Aus dem Institut für Biochemie
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

Structure investigations of plasticity-related gene 2 (PRG2).

zur Erlangung des akademischen Grades
Doctor of Philosophy (PhD)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von
Fatih Mehmet Ipek
aus Istanbul

Datum der Promotion: 29 November 2024

Vorwort

Attention: Data in my monograph are already published in Brosig, A., Fuchs, J., Ipek, F., Kroon, C., Schrötter, S., Vadhvani, M., Polyzou, A., Ledderose, J., van Diepen, M., Holzhütter, H.G. and Trimbuch, T., 2019. The axonal membrane protein PRG2 inhibits PTEN and directs growth to branches. Cell reports, 29(7), pp.2028-2040.

Contents

List of figures

List of tables

Abbreviations

°C	Degrees Celsius
2D	Two dimensional
3D	Three dimensional
A8-35	Amphipol
aa	Amino acid
APS	Ammonium persulfate
BN	Blue Native PAGE
BSA	Bovine Serum Albumin
C-domain	Intracellular domain
CO2	Carbon dioxide
cryo-EM	Cryogenic electron microscopy
Cy5	Cyanine
DAPI	4′,6-diamidino-2-phenylindole
DDM	n-Dodecyl−β−D-Maltopyranoside
dH2O	distilled water
DLS	Dynamic Light Scattering
DNAse	deoxyribonuclease
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
FBS	Fetal Bovine Serum
FPLC	Fast protein liquid chromatography
HEK	Human embryonic kidney
HRP	Horseradish peroxidase
IP	Immunoprecipitation
IPTG	Isopropyl-β-D-thiogalactoside
IR	Infrared
Kd	Dissociation constant
kDa	Kilodalton
LB	Lysogeny broth
LC-MS/MS	Liquid chromatography-mass spectrometry

LPP	Lipid phosphate phosphatase
M	Protein Molecular Weight Marker
MALDI-MS	Matrix Assisted Laser Desorption/Ionization Mass Spectrometry
min	Minute
ml	Milliliter
mM	Millimolar
MST	Microscale Thermophoresis
Nano DSF	Nano differential scanning fluorimetry
NHS	N-Hydroxysuccinimide
N1E-115	mouse neuroblastoma cells
nl	Nanoliter
nm	Nanometer
OD	Optical density
PAGE	polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PEI	Polyethylenimine
PFA	Paraformaldehyde (PFA) (Sigma Aldrich)
PI(3,4,5)P3	Phosphatidylinositol (3,4,5)-trisphosphate
PI(4,5)P2	Phosphatidylinositol (4,5)-bisphosphate
PLA	Proximity ligation assay
PRG2	Plasticity-related gene 2
PRG2ΔC	Plasticity-related gene (residues 1-438)
PRG2ΔN	Plasticity-related gene (residues 285-716)
PTEN	Phosphatase and tensin homolog
PVDF	Polyvinylidene difluoride
rpm	Rotation Per Minute
SDS	sodium dodecyl sulfate
Sec	second
SEC	Size exclusion chromatography
shRNA	Short hairpin RNA
SLB	Laemmli Sample Loading Buffer
SN	Supernatant
TEMED	Tetramethylethylenediamine

TFA	Trifluoroacetic acid
TL	Total cell lysate
TM	Transmembrane
Tm	Melting temperature
V	Volt
WB	Western blot
µg	Microgram

Abstract

Human plasticity-related gene 2 (*PRG2*) encodes a 718 amino acid (aa) long, 76 kDa molecular weight amphipathic integral membrane protein with six hydrophobic transmembrane (TM) spanning helices and 440 aa long hydrophilic intracellular domain (C-domain) with unknown function.

PRG2 protein is a member of the lipid phosphate phosphatase (LPP) family, which is thought to regulate bioactive lipid mediators. To date, five members have been identified (PRG1-PRG5). PRG2 shows a close homology to PRG1 with 49% amino acid identity and 63% similarity. Similar to PRG1, PRG2 contains a long hydrophilic C-terminal cytoplasmic extension; however, PRG2 contains an unusual accumulation of 20 glutamic acid residues ('poly-E-box') within the cytoplasmic C-domain.

It has been shown that PRG2 expression in the rodent brain is associated with a time window spanning from late prenatal to early postnatal stages, which are characterized by complex maturation of neuronal morphology. Silencing PRG2 expression with short hairpin RNA (shRNA) reduces the number of filopodia branches along the axon in primary cortical neurons, whilst overexpression of PRG2 in N1E-115 mouse neuroblastoma cell line and primary cortical neurons induce the number of filopodial membrane protrusions.

It was further shown that PRG2 spatially and temporally inhibits Phosphatase and tensin homolog (PTEN) activity by direct interaction with PTEN, and/ or binding to phosphatidylinositol (3,4,5)-trisphosphate (PI(3,4,5)P3) in the inner leaflet of the plasma membrane, therefore restricting the availability/accessibility of PTEN to its substrate, leading to local nanodomains of PI(3,4,5)P3 accumulations along the axonal the plasma membrane.

In summary, my data demonstrate that the intracellular domain of PRG2 directly interacts with PTEN and phosphoinositides (PIPs). We further identified that the poly-E box is required for PRG2-PTEN and PIPs binding. Also, biochemical data revealed the self-association ability of PRG2 as well as the possibility to form higher-order multimers

with itself or other PRG family members. Another novel finding of this thesis is that PRG2 C-domain alone is sufficient for PRG2 homomultimerization.

Zusammenfassung

Human plasticity-related gene 2 (PRG2) codiert für ein amphipathisches integrales Membranprotein, welches aus 718 Aminosäuren besteht und ein Molekulargewicht von 76 kDa aufweist. Das PRG2-Protein ist aus sechs hydrophoben Transmembranhelizes (TM) und einer hydrophilen intrazellulären C-Domäne aufgebaut, welche aus 440 Aminosäuren besteht und deren Funktion unbekannt ist.

Das PRG2-Protein gehört zur Familie der Lipidphosphat-Phosphatasen (LPP), von welcher angenommen wird, dass sie bioaktive Lipidmediatoren reguliert. Bis heute wurden fünf Mitglieder identifiziert (PRG1–PRG5). PRG1 und PRG2 weisen mit 49 % identischen Aminosäuren und einer Sequenzähnlichkeit von 63 % eine hohe Homologie auf. Ähnlich wie PRG1 verfügt PRG2 über eine lange intrazelluläre Domäne am C-Terminus, wobei in der C-Domäne von PRG2 eine ungewöhnliche Häufung von 20 Glutaminsäureresten auftritt, die als Poly-E-Box bezeichnet wird.

Es wurde gezeigt, dass die Expression von PRG2 im Nagetierhirn im Zeitraum später pränataler bis früher postnataler Entwicklungsstufen stattfindet, welche durch komplexe Reifungsprozesse der neuronalen Morphologie gekennzeichnet sind. Während die Repression der Expression von PRG2 mit shRNA (short hairpin RNA, kurze Haarnadel-RNA) die Anzahl an Filopodienzweigen entlang des Axons in primären kortikalen Neuronen reduziert, wird durch Überexprimierung von PRG2 sowohl in der N1E-115-Maus-Neuroblastom-Zelllinie als auch in primären kortikalen Neuronen die Anzahl an Ausstülpungen in der Filopodienmembran erhöht.

Außerdem konnte gezeigt werden, dass das PRG2-Protein die Aktivität des Enzyms PTEN (Phosphatase and tensin homolog) räumlich und zeitlich durch eine direkte Interaktion mit PTEN und/ oder die Bindung von Phosphatidylinositol(3,4,5)-triphosphat (PI(3,4,5)P3) an der Innenseite der Zellmembran inhibiert und dadurch die Verfügbarkeit bzw. Zugänglichkeit von PTEN für sein Substrat verhindert; in der Folge dieser Inhibierung entstehen lokale Nanodomänen von akkumuliertem PI(3,4,5)P3 entlang der Axonmembran.

Zusammengefasst zeigt diese Arbeit, dass die intrazelluläre Domäne von PRG2 direkt mit PTEN und Phosphoinositiden (PIP) interagiert. Außerdem geht aus den Daten hervor, dass die Poly-E-Box für die Bildung des PRG2-PTEN-Komplexes sowie die PIP-Bindung erforderlich ist. Zusätzlich zeigen biochemische Daten, dass PRG2 über die Fähigkeit der Selbstassoziation verfügt sowie über die Möglichkeit der Ausbildung von Multimeren höherer Ordnung sowohl mit sich selbst als auch mit anderen Mitgliedern der PRG-Familie. Eine weitere neue Erkenntnis dieser Arbeit besteht darin, dass die C-Domäne des PRG2-Proteins alleine für eine Homomultimerbildung ausreicht.

1. Introduction

1.1 Membrane proteins

Integral membrane proteins (IMPs) are embedded in the different phospholipid membrane bilayer present in all eukaryotic cells. Present estimations propose that nearly one third, between 15 and 32% of open reading frames in the sequenced human genome encode IMPs [1–4]. As an example, the human protein database comprises 30057 proteins, and, taking the most concise IMP prediction method with a consensus of 13% of proteome into account, this estimation leads to approximately 3900 IMP existing in the human proteome [5]. Membrane protein families are defined by amino acid sequence similarities; however, individual members of a protein family can be functionally and structurally distinct. On the other hand, diverse amino acid sequences of different membrane protein family members can form similar structures [6].

Functionally, IMPs perform a range of essential key tasks in the cell for survival includes movement of molecules like nutrients, ions, the release of toxins, and waste products across phospholipid bilayers, as well as participating in cell signaling and motility [4]. IMPs are associated with heart disease, cystic fibrosis, depression, obesity, cancer, and many others [7-8], even a single missense mutation on the TM domain helix-helix interaction side can lead to misfolding and loss of function [9]. Approximately 60% of drugs target IMPs to achieve their therapeutic efficacy [10].

Understanding IMPs through structural, biochemical, and biophysical analyses are crucial for the development of new therapeutic approaches in order to build up a meticulous picture of how particular membrane proteins function at the molecular level. Since the first IMP structure has been solved by X-ray crystallography in 1985 [11], more than 943 unique membrane protein structures have been solved using the same method (http://blanco.biomol.uci.edu/mpstruc, Stephen White Lab at UC Irvine) (**Figure 1**).

1.2 Solubilization of membrane proteins

In recent years, the available techniques for protein expression, purification, and characterization were optimized for water-soluble proteins in an aqueous environment. IMPs exist in a lipid environment of phospholipid membrane, and, based on their

intrinsically hydrophobic characteristic, exhibit low solubility in aqueous environments. The poor water solubility of IMPs creates a challenge to successful *in vitro* protein characterization. Adding a detergent is one of the key strategies to solubilize the phospholipid membrane by forming a soluble complex with lipids and membrane proteins [12].

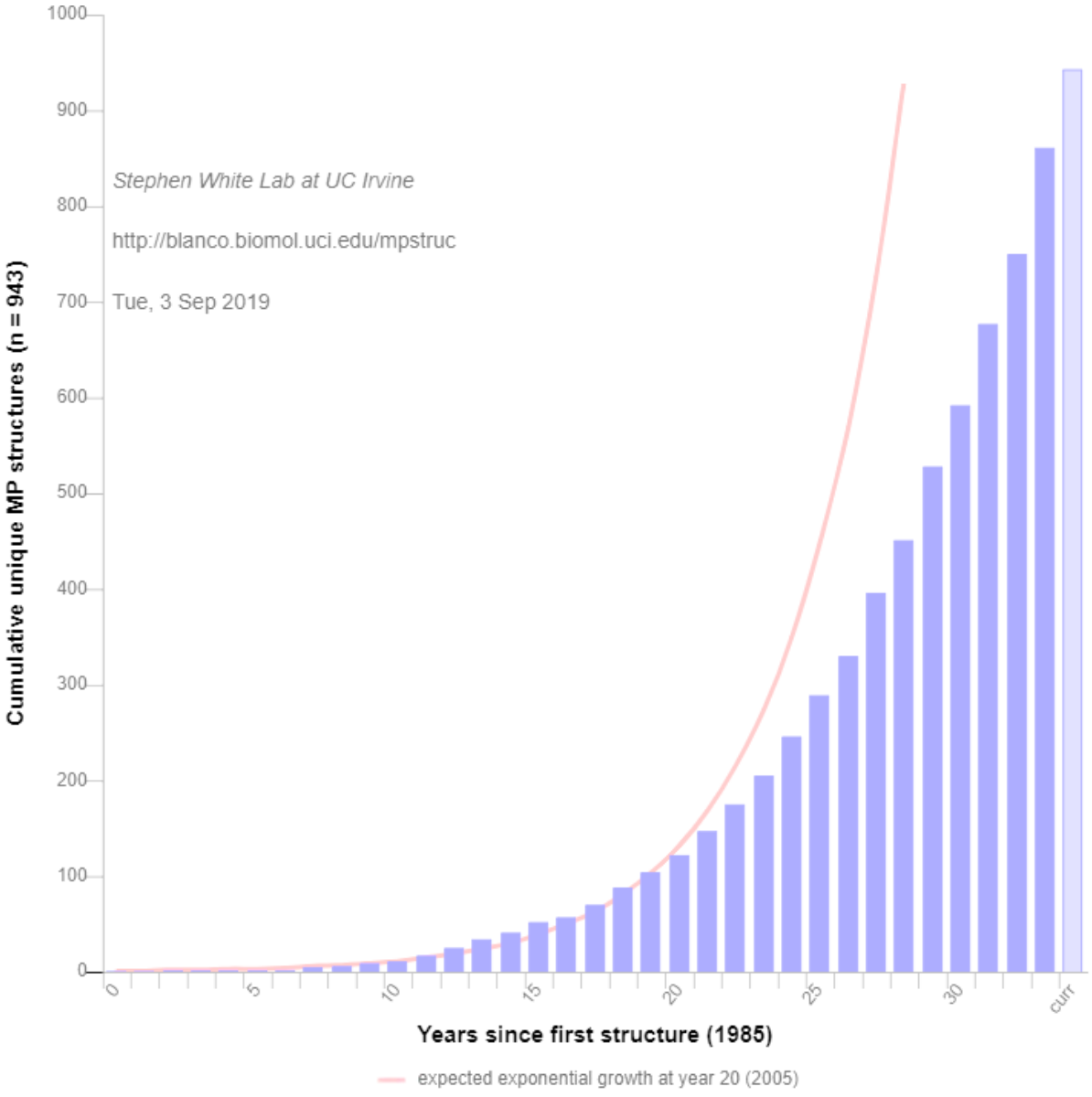


Figure 1: IMP structure determination statistics.
Cumulative number of unique IMP structures solved by X-ray crystallography. The orange curve shows the best exponential fit, illustrating that the increase in the number of unique IMP structures does not follow an exponential growth. The data on unique IMPs were collected on September 3, 2019 from PDB and from MPSTRUCT database (http://blanco.biomol.uci.edu/mpstruc/).

Detergent molecules form micelle structures, which encircles the membrane protein and provide an environment like the natural lipid surroundings. Nevertheless, solubilization with detergents is a harsh process that must be carefully optimized to avoid protein loss, inactivation, destabilization, aggregation and/or protein unfolding. Frequently, IMPs are prone to denature and/or aggregation. Therefore, solubilization is one of the most critical steps in handling IMPs.

Finding suitable detergents and buffer conditions, which support optimal protein stability without loss of protein function is an empirical and time-consuming process of trial and error [13]. Some high-throughput methods have been developed to aid in this search like ultracentrifugation [13], differential scanning fluorimetry (DSF) method [14], light-scattering assay [15], differential filtration assay [16], and insoluble fraction detergent screen complemented with dot-blot analysis as implemented in this thesis.

Detergents are a structurally diverse group of molecules that share a common amphipathic character, along with the ability to exist in solution either as a monomer or, as their concentration increases, or as self-associated micelles. The threshold of a detergent to form micelles depends on concentration, temperature, pH, and charge of the detergent, which is referred to as the critical micellar concentration (CMC) [17-19]. It is essential to know the CMC of a detergent being used to purify an IMP because the membrane protein will only stay in solution if the detergent concentration remains at, or above, the CMC [18], and detergent associates with the hydrophobic transmembrane domain of IMPs to create water-soluble protein-detergent complexes [19]. There are alternative methods developing for solubilizing IMPs, including the use of amphipols [20], bicelles [21] and nanodiscs [22], substituting the exterior lipid facing hydrophobic residues of the protein for hydrophilic ones [23]. In this study, we exchanged detergent with amphipol in order to prevent protein aggregation in the SEC and maintain biologically active conformation. Amphipols are long amphipathic polymers, which are not able to strip IMPs from their native membranes [24]. Therefore, the initial detergent solubilization step is mandatory before the detergent exchange with amphipols by using biobeads to absorb and eliminate detergents for effective substitution. In this study we used A8-35, where the letter “A” stands for “anionic”, the first number refers to the

average apparent molecular weight which is 8 kDa and the second number represents the percentage of free carboxylic groups.

1.3 Over-expression of membrane proteins

There are other key difficulties that add to our lack of detailed structural and functional understanding of membrane proteins. One of them is the low abundance of IMP in cells, which is due to the fact that their presence is limited with the occupancy of the physical capacity of the cell membrane. Another difficulty is the relatively low expression of IMPs, due to insufficient membrane insertion and folding or due to missing posttranslational modifications. Similarly, over-expression of membrane proteins can be toxic to the cell. Therefore, testing a wide range of eukaryotic cell lines to investigate the highest level of protein expression, avoiding any toxic gain of function, misfolding, aggregation, degradation, or mislocalization is an essential preparatory step before actual purification. As an example, in this study, we gain the advantage of using the HEK293S GnTI$^-$ cell line was prepared by mutagenesis by using ethyl methanesulfonate, which resulted in ablation of N-acetyl-glucosaminyltransferase I (GnTI) activity. Therefore, this cell line is missing complex N-glycans post-translational modification [25]. The main reason to establish this cell line is to minimize the large diversity of glycosylation, which interferes with the purification as well as the crystallization process. Moreover, this cell line also increases the yield of protein expression and leads to preferential IMPs localization in the plasma membrane. It has been successfully used to overexpress a variety of mammalian G protein-coupled receptors [25]. In our study, we also achieved the highest protein expression by using this cell line.

1.4 Dynamic light scattering

Dynamic Light Scattering (DLS) is a technique for measuring the size of various biomolecular particles including proteins, nucleic acids, and complexes of protein-protein, protein-nucleic acid, or protein-small molecule interactions, polymers, micelles, and vesicles [26]. Measurements are instigated in an aquatic environment, performed at certain a temperature by using Brownian motion, and relates this to the size of the particles. Brownian motion is the random movement of particles due to the bombardment by the solvent molecules. The larger the particle, the slower the Brownian

motion is. Whereas, smaller particles are hit further by the solvent molecules and move more rapidly.

DLS essentially measures fluctuations in scattered light intensity due to diffusing particles, therefore, if the system is monodispersed, the diffusion coefficient of the particles can be used to determine the size of the particle. In contrast, a polydispersed system would show multiple particle populations. Distribution of the particle population at different diameters can be also measured by this method. In our study, we used this technology to investigate the homogeneity of our purified protein in terms of the number and relative distribution of the different classes of IMP complexes. This method also allowed us to calculate the diameter of our purified protein complex.

