ACIBADEM MEHMET ALI AYDINLAR UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

COMPUTATIONAL AND EXPERIMENTAL CHARACTERIZATION OF PEPTIDE INHIBITORS TARGETING THE AURORA KINASE A AND N-MYC COMPLEX

LARA NASERIKHOJASTEH
MASTER THESIS

DEPARTMENT OF BIOSTATISTICS AND BIOINFORMATICS

SUPERVISOR
Prof. Emel Timuçin

ISTANBUL-2023

ACIBADEM MEHMET ALI AYDINLAR UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

COMPUTATIONAL AND EXPERIMENTAL CHARACTERIZATION OF PEPTIDE INHIBITORS TARGETING THE AURORA KINASE A AND N-MYC COMPLEX

LARA NASERIKHOJASTEH
MASTER THESIS

DEPARTMENT OF BIOSTATISTICS AND BIOINFORMATICS

SUPERVISOR
Prof. Emel Timuçin

ISTANBUL-2023

DECLARATION

I declare that this thesis work is my own work, I had no unethical behavior at any stages from the planning to the writing of the thesis, I obtained all the information in this thesis in accordance with academic and ethical rules, I cited all the information and comments that were not obtained with this thesis work, and I provided resources in the list of references. I also declare that there was no violation of any patents and copyrights during the study and writing of this thesis.

06.07.2023

Lara Naserikhojasteh

PREFACE AND AKNOWLEDGEMENT

I would like to express my sincere gratitude and appreciation to all those who have contributed to the completion of this master's thesis. I thank the Scientific and Technological Research Council of Turkey (TÜBİTAK) (grant number: 119Z247). First and foremost, I am deeply grateful to my thesis advisor, Prof. Dr. Emel Timuçin, for her guidance and unwavering support throughout the research process. Her invaluable insights and constructive feedback have been instrumental in shaping this thesis. I would also like to extend my thanks to the Prof. Dr. Uğur Sezerman for his passion for teaching and dedication to fostering intellectual growth have been a source of inspiration. I am indebted to my Late Professor Dr. Özgür Gül, who believed in me and prepared me to the grad school while I was an undergrad student. Working in his lab contributed my academic knowledge greatly and I will always be grateful for his guidance. I can only hope I make him proud with this thesis.

I find myself humbled and immensely grateful for the support of my remarkable colleagues, Ali Sinan Kara, İlayda Amanoğlu and Zeynep Kavalcı. This academic journey would not have been the same without their genuine care and encouragement. Beyond our academic pursuits, the bonds we formed as friends have made every step of this journey enjoyable and memorable. Your laughter, encouragement, and understanding during moments of stress have been a constant source of strength.

Lastly, my heartfelt appreciation goes to my mother, Serpil Naserikhojasteh, my father, Navid Naserikhojasteh and my friends Ezgisu Yeniocak and Egemen Ton for their support, understanding, and encouragement throughout this journey. Their belief in my abilities has been a constant source of motivation, and I am grateful for their love and encouragement. Completing this master's thesis has been a challenging yet rewarding experience, and I am grateful for the collective support I have received along the way. Thank you all for your contributions, both big and small, in helping me accomplish this milestone in my academic journey.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND SYMBOLS

A	Alanine, ALA
˚A	Angstrom
AMBER	Assisted Model Building with Energy Refinement
AurA	Aurora Kinase A
BSA	Bovine Serum Albumin
C	Cysteine, Cys
CHARMM	Chemistry at HARvard Molecular Mechanics
Cl-	Chloride Ion
Cα	Carbon Alpha Atom
D	Aspartic acid, Asp
DBD	DNA Binding Domain
E	GLU, Glutamic acid
EDTA	Ethylenediaminetetraacetic Acid
F	Phenylalanine, Phe
G	Glycine, Gly
GFP	Green Fluorescent protein
GST	Glutathione S-transferases
H	Histidine, HIS
H	Hydrogen
I	Isoleucine, Ile
IMAC	Immobilized Metal Affinity Chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
K	Kelvin
K	Lysine, Lys
L	Leucine, Leu
LB	Luria-Bertani
M	Methionine, Met
MD	Molecular Dynamics
mM	Milimolar
Myc	Myelocytomatosis Oncogene

MYCN	Neuroblastoma-Myelocytomatosis viral gene
N	Asparagine, Asn
Na+	Sodium Ion
NaCl	Sodium Choloride
NAMD	Nanoscale Molecular Dynamics
Nm	Nanometer
NPT	Constant Number of Particles, Pressure and Temperature
ns	Nanosecond
NVT	Constant number of particles, volume and temperature
OD	Optical Density
P	proline, Pro
PBS	Phosphate Buffered Saline
PDB	Protein Data Bank
ps	Picosecond
Q	Glutamine, Gln
R	Arginine, ARG
RGYR	Radius of Gyration
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
S	Serine, Ser
SDS	Sodium Dodecly Sulfate
STK6	Serine/Thyrosine Kinase 6
T	Threonine, Thr
TAD	Transactivation Domain
TBS	Tris Buffered Saline
TRUBA	Turkish National Science e-Infrastructure
TTK	Threonine thyrosine kinase
V	Valine, Val
V/V	Volume/Volume
VMD	Visual Molecular Dynamics
W	Tryptophan, Trp
W/V	Weight/Volume

Y	Tyrosine, Tyr

LIST OF FIGURES

LIST OF TABLES

ÖZET

Aurora Kinaz A Ve N-Myc Kompleksi'ne Yönelik Peptit İnhibitörlerinin Hesaplamali Ve Deneysel Karakterizasyonu

Myc proteinleri, transkripsiyonu hem aktive ederek hem de baskılayarak gen ekspresyonunu etkileyen transkripsiyon faktörleridir. İnsan malignitelerinde, Myc protein ailesinin üç üyesi anormal ifadeye sahip olabilir. Hücresel Myc (c-Myc), viral Myc (v-Myc) onkoproteinin hücresel homologudur ve çeşitli insan malignitelerinde sıklıkla düzensizdir. N-Myc ve L-Myc'nin nöroblastomda güçlendirilmiş MYCN ve küçük hücreli akciğer kanserinde MYCL genlerinin ürünleri olduğu keşfedilmiştir. Bu kanser hücrelerinde N-Myc birikimi için, N-Myc'nin tamamen yok edilmediği protein kinaz Aurora-A ile bir kompleksin oluşturulması gereklidir. Aurora kinazlar, mitozun başlaması ve gelişmesi için gerekli olan serin/treonin kinazlardır. Bu çalışmada, Aurora A/N-Myc karışan peptit ile ilgili önceki çalışmalardan elde edilen bulgular ayrıca GST aşağı çekme testi ile analiz edildi, ifade edildi ve karakterize edildi. Bir önceki çalışmanın yarı rasyonel tasarımı, in-silico mutajenez yoluyla rasyonel tasarıma dönüştürüldü ve MD simülasyonları yapıldı. Ayrıca RGYR (Radius of gyration), RMSF (Root mean square fluctuation) ve Pairwise RMSD (Root mean square deviation) analizi yapılmıştır. Bu tez, protein ekspresyonundaki zorlukları vurguladı ve başarısızlığın nedenlerini değerlendirdi, GST aşağı çekme testinin özgüllüğünü ve duyarlılığını göstererek, gerçek etkileşimleri spesifik olmayan bağlanmadan ayırt etmeye izin vererek, AurA/N-Myc bağlanma yüzey etkileşimi hakkında değerli bilgiler sağladı. Ayrıca, sistematik hesaplamalı simülasyonlar yoluyla, rasyonel tasarlanmış peptitlerin etkileşimlerine ilişkin değerli iç görüler elde edildi; bu, vahşi tipe kıyasla P2 ve AurA/N-Myc bağlanma yüzeyi arasında gelişmiş kararlılık ve daha güçlü bağlanma afinitesine işaret etmektedir.

Anahtar Sözcükler: Sentetik peptidler, Moleküler dinamik simülasyonu, Serbest enerji hesaplamaları, Protein mühendisliği, AurA/N-Myc bağlanma yüzeyi

ABSTRACT

Computational And Experimental Characterization of Peptide Inhibitors Targeting the Aurora Kinase A And N-Myc Complex

Myc proteins are transcription factors affecting gene expression by both activating and repressing transcription. In human malignancies, three members of the Myc protein family can have abnormal expression. Cellular Myc (c-Myc) is the cellular homolog of the viral Myc (v-Myc) oncoprotein, and it is frequently unregulated in a variety of human malignancies. N-Myc and L-Myc were discovered to be the products of the amplified genes MYCN in neuroblastoma and MYCL in small cell lung cancer, accordingly. Generation of a complex with the protein kinase Aurora-A in which the N-Myc is not fully destroyed is required for N-Myc accumulation in cancer cells. Aurora kinases are serine/threonine kinases that are required for mitosis to begin and develop. In this study, findings from previous studies of Aurora A/N-Myc interfering peptide are further analyzed, expressed, and characterized by GST pull-down assay. Semi-rational design of the previous study, converted into a rational design via *in-silico* mutagenesis and MD simulations were performed. In addition, RGYR (Radius of gyration), RMSF (Root mean square fluctuation) and Pairwise RMSD (Root mean square deviation) analysis were performed. This thesis highlighted challenges in protein expression and evaluated the causes of failure, showed specificity and sensitivity of the GST pull-down assay allowing to discern genuine interactions from nonspecific binding, providing valuable insights into AurA/N-Myc binding surface interaction. Furthermore, through systematic computational simulations, valuable insights into the interactions of the rational designed peptides were gained, pointing enhanced stability and stronger binding affinity between the P2, and the AurA/N-Myc binding surface compared to wild-type.

Keywords: Synthetic peptides, Molecular dynamics simulation, Free energy calculations, Protein engineering, AurA/N-Myc binding surf.

1 INTRODUCTION AND AIM

1.1 What Are Aurora Kinases?

Aurora kinase is a type of enzyme that plays a crucial role in the process of cell division, specifically in the separation of chromosomes during mitosis. There are three main types of Aurora kinase: Aurora A, Aurora B, and Aurora C (1). Aurora kinase A, also known as Aurora-A or STK6, is a serine/threonine protein kinase that plays a critical role in regulating mitosis and cell division since all eukaryotes require the activity of the family of serine/threonine kinases for the completion of mitosis (2).

1.2 Types of Aurora Kinase

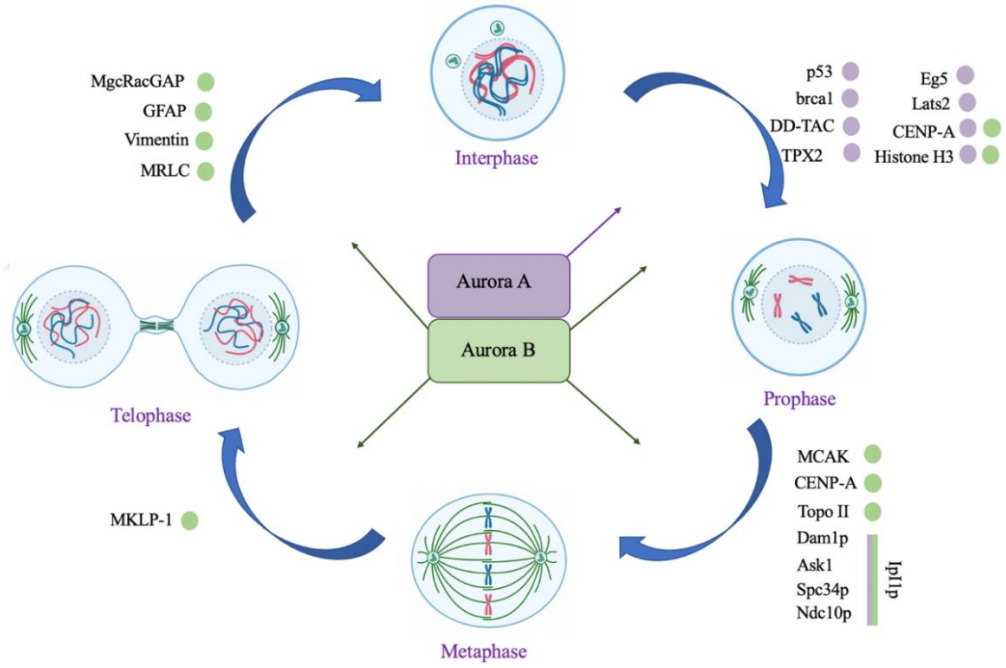

Figure 1: All phases of mitosis in which Aurora A and B are involved.

Aurora A is primarily localized to the centrosomes and is involved in the formation and maturation of the spindle poles, which are essential for the proper alignment and separation of chromosomes during cell division. In addition to its role in mitosis, Aurora-A also regulates other cellular processes, including cell proliferation, apoptosis, and DNA damage response (2). Aurora-A is oftentimes overexpressed in cancer cells, where it promotes unchecked cell proliferation and

tumor development. Therefore, Aurora-A has become a desirable target for cancer therapy. To prevent Aurora-A from acting abnormally and to stop the pathways that lead to abnormal cell turnover in cancer cells, researchers are looking for small chemical inhibitors that specifically target the protein (3). Dysregulation of Aurora-A has been implicated in the development of several types of cancer, including breast, lung, colon, and pancreatic cancer (4).

Aurora B plays a critical role in the regulation of chromosomal segregation and cytokinesis during mitosis. It is also involved in the regulation of spindle checkpoint signaling and maintaining genomic stability. Aurora B is essential for cell division and its inhibition can induce mitotic arrest and cell death (5). Aurora C is primarily expressed in germ cells and is involved in the regulation of meiosis. It is also expressed in some cancer cells and may play a role in the regulation of cell division in these cells. The exact function of Aurora-C is still under investigation (6). Researchers have focused on Aurora A and B as they have been shown to be important for the proper alignment and separation of chromosomes during cell division (7).

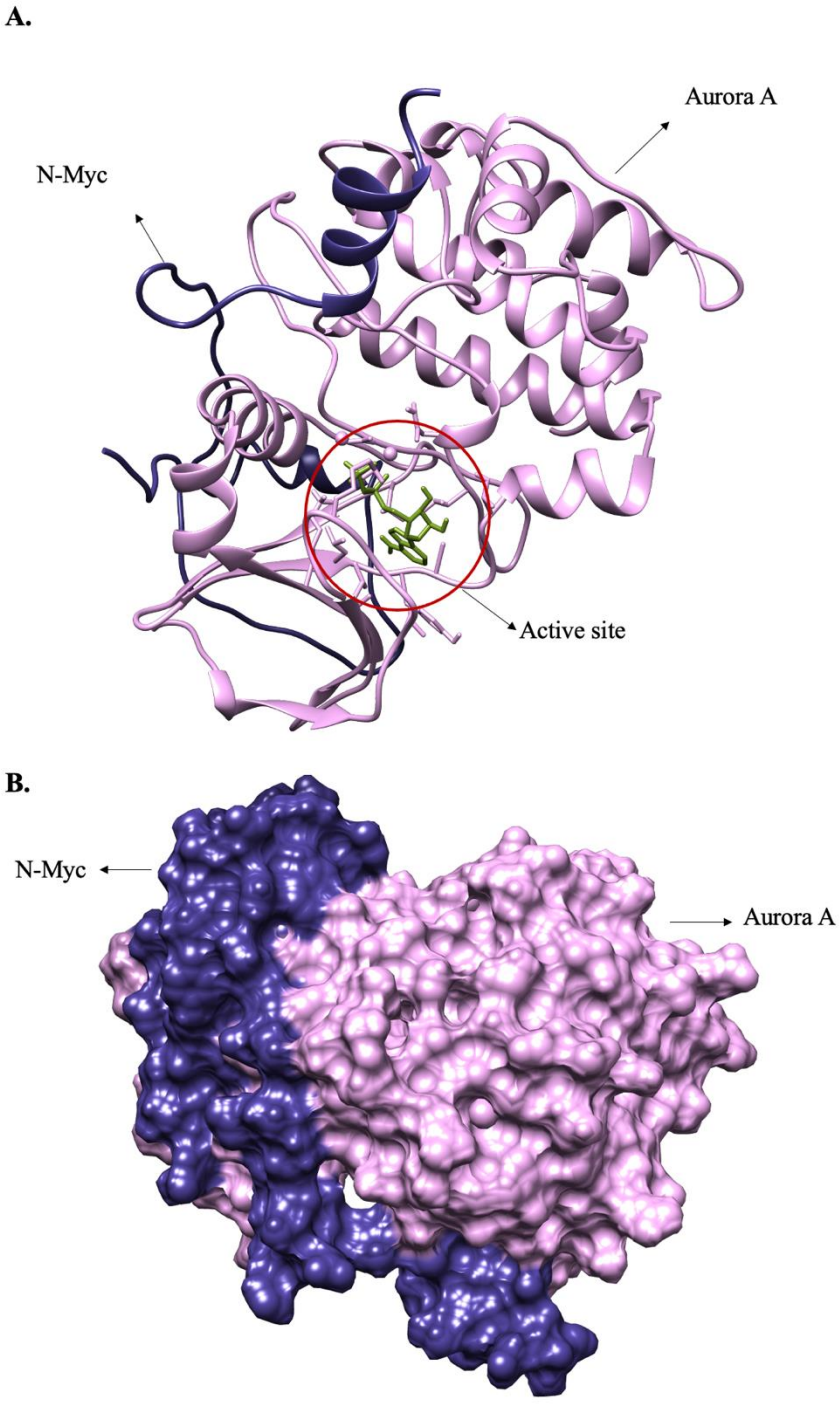

Figure 2: Chimera snapshots of the Aurora Kinase A (PDB ID: 5G1X, Resolution: 1.72 Å). AurA: New cartoon, plum. N-Myc: New cartoon, slate blue. (A) Active site of the protein is pointed out with red. (B) Surface is shown for both AurA and N-Myc.

