5-2021

The Biochemical Characterization of AZA197 and a Ras Related Protein Cdc42

Alix Montoya-Beltran
University of Arkansas, Fayetteville

The Biochemical Characterization of AZA197 and a Ras Related Protein Cdc42

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Cell and Molecular Biology

by

Alix Montoya-Beltran
University of Arkansas
Bachelor of Science in Chemistry, 2016

May 2021
University of Arkansas

This thesis is approved for recommendation to the Graduate Council.

_____
Paul D. Adams Ph.D.
Chair Committee

_____
Matt McIntosh Ph.D.
Committee member

_____
Ines Pinto Ph.D.
Committee member

_____
Suresh Kumar Thallapuranam Ph.D.
Committee member

Abstract

Eukaryotic cells contain an extensive amount of GTP/GDP binding proteins. Proteins known as Ras GTPase primary function as a binary switch, where they cycle from an on and off state when GTP or GDP are bound, respectively. They are known to play a critical role in many cellular functions where a dysregulation could potentially lead to oncogenic behavior or other malignancies. In our laboratory, our focus is the study of a Ras related protein Cell division control 42 homolog (Cdc42) which belongs to the Rho subfamily. Cdc42 plays a critical role in many biological signaling processes; therefore, its uncontrol gene expression has been known to play a pivotal role in cancer cell progression and metastasis. A single point mutation was done on Cdc42 where a Threonine(T) was replaced to an Alanine(A) at position 35 (Cdc42(T35A)), this demonstrated a decrease in the flexibility in a region that is important for the interactions of downstream effector proteins.[1] AZA197 has been demonstrated to show selectivity to Cdc42 and inhibit colon rectal tumor growth.[2] The purpose of this research was to successfully synthesize the small molecule, AZA197, and study its interaction with Cdc42. Fluorescence spectroscopy was used in both characterizations of Cdc42(WT) and Cdc42(T35A) interactions with AZA197. The data suggest that in the presence of AZA197 Cdc42(WT) is more resistant to protein denaturation and has an increase in the Gibbs free energy of unfolding in the absence of GnHCl. Results also highlighted that AZA197 also interacts with Cdc42(T35A) and may have a destabilizing effect on this Cdc42 variant.

Acknowledgements

I would like to take this opportunity to express my gratitude to my thesis advisor, Dr. Paul D. Adams in the Department of Chemistry and the Program of Cellular and Molecular Biology at the University of Arkansas. During my time studying working under him, I have had his unconditional support, guidance, and mentorship. From the moment I first walked into his office to inquire about joining his lab for my master's program I felt like he believed in me. Over the past couple of years, he has truly made me believe in myself and my capabilities, helping me feel that I belong in this field.

Furthermore, I would like to thank my committee members/advisors: Dr. Matt McIntosh; Dr. Ines Pinto; and last but not least, Dr. T. K. Suresh Kumar. For being part of my committee, by taking the time to read my thesis and attending my defense.

I also want to thank my mother, Reina Beltran Almaraz, for never giving up and always encouraging me throughout my years of study. For always making sure my brother and I had a better opportunity in life by immigrating to this country that we now call home, that has given us the opportunity to accomplish the American Dream.

Also, a special thanks to my support group; friends and family. For my dear friend Colette Robinson who understands the struggles as a graduate student can face but also who has helped me and supported me not only academically but personally and unconditionally. To Ruby & Jesus Monter and Yeni Beltran for always motivating and supporting me emotionally.

Table of Contents

Abbreviations

BME	β-mercaptoethanol
CD	Circular dichroism
Cdc42	Cell division cycle 42
Cdc42(T35A)	Cell division cycle 42 in which Threonine at residue 35 has been mutated to Alanine
CV	Colum Volume
DNAse	Deoxyriboribonuclease I, from bovine pancreas
EDTA	Ethylenediaminetetraacetic acid
FPLC	Fast Protein Liquid Chromatography
GAP	GTPase Activating Protein
GDI	Guanine dissociation inhibitor
GDP	Guanosine diphosphate
GEF	Guanine Exchange Factors
GMPPNP	β,γ-Methyleneguanosine 5′-triphosphate sodium salt; a nonhydrolyzable GTP analog
GTP	Guanosine triphosphate
GTPase	GTP hydrolyzing protein
GnHCl	Guanidinum hydrochloride
His_6	Hexahistidine tag; affinity tag
IPTG	Isopropyl-β-d-thiogalactopyranoside
ITC	Isothermal titration calorimetry

LB	Lysogeny Broth
NMR	Nuclear magnetic resonance
PAK	p-21 activated kinase
PBD46	46 amino acid derivative of PAK-21 binding domain
PBS	Phosphate buffered saline
PDB	Protein Data Bank
P_i	Inorganic phosphate
PMSF	Phenylmethanesulfonuylfloride
RBD	Ras binding domain
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TCA	Trichloroacetic acid
WT	Wild type

Chapter I-Introduction

Ras proteins

Ras or Rat Sarcoma genes were first identified and characterized as transduced oncogenes.[3] The Ras protein family is composed of small molecular weight GTPase proteins that contribute to a series of diverse biological roles upon interaction with specific regulatory and effector proteins. GTPase proteins primarily function as molecular switches, as they cycle between Guanosine Triphosphate (GTP) and Guanosine Diphosphate (GDP), known as the on and off states, respectively. The Human Ras superfamily contains 153 protein members, that contribute to various signal transduction pathways, making them a perfect candidate to study protein-protein interactions.[4, 5] They also have the potential to be utilized as targets for therapeutic treatments such as cancer or other diseases.

The Ras superfamily are a GTPase protein characterized by a Gbox domain that is found on the N-terminal side that is responsible for the GTP binding and hydrolysis biochemical activity.[3, 4, 6] The "G box" is separated into five different segments, G1-G5, involving the GTP/GDP binding region, and an active site for the common mechanism of GTPase proteins. The G1 box is known as the P-loop, a purine nucleotide binding region.[3, 6]The G2 box has a Threonine residue at position 35, that is highly conserved among the Ras superfamily.[3] The G3 box sequence has an aspartic acid residue, responsible for participating in the binding of a nucleotide-associated Mg^{2+} ion.[3, 6] The G4 box plays a role in the hydrogen bonding of the guanine ring, providing the stabilizing interactions within the G1 box.[3] The G5 box is the least conserved among the superfamily and it's believed to be involved in indirect associations with the guanine nucleotide.[3]

Rho Proteins

Rho proteins are a subfamily within Ras proteins that contain a highly conserved region within the G box.[7] Nonetheless, some members of this group are uniquely differentiated by containing a specific sequence among all Ras superfamily proteins.[3, 8] They are low molecular weight proteins that are mainly found in the cytoplasm.[3, 8] Rho proteins are divided into six different families based on sequence similarities and are further characterized into typical and atypical GTPases depending upon its mode of regulation.[5]

The Rho GTPases family contributes to cellular processes such as: organization of actin and microtubules in the cytoskeleton, regulation of gene expression, vesicle trafficking, cell cycle progression, cell morphogenesis, cell polarity, and cell migration.[5, 9, 10] In addition, Rho proteins are involved in inflammation process, wound repairs, and cancer progression.[9]

Specifically, the human Rho family is composed of 23 proteins that are divided into 8 categories depending on their type of regulation, typical or atypical.[11] Typical regulation refers to the GTPase proteins dependency on Guanine nucleotide exchange factors (GEFs) and GTPase activating proteins (GAPs).[9] Atypical regulation of GTPase protein do not depend on regulatory proteins but are believed to use post-translational modifications for its regulation and expression.[9]

Typical regulation of Rho proteins rely on three main effector proteins: GAPs, GEFs, and Guanine nucleotide dissociation inhibitors (GDIs).[5, 8, 9] As mentioned above, GTPase proteins cycle between an active and inactive state, GTP and GDP, respectively. GEFs are proteins that facilitate the exchange of GDP (inactive) to GTP (active). GAPs are responsible for facilitating the hydrolysis of GTP (active to inactive) by providing an essential catalytic group.[8, 9, 12] The other regulator protein, GDIs, are responsible for the activation regulation and control the

location of some Rho GTPases within the cell, either in the cytoplasm or on the plasma membrane.[8, 9, 12] Once a GTPase protein is bound to GTP they have a greater affinity for binding to downstream effector proteins.

Rho GTPases play a significant role in the regulation of cell migration, proliferation, cell survival, and apoptosis.[9] All of this is necessary for the maintenance of normal tissues, but a slight dysregulation could lead to the progression of cancer. The process of oncogenesis, the initiation of cancer cells from normal cells, is a multi-step process from environmental and genetic factors. This results in an imbalance of certain cellular processes such as, cell proliferation or the inhibition of apoptosis leading to the uncontrolled replication of cells i.e., immortalization. Cell division control 42 (Cdc42), Ras homolog gene family A (RhoA), and Ras related C3 botulinum substrate 1(Rac1) are among some of the Rho subfamily proteins that have been identified to play a role in the progression and metastasis of different human tumors.[9]

Cdc42 (WT and T35A)

Cdc42 is a member of the typical Rho GTPase family known to play a significant role in cancer progression and is tissue specific. For example, in non-small lung cancer, colorectal, adenocarcinoma, melanoma, lung, and breast cancer it has been upregulated, and for kidney, central nervous system, and adrenal gland cancers it is downregulated.

Cdc42 is a small GTPase protein that has a molecular weight of approximately 21kDa and its secondary sequence is composed of 6 β sheets and 5 α helices.[13, 14]

This protein contains 11 homologs ranging in molecular size from 190 to 193 amino acids.[10] It contains a particular domain unique among other Ras proteins, known as the Rho-insert domain (RI) composed of 13 amino acids which serve as a recognition region for downstream effectors.[14]

In humans, there are two major isoforms of Cdc42 and they are composed of two main regions known as the Switch I (residues 28-41) and Switch II (residues 57-74).[1, 7] The Switch I and Switch II regions are highly conserved regions among Cdc42 isoforms.[1, 7] These two regions have demonstrated internal dynamics that allow them to interact with several regulatory and effector proteins efficiently. Cdc42 can interact with effector/adaptor proteins such as kinases, phospholipases, and actin associated proteins. Cdc42 undergoes post-translational modification at the C-terminal side on the Cys residue pivotal for its attachment to the plasma membrane.[14] Downstream effector proteins contain a Cdc42/Rac interactive binding domain (CRIB domain) important for the binding of downstream effector proteins with Cdc42 The specificity of the effector proteins interaction with Cdc42 is believed to take place with different residues within the Switch I region.[15] This leads to a competition between effector and regulatory proteins, i.e., different interactions at the different times in the cell cycle.