1.5 Microscale Thermophoresis

Microscale thermophoresis (MST) is a technology for biophysical analysis of interactions between proteins [27], DNA [28], RNA [29], peptides [30], lipids [31], and ions in an aquatic, immobilization-free environment under native conditions. The term MST refers to the motion of molecules in microscopic temperature gradients. It allows for quantitative analysis of biomolecular interactions.

In the standard MST assay, one binding partner is fluorescently labeled, either through fluorescent-protein tagging or through crosslinking using a chemical dye carrying a reactive N-hydroxysuccinimide (NHS) group. As N- or C-terminal tagging of protein with a fluorescent protein marker may hinder protein structure and function, chemical labeling methods are currently favored. Therefore, in this thesis, we used the NHS chemical labeling method to perform the MST experiment.

In a typical MST assay, infrared (IR) laser is used for local heating at the exact spot where fluorescence intensity is measured. The IR radiation is absorbed by the water molecules of the sample buffer to create a temperature gradient. Heated up sample gains thermodynamic energy and moves further from the center of the heat source to colder surrounding to maintain lower thermodynamic energy for gaining back high level of order by lowering energy state, which is known as thermophoresis. This directed motion of molecules in temperature gradient is highly sensitive to all types of binding-

induced changes of molecular properties, size, charge, hydration shell or molecular conformation. Therefore, any molecular interaction will directly impact on a particular thermophoretic behavior of a fluorescently labeled molecule. Serial dilution of a ligand can provide a dose-dependent retention on the movement speed of labeled molecule in a defined time frame, and this data translates into a precise and quantitative disassociation constant (Kd) value. In our project, we implement this method to confirm PRG2-PTEN, to evaluate PRG2-PIPs interaction, and to measure Kd values, as well as to further investigate the potential binding domain of PRG2 to PTEN and PIPs.

1.6 Thermal Shift Assay

Since the stability of proteins is a critical factor during the preparation of protein samples at the native condition for structural analysis, designing a high throughput-screening platform for the thermal stability of proteins is essential. Thermal Shift Assay (TSA) was developed for testing the thermal stability of proteins in a variety of conditions, including pH, salts, additives, drugs, mutations, or following supplementation with other proteins or nucleic acids. The main scope of this approach is to characterize the optimum condition of a protein to maintain the most stable condition by achieving the highest degree of melting temperature (Tm).

The simplest and most commonly used method of TSA is the so-called Thermofluor assay. This assay requires fluorescent dye SYPRO Orange as a probe, which is selectively active in a hydrophobic, non-polar environment, whilst being quenched in aqueous polar solution. As the temperature is increased, the protein begins to unfold, and its hydrophobic core becomes exposed. This conformational shift induces the binding of the SYPRO Orange dye on hydrophobic domains of the unfolding proteins [32] and increases the fluorescent signal until all protein molecules are completely denatured [32,33]. As the SYPRO Orange dye binds on hydrophobic domains and detergent micelles, this dye is incompatible with IMPs. Therefore, a special procedure was developed for measuring the stability of solubilized IMPs, making use of a thiol-reactive fluorescent probe 7-diethylamino-3-(4′- maleimidylphenyl)-4-methylcoumarin (CPM) [34]. CPM has a low fluorescence in aqueous solution, which increases dramatically when it reacts with the side chain of free cysteine residues that become accessible to solution upon protein unfolding. This method is applicable as long as IMP

has cysteine in the hydrophobic core. Differential scanning fluorimetry (DSF) has been developed as an alternative, which is a fluorescent reporter dye-free method based on measuring protein unfolding by intrinsic fluorescence. In this assay, UV light is used to monitor tryptophan, tyrosine, and phenylalanine fluorescence to monitor protein unfolding. Similar to the CMP method, DSF is also limited to proteins that carry tryptophan, tyrosine, and phenylalanine residues in the core. We performed the DSF method to optimize the buffer condition for the purification of the native PRG2 complex.

1.7 UV-Vis spectrophotometry

Quantitative study of protein-protein and protein-ligand interactions in solution requires an accurate determination of protein concentrations. The given concentration of purified protein samples can be quantified by direct absorbance of aromatic chains on the amino acid tryptophan and tyrosine at 280 nm using a NanoDrop device. In this assay, the absorbance of light at 280nm (A280) can be exploited to extract the protein concentration (c) by applying the Beer-Lambert equation $A280 = c * \varepsilon * b$ (ε is the wavelength-dependent protein extinction coefficient, b is the pathlength). Each pure protein has a unique extinction coefficient. For precise results, the correct protein extinction coefficient ε must be considered, which can be estimated with ±5% accuracy using the primary amino acid sequence information [35].

1.8 Liquid chromatography

Liquid chromatography (LC) is an analytical technique that can be used for protein purification. It has a simple and modular setup consisting of a column that holds a stationary phase in equilibrium with a buffer solvent. The type of stationary phase determines the specific kind of separation. The column material can consist of porous particles for gel filtration (GF) also called size-exclusion chromatography (SEC), charged particles for ion-exchange chromatography (IEX), coated with ligand or antibody for immunoprecipitation (IP) also known affinity chromatography (AC), hydrophobic surfaces for reversed-phase chromatography (RPC), or metals like nickel for immobilized metal ion affinity chromatography (IMAC) to pull down histidine-tagged proteins. The protein sample is loaded onto the top of the column followed by buffer. Based on the specific nature of the protein sample, the sample mixture passes through the column at different rates due to differences in their partitioning behavior between the

mobile liquid phase and the stationary phase [36]. Therefore, the protein of interest is separated, and aliquots of the column effluent collected as fractions in time. Protein sample can be traced by either intrinsic fluorescence absorption at A280 UV light spectrum, or via fluorescent protein tag or chemical dye. This method also provides a rough information about the molecular weight of the protein of interest, if there is a known molecular weight marker calibration curve generated under the same temperature, back pressure, flow rate, and buffer conditions.

1.8.1 Size exclusion chromatography

Size exclusion chromatography (SEC) is a technique designed to separate biomolecules in aqueous systems based on the size and globular topology of the molecules. To perform a separation, the resin is packed into a column to form a packed bed. SEC resins consist of a fixed size porous matrix of chemically and physically stable spherical particles with properties that minimize the adsorption of proteins. The packed bed is equilibrated with a specific buffer, which fills the pores of the matrix and the space between the particles. Protein samples loaded onto the column are separated based on molecular size. Initially, insoluble aggregates molecules that are larger than the largest pores of the matrix are eluted together in the void volume without any retention by stacking in the porous surface of the column material. After this, big protein and large complexes eluate, followed by mid-size proteins, and finally, small proteins and peptides. SEC can be performed as an initial stage of purification, which is called the preparative scale of SEC, in order to fractionate different sizes of a complex mixture of proteins. SEC can also be performed directly after IEX, IMAC, or AC, which is known as the analytical SEC as the final step of purification to characterize the sample quality in terms of stability, homogeneity, and purity. Moreover, SEC also can be used for desalting or buffer exchange. In this study, we used SEC during both preparative and analytical protein purification steps as well as to exchange buffer and for desalting of samples.

1.8.2 Ion exchange chromatography

Ion exchange chromatography (IEX) is a separation method that is based on adsorption and reversible binding of charged protein to oppositely charged groups attached to an insoluble matrix. The isoelectric point (pI) is defined as the pH value at which the overall

charge of a given protein is equal to zero. When exposed to a pH below its pI, the biomolecule will carry a positive charge and will bind to a cation exchanger; at a pH above its pI, the protein will carry a negative charge and will bind to an anion exchanger [37]. Elution of target protein is performed by linear or gradient increase of buffer salt concentration. This procedure neutralizes the net charge of the column and therefore leads to the elution of bound proteins. After IEX, to characterize the elution point of the protein of interest, all eluted fractions must be analyzed by SDS-PAGE or WB. Because of the pI of our purified protein complex, we used anion exchange chromatography in this study.

1.8.3 Affinity chromatography

Affinity chromatography is the most selective, flexible, and intricate form of liquid chromatography that uses a biologically related agent as the stationary phase [38-40]. This stationary phase can be coated with a particular ligand, which demonstrates a high binding affinity of a target protein, such as an antibody for an antigen, a substrate for an enzyme, a hormone for a receptor, or a small peptide-like FLAG or 1D4 as used in this thesis. In addition to this, the histidine-tagged (6xHis) protein can be pull downed by using the so-called immobilized metal affinity chromatography (IMAC) method [41,42]. The chelator most commonly used as a ligand for IMAC is nitrilotriacetic acid (NTA) which binds on divalent metal ions such as Cu^{2+}, Ni^{2+}, Co^{2+}, or Zn^{2+}. The most commonly used metal ion is Ni^{2+} for IMAC purification. Target protein can selectively bind on this particular stationary phase of the column and all the rest of other biomolecules leave the column by extensive wash. Then a special reagent with higher affinity to the column has been used to eluate protein of interest, for example 3xFLAG peptide, higher molarity of 1D4 peptide, or imidazole for FLAG, 1D4, or His tagged protein elution respectively.

1.9 Electron microscopy

To understand a biological process, one of the most essential methods is to observe the process in the greatest detail possible. For example, to understand how internal a combustion engine is able to move a vehicle, it is necessary to see a detailed 3D map of every individual component and its specific position on a blueprint. Therefore, in this example, seeing is understanding. Biological processes do not represent an exception,

as the analyses of the 3D structure of each single component, and its spatial-temporal role is essential. Structural biology is filling the fundamental need to analyze processes at the highest possible resolution, by using a range of biophysical tools like X-ray crystallography, Nuclear magnetic resonance (NMR) spectroscopy, IR spectroscopy, or the emerging field of Cryogenic electron microscopy (cryo-EM). By gaining the advantages of these tools, detailed information on molecular structure and dynamics can be dissected in atomic or near-atomic resolution depth. At the cellular level, optical microscopy reveals the distribution and intramolecular dynamics of molecules tagged with fluorophores. Electron microscopy (EM) overlaps with this approach, covering a broad range from atomic to cellular structures. The development of negative stain and cryogenic methods has enabled EM imaging to provide snapshots of biological molecules trapped via heavy metal ion fixation [43], or shock frozen in liquid ethane and embedded in vitreous ice provides a close to native, hydrated state [44,45].

1.9.1 Single-particle electron microscopy

Single-particle electron microscopy (EM) can provide 3D structural information from images of individual molecules for a large variety of biological molecules. Ranging from small proteins with molecular weights of less than 0.1 MDa to large macromolecular assemblies with molecular weights higher than 10 MDa weight (**Figure 2**), Single-particle EM can generate structural information without the need to produce crystals. Another advantage is the relatively small abundance of a protein necessary for resolving its structures, as nanograms of pure samples are sufficient for single-particle EM. In this technique, to reconstruct a 3D volume, multiple viewing angles of the same object must be covered. To this end, a sample must contain multiple copies of the object in different orientations.

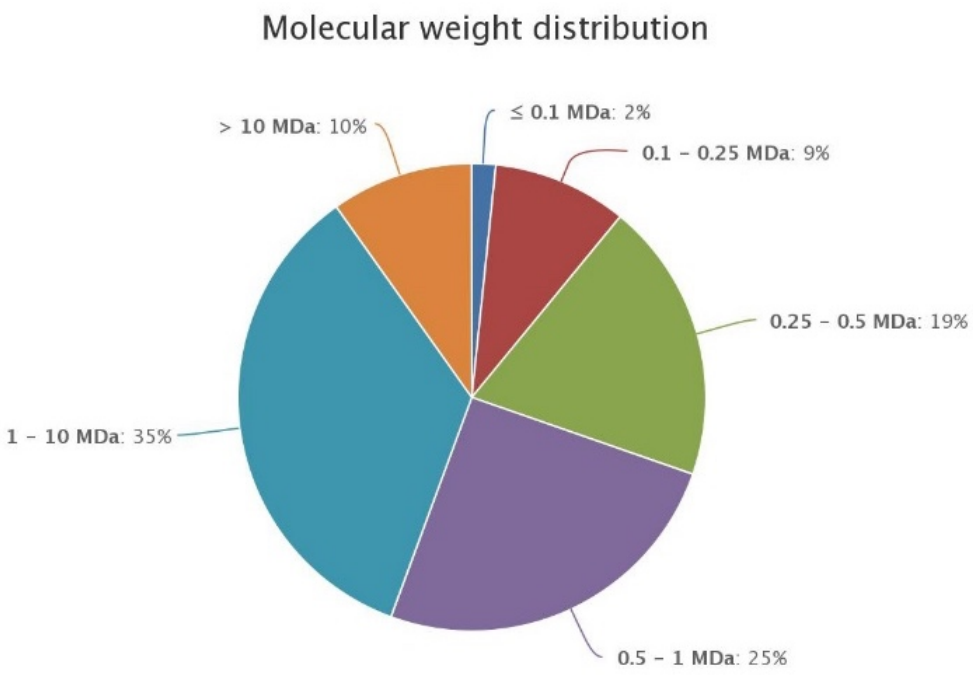


Figure 2: Molecular weight distribution of single-particle electron microscopy released maps.
Single particle EM is a powerful tool to study protein structures in a range of less than 100 kDa to higher than 10 MDa weight. Only 2% of entire database consists of smaller than 100 kDa molecular weight indicates existing limitation of this method at current stage. The data collected on June 10, 2019 from EMBL Protein Data Bank (https://www.ebi.ac.uk/pdbe/emdb/index.html/).

Since the production of the initial single-particle EM structure of actomyosin structures from sectioned muscle [46], and structure of the E. coli 70S ribosome were published in 2000 [47], there has been a drastic increase in number of identity protein 3D structure map. Up until July 2019, there are 8627 unique protein structure has been revealed (**Figure 3**). In parallel, resolution power has increased from 20 Å in 2000 to lower than 2 Å in 2019 (**Figure 4**). The highest resolution achieved so far is the 1.54 Å resolution structure of apoferritin [48].

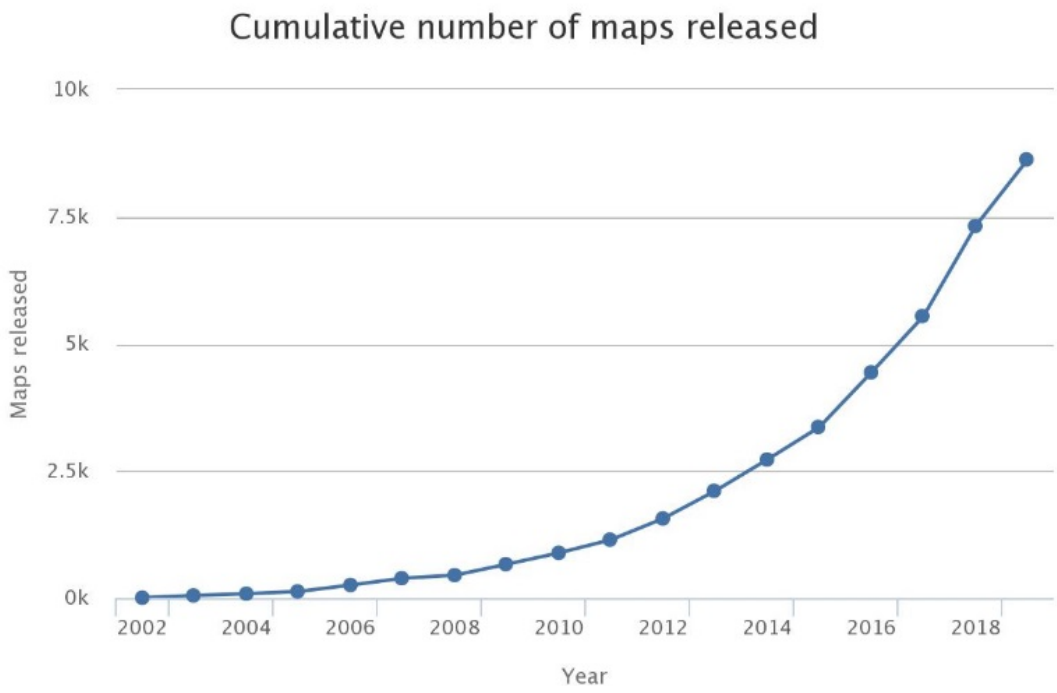


Figure 3: Cumulative number of unique protein structure maps solved by single-particle electron microscopy.
Since early 2000 there is an exponential growth in number of identified protein structure by using single-particle EM approach. The data collected on June 10, 2019 from EMBL Protein Data Bank (https://www.ebi.ac.uk/pdbe/emdb/index.html/).

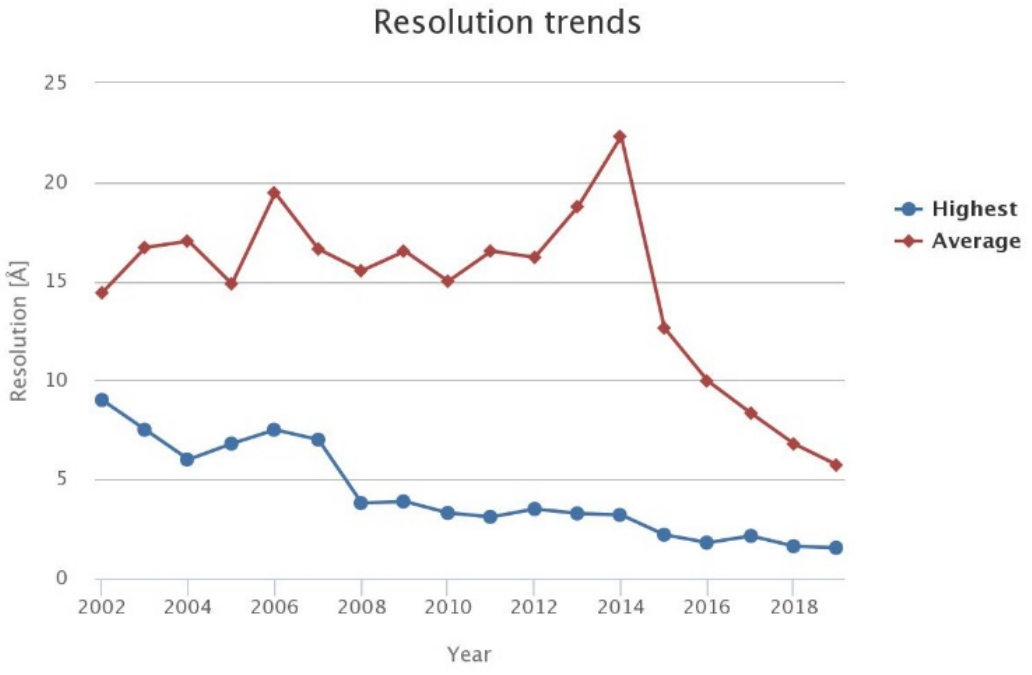


Figure 4: Resolution trends of identified protein structure by single-particle EM.
Dramatic increase in single-particle EM resolution in two decades. Blue line indicates the highest level of resolution power, and red line shows average resolution level in each year. The data collected on June 10, 2019 from EMBL Protein Data Bank (https://www.ebi.ac.uk/pdbe/emdb/index.html/).