Abnormalities in Aurora kinase activity have been linked to various types of cancer and other diseases, making it an important target for drug development (8). Abnormalities in Aurora kinase regulation can have significant impacts on cell

division, leading to a variety of disease states. For example, overexpression of Aurora kinase A has been linked to the development of many types of cancer, including breast, lung, and pancreatic cancer. Overexpression of Aurora kinase B has also been linked to cancer, as well as other diseases such as Alzheimer's and Parkinson's (9). In addition, mutations in genes encoding Aurora kinases or their associated regulatory proteins can cause various developmental disorders as well as cancers. For example, mutations in the TTK gene, which encodes a protein involved in regulating Aurora kinase activity, have been associated with a rare form of congenital heart disease called atrial septal defect. Mutations in the Aurora B gene have been associated with a form of autoimmune disease called systemic lupus erythematosus (10). Therefore, there is a reason for researchers to focus on this area, and inhibitors of Aurora kinase activity have shown promise as potential treatments for various types of cancer and are currently being tested in clinical trials. In addition, more research is needed to fully understand the complexity of Aurora kinase regulation and its role in various diseases, and to increase studies. Aurora kinase inhibitors are a class of drugs that work by blocking the activity of Aurora kinase and are being investigated as potential treatments for cancer. These inhibitors have shown promising results in preclinical and clinical trials, and several Aurora kinase inhibitors are currently approved by regulatory agencies for cancer treatment (11).

1.3 The Myc Family and N-Myc

The N-Myc protein is a member of the Myc family of transcription factors that regulate cell growth and differentiation. It is known to play a critical role in the development and progression of several types of cancer, including neuroblastoma, a pediatric cancer of the nervous system (12). The N-Myc protein forms a complex with other proteins, to regulate gene expression. The N-Myc/Max complex binds to specific DNA sequences known as E-boxes to activate the expression of genes involved in cell proliferation.The activity of the N-Myc complex is tightly regulated in normal cells, but in cancer cells, it can become dysregulated, leading to uncontrolled cell proliferation and tumor growth (13). The overexpression of N-Myc is a common feature of neuroblastoma and other types of cancer, and is associated with a poor

prognosis. Understanding the mechanisms by which the N-Myc complex regulates gene expression is an area of active research, and targeting the complex is a potential strategy for the development of new cancer therapies (14).

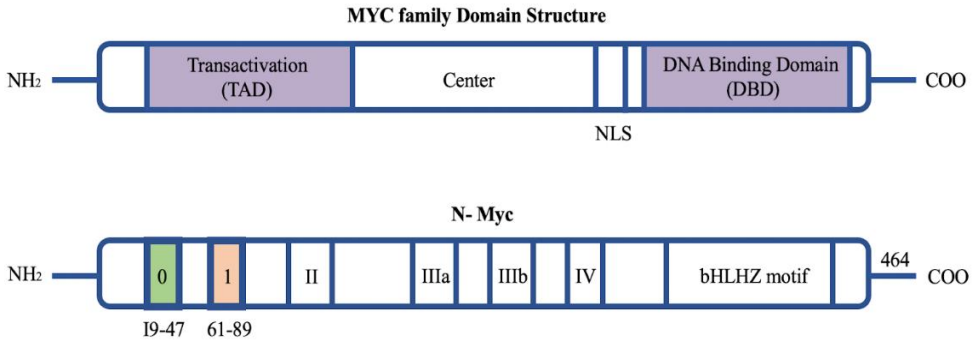

Figure 3: Represantation of MYC family domain structure.

MYCN (Neuroblastoma-Myelocytomatosis viral gene) located on the short arm (2p24.3) of chromosome 2 (encodes the N-Myc protein) is found in a gene family consisting of 3 members. This gene was first identified as an oncogene in 1983 when it was found in genomic amplified form in a group of human neuroblastoma cell lines (15). In addition to being seen in different human cancers (16), MYCN gene amplification is found in 20% of neuroblastoma cases (17–19). Unlike other family members, cellular MYC (encodes C-Myc protein) and MYCL (encodes L-Myc protein), MYCN is a gene that is expressed by nervous system cells during embryogenesis and plays a role in neural crest development (20,21). MYCN expression decreases markedly after embryogenesis and is lost in adult tissue (22). Under normal conditions, expression of MYCN gene is tightly controlled like other members of the family (23); In conditions such as genomic amplification, chromosomal translocation or mutagenesis in signaling pathways, high N-Myc protein levels result in the formation of tumors that are difficult to treat (24). In this background, the N-Myc protein is an essential target for therapeutic approaches to tumors in which it is overexpressed (25).

Although it has been known as an important cancer treatment target since it was first described, compounds that directly target N-Myc are very limited. Like other Myc

family members, N-Myc has been described as untargetable (25). The recognition of this family as non-targetable is primarily related to their being transcription factors. Even though transcription factors constitute 20% of all oncogenes, they are considered non-targetable (26,27). The main observations that led to this paradigm relate to the structural features and cellular location of transcription factors. Except for those inducible with ligands, transcription factors lack the enzymatic function to facilitate pharmacological targeting strategies and thus pocket regions that can be directly targeted. In addition, they are unlikely to be targeted by antibody-based therapeutic approaches, which are difficult to access intracellularly. Therefore, approaches that directly target transcription factors, including N-Myc, are inconclusive (28).

1.4 Aurora Kinase A and N-Myc

There is evidence to suggest that the Aurora kinase A and N-Myc complex are functionally linked in the regulation of cell proliferation and tumor growth. Aurora kinase A has been shown to be a downstream target of N-Myc, and its expression is upregulated in N-Myc-overexpressing cells (29). The dysregulation of both Aurora kinase A and N-Myc is commonly observed in neuroblastoma and other types of cancer, and targeting the complex is being investigated as a potential therapeutic strategy. Small molecule inhibitors of Aurora kinase A, as well as other components of the N-Myc complex, are being developed and tested in preclinical and clinical studies for the treatment of cancer. Inhibition of Aurora A has been shown to reduce the stability and activity of N-myc, resulting in decreased cell proliferation and tumor growth (30).

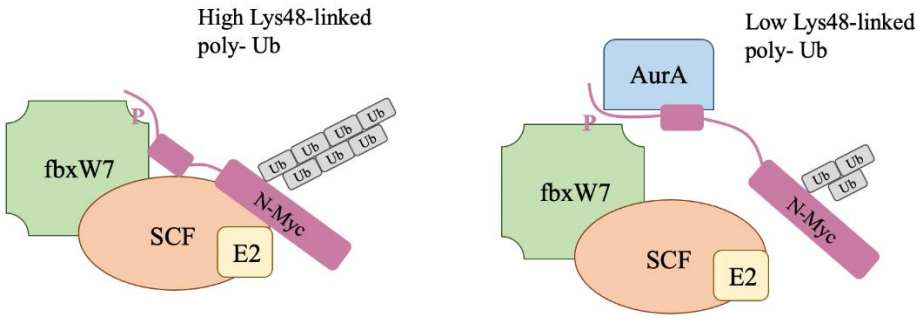

Figure 4: A representation to explain how the interaction between Aurora-A and N-Myc residues 61–89 may reduce ubiquitination to stabilize N-Myc (31).

Although its direct targeting is challenging, it is possible to inhibit N-Myc indirectly (21,32). Indirect inhibition strategies include stopping the protein-protein interactions established by N-Myc (33). Protein-protein interactions established by the Myc family occur through TAD (34,35). In this domain, there are 3 conserved regions (MB0-I-II) within the Myc family, called "Myc Box (MB)". Although this region is irregular in the free state, it can take on a temporary secondary structure during protein-protein interactions (36,37). Protein-protein interactions over these regions regulate the stability of Myc proteins, including N-Myc (38). For example, SCFFbxW7, the E3 ubiquitin ligase, modifies N-Myc at the K48 position with ubiquitin, triggering its proteosomal degradation (39).Although N-Myc levels are normally controlled in this way bySCFFbxW7 , this process is blocked by the binding of a serine/theronin kinase (AURKA gene product) to N-Myc, particularly in neuroblastoma cells, and N-Myc protein stability increases (33). According to the results of the recent research, it was seen that AurA and SCFFbxW7 N-Myc competed to bind to the MBI (61-89 (amino acid positions)) (Figure 4) position (31). Thus, inhibiting the AurA surface of the AurA/N-Myc complex would be a viable strategy to pave the way for blocked proteasomal degradation. No compound has yet been found that directly inhibits AurA/N-Myc interactions because the complex structure formed by AurA and an N-Myc region belonging to the 26 amino acid long MBI motif has been solved recently (31). Thanks to this 3-dimensional structure, a surface for N-Myc interaction was found on the AurA surface.

1.5 Aim of The Thesis

The aim of this thesis is to design compounds that will bind to the AurA surface of the AurA/N-Myc complex and validate trough computational and experiment methods. The 3D complex structure of AurA/N-Myc has not been the subject of any drug design study before. In these respects, aiming to find compounds that will target the AurA/N-Myc binding surface, and the compounds intended to be revealed is unique to this thesis and have the potential to have great impact, both scientifically and technologically.

2 BACKGROUND

The myc family is a family of proteins of the "basic helix-loop-helix-zipper" (bHLHZ) class, 364-to-464 aa long, composed of 3 different transcription factors. This family carries the transactivation domain (TAD) at the amino end and the domain responsible for DNA binding and dimerization (DBD) at the carboxy end. Of these domains, only DBD (83 aa long) has a stable 3D structure; the remaining region is intrinsically disordered, as reported by circular dichroism and NMR spectroscopy (36,37). Therefore, in addition to being transcription factors, the observed irregularity in Myc family structures makes it impossible to target this family directly. In addition, although it has a distinct 3-dimensional structure, it is not possible to specifically target DBD because it contains a highly conserved bHLHZ motif in transcription factors (40).

AurA is a member of the Aurora kinase family that plays a role in the regulation of the G2/M checkpoint of the cell cycle. Like N-Myc, AurA can be found amplified in human tumors (41). Thus, there are currently many AurA small molecule inhibitors for inhibiting kinase activity. It has been shown that the catalytic activity of AurA does not affect the degradation or stability of N-Myc (31). In parallel, the currently known AurA kinase inhibitors do not have the potential to directly affect the formation of the AurA/N-Myc complex (39,42). Because when the AurA/N-Myc complex structure is examined, it is seen that the AurA active site and the binding surfaces of N-Myc are in separate regions from each other. Despite this difference, it has been reported that a class of AurA kinase inhibitors (for example, MLN8237/alisertib and CD532) can destabilize the complex formed with N-Myc because it allosterically alters the AurA structure (30). In addition to this observation, in vitro studies have shown that these inhibitors can induce N-Myc degradation (29), and even a combination of drugs containing alisertib regressed in a group of neuroblastoma cases (43). These positive results reiterated that targeting the AurA/N-Myc complex will be an essential strategy for the treatment of tumors with increased N-Myc expression (44). However, existing AurA kinase inhibitors cannot be used as inhibitors of the AurA/N-Myc complex as they do not directly target the AurA/N-Myc interaction surface.

3 MATERIALS AND METHODS

3.1 Semi-Rational Design

As seen in the workflow (Figure 5), one peptide is selected that is used in this thesis, was established by the members of the Timucin Lab, In the previous studies. The selected sequence is: (LEFDSLQMCFYDDEDDWYFGGRDSTPPGEDIWDKFELLPTPPMSPFRGFWEWSSEPPSWVTEMLLENELWG) with Δ G -70.587, SD 12.9972, ΔΔG-46.9835 scores as it is seen in the Table1.

Table 1: Selected peptide sequence with semi-rational design

ID	Sequence		Δ G	SD	Δ Δ G
5	**LEFDSLQMCFYDDEDDWYFGGRDS TPPGEDIWDKFELLPTPPMSPFRGF WEWSSEPPSWVTEMLLENELWG**		**-70.587**	**-12.997**	**-46.984**

```
CLUSTAL O(1.2.4) multiple sequence alignment


semi_rational_peptide      LEFDSLQMCFYDDEDDWYFGGRDSTPPGEDIWDKFELLPTPPMSPFRGFWEWSSEPPSWV 60
wild-type                  ----SLQPCFYPDEDDFYFGGPDSTPPGEDIWKKFELLPTPPLSPSRGFAEXSSEPPSWV 56
                               *** *** ****:**** **********.*********:** *** * ********

semi_rational_peptide      TEMLLENELWG  71
wild-type                  TEMLLENEL--  65
                           *********
```

Figure 5: Multiple sequence alignment of semi-rational designed peptide and the wild-type sequence.

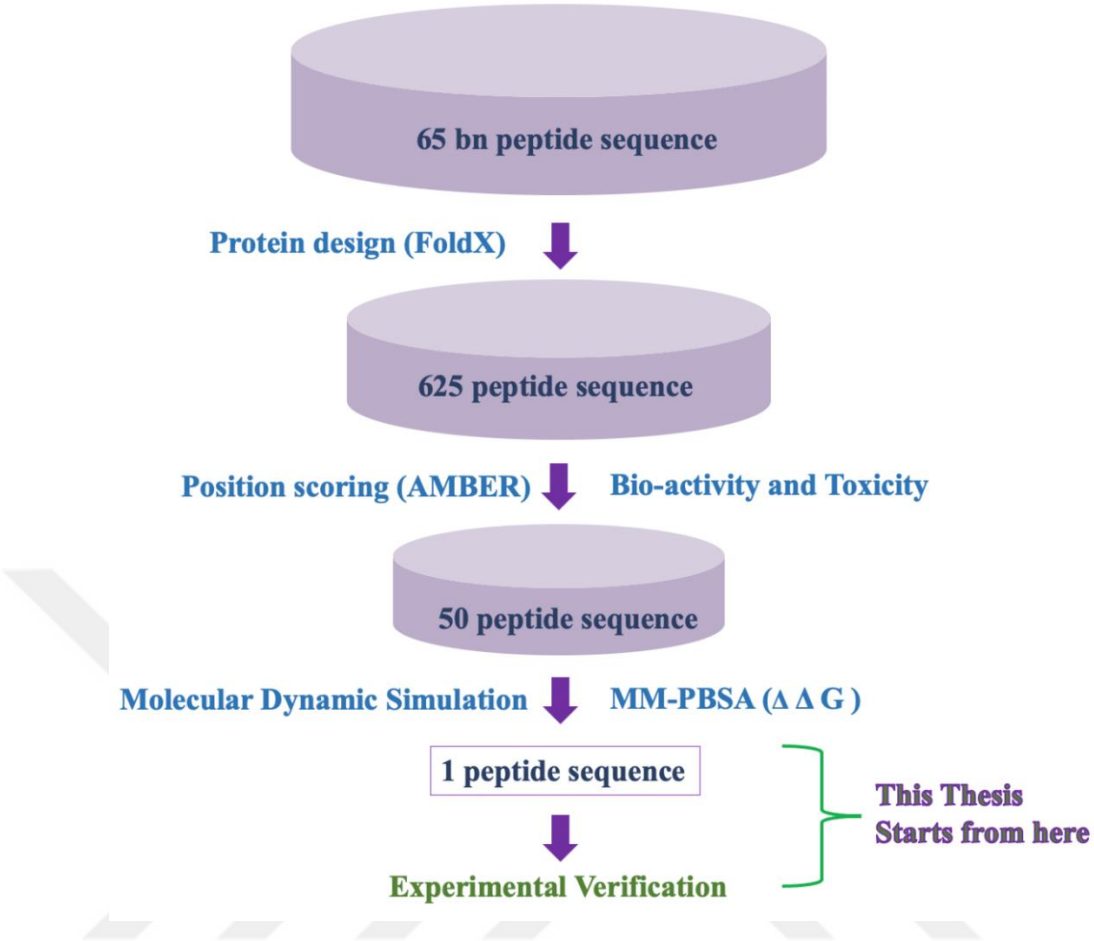

Figure 6: Workflow of the semi-rational peptide design from previous studies.

3.2 Experimental Methods

3.2.1 Plasmid isolation

Recombinant plasmids (i) AurA-GFP and (ii) AurA-GPF-mCherry pET21 obtained from Addgene as stab-culture of DH5α bacterial strain were first spread on LB-agar (100 µg/ml ampicillin) medium and continued for 16-18 hours. It was incubated at 37°C. Single colonies selected from this medium were transferred to liquid medium and incubated in a shaker at 37°C for approximately 16 hours. Plasmid isolation from the obtained bacterial cultures was performed with the help of Qiagen Miniprep kit. The isolated plasmids were analyzed spectroscopically with Nanodrop measurements in terms of concentration and purity, and necessary glycerol stocks were made. Concentration values were determined as 50-100 ng/µl.