A single point mutation on Cdc42 wild type (WT) was performed, replacing a Threonine with an Alanine at position 35 known as Cdc42(T35A) to study Cdc42 backbone flexibility and the overall stability of the protein.[7] This mutation was done inside the Switch I region that has been recognized to play a role in interacting with downstream effector proteins, upon Cdc42 being active.[1] This single point mutation demonstrated no destruction in the overall tertiary structure of the protein and had no effect on its activity.[7] Cdc42(T35A) is believed to reduce the overall flexibility of the Switch I region.[1] This point mutation is significant because Threonine at position 35 is highly conserved among all Ras family proteins.[1, 7] The point mutation not only demonstrated a reduce flexibility with in the Switch I region but a weak interaction was observed from a protein known as PBD46.[1, 16] This protein is a partial segment of p21-activated serine/threonine kinase (PAK1) that interacts with active Cdc42 it was demonstrated to reduce its

interaction in comparison with the Cdc42(WT).[1, 16, 17] In an in vivo study performed in mammals, Cdc42 has been proven to play a significant role in the regulation of the bones and the cardiovascular, immune, and nervous systems.[18] Cdc42 has the potential to be studied as a model for structure-function relationship of cell signaling processes such as actin dynamics, migration, survival, and cell cycle progression.[8, 19, 20] Additionally, its important role in the progression of cancer and other diseases has led current research to focus on small molecules and microRNAs as new candidates for treatment.[21-23]

Cdc42 in Cancer

Cancer develops when an accumulation of genetically and epigenetic changes occur on tumor suppressor genes and oncogenes.[10, 11, 24] Targeted therapeutic treatment depends on the specific molecular mechanism responsible for the progression of the disease. Cancer cells are unique in their ability to alter the intermediate metabolism. They are able to alter the glycolysis pathway by the phosphorylation of the PKM2 enzyme that leads to the formation of lactic acid instead of acetyl CoA creating an acidic environment.[10] The acidic environment has also been related to an increase in tumor progression.[10] Cdc42 is usually not mutated in cancer cells but rather overexpressed or hyperactivated, and in human cancers they are mainly activated due to an upregulation of GEF's.[11, 12, 25, 26]

Cdc42 in its active state has been known to interact with several downstream proteins such as PAK, PI3K-AKT, and Raf Merk-Erk pathways.[27] Therefore, recent efforts have been made to find inhibitory candidates between specific GEF's and Cdc42.[27] ZCL278 is a small molecule demonstrated to inhibit the protein intersecting (ITSN), a specific Cdc42 GEF, leading to a reduction in cell mobility and the disruption of the Golgi organization.[28] Another successful derivative from the ZCL compounds is ZCL367 that resulted in a halt of cancer progression in

Human A56 lung cancer cells and its high selectivity of Cdc42 in comparison to other Rho proteins.[27]

Cdc42 can be used as a molecular marker for the identification of early-stage cancers such as for the early detection of pancreatic ductal adrenal carcinoma, PDAC.[26] Pancreatic cancer is known to have no preliminary signs and symptoms, and unfortunately this leads to its diagnosis at a later stage with poor prognosis.[26] In a study it was found that there is a direct relationship between pancreatic carcinogenesis and Cdc42 overexpression.[26]

Another recent therapeutic discovery was the use of a small non-coding microRNA (miRNA), known to assist in post-transcriptional regulations that play a role in the physiological and pathological processes.[22, 24, 29] miRNA can serve as therapeutic agents or biomarkers for certain cancers. One example is the MicroRNA-548J which has been identified to be overexpressed in breast cancer cell lines, known as invasion-inducing miRNA.[22] MicroRNA-548J was demonstrated to activate Tensin1 that is related to the activation of Cdc42.[22] The selective inhibitor known as ML141 was demonstrated to suppress microRNA-548J mediated cell invasion.[22] Furthermore, it was demonstrated that microRNA-548J and Tensin1 have a strong correlation in regards to breast cancer cell invasion and metastasis.[22]

AZA197

A small organic molecule known as AZA197 was found to suppress tumor growth of colon rectal cancer (CRC), and has demonstrated a selectivity of Cdc42 over other Rho proteins.[2] This highlighted the potential of Cdc42 as a candidate for targeted therapies to treat malignancies related to its overexpression.[2]

Cdc42 has been known to play a key role in the progression of cancer even though it is tissue specific. One downstream signal that Cdc42 tends to interact with is PAK1.[2] The

overexpression of PAK1 has been associated to be an aggressive sign of the CRC disease progression.[2] Additionality, Cdc42 has also been proven to be overexpressed suggesting a potential role in the tumor development.[2] In a study using cell lines, SW620 (highly invasive colorectal cancer cells) and HT-29 (colorectal cancer cells), it was demonstrated that after being treated with AZA197 there was a reduction in cell migration and invasion.[2] It was demonstrated in the in vivo portion of the study that SW620 colon cancer xenografts had reduced tumor growth after being treated with 10 μM AZA197 for twenty days as compared to the control xenografts that had not been treated with the small molecule.[2]

The synthesis of AZA197 was based on a chemical compound from the National Cancer Institute Chemical Database that was used to target against the Rho GTPase protein, Rac A.[2] Using a free docking database known as ZINC, 17 small organic molecules were synthesized by a company called SPECS located in the Netherlands.[2] All seventeen molecules were carefully evaluated and only AZA197 exhibited strong inhibitory activity in SW620 cells.[2] The study demonstrated that based on a cytotoxicity assay the optimum concentration of AZA197 is 10 μM. Another assay performed was to identify if the GTPases; Rho, Rac1, and Cdc42 would interact with the small molecule in both cell lines.[2] This demonstrated that AZA197 specifically and significantly downregulates Cdc42 activity in a dosage dependent manner, while there was no significant effect on Rac1 or RhoA.[2] Next it was determined whether or not AZA197 hinders the interaction between Cdc42 and the Dbs protein, a particular Cdc42 GEF protein.[2] The experiment concluded that AZA197 blocks the interaction of Dbs and Cdc42, halting the activation of Cdc42 resulting in its inhibition.[2] An in vivo experiment was also performed treating mice with SW620 xenografts in the presence and absence of the small molecule for

twenty days.[2] The results reveled that in the mice that were treated with the small molecule the tumor weight was suppressed and resulted in prolonged animal survival.[2]

Chapter II-Experimental Methods

Growth, Expression, and Harvest

Cdc42(WT) and Cdc42(T35A) constructs are overexpressed using a (N-His_6) tag from pET15b in Escherichia Coli strain, BL21(DE3), in LB media. In a 500 mL Erlenmeyer flask containing 100 mL of LB media, 100 uL of ampicillin (AMP), and a 1 mL bacterial glycerol stock, containing 500 uL of 50% glycerol and 500 uL of the overnight culture, was used to make an overnight culture growth that was left shaking for 15 hours at a constant speed and temperature of 37°C. Upon completion, the overnight culture is used to inoculate a 6 L Erlenmeyer flask containing 1.5 L of LB media and 750 uL of AMP utilizing the same conditions. For the next couple of hours, the culture was monitored until the optimum optical density (OD) was reached, 0.4-0.6. A Varian Cary 50 bioUV-vis Spectroscopy was used to measure the OD of the culture at a wavelength of 600 nm. Once the optimal OD was reached, the culture was induced with 0.5 mM Isopropyl-beta-D-thioglactopyranoside (IPTG) and was left to express for four hours in the shaker under the same conditions. After expression, the cells are harvest by centrifugation at 6,500 rpm for 30 minutes at 4ºC. The cell pellet is collected, and the supernatant is discarded. The cell pellet is resuspended with 1x PBS Buffer solution (pH 7.4) and centrifuged again at 4,500 rpm for 12 minutes to washout left over LB media. The supernatant is again discarded, and the cell pellet is stored at –80ºC.

Protein Purification

A 5 mL His-Trap FF column was cleaned and regenerated by the following protocol: 5 cv (1 column volume (cv) = 5 mL) of 20% ethanol, 10 cv of deionized water, 3 cv of 1x stripping

buffer, 5 cv of deionized water, 3 cv of 1x binding buffer, 2 cv of 1x charging buffer, and 4 cv of deionized water.

A cell pellet was left to thaw on ice and resuspended in lysis buffer. The lysis buffer contains: 45 mL of Tris/NaCl binding buffer (40 mM imidazole, 100 mM NaCl, 50 mM Tris (pH 8.0), and 10 mM $MgCl_2$), 45 mg of lysozyme, 90 uL of phenylmethylsulfonyl fluoride (PMSF), 160 uL of Halt protease, and a pinch of Bovine DNAse. The cell pellet on ice undergoes a sonication cycle, where the sample was sonicated every other ten seconds for five minutes. The cell debris and soluble components are separated by centrifugation at 19,000 rpm for 35 minutes. The supernatant is collected and filtered and left on ice to prevent proteins degradation.

The proteins are purified by affinity chromatography using a Fast Protein Liquid Chromatography (FPLC) machine. Using the ÄKTA Start system (FPLC, GE Healthcare) with a 5 mL His-Trap FF column, lysis buffer, binding buffer, and elution buffer were all added into their respective place on the FPLC for the protein purification process The column is first equilibrated with binding buffer, and upon equilibration the lysis buffer will be passed through the column. The lysis buffer interacts with the column and the protein of interest will bind to the column while everything else passes through the column. The column is then washed with binding buffer to remove unwanted proteins. An increasing linear gradient of elution buffer starts to run through the column until 100% of elution buffer has been reached. Allowing for the bound protein to come off the column and into the collection tubes. Once this takes place the column is re-equilibrated. After the cycle is completed, the column is removed, and everything is cleaned according to manufacture protocol. Figure 2.1 represents an example of a chromatogram from Cdc42(T35A) after the purification process is complete. The binding phase is illustrated by the first peak between 40-65 mL. In which the lysis buffer passes through the column, where all

unwanted proteins will elute and the protein of interested will remain in the column. The washing phase is between 65-115 mL, where binding buffer is passed through the column to make sure any unwanted protein is eluted out of the column. The Elution phase is represented from 115-160 mL, where at first a linear gradient of increasing concentration of elution buffer will be added to the column to remove the protein of interest. To make sure the protein of interest is unbound a 100% elution buffer solution will be added to the column. In this case fractions 9-11 where collected. A Chromatograph of the purification of Cdc42(WT)is shown in figure 2.2. The binding phase of the purification process was observed between 0-30 mL. The washing out phase is demonstrated in the 30-80 mL to ensure the unwanted protein was completely washed out. The elution phase is demonstrated in the 80-120 mL, where Cdc42(WT) was eluted out, in this case fractions 10-16 where collected.