1.9.2 Negative stain electron microscopy

Negative staining is an important technique in single-particle EM. For small molecules below 100 to 200 kDa weight, this method is ideal to generate sufficient contrast with high signal to noise ratio for detection. The simplest method for examining size, shape, symmetry, homogeneity, and suitable buffer condition of an isolated particles is negative staining [43], in which a droplet of the heavy metal stain coats the surface of purified sample suspension that is spread on a graphene carbon EM grid [49]. Electron dense heavy metal like uranyl acetate provides the highest contrast. However, as uranyl acetate has a low pH level from 4.2 to 4.5, some particles conserve better structure with more neutral or high pH metal stains from 7 to 9, like tungsten or molybdenum salts [50]. The method is called negative staining because the macromolecular shape is seen by exclusion rather than by binding of the particular stain to the molecules. Under the applied high vacuum pressure and exerted high electron energy on the grid, biological samples during this process are burned and vanish. What remains on the grid after the procedure are the clear negative footsteps of the sample in 2D. As the rest of the grid is still covered with heavy metal salt, lack of stain covered area defines the information of the surface topology of the particle. Although this method is quick and simple, it also has some limitations. Heavy metal stain can squish some flanking domains, it can break up weak interactions, flatten the sample, or it can distort structure. Frequently, the stain results in a single or very few preferred sample orientations on the grid due to applied weight force and charge. This can hinder on initial 3D model construction from 2D projection data sets. Also, as heavy metal penetration power is low, it can decrease the resolution of a negatively stained sample to 20 Å. Therefore, this method is chosen as a first step to evaluate sample quality and homogeneity. In case a particle population turns out heterogenic during negative stain analyses, proceeding further to performing vitrification for cryo-EM is not advised. In this case, single-particle cryo-EM image processing will not generate enough structural information to generate 3D structures. However, the negative stain method is a powerful first assessment for obtaining structural data for biomolecules where no information is available. Initial low-resolution 3D models can be used as a mask to build up a higher resolution electron density map by using cryo-EM imaging and data analysis.

1.10 Plasticity-related genes

Plasticity-related genes (PRGs), also known as lipid phosphate phosphatase-related proteins (LPPRs), are a family of integral membrane proteins that possess six-transmembrane domains. They are classified as a brain-enriched subclass of the lipid phosphate phosphatase (LPP) superfamily. Five members (PRG1–PRG5) having been identified to date [51-55]. The first identified member, PRG1, was identified as a protein upregulated following lesions to the adult hippocampal, and expression of this protein functionally increases axonal outgrowth during the early stage of brain development, and following axon injury in adulthood [51]. Therefore, the name of the protein relates to its ability to increase neuronal (structural) plasticity. The other family members were identified by in silico analysis [55]. PRGs show high homology with other members of the LPP superfamily of proteins, yet PRGs lack the bioactive lipid ecto-phosphatase activity within their extracellular loops. This LPP characteristic partakes as phosphatase that targets hydrolysis of bioactive lipid substrates, such as lysophosphatidic acid and sphingosine 1-phosphate [56], in different extracellular signaling events. PRGs demonstrate a restricted spatiotemporal expression profile, with an expression pattern that is associated with short periods during specific stages in development. Likewise, they show restricted spatial expression largely limited to the central nervous system (CNS) [57]. For example, the expression of PRG2 in rat brain lysate is restricted from the late prenatal stage to the early postnatal stage in rodents (E15 to P21), whilst PRG1 is limited to postnatal stages (P7 to P21) [58]. The PRG2 expression profile matches with the development stages of neuronal outgrowth and the formation of axon and dendritic branches. During development, axons in the CNS elongate through the extracellular space over long distances and form contacts with their target, facilitating the establishment of synapses. It has been shown that overexpression of different members of the plasticity-related gene (PRG1-5) family proteins induce the formation of filopodia [59,60], as well as branches. Moreover, some PRGs are upregulated in the adult brain in rodents following axonal injuries, suggesting that PRGs play additional roles in regenerating axons [51,52]. The exact functional roles of PRGs and their detailed molecular mechanisms nor their structural blueprint is currently unclear.

1.10.1 Multimerization of PRGs

PRG family members show structural and sequence similarity with six transmembrane spanning domains connected by intra- and extracellular loops, yet each member has a unique structurally highly disordered C-terminus ends. In particular, PRG1 and PRG2 have a long (~400 amino acids) C-termini, whereas PRG3, PRG4, and PRG5 have a very short (40-50 amino acids) C-termini (**Figure 5**). Although there is a high similarity in PRG1 and PRG2 C-termini, PRG2 has a highly acidic and negatively charged unique sequence of 20 glutamic acid residues ('poly-E-box') with an unknown function (**Figure 6**).

PRGs are able to form higher-order complexes by interaction with itself to form homo or their family members to form heteromeric complexes. It has been shown the interaction of PRG3 with PRG1, PRG2, and PRG5 in overexpression of tagged PRG variants co-immunoprecipitation experiments in Neuro2A cells [61]. High order multimerization capability of PRGs could exert specific physiological functions which have not been postulated yet.

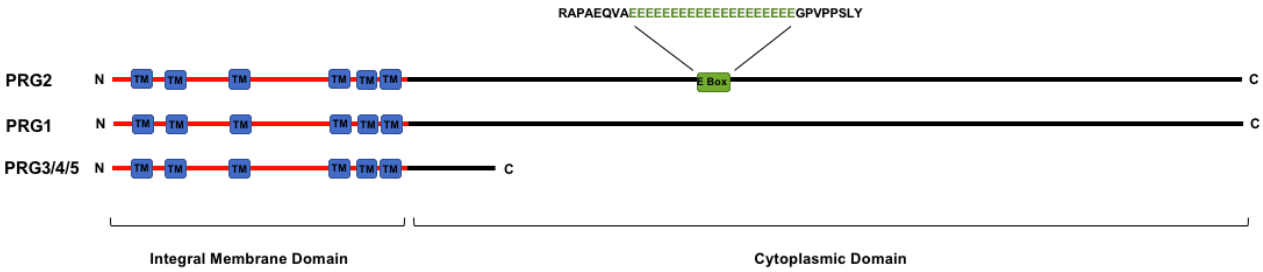


Figure 5: Schematic of the protein family of PRGs.
The protein family of PRGs includes 5 members (PRG1-5). These integral membrane proteins share a common domain structure consisting of six transmembrane domains (blue boxes) connected by intra- and extracellular loops (red line). The length of the intracellular cytoplasmic domain varies (black line). The highly homologous PRG1 and PRG2 possess a long cytoplasmic tail of 400 amino acids. Whereas PRG3-5 display a rather short cytoplasmic tail of 40-50 residues. Unique upon PRGs is a highly acidic stretch within the intracellular domain of PRG2, displayed by a sequence of 20 glutamic acid residues (displayed in green).

>NP_859009.2 phospholipid phosphatase-related protein type 3
[Mus musculus]

MLAMKEKNKTPKDSMTLLPCFYFVELPIVASSIVSLYFLELTDLFKPAKVGFQCYDRALSMPYVETNEELIPLLMLLSLAFAAPAASIMVGEGMVYCLQSRLWGRGPGGVEGSINAGGCNFNSFLRRTVRFVGVHVFGLCATALVTDVIQLATGYHTPFFLTVCKPNYTLLGTSCESNPYITQDICSGHDTHAILSARKTFPSQHATLSA*FAAVYVSMYFNAVISDTTKLLKPILVFAFAIAAGVCGLTQITQYRSHPVDVYAGFLIGAGIAAYLACHAV*GNFQAPPAEKVPTPAPAKDALRALTQRGHESMYQQNKSVSTDELGPPGRLEGVPRPVAREKTSLGSLKRASVDVDLLAPRSPMGKEGMVTFSNTLPRVSTPSLDDPARRHMTIHVPLDASRSRQLIGEWKQKSLEGRGLGLPDEASPVHLRAPAEQVAEEEEEEEEEEEEEEEEEEGPVPPSLYPTVQARPGLGPRVILPPRPGPQPLVHIPEEGVQAGAGLSPKSSSSSVRAKWLSMAEKGGGPVAVAPSQPRVANPPRLLQVIAMSKAAGGPKAETASSSASSDSSQYRSPSDRDSASIVTIDAHAPHHPVVHLSAGSTPWEWKAKVVEGEGSYELGDLARGFRSSCKQPGMGPGSPVSDVDQEEPRFGAVATVNLATGEGLPPPGASEGALGAGSRESTLRRQVGGLAEREVEAEAESYYRRMQARRYQD

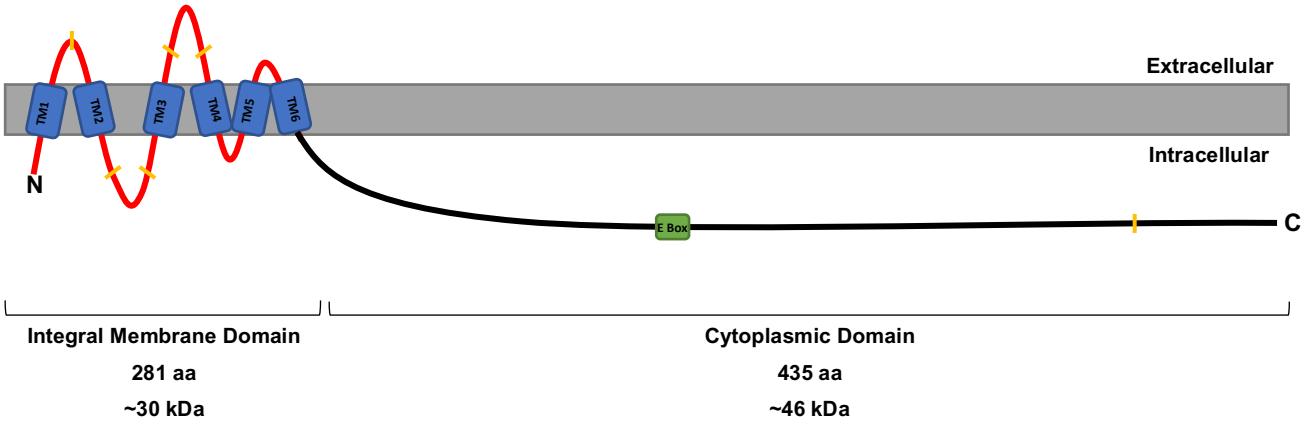


Figure 6: Sequence of mouse PRG2 protein.
A Amino acid sequence of mouse PRG2 protein taken from NCBI database. Red letters indicate integral membrane loops, blue underline sequence shows 6 trans membrane domains, black letters indicate long flanking cytoplasmic domain with 20 amino acid long poly glutamic acid stretch in green. Cysteines are showed in yellow boxes. **B** Schematic illustration of PRG2 protein topology embedded in the plasma membrane. Blue boxes are transmembrane domains. Green box is poly-E-box with 20 glutamic acid residues. Yellow lines represent cysteine positions.

1.10.2 Molecular interaction of PRG2, PTEN, and PIPs

In addition to multimerization, PRG2 has been identified as an interactor of tumor suppressor protein phosphatase and tensin homolog deleted on chromosome 10 (PTEN) in neurons by using mass spectrometry [61] and has been identified as a binding partner in rat brain lysates as well as in cortical neuron cultures by performing co-immunoprecipitation [62]. In this thesis, we further validate PRG2-PTEN interaction *in vitro* with overexpressed and purified PRG2 from HEK cells, narrow down direct binding domain (interaction site), and measure molecular affinities by using MST assay. In this

thesis, we also studied direct molecular interaction of PRG2-PIPs, to elucidate the role of PRG2 in local PIPs level.

PTEN is a lipid and protein phosphatase, which is mainly present in the cytosol. PTEN was initially identified as a protein tyrosine phosphatase (PTP) based on its sequence homology in the catalytic domain to members of the PTP family [63-65]. However, PTEN functions predominantly via transient localization to the plasma membrane, by dephosphorylating PI(3,4,5)P3 to PI(4,5)P2. Thereby it antagonizes the PI3K pathway and it opposes downstream PI3K–AKT–mTOR signaling pathways that coordinate cell proliferation, growth, survival, and metabolism [65, 66] (**Figure 7**). PIP3 can also act as a secondary messenger regulating cell polarity and migration [67,68]. For example, PTEN not only negatively controls cell motility through its lipid phosphatase activity by down-regulating Rac1 and Cdc42 in mouse fibroblast lines, and employs its tumor suppressor function in tumor progression and metastasis [67], but also contributing to the establishment of a PIP3–PIP2 gradient that has been shown to regulate cell polarity and migration in amoeba *Dictyostelium discoideum* [68], and similarly in neurons during axon filopodia initiation and branch formation [62]. Currently, it has been postulated that PRG2 spatially and temporally inhibits PTEN activity by direct interaction with PTEN [62].

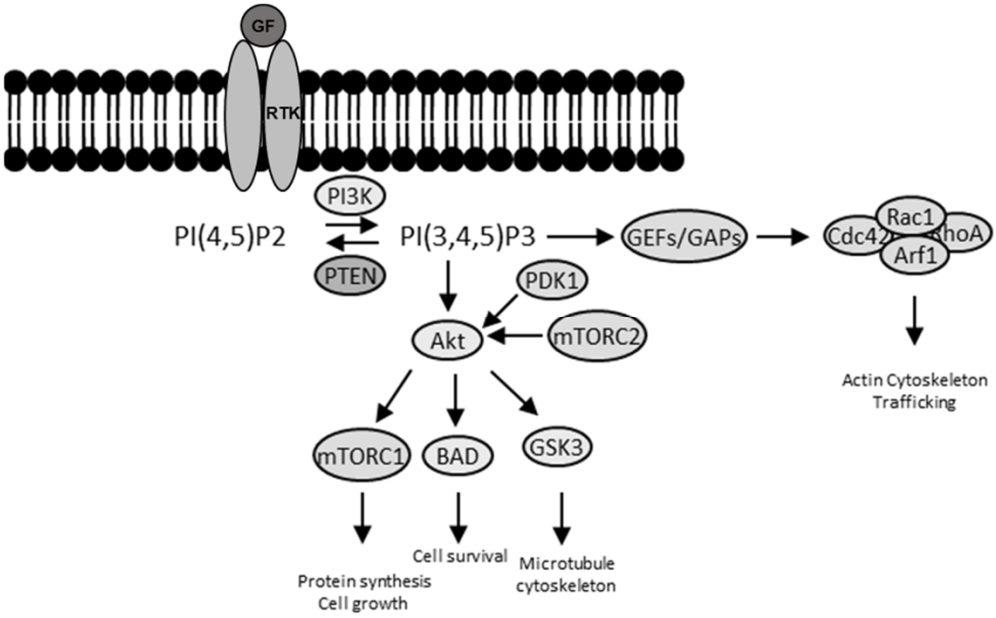


Figure 7: PI3K/PTEN signaling pathway.

PI3K is activated upon growth factor (GF) binding to the receptor tyrosine kinase (RTK). Active PI3K synthesizes PI(3,4,5)P3 from PI(4,5)P2. In turn PI(3,4,5)P3 mediates the activation of Akt by PDK1 and mTORC2. Subsequently, processes, such as protein synthesis, cell survival are regulated by downstream signaling of phosphorylated Akt and GTPases. PI3K is directly antagonized by PTEN, which functions by dephosphorylating PI(3,4,5)P3 to PI(4,5)P2 [69].

1.11 Aim of study

The main aim of this study was to establish a pipeline to overexpress and to purify PRG2 complexes under near-native conditions, which facilitated the analyses of the supercomplex organization of PRG2. Purified complexes were investigated by using negative stain, and later by cryogenic electron microscopy (cryo-EM) to gain initial insight into the macromolecular structure. In order to maintain the PRG2 complex under near-native condition, we established a purification strategy that mimicked the physiological lipid membrane environment using zwitterionic detergent, which was substituted with the amphipathic lipid molecule Amphipol (A8-25). The structural information will provide us to understand PRG complex formation in a native-like environment and increase the understanding physiological of the role of the PRG2 complex.

In addition to this, purified native PRG2 protein complexes were used for molecular interaction studies. Using MST assay, we characterized and measured the biochemical affinity of PRG2 to PTEN, as well as from PRG2 to PI(3,4,5)P3 and PI(4,5)P2. This characterization is vital to support our hypothesis that inhibition of PTEN activity by direct interaction with PRG2, could lead local inhibition of PTEN in a close vicinity of PRG2 accommodating nanodomains of the plasma membrane to initiate PI(3,4,5)P3 accumulations in these nanodomains. We also encounter direct interaction of PRG2 with PI(3,4,5)P3 in vitro MST assay. This result could speculate that PRG2 C-term could interact with PI(3,4,5)P3 in the inner leaflet of the plasma membrane which can create steric hindrance for PTEN enzymatic activity to dephosphorylate PI(3,4,5)P3. Particularly, PI(3,4,5)P3 domains are described to pave the way of F-actin patch formation, which in turn drives the initiation of axonal filopodia protrusion before microtubules invasion to mature filopodia into collateral branches. Eventually, this accumulation may lead to the formation of axonal filopodia at these focal adhesion points, which serve as precursors for collateral branches.