3.2.2 Transformation

Two recombinant plasmids pAuroraA-GFP-AURKA-mCherry (Addgene Plasmid #99878) and pAuroraA-GFP-AURKA (Addgene Plasmid #99877) carrying the gene encoding the human Aurora A Kinase enzyme (AURKA), both XL-1B and BL21(DE3)) were transferred to E. coli strains. For this, cells that were previously made competent and stored at -80°C were used. Approximately 1 µl (approximately 50-100 ng of plasmid DNA) from the plasmids isolated by following the calcium chloride assisted chemical transformation method was mixed with 200 µl of competent cell solution and the mixture was incubated on ice for 5 minutes. Then, LB broth media was added to this mixture and incubated at 370C with shaking for approximately 45-60 minutes. The cells were then precipitated and some of the supernatant was discarded to increase the cell density and spread on LB-agar (100 µg/ml ampicillin) medium (45). This medium was incubated at 370C for 16-18 hours. Reproducing single colonies were selected. Glycerol stocks were taken from these colonies and stored at -80°C, and plasmid stocks were stored at -20°C. Glycerol stock was taken for plasmid propagation from selected colonies from strain XL-1 blue. The BL21 strain was also stockpiled for use in protein expression.

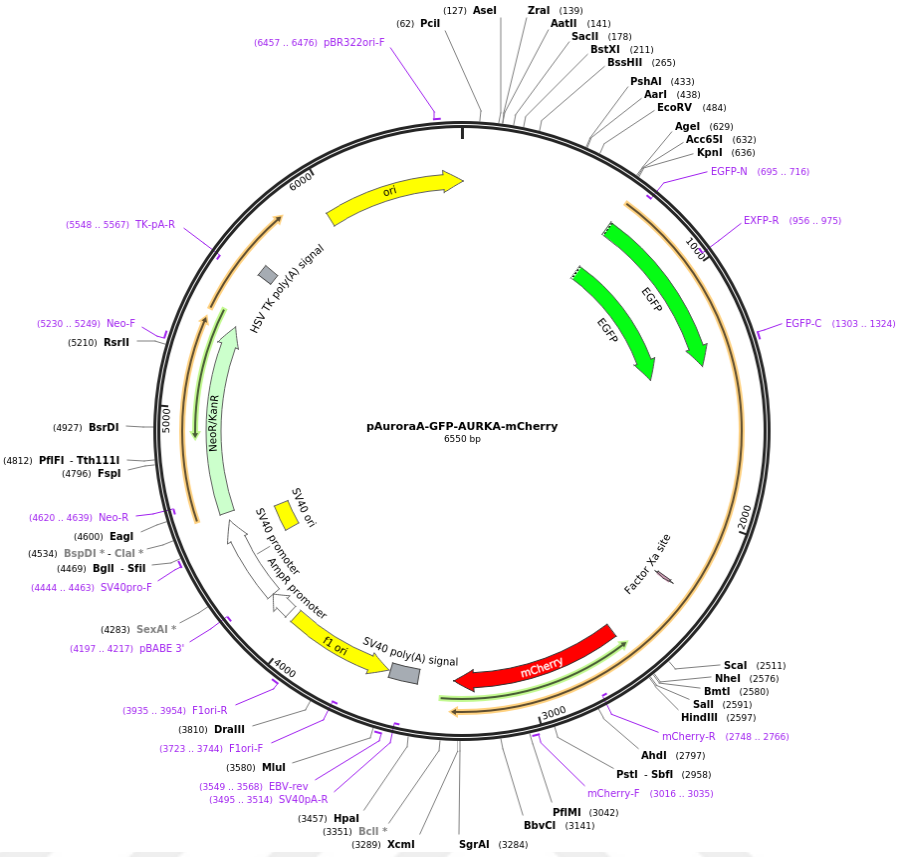

Figure 7: Plasmid map of pAuroraA-GFP-AURKA-mCherry (Addgene Plasmid #99878).

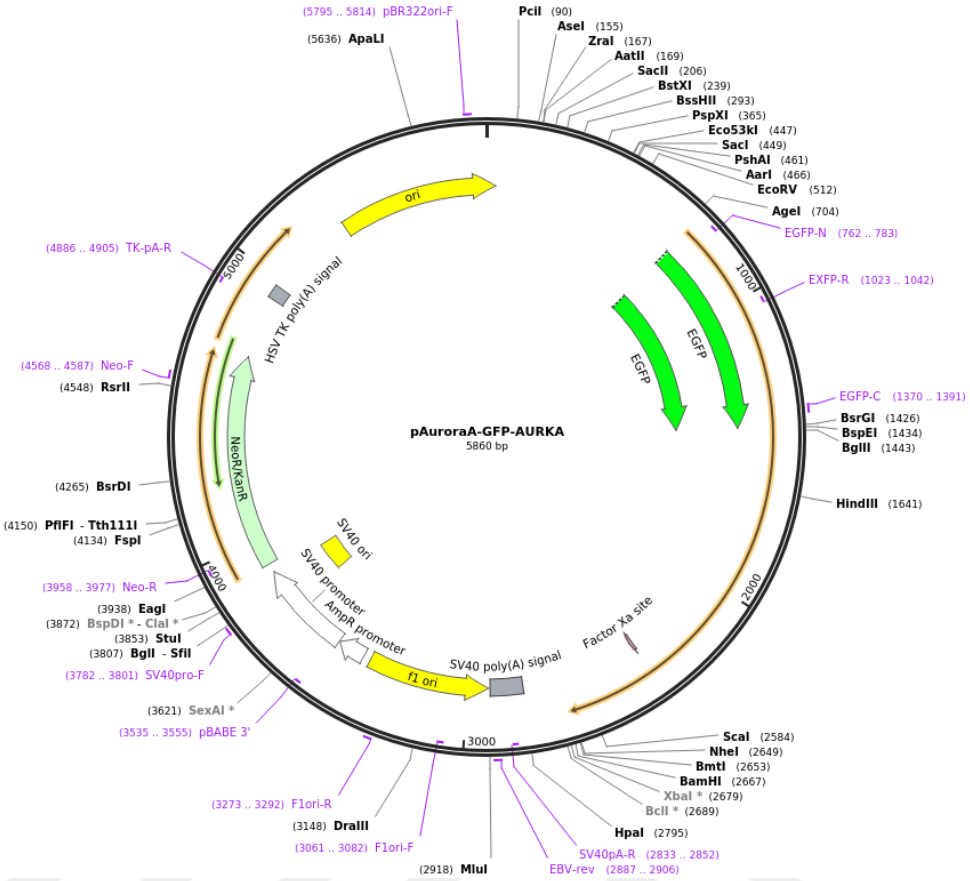

Figure 8: Plasmid map of pAuroraA-GFP-AURKA (Addgene Plasmid #99877).

3.2.3 Recombinant protein expression

Two different induction methods were tried for recombinant AurA expression. One of them is auto-induction and the other is IPTG induction. Since the IPTG method is more established and high efficiency, it was tried at the first stage. However, Since the autoinduction method were available and more cost effective it is decided to continue with auto-induction (46). The method followed in these experiments, respectively, is given below.

3.2.4 IPTG induction

Nearly all the colonies obtained in the previous transformation step were scanned via pipet tip and seeded into a cell culture tube containing 5 ml liquid LB medium and 100 μg/ml ampicillin. This inoculum was grown at 37°C with 200 rpm shaking for 16 hours. Then, 1 ml of this culture was transferred to 50 ml of liquid LB-ampicillin medium and its optical density (OD) at 600 nm was followed by spectrophotometer (Varioskan Flash, Thermo Fisher) and incubated at 37°C with 200 rpm shaking. In this step, more than 50 ml of cell solution has been taken in experiments. To increase the amount of protein produced, the volume of induction medium was increased up to 200 ml. The culture was induced by IPTG, with an OD measuring between 0.6--0.8. In protein expression experiments, 1 ml sampling was done before and after induction, and the sampled cell pellets were stored at -20°C for a short time and at -80°C for a long time. IPTG induction was sustained for a maximum of 6 hours.

3.2.5 Auto-induction

The medium known as auto-induction medium was prepared and sterilized to contain 10 g/L peptone, 5 g/L yeast extract. The solutions named 50xM and 50x5052 and 500x $MgSO_4$, which were sterilized in the same way and the contents of which are given in Table 2, were added to the first medium as 1x and an auto-induction medium was obtained. buffered. A few colonies obtained in the previous transformation step were seeded into a cell culture tube containing 5 ml liquid LB medium and 100 μg/ml ampicillin. This inoculum was grown at 37°C with 200 rpm shaking for 3 hours. This

entire culture was then added to a 50 ml auto-induction broth culture and incubated overnight (16 hours) at 37°C with 200 rpm shaking. This obtained pellet was centrifuged at 5000 rpm for 20 minutes at 4°C and stored at -80°C if the protein purification step was not immediately proceeded.

Table 2: Ingredient list of The Auto-induction media.

ZYM-5052 Medium Stock Solution	**Ingredients**	**Amount**
ZY	Tryptone	10g
	Yeast Extract	5g
	DDW	958ml
50x M	$Na_2HPO_4 – 7H_2O$	335g
	KH_2PO_4	170g
	NH_4Cl	134g
	Na_2SO_4	35.5g
	DDW	Up to 700ml
50x 5052	Glycerol	250g
	Glucose	25g
	Lactose	100g
	DDW	730ml
500x $MgSO_4$	$MgSO_4 – 7H_2O$	24.65g
	DDW	87ml

3.2.6 Protein extraction

Both sonication and detergent-based methods were attempted for the lysis of cell pellets obtained after induction (and maintained at -80 °C). For sonication-based lysis, the obtained pellet was resuspended in a solution containing Tris-Cl (50 mM, pH 8.0), NaCl (500mM), glycerol (10% v/v) used as sonication buffer in IMAC binding buffer. 50 ml of obtained E. coli pellets were added to 1 ml of buffer. Cell suspensions were kept on ice during sonication. The frequency for lysing the cells was adjusted not to exceed 30 kHz. Sonication was continued for a total of 5 minutes in the form of a cycle of 30 seconds of sonication and then 30 seconds of waiting. Detergent-based lysis was performed as described with the aid of BPER (Bacterial Protein Extraction Reagent, Thermo Fisher). 50 ml of E. coli pellets were added with 1 ml of B-PER and incubated for 30 minutes at room temperature in an orbital mixer. As a result of both lysis, the extract was centrifuged at high speed at 13,000 rpm for 10 minutes, and the supernatant of the extract was obtained as a soluble fraction and purification steps were progressed. Then, the solution was separated into total (T), soluble (S) and insoluble (I) fractions. S fraction was obtained as supernatant after centrifugation. Fraction I was obtained as pellet which precipitated after this process. From these fractions, samples in the same volume were separated according to their molecular size by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) method. The gel concentration used for SDS-PAGE is 12%, and the gels run at 100 Volts for 1 hour were stained with coomassie staining and protein expression determinations were made. In both lysis steps, the serine protease inhibitor PMSF (phenylmethylsulfonyl fluoride) was added to the lysis buffer solution at a final concentration of 1mM. The soluble fraction was passed into IMAC binding buffer and used for purification. These solutions were stored at -20°C.

3.2.7 Protein purification

Recombinantly expressed AurA protein(s) were purified by immobilized metal (nickel) affinity chromatography (IMAC) with the aid of poly-histidine tags for purification. For this, purification was carried out with the help of both Ni-coated beads (Ni-NTA Resin, Thermo Fisher) and His-trap (Thermo Fisher) columns, known

as the batch method. In this step, the recommended buffer contents were used for purification by the Ni-based affinity method. The contents of the binding, washing and elution buffers are as follows: binding solution (20mM sodium phosphate, 300mM NaCl –PBS, 10mM imidazole pH 7.4), wash solution (25mM imidazole in PBS; pH 7.4) and elution solution (300mM in PBS) imidazole pH 7.4). If necessary, both the resin and columns were cleaned by removing nickel. In this procedure, the column was washed with 5-to-10 column volumes with nickel dissolving buffer (50mM EDTA, 500mM NaCl). Then, 5 column volumes were washed with distilled water, 3 column volumes were loaded with 0.1 M NiSO4 nickel, and then washed again in the same way and stored in 20% ethanol at 4°C. The protocol followed for protein purification was applied in the same way for purification from resin or column. First, 1 volume of resin/column was washed in 2 volumes of binding solution. The extract and resin were then incubated in an orbital mixer for 1 hour at room temperature. After incubation, the resin and extract were separated by centrifugation at 500 rpm for 2 minutes and the supernatant was sampled. The resin was washed 3 times with 3 volumes of washing solution and sampled in the supernatants that are obtained. Then, at least 2 times, the proteins bound to the resin were separated with 1 volume of elution solution. The extract was loaded to the columns at a rate not exceeding 0.5 ml/min and similarly, it was washed with 3 volumes of washing solution and the proteins showing nickel affinity were purified with the elution solution. Color change was monitored by the Bradford test during both washing and elution. As a result, 5 fractions showing the most significant color change from 6 consecutive elution fractions of both counters were analyzed by SDS-PAGE to evaluate their purity. For SDS-PAGE, 12% gel was run at 100 Volt for 1 hour at constant potential (Green et al., 2012) and then their purity was analyzed by staining with coomassie staining.

3.3 Characterization

3.3.1 Tris-tricine gel analysis

The gels in which the peptides will be analyzed were prepared as a 16% tris-tricine gel (47). The separation gel contains purified water, 16% acrylamide, Tris-Cl buffer, glycerol. Solution containing 0.1M Tris, 0.1M Tricin and 0.1% SDS was added as

cathode buffer while running the tricine gel. 0.2M Tris-Cl solution adjusted at pH8.9 was used as the anode buffer (47). Unlike SDS-PAGE electrophoresis, tris-tricine gels were run on ice and at low voltage (not exceeding 70 V), and a low range size marker was also used as a marker. Gel was stained via Coomassie staining method. The stain and de-stain that is used in this method prepared according to the Table 3 and Table 4.

Table 3: SDS-PAGE gel stain solution ingredients.

SDS-PAGE Stain	**Ingredient Amount**
Coomassie Blue BB R-250	**2g**
Methanol	**500ml**
↓ Filter Though 3MM Filter Paper ↓	
Acetic Acid	**100ml**
DDW	**400ml**

Table 4: SDS-PAGE gel de-stain solution ingredients.

SDS-PAGE De-stain	Ingredient Amount
Methanol	**400ml**
Acetic Acid	**140ml**
DDW	**1460ml**

3.3.2 Protein concentration, buffer exchange and storage

The protein samples were transferred to centrifuge filter tubes. Filtered tubes were placed vertically on the centrifuge wall and centrifuged at 10,000 rpm for 8-10 minutes. The same procedure was performed for the remaining protein samples. After the last centrifugation, 1x TBS was transferred into the filter tube and after the liquid in the tube was pipetted several times against the filter wall, all the liquid was transferred to a new tube. Glycerol as much as the amount of liquid in the tube was added into the tube and stored at -20C.

3.3.3 Bradford assay

A range of dilutions were made for the unknown concentration of the protein sample (1, 1/10, 1/100, 1/1000). Duplicates of each sample were prepared. To get the calibration, various duplicated volumes of bovine serum albumin (BSA) standard and a blank were prepared. When all the samples were ready, assay reagent added and absorption at A595 were measured.

3.3.4 GST pull-down assay of AurA and peptides

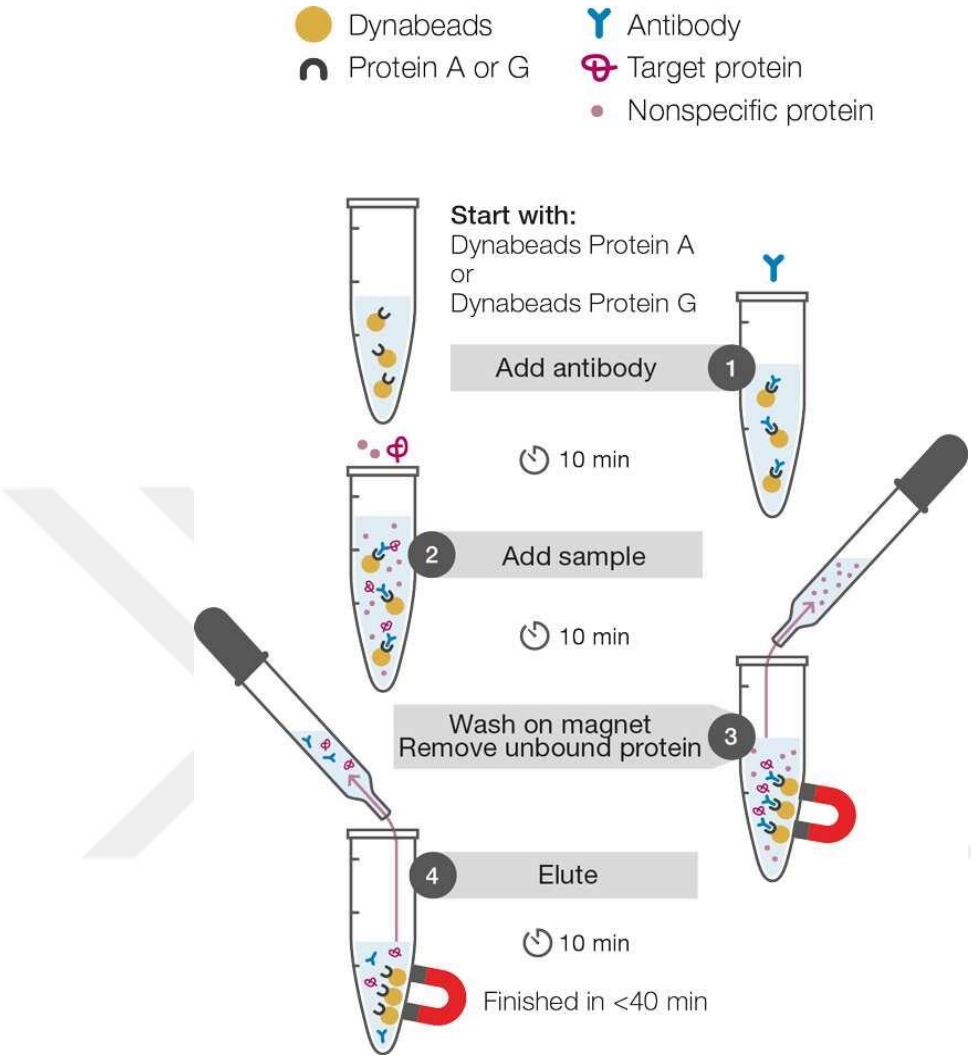

Figure 9: Illustration of GST pull-down assay working principle.

