VAIK motif), the aspartic acid pre activation loop (D in the DFG motif), and the aspartic acid in the catalytic loop (D in the HxD motif).

Kinases catalyze the phosphorylation reaction by attaching phosphate onto hydroxyl groups of substrates such as proteins, lipids, or sugars. The opposite are phosphatases which remove a phosphate group from the substrate to maintain homeostasis. Kinase structure and function are well conserved. To carry out their enzymatic mechanism, kinases need to recognize the substrate to phosphorylate and the ATP (adenosine triphosphate) to transfer the gamma phosphate (Figure 1.3).

Protein kinases are formed by a bilobal fold. The N-terminal lobe (N-lobe) has five β strands, named 1 through 5, and the αC helix. The C-terminal lobe (C-lobe) is mostly helical. The ATP molecule binds in the pocket between the lobes (canonical ATP pocket). The PKA was the first protein kinase to be crystallized [54]. The N-lobe experiences structural rearrangements when the kinase is active or inactive. It is possible to find two important sequence motifs within the first three β strands. The first motif is the Glycine rich loop between β1 and β2 (GxGxxG). It is the most flexible part of the N-lobe because of the several glycines. This loop folds over the nucleotide and

anchors the phosphates of ATP. In some kinases, it can be found the P-loop, also known as Walker-A motif (GxxxxGKT/S) instead of the glycine rich loop. The main difference between the two is that the P-loop does not have any contact with the adenosine of ATP whereas the Glycine rich loop connects two β strands which accommodate the adenosine. The other motif is the VAIK (more generally AxK) located in the β3 strand. The conserved lysine residue helps to couple the phosphatases of ATP to the αC helix. In most of the kinases, the conformation of αC helix is a key region that determines the active or inactive state of the enzyme. In the active conformation, it is rotated toward the ATP pocket, allowing the lysine located in the β3 strand to form a salt bridge with the conserved glutamate in the αC helix (αC in). Disruption of the salt bridge is a strong indicator of inactivity where the αC helix is usually outward the ATP pocket (αC out) [55–57]. In contrast, the C-lobe is more rigid and has less flexibility with the exception of the activation loop. This loop starts right after the DFG motif (aspartic acid, phenylalanine, glycine) and includes a phosphorylation site for most of the kinases. When the activation loop is phosphorylated it adopts an extended conformation that stabilizes the active state of the protein. In the active conformation the aspartic acid of the DFG motif points into the cavity to coordinate Mg^{2+} and ATP, known as the DFG-in state. In the DFG-out conformation, the aspartate points out from the ATP pocket and the phenylalanine prevents the binding of the ATP molecule. In this inactive state (DFG-out) the phenylalanine switching opens a hydrophobic pocket between the active site and the αC helix. Changes in the activation loop are linked to conformational changes of the αC helix [58]. Moreover, the C-lobe serves as a binding region for substrates

Kinases also conserve two clusters of hydrophobic residues which are very important for kinase regulation [56]. These clusters are also known as spines. The regulatory or R-spine is formed by four hydrophobic residues located in the β4 strand (leucine), the αC helix (leucine), the activation loop (phenylalanine), and the catalytic loop (tyrosine). The catalytic or C-spine comprises residues from both lobes as well but in this case, the adenine ring of ATP is also included. The residues are in the β2 (valine), β3 (alanine, from the AxK motif), β7 (two leucines), D-helix (methionine), and F-helix (leucine and methionine). These spines are disrupted in inactive conformations, making them very dynamic modulators [56]. The gatekeeper residue is found at the end of the β5 strand, between the R and C-spine. It is the first residue of the hinge connecting the C and N-lobe. It is usually a very hydrophobic amino acid as leucine, methionine, or phenylalanine although in some kinases can be smaller as threonine or valine.

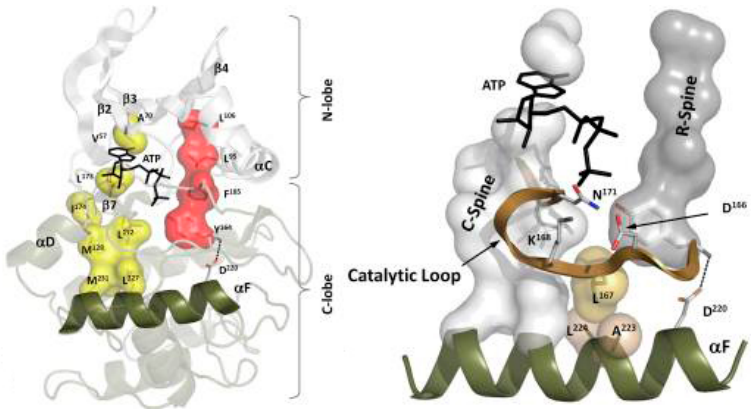


Figure 1.4. Conserved spines and key residues in kinases. The R-spine and C-spine are linked to the F-helix. Figure extracted from [56] with permission from Elsevier.

The gatekeeper residue regulates the accessibility to the back ATP pocket, which contributes to the selectivity of kinases for small molecules [59].

1.3.2 Kinases: one of the main targets in the 21st century

The Kinase family is one of the most frequent protein families implicated with cancer in many different functions such as proliferation or metastasis [60]. This is due to the great important role of phosphorylation in cell signaling. The dysregulation in the phospho-dephosphorylation state in many signaling pathways, e.g. Mitogen-activated protein kinase (MAPK) or Cyclin-dependent kinase, triggers a cascade of activation-inhibition of the cell cycle [61]. MAP kinase signaling has an important role in cancer growth and progression, where cyclins are the key regulators of the cell cycle [61]. That is why kinases are one of the most targeted proteins, and kinase modulators have been on the top list of therapeutic usage for cancer in the last decades. But before they became so popular, there was a time when Kinases were considered undruggable. The idea behind this was that since all kinases were known to recognize ATP in the same way, how would it be possible to design a drug targeting specific kinases? On top of that, ATP is a natural small molecule with a millimolar concentration in the cell, how is a drug going to compete against that? [62]. These questions and more have been solved over the years thanks to research and serendipity.

For example, some kinase inhibitors have proved to be more efficient thanks to the lack of specificity. Apart from inhibiting the main kinase target, additional kinases are targeted giving an improved activity. The simultaneous inhibition of several kinases can prevent drug resistance or even can be used to treat other cancers. Imatinib was the first kinase inhibitor approved by the FDA in 2001. It inhibits the Abelson (ABL) kinase for chronic myeloid leukemia (CML) being more efficient during the chronic phase of CML. The success of imatinib is that it transformed CML from a fatal disease to a manageable condition. Moreover, it is also effective in treating other cancers such as GIST (gastrointestinal stromal tumor) or HES (hyper-eosinophilic syndrome) [63, 64]. Before imatinib, other kinase inhibitors were approved although without knowing their mechanism of action. The first kinase inhibitor approved was fasudil in Japan in 1995 for cerebral vasospasm. Then, sirolimus (rapamycin) an inhibitor of the kinase TORC1 was approved in 1999 to prevent organ rejection in renal transplants [63, 64]. The success of imatinib leads the path to focus on kinase inhibitors in academia and the pharmaceutical industry. At the time of writing this thesis, there are 76 Kinase inhibitors approved by the FDA for the treatment of a wide variety of diseases [64]. As part of kinase inhibitor history, it is worth mentioning Sugen company which led to the discovery of many kinases that even today are still named according to the company (e.g Sgk269 or Sugen kinase 269). The first kinase classification was published by Munnings and colleagues from Sugen in 2002, known as Kinome (Figure 1.5) [65]. There are eight main groups of eukaryotic protein kinases: AGC (protein kinase A, G, and C); CaMK (calcium/ calmodulin-dependent kinases); CMGC (cyclin-dependent kinases, MAP kinases, glycogen synthase kinases, CDC- like kinases); TK (tyrosine kinases); STE (homologs of sterile 7); CK1 (casein kinases); TKL (tyrosine kinase-like); and the RGC (receptor guanylate cyclase) groups [66].

Type of Kinase inhibitor

Depending on where and how the small molecules bind, they can be classified into several groups (Figure 1.6). The main pocket is the canonical ATP binding pocket between the N and C lobes, which can be divided into several parts. The front pocket or front clef, gate area, and a back cleft. The front cleft covers the glycine rich loop, the adenine binding pocket, the hinge residues, and the catalytic loop. There is also a hydrophobic pocket or back pocket which includes the gate area and back cleft. Apart from the canonical ATP pocket, we can find other pockets in the kinase domain, such as allosteric (e.g.

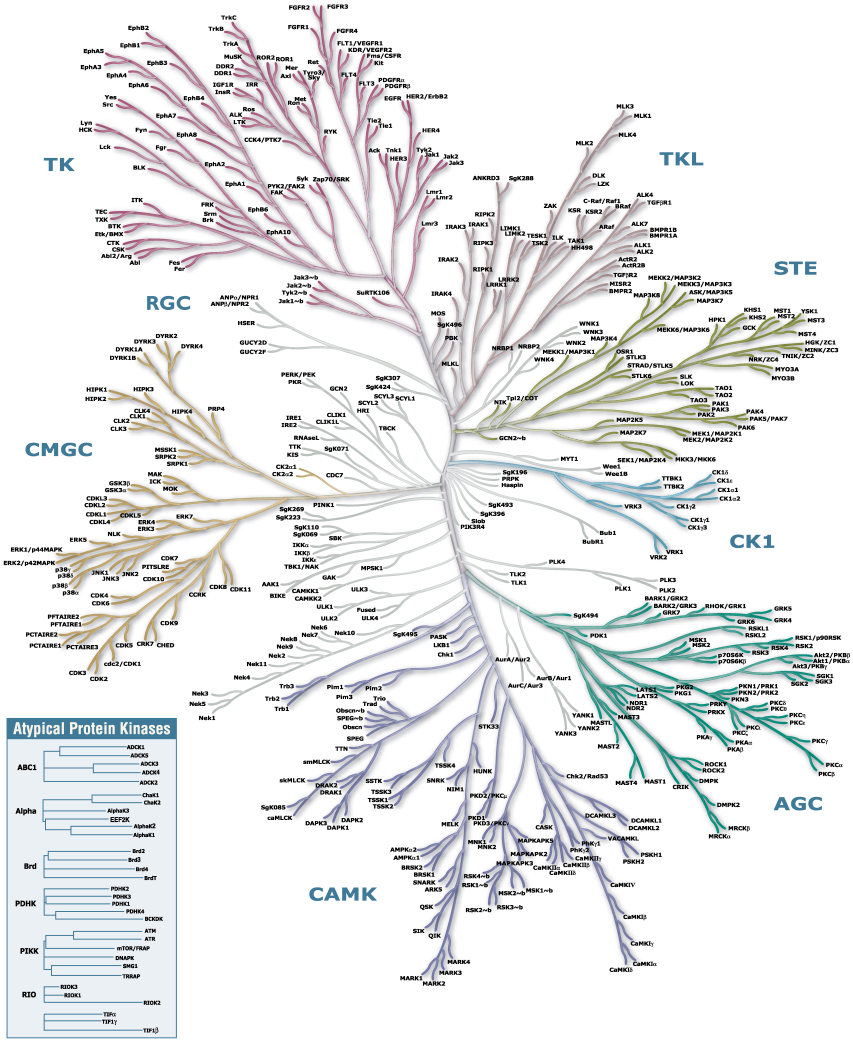


Figure 1.5. The human kinome. Illustration reproduced courtesy of Cell Signaling Technology, Inc. (www.cellsignal.com).

PDK1 Interacting Fragment or PIF pocket) or in other domains nearby (e.g. Pleckstrin homology (PH) domains). Small molecule kinase inhibitors can be classified into two main classes, reversible and irreversible inhibitors [67].

Kinase Reversible inhibitors

Given the plasticity and complexity of kinases, this type of inhibitor can be subdivided into six groups. Where type I, I ½ and II are ATP competitive and III, IV, and V do not necessarily bind in the canonical ATP binding pocket.

Type I. Inhibitors bind the canonical ATP pocket within the front cleft in an active state. The small molecule competes with the ATP binding in the active conformation with the DFG in and αC helix in. There is no occupancy in the back cleft and the activation segment is open (out). The small molecule mimics the interaction made by the ATP. The ATP forms two hydrogen bonds with the backbone residues of the hinge region, the first carbonyl backbone residue, and the NH group of the third hinge residue. To gain some selectivity, the gatekeeper residue and poorly conserved residues have been used to bias the design of the type I inhibitor. Despite the highly conserved ATP pocket, this type of inhibitor is one of the major classes of kinase inhibitors clinically approved [63]. Examples for this type of approved kinase inhibitor are gefitinib, erlotinib, dasatinib, and sunitinib where the suffix -tinib indicates tyrosine kinase inhibition.

Type I ½ . Binds in the canonical ATP pocket in an inactive state. It is a subtype of type I where the kinase adopts a DFG-in conformation, but with the C-helix variable (in or out). It has been also subclassified into two subtypes, subtypes A where the small molecule extends into the back cleft, and B where the compounds do not extend. Vemurafenib was discovered as a highly specific BRAFV600 kinase inhibitor with selectivity against melanoma cells [68]. On the other hand, dasatinib can also inhibit Lyn tyrosine kinase as type I ½ which is approved for chronic myelogenous leukemia (CML) acute lymphoblastic leukemia [69].

Type II. Binds in the canonical ATP pocket in an inactive state (DFG-out). It can be also subclassified as type I ½ into subtypes A and B, depending on the occupancy of the back cleft. When the kinase inhibitor extends into the back cleft (type A) the residence time is longer whereas binding to the front cleft and gate area only is associated with shorter residence times [69]. The inactive form of the kinases was considered more selective because of their structural

differences compared with the generic active state (ATP bound). However, this was proved to be false and there is no inherent selectivity advantage of type II inhibitors [70]. Due to the inactive state of the enzyme, type II needs specific assays to distinguish between active and inactive states, for example, fluorescent-based assays can help to identify the kinase conformation. The binding of imatinib to Abl with the DFG-D out con-figuration while extending into the back cleft was the first example of type II inhibition [71]. Sorafenib is an example of a non-selective type II inhibitor that can inhibit several kinases including B-Raf, CDK8, and the VEGFR1/2/3. This drug is approved for the treatment of hepatocellular, renal cell, and differentiated thyroid carcinomas [69].

Type III. Allosteric inhibitors bind next to the canonical ATP pocket. This type of inhibitor is not ATP competitive since ATP cannot prevent the interaction of the small molecule with the protein. Type III class binds in the allosteric back pocket stabilizing the inactive form of the kinase. The structural conformation of the DFG motif is variable but with the αC helix out, due to the occupancy of the back cleft. Type III has no physical contact with the hinge and shows the highest degree of selectivity by exploiting unique binding sites [72]. As an example of this type of inhibitor, we can find Trametinib and cobimetinib, both MEK inhibitors that have been approved for the treatment of melanoma bearing mutant BRAF V600E [69].

Type IV. Allosteric inhibitor binding away from the canonical ATP pocket. This type of inhibitor offers the opportunity to alter the kinase activity by disrupting upstream activators or downstream substrates by not competing with ATP. Some of the allosteric pockets found in the kinase domain are at the C-lobe (c-Abl myristoyl pocket, Docking site for ERK) [73, 74], at the interface between the kinase domain and other close domains (AKT/PKB and the PH domain) [75], and the N-lobe (PIF pocket of AGC protein kinases, ANS pocket in CDK2) [76, 77]. GNF-2, an antagonist of BCR-Abl, is an example of a type IV inhibitor that binds to the myristoyl pocket and stabilizes the inactive enzyme form [73].

Type V. Binds two distinct binding sites (bivalent). Bifunctional inhibitor that combines type I inhibitors via a chemical linker with a compound that targets a remote secondary binding site. Thanks to targeting two different regions, they offer a greater potential for increasing selectivity and potency. There are two main classes: bisubstrates and bivalent inhibitors. Bisubstrates inhibitors consist of two components, one targeting the ATP pocket and the other the

protein substrate binding domain [78, 79]. In 1991, the first serine/threonine-protein kinase C (PKC) directed bisubstrate inhibitor was developed. It consists of a pseudosubstrate peptide-specific PKC linked to an ATP competitive inhibitor [80]. Bivalent inhibitors are more general and can target any region outside of the active site. As an example, a bivalent inhibitor of the MAPK c-Jun N-terminal kinase 1 (JNK1) was generated by linking an ATP-competitive inhibitor of the JNK family, to a peptide that targets JNK1's docking groove. The bivalent inhibitor displayed a 20000-fold increase in JNK1 inhibition compared to the ATP competitive inhibitor alone [79, 81]. One of the major drawbacks of this type of inhibitor is how to transport them inside the cellular environment [82].

Kinase irreversible inhibitors

Type VI. This class can bind covalently to the kinase domain and can interact with the ATP or other regions in the kinase. These compounds contain an electrophilic group or warhead that react primarily with nucleophilic amino acid (e.g. cysteine, lysine, aspartic acid, and tyrosine) to form the covalent bond. One of the most common nucleophilic addition is the Michael reaction which represents the addition of a nucleophile to an alpha,beta-unsaturated carbonyl compound [83, 84]. The non-covalent interactions are also important to orient the small molecule in a productive position that allows the covalent bond formation. There are two types of covalent bond formation: irreversible and reversible. The irreversible is the normal formation of the covalent bond, whereas electron-deficient olefins (alkane) can be tuned to react with a nucleophile in a rapidly reversible manner [85]. α-Cyanoacrylamides have been used to create potent and selective kinase inhibitors that act by a covalent, reversible mechanism [86].

The advantages of covalent small molecules compared to non-covalent inhibitors are selectivity, potency, and longer duration of action. However, safety concerns and side effects of promiscuous reactive electrophiles make them less attractive [87]. Afatinib and ibrutinib are examples of approved covalent inhibitors, both containing an acrylamide group as the electrophile for covalent bond formation.

As stated before, kinases are very dynamic entities and therefore it is not surprising to see some inhibitors acting as type I or II depending on the kinase. For example, sunitinib can bind as type I ½ with CDK2 and type II with VEGFR or bosutinib which can bind as type I with Src or type II with Abl.

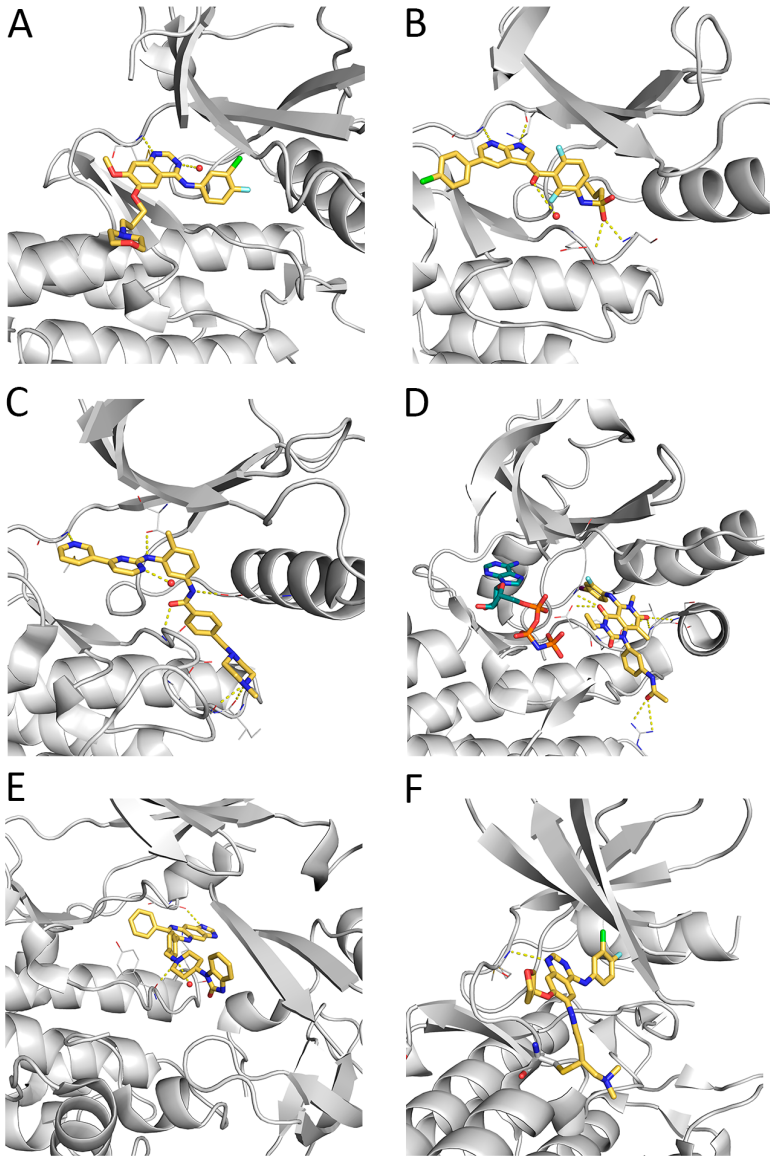


Figure 1.6. Binding mode of kinase inhibitors. A) Type I (PDB code 2ITY, gefitinib). B) Type 1 ½ (PDB code 3OG7, vemurafenib). C)Type II (PDB code 1IEP, imatinib). D) Type III (PDB code 7JUR, trametinib). E) Type IV (PDB code 3O96, AKT inhibitor VIII). F) Type VI (PDB code 4G5J, afatinib).

1.4 Drug Discovery Approaches to study protein-ligand interaction

Several strategies have been developed over the years to find hits for a new target or disease. Phenotypic screening or target-based assays are chosen as starting points depending on the target, disease, and the number of resources. On the other hand, computational techniques help to reduce the time and expense involved in the discovery and optimization of a new drug, known as Computer-Aided Drug Design (CADD). Depending on the amount of information gathered around the target it will be possible to follow alternative strategies. For example, when structural information of the target is known, Structure-Based Drug Discovery (SBDD) is applied. In other cases where there is a lot of collected data, other methods such as statistics or machine learning can be applied to discover and design new drugs.

Interfering with the function of proteins is the main strategy in the design and development of new drugs. The pharmaceutical industry has been applying high-throughput screening (HTS) of large libraries of compounds (usually higher than 10,000 compounds per day) to increase the success rate of finding hits for new targets. A variety of strategies have been used for HTS assay, from the measurement of catalytic activity of an isolated enzyme or a complex (e.g. kinases, GPCRs), cellular extract (e.g., to study ubiquitination and degradation), or phenotypic assay (e.g. selective killing cancer cells) [88].

Ideally, HTS is used as a primary screen and in parallel with other orthogonal techniques to discriminate false positives derived by the design of a single assay. For example, in a fluorescent-based assay, a fluorescent compound will give a signal and can be interpreted as a false positive. Therefore, applying different methods to the same target or disease increases the success and robustness of identifying new hits or starting points to develop new chemical probes or drugs through medicinal chemistry. Characterizing the interactions of the small molecule with the protein target is crucial at the initial stages of designing new drugs. It can help to elucidate the structure-activity relationship of the small molecule with its target and therefore improve not only the affinity but also the activity of the compound. As a first step, it is necessary to find compounds able to interact with the target protein. For that purpose, several techniques from experimental to computational have been developed.

Experimental techniques

There are several ways of experimentally characterizing the interaction between small molecules and proteins. It can be split into biophysical or cell-based depending on the type of assay. Biophysical methods allow us to directly measure the binding and even to calculate the Kd or its thermodynamics properties (e.g., H or S). They can be applied to almost any protein class if the protein is possible to purify, so they do not require to know the protein's function or design an enzymatic assay. However, the detection of protein-small molecule interaction using purified proteins does not entail any binding of the compound to the protein inside the cell. It is crucial to measure the binding of the small molecule to its target in the cell and to calculate if the compound reaches its target in the cell, i.e target engagement. Cell-based methods are highly important since without them it is difficult to attribute pharmacological effects to perturbation of protein of interest against other mechanisms as side effects [89, 90]. Between the biophysical methods most used, we can find differential scanning fluorimetry (DSF), surface plasmon resonance (SPR), isothermal titration calorimetry (ITC), nuclear magnetic resonance spectroscopy (NMR), and protein X-ray crystallography.

Differential Scanning Fluorimetry

The interaction between a ligand, such as a small molecule, and a protein will induce conformational changes translated into a stabilization or destabilization of the protein. As stated in the section section 1.2 the stability of a protein is related to its Gibbs free energy and is temperature-dependent. Proteins are in thermodynamic equilibrium between folded and unfolded states, the temperature at which the concentrations of folded and unfolded proteins are equal is defined as the melting temperature (Tm) [91]. Differential scanning fluorimetry (DSF), also known as ThermoFluor or Thermal Shift Assay, allows us to measure the Tm of a protein through monitoring thermal unfolding in the presence of a fluorescent dye [92]. By increasing the temperature, the protein starts to unfold and expose the hydrophobic core. The fluorescent dye is highly fluorescent in a nonpolar environment and starts to emit a signal upon binding to the hydrophobic surfaces. Therefore, it is possible to determine the Tm by monitoring the fluorescent signal (Figure 1.7). A change in the Tm may indicate a change in the equilibrium where a small molecule could stabilize or destabilize the protein. If the ligand binds the native protein will stabilize and increase the Tm, where binding more tightly to the unfolded will decrease the Tm [93].

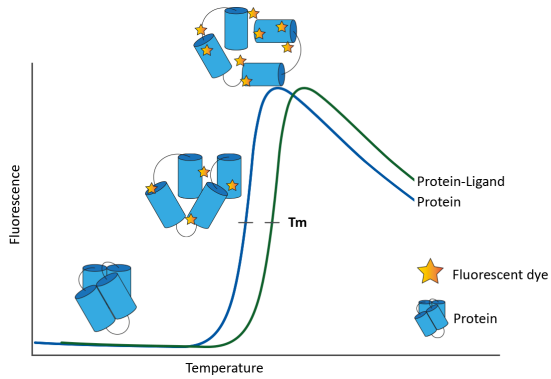


Figure 1.7. Example of a typical melt curve from a DSF assay and the thermal shift when adding a small molecule (ligand).

One of the limitations of the technique is its temperature dependence, making it difficult to calculate the thermodynamics of the protein-ligand interaction. A Tm is not unique to a specific binding affinity, different entropic and enthalpic contributions can produce the same change in Tm [91]. That is why DSF is mainly used as a qualitative assay rather than a quantitative assay [94]. Compounds with fluorescent properties are another restriction of the technique since they could interfere with the dye, giving false positives. Another limitation for ligand screening is the compounds with fluorescent properties since they could interfere with the dye, giving false positives.

Despite their limitations, DSF has been extensively used because of being easy to implement, even for HTS, cheap, and enables to obtain hits on early stages of drug discovery. This method has been also applied in protein buffer optimization, finding optimal conditions for storage, assay screening, and crystallization to help to increase the success rate [95, 96].

Surface Plasmon Resonance

SPR is an optical technique that records the angular shift of polarized light reflected from a metal film due to ligand binding to an immobilized target on a sensor. It is a label-free technique able to detect biomolecular interactions in real-time. SPR is a phenomenon that occurs in thin conducting films, usually gold, at an interface between media of different refractive index, high (e.g

sensor surface) and low refractive index (e.g buffer solution). A beam of polarized monochromatic light is shone through a prism, which makes the light be reflected at the conducting surface under total internal reflection conditions. The light incident on the reflecting interface penetrates some distance into the medium of the lower refractive index, called evanescent wave formation. Under a specific angle of incidence, the light excites electron charge density waves, also known as surface plasmon, in the conducting film. The absorption of energy via the evanescent wave field reduced the intensity of the reflected light. Changes at the surface of the sensor cause changes in the refractive index which can be measured as a shift in the resonance angle. The evanescent wave can extend within 100-200 nm of the sensor surface and decays from it. Therefore, if something binds near the surface it is possible to detect a change of the refractive index, which is measured as a change in resonance angle in real-time and can be plotted as a response unit (RU) against time, known as sensorgram (Figure 1.8) [95]. There are a wide variety of sensor surfaces depending on the use and purpose. Usually, a dextran matrix is covalently attached to the conducting surface, to allow immobilization of proteins. But there are also dextran matrices with streptavidin, a purified protein that recognizes biotinylated proteins or sensors with a hydrophobic surface for capturing lipids. Then, the small molecule or ligand can flow through the sensor surface and if binding occurs it can be monitored in the sensorgram (Figure 1.8). SPR allows us to calculate the kinetics (k_{on} and k_{off}) and affinity (K_d) of the binding event. As it is a label-free detection method and easy to automatize, together with low consumption of material makes SPR ideal for HTS. SPR can be applied to most types of proteins although the protein needs to be immobilized, which can be problematic depending on the protein [97].

Isothermal Titration Calorimetry.

ITC compares the temperature differences between a reference and receptor solution to quantify the energy associated with a chemical reaction, in this case, protein-ligand interaction. It is the only technique that allows us to measure the thermodynamics of the reaction, the entropy change (ΔS) and enthalpy change (ΔH). Thermodynamics factors (H, S) can be used as a guide for molecular design since they are related to structural parameters and can improve structure affinity [98]. Thanks to that, ITC has been considered one of the most important techniques to study protein-ligand interaction. The experiment consists of injecting one component of the reaction, usually the small molecule, into a temperature-controlled stirred cell with the other component,

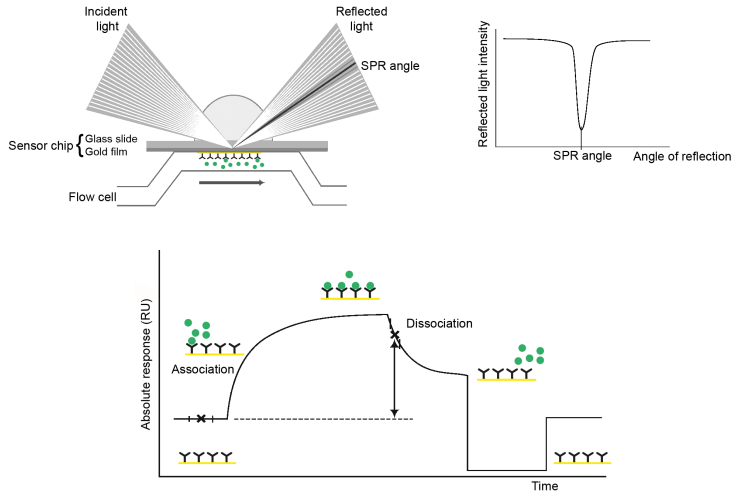


Figure 1.8. Schematic illustration of SPR detection principle and a sensorgram, figure modified from Biacore, Sensor Surface Handbook [97].

the protein. The small molecule is titrated into the cell until full saturation is achieved, allowing measurement of the heat released or used in the binding event and estimation of the affinity and stoichiometry (Figure 1.9) [95]. The main advantages of this technique are the amount and quality of the information provided together without the need of labeling the protein. On the contrary, the big amount of material needed makes this technique a very low throughput method and more suited for optimization of the compound.

Cell-based

Unfortunately, if a compound binds to the target in a biophysical assay it is not possible to guarantee the same behavior inside the cell. That is why alternative methods need to be performed to confirm that the compound interacts with its target in a living system, known as target engagement. Several methods have been developed to study cellular target engagement, which can be classified depending on if they require modification of the small molecule, the target protein, both or even none (label-free approach) [90]. Direct assessment of protein-ligand interaction can be performed by using fluorescence or bioluminescence resonance energy transfer measurements (FRET and BRET).

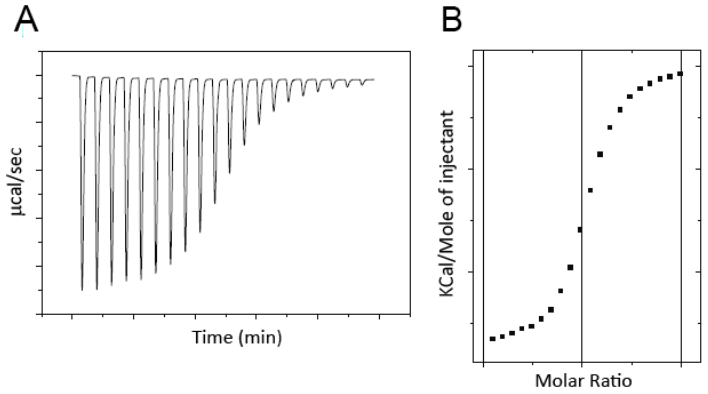


Figure 1.9. Raw ITC data showing an exothermic binding reaction. Figure from [99] with permission of Springer.

It is possible to modify the protein by adding a fluorescent or luminescence tag and the small molecule with a fluorophore. Other approaches can modify only the small molecule, such as affinity-based chemical proteomics for covalent ligands. In this case, the reactive compound is modified so it can get attached to a functional tag (e.g., biotin) that enables the affinity purification of the covalently bound protein for later identification with mass spectrometry analysis. However, modifications of the target protein and the small molecule might alter the results by affecting the protein-ligand interaction. As an alternative, cellular thermal shift assay (CETSA) was developed to perform target engagement without modification of the protein or small molecule. CETSA is based on the same biophysical principle as DSF, binding of a small molecule to a protein can lead to thermal stabilization or destabilization resulting in a shift of its melting temperature. In a CETSA, cells are heat treated followed by precipitation of the protein and quantification by immunoblotting [100]. One disadvantage of this method is the need for suitable antibodies, but it can be overcome using quantitative mass spectroscopy [101].

Computational methods

Over the last decades, CADD has been playing an important role in the drug discovery field by reducing costs and time. It has been applied to different purposes, from virtual screening, hit to lead optimization to even the design

of novel compounds. Virtual screening consists of performing HTS in silico by filtering large compound libraries to obtain a subgroup of predicted active compounds that will later be tested experimentally. The aim is to increase the success rate of finding a hit by reducing the number of compounds tested experimentally. Hit-to-lead optimization can be made by increasing the affinity or optimizing other properties such as ADMET. Design novel compounds can be made through combining fragments or by modifying functional groups to obtain new compounds [102, 103]. CADD can be classified as structure-based (SB) and ligand-based (LB) depending on the type of data available for the target. LB methods use the information about active-inactive compounds while SB is based on the protein structure of interest. To study protein-ligand interaction computationally, we can find several SBDD methods depending on the flexibility and computational effort, where two of the most known are molecular docking and molecular dynamics [102].

Molecular Docking

The objective of protein-ligand docking is to predict the binding mode and affinity of a ligand within the protein. The docking process involves the prediction of the ligand conformation and orientation, known as pose, regarding the protein [104]. That is why, the structural information of the target is crucial, and its quality will also determine the success of the prediction. The identification of molecular features to improve the affinity and therefore the potency of the ligand is not an easy task. To identify the correct binding, molecular docking requires two steps, first the exploration of the conformational space and second the accurate evaluation of the pose by a scoring function [105]. In the exploration of the conformational space or sampling process, structural parameters of the ligands (e.g. torsional, translational, and rotational degrees of freedom) are modified by a search algorithm. The main objective is to explore a wide range of conformational states in a reasonable amount of time. We can find two classes of algorithms, systematic or stochastic. Systematic search algorithms gradually perform variations in the structural parameters until a local or global minimum is reached. The main disadvantages are that it can get stuck in a local minimum and for bigger ligands, it is necessary to limit the search space to avoid combinatorial explosion. This problem has been solved by fragmentation, where the ligand is gradually built in the binding site [103, 105]. The stochastic search introduces random modifications to the values of the degrees of freedom of the system, the structural parameters. This method increases the probability of finding a global minimum by generating ensembles of molecular conformations which can populate a broad range

of the energy landscape. Some of the most known stochastic algorithms are Monte Carlo methods and genetic algorithms [103, 105]. The accurate evaluation of the binding mode is key to successfully selecting the correct poses and rank ligands. Depending on how they are evaluated, we can differentiate several scoring functions: force-field, empirical and knowledge-based. Force-field based scoring functions calculate the binding energy by the sum of intramolecular (internal ligand energy) and intermolecular (receptor-ligand interaction) components. The force-field parameters are usually parametrized by quantum chemical calculations or experimental data. Empirical scoring functions fit parameters to experimental data as a sum of several parametrized functions coefficients are obtained from regression analysis from experimental data. The main disadvantage of this method is its dependence on the data used to perform the fitting. Knowledge-based scoring functions are designed to reproduce experimental structures rather than binding energies. Another strategy consists of the combination of multiple scoring functions leading to the so-called consensus scoring citesalmaso2018bridging. Docking is mainly used to perform virtual screening on large libraries of compounds by ranking them and helping to classify the compounds into possible active or inactive. It also proposes a structural hypothesis of how the ligands bind. The main advantage is to reduce cost, speed up the discovery of new ligands, and help in the design of new drugs. However, the main disadvantage is the lack of flexibility, usually, a static structure of the protein is used, which could neglect important features of the binding event. To avoid this problem, it is recommended to use several structures (ensembles) for very flexible proteins. Other methods, such as molecular dynamics can help to obtain an ensemble of structures or to study the protein-ligand interaction with more flexibility.

Molecular Dynamics

Molecular Dynamics (MD) consists in applying Newton's second law to calculate the position and speed of each atom in the system under study [106].

$$F = m \cdot a$$

Where f is the net force acting on the particle i, m is the mass, and a the acceleration. It is possible to determine the acceleration of the system by knowing the force on each atom. The force can be expressed as the gradient of the potential energy.

$$F = -\frac{dV}{dx}$$

Where V is the potential energy of the system and x is the position of the particle i. Combining these two equations, we can now calculate the acceleration of the system.

$$a = -\frac{dV}{dx} \cdot m$$

The potential energy is approximated by a simple function, known as a force field. The force field, as described before, it is the sum of bonded and nonbonded energy terms:

$$V = \sum_{i}^{bonds} \frac{k_{l,i}}{2}(l_i - l_{0,1})^2 + \sum_{i}^{angles} \frac{k_{\alpha,i}}{2}(\alpha_i - \alpha_{0,1})^2 + \sum_{i}^{torsions} \left\{ \sum_{k}^{M} \frac{V_{ik}}{2}[1 + cos(n_{ik} \cdot \theta_{0,ik})] \right\} + \sum_{i,j}^{pairs} \varepsilon_{ij} \left[\left(\frac{r_{0,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{0,ij}}{r_{ij}} \right)^{6} \right] + \sum_{i,j}^{pairs} \frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon_r r_{ij}}$$

where the first three terms represent bonded interactions (chemical bonds, atomic angles, and torsions or dihedral angles) and the last two nonbonded interactions (van der Waals and electrostatic). Bonds and angles can be described using harmonic potential, and dihedral angles (torsions) are modeled using a sinusoidal function (cosine). The van der Waals interactions are modeled using the Lennard-Jones potential and the electrostatics using Coulomb potential [107, 108]. Although force fields have helped to speed the simulation of large biomolecular systems, due to a large number of atoms, these equations are not an easy task to solve analytically. That is why they must be solved numerically. Numerical methods split the integration of the equation of motion into discrete time intervals, known as time-steps (dt). A small time-step is chosen, usually, 1 or 2 fs, since forces depend upon the atomic position that changes

over time [106]. Several force fields are commonly used in MD simulations for biomolecular simulations (protein alone or protein-ligand), including AMBER [109], CHARMM [110], and GROMOS [111]. However, small molecules usually require their own parametrization due to their large structural variability. Although general force-fields have been developed to help simplify it, as GAFF [112].

MD simulations allow us to obtain a trajectory of the system which contains the conformational properties, and it changes with(over) time. Several properties can be calculated from the MD trajectories, including free energy and kinetics measures. That is why, MD is widely used in several fields, from material science, chemical physics to drug discovery. Thanks to MD, protein-ligand interactions can be analyzed in a more flexible environment and even with explicit solvent. Being able to simulate the binding and unbinding events is key for designing and improving small molecule's affinity and efficacy for drug discovery purposes.

However, MD has some limitations in the force fields and the length of the simulation [108]. First, force fields require optimization and refinement to properly describe the occurring events. As an example, MD simulations cannot be used to study chemical reactivity, since chemical bonds cannot be broken or formed during MD. Second, time-scale limitation due to the computational work needed to calculate real-time events. Standard MD simulations allow the calculation of microseconds in a cost-efficient way depending on the system. Whereas biological processes such as protein-ligand interaction, enzyme catalysis, or conformational changes occur on time scales from milliseconds to minutes [106]. Moreover, biological systems are known to have many local minima frequently separated by high energy barriers that are difficult to sample for conventional MD simulations [113]. To overcome the time and sample limitation by classical MD simulations, alternative methods have been developed. Some examples are the coarse-grained MD simulations [114], umbrella sampling [115], replica-exchange [116], metadynamics [117], accelerated MD [118], steered MD [119] and much more.

1.5 Pseudokinases: new targets for drug discovery

As the name implies, pseudokinase refers to a subfamily of the kinase family where these proteins lack the ability to phosphorylate other substrates. Even though kinases are one of the most studied families of proteins, pseudokinases

have remained much more elusive. Initially, inactive kinases were classified as pseudokinases by sequence analysis based on their lack of at least one of the catalytic residues [65, 120]. However, not all kinases that lack a catalytic residue are inactive. For example, the kinase WNK (with-no-lysine kinase), as its name suggests, lacks the catalytic lysine residue in the VAIK motif and is catalytically active [121]. The lysine in strand β3 at the N lobe is replaced by cysteine. However, a lysine from strand β2 is placed in the active site allowing the interaction with ATP and therefore the kinase activity [122]. Conversely, the BUBR1 pseudokinase, which has the essential catalytic residues but lacks the Glycine-rich loop, is unable to phosphorylate and bind ATP [123]. These examples show the diversity of kinases, and more particularly pseudokinases, and how proteins with very similar structures have very different functions. Although most of the pseudokinases cannot phosphorylate other substrates, this does not automatically imply the inability to bind or catalyze ATP and derivatives. In 2014, James M. Murphy and colleagues published an article showing a methodology to subclassify pseudokinases based on their nucleotide-binding properties [124]. They proposed the thermal shift assay DSF to benchmark the binding and catalytic potential of pseudokinases. Thanks to that, they subclassify pseudokinases in four classes: class 1 - devoid of nucleotide or cation binding, class 2 - nucleotide binding, class 3 - cation binding, and class 4 - nucleotide and cation binding (**Table** A1).

Approximately 10% of a typical vertebrate kinome are pseudokinases, which are very well conserved between species, suggesting conserved signaling roles. Pseudokinases are present among all kingdoms of life, it is possible to find orthologous in other eukaryotic proteomes (e.g., flies, plants) which supports the theory that pseudokinases are part of ancient genetic lineages [125]. Pseudokinases perform several essential functions despite their lack of catalytic activity. Some of the known functions are based on one-to-one cases which cannot always extrapolate between them. Since pseudokinases cannot phosphorylate as a canonical kinase, they act as scaffolds or interact with other proteins as modulators, competitors, and/or spatial anchors [126].

One of the best-known functions is the allosteric regulation of conventional kinases. Some kinases can regulate themselves through allosteric interaction with another kinase or pseudokinase. This dimerization event can regulate the switch between active or inactive kinase by modulating the position of key regulatory elements, such as the αC helix [127]. Pseudokinases have been characterized interacting with kinases in several distinct modes of dimerization:

head-to-tail (HER3 pseudokinase: EGFR kinase; PDB code 4RIW), parallel side to side (MLKL pseudokinase: RIPK3 kinase; PDB code 4M69), transverse side to side (Tyk2-b kinase: Tyk2 JH2 pseudokinase; PDB code 4OLI), and even heterotrimeric complex (Stradα pseudokinase: LKB1 kinase: MO25α calcium-binding protein; PDB code 2WTK). This highlights the plasticity of pseudokinases to interact with the kinase domain. Even some pseudokinases have been reported to regulate the activities of other enzymes, as the pseudokinase of guanylyl cyclase–A and guanylyl cyclase–B regulate the activity of the tandem guanylyl cyclase domains or the VRK3 pseudokinase binding and activating the VHR phosphatase [128, 129].

Other functions of pseudokinases include acting as a signal or molecular switch, as in the phosphorylation of MLKL pseudokinase by RIPK3 kinase which alters the conformation of the MLKL to cause necroptosis [130]. MLKL pseudokinase can function also as a protein interaction domain, where the pseudokinase domain is regulated by binding to the RIPK3 and HSP90:Cdc37 co-chaperones [131]. Pseudokinases can act as signaling scaffolds by binding components of a signaling cascade. As an example, Tribbles (TRIB) pseudokinases nucleate assembly of a complex between C/EBPα and an E3 ubiquitin ligase COP1 [132, 133]. Through this mechanism, they can control subcellular localization and connect components to continue a specific signal. Another example is BUBR1 pseudokinase, which is an essential component of the mitotic checkpoint required for normal mitosis progression [123]. A balance of phospho-signaling is critical for the correct exit from mitosis. The pseudokinase domain of BUBR1 promotes the recruitment of PP2A-B56 phosphatase crucial for this event.

1.5.1 Pseudokinases and disease

Bearing all of this in mind, pseudokinases should not be considered remnants of evolution. Pseudokinases can be fundamental in several signaling processes across species. Although the functions of most of the pseudokinases still remain unclear, dysregulation of pseudokinases is linked to diseases such as cancer, metabolic, neurological, and autoimmune [134]. We can find examples of pseudokinases playing important roles in very different signaling and metabolic pathways. Here, I will describe some of the best known cases for which even some modulators (e.g, small molecule, antibody) have been developed [134, 135].

JAKs

The Janus Kinase (JAK) family of non-receptor tyrosine kinases (JAK 1-3 and tyrosine kinase 2, TYK2) have an N-terminal FERM domain, an SH2-like domain, a pseudokinase domain Janus Homology 2 (JH2), and C-terminal tyrosine kinase domain (JH1). JH1 is activated via type I/II cytokine receptors which regulate inflammation, immunity, and even infections [136]. JH2 lacks the catalytic aspartic in the HRD motif and has degraded the Gly-rich loop (Table A1). Interestingly, all the JH2 domains can bind ATP but the JH2 of JAK3 binds ATP in a cation-independent manner [137]. JH1 is regulated by JH2 upon binding and stabilizing an inactive conformation [138]. Mutations in the JH2 domains of JAK1, JAK2, or JAK3 result in the activation of JH1 and are associated with diseases, including various types of leukemia [139]. As an example, the V617F point mutation in the JH2 domain of JAK2 is detected in more than 95% of patients with polycythemia vera (PV), and in 50% with essential thrombocythemia (ET). Both diseases are a type of myeloproliferative neoplasms, a type of blood cancer [140]. V617F is found in the JH2 domain near the JH1-JH2 interface which facilitates the activation of JH1 [141]. To date, only three JAK inhibitors have been approved by the FDA, fedratinib, ruxolitinib, and tofacitinib, all of them targeting the JH1 kinase domain. As an alternative approach, there is a compound targeting the pseudokinase domain JH2 of TYK2, deucravacitinib (BMS-986165), which is currently in phase III clinical trials for psoriasis [142, 143].

HER

Human epidermal growth factor receptor 3 (HER3) is a member of the EGFR family of receptor tyrosine kinases (RTKs). The EGFR family is formed by four transmembrane receptors (EGFR, HER2, HER3, and HER4), where HER3 lacks intrinsic kinase activity [144]. The HER3 pseudokinase domain lacks the canonical aspartic in the HRD motif but still can bind nucleotides and cations (Table A1). HER3/HER2 heterodimer stimulates allosteric trans-activation and activates the PI3K/AKT signaling cascade through direct HER3 binding to the p85 subunit of PI3K [145]. The PI3K/AKT signaling pathway is one of the most frequent networks activated in human cancer [146]. HER3 overexpression and mutations are associated with several cancers like ovarian, prostate, colon, breast, or melanoma [147]. That is why HER3 is considered a potential drug target due to its role in the pro-survival pathway in cellular proliferation and drug resistance to some therapies [148]. Targeting HER3 with antibodies and the antibody-drug conjugate are the only strategies currently being examined in clinical studies for cancer patients [149]. Alternative

strategies have been developed to selectively target HER3 through bifunctional molecules, where one end binds to the protein of interest while the other end hijacks the cellular mechanism to induce the degradation of the target protein [150]. An ATP competitive ligand of HER3 that forms a covalent bond to cys721 was modified with a hydrophobic tag which induces the degradation by the proteasome [151].

MLKL

The pseudokinase domain of mixed lineage kinase domain-like protein (MLKL) contains a degraded glycine-rich loop and lacks the catalytic residues in the DFG and HRD motif. Despite this, MLKL can bind ATP without cation (Table A1). Like JAK, the MLKL pseudokinase domain regulates the activity of the N-terminal domain, the four-helix bundle (4HB). The pseudokinase domain of MLKL interacts with the 4HB domain via an interface centered around the αC helix in an inhibited conformation [152]. MLKL is a key factor for necroptotic cell death by triggering the tumor necrosis factor (TNF)-α through the phosphorylation of two activation loop residues of MLKL by the kinase RIP3 [130]. The phosphorylation of the pseudokinase domain releases the 4HB domain which causes MLKL oligomerization and localization to the cell membrane and induces permeabilization [153]. The loss of membrane integrity triggers necroptotic cell death, which may have evolved as a line of defense against intracellular infection, recent studies implicate it in a variety of disease states (e.g., myocardial infarction, stroke, atherosclerosis, inflammatory bowel disease) [154]. Targeting necroptosis has emerged as an attractive concept and the development of MLKL inhibitors has been found to affect its function [153]. However, another research found that MLKL selective compounds targeting the ATP binding site have no activity in rescuing cells from necroptosis [155]. However, targeting the cys86 of MLKL with covalent compounds demonstrates to inhibit the oligomerization and translocation of MLKL to the cell membrane [156].