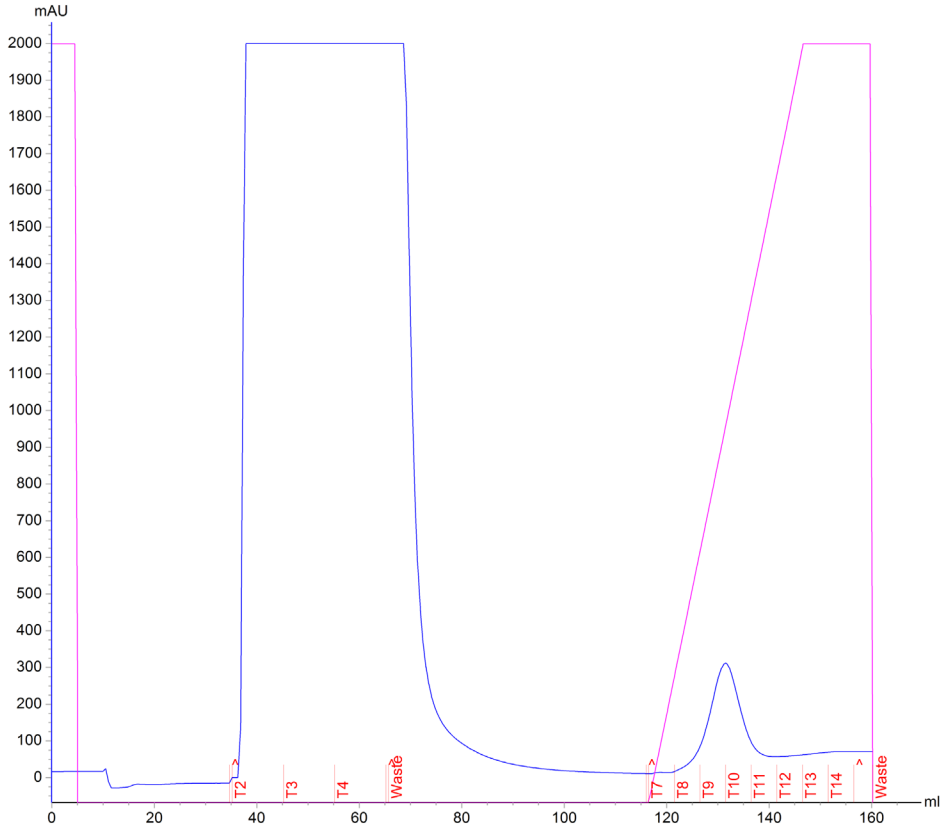

Figure 2.1. Depicts a chromatogram from the purification of Cdc42(T35A), in this case fractions 9-11 were collected. The y-axis is a measure of the absorbance units (mAU), and the x-axis

represents the volume (mL) used. The blue line depicts the absorbance units, and the pink line represents the gradual increase of elution buffer used in order to unbind the protein of interest from the column.

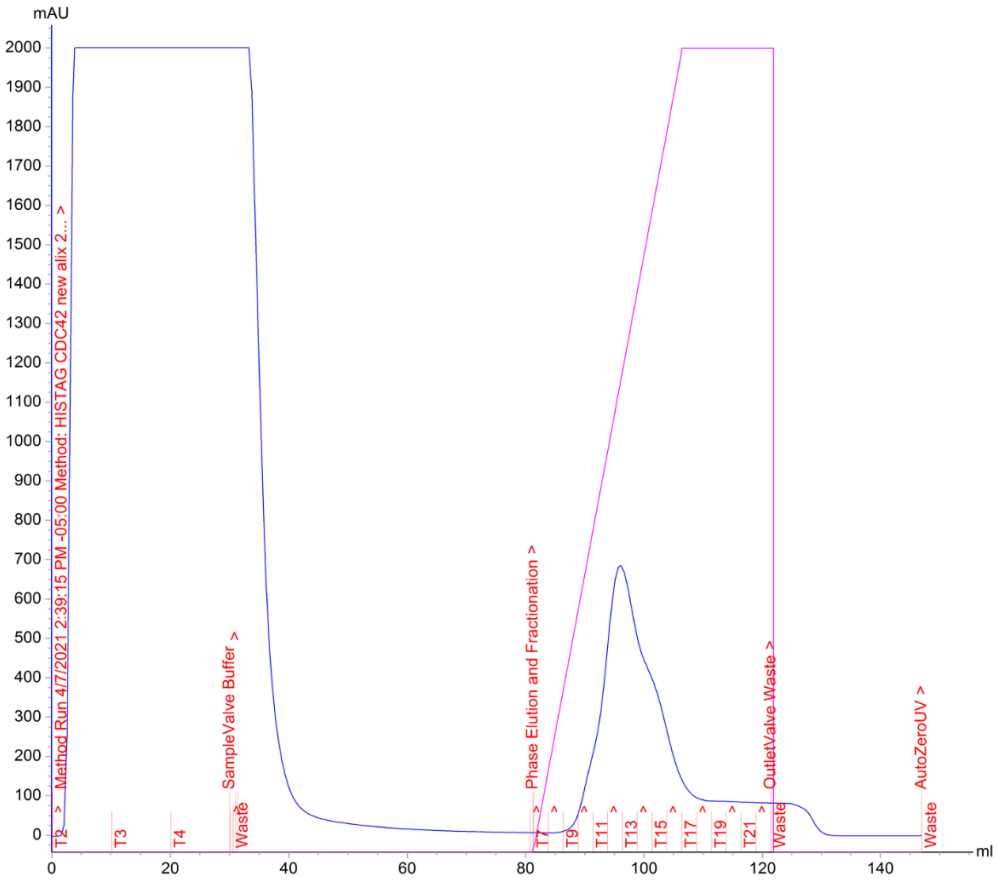

Figure 2.2. Affinity chromatography showing the purification of Cdc42(WT). The blue line depicts the absorbance units (mAU), and the pink line represents the gradual increase of elution buffer used in order to unbind the protein of interest from the column.

Once the purification process was completed the collection tubes were collected and placed in ice. A 20 uL sample from each collection tube was collected and placed into a 1.5 mL microcentrifuge tube with 10 uL of 2X dye to make gel samples to run an SDS-Page. Once the gel samples were obtained the collection fraction tubes from the purification were combined and added into a dialysis bag with a molecular cut off weight of 4-6 kDa, which was then placed in dialysis buffer (100 mM NaCl, 20 mM Tris, and 10 mM $MgCl_2$ (pH of 8.0)) overnight at 4°C.

Gel Electrophoresis

The presence and condition of the protein was confirmed by SDS-Page gel electrophoresis. The gel was composed of a 15% bis acrylamide mixture. The 1.5 mL microcentrifuge tubes collected from the purification were used, and the gel was run according to protocol. Once completed it was washed with water, stained for 10 minutes, and destained overnight, then washed with distilled water and scanned. Figure 2.3 reveals a gel obtained from two different purifications. The first band represents the molecular weight ladder, where is a set of weights standards to identify approximate weights of proteins. The Lys or Lysis buffer is the sample obtained before the purification. The flow or flow through represent the washing phase of the FPLC, where all of the proteins not bound to the column are eluted out (similar to the peak between 65-115 mL in figure 2.1). The 10-14 columns are the actual sample gels collected form the elution phase.

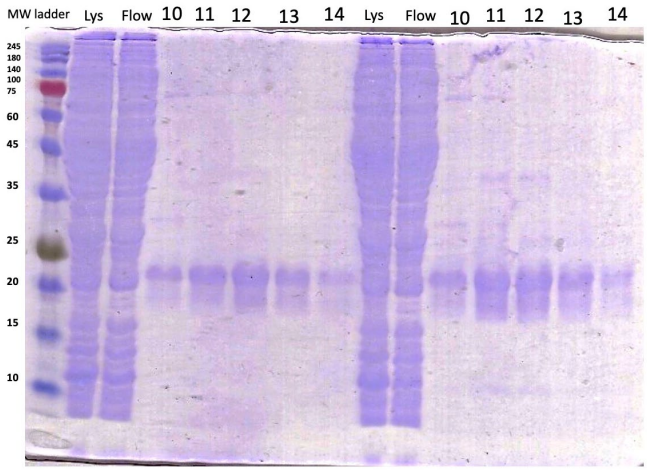

Figure 2.3. SDS-PAGE electrophoresis confirming the presence of a Cdc42(WT) based on the molecular weight. Two purifications were performed independently from each other. The Lys represents the sample obtained before the purification. The flow is the sample obtained during the binding phase of the purification, and the numbers correspond to the fractions collected during the elution phase of the purification, in this particular case fractions 10-14 were collected.

Nucleotide Exchange

A nucleotide exchange assay was performed to exchange the GTP to a nonhydrolyzable form of GTP. The protein was transferred from the dialysis bag into a 15 mL Falcon tube, and the concentration was checked with the Bradford protein assay and later corrected using the correct molar absorptivity and molecular weight depending on the protein of interest. Once the concentration of the protein was obtained, it was then incubated with the GTP analog, β,γ-Methyleneguanosine 5′-triphosphate sodium salt. The GTP analog serves to conserve the protein in its active state. In addition to the GTP analog, 0.5 mM of EDTA is added and left for an incubation period of 3.5 hours, at a temperature of 4^{o}C, on a gentle shaker. A desalting column, PD-10, was equilibrated with 50 mL of 1X PBS buffer (pH 7.4) and the protein was passed through the column and collected in 1 mL microcentrifuge tubes. Each collection tube was analyzed using a Bradford Assay to determine its location. Protein Fractions were obtained and 0.1 M $MgCl_2$ is added to the sample and the concentration was checked and stored at -80^{o}C.

Chapter III-Experimental Results

Synthesis of AZA197

The synthesis of AZA197 was done in collaboration with the McIntosh Group. The synthesis involved a two-step reaction. Figure 3.1 illustrate the first step of the reaction synthesis, where Tryptamine (2 molar equivalence) and triethylamine (1 molar equivalence) in acetonitrile were mixed in a round bottom flask. Once in solution dichloro-6-methylpyrimidine (1 molar equivalence) was added. The reaction ran at 70^{o}C for 16 hours. The reaction was concentrated and both isomeric products were isolated using a TLC Flash Chromatograph AKROS machine. Products were eluted using 60:40 hexane: ethyl acetate, the compounds after

being concentrated were a yellow-brown solid. Table 3.1 demonstrates the optimizations used in order to obtain the first reaction.