AurA recombinant protein was dissolved in 50 mM Tris (pH 7.5), 150 mM NaCl, 5 mM MgCl2, 5mMM- mercaptoethanol, 0.02% Tween-20 buffer solution (31). Protein G coated magnetic beads (Dynabeads, Invitrogen) were used in immunoprecipitation experiments. For preparation of magnetic beads, the tube was briefly resuspended by vortex. 50 µL of beads were separated from the liquid with a magnetic rack and washed once in a binding solution (PBS pH 7.4, 0.02% Tween 20). Then, for GFP-tagged AurA produced in E. coli, 10 µg of Anti-GFP antibody (Abcam, ab290) or externally purchased GST-labeled AurA anti-GST antibody (Abcam, ab92) added in the same binding solution resulting in total of 200 µL antibody solution. This solution was added to the washed magnetic beads and the beads were incubated at

room temperature on an orbital shaker at different time intervals. Incubation with the antibody was terminated with the aid of a magnet and the antibody solution was separated from the beads. Accompanying this incubation, in a separate tube, the recombinant AurA and peptides were incubated in the binding solution. In order to optimize the AurA and peptide binding conditions, different concentrations of peptide and AurA were used, and the incubation time was changed. After AurA-peptide incubation, this solution was added to the antibody-loaded magnetic beads and incubated again for 45 min at room temperature. The beads were then separated from the liquid and eluted with the binding elution solution (50 mM glycine pH 2.8) in the beads. The elution result of tris-tricine was analyzed by SDS-PAGE according to their size.

3.4 Computational Methods

3.4.1 FoldX and MM-PBSA calculations

FoldX calculates protein-protein interactions by utilizing an empirical force field. The contributions of the N-Myc MBI motif amino acids to the AurA binding free energy in this complex were analyzed with the FoldX PSSM instruction (48).The PSSM command, also includes the Analyze command, operates by converting the selected amino acids into all possible amino acid combinations. The contribution of the change in the relevant position was determined by calculating the difference (ΔΔG) between the binding free energy in mutant peptide-bearing complexes (ΔGmt) and the AurA/N-Myc binding free energy (ΔGnt) in the native complex. FoldX results are given as "mean ± standard error" and ΔΔG values less than -0.5 were chosen to be used in the next step. By MM-PBSA method, the binding free energy of mutations in the selected peptide sequence is calculated and ranked. In this context, the MM-PBSA methodology has previously been used for the analysis of protein-peptide complexes (48,49). During the MM/PBSA calculation, the parameters of the model were chosen to be obtained from the publication of this tool and to be the same as the parameters used in the preliminary study (50). MM-PBSA is basically an operational method that combines molecular mechanics and continuous solvents to calculate the bond free energy (51).

3.4.2 System building and *in-silico* mutagenesis

System for Molecular Dynamic Simulations was prepared using CHARMM-GUI Input generator, solution builder tool (52). Rename to ADP and mutation options were set within the PDB manipulation options, mutations were added according to the table, water box type selected as rectangular, and Na^+ and Cl^- ions were added. Temperature of the system determined as 303.15 K Input generation option selected as NAMD and default parameters were selected as CHARMM36 force field (52). System input files were visualized via Visual Molecular Dynamics (VMD) (53).

Table 5: Sequences of peptides after substitutions

Peptide	**Sequence**	**Number of Substitutions**
WT (5G1X)	SLQPCFYPDEDDFYFGGPDSTPPGEDIWKKFE LLPTPPLSPSRGFAEXSSEPPSWVTEMLLENEL	0
Semi-rational peptide	LEFDSLQMCFYDDEDDWYFGGRDSTPPGEDI WDKFELLPTPPMSPFRGFWEWSSEPPSWVTE MLLENELWG	9
P1	SFQPCFYRDEDDFYFGGPDSEPPGEDEWKKFE LLPEPPYSFSDGFAEXSSEPPSWITEFLLEWEL	12 (9 + 3)
P2	SFQPCFRPDEDDFYFGGPDSEPPGEDEWKKFE LLPEPPYSFSDGFAEXSSEPPSWITEFLLEEEL	12 (9 + 3)
P3	SFQPCFYRDEDDFYFGGPDSEPPGEDEWKKFE LLPEPPYSFSDGFWEXSSEPPSWITEFLLEEEL	13 (9 + 4)
P4	SFQPCFRPDEDDFYFGGPDSEPPGEDEWKKFE LLPEPPYSFSDGFWEXSSEPPSWITEFLLEEEL	13 (9 + 4)

3.4.3 Molecular dynamic simulations

Molecular dynamics (MD) simulations of the systems that were prepared in the previous steps were carried out with TRUBA (Turkish National Science e-Infrastructure) and analyzed by MD simulations using parameters previously optimized by us and other research groups. Accordingly, all systems were simulated with NPT thermodynamic ensemble for 100 ns, under 303.15 K temperature and 1 atm

pressure, with periodic boundary conditions and CHARMM36 force field (52) with NAMD molecular dynamics algorithm (54). All simulation trajectories and structures were visualized with Visual Molecular Dynamics (VMD) (53). The stability of the simulations was evaluated according to the root mean square displacement (RMSD) of the backbone atoms (C, N, Cα, O).

4 RESULTS

4.1 Recombinant AurA expression

4.1.1 IPTG expression and protein solubility assay

The clone used for AurA expression was obtained from Addgene (Clone numbers: 99875, 99876) as cloned into a vector allowing for bacterial expression. Both IPTG and Auto-induction expressions methods were applied to decide the optimal expression method. As a result, it was decided to also test the Auto-induction methodology. For both pAurora-GFP-AURKA-mCherry and pAurora-GFP-AURKA, soluble lanes are less dense than insoluble lanes. The molecular weight of Aurora and mCherry together (pAurora-GFP-AURKA-mCherry) is about ~100 kDa, Aurora and GFP together (pAurora-GFP-AURKA) is ~70 kDa. There were no significant bands observed at 70 kDa and 100 kDa range.

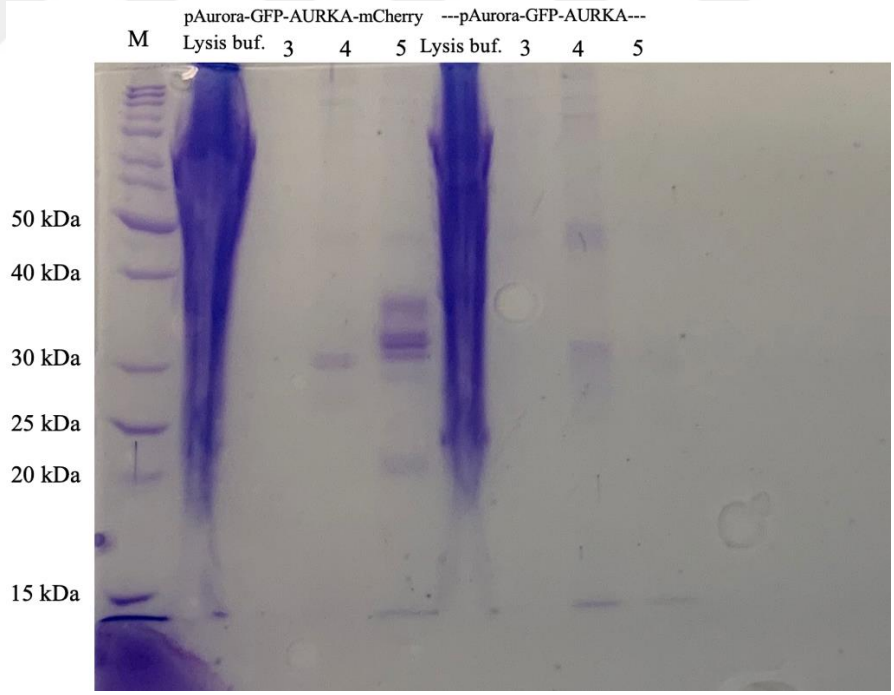

Figure 10: SDS-PAGE gel picture of expression with IPTG method

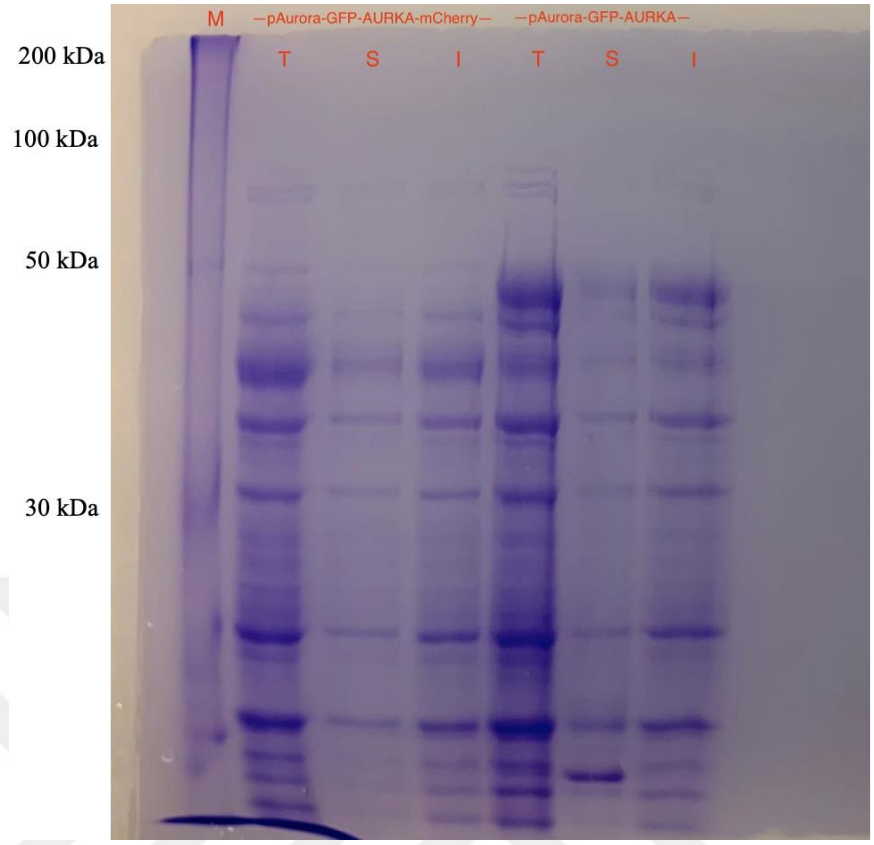

Figure 11: SDS-PAGE gel analysis of pAurora-GFP-AURKA-mCherry and pAurora-GFP-AURKA. (M: protein ladder, T: Total samples, S: Soluble samples, I: Insoluble samples)

4.1.2 Auto-induction expression

As it is seen in Figure 11, after the expression step, pinkish color is observed in the cell suspension. This result can be interpreted as a possibly successful mCherry expression but for the whole protein, SDS-PAGE analysis is performed in the next step to decisively show the right sized protein expression (Figure 15). In addition to that, as it is seen in Figure 12, cell suspension samples were evaluated under UV light to detect the activity of the mCherry and GFP (Green fluorescent protein) and showed fluorescent activity before the purification step.

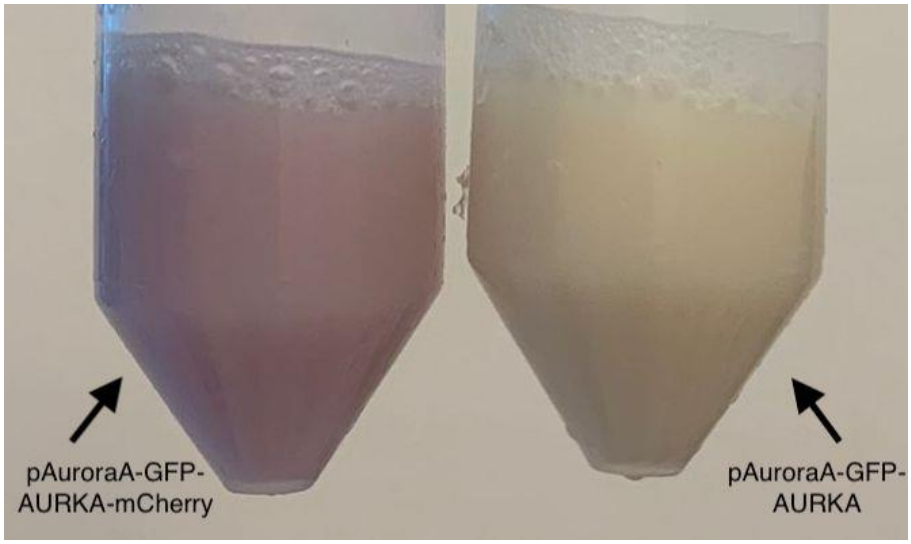

Figure 12: Cell suspension of pAuroraA-GFP-AURKA-mCherry and pAuroraA-GFP-AURKA

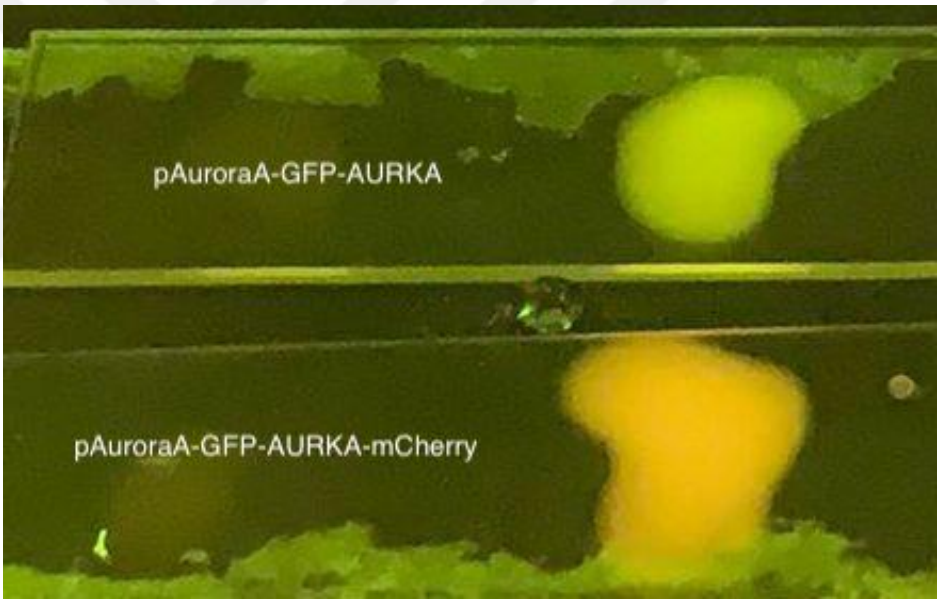

Figure 13: pAuroraA-GFP-AURKA-mCherry and pAuroraA-GFP-AURKA under UV light.

4.2 Purification SDS-PAGE Analysis

The purification step was carried out by affinity chromatography over the poly-histidine tag, as mentioned in the method. When the samples were loaded in the purification column, mCherry protein showed a significant color as the previous steps (Figure 14). During the purification procedure, the wash step controlled with the Bradford solution and continued after the reagent turned into a light blue color. Chromatography columns were monitored during the washing step and pinkish color remained in the columns until the elution step. The pinkish color is observed in the eluted fraction tube as well as it is seen in Figure 15. Molecular weight of Aurora and mCherry together (pAurora-GFP-AURKA-mCherry) is about ~100 kDa, Aurora and GFP together (pAurora-GFP-AURKA) is ~70 kDa (55). SDS-PAGE gel analysis picture (Figure 16) shows that our target Aurora protein is not present in the purified samples although it is predicted that mCherry protein is in fact present in the elution solution (Figure 15). A band can be seen around ~30kDa that can be observed in wells 4 and 5 for sample pAurora-GFP-AURKA-mCherry. This is similar to the size of native mCherry (~26 kDa) protein. Because of this, it can be interpreted as a breakage between the fused proteins might have happened here. There were no bands around 100kDa which is the expected size of the fused AURORA-GFP-mCherry protein. Broken Aurora-GFP was suspected to be lost in the wash procedure.

