2019

In silico screening of a chemical library based on pharmacophores for potential RAF-protein kinase inhibitors.

JURRIAAN VAREKAMP
S3087034

29-11-2019 | Rijksuniversiteit Groningen

Inhoudsopgave

Introduction

Cancer

Cells in the human body reproduce and die to make space for the new cells. However, when cell reproduction goes wrong, cancer cells can be generated. Cancer is the name used for diseases, which induce the uncontrolled and abnormal growth of cells. The induction of uncontrolled and abnormal cell growth can be due to numerous reasons. One of those reasons can be a virus, e.g. the human papillomavirus (HPV), or a bacterium like *Helicobacter pylori (H. pylori).* However, cancer can also be induced by the own human body through a fault in the DNA. This can be due to one wrong inserted nucleotide or a whole gene relocation. Even though most of the time the cells with damaged DNA strands will be disposed off or killed by the immune system of the body, this will not always be the case, and when the immune system is not able to dispose or kill the cancer cell, it will grow and form into a tumour. (1)

Cancer survival rate for the first 5 years after diagnosis is around 65% for most people. This is a substantial increase from around 35% in the 1960s. However, around 35% of people still pass away within 5 years after diagnosis. There are many different types of cancers and in order to treat these cancers efficiently, a system is used to determine the severity of the cancer. There are multiple systems to characterize the severity of the cancer. The most commonly known system is through TNM staging. The size and extent of the primary tumour refer to T, N is the absence or presence of cancer in nearby lymph nodes and M indicates if there are any metastases of the primary tumour. Subsequently, the T, N and M are divided into 5 stages. These are the stages 0, I, II, III, IV. With 0 having no abnormal growth being measured, I being early-stage cancer and IV being the most advanced stage for the cancer. This means the first stage is the least severe and the last stage is the most severe. The induction of cell proliferation and the inhibition of cell apoptosis by kinases in cancer is one of the main reasons the cell keeps growing uncontrollably. (1)
The sheer vastness of the number of different types of cancers and their ability to gain resistance against existing drugs means that many different treatments against cancer need to be researched and refined continuously in order to optimize treatment and patient care.

Kinases

Kinases are enzymes that phosphorylate other molecules as seen in figure 1. They do this by carrying a phosphate from an ATP molecule to various substrates by binding it to the hydroxyl groups on these substrates. These hydroxyl groups are mostly present in tyrosine, serine, and threonine. Protein kinases phosphorylate either the amino acids; tyrosine (tyrosine-specific protein kinases; TPKs) or serine-threonine (serine-/threonine-specific kinases; STPKs) and in some cases even both tyrosine and threonine (dual-specificity protein kinases) in eukaryotic cells like in humans. Abnormal regulation of the phosphorylation of protein kinases in eukaryotic cells can lead to a multiple of disease in human. Such as inflammatory diseases, cardiovascular diseases, and even cancer. This is due to the fact that kinases regulate multiple pathways that lead to cell proliferation or inhibit cell apoptosis. The most important of these pathways are the RAS-, ERK- and MEK-pathways. (2)(3)

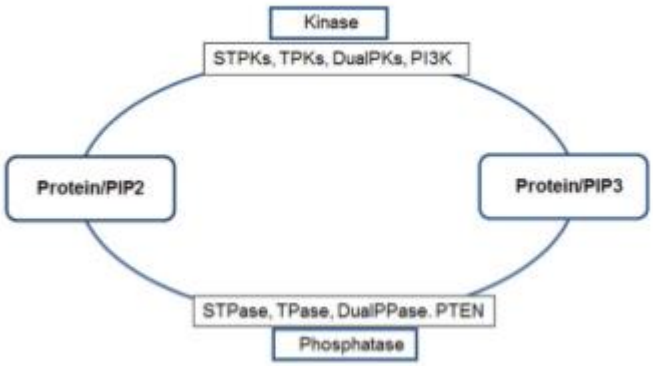

Figure 1: Phosphorylation and dephosphorylation by respectively different kinases and different phosphates. (2)

The catalytic domain of protein kinases is highly conserved in their structures. Meaning the structures in this domain are almost identical in most if not every kinase despite the function of the kinase. Although most active kinases share the same conformation, inactive kinases do not. Inactive kinases have a wide range of varied conformations between them. Some of these conformations expose pockets that can be targeted by kinase inhibitors. These pockets are mostly referred to as DFG-out and C-helix-out. The DFG-out conformation leads to partial blocking of the active site. This means ATP is not able to bind to the active site anymore. Therefore, the kinase cannot activate and will not be able to phosphorylate any protein. (2)

In contrast, with DFG-out ATP is still able to bind to kinases with a C-helix-out. However, the catalytic activity will be reduced due to disruption of the catalytic machinery. (3) A very important part of kinases is the gatekeeper region because this region differs between most of the kinases and mutation of this region can lead to resistance against kinase inhibitor. Another important region is the hinge region close to the gatekeeper. Protein kinase inhibitors bind to this region to get a foothold and therefore better binding affinity with the kinase. (2)