STRADα

STRADα lacks all the catalytic residues that are usually found on kinases, the B3Lysine and the aspartic at the DFG and HRD motif. Moreover, the glycine-rich loop is also degraded (Table A1). Even though STRADα cannot phosphorylate, it binds to ATP without cations [157]. STRADα is an allosteric regulator of the tumor suppressor LKB1 that controls the activity of the AMP-activated protein kinase (AMPK) family. STRADα forms a het-

erotrimeric complex with LKB1 and the auxiliary scaffolding protein MO25α (Calcium-binding protein 39 α). STRADα and MO25α promote the active conformation of LKB1 which depends on ATP binding to STRADα [158]. This heterotrimeric complex translocates LKB1 from the nucleus to the cytoplasm, which may be important for downstream signaling partners [159]. LKB1 malfunction is related to several human cancers and the rare genetic disorder Peutz-Jeghers Syndrome caused by mutations in LKB1 [158]. LKB1 function consists of phosphorylating AMPK which can be activated by metformin, medication for the treatment of type 2 diabetes. All of this makes STRADα an interesting target for diseases such as cancer and diabetes by regulating LKB1 [135]. Recently, STRADα was found to bind kinase inhibitor type II, which has the potential to modulate this pseudokinase in a similar way to conventional kinases [160].

RTK family pseudokinases

Pseudokinases from the RTKs (Receptor Tyrosine Kinases), ROR1, ROR2, protein tyrosine kinase 7 (PTK7), and RYK are all involved in Wnt signaling [161]. Wnt signaling regulates key cellular functions including development and stemness and is also involved in cancer [162]. All of them lack the ability to bind ATP and therefore phosphorylate (Table A1). Despite that, these pseudokinases have been implicated in several cancers and developmental disorders. Higher expression levels of these pseudokinases have arisen their oncogenic role in cell survival and therapy resistance [163]. Knockdown of ROR1 was found to reduce cell proliferation in multiple cancer cell lines and even inhibited lung adenocarcinoma in xenograft models [164, 165]. Knockdown of ROR2 inhibits tumor growth of osteosarcoma cells in vitro and renal cell carcinoma both in vitro and in vivo [166, 167]. ROR2 expression is associated with more aggressive disease states [168]. Inhibition of PTK7 signaling will be beneficial in lung adenocarcinoma and its knockdown reduces cell viability and increases apoptosis in lung cancer cells [169, 170]. RYL is associated with gastric cancer tumorigenesis and liver metastasis. RYL knockdown inhibited migration and suppressed tumorigenesis [171]. Several strategies have been developed to target these pseudokinases. For example, ROR1 approaches include monoclonal antibodies, antibody-drug conjugate, chimeric antigen receptor T-cells (CAR T cells), and even a small molecule [172–175]. Recently it was published a paper where they probe binding of ROR1 to ponatinib, a multi-targeted kinase inhibitor approved by the FDA for chronic myeloid leukemia and Philadelphia chromosome-positive acute lymphoblastic leukemia

[176]. PTK7 has been targeted with an antibody-drug conjugate which has entered clinical trials [177, 178].

Kinase suppressor of Ras 1 and 2 (KSR1/2)

Kinase suppressor of RAS 1 (KSR1) and KSR2 are pseudokinases in the RAF kinase family. Although both pseudokinases have degraded the glycine-rich loop and lack the lysine in the VAIK motif, they can bind ATP and even phosphorylate MEK [179, 180]. KSR1 and KSR2 function as scaffolds coordinating the assembly of Raf-MEK-ERK complexes and promoting activation of the MAPK pathway [180]. KSR can activate RAF through heterodimerization and regulates MEK phosphorylation. KSR functions as a scaffold by regulating RAF to induce a conformational change of MEK [180–182]. KSR are key regulators of the MAPK pathway which play a vital role in cell proliferation and its alteration is implicated in cancer [183]. KSR1 is implicated in Ras-driven cancers whose alterations contribute to almost 30% of all human cancers [184, 185]. Mutations in KSR2 that disrupt the Raf-MEK-ERK pathway have been associated with obesity, and insulin resistance [186]. KSR2 has been targeted with a small molecule that stabilizes the inactive state and antagonizes oncogenic Ras signaling [187]. Moreover, an FDA-approved MEK inhibitor, trametinib, was proved to directly engage KSR at the MEK interface. This highlights KSR as a key player in the mechanism of action for MEK inhibitors. A derivative of trametinib was developed, named trametiglue which targets both KSR-MEK and RAF-MEK and suggests it as an effective strategy to overcome limitations of MEK inhibitors [188].

CASK

The calcium/calmodulin-dependent serine protein kinase (CASK) is a member of the membrane-associated guanylate kinase (MAGUK) family. This pseudokinase has degraded the glycine-rich loop and lacks the aspartic in the DFG motif (Table A1). Despite this, CASK binds ATP and is even able to phosphorylate itself and its substrate in the absence of ions [189]. CASK contributes to neural development and regulation of gene expression neurological diseases [190]. The CASK overexpression has been associated with colorectal and pancreatic cancer patient survival and disease progression with poor prognosis [191, 192]. CASK has been shown to be able to bind several kinase inhibitors, such as staurosporine (broad-spectrum protein kinase inhibitor [193]) and AT7519 (a potent inhibitor of CDKs) [194]. Moreover, a chemical probe has been specifically designed to target CASK to help in the development of

new specific inhibitors and to better understand the biological roles of CASK [195].

ULK4

The ULK (UNC51-like) family of kinases have 5 members in mammals, ULK1, ULK2, ULK3, ULK4, and STK36. They share a conserved N-terminal kinase domain that is homologous to the C.elegans UNC-51 and the yeast Atg1. ULK4 is linked to neurological development and has been linked with schizophrenia, autism, and depression [196]. ULK4 is a neuron-specific protein, 142 KDa consisting of an N-terminal pseudokinase domain, and five predicted C-terminal HEAT repeats involved in scaffolding and interacting with other proteins. ULK4 binds ATP without Mg^{2+} and just conserves the HRD motif and the glycine-rich loop (Table A1). ULK4 has been crystallized with ATP revealing an unusual binding mode where the canonical VAIK motif is compensated by another lysine at αC helix [197]. Lazarus et al found small molecules able to interact with ULK4 with micromolar affinity and identified several scaffolds. Moreover, they could even be co-crystallized in the presence of an ATP-competitive inhibitor.

IRAK3

Interleukin-1 receptor associated kinases (IRAKs) are a family of kinases with an important function in innate immune signaling. IRAK3 pseudokinase is thought to be a negative regulator of innate immune signaling and mutations in IRAK3 are associated with asthma and cancer [198]. IRAK3 does not bind nucleotides or cations but retains the DFG motif and was structurally solved with a dimeric closed "active-state" conformation [198]. In addition, it was shown how IRAK3 has a low affinity for ATP but a high affinity for ATP competitive inhibitors like staurosporine.

1.6 Tribbles family of pseudokinases

This thesis is focused on a specific family of pseudokinases, Tribbles (TRIB). There are three TRIB pseudokinases (TRIB1, TRIB2, and TRIB3) in humans, which are members of the CaMK superfamily of kinases. This family of pseudokinases has an important role in controlling immunity, metabolism, and cancer through protein-protein interactions (PPIs) [199]. TRIB derives from the Drosophila melanogaster (fruit fly) gene of the same name. TRIB gene

was identified in Drosophila mutational screens for genes that control cell proliferation and migration. Interestingly, three different groups independently identified the TRIB gene in Drosophila in the year 2000 [200–202]. Thomas C. Seher and Maria Leptin identified a mutation in the TRIB gene that causes embryonic gastrulation defects. Because of the tradition in science, especially in genetics using Drosophila, researchers who find something new get the privilege of naming it. Historically in Drosophila, gene names are related to the phenotype caused by the mutation of the gene. That is why the mutated gene causing premature division in mesodermal cells resembled the fictional small round organism that proliferates uncontrollably in contact with water from Star Trek, known as tribbles [200]. Jorg Grosshans and Eric Wieschaus identified two genes, frühstart and TRIB, as novel mitotic inhibitors that act during Drosophila gastrulation [201]. Juan Mata et al. showed that TRIB induces degradation of the phosphatase that regulates the entry into mitosis [202].

Although the mammalian orthologs of Drosophila TRIB were identified before, the TRIB name remains as the reference. By using dog thyroid cells in 1997, Wilkin et al. found an uncharacterized protein, C5FW (clone 5 Françoise Wilkin) which levels are induced in the mitogenic pathway. This protein shares 95% with its human ortholog, TRIB2. They also cloned human TRIB1, calling it C8FW [203]. In the year 1999, TRIB3 was identified because its expression was enhanced during the programmed death of neuronal rodent cells deprived of NGF (Nerve Growth Factor) [204]. They named this new gene as NIPK (Neuronal cell death Inducible Putative Kinase). About the same time, two different groups identified the human ortholog of the rat NIPIK. Wu et al. cloned NIPK, which they designated SINK (p65-interacting inhibitor of NF-kB) because it interacts with p65 and inhibits NF-kB dependent transcription [205].

Bowers et al identified the human ortholog of NIPK as SKIP3, which is overexpressed in multiple human tumors and is regulated by hypoxia [206]. Taxonomic analysis of TRIB-related sequences shows that they are almost exclusive to the animal kingdom (metazoans). Moreover, they are absent in non-metazoan eukaryote kinomes (e.g. fungi, plants) [199]. TRIB2 is proposed as the most ancestral among the TRIB family. TRIB2 sequence can be detected in the oldest metazoans, such as sponges and cnidarians (e.g. jellyfish). On the other hand, TRIB1 and TRIB3 appear in more recently evolved metazoans where TRIB2 is also present. An explanation of this could be that TRIB1 and TRIB3 appear after TRIB2 gene duplication during the diversification of vertebrates from invertebrates. Interestingly, orthologues of TRIB3 appear mostly

in mammals which could be driven by the need for new regulatory elements in more evolved organisms [199].

1.6.1 TRIB structure and function

The mammalian TRIB family consists of three genes, TRIB1 (C8FW), TRIB2 (C5FW), and TRIB3 (NIPK, SINK, SKIP3). TRIB family of proteins lack catalytic function, TRIB1 is unable to bind ATP although TRIB2 and TRIB3 have shown a very low ATP affinity and phosphotransferase activity in vitro [207]. TRIB pseudokinases are involved in a wide variety of processes and functions because of the ability to interact and modulate kinases (e.g. AKT/PKB), phosphatases (e.g. CDC25), E3 ligases (e.g. COP1, SIAH1), and transcription factors (e.g. C/EBPα). TRIB pseudokinases have the same overall tertiary structure as the eukaryotic protein kinase domain but lack some of the main motifs needed for its catalytic function, such as the DFG motif and a modified glycine-rich loop. TRIB family have a disordered N-terminal PEST region that is rich in P (proline), E (glutamate), S (serine), and T (threonine), which is associated with target proteins for proteolytic degradation [208]. Following the N-terminal is the pseudokinase domain and a C-terminal tail which contains two conserved motifs. The MEK motif is involved in the binding of MEK components of the MAPK (Mitogen Activated Protein Kinase) cascade and the COP1 binding domain, which is essential for the interaction with the E3 ubiquitin ligase COP1. When comparing TRIB sequence with a canonical kinase, instead of a DFG motif, TRIB is defined by an SLE motif except for TRIB3 with NLE. TRIB pseudokinases conserve the β3 lysine and TRIB1 has a glycine instead of an arginine in the catalytic motif (Table A1 and figure 1.10).

Only two crystal structures of the TRIB1 pseudokinase domain have been solved showing an atypical kinase fold. Both structures reveal the conformational changes that TRIB1 undergoes with and without substrate (Figure 1.11). The first structure solved was the apo form (no substrate-bound) where there is an intramolecular interaction between the COP1 binding motif at the C-terminal tail and the αC helix of the pseudokinase domain. Interestingly, the crystal structure shows a deprecated N-terminal lobe and ATP binding site with the TYR134 pointing towards the ATP pocket, which is consistent with the inability of TRIB1 to bind ATP. TRIB1 contains a bent αC helix very different from conventional kinases. The αC helix is a very dynamic regulatory element in kinases with a strategically important position between the N

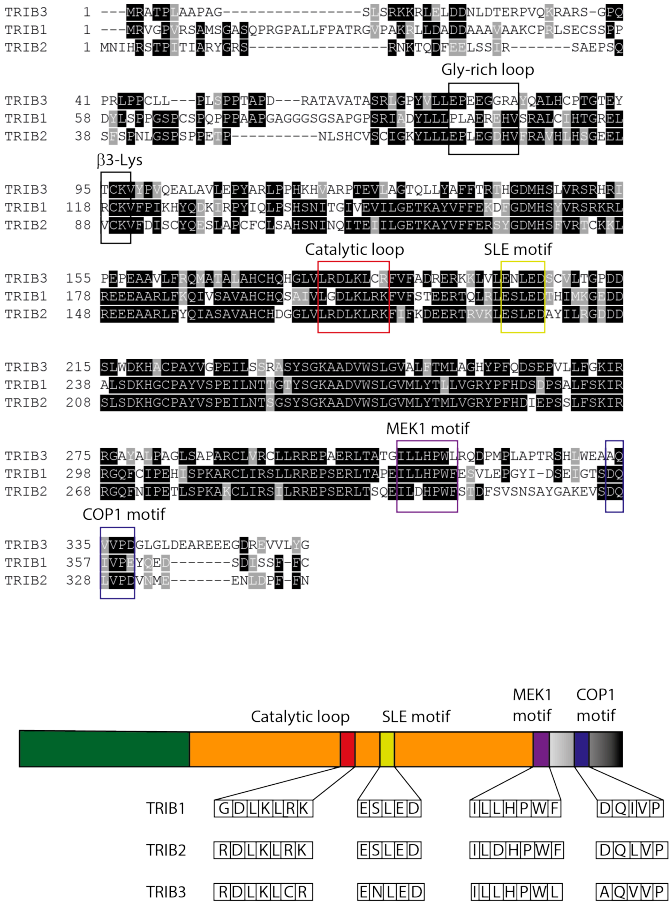


Figure 1.10. Multiple sequence alignment of TRIB1, TRIB2 and TRIB3 with key motifs and regions highlighted in colors. Figure adapted from [209]

and C-lobe [56]. In contrast, the C-lobe is much more conserved compared to conventional kinases.

TRIB1 also contains a degraded glycine-rich loop (linking β1 and β2) lo-

cated above the canonical ATP binding pocket. TRIB glycine-rich loop is three residues shorter than conventional kinases and incorporates a proline near the end of β1. The degraded glycine-rich loop together with the bend αC helix contributes to a wide opening of the ATP-binding pocket. Another feature unique to TRIB is the SLE motif (NLE in TRIB3) instead of the DFG motif in charge of coordinating Mg^{2+} and completing the hydrophobic regulatory spine in canonical kinases. In this case, the SLE motif is stabilized by the conserved β3 lysine forming a hydrogen bond with the backbone of the leucine (SLE; Leu226). Generally, the β3 lysine forms a salt bridge with glutamate from the αC helix. Moreover, an analysis of all human protein kinases reveals that TRIB is unique having an acidic residue preceding the DFG position. All together suggest the inability of TRIB1 to bind ATP by a stable inactive conformation [210].

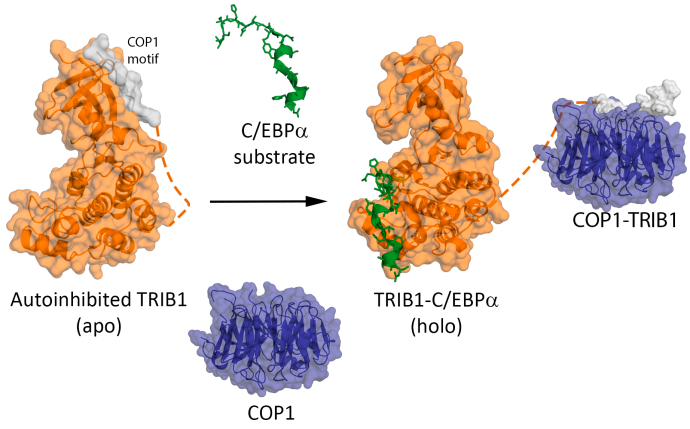


Figure 1.11. TRIB1 conformational change upon C/EBPα binding. TRIB1 is able to bind C/EBPα which triggers conformational changes that release the autoinhibition of TRIB1, favoring the TRIB1-COP1 complex. Then, C/EBPα substrate is ubiquitinated and degraded by the proteasome-dependent pathway.

TRIB pseudokinases regulate C/EBP (CCAAT-enhancer binding proteins) transcription factors by recruiting them to the E3 ubiquitin ligase COP1 for ubiquitination. Thanks to the solved crystal structure of TRIB1 with C/EBPα (holo form) we can understand how C/EBPα binding triggers an allosteric modulation that allows the release of the C-terminal tail. Upon binding to

C/EBPα, the TRIB1 activation loop adopts a more ordered conformation between the N and C-terminal lobes in contrast with the autoinhibitory form (apo state). The leucine from the SLE motif moves toward the αC helix, like the DFG-in conformation in conventional kinases. There is a rotation of the αC helix allowing a rearrangement of the leucine from the SLE motif and the TYR134 (Figure 1.12). The conformational changes in the SLE motif resemble the DFG-in (SLE-in) and DFG-out (SLE-out) in normal kinases. In the SLE-out state (apo or autoinhibited), the activation loop blocks the opening of the canonical ATP binding pocket. Whereas in the SLE-in state (holo or active state), the activation loop is reordered along with the residues nearby the SLE motif thus opening the ATP binding pocket. Such conformational changes allow TRIB to interact with C/EBP on one side and release the COP1 binding motif from the αC helix [211]. As stated before, the TRIB C-terminal tail contains two binding motifs, the HPW[F/L] (MEK1 motif) near the N-terminal of the C-tail and the DQXVP[D/E] (COP1 motif) at the C-terminal of the C-tail. The interaction between the C-tail of TRIB1 and the WD40 domain of COP1 has been characterized through X-ray crystallization [212]. The structure reveals how COP1 recognizes TRIB1 using the top face of its WD40 β propeller domain with nanomolar affinity. The valine and proline residues of the COP1 binding motif (DQIVPE) are the most important residues for the interaction (Figure 1.12). The HPW[F/L] motif is involved in binding MAPK-kinases (MAPKK) and modulation of the MPAK/ERK signal transduction pathway [213]. Kiss-Toth et al. showed that phosphorylation of ERK is enhanced by TRIB1 overexpression. TRIB1 mutants lacking the HPW[F/L] motif or having the tryptophan mutant lost the binding activity [214].

Although there is no public structural information for the other domains or any domains of TRIB2 and TRIB3, there is a high degree of similarity between the pseudokinase domain of this family (TRIB1-TRIB2: 71%; TRIB1-TRIB3: 54%; TRIB2-TRIB3: 54%). Another pseudokinase identified as a distant counterpart of the TRIB family is STK40 (TRIB1-STK40:39%). The structure of STK40 along with ATP-binding assay showed the inability of this pseudokinase to bind ATP [215]. Similar to TRIB, STK40 regulates C/EBP proteins and activates the MAPK pathway [215, 215]. TRIB proteins are mainly located in the nucleus, TRIB1 and TRIB3 are expressed in the nucleus while TRIB2 is mostly expressed in the cytoplasm [216].

1.6.2 TRIB biology and disease

TRIB pseudokinases can regulate protein substrates by modulating the ubiquitin-proteasome system as has been demonstrated for TRIB1 mediated degradation of C/EBPα by COP1 or TRIB3 mediated inhibition of PML-RARα ubiquitination [217].

Another important function of the TRIB family is the modulation of AKT/PKB signaling. Although the molecular mechanisms need to be fully understood since different TRIB family members have different roles. TRIB2 has been shown to activate AKT/PKB in cancer cells and promote drug resistance [218], conversely, TRIB3 has been shown to inhibit AKT/PKB phosphorylation in the liver and neuronal cells [218–220]. Active AKT/PKB is a principal target of insulin signaling since it regulates the translocation of glucose transport to the cell membrane and increases glucose uptake. The role of TRIB3 as a negative regulator of AKT/PKB could lead to insulin resistance. The interaction with COP1 E3 ligase can stimulate lipolysis by the degradation of acetyl-coenzyme A carboxylase (ACC) which regulates fatty acid synthesis [221].

TRIB proteins also regulate hematopoiesis, the formation of blood cellular components with different functions. TRIB1 is highly expressed in myeloid lineage, TRIB2 in the lymphoid lineage, and TRIB3 uniformly across all hematopoietic cells [222, 223]. TRIB1 is critical for the differentiation of M2-like macrophages and eosinophils, which are associated with responses to anti-inflammatory reactions and tumor progression [224]. Studies in mouse models have demonstrated that TRIB1 regulates macrophages and TRIB1 deficiency increases plasma cholesterol and triglyceride levels. Myeloid specific knockout of TRIB1 leads to decreased atherosclerosis, one of the primary causes of stroke, myocardial infarction, and cardiac death [225]. C/EBP transcription factors regulate cellular proliferation and differentiation, metabolism, inflammation, and numerous other responses [226].

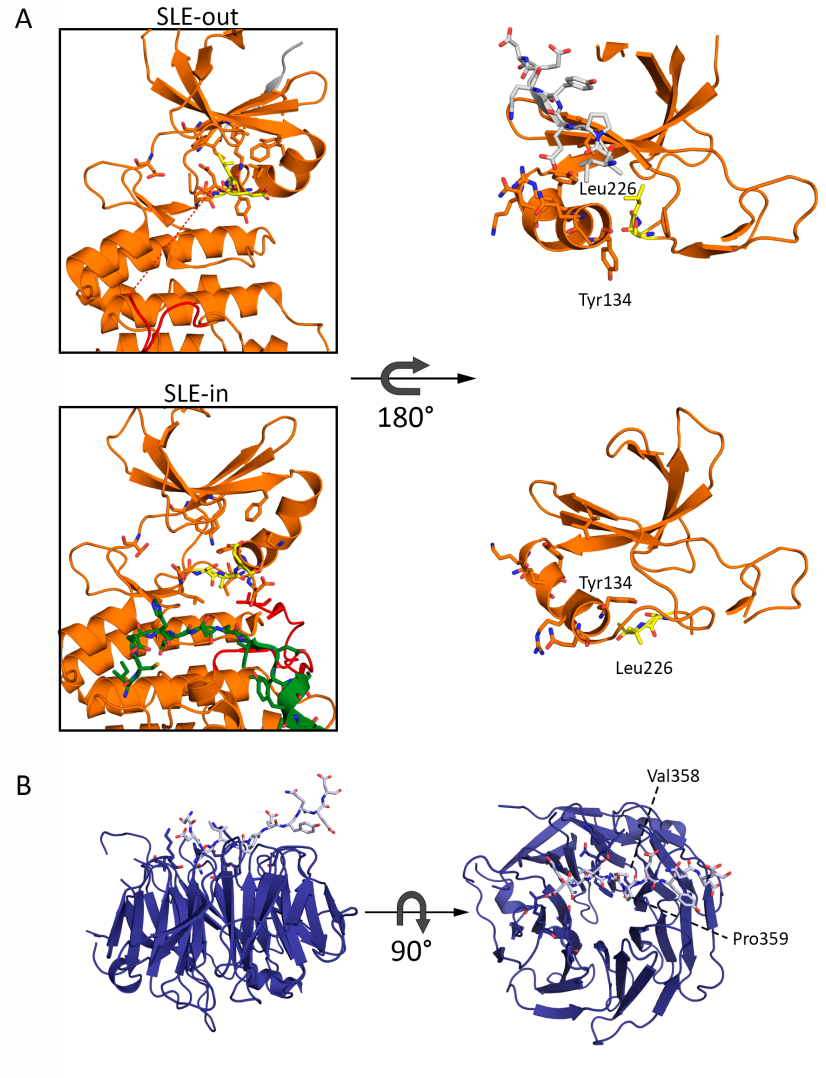


Figure 1.12. TRIB1 structures. A) Comparison of structures of autoinhibited TRIB1 SLE-out(PDB code 5CEM) and TRIB1-C/EBPα SLE-in (PDB code 6DC0) with a simplify view of the N-lobe reorganization (αC-helix, Tyr134 and Leu226). SLE motif in yellow, COP1 binding motif in grey, activation loop in red. B) COP1 binding motif (DQIVEPY) represented as sticks in grey with the COP1 WD40 domain in blue.

Since TRIB can regulate C/EBPs, it supports the relevance of TRIB in cell development and differentiation. TRIB1 and TRIB2 mediate the degradation of C/EBPα, which blocks neutrophil differentiation and increases the differentiation of eosinophils, monocytes, and macrophages [222]. There are two isoforms of C/EBPα, the fully translated protein C/EBPα p42 (42 KDa) and an N-terminal truncated protein C/EBPα p30 (30 KDa). This is because the mRNA encoding C/EBPα contains alternative translation initiation sites [227]. Only the isoform p42 is degraded through TRIB1 and TRIB2.

Depending on the member of the TRIB family and the cellular context, TRIB can function as oncogenic or tumor suppressors. For example, the degradation of the C/EBPα p42 isoform causes an increase in the p30 isoform, a mechanism seen in acute myeloid leukemia (AML) [227]. TRIB1 overexpression is associated with increased cell self-renewal through MEK1/ERK activity, required for AML development [213]. R107L mutation TRIB1 increases ERK phosphorylation and C/EBPα degradation in human Down syndrome-related acute megakaryocytic leukemia [228]. A rare and aggressive subtype of AML, acute promyelocytic leukemia is regulated by TRIB3 which stabilizes the oncoprotein PML-RARα [217]. There is evidence of the tumor suppressor role of TRIB2 in AML via activation of p38 stress signaling from the MAPKK [229]. The identification of TRIB association with AML makes TRIB family a promising therapeutic target in AML.

TRIB are also known to modulate MAPKK which dysregulation has been identified with cancer, obesity, diabetes, and inflammatory diseases [214, 230]. TRIB1 negatively regulates the tumor suppressor activity of p53 by enhancing Histone deacetylase 1 (HDAC1) mediated p53 deacetylation and decreases DNA binding of p53 [231]. TRIB1 is associated with the development and progression of several cancers such as breast, prostate, colorectal, liver, gastric, and glioma [232]. The oncogenic activity of TRIB2 is linked to its overexpression in melanoma and colorectal cancer where TRIB2 is inversely correlated with patient survival [218]. TRIB2 is highly expressed in human T cell acute lymphoblastic leukemia (T-ALL) and contributes to cell survival by regulating the XIAP anti-apoptotic gene in T-ALL cells [233]. In contrast, it was found that TRIB2 knockdown in murine T-ALL did not affect cell growth or survival and acts as a tumor suppressor [234]. TRIB3 promotes lymphoma by suppressing ubiquitination and therefore degradation of MYC [235]. Breast cancer stem cells (BCSCs) are the principal cause of breast cancer metastasis and recurrence after chemotherapy and radiotherapy. Elevated TRIB3 expression supports BCSCs by interacting with AKT/PKB to interfere with the

FOXO1-AKT/PKB interaction and suppress FOXO1 phosphorylation [236]. TRIB3 has been shown to regulate neurological disorders like Parkinson's, characterized by the progressive loss of several neuronal populations. In this disease, TRIB3 expression is elevated and contributes to neuron degeneration and death, in part by diminishing the expression of Parkin protein [237]. The most common cause of dementia is Alzheimer's disease, a progressive neurodegenerative disorder that involves a gradual deterioration. The two indicatives of the disease include plaques of β-amyloid and neurofibrillary tangles of hyperphosphorylated Tau. Experimental observation indicates that TRIB3 promotes neuronal death by both apoptosis and autophagy in response to β-amyloid [238]. Another study identified a link between TRIB3 and the pattern of gray matter loss in Alzheimer's disease [239]. It is not surprising that, given the high number of processes where TRIB proteins are involved, the dysregulation of TRIB can lead to several diseases such as cancer, metabolic syndromes, neurodegenerative and inflammatory disorders [240].

1.6.3 Approaches to study protein-protein interactions

Proteins rarely act alone and least of all proteins without catalytic function, like the pseudokinases. A critical step towards understanding the complex networks in the cell is to correctly elucidate protein-protein interactions. It is important to define properly what protein-protein interaction means since not all interactions have a biological context or are directly related. According to De Las Rivas et al. protein-protein interactions are defined as "specific physical contacts between protein pairs that occur by selective molecular docking in a particular biological context" [241]. To date, experimental and computational methods help to clarify how proteins can interact with each other.

Experimental methods

There are several methods based on laboratory experiments to determine protein-protein interaction. Of course, the interaction can be binary, between protein pairs, or co-complex where the technique measures physical interactions among groups of proteins. There is no standard or gold method, each method has its limitations and advantages. Therefore, it is recommended to use more than one approach to identify protein-protein interactions. Experimental methods can be classified based on the technique and how the interaction is measured.

- **Biochemical**
  - Affinity Chromatography

Co-immunoprecipitation (Co-IP). In this approach, an antibody, specific for the protein of interest (POI) or a tag expressed with the protein, is used to separate the POI from a molecular mixture or a cell lysate to capture the complex. The partners that bind to the POI can then be eluted and identified [242]. This technique detects complexes but not binary interactions. Only thigh interactions are recognized and transient or weak may be missed. Moreover, the mixture during the cell lysis is a potential source of false positives. On the other hand, Co-IP is fast and relatively easy in comparison with other techniques.

Pulldown. A specific affinity chromatography method where the POI is bound to a column, often via an affinity tag expressed as a protein fusion (e.g., GST, HIS tag) or chemically linked to the POI. Pulldown differs from Co-IP in that it is not based on an antigen-antibody interaction. The presence of a tag may influence the results.

- **Biophysical**

Mass spectrometry. An analytical technique based on the quantification of the mass-to-charge ratio of charged particles. Ion mobility mass spectrometry (IM-MS) analysis is performed by first ionizing the protein complex of interest. After ions are injected into a region containing neutral gas at controlled pressure and under the influence of a relatively weak electric field, ions undergo IM separation. Following that, ions are sampled by a mass spectrometer and analyzed according to their mass-to-charge (m/z) ratio [243]. Affinity purification combined with MS (AP-MS) is one of the most used techniques to detect protein-protein interactions. In AP-MS, the POI is immobilized in a matrix as bait. Then a protein mixture is passed through the matrix so finally, mass spectrometry is used to identify the captured proteins. This approach has similar advantages and disadvantages to affinity chromatography but depending on the sensitivity of the MS, this method can detect very small concentrations.

Fluorescent technology. A technique based upon the measurement of the emission of one or more photons by a molecule activated by the absorption of a quantum of electro-magnetic radiation. Examples include fluorescent polarization (FP), where a fluorescent-labeled molecule is excited with polarized

light and is based on the measurement of molecule rotation. We can detect once a fluorescent-labeled molecule (e.g. protein or small molecule) binds to a target protein [244]. Another example is FRET (Förster Resonance Energy Transfer) a technique able to inform about interactions between proteins. FRET relies on the transfer of energy from a fluorophore-donor excited at a specific wavelength to a fluorophore acceptor. It can take place when the two fluorophores are situated at a distance lower than 10 nm [245].

Biosensor. Surface Plasmon Resonance (SPR), this method measures the formation of the complex by monitoring changes in the resonance angle of light impinging on a gold surface as a result of changes in the refractive index of the surface. A ligand of interest (small molecule, peptide, protein, sugar, lipid, nucleic acid) is immobilized on a gold surface, and the interacting partner is injected in buffer flow over it. (see more information in section 2.3.1) [246]. X-ray. Technique based on the analyses of a diffraction pattern generated by a crystal. (see more information in section 2.3.1).

- **Protein Complementation Assay (PCA)**
  - Transcriptional complementation.

Two-hybrid. The classical yeast two-hybrid (Y2H) system is a method that uses transcriptional activity as a measure of protein-protein interaction. The technique is based on a transcription factor, which is split into two parts, the DNA-binding domain (BD) and the DNA-activation domain (AD). The BD and AD domains are fused to two POIs. If the two POIs fused with the domains bind to each other when expressed in yeast cells, the transcription machinery becomes activated, and a reported gene is turned on. Three hybrid systems can also be used, where a third participant is shown to be necessary for the binding of the BD and AD [247]. This technique offers several advantages, Y2H is inexpensive, fast, and provides binary interaction. Although, some interactions would occur between POIs not normally present in the same cellular location, and a yeast protein may act as a bridge for the interaction. An important aspect is a necessity for a clone library.

  - Split luciferase complementation.

This PCA approach is based on the formation of a complex when two non-active fragments of a luciferase (reporter protein) are brought together when

fused to two POIs. An example of this technique is NanoBit, a two-subunit system based on a luciferase, a large Bit (17 KDa), and a Small Bit (11 amino acids) [248]. The POIs are expressed along with the subunits, the interaction brings the subunit into proximity to form a functional enzyme that generates a luminescent signal. Such a small tag would minimize steric conflicts on fusion partners. NanoBit is optimized for high conformational stability and low intrinsic affinity (Kd = 190 M) since the binding affinity of the luciferase fragments could bias the behavior of their fusion proteins. The interaction of the fragments is reversible so the system can be used to detect transient interactions. Advantages over other protein fragment systems include greater sensitivity, fusion to a small peptide, real-time measurements in their proper subcellular localization [249]. PCAs can be used to study pharmacological modulation of protein interactions. This technique allows testing compounds by measuring the spatial and temporal changes in protein complexes in response to modulators (e.g., small molecules, peptides). Moreover, PCAs allow high-throughput experiments and the ability to work with cell lines.

Computational

Determination of protein complexes using computational methods can be divided based on the types of data used as a reference. Some computational methods are based on the sequence (coevolution, motifs) where multiple alignments of orthologous sequences in the same species and phylogenetic trees are built [250]. Other methods are based on the 3D structure of proteins, generally known as protein-protein docking. If we know the structure of the potentially interacting proteins is possible to predict the potential interaction interface. However, these methods are usually very challenging since proteins undergo conformational changes, posttranslational modifications that could not be included in a model, and other factors. That is why Template-Based modeling (TBM) offers a promising alternative to protein-protein docking if 3D structures of complexes formed by similar proteins are available. Under this framework, several strategies have been used to model protein complexes like global superimposition, dimeric threading, direct homology modeling of the complex, and interface structure alignments [251]. Alternatively, machine learning methods have been developed based on databases of experimentally proven interactions and other additional biological properties. Even using only sequence information to generate accurate protein-protein complexes [23, 252].

2 Objectives

TRIB pseudokinases have been described as scaffolds for diverse proteins in many different signaling pathways. However, there is still much to learn about how TRIB function. Consequently, the aim of this thesis is to elucidate TRIB mechanisms by applying computational and experimental techniques.

More specifically, the objectives are as follows:

- To discover small molecules that could interact with TRIB1 and potentially modulate its function. The discovery of a chemical probe for this unique family of pseudoenzymes will help to understand better their activity.
- To identify the domain and key residues of TRIB3 that modulate AKT/PKB phosphorylation.
- To identify the region of TRIB3 interacting with SIAH1 to unravel how TRIB3 could be ubiquitinated.

3 Materials and Methods

3.1 Computational techniques

3.1.1 Computational modeling of TRIB1 with small molecule

A model of TRIB1 was constructed with the crystal structure of Human STK10 bound to GW683134A (PDB id 6EIM) and the apo structure of TRIB1 (PDB id 5CEM) using the automodel function in MODELLER 9.18 [253, 254]. The activation loop of TRIB1 (residues 224-259) was modeled by first removing the residues E224-D228, and A238-T259. Water molecules surrounding the ligand from STK10 were kept during model building. We removed the residues in the activation loop (residues 224-259). A total of 10 models were generated and the model with the best Discrete Optimized Protein Energy (DOPE) score was selected for further analysis.

1 microsecond MD simulation was performed with GROMACS2018 [255]. The Amber ff14SB and the General Amber Force Field (GAFF) set of parameters were employed for modeling receptors and ligands, respectively [256, 112]. The MD simulation was carried out in explicit solvent using the Single Point Charge (SPC) water model, and ions were added to an approximate final concentration of 0.15 M NaCl with the imposition of periodic boundary conditions via a dodecahedral box. Electrostatic interactions were calculated by the particle-mesh Ewald (PME) method using constant pressure and temperature conditions. The temperature was kept constant at 300 K using a Berendsen thermostat with a 0.1 picosecond (ps) coupling constant, and the pressure at 1.0 bar using the Berendsen barostat with a 0.5 ps time coupling constant. Van der Waals and short-range Coulomb interactions were truncated at 9Å. The complex was solvated with a minimum distance of 1.2 nm from the surface of the complex to the edge of the box. The time step employed was 2 femtoseconds (fs) with all bond length constrained with LINCS algorithm [257]. The productive simulations were carried out at an isothermal-isobaric ensemble (NPT) for at least 1 μs.

3.1.2 Virtual Ligand Screening

A subset library of 71651 compounds from MOLPORT "in stock" containing diaryl urea as a substructure (SMILE: O=C([NH]c1ccccc1)[NH]c2ccccc2) were prepared for docking with LigPrep (Schrödinger) [258]. The substructure was selected based on the compound GW683134A. The library was prepared from 2D to 3D so at most eight stereosimores, six tautomers and eight ring conformers would be generated and with a probable ionization sates within the pH range of six to eight. Then, the prepared ligands were docked into the pseudo-canonical pocket of TRIB1. The cavity was defined using the reference ligand method from rDock which define a docking volume around the reference GW683134A ligand. The structure of TRIB1 was selected by clustering the ligand from the MD trajectory using GROMACS with the linkage method and a cutoff of 0.1 nm. Only one cluster was found, and the centroid was selected for the virtual screening. Several pharmacophores were applied as restraints for virtual screening. Among the main interactions to be conserved between GW683134A and TRIB1, we selected the hydrogen bonds involving the side chain of the SER225, and the main chain amine group of the ASP163 and carbonyl group of LEU223. The rDock software [259] was used with the standard protocol along with the pharmacophoric restraints to filter the molecules. The scoring function of rDock is a weighted sum of intermolecular (Sinter), ligand intramolecular (Sintra), site intramolecular (Ssite), and external restraint terms, if provided (Srestraint). When comparing different ligands, we used Sinter because the intramolecular scores lack an absolute scale since they have not been referenced to the lowest-energy conformation of each ligand. Next, top score molecules from the pharmacophore models were grouped and sorted by their Sinter to finally by visual inspection select 41 molecules for further proceedings.

3.1.3 Mutatex

MutateX is an automated pipeline engine that handles input preparation, performs parallel runs with FoldX (http://foldxsuite.crg.eu) and outputs publication-ready figures [260]. Mutatex allowed us, by using the FoldX method, to predict changes in free energy ($\Delta\Delta$G) of folding upon mutation as the difference between the estimated free energy of folding of the mutant and the reference wild type variant. FoldX uses an empirical free energy function, which includes terms for Van der Waals interactions, solvation free energies, water bridges, hydrogen bonding, electrostatics and entropy changes upon folding for main-chain and side-chains. We used as input structure the same as for

the virtual screening, the centroid from the biggest cluster in the MD trajectory. Mutatex calculated the stability of mutating a residue to any of the 20 amino acids for TRIB1.

3.1.4 Computational modeling of TRIB3 C-tail in complex with AKT

The interaction between TRIB3 C-tail and AKT Kinase Domain (KD) was built with Rosetta (release 2020.03) using the FlexPepDock ab-initio protocol [261]. Protein-peptide docking was carried out between the AKT KD (residues 144-450) and selected regions of the C-terminal tail of TRIB3 in ranges of 15 residues. AKT KD template PDB id 4GV1 was selected based on its high resolution. The C-terminal tail of AKT (residues 457-477) was removed to allow docking on the groove over the αC helix. The FlexPepDock ab-initio protocol was first evaluated over a linear C-terminal tail of AKT (res 463-477) to check the accuracy of predicting its binding mode. A preliminary step of the FlexPepDock ab-initio protocol is the generation of fragment libraries for the peptide sequence. The fragment libraries of three and nine residues were generated with the Robetta server [262] using the peptide sequence of AKT and TRIB3 C-tail as reference for each case. We selected the last 28 residues from the C-tail of TRIB3 and split them in three fragments (residues 331-345, 338-352 and 344-358). The peptides were initially positioned approximately 10 Å away from the proposed binding pocket, the αC helix of AKT. Then, the input structures were prepacked to remove internal clashes in the protein and the linear peptide. We generated 100,000 models from each starting structure. The docking simulations were evaluated with the beta_nov16 score function. The top scoring 500 models were clustered using the Rosetta cluster application with a cluster radius cutoff of 3 Å peptide backbone atom RMSD. The models were ranked according to the reweighted score, in which interface residues are given double weight, and peptide residues are given triple weight. We considered the best model as the lowest structure with the biggest cluster in the top 10 lowest models.

3.1.5 Computational modeling of TRIB3$^{328-340}$ in complex with SIAH1

The interaction proposed between SIAH1 and TRIB3 was constructed using a direct homology modeling method, template-based modeling (TBM). Several

models of the SIAH1- TRIB3328-340 were constructed using different templates. In all the cases the program used to create the homology models was MODELLER 9.18 [253, 254]. Three complexes were built using several templates: Complex 1 (PDB ids 4I7D with Zn, 4X3G and 2A25 with H20) where only one template (4I7D) was used for TRIB3; complex 2 (PDB ids 4I7D with Zn, 4X3G and 2A25 with H20) using several templates for SIAH and TRIB3; complex 3 (PDB id 5WZZ with Zn and H20) with only one template for both. In each of the cases, the best model was selected based on the DOPE score given by MODELLER.

3.2 Experimental techniques

3.2.1 Protein Constructs

The proteins used in this thesis were cloned into a modified pET-28 plasmid including an N-terminal hexa-Histidine tag, a ligation independent cloning (LIC) site and a Human Rhinovirus 3C (3C) protease cleavage site between the Histidine tag and the protein of interest. The plasmid contains antibiotic resistance cassette to kanamycin for the selection of successfully transformed bacteria. Protein constructs with the pET-28 plasmid:

- TRIB1 pseudokinase domain with tail (84-372)
- TRIB1 (84-372)-D163I
- TRIB1 (84-372)-S225F
- TRIB1 PKD (84-343)
- SIAH1 SBD (residues 91-282)

The proteins expressed and purified on this thesis were provided by Peter Mace's laboratory (Otago University) and TRIB1 mutants (D163I and S225F) were provided by Miguel Hernandez Quiles (UMC Utrecht).

3.2.2 Generation of competent E.coli cells

Single colonies of BL21 E. coli strain were inoculated onto 15 mL of Lysogeny broth (LB) media in a 50 mL sterile falcon. They were placed in a shaker

incubator at 37 ºC for 4-6 hours until the OD600 (Optical Density measured at a wavelength of 600nm) reached 0.6. The cells were then placed on ice for 10 minutes and subsequently centrifuged at 6000 rpm for 3 minutes at 4 ºC. The supernatant was decanted, and the bacteria was resuspended in 10 mL cold 0.1 M CaCl and incubated for 20 minutes on ice. The bacteria were centrifuged again at 6000 rpm for 3 minutes at 4ºC. The supernatant was discarded, and the bacteria was resuspended in 5 mL cold 0.1 M CaCl and 15% (v/v) glycerol. Then, the competent bacteria were aliquoted at a volume of 200-300 uL into autoclaved 1.5 mL Eppendorf's tubes and stored at -80ºC.

3.2.3 Protein Expression

All recombinant proteins were expressed in Escherichia coli BL21 (DE3) cells. The bacterial constructs were transformed by adding 0.5 µL of plasmid into 20µL of competent BL21. The cells were incubated for 10 minutes on ice and then heat shocked at 42 ºC for 45 seconds. After that, 350 µL of LB was added and incubated in a shaker at 37 ºC for 45 minutes, subsequently plating the cells on the plate bearing appropriate antibiotic for the plasmid in a 37 ºC incubator overnight. The next day, 3 mL of sterile LB was added to each plate in order to resuspend colonies in the LB. The appropriate antibiotic was added to the final volume of culture, 100 μL to 100 mL and 1 mL to 1L. Depending on the final volume of culture, 100 L or 800-1000 L of resuspended cells were added to 100 mL or 1 L, respectively. The cells were incubated with shaking at 37 ºC monitoring the OD600 (Optical Density measured at a wavelength of 600nm) till they reached approximately 0.5-0.8. Then, the cultures were placed at 18 ºC shaker in a cold room where, after 40 minutes for temperature to come down, it was added 20 µL or 200 µL IPTG (1M) for 100 mL or 1 L culture respectively. Further on the cells were left shaking at 18 ºC overnight. The next morning cells were harvested by centrifuging at 4000 rpm for 20 minutes at 4 ºC (Beckman Coulter J6-MC with a JA-8.1 rotor). The pellet was resuspended in 1.2 mL or 10 mL of freeze buffer (50 mM Tris (pH 8.0), 300 mM NaCL) per 100 mL or 1L previous freezing at -20 ºC until use.

3.2.4 Protein Purification

Cells pellets were defrosted at room temperature and resuspended in purification buffer (50 mM Tris (pH 8.0), 300 mM NaCl, 10% (v/v) glycerol and 10% (w/v) sucrose). To lyse the cells, 20 l lysozyme (25 mg/ml) and 5 l of DNAse was added before using the homogenizer Avestin Emulsiflex-C5 at

15.000 psi (pounds per square inch). In some cases, cells were lysed using another method, sonication. In those cases, the cells were sonicated for 100 second intervals on ice (10 seconds on, 30 seconds off) at an amplitude of 35%. Proteins were purified by using the complementary resin of the tag linked. The constructs with a histidine tag were purified by a Ni2+ affinity chromatography (HIS-Select resin, Sigma-Aldrich) pre-equilibrated with a purification buffer. A purification buffer containing 300 mM imidazole was used to elute the proteins. Then proteins were pulled together and digested overnight with 3C protease (remove N-terminal histidine tag) and 2 mM dithiothreitol. Columns were pre equilibrated with buffer (20 mM Tris-HCl (pH 8.0), 200 mM NaCl, 1 mM EDTA). Proteins were concentrated using concentrators of appropriate molecular weight cut off (Amicon Ultra Millipore). Eluted proteins were further purified by size-exclusion chromatography using Superdex 75 16/60 or Superdex 200 26/60 (GE Life Sciences) columns on an ÄKTA systems with correspondent buffer 10 mM Hepes (pH 7.6), 300 mM NaCl, and 0.5 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP). SDS-PAGE gel was used to check samples results. The core peak fractions were pooled and snap-frozen for storage at -80°C

Quantification of protein concentration

Protein concentration was calculated by measuring the absorbance at 280 nm with the Nanodrop ND-1000 spectrophotometer. The absorbance was adjusted with the extinction coefficient of the protein that is calculated from the amino acid sequence of the protein by the ProtParam tool from Expert Protein Analysis System (ExPASy) [263].

In the case of SIAH1 SBD a Bradford assay was also used to quantify the protein concentration. Bovine serum albumin (Thermo Scientific) was employed to produce protein standards and diluted to concentrations ranging between 0 and 500 g/mL and used to generate a standard curve in a quantification assay. 200 L of Coomassie Plus Protein Assay Reagent (Thermo Scientific), was mixed with 5 L of sample. The absorbance at 595 nm was detected in a spectrophotometer in a 96-well plate. The absorbance of BSA standards were measured relative to a blank (BSA and buffer) and a standard curve was plotted to allow estimation of protein concentration of unknown samples.

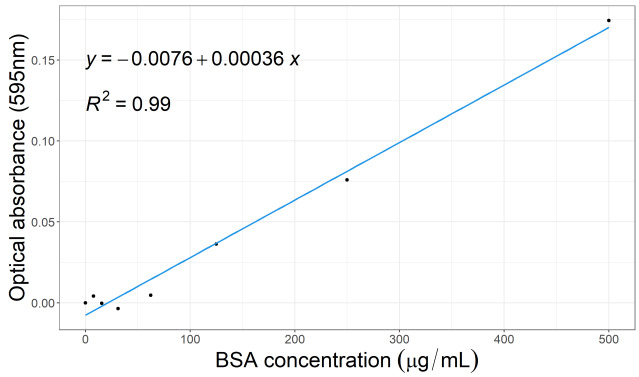


Figure 3.1. Bradford Assay Standard Curve for determination of protein concentration. Increasing concentrations of BSA were incubated with the Coomassie based reagent before reading the optical absorbance at 595 nm in a spectrophotometer. A standard curve was plotted with R.

3.2.5 SDS-PAGE

SDS-PAGE (sodium dodecyl sulphate–polyacrylamide gel electrophoresis) was performed in order to visualize the purity and molecular weight of purified proteins. Proteins were denatured in 5x SDS sample buffer (0.25 M Tris-HCl pH 6.8, 0.25% bromophenol blue, 500 mM DTT, 5% SDS and 50% glycerol) to a final concentration of 1% SDS and heated to 90°C for 2 minutes before being subjected to electrophoresis. To detect proteins predicted to be between 32 and 80 kDa 10% polyacrylamide gels were used. Polyacrylamide gels resolve proteins for 60-80 minutes with an electric field strength of 20 Vcm-1 in the running buffer (25mM Tris-HCl, 190mM glycine and 0.1% SDS). Subsequently, the resolving gel, containing the proteins, was immersed, and incubated in Coomassie stain (0.2% Brilliant blue R-250, 7.5% acetic acid, 50% methanol) for approximately 1 hour with agitation. The stained gel was then washed for 16-24 hours using destain buffer (H20, methanol and acetic acid 50/40/10 v/v/v) until a desired level of background was reached.