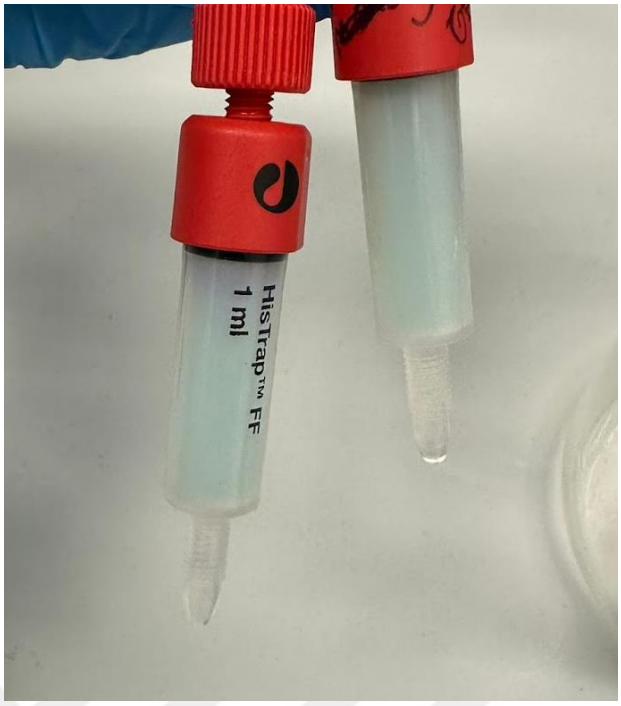

Figure 14: His-tag affinity columns during purification procedure.

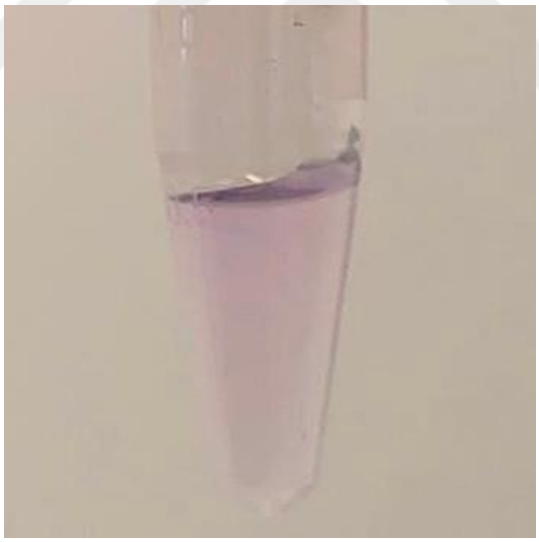

Figure 15: mCherry protein fraction after elution step.

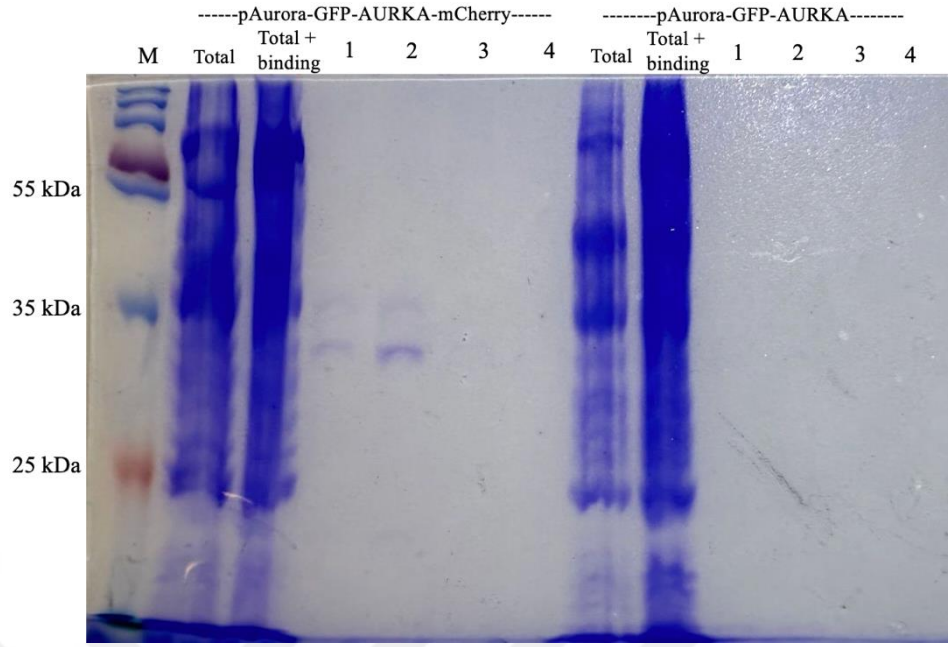

Figure 16: SDS gel picture of pAuroraA-GFP-AURKA-mCherry and pAuroraA-GFP-AURKA total protein samples before purification and elution fractions.

4.3 GST Pull-Down Assay of AurA and Peptides

As it is seen in the Tris-tricine gel (Figure 17), commercially purchased Aurora A Kinase and target peptide's concentrations were tested to determine the optimal concentration to use in the GST pull-down assay. As it is seen in Figure 18, magnetic beads (Dynabeads, Invitrogen) were tested with the antibodies, Aurora, and peptide to adjust the assay incubation time and protein concentration. Lastly, the result of the GST pull-down assay can be seen in the tris-tricine gel (Figure 19), and the lanes of the procedure as input, wash, elution of the antibody-peptide sample and elution for the both antibody-AurA-control peptide and antibody-AurA-peptide.

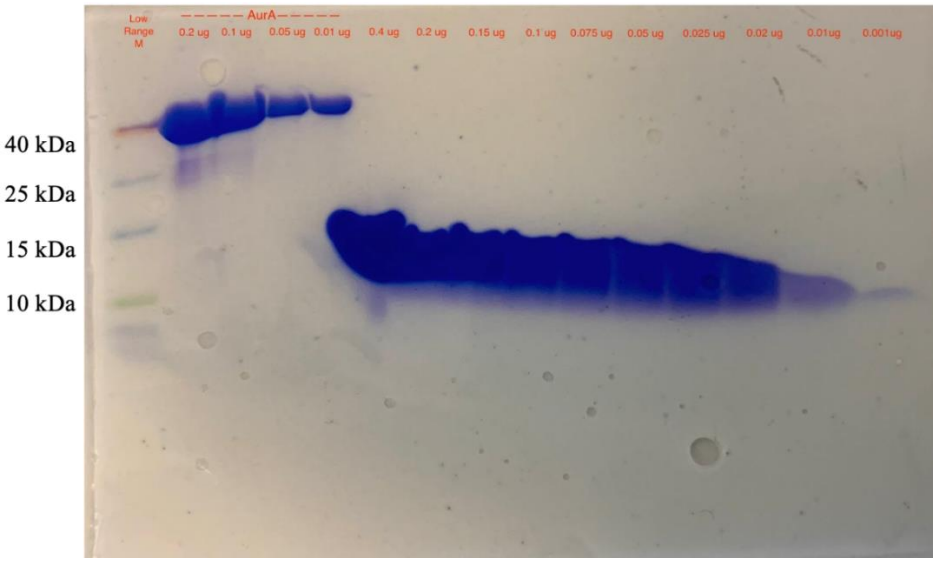

Figure 17: Tris-tricine gel picture of commercially purchased Aurora Kinase A and peptide.

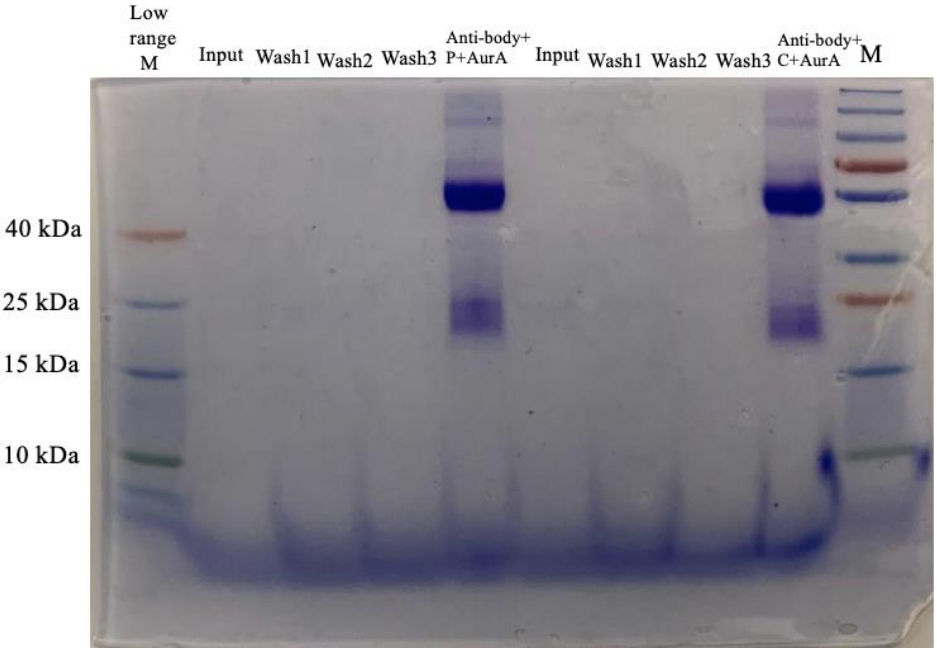

Figure 18: Tris-tricine gel picture of first GST pull-down assay.

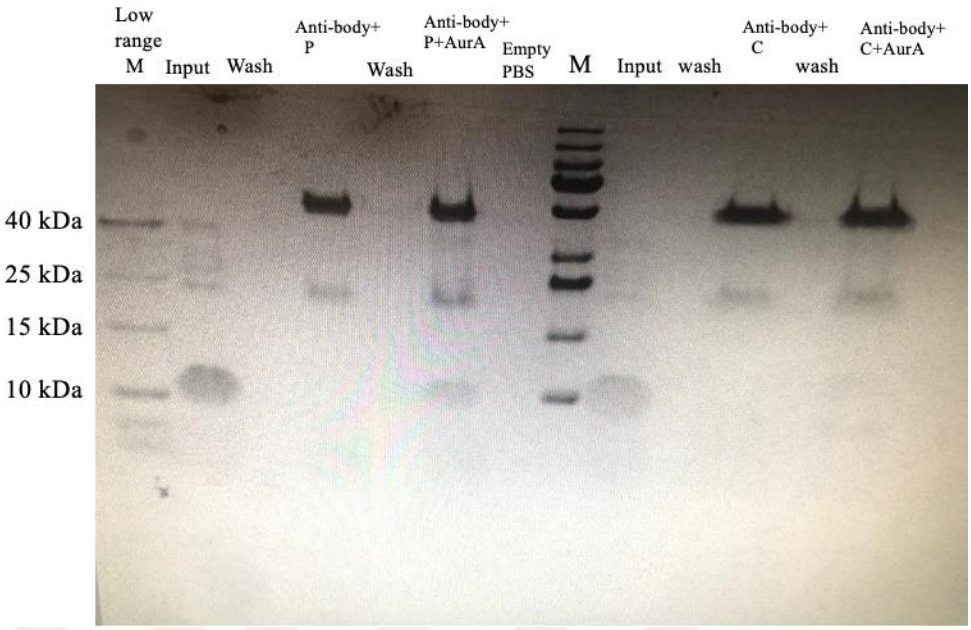

Figure 19: Tris-tricine gel picture of last GST pull-down assay.

4.4 FoldX and MM-PBSA Calculations

The heatmap constructed with python for the output files of FoldX (48) is shown in the Figure 20. According to the heatmap, stability decreases as the blue shade gets darker and increases as the blue shade gets lighter. Since there were not enough mutations greater than 0, it is decided to include the range between 0 and -5 according to the binding free energy scores. The heatmap generated from the binding free energy (ΔΔG) calculations (Figure 20), is used for determining the amino acid mutation that will best fit the stabilization criteria and final mutation list is constructed also considering the interactions within 4 angstrom amino acid residues.

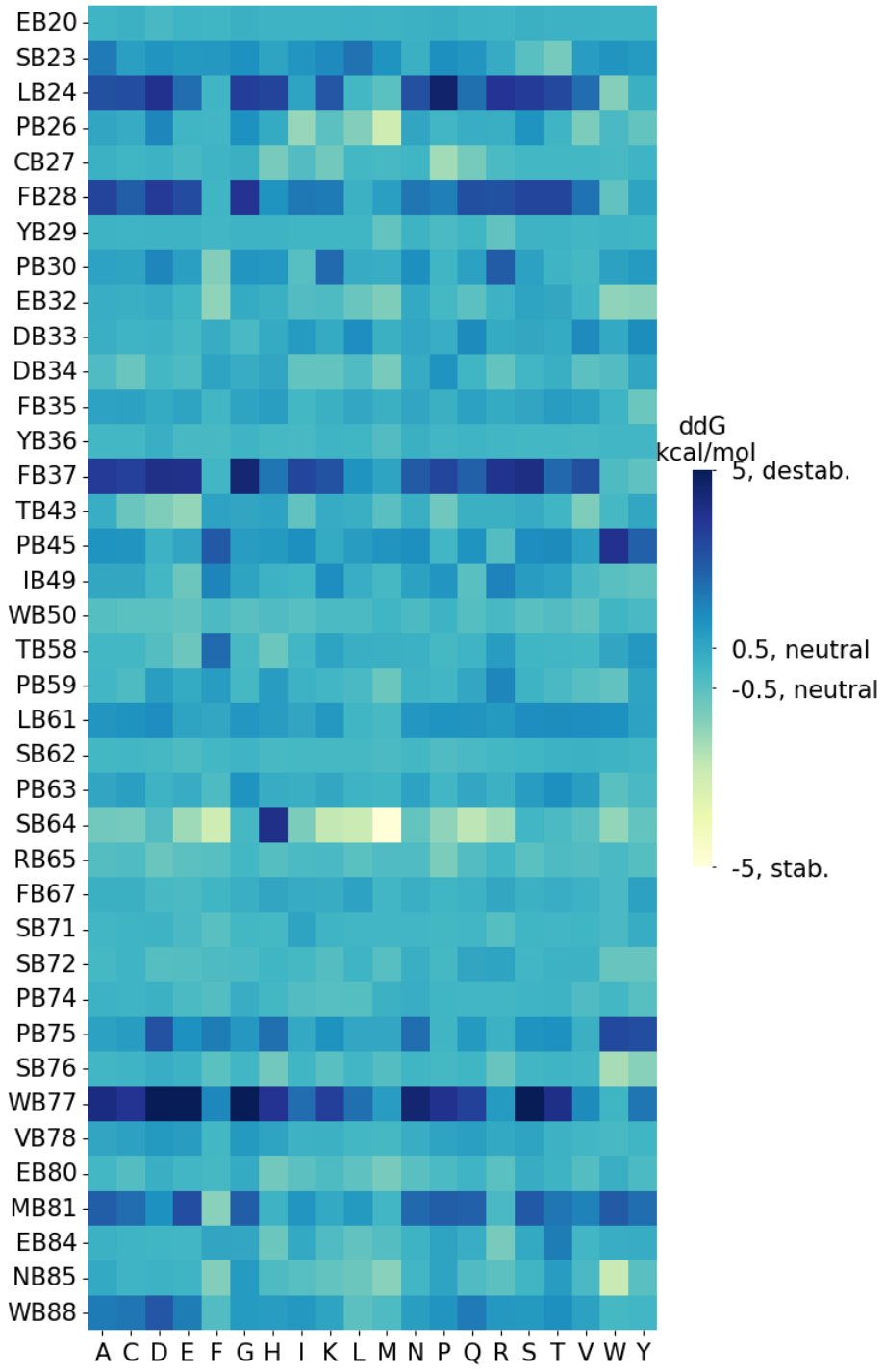

Figure 20: Heatmap constructed from the FoldX calculations of ΔΔG.

4.5 System Building and *in-silico* Mutagenesis

With CHARMM-GUI, 4 different system input files were prepared. As it is seen in the Figure 21, AurA/N-Myc in the water box is visualized by VMD (53). Two of the input files consist of 11 mutations at and two of the input files consist of 12 mutations that were aligned with the ClustalW, and the output option were chosen as ClustalW character count (56) shown in the Figure 22. Also all of the mutation locations represented in the (Figure 23 - 29), again prepared with VMD (53).

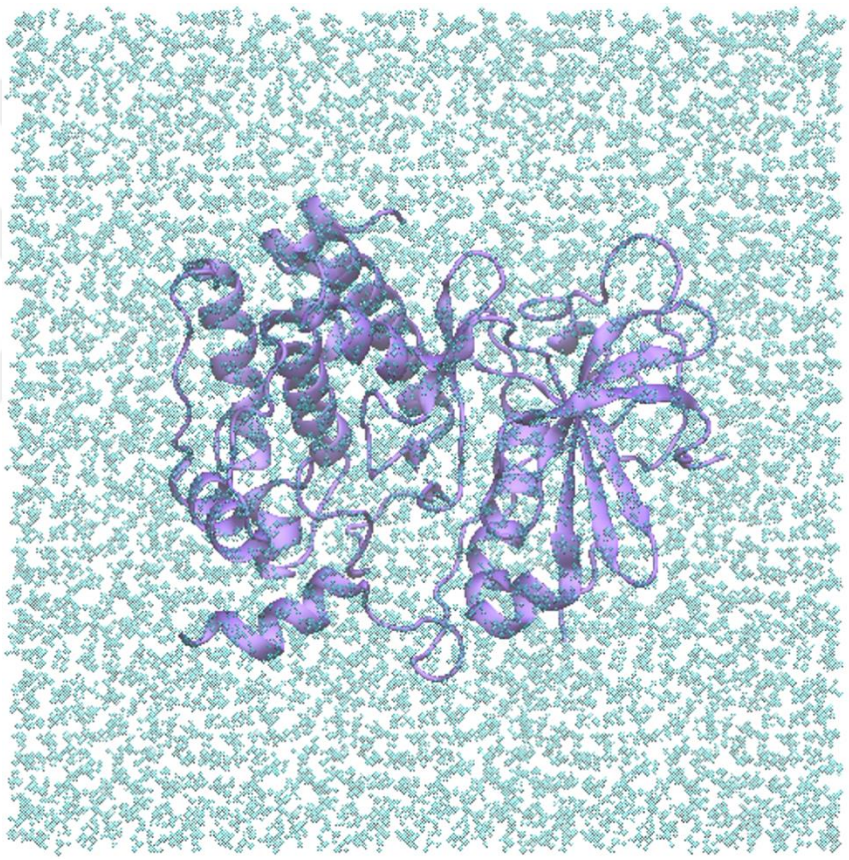

Figure 21: VMD Snapshot of the MD input files in water box in new cartoon, violet.