Kinases can be inhibited in many different ways. One of these ways is through covalent bonding. This can be irreversible binding but also reversible depending on the reactivity of the inhibitor. Inhibiting kinases by covalent binding happens mostly through binding Lysine or a Cysteine active site in near proximity of the ATP binding site. However binding of covalent inhibitors can be prevented by the kinase when the Cysteine active site is mutated. This is especially interesting because the mutation of Cysteine has no effect on the structure or activity of the kinase. Another way to inhibit kinases is by non-covalent binding. Non-covalent binding of protein kinases can occur by binding of the hinge region or another region. This is further classified in type-1, type-2 and type-3 reversible kinase inhibitors. (2)

Most of the type-1 inhibitors bind to active kinase conformations. Type-1 inhibitors gain most of their selectivity for kinases through the gatekeeper region, because this region is different in various kinases. Type-1.5 is a subgroup of type-1 inhibitors, which bind to inactive kinase conformations (e.g. vemurafenib against B-rapidly accelerated fibrosarcoma, BRAF). In the case of vemurafenib, the drug induces the C-helix into an C-helix-out conformation. This leads to disruption of ion-pairing between Lys and Glu in the C-helix. Even though BRAF still has a DFG-in conformation, the kinase cannot be activated due to the C-helix-out conformation. (2)

Type-2 inhibitors bind to protein kinases, which have the inactive conformation. Type-2 inhibitors also have contact with the hinge region of the protein kinase. These type of protein inhibitors are also ATP competitive inhibitors as mentioned before. Type-2 inhibitors bind the inactive DFG-out conformation. In contrary to the type-1.5 inhibitors, since they bind the DFG-in inactive conformation with the C-helix-out. The binding of a type-2 inhibitor makes sure that the

DFG-in confirmation cannot form. Because of this, the protein kinase will not be activated. (2)

The type-3 kinase inhibitors are the last group of kinase inhibitors. This group exists mostly of non-competitive ATP inhibitors and have no contact with the hinge region. However, some of the type-3 inhibitors compete indirectly with ATP by binding and locking-in an inactive conformation of the protein kinase to which ATP is unable to bind. Type-3 inhibitors work mostly by binding to an inactive kinase conformation. By binding to an inactive kinase conformation, the active conformation cannot be achieved. In addition, because type-3 inhibitors bind to other regions than type-1 and type-2, which bind to mostly conserved regions, they have a better selectivity for the kinase they are inhibiting. (2)

One of the ways kinases may become resistant to drugs is by structural changes to the kinase. It has been revealed that mutations in the structure of the gatekeeper region can lead to drug resistance. The gatekeeper region is an important part for the binding affinity of the kinase inhibitor to the protein kinase. Structural changes in the gatekeeper region may lower the affinity of the inhibitor and therefore induce drug resistance. (4) Another way kinases are resistant to drugs is because of dimerization of the kinase. Dimerization of kinases, especially with BRAF, lead to the drug being unable to inhibit the ERK-pathway. Because of this cell proliferation will still be increased and cell apoptosis will still be decreased. (5)

RAF-protein kinases and their roles in cancer

Cell proliferation is managed by many different pathways. One of which is the RAF-pathway as seen in the figure below (figure 2).