3.2.6 Differential Scanning Fluorimetry

The 41 molecules selected through the virtual screening were prepared at 10 mM stock concentration in dimethyl sulfoxide (DMSO). All the compounds

were initially tested at a final concentration of 40 M. TRIB1 (84 to 372) was diluted for use at 5 M with or without C/EBP peptide (40 M) with DSF buffer [10 mM Hepes (pH 7.6), 300 mM NaCl, and 0.5 mM TCEP]. The protein was incubated in a 96-well plate at room temperature for 30 min. SYPRO Orange was then diluted with a DSF buffer for use at 5X as final concentration and was pipetted into each well. For every assay at least 3 reaction mixtures were made for each condition. The 96-well plate was then covered with polymerase chain reaction film and centrifuged for 5 min before being measured on a Roche LightCycler 480 using the default SYPRO Orange protein programme. The data were initially condensed using R version 3.4.3. Data were then analyzed in Microsoft Excel using the CS example DSF Analysis v3.0 template provided by the SGC [264, 91] and Boltzmann fitting in GraphPad Prism 8. Each assay was performed in triplicate unless stated otherwise. For dose-dependent thermal-shift assays a compound range was prepared by serial dilution in DMSO and added directly into the assay to the appropriate final concentration. The thermal profiles were normalized so that the lowest value was assigned as 0 and the largest fluorescence reading as 1.

3.2.7 Surface Plasmon Resonance

Surface Plasmon Resonance (SPR) assay was performed using Biacore T200 SPR biosensor instrument (GE Healthcare, Uppsala, Sweden) at 30 ºC. Prior to the start of a new experiment, the instrument was cleaned using Biacore maintenance kit and a new CM5 chip was inserted and prepared according to the instructions displayed by the software. TRIB1 (84-372), TRIB1 (84-372)-D163I and TRIB1 (84-372)-S225F were immobilized on one CM5 sensor chip one on each channel, using standard amine coupling procedure. The carboxymethyl dextran matrix of the sensor chip was activated with 0.1 M N-hydroxysuccinimide and 0.4 M 1-ethyl-3-(3 (dimethylamino)propyl)carbodiimide hydrochloride at a flow rate of 15 L/min for approximately 7 min. The immobilization was performed with 10 g/mL of TRIB1 constructs in 10 mM sodium acetate (pH 5.0) at a flow rate of 5 L/min. Unreacted activated groups of the dextran matrix were deactivated by injection of 1M ethanolamine hydrochloride for 7 min. The assay was performed using phosphate buffered saline (PBS; 10mM phosphate, pH 7.4, 150mM NaCl) as running buffer. Interaction assays were performed in a running buffer consisting of 1xPBS, 0.05% (v/v) tween 20, and 5% (v/v) DMSO. To collect kinetic binding data, the compounds were prepared at 10mM stock solution in DMSO and diluted with 1.05 PBS 0.05% (v/v) tween20. The compounds were injected at different concentrations with a flow rate of 60 L/min, the association

was monitored for 60s while the dissociation for 160s. We used the following formula to calculate the amount of protein to immobilize in the surface of the chip:

$$R_L = \frac{MW_L}{MW_A} \times \frac{R_{MAX}}{S} \tag{3.1}$$

Formula 3.1. Formula applied to calculate the amount of protein to immobilize on the surface of the chip considering the molecular weight of the ligand (protein) and the analytes (small molecules. RL: response level (RU) of immobilized ligand; MWL: molecular weight of ligand; MWA:molecular weight of analyte; Rmax: maximum binding capacity; S:number of binding sites per ligand.

In general, a total analyte response of maximal 100 RU is used as reference (Rmax). Therefore, using the formula X and an average library weight of 450 Da with one binding site per ligand, we decided to immobilize at least 7400 RU to get a maximum theoretical response of 100 RU. To optimize the experiment, TRIB1 (84-372) was immobilized at high (7400 RU) and low (1800 RU) density values. The high-density option gave the best results, and it was chosen as the default condition to test the other TRIB1 constructs. The CM5 chip has several channels where the channel 1 was used as a reference surface where no protein is immobilized. The Biacore T200 evaluation software 2.0 was used for data analysis. Corrections for minor differences between protein surfaces and reference surface interactions with DMSO were introduced by using a series of solvents standards (solvent correction buffer containing 3%, 3.7%, 4.4%, 5.1%, 5.8%, 6.6%, 7.3%, and 8% DMSO). Finally, signals were corrected for background by subtracting signals from an average of two blank injections from those of compound injections (blank subtraction).

3.2.8 Isothermal Titration Calorimetry

ITC experiments were performed at several temperatures (25-30°C) using an Affinity ITC from TA instruments (Lindon, UT, USA). SIAH1 SBD (91-282), TRIB3 328-340 (Mimotopes) and SIAH1 SBD were prepared using a matched buffer consisting of 20 mM Tris at pH 8.0, 100 mM NaCl. Before performing the ITC experiment, we predicted that the binding affinity of TRIB3 328-340 will would be similar to the AXIN1 375-394 [265]. That is why we decided to

use similar protein and peptide concentrations as the ones used for AXIN1 375-394 and SIAH1 SBD. TRIB3 328-340 (800-1200 M) was injected 20 times with a volume of 2.5 l into 350 l of SIAH1 SBD (30- 50 M). The concentration of SIAH1 SBD was verified by absorbance and the Bradford assay. It is extremely important that the analyte concentration is known precisely since errors in these values will affect the determination of the binding affinity and enthalpy. All the samples were degassed, and the temperature was adjusted prior to filling the cell and injection syringe in a vacuum degassing accessory.

3.2.9 Crystallization screens

We attempted to co-crystallize SIAH1-TRIB3 328-340, and the B1 compound with TRIB1 (84-372) fused to the C/EBP (kindly gifted by the lab of Peter Mace at the University of Otago) or TRIB1 PKD (84-343). High throughput screens were performed in a 96-well plate HRC 2-well plate (Hampton Research) using the commercially available 96 reagent crystallization conditions screens SaltRx HT™, Index HT™ and PEG/ION from Hampton Research. The crystallization screens were performed using the sitting-drop vapor diffusion method with the commercial screens described. The setup was performed with a Mosquito® crystallisation robot (TTP LabTech Ltd.), which plated the purified protein and crystallisation reagent in ratios of either 1:1, 1:2 or 2:1 across the entire screen. Each commercial screen was set up in an identical manner and then the plates were incubated at 18 °C in the Rock Imager® 1000 (Formulatrix).

3.2.10 Protein Complementation Assay

The NanoLuc Binary Technology (NanoBiT™) system from Promega was used to quantify PPIs among TRIBs in live cells. This method consists of 2 small subunits (large LgBiT and small SmBiT) that are fused to TRIB3 and the protein of interest. Once the target proteins interact within the cell, the subunits come together and form an active enzyme that generates a bright luminescent signal. This technology allows real-time kinetic analysis in live cells and detection at low expression levels of the interacting proteins [248]. AKT-C terminal SmBiT and TRIB3-N terminal LgBiT plasmids were constructed at the University of Sheffield. This selection of constructs was obtained after several trials with all the combinations possible (up to 8 possible combinations depending on which part of the N or C-terminal part of the protein the BiT is linked). Constructions and controls were transfected to HEK293T cells in

96-well plates with X-tremeGEN ™ 9 DNA Transfection Reagent from Roche. After 24h the old medium was removed and replaced with OPTI-MEM™ with the Nano-Glo Live Cell Assay System, a nonlytic detection reagent that contains the cell-permeable substrate furimazine. The luminescence was measured using Thermo Scientific™ Varioskan™ LUX. The signal from a known PPI pair is expected to be at least tenfold higher than the corresponding control, if the signal is less this could be produced by a nonspecific interaction between fusion partners.

4 Results
TRIB1 & small molecules

4.1 Background

Tribbles (TRIB) family of pseudokinases have started to gather a lot of attention as we have been able to understand their importance. TRIB are a family of three pseudokinases with an important role controlling immunity, metabolism, and cancer through protein-protein interactions (PPIs) [199]. Unlike kinases where it is possible to regulate their function, pseudokinases are much more difficult to target because of their non catalytic activity. Among the three members of TRIB family (TRIB1, TRIB2, and TRIB3), only TRIB2 and presumably TRIB3 can bind ATP very weakly in the absence of metal ions and even autophosphorylate in an in vitro kinase assay [207]. In contrast, TRIB1 is unable to bind ATP but could potentially bind small molecules [211]. Jamieson and colleagues showed that when TRIB1 binds its natural substrate C/EBPα, there is a conformational change on the canonical ATP binding pocket (Figure 4.1). TRIB1 has been crystallized in two conformations, autoinhibited (apo) or in complex with C/EBPα (holo). The holo structure shows how C/EBPα triggers allosteric changes on the pseudokinase domain allowing the opening of the canonical ATP pocket. There are crucial rearrangements of the activation loop and the αC helix. TRIB1 has a SLE motif instead of the DFG motif in charge of coordinating Mg^{2+} and completing the hydrophobic regulatory spine typical of canonical kinases. In the apo form the SLE motif is stabilized by the conserved 3 lysine forming a hydrogen bond with the backbone of the leucine (SLE; Leu226). This apo conformation shows the canonical ATP pocket completely locked, which is in accordance with TRIB1 inability to bind ATP.

Two different groups of researchers used the Published Kinase Inhibitor Set (PKIS) which is composed of approximately 370 ATP competitive inhibitors as a first screen library against TRIB [266]. Using a thermal shift assay, they could find compounds able to interact with TRIB1 and TRIB2. Moreover, TRIB2 was found to bind covalently to afatinib, a type VI kinase inhibitor of the ErbB family of tyrosine kinases approved by the FDA in 2013 [267]. Unlike kinases

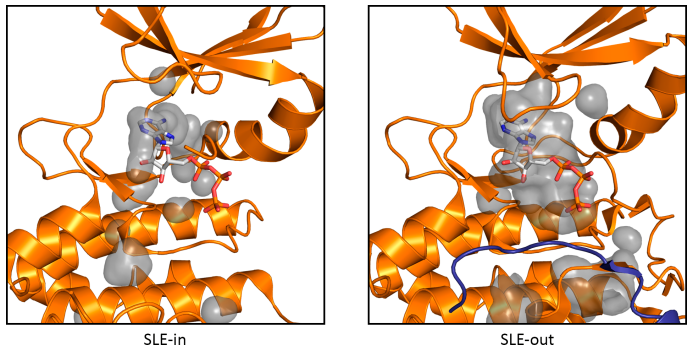


Figure 4.1. Representation of the available binding cavity (grey) of TRIB1 (orange) in the apo (SLE-out; PDBid 5CEM) and the holo (SLE-in; PDBid 6DC0) conformation when bound to C/EBPα (blue). Figured modified from [211].

where it is possible to regulate their function, pseudokinases remain unclear whether it is possible or not to regulate their non catalytic function with small molecules. Kinases are well known to have two main states, active (DFG-in) or inactive (DFG-out), depending on whether they are interacting with ATP or not, respectively. Conformational dynamics of kinases are well known thanks to all the structural information obtained through X-ray or NMR experiments over the last decades. It is clear the importance of the canonical ATP pocket to modulate the conformational changes on the kinase structure [56, 268]. Consequently, we should not be surprised that similar mechanisms could be seen on the pseudokinases. Although there is less structural evidence of the pseudokinase flexibility, still there are some examples where we can see this behavior. For example, it was described the structural and dynamic similarities of the pseudokinase domains from the Wnt-binding receptor tyrosine kinases (PTK7, ROR1, ROR2, and RYK) with respect to canonical kinases [176]. Lemmon and colleagues solved several structures of pseudokinases and found how the activation loop in the inactive conformation mirrors the autoinhibited state of insulin receptor kinase. They found important domain plasticity and different interactions that substitute for the absent canonical kinase motifs. We hypothesized that TRIB1 could bind small molecules in a similar way to canonical kinases. By using rational design, we applied molecular modeling techniques such as molecular dynamics, pharmacophore modeling, and molecular docking to find molecules able to bind and potentially modulate TRIB1. Kinase inhibitors

from the PKIS screening with DSF were used as preliminary hits to discover new chemical probes for TRIB. Several experimental assays were performed to validate the computationally predicted small molecules. DSF was used as the primary screen to detect new small molecules and SPR to verify and validate the interaction. The objectives of the work described in this chapter were to discover new small molecules derived from kinase inhibitors able to interact with TRIB1. Moreover, it was important to elucidate whether the compounds could trigger some conformational changes and modulate TRIB1 activity.

4.2 Discovery of novel TRIB1 small molecules

A specific in silico workflow was applied to develop a computer-based prediction model for TRIB1. First, a homology model was created to dock the preliminary hit GW683134A into the canonical ATP pocket of TRIB1. GW683134A is a benzimidazole-urea based inhibitor, which had been developed to target VEGFR-2 (KDR) and TIE-2 kinase receptors [269]. The next step was to perform a MD simulation of the ligand-protein structure to address the structural flexibility and stability of the interaction. Analysis of the MD trajectory allowed us to select a conformation for the molecular docking to perform the virtual screening. Selected compounds were experimentally validated to interact with TRIB1 using biophysical assays. To further validate the proposed binding mode, we generated single point mutations on the canonical ATP binding pocket of TRIB1. Moreover, to investigate whether our best compound has effects in kinases beyond TRIB1, it was screened against a panel of kinases.

4.2.1 Homology Modeling and Molecular Dynamics

A homology model of TRIB1 with GW683134A, a potential hit from the PKIS screening, was generated based on the unique crystal structure of GW683134A which is bound to the kinase STK10. The kinase domain of STK10 shares a sequence similarity of 24% with the pseudokinase domain of TRIB1. STK10 conserves all the canonical kinase motifs, the DFG, the lysine at the B3 strand and the aspartic acid in the catalytic loop (D in the HxD motif). GW683134A binds in the canonical ATP pocket of STK10 with a DFG-out conformation and the αC helix in, which correspond to an inactive state of the protein. The computationally predicted model was generated as described in the Methods section 3.1 (Figure 4.2). Given the low similarity on sequence and conserved motifs, the proposed binding mode cannot be analyzed without further optimization.

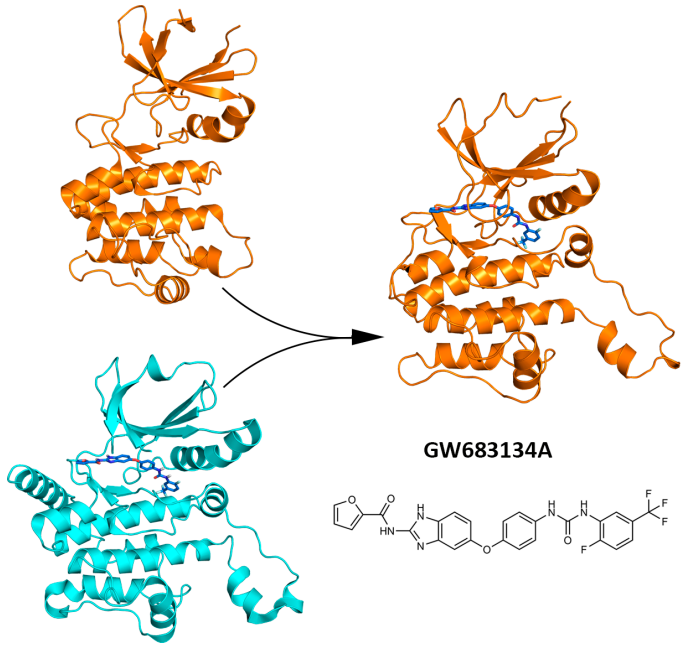


Figure 4.2. Schematic representation of the model of GW683134A bound to TRIB1. The apo structure of TRIB1 (orange; PDBid 5cem) and the GW683134A (blue) bound to STK10 (cyan; PDBid 6eim) were used as templates to create the homology model.

To evaluate the quality of the proposed binding mode a molecular dynamic (MD) simulation was run using the homology model with the best DOPE score. A MD simulation of 1200 ns was performed to effectively assess the dynamic stability of the compound. The root mean square deviation (RMSD) of the atomic positions of TRIB1 and GW683134A shows two different trends (Figure 4.3). Smaller RMSD variations, around 1 A, shows the stability of the ligand binding whereas the protein needs some rearrangements and conformational changes given the higher RMSD values around 6 A. The root mean square fluctuation (RMSF) of TRIB1 was calculated to elucidate which regions are more flexible over the simulation. As seen in figure 4.3, the residues from the activation loop (residues 225-260) showed higher RMSF compared with the rest of the domain, as well as the residues at the G-loop and the C-tail.

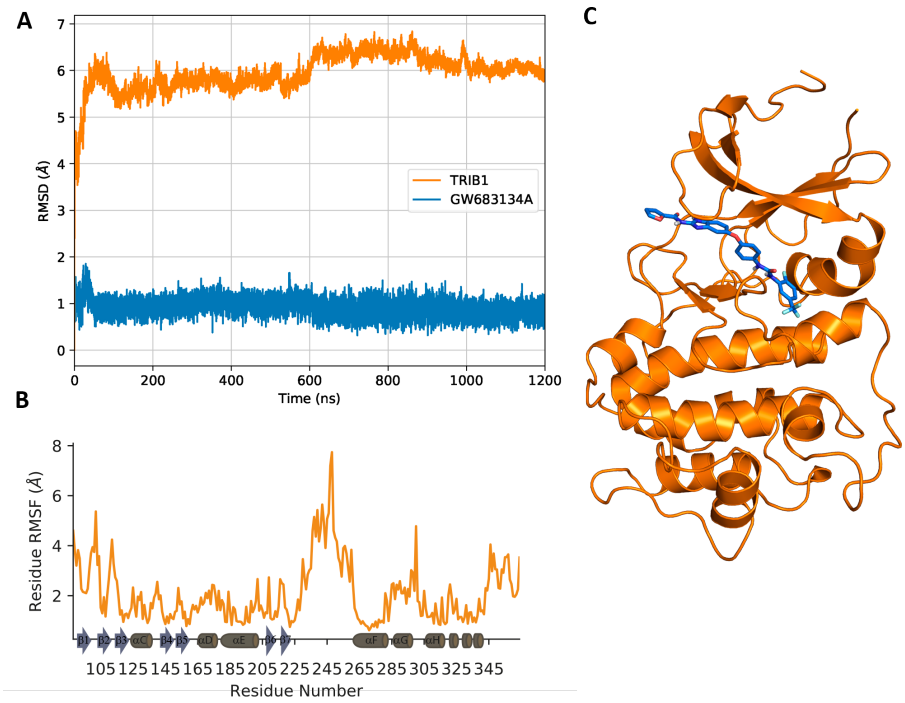


Figure 4.3. Computational modeling of TRIB1 (orange) with GW683134A (blue). A) Root Mean Square Deviation (RMSD) of the GW683134A over the whole simulation (1200 ns). B) Root Mean Square Fluctuation (RMSF) of each residue in TRIB1 over the course of the simulation (1200 ns). The secondary structure of TRIB1 is shown schematically on the x axis. C) Selected frame from the MD of the homology modeling of TRIB1 with GW683134A.

Residues in secondary structures as β-sheets and α helix have lower RMSF values compared with flexible loops. The residues around 5 Åof GW683134A have low fluctuations except for the GLU101, ARG102, and GLU103 in the G-loop with values higher than 3 Å(Figure A.1).

4.2.2 Molecular docking: pharmacophore identification, virtual screening and in silico mutations

After the MD simulation, MD frames were clustered together if their RMSD values were below 1 Å using the linkage method from GROMACS. Given the stability of the ligand over the MD simulation, all of the structures clustered in one group. The cluster centroid was selected as input structure to perform the virtual screening. Based on the analysis of the interactions involved between the ligand and the protein, three specific pharmacophores were created to help the identification of new small molecules. These pharmacophores were considered essential for the ligand-target interaction. Among the main interactions to be conserved between GW683134A and TRIB1, we selected the hydrogen bonds involving the side chain of the SER225, and the main chain amine group of the ASP163 and carbonyl group of LEU223 (Figure 4.4). A subset of 71651 compounds from MOLPORT with a similar scaffold to GW683134A were prepared for docking with LigPrep (Schrodinger) to generate tautomeric and ionization states at pH 7. The prepared ligands were docked into the canonical ATP pocket of TRIB1. The rDock software was used with the standard protocol along with different combinations of the pharmacophore restraints to filter the molecules for which the cutoff was set to 1. The results were filtered using the rDock SCORE.INTER which evaluates the intermolecular interactions between the ligand and the protein with a value lower than -20. After applying those filters, we obtained 16759 compounds. Finally, the compounds were sorted by SCORE.INTER and the best 250 compounds were visually inspected for the final selection. The 41 selected compounds were further validated using DSF as the primary screen.

To guide which residues should be selected for single point mutation, we used a python wrapper called Mutatex which uses the FoldX empirical force field to predict changes in stability and interactions energies [260]. Mutatex allows us to predict $\Delta\Delta$G values in silico mutagenesis for all possible aminoacids in our structure and, more specifically, residues inside the canonical ATP pocket and near the small molecule (Figures A.1 and A.5). We expected that the side chains of D163 and S225, which are located inside the canonical ATP binding pocket, will be directly interacting with the compounds. Therefore, we decided to mutate D163 to isoleucine and S225 to phenylalanine based on the Mutatex's results and the interactions we want to disrupt (Figure 4.5).

The $\Delta\Delta$G values for the proposed mutations are close to 0 kcal/mol in the mutational scan. Values close to 0 kcal/mol indicate that the proposed muta-

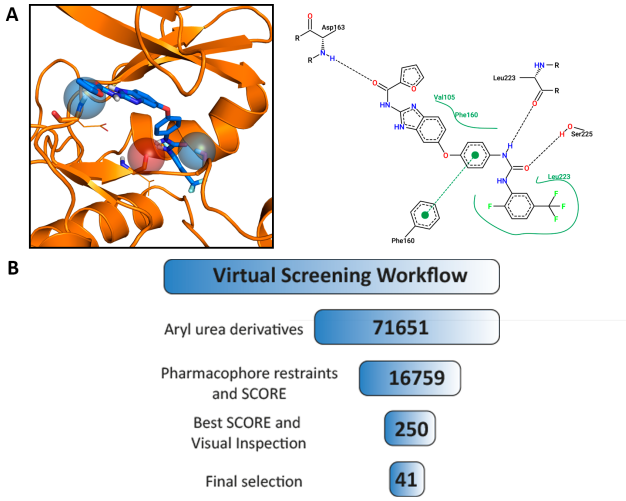


Figure 4.4. Depict of the pharmacophore restraints and the virtual screening workflow. A) GW683134A in the canonical ATP binding pocket of TRIB1 and a 2D interaction diagram by PoseView [270]. B) The small molecule compound database was docked to the canonical ATP binding site on TRIB1. The rDock software was used with three different pharmacophores, two hydrogen bond donors (blue spheres) and one hydrogen bond acceptor (red sphere). The compounds successfully docked were scored by the SCORE.INTER and the compounds with higher scores were selected for visual inspection and 41 small molecule hits were selected.

tion will not change the stability of the whole protein and therefore will only engage the canonical ATP pocket conformation. On the other hand, positive values indicate that the mutant is less stable than the native conformation and negative values the opposite, the mutant is more stable than the wild type.

4.2.3 Protein Expression and Purification

The gene coding for TRIB1 was kindly provided by Peter Mace at the University of Otago (New Zealand) and TRIB1 mutants (D163I and S225F) were provided by Miguel Hernandez Quiles at the UMC Utrecht. All constructs, TRIB1 (84-372), TRIB1ΔCt (84-343), TRIB1 (84-372)-D163I and TRIB1 (84-372)-S225F were expressed in Escherichia coli BL21 (DE3) cells and purified

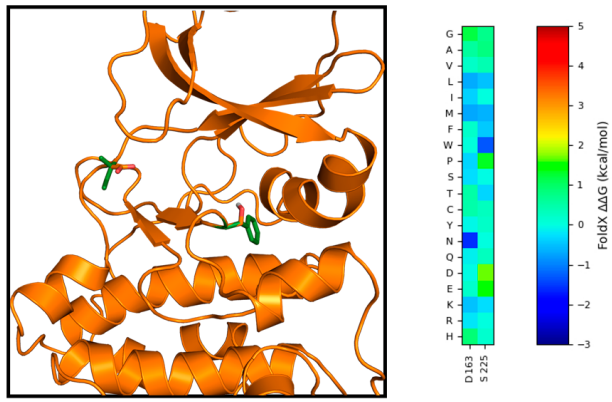


Figure 4.5. TRIB1 structure showing the mutated residues D163I and S225F in green at the canonical ATP binding pocket of TRIB1 along with the ΔΔG values predicted with Mutatex.

using nickel affinity. After elution with 300 mM imidazole, most of the protein eluted in the first fractions. Fractions from nickel affinity chromatography were polled together and digested overnight with 3C protease. Then, the protein was analyzed by size exclusion chromatography (SEC) as described in the method section. TRIB1 (84-372) eluted from the SEC as a single peak and is very symmetric compared to the other constructs (Figure A.4). The molecular weight of TRIB1 (84-372) is 33,115 which is consistent with the weight of the sample by SDS-PAGE (Figure 4.6) and mass spectrometry (figure A.9).

All the proteins used for biophysical assays and structural studies were purified without the His tag. TRIB1 (84-372) was used for DSF and SPR whereas the TRIB1ΔCt (84-343) was only used for crystallization trials which ended without positive results. TRIB1 mutants were only tested for SPR. TRIB1 (84-372) and TRIB1ΔCt (84-343) final yields were approximately 10 mg/L of cell culture. In contrast, the mutant's yield was very low (0.3-0.5 mg/L). SPR needs a very low amount of protein, while DSF and crystallization trials require a considerable amount of protein. Purified proteins were aliquoted and snap frozen in liquid nitrogen for storage at -80 ºC.

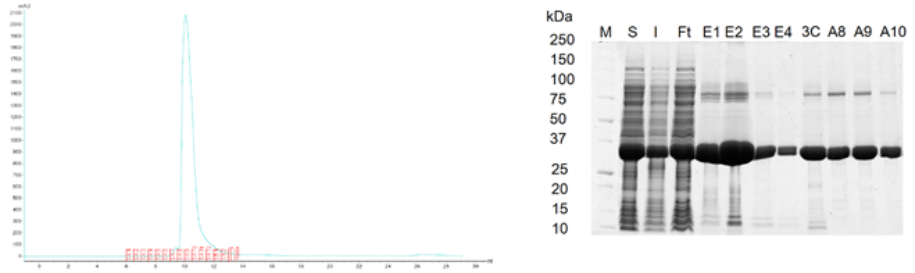


Figure 4.6. Size exclusion chromatography and SDS-PAGE analysis of TRIB1 (84-372). TRIB1 (84-372) from cleared bacterial lysate was bound to a nickel column and eluted with 300 mM imidazole. The peak corresponds to TRIB1 with an absorbance at 280 nm (mAU). Fractions were collected and 5 μl of sample was boiled with 2 X SDS sample buffer and subjected to electrophoresis on a 10% SDS-PAGE polyacrylamide gel. Samples from the bacterial lysate (Soluble), the cell pellet (Insoluble), the lysate collected after the nickel column (flow through), the elution steps (E1, E2, E3, E4), enzyme cleavage (3C) and fractions A8-A10 were analyzed alongside.

4.2.4 Differential Scanning Fluorimetry

TRIB1 has been screened by DSF using the 41 small molecules filtered through the in silico workflow. The small molecule library screening was performed as described in section 3.2.6, following the same procedure as Jamieson and colleagues when they tested the PKIS on TRIB1 [211].

The 41 compounds were purchased through MOLPORT and the C/EBPα peptide was kindly gifted by Peter Mace laboratory. To discover new potential ligands for TRIB1, we screened the 41 compounds against TRIB1 (84-372) in a protein:compound ratio of 1:8. We also tested the effect of C/EBPα degron in combination with the small molecules on the melting temperature of TRIB1 (84-372). In order to reach a preliminary list of stabilizing compounds, a cut-off of ΔTm higher than 1.5 ^{o}C was used to discriminate compounds for further screens (Figure 4.7). None of the compounds gave higher negative values to be considered a destabilizing compound.

Then, a subset of 14 compounds was tested several times with the final cut-off of ΔTm higher than 2^{o}C to define a compound as a hit. The top stabi-

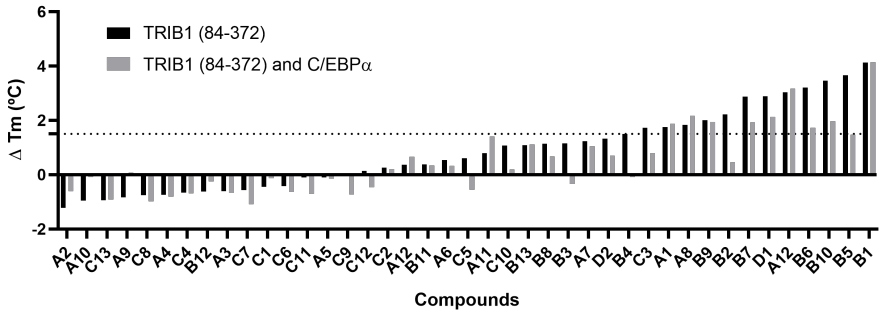


Figure 4.7. Single thermal shift results from the DSF assay on all the 41 small molecules. Tm was calculated as the difference between the compound and the control (only DMSO and protein). The compounds are sorted by ΔTm without the C/EBPα peptide.

lizing compound was B1 with a ΔTm higher than 4 ºC in the two conditions tested. There was no significant preference for either TRIB1 (84-372) alone or in presence of C/EBPα conformations.

To determine that the TRIB1 thermal stabilization was dependent on compound concentration, a dose-response experiment was conducted on the most promising small molecules. The experiments were performed with 2-fold serial dilution and all the melting curves were normalized to be compared. The most promising compounds were tested with and without the C/EBPα. The results are shown in descending order from the most stabilizing compounds to the lowest from the subset of 14 compounds. B1 appears to have the best effect on TRIB1 stabilization at several concentrations with all the melting curves with the same denaturation pattern. At 10 μM B1 shown TRIB1 stabilization (ΔTm 2.21 ºC) whereas D1 (ΔTm 1.65 ºC) or A12 the effect (ΔTm 1.17ºC) at the same concentration is much smaller (Figure 4.10). At higher concentration, the melting curves of D1 and A12 do not conserve the same pattern, being more pronounced at 40 μM with A12. None of the compounds at lower concentrations showed a significant preference for the protein alone or in presence of the peptide

The rest of compounds from the subset of 14 were least effective and have a lower concentration dependent effect. Only compounds B5, B7, B6 and B2

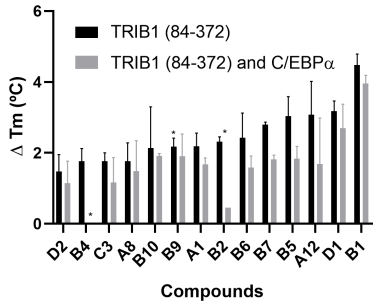


Figure 4.8. Thermal shift results from the DSF assay on the 14 compounds that potentially bind TRIB1 at 40 μM. Tm was calculated as the difference between the compound and the control (only DMSO and protein). Values shown are the mean $\pm$ S.D (n = 3) where the $*$ are those compounds with n $\leq$ 2.

have a small dose-dependent effect but without the same denaturation pattern. The compound B3 was included given the similarity with B6. At 4 μM the compounds B3 and B4 displayed a high shift on the melting curve of the protein, whereas there was a very low or absence of dose-dependent effect with the rest of compounds (A1, B9, B10, A8, C3, and D2).

4.2.5 Surface Plasmon Resonance

TRIB1 (84-372), TRIB1 (84-372)-D163I and TRIB1 (84-372)-S225F were immobilized on the sensor chip's surface using the protocol as described in section 3.2.7. The immobilization levels of TRIB1 (84-372) and TRIB1 (84-372)-D163I were approximately 8300, whereas for TRIB1 (84-372)-S225F was approximately 12000. According to the amount of protein immobilized in the chip and the average molecular weight of the library, approximately 100 RU for TRIB1 (84-372) and TRIB1 (84-372)-D163I and 180 RU for TRIB1 (84-372)-S225F were determined as maximum response for a ligand of 450 Da.

To calculate the affinity for a given interaction we analyzed the level of steady state or equilibrium binding by analysis of the amount of complex against analyte concentration. The affinity can also be derived from kinetics but in our case, it was not possible because of the very fast k_{on} and k_{off}. The C/EBPα peptide was tested as a positive control but produced an unexpected sensorgram of negative values. However, the shape of the sensorgram

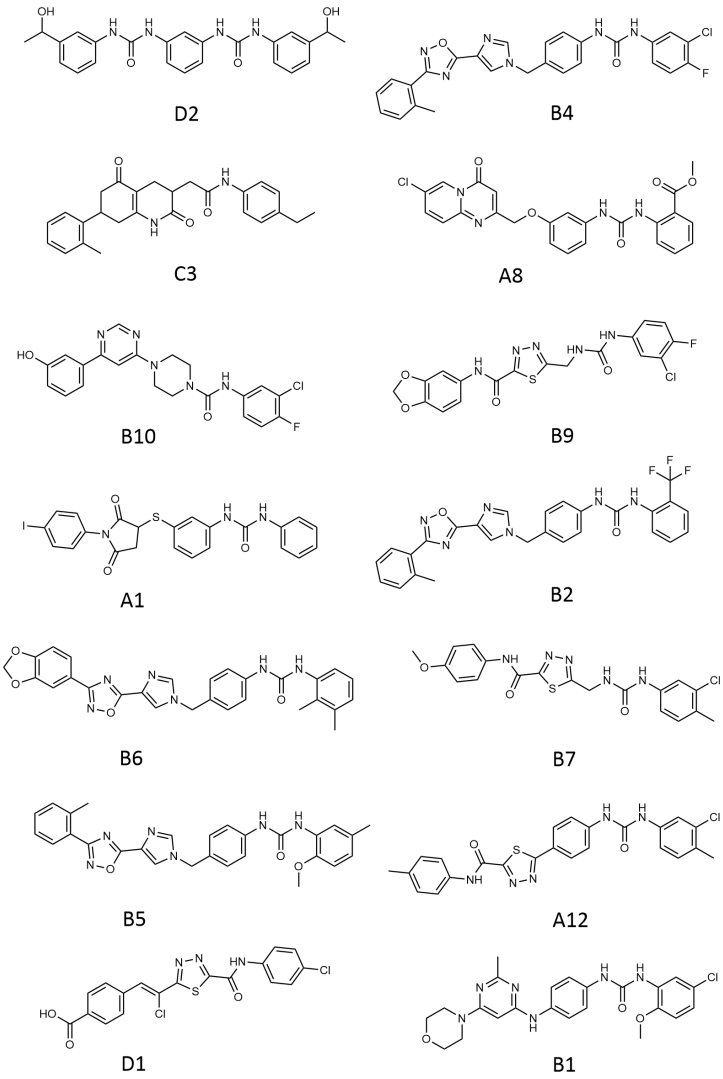


Figure 4.9. Chemical structures of the subset of 14 compounds that potentially bind TRIB1.

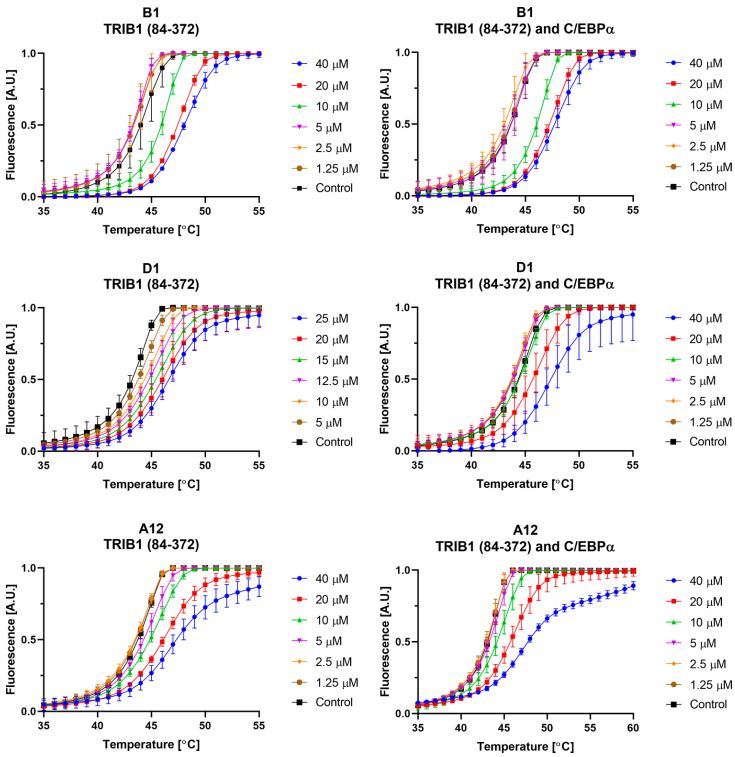


Figure 4.10. Dose response of the DSF from the three best compounds (A12, D1 and B1). Several concentrations were used ranging from 1.25 μM and going up two-fold to 40 μM, except for D1 without C/EBPα. Values shown are from a single experiment with 3 technical replicates.

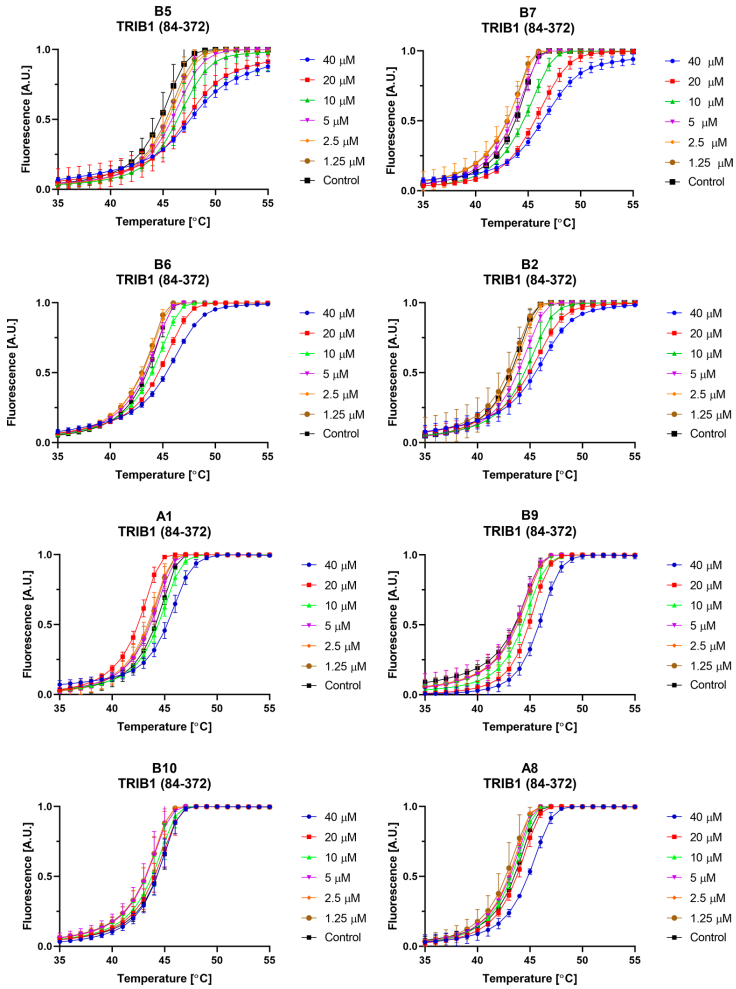


Figure 4.11. Dose response of the DSF from the compounds B5, B7, B6, B2, A1, B9, B10, A8. Several concentrations were used ranging from 1.25 μM and going up two-fold to 40 μM. Values shown are from a single experiment with 3 technical replicates.

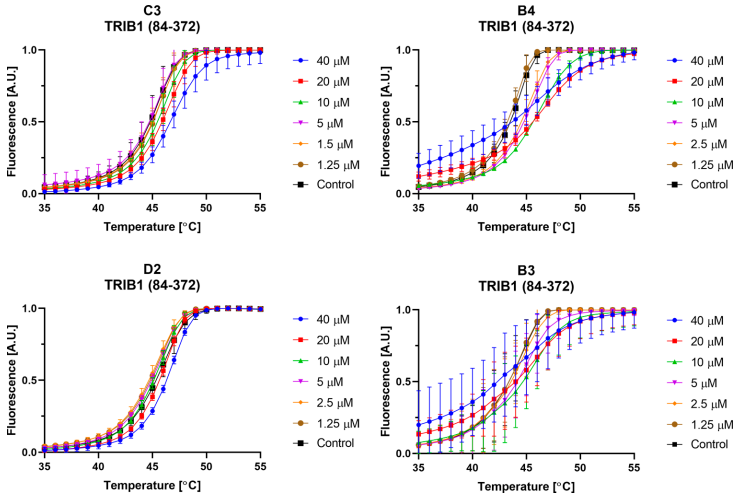


Figure 4.12. Dose response of the DSF from the compounds C3, B4, D2 and B3. Several concentrations were used ranging from 1.25 μM and going up two-fold to 40 μM. Values shown are from a single experiment with 3 technical replicates.

showed a normal behavior with a progressive response during the injection of the peptide and a slow dissociation after the injection. The figure 4.13. shows the sensorgrams based on the dose response curve of the C/EBPα peptide injections at 7 different concentrations ranging from 500 nM and going up to 100 μM. The K_d has been previously calculated with a value of 11 μM as measured by ITC of MBP-fused C/EBPα peptide injected into TRIB1 (84-372) [211]. We obtained a value of approximately 2 μM with small differences between the TRIB1 mutants and wild-type (Table 4.1).

Given the results with the C/EBPα peptide, the compound B1 was used as a positive control to validate and develop an SPR assay to detect interactions between ligands and TRIB1. To find better compounds, we purchased analogs commercially available from the MOLPORT database with a permissive chemical similarity threshold of 80% (Tanimoto coefficient) along with two fragments (F1 and F2) and a new scaffold similar to other hits from PKIS (Figure 4.14 and A.14). In addition, previously tested compounds through dose-response DSF were also tested with SPR (Figures A.11,A.12,A.13), where most of them

have a much lower K_d than B1 (Table A3), as for example the compound A12 has a K_d of 105μM. The compound D1 was confirmed to be a promiscuous binder (higher RU than the expected Rmax) and it was not possible to saturate the protein to determine a steady state K_d. The compounds B4 and B5 did not show a binding event and they were classified as not binders. The compound B3 showed a K_d of 322 μM although the steady state was not stable at higher concentration than 100 μM (Table A3 and figure A.13).

All the compounds from which the K_d was calculated are bound reversibly, consequently dissociating to baseline. From the analogs derived from B1, three compounds (B1.1, B1.2 and B1.5) have a very similar affinity compared to B1 but B1.2 have the better affinity with a K_d of approximately 10 μM (Figure AA.10). The SPR results suggested that removing the methoxy at the ortho-substituent in the right-side phenyl ring (R2) was more favorable. Keeping the chlorine in meta (B1.2) improved potency and was more favorable than trifluoromethyl moiety (B1.5) and the methoxy (B1.3) (Table 4.1). The trifluoromethyl moiety was more favorable than the methoxy increasing the affinity by 3-fold (Table 4.1). Introduction of a methoxy at the ortho and meta-substituent (B1.4) showed a steady state with no equilibrium state (plateau) at high concentrations and a chi square bigger than only methoxy at ortho (B1.1) or meta-substituent (B1.3) (Figure 4.19 and Table 4.1).

The region predicted to be solvent exposed (R1) can also modulate the binding affinity. The morpholine at the front region seems to be less favorable than the piperidine when there is one methoxy at the ortho-substituent in R2, as seen when comparing the sensorgram of B1.1 against B1.11 (Figure 4.15 and 4.26). Introduction of a cyclobutadiene with a secondary amine as linker decreased the affinity by approximately 5-fold (B1.10) compared to B1 (Table 4.1). Bulkier elements at position R2, as the benzo-1,4-dioxane in R2 (B1.9) gave very low response compared to B1 and it was not possible to calculate a reliable binding affinity (Figure 4.24 and Table 4.1). Interestingly, when replacing R2 with a hydrogen but keeping the piperidine at R1 (B1.13) decreased the binding affinity by 10-fold compared to B1 (Figure 4.28 and Table 4.1). In contrast, when keeping the benzene and removing the substituents in R2 (B1.6 and B1.12) resulted in a significant drop in signal (Figure 4.21 and 4.27). Similar results were obtained by introducing a methoxy at the para-substituent in R2 (B1.7 and B1.8).

The fragment F1 did not saturate at 500 μM and due to problems with solubility it was not possible to test higher concentrations. Moreover, there

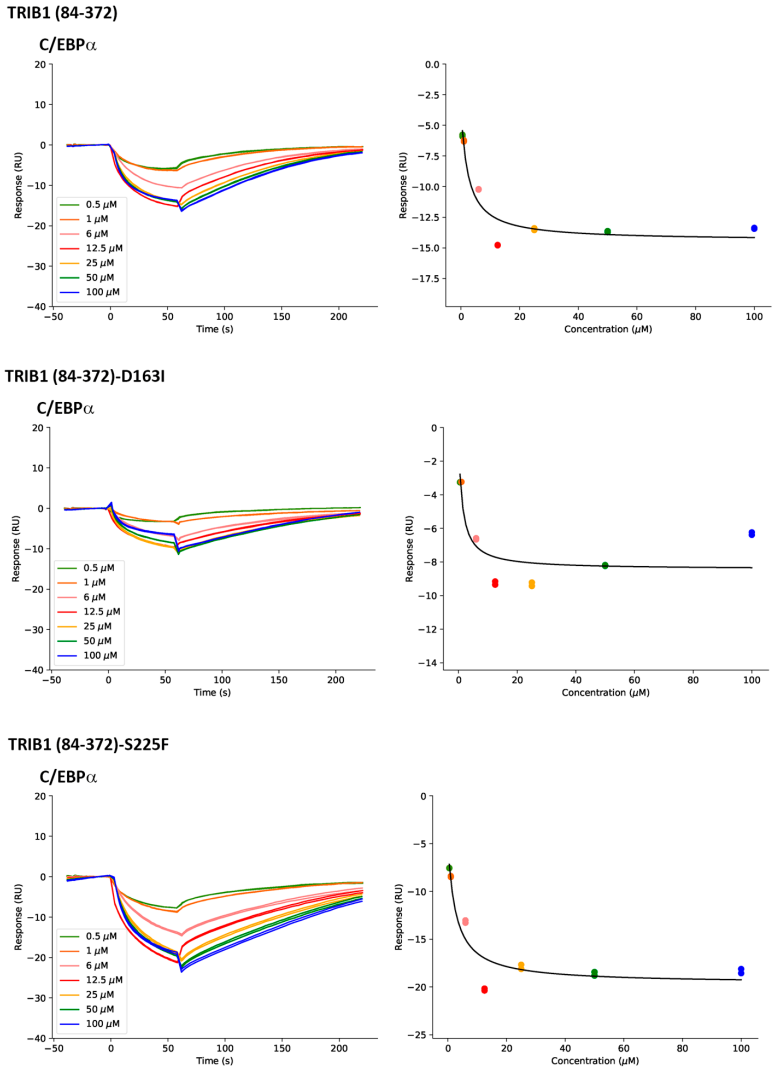
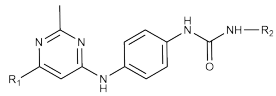
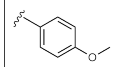
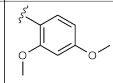
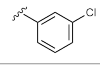
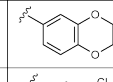
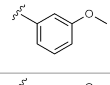
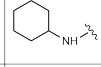
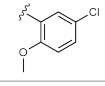
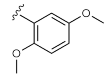
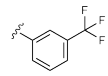


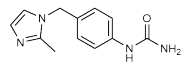
Figure 4.13. Negative sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of C/EBPα against TRIB1 (84-372), TRIB1 (84-372)-D163I and TRIB1 (84-372)-S225F. Values shown are from experimental duplicates.

A

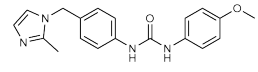


Compound	R_1	R_2	Compound	R_1	R_2
B1			B1.7		
B1.1			B1.8		
B1.2			B1.9		
B1.3			B1.10		
B1.4			B1.11		
B1.5			B1.12		
B1.6			B1.13		

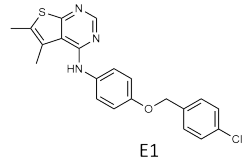
B



F1



F2



E1

Figure 4.14. Analogs series of compound B1 and fragments. A) Chemical structure of the core scaffold of B1 and analogs. B) Chemical structure of the two fragments (F1 and F2) and the new scaffold (E1).

is no dissociation of the fragment at lower concentration as 12.5 μM (Figure 4.30). The fragment F2 showed a K_d of approximately 100uM with a Rmax a bit higher than the expected maximum response (100 RU for TRIB1 and TRIB1-D163I; 180 RU for TRIB1-S225F). Unfortunately, the new scaffold E1 obtained from similarity with GW68462B and GSK-300014 from PKIS did not show an equilibrium state at any of the tested concentrations (Figure 4.30). All the compounds had a fast-on, fast-off binding, which did not permit a reliable kinetics analysis. Moreover, none of the compounds showed a significant difference for any of the TRIB1 constructs, wild-type, or mutants. As expected, the TRIB1 (84-372)-S225F showed higher values of RU given the difference of protein immobilized. The spikes at the start and end of association were removed due to the change in buffer components for all the compounds.