2. Material and Method

2.1 Materials

2.1.1 Chemicals

Table 1: List of Chemicals

Name	Company	Catalog number
0,22 µm pore size centrifugal filter	Millipore	UFC30GV00
150 cm^2 culture flask	Sigma-Aldrich	CLS3291
1-Step Transfer Buffer	Thermo Fisher Scientific	84731
1x Phosphate-buffered saline	Thermo Fisher Scientific	10010023
293 SFM II	Thermo Fisher Scientific	11686029
3xFLAG peptide	Sigma-Aldrich	F4799
A8-35	Anatrace	A835
Acrylamide	Bio-Rad	1610140
Ampicillin	Thermo Fisher Scientific	11593-027
APS	Thermo Fisher Scientific	17874
Bio-Beads	Bio-Rad	1523920
BSA	Sigma-Aldrich	A2058
Centrifugal Filters	Amicon	UFC503096
Centrifuge Columns, 0.8 mL	Thermo Fisher Scientific	89869
Coomassie	Thermo Fisher Scientific	LC6065
DAPI	Sigma-Aldrich	D9542
DDM	Anatrace	D310
DNAse	Roche	10104159001
DTT	Sigma-Aldrich	D0632
Duolink In Situ Red Starter Kit Mo/Rab	Sigma-Aldrich	DUO92101
ECL	GE Healthcare Life Sciences	RPN2236
EDTA	Sigma-Aldrich	EDS
EDTA free, Protease Inhibitor Cocktail	Roche	11873580001

Ethanol	Thermo Fisher Scientific	64-17-5
Fetal Bovine Serum	Merck	TMS-013-B
FLAG M2 monoclonal mouse affinity gel	Sigma-Aldrich	A2220
Fos Choline-14	Anatrace	F312
FreeStyle 293 Expression Medium	Thermo Fisher Scientific	12338018
Glutamax	Thermo Fisher Scientific	35050061
Glycine	Sigma-Aldrich	G7126
HEPES	Sigma-Aldrich	H3375
HiTrap HP Q anion exchange 5ml column	GE Healthcare Life Sciences	17115401
Hybridoma-SFM	Thermo Fisher Scientific	12045076
Imidazole	Sigma-Aldrich	I0250
Isopropyl-β-D-thiogalactoside	Sigma-Aldrich	I6758
JBS Floppy-Choppy	Jena Bioscience	CO-110
JBScreen Thermofluor Fundament	Jena Bioscience	CS-332
JBScreen Thermofluor Specific	Jena Bioscience	CS-333
Laemmli Sample Loading Buffer	Thermo Fisher Scientific	B0008
Lipofectamin 2000 Transfection Reagent	Thermo Fisher Scientific	11668027
Methanol	Thermo Fisher Scientific	67-56-1
Monolith NT.115 Capillaries	Nano-Temper Technologies	MO-K002
Monolith NT.115 Hydrophobic Capillaries	Nano-Temper Technologies	MO-KO23
NaCl	Sigma-Aldrich	S9888
NativeMark Unstained Protein Standard	Thermo Fisher Scientific	LC0725
NativePAGE 3-12% Bis-Tris Protein Gel	Thermo Fisher Scientific	BN1001
Ni Sepharose 6 Fast Flow HisPrep beads	GE Healthcare Life Sciences	17531802
Ni Sepharose HisTrap HP 5 ml column	GE Healthcare Life Sciences	17524802
Nitrocellulose membrane	GE Healthcare Life Sciences	10600000
Normal goat serum	Jackson Immuno Research	AB_2336990
NuPAGE 4-12% Bis-Tris Protein Gel	Thermo Fisher Scientific	NP0321
NuPAGE Antioxidant	Thermo Fisher Scientific	NP0005
NuPAGE Sample Reducing Agent	Thermo Fisher Scientific	NP0004
Nutrient Mixture F-12	Thermo Fisher Scientific	11320033
Paraformaldehyde	Sigma-Aldrich	158127

Penicillin-Streptomycin	Thermo Fisher Scientific	15140122
Pierce Silver Stain Kit	Thermo Fisher Scientific	24612
Polyethylenimine	Sigma-Aldrich	408727
PTEN recombinant protein	Origene	NM_000314
PureCube Rho1D4 agarose affinity resins	Cube Biotech	33102
PVDF membrane	Millipore	IPVH00010
Rho1D4 peptide	Cube Biotech	16201
SDS	Thermo Fisher Scientific	AM9822
SeeBlue Plus2 Protein Standard	Thermo Fisher Scientific	LC5925
Sucrose	Sigma-Aldrich	S9378
Superose 12 3.2/300 column	GE Healthcare Life Sciences	29-0362-25
TEMED	Bio-Rad	1610801
Tris-HCl	Sigma-Aldrich	93363
Trypsin	Promega	V5111
TurboFect Transfection Reagent	Thermo Fisher Scientific	R0531
Tween-20	Bio-Rad	1610781
Ultra-thin Carbon Film Grid	Electron Microscopy Sciences	CF300Cu-UL

2.1.2 Buffers and solutions

Table 2: List of buffers and solutions

Solution	Ingredients
Resuspension Buffer (Section 2.1.2)	20 mM HEPES 500 mM NaCl dH_2O pH 7.5
Blocking buffer (Section 2.1.3)	5% NGS 1x PBS
Lysis buffer (Section 2.3.1.1)	20 mM HEPES pH 7.4 150 mM NaCl 1 % (w/v Fos Choline-14) cOmplete EDTA-free Protease Inhibitor Cocktail dH_2O
SDS PAGE Buffer	0,5 M Tris-HCl 0,4% SDS dH_2O pH 6,8
Running gel (12%)	4.1 ml dH_2O 3.9 ml Acrylamide (30%) 2.5 ml SDS PAGE Buffer 10 µl TEMED 32 µl APS (10%)
Stacking gel (4%)	6.1 ml dH2O 1.3 ml Acrylamide (30%) 2.5 ml SDS PAGE Buffer 10 µl TEMED 100 µl APS (10%)

TBS-T (10x)	50mM Tris-HCl 150 mM NaCl 0,05% Tween-20 dH_2O pH 7,4
Transfer buffer (1x)	20 mM Tris-HCl 190 mM Glycine 20% Methanol dH_2O
Resuspension Buffer (Section 2.3.5.2)	20 mM HEPES 500 mM NaCl 20 mM imidazole cOmplete EDTA-free Protease Inhibitor Cocktail dH_2O pH 7.5
Elution Buffer (Section 2.3.5.2)	20 mM HEPES 500 mM NaCl 250 mM imidazole 5 mM DTT dH_2O pH 8.5
Equilibration Buffer (Section 2.3.5.2)	50 mM Tris 100 mM NaCl dH_2O pH 7.5
High Salt Buffer (Section 2.3.5.2)	50 mM Tris 1 M NaCl

2.1.3 Antibodies

Table 3: Primary antibodies

Name	Company	Catalog Number
GFP rabbit	Abcam	ab6556
FLAG M2 monoclonal mouse antibody	Sigma	F3165
6x-His Tag antibody	Thermo Fisher Scientific	MA1-21315-HRP
PTEN rabbit monoclonal mouse antibody	Cell Signaling	9559

Table 4: Secondary antibodies

Name	Company	Catalog Number
Alexa Fluor 647 donkey anti-mouse	Jackson Immuno Research	715-605-150
Alexa Fluor 488 goat anti-rabbit	Jackson Immuno Research	711-545-152
IgG (H+L HRP goat anti-mouse)	Thermo Fisher Scientific	62-6520

2.1.4 Devices

Table 5: Devices

Name	Company
100 kV Philips CM100 Electron Microscope	Philips
ÄKTA FPLC	GE ÄKTA pure & analytical
Amersham Imager 600	GE Healthcare Life Sciences
DynaPro NanoStar	Wyatt Technology
Monolith NT.115	Nano-Temper Technologies

NanoDrop	Thermo Fisher Scientific
Orbitrap Velos mass spectrometer	Thermo Scientific
Power Blotter system	Thermo Fisher Scientific
Prometheus NT.48	Nano-Temper Technologies
Reversed-phase Ultimate 3000 LC	Thermo Scientific
SP5 Laser Scanning Microscope	Leica

2.1.5 Software

Table 6: Software

Name	**Company**
e2boxer software	National Institutes of Health
ImageJ software	National Institutes of Health
Mascot Server Version 2.5.0	Matrix Science
MO.Affinity Analysis software	Nano-Temper Technologies
UNICORN 5 software	GE Healthcare Life Sciences

2.2 Methods

2.2.1 Cell culture

2.2.1.1 Cell culture and transient transfection of cell lines (HEK293S GnTI$^-$, N1E-115, COS-7)

Cells were cultured in Dulbecco's Modified Eagle Medium: Nutrient Mixture F-12 (Thermo Fisher Scientific) with 1% Penicillin-Streptomycin (Thermo Fisher Scientific) 1% Glutamax (Thermo Fisher Scientific) and 10% Fetal Bovine Serum (FBS) (Merck). 12 million cells were seeded and transfected one day after plating with mouse PRG2

(NCBI Reference Sequence: NP_859009.2, full residues 1-716) containing C-terminal FLAG and 1D4 tags, PRG2ΔC (residues 1-438) containing a C-terminal FLAG tag, PRG2ΔN (residues 285-716) containing a 6xHistidine tag at C-terminal, and GFP control were cloned in pMT4 vector using [1mg/ml] sterile polyethylenimine (PEI) with 2:1 (v/v) ratio of PEI:DNA, Turbofect (Thermo Fisher Scientific) and Lipofectamine 2000 (Thermo Fisher Scientific) for transient transfection followed by the supplier's protocol. Cell lines were cultured at 37 °C in a 5 % CO2 atmosphere in T150 flasks (Corning) for two days.

2.2.1.2 Bacterial culture and transformation (E. coli BL21 (DE3), LOBSTR-E. coli BL21(DE3))

The intracellular domain of mouse PRG2 (PRG2ΔN) (residues 285-716) containing a 6xHistidine tag at C-terminal were subcloned into the prokaryotic vector pPAL. The recombinant plasmid was transformed into E. coli BL21 (DE3), or LOBSTR-E. coli BL21(DE3) strain using a standard heat shock protocol (42°C for 20 seconds) for protein expression. The transformed E. coli strains were cultured at 37°C in 100 ml Lysogeny broth (LB) medium containing 100 µg/ml ampicillin (Thermo Fisher Scientific)
for overnight at 180 rpm. 30 ml of this culture was transferred in 400 ml LB medium containing ampicillin and incubated at 37°C with 180 rpm until the culture reached an optical density of 0.6 at a wavelength of 600 nm (OD600≈0.6). Then, isopropyl-β-D-thiogalactoside (IPTG) (Sigma-Aldrich) was added to a final concentration of 0.1 mM to the cultures. After further overnight incubation at 18°C with 180 rpm shaking, cells were collected by 6,500 rpm 15 min centrifugation and resuspended in 25 ml of 20 mM HEPES, 500 mM NaCl at pH 7.5 buffer, and stored at -40°C.

2.2.1.3 Proximity ligation assay

N1E-115 cells were fixed with 4% Paraformaldehyde (PFA) (Sigma-Aldrich) containing 15% Sucrose (Sigma-Aldrich) solution for 30 minutes at room temperature. Cultures were then washed with sterile 1x Phosphate-buffered saline (PBS) buffer (Thermo Fisher Scientific) and blocked with 5% normal goat serum (Jackson Immuno Research) for 1 hour at room temperature. Cells were probed by 1:100,000 dilutions of anti-FLAG (Sigma-Aldrich, #F3165) and anti-GFP (Abcam, # ab6556) primary antibodies in blocking buffer for overnight at 4°C. Cells were then washed with sterile 1x PBS (Thermo Fisher Scientific) for 10 minutes at room temperature three times and probed with 1:20

dilution of Duolink PLA reagent (Sigma-Aldrich). Further immunocytochemistry staining was undertaken to enhance FLAG and GFP tagged proteins detection with the same set of FLAG and GFP primary antibodies with 1:2,000 dilution rate, and coupled with secondary antibodies Alexa Fluor 647 donkey anti-mouse (Jackson Immuno Research, # 715-605-150) and Alexa Fluor 488 goat anti-rabbit (Jackson Immuno Research, #711-545-152) at 1:1,000 dilution. Nuclear staining was performed using 1:10,000 dilution of 4′,6-diamidino-2-phenylindole (DAPI) (Sigma-Aldrich) for 5 min incubation at room temperature.

2.2.2 Image acquisition and analysis

Confocal images of PLA probed N1E-115 cells were acquired using the Leica TCS SP5 Laser Scanning Microscope. Images were analyzed using ImageJ software (National Institutes of Health, USA).

2.2.3 Biochemistry

2.2.3.1.1 Preparation of mammalian cell lysate for western blotting

Cultured cell lines were washed with 1x PBS, and harvested by scraping and centrifugation at 4,000g. All extraction procedures were carried out at 4°C. The cell pellets were washed twice with 1x PBS and lysed in 20 mM HEPES pH 7.4, 150 mM NaCl, 1 % (w/v) Fos Choline-14 (Anatrace) supplemented with a cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail (Roche) by gentle agitation for 1 hour. After extraction, the supernatant was collected after a 40 min centrifugation at 14,000 g, mixed with 5x Laemmli Sample Loading Buffer (SLB) (Thermo Fisher Scientific) with 1% sodium dodecyl sulfate (SDS) (Thermo Fisher Scientific), and loaded onto a home cast 12% SDS polyacrylamide gel electrophoresis (PAGE). The protein concentrations of supernatants were determined using the NanoDrop Ultraviolet-visible spectrophotometers (Thermo Fisher Scientific) at 280 nm wavelength.

2.2.3.1.2 Preparation of bacterial cell lysate for western blotting

Frozen cells were thawed, supplemented with a cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail and 1 ml deoxyribonuclease (DNAse) (3mg/ml). Then cells were sonicated 3 times with 20 second pause at 1 min 6 cycle 10% with mid amplitude on ice.

The lysate was centrifuged at 6,500 rpm at 4°C for 10 min. The supernatants were collected and mixed with 5x Laemmli SLB, heated at 95°C for 3 min, and loaded onto a 12% SDS PAGE. The protein concentration in the supernatant was determined using NanoDrop.

2.2.3.2 SDS-polyacrylamide gel electrophoresis (SDS PAGE)

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis is an analytical method to separate proteins in a gel matrix according to their molecular weight. SDS PAGE consisted of a stacking gel (**Table 2**) with 4% acrylamide and a running gel (**Table 2**) with 12% acrylamide. Lysates were loaded onto the gel and an electric field was applied, first 80 volts (V) for 30 min followed by 140 V until the blue front ran out of the gel. Each gel was run with PAGERuler Prestained Protein Ladder (Thermo Fisher Scientific) as a reference.

2.2.3.3 Western blotting

After gel electrophoresis, proteins were immobilized on a membrane and detected by specific antibodies. Here, proteins were transferred from the polyacrylamide matrix to a nitrocellulose membrane using a Power Blotter system (Thermo Fisher Scientific) according to the user manual for high molecular weight proteins provided by the supplier. Post-transfer, membranes were blocked with 5% Bovine Serum Albumin (BSA) (Sigma) in TBS-T buffer for 1 hour at room temperature at 25 rpm rotor. Overnight incubation was undertaken using a 1:5,000 dilution of FLAG M2 monoclonal mouse antibody (Sigma, # F3165) or 1:20,000 dilution of 6x-His Tag Antibody (Thermo Fisher Scientific, # MA1-21315-HRP) in TBS-T buffer at 4°C at 25 rpm rotor. Blots were then washed three times with TBS-T buffer, and incubated for 1 hour at room temperature with 1:10,000 dilution of goat anti-mouse IgG (H+L) secondary antibody coupled with horseradish peroxidase (HRP) (Thermo Fisher Scientific, # 62-6520) at 25 rpm rotor for FLAG-tagged proteins. Washed three times with TBS-T buffer and enhanced chemiluminescence (ECL) solution was added on the blot. ECL exposure was detected by Amersham Imager 600 (GE Healthcare).

2.2.3.4 Dot blotting

2.2.3.4.1 Dot blotting for detection of purified protein in SEC fractions

2 µl of a sample fraction derived from the size exclusion chromatography (SEC) was spotted on nitrocellulose membrane and blocked with 5% BSA in TBS-T buffer for 1 hour at room temperature at 25 rpm shaker. Then, proteins were detected and visualized using primary and secondary HRP antibodies as explained above in the western blotting section (**2.2.3.3**). Protein detection was performed using ECL solution and exposure was detected by Amersham Imager 600.

2.2.3.4.2 Dot blotting for PRG2-PTEN interaction

2 µg protein or peptide sample spotted onto a nitrocellulose membrane were blocked with 5% BSA in TBS-T buffer for 1 hour at room temperature at 25 rpm shaker. Next, the membranes were incubated for 2 hours with 5 µg/ml human recombinant PTEN protein (Origene NM_000314) diluted in TBS-T buffer at 25 rpm shaking at room temperature. Blots were then washed three times with TBS-T buffer and incubated for 1 hour at room temperature with 1:2,500 dilution of PTEN rabbit monoclonal mouse antibody (Cell Signaling, #9559) in TBS-T buffer at 25 rpm shaker at room temperature. Following three further washes with TBS-T buffer, blots were then incubated for 1 hour with 1:10,000 dilution of goat anti-mouse IgG secondary antibody coupled with HRP at 25 rpm shaking at room temperature. Then, ECL solution was added on the blot. ECL exposure was detected by Amersham Imager 600.

2.2.3.5 Protein purification

2.2.3.5.1 Protein purification from mammalian cell lysate with FLAG or 1D4 Immunoprecipitation (IP)

For FLAG or 1D4 IP, HEK293S GnTI- cells were plated in T150 flasks and transfected with FLAG 1D4 tagged plasmid. 24 h post-transfection, cells were lysed and total cell lysate (TL) collected as described in section 2.3.1.1. 20 µL of a TL fraction was kept for Western blot (WB) analyses. 20 µL FLAG M2 monoclonal mouse antibody affinity gel (Sigma-Aldrich # A2220) or PureCube Rho1D4 agarose affinity resins (Cube Biotech # 33102) was used per condition. Before the IP, beads were washed twice with lysis buffer

by centrifugation for 2 min at 1000 rpm. All centrifugation steps were conducted at 4°C. Beads were resuspended in lysis buffer in the same volume, respectively. IP was undertaken to rotate and overnight at 4°C. The beads/lysate mixture was spun down, and supernatants were transferred to a fresh tube and kept as supernatant (SN) that were probed later to monitor IP efficiency. Beads were washed 3x with lysis buffer containing 0.05% Fos Choline-14. Elution of beads was performed by the addition of 200 µM 3xFLAG peptide (Sigma-Aldrich # F4799) or Rho1D4 peptide (Cube Biotech # 16201) overnight. Immunoprecipitated samples and inputs were loaded onto a 12% SDS PAGE gel and analyzed by WB using the FLAG M2 monoclonal mouse antibody. Subsequent to collecting IP eluates, detergent was exchanged with amphipol (A8-35) (Anatrace). For the substitution process, 1 % (w/v) A8-35 was added and incubated for 1 hour. After the initial incubation, active Bio-Beads (Bio-Rad) were added, and the removal of detergent was carried out by overnight incubation. Eluates in A8-35 were concentrated with 100 kDa cut off Ultra-0.5 mL Centrifugal Filters (Amicon) and loaded onto a SEC, which was equilibrated with 100 mM Tris pH 7.4, 150 mM NaCl. SEC fractions were spotted for dot blot analysis using the FLAG antibody to detect the highest enrichment of purified protein among all SEC fractions.

2.2.3.5.2 Protein purification from mammalian cell lysate with His pull-down assay

For His IP, HEK293S GnTI- cells were plated in T150 flasks and transfected with His tagged plasmids. 24 h post-transfection cells were harvested and resuspended with 20 mM HEPES, 500 mM NaCl, 20 mM imidazole (Sigma-Aldrich) at pH 7.5 buffer supplemented with a cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail (Roche). Then, cells were sonicated 3 times with 30 second pauses at 30 sec 6 cycle 10% with mid amplitude on ice. The lysates were then centrifuged at 6,500 rpm for 10 min at 4°C. The supernatants were collected and saved on ice for purification using Ni Sepharose 6 Fast Flow HisPrep beads (GE Healthcare Life Sciences). Nickel beads were equilibrated with 20 mM HEPES, 500 mM NaCl, 20 mM imidazole at pH 7.5 buffer, and mixed with total protein cell lysate samples. Pull-down assays were undertaken by rotating overnight at 4°C. Afterward, the beads/lysate was spun down. Beads were washed 3x with 20 mM HEPES, 500 mM NaCl, 20 mM imidazole at pH 7.5 buffer. Elution of bound protein was performed by incubating beads with 20 mM HEPES, 500 mM NaCl, 250 mM imidazole, 5 mM DTT at pH 8.5 buffer for 2 hours at 4°C on an orbital shaker.