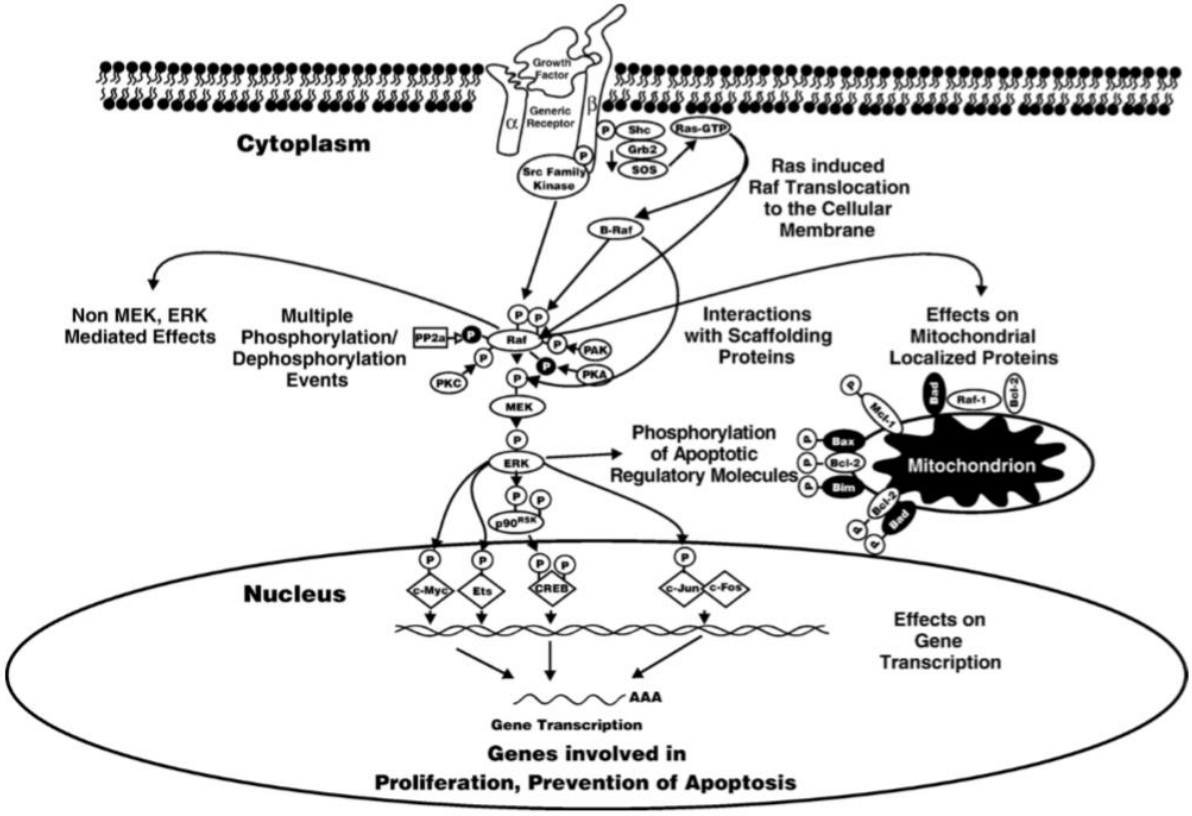

Figure 2: RAF-signaling pathway with BRAF in human cells. (6)

The RAF-pathway induces the activation of RAS, which in turn activate the kinases ARAF, BRAF, and CRAF. These protein kinases have a globally similar structure as depicted in the figure below (figure 3). Their structure consists of a regulatory domain that is activated by RAS when bound to the RAS-

binding domain (RBD). When activated the cysteine-rich domain (CRD) is relieved and will no longer inhibit the kinase domain, meaning the kinase is activated. However, the kinase domain of BRAF has, unlike other kinases, an atypical conformation. In this conformation, an intra-molecular protein-protein interaction is made between the activation segment and the glycine-rich loop. This conformation leads to an inactive state of the protein kinase BRAF. However, some mutations can disrupt this interaction between the activation segment and the glycine-rich loop. One of these mutations is the Valine amino acid at position 600, BRAFV600E. This mimics phosphorylation which would normally take place at residues surrounding V600 resulting in folding of the kinase into the active conformation. (7)(8)

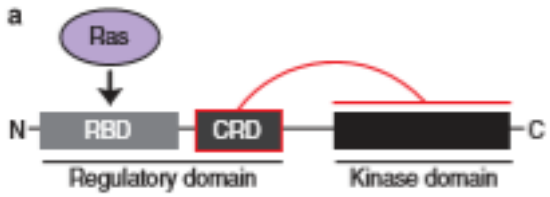

Figure 3: General structure of RAF-protein kinases with the blunt red arrow indicating inhibition (a.k.a. a negative feedback loop). (7)

The activation of these protein kinases leads to the induction of growth-promoting-genes (GPG). Growth-promoting-genes either induce cell proliferation directly or inhibit cell death by preventing apoptosis. Most of the time these protein kinases get activated and deactivated when the equilibrium shifts. However, a mutation in the BRAF-gene, in particular, creates BRAFV600E, which induces the activation of the kinase, as mentioned above, and is highly insensitive to the negative feedback mechanisms that are in place. This means that in contrary to normal RAF-protein kinases, BRAFV600E cannot be stopped by normal negative feedback loops that exist in the human body. Therefore, BRAFV600E will continuously trigger the ERK- and MEK-pathways, leading to the transcription of growth-promoting-genes. This leads to uncontrolled cell growth since BRAFV600E is highly insensitive to the negative feedback loop. BRAFV600E can also induce evasion of the immune system in cancer cells. BRAFV600E does this primarily through activation and promotion of interleukin-10 and -6 in cancer cells, but also through a decreased expression of melanoma antigens by the MEK/ERK-dependent pathway. Because of this, the cancer cells will not be recognized by T-lymphocytes. Even though inhibiting the MEK-pathway decreases the effectiveness of T-lymphocytes, inhibiting BRAF does not. Meaning that inhibition of BRAF does not jeopardize the cancer treatment by increasing the immune evasion of the cancer cells. BRAFV600E can also facilitate cancer cell migration. BRAFV600E increase the production of interleukin-8 as well in cancer cells. In its turn, interleukin-8 produces polymorphonuclear leukocytes. These leukocytes enable the trans-endothelial passage of cancer cells by binding to these cells. (6)(7)(8)(9)

Other factors like metabolic stress in the form of a lack of sustenance in cancer cells must be bypassed for that cancer cell to continue growing and proliferation. BRAFV600E also has a solution for this problem. Through activation of the ERK-pathway by BRAFV600E, the human tumour suppressor LKB1 will be phosphorylated. Due to this phosphorylation LKB1 is no longer able to activate the AMP protein kinase. AMP protein kinase is needed to halt the synthesis of proteins. Because of this, the cancer cells will still be able to proliferate, even though there are not enough nutrients in its surroundings. This gives cancer cells a competitive edge over other (non-cancer) cells in nutrient deprived areas. Studies have shown that mutations in BRAF are present in 40-70% of

melanomas. Of all the mutations, is the BRAFV600E mutation with a valine amino acid at position 600 the most common with a BRAF-mutation frequency of 95% in melanomas. (6)(7)(8)(9)