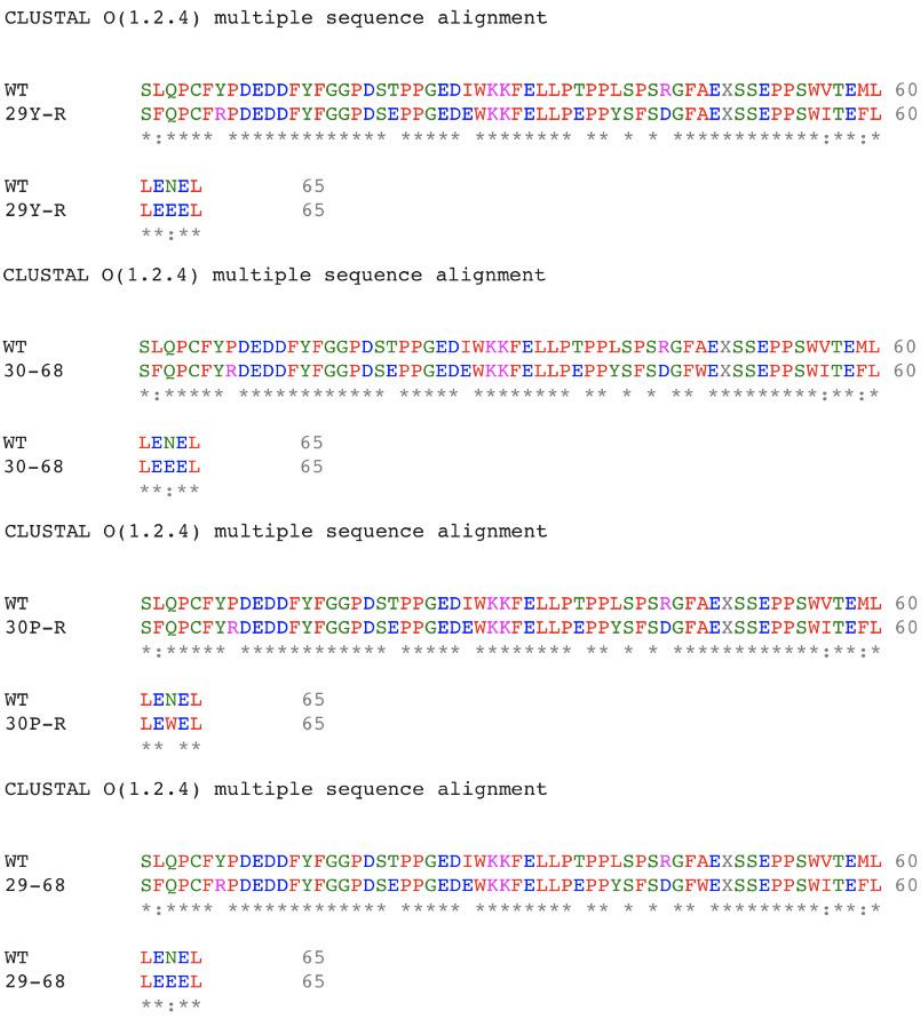

Figure 22: Sequence alignment of 4 rational designed peptide sequences of MB0 with Wild type sequence.

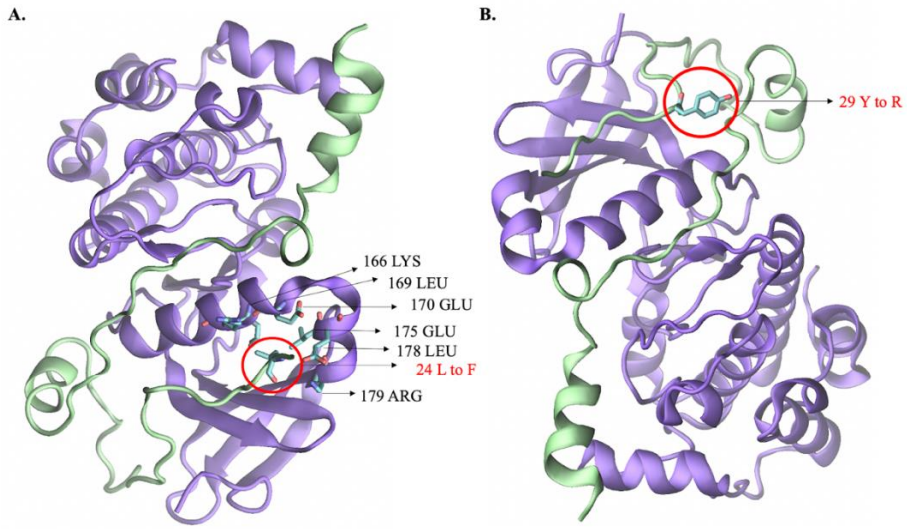

Figure 23: Snapshots of the single mutation and the residues within 4 angstroms of the mutated sequences. AurA: New cartoon, Violet. N-Myc: New cartoon, lime. Mutated residue: Licorice, Name. (A) Residue 24 pointed out with red and 4 angstrom interactions pointed out black. (B) Residue 29 pointed out with red and 4 angstrom interactions pointed out black.

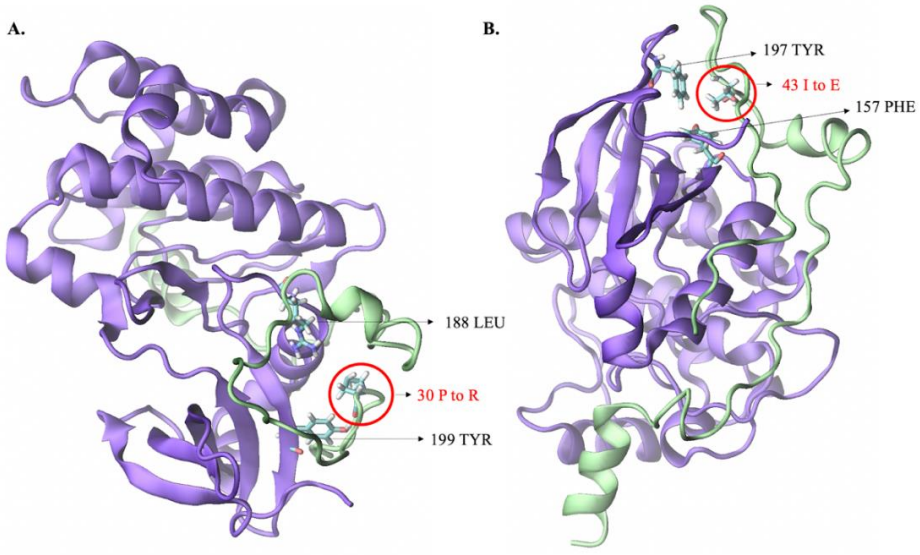

Figure 24: Snapshots of the single mutation and the residues within 4 angstroms of the mutated sequences. AurA: New cartoon, Violet. N-Myc: New cartoon, lime. Mutated residue: Licorice, Name. (A) Residue 30 pointed out with red and 4 angstrom interactions pointed out black. (B) Residue 43 pointed out with red and 4 angstrom interactions pointed out black.

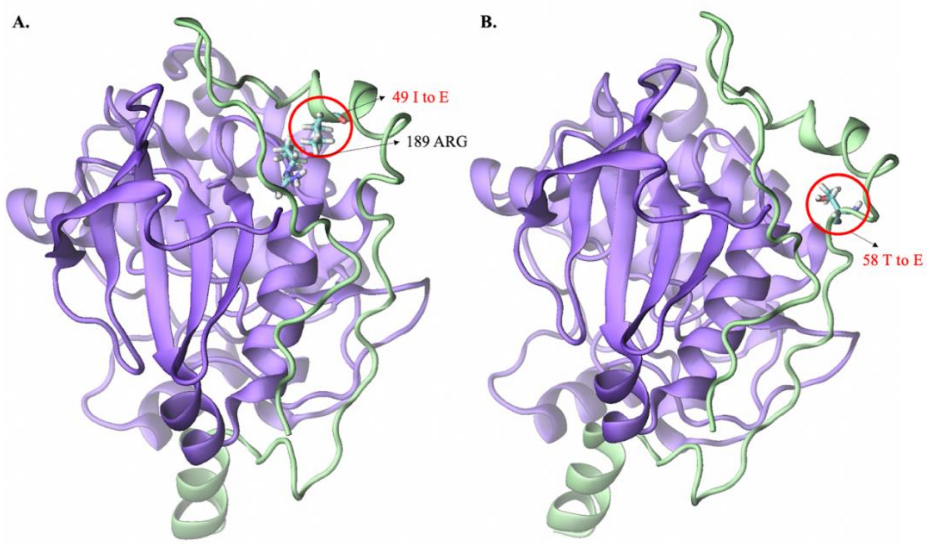

Figure 25: Snapshots of the single mutation and the residues within 4 angstroms of the mutated sequences. AurA: New cartoon, Violet. N-Myc: New cartoon, lime. Mutated residue: Licorice, Name. (A) Residue 49 pointed out with red and 4 angstrom interactions pointed out black. (B) Residue 58 pointed out with red and 4 angstrom interactions pointed out black.

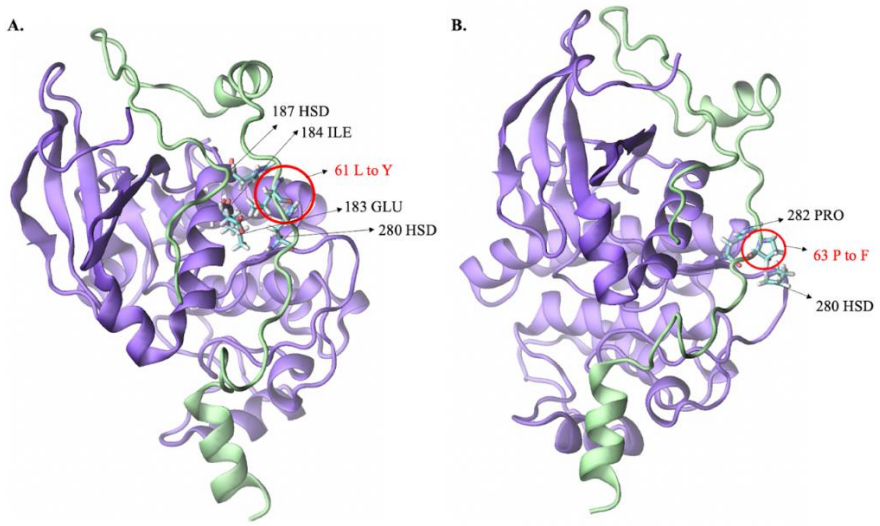

Figure 26: Snapshots of the single mutation and the residues within 4 angstroms of the mutated sequences. AurA: New cartoon, Violet. N-Myc: New cartoon, lime. Mutated residue: Licorice, Name. (A) Residue 61 pointed out with red and 4 angstrom interactions pointed out black. (B) Residue 63 pointed out with red and 4 angstrom interactions pointed out black.

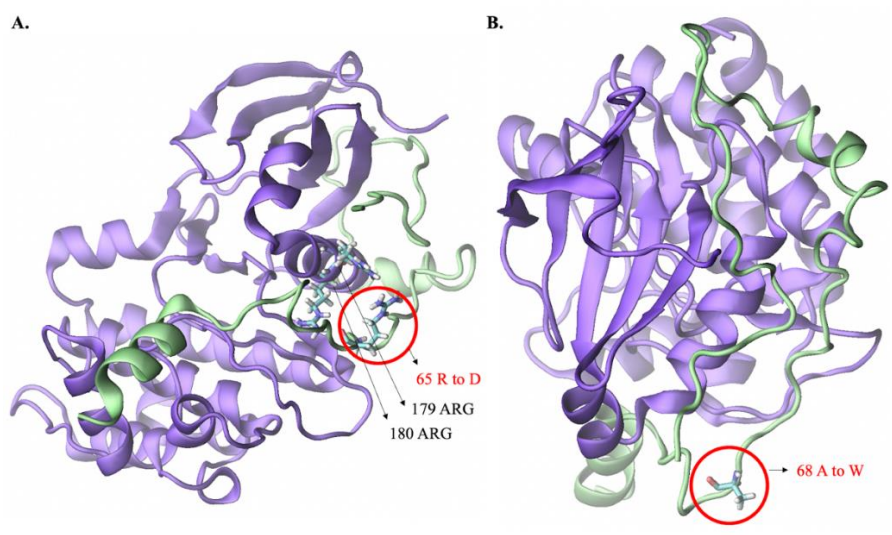

Figure 27: Snapshots of the single mutation and the residues within 4 angstroms of the mutated sequences. AurA: New cartoon, Violet. N-Myc: New cartoon, lime. Mutated residue: Licorice, Name. (A) Residue 65 pointed out with red and 4 angstrom interactions pointed out black. (B) Residue 68 pointed out with red and 4 angstrom interactions pointed out black.

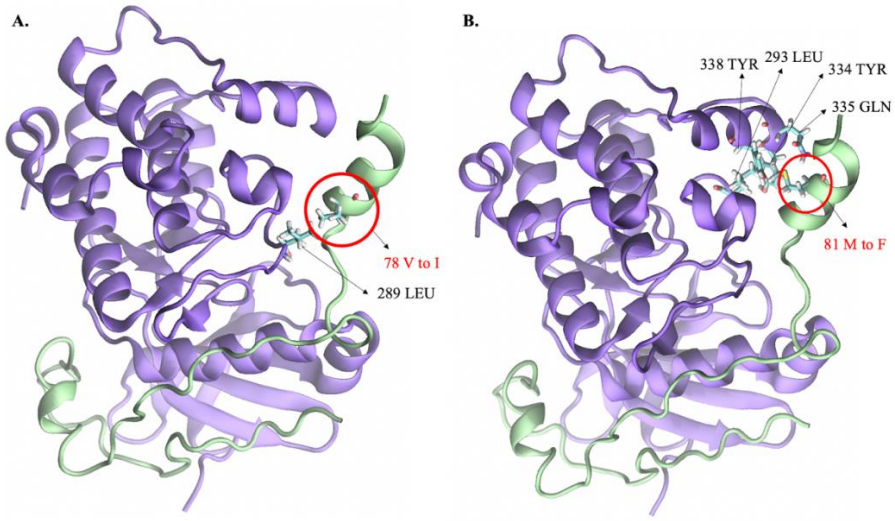

Figure 28: Snapshots of the single mutation and the residues within 4 angstroms of the mutated sequences. AurA: New cartoon, Violet. N-Myc: New cartoon, lime. Mutated residue: Licorice, Name. (A) Residue 78 pointed out with red and 4 angstrom interactions pointed out black. (B) Residue 81 pointed out with red and 4 angstrom interactions pointed out black.

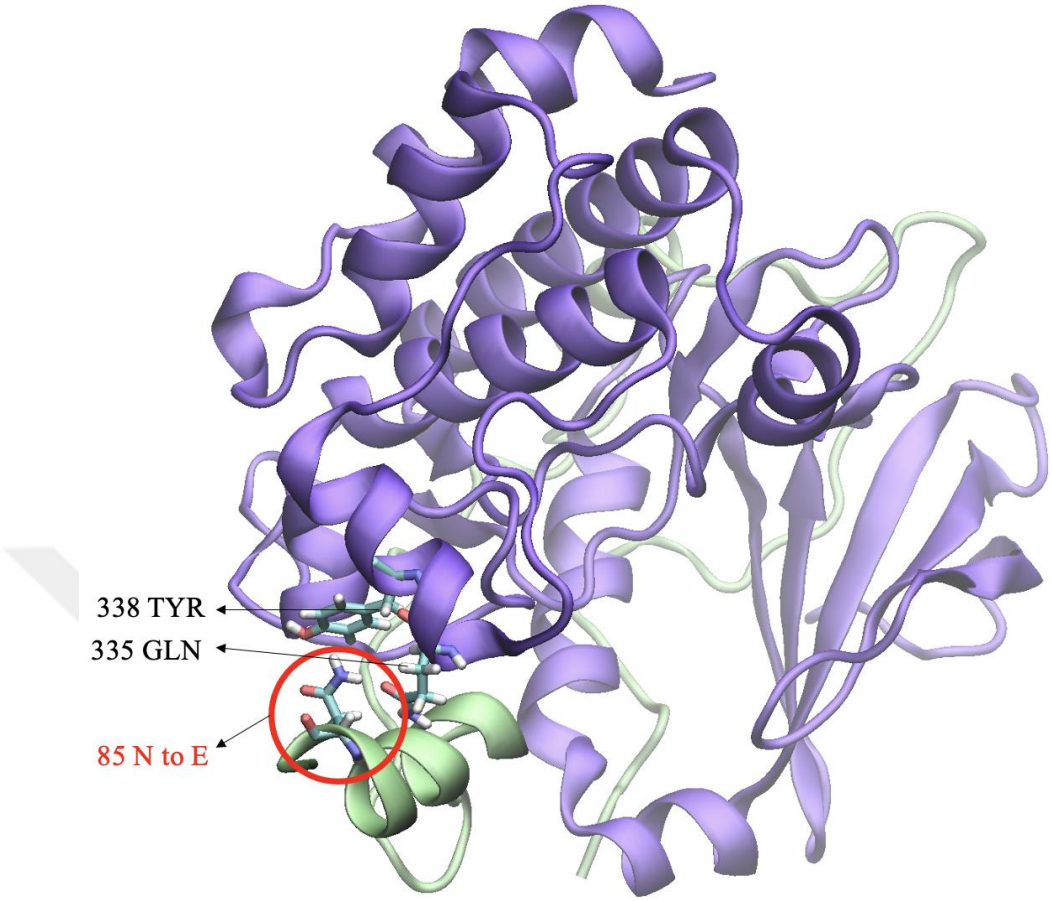

Figure 29: Snapshots of the single mutation and the residues within 4 angstroms of the mutated sequences. AurA: New cartoon, Violet. N-Myc: New cartoon, lime. Mutated residue: Licorice, Name. Residue 85 pointed out with red and 4 angstrom interactions pointed out with black.