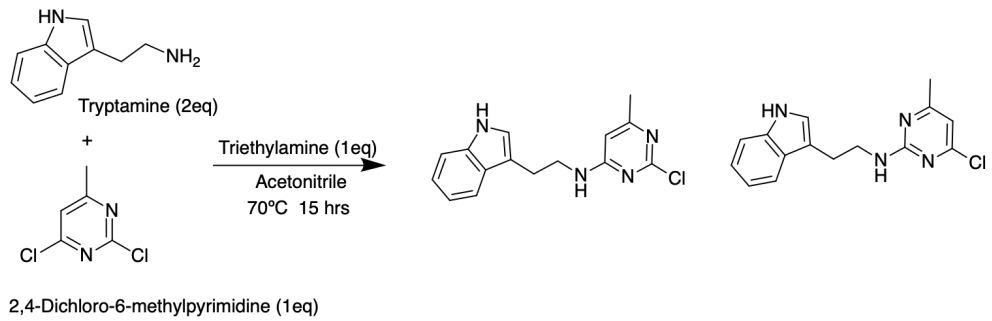

Figure 3.1. Illustrates the first step reaction synthesis.

Table 3.1 Shows the reactions done to optimize the first reaction.

Tryptamine (equivalence)	2-4-Dichloro-6-methylpyrimidine (equivalence)	Temperature	Time	Solvent	Successful
1	2	30ºC	17hrs	THF	No
1	2	70ºC	17hrs	THF	No
1	2	70ºC	15hrs	Triethylamine and acetonitrile	Yes

Figures 3.2-3.5 illustrate the proton NMRs obtained for the reaction. Figure 3.2 is a proton NMR performed after adding Tryptamine and 2,4-Dichloro-6methylpyrimidine in order to obtain a preliminary NMR of the reactants. Figure 3.3 is a proton NMR obtained after the reaction was completed and concentrated before the separation of the two isomers. Figures 3.4 and 3.5 are the proton NMRs after both constitutional isomers where successfully separated and concentrated. Figure 3.4 illustrates the isomer of interest, a successfully synthesized AZA197, ^{1}H NMR ($CDCl_3$): 2.3(s,3H), 3.08-3.12(t,2H), 3.668(s,2H), 6.01(s,1H), 7.071(s,1H),7.13-7.16(m,1H), 7.2-7.22(m,1H),7.43-7.405(d,1H),7.64-7.62(d,1H). Figure 3.5 depicts the unwanted isomer ^{1}H NMR ($CDCl_3$): 2.3(s,3H), 3.06-3.08(t,2H), 3.7-3.8(t,2H), 6.45(s,1H), 7.071(s,1H),7.13-7.16(m,1H), 7.2-7.22(m,1H),7.37-7.39(d,1H),7.68-7.7(d,1H).

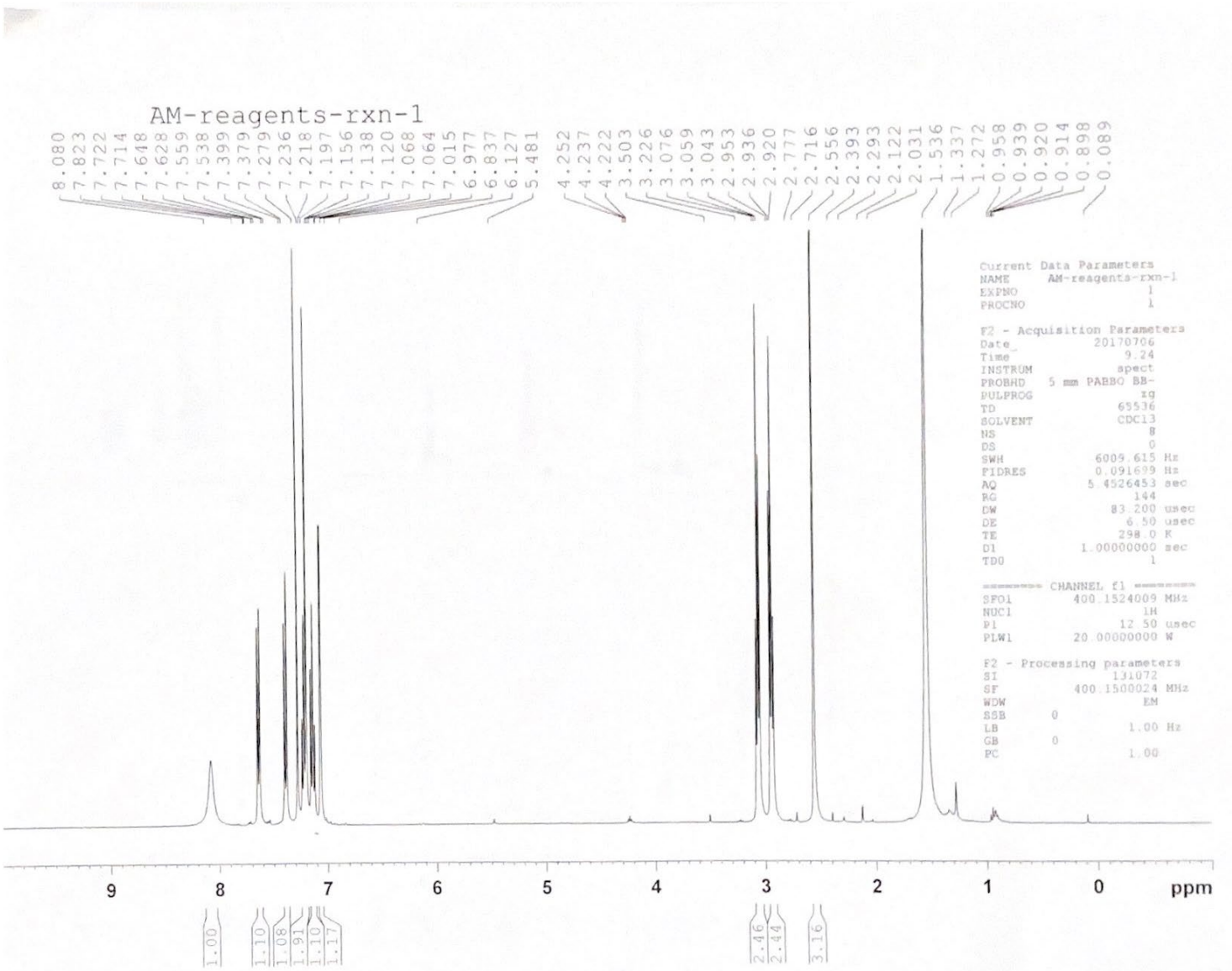

Figure 3.2. Illustrates a Proton NMR of the two reagents added together. Tryptamine and 2-4-Dichloro-6-methylpyrimidine.

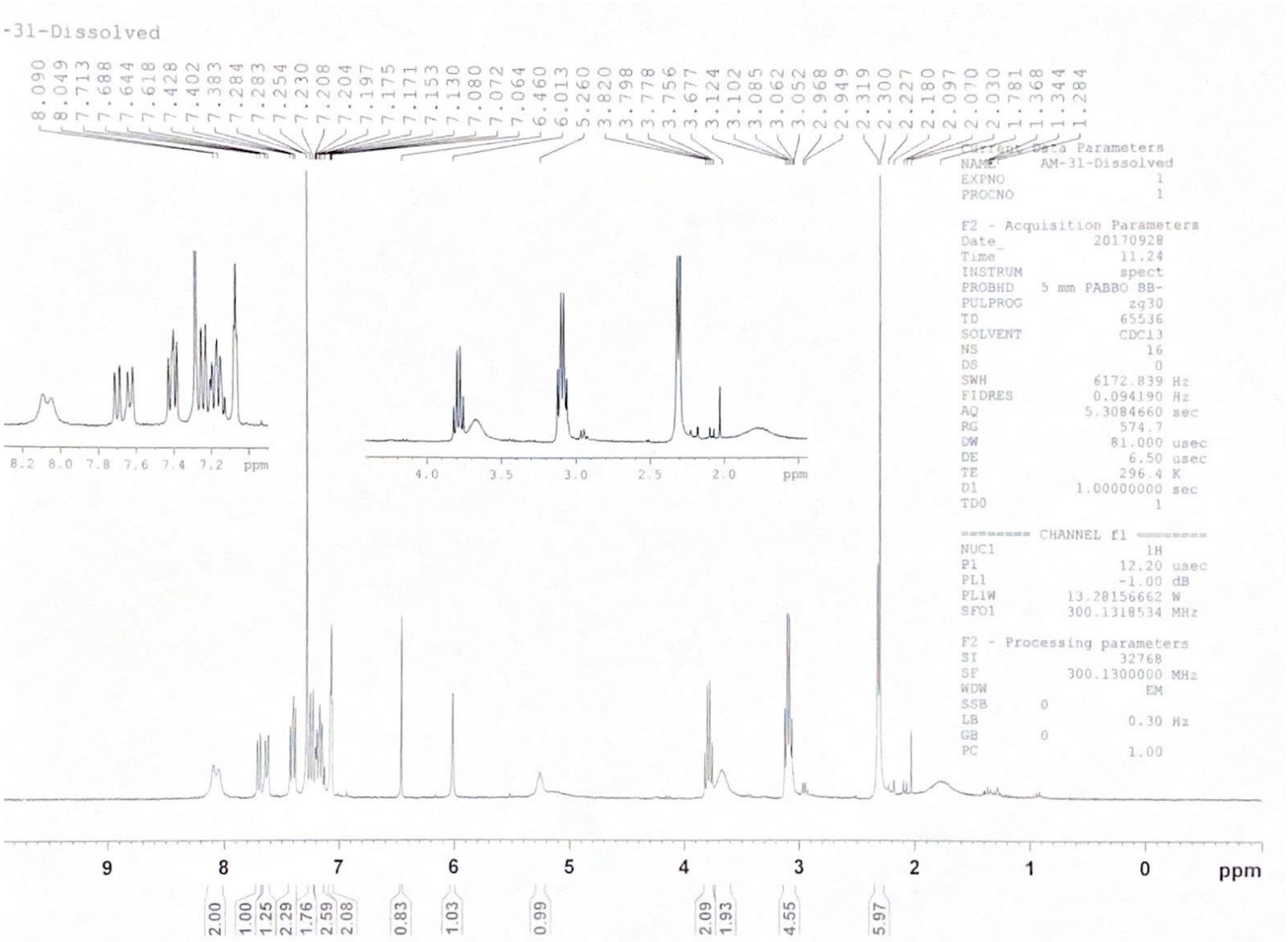

Figure3.3. Illustrates the proton NMR after the reaction was completed and concentrated. Both Isomers are still not separated.