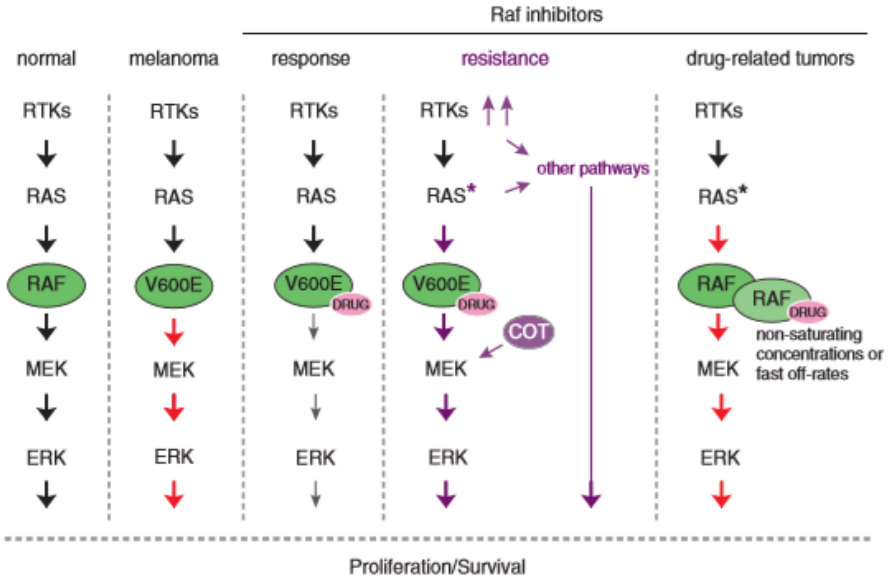

Figure 4: RAS-pathway under different circumstances (circumstances are depicted in the top of the picture). COT is an synonym for Mitogen-activated protein kinase kinase kinase 8 (MAP3K8), which is able to induce resistance against RAF-inhibitors by activating the MEK-pathway. (7)

CRAF (a.k.a. RAF-1) is known for its capabilities as a partner and activator in the MEK/ERK pathways. CRAF has some mutated forms that can induce cell proliferation continuously just as BRAF does. Although, the percentage of CRAF mutated forms that are present in cancers with is low, 1%. However, it is described that overexpression of CRAF is present in a lot of cancers. This overexpression of CRAF may be the reason why cancer cells are unable to undergo apoptosis because CRAF is able to bind to and inhibit stress-induced kinase ASK-1. Differentiation of cells lead to the inhibition of proliferation signals. Therefore, if a cancer cell were to receive differentiation signals the cell would also inhibit most of its proliferation signals. However, CRAF is known to maintain the undifferentiated cells in some cancers. CRAF is able to activate ERK/MEK-dependent and –independent pathways contributing to the undifferentiated state of the cancer cells or bind directly to Rok-α, what leads to inhibition of cell migration, apoptosis, and differentiation. It is clear that CRAF plays an important role in many cancers and that inhibition of CRAF may lead to less proliferation and cell growth. (7) It is also described that other mutations of BRAF may initiate activation of CRAF-protein kinases in the cells. This would also lead to overexpression of CRAF and therefore continuous activation of the MEK- and ERK-pathways. Apart from this CRAF is also known to be able to induce cell survival independent from the MEK-pathway. (7)(8)

RAF-protein kinases drug resistance against cancer treatment

While it is known that mutations in the gatekeeper region do induce resistance against multiple inhibitors since the drug has a lower affinity to BRAF with a mutated gatekeeper. (4) It is also true that these mutations in the gatekeeper area have not been observed and therefore do not pose a threat to BRAF-inhibitors yet. There is also a strong indication that the resistance of BRAF to inhibitors like vemurafenib is not due to the inability of the inhibitor to bind BRAF as would be expected. It is most likely that resistance will occur when BRAF-inhibitors are unable to inhibit the ERK-pathway. This can happen when RAS-GTP levels increase and when RAS-independent dimerization is induced. (5) Dimerization of RAF-protein kinases happens mostly in cells

neighbouring the cancer cells that do not have the BRAFV600E mutation. Due to this dimerization of RAF-protein kinases the ERK- and MEK-pathway will still be activated, since only one component of the dimer needs to be active to activate these pathways. This paradoxically means that RAF-kinase inhibitors may lead to induction of the ERK- or MEK-pathway instead of inhibiting it as depicted in figure 4. Due to this reason, the RAF-inhibitors, especially BRAF-inhibitors, need to be given at saturated concentrations. The repercussions are potentially huge if this does not happen. (e.g. benign skin tumours.) (7)

Structure-Based drug design and in silico screening (molecular docking)

Drug design is nothing new. Normally an x-ray crystal structure would have to be made of the target, in this case, a protein, in order to acquire knowledge about the mode of action of the target. This would, in turn, be used to determine which molecules could act upon the molecule either to inhibit or induce a reaction. It is also possible to make crystal structure of a drug and its target molecule in order to see where the two bind together and why they would bind together by uncovering it's mode of action. (10) Nowadays, it is possible to reconstruct these structures of the target molecule on the computer with the knowledge obtained from x-ray crystallography and nuclear magnetic resonance (NMR). Because of this, it is also possible to screen known drugs or molecules against this structure to potentially use this drug or molecule against the target. This is called in silico screening or virtual screening (VS). In silico screening is one of the many ways to use a computer in structure-based drug design (SBDD) and happens most of the time through a stepwise process with molecular docking. The first step is virtually recreating the known three-dimensional structure of a molecular target. Afterwards, molecular modelling takes place in order to design and then synthesize new ligand compounds, if new bioactive small-molecules are discovered. When the potentially useful compounds are synthesized, they will be evaluated on their biological properties with multiple experiments. When the compound is active a ligand-receptor complex can be constructed that can help by understanding the process of molecular recognition by identifying intermolecular features. This way structural information can be correlated to the biological activity data. With this new information, the molecular modelling starts anew to create compounds with stronger binding affinity to the target binding site. (11)