4.6 Molecular Dynamic Simulations

Graph of RGYR (radius of gyration), is shown in Figure 30. The green line represents P1, red line represents P2, blue line represents P3, and the purple line represents P4. Sequences of these peptides were given in Table 5. The graph includes all 100 ns of the simulation, starting from time 0.

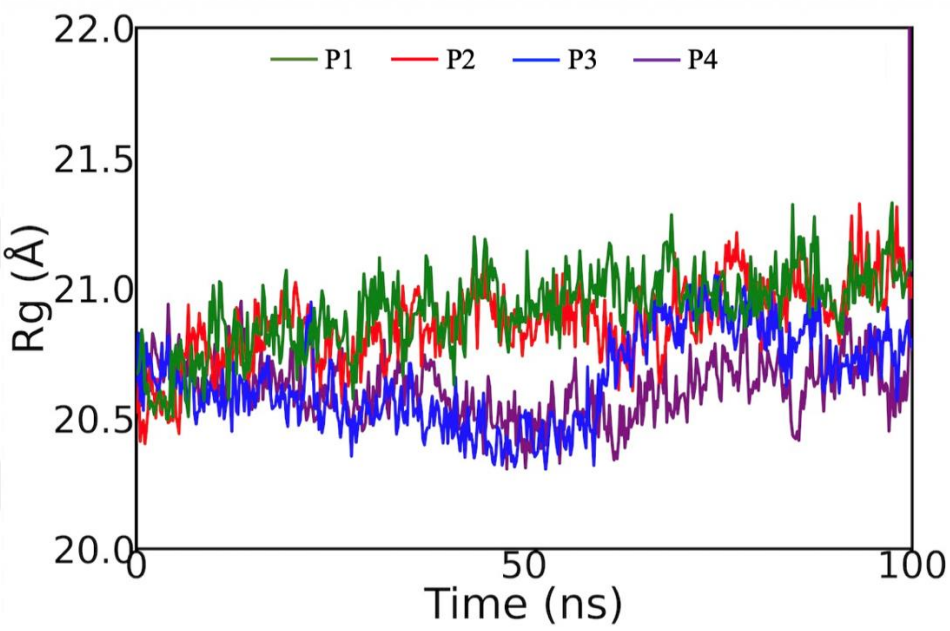

Figure 30: The graph of RGYR from MD simulations. Green line represents P1, red line represents P2, blue line represents P3, and the purple line represents P4.

Graph of RMSF (root mean square fluctuation), is shown in Figure 31. The green line represents P1, red line represents P2, blue line represents P3, and purple line represents P4. Sequences of these peptides were given in Table 5. Aurora (residue 126-234) and the peptide (residue 23-87) analyzed in separate graphs to provide detailed visualization. Again, graph of RMSD of the semi-rational designed peptide can be seen in Figure 32. The semi rational designed peptide has a length of 70 amino acid residues (19-89), making this sequence 6 amino acid longer than the rational designed peptides (23-87).

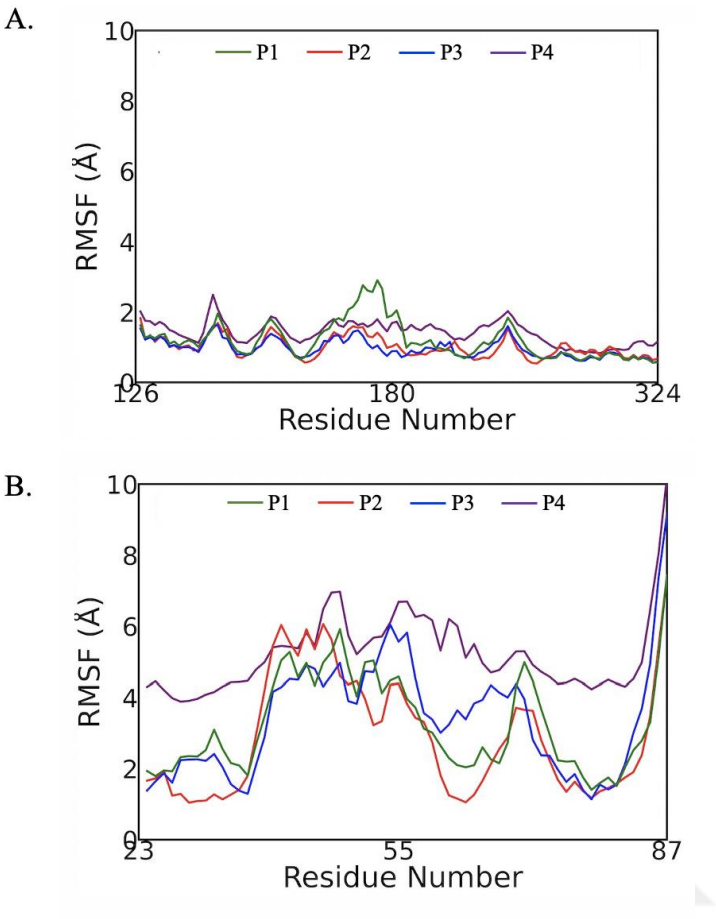

Figure 31: The graph of RMSF from MD simulations. (A) represents Aurora (residues from 126-234). (B) represents the peptide (residues from 23-87). The green line represents P1, red line represents P2, blue line represents P3, and purple line represents P4.

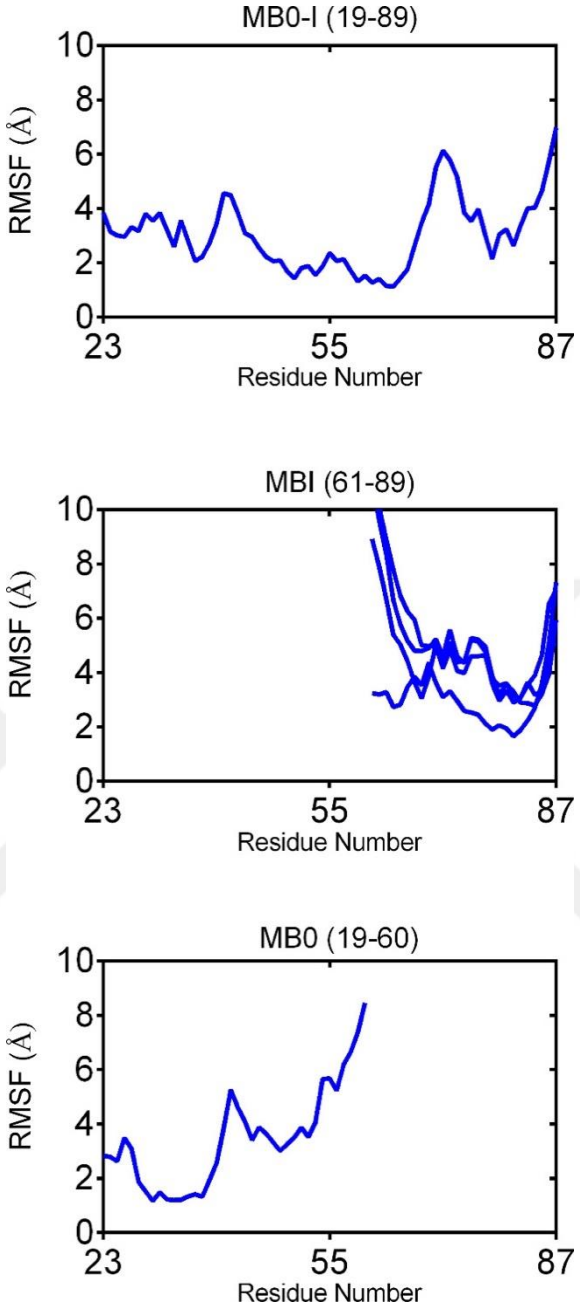

Figure 32: The graph of RMSF of semi-rational designed peptide from MD simulations.

The graphs of Pairwise RMSD (Root mean square deviation) for the wild-type peptide and the mutated peptides (P1, P2, P3, P4) are shown in Figure 33. Sequences of these peptides were given in Table 5. The graph includes all 100 ns of the simulation, starting from time 0. As the color changes from dark blue to yellow, conformational changes in the protein increases.

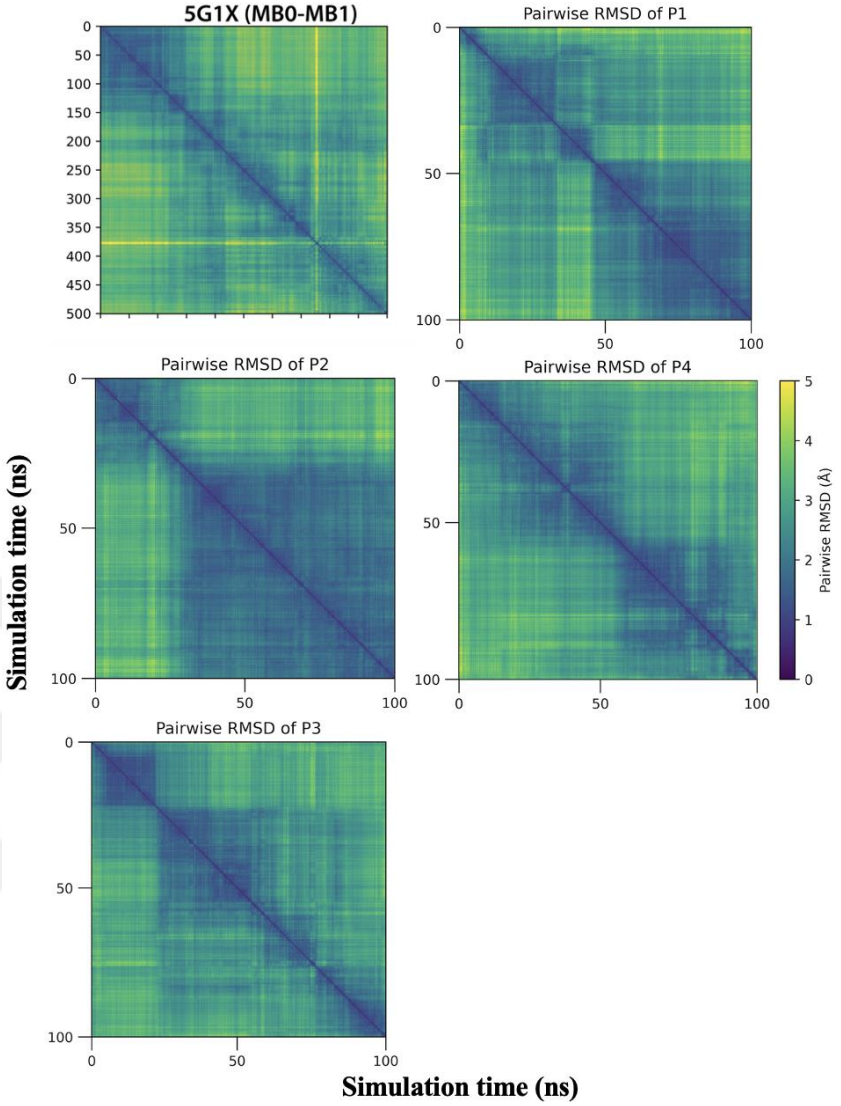

Figure 33: The graphs of Pairwise RMSD (Root mean square deviation) for the wild-type peptide and the mutated peptides (P1, P2, P3, P4).

5 DISCUSSION

5.1 Results of IPTG Expression and Solubility Assay

As mentioned before, there were no significant protein bands observed at the ~70kDa and ~100kDa range. It was decided to proceed with Auto-induction methodology since with the IPTG expression method, target proteins could not be detected as it is seen in Figure 9. Auto-induction is a simpler and cheaper expression method compared to IPTG expression and it can produce slightly more cell volume than IPTG induction (46). Increasing the expression time can increase the cell volume and can also affect the quantity of the target protein. SDS-PAGE analysis of the solubility assay of IPTG expression is given in Figure 10. As expected, the insoluble protein amount is higher than soluble amount when total protein is referenced (57). At this point, it was decided that the Auto-induction expression method should be also tried.

5.2 Results of Expression by Auto-Induction

There are many benefits of working with a fluorescent protein during the expression process. First and one of the most important functions of a fluorescent protein is the ability to visually track the real-time presence during protein expression and purification procedures. Another benefit is by evaluating the color intensity, researchers can assess the quantity of the expressed protein. In addition to that, fluorescent proteins like mCherry are also compatible with many of the experimental techniques (58). Hence, this result gives an important insight into the status of the protein before the purification procedure, since the pink color of the mCherry is present all through the expression procedure.

5.3 Results of Purification of Recombinant Kinase

It is suspected that several reasons could have caused this issue during the procedure causing a split between mCherry and AurA protein. The first reason that is

considered lies within the extraction methodology of the protein. There are multiple protein extraction methods that are available. In this thesis, sonication which is a physical extraction method and B-PER (Bacterial protein extraction reagent) which is a chemical extraction method are utilized as cell lysis methods (59). Another reason could be the stability of the Aurora A protein, since it is a human protein, it could be very sensitive to the external working environment while handling the samples during the procedures (60). Additionally, composition of the binding, elution and wash buffers which are given in appendix. Adjustments of the concentrations of imidazole, pH of the buffer and salt concentration in the purification buffers plays a crucial role in the purification procedure (61) and determination of the optimum range of the buffer composition for a specific protein could have taken too much time for this thesis, so it is decided to purchase the Aurora protein commercially to maintain the timeline of the thesis.

5.4 Results of GST Pull-Down Assay of AurA and Peptides

In the tris-tricine gel picture (Figure 17) it is seen that, commercially purchased Aurora kinase A (~70 kDa) is above of the 40 kDa band and the peptide sample is between 15 kDa and 10 kDa bands. According to Figure 17, it is decided to proceed with the concentration range of 0.05 ug and 0.01 ug for Aurora kinase A and 0.01 ug for the peptide to carry out the GST pull-down assay. First, binding between magnetic beads and antibody+Aurora+peptide is tested to adjust the incubation period (Figure 18). It has been decided to increase the incubation time 1 hour to 3 hours to increase the binding efficiency. It is usually traditional to see antibodies as two separate bands in the gel pictures of coimmunoprecipitation (Co-Ip) assays (62). As the result of antibody-magnetic bead binding assay, this interaction and incubation time can be interpreted as enough for the next step which is the pull-down assay of antibody, AurA, and peptide. Finding from Figure 19 may help to understand and determine the condition of the samples in each step of the procedure. As control for the interaction between only peptide and the antibody, one of the samples did not include Aurora A kinase. This result confirmed there were no binding between antibody and the peptide directly. In contrast to antibody-peptide lane, it is observed that there is a light binding

between Aurora kinase A-peptide and antibody although part of the peptide is lost in the input lane. These findings further support the idea of a promise with the semi-rational peptide sequence that was designed in prior studies. However, some questions remained unanswered for the weak binding of the peptide. Thus, it is suggested that generating a rational design for the peptide sequence could solve this issue.

5.5 FoldX and MM-PBSA Calculations

It is important to bear in mind the properties of the surrounding amino acids could affect greatly the mutated amino acid. As an example, glutamate and aspartate usually can be detected in binding and active sites of the protein (63). Hence, the final decisions for choosing the substitutions were made considering both binding free energy ($\Delta\Delta G$) calculations and interactions between other close located amino acids.

5.6 Comparison of Peptide Sequences

According to the protein sequence alignment that were performed with ClustalW (56), two dots represent the conserved mutations meaning the mutated amino acid has similar biochemical properties. Non-conserved mutations represented with space classified as difference between biochemical properties such as hydrophobicity, charge, and size. As it is seen in Figure 22, amino acid substitutions between similar amino acid residues (in terms of classification) have unsignificant effect and substitutions between distinct amino acid residues have significant effect on the protein. Considering all available information, it is decided to proceed with the MD simulations using all four peptide sequences listed in Table 5.