Eluates were then concentrated by a 30 kDa cut off Amicon Ultra-14 ml centrifugal filter and loaded for SEC on a Superose 12 3.2/300 column (GE Healthcare Life Sciences) pre-equilibrated with 20 mM HEPES, 500 mM NaCl, 5 mM DTT at pH 8.5 buffer. Fractions were loaded onto a 12% SDS PAGE and stained with Coomassie (Bio-Rad). Best fractions were used for negative stain EM analyses.

2.2.3.5.3 Protein purification from bacterial cell lysate with His pull-down assay

Frozen cells were thawed, supplemented with a cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail and 1ml DNAse (3mg/ml). Then cells were sonicated 3 times with 20 second pause at 1 min 6 cycles 10% with mid amplitude on ice. The lysate was centrifuged at 6,500 rpm at 4°C for 10 min, and supernatants were collected and saved on ice for purification using Ni Sepharose High-Performance HisTrap HP 5 ml column (GE Healthcare Life Sciences). HisTrap columns were equilibrated with 20 mM HEPES, 500 mM NaCl, 20 mM imidazole at pH 7.5 buffer, and samples were loaded onto columns. Elution of bound protein was undertaken using a linear gradient of 20 mM to 500 mM imidazole in 20 mM HEPES, 500 mM NaCl at pH 7.5 by using ÄKTA FPLC. Fractions were loaded onto a 12% SDS PAGE gel and stained with Coomassie to investigate fractions, which carried eluted His tagged proteins. Protein-enriched fractions were pooled, and the eluate was concentrated by 30 kDa cut off Amicon Ultra-14 ml centrifugal filter. Subsequently, pooled fractions were loaded for SEC onto home casted resin columns, pre-equilibrated with 50 mM Tris, 100 mM NaCl at pH 7.5 to eliminate imidazole that leads to protein precipitation. All collected fractions were loaded onto a 12% SDS PAGE gel and stained with Coomassie to identify enriched fractions. Selected fractions were pooled and loaded on a HiTrap HP Q 5ml anion exchange column (GE Healthcare), which was pre-equilibrated with 50 mM Tris, 100 mM NaCl at pH 7.5. Elution of bound protein was performed using a stepwise gradient of 100 mM to 1 M NaCl concentration in 50 mM Tris at pH 7.5 by using ÄKTA FPLC. Fractions were loaded onto a 12% SDS PAGE gel and stained with Coomassie to identify the enriched fractions. Fractions were then pooled and concentrated by 30 kDa cut off Amicon Ultra-0.5 ml centrifugal filter. Samples with a final concentration of 10 mg/ml were used for crystallization screen. The remaining protein was purified by SEC on a Superose 12 3.2/300 column (GE Healthcare) pre-equilibrated with 50 mM Tris, 100 mM NaCl at pH

7.5. Fractions were loaded onto a 12% SDS PAGE and stained with Coomassie. Purest fractions were used for negative stain EM.

2.2.3.6 Blue native PAGE

Following purification, the molecular weights of the native protein complexes were validated by blue native polyacrylamide gel electrophoresis (BN-PAGE) and western blot using anti-FLAG antibody. BN-PAGE was performed with the NativePAGE Novex Bis-Tris gel system (Thermo Fisher Scientific) according to the manufacturer's protocol. NativePAGE Sample Buffer (4X) was added to 20 µl of purified protein and loaded onto a 3-12% Bis-Tris gel. BN-PAGE was performed at 150 V for 1 h and 250 V for several hours in the cold room until the blue front ran out. Afterward, the gel was blotted on a polyvinylidene difluoride (PVDF) membrane (Immobilon-P 0.45 µm, Millipore) using the Power Blotter system (Thermo Fisher Scientific) according to the user manual for high molecular weight proteins. The membrane was then blocked, labeled, and detected as explained in the western blotting section (**2.2.3.3**).

2.2.4 Single particle negative stain electron microscopy (EM) study

Purified protein samples were used for negative stain EM investigation. Initially, samples were spun down for 10 minutes at 4°C at 14,000 g to eliminate any aggregates. Then, 3.5 µl of sample was loaded on a carbon film-coated EM grid, which was electrostatically discharged. After incubation of 45 seconds, the samples were blotted out with Whatman filter paper, and 3,5 µl 2% heavy metal salts uranyl acetate negative staining solution was applied to the grid. After 45 second incubation, the stain was blotted, and the grid dried for 10 min before subjecting to 100 kV Philips CM100 EM equipped with a 1kx1k Fastscan CCD camera (TVIPS) and an Oxford cryo holder. Micrograph images were taken and used for further selection of potential particles by using e2boxer software (NIH). Initially, we picked 10 to 20 particles as a reference by using the Swarm function of the image analysis software to generate automatic particle picking. We set the box size into 128 pixel, and proximity threshold setting to one-fourth of the box size. Next, we went through all chosen particles manually to eliminate bad particles, which looks like aggregate and staining artifacts. Eventually, we ended up with particles to generate class averages. To do so, we performed particle centering and tilting by applying a

planar angle. Particles were clustered based on their image signal to noise ratio. These 2D particle classes were used to attempt to build up the initial 3D model generation.

2.2.5 Liquid chromatography-mass spectrometry (LC-MS/MS)

All mass spectrometry experiments made in the Free University of Berlin under the supervision of Dr. Benno Kropka. For LC-MS/MS analysis purified protein samples were separated onto 4-12% NuPAGE Bis-Tris Gels (Thermo Fisher Scientific) and proteins were stained by Coomassie Brilliant Blue. Stained gel bands were excised, and in-gel digested with trypsin. In brief, gel bands were washed with 50% (v/v) acetonitrile in 50 mM ammonium bicarbonate, shrunk by dehydration in acetonitrile, and dried in a vacuum centrifuge. The dried gel pieces were incubated with ~50 ng trypsin (sequencing grade modified, Promega) in 25 µL of 50 mM ammonium bicarbonate at 37 °C overnight. To extract the peptides, 25 µL of 0.5% (v/v) trifluoroacetic acid (TFA) in acetonitrile was added and the extract was dried under vacuum. Peptides were reconstituted in 10 µl of 0.1% (v/v) TFA, 5% (v/v) acetonitrile and analyzed by a reversed-phase capillary nano liquid chromatography system (Ultimate 3000, Thermo Scientific) connected to an Orbitrap Velos mass spectrometer (Thermo Scientific). Samples were desalted on a trap column (PepMap100 C18, 3 µm, 100 Å, 75 µm i.d. × 2 cm, Thermo Scientific) using a mobile phase of 0.05% TFA, 2% acetonitrile in water. After switching the trap column inline, LC separations were performed on a capillary column (Acclaim PepMap100 C18, 2 µm, 100 Å, 75 µm i.d. × 25 cm, Thermo Scientific) at an eluent flow rate of 300 nl/min. Mobile phase A contained 0.1% formic acid in water, and mobile phase B contained 0.1% formic acid in acetonitrile. The column was pre-equilibrated with 3 % mobile phase B followed by an increase of 3–50% mobile phase B in 50 min. Mass spectra were acquired in a data-dependent mode utilizing a single MS survey scan (m/z 350-1500) with a resolution of 60,000 in the Orbitrap, and MS/MS scans of the 20 most intense precursor ions in the linear trap quadrupole. Identification of proteins was performed using the Mascot Daemon and Mascot Server (version 2.5.0, Matrix Science). Raw data were searched against the SwissProt human database (May 2017) including the sequence of tagged protein that was manually added. A maximum of two missed cleavages was allowed and the mass tolerance of precursor and sequence ions was set to 10 ppm and 0.35 Da, respectively. Oxidation, propionamide, and acetylation were

used as variable modifications. A significance threshold of 0.05 was used based on decoy database searches and a peptide ion score cut-off of 20 was applied.

2.2.6 Microscale Thermophoresis (MST)

To test the function of the purified native PRG2 complex, we performed MST by using Monolith NT.115 (Nano-Temper Technologies) to study PTEN and PI(4,5)P2 , PI(3,4,5)P2 binding.

Purified protein complexes were concentrated to 20 µM by filter centrifugation and fluorescently labeled with Monolith NT-647 Protein Labeling kit RED-NHS (Nano-Temper Technologies) according to the user manual. Access dye was eliminated, and protein purified by SEC on a Superose 12 3.2/300 column (GE Healthcare) pre-equilibrated with 20 mM HEPES pH 7.4, 150 mM NaCl, 0.05 % (v/v) Tween 20. The protein peak was collected to perform MST measurements.

2.2.6.1 MST with PTEN

MST was measured individually with 20 nM fluorescently labelled purified protein sample in 20 mM HEPES pH 7.4, 150 mM NaCl, 0.05 % (v/v) Tween 20. The samples were added to standard capillaries (Nano-Temper Technologies), and MST was measured with 20 % IR laser power, using medium MST power settings. Sixteen-fold serial dilution of 2.5 µM PTEN (ORIGENE) ligand in 25 mM Tris-HCl pH 7.3, 100 mM glycine, 10 % (v/v) glycerol was prepared, mixed with a constant concentration of labeled purified protein. The protein mixtures were then loaded into capillaries and analyzed in the instrument by subsequent scanning of each capillary. Data analyzes were performed using the Nano-Temper analysis software. The Kd constants between a protein and its ligand can be calculated using the standard fitting model derived from the law of mass action.

2.2.6.2 MST with Phosphatidylinositols

MST measurements with PI(4,5)P2 and PI(3,4,5)P2 were performed in the same way except standard capillaries replaced by hydrophobic capillaries (Nano-Temper Technologies) to prevent surface adsorption of phosphatidylinositols. Fluorescently labeled purified protein sample concentrations were fixed to 20 nM, and the highest

PI(4,5)P2 and PI(3,4,5)P2 concentrations were 584 to 510 µM respectively for sixteen-fold serial dilution in 20 mM HEPES pH 7.4, 150 mM NaCl, 0.05 % (v/v) Tween 20. MST was measured with 20% IR laser power, with high MST power settings.

2.2.7 Stability assay with NanoDSF

We perform a stability assay by using nano differential scanning fluorimetry (Nano DSF) Prometheus NT.48 (Nano-Temper Technologies) to characterize thermal unfolding, chemical denaturation and aggregation properties of purified proteins. In addition to this, we also performed a systematic buffer screen by using the thermostability assay to identify the best buffer condition, which provided the highest melting temperature (Tm) of purified proteins. We used a commercial buffer screen kits, JBScreen Thermofluor FUNDAMENTAL and SPECIFIC (Jena Bioscience), which covers a range of pH 4 to 10, and 0 to 1 M NaCl in 100 mM Citric acid : HEPES : CHES in molar ratio 2:3:4, and 100 mM Malonic acid : Imidazole : Boric acid in molar ratio 2:3:3. 15 µl of 1.84 mg/ml concentration of PRG2 pull-down was mixed with 15 µl of individual buffer by gently pipetting. Samples were spun down for 10 minutes at 4°C at 14,000g to eliminate any aggregates and loaded into standard capillaries (Nano-Temper Technologies) and subjected to NanoDSF for thermostability assays.

2.2.8 Dynamic Light Scattering (DLS)

We also performed DLS to measure the native monodispersity of PRG2ΔN eluates isolated from Lobstr cell strains with 0,1 mg/ml concentrations. At first, samples were spun down for 10 minutes at 4°C at 14,000g to eliminate any aggregates. Then the sample was filtered through a 0.22 µm nitrocellulose membrane. Then, 4 µl of sample was loaded on disposable cuvette, and measured with DynaPro NanoStar (Wyatt Technology).

2.2.9 Limited Proteolysis Assay

We performed limited proteolysis assay to characterize the inner core region of the protein complex. This was achieved by treating purified protein complexes with restricted concentration of proteolytic enzymes at limited incubation times and temperatures. During this process, enzymatic digestion takes place on the accessible hydration shells of the protein complexes, whilst the core of the complex remains intact.

For this experiment, we incubated 2 µg PRG2ΔN eluate from Lobstr strain with a serial dilution of Chymotrypsin from 1 to 0.004 mg/ml (Jena Bioscience). After 30 min incubation at 22°C in the water bath, reactions were quenched with 4x SLB supplemented with protease inhibitor cocktail (Roche). After boiling of the sample for 3 min at 95°C, we load them onto a 12% SDS PAGE and stained with Coomassie. Bands were excised for LC-MS analysis.

3. Results

The main scope of this thesis is to investigate the molecular structure of an integral membrane protein PRG2 in near-native complex. This section will guide the reader about the chronological order of an established pipeline for overexpression and purification of the native PRG2 protein complex. As each protein has an individual biophysical and biochemical entities, the aforementioned pipeline requires empirical multistep tests, systematic screening, optimizations, and quality control checks. Reader can also find the structural investigation of the purified complex via negative stained EM. In addition to that, the functionality of the purified PRG2 complex has been examined via biophysical assay as a proof of concept for maintaining native 3D protein structure.

Result section divided into two parts based on the hydrophobicity of the complex. In the first part, the reader can find expression, purification, and characterization of the amphipathic full PRG2 protein complex which was solubilized in detergent, and in the second part structurally highly disordered cytoplasmic domain of PRG2 protein which is hydrophobic and water-soluble.

3.1 Full length PRG2 and integral membrane domain of PRG2 (PRG2ΔC)

3.1.1 Selection of the eukaryotic cell line

Using the right expression system for structural biology is the key to success. Protein solubility, functionality, purification speed, and yield are often crucial factors to consider when choosing an expression system. The use of mammalian cell expression systems is becoming increasingly popular for the overexpression of mammalian membrane proteins for structural studies [70]. The main advantage of using mammalian cell lines to overexpress recombinant protein is to introduce accurate protein folding and maintain post-translational modifications (PTMs). This is also critical for membrane protein guidance to proper localization. After the protein synthesized in the endoplasmic reticulum, only properly folded and assembled proteins guide to the plasma membrane [71]. Therefore, using mammalian cell line provides indispensable PTMs and molecular chaperones, which are vital for maintaining the functionality of a protein.

There is no single recipe to follow for recombinant protein expression. Thus, it is essential to test various cell lines to find out an optimum protein expression level. The most commonly used mammalian cell lines in research for protein expression are human embryonic kidney 293 cells (HEK293) and African green monkey kidney fibroblast-like cell line COS [72-74]. COS are fibroblast-like cell lines derived from monkey kidney tissue. COS-1 cells are obtained by immortalizing CV-1 fibroblast cells with a version of the SV40 virus that can produce large T antigen but has a defect in genomic replication [75]. The HEK293 cell line was created when human embryonic kidney cells were transformed with fragments of adenovirus type 5 DNA, resulting in a stable cell line that replicated rapidly, tolerated reduced serum concentrations in the media, and was easily transfected [76]. A suspension adapted Human Embryonic Kidney 293 (HEK293S) cell line lacking N-acetylglucosaminyltransferase I (GnTI) has been successfully employed to overexpress a wide variety of mammalian membrane protein like tetraspanin, and of G protein-coupled receptors such as β2-adrenergic receptor and rhodopsin [77–80]. As this cell line has been successfully used to overexpress a wide variety of mammalian membrane proteins, we also tested HEK293S GnTI- cell line to overexpress integral membrane protein PRG2. Moreover, as PRG2 mainly expressed in neuron, I also tested mouse neuroblastoma N1E-115 cells to examine the expression level of transiently transfected PRG2.

Accordingly, tested N1E-115 neuroblastoma, COS-1, and HEK293S GnTI- cell lines to transiently transfect PRG2, cytoplasmic tail deleted PRG2ΔC and GFP control. Two days after transient transfection, cell pellets were lysed in 150 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 % (w/v) DDM supplemented with protease inhibitor. An equal volume of total cell lysate was loaded onto a 12% SDS PAGE for WB analysis. The result indicates us that HEK293S GnTI- cell has the highest expression level for both PRG2 and PRG2ΔC. I observed the PRG2 band between 70 to 100 kDa molecular weight marker, and PRG2ΔC under 55 kDa molecular weight marker (**Figure 8**). The expected molecular size of full PRG2 and integral membrane domain PRG2ΔC are 76 kDa and 42 kDa respectively. Consequently, conduct further studies to establish a pipeline to overexpress native PRG2 complex in HEK293S GnTI- cell line transiently.

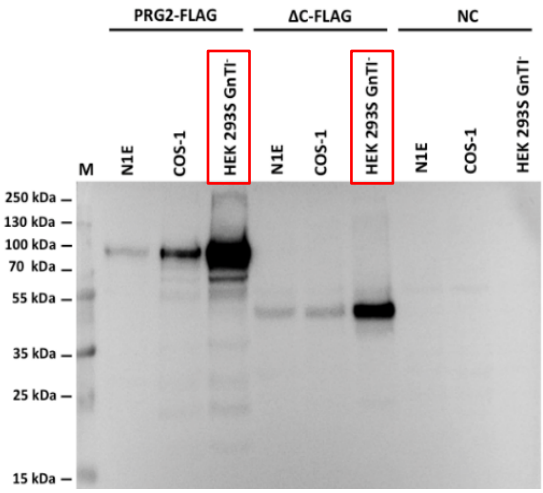


Figure 8: Selection of the eukaryotic cell line.
Western blot image of total cell lysate of FLAG tagged PRG2, PRG2ΔC, and GFP without FLAG tag were overexpressed in N1E-115 neuroblastoma, COS-1, and HEK293S GnTI$^-$ cells. HEK293S GnTI$^-$ cells had the highest expression level, indicated in red boxes. M, protein molecular weight marker. NC, negative control with GFP expression.

3.1.2 Selection of the transient transfection method

NIH National Library of Medicine defines transfection as the process of deliberate introducing naked or purified nucleic acids into eukaryotic cells to produce genetically modified cells. The main purpose of this analytical tool to study the function of a gene and use cells as a manufacturing factory to produce recombinant protein [81]. There are two methods exist in cells either stably or transiently depending on the nature of the genetic materials [82].

For stable transfection, introduced genetic materials are not only entered into cells, but also integrated into the host genome by homologous recombination, and sustain transgene expression even after host cells replicate [83]. Transgene usually incorporated with the antibiotic resistance gene to select transfected cells. Also, genes can be directed in the specific locus in the genome to land in by using site-directed recombination enzymes like Cre and Flp recombinases [83]. Stable cell line generation has several bottlenecks. The frequency of stable genome integration upon transfection with plasmid vectors is low. Integrated transgene can randomly land on and generate negative positional effect on genome, can alternate or halt specific gene expression. Most integrated transgenes will be silenced or land on untranslated locus [84].

Whereas, for transient transfection, cells express the foreign gene but do not integrate it into their genome. Thus, genetic materials can be lost by environmental factors and cell division. In this method, transgene incorporates with antibiotic resistance reporter gene for the elimination of untransfected cells.

The choice of stable or transient transfection depends on the objective of the experiment. At first glance, stable cell line creation looks advantageous for long-term projects, however, the risk of unexpected mutagenesis, the silence of gene expression, alteration of genomic integrity of a stable cell line by being in culture for long passages can be critical for protein structure. Therefore, in this study, the transient transfection method has been chosen.