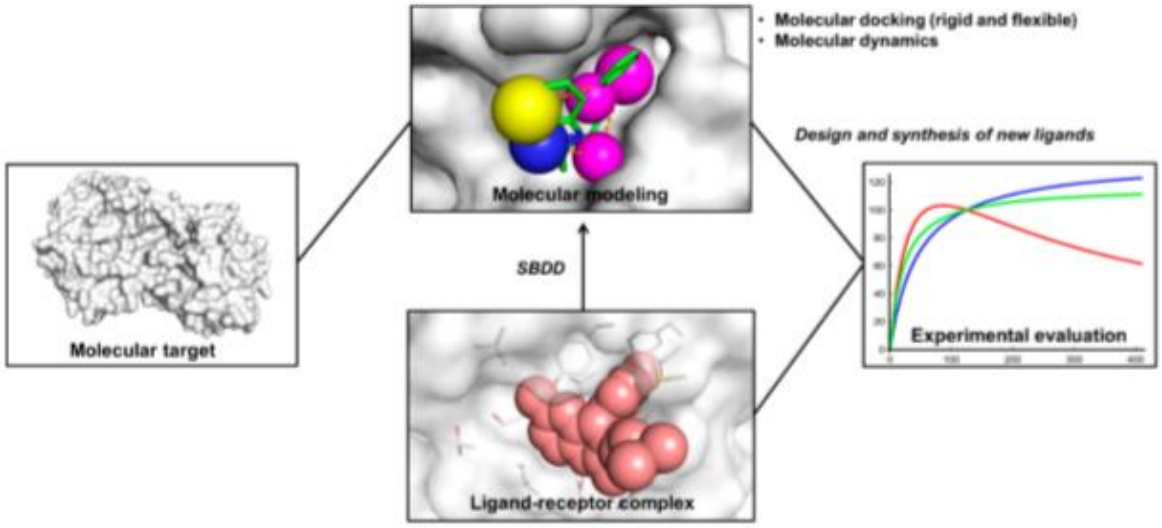

Figure 5: The steps of structure-based drug design (SBDD). (11)

One of the most used techniques in SBDD is molecular docking. In the molecular docking process, a ligand is docked to a receptor. To be precise the ligand is docked to the binding cavity of the receptor. This way multiple conformations of the ligand are researched. When this is done the

most likely binding conformation is identified with the corresponding intermolecular interactions. In order to identify the most likely binding conformations, two tasks must be performed by the program. The first task is exploring a conformational space that represents multiple potential binding modes. The second task is giving an accurate prediction interaction energy that corresponds with the predicted binding conformations. Molecular docking has the ability to predict the structure of small-molecule ligands within the target binding site with high accuracy. Molecular docking became a useful tool to investigate important molecular events in a convenient way. This includes ligand binding modes and their corresponding intermolecular interactions that stabilize the ligand-receptor complex. In addition, algorithms of molecular docking are able to do quantitative predictions of binding energetics. The algorithms are able to give rankings of docked compounds with these predictions. These rankings are based on the binding affinity of ligand-receptor complexes. The molecular docking programs are able to do this continuously until a solution of minimum energy is reached. (11)

In order to reach the solution with the minimal amount of energy, a conformational search is performed by the molecular docking program. In order to achieve this, the program modifies the structural parameters of the ligands to uncover which has the lowest energy solution. Molecular docking programs can do this in two separate ways; systematic search methods or stochastic methods. The systematic search method executes gradual changes in the conformations of the ligands by small variations in the parameters of the structure. The program does this by sampling the energy landscape of the conformational space. After multiple cycles, the program reaches a minimum energy solution associated with the binding mode that is most likely. The downside to the systematic search method is that it might reach a local minimum instead of a global minimum. However, this can be dealt with by executing multiple searches from different starting points in the energy landscape. The stochastic method, on the other hand, generates groups of molecular conformations and sampling a wider range of the energy landscape. This way the conformational searches are carried out by promoting random modifications to the structural parameters of the ligand. With this method, the probability of finding a global minimum is increased because a wider range of the energy landscape is covered. (11)

As mentioned before virtual screening is one method in SBDD to search for promising compounds in already established databases. Virtual screening can be divided into two main approaches. One important approach in virtual screening is structure-based virtual screening (SBVS). Another approach is ligand-based virtual screening (LBVS). In SBVS a target binding site is selected and compounds available in the database are docked into that target binding site. SBVS then ranks the docked compounds and does a prediction of the binding mode. SBVS exist out of four steps. Step one is the preparation of the molecular target. The second step is the selection of the compound database. The third step is molecular docking and the final step is post docking analysis of the structure. Most of the time multiple structures are available of the target compound. When this is the case, structural resolution and conformational changes are essential details to take into consideration when selecting the most suitable structure. Afterwards, a preparation procedure takes place to fully prepare for molecular docking studies. First hydrogen atoms are added and water molecules are removed, except for those important for the interaction between the target and potential binding compound, then the correct tautomerization and protonation states of the target binding site residues are specified and partial charges are calculated. LBVS, on the other hand, is founded on the exploration of molecular properties gathered from active compounds. The molecular properties are gathered and applied as a filter to sieve through the database. The primary usages for database filtering methods is to select compounds that can be used for experimental evaluation or

to reduce the chemical space that needs to be explored in further screening. Another use for LBVS is to collect structural features from ligands that are known to generate pharmacophore models. (11)