Analyte	TRIB1 (84-372)			TRIB1 (84-372)-D163I			TRIB1 (84-372)-S225F		
	K_d (μM)	Rmax (RU)	χ^2	K_d (μM)	Rmax (RU)	χ^2	K_d (μM)	Rmax (RU)	χ^2
C/EBPα	2.42	-10.8	1.22	1.27	-7.8	1.67	2.74	-14.75	2.85
B1	26.8	24.5	1.84	41.1	20.8	1.56	31.9	38.45	5.46
B1.1	43.8	34.3	0.88	36.7	28.78	0.457	40.3	53.21	1.81
B1.2	9.66	18.47	1.85	9.60	14.66	1.97	9.47	27.9	6.38
B1.3	104	127.5	3.89	110	109.7	3.2	97.4	127.5	3.89
B1.4	12.3	16.3	7.65	9.16	15.58	7.71	11.3	25.7	19.7
B1.5	25.98	25.95	2.7	25.8	23.38	1.65	22.99	42.51	5.88
B1.6	-	-	-	-	-	-	-	-	-
B1.7	-	-	-	-	-	-	-	-	-
B1.8	-	-	-	-	-	-	-	-	-
B1.9	-	-	-	-	-	-	-	-	-
B1.10	174	90.02	0.972	183	86.5	0.561	156.5	136.2	1.99
B1.11	-	-	-	-	-	-	-	-	-
B1.12	-	-	-	-	-	-	-	-	-
B1.13	354.9	112.6	7.09	313.5	73.28	6.97	381.1	181.9	18.2
F1	-	-	-		N.D.			N.D.	
F2	110	140	28.5	146	113.5	10.5	107	219.8	77.9
E1	-	-	-		N.D.			N.D.	

Table 4.1. K_d (dissociation constant), Rmax (maximum response) and χ^2 (Chi square) of C/EBPα, B1 and B1 analogs. N.D.: Not determined.

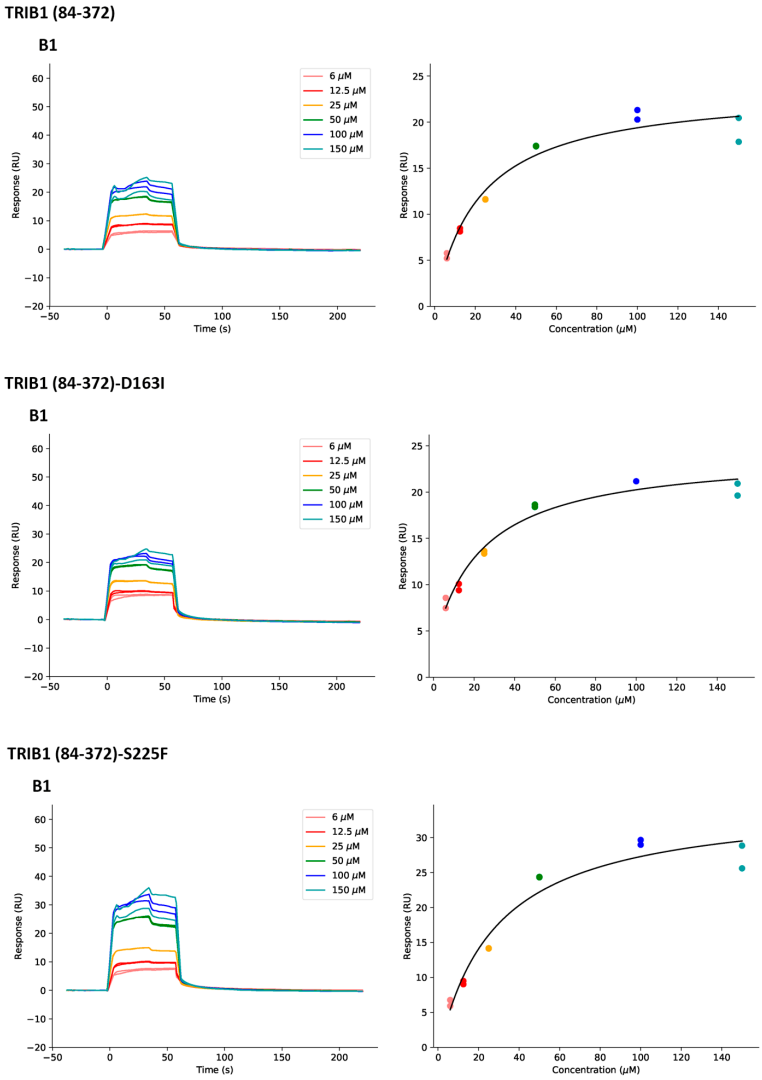


Figure 4.15. Sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of compound B1 against TRIB1 (84-372), TRIB1 (84-372)-D163I and TRIB1 (84-372)-S225F. Values shown are from experimental duplicates and sensorgrams are plotted after removing spikes due to buffer changes.

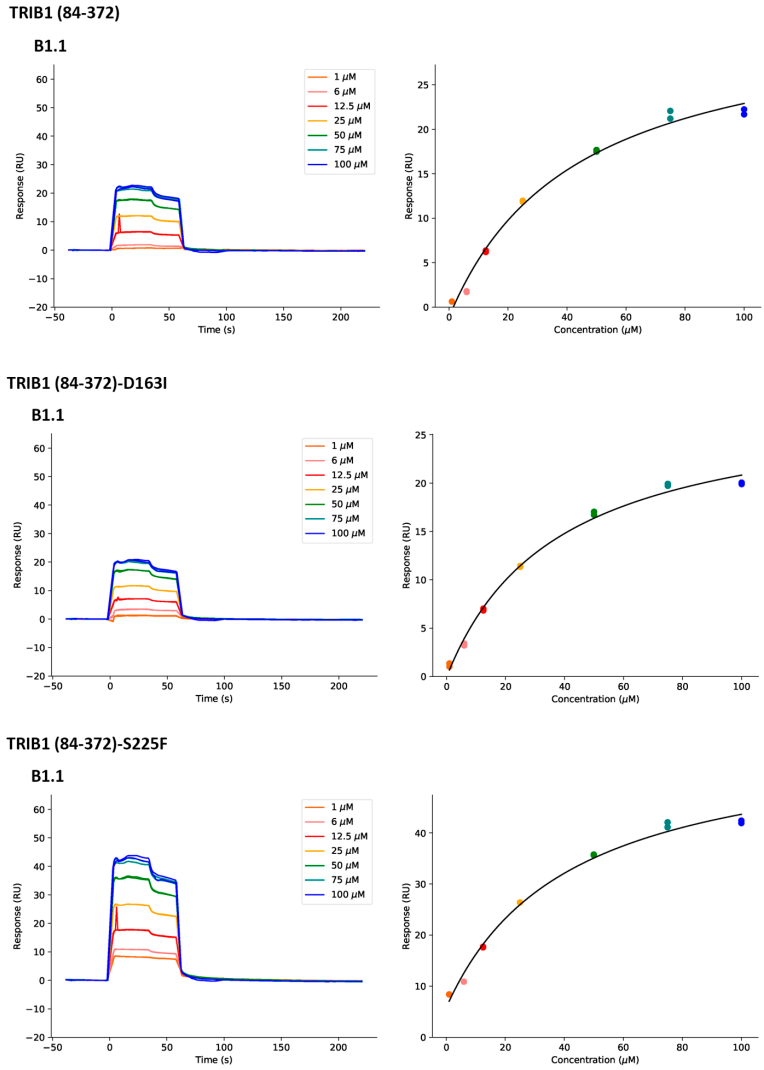


Figure 4.16. Sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of compound B1.1 against TRIB1 (84-372), TRIB1 (84-372)-D163I and TRIB1 (84-372)-S225F. Values shown are from experimental duplicates and sensorgrams are plotted after removing spikes due to buffer changes.

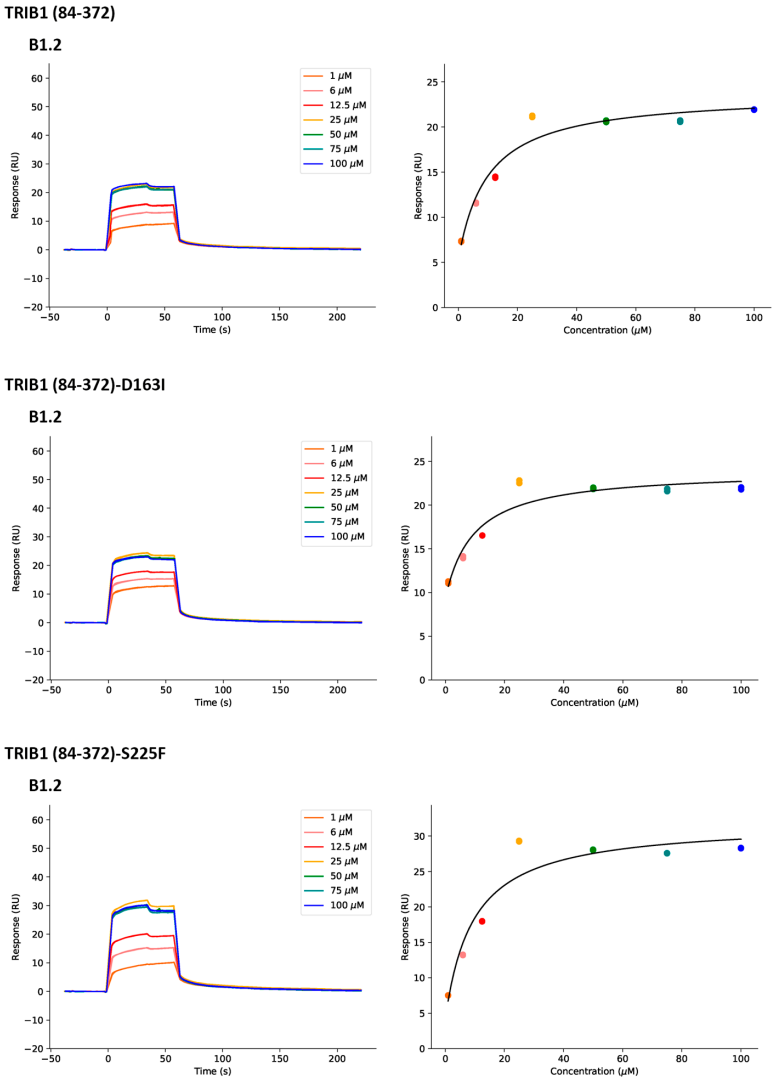


Figure 4.17. Sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of compound B1.2 against TRIB1 (84-372), TRIB1 (84-372)-D163I and TRIB1 (84-372)-S225F. Values shown are from experimental duplicates and sensorgrams are plotted after removing spikes due to buffer changes.

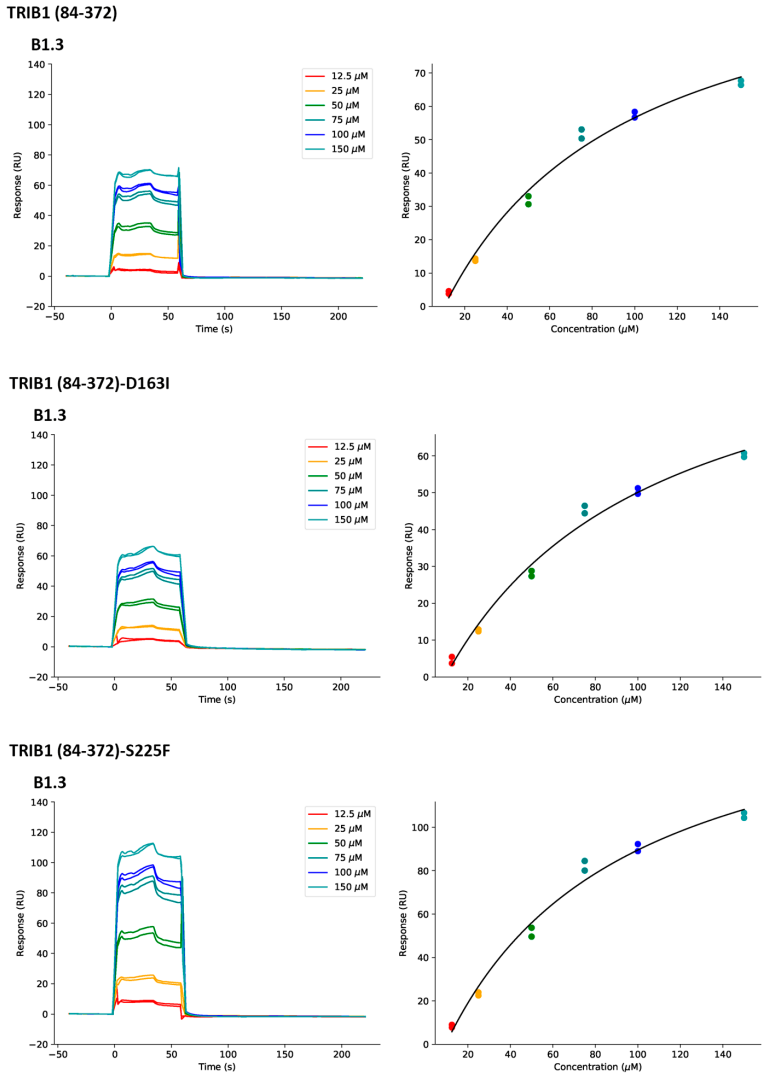


Figure 4.18. Sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of compound B1.3 against TRIB1 (84-372), TRIB1 (84-372)-D163I and TRIB1 (84-372)-S225F. Values shown are from experimental duplicates and sensorgrams are plotted after removing spikes due to buffer changes.

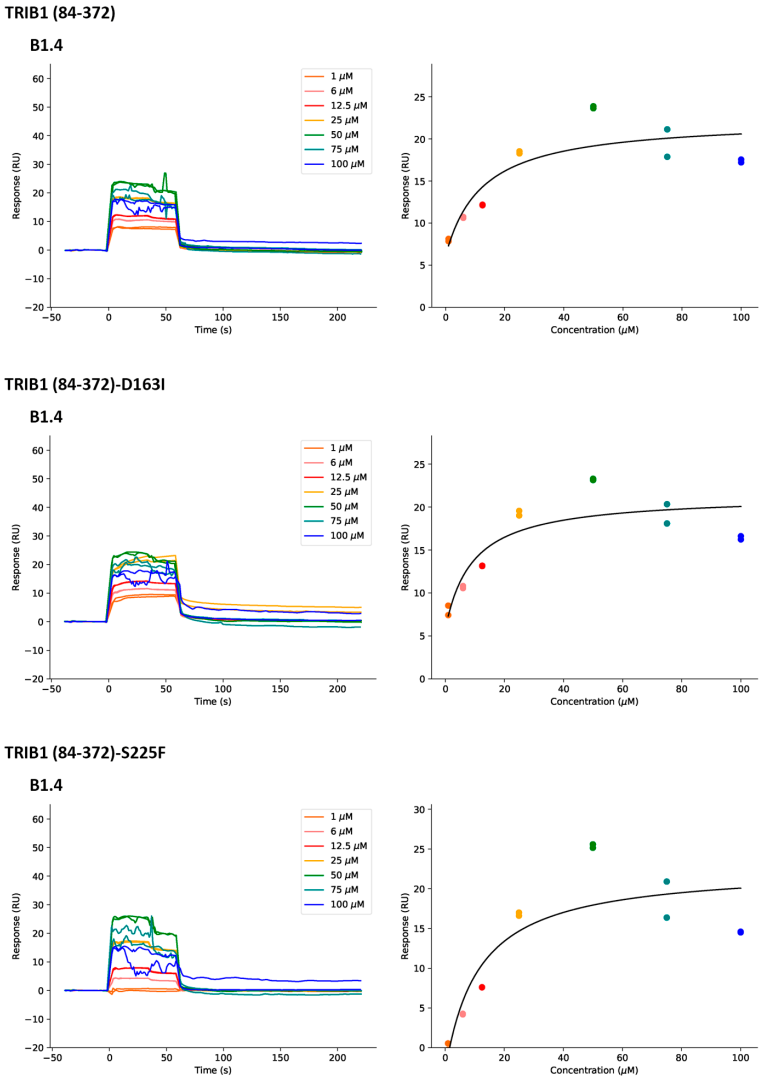


Figure 4.19. Sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of compound B1.4 against TRIB1 (84-372), TRIB1 (84-372)-D163I and TRIB1 (84-372)-S225F. Values shown are from experimental duplicates and sensorgrams are plotted after removing spikes due to buffer changes.

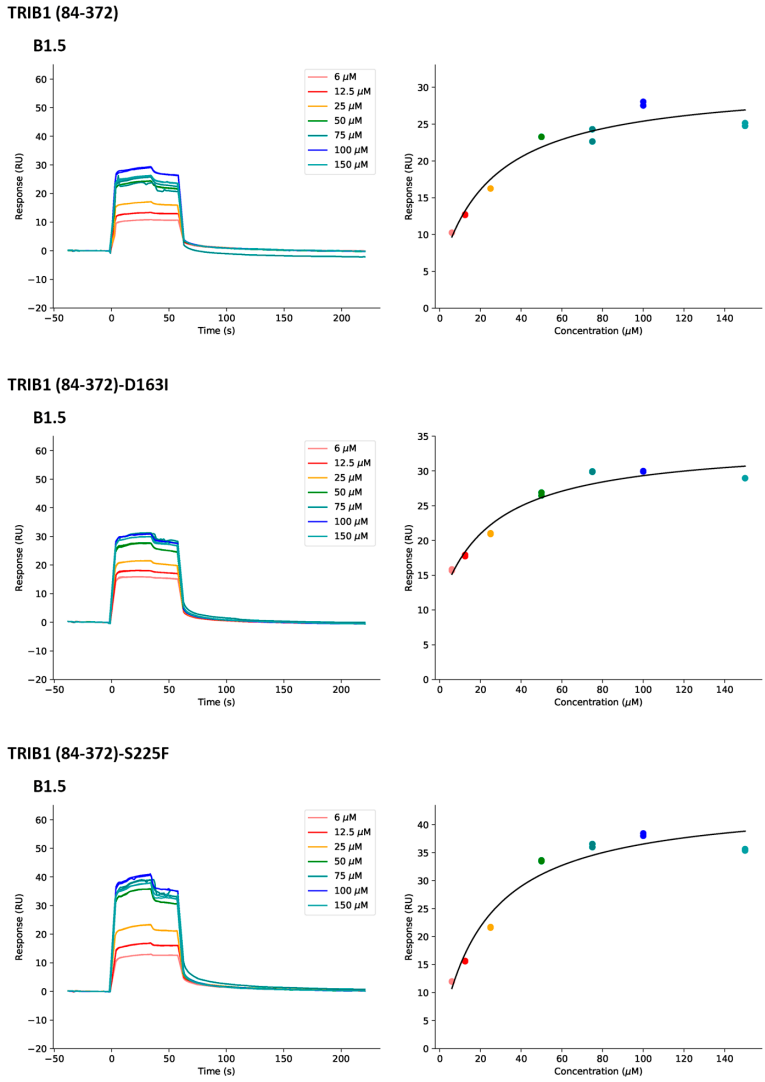


Figure 4.20. Sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of compound B1.5 against TRIB1 (84-372), TRIB1 (84-372)-D163I and TRIB1 (84-372)-S225F. Values shown are from experimental duplicates and sensorgrams are plotted after removing spikes due to buffer changes.

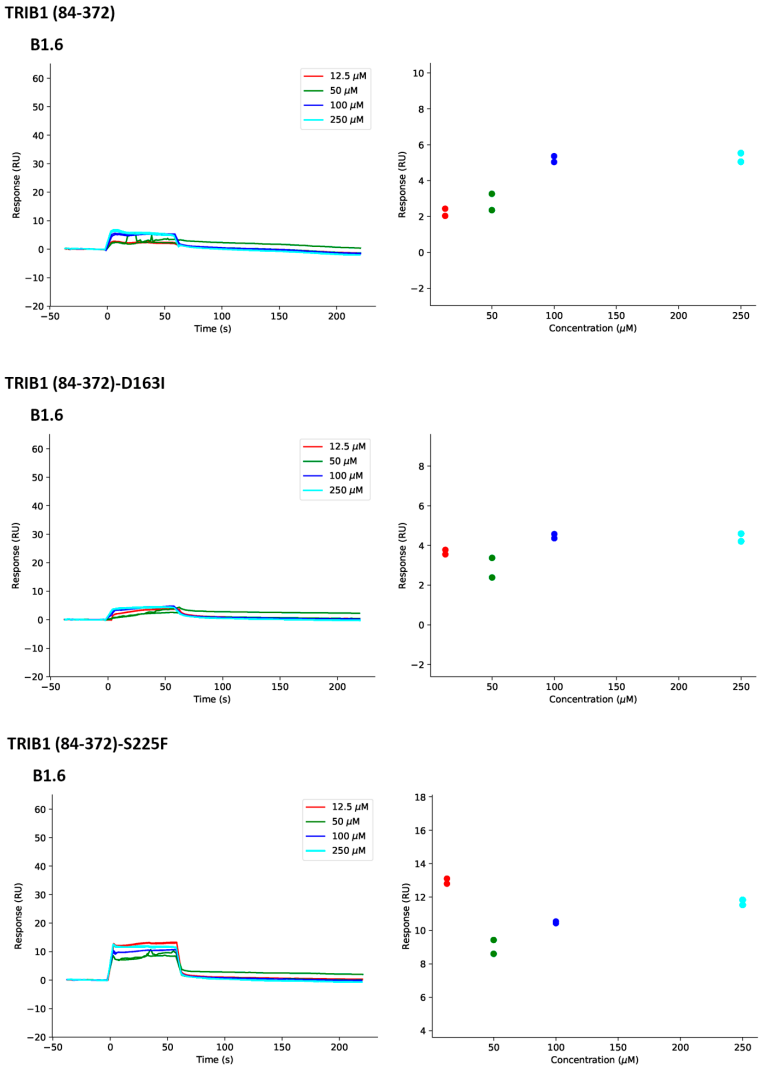


Figure 4.21. Sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of compound B1.6 against TRIB1 (84-372), TRIB1 (84-372)-D163I and TRIB1 (84-372)-S225F. Values shown are from experimental duplicates and sensorgrams are plotted after removing spikes due to buffer changes.

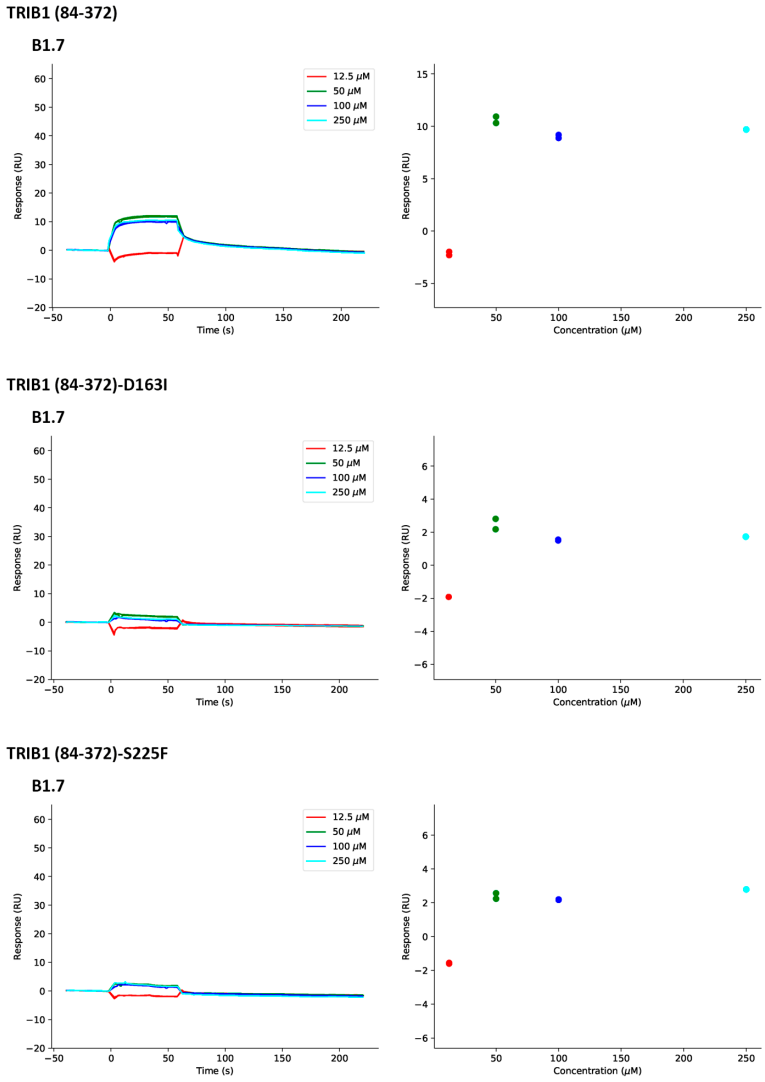


Figure 4.22. Sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of compound B1.7 against TRIB1 (84-372), TRIB1 (84-372)-D163I and TRIB1 (84-372)-S225F. Values shown are from experimental duplicates and sensorgrams are plotted after removing spikes due to buffer changes.

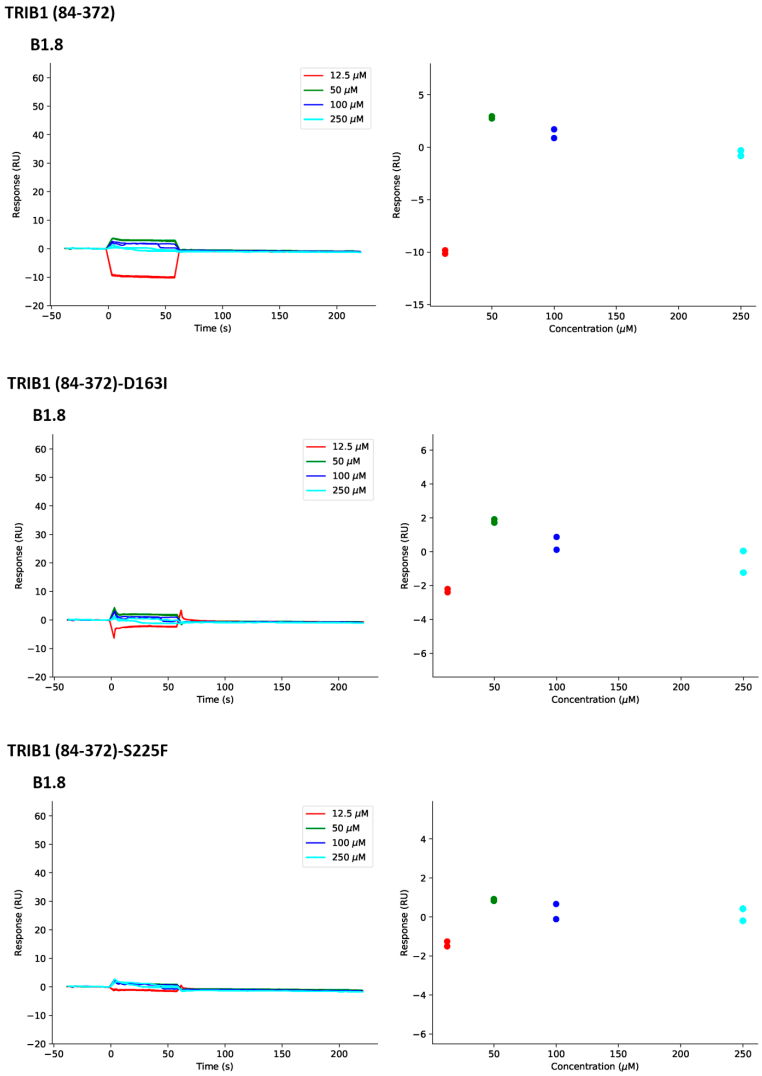


Figure 4.23. Sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of compound B1.8 against TRIB1 (84-372), TRIB1 (84-372)-D163I and TRIB1 (84-372)-S225F. Values shown are from experimental duplicates and sensorgrams are plotted after removing spikes due to buffer changes.

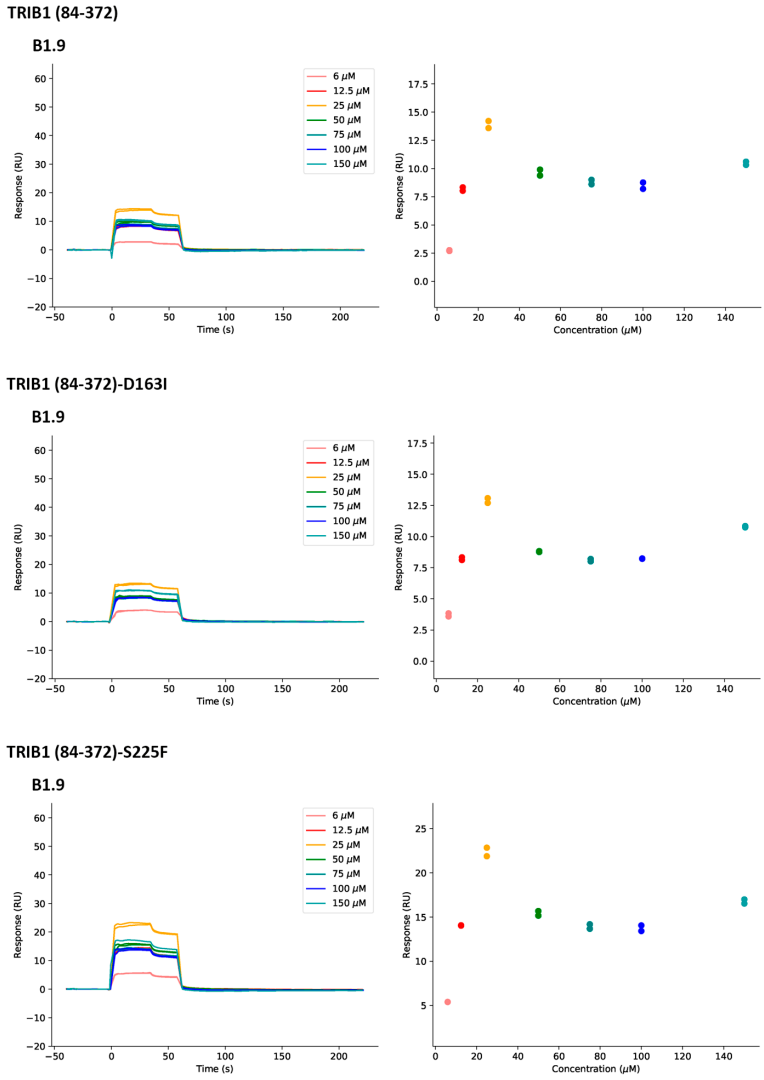


Figure 4.24. Sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of compound B1.9 against TRIB1 (84-372), TRIB1 (84-372)-D163I and TRIB1 (84-372)-S225F. Values shown are from experimental duplicates and sensorgrams are plotted after removing spikes due to buffer changes.

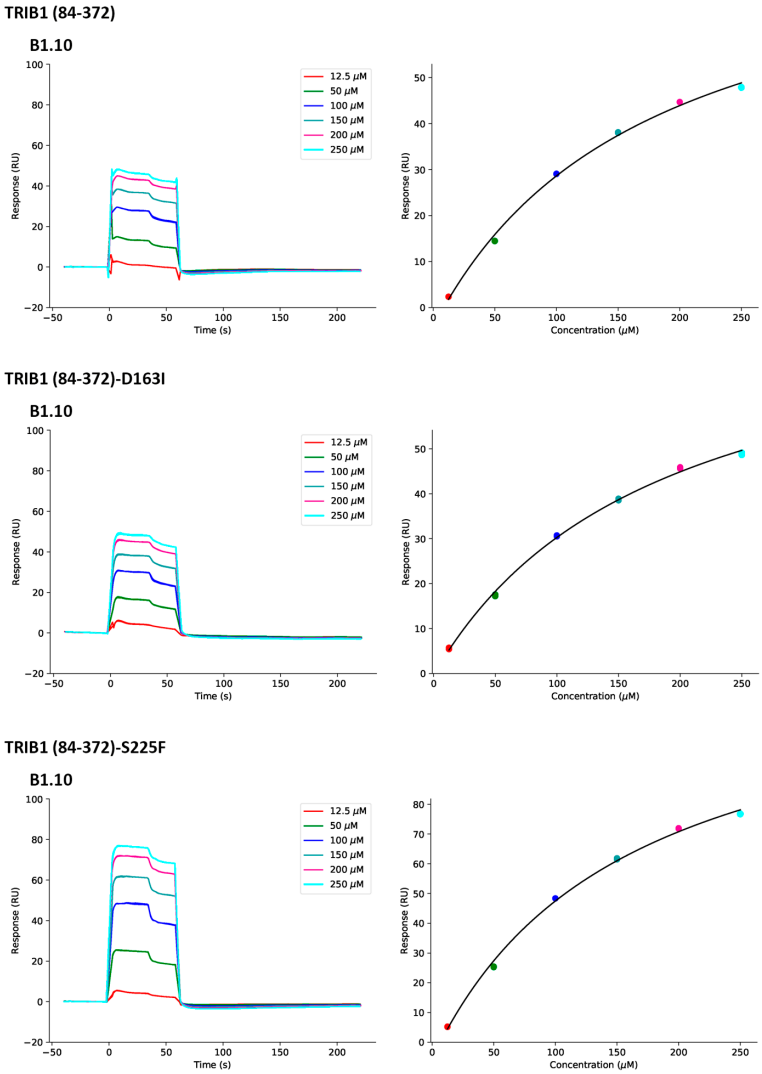


Figure 4.25. Sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of compound B1.10 against TRIB1 (84-372), TRIB1 (84-372)-D163I and TRIB1 (84-372)-S225F. Values shown are from experimental duplicates and sensorgrams are plotted after removing spikes due to buffer changes.

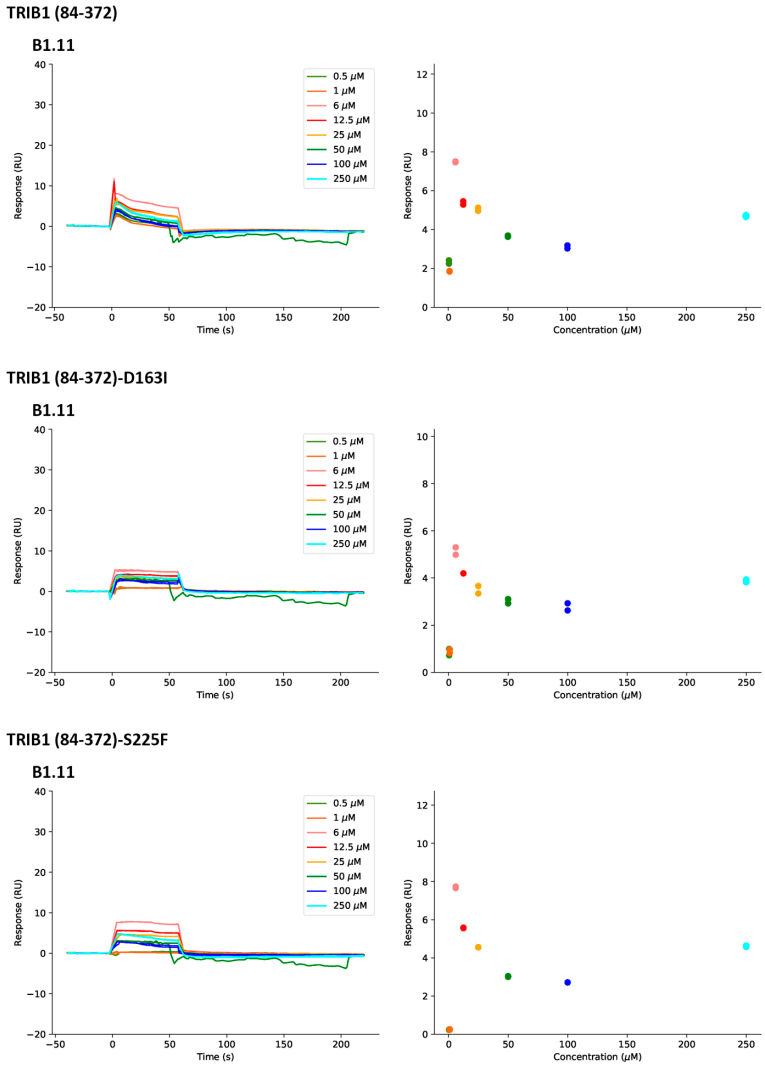


Figure 4.26. Sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of compound B1.11 against TRIB1 (84-372), TRIB1 (84-372)-D163I and TRIB1 (84-372)-S225F. Values shown are from experimental duplicates and sensorgrams are plotted after removing spikes due to buffer changes.

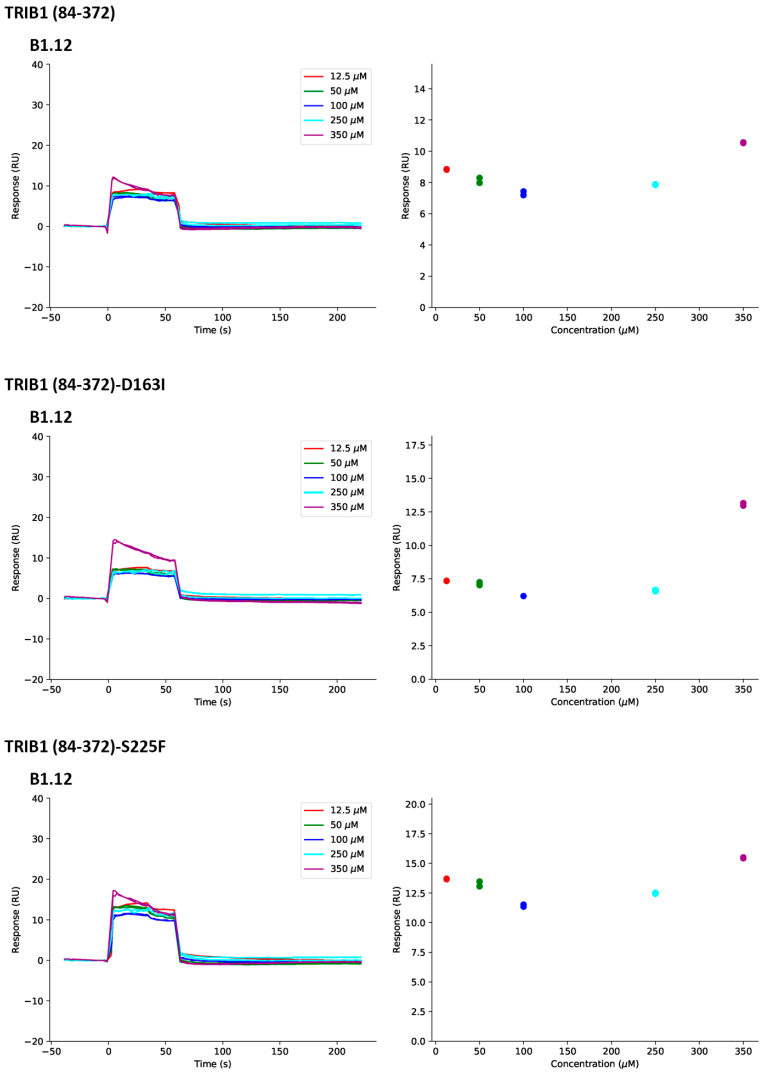


Figure 4.27. Sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of compound B1.12 against TRIB1 (84-372), TRIB1 (84-372)-D163I and TRIB1 (84-372)-S225F. Values shown are from experimental duplicates and sensorgrams are plotted after removing spikes due to buffer changes.

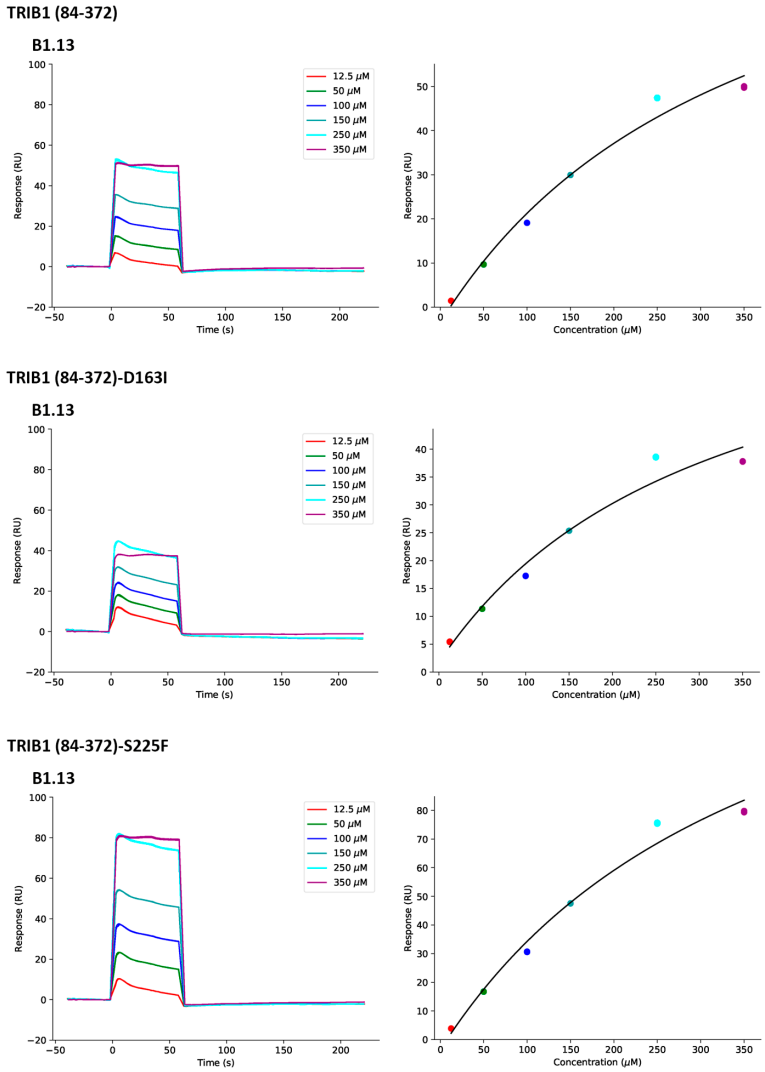


Figure 4.28. Sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of compound B1.13 against TRIB1 (84-372), TRIB1 (84-372)-D163I and TRIB1 (84-372)-S225F. Values shown are from experimental duplicates and sensorgrams are plotted after removing spikes due to buffer changes.

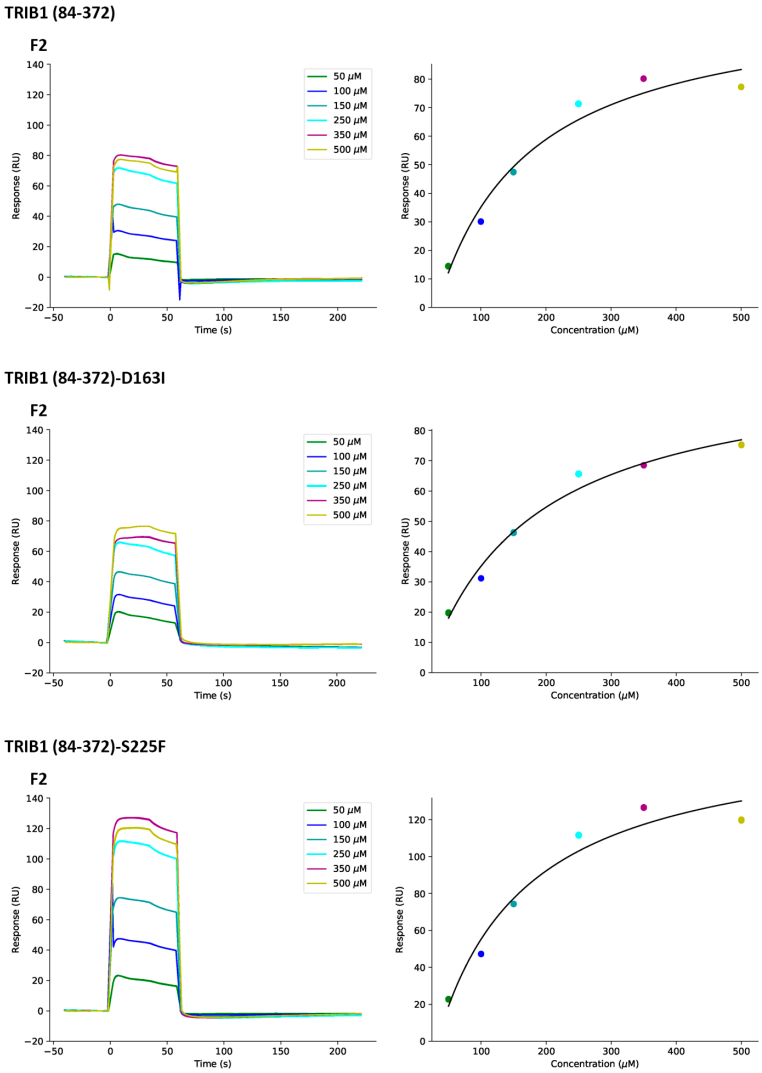


Figure 4.29. Sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of fragment F2 against TRIB1 (84-372), TRIB1 (84-372)-D163I and TRIB1 (84-372)-S225F. Values shown are from experimental duplicates and sensorgrams are plotted after removing spikes due to buffer changes.

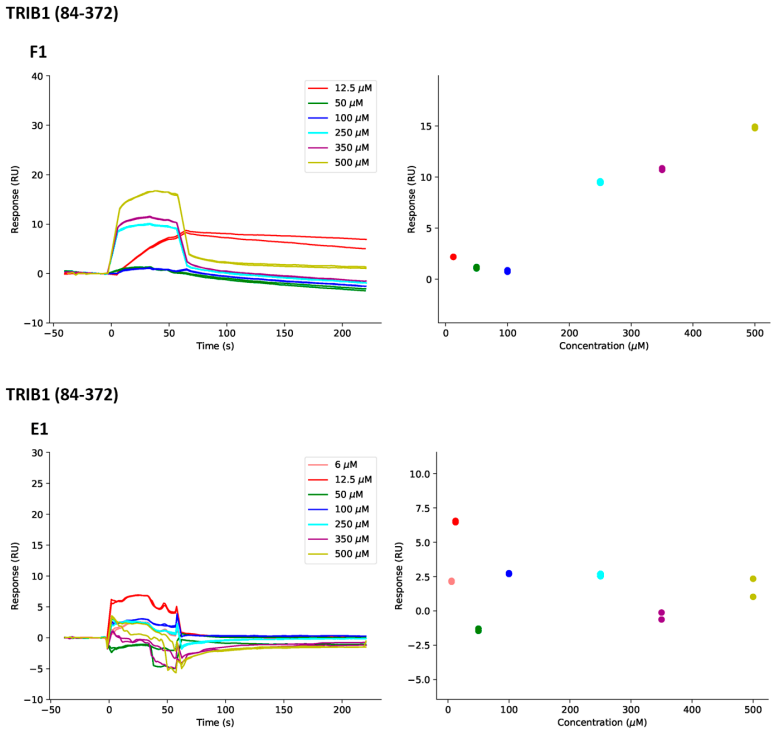


Figure 4.30. Sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of fragment F1 and E1 against TRIB1 (84-372). Values shown are from experimental duplicates and sensorgrams are plotted after removing spikes due to buffer changes

4.2.6 Panel of kinases

To probe whether our best compound has effects in kinases beyond the pseudokinase TRIB1, it was screened against a panel of kinases. We tested the specificity of the best hit, compound B1.2, against a panel of serine/threonine tyrosine, and lipid kinase from the University of Dundee (Figure 4.31). The panel contains 140 kinases plus 15 lipid kinases from the 555 possible [53], where only 11 kinases were inhibited at least more than 25% when tested with compound B1.2 at 10 μM (Table 4.2). Among the inhibited kinases, only four out of the nine families (AGC, CAMK, CK1, CMGC, Other, RGC, STE, TK and TKL) contain at least one kinase.

Kinase	Family	% Activity	SD	% Inhibition
CHK2	CAMK	17	4	83
EPH-B2	TK	42	8	58
EPH-B3	TK	53	15	47
p38d MAPK	CMGC	56	5	44
TrkA	TK	57	6	43
p38g MAPK	CMGC	64	13	36
DYRK1A	CMGC	68	7	32
GCK	STE	69	10	31
DDR2	TK	69	8	31
NUAK1	CAMK	70	8	30
PIK4Cb	Lipid Kinases	75,2	1,0	24,77

Table 4.2. List of kinases inhibited by B1.2 at 10 μM. The percentage of activity was calculated as the mean of two experiments. SD: Standard deviation. The percentage of inhibition is calculated by subtracting the percentage of activity to 100.