Chemical transfection methods are the most widely used methods and were the first to be used to introduce foreign genes into mammalian cells [85]. Chemical methods commonly use cationic polymer, calcium phosphate, cationic lipid [85–87]. The underlying principle of chemical methods are similar. Positively charged cationic chemicals form complexes with the negatively charged DNA molecule, allowing their fusion to cellular membranes. This complex is able to enhance gene delivery in diverse cell types and tissues. The exact mechanism of how nucleic acid/chemical complexes pass through the cell membrane is unknown, but it is believed that endocytosis and phagocytosis are involved in the process [88]. Transfected DNA must be delivered to the nucleus to be expressed and again the translocation mechanism to the nucleus is not known.

In this section, commercially available cationic lipid-based transient transfection reagents Turbofect and Lipofectamine 2000 from Thermo Fisher Scientific have been tested. In addition to them, I prepared cationic polymer polyethylenimine (PEI) transient transfection reagent [89].

After the selection of cell line, I tested the abovementioned three positively charged chemicals to express PRG2, and PRG2ΔC in HEK293S GnTI- cells using PEI, Turbofect and Lipofectamine 2000 for transient transfection efficiencies and protein expression levels. Cells were cultured and harvested, and total cell lysates were extracted. An equal

volume of total cell lysate was loaded onto 12% SDS PAGE for WB with the FLAG antibody. The result indicates that all three transfection methods were efficient (**Figure 9**). For small-scale optimization experiments, all three methods can be used, however, the PEI method is cost-effective for large-scale adherent cells and suspension cultures. As a result, we find PEI transient transfection procedure to be more cost-effective than lipid-based commercial transient transfection methods, and quicker than the more traditional route of generating stable cell lines.

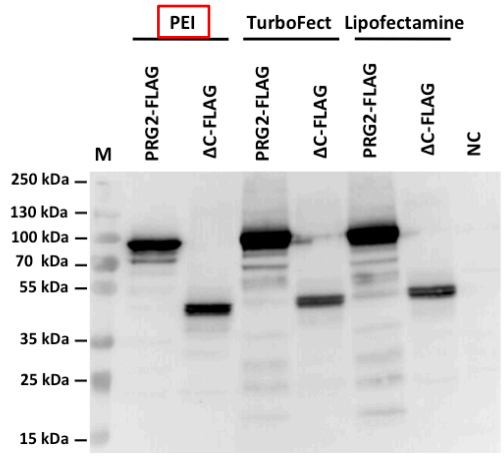


Figure 9: Selection of the transient transfection method.
Western blot image of FLAG tagged PRG2 and PRG2ΔC overexpression in HEK293S GnTI- cells with PEI, Turbofect and Lipofectamin 2000 transient transfection. All three transfection methods were efficient. Chosen transient transfection method PEI is indicated in the red box. M, protein molecular weight marker. NC, negative control with GFP expression.

3.1.3 Determination of the transient transfection ratio

Apart from selected cell type and preferred transfection method, actual transfection efficiency is primarily related to nucleic acid versus transient transfection reagent ratio (v/v), temperature of the solution, pH level, confluency of cells, and the cell membrane condition [88]. Meticulous attention is required to establish the chemical transfection-to-DNA ratio for each cell type and DNA molecule. An excess of cationic transfection reagent could be toxic for a particular cell line. An empirical observation is needed to find out a low amount of reagent as a vehicle to carry a sufficient amount of plasmid DNA into the cell. Additional care is needed during the coupling procedure, as it needs to be done very slowly with the addition of progressive amounts of DNA to wrap each PEI particle with the largest amount of plasmid DNA. In this way, optimal transient transfections can be facilitated.

We tested 1:1 and 3:1 (v/v) ratio of PEI:DNA to transfect 80% confluent HEK293S GnTI- cells with GFP control vector as a marker to evaluate transient transfection efficiency. Cells were cultured for two days post-transfection and imaged with a fluorescent microscope. Our findings suggest to us that 3:1 (v/v) ratio of PEI:DNA has higher transfection efficiency (**Figure 10**).

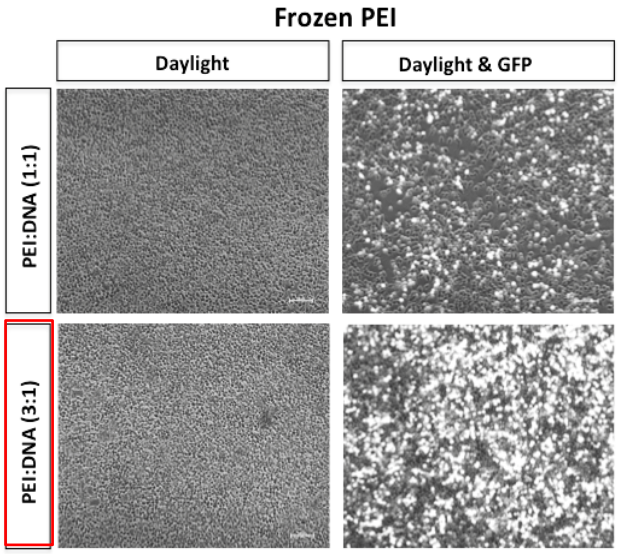


Figure 10: Determination of the transient transfection ratio.
Phase contrast and fluorescent microscopic images of 1:1 and 3:1 (v/v) ratio of PEI:DNA to express GFP in HEK293S GnTI$^-$ cells. 3:1 (v/v) ratio of PEI:DNA shows higher transfection efficiency, indicates in the red box (10X objective, scale bar = 100 µm).

3.1.4 Determination of the incubation time

The optimal time for harvesting transiently transfected HEK293S GnTI- cells for recombinant protein expression should be determined empirically for each protein. Several factors can alter the answer. Notably, the strength of the overexpression promoter of the plasmid, protein half-life, and any required post-translational modifications can yield different answers. A conventional way to determine an optimum time window for high protein yield without aggregation nor degradation is to assess small scale overexpression analysis using a time course to determine the optimal time for each protein of interest. This timeline is usually between 12 hours up to 72 hours after post-transfection. Therefore, after the selection of the cell line, and the transient transfection method, reagent with an optimum transfection ratio, I tested 24, 48, and 72

hours duration of incubation times to find out the highest protein expression level of PRG2 and PRG2ΔC. I observed that two days for PRG2 and one day for PRG2ΔC are the best for harvesting cells after the transient transfection (**Figure 11**).

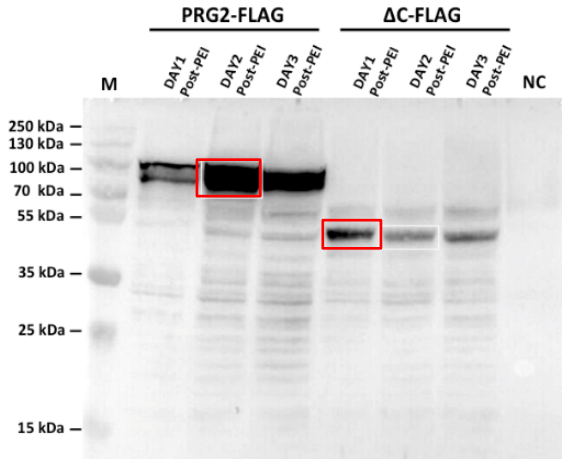


Figure 11: Determination of the incubation time.
Western blot image of FLAG tagged PRG2 and PRG2ΔC overexpression in HEK293S GnTI$^-$ cells with PEI transient transfection. Cell were harvested at day one, two, and three after transient transfection. Two days for PRG2, and one day for PRG2ΔC incubation after the transfection have the highest level of expressions which were shown in red boxes. M, protein molecular weight marker. NC, negative control with GFP expression.

3.1.5 Selection of the optimum growth condition

The selected method of growth condition is strategically important to achieve the desired level of protein yield. In example, using culture flasks yield in a few micrograms of purified recombinant proteins from cultures of adherent cells. With an expanding need for milligrams of protein for structural biology, the static tissue culture dish has been replaced by large-scale systems like cell suspension. Therefore, researchers who need high concentration of transgenic protein are encouraged to switch to a suspension-adapted cell line [90].

Adaptation from an adherent monolayer to suspension cultures is required several weeks of passaging cells with stepwise blending of adherent cell culture medium DMEM with suspension growth medium [91]. In each subsequent passage, the ration of suspension media volume increased. Serum-free cell suspension growth medium which lacks the calcium and contains a non-ionic detergent to prevent cell adherence and to

protect cells against high levels of shear stress introduced by sparging and stirring in the suspension culture [92].

Chosen HEK293S GnTI- cells are suitable to grow as an adherent monolayer in disposable flask, and cell suspension under the small-scale volume (less than 1000 ml) in a magnetic steering bottle in the traditional laboratory-scale tissue culture [70].

In this section, I have tested both growth conditions, adherent cell culturing in the flask, and cell suspension in the stirring bottle with commercial growth medium FreeStyle 293 Expression Medium, and 293 SFM II (Thermo Fisher Scientific). After PEI transient transfection to overexpress PRG2, cells were cultured for two days, harvested, and total cell lysates were extracted. An equal volume of total cell lysate with or without 1:10 dilution (v/v) with wash buffer was loaded onto 12% SDS PAGE for WB with the FLAG antibody. Our result indicates that there was no PRG2 expression in the cell suspension system (**Figure 12**). Although I tested ideal growth mediums, various rpm rates, and stepwise adaptation of cells to suspension medium, inevitable drastic cell death and cell aggregation in suspension growth conditions were observed over the course of incubation time. Thus, I chose to proceed with adherent cell culture method to overexpress PRG2 by scaling up the number of flasks to increase the final protein yield (**Figure 12**). Besides, aggregation of PRG2 detected when it was loaded in higher concentrations as observed in the above 250 kDa molecular weight marker (**Figure 12**). This is a clear indication for PRG2 is in favor to make a bigger multimeric complex in dose-dependent manner. Hence, keeping purified PRG complex in low concentration is vital to prevent undesired aggregations.

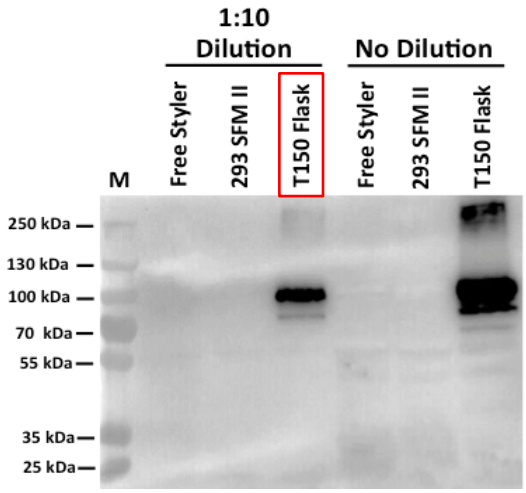


Figure 12: Selection of the optimum growth condition.
Western blot image of FLAG tagged PRG2 overexpression in HEK293S GnTI$^-$ cells with PEI transient transfection in bioreactor as cell suspension or in flask as adherent cell. Cell were harvested at two days after transient transfection. Adherent cell in flask condition is effective to express PRG2. M, protein molecular weight marker.

3.1.6 Initial buffer optimization

Once I have accomplished the production of recombinant mouse PRG2 protein in HEK293S GnTI- cells, and performed immunoprecipitation approach successfully by using C-terminal FLAG tag, the next step is further purification, which is essential to isolate PRG2 protein from the surrounding contaminants while keeping it soluble, active, and in the near-native complex. The amount of total protein that goes into a solution under specified conditions refers to protein solubility [93]. Protein structure, amino acid composition, pH, ionic strength, solution composition, temperature, duration of extraction and many other extrinsic factors determine protein solubility.

Before embarking on extensive purification studies, we need to perform an initial buffer condition screen by testing a range of pH and salt concentration to find out optimum buffer formulation to keep protein complex soluble, biologically active, and in native complex form.

The pH range of a buffer is based on its pKa, defined as the pH at which 50% of the molecules are in their acidic form, and 50% are in their basic form. A general rule regarding buffers is that the pH of the buffer solution should be within 1.0 pH unit of the

pKa to provide appropriate buffering capacity [94]. This ensures that there is a sufficient amount of the molecule in both its acidic and basic forms to neutralize the solution in case of H+ or OH- influx. Thus, buffers prevent pH changes that could negatively affect protein stability. Given the fact that proteins have minimum solubility in solutions at the pH that corresponds to their pI, chose of 1 or 2 pH unit of the pI is necessary to avoid protein precipitation.

The mechanism of the ionic strength effect on protein solubility involves solvation, electrostatic and salting in and salting-out phenomena [95]. Low concentrations of neutral salts at molarities of the order of 0.1-1.0 M may increase the solubility of proteins. In salt solutions, protein solubility initially increases (salting in) and after a maximum of solubility, it starts to decrease (salting out). An increase in salt concentration greater than 1.0 M caused a decrease in protein solubility. Water molecules are strongly bound to the salt and there is competition between the salt ions and the protein molecules for the water molecules [96].

In this section, various pH and salt conditions have tasted in Tris buffer to extract native PRG2 complex by immunoprecipitation (IP). In this assay, I have tested 100 mM Tris, 2 % (w/v) n-Dodecyl−β−D-Maltopyranoside (DDM) detergent in lysis buffer, and 0.01 % DDM to maintain above CMC in wash and elution buffer with a range of pH level of 5, 6, 7, 8, 9, and salt concentration at 100, 200, 300, 400, and 500 mM. Cells were cultured for two days, harvested, and total cell lysates were extracted. FLAG-tagged PRG2 pull downed with FLAG coated beads and eluted with 3xFLAG peptide. An equal volume of IP eluate was loaded onto a 12% SDS PAGE for WB with FLAG antibody and silver staining. Our results show that there is no difference in the recovery of PRG2 in any pH levels and salt concentrations, accept at pH 5 (**Figure 13**). Since the isoelectric point of PRG2 is 5.7, we observed less recovery duo to possible protein precipitation. As there is no substantial effect of salt concentration and pH level on eluate yield and band profile, I proceeded with pH 7.4 and 150 mM salt concentration.

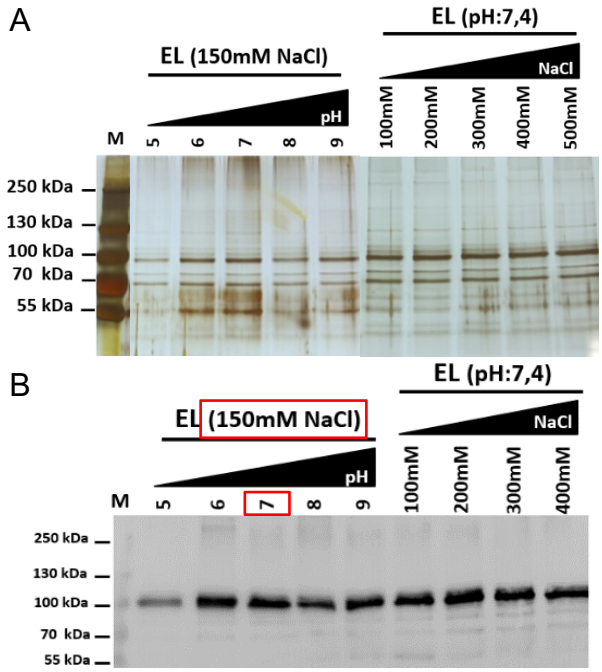


Figure 13: Initial buffer optimization.
A Silver stain and **B** western blot images of FLAG tagged PRG2 eluate after FLAG immunoprecipitation. Except at pH 5, all buffer conditions yielded similar level of PRG2 eluate, and protein band profile. Chosen salt concentration and pH level were indicated in red boxes. M, protein molecular weight marker. EL, eluate.

3.1.7 Protein solubility test of PRG2

One of the most critical and initial steps in membrane protein sample preparation is the solubilization of the membrane to extract the protein of interest in an active form in the native complex with maintaining physiological structure for purification and further structural analysis [97]. While detergents are necessary to extract and maintain the protein outside of the lipid environment, they are also denaturants. Thus, using mild and relatively non-denaturing nonionic detergents are vital. These detergents maintain protein-protein interactions while braking lipid–lipid interactions and lipid-protein interactions without affecting the protein's structural features, such that it can be isolated in its biologically active form [98]. Such as n-dodecyl-β-d-maltoside (DDM) is the most commonly used detergent in membrane protein purification [99]. Therefore, we used DDM to establish our purification pipeline optimization. At this stage, we investigate the solubilization power of this detergent by loading fractions of sample from each step of IP assay onto 12% SDS PAGE. WB results show that FLAG pulldown assay works

effectively, however, the vast majority of PRG2 was remained in the insoluble pellet fraction (**Figure 14**). This is a clear indication that DDM has suboptimal solubilization strength to strip PRG2 complex from HEK293S GnTI- cell membrane. As a result, I proceed with screening commercially available mild detergents to find out an alternative detergent with higher solubility strength.

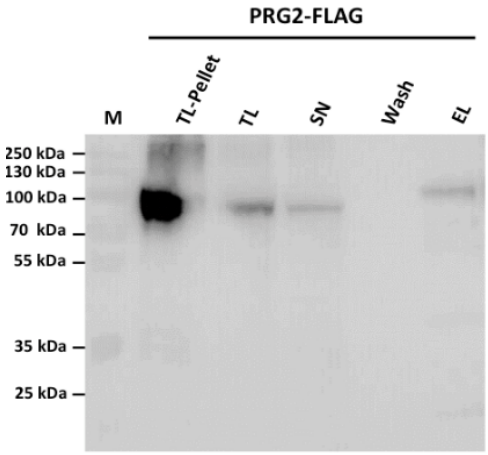


Figure 14: PRG2 solubility test.
Western blot images of FLAG tagged PRG2 eluate after FLAG immunoprecipitation. Majority of PRG2 was in the insoluble pellet fraction. TL-Pellet, pellet fraction after the initial centrifugation to eliminate insoluble proteins and cell debris. M, protein molecular weight marker. TL, total cell lysate. SN, unbounded proteins after FLAG IP. Wash, third washing step of FLAG agarose to eliminate unspecific interaction. EL, eluate.

3.1.8 Detergent screening of PRG2

No single detergent can be the best detergent suitable for solubilization of all membrane proteins, as the molecular principles of solubilization remain unclear [100]. As finding out the most suitable detergent requires on empirical, time-consuming try and error approach, a systematic detergent screening assay is needed. In this section, I established a systematic screening method to test 88 state of the art non-ionic and zwitterionic mild detergents (**Table 7**) in 96 well plate scale at once in order to increase solubilization level of native PRG2 complex.

This assay assumes that non-solubilized membranes remain in the total cell lysate (TL) fraction. Insoluble pellet fraction was homogenized by Dounce homogenizer with ice-cold 100 mM Tris pH 7.4, 150 mM NaCl, aliquoted, and solubilized with 1 % (v/v) of each detergent individually by gentle pipetting up and down. The insoluble fraction was

removed by centrifugation, and an equal volume of supernatant with solubilized PRG2 was spotted on nitrocellulose membrane, blocked with BSA, and coupled with FLAG antibody for detection. Densitometric analysis of dot blot shows that Fos-Choline-14 is fifty-fold higher solubilization rate than DDM (**Figure 15**). Therefore, I concluded to switch from DDM to Fos-Choline-14 detergent for solubilization.