Problem

Known drugs against RAF-protein kinase inhibitors

Most protein kinase inhibitors are forming hydrogen bonds with the hinge region of these protein kinases. Mostly, these are one or two hydrogen bonds that will form between the hinge region of the protein kinase and the inhibitor, but some have been observed forming three hydrogen bonds. However, none of the inhibitors have been observed forming four or more hydrogen bonds with the protein kinase. All these protein kinase inhibitors have an imidazole-like structure. These imidazoles are able to make hydrogen bonds close to or with the gatekeeper region in the protein kinase because of the nitrogen that is present. (13)

One of the drugs that work against BRAF and/or CRAF is sorafenib. Sorafenib competes with ATP for the binding site in BRAF-molecules. Crystal structures of sorafenib and BRAF have shown that the distal pyridyl ring of sorafenib interacts with three amino acids in the ATP adenine binding pocket. These ATP adenine binding pockets are also present in CRAF and some mutational isoforms for RAF. Meaning that sorafenib is also able to inhibit other RAF isoforms by stabilizing the inactive conformation. Because of this, it is reasonable to assume that sorafenib is able to delay the growth of human melanoma cells in mice. (8)

PLX4720 is a BRAF-inhibitor as well. PLX4720 is able to facilitate the displacement of the Phe595 residue present in BRAF. This displacement enables the DFG motif to be out of the ATP binding pocket (a.k.a. DFG-out conformation as mentioned before). The DFG-motif exists of Asp-Phe-Gly amino acids. However, occupancy is low in this state (60%) when compared to the occupancy in the DFG-in conformation (100%), meaning that PLX4720 has more affinity for the active conformation than the inactive conformation. This is highly likely due to the sulphonamide group of PLX4720 that forms a hydrogen bond with the NH group of the main chain in Asp594. At the same time oxygen atoms on sulphonamide form hydrogen bonds with the side chain of Lys483 and the backbone of Phe594. PLX4720 binds to the kinase domain that is near the hinge region. More specifically, PLX4720 is able to bind between the C and N lobes in the kinase domain. The hinge region also overlaps with the binding site of ATP. The azaindole that is present in PLX4720 interacts with the hinge region of BRAF through hydrogen bonds, making sure the compound is anchored in its place. The most important reason that PLX4720 is specifically inhibiting the active conformation of RAF-kinases is the propyl group, because the propyl group is able to bind RAF-selective pockets in the active conformation. These RAF-selective pockets are almost only found in RAF-protein kinases and very few other kinases. As long as inhibitors have a higher selectivity for the active conformation

and the associated RAF-selective pocket than the inactive conformation, they should have a higher affinity for RAF-protein kinases than other non-RAF kinases. (14)

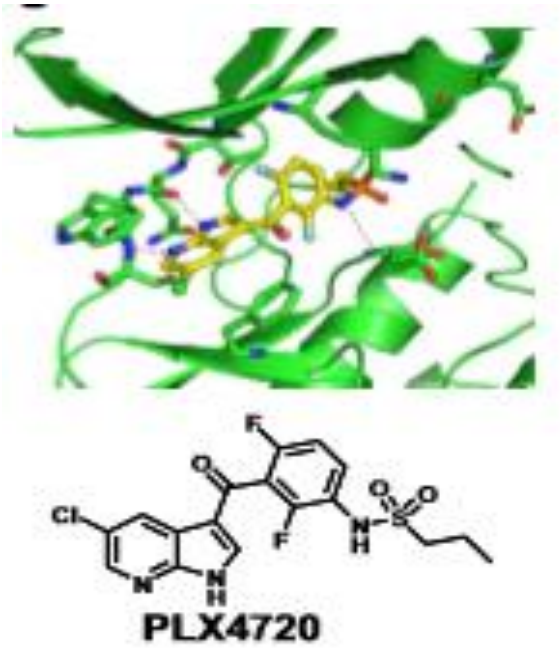

Figure 6: Three-dimensional structure of the PLX4720 and BRAF complex on the top and a two-dimensional structure of PLX4720 on the bottom. (14)