The kinase most inhibited is Checkpoint kinase 2 (CHK2) from the CAMK family, a tumor suppressor which is required for the induction of cell cycle arrest and apoptosis by DNA damage [269]. Another member of the CAMK family, NUAK1 is inhibited around 30%. The tyrosine kinase (TK) family is the most overrepresented family from the set of inhibited kinases. Three receptors TK, EPH-B2, EPH-B3 and TrkA are inhibited more than 40% and the TK DDR2 around 30%. Three members of the CMGC family were inhibited at least 30%, two p38 MAPK (p38d-MAPK13 and p38g -MAPK12) and DYRK1A. Among the 15 lipid kinases tested, only PIK4CB was inhibited at least 25%. Most of

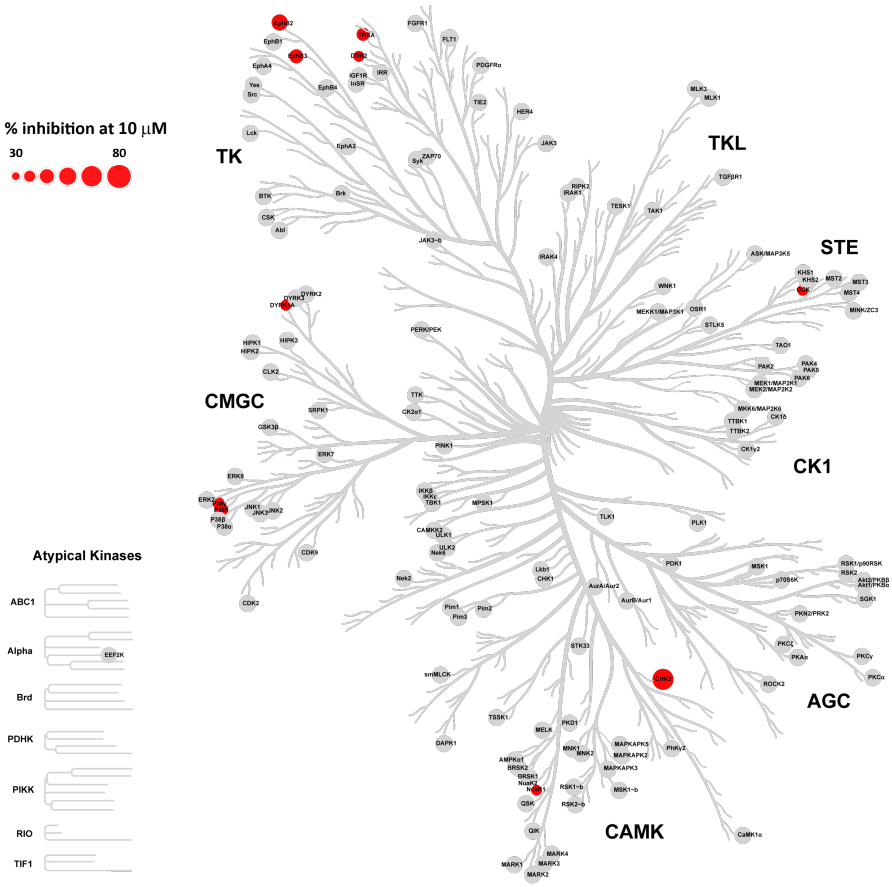


Figure 4.31. Kinome inhibition profile at 10 μM of compound B1.2. Inhibition is shown with a red circle proportional to the strength of the inhibition. Grey circles are kinases where the compound inhibits less than 25%. Coral was used to make the kinome tree [271].

the lipid kinases were not affected by the compound B1.2, where the next lipid kinase most inhibited is PI3Kα with an inhibition of around 23% (Table A5).

5 Results TRIB3-AKT/PKB

5.1 Background

The interaction between AKT/PKB and TRIB3 is already well known although the specific mechanism on how they interact remains unclear. The AKT/PKB family belongs to the AGC superfamily of serine-threonine kinases and consists of three family members: AKT1/PKBα, AKT2/PKBβ, and AKT3/PKBγ. AKT/PKB is crucial in multiple cell processes like metabolism, proliferation, and survival [272]. Dysregulation of AKT/PKB signaling is associated with cancer, diabetes, cardiovascular and neurological diseases. That is why, there has been an increasing interest to understand and modulate AKT/PKB activity [273].

From a structural point of view, AKT/PKB contains three domains. The N-terminal pleckstrin homology (PH) domain, which interacts with the lipid membrane through PIP3 (phosphatidylinositol (3,4,5)-triphosphate), recruits AKT/PKB to the plasma membrane. Following the PH domain, a central kinase domain belongs to the AGC superfamily and a C-terminal domain that contains the hydrophobic motif. Once AKT/PKB is recruited to the plasma membrane is phosphorylated by mTORC2 kinase at Ser473 and PDK1 kinase at Thr308. Another constitutive phosphorylation site is Thr450 in the turn motif which is essential for AKT folding and stability [274]. AKT/PKB may be phosphorylated on Ser477 and Thr479 by Cdk2-cyclin A. The Thr308 is in the activation loop and the three other phosphorylation (Ser437, Ser477, and Thr479) are in the C-terminal tail [275]. The phosphorylated residues are protected from dephosphorylation if AKT/PKB remains bound to the membrane. Upon termination of PIP3 signal by lipid phosphatases such as PTEN (phosphatase and tensin homolog), dissociation from the membrane leads to inhibition of the kinase domain by the PH domain and dephosphorylation of Ser473 by PHLPP (PH domain leucine-rich repeat containing protein phosphatase) [276] and the activation loop by PP2A [277] (Figure 5.1).

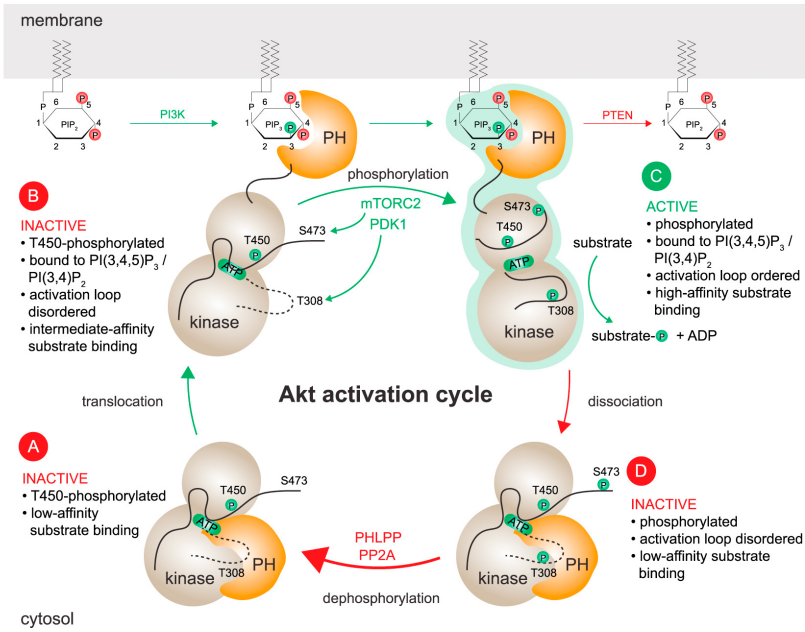


Figure 5.1. Conformational cycle of AKT/PKB. Autoinhibited AKT/PKB is characterized by a PH-in conformation in which the PH domain is sequestered by interacting with the kinase domain. Upon recruitment to the membrane, the kinase domain exposes two phosphorylation sites, one in the activation loop (Thr308) and the other in the hydrophobic motif (Ser473). These phosphorylation lead to conformational changes allowing the activation of AKT/PKB. Dissociation from the membrane leads to inhibition of the kinase by the PH domain and dephosphorylation. AKT/PKB activation and activity are restricted to the membrane and the inactive form is mainly in the cytosol. Figure from [278] with permission from Elsevier.

The interaction between TRIB3 and AKT/PKB was first detected during a yeast two-hybrid screen designed to identify AKT-binding proteins [219]. According to these results, TRIB3 binds to the kinase domain of AKT/PKB in the activation loop and inhibits its phosphorylation at both Thr308 and Ser473 by growth factors and insulin [219, 279]. In another article was proved the lack of evidence of TRIB3 as an inhibitor of AKT/PKB [280], but as they declared in their publication "an important difference between the two studies relates to the cell systems used, primary hepatocytes in the present study and tumoral HEK-293 cells by Du et al". An essential element of the protein-protein interaction is the biological context. Not all possible interactions will occur in any cell at any time. Instead, interactions depend on cell type, cell cycle phase and state, developmental stage, environmental conditions, protein modifications (e.g., phosphorylation), presence of cofactors, and presence of other binding partners [241]. Interestingly, it was also hypothesized that TRIB3 binds to mTORC2 [281]. mTOR is a conserved serine/threonine kinase modulated by growth factors and cellular energy status. It forms two distinct molecular complexes known as mTOR complex 1 (mTORC1) and mTORC2. mTORC1 regulates growth, autophagy, survival, and metabolism, whereas the role of mTORC2 in cellular biology is not completely understood. Sarbassov et al demonstrated that the Rictor component in mTORC2 complex directly phosphorylates AKT at Ser 473 residue [282]. In addition, Borsting et al observed differential effects of TRB3 on phosphorylation of AKT at Ser 473 and Thr 308, they hypothesized that TRB3 binds to mTORC2 [281]. They reported that TRIB3 may bind to both Rictor and mTOR, and not directly to AKT. It was not until the end of 2019 when it was identified that the αC helix of AKT1/PKBα (residues 192-204) interacts with TRIB3 [236]. They performed several co-immunoprecipitations (co-IPs) and even calculated the kinetic interaction of the αC helix of AKT1 and TRIB3 through SPR. Moreover, they demonstrated that the αC helix peptide disrupts the TRIB3-AKT1/PKBα interaction.

Bearing all of this in mind, we need to be very careful and to evaluate in detail all the research done to prove a reliable interaction between TRIB3 and AKT/PKB. One of the best-known functions of pseudokinases is the modulation of kinases. Dimerization has already been proposed as an allosteric mechanism for kinases [283, 284]. In several kinases, the conformation of αC helix is a key hub on which regulatory inputs converge to induce catalytic switching [127]. Several modes of Kinase-pseudokinase heterodimerization have been reported with very different αC helix orientation (Figure 5.3). Some of the

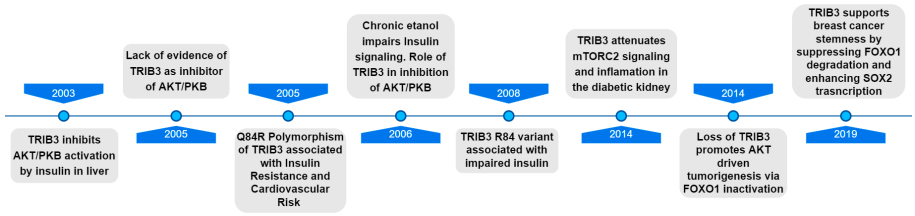


Figure 5.2. Timeline of the most relevant discoveries about AKT/PKB and TRIB3 interaction.

modes are head to tail (e.g., HER3 pseudokinase-EGFR kinase), transverse side to side (e.g., Tyk2-b pseudokinas-Tyk2 kinase), heterotrimeric complex (e.g., STRADα pseuokinase- LKB1 kinase), parallel side to back (e.g., MLKL pseudokinase- RIPK3 kinase). This shows the high flexibility of pseudokinases and kinases to interact with each other and the complexity of it. Therefore, it is not a surprise that the αC helix of AKT1 is a key region for the interaction with TRIB3. Moreover, when the Serine 473 located in the C-terminal is phosphorylated it allows a conformational change by interacting with the αC helix (Figure 5.1). Here we hypothesized that the C-tail of TRIB3 can regulate the phosphorylation of the S473 by interacting with the αC helix groove on the AKT1/PKBα N-lobe. We have tried to identify the mechanism of interaction between the kinase domain of AKT1/PKBα and the pseudokinase domain of TRIB3.

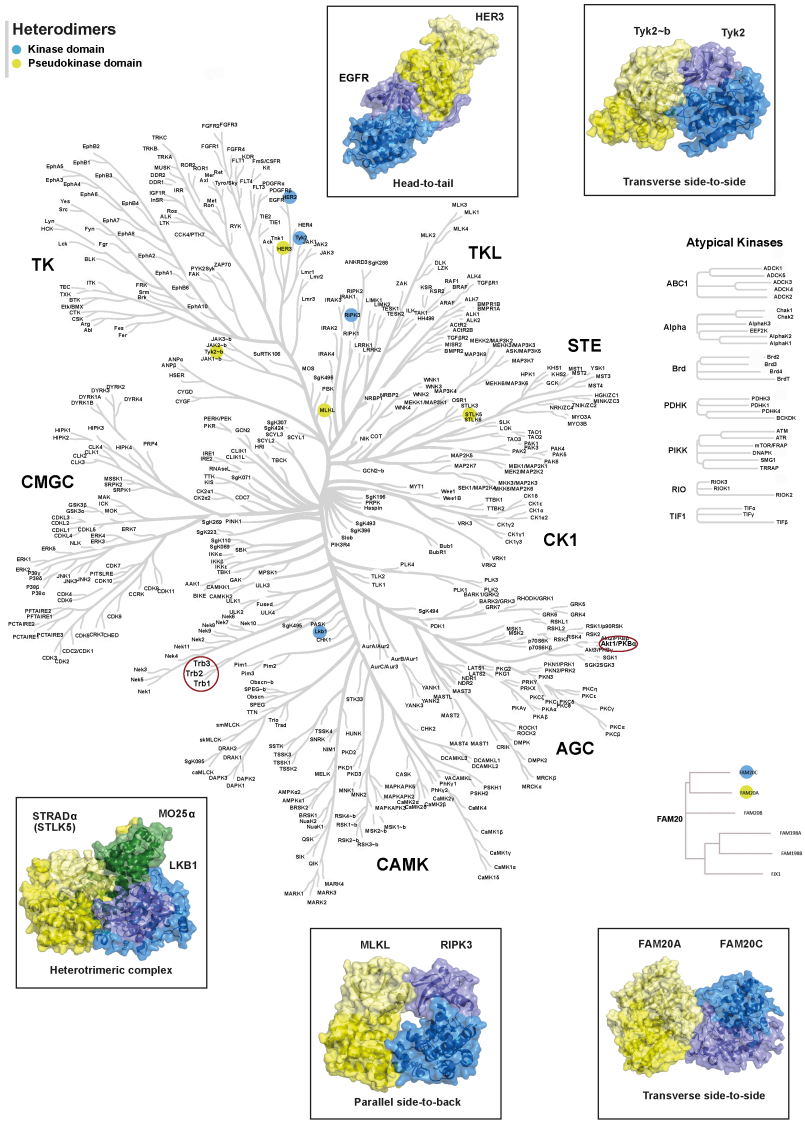


Figure 5.3. Kinase-pseudokinase heterodimers present in the kinome together with the Fam20 family of atypical kinases. Dendrogram showing the main families found in the human kinome, TRIBs family and AKT/PKB are highlighted with a circle. Some of the modes are head to tail (e.g., HER3 pseudokinase-EGFR kinase PDB code 4RIW), transverse side to side (e.g., Tyk2-b pseudokinas-Tyk2 kinase PDB code 4OLI, Fam20A-FAM20C PDB code 5YH3), heterotrimeric complex (e.g., STRADalpha pseuokinase- LKB1 kinase PDB code 2WTK), parallel side to back (e.g., MLKL pseudokinase- RIPK3 kinase PDB code 4M69). Kinase domains are shown in blue and pseudokianses in yellow with the N-terminal as purple and light yellow respectively.

5.2 Mapping of the TRIB3-AKT/PKB interaction

The interaction between TRIB3 and AKT was measured with cellular assays using PCA and computational modeling with Rosetta.

5.2.1 Protein Complementation Assay of AKT/PKB with TRIB3

The PCA was done in collaboration with Miguel Hernandez Quiles from UMC Utrecht. To determine how TRIB3 can interact with AKT, we tested several constructs mapping different regions of the TRIB3 sequence. We also included the Q84R mutation, since it was described as a gain of function amino acid substitution being a stronger inhibitor of insulin mediated AKT activation [285].

Based on this assay, there is no significant difference between the Gln84 and the Arg84 although the interaction was higher with the Arg mutant. The TRIB3 construct without the N-terminal or PEST domain showed similar values to the 84Q variant. Interestingly, when we removed the C-terminal tail of TRIB3 (residues 316-358) the interaction of TRIB3-AKT was reduced dramatically. Even only the C-terminal region of TRIB3 can interact with the full length of AKT (Figure 5.4). These data suggest that although the two proteins interact through their kinase and pseudokinase domain, the C-terminal region of TRIB3 is not only able to interact with AKT but also plays an important role in the interaction. To further interrogate the role of the C-terminal tail of TRIB3-AKT interaction, we tested how it was influenced by the phosphomimic mutation of Ser 472, a substitution of Asp for Ser to mimic Ser 473 phosphorylation. We found that the phosphomimic decreased the interaction between the TRIB3 C-terminal region and AKT (Figure 5.4.C).

5.2.2 Computational modeling of TRIB3 C-tail in complex with AKT1/PKBα

To gain structural insights into the interaction of TRIB3 C-tail with AKT1/PKBα, we used the Rosetta FlexPepDock ab-initio peptide docking protocol to model AKT1/PKBα bound to a segment of TRIB3 C-tail. In all cases, we calculated 100,000 conformations starting from an extended peptide within 10 Å from the proposed binding site, the groove at the αC helix of AKT/PKB. First, we tested the accuracy and precision of the Rosetta FlexPepDock ab-initio protocol by applying it to the C-tail of AKT1/PKBα (residues

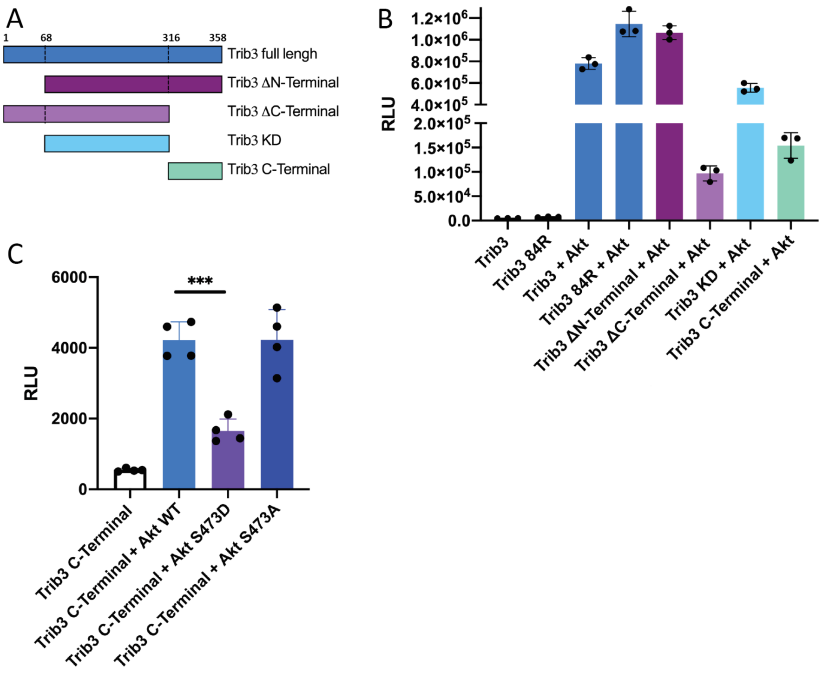


Figure 5.4. TRIB3-AKT interaction. A) Diagram of TRIB3 constructs used to map the AKT-TRIB3 interaction. B and C) After the co-transfection of the indicated NanoBit (AKT-C terminal SmBiT and TRIB3-N terminal LgBiT) the luciferase activity was measured as described in section 3.2.10.

463-SERRPHFPQFSYSAS-477) and the AKT1/PKBα kinase domain (KD; residues 144-450) (Figure 5.5). For analysis of the various protein-peptide docking screens, scattering plots were constructed from each model with the Rosetta score against the RMSD with respect to the reference structure. In the case of AKT the reference model is the native structure and for TRIB3 is the best model coming from the biggest cluster within the top 10 lowest energy structures.

Validation of the protocol

We tested the protocol with the C-tail of AKT as a retrospective case using the known native structure as the reference model of the interaction (PDB 4GV1). The C-tail of AKT is already crystallized interacting with its own kinase domain. The hydrophobic motif of AKT1/PKBα, residues from 469-FPQFSY-474, binds tightly to the αC helix when Ser473 is phosphorylated. The reference crystallographic structure contains a phosphomimic mutation of Ser472, which was also included in the C-tail of AKT.

We were able to generate models for the binding of the hydrophobic motif with residues close to near-native, as the predicted peptide conformation deviates by $\leq$ 2 Å RMSD. Especially for the Tyr474 which deviates by less than 1 Å in the best model (Figure 5.5.B). In contrast, residues located at the terminal showed a much higher RMSD (e.g., Ser477), given the high flexibility of the peptide. The best docking model corresponds to the representative model selected according to the best energy score of the biggest cluster in the top 10 ranked by energy. A successful sampling and ranking are considered when a near native model is ranked among the ten lowest ranking clusters. We expect to find the best result within the biggest cluster among the ten lowest. Six out of the ten best ranked models showed an overall RMSD lower than 2 Å (Figure 5.5.A).

Several protein-peptide docking were carried out between the AKT1/PKBα KD and the selected region of TRIB3 C-tail in segments of 15 residues (Figure 5.6). The length of the peptide was selected based on the maximum length used to validate the FlexPepDock protocol. In all the three segments the best model appeared docked in the groove over the αC helix. Although, according to the docked models, the last 15 residues of the C-tail seem to be more flexible. The sampling of structures showed a larger RMSD distribution of the 10 top structures in comparison with the other two segments, which showed a more convergence solution to a low energy structure (funnel effect).

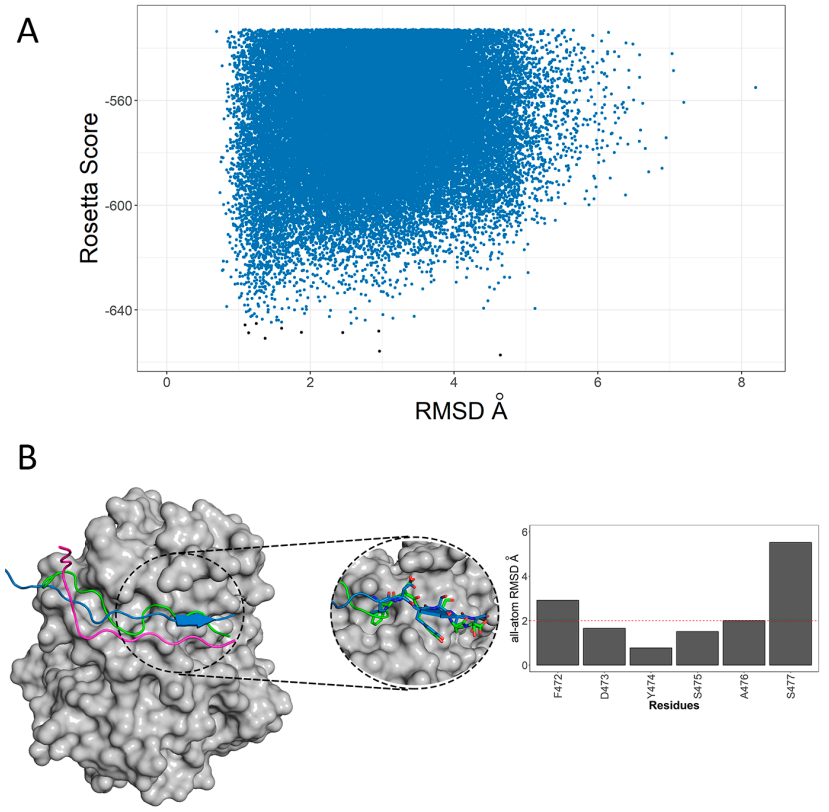


Figure 5.5. Protein-peptide docking structure of AKT KD with its own C-tail (463-477). A) Rosseta score energy landscape plots where each model is a blue circle according to its RMSD and its energy score. The top 10 lowest energy clusters created from the top 500 scoring models are shown as black circles. B) The peptide starting orientation (magenta), the best model (blue), and the native peptide (green) are shown in complex with AKT KD (grey). A detailed view of the best result showing the peptide side chains compared with the crystallized region along with the RMSD of residues 472-477.

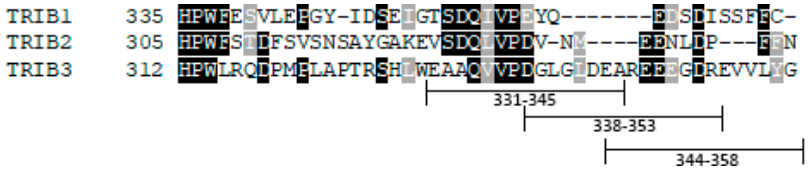


Figure 5.6. Multiple sequence alignment of C-terminal tail of TRIB1-3.

Our results converged into distinct conformations of the peptide among the top-scoring models. The best model from the first segment (residues 331-345) docked the COP1-binding motif in the same region as the hydrophobic motif located in the C-tail of AKT. The COP1 binding motif contains several hydrophobic residues (Val335-Val336-Pro337) although it has a very flexible region following the motif with two glycine residues (Gly339-Leu340-Gly341-Leu-342). The Val335 appears docked in the same place as the Phe472 from the C-tail of AKT1/PKBα. Among the main interactions, there are two main chain hydrogen bonds between Ser216-AKT1 and Asp338-TRIB3 and one between Lys214-AKT1 and Lue340-TRIB3 (Figure 5.7).

In the best model from the second segment (residues 338-352), the Leu342 appears docked in a hydrophobic groove where the Phe472 from the C-tail of AKT1/PKBα is also crystallized. Interestingly, the glutamic motif Glu347-Glu-348-Glu349 from TRIB3 C-tail appeared docked in a region close to a basic patch (Lys142-His143-Arg144) in AKT1/PKBα (Figure SX). Similar to the previous model, there are two main chain hydrogen bonds involving the residues Ser216-AKT1 and Arg346-TRIB3 and another with Lys214-AKT1 and Glu347 (Figure 5.8).

The best model from the third segment (residues 344-358) also docked the glutamic motif EEE close to the basic patch of AKT1/PKBα (Figure 5.9). In this case, there are no hydrophobic interactions with the groove over the αC helix of AKT1. Also, there are two main chain hydrogen bonds involving the residues Ser216-AKT1 and Arg346-TRIB3 and another with Lys214-AKT1 and Glu348.

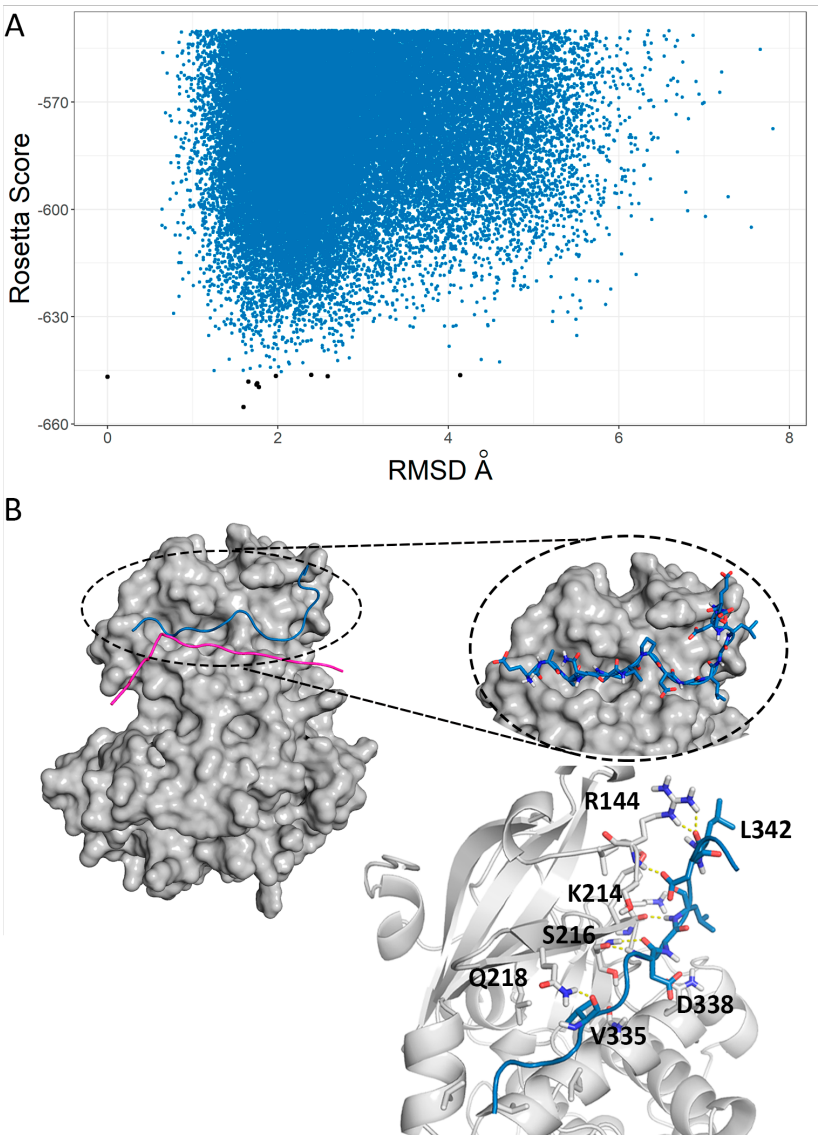


Figure 5.7. Protein-peptide docking structure of AKT KD with the best model of TRIB3 331-345. A) Rosseta score energy landscape plots where each model is a blue circle according to its all atom RMSD and its energy score. The top 10 lowest energy clusters created from the top 500 scoring models are shown as black circles. Using as reference the best result from the biggest cluster among the ten lowest. B) The peptide starting orientation (magenta) and the best model (blue) are shown in complex with AKT KD (grey). A detailed view of the best result showing the peptide side chains interacting with the αC helix of AKT KD.

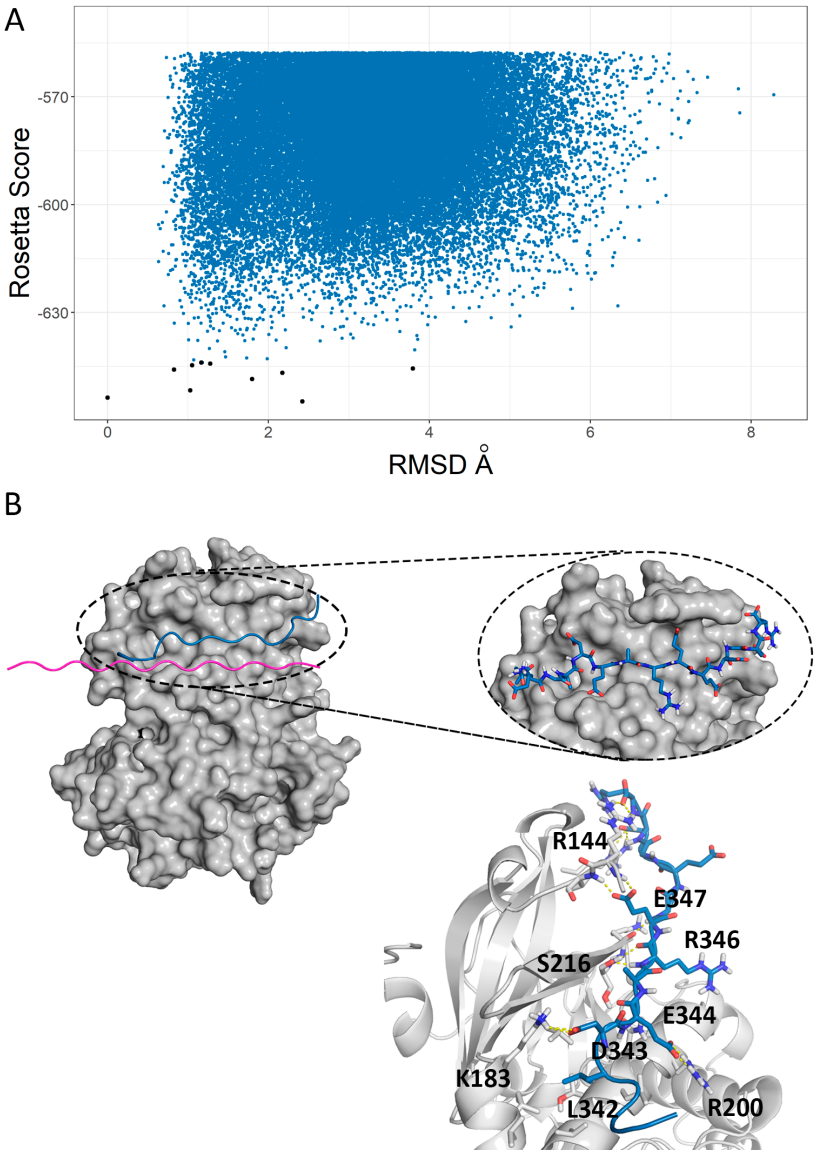


Figure 5.8. Protein-peptide docking structure of AKT KD with the best model of TRIB3 338-352. A) Rosseta score energy landscape plots where each model is a blue circle according to its all atom RMSD and its energy score. The top 10 lowest energy clusters created from the top 500 scoring models are shown as black circles. Using as reference the best result from the biggest cluster among the ten lowest. B) The peptide starting orientation (magenta) and the best model (blue) are shown in complex with AKT KD (grey). A detailed view of the best result showing the peptide side chains interacting with the αC helix of AKT KD.

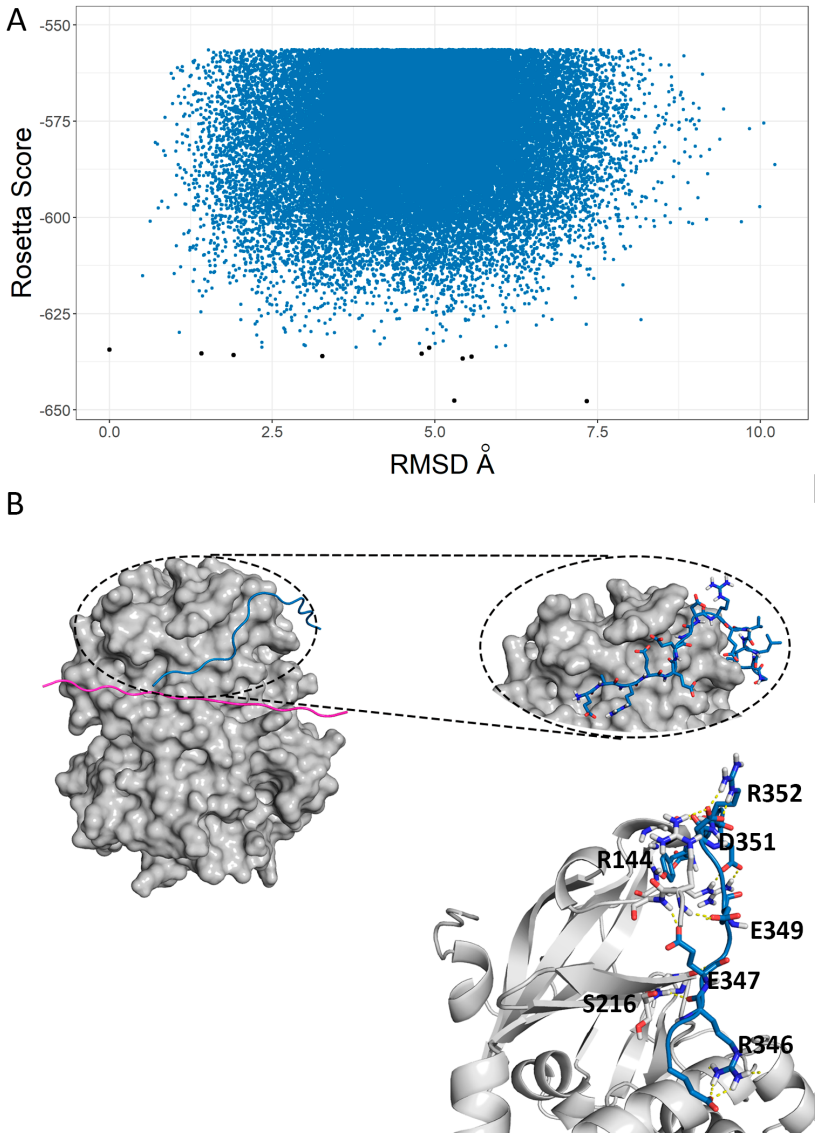


Figure 5.9. Protein-peptide docking structure of AKT KD with the best model of TRIB3 344-358. A) Rosseta score energy landscape plots where each model is a blue circle according to its all atom RMSD and its energy score. The top 10 lowest energy clusters created from the top 500 scoring models are shown as black circles. Using as reference the best result from the biggest cluster among the ten lowest. B) The peptide starting orientation (magenta) and the best model (blue) are shown in complex with AKT KD (grey). A detailed view of the best result showing the peptide side chains interacting with the αC helix of AKT KD.

6 Results TRIB3-SIAH1

6.1 Background

Despite the diverse roles of TRIB3 in different signaling pathways, there is still a lack of knowledge on TRIB3 regulation. TRIB3 upregulation is well-known to be triggered by cellular stress. There is evidence that endoplasmic reticulum (ER) stress, hypoxia, amino acid deficiency, glucose levels, or free fatty acid excess are responsible for TRIB3 upregulation [286]. However, post-translational modifications remain more elusive. For example, high-throughput proteomic analyses followed by mass-spectrometry confirmed the presence of seven phosphorylation sites on TRIB3 protein. Unfortunately, there is no data associated with a functional significance [287]. Another type of post-translational modification is ubiquitination, which regulates protein degradation by the 26S proteasome, a large multi-subunit protease. The ubiquitin-proteasome system (UPS) is the main proteolytic pathway of eukaryotic cells [288]. Ubiquitination is a multistep process requiring the activity of three classes of enzymes to transfer the ubiquitin molecule (76 amino acids). Initially, the ubiquitin-activating enzyme, E1, which activates the ubiquitin molecules and transferred to the ubiquitin-conjugating enzyme, E2, and is finally transferred to the substrate that is specifically bound to the ubiquitin-protein ligase, E3. Ubiquitin transfer can be either directly to the E3 bound substrate or through an additional step E3-ubiquitin. Ubiquitination is a reversible process and ubiquitin can be removed by deubiquitinating enzymes (DUB) [289]. There are different types of ubiquitination, monoubiquitination and polyubiquitination, leading to different cellular outcomes. For example, monoubiquitination can regulate DNA repair and gene expression and polyubiquitination generally targets proteins for proteasomal degradation [290].

Historically, there have been two major types of E3 according to their structure and mechanism, the RING (Really Interesting New Gene) and the HECT (Homologous to the E6AP carboxyl terminus). The RING E3s are characterized by their RING or U-box fold catalytic domain, which promotes direct

ubiquitin transfer from an E2 to a substrate. Distinct from the RING, the HECT forms an intermediate with ubiquitin before the modifier is ligated to the substrate [291]. The discovery of RING-IBR-RING (RBR) proteins as a unique family of RING-HECT hybrid E3s defined a third class of ubiquitin ligases distinct from the RING and HECT types [292].

E3 ligases can recognize their substrate through two different mechanisms, via short linear sequence or recognition of protein domains [293]. First it was discovered how ubiquitinated proteins interact with their E3 ligases through a short linear motif called degron. These binding motifs can suffer post-translational modifications which can improve the binding, like the Skp1-Cullin-F-box (SCF) E3 ligase which recognize phosphorylated degrons [293]. The opposite can also be the case, where phosphorylation disrupts the interaction between degron and ubiquitin ligase [294]. Recognition of protein domains offers a more distinct surface for E3 rather than a linear motif. As in the case of the ubiquitination of the telomere repeat binding protein TRF1 by the F-box of SCF [295]. TRIB3 protein turnover is regulated by at least two E3 ligases, APC/C (anaphase-promoting complex/cyclosome) and SIAH1 (Seven in Absentia Homolog 1). APC/C is used at the steady-state and SIAH1 response to genotoxic stress [296, 297]. APC/C is a ubiquitin RING E3 ligase complex which regulates the progression of the cell cycle. Ohoka et al. found a D-box motif (XXX), which is a sequence motif recognized by APC/C, being critical for TRIB3 ubiquitination. APC/C requires one or two WD40-domain proteins, Cdc20 or Cdh1 as activators. In the case of TRIB3, APC/C-Cdh1 regulates TRIB3 stability [296]. On the other hand, little is known about how SIAH1 interacts with TRIB3 and forms the E3 ligase complex. SIAH1 was identified as a novel interactor of TRIB3 through the course of a Yeast two hybrid screen (Y2H). They also used co-immunoprecipitation to confirm the interaction in mammalian cells [297]. SIAH (Seven In Absentia Homolog) proteins belong to the RING E3 ligase family. The seven in absentia (SINA) gene was identified in Drosophila eye [298]. The human genome contains two SINA homologs, SIAH1 and SIAH2 and both encode functional proteins. Interestingly, a new member of the family, SIAH3, was found in a limited subset of cancer cell lines [299]. The SIAH family of proteins are highly conserved across species but unlike SIAH1 and SIAH2, SIAH3 lacks a functional RING domain [300]. Human SIAH1 and SIAH2 proteins share 85% amino acids homology, whereas SIAH3 is 55% homologous. SIAH1 and SIAH2 differ mainly at the N-terminal where SIAH2 contains 20 residues extra. SIAH1 and SIAH2 proteins are localized in the nucleus and the cytoplasm where they control the degradation of

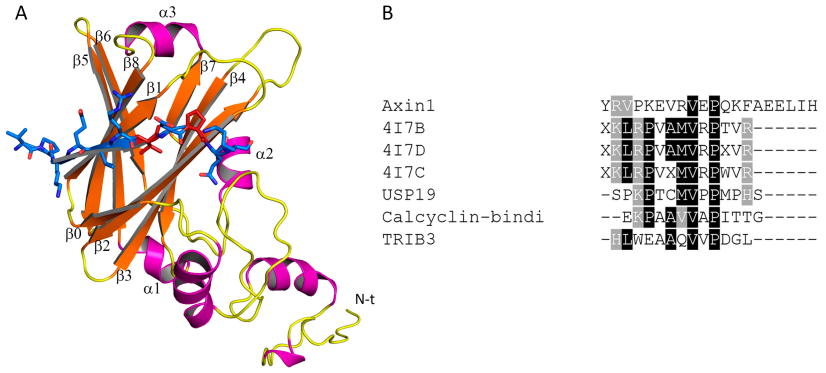


Figure 6.1. SIAH1 monomer in complex with Axin1 and sequence alignment of SIAH1 crystallized substrates and TRIB3 C-tail segment. A) Ribbon diagram showing the SBD and two zinc-finger subdomains. The α helices and β strands are coloured purple and orange, respectively. The N- terminal end is labeled, as well as the secondary-structure elements. Axin1 peptide is shown in blue with the VxP motif residues in red (PDB id 5WZZ). B) . The residues displayed with an X are modified residues that do not correspond with standard amino acids. The names refer to SIAH1 inhibitors (synthetic peptides (PDB codes 4I7B,4I7C,4I7D) and substrates as Calcyclin-binding protein (PDB code 2a25), USP19 (PDB code 4x3g),Axin1 (PDB code 5WZZ) and the proposed sequence of TRIB3.

proteins involved in a myriad of functions like transcriptional regulation, cell migration and neuronal-associated functions [301]. Because of the high similarity between SIAH1 and SIAH2, they share some ubiquitination substrates (e.g. ACK1) [302]. Despite this, they also have specific substrates. Indeed, the expression of SIAH1 and SIAH2 is differentially regulated, SIAH1 is induced by p53 in response to genomic stress, while SHIA2 is induced by hypoxia and estrogens [303].

SIAH proteins consist of an N-terminal RING domain, responsible for the interaction with the E2 conjugating to form the active complex for ubiquitination, two zinc finger subdomains and a C-terminal substrate binding domain (SBD) (Figure 6.1). The SBD is primarily responsible for dimer formation

[304]. The C-terminus can be considered as a substrate- and cofactor interaction domain because it interacts with several proteins [305, 306].

Most of the known SIAH1 binding proteins contains the sequence PxAxVxP where x is not conserved. The VxP motif was found to make the greatest binding contributions in some of the SIAH1 substrate, although not in all proteins [307]. The crystallized degrons peptides adopt an extended conformation, packing against the β1 and β2 strands of the β sandwich of the SIAH SBD (Figure X) [308]. SIAH1 can ubiquitinate alone, as in the case of β-catenin or when it forms an E3 ligase complex [309]. The ubiquitination process is more efficient when SIAH1 forms the complex with SIAH-interacting protein (SIP), the adaptor protein Skp1 that is common to the SCF complex and the F-box protein Ebi which binds B-catenin [310]. Therefore, not all SIAH binding proteins are targeted for degradation. Some proteins act as cofactors helping to recruit substrates and others can function as negative regulators [311].

Given the importance of regulating TRIB3 protein levels, we decided to investigate more in detail TRIB3-SIAH1 interaction. In this thesis, TRIB3 is proposed to bind to the SBD domain. We proposed this binding hypothesis based on the similarity between a fragment in TRIB3 and a consensus motif of several crystallized SIAH1 substrates (Figure 6.1.B). The VxP motif is also found in TRIB3 in a region previously seen to interact with the E3 ligase COP1, also known as the COP1 binding domain [212]. The COP1 binding motif of Tribbles is in the C-tail, an expected disordered region.

The main objective of the work described in this chapter was to identify and characterize the region of TRIB3 interacting with SIAH1. We have investigated how TRIB3 could interact with SIAH1 by applying a combination of computational (protein-peptide docking) and biophysical (isothermal titration calorimetry) techniques.

6.2 TRIB3-SIAH1 interaction study

6.2.1 Computational modeling of TRIB3 328-340 in complex with SIAH1

To modulate the interaction between TRIB3 and SIAH1, three models of the SIAH1-TRIB3 328-340 complex were created based on homology modeling using MODELLER. We selected the residues 328 to 340 from TRIB3

(HLWEAAQVVPDGL) as the region interacting with SIAH1 (Fig X). The models were selected based on the lowest DOPE score from MODELLER. The three models are shown in figure X, and as can be seen the variations between them are minimal except for the N and C-terminal ends. The TRIB3 residues considered as main contributors of the SIAH1 binding domain, valine 335 (V335) and proline 337 (P337), have the same position in all models. The main interactions stabilizing the SIAH1 SBD and TRIB3 peptide are from the residues close to the VxP motif. V335 and P337 are both interacting with a hydrophobic pocket and stabilized by hydrogen bonds between backbone atoms of nearby residues. In addition, V335 forms hydrogen bonds with the carboxyl backbone of leucine 166 and the amine backbone of threonine 168 from SIAH1. Alanine 333, which is also conserved compared with other SIAH1 substrates (Fig X) forms hydrogen bonds with the carbonyl backbone of valine 164 and the amine backbone of leucine 166. Several hydrogen bonds are formed between the main chains of SIAH1 and TRIB3 stabilizing the peptide between the sheets of the SBD.

6.2.2 Expression and purification of SIAH1

The gene coding for SIAH1 SBD (residues 91-282) was obtained from a commercial source and inserted into a pET-28 vector for protein expression. The protein was expressed in E.coli BL21 (DE3) cells and after cell lysis, nickel affinity chromatography was used to purify His-SIAH1 SBD. After elution with 300 mM imidazole, the majority of His-SIAH1 SBD eluted in the first and second fractions (Figure X). Fractions E1 and E2 from nickel affinity chromatography were polled together and digested overnight with 3C protease. Eluted SIAH1 SBD was analyzed by size exclusion chromatography (SEC) as described in the methods section. SIAH1 eluted from the SEC as a symmetric peak (Figure X). The molecular weight of SIAH1 SBD is 21,571 Da which is consistent with the weight of the sample by SDS-PAGE (Figure X). The purification yields were approximately 10 mg per liter of bacterial culture, calculated using a NanoDrop spectrophotometer. The absorbance was adjusted with the extinction coefficient of the protein that is calculated from the amino acid sequence of the protein by the ProtParam tool from Expert Protein Analysis System (ExPASy) [263]. Purified protein was aliquoted and snap frozen in liquid nitrogen for storage at -80 ºC.

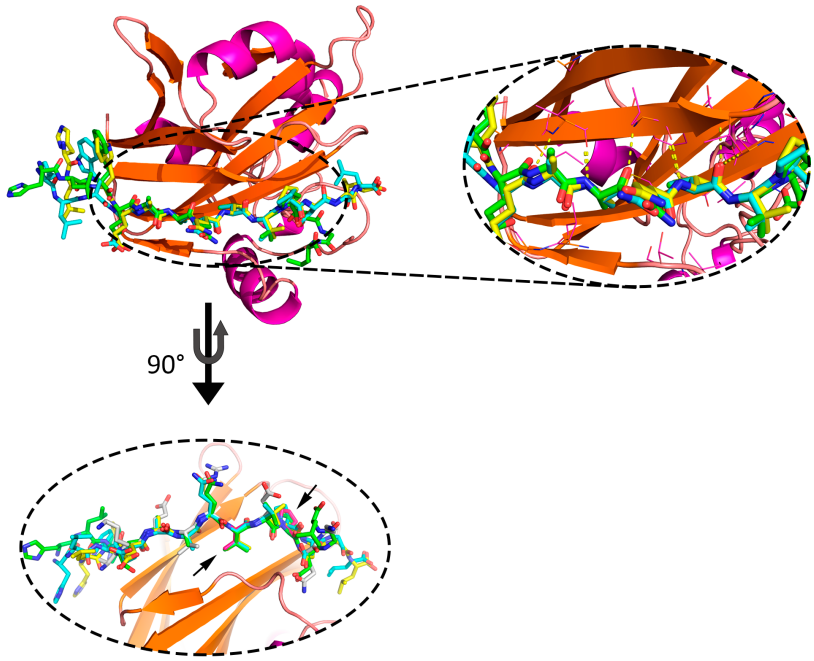


Figure 6.2. Protein-peptide interface of SIAH1 SBD and TRIB3 328-340 peptide. SIAH1-SBD is shown as a ribbon diagram with the proposed binding mode of TRIB3 peptides. Superposition of the 3 models obtained using several templates (complex 1 in yellow, complex 2 in blue and complex 3 in green). Zoom view depicting some of the residues forming the intermolecular hydrogen bonding (yellow dashed line), where nitrogen and oxygen atoms are colored blue and red, respectively.