The Fos-Choline detergent family has a well-known positive effect in preventing aggregation. However, currently, it is a debate in the field whether they have denaturing effects on membrane protein structure [101-103]. As this detergent is controversial in the field, I decided to use it merely at the initial extraction and solubilization step in our purification pipeline.

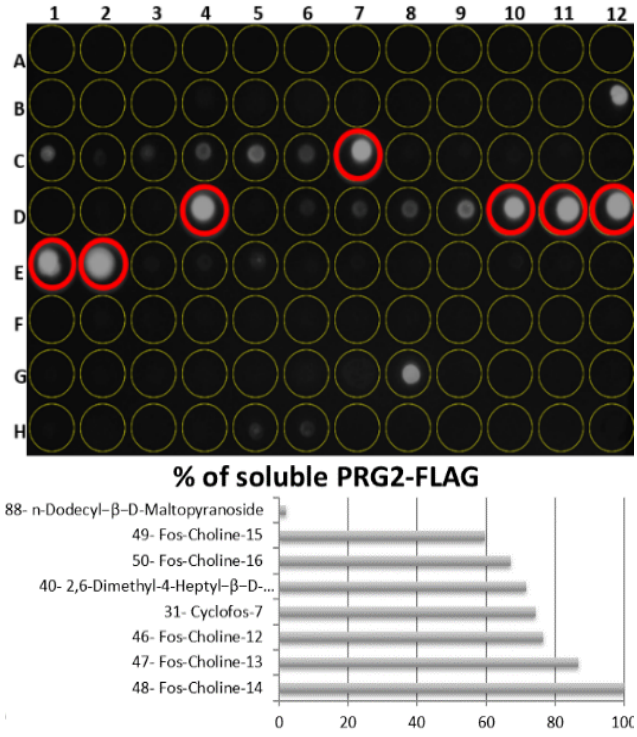


Figure 15: PRG2 detergent screen.
Dot blot images of insoluble fractions of FLAG tagged PRG2 resolubilized with a range of mild detergents. Seven of the highest solubilizing detergents are depicted in red circle, and densitometric values shown in graph including DDM control. Fos-Choline-14 has the highest recovery yield among the others.

Table 7: List of detergents and densitometric analysis result.

Maximum exposure level set as 100%, and all signals were normalized accordingly.

#	Detergent	% of soluble PRG2-FLAG
48	Fos-Choline-14	100
47	Fos-Choline-13	87
46	Fos-Choline-12	77
31	Cyclofos-7	74
40	2,6-Dimethyl-4-Heptyl−β−D-Maltopyranoside	72
50	Fos-Choline-16	67
49	Fos-Choline-15	60
24	Deoxycholic Acid Sodium Salt	49
80	Sodium Dodecanoyl Sarcosine	33
45	Fos-Choline-11	22
29	Cyclofos-5	19
25	Sodium Cholate	18
28	Cyclofos-4	11
44	Fos-Choline-10	11
53	Fos-Choline-Unsat-11-10	10
30	Cyclofos-6	10
43	Fos-Choline-9	9
27	Cyclofos-3	8
42	Fos-Choline-8	5
79	n-Dodecyl-N,N-Dimethylglycine	4
78	n-Decyl-N,N-Dimethylglycine	4
26	Cyclofos-2	4
52	Fos-Choline-ISO-11	3
34	CYMAL-2	3
58	n-Hexyl−β−D-Glucoside	3
33	CYMAL-1	3
54	Fos-Mea-8	3
67	n-Undecyl−α−D-Maltoside	3

57	PMAL-C8	3
66	n-Decyl−β−D-Maltoside	3
55	Fos-Mea-10	2
68	n-Undecyl−β−D-Maltoside	2
51	Fos-Choline-ISO-9	2
18	Anzergent 3-12	2
70	n-Tridecyl−β−D-Maltoside	2
35	CYMAL-3	2
69	n-Dodecyl−α−D-Maltoside	2
59	n-Heptyl−β−D-Glycoside	2
81	C8E6	2
16	Anzergent 3-8	2
83	C10E5	2
17	Anzergent 3-10	2
65	n-Decyl−α−D-Maltoside	2
32	CYGLU-3	2
21	Big CHAP, Deoxy	2
56	MEGA-8	2
76	n-Undecyl−β−D-Thiomaltoside	2
88	n-Dodecyl−β−D-Maltopyranoside	2
77	n-Dodecyl−β−D-Thiomaltoside	2
82	C12E8	2
87	n-Dodecyl-N,N-Dimethylamine-N-Oxide (DDAO)	2
86	n-Tetradecyl-N,N-Dimethylamine-N-Oxide (TDAO)	2
64	n-Nonyl−β−D-Maltoside	2
71	n-Tetradecyl−β−D-Maltoside	2
60	n-Octyl−β−D-Glucoside	1
84	C8E4	1
62	n-Hexyl−β−D-Maltoside	1
19	Anzergent 3-14	1
72	n-Heptyl−β−D-Thioglucoside	1
75	n-Decyl−β−D-Thiomaltoside	1

41	2-Propyl-1-Pentyl Maltopyranoside	1
63	n-Octyl−β−D-Maltoside	1
36	CYMAL-4	1
39	CYMAL-7	1
74	n-Nonyl−β−D-Thiomaltoside	1
20	Big CHAP	1
9	Anapoe-C12E9	1
8	Anapoe-C12E8	1
85	Sucrose Monododecanoate	1
22	CHAPS	1
38	CYMAL-6	1
23	CHAPSO	1
61	n-Nonyl−β−D-Glucoside	1
73	n-Octyl−β−D-Thiomaltoside	1
11	Anapoe-C13E8	1
12	Anapoe-X-100	1
6	Anapoe-C10E6	1
15	Anapoe-X-405	1
10	Anapoe-C12E10	1
3	Anapoe-35	1
5	Anapoe-80	0
7	Anapoe-C10E9	0
4	Anapoe-58	0
13	Anapoe-X-114	0
14	Anapoe-X-305	0
37	CYMAL-5	0
2	Anapoe-20	0
1	Anameg-7	0

3.1.9 Detergent exchange with Amphipol

Next an analytical Size Exclusion Chromatography (SEC) was performed to check the quality of the sample to evaluate protein stability, complex formation, tendency to aggregate and quantity of aggregates.

After the IP with Fos-Choline-14, eluate was concentrated, and loaded onto a pre-equilibrated SEC with 100 mM Tris pH 7.4, 150 mM NaCl, 0.01 % (w/v) Fos-Choline-14 for purification. I encountered that the sample was precipitated during the SEC process, therefore, I could not collect any PRG2 after the SEC. As was mentioned in the previous section, the Fos-Choline detergent family is ideal for native membrane protein extraction from cellular lipid bilayer mildly, yet it is stringent enough to alternate physiological molecular structure. Given the fact that I consider finding out an alternative solubilization reagent which provides regaining and maintenance of functional physiological molecular structure of PRG2 protein complex.

Recently, high-resolution structures of a number of different membrane proteins were obtained using cryo-EM from samples containing amphipol A8-35 [104-107]. That carries hydrophobic and hydrophilic groups and is effective at keeping membrane proteins water-soluble [108]. They also exhibit a high affinity for hydrophobic particles. Because of these unique properties, using amphipols to surround membrane proteins results in minimal extra density around the proteins' hydrophobic domains compared to detergents.

Consequently, Fos-Choline-14 used in sample preparation was exchanged with the amphipathic polymer, amphipol A8-35. Subsequently, after collection of FLAG-tagged PRG2 eluate, detergent has been exchanged with amphipol, concentrated and loaded onto a SEC which was equilibrated with 100 mM Tris pH 7.4, 150 mM NaCl. We could observe a differential peak on SEC chromatogram from B4 to B12 fractions (**Figure 16B**) comparing detergent-solubilized SEC. In order to evaluate the presence of PRG2 in SEC fractions, we performed dot blot analysis with an equal volume of each fraction was spotted on nitrocellulose membrane, blocked with BSA, and coupled with FLAG antibody for detection. I could only detect FLAG signals from amphipol exchanged PRG2 at B4 to C7 fractions (**Figure 16B**). These results displayed a high potential for

the utilization of this amphipolic molecule for the stabilization and structural analysis of the native PRG2 complex.

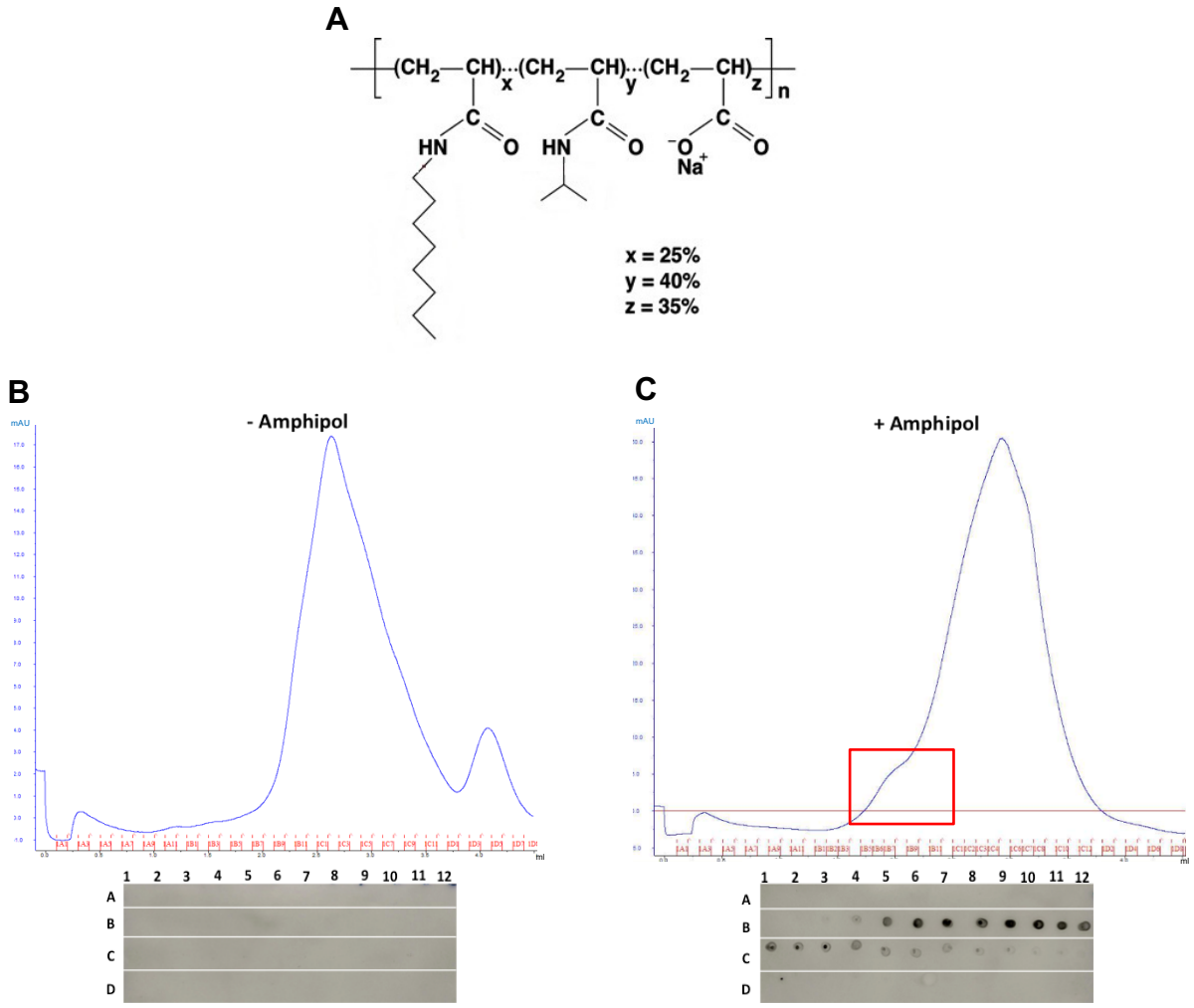


Figure 16: Detergent exchange with Amphipol.
A Chemical formulation of amphipol A8-35, and Dot blot images of size-exclusion chromatography fractions of FLAG tagged PRG2 eluate with **B** Fos-Choline-14 versus **C** Amphipol. Red box shows PRG2 positive SEC fractions corresponding to 280 nm absorption wavelength chromatogram peak.

3.1.10 Systematic buffer and additive screening of PGR2 with NanoDSF

Protein homogeneity, stability, and solubility are factors that are highly correlated with the native membrane protein purification for structural investigation Buffer condition with specific pH and ionic strength, and additive mono and divalent cations providing the local solvent environment keeps the protein stable and minimizes aggregation.

In this section, it was performed a systematic buffer and additive screen by using fluorescence dye-free, NanoDSF thermal stability assay to find out the ideal buffer and supplementary ion additives which could increase the thermostability of purified native PRG2 complex. In this assay, protein's melting temperature (Tm) is used as a reporter for protein stability and determined by a thermal shift assay monitoring the unfolding of a protein in a temperature-dependent manner. The higher the Tm, the higher is the thermostability of the protein in that specific environment.

I tested buffers with a range of pH levels from 4 to 10, and 0 to 1 M NaCl in 100 mM Citric acid:HEPES:CHES in molar ratio 2:3:4, and 100 mM Malonic acid:Imidazole:Boric acid in molar ratio 2:3:3. FLAG-tagged PRG2 pull downed with FLAG coated beads, and eluted with 3xFLAG peptide in 100 mM Tris pH 7.4, 150 mM NaCl, 0.01 % (w/v) Fos-Choline-14. Detergent exchanged with amphipol A8-35. An equal volume of purified protein mixed with each tested buffer and subjected for NanoDSF for thermostability assay. Tm value of PRG2 in Tris buffer is 47.9 °C. We observed the highest Tm values of 48.9 °C and 48.8 °C at pH 7 in 30 mM HEPES buffer with 50 or 150 mM NaCl concentration with or without 20 mM LiCl respectively (**Table 8**). Based on this result, I decided to conduct upcoming experiments in 20 mM HEPES, 150 mM NaCl at pH 7.4 buffer condition instead of 100 mM Tris pH 7.4, 150 mM NaCl.

Table 8: Systematic Buffer Screen with NanoDSF.
Table shows melting temperature of pull downed PRG2 values in a range of buffer conditions. Dark green color shows the highest Tm values in this assay.

100mM Malonic acid:Imidazole:Boric acid

pH	4	4	4	4	4
NaCl	150mM	150mM	150mM	150mM	150mM
Additive		20mM MgSO4	20mM FeCl3	20mM ZnCl2	20mM MnCl2
Tm	29.0	29.2	45.0	28.2	30.1
pH	5	5	5	5	5
NaCl	150mM	150mM	150mM	150mM	150mM
Additive		20mM MgSO4	20mM FeCl3	20mM ZnCl2	20mM MnCl2
Tm	43.3	42.3	43.8	43.4	44.5
pH	6	6	6	6	6
NaCl	150mM	150mM	150mM	150mM	150mM
Additive		20mM MgSO4	20mM FeCl3	20mM ZnCl2	20mM MnCl2
Tm	48.2	48.1	48.2	46.7	47.9
pH	7	7	7	7	7
NaCl	150mM	150mM	150mM	150mM	150mM
Additive		20mM MgSO4	20mM FeCl3	20mM LiCl	20mM KCl
Tm	48.9	48.5	48.2	48.8	48.3
pH	8	8	8	8	8
NaCl	150mM	150mM	150mM	150mM	150mM
Additive		20mM MgSO4	20mM FeCl3	20mM LiCl	20mM KCl
Tm	48.2	48.4	47.8	48.4	48.5
pH	9	9	9	9	9
NaCl	150mM	150mM	150mM	150mM	150mM
Additive		20mM MgSO4	20mM FeCl3	20mM LiCl	20mM KCl
Tm	47.7	48.2	47.8	47.8	48.4
pH	10	10	10	10	10
NaCl	150mM	150mM	150mM	150mM	150mM
Additive		20mM MgSO4	20mM FeCl3	20mM LiCl	20mM KCl
Tm	47.5	37.7	47.8	47.9	48.1

100mM Citric acid:HEPES:CHES

4	4	4	4	4
150mM	150mM	150mM	150mM	150mM
	20mM MgSO4	20mM CaCl2	20mM ZnCl2	20mM MnCl2
35.5	35.2	36.6	33.1	25.0
5	5	5	5	5
150mM	150mM	150mM	150mM	150mM
	20mM MgSO4	20mM CaCl2	20mM ZnCl2	20mM MnCl2
43.2	43.7	43.0	40.5	42.4
6	6	6	6	6
150mM	150mM	150mM	150mM	150mM
	20mM MgSO4	20mM CaCl2	20mM ZnCl2	20mM MnCl2
47.2	47.4	46.4	47.4	47.2
7	7	7	7	7
150mM	150mM	150mM	150mM	150mM
	20mM MgSO4	20mM CaCl2	20mM LiCl	20mM KCl
48.4	48.4	48.1	48.4	48.2
8	8	8	8	8
150mM	150mM	150mM	150mM	150mM
	20mM MgSO4	20mM CaCl2	20mM LiCl	20mM KCl
48.0	47.7	47.7	48.0	47.8
9	9	9	9	9
150mM	150mM	150mM	150mM	150mM
	20mM MgSO4	20mM CaCl2		20mM KCl
47.6	47.6	47.8	47.3	48.0
10	10	10	10	10
150mM	150mM	150mM	150mM	150mM
	20mM MgSO4	20mM CaCl2	20mM LiCl	20mM KCl
46.8	46.5	47.9	46.8	46.7

100mM Malonic acid:Imidazole:Boric acid

pH	4	4	4	4	4	4
NaCl		50mM	125mM	250mM	500mM	1M
Additive						
Tm	47.7			29.1	28.7	
pH	5	5	5	5	5	5
NaCl		50mM	125mM	250mM	500mM	1M
Additive						
Tm	40.3	41.3	40.8	42.4	42.1	41.6
pH	6	6	6	6	6	6
NaCl		50mM	125mM	250mM	500mM	1M
Additive						
Tm	47.6	47.6	47.8	47.9	47.7	47.2
pH	7	7	7	7	7	7
NaCl		50mM	125mM	250mM	500mM	1M
Additive						
Tm	48.1	48.9	48.6	48.6	47.9	47.7
pH	8	8	8	8	8	8
NaCl		50mM	125mM	250mM	500mM	1M
Additive						
Tm	Handling Mistak	48.4	48.5	48.4	47.9	47.4
pH	9	9	9	9	9	9
NaCl		50mM	125mM	250mM	500mM	1M
Additive						
Tm	48.4	48.2	47.7	48.1	47.7	46.2
pH	10	10	10	10	10	10
NaCl		50mM	125mM	250mM	500mM	1M
Additive						
Tm	48	47.5	47.5	47.3	47	46.1

3.1.11 Detection of the size and stoichiometry of native PGR2 complex

After establishing the pipeline to purify near-native PRG2 complex, we need to evaluate the molecular weight of the complex, homogeneity of the complex, and stoichiometry if there is multiple population of multimeric complexes.