BAY43-9006 is also a BRAF-inhibitor and is able to form multiple interactions with BRAF. A three-dimensional structure of the BAY43-9006 and BRAF complex is shown in figure 7. The groups from BRAF that form interactions with BAY43-9006 are conserved and present in CRAF as well. The interactions that are formed are three hydrogen bonds from the distal pyridyl ring of BAY43-9006 with three aromatic groups of BRAF. One of the aromatic groups that interact with the inhibitor are Trp530 of the hinge region. Another one is the Phe582 groups, which is present at the end of the catalytic loop. The last group is Phe594 of the DFG motif. In addition, the Phe594 residue also interacts with the central phenyl ring of the inhibitor. Other residues also interact with the central phenyl ring of the inhibitor like the aliphatic side chains of Lys482, Leu513, and Thr528. Furthermore, the urea group that is present in the inhibitor also interacts with BRAF by forming two hydrogen bonds. One hydrogen bond is made by the amide nitrogen atom that interacts with the carboxylate side chain of Glu500 in the catalytic domain. The other hydrogen bond is made by the carbonyl on the inhibitor that interacts with the nitrogen of the main chain in Asp593 of the DFG-motif. A nitrogen atom instead of a carbon atom in the ring of the pyridyl moiety increases the affinity of the inhibitor to the protein kinase 5-fold. Not only van der Waals interactions are present, although dominant when BAY43-9006 is inhibiting BRAF. It has been reported that BAY43-9006 also has polar interactions contributing to the formation of the BAY43-9006-BRAF complex. Due to the interaction of BAY43-9006 with the phenyl ring of Phe594 is BAY43-9006 able to procure an inactive conformation of the DFG-motif of BRAF. In addition, BAY43-9006 contributes to the inactive

conformation is due to the trifluoromethyl phenyl moiety that is inserted into the site that Phe594 would have occupied if BRAF were in the active state. (15)

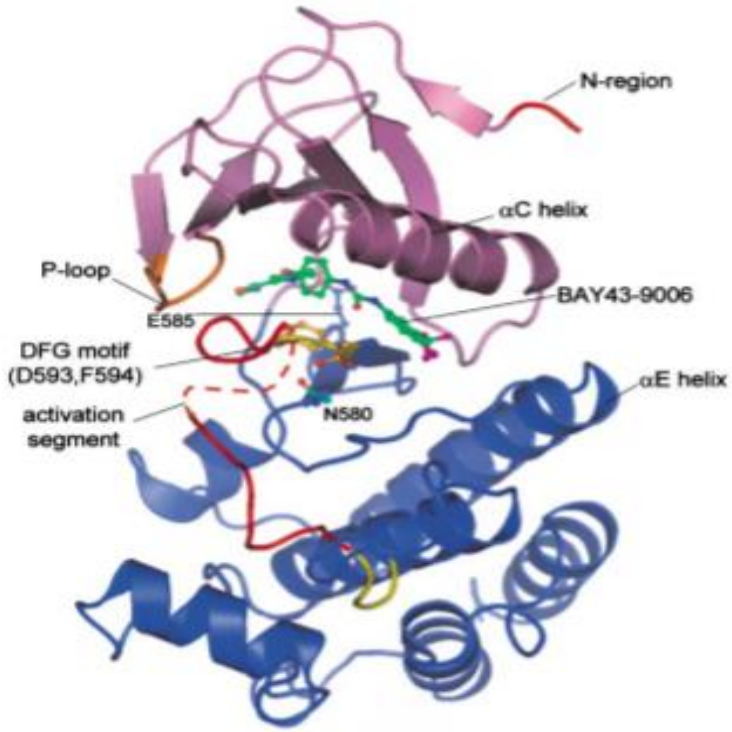

Figure 7: Three-dimensional structure of BAY43-9006 interacting with BRAF-kinase domain. (15)

Challenges faced by molecular docking and SBDD

As mentioned before conformational changes due to binding between receptor and ligand must be considered when SBDD is being performed. Because of this, it is important to take flexibility of the target binding site into consideration in molecular docking. One way to take flexibility into consideration is with the use of ensemble docking. Ensemble docking is when the ligand is being docked onto multiple protein conformations instead of one single conformation. Another way to address the flexibility issue is by using the molecular dynamics technique. Molecular dynamics techniques are able to determine the speed and position of each atom by using Newton's laws of motion. At first, the atoms are configured in such a way that the pressure and temperature of the real system is duplicated. Then the forces that act on every atom are simulated. With this, the velocity and position of every atom can be determined through time. After multiple simulations, the molecular trajectories will be integrated per time interval. Molecular interaction potentials determine the forces that have an effect on the system. These molecular interaction potentials are characterized by experimental data or quantum chemical calculations. With the MD technique, it is also possible to check stability of suggested ligand-receptor complexes by molecular docking. To check the stability, the generated conformations are given an RMSD value. When the RMSD value differs more than the RMSD value of the correlated docking solution, the complex can be viewed as unstable. Most of the time a ligand is able to stabilize multiple configurations of the target receptor, where the equilibrium moves toward the minimum energy configurations. Simulations with MD techniques are able to create these alternative configurations associated with the ligand-induced conformations under study. (11)(12)

One major challenge in SBDD and molecular docking is simulating and predicting covalent bonds. This is because most docking programs or algorithms are made to examine non-covalent

bonds and not much attention has been given to covalent bonds. The interactions of covalent and non-covalent bonds are very different, especially concerning the thermodynamics involved in the bonds. One approach to this issue is the use of mimicking the covalent bond by defining one atom in the ligand and one atom in the protein as “link atoms” to overlay the ligand link atom on the protein link atom. By doing this the geometry of the covalent bond can be evaluated by scoring functions and specific terms. Another approach is the use of large-scale covalent virtual screening. In this approach a covalent attachment point is defined and the conformational space of the ligand around the defined covalent bond is explored. A default scoring ranks each conformation. (11)(12)