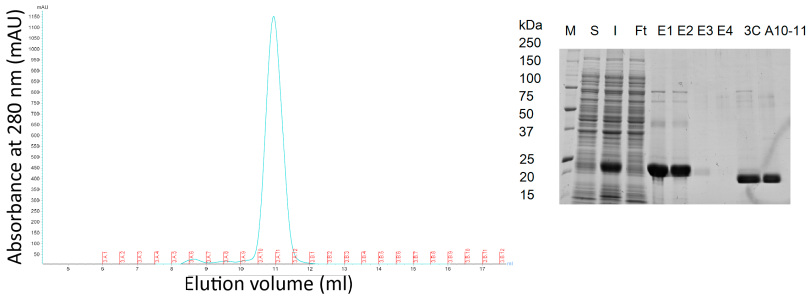


Figure 6.3. Size exclusion chromatography and SDS-PAGE analysis of SIAH1 SBD. SIAH1 SBD from cleared bacterial lysate was bound to a nickel column and eluted with 300 mM imidazole. Fractions were collected and 5 μl of sample was boiled with 2 X SDS sample buffer and subjected to electrophoresis on a 10% SDS-PAGE polyacrylamide gel. Samples from the bacterial lysate (Soluble), the cell pellet (Insoluble), the lysate collected after the nickel column (flow through), the elution steps (E1, E2, E3, E4), enzyme cleavage (3C) and fractions A10-A11 were analyzed alongside.

6.2.3 Isothermal Titration Calorimetry of SIAH1 SBD and TRIB3 328-340

The interaction between SIAH1 SBD and TRIB3 328-340 was measured several times at different concentrations. The concentration of SIAH1 SBD was checked with a Bradfford assay (see section 3.2.4. SIAH1 SBD was dialyzed in buffer 20 mM Tris at pH 8.0 10 mM NaCl before the ITC experiment, and the same buffer was used to prepare TRIB3 320-340 sample. The analysis software calculates the area of each peak by integrating the power (μcal/s) over time (s). To analyze the data it is necessary to do a baseline inspection and subtraction from the raw thermogram. Since we know the protein and peptide concentrations, the integrated heat of each injection is normalized by the molar concentration of the injected peptide versus the peptide-protein molar ratio (Figure 6.4). The isothermal titration calorimetry (ITC) analysis did not produce satisfactory results. According to the ITC titration, there was no heat change between the two constructs and it was not possible to calculate any thermodynamic parameter (Figure 6.4).

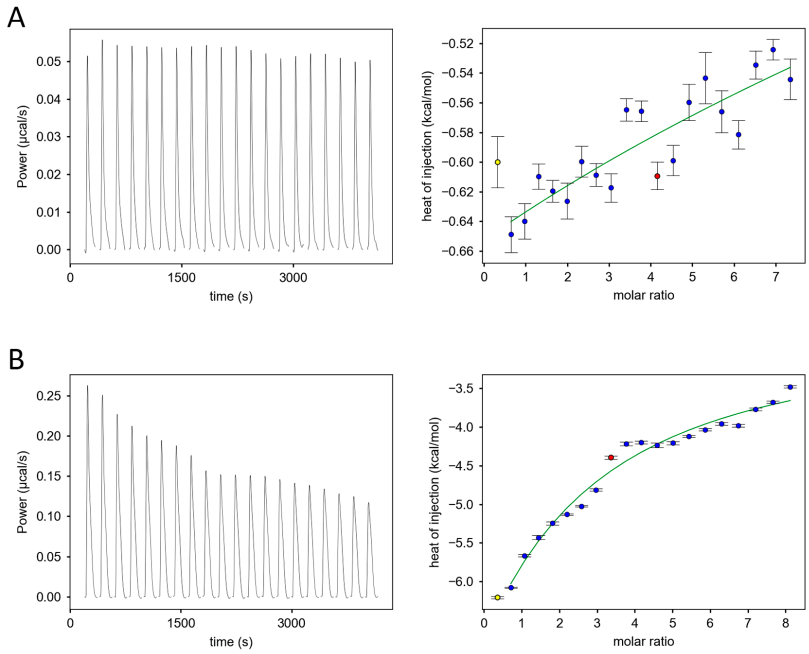


Figure 6.4. ITC analysis of TRIB3 328-340 and SIAH SB. A) 1200 μM TRIB3 328-340 injected into 50 μM SIAH1 SBD. B) 800 μM TRIB3 328-340 injected into 30 μM SIAH1 SBD. In both cases, the differential heating power versus time is on the left and the integrated and normalized heat of reaction versus the molar ratio is on the right. The thermograms are reconstructed by singular value decomposition (SVD), a computational method for peak-shape analysis. Yellow circles represent the first injection and heats excluded from the analysis. Red circles are the transition points defined by the software (NITPIC) and the green line is the fit to the isotherm.

7 Discussion

Proteins are essential elements for life, the cell contains thousands of different proteins carrying out most of the functions and contribution to cellular homeostasis. Therefore, understanding their role can facilitate the development of treatments and disease prevention. The regulation of proteins has been possible thanks to the progress made on finding drugs (i.e chemical compounds like small molecules) or biological products (e.g., hormones, vaccines, monoclonal antibodies, gene therapy) to target them. Unfortunately, not all the proteins are druggable and only a 4% of the human proteome is estimated to have a chemical probe [312]. Kinases are among the most targeted proteins in the drug discovery field. Interestingly, their homologous pseudokinases have remained much more elusive, probably due to the additional difficulty of targeting a protein with no catalytic function. However, dysregulation of pseudokinases is linked to several human pathologies such as cancer, diabetes, cardiovascular or even neurodegenerative diseases [126]. TRIB pseudokinases (TRIB1, TRIB2, and TRIB3) have emerged as an interesting target for several diseases and even resistance to anti-cancer therapy [313, 218]. Here, we sought to target the pseudokinase TRIB1 with small molecules to modulate its activity. Small molecules capable of modifying TRIB signaling might have utility as drugs in cancers or other diseases where TRIBs overexpression contributes to the maintenance or development of the disease. Moreover, the discovery of a chemical probe for this unique family of pseudoenzymes will help to better understand their activity. TRIB proteins function through two main mechanisms of action. This family of proteins lacks its catalytic activity and has been shown to serve as scaffolds. TRIB can interact with substrates using the same mechanism as kinases, given the similar fold, and control their E3 ligase ubiquitination thanks to a unique C-terminal tail. The second still needs to be elucidated on how TRIB regulates other kinases like AKT or MAPK through protein-protein interactions. In this thesis, we found new TRIB1 binding small molecules using computational and experimental techniques as described in chapter 4. First, thanks to the screening of the small molecule library PKIS by Jamieson and colleagues, we were able to identify diphenylurea as potential scaffold binding to TRIB1. Thereafter, we built a homology modeling to dock the structure into

the canonical ATP pocket of TRIB1. Then, we evaluated the binding mode through MD to obtain a more accurate prediction of the interaction between the small molecule GW683134A and TRIB1. We can see how the RMSD of the compound is around 1 Å with a small shift of 2 Å at the beginning of the simulation which indicates the rearrangement of the small molecule in the canonical ATP binding pocket of TRIB1 (Figure X). Thanks to that, we could optimize the pose and be more confident about the proposed binding mode, albeit taking account of the limits of the technique (e.g simulation time, parameterization). The binding of the compound resembles a type-II kinase inhibitor where the small molecule binds in the canonical ATP pocket in an inactive state (DFG-out) exploring the hydrophobic back pocket.

TRIB1 is a very particular case for investigating the binding of small molecules to the ATP binding site. This is because TRIB1 lacks key residues in the ATP pocket that bind to the nucleotide or cations in canonical kinases. In TRIB1, the SLE motif replaces the DGF motif and has a unique acidic glutamic residue (GLU224) preceding the SLE, which stabilizes the autoinhibited state of TRIB1. This residue is conserved in all TRIB pseudokinases. Moreover, the catalytic loop lacks the histidine of the HRD motif found in almost all kianses, which is important for the interaction with the metal ions. In addition, the hydrophobic back pocket contains several bulky and hydrophobic residues, like the gatekeeper PHE160, the TYR134, the LEU223 and the ILE131. In most of the type II inhibitors, there is usually a glutamic acid in the αC helix interacting with the aspartic acid from the DFG motif. Whereas in TRIB1 both residues are replaced by an isoleucine (ILE131) in the αC helix and the serine (SER225) from the SLE motif.

To find TRIB1 binding small molecules, we selected a library of compounds with similar scaffolds with respect to diphenyl urea to filter through virtual screening. We based the interaction on the proposed binding mode of GW683134A obtained through MD. Three specific pharmacophores were created to help the identification of new binding small molecules, two hydrogen bond donors (side chain of SER225 and the main chain amine group of ASP163) and one hydrogen bond acceptor (carbonyl group of LEU223). We selected compounds able to form a hydrogen bond with at least two of the three pharmacophores. The idea behind this selection is to find molecules able to interact with the SLE motif and/or the hinge region (ASP163) to have certain flexibility. As a result, we selected 41 compounds to test in a thermal shift assay, DSF (Figure X). Given the lack of TRIBs catalytic activity, DSF allows us

to measure thermal stabilization of compounds against TRIB1 and determine ligand binding abilities. Usually, if a compound can increase protein stability by at least 2 degrees is considered a positive hit. To further discriminate between the potential hits, the compounds were tested at several concentrations and in the presence or absence of C/EBPα. When TRIB1 binds C/EBPα, there is a conformational change on the canonical ATP binding pocket, and we wanted to inspect if any of the compounds have a TRIB1 conformational preference. Unfortunately, none of the compounds seems to have a choice of conformation but 10 out of 41 compounds were able to increase TRIB by at least 2 C (Figure X). DSF has the limitations of being a temperature dependence technique, which could interfere with the conformational changes of TRIB1 and therefore the binding of the small molecules. It is therefore highly important to determine that the thermal stabilization is dependent on compound concentration and not to artifacts. Consequently, we performed dose-response experiments to analyze the melting curves and to discard compounds not able to stabilize the protein at lower concentrations. The ideal melting curve is flat before the proteins start to unfold and then it has a regular sigmoidal shape for the unfolding. The most common issues are high background fluorescence and multiple transitions due to the temperature dependence technique giving unusually shaped curves.

That is why we tested the best compounds with SPR as an orthologous assay to validate their interaction and quantify their binding affinity. One of the most important advantages of SPR is the possibility to obtain thermodynamic (K_d) and kinetic (k_{on} and k_{off}) details of the binding event. To set-up the experiment, we used as controls the compounds from the DSF and the C/EBPα peptide. Moreover, trying to predict the binding mode of the compounds, we tested two TRIB1 variants with one mutation in the canonical ATP binding pocket, D163I and S225F separately. Both mutations were selected based on their localization and potential interaction with the proposed small molecules. In addition, the mutations were computationally predicted to not compromise the overall folding of TRIB1. When mutating an amino acid is important to discriminate that the effect produced is not affecting the overall protein folding. According to Mutatex, none of the selected mutations will affect the protein stability. Of course, we cannot ignore that those mutations will affect the conformational state of TRIB1 and therefore we expected to interfere with the binding of the small molecules but not C/EBPα since they are located at the ATP pocket. As shown in chapter 4, the C/EBPα peptide produced an unexpected sensorgram of negative values but the shape of the sensorgram

showed a normal trend with a progressive response during the injection of the peptide and a slow dissociation after the injection (Figure X). A dose dependent negative SPR signal has been proposed before to be predictive of the allosteric structural changes undergone during protein ligand interaction [314, 315]. All of this is in accordance with the crystallographic information of TRIB1 bound to C/EBPα and how there is a marked conformational change upon substrate binding. As expected, none of the TRIB1 mutants affected the binding of C/EBPα since they are not located at the binding region. Therefore, we could discard that the mutations affect the protein stability and the overall folding of TRIB1.

The SPR screening of the filtered compounds from DSF validated B1 as the best compound with an affinity of approximately 30 M. In addition, we tested some B1 analogs and discovered some selectivity over the substituents predicted to interact at the TRIB1 hydrophobic back pocket. The introduction of a methoxy at the ortho or para-substituent showed a decreased or complete abolition of binding affinity. This is in accordance with the proposed binding mode and how substituents at those positions cannot bind due to clashes with residues in the αC helix. We also tested some variants on the proposed solvent exposed region of the compounds, which could help in the design of more potent and specific hits. Interestingly, no signal was recorded when some analogs were tested which classified them as inactive analogs and negative controls. Unfortunately, none of the mutations gave significant difference which could be explained by the high structural flexibility of TRIB1, and the high plasticity observed in the canonical ATP pocket. Conformational dynamics of kinases are well known thanks to all the structural information obtained through X-ray or NMR experiments over the last decades. Although there is less structural evidence of the pseudokinase flexibility, still there are some examples where we can see this behavior. For example, it was described the structural and dynamic similarities of the pseudokinase domains from the Wnt-binding receptor tyrosine kinase with respect to canonical kinases [269]. Moreover, they identified type II inhibitors for ROR1 and even characterized the crystal structure. ROR1 is a pseudokinase that does not bind nucleotides or cations although it keeps the DFG motif. Other pseudokinases with unusual DFG motif but able to bind ATP have been also reported to interact with small molecules, like CASK as type I inhibitors (GFG-in) and STRADα as type II (GLR-out) [160, 195].

Although kinases are one of the most targeted proteins in drug discovery, many of the kinase inhibitors are promiscuous and often inhibit multiple ki-

nases in key signaling pathways. We can find examples of very promiscuous compounds targeting a substantial fraction of kinases as dasatinib or sunitinib [316]. However, in some cases inhibiting a high number of kinases resulted in improved antitumor activity. Unfortunately, the off targets usually result in toxicity which limits the dose and prevents the original target. That is why the characterization of kinase drug selectivity has become an important prerequisite to develop a kinase inhibitor [317].

Given the high similarity of kinases and pseudokinases it will be interesting to evaluate specificity between these families. Hence, we tested the specificity of the best hit, B1.2 against a panel of serine/threonine tyrosine, and lipid kinase from the University of Dundee (Figure X.) The analog B1.2 showed a greater binding affinity for TRIB1 with a K_d of 10 μM. Even though the panel contains 140 kinases from the 555 possible, only 11 kinases were inhibited at least more than 25% when tested with the compound B1.2 at 10 μM. These results offer a confident starting point to develop specific small molecules targeting TRIB1. Interestingly, the most inhibited kinase CHK2 has been already crystallized with a diphenylurea based small molecule. Of course, data from a single small panel is not enough for a robust measure of selectivity but it can be seen as a first, encouraging development. The diphenyl urea scaffold is very flexible on the binding pocket of kinases, being an interesting chemical moiety to explore for non-catalytic enzymes. Additionally, given the evidence that pseudokinases can interact with small molecules in a similar way to normal kinases, it will be intriguing to evaluate the potential side effects of kinases inhibitors over pseudokinases. Notably, the scaffolding function of kinases and pseudokinases could massively influence the clinical efficacy and even resistance of kinase inhibitor drugs.

Overall, these results showed that TRIB1 can bind small molecules and offers the possibility of designing a potent and selective TRIB1 ligand. Starting with an undruggable target and being able to find small molecules interacting with it should be encouraged scaling up chemical probes for the human proteome. Moreover, the development of a selective TRIB1 ligand will help to elucidate the mechanism of how this family of pseudokinases work and potentially lead to new treatments. The discovery of ligands for scaffold proteins could generate new strategies for modulating proteins. The promising power of targeting pseudoenzymes by modulating PPIs open an interesting field of possibilities where this approach can be applied. As in the case of PROTACs (Proteolysis Targeting Chimera) or molecular glues, which hijack E3 ligase to

induce ubiquitination and subsequent proteasomal degradation of a target protein. Small molecules able to induce target degradation is just the tip of the iceberg of the new chemical modalities that are about to come. For example, other post-translational modifications like phosphorylation have been shown to be feasible (e.g. PhosTAC) and will lead the way for new chemical modalities with huge potential [318].

How TRIB regulates other kinases like AKT/PKB, or MAPK through PPIs still need to be elucidated. As described in chapter 5, TRIB3 can regulate AKT/PKB phosphorylation in the liver although the specific mechanism is still unknown. The αC helix of AKT/PKB has been already proposed to interact with TRIB3 using Co-IP and SPR [ref]. We hypothesize that the C-terminal tail of TRIB3 can interact with the N-lobe of AKT and therefore modulate its activity. In collaboration with Prof. Eric Kalkhoven from UMC in Utrecht we tried to elucidate how these two proteins interact. The Protein Complementation Assay (PCA) was used to verify experimentally the interaction of TRIB3 and AKT1. Since PPIs are very context dependent, the PCA allows us to measure the interaction between protein complexes in an endogenous background. One of the disadvantages of the technique is to be sure that the interaction is driven by the proteins and not the presence of additional partners. Bearing that in mind, we selected as a reporter system of the interaction an engineered luciferase composed of a large bit (LgBit, 17.6 kDa) and a small bit (SmBit, 11 amino acids). The LgBit and SmBit are fused to our protein of interests (POIs) and when expressed, the interaction of the two POIs form a functional enzyme that generates a signal. The luciferase subunits interact very weakly (K_d = 190 μM) so the driven force of the interaction only occurs upon the binding of the POIs. The interaction of the LgBit and SmBit is reversible, therefore the system can be used to detect more transient or rapid dissociating proteins. One advantage of this technique is to measure the interaction inside the cell and even real-time analysis of protein interaction dynamics. In contrast, this assay needs time to optimize depending on the POIs to obtain a strong and reliable signal.

Although the interaction between TRIB3 and AKT/PKB is well known there is a lot of controversy on how these two proteins interact. TRIB3 has been shown to regulate the phosphorylation of Ser473 in AKT1/PKBα. That is why, we have been trying to elucidate the specific mechanism on how these two proteins can interact. First, we measured the interaction between full length AKT1/PKBα and TRIB3 including the kinase and pseudokinase domain respectively. We assume the main interaction of these two proteins is driven by

the kinase and pseudokinase domain, where the C-tail of TRIB3 is regulating the phosphorylation of AKT1/PKBα. To test our hypothesis, we measured the interaction of only the C-tail of TRIB3 in presence of AKT1 with and without the phosphomimic at Ser473. The mutation of Ser473 to Asp has been suggested to be a phospho-mimetic in Akt1 [319]. Our results showed how the phosphomimic Asp473 abolishes the interaction of the C-tail of TRIB3 with AKT. When Ser473 is phosphorylated the AKT1 C-terminal region interacts with the N lobe of the kinase domain, more specifically with the αC helix. Consistent with this idea, the interaction of TRIB3 with the AKT1 αC helix has been already validated through SPR [236]. These results indicate that the C-tail of TRIB3 could interact in the same region of the AKT C-tail when Ser473 is not phosphorylated. Then, a molecular docking study was carried out employing FlexPepDock from Rosetta to evaluate the interaction between TRIB3 and AKT/PKB. The aim is to visualize and analyze the interactions among the C-tail of TRIB3 and AKT. We decided to use the last part of the C-tail of TRIB3 (residues 331-358) since it is already known to bind other proteins (e.g E3 ligases) and is very flexible. Moreover, we already tested some mutations at the PxxP motif (321Pro-322Leu-323Ala-324Pro) in the C-terminal of TRIB3 without affecting the interaction with AKT1. The FlexPepDock ab-initio protocol assumes that the peptide is located at the vicinity of the binding site but does not assume anything about the initial peptide backbone conformation. We decided to split the last 28 residues of the C-tail of TRIB3 based on the limitations to successfully calculate larger peptides due to a large number of peptide conformations and degrees of freedom that need to be sampled. Indeed, the FlexPepDock protocol was benchmarked with peptides length varies between 5 and 13 amino acids, with up to 52 rotatable bonds [261].

From the three fragments of the C-tail of TRIB3, we found that the lowest models from the residues 338-352 converged more than the other two fragments. Among the 10 lowest energy clusters we found very similar structures which could correlate to a more precise binding mode. Interestingly, the COP1 binding motif (residues 334-338) showed a plausible binding mode but with less convergence in comparison with the residues 338-352. Moreover, there is a small motif of three glutamic acids (EEE) between the residues 338-352, which overlap among the lowest structures. That motif is also present in the residues 344-358 and the best model, the biggest cluster among the lowest models, has that motif with a very similar binding mode. An explanation is that the negatively charged groups form strong electrostatic interactions with a conserved basic region in the AKT N lobe. There is a basic patch of linker

residues (Lys142-His143-Arg144) in AKT where the interaction of phosphorylated Ser473 appears to help displace the AKT PH domain thereby activating AKT1 [275]. It is important to notice that due to the complexity of the system under study, we did not include the PH domain of AKT1/PKBα in the calculations. There is no crystal structure of full-length AKT without allosteric inhibitors or active AKT including the PH domain.

Our experimental and computational analysis help to explain the complex mechanism of interaction between these two proteins. Indeed, suggest a new potential role of TRIB3 C-tail that will lead to a better understanding of this pseudoenzyme and point to new questions. From a future perspective, it will be interesting to test mutations on the basic patch of AKT1 and how that could interfere with the scaffolding role of TRIBs. Moreover, the complementation reporter assay will allow us to screen small molecules or peptides to modulate the interaction and later the activity of AKT1.

TRIB3 interacts with the E3 ligase SIAH1, although the specific domains or regions are still undetermined. In chapter 6, we proposed that TRIB3 binds the Substrate Binding Domain (SBD) of SIAH1. This hypothesis is based on the similarity between a fragment in TRIB3 and a consensus motif of several crystallized SIAH1 inhibitors (synthetic peptides (PDB ids 4I7B, 4I7C, 4I7D) and protein (Calcyclin-binding protein (PDB id 2A25), USP19 (PDB id 4X3G), Axin1 (PDB id 5WZZ)) where all of them conserve a VxP binding motif. This motif is also present in the C-tail of TRIB3.

The homology modeling of SIAH1-TRIB3 328-340 was performed using several templates to address the flexibility of the SBD of SIAH1. The variations between the models are on the terminal parts of the peptide (residues HIS-328 and TRP-330 on the N-terminal part or the ASP-338, GLY-339, and LEU-340 on the C-terminal part of the TRIB3 fragment). The TRIB3 residues considered as main contributors of the SIAH1 binding domain, VAL-335, and PRO-337, have the same position in all the models. This is expected as the valine and proline are part of the binding motif, and the terminal parts of the peptides are more flexible. Unfortunately, when measuring heat transfer during ITC experiments of SIAH1 and TRIB3 328-240, we could not obtain any heat exchange during the experiments. We tried several conditions, taking as reference the ITC between SIAH1 and AXIN1 [265] but we could not get satisfactory results. We even tried to co-crystallize the TRIB3 peptide with the SIAH1 SBD using several conditions, mostly based on the crystals of SIAH1 SBD. One reason for the lack of interaction could be the size of the peptide since

the AXIN1 peptide used in a crystal structure is longer (20 residues) than the TRIB3 peptide tested (13 residues). Of course, we do not know how the isolated peptide will fold, and maybe the pseudokinase domain of TRIB3 is necessary for the interaction. The C-tail of TRIBs is very flexible to not even be able to have a determined 3D structure according to the X-ray structures of TRIB1. This will be in agreement with the dynamic and flexibility to interact with different proteins and have distinctive functions. Another important aspect is that the interaction of TRIB3 with SIAH1 was identified with Co-IP, which could detect complexes of several proteins rather than a binary interaction [297]. Therefore, it will be interesting to test the full length of TRIB3 along with other experimental assays (e.g., cell-based assay as PCA) as an alternative approach. TRIBs are very dynamic and complex to study

8 Conclusions

Pseudokinases have emerged as a new target in the field of drug discovery since it was identified that non enzymatic proteins can regulate and control normal physiology and disease. In this thesis I have described the work that I have done during my PhD with the challenge of finding small molecules for an undruggable target and shed light on their scaffolding roles.

We have been able to discover new small molecules interacting with TRIB1. We have used several techniques, showing the great advantage of computational methods to increase our possibilities of success. Moreover, we validated our assumptions using two biophysical assays. First, we applied DSF to find new compounds so later we confirmed with an orthogonal technique, SPR. We have calculated the K_d and tested several small molecules variants to verify our proposed binding mode. When developing small molecules for kinases it is important to always test off-targets due to their high similarity within the family. That is why we tested our best compound with a small panel of kinases to evaluate its specificity. Surprisingly, the compound seems to inhibit very few kinases. In this way, our findings lead the way towards a better strategy for the development of TRIBs ligands. This will lead to a better understanding of TRIBs function and potentially to even develop new treatments.

In addition, we have studied more in detail two interesting PPIs related to TRIB3. Since the interaction of TRIB3 with AKT1/PKBα was found in 2003, the interest to this family of proteins increased considerably. However, the specific mechanism of how TRIB3 can regulate the activity of AKT remains unknown. Here, we have proposed a potential binding mode between the TRIB3 C-tail and AKT1/PKBα which explains how this pseudokinase regulates AKT phosphorylation. On the other hand, the interaction between SIAH1 and TRIB3 was not possible to confirm but it will help in the design of future assays to validate TRIB3 interactions.

The contributions presented in this thesis have been possible thanks to fruitful collaborations. The challenges that a potential undruggable target offers are a great framework for interdisciplinary fields such as computational

and experimental to transform those targets from difficult to drug into expected to be drugged.

Bibliography

[1] Alchemy (n.). *Online Etymology Dictionary*. URL `https://www.etymonline.com/word/alchemy`.

[2] Protein (n.). *Online Etymology Dictionary*. URL `https://www.etymonline.com/word/protein`.

[3] Ian W Hamley. *Introduction to Peptide Science. (p.1)*. John Wiley & Sons, 2020.

[4] Jeremy M Berg, John L Tymoczko, Lubert Stryer, et al. Biochemistry. section 1.3, chemical bonds in biochemistry., 2002.

[5] Katharina Wendler, Jens Thar, Stefan Zahn, and Barbara Kirchner. Estimating the hydrogen bond energy. *The Journal of Physical Chemistry A*, 114(35):9529–9536, 2010.

[6] C Nick Pace, J Martin Scholtz, and Gerald R Grimsley. Forces stabilizing proteins. *FEBS letters*, 588(14):2177–2184, 2014.

[7] Peter E Wright and H Jane Dyson. Intrinsically disordered proteins in cellular signalling and regulation. *Nature reviews Molecular cell biology*, 16(1):18–29, 2015.

[8] Robin Van Der Lee, Marija Buljan, Benjamin Lang, Robert J Weatheritt, Gary W Daughdrill, A Keith Dunker, Monika Fuxreiter, Julian Gough, Joerg Gsponer, David T Jones, et al. Classification of intrinsically disordered regions and proteins. *Chemical reviews*, 114(13):6589–6631, 2014.

[9] Cyrus Chothia and Arthur M Lesk. The relation between the divergence of sequence and structure in proteins. *The EMBO journal*, 5(4):823–826, 1986.

[10] Kristoffer Illergård, David H Ardell, and Arne Elofsson. Structure is three to ten times more conserved than sequence—a study of structural response in protein cores. *Proteins: Structure, Function, and Bioinformatics*, 77(3):499–508, 2009.

[11] Robert Zwanzig, Attila Szabo, and Biman Bagchi. Levinthal's paradox. *Proceedings of the National Academy of Sciences*, 89(1):20–22, 1992.

[12] Helen M Berman, John Westbrook, Zukang Feng, Gary Gilliland, Talapady N Bhat, Helge Weissig, Ilya N Shindyalov, and Philip E Bourne. The protein data bank. *Nucleic acids research*, 28(1):235–242, 2000.

[13] Gerard J Kleywegt and T Alwyn Jones. [11] model building and refinement practice. *Methods in enzymology*, 277:208–230, 1997.

[14] Toshihiko Sugiki, Naohiro Kobayashi, and Toshimichi Fujiwara. Modern technologies of solution nuclear magnetic resonance spectroscopy for three-dimensional structure determination of proteins open avenues for life scientists. *Computational and structural biotechnology journal*, 15:328–339, 2017.

[15] Jean-Paul Renaud, Ashwin Chari, Claudio Ciferri, Wen-ti Liu, Hervé-William Rémigy, Holger Stark, and Christian Wiesmann. Cryo-em in drug discovery: achievements, limitations and prospects. *Nature reviews Drug discovery*, 17(7):471–492, 2018.

[16] Ewen Callaway. The protein-imaging technique taking over structural biology. *Nature*, 578(7794):201–201, 2020.

[17] Brian Kuhlman and Philip Bradley. Advances in protein structure prediction and design. *Nature Reviews Molecular Cell Biology*, 20(11):681–697, 2019.

[18] Burkhard Rost. Twilight zone of protein sequence alignments. *Protein engineering*, 12 (2):85–94, 1999.

[19] Andras Fiser. Template-based protein structure modeling. *Computational biology*, pages 73–94, 2010.

[20] Jooyoung Lee, Peter L Freddolino, and Yang Zhang. Ab initio protein structure prediction. In *From protein structure to function with bioinformatics*, pages 3–35. Springer, 2017.

[21] Andriy Kryshtafovych, Torsten Schwede, Maya Topf, Krzysztof Fidelis, and John Moult. Critical assessment of methods of protein structure prediction (casp)—round xiii. *Proteins: Structure, Function, and Bioinformatics*, 87(12):1011–1020, 2019.

[22] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.

[23] Minkyung Baek, Frank DiMaio, Ivan Anishchenko, Justas Dauparas, Sergey Ovchinnikov, Gyu Rie Lee, Jue Wang, Qian Cong, Lisa N Kinch, R Dustin Schaeffer, et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, 373(6557):871–876, 2021.

[24] Andrew W Senior, Richard Evans, John Jumper, James Kirkpatrick, Laurent Sifre, Tim Green, Chongli Qin, Augustin Žídek, Alexander WR Nelson, Alex Bridgland, et al. Improved protein structure prediction using potentials from deep learning. *Nature*, 577 (7792):706–710, 2020.

[25] Emil Fischer. Einfluss der configuration auf die wirkung der enzyme. *Berichte der deutschen chemischen Gesellschaft*, 27(3):2985–2993, 1894.

[26] DE Koshland Jr, RAY WJ Jr, and ERWIN MJ. Protein structure and enzyme action. In *Federation proceedings*, volume 17, pages 1145–1150, 1958.

[27] Peter Csermely, Robin Palotai, and Ruth Nussinov. Induced fit, conformational selection and independent dynamic segments: an extended view of binding events. *Nature Precedings*, pages 1–1, 2010.

[28] Xing Du, Yi Li, Yuan-Ling Xia, Shi-Meng Ai, Jing Liang, Peng Sang, Xing-Lai Ji, and Shu-Qun Liu. Insights into protein–ligand interactions: mechanisms, models, and methods. *International journal of molecular sciences*, 17(2):144, 2016.

[29] Tammy Nolan, Nidhi Singh, and Christopher R McCurdy. Ligand macromolecule interactions: theoretical principles of molecular recognition. *Ligand-Macromolecular Interactions in Drug Discovery*, pages 13–29, 2010.

[30] Remo Perozzo, Gerd Folkers, and Leonardo Scapozza. Thermodynamics of protein–ligand interactions: history, presence, and future aspects. *Journal of Receptors and Signal Transduction*, 24(1-2):1–52, 2004.

[31] David M Ford. Enthalpy- entropy compensation is not a general feature of weak association. *Journal of the American Chemical Society*, 127(46):16167–16170, 2005.

[32] Caterina Bissantz, Bernd Kuhn, and Martin Stahl. A medicinal chemist's guide to molecular interactions. *Journal of medicinal chemistry*, 53(14):5061–5084, 2010.

[33] Renato Ferreira de Freitas and Matthieu Schapira. A systematic analysis of atomic protein–ligand interactions in the pdb. *MedChemComm*, 8(10):1970–1981, 2017.

[34] Paul Muller. Glossary of terms used in physical organic chemistry (iupac recommendations 1994). *Pure and Applied Chemistry*, 66(5):1123, 1994.

[35] Andrew M Davis and Simon J Teague. Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. *Angewandte Chemie International Edition*, 38(6):736–749, 1999.

[36] Ian K McDonald and Janet M Thornton. Satisfying hydrogen bonding potential in proteins. *Journal of molecular biology*, 238(5):777–793, 1994.

[37] Sanjay Sarkhel and Gautam R Desiraju. N-H…O, O-H…O, and C-H…O hydrogen bonds in protein–ligand complexes: Strong and weak interactions in molecular recognition. *Proteins: Structure, Function, and Bioinformatics*, 54(2):247–259, 2004.

[38] Yuan Zhao, Jue Li, Hui Gu, Dongqing Wei, Yao-chang Xu, Wei Fu, and Zhengtian Yu. Conformational preferences of π–π stacking between ligand and protein, analysis derived from crystal structure data geometric preference of π–π interaction. *Interdisciplinary Sciences: Computational Life Sciences*, 7(3):211–220, 2015.

[39] Zachary S Hendsch and Bruce Tidor. Do salt bridges stabilize proteins? a continuum electrostatic analysis. *Protein Science*, 3(2):211–226, 1994.

[40] Michael Harder, Bernd Kuhn, and François Diederich. Efficient stacking on protein amide fragments. *ChemMedChem*, 8(3):397–404, 2013.

[41] Gabriella Cavallo, Pierangelo Metrangolo, Roberto Milani, Tullio Pilati, Arri Priimagi, Giuseppe Resnati, and Giancarlo Terraneo. The halogen bond. *Chemical reviews*, 116 (4):2478–2601, 2016.

[42] Tapan K Chaudhuri and Subhankar Paul. Protein-misfolding diseases and chaperone-based therapeutic approaches. *The FEBS journal*, 273(7):1331–1349, 2006.

[43] P Andrew Futreal, Lachlan Coin, Mhairi Marshall, Thomas Down, Timothy Hubbard, Richard Wooster, Nazneen Rahman, and Michael R Stratton. A census of human cancer genes. *Nature reviews cancer*, 4(3):177–183, 2004.

[44] Chris Finan, Anna Gaulton, Felix A Kruger, R Thomas Lumbers, Tina Shah, Jorgen Engmann, Luana Galver, Ryan Kelley, Anneli Karlsson, Rita Santos, et al. The druggable genome and support for target identification and validation in drug development. *Science translational medicine*, 9(383):eaag1166, 2017.

[45] Rita Santos, Oleg Ursu, Anna Gaulton, A Patrícia Bento, Ramesh S Donadi, Cristian G Bologa, Anneli Karlsson, Bissan Al-Lazikani, Anne Hersey, Tudor I Oprea, et al. A comprehensive map of molecular drug targets. *Nature reviews Drug discovery*, 16(1): 19–34, 2017.

[46] Andrew L Hopkins and Colin R Groom. The druggable genome. *Nature reviews Drug discovery*, 1(9):727–730, 2002.

[47] Andreas P Russ and Stefan Lampel. The druggable genome: an update. *Drug discovery today*, 10(23-24):1607–1610, 2005.

[48] FDA. Step 3: Clinical research. URL `https://www.fda.gov/patients/drug-development-process/step-3-clinical-research`. [Online; accessed December-2021].

[49] ChEMBL. Pubchem classification browser. URL `https://pubchem.ncbi.nlm.nih.gov/classification/#hid=72`. [Online; accessed December-2021].

[50] David Mendez, Anna Gaulton, A Patrícia Bento, Jon Chambers, Marleen De Veij, Eloy Félix, María Paula Magariños, Juan F Mosquera, Prudence Mutowo, Michał Nowotka, et al. Chembl: towards direct deposition of bioassay data. *Nucleic acids research*, 47 (D1):D930–D940, 2019.

[51] Favour Danladi Makurvet. Biologics vs. small molecules: Drug costs and patient access. *Medicine in Drug Discovery*, 9:100075, 2021.

[52] Alexey G Murzin, Steven E Brenner, Tim Hubbard, and Cyrus Chothia. Scop: a structural classification of proteins database for the investigation of sequences and structures. *Journal of molecular biology*, 247(4):536–540, 1995.

[53] Georgi K Kanev, Chris de Graaf, Iwan JP de Esch, Rob Leurs, Thomas Würdinger, Bart A Westerman, and Albert J Kooistra. The landscape of atypical and eukaryotic protein kinases. *Trends in pharmacological sciences*, 40(11):818–832, 2019.

[54] Daniel R Knighton, Jianhua Zheng, Lynn F Ten Eyck, Victor A Ashford, Nguyen-Huu Xuong, Susan S Taylor, and Janusz M Sowadski. Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science*, 253 (5018):407–414, 1991.

[55] Wolfram Hemmer, Maria McGlone, Igor Tsigelny, and Susan S Taylor. Role of the glycine triad in the atp-binding site of camp-dependent protein kinase. *Journal of Biological Chemistry*, 272(27):16946–16954, 1997.

[56] Susan S Taylor and Alexandr P Kornev. Protein kinases: evolution of dynamic regulatory proteins. *Trends in biochemical sciences*, 36(2):65–77, 2011.

[57] Lorenzo Palmieri and Giulio Rastelli. αc helix displacement as a general approach for allosteric modulation of protein kinases. *Drug Discovery Today*, 18(7-8):407–414, 2013.

[58] Jennifer E Kung and Natalia Jura. Structural basis for the non-catalytic functions of protein kinases. *Structure*, 24(1):7–24, 2016.

[59] Fabio Zuccotto, Elena Ardini, Elena Casale, and Mauro Angiolini. Through the "gatekeeper door": exploiting the active kinase conformation. *Journal of medicinal chemistry*, 53(7):2681–2694, 2010.

[60] Ali Torkamani and Nicholas J Schork. Prediction of cancer driver mutations in protein kinases. *Cancer research*, 68(6):1675–1682, 2008.

[61] Vishakha Singh, Mahendra Ram, Rajesh Kumar, Raju Prasad, Birendra Kumar Roy, and Kaushal Kumar Singh. Phosphorylation: implications in cancer. *The protein journal*, 36(1):1–6, 2017.

[62] Philip Cohen and Dario R Alessi. Kinase drug discovery–what's next in the field? *ACS chemical biology*, 8(1):96–104, 2013.

[63] Doriano Fabbro, Sandra W Cowan-Jacob, and Henrik Moebitz. Ten things you should know about protein kinases: Iuphar r eview 14. *British journal of pharmacology*, 172 (11):2675–2700, 2015.

[64] Philip Cohen, Darren Cross, and Pasi A Jänne. Kinase drug discovery 20 years after imatinib: Progress and future directions. *Nature reviews drug discovery*, 20(7):551–569, 2021.

[65] Gerard Manning, David B Whyte, Ricardo Martinez, Tony Hunter, and Sucha Sudarsanam. The protein kinase complement of the human genome. *Science*, 298(5600): 1912–1934, 2002.

[66] Misty M Attwood, Doriano Fabbro, Aleksandr V Sokolov, Stefan Knapp, and Helgi B Schiöth. Trends in kinase drug discovery: Targets, indications and inhibitor design. *Nature Reviews Drug Discovery*, 20(11):839–861, 2021.

[67] Robert Roskoski Jr. Properties of fda-approved small molecule protein kinase inhibitors. *Pharmacological research*, 144:19–50, 2019.

[68] Hong Yang, Brian Higgins, Kenneth Kolinsky, Kathryn Packman, Zenaida Go, Raman Iyer, Stanley Kolis, Sylvia Zhao, Richard Lee, Joseph F Grippo, et al. Rg7204 (plx4032), a selective brafv600e inhibitor, displays potent antitumor activity in preclinical melanoma models. *Cancer research*, 70(13):5518–5527, 2010.

[69] Robert Roskoski Jr. Classification of small molecule protein kinase inhibitors based upon the structures of their drug-enzyme complexes. *Pharmacological research*, 103: 26–48, 2016.

[70] Zheng Zhao, Hong Wu, Li Wang, Yi Liu, Stefan Knapp, Qingsong Liu, and Nathanael S Gray. Exploration of type ii binding mode: A privileged approach for kinase inhibitor focused drug discovery? *ACS chemical biology*, 9(6):1230–1241, 2014.

[71] Robert Roskoski Jr. Properties of fda-approved small molecule protein kinase inhibitors: A 2021 update. *Pharmacological research*, 165:105463, 2021.

[72] Connie S Lebakken, Laurie J Reichling, Jason M Ellefson, and Steven M Riddle. Detection of allosteric kinase inhibitors by displacement of active site probes. *Journal of Biomolecular Screening*, 17(6):813–821, 2012.

[73] Francisco J Adrián, Qiang Ding, Taebo Sim, Anastasia Velentza, Christine Sloan, Yi Liu, Guobao Zhang, Wooyoung Hur, Sheng Ding, Paul Manley, et al. Allosteric inhibitors of bcr-abl–dependent cell proliferation. *Nature chemical biology*, 2(2):95–102, 2006.

[74] Kenneth M Comess, Chaohong Sun, Cele Abad-Zapatero, Eric R Goedken, Rebecca J Gum, David W Borhani, Maria Argiriadi, Duncan R Groebe, Yong Jia, Jill E Clampit, et al. Discovery and characterization of non-atp site inhibitors of the mitogen activated protein (map) kinases. *ACS Chemical Biology*, 6(3):234–244, 2011.

[75] Wen-I Wu, Walter C Voegtli, Hillary L Sturgis, Faith P Dizon, Guy PA Vigers, and Barbara J Brandhuber. Crystal structure of human akt1 with an allosteric inhibitor reveals a new mode of kinase inhibition. *PloS one*, 5(9):e12913, 2010.

[76] T Justin Rettenmaier, Jack D Sadowsky, Nathan D Thomsen, Steven C Chen, Allison K Doak, Michelle R Arkin, and James A Wells. A small-molecule mimic of a peptide docking motif inhibits the protein kinase pdk1. *Proceedings of the National Academy of Sciences*, 111(52):18590–18595, 2014.

[77] Stephane Betzi, Riazul Alam, Mathew Martin, Donna J Lubbers, Huijong Han, Sudhakar R Jakkaraj, Gunda I Georg, and Ernst Schönbrunn. Discovery of a potential allosteric ligand binding site in cdk2. *ACS chemical biology*, 6(5):492–501, 2011.

[78] Kurt J Cox, Carolyn D Shomin, and Indraneel Ghosh. Tinkering outside the kinase atp box: allosteric (type iv) and bivalent (type v) inhibitors of protein kinases. *Future medicinal chemistry*, 3(1):29–43, 2011.

[79] Carrie M Gower, Matthew EK Chang, and Dustin J Maly. Bivalent inhibitors of protein kinases. *Critical reviews in biochemistry and molecular biology*, 49(2):102–115, 2014.

[80] A Ricouart, JC Gesquiere, A Tartar, and C Sergheraert. Design of potent protein kinases inhibitors using the bisubstrate approach. *Journal of medicinal chemistry*, 34 (1):73–78, 1991.

[81] John L Stebbins, Surya K De, Petra Pavlickova, Vida Chen, Thomas Machleidt, Li-Hsing Chen, Christian Kuntzen, Shinichi Kitada, Michael Karin, and Maurizio Pellecchia. Design and characterization of a potent and selective dual atp-and substrate-competitive subnanomolar bidentate c-jun n-terminal kinase (jnk) inhibitor. *Journal of medicinal chemistry*, 54(18):6206–6214, 2011.

[82] Seungbeom Lee, Jisu Kim, Jeyun Jo, Jae Won Chang, Jaehoon Sim, and Hwayoung Yun. Recent advances in development of hetero-bivalent kinase inhibitors. *European Journal of Medicinal Chemistry*, 216:113318, 2021.

[83] Phillip A Schwartz and Brion W Murray. Protein kinase biochemistry and drug discovery. *Bioorganic chemistry*, 39(5-6):192–210, 2011.

[84] Péter Ábrányi-Balogh, László Petri, Tímea Imre, Péter Szijj, Andrea Scarpino, Martina Hrast, Ana Mitrović, Urša Pečar Fonovič, Krisztina Németh, Hélène Barreteau, et al. A road map for prioritizing warheads for cysteine targeting covalent inhibitors. *European journal of medicinal chemistry*, 160:94–107, 2018.

[85] Iana M Serafimova, Miles A Pufall, Shyam Krishnan, Katarzyna Duda, Michael S Cohen, Rebecca L Maglathlin, Jesse M McFarland, Rand M Miller, Morten Frödin, and Jack Taunton. Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles. *Nature chemical biology*, 8(5):471–476, 2012.

[86] Kristine Senkane, Ekaterina V Vinogradova, Radu M Suciu, Vincent M Crowley, Balyn W Zaro, J Michael Bradshaw, Ken A Brameld, and Benjamin F Cravatt. The proteome-wide potential for reversible covalency at cysteine. *Angewandte Chemie*, 131 (33):11507–11511, 2019.

[87] Ayah Abdeldayem, Yasir S Raouf, Stefan N Constantinescu, Richard Moriggl, and Patrick T Gunning. Advances in covalent kinase inhibitors. *Chemical Society Reviews*, 49(9):2617–2687, 2020.

[88] James Inglese, Ronald L Johnson, Anton Simeonov, Menghang Xia, Wei Zheng, Christopher P Austin, and Douglas S Auld. High-throughput screening assays for the identification of chemical probes. *Nature chemical biology*, 3(8):466–479, 2007.

[89] Gabriel M Simon, Micah J Niphakis, and Benjamin F Cravatt. Determining target engagement in living systems. *Nature chemical biology*, 9(4):200–205, 2013.

[90] Marc Schürmann, Petra Janning, Slava Ziegler, and Herbert Waldmann. Small-molecule target engagement in cells. *Cell chemical biology*, 23(4):435–441, 2016.

[91] Frank H Niesen, Helena Berglund, and Masoud Vedadi. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature protocols*, 2(9):2212–2221, 2007.

[92] Michael W Pantoliano, Eugene C Petrella, Joseph D Kwasnoski, Victor S Lobanov, James Myslik, Edward Graf, Ted Carver, Eric Asel, Barry A Springer, Pamela Lane, et al. High-density miniaturized thermal shift assays as a general strategy for drug discovery. *Journal of biomolecular screening*, 6(6):429–440, 2001.

[93] Travis T Waldron and Kenneth P Murphy. Stabilization of proteins by ligand binding: application to drug screening and determination of unfolding energetics. *Biochemistry*, 42(17):5058–5064, 2003.

[94] Nan Bai, Heinrich Roder, Alex Dickson, and John Karanicolas. Isothermal analysis of thermofluor data can readily provide quantitative binding affinities. *Scientific reports*, 9(1):1–15, 2019.

[95] Alessio Ciulli. Biophysical screening for the discovery of small-molecule ligands. In *Protein-Ligand Interactions*, pages 357–388. Springer, 2013.

[96] Kai Gao, Rick Oerlemans, and Matthew R Groves. Theory and applications of differential scanning fluorimetry in early-stage drug discovery. *Biophysical reviews*, 12(1): 85–104, 2020.

[97] Biacore, A. B. Sensor Surface Handbook. *BR-1005-71 Edition AB*, 2005.

[98] Stephanie Leavitt and Ernesto Freire. Direct measurement of protein binding energetics by isothermal titration calorimetry. *Current opinion in structural biology*, 11(5):560–566, 2001.

[99] Luminita Damian. Isothermal titration calorimetry for studying protein–ligand interactions. In *Protein-ligand interactions*, pages 103–118. Springer, 2013.

[100] Jakub Stefaniak and Kilian VM Huber. Importance of quantifying drug-target engagement in cells. *ACS Medicinal Chemistry Letters*, 11(4):403–406, 2020.

[101] Kilian VM Huber, Karin M Olek, André C Müller, Chris Soon Heng Tan, Keiryn L Bennett, Jacques Colinge, and Giulio Superti-Furga. Proteome-wide drug and metabolite interaction mapping by thermal-stability profiling. *Nature methods*, 12(11):1055–1057, 2015.

[102] Gregory Sliwoski, Sandeepkumar Kothiwale, Jens Meiler, and Edward W Lowe. Computational methods in drug discovery. *Pharmacological reviews*, 66(1):334–395, 2014.

[103] Veronica Salmaso and Stefano Moro. Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: An overview. *Frontiers in pharmacology*, page 923, 2018.

[104] Douglas B Kitchen, Hélène Decornez, John R Furr, and Jürgen Bajorath. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature reviews Drug discovery*, 3(11):935–949, 2004.

[105] Leonardo G Ferreira, Ricardo N Dos Santos, Glaucius Oliva, and Adriano D Andricopulo. Molecular docking and structure-based drug design strategies. *Molecules*, 20 (7):13384–13421, 2015.

[106] Marco De Vivo, Matteo Masetti, Giovanni Bottegoni, and Andrea Cavalli. Role of molecular dynamics and related methods in drug discovery. *Journal of medicinal chemistry*, 59(9):4035–4061, 2016.

[107] Michael Levitt, Miriam Hirshberg, Ruth Sharon, and Valerie Daggett. Potential energy function and parameters for simulations of the molecular dynamics of proteins and nucleic acids in solution. *Computer physics communications*, 91(1-3):215–231, 1995.

[108] Jacob D Durrant and J Andrew McCammon. Molecular dynamics simulations and drug discovery. *BMC biology*, 9(1):1–9, 2011.

[109] Wendy D Cornell, Piotr Cieplak, Christopher I Bayly, Ian R Gould, Kenneth M Merz, David M Ferguson, David C Spellmeyer, Thomas Fox, James W Caldwell, and Peter A Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, 117(19):5179–5197, 1995.

[110] Bernard R Brooks, Robert E Bruccoleri, Barry D Olafson, David J States, S a Swaminathan, and Martin Karplus. Charmm: a program for macromolecular energy, minimization, and dynamics calculations. *Journal of computational chemistry*, 4(2):187–217, 1983.

[111] Markus Christen, Philippe H Hünenberger, Dirk Bakowies, Riccardo Baron, Roland Bürgi, Daan P Geerke, Tim N Heinz, Mika A Kastenholz, Vincent Kräutler, Chris Oostenbrink, et al. The gromos software for biomolecular simulation: Gromos05. *Journal of computational chemistry*, 26(16):1719–1751, 2005.