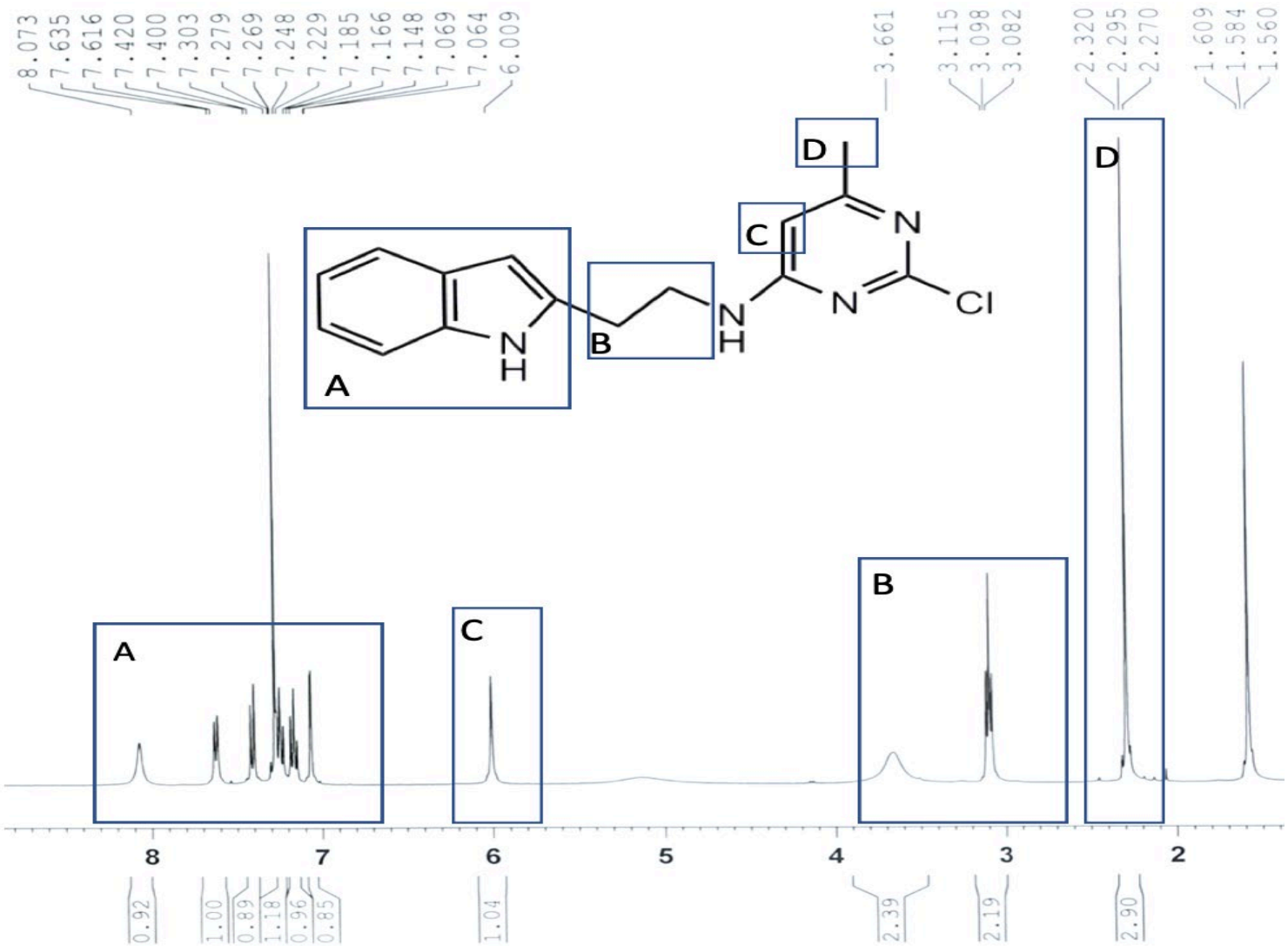

Figure 3.4. Illustrates the isomer of interest after its separation and concentration. The main difference between the constitutional isomers are within the peaks on box C and B.

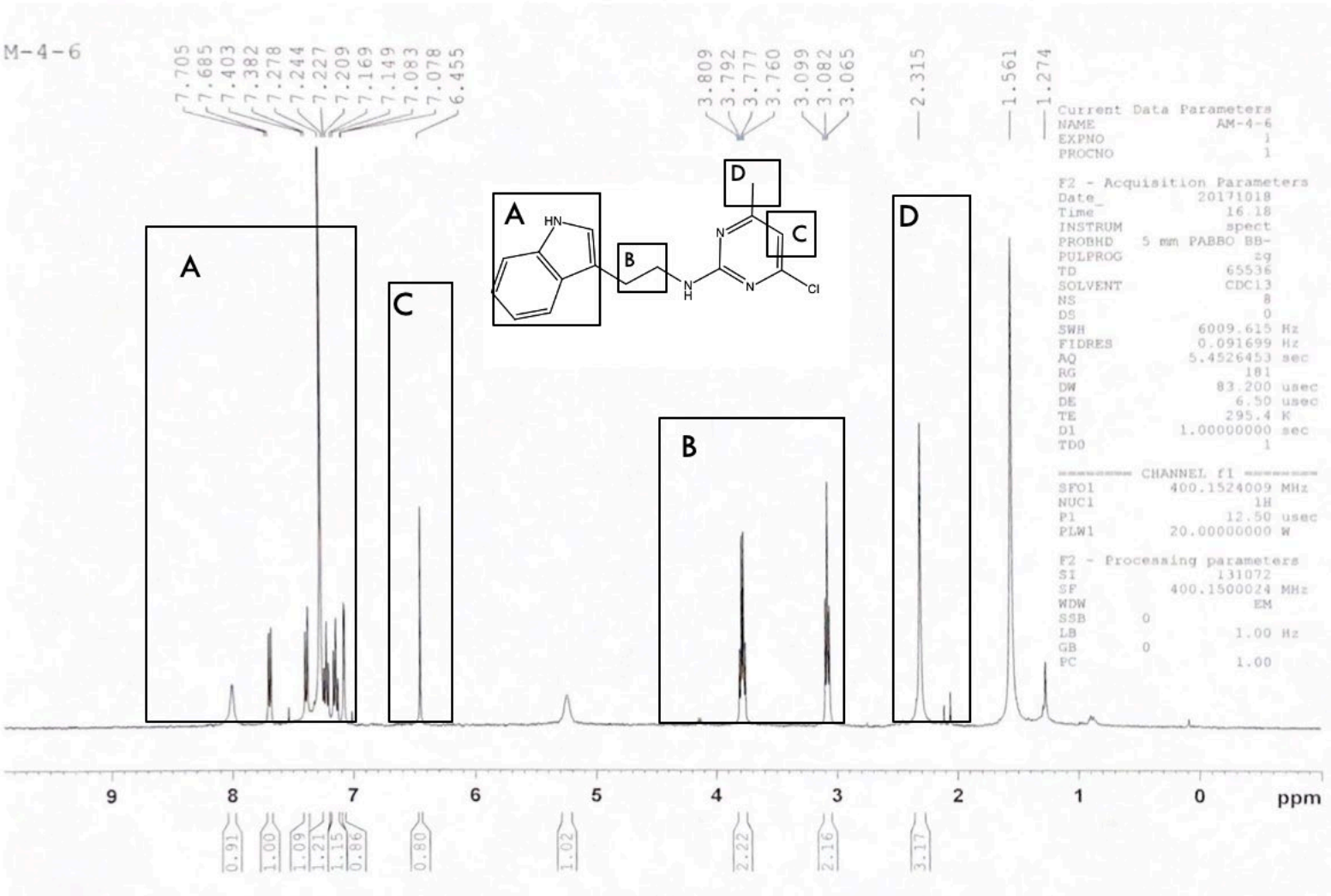

Figure 3.5. Illustrates the unwanted isomer after its separation and concentration. The main difference between the constitutional isomers are within the peaks on box C and B.

The Mass spectra was done after successfully isolating both isomers to confirm that both had the corresponding molecular weight of the compounds, which is 286.76 g/mol, and the composition of $C_{15}H_{15}N_4Cl$. Figures 3.6 and 3.7 illustrate the mass spectra obtained from the wanted isomer and the unwanted isomer, respectively.

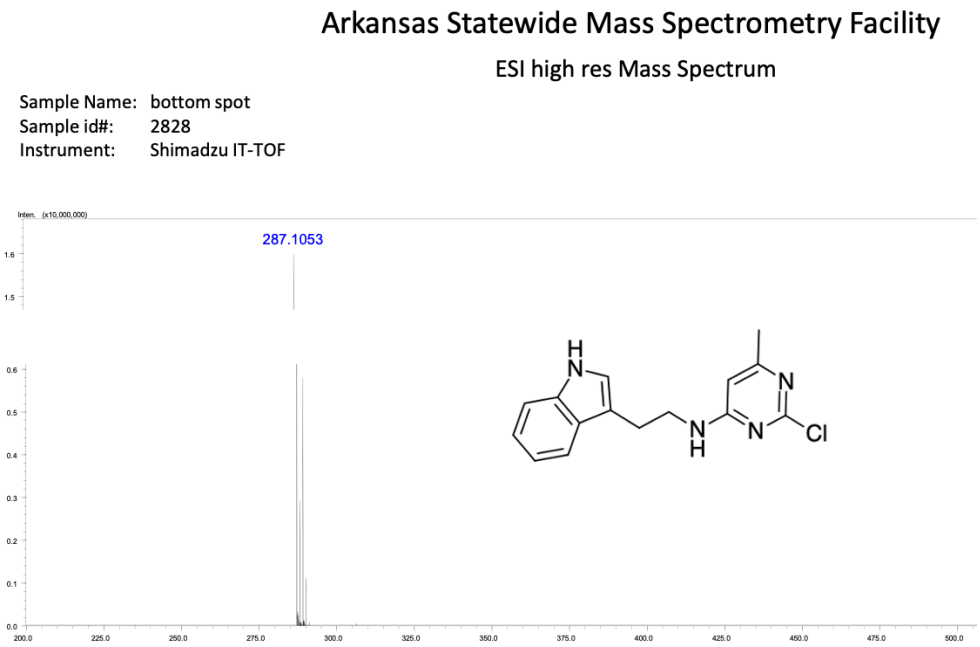

Figure 3.6. Illustrates the ESI Mass spectrum obtained for the isomer of interest.