Another challenge in SBDD and molecular docking is the presence of water molecules. Water molecules are involved in ligand-receptor recognition in 65% of the ligand-protein complexes acquired with x-ray crystallography. Water molecules can stabilize ligand-protein or small molecule-protein complexes by the formation of water bridges or a hydrogen-bond network. Correctly predicting the free binding energy that is generated by the ligand when displacing water molecules is very difficult. However, to distinguish between displaceable water molecules and structural water molecules, strategies can be used that estimate this free binding energy for a water molecule. Other strategies that can compensate for the presence of water includes the analysis of geometric parameters surrounding each of the water molecules. (11)(12)

Research goal

In conclusion, the features to be attentive of in potential RAF-protein kinase inhibitors are the following; first imidazole or azole-like structures are preferred to bind to the ATP binding sites. This can prevent ATP from binding resulting in inhibition of the RAF protein kinase. In addition, nitrogen atoms in side groups help the binding affinity of the inhibitor by binding other residues in the protein kinase leading to more interactions between the inhibitor and the protein kinase. These are some features to keep in mind when looking for molecular compounds that are able to inhibit RAF-protein kinases. In order to search for these potential RAF-protein kinase inhibitors, an extensive chemical library based on pharmacophores will be used. The chemical database of pharmacophore models is accessible by computer and will be thoroughly screened to search for potential RAF-protein kinase inhibitors. After the virtual screening of this database, the hits provided by the search algorithm will be analysed to determine which of the compounds are most feasible as RAF-protein kinase inhibitor. When analysing the potential RAF-protein kinase inhibitors, it should be taken into account that the flexibility of the target binding site may change. This means that the conformation of the target site may change in such a way that the complex between the target and potential compound will break. Other factors are the presence of water molecules and predicting whether or not covalent bonds will be formed. After analysis of the compounds, a list consisting of the most promising compounds will be made for potential further research.

Time chart

Activity	Week 1 (2/12/19 - 6/12/19)	Week 2 (9/12/19 - 13/12/19)	Week 3 (16/12/19 - 20/12/19)	Week 4 (6/01/20 - 10/01/20)	Week 5 (13/01/20 - 17/01/20)	Week 6 (20/01/20 - 17/01/20)
Tutorials on Chemdraw, Bobel and Vina						
Tutorial on Smiles						
Making a compound library with Chemdraw, Bobel, and Vina and completing the library						
Thesis presentation						
Drawing non-enantiomer compounds						
Drawing enantiomer compounds						
Making 3D-structures of the drawn compounds with Bobel						
Data analysis from the compounds found in the database						
Project presentation						
Report writing						
Finishing the report						

References

1. American Cancer Society. Facts & Figures 2019. American Cancer Society. 2019.
2. Fabbro D, Cowan-Jacob SW, Moebitz H. Ten things you should know about protein kinases: IUPHAR Review 14. British Journal of Pharmacology. 2015.
3. Xing L, Klug-Mcleod J, Rai B, Lunney EA. Kinase hinge binding scaffolds and their hydrogen bond patterns. Bioorganic Med Chem. 2015;
4. Whittaker S, Kirk R, Hayward R, Zambon A, Viros A, Cantarino N, et al. Gatekeeper mutations mediate resistance to BRAF-targeted therapies. Sci Transl Med. 2010;
5. Poulikakos PI, Persaud Y, Janakiraman M, Kong X, Ng C, Moriceau G, et al. RAF inhibitor resistance is mediated by dimerization of aberrantly spliced BRAF(V600E). Nature. 2011;
6. McCubrey JA, Steelman LS, Chappell WH, Abrams SL, Wong EWT, Chang F, et al. Roles of the Raf/MEK/ERK pathway in cell growth, malignant transformation and drug resistance. Biochimica et Biophysica Acta - Molecular Cell Research. 2007.
7. Maurer G, Tarkowski B, Baccarini M. Raf kinases in cancer-roles and therapeutic opportunities. Oncogene. 2011.
8. Wellbrock C, Hurlstone A. BRAF as therapeutic target in melanoma. Biochemical Pharmacology. 2010.
9. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. Vol. 10, Nature Medicine. 2004. p. 789–99.

10. Huxford T. X-Ray Crystallography. In: Brenner's Encyclopedia of Genetics: Second Edition. 2013.

11. Ferreira LG, Dos Santos RN, Oliva G, Andricopulo AD. Molecular docking and structure-based drug design strategies. Molecules. 2015.

12. Guedes IA, Pereira FSS, Dardenne LE. Empirical scoring functions for structure-based virtual screening: Applications, critical aspects, and challenges. Vol. 9, Frontiers in Pharmacology. 2018.

13. Wang Y, Sun Y, Cao R, Liu D, Xie Y, Li L, et al. In Silico Identification of a Novel Hinge-Binding Scaffold for Kinase Inhibitor Discovery. J Med Chem. 2017;60(20):8552–64.

14. Tsai J, Lee JT, Wang W, Zhang J, Cho H, Mamo S, et al. Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. Proc Natl Acad Sci U S A. 2008;

15. Wan PTC, Garnett MJ, Roe SM, Lee S, Niculescu-Duvaz D, Good VM, et al. Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. Cell. 2004;