[112] Junmei Wang, Romain M Wolf, James W Caldwell, Peter A Kollman, and David A Case. Development and testing of a general amber force field. *Journal of computational chemistry*, 25(9):1157–1174, 2004.

[113] Rafael C Bernardi, Marcelo CR Melo, and Klaus Schulten. Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1850(5):872–877, 2015.

[114] Sebastian Kmiecik, Dominik Gront, Michal Kolinski, Lukasz Wieteska, Aleksandra Elzbieta Dawid, and Andrzej Kolinski. Coarse-grained protein models and their applications. *Chemical reviews*, 116(14):7898–7936, 2016.

[115] Johannes Kästner. Umbrella sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(6):932–942, 2011.

[116] Yuji Sugita and Yuko Okamoto. Replica-exchange molecular dynamics method for protein folding. *Chemical physics letters*, 314(1-2):141–151, 1999.

[117] Alessandro Barducci, Massimiliano Bonomi, and Michele Parrinello. Metadynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(5):826–843, 2011.

[118] Donald Hamelberg, John Mongan, and J Andrew McCammon. Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *The Journal of chemical physics*, 120(24):11919–11929, 2004.

[119] Barry Isralewitz, Mu Gao, and Klaus Schulten. Steered molecular dynamics and mechanical functions of proteins. *Current opinion in structural biology*, 11(2):224–230, 2001.

[120] Jérôme Boudeau, Diego Miranda-Saavedra, Geoffrey J Barton, and Dario R Alessi. Emerging roles of pseudokinases. *Trends in cell biology*, 16(9):443–452, 2006.

[121] James A McCormick and David H Ellison. The wnks: atypical protein kinases with pleiotropic actions. *Physiological reviews*, 91(1):177–219, 2011.

[122] Xiaoshan Min, Byung-Hoon Lee, Melanie H Cobb, and Elizabeth J Goldsmith. Crystal structure of the kinase domain of wnk1, a kinase that causes a hereditary form of hypertension. *Structure*, 12(7):1303–1311, 2004.

[123] Saskia JE Suijkerbuijk, Teunis JP van Dam, G Elif Karagöz, Eleonore von Castelmur, Nina C Hubner, Afonso MS Duarte, Mathijs Vleugel, Anastassis Perrakis, Stefan GD Rüdiger, Berend Snel, et al. The vertebrate mitotic checkpoint protein bubr1 is an unusual pseudokinase. *Developmental cell*, 22(6):1321–1329, 2012.

[124] James M Murphy, Qingwei Zhang, Samuel N Young, Michael L Reese, Fiona P Bailey, Patrick A Eyers, Daniela Ungureanu, Henrik Hammaren, Olli Silvennoinen, Leila N Varghese, et al. A robust methodology to subclassify pseudokinases based on their nucleotide-binding properties. *Biochemical Journal*, 457(2):323–334, 2014.

[125] Annie Kwon, Steven Scott, Rahil Taujale, Wayland Yeung, Krys J Kochut, Patrick A Eyers, and Natarajan Kannan. Tracing the origin and evolution of pseudokinases across the tree of life. *Science signaling*, 12(578):eaav3810, 2019.

[126] Veronika Reiterer, Patrick A Eyers, and Hesso Farhan. Day of the dead: pseudokinases and pseudophosphatases in physiology and disease. *Trends in cell biology*, 24(9):489–505, 2014.

[127] Hugo Lavoie, John J Li, Neroshan Thevakumaran, Marc Therrien, and Frank Sicheri. Dimerization-induced allostery in protein kinase regulation. *Trends in biochemical sciences*, 39(10):475–486, 2014.

[128] Aaron B Edmund, Timothy F Walseth, Nicholas M Levinson, and Lincoln R Potter. The pseudokinase domains of guanylyl cyclase–a and–b allosterically increase the affinity of their catalytic domains for substrate. *Science signaling*, 12(566):eaau5378, 2019.

[129] Tae-Hong Kang and Kyong-Tai Kim. Negative regulation of erk activity by vrk3-mediated activation of vhr phosphatase. *Nature cell biology*, 8(8):863–869, 2006.

[130] James M Murphy, Peter E Czabotar, Joanne M Hildebrand, Isabelle S Lucet, Jian-Guo Zhang, Silvia Alvarez-Diaz, Rowena Lewis, Najoua Lalaoui, Donald Metcalf, Andrew I Webb, et al. The pseudokinase mlkl mediates necroptosis via a molecular switch mechanism. *Immunity*, 39(3):443–453, 2013.

[131] AV Jacobsen, KN Lowes, MC Tanzer, IS Lucet, JM Hildebrand, EJ Petrie, MF Van Delft, Z Liu, SA Conos, JG Zhang, et al. Hsp90 activity is required for mlkl oligomerisation and membrane translocation and the induction of necroptotic cell death. *Cell death & disease*, 7(1):e2051–e2051, 2016.

[132] Priya H Dedhia, Karen Keeshan, Sacha Uljon, Lanwei Xu, Maria E Vega, Olga Shestova, Meirav Zaks-Zilberman, Candice Romany, Stephen C Blacklow, and Warren S Pear. Differential ability of tribbles family members to promote degradation of c/ebpα and induce acute myelogenous leukemia. *Blood, The Journal of the American Society of Hematology*, 116(8):1321–1328, 2010.

[133] Jennifer E Kung and Natalia Jura. The pseudokinase trib 1 toggles an intramolecular switch to regulate cop 1 nuclear export. *The EMBO journal*, 38(4):e99708, 2019.

[134] Dominic P Byrne, Daniel M Foulkes, and Patrick A Eyers. Pseudokinases: update on their functions and evaluation as new drug targets. *Future medicinal chemistry*, 9(2): 245–265, 2017.

[135] Jennifer E Kung and Natalia Jura. Prospects for pharmacological targeting of pseudokinases. *Nature Reviews Drug Discovery*, 18(7):501–526, 2019.

[136] Arian Laurence, Marko Pesu, Olli Silvennoinen, and John O'Shea. Suppl 2: Jak kinases in health and disease: An update. *The open rheumatology journal*, 6:232, 2012.

[137] Juuli Raivola, Henrik M Hammarén, Anniina T Virtanen, Vilasha Bulleeraz, Alister C Ward, and Olli Silvennoinen. Hyperactivation of oncogenic jak3 mutants depend on atp binding to the pseudokinase domain. *Frontiers in oncology*, 8:560, 2018.

[138] Daniela Ungureanu, Jinhua Wu, Tuija Pekkala, Yashavanthi Niranjan, Clifford Young, Ole N Jensen, Chong-Feng Xu, Thomas A Neubert, Radek C Skoda, Stevan R Hubbard, et al. The pseudokinase domain of jak2 is a dual-specificity protein kinase that negatively regulates cytokine signaling. *Nature structural & molecular biology*, 18(9): 971–976, 2011.

[139] Juuli Raivola, Teemu Haikarainen, Bobin George Abraham, and Olli Silvennoinen. Janus kinases in leukemia. *Cancers*, 13(4):800, 2021.

[140] E Joanna Baxter, Linda M Scott, Peter J Campbell, Clare East, Nasios Fourouclas, Soheila Swanton, George S Vassiliou, Anthony J Bench, Elaine M Boyd, Natasha Curtin, et al. Acquired mutation of the tyrosine kinase jak2 in human myeloproliferative disorders. *The Lancet*, 365(9464):1054–1061, 2005.

[141] Rajintha M Bandaranayake, Daniela Ungureanu, Yibing Shan, David E Shaw, Olli Silvennoinen, and Stevan R Hubbard. Crystal structures of the jak2 pseudokinase domain and the pathogenic mutant v617f. *Nature structural & molecular biology*, 19 (8):754–759, 2012.

[142] Stephen T Wrobleski, Ryan Moslin, Shuqun Lin, Yanlei Zhang, Steven Spergel, James Kempson, John S Tokarski, Joann Strnad, Adriana Zupa-Fernandez, Lihong Cheng, et al. Highly selective inhibition of tyrosine kinase 2 (tyk2) for the treatment of autoimmune diseases: discovery of the allosteric inhibitor bms-986165. *Journal of medicinal chemistry*, 62(20):8973–8995, 2019.

[143] Yu Chang, Shilin Xu, and Ke Ding. Tyrosine kinase 2 (tyk2) allosteric inhibitors to treat autoimmune diseases, 2019.

[144] Susan L Sierke, Kunrong Cheng, Hong-Hee Kim, and John G Koland. Biochemical characterization of the protein tyrosine kinase homology domain of the erbb3 (her3) receptor protein. *Biochemical Journal*, 322(3):757–763, 1997.

[145] Atsushi Suenaga, Naoki Takada, Mariko Hatakeyama, Mio Ichikawa, Xiaomei Yu, Kentaro Tomii, Noriaki Okimoto, Noriyuki Futatsugi, Tetsu Narumi, Mikako Shirouzu, et al. Novel mechanism of interaction of p85 subunit of phosphatidylinositol 3-kinase and erbb3 receptor-derived phosphotyrosyl peptides. *Journal of Biological Chemistry*, 280(2):1321–1326, 2005.

[146] Gerta Hoxhaj and Brendan D Manning. The pi3k–akt network at the interface of oncogenic signalling and cancer metabolism. *Nature Reviews Cancer*, 20(2):74–88, 2020.

[147] Rosalin Mishra, Hima Patel, Samar Alanazi, Long Yuan, and Joan T Garrett. Her3 signaling and targeted therapy in cancer. *Oncology reviews*, 12(1), 2018.

[148] Ningyan Zhang, Yujun Chang, Adan Rios, and Zhiqiang An. Her3/erbb3, an emerging cancer therapeutic target. *Acta biochimica et biophysica Sinica*, 48(1):39–48, 2016.

[149] Wolfgang Jacob, Ian James, Max Hasmann, and Martin Weisser. Clinical development of her3-targeting monoclonal antibodies: perils and progress. *Cancer treatment reviews*, 68:111–123, 2018.

[150] Ashton C Lai and Craig M Crews. Induced protein degradation: an emerging drug discovery paradigm. *Nature reviews Drug discovery*, 16(2):101–114, 2017.

[151] Ting Xie, Sang Min Lim, Kenneth D Westover, Michael E Dodge, Dalia Ercan, Scott B Ficarro, Durga Udayakumar, Deepak Gurbani, Hyun Seop Tae, Steven M Riddle, et al. Pharmacological targeting of the pseudokinase her3. *Nature chemical biology*, 10(12): 1006–1012, 2014.

[152] Emma J Petrie, Jarrod J Sandow, Annette V Jacobsen, Brian J Smith, Michael DW Griffin, Isabelle S Lucet, Weiwen Dai, Samuel N Young, Maria C Tanzer, Ahmad Wardak, et al. Conformational switching of the pseudokinase domain promotes human mlkl tetramerization and cell death by necroptosis. *Nature communications*, 9(1):1–15, 2018.

[153] Joanne M Hildebrand, Maria C Tanzer, Isabelle S Lucet, Samuel N Young, Sukhdeep K Spall, Pooja Sharma, Catia Pierotti, Jean-Marc Garnier, Renwick CJ Dobson, Andrew I Webb, et al. Activation of the pseudokinase mlkl unleashes the four-helix bundle domain to induce membrane localization and necroptotic cell death. *Proceedings of the National Academy of Sciences*, 111(42):15072–15077, 2014.

[154] Andreas Linkermann and Douglas R Green. Necroptosis. *New England Journal of Medicine*, 370(5):455–465, 2014.

[155] Bin Ma, Doug Marcotte, Murugan Paramasivam, Klaus Michelsen, Ti Wang, Andrea Bertolotti-Ciarlet, John Howard Jones, Ben Moree, Margaret Butko, Joshua Salafsky, et al. Atp-competitive mlkl binders have no functional impact on necroptosis. *PLoS One*, 11(11):e0165983, 2016.

[156] Bo Yan, Lei Liu, Shaoqiang Huang, Yan Ren, Huayi Wang, Zhenglin Yao, Lin Li, She Chen, Xiaodong Wang, and Zhiyuan Zhang. Discovery of a new class of highly potent necroptosis inhibitors targeting the mixed lineage kinase domain-like protein. *Chemical Communications*, 53(26):3637–3640, 2017.

[157] Elton Zeqiraj, Beatrice Maria Filippi, Simon Goldie, Iva Navratilova, Jérôme Boudeau, Maria Deak, Dario R Alessi, and Daan MF van Aalten. Atp and mo25α regulate the conformational state of the stradα pseudokinase and activation of the lkb1 tumour suppressor. *PLoS biology*, 7(6):e1000126, 2009.

[158] Elton Zeqiraj, Beatrice Maria Filippi, Maria Deak, Dario R Alessi, and Daan MF Van Aalten. Structure of the lkb1-strad-mo25 complex reveals an allosteric mechanism of kinase activation. *Science*, 326(5960):1707–1711, 2009.

[159] Jérôme Boudeau, Annette F Baas, Maria Deak, Nick A Morrice, Agnieszka Kieloch, Mike Schutkowski, Alan R Prescott, Hans C Clevers, and Dario R Alessi. Mo25α/β interact with stradα/β enhancing their ability to bind, activate and localize lkb1 in the cytoplasm. *The EMBO journal*, 22(19):5102–5114, 2003.

[160] Ryan HB Smith, Zaigham M Khan, Peter Man-Un Ung, Alex P Scopton, Lisa Silber, Seshat M Mack, Alexander M Real, Avner Schlessinger, and Arvin C Dar. Type ii binders targeting the "glr-out" conformation of the pseudokinase stradα. *Biochemistry*, 60(4):289–302, 2021.

[161] Jennifer Green, Roel Nusse, and Renée van Amerongen. The role of ryk and ror receptor tyrosine kinases in wnt signal transduction. *Cold Spring Harbor Perspectives in Biology*, 6(2):a009175, 2014.

[162] T Zhan, N Rindtorff, and Michael Boutros. Wnt signaling in cancer. *Oncogene*, 36 (11):1461–1473, 2017.

[163] Hanna Karvonen, Robert Perttilä, Wilhelmiina Niininen, Harlan Barker, and Daniela Ungureanu. Targeting wnt signaling pseudokinases in hematological cancers. *European Journal of Haematology*, 101(4):457–465, 2018.

[164] Alessandra Gentile, Luca Lazzari, Silvia Benvenuti, Livio Trusolino, and Paolo M Comoglio. Ror1 is a pseudokinase that is crucial for met-driven tumorigenesis. *Cancer research*, 71(8):3132–3141, 2011.

[165] Tomoya Yamaguchi, Kiyoshi Yanagisawa, Ryoji Sugiyama, Yasuyuki Hosono, Yukako Shimada, Chinatsu Arima, Seiichi Kato, Shuta Tomida, Motoshi Suzuki, Hirotaka Osada, et al. Nkx2-1/titf1/ttf-1-induced ror1 is required to sustain egfr survival signaling in lung adenocarcinoma. *Cancer cell*, 21(3):348–361, 2012.

[166] Kazuhito Morioka, Chizu Tanikawa, Kensuke Ochi, Yataro Daigo, Toyomasa Katagiri, Hirotaka Kawano, Hiroshi Kawaguchi, Akira Myoui, Hideki Yoshikawa, Norifumi Naka, et al. Orphan receptor tyrosine kinase ror2 as a potential therapeutic target for osteosarcoma. *Cancer science*, 100(7):1227–1233, 2009.

[167] Tricia M Wright, A Rose Brannon, John D Gordan, Amanda J Mikels, Cicely Mitchell, Shufen Chen, Inigo Espinosa, Matt van de Rijn, Raj Pruthi, Eric Wallen, et al. Ror2, a developmentally regulated kinase, promotes tumor growth potential in renal cell carcinoma. *Oncogene*, 28(27):2513–2523, 2009.

[168] Zufan Debebe and W Kimryn Rathmell. Ror2 as a therapeutic target in cancer. *Pharmacology & therapeutics*, 150:143–148, 2015.

[169] Ron Chen, Purvesh Khatri, Pawel K Mazur, Melanie Polin, Yanyan Zheng, Dedeepya Vaka, Chuong D Hoang, Joseph Shrager, Yue Xu, Silvestre Vicent, et al. A meta-analysis of lung cancer gene expression identifies ptk7 as a survival gene in lung adenocarcinoma. *Cancer research*, 74(10):2892–2902, 2014.

[170] Silvia Gärtner, Angela Gunesch, Tatiana Knyazeva, Petra Wolf, Bernhard Högel, Wolfgang Eiermann, Axel Ullrich, Pjotr Knyazev, and Beyhan Ataseven. Ptk 7 is a transforming gene and prognostic marker for breast cancer and nodal metastasis involvement. *PloS one*, 9(1):e84472, 2014.

[171] Yongan Fu, Yilin Chen, Jinghua Huang, Zongda Cai, and Yangqiang Wang. Ryk, a receptor of noncanonical wnt ligand wnt5a, is positively correlated with gastric cancer tumorigenesis and potential of liver metastasis. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 318(2):G352–G360, 2020.

[172] Michael Y Choi, George F Widhopf II, Christina CN Wu, Bing Cui, Fitzgerald Lao, Anil Sadarangani, Joy Cavagnaro, Charles Prussak, Dennis A Carson, Catriona Jamieson, et al. Pre-clinical specificity and safety of uc-961, a first-in-class monoclonal antibody targeting ror1. *Clinical Lymphoma Myeloma and Leukemia*, 15:S167–S169, 2015.

[173] Michael Y Choi, George F Widhopf II, Emanuela M Ghia, Reilly L Kidwell, Md Kamrul Hasan, Jian Yu, Laura Z Rassenti, Liguang Chen, Yun Chen, Emily Pittman, et al. Phase i trial: cirmtuzumab inhibits ror1 signaling and stemness signatures in patients with chronic lymphocytic leukemia. *Cell Stem Cell*, 22(6):951–959, 2018.

[174] Lars Wallstabe, Claudia Göttlich, Lena C Nelke, Johanna Kühnemundt, Thomas Schwarz, Thomas Nerreter, Hermann Einsele, Heike Walles, Gudrun Dandekar, Sarah L Nietzer, et al. Ror1-car t cells are effective against lung and breast cancer in advanced microphysiologic 3d tumor models. *JCI insight*, 4(18), 2019.

[175] Xuesha Liu, Wenchen Pu, Huaiyu He, Xin Fan, Yuanyuan Zheng, Jian-Kang Zhou, Rui Ma, Juan He, Yuzhu Zheng, Ke Wu, et al. Novel ror1 inhibitor ari-1 suppresses the development of non-small cell lung cancer. *Cancer Letters*, 458:76–85, 2019.

[176] Joshua B Sheetz, Sebastian Mathea, Hanna Karvonen, Ketan Malhotra, Deep Chatterjee, Wilhelmiina Niininen, Robert Perttilä, Franziska Preuss, Krishna Suresh, Steven E Stayrook, et al. Structural insights into pseudokinase domains of receptor tyrosine kinases. *Molecular cell*, 79(3):390–405, 2020.

[177] Marc Damelin, Alexander Bankovich, Jeffrey Bernstein, Justin Lucas, Liang Chen, Samuel Williams, Albert Park, Jorge Aguilar, Elana Ernstoff, Manoj Charati, et al. A ptk7-targeted antibody-drug conjugate reduces tumor-initiating cells and induces sustained tumor regressions. *Science translational medicine*, 9(372):eaag2611, 2017.

[178] Michael L Maitland, Jasgit C Sachdev, Manish R Sharma, Victor Moreno, Valentina Boni, Shivaani Kummar, Erica Stringer-Reasor, Nehal Lakhani, Allison R Moreau, Dawei Xuan, et al. First-in-human study of pf-06647020 (cofetuzumab pelidotin), an antibody–drug conjugate targeting protein tyrosine kinase 7, in advanced solid tumors. *Clinical Cancer Research*, 27(16):4511–4520, 2021.

[179] Jeremy A Goettel, Dongchun Liang, Valda C Hilliard, Karen L Edelblum, Matthew R Broadus, Kathleen L Gould, Steven K Hanks, and D Brent Polk. Ksr1 is a functional protein kinase capable of serine autophosphorylation and direct phosphorylation of mek1. *Experimental cell research*, 317(4):452–463, 2011.

[180] Damian F Brennan, Arvin C Dar, Nicholas T Hertz, William CH Chao, Alma L Burlingame, Kevan M Shokat, and David Barford. A raf-induced allosteric transition of ksr stimulates phosphorylation of mek. *Nature*, 472(7343):366–369, 2011.

[181] Thanashan Rajakulendran, Malha Sahmi, Martin Lefrançois, Frank Sicheri, and Marc Therrien. A dimerization-dependent mechanism drives raf catalytic activation. *Nature*, 461(7263):542–545, 2009.

[182] Hugo Lavoie, Malha Sahmi, Pierre Maisonneuve, Sara A Marullo, Neroshan Thevakumaran, Ting Jin, Igor Kurinov, Frank Sicheri, and Marc Therrien. Mek drives braf activation through allosteric control of ksr proteins. *Nature*, 554(7693):549–553, 2018.

[183] Libero Santarpia, Scott M Lippman, and Adel K El-Naggar. Targeting the mapk–ras–raf signaling pathway in cancer therapy. *Expert opinion on therapeutic targets*, 16(1): 103–119, 2012.

[184] Beth K Neilsen, Danielle E Frodyma, Robert E Lewis, and Kurt W Fisher. Ksr as a therapeutic target for ras-dependent cancers. *Expert opinion on therapeutic targets*, 21(5):499–509, 2017.

[185] Ryan C Gimple and Xiuxing Wang. Ras: Striking at the core of the oncogenic circuitry. *Frontiers in oncology*, 9:965, 2019.

[186] Laura R Pearce, Neli Atanassova, Matthew C Banton, Bill Bottomley, Agatha A van der Klaauw, Jean-Pierre Revelli, Audrey Hendricks, Julia M Keogh, Elana Henning, Deon Doree, et al. Ksr2 mutations are associated with obesity, insulin resistance, and impaired cellular fuel oxidation. *Cell*, 155(4):765–777, 2013.

[187] Neil S Dhawan, Alex P Scopton, and Arvin C Dar. Small molecule stabilization of the ksr inactive state antagonizes oncogenic ras signalling. *Nature*, 537(7618):112–116, 2016.

[188] Zaigham M Khan, Alexander M Real, William M Marsiglia, Arthur Chow, Mary E Duffy, Jayasudhan R Yerabolu, Alex P Scopton, and Arvin C Dar. Structural basis for the action of the drug trametinib at ksr-bound mek. *Nature*, 588(7838):509–514, 2020.

[189] Konark Mukherjee, Manu Sharma, Henning Urlaub, Gleb P Bourenkov, Reinhard Jahn, Thomas C Südhof, and Markus C Wahl. Cask functions as a mg2+-independent neurexin kinase. *Cell*, 133(2):328–339, 2008.

[190] Yi-Ping Hsueh. The role of the maguk protein cask in neural development and synaptic function. *Current medicinal chemistry*, 13(16):1915–1927, 2006.

[191] Jin-Lai Wei, Zhong-Xue Fu, Min Fang, Qiu-Yuan Zhou, Qing-Ning Zhao, Jin-Bao Guo, Wei-Dong Lu, and Hao Wang. High expression of cask correlates with progression and poor prognosis of colorectal cancer. *Tumor Biology*, 35(9):9185–9194, 2014.

[192] Jun Qu, Yeqi Zhou, Yuan Li, Jinsong Yu, and Wei Wang. Cask regulates notch pathway and functions as a tumor promoter in pancreatic cancer. *Archives of Biochemistry and Biophysics*, 701:108789, 2021.

[193] Anna Stepczynska, Kirsten Lauber, Ingo H Engels, Ottmar Janssen, Dieter Kabelitz, Sebastian Wesselborg, and Klaus Schulze-Osthoff. Staurosporine and conventional anticancer drugs induce overlapping, yet distinct pathways of apoptosis and caspase activation. *Oncogene*, 20(10):1193–1202, 2001.

[194] Mindy I Davis, Jeremy P Hunt, Sanna Herrgard, Pietro Ciceri, Lisa M Wodicka, Gabriel Pallares, Michael Hocker, Daniel K Treiber, and Patrick P Zarrinkar. Comprehensive analysis of kinase inhibitor selectivity. *Nature biotechnology*, 29(11):1046–1051, 2011.

[195] Nadine Russ, Martin Schröder, Benedict-Tilman Berger, Sebastian Mandel, Yagmur Aydogan, Sandy Mauer, Christian Pohl, David H Drewry, Apirat Chaikuad, Susanne Müller, et al. Design and development of a chemical probe for pseudokinase ca2+/calmodulin-dependent ser/thr kinase. *Journal of Medicinal Chemistry*, 64(19): 14358–14376, 2021.

[196] Susmita Khamrui, Peter MU Ung, Cody Secor, Avner Schlessinger, and Michael B Lazarus. High-resolution structure and inhibition of the schizophrenia-linked pseudokinase ulk4. *Journal of the American Chemical Society*, 142(1):33–37, 2019.

[197] Franziska Preuss, Deep Chatterjee, Sebastian Mathea, Safal Shrestha, Jonathan St-Germain, Manipa Saha, Natarajan Kannan, Brian Raught, Robert Rottapel, and Stefan Knapp. Nucleotide binding, evolutionary insights, and interaction partners of the pseudokinase unc-51-like kinase 4. *Structure*, 28(11):1184–1196, 2020.

[198] Sven M Lange, Marina I Nelen, Philip Cohen, and Yogesh Kulathu. Dimeric structure of the pseudokinase irak3 suggests an allosteric mechanism for negative regulation. *Structure*, 29(3):238–251, 2021.

[199] Patrick A Eyers, Karen Keeshan, and Natarajan Kannan. Tribbles in the 21st century: the evolving roles of tribbles pseudokinases in biology and disease. *Trends in cell biology*, 27(4):284–298, 2017.

[200] Thomas C Seher and Maria Leptin. Tribbles, a cell-cycle brake that coordinates proliferation and morphogenesis during drosophila gastrulation. *Current Biology*, 10(11): 623–629, 2000.

[201] Jörg Großhans and Eric Wieschaus. A genetic link between morphogenesis and cell division during formation of the ventral furrow in drosophila. *Cell*, 101(5):523–531, 2000.

[202] Juan Mata, Silvia Curado, Anne Ephrussi, and Pernille Rørth. Tribbles coordinates mitosis and morphogenesis in drosophila by regulating string/cdc25 proteolysis. *Cell*, 101(5):511–522, 2000.

[203] Fraçoise Wilkin, Nathalie Suarez-Huerta, Bernard Robaye, Julien Peetermans, Frederick Libert, Jacques E Dumont, and Carine Maenhaut. Characterization of a phosphoprotein whose mrna is regulated by the mitogenic pathways in dog thyroid cells. *European Journal of Biochemistry*, 248(3):660–668, 1997.

[204] Keiko Mayumi-Matsuda, Satoshi Kojima, Harukazu Suzuki, and Tsuneaki Sakata. Identification of a novel kinase-like gene induced during neuronal cell death. *Biochemical and biophysical research communications*, 258(2):260–264, 1999.

[205] Min Wu, Liang-Guo Xu, Zhonghe Zhai, and Hong-Bing Shu. Sink is a p65-interacting negative regulator of nf-κb-dependent transcription. *Journal of Biological Chemistry*, 278(29):27072–27079, 2003.

[206] Alex J Bowers, Sheila Scully, and John F Boylan. Skip3, a novel drosophila tribbles ortholog, is overexpressed in human tumors and is regulated by hypoxia. *Oncogene*, 22(18):2823–2835, 2003.

[207] Fiona P Bailey, Dominic P Byrne, Krishnadev Oruganty, Claire E Eyers, Christopher J Novotny, Kevan M Shokat, Natarajan Kannan, and Patrick A Eyers. The tribbles 2 (trb2) pseudokinase binds to atp and autophosphorylates in a metal-independent manner. *The Biochemical journal*, 467(1):47, 2015.

[208] Martin Rechsteiner and Scott W Rogers. Pest sequences and regulation by proteolysis. *Trends in biochemical sciences*, 21(7):267–271, 1996.

[209] Takashi Yokoyama and Takuro Nakamura. Tribbles in disease: Signaling pathways important for cellular function and neoplastic transformation. *Cancer science*, 102(6): 1115–1122, 2011.

[210] James M Murphy, Yoshio Nakatani, Sam A Jamieson, Weiwen Dai, Isabelle S Lucet, and Peter D Mace. Molecular mechanism of ccaat-enhancer binding protein recruitment by the trib1 pseudokinase. *Structure*, 23(11):2111–2121, 2015.

[211] Sam A Jamieson, Zheng Ruan, Abigail E Burgess, Jack R Curry, Hamish D McMillan, Jodi L Brewster, Anita K Dunbier, Alison D Axtman, Natarajan Kannan, and Peter D Mace. Substrate binding allosterically relieves autoinhibition of the pseudokinase trib1. *Science signaling*, 11(549):eaau0597, 2018.

[212] Sacha Uljon, Xiang Xu, Izabela Durzynska, Sarah Stein, Guillaume Adelmant, Jarrod A Marto, Warren S Pear, and Stephen C Blacklow. Structural basis for substrate selectivity of the e3 ligase cop1. *Structure*, 24(5):687–696, 2016.

[213] Takashi Yokoyama, Yohei Kanno, Yukari Yamazaki, Tomoko Takahara, Satoshi Miyata, and Takuro Nakamura. Trib1 links the mek1/erk pathway in myeloid leukemogenesis. *Blood, The Journal of the American Society of Hematology*, 116(15):2768–2775, 2010.

[214] Endre Kiss-Toth, Stephanie M Bagstaff, Hye Y Sung, Veronika Jozsa, Clare Dempsey, Jim C Caunt, Kevin M Oxley, David H Wyllie, Timea Polgar, Mary Harte, et al. Human tribbles, a protein family controlling mitogen-activated protein kinase cascades. *Journal of Biological Chemistry*, 279(41):42703–42708, 2004.

[215] Izabela Durzynska, Xiang Xu, Guillaume Adelmant, Scott B Ficarro, Jarrod A Marto, Piotrek Sliz, Sacha Uljon, and Stephen C Blacklow. Stk40 is a pseudokinase that binds the e3 ubiquitin ligase cop1. *Structure*, 25(2):287–294, 2017.

[216] Hongtao Guan, Aban Shuaib, David Davila De Leon, Adrienn Angyal, Maria Salazar, Guillermo Velasco, Mike Holcombe, Steven K Dower, and Endre Kiss-Toth. Competition between members of the tribbles pseudokinase protein family shapes their interactions with mitogen activated protein kinase pathways. *Scientific reports*, 6(1): 1–12, 2016.

[217] Ke Li, Feng Wang, Wen-Bin Cao, Xiao-Xi Lv, Fang Hua, Bing Cui, Jiao-Jiao Yu, Xiao-Wei Zhang, Shuang Shang, Shan-Shan Liu, et al. Trib3 promotes apl progression through stabilization of the oncoprotein pml-rarα and inhibition of p53-mediated senescence. *Cancer cell*, 31(5):697–710, 2017.

[218] Richard Hill, Patricia A Madureira, Bibiana Ferreira, Inês Baptista, Susana Machado, Laura Colaco, Marta Dos Santos, Ningshu Liu, Ana Dopazo, Selma Ugurel, et al. Trib2 confers resistance to anti-cancer therapy by activating the serine/threonine protein kinase akt. *Nature communications*, 8(1):1–9, 2017.

[219] Keyong Du, Stephan Herzig, Rohit N Kulkarni, and Marc Montminy. Trb3: a tribbles homolog that inhibits akt/pkb activation by insulin in liver. *Science*, 300(5625):1574–1577, 2003.

[220] N Zareen, SC Biswas, and LA Greene. A feed-forward loop involving trib3, akt and foxo mediates death of ngf-deprived neurons. *Cell Death & Differentiation*, 20(12): 1719–1730, 2013.

[221] Ling Qi, Jose E Heredia, Judith Y Altarejos, Robert Screaton, Naomi Goebel, Sherry Niessen, Ian X MacLeod, Chong Wee Liew, Rohit N Kulkarni, James Bain, et al. Trb3 links the e3 ubiquitin ligase cop1 to lipid metabolism. *science*, 312(5781):1763–1766, 2006.

[222] Karen Keeshan, Yiping He, Bas J Wouters, Olga Shestova, Lanwei Xu, Hong Sai, Carlos G Rodriguez, Ivan Maillard, John W Tobias, Peter Valk, et al. Tribbles homolog 2 inactivates c/ebpα and causes acute myelogenous leukemia. *Cancer cell*, 10(5):401–411, 2006.

[223] Mara Salomé, Lisa Hopcroft, and Karen Keeshan. Inverse and correlative relationships between tribbles genes indicate non-redundant functions during normal and malignant hemopoiesis. *Experimental Hematology*, 66:63–78, 2018.

[224] Takashi Satoh, Hiroyasu Kidoya, Hisamichi Naito, Masahiro Yamamoto, Naoki Takemura, Katsuhiro Nakagawa, Yoshichika Yoshioka, Eiichi Morii, Nobuyuki Takakura, Osamu Takeuchi, et al. Critical role of trib1 in differentiation of tissue-resident m2-like macrophages. *Nature*, 495(7442):524–528, 2013.

[225] Jessica M Johnston, Adrienn Angyal, Robert C Bauer, Stephen Hamby, S Kim Suvarna, Kajus Baidžajevas, Zoltan Hegedus, T Neil Dear, Martin Turner, Cardiogenics Consortium, et al. Myeloid tribbles 1 induces early atherosclerosis via enhanced foam cell expansion. *Science advances*, 5(10):eaax9183, 2019.

[226] Dipak P Ramji and Pelagia Foka. Ccaat/enhancer-binding proteins: structure, function and regulation. *Biochemical Journal*, 365(3):561–575, 2002.

[227] Kristian Reckzeh and Jörg Cammenga. Molecular mechanisms underlying deregulation of c/ebpα in acute myeloid leukemia. *International journal of hematology*, 91(4):557–568, 2010.

[228] Takashi Yokoyama, Tsutomu Toki, Yoshihiro Aoki, Rika Kanezaki, Myoung-ja Park, Yohei Kanno, Tomoko Takahara, Yukari Yamazaki, Etsuro Ito, Yasuhide Hayashi, et al. Identification of trib1 r107l gain-of-function mutation in human acute megakaryocytic leukemia. *Blood, The Journal of the American Society of Hematology*, 119(11):2608–2611, 2012.

[229] Mara Salomé, Aoife Magee, Krisha Yalla, Shahzya Chaudhury, Evgenia Sarrou, Ruaidhrí J Carmody, and Karen Keeshan. A trib2-p38 axis controls myeloid leukaemia cell cycle and stress response signalling. *Cell death & disease*, 9(5):1–17, 2018.

[230] Michael C Lawrence, Arif Jivan, Chunli Shao, Lingling Duan, Daryl Goad, Elma Zaganjor, Jihan Osborne, Kathleen McGlynn, Steve Stippec, Svetlana Earnest, et al. The roles of mapks in disease. *Cell research*, 18(4):436–442, 2008.

[231] Chiharu Miyajima, Yasumichi Inoue, and Hidetoshi Hayashi. Pseudokinase tribbles 1 (trb1) negatively regulates tumor-suppressor activity of p53 through p53 deacetylation. *Biological and Pharmaceutical Bulletin*, 38(4):618–624, 2015.

[232] Hamish D McMillan, Karen Keeshan, Anita K Dunbier, and Peter D Mace. Structure vs. function of trib1—myeloid neoplasms and beyond. *Cancers*, 13(12):3060, 2021.

[233] Shi Hao Tan, Alice Wei Yee Yam, Lee N Lawton, Regina Wan Ju Wong, Richard A Young, A Thomas Look, and Takaomi Sanda. Trib2 reinforces the oncogenic transcriptional program controlled by the tal1 complex in t-cell acute lymphoblastic leukemia. *Leukemia*, 30(4):959–962, 2016.

[234] Sarah J Stein, Ethan A Mack, Kelly S Rome, Kostandin V Pajcini, Takuya Ohtani, Lanwei Xu, Yunlei Li, Jules PP Meijerink, Robert B Faryabi, and Warren S Pear. Trib2 suppresses tumor initiation in notch-driven t-all. *PLoS One*, 11(5):e0155408, 2016.

[235] Ke Li, Feng Wang, Zhao-na Yang, Ting-ting Zhang, Yu-fen Yuan, Chen-xi Zhao, Zaiwuli Yeerjiang, Bing Cui, Fang Hua, Xiao-xi Lv, et al. Trib3 promotes myc-associated lymphoma development through suppression of ube3b-mediated myc degradation. *Nature communications*, 11(1):1–20, 2020.

[236] Jin-mei Yu, Wei Sun, Zhen-he Wang, Xiao Liang, Fang Hua, Ke Li, Xiao-xi Lv, Xiao-wei Zhang, Yu-ying Liu, Jiao-jiao Yu, et al. Trib3 supports breast cancer stemness by suppressing foxo1 degradation and enhancing sox2 transcription. *Nature communications*, 10(1):1–16, 2019.

[237] Pascaline Aimé, Xiaotian Sun, Neela Zareen, Apeksha Rao, Zachary Berman, Laura Volpicelli-Daley, Paulette Bernd, John F Crary, Oren A Levy, and Lloyd A Greene. Trib3 is elevated in parkinson's disease and mediates death in parkinson's disease models. *Journal of Neuroscience*, 35(30):10731–10749, 2015.

[238] Suraiya Saleem and Subhas Chandra Biswas. Tribbles pseudokinase 3 induces both apoptosis and autophagy in amyloid-β-induced neuronal death. *Journal of Biological Chemistry*, 292(7):2571–2585, 2017.

[239] Marco Lorenzi, Andre Altmann, Boris Gutman, Selina Wray, Charles Arber, Derrek P Hibar, Neda Jahanshad, Jonathan M Schott, Daniel C Alexander, Paul M Thompson, et al. Susceptibility of brain atrophy to trib3 in alzheimer's disease, evidence from functional prioritization in imaging genetics. *Proceedings of the National Academy of Sciences*, 115(12):3162–3167, 2018.

[240] Laura Richmond and Karen Keeshan. Pseudokinases: a tribble-edged sword. *The FEBS Journal*, 287(19):4170–4182, 2020.

[241] Javier De Las Rivas and Celia Fontanillo. Protein–protein interactions essentials: key concepts to building and analyzing interactome networks. *PLoS computational biology*, 6(6):e1000807, 2010.

[242] Eric M Phizicky and Stanley Fields. Protein-protein interactions: methods for detection and analysis. *Microbiological reviews*, 59(1):94–123, 1995.

[243] Brandon T Ruotolo, Justin LP Benesch, Alan M Sandercock, Suk-Joon Hyung, and Carol V Robinson. Ion mobility–mass spectrometry analysis of large protein complexes. *Nature protocols*, 3(7):1139–1152, 2008.

[244] Yuhong Du. Fluorescence polarization assay to quantify protein-protein interactions in an hts format. In *Protein-Protein Interactions*, pages 529–544. Springer, 2015.

[245] Anca Margineanu, Jia Jia Chan, Douglas J Kelly, Sean C Warren, Delphine Flatters, Sunil Kumar, Matilda Katan, Christopher W Dunsby, and Paul MW French. Screening for protein-protein interactions using förster resonance energy transfer (fret) and fluorescence lifetime imaging microscopy (flim). *Scientific reports*, 6(1):1–17, 2016.

[246] Matthew A Cooper. Optical biosensors in drug discovery. *Nature reviews Drug discovery*, 1(7):515–528, 2002.

[247] Anna Brückner, Cécile Polge, Nicolas Lentze, Daniel Auerbach, and Uwe Schlattner. Yeast two-hybrid, a powerful tool for systems biology. *International journal of molecular sciences*, 10(6):2763–2788, 2009.

[248] Andrew S Dixon, Marie K Schwinn, Mary P Hall, Kris Zimmerman, Paul Otto, Thomas H Lubben, Braeden L Butler, Brock F Binkowski, Thomas Machleidt, Thomas A Kirkland, et al. Nanoluc complementation reporter optimized for accurate measurement of protein interactions in cells. *ACS chemical biology*, 11(2):400–408, 2016.

[249] Montse Morell, Salvador Ventura, and Francesc X Avilés. Protein complementation assays: approaches for the in vivo analysis of protein interactions. *FEBS letters*, 583 (11):1684–1691, 2009.

[250] Sergey Ovchinnikov, Hetunandan Kamisetty, and David Baker. Robust and accurate prediction of residue–residue interactions across protein interfaces using evolutionary information. *elife*, 3:e02030, 2014.

[251] Li C Xue, Joao PGLM Rodrigues, Drena Dobbs, Vasant Honavar, and Alexandre MJJ Bonvin. Template-based protein–protein docking exploiting pairwise interfacial residue restraints. *Briefings in bioinformatics*, 18(3):458–466, 2017.

[252] Richard Evans, Michael O'Neill, Alexander Pritzel, Natasha Antropova, Andrew W Senior, Timothy Green, Augustin Žídek, Russell Bates, Sam Blackwell, Jason Yim, et al. Protein complex prediction with alphafold-multimer. *BioRxiv*, 2021.

[253] Andrej Šali and Tom L Blundell. Comparative protein modelling by satisfaction of spatial restraints. *Journal of molecular biology*, 234(3):779–815, 1993.

[254] Benjamin Webb and Andrej Sali. Protein structure modeling with modeller. *Protein Structure Prediction. Methods in Molecular Biology (Methods and Protocols)*, 1137: 1–15, 2014.

[255] Mark James Abraham, Teemu Murtola, Roland Schulz, Szilárd Páll, Jeremy C Smith, Berk Hess, and Erik Lindahl. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1:19–25, 2015.

[256] James A Maier, Carmenza Martinez, Koushik Kasavajhala, Lauren Wickstrom, Kevin E Hauser, and Carlos Simmerling. ff14sb: improving the accuracy of protein side chain and backbone parameters from ff99sb. *Journal of chemical theory and computation*, 11(8):3696–3713, 2015.

[257] Berk Hess, Henk Bekker, Herman JC Berendsen, and Johannes GEM Fraaije. Lincs: a linear constraint solver for molecular simulations. *Journal of computational chemistry*, 18(12):1463–1472, 1997.

[258] LLC Schrödinger. Release, 2016-3. *Schrödinger LLC, New York, NY*, 2016.

[259] Sergio Ruiz-Carmona, Daniel Alvarez-Garcia, Nicolas Foloppe, A Beatriz Garmendia-Doval, Szilveszter Juhos, Peter Schmidtke, Xavier Barril, Roderick E Hubbard, and S David Morley. rdock: a fast, versatile and open source program for docking ligands to proteins and nucleic acids. *PLoS computational biology*, 10(4):e1003571, 2014.

[260] Matteo Tiberti, Thilde Terkelsen, Tycho Canter Cremers, Miriam Di Marco, Isabelle da Piedade, Emiliano Maiani, and Elena Papaleo. Mutatex: an automated pipeline for in-silico saturation mutagenesis of protein structures and structural ensembles. *BioRxiv*, page 824938, 2019.

[261] Barak Raveh, Nir London, Lior Zimmerman, and Ora Schueler-Furman. Rosetta flexpepdock ab-initio: simultaneous folding, docking and refinement of peptides onto their receptors. *PloS one*, 6(4):e18934, 2011.

[262] David E Kim, Dylan Chivian, and David Baker. Protein structure prediction and analysis using the robetta server. *Nucleic acids research*, 32(suppl_2):W526–W531, 2004.

[263] Elisabeth Gasteiger, Christine Hoogland, Alexandre Gattiker, Marc R Wilkins, Ron D Appel, Amos Bairoch, et al. Protein identification and analysis tools on the expasy server. *The proteomics protocols handbook*, pages 571–607, 2005.

[264] Masoud Vedadi, Frank H Niesen, Abdellah Allali-Hassani, Oleg Y Fedorov, Patrick J Finerty, Gregory A Wasney, Ron Yeung, Cheryl Arrowsmith, Linda J Ball, Helena Berglund, et al. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proceedings of the National Academy of Sciences*, 103(43):15835–15840, 2006.

[265] Lei Ji, Bo Jiang, Xiaomo Jiang, Olga Charlat, Amy Chen, Craig Mickanin, Andreas Bauer, Wenqing Xu, Xiaoxue Yan, and Feng Cong. The siah e3 ubiquitin ligases promote wnt/β-catenin signaling through mediating wnt-induced axin degradation. *Genes & development*, 31(9):904–915, 2017.

[266] Jonathan M Elkins, Vita Fedele, Marta Szklarz, Kamal R Abdul Azeez, Eidarus Salah, Jowita Mikolajczyk, Sergei Romanov, Nikolai Sepetov, Xi-Ping Huang, Bryan L Roth, et al. Comprehensive characterization of the published kinase inhibitor set. *Nature biotechnology*, 34(1):95–103, 2016.

[267] Rosselle T Dungo and Gillian M Keating. Afatinib: first global approval. *Drugs*, 73 (13):1503–1515, 2013.

[268] Christopher L McClendon, Alexandr P Kornev, Michael K Gilson, and Susan S Taylor. Dynamic architecture of a protein kinase. *Proceedings of the National Academy of Sciences*, 111(43):E4623–E4631, 2014.

[269] Masaichi Hasegawa, Naohiko Nishigaki, Yoshiaki Washio, Kazuya Kano, Philip A Harris, Hideyuki Sato, Ichiro Mori, Rob I West, Megumi Shibahara, Hiroko Toyoda, et al. Discovery of novel benzimidazoles as potent inhibitors of tie-2 and vegfr-2 tyrosine kinase receptors. *Journal of medicinal chemistry*, 50(18):4453–4470, 2007.

[270] Rainer Fährrolfes, Stefan Bietz, Florian Flachsenberg, Agnes Meyder, Eva Nittinger, Thomas Otto, Andrea Volkamer, and Matthias Rarey. Proteins plus: a web portal for structure analysis of macromolecules. *Nucleic Acids Research*, 45(W1):W337–W343, 2017.

[271] Kathleen S Metz, Erika M Deoudes, Matthew E Berginski, Ivan Jimenez-Ruiz, Bulent Arman Aksoy, Jeff Hammerbacher, Shawn M Gomez, and Douglas H Phanstiel. Coral: Clear and customizable visualization of human kinome data. *Cell systems*, 7 (3):347–350, 2018.

[272] George Mihai Nitulescu, Maryna Van De Venter, Georgiana Nitulescu, Anca Ungurianu, Petras Juzenas, Qian Peng, Octavian Tudorel Olaru, Daniela Grădinaru, Aristides Tsatsakis, Dimitris Tsoukalas, et al. The akt pathway in oncology therapy and beyond. *International Journal of Oncology*, 53(6):2319–2331, 2018.

[273] Ingeborg Hers, Emma E Vincent, and Jeremy M Tavaré. Akt signalling in health and disease. *Cellular signalling*, 23(10):1515–1527, 2011.

[274] Valeria Facchinetti, Weiming Ouyang, Hua Wei, Nelyn Soto, Adam Lazorchak, Christine Gould, Carolyn Lowry, Alexandra C Newton, Yuxin Mao, Robert Q Miao, et al. The mammalian target of rapamycin complex 2 controls folding and stability of akt and protein kinase c. *The EMBO journal*, 27(14):1932–1943, 2008.

[275] Nam Chu, Thibault Viennet, Hwan Bae, Antonieta Salguero, Andras Boeszoermenyi, Haribabu Arthanari, and Philip A Cole. The structural determinants of ph domain-mediated regulation of akt revealed by segmental labeling. *Elife*, 9:e59151, 2020.

[276] Tianyan Gao, Frank Furnari, and Alexandra C Newton. Phlpp: a phosphatase that directly dephosphorylates akt, promotes apoptosis, and suppresses tumor growth. *Molecular cell*, 18(1):13–24, 2005.

[277] Mirjana Andjelković, Teresa Jakubowicz, Peter Cron, Xiu-Fen Ming, Jeung-Whan Han, and Brian A Hemmings. Activation and phosphorylation of a pleckstrin homology domain containing protein kinase (rac-pk/pkb) promoted by serum and protein phosphatase inhibitors. *Proceedings of the National Academy of Sciences*, 93(12): 5699–5704, 1996.

[278] Michael Ebner, Iva Lučić, Thomas A Leonard, and Ivan Yudushkin. Pi (3, 4, 5) p3 engagement restricts akt activity to cellular membranes. *Molecular Cell*, 65(3):416–431, 2017.

[279] Keyong Du and Philip N Tsichlis. Regulation of the akt kinase by interacting proteins. *Oncogene*, 24(50):7401–7409, 2005.

[280] Patrick B Iynedjian. Lack of evidence for a role of trb3/nipk as an inhibitor of pkb-mediated insulin signalling in primary hepatocytes. *Biochemical Journal*, 386(1):113–118, 2005.

[281] Emily Borsting, Shalin V Patel, Anne-Emilie Declèves, Sarah J Lee, Qazi M Rahman, Shizuo Akira, Joe Satriano, Kumar Sharma, Volker Vallon, and Robyn Cunard. Tribbles homolog 3 attenuates mammalian target of rapamycin complex-2 signaling and inflammation in the diabetic kidney. *Journal of the american society of nephrology*, 25(9):2067–2078, 2014.

[282] Dos D Sarbassov, David A Guertin, Siraj M Ali, and David M Sabatini. Phosphorylation and regulation of akt/pkb by the rictor-mtor complex. *Science*, 307(5712): 1098–1101, 2005.

[283] Steven Pelech. Dimerization in protein kinase signaling. *Journal of Biology*, 5(5):1–7, 2006.

[284] Boris N Kholodenko. Drug resistance resulting from kinase dimerization is rationalized by thermodynamic factors describing allosteric inhibitor effects. *Cell reports*, 12(11): 1939–1949, 2015.