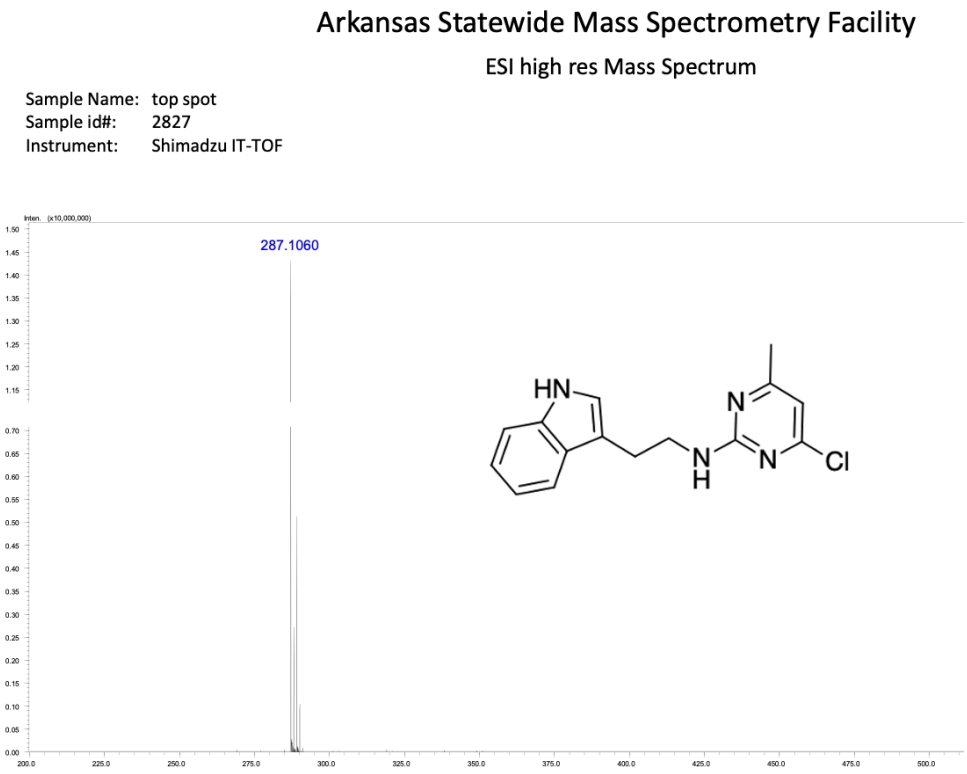

Figure 3.7. Illustrates the ESI Mass Spectrum obtained from the isomer that was not needed.

For the second step of the reaction synthesis Isomer B and 4-amino-1-diethylaminopentane were reacted in a 1:1 molar ratio without any solvent as depicted in Figure 3.8.

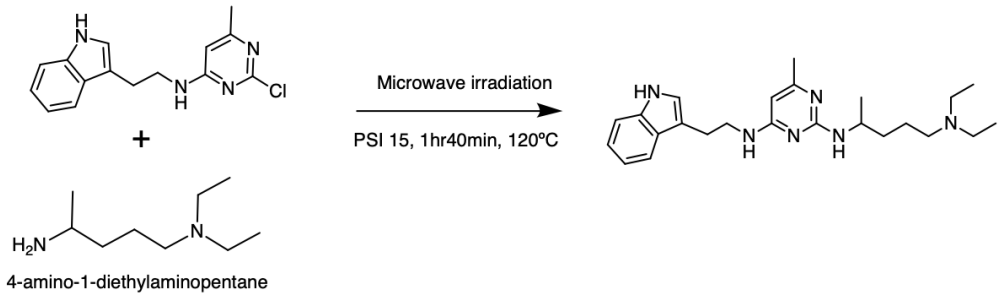

Figure 3.8. A visual representation of the second step of the reaction synthesis This was performed after isomer B has been successfully isolated.

The reaction was performed using a CEM Discover microwave machine under the following settings: pressure 15 psi, temperature 120ºC, and reacted for one hour and 45 minutes. The product was a brown solid with a 62% yield. Table 3.2 represents the optimization conditions for the second reaction before the final reaction conditions was established.

Table 3.2. Optimization conditions for the second reaction.

Isomer B (equivalence)	4-amino -1- diethylaminopentane (equivalence)	Temperature	Time	Pressure (PSI)		Solvent	Successful
1	1.1	130ºC	17hrs	N/A	Oil bath	No	No
1	1.2	60ºC	18hrs	N/A	Oil bath	Triethylamine and ethanol	No
1	1.5	70ºC	18hrs	N/A	Oil bath	No	No
1	1.5	78ºC	20min	30	Microwave irradiation	No	No
1	1.5	78ºC	20min	20	Microwave irradiation	No	No
1	1.5	120ºC	30min	20	Microwave irradiation	No	No
1	1.5	130ºC	30min	25	Microwave irradiation	No	No
1	1.5	130ºC	90min	25	Microwave irradiation	No	No
1	1	120ºC	1hr40min	15	Microwave irradiation	No	Yes

Fluorescence experiments

Using the intrinsic Tryptophan (Trp) in Cdc42(WT) and Cdc42(T35A), fluorescence spectroscopy was performed. This was done via a protein denaturation experiment by performing an increasing concentration gradient of Guanidine Hydrochloride (GnHCl) in the presence and absence of 10 μM AZA197. The data was collected using a Hitachi f2500 fluorometer with an excitation wavelength of 280 nm and an emission collection range of 300-400 nm.

Using the data collected the fraction of denaturation (F_d), equilibrium constant (K_{eq}), denaturation concentration (Cm), and Gibbs free energy (ΔG) were determined. F_d is obtain by equation 1.

$$F_d = \frac{F - F_n}{F_D - F_n} \qquad \text{(Eq.1)}$$

F corresponds to the relative fluorescence reading at any given point. F_n is the fluorescence reading of the native protein before any denaturant is added. F_D, corresponds to the fluorescence reading of the completely denatured protein. The equilibrium constant is obtained by equation 2.

$$K_{eq} = \frac{F_d}{1 - F_d} \qquad \text{(Eq.2)}$$

The use of the equilibrium constant can be used to solve for the Gibbs free energy of protein unfolding, equation 3.

$$\Delta G_d = -RTLn(K_{eq}) \qquad \text{(Eq.3)}$$

The ideal gas constant (R) used was 8.314 J mol^{-1} K^{-1}, and 298 K was the temperature (T) used for the determination of the Gibbs free energy. The use of equation 3 allows us to use equation (4) to find the Gibbs free energy of the protein unfolding in the absence of no protein denaturant ($\Delta G_{d,H_2O}$).

$$\Delta G_d = \Delta G_{d,H_2O} - m[GnHCl] \qquad \text{(Eq.4)}$$

To solve for ΔG_{dH2O} (y-intercept), a derivative of the slope-intercept equation is used. The slope is represented by m[GnHCl] and changes with respect to increasing [GnHCl]. All denaturation experiments were performed in triplicates in the presence and absence of 10 μM AZA197 and 100 mL of 1 mg/mL of protein.

Table 3.1 and Table 3.2 demonstrate the values obtained from the fluorescence experiments in the presence and absence of the small molecule AZA197 with increasing concentrations of GnHCl. The fluorescence intensity was normalized by using Eq. 1 to find the F_d for each point. Since each sample was triplicated the F_d represented is the average of the samples and the standard error was obtained in order to used them for error values. The F_d was then used to calculate the equilibrium constant (K_{eq}) followed by the determination of Gibbs free energy (ΔG_d). Figure 3.9 and Figure 3.10 are a graphical representation of the fraction of denaturation from Tables 3.3 and Table 3.4, respectively with increasing concentration of GnHCl in the presence and absence of AZA197.

Table 3.3 Demonstrates the data used to graph the fractions of denaturation with increasing concentration of GnHCl of Cdc42(WT) in the absence and presence of 10 μM AZA197, respectively.

[GnHCl]	Cdc42(WT) Fd	Standard Error	Cdc42(WT) + AZA197 Fd	Standard Error
0	0	0	0	0
1	0.3255	0.05977	0.33572	0.04047
2	0.75437	0.17948	0.556	0.11859
2.5	0.65448	0.09153	0.61069	0.11734
3	0.66419	0.05121	0.59782	0.0934
3.5	0.69207	0.03227	0.62767	0.0496
4.5	0.7862	0.05696	0.77871	0.08101
5	0.84542	0.0241	0.93273	0.03363
5.4	1	0	1	0

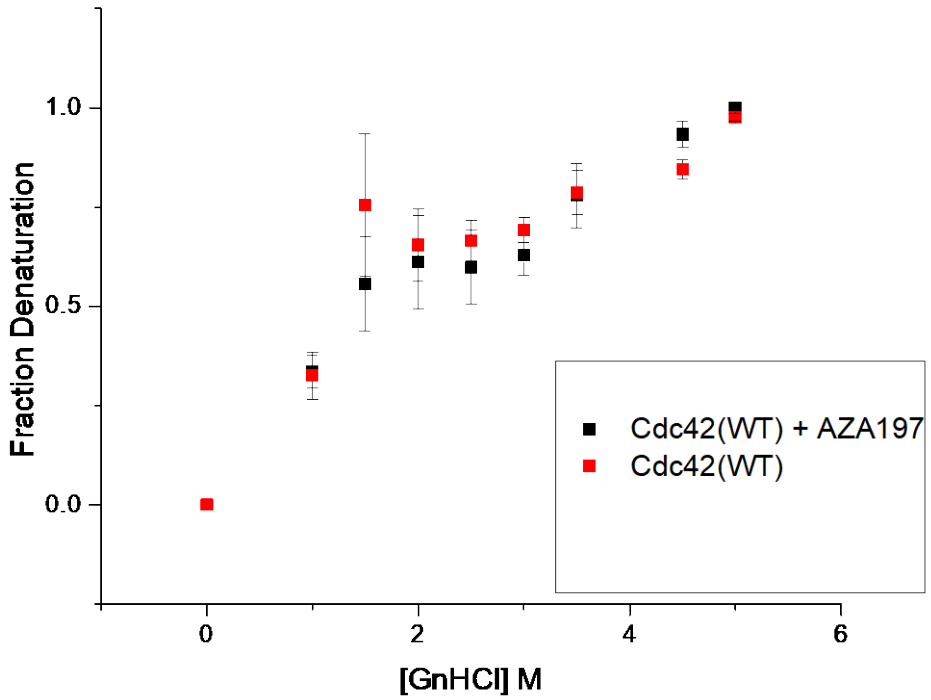

Figure 3.9 The fraction of denaturation in the absence (red) and presence (black) of 10 μM AZA197 with increasing concentrations of denaturant (GnHCl) for Cdc42(WT). The numerical values used can be found on Table 3.1. The data was analyzed using Origin software.

Table 3.4. Demonstrates the data used to graph the fraction of denaturation with increasing concentration of GnHCl of Cdc42(T35A) in the absence and presence of 10 μM AZA197, respectively.