5.7 Analysis of Molecular Dynamic Simulations

The radius of gyration analysis is a technique used to characterize the compactness or the overall size of a biomolecule or a macromolecular structure over time during the simulation. The radius of gyration (RGYR) for a molecule is a measure of its spatial extent and is calculated as the root mean square distance of the atoms from their center

of mass. It quantifies how much the atoms of the molecule are spread out or confined in space (64). According to the Figure 30, P2 (the red line) has stored its compactness the most by remaining more still throughout the simulation time in comparison to the other peptide sequences. RMSF stands for Root Mean Square Fluctuation, and it is a common analysis technique used in Molecular Dynamics (MD) simulations to examine the flexibility and dynamics of biomolecules such as proteins or nucleic acids. The RMSF analysis calculates the average deviation or fluctuation of each atom in a biomolecular system from its mean position over the course of an MD trajectory. It provides valuable insights into the local flexibility and conformational changes of the molecule (46). Comparing the lines in the graph (Figure 31), it can be said that P2 (the red line), has a notable decrease in movement, making it the most promising peptide sequence of them all. Also, RMSF graph of semi-rational designed peptides (Figure 32), provides an insight about the improved binding efficiency of the P2 as well. Pairwise RMSD analysis is a technique used in Molecular Dynamics (MD) simulations to assess the structural similarity or dissimilarity between different frames or conformations of a biomolecular system. It calculates the Root Mean Square Deviation (RMSD) between pairs of structures from an MD trajectory (65). The graph of the wild-type protein (5G1X), taken as a control and the graph of all four peptide sequences are given in Figure 33. From this data, when comparing the wild-type graph to substituted peptides, it can be seen that, all of the peptides (P1, P2, P3, P4) demonstrated decreased movement in regard to the wild-type sequence because they consist darker blue to green when the wild type is also examined. Especially, in respect to others, P2 seems to remain more stable yet again. Overall, the result of the MD simulations suggests that P2 (from Table 5), could demonstrate valuable outcomes for targeting the AurA/N-Myc binding surface.

6 CONCLUSION

In conclusion, the failure in the expression of Aurora A kinase during laboratory experiments is a multifaceted issue that can arise from various factors and challenges. This thesis has highlighted some of the common reasons behind unsuccessful protein expression attempts, including issues with the improper culture conditions and the delicacy of the human protein. It is evident that achieving successful protein expression requires meticulous optimization and troubleshooting that could have taken too long and jeopardize the timeline of the thesis. The successful GST pull-down assay conducted with the commercially purchased Aurora has proven to be a powerful and reliable technique for studying protein-protein interactions. By utilizing the glutathione-S-transferase (GST) tag as an affinity handle, it was able to capture the GST-tagged bait protein along with its interacting partners which are Aurora A kinase and N-myc. The specificity and sensitivity of the assay allowed to discern genuine interactions from nonspecific binding, providing valuable insights into AurA/N-Myc binding surface interaction. The data obtained from this assay not only contributes to our understanding of protein interactions but also laid the foundation for further investigations into rational design of the peptide that can interact with the AurA/N-Myc binding surface more efficiently, directing the thesis into the computational experiments. The Molecular Dynamics (MD) simulation graphs presented in this thesis provide compelling evidence supporting the rational design of a peptide (P2) that exhibits improved binding to the target surface. Through systematic computational simulations, valuable insights have been gained into the dynamic behavior and intermolecular interactions of the designed peptides. The analyses indicate enhanced stability and stronger binding affinity between the P2, and the AurA/N-Myc binding surface compared to the wild-type. The successful rational design highlights the effectiveness of computational approaches in tailoring peptide properties for specific applications. Overall, these findings have significant implications for the development of novel biomaterials, drug delivery systems, and targeted therapies, where improved binding interactions are crucial for enhanced performance and efficiency.

7 REFERENCES

1. Willems E, Dedobbeleer M, Digregorio M, Lombard A, Lumapat PN, Rogister B. The functional diversity of Aurora kinases: a comprehensive review. Cell Div. 2018 Dec;13(1):7.
2. Naso FD, Boi D, Ascanelli C, Pamfil G, Lindon C, Paiardini A, et al. Nuclear localisation of Aurora-A: its regulation and significance for Aurora-A functions in cancer. Oncogene. 2021 Jun 10;40(23):3917–28.
3. D'Assoro AB, Haddad T, Galanis E. Aurora-A Kinase as a Promising Therapeutic Target in Cancer. Front Oncol [Internet]. 2016 Jan 6 [cited 2023 Jul 2];5. Available from: http://journal.frontiersin.org/Article/10.3389/fonc.2015.00295/abstract
4. Lin X, Xiang X, Hao L, Wang T, Lai Y, Abudoureyimu M, et al. The role of Aurora-A in human cancers and future therapeutics.
5. Krenn V, Musacchio A. The Aurora B Kinase in Chromosome Bi-Orientation and Spindle Checkpoint Signaling. Front Oncol [Internet]. 2015 Oct 16 [cited 2023 Jul 2];5. Available from: http://journal.frontiersin.org/Article/10.3389/fonc.2015.00225/abstract
6. Fellmeth JE, Gordon D, Robins CE, Scott RT, Treff NR, Schindler K. Expression and characterization of three Aurora kinase C splice variants found in human oocytes. Mol Hum Reprod. 2015 Aug;21(8):633–44.
7. Bavetsias V, Linardopoulos S. Aurora Kinase Inhibitors: Current Status and Outlook. Front Oncol [Internet]. 2015 Dec 21 [cited 2023 Jul 2];5. Available from: http://journal.frontiersin.org/Article/10.3389/fonc.2015.00278/abstract
8. Tang A, Gao K, Chu L, Zhang R, Yang J, Zheng J. Aurora kinases: novel therapy targets in cancers. Oncotarget. 2017 Apr 4;8(14):23937–54.
9. Yan M, Wang C, He B, Yang M, Tong M, Long Z, et al. Aurora-A Kinase: A Potent Oncogene and Target for Cancer Therapy: AURORA-A KINASE IN CANCER. Med Res Rev. 2016 Nov;36(6):1036–79.
10. Nguyen HG, Chinnappan D, Urano T, Ravid K. Mechanism of Aurora-B Degradation and Its Dependency on Intact KEN and A-Boxes: Identification of an Aneuploidy-Promoting Property. Molecular and Cellular Biology. 2005 Jun 1;25(12):4977–92.
11. Katayama H, Sen S. Aurora kinase inhibitors as anticancer molecules. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms. 2010 Oct;1799(10–12):829–39.
12. Ruiz-Pérez MV, Henley AB, Arsenian-Henriksson M. The MYCN Protein in Health and Disease. Genes. 2017 Mar 30;8(4):113.
13. Farina A, Faiola F, Martinez E. Reconstitution of an E box-binding Myc:Max complex with recombinant full-length proteins expressed in Escherichia coli. Protein Expression and Purification. 2004 Apr;34(2):215–22.
14. Dang CV. MYC on the Path to Cancer. Cell. 2012 Mar;149(1):22–35.
15. Schwab M, Alitalo K, Klempnauer KH, Varmus HE, Bishop JM, Gilbert F, et al. Amplified DNA with limited homology to myc cellular oncogene is shared by human neuroblastoma cell lines and a neuroblastoma tumour. Nature. 1983 Sep;305(5931):245–8.
16. Rickman DS, Schulte JH, Eilers M. The Expanding World of N-MYC–Driven Tumors. Cancer Discovery. 2018 Feb 1;8(2):150–63.
17. Brodeur GM, Seeger RC, Schwab M, Varmus HE, Bishop JM. Amplification of N- *myc* in Untreated Human Neuroblastomas Correlates with Advanced Disease Stage. Science. 1984 Jun 8;224(4653):1121–4.
18. Seeger RC, Brodeur GM, Sather H, Dalton A, Siegel SE, Wong KY, et al. Association of Multiple Copies of the N- *myc* Oncogene with Rapid Progression of Neuroblastomas. N Engl J Med. 1985 Oct 31;313(18):1111–6.
19. Schwarz N, Bless H, Strack F, Klumpp G, Rittenauer-Schatka H, Simons A. Ease of Retrieval as Information: Another Look at the Availability Heuristic.
20. Swartling FJ, Savov V, Persson AI, Chen J, Hackett CS, Northcott PA, et al. Distinct Neural Stem Cell Populations Give Rise to Disparate Brain Tumors in Response to N-MYC. Cancer Cell. 2012 May;21(5):601–13.
21. Beltran H, Tomlins S, Aparicio A, Arora V, Rickman D, Ayala G, et al. Aggressive Variants of Castration-Resistant Prostate Cancer. Clinical Cancer Research. 2014 Jun 1;20(11):2846–50.

22. Huang M, Weiss WA. Neuroblastoma and MYCN. Cold Spring Harbor Perspectives in Medicine. 2013 Oct 1;3(10):a014415–a014415.
23. Soucek L, Evan GI. The ups and downs of Myc biology. Current Opinion in Genetics & Development. 2010 Feb;20(1):91–5.
24. Vervoorts J, Lüscher-Firzlaff J, Lüscher B. The Ins and Outs of MYC Regulation by Posttranslational Mechanisms. Journal of Biological Chemistry. 2006 Nov;281(46):34725–9.
25. Beltran H. The N-myc Oncogene: Maximizing its Targets, Regulation, and Therapeutic Potential. Molecular Cancer Research. 2014 Jun 1;12(6):815–22.
26. Wang C, Zhang J, Yin J, Gan Y, Xu S, Gu Y, et al. Alternative approaches to target Myc for cancer treatment. Sig Transduct Target Ther. 2021 Mar 10;6(1):117.
27. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, et al. The Human Transcription Factors. Cell. 2018 Feb;172(4):650–65.
28. Soucek L, Whitfield J, Martins CP, Finch AJ, Murphy DJ, Sodir NM, et al. Modelling Myc inhibition as a cancer therapy. Nature. 2008 Oct;455(7213):679–83.
29. Brockmann M, Poon E, Berry T, Carstensen A, Deubzer HE, Rycak L, et al. Small Molecule Inhibitors of Aurora-A Induce Proteasomal Degradation of N-Myc in Childhood Neuroblastoma. Cancer Cell. 2013 Jul;24(1):75–89.
30. Gustafson WC, Meyerowitz JG, Nekritz EA, Chen J, Benes C, Charron E, et al. Drugging MYCN through an Allosteric Transition in Aurora Kinase A. Cancer Cell. 2014 Sep;26(3):414–27.
31. Richards MW, Burgess SG, Poon E, Carstensen A, Eilers M, Chesler L, et al. Structural basis of N-Myc binding by Aurora-A and its destabilization by kinase inhibitors. Proc Natl Acad Sci USA. 2016 Nov 29;113(48):13726–31.
32. Posternak V, Cole MD. Strategically targeting MYC in cancer. F1000Res. 2016 Mar 29;5:408.
33. Otto T, Horn S, Brockmann M, Eilers U, Schüttrumpf L, Popov N, et al. Stabilization of N-Myc Is a Critical Function of Aurora A in Human Neuroblastoma. Cancer Cell. 2009 Jan;15(1):67–78.
34. Helander S, Montecchio M, Pilstål R, Su Y, Kuruvilla J, Elvén M, et al. Pre-Anchoring of Pin1 to Unphosphorylated c-Myc in a Fuzzy Complex Regulates c-Myc Activity. Structure. 2015 Dec;23(12):2267–79.
35. Jaenicke LA, Von Eyss B, Carstensen A, Wolf E, Xu W, Greifenberg AK, et al. Ubiquitin-Dependent Turnover of MYC Antagonizes MYC/PAF1C Complex Accumulation to Drive Transcriptional Elongation. Molecular Cell. 2016 Jan;61(1):54–67.
36. Fladvad M, Zhou K, Moshref A, Pursglove S, Säfsten P, Sunnerhagen M. N and C-terminal Sub-regions in the c-Myc Transactivation Region and their Joint Role in Creating Versatility in Folding and Binding. Journal of Molecular Biology. 2005 Feb;346(1):175–89.
37. Pineda-Lucena A, Ho CSW, Mao DYL, Sheng Y, Laister RC, Muhandiram R, et al. A Structure-based Model of the c-Myc/Bin1 Protein Interaction Shows Alternative Splicing of Bin1 and c-Myc Phosphorylation are Key Binding Determinants. Journal of Molecular Biology. 2005 Aug;351(1):182–94.
38. Farrell AS, Sears RC. MYC Degradation. Cold Spring Harbor Perspectives in Medicine. 2014 Mar 1;4(3):a014365–a014365.
39. Welcker M, Orian A, Jin J, Grim JA, Harper JW, Eisenman RN, et al. The Fbw7 tumor suppressor regulates glycogen synthase kinase 3 phosphorylation-dependent c-Myc protein degradation. Proc Natl Acad Sci USA. 2004 Jun 15;101(24):9085–90.
40. Nair SK, Burley SK. X-Ray Structures of Myc-Max and Mad-Max Recognizing DNA: Molecular Bases of Regulation by Proto-Oncogenic Transcription Factors.
41. Marumoto T, Hirota T, Morisaki T, Kunitoku N, Zhang D, Ichikawa Y, et al. Roles of aurora-A kinase in mitotic entry and G2 checkpoint in mammalian cells: Aurora-A is required for G2-M transition. Genes to Cells. 2002 Nov;7(11):1173–82.
42. Adhikary S, Eilers M. Transcriptional regulation and transformation by Myc proteins. Nat Rev Mol Cell Biol. 2005 Aug;6(8):635–45.
43. Dubois B, Hampel H, Feldman HH, Scheltens P, Aisen P, Andrieu S, et al. Preclinical Alzheimer's disease: Definition, natural history, and diagnostic criteria. Alzheimer's & Dementia. 2016 Mar;12(3):292–323.
44. Chen H, Liu H, Qing G. Targeting oncogenic Myc as a strategy for cancer treatment. Sig Transduct Target Ther. 2018 Feb 23;3(1):5.
45. Green R, Rogers EJ. Transformation of Chemically Competent E. coli. In: Methods in Enzymology [Internet]. Elsevier; 2013 [cited 2023 Jul 6]. p. 329–36. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780124186873000288

46. Studier FW. Protein production by auto-induction in high-density shaking cultures. Protein Expression and Purification. 2005 May;41(1):207–34.
47. Schägger H. Tricine–SDS-PAGE. Nat Protoc. 2006 Jun;1(1):16–22.
48. Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L. The FoldX web server: an online force field. Nucleic Acids Research. 2005 Jul 1;33(Web Server):W382–8.
49. Obiol-Pardo C, Rubio-Martinez J. Comparative Evaluation of MMPBSA and XSCORE To Compute Binding Free Energy in XIAP−Peptide Complexes. J Chem Inf Model. 2007 Jan 1;47(1):134–42.
50. Kumari R, Kumar R, Open Source Drug Discovery Consortium, Lynn A. *g_mmpbsa* —A GROMACS Tool for High-Throughput MM-PBSA Calculations. J Chem Inf Model. 2014 Jul 28;54(7):1951–62.
51. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, et al. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. Acc Chem Res. 2000 Dec 1;33(12):889–97.
52. Huang J, MacKerell AD. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. J Comput Chem. 2013 Sep 30;34(25):2135–45.
53. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. Journal of Molecular Graphics. 1996 Feb;14(1):33–8.
54. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. J Comput Chem. 2005 Dec;26(16):1781–802.
55. Bertolin G, Sizaire F, Herbomel G, Reboutier D, Prigent C, Tramier M. A FRET biosensor reveals spatiotemporal activation and functions of aurora kinase A in living cells. Nat Commun. 2016 Sep 14;7(1):12674.
56. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl Acids Res. 1994;22(22):4673–80.
57. Trimpin S, Brizzard B. Analysis of insoluble proteins. BioTechniques. 2009 Apr;46(5):321–6.
58. Ransom EM, Ellermeier CD, Weiss DS. Use of mCherry Red Fluorescent Protein for Studies of Protein Localization and Gene Expression in Clostridium difficile. Kivisaar M, editor. Appl Environ Microbiol. 2015 Mar;81(5):1652–60.
59. Bhaduri S, Demchick PH. Simple and rapid method for disruption of bacteria for protein studies. Appl Environ Microbiol. 1983 Oct;46(4):941–3.
60. Sunderhaus A, Imran R, Enoh E, Adedeji A, Obafemi T, Abdel Aziz MH. Comparative expression of soluble, active human kinases in specialized bacterial strains. Barnwal RP, editor. PLoS ONE. 2022 Apr 19;17(4):e0267226.
61. Kimple ME, Brill AL, Pasker RL. Overview of Affinity Tags for Protein Purification. CP Protein Science [Internet]. 2013 Aug [cited 2023 Jul 6];73(1). Available from: https://onlinelibrary.wiley.com/doi/10.1002/0471140864.ps0909s73
62. Lee C. Coimmunoprecipitation Assay. In: Rosato E, editor. Circadian Rhythms [Internet]. Totowa, NJ: Humana Press; 2007 [cited 2023 Jul 6]. p. 401–6. (Walker JM, editor. Methods in Molecular Biology; vol. 362). Available from: http://link.springer.com/10.1007/978-1-59745-257-1_31
63. Furukawa H, Simorowski N, Michalski K. Effective production of oligomeric membrane proteins by EarlyBac-insect cell system. In: Methods in Enzymology [Internet]. Elsevier; 2021 [cited 2023 Jul 6]. p. 3–19. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0076687920303852
64. Rhodes G. Other Diffraction Methods. In: Crystallography Made Crystal Clear [Internet]. Elsevier; 2006 [cited 2023 Jul 19]. p. 211–35. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780125870733500118
65. Kuzmanic A, Zagrovic B. Determination of Ensemble-Average Pairwise Root Mean-Square Deviation from Experimental B-Factors. Biophysical Journal. 2010 Mar;98(5):861–71.

8 CURRICULUM VITAE