[285] Sabrina Prudente and Vincenzo Trischitta. The trib3 q84r polymorphism, insulin resistance and related metabolic alterations. *Biochemical Society Transactions*, 43(5): 1108–1111, 2015.

[286] Bojana Stefanovska, Fabrice André, and Olivia Fromigué. Tribbles pseudokinase 3 regulation and contribution to cancer. *Cancers*, 13(8):1822, 2021.

[287] from https://www.phosphosite.org/uniprotaccaction?id=q96ru7 accessed on january 2022, 2013.

[288] Alan L Schwartz and Aaron Ciechanover. Targeting proteins for destruction by the ubiquitin system: implications for human pathobiology. *Annual review of pharmacology and toxicology*, 49:73–96, 2009.

[289] Jeanine A Harrigan, Xavier Jacq, Niall M Martin, and Stephen P Jackson. Deubiquitylating enzymes and drug discovery: emerging opportunities. *Nature reviews Drug discovery*, 17(1):57–78, 2018.

[290] Martin Sadowski and Boris Sarcevic. Mechanisms of mono-and poly-ubiquitination: Ubiquitination specificity depends on compatibility between the e2 catalytic core and amino acid residues proximal to the lysine. *Cell division*, 5(1):1–5, 2010.

[291] Raymond J Deshaies and Claudio AP Joazeiro. Ring domain e3 ubiquitin ligases. *Annual review of biochemistry*, 78, 2009.

[292] Dawn M Wenzel, Alexei Lissounov, Peter S Brzovic, and Rachel E Klevit. Ubch7 reactivity profile reveals parkin and hhari to be ring/hect hybrids. *Nature*, 474(7349): 105–108, 2011.

[293] Ning Zheng and Nitzan Shabek. Ubiquitin ligases: structure, function, and regulation. *Annual review of biochemistry*, 86:129–157, 2017.

[294] Liam J Holt. Regulatory modules: Coupling protein stability to phopshoregulation during cell division. *FEBS letters*, 586(17):2773–2777, 2012.

[295] Zhixiong Zeng, Wei Wang, Yuting Yang, Yong Chen, Xiaomei Yang, J Alan Diehl, Xuedong Liu, and Ming Lei. Structural basis of selective ubiquitination of trf1 by scffbx4. *Developmental cell*, 18(2):214–225, 2010.

[296] Nobumichi Ohoka, Satoshi Sakai, Kikuo Onozaki, Makoto Nakanishi, and Hidetoshi Hayashi. Anaphase-promoting complex/cyclosome-cdh1 mediates the ubiquitination and degradation of trb3. *Biochemical and biophysical research communications*, 392 (3):289–294, 2010.

[297] Ying Zhou, Lu Li, Qiongming Liu, Guichun Xing, Xuezhang Kuai, Jing Sun, Xiushan Yin, Jian Wang, Lingqiang Zhang, and Fuchu He. E3 ubiquitin ligase siah1 mediates ubiquitination and degradation of trb3. *Cellular signalling*, 20(5):942–948, 2008.

[298] Richard W Carthew and Gerald M Rubin. seven in absentia, a gene required for specification of r7 cell fate in the drosophila eye. *Cell*, 63(3):561–577, 1990.

[299] Christiane M Robbins, Waibov A Tembe, Angela Baker, Shripad Sinari, Tracy Y Moses, Stephen Beckstrom-Sternberg, James Beckstrom-Sternberg, Michael Barrett, James Long, Arul Chinnaiyan, et al. Copy number and targeted mutational analysis reveals novel somatic events in metastatic prostate tumors. *Genome research*, 21(1): 47–55, 2011.

[300] Ian J Pepper, Robert E Van Sciver, and Amy H Tang. Phylogenetic analysis of the sina/siah ubiquitin e3 ligase family in metazoa. *BMC evolutionary biology*, 17(1):1–16, 2017.

[301] Jianfei Qi, Hyungsoo Kim, Marzia Scortegagna, and Ze'ev A Ronai. Regulators and effectors of siah ubiquitin ligases. *Cell biochemistry and biophysics*, 67(1):15–24, 2013.

[302] M Buchwald, K Pietschmann, P Brand, A Günther, NP Mahajan, T Heinzel, and OH Krämer. Siah ubiquitin ligases target the nonreceptor tyrosine kinase ack1 for ubiquitinylation and proteasomal degradation. *Oncogene*, 32(41):4913–4920, 2013.

[303] Anupriya Gopalsamy, Thilo Hagen, and Kunchithapadam Swaminathan. Investigating the molecular basis of siah1 and siah2 e3 ubiquitin ligase substrate specificity. *PloS one*, 9(9):e106547, 2014.

[304] Vadim Rimsa, Thomas C Eadsforth, and William N Hunter. Two high-resolution structures of the human e3 ubiquitin ligase siah1. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, 69(12):1339–1343, 2013.

[305] Galina Polekhina, Colin M House, Nadia Traficante, Joel P Mackay, Frédéric Relaix, David A Sassoon, Michael W Parker, and David DL Bowtell. Siah ubiquitin ligase is structurally related to traf and modulates tnf-α signaling. *Nature structural biology*, 9 (1):68–75, 2002.

[306] Gang Hu and Eric R Fearon. Siah-1 n-terminal ring domain is required for proteolysis function, and c-terminal sequences regulate oligomerization and binding to target proteins. *Molecular and cellular biology*, 19(1):724–732, 1999.

[307] Colin M House, Ian J Frew, Huei-Luen Huang, Gerhard Wiche, Nadia Traficante, Edouard Nice, Bruno Catimel, and David DL Bowtell. A binding motif for siah ubiquitin ligase. *Proceedings of the National Academy of Sciences*, 100(6):3101–3106, 2003.

[308] Colin M House, Nancy C Hancock, Andreas Möller, Brett A Cromer, Victor Fedorov, David DL Bowtell, Michael W Parker, and Galina Polekhina. Elucidation of the substrate binding site of siah ubiquitin ligase. *Structure*, 14(4):695–701, 2006.

[309] Yoana N Dimitrova, Jiong Li, Young-Tae Lee, Jessica Rios-Esteves, David B Friedman, Hee-Jung Choi, William I Weis, Cun-Yu Wang, and Walter J Chazin. Direct ubiquitination of β-catenin by siah-1 and regulation by the exchange factor tbl1. *Journal of Biological Chemistry*, 285(18):13507–13516, 2010.

[310] Shu-ichi Matsuzawa and John C Reed. Siah-1, sip, and ebi collaborate in a novel pathway for β-catenin degradation linked to p53 responses. *Molecular cell*, 7(5):915–926, 2001.

[311] Clarissa L Waites, Sergio A Leal-Ortiz, Nathan Okerlund, Hannah Dalke, Anna Fejtova, Wilko D Altrock, Eckart D Gundelfinger, and Craig C Garner. Bassoon and piccolo maintain synapse integrity by regulating protein ubiquitination and degradation. *The EMBO journal*, 32(7):954–969, 2013.

[312] Asher Mullard. A probe for every protein. *Nature Reviews Drug Discovery*, 18(10): 733–737, 2019.

[313] Daniel M Foulkes, Dominic P Byrne, Fiona P Bailey, and Patrick A Eyers. Tribbles pseudokinases: novel targets for chemical biology and drug discovery? *Biochemical Society transactions*, 43(5):1095–1103, 2015.

[314] Céline Crauste, Nicolas Willand, Baptiste Villemagne, Marion Flipo, Eve Willery, Xavier Carette, Martin Moune Dimala, Anne-Sophie Drucbert, Pierre-Marie Danze, Benoit Deprez, et al. Unconventional surface plasmon resonance signals reveal quantitative inhibition of transcriptional repressor ethr by synthetic ligands. *Analytical biochemistry*, 452:54–66, 2014.

[315] H Bonnet, L Coche-Guérente, E Defrancq, N Spinelli, A Van der Heyden, and J Dejeu. Negative spr signals during low molecular weight analyte recognition. *Analytical Chemistry*, 93(8):4134–4140, 2021.

[316] Mazen W Karaman, Sanna Herrgard, Daniel K Treiber, Paul Gallant, Corey E Atteridge, Brian T Campbell, Katrina W Chan, Pietro Ciceri, Mindy I Davis, Philip T Edeen, et al. A quantitative analysis of kinase inhibitor selectivity. *Nature biotechnology*, 26(1):127–132, 2008.

[317] Stefan Knapp. New opportunities for kinase drug repurposing and target discovery, 2018.

[318] Po-Han Chen, Zhenyi Hu, Elvira An, Ifunanya Okeke, Sijin Zheng, Xuanmeng Luo, Angela Gong, Saul Jaime-Figueroa, and Craig M Crews. Modulation of phosphoprotein activity by phosphorylation targeting chimeras (phostacs). *ACS Chemical Biology*, 16 (12):2808–2815, 2021.

[319] Dario R Alessi, Mirjana Andjelkovic, Barry Caudwell, Peter Cron, Nick Morrice, Philip Cohen, and Brian A Hemmings. Mechanism of activation of protein kinase b by insulin and igf-1. *The EMBO journal*, 15(23):6541–6551, 1996.

[320] LLC Schrödinger. The pymol molecular graphics system, version 1.7.2.1. *Schrödinger LLC*, 2015.

Appendix

Appendix Supplementary Information

Table A1. Classes of Pseudokinases. Table updated and modified from [135]. The * indicates that the crystal structure contains a ligand. PDBid color correspond to one organism: black-homo sapiens, red-rattus novergicus or mus musculus, yellow-drosophila melanogaster, orange- Chaetomium thermophilum, green-saccharomyces cerebisiae, purple-sus scrof

Pseudokinase	Degrader glycine rich loop	DFG motif	VAIK motif	HRD motif	Crystal structures	NMR	Crystal Ligands
Class 1							
BIR2 (A.thaliana)	Y	Y	Y	N	4L68	-	-
BUBR1	Y	Y	Y	Y	6JKK,6JKM	-	-
GCN2	Y	Y	N	N	-	-	-
IRAK3	N	Y	Y	N	6RUU, 6ZIW	-	-
MviN (M.tuberculosis)	Y	N	N	N	3OTV, 3OUK, 3OUN, 3UQC	-	-
NRBP1	Y	N	N	N	-	-	-
Pragmin	Y	N	Y	Y	5VE6, 6EWX	-	-
PTK7	N	N	Y	Y	6VG3	-	-
ROP8 (T.gondii)	Y	N	N	N	3BYV	-	-
ROR1	N	Y	Y	Y	6TU9*, 6BA5, 6BAN	5Z55	Y (0LI, Ponatinib)
ROR2	N	Y	Y	Y	3ZZW, 4GT4, 6OSH, 6OSN, 6OSV	-	-
RYK	N	Y	Y	Y	6TUA	-	-
SCYL1	Y	N	N	N	-	-	-
STK40	Y	N	Y	Y	5L2Q	-	-
TITIN	N	N	Y	Y	4JNW	1TKI	-
TRIB1	Y	N	Y	Y	5CEK, 5CEM, 6DC0, 5IGO, 5IGQ	-	-
VRK3	Y	N	Y	N	2JII	-	-
Class2							
CASK	N	N	Y	Y	3C0G,3C0H*, 3C0I*, 3MFR*, 6LNM	-	Y (3'-AMP, 5'-AMP, ANP**)
EPHB6	N	N	N	N	-	-	-
FAM20A	Y	Y	Y	Y	5WRR, 5WRS*, 5YH2*, 5YH3*	-	Y (ATP)
MLKL	Y	N	Y	N	4M67, 4MWI, 5KNJ*, 5KO1*, 6BWK(T357E/S358E), 6LK6(T357AS358A), 6O5Z*	-	Y (6UX, 6UY, LN4)
SRATDalpha	N	N	N	N	2WTK*, 3GNI*	-	Y (ANP, ATP)
TRIB2	Y	N	Y	Y	-	-	-
TRIB3	Y	N	Y	Y	-	-	-
ULK4	N	N	N	Y	6TSZ*, 6U5L*	-	Y (AGS, 3RJ)
Class 3							
PEAK1	Y	N	Y	Y	6BHC	-	-
ROP2 (T.gondii)	Y	N	N	Y	2W1Z,3DZO	-	-
Class4						-	-
ADCK3	Y	Y	Y	Y	4PED, 5I35*	-	Y(ANP)
ANP-A	N	Y	Y	N	-	-	-
BSK8 (A.thaliana)	Y	N	Y	N	4I92, 4I93, 4I94*	-	Y(ANP)
HER3	N	Y	Y	N	3KEX*,3LMG*, 4RIW*, 4RIX (Q790R)*, 4RIY(E909G)*, 6OP9*	-	Y(ANP, ADP, DB8 (Bosutinib))
HSER (GUCY2C)	Y	Y	Y	N	-	-	-
ILK	Y	Y	Y	N	3KMU,3KMW*, 3REP*, 6MIB(L207W)	-	Y (ATP)
IRAK2	Y	N	Y	N	3MOP	-	-
JAK1 JH2	N	Y	Y	N	4L00, 4L01 (V685F)	-	-
JAK2 JH2	N	Y	Y	N	4FVP, 4FVQ*, 4FVR*, 514N (E596A/V617F)*, 5USZ*, 5UT0*, 5UT1*, 5UT2*, 5UT3*, 5UT4*, 5UT5*, 5UT6*, 5WLJ*, 5WIK*, 5WIL*, 5WIM*, 5WIN*, 6BRW*, 6BS0*, 6BSS*, 6D2I (V617F)*, 6G3C (V617F)*, 6M9H*, 6OAV*, 6OBB*, 6OBF*, 6OBL*, 6OCC*, 6XJK*	-	Y(ATP, SKE(JNJ-7706621), 7DZ (BI-D1870), 3YT(PRT062607), IK1 (IKK-2 Inhibitor VI), DQX (NVP-BSK805), 2HB (GLPG0634), 8MY (diaminopyrimidine), AQG (NU6140), 584 (BI-D1870), YDJ (AZD7762), 35R (AT9283), 5BS (XMU-MP-1), E4V, 4SP (NU6102), EKT, J9D, M3A, M3Y, M4G, M4P,M57, V4D
KSR1	N	Y	N	Y	7JUW*, 7JUX*, 7JUY*, 7JUZ*, 7JV0*, 7JV1*	-	Y (ANP, QOM (trametinib)***, VKG (APS-9-95-1) ***
KSR2	N	Y	N	Y	2Y4I*, 5KKR*, 7JUQ*, 7JUR*, 7JUS*, 7JUT*, 7JUU*, 7JUV*	-	Y(ATP , 6U7 (APS-2-79), ADP, QOM (trametinib)***, VKG (APS-9-95-1) ***
PAN3	Y	N	Y	N	4BWK*, 4BWX*(7 mutations**), 4CZY*,4BWP*, 4CYI*, 4CYJ*, 4XR7	-	Y (AGS, AN2, ATP
POMK	Y	Y	N	Y	5GZ8, 5GZ9*, 5GZA(D.rerio)*	-	Y(ANP, ADP)
RNase L	Y	Y	Y	Y	4OAU*, 4OAV*, 4O1O, 4O1P*, 6M11*,6M12*, 6M13*	-	Y (ADP, ACP, ANP, B49(Sunitinib), J60 (SU11652), BWC(Toceranib)
ROP5B (T.gondii)	N	Y	Y	N	3Q5Z, 3Q60*, 4LV5*	-	Y (ATP, ADP)
STKLD1	N	Y	Y	N	-	-	-
TYK2 JH2	N	Y	Y	N	3ZON*, 4OLI*, 4WOV*, 5C01, 5C03*, 5TKD*, 6NSL*, 6NZE*, 6NZF*, 6NZH*, 6NZP*, 6NZQ*, 6NZR*, 7AX4*, 7K7O*, 7K7Q*	-	Y (IK1, 2TT, 3SM (BMS-066), AGS, 7GL, KZJ, L8Y, L91, L9A, LB7, LB4, LAJ, NM7, VZJ, VZG

Table A1. Classes of pseudokinases.

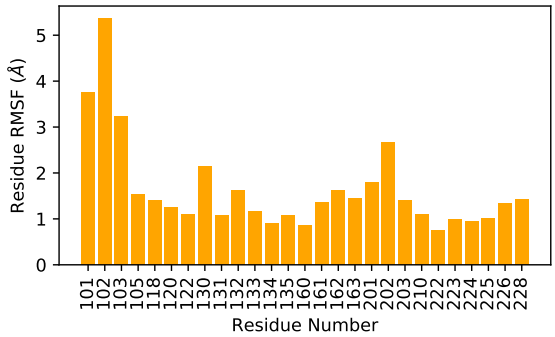


Figure A.1. RMSF of residues at 5 Åfrom GW683134A in the canonical ATP binding pocket of TRIB1.

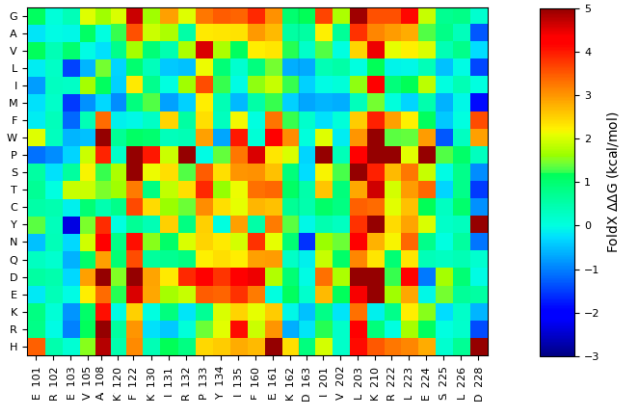


Figure A.2. Heatmap of the mutational scan corresponding to residues at 5 Åfrom GW683134A in the canonical ATP binding pocket.

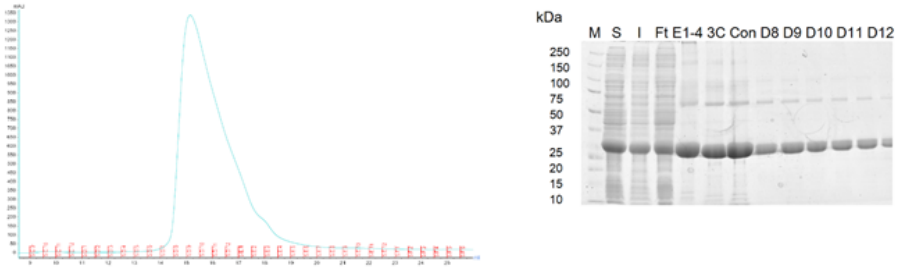


Figure A.3. Size exclusion chromatography and SDS-PAGE analysis of TRIB1ΔCt (84-343).

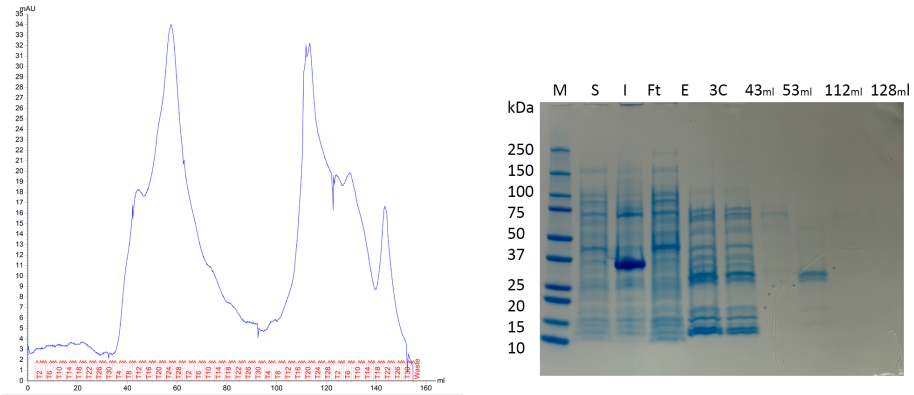


Figure A.4. Size exclusion chromatography and SDS-PAGE analysis of TRIB1 (84-372)-D163I

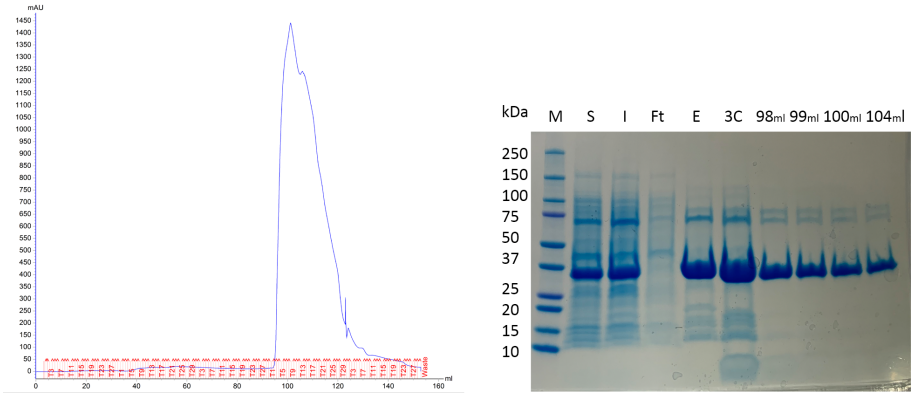


Figure A.5. Size exclusion chromatography and SDS-PAGE analysis of TRIB1 (84-372)-S225F

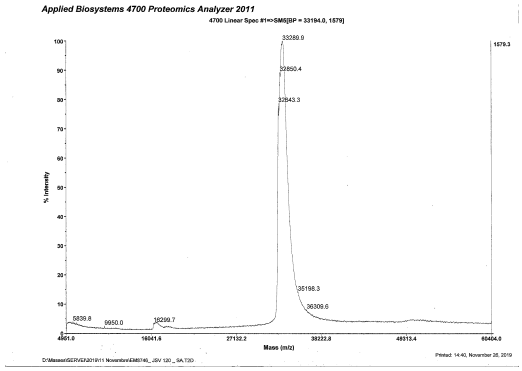


Figure A.6. Matrix-assisted Laser Desorption/Ionization (MALDI) of TRIB1 (84-372) (33 KDa).

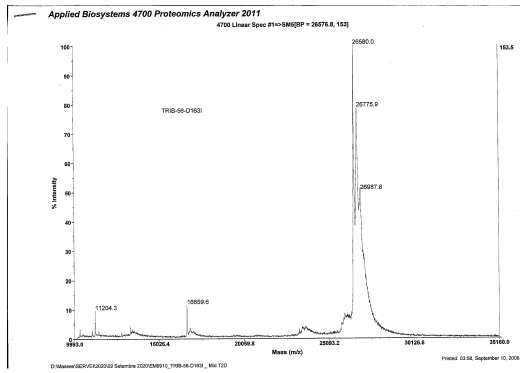


Figure A.7. Matrix-assisted Laser Desorption/Ionization (MALDI) of TRIB1 (84-372)-D163I (33 KDa).

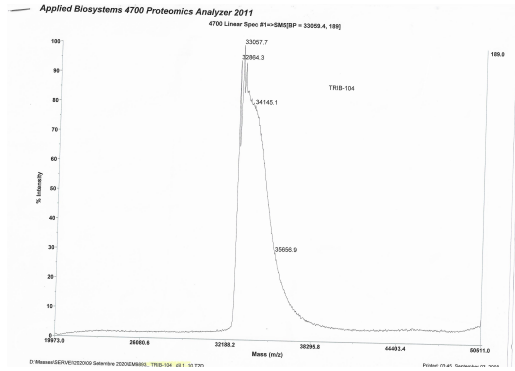


Figure A.8. Matrix-assisted Laser Desorption/Ionization (MALDI) of TRIB1 (84-372)-S225F (33 KDa).

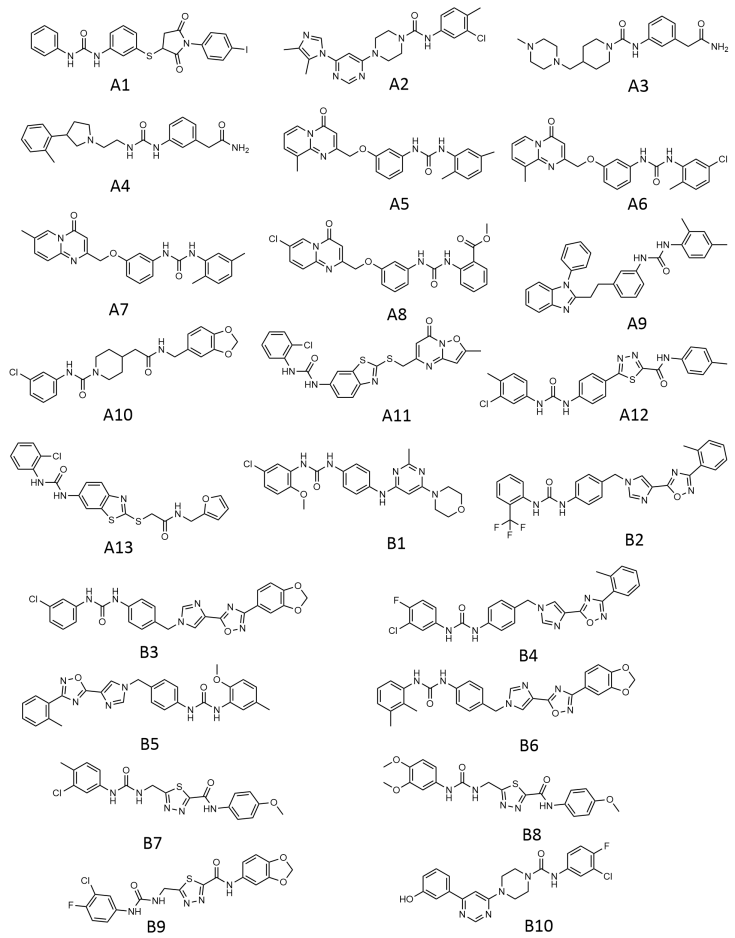


Figure A.9. Chemical structure of the small molecules filtered through the in silico workflow and purchased at MOLPORT. Compounds A1 to B10.

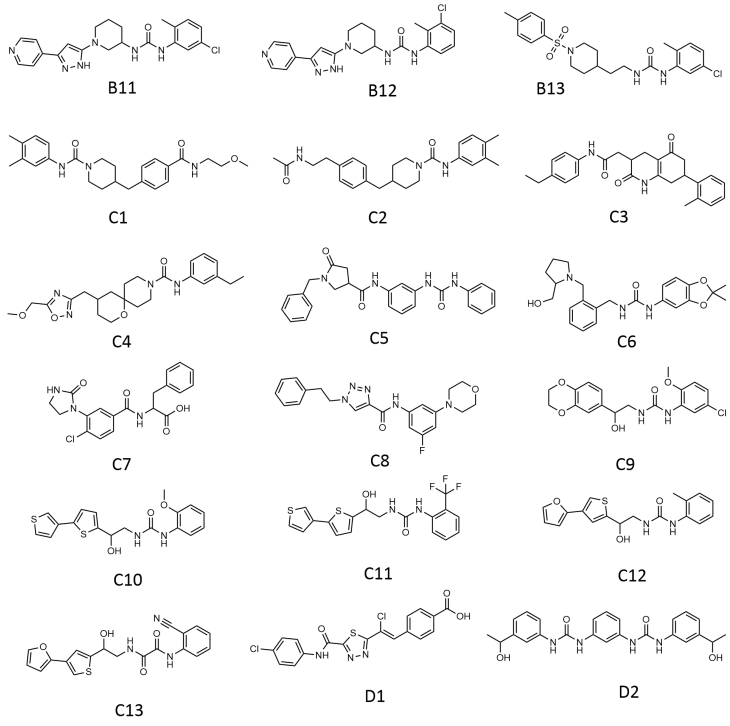


Figure A.9. (Continued). Chemical structure of the small molecules filtered through the in silico workflow and purchased at MOLPORT.Compounds B11 to D2.

Compound	SCORE	SCORE. INTER	SCORE. INTER.VDW	SCORE. INTER.norm	SCORE. INTRA	SCORE. INTRA.norm	SCORE. RESTR	SCORE. RESTR.norm	SCORE. norm
B4	-59,60	-39,46	-42,49	-1,10	3,38	0,09	0,97	0,03	-1,66
B3	-51,93	-39,22	-39,09	-1,06	3,52	0,10	0,93	0,03	-1,40
B5	-58,88	-39,07	-42,80	-1,06	1,79	0,05	0,99	0,03	-1,59
B6	-57,46	-38,21	-37,44	-1,01	2,51	0,07	0,93	0,02	-1,51
B2	-58,48	-38,10	-42,36	-1,00	1,97	0,05	0,99	0,03	-1,54
C3	-54,15	-37,32	-39,71	-1,20	3,87	0,12	0,99	0,03	-1,75
D2	-48,68	-36,97	-34,30	-1,16	8,16	0,26	0,77	0,02	-1,52
A10	-55,86	-36,70	-40,05	-1,22	3,48	0,12	0,79	0,03	-1,86
A6	-54,72	-36,17	-41,43	-1,13	4,75	0,15	0,82	0,03	-1,71
B10	-57,02	-35,94	-39,45	-1,16	1,59	0,05	0,96	0,03	-1,84
C2	-51,45	-35,91	-41,93	-1,20	6,80	0,23	0,84	0,03	-1,72
B13	-49,74	-35,89	-42,62	-1,20	7,17	0,24	0,76	0,03	-1,66
A13	-56,15	-35,64	-41,22	-1,15	0,10	0,00	0,99	0,03	-1,81
C10	-52,55	-35,61	-37,59	-1,42	6,32	0,25	0,96	0,04	-2,10
C11	-49,97	-35,57	-37,31	-1,32	8,42	0,31	0,83	0,03	-1,85
A5	-52,91	-35,36	-40,12	-1,11	4,97	0,16	0,80	0,03	-1,65
A9	-53,74	-35,33	-42,14	-1,01	5,89	0,17	0,87	0,02	-1,54
A8	-49,62	-35,30	-42,19	-1,04	5,38	0,16	0,97	0,03	-1,46
A7	-53,61	-35,23	-40,86	-1,10	4,72	0,15	0,97	0,03	-1,68
D1	-50,68	-35,22	-37,10	-1,30	5,44	0,20	0,89	0,03	-1,88
A12	-52,57	-35,04	-38,08	-1,06	5,90	0,18	0,88	0,03	-1,59
C9	-50,13	-34,84	-35,10	-1,34	6,45	0,25	0,77	0,03	-1,93
A4	-54,47	-34,72	-37,44	-1,24	3,53	0,13	0,81	0,03	-1,95
C7	-51,05	-34,67	-33,83	-1,28	6,53	0,24	0,96	0,04	-1,89
A1	-54,52	-34,65	-41,94	-1,12	2,18	0,07	0,98	0,03	-1,76
C5	-56,99	-34,58	-42,58	-1,08	0,87	0,03	0,86	0,03	-1,78
A2	-57,22	-34,51	-39,98	-1,15	1,40	0,05	0,95	0,03	-1,91
C4	-51,96	-34,46	-40,25	-1,11	4,18	0,13	0,96	0,03	-1,68
A3	-54,39	-34,29	-38,12	-1,27	3,86	0,14	0,93	0,03	-2,01
C1	-47,45	-34,26	-40,90	-1,11	5,12	0,17	0,87	0,03	-1,53
B11	-53,96	-34,25	-39,37	-1,18	3,14	0,11	0,89	0,03	-1,86
B9	-50,51	-34,18	-37,42	-1,14	5,48	0,18	0,83	0,03	-1,68
C13	-51,93	-34,15	-36,14	-1,26	5,40	0,20	0,99	0,04	-1,92
C6	-46,22	-34,14	-32,29	-1,14	4,78	0,16	0,97	0,03	-1,54
C8	-46,96	-34,09	-36,59	-1,18	10,37	0,36	0,94	0,03	-1,62
B12	-52,87	-34,07	-38,65	-1,17	3,16	0,11	0,84	0,03	-1,82
B7	-47,75	-33,99	-33,42	-1,17	8,83	0,30	0,76	0,03	-1,65
C12	-51,07	-33,93	-34,62	-1,41	7,22	0,30	0,92	0,04	-2,13
B1	-50,17	-33,83	-38,53	-1,03	5,09	0,15	0,72	0,02	-1,52
A11	-55,55	-33,82	-39,02	-1,02	1,10	0,03	0,98	0,03	-1,68
B8	-49,65	-33,79	-33,45	-1,09	6,92	0,22	0,81	0,03	-1,60

Table A2. Docking scores of DSF selected compounds

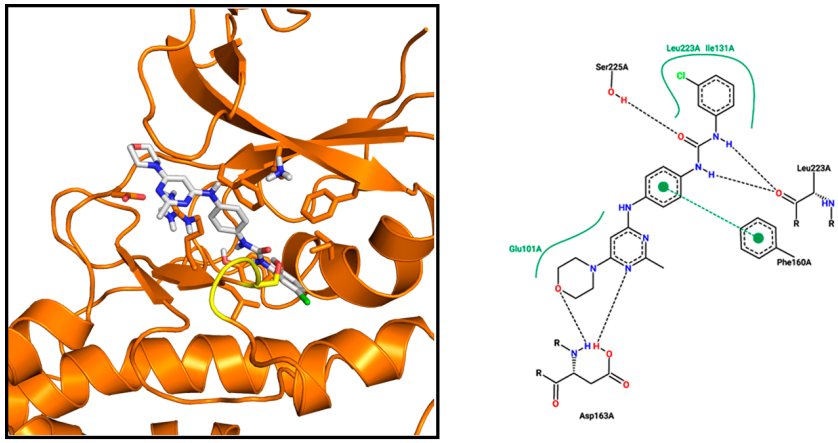


Figure A.10. Compound B1.2 (grey) in the canonical ATP binding pocket of TRIB1(orange) and a 2D interaction diagram by PoseView. [270]

Table A3. K_d (dissociation constant), Rmax (maximum response) and χ^2 (Chi square) of D1, A12, B5, B7, B6, B2, B4 and B3.

	TRIB1 (84-372)		
Compound	Kd (M)	Rmax (RU)	χ^2
D1	-	-	-
A12	105.6	24.35	2.15
B5	-	-	-
B7	287.3	26.99	2.86
B6	342.4	35.5	0.522
B2	2216	145	0.466
B4	-	-	-
B3	322	37	0.991

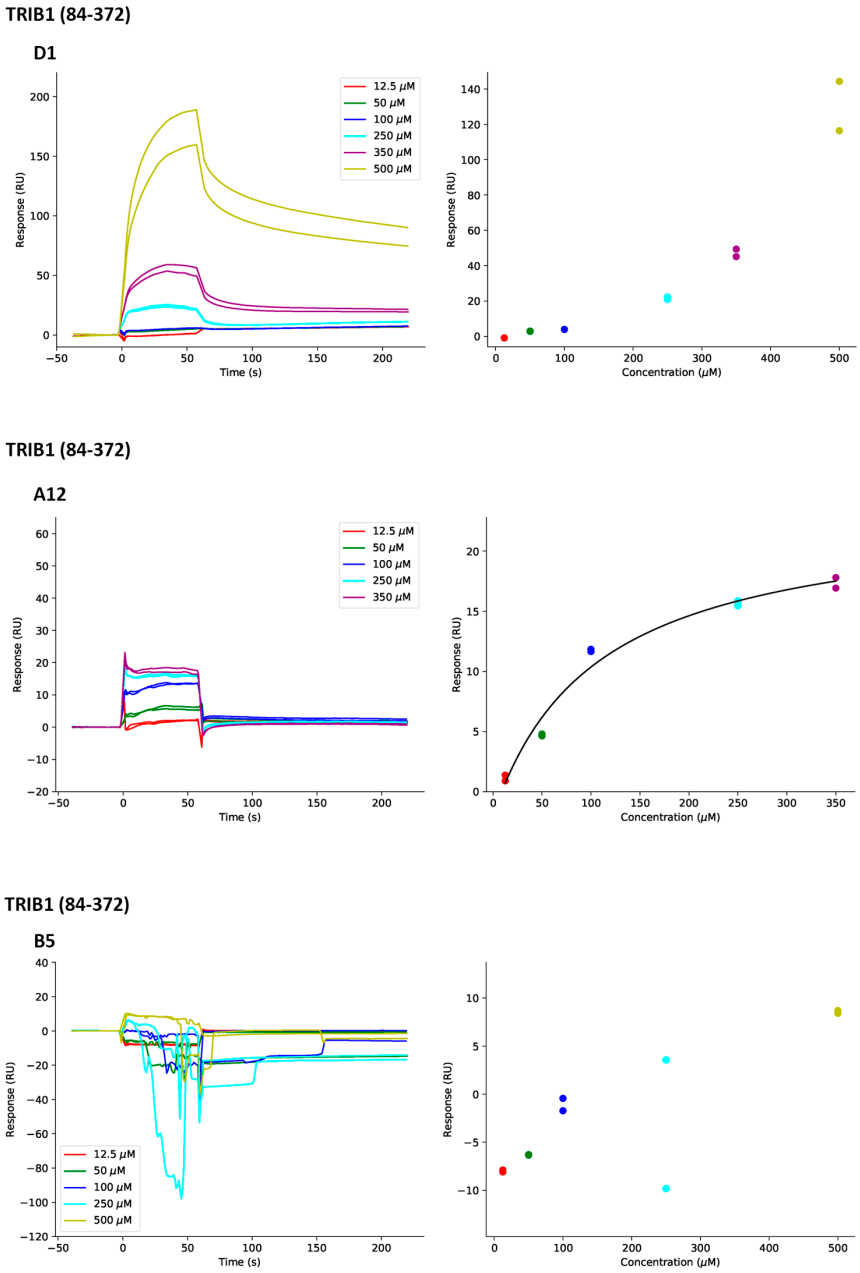


Figure A.11. Sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of compound D1, A12, B5 against TRIB1 (84-372). Values shown are from experimental duplicates and sensorgrams are plotted after removing spikes due to buffer changes.

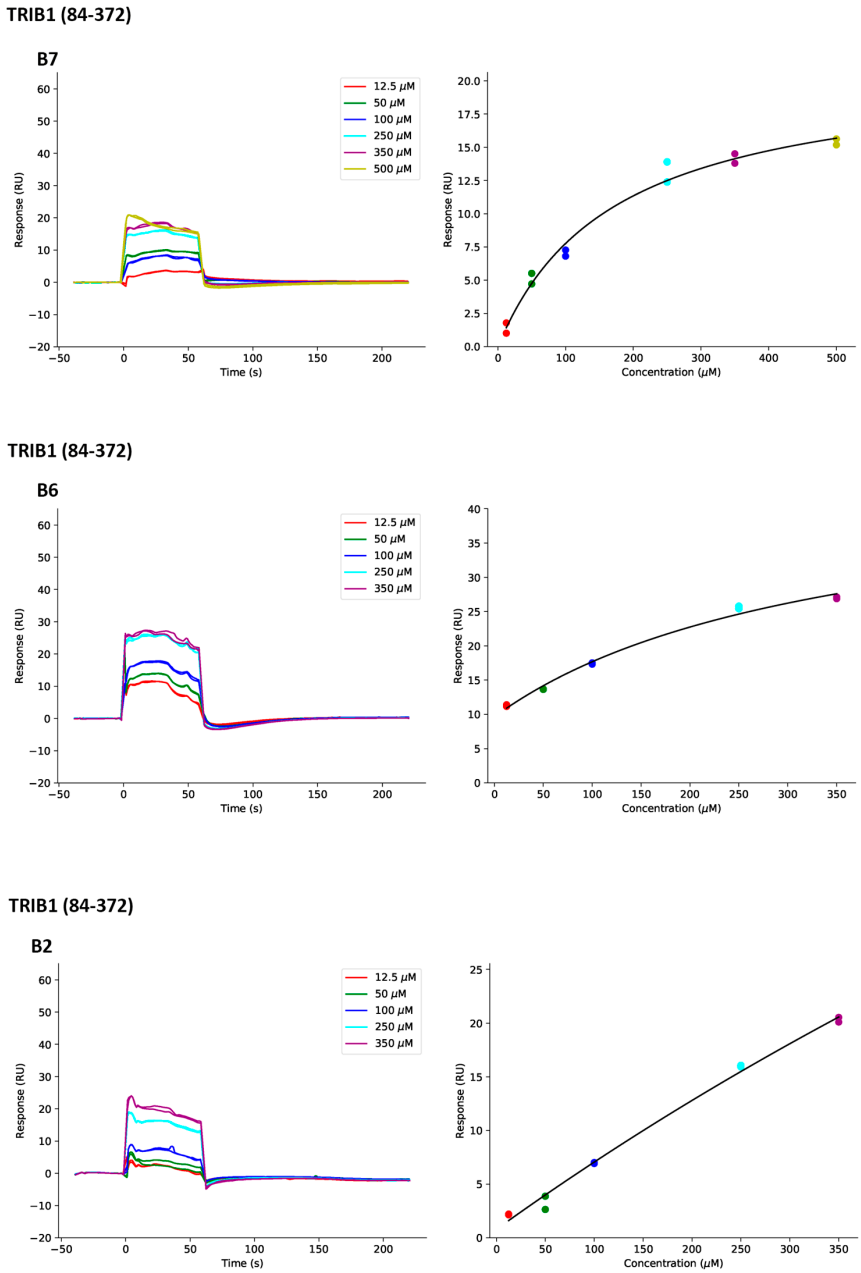


Figure A.12. Sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of compound B7, B6, and B2 against TRIB1 (84-372). Values shown are from experimental duplicates and sensorgrams are plotted after removing spikes due to buffer changes.

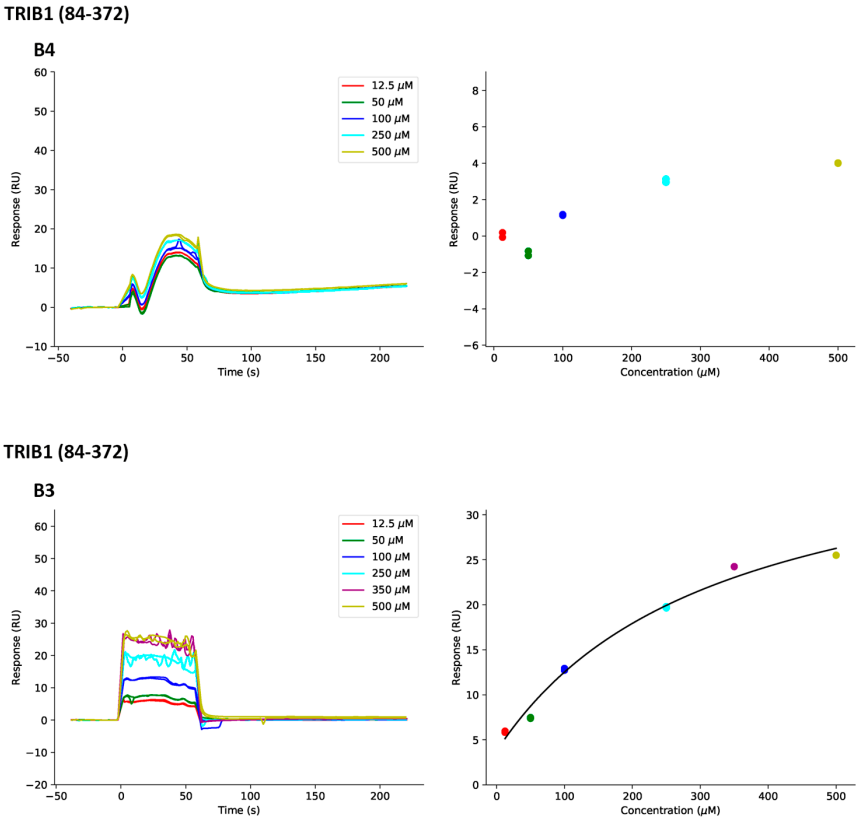


Figure A.13. Sensorgrams (left) and corresponding plot of steady state response against concentration (right) for determination of binding affinity of compound B4 and B3 against TRIB1 (84-372). Values shown are from experimental duplicates and sensorgrams are plotted after removing spikes due to buffer changes.

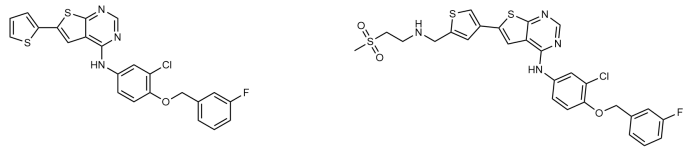


Figure A.14. Chemical structure of GW68462B and GSK-300014 from PKIS. Both compounds were used in the DSF from [211] giving a ΔTm of approximately 2 ºC.

Table A4. List of all kinases tested with B1.2 at 10 μM. The percentage of activity was calculated as the mean of two experiments. SD: Standard deviation.

Kinase	**Mean % activity**	**SD**
ABL	101	12
AMPK (hum)	95	4
ASK1	113	8
Aurora A	108	4
Aurora B	115	13
BRK	104	1
BRSK1	96	11
BRSK2	97	10
BTK	80	26
CAMK1	104	9
CAMKKb	91	10
CDK2-Cyclin A	124	31
CDK9-Cyclin T1	138	50
CHK1	102	0
CHK2	17	4
CK1 2	118	2
CK1	90	26
CK2	92	1
CLK2	87	4
CSK	122	28
DAPK1	125	5
DDR2	69	8
DYRK1A	68	7
DYRK2	95	2
DYRK3	79	17
EF2K	94	9
EIF2AK3	124	6
EPH-A2	99	22
EPH-A4	113	13
EPH-B1	96	31
EPH-B2	42	8
EPH-B3	53	15
EPH-B4	113	27
ERK1	108	0
ERK2	88	4
ERK5	112	5
ERK8	99	11
FGF-R1	103	16
GCK	69	10
GSK3b	109	21
HER4	83	2
HIPK1	118	31
HIPK2	109	12
HIPK3	84	8
IGF-1R	130	24
IKKb	100	7

Table A4. (Continued).

Kinase	Mean % activity	SD
IKKe	94	22
IR	82	21
IRAK1	110	8
IRAK4	98	4
IRR	125	19
JAK3	106	3
JNK1	88	4
JNK2	92	10
JNK3	100	29
Lck	105	11
LKB1	96	2
MAP4K3	81	2
MAP4K5	91	4
MAPKAP-K2	127	9
MAPKAP-K3	94	20
MARK1	87	3
MARK2	85	4
MARK3	85	9
MARK4	108	13
MEKK1	84	6
MELK	105	24
MINK1	120	9
MKK1	109	13
MKK2	82	7
MKK6	99	1
MLK1	81	3
MLK3	108	2
MNK1	108	15
MNK2	113	9
MPSK1	95	2
MSK1	113	17
MST2	93	10
MST3	101	13
MST4	86	0
NEK2a	105	19
NEK6	107	15
NUAK1	70	8
OSR1	96	1
p38a MAPK	81	32
p38b MAPK	105	2
p38d MAPK	56	5
p38g MAPK	64	13
PAK2	89	8
PAK4	115	4
PAK5	80	2
PAK6	112	7
PDGFRA	85	4

Table A4. (Continued).

Kinase	Mean % activity	SD
PDK1	104	31
PHK	122	25
PIM1	87	12
PIM2	106	6
PIM3	83	5
PINK	90	0
PKA	109	8
PKBa	100	7
PKBb	114	16
PKCa	116	11
PKCz	110	2
PKC	107	12
PKD1	100	21
PLK1	81	7
PRAK	99	27
PRK2	102	4
RIPK2	91	2
ROCK 2	103	28
RSK1	95	41
RSK2	117	3
S6K1	120	13
SGK1	87	13
SIK2	106	1
SIK3	108	0
SmMLCK	98	14
Src	79	10
SRPK1	106	15
STK33	110	2
SYK	101	23
TAK1	118	19
TAO1	102	14
TBK1	128	22
TESK1	128	32
TGFBR1	90	37
TIE2	106	11
TLK1	98	9
TrkA	57	6
TSSK1	94	1
TTBK1	83	6
TTBK2	103	1
TTK	96	5
ULK1	104	4
ULK2	121	18
VEG-FR	82	16
WNK1	130	3
YES1	115	18
ZAP70	103	7

.

Table A5. List of Lipid kinases tested with B1.2 at 10 μM. The percentage of activity was calculated as the mean of two experiments. SD: Standard deviation.

Lipid kinase	**Mean % activity**	**SD**
CHKa	103,7	0,2
CHKb	97,7	0,5
DGK beta	98,9	0,6
DGK gamma	99,6	7,2
DGK zeta	93,5	8,8
PI3K alpha	77,2	2,0
PI3K beta	86,3	1,8
PI3K E545K+p85	80,9	1,9
PI3K gamma	102,6	0,2
PI3Ka E524K + p85	105,6	4,3
PI4K2a	89,3	1,7
PIK4Cb	75,2	1,0
PIP5K2a	97,6	6,9
SPHK1	98,5	5,4
SPHK2	99,2	0,4

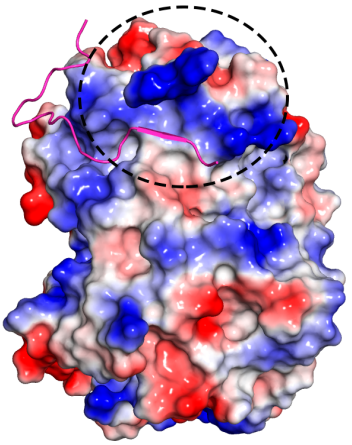


Figure A.15. Electrostatic surface potential of AKT KD. The C-terminal tail of AKT (magenta) is shown close to the basic patch in the N-lobe, inside the dashed dot circle. The electrostatic potencial was calculated with the APBS (Adaptive Poisson-Boltzmann Solver) plugin of PyMOL [320]. The potentials are on a red–white–blue color map for negative-neutral-positive charged residues.