[GnHCl]	Cdc(T35A) Fd	Standard Error	Cdc42(T35A)+ AZA197 Fd	Standard Error
0	0	0	0	0
1	0.71974	0.01656	0.64858	0.03433
1.5	0.80246	0.00345	0.73959	0.02023
2	0.8325	0.01494	0.74068	0.05578
2.5	0.90647	0.01454	0.78292	0.03707
3	0.91419	0.01798	0.81606	0.0279
3.5	0.94599	0.0187	0.83875	0.04071
4.5	0.97234	0.01332	0.92051	0.02639
5	1	0	0.97471	0.00943

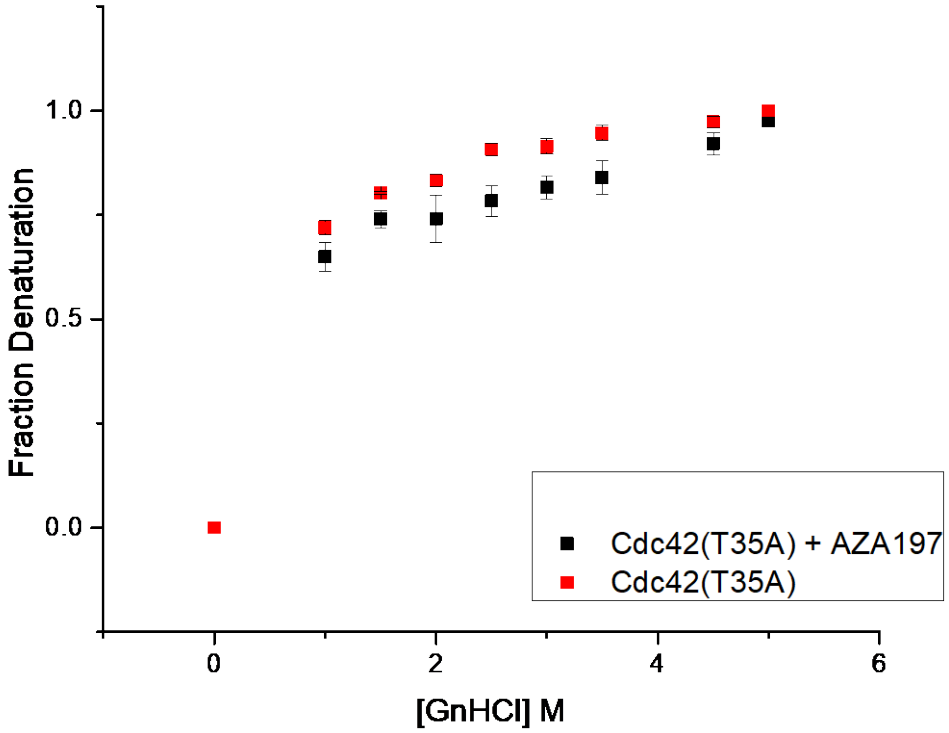

Figure 3.10. The fraction of denaturation in the absence (red) and presence (black) of 10 µM AZA197 with increasing concentrations of denaturant (GnHCl) for Cdc42(T35A). The numerical values used can be found on Table 3.2. The data was analyzed using Origin software.

The Cm is the concentration of the denaturant were the folded and unfolded proteins are in equilibrium. The Fd of 0.5 was used to determine the C_m values for each data set. As seen in Table 3.5 the C_m values of Cdc42(WT) and Cdc42(T35A) implies that there is no significant change in the presence and absence of AZA197. This suggest that AZA197 does not have an effect on the overall stability of the proteins.

Table 3.5 C_m values in the presence and absence of AZA197 for Cdc42(WT) and Cdc42(T35A).

	C_m
Cdc42(WT)	1.49
Cdc42(WT) + AZA197	1.37
Cdc42(T35A)	0.69
Cdc42(T35A) + AZA197	0.75

Tables 3.6 and Table 3.7 Exhibit the Gibbs free energy in respects to increasing concentration of GnHCl in the presence and absence of 10 µM AZA197 for Cdc42(WT) and

Cdc42(T35A), respectively. Figure 3.11 and figure 3.12 are a graphical representation of the Gibs free energy of the Tables 3.6 and 3.7, respectively.

Table 3.6. The Gibbs free energy of protein unfolding in the absence of the denaturant for Cdc42(WT)

[GnHCl]	ΔGd Cdc42(WT)	ΔGd Cdc42(WT) + AZA197
1	1805.1416	1692.172
1.5	-335.7254	-1116.369
2.5	-1582.294	-982.8852
3	-1689.776	-1294.915
3.5	-2006.36	-3119.765
4.5	-3226.281	-6520.107

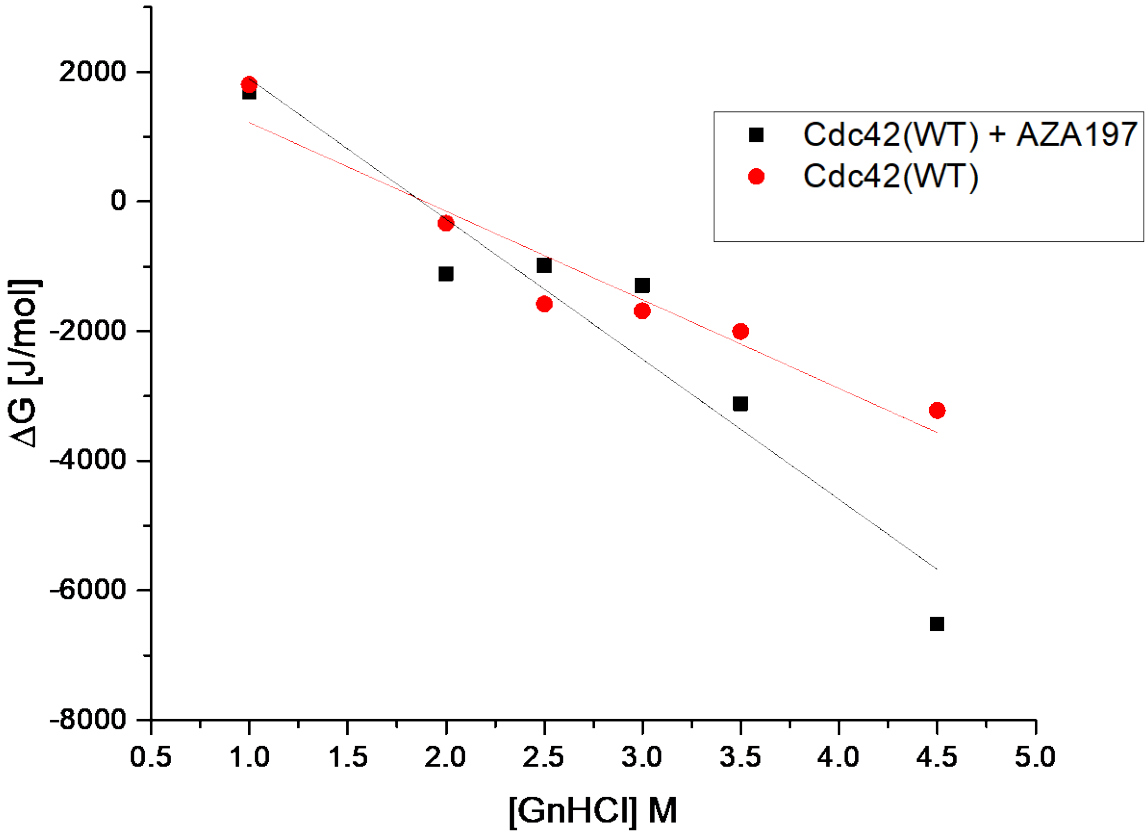

Figure 3.11. The denatured Gibbs free energy (ΔG_d) of Cdc42(WT)in the presence and absence of AZA197 in respect to increasing concentrations of GnHCl. The data points were derived from the F_d. The data was analyzed using Origin software.

Table 3.7. The Gibbs free energy of the protein unfolding in the absence of the denaturant. For Cdc42(T35A).

[GnHCl]	ΔGd Cdc(T35A)	ΔGd Cdc(T35A)+ AZA197
1	-969.7889	-1518.399
2	-2336.356	-2600.4174
2.5	-3973.171	-3178.3404
3	-5626.093	-3691.4551
3.5	-5860.508	-4085.7334
4.5	-7092.292	-6068.7436

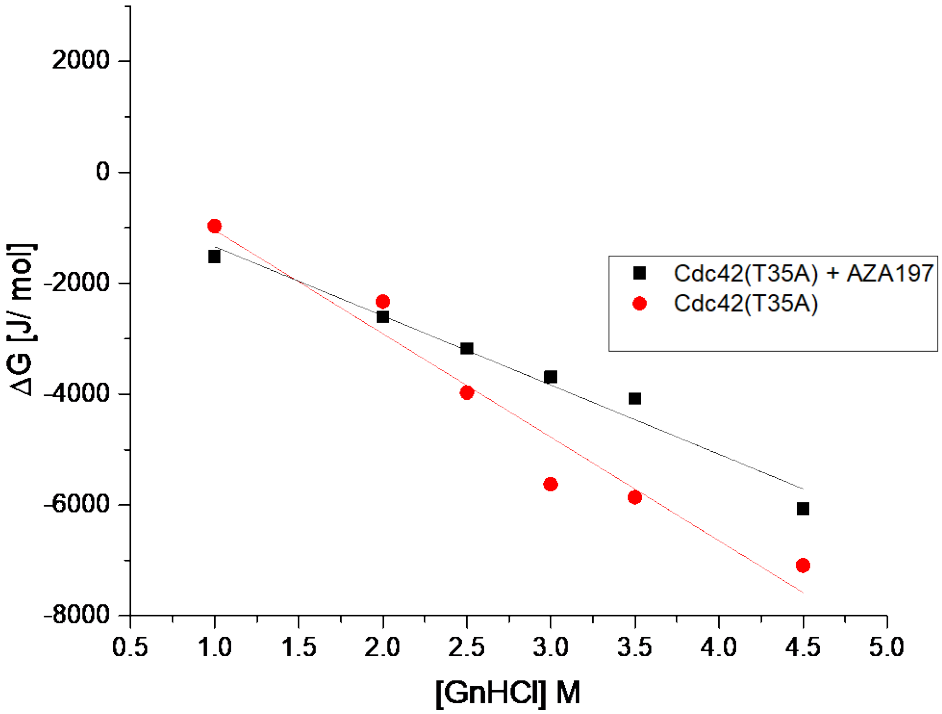

Figure 3.12. The denatured Gibbs free energy (ΔG_d) of Cdc42(T35A)in the presence and absence of AZA197 in respect to increasing concentrations of GnHCl. The data points were derived from the F_d. The data was analyzed using Origin software.

For Cdc42 (WT) K_{eq} and ΔG_d have a significant difference suggesting that in the presence of the small molecule seems to have an influence. As its signified in Table 3.8, in the presence of the small molecule there is an increase in $\Delta G_{d,H2O}$ and the slope, meaning that in the presence of AZA197 it unfolds at a slower rate. For Cdc42(T35A), the small molecule seems to have an opposite influence on unfolding relative to the wild type. There is a significant change in

the F_d, K_{eq}, and ΔG_d. Suggesting that AZA197 destabilizes Cdc42(T35A), not interacting in the same manner as compared to the wild type.

Table 3.8. The Gibbs free energy of unfolding in the absence of no protein denaturant and its corresponding slope determined by Eq.4

	$DG_{d,H2O}$ (kJ/mol)	**m[GnHCl] (kJ/L)**
Cdc42(WT)	2.587	-1.367
Cdc42(WT)+AZA197	4.057	-2.162
Cdc42(T35A)	0.825	-1.867
Cdc42(T35A) +AZA197	-0.09134	-1.248

Chapter IV-Conclusion

The biophysical characterizations of the small molecule, AZA197, and Cdc42 are important for its potential use as an anticancer drug therapy. It has been demonstrated that Cdc42 is tissue specific, and it can be either be upregulated or downregulated in cancer. In those malignancies that have been associated with Cdc42 upregulation, small molecules have been used to inhibit the interaction with downstream effector proteins, reducing its overactivity.

The discovery of the small molecule AZA197 demonstrates its potential use as an anticancer drug therapy as it has high selectivity to Cdc42 among other Rho related proteins. Nonetheless, to understand Cdc42 and AZA197 interactions at a molecular level, more biophysical characterizations need to be performed. AZA197 was synthesized and its interactions with Cdc42(WT) and Cdc42(T35A) were studied with increasing concentrations of GnHCl.

This biophysical characterization demonstrated that Cdc42(WT) is more stable in the presence of the small molecule under increasing concentrations of the denaturant. The Gibbs free energy of the protein unfolding was determined via fluorescence experiments, and an increase was demonstrated in the presence of AZA197.

This suggested that AZA197 helps stabilize Cdc42(WT). In contrast, the fluorescence denaturation experiments with Cdc42(T35A) in the presence and absence of the small molecule

has a negative effect making it more destabilized in comparison to Cdc42(WT). This data suggested that Cdc42(T35A) has no positive interaction with AZA197. Cdc42(T35A) is in the Switch I region that is highly conserved in Cdc42(WT), where its flexibility allows for its interactions with downstream effector proteins once the protein is in the active state. The data suggest that AZA197 binds to the Switch I region which could lead to a potential inhibition to its overall activity.

However, to better understand the molecular interactions of Cdc42(WT) and Cdc42(T35A) with AZA197 more experiments should be performed. Future experiments that could give a better understanding of its interactions would be a time-dependent fluorescence denaturation experiment, Isothermal Calorimetry (ITC), NMR, and GTP hydrolysis. These would allow for the understanding of the interactions between Cdc42 and AZA197 at a molecular level to determine AZA197's potential as an anticancer drug therapy.

References

1. Chandrashekar, R.; Salem, O.; Krizova, H.; McFeeters, R.; Adams, P. D., A switch I mutant of Cdc42 exhibits less conformational freedom. *Biochemistry* **2011,** *50* (28), 6196-207.

2. Zins, K.; Gunawardhana, S.; Lucas, T.; Abraham, D.; Aharinejad, S., Targeting Cdc42 with the small molecule drug AZA197 suppresses primary colon cancer growth and prolongs survival in a preclinical mouse xenograft model by downregulation of PAK1 activity. *J Transl Med* **2013,** *11*, 295.

3. Colicelli, J., Human RAS superfamily proteins and related GTPases. *Sci STKE* **2004,** *2004* (250), RE13.

4. Wennerberg, K.; Rossman, K. L.; Der, C. J., The Ras superfamily at a glance. *J Cell Sci* **2005,** *118* (Pt 5), 843-6.

5. Hodge, R. G.; Schaefer, A.; Howard, S. V.; Der, C. J., RAS and RHO family GTPase mutations in cancer: twin sons of different mothers? *Crit Rev Biochem Mol Biol* **2020,** *55* (4), 386-407.

6. Li, G.; Zhang, X. C., GTP hydrolysis mechanism of Ras-like GTPases. *J Mol Biol* **2004,** *340* (5), 921-32.

7. Adams, P. D.; Oswald, R. E., NMR assignment of Cdc42(T35A), an active Switch I mutant of Cdc42. *Biomol NMR Assign* **2007,** *1* (2), 225-7.

8. Jansen, S.; Gosens, R.; Wieland, T.; Schmidt, M., Paving the Rho in cancer metastasis: Rho GTPases and beyond. *Pharmacol Ther* **2018,** *183*, 1-21.

9. Haga, R. B.; Ridley, A. J., Rho GTPases: Regulation and roles in cancer cell biology. *Small GTPases* **2016,** *7* (4), 207-221.

10. Qadir, M. I.; Parveen, A.; Ali, M., Cdc42: Role in Cancer Management. *Chem Biol Drug Des* **2015,** *86* (4), 432-9.

11. Aguilar, B. J.; Zhou, H.; Lu, Q., Cdc42 Signaling Pathway Inhibition as a Therapeutic Target in Ras- Related Cancers. *Curr Med Chem* **2017,** *24* (32), 3485-3507.

12. Humphries-Bickley, T.; Castillo-Pichardo, L.; Hernandez-O'Farrill, E.; Borrero-Garcia, L. D.; Forestier-Roman, I.; Gerena, Y.; Blanco, M.; Rivera-Robles, M. J.; Rodriguez-Medina, J. R.; Cubano, L. A.; Vlaar, C. P.; Dharmawardhane, S., Characterization of a Dual Rac/Cdc42 Inhibitor MBQ-167 in Metastatic Cancer. *Mol Cancer Ther* **2017,** *16* (5), 805-818.

13. Mott, H. R.; Owen, D.; Nietlispach, D.; Lowe, P. N.; Manser, E.; Lim, L.; Laue, E. D., Structure of the small G protein Cdc42 bound to the GTPase-binding domain of ACK. *Nature* **1999,** *399* (6734), 384-8.

14. Abramovitz, A.; Gutman, M.; Nachliel, E., Structural coupling between the Rho-insert domain of Cdc42 and the geranylgeranyl binding site of RhoGDI. *Biochemistry* **2012,** *51* (2), 715-23.

15. Johnson, D. I., Cdc42: An essential Rho-type GTPase controlling eukaryotic cell polarity. *Microbiol Mol Biol Rev* **1999,** *63* (1), 54-105.

16. Guo, W.; Sutcliffe, M. J.; Cerione, R. A.; Oswald, R. E., Identification of the binding surface on Cdc42Hs for p21-activated kinase. *Biochemistry* **1998,** *37* (40), 14030-7.

17. Chandrashekar, R.; Adams, P. D., Prospective Development of Small Molecule Targets to Oncogenic Ras Proteins. *Open J Biophys* **2013,** *3* (4), 207-211.

18. Melendez, J.; Grogg, M.; Zheng, Y., Signaling role of Cdc42 in regulating mammalian physiology. *J Biol Chem* **2011,** *286* (4), 2375-81.

19. Maldonado, M. D. M.; Dharmawardhane, S., Targeting Rac and Cdc42 GTPases in Cancer. *Cancer Res* **2018,** *78* (12), 3101-3111.

20. Lin, Y.; Zheng, Y., Approaches of targeting Rho GTPases in cancer drug discovery. *Expert Opin Drug Discov* **2015,** *10* (9), 991-1010.

21. Li, Y.; Zhu, X.; Xu, W.; Wang, D.; Yan, J., miR-330 regulates the proliferation of colorectal cancer cells by targeting Cdc42. *Biochem Biophys Res Commun* **2013,** *431* (3), 560-5.

22. Zhan, Y.; Liang, X.; Li, L.; Wang, B.; Ding, F.; Li, Y.; Wang, X.; Zhan, Q.; Liu, Z., MicroRNA-548j functions as a metastasis promoter in human breast cancer by targeting Tensin1. *Mol Oncol* **2016,** *10* (6), 838-49.

23. Humphreys, K. J.; McKinnon, R. A.; Michael, M. Z., miR-18a inhibits CDC42 and plays a tumour suppressor role in colorectal cancer cells. *PLoS One* **2014,** *9* (11), e112288.

24. Xiao, X. H.; Lv, L. C.; Duan, J.; Wu, Y. M.; He, S. J.; Hu, Z. Z.; Xiong, L. X., Regulating Cdc42 and Its Signaling Pathways in Cancer: Small Molecules and MicroRNA as New Treatment Candidates. *Molecules* **2018,** *23* (4).

25. Valdés-Mora, F.; Locke, W. J.; Bandrés, E.; Gallego-Ortega, D.; Cejas, P.; García-Cabezas, M. A.; Colino-Sanguino, Y.; Feliú, J.; Del Pulgar, T. G.; Lacal, J. C., Clinical relevance of the transcriptional signature regulated by CDC42 in colorectal cancer. *Oncotarget* **2017,** *8* (16), 26755-26770.

26. Yang, D.; Zhang, Y.; Cheng, Y.; Hong, L.; Wang, C.; Wei, Z.; Cai, Q.; Yan, R., High Expression of Cell Division Cycle 42 Promotes Pancreatic Cancer Growth and Predicts Poor Outcome of Pancreatic Cancer Patients. *Dig Dis Sci* **2017,** *62* (4), 958-967.

27. Aguilar, B. J.; Zhao, Y.; Zhou, H.; Huo, S.; Chen, Y. H.; Lu, Q., Inhibition of Cdc42-intersectin interaction by small molecule ZCL367 impedes cancer cell cycle progression, proliferation, migration, and tumor growth. *Cancer Biol Ther* **2019,** *20* (6), 740-749.

28. Friesland, A.; Zhao, Y.; Chen, Y. H.; Wang, L.; Zhou, H.; Lu, Q., Small molecule targeting Cdc42-intersectin interaction disrupts Golgi organization and suppresses cell motility. *Proc Natl Acad Sci U S A* **2013,** *110* (4), 1261-6.

29. Zhang, M.; Guo, W.; Qian, J.; Wang, B., Negative regulation of CDC42 expression and cell cycle progression by miR-29a in breast cancer. *Open Med (Wars)* **2016,** *11* (1), 78-82.