To do so, we choose to perform native polyacrylamide gel electrophoresis (Native PAGE) analysis where proteins that are kept in their folded state with conserved physical protein-protein interaction. In this method, electrophoretic mobility depends not only on the charge-to-mass ratio but also on the physical shape and size of the protein [109,110]. Hence, it is a very common method to determine native protein complex masses and oligomeric states [111] and the stoichiometry of multiprotein complexes [112].

I performed Blue Native (BN) PAGE WB. FLAG-tagged PRG2 solubilized with Fos-Choline-14, pull downed with FLAG coated beads, and eluted with 3xFLAG peptide. Detergent exchanged with amphipol, and the eluate was concentrated and loaded onto a 3-12% Native PAGE Bis-Tris Gel for Coomassie stain, and WB with FLAG antibody. I observed multiple levels of complex formation in addition to the most abundant monomeric form of PRG2. Other distinct classes were observed in dimeric, tetrameric, and octameric stages of PRG2 multimer classes under the native condition (**Figure 17**).

Multimerization of purified native PRG2 complex is an expected result. Multiple membrane domains of integral transmembrane proteins involved in oligomerization in the cell membrane, providing functional implications in the transduction of signaling events and resulting cellular responses. For example, oligomerization between 7-transmembrane GPCRs is acknowledged to modulate the pharmacological characteristics of the receptors and influence their coupling to G proteins [111,112]. Likewise, transient receptor potential (TRP) proteins, which are 6-transmembrane ion channels assemble in homo- and hetero-tetrameric complexes to enhance ion permeation property [113,114].

This result also indicates that our extraction and purification pipeline is mild enough to keep the complex intact, yet strong enough to extract it from the cellular membrane. As I observed native PRG2 forms higher order of homomeric complexes, subsequent physical separation step is indispensable to distinguish each of these multimeric complex classes. In line with our previous findings in section 3.1.5, **Figure 12**, I have seen aggregates on the top of the gel, where the high concentration of PRG2 complex

gain tendency to aggregates. Thus, keeping purified PRG2 complex in low concentration is critical.

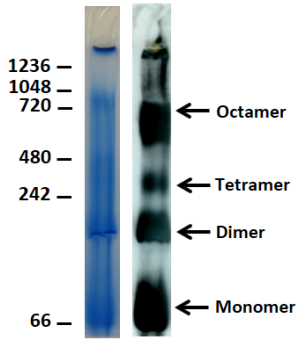


Figure 17: Detection of native PGR2 complex size and stoichiometry.
Coomassie stained Blue Native PAGE and WB image of PRG2 eluate shows a monomeric PRG2 as well as range of native PRG2 complex multimers.

3.1.12 Detection of native PGR2 complex size and stoichiometry after SEC

As mentioned in the previous section, multiple classes of native PRG2 complexes needed to be separated from each other. There are two main tools exist to achieve this goal, namely analytical SEC, and analytical ultracentrifugation (AUC). Both methods can simultaneously conduct the separation and analysis of protein complexes in solution by maintaining their native structure [115-117]. As we have analytical SEC in-house, I use this method to attempt for the separation of distinct clusters of native PRG2 complex classes. Next it was performed our purification pipeline by solubilizing FLAG-tagged PRG2 with Fos-Choline-14, subsequently pull downed with FLAG coated beads, and eluted with 3xFLAG peptide. Detergent exchanged with amphipol, and the eluate was concentrated, and 200 µg protein loaded onto a SEC for the final purification step to see if I could separate individual classes of native PRG2 multimers. PRG2 signal was started on B4 fractions and ended on C4 fraction as we observed in dot blot analysis (**Figure 18A**). Since I could distinguish a shoulder peak starts at A4 fraction and maximizes at B7 fraction, we focused on these fractions for further analysis. B3 to B12 fractions were loaded onto 3-12% BN-PAGE for WB analysis to see which multimeric classes present in these individual fractions. WB shows us a smear of complex from dimeric to tetrameric complex sizes (**Figure 18B**). As I used all B6 fraction to concentrate and reload in SEC, to test stability and if possible further separation to minimize smear pattern on BN-PAGE analysis, I could not use this fraction for dot blot and WB analysis. Unfortunately, the

sample concentration of an individual B6 fraction was under the detection limit of SEC, therefore I could not obtain a chromatogram peak with B6 fraction which was loaded on SEC (data has not been shown here). This result is an indication for the destabilization of the complex as we could not detect a single Gaussian distribution curve at B6 fraction. The main finding of this experiment is that native PRG2 complexes are not stable enough as they create a broad range of molecular complex size smear after SEC. In order to tackle this bottleneck, I tested chemical crosslinking with formaldehyde and glutaraldehyde before subjecting a sample in gel filtration, which leads to precipitation of native PRG2 complex (data has not been shown here). Additionally, I performed intramolecular chemical crosslinking combined with density gradient ultracentrifugation method GraFix (Gradient Fixation) to stabilize and separate PRG2 complexes [118]. In this method, we could not collect distinct classes of PRG2 complex in any fractions, instead, we observed that all fractions were positive for PRG2 in dot-blot analysis (data has not been shown here).

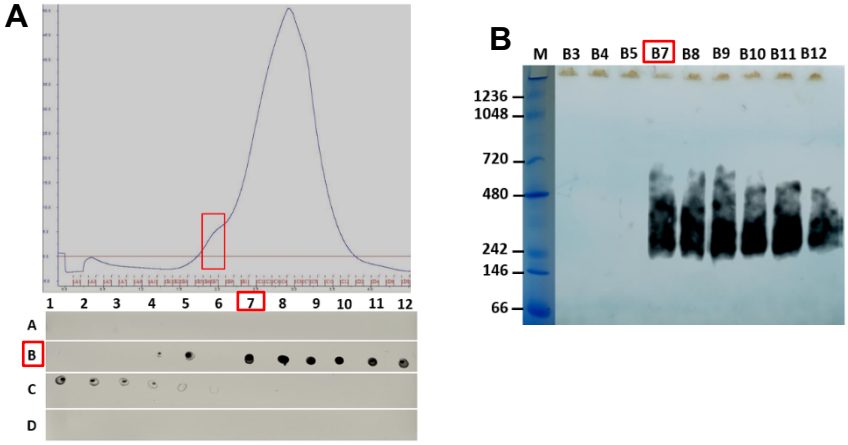


Figure 18: Detection of native complex sizes after SEC.
A SEC chromatogram shows 280 nm wavelength of absorption. Dot blot analysis shows PRG2 signal for each fraction. **B** Native PAGE WB image shows stages of multimers. B7 fraction is depicted in red boxes, which is the maxima of PRG2 peak. M, protein molecular weight marker.

3.1.13 Negative-stain transmission electron microscopy of the PGR2 complex

Although I observed PRG2 complex instability after analytical gel filtration, I proceed to test the peak B7 fraction (**Figure 18A**) for initial negative stain EM investigation. As this method is informative about the molecular morphology, organization, and heterogeneity of the sample, we could gain deeper knowledge on sample quality. Also, negative stan EM is a powerful method to study heterogeneous samples as the image contrast for

individual particles is adequate to conduct proper image classification [119,43]. Conversely, this method has its pitfall, it can generate artifacts induced by heavy metal stain-protein interactions, such as aggregation, molecular dissociation, flattening and stacking [120-122]. Eventually, with the great help of Jörg Bürger, B7 fraction was loaded on a carbon-coated film EM grid, which was electrostatically discharged. Stained with 2% heavy metal salts uranyl acetate negative staining solution, and subjected to EM image acquisition (**Figure 19A**). There were 225 micrograph images were taken. I gained image processing and analysis expertise of Dr. Justus Loerke, and 170 images were used for further picking potential particles by using e2boxer software (**Figure 19B**). In the end, 13841 particles were used to generate particle class averages. In order to generate this particle class averages, we performed particle centering and tilting by planar angle. Next, particles were clustered based on their image signal to noise ratio into 138 classes (**Figure 19C**). EM chromatograms show a high-quality image of successful negative stain with adequate contrast between particle to grid background and optimum particle density. The majority of particles fall in eight-fold symmetric bagel shaped particles with 12 nm size in diameter. These particles could be tetramer of PRG2 homomultimer with a top angle view. All 138 particle classes were used to build up the initial 3D model generation. However, since there is a wide range of particle heterogeneity between class averages, an initial low-resolution 3D model generation attempt could not be achieved.

Together with Jörg, we also performed direct cryogenic freezing of purified native PRG2 complex to eliminate any possible artifacts that may arise due to negative staining. Unfortunately, these grids were not subjected to cryo-EM image acquisition as we did not achieve an ideal sample homogeneity in our established purification pipeline. Nevertheless, these grids are kept in cryo storage and can be subjected to single particle cryo EM analysis in the future.

Moreover, we also performed negative stain EM imaging with the purified sample omitting the SEC step. In this experiment, we observed higher heterogeneity of particles on the grid (data has not been shown here).

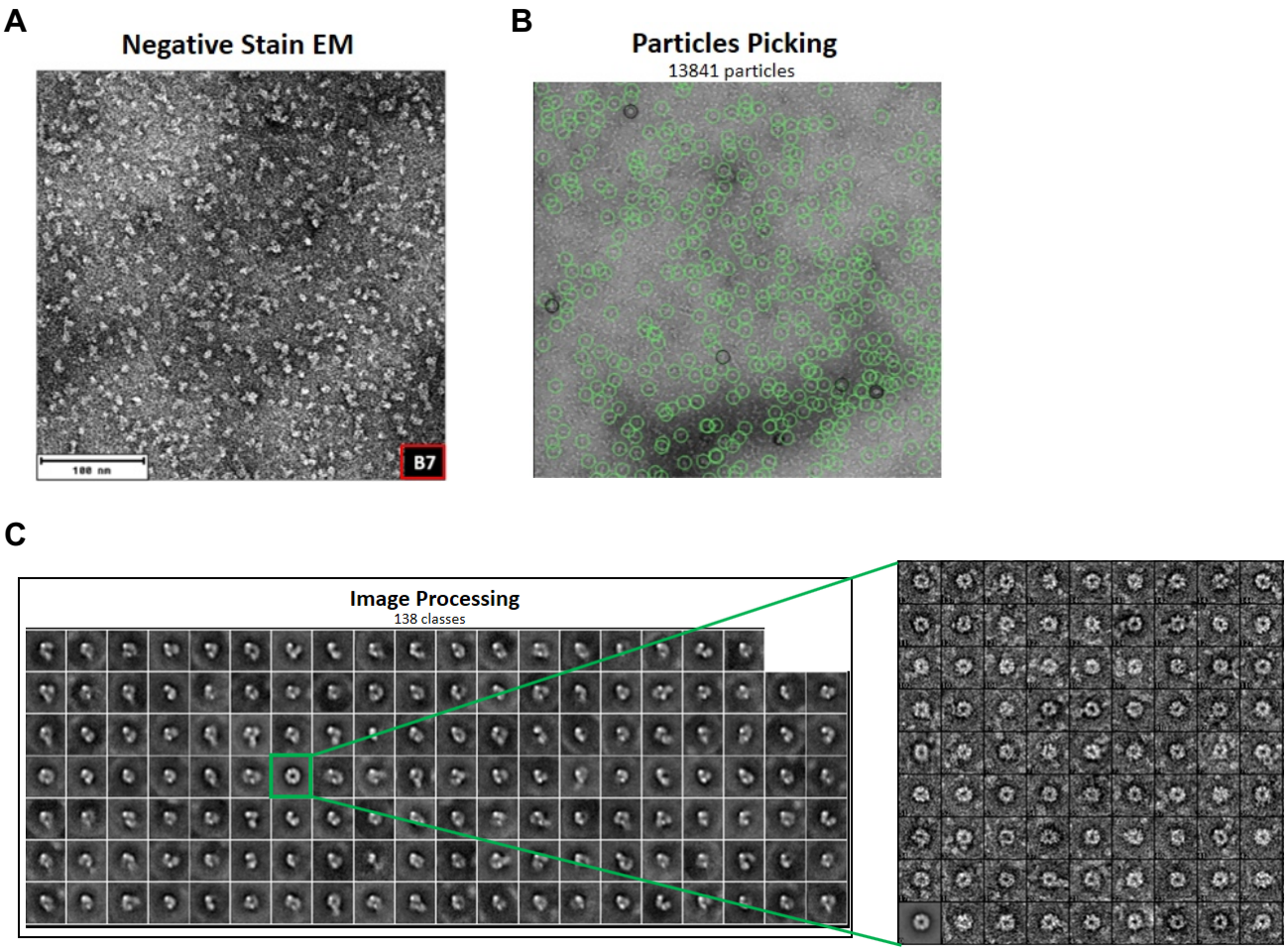


Figure 19: Single particle negative stain EM study.
A An example of a negative stain EM micrograph for particle picking. **B** An example of picked particles in green circles. **C** Clusters of picked particles with different size, shape, and orientation. The most populated class of picked particles is shown in green box which are in bagel shape form with eight-fold symmetry.

3.1.14 Mass Spectrometry analysis

There were multiple classes of purified native PRG2 complexes in native PAGE experiment before SEC in section 3.1.11, and I could not effectively separate distinct classes with analytical SEC in section 3.1.12. These results were an indicator of our diverse particle heterogeneity on negative stain EM micrographs in section 3.1.13. As all these results are complimentary of sample heterogeneity, further investigation has been needed to address which proteins are involved in each of these protein complexes. Therefore, together with Dr. Benno Kuropka, we performed LC-ESI-MS analysis on purified native PRG2 protein complex, which has not been subjected to final analytical SEC, and directly loaded onto denaturing NuPAGE 4-12% Bis-Tris gel. WB and Coomassie gel show a band above 71 kDa weight which corresponds to 76 kDa

monomeric PRG protein under denaturing conditions (**Figure 20A**). Apart from this expected result, I have also seen 150 kDa dimmer and some bigger weight aggregates (**Figure 20A**). Moreover, cleaved PRG2 fragments have been observed (**Figure 20A**). Therefore, I decided to excise and analyze seven bands on Coomassie-stained NuPAGE as indicated in Figure 20A to reveal whether all these bands are PRG2 protein, some other interacting proteins or possible contaminants. LC-ESI-MS analysis revealed that all seven bands were predominantly enriched with PRG2. Besides, band number 4 and 5 were also enriched with heat shock protein HSPA5 and HSPA1A in the fraction of catalytic PRG2 fragments (**Figure 20B**).

As we identified PRG2 peptides as the highest identified hits for all seven bands, the overall result indicates that the native PRG2 complex is pure. Also, band number 1 is a dimer of PRG2 even under denaturing the NuPAGE condition. Whereas, band number 4 and 5 also identified with heat shock protein and molecular chaperon. These proteins are usual suspects in heterologous protein overexpression studies to maintain protein expression quality with proper folding, assembly, translocation, and guiding proteins to degradation route. The main finding of this experiment is the purification pipeline yields high purity of native PRG2 protein complex; however, the purified complex consists of not only full PRG2 but also fragments of it. This finding is parallel with our observations on EM analysis as we detected a wide range of multimeric particles.

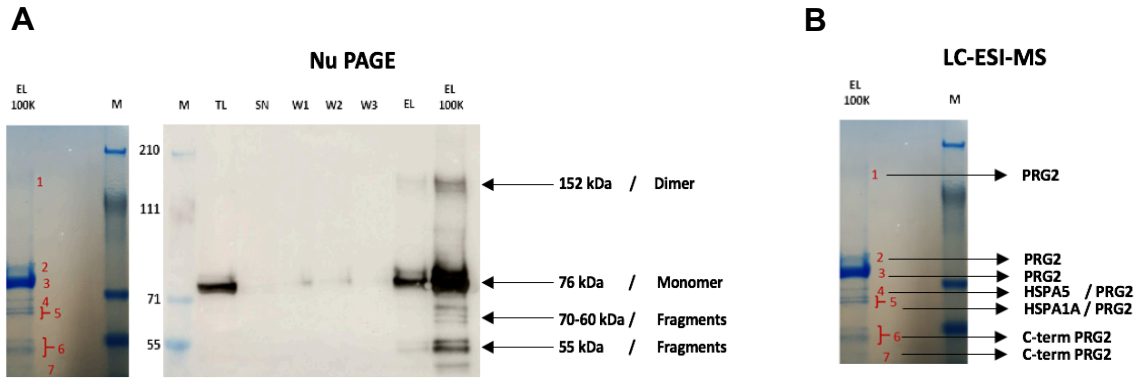


Figure 20: Mass Spectrometry analysis. Nu and Native PAGE.
A Coomassie and WB image of PRG2 eluate shows stages of multimers and catalytic fragments in denaturing Nu PAGE. Red numbers show which bands were picked for MS analysis on Coomassie stained gel. **B** Black arrows show identified proteins from LC-MS analysis. M, protein molecular weight marker. TL, total cell lysate. SN, unbounded proteins after FLAG IP. W1, W2, W3, first, second, third washing step of FLAG agarose to eliminate unspecific interaction. EL, eluate. EL 100K, concentrated eluate.

3.1.15 Proximity ligation assay study on PRG2 homo-, and PRG2/PRG3 heteromerization

A different experimental paradigm to analyze multimerization was used to confirm interaction. Here we tested PRG2 homomerization and possible heteromerization formation with another PRG2 family protein PRG3. It has been shown that PRG2 interacts with PRG3 in co-transfected Neuro2A cells using co-immunoprecipitation coupled to mass spectrometry [59]. In this section, we performed Proximity ligation assay (PLA) in mouse neuroblastoma N1E-115 cells. Additionally, this assay allows us to further explore the localization of PRG2 homomerization and multimerization in the cell.

The PLA is an antibody-based assay, combining proximity ligation with rolling circular amplification. The PLA probes consist of species-specific antibodies coupled to linear oligonucleotide sequences that in the proximity and the presence of a ligase will catalyze the ligation between the two ends in proximity. Rolling circular amplification is directed by a polymerase and the addition of fluorescence-labeled oligonucleotides complementary to the PLA probes only if the distance between these probes is less than 40nm [123].

To do so, cells were transfected with GFP and FLAG-tagged proteins, fixed, blocked, and probed by FLAG and GFP antibodies. After extensive washing, cells were probed with secondary antibodies called Proximity Ligation Assay (PLA) probes, each coupled to an oligonucleotide, recognize the two primary antibodies. Ligation takes place in case two complementary PLA probes are in close vicinity. Circular DNA is amplified by a DNA polymerase. DNA probes coupled to fluorochromes hybridize to amplicons and lead to the formation of intracellular red fluorescent dots specific to the interaction of the two FLAG and GFP tagged proteins. In order to control for the specificity of this approach, I undertook two sets of control experiments. In the first, cells were co-transfected with GFP vehicle and FLAG-tagged PRG2, whilst in the second GFP vehicle and empty plasmid. The PLA in these two control experiments resulted in no puncta, indicating the specificity of the approach in detecting PRG2 homomerization and PRG2-PRG3 heteromerization. The maximum intensity projection of confocal images of red PLA signal on PRG2-PRG2 and PRG2-PRG3 multimeric interaction have been observed on

the plasma membrane (**Figure 21**). Taken together, functional cooperation between PRG protein family members could exert a specific role in the cell. Further functional studies are required to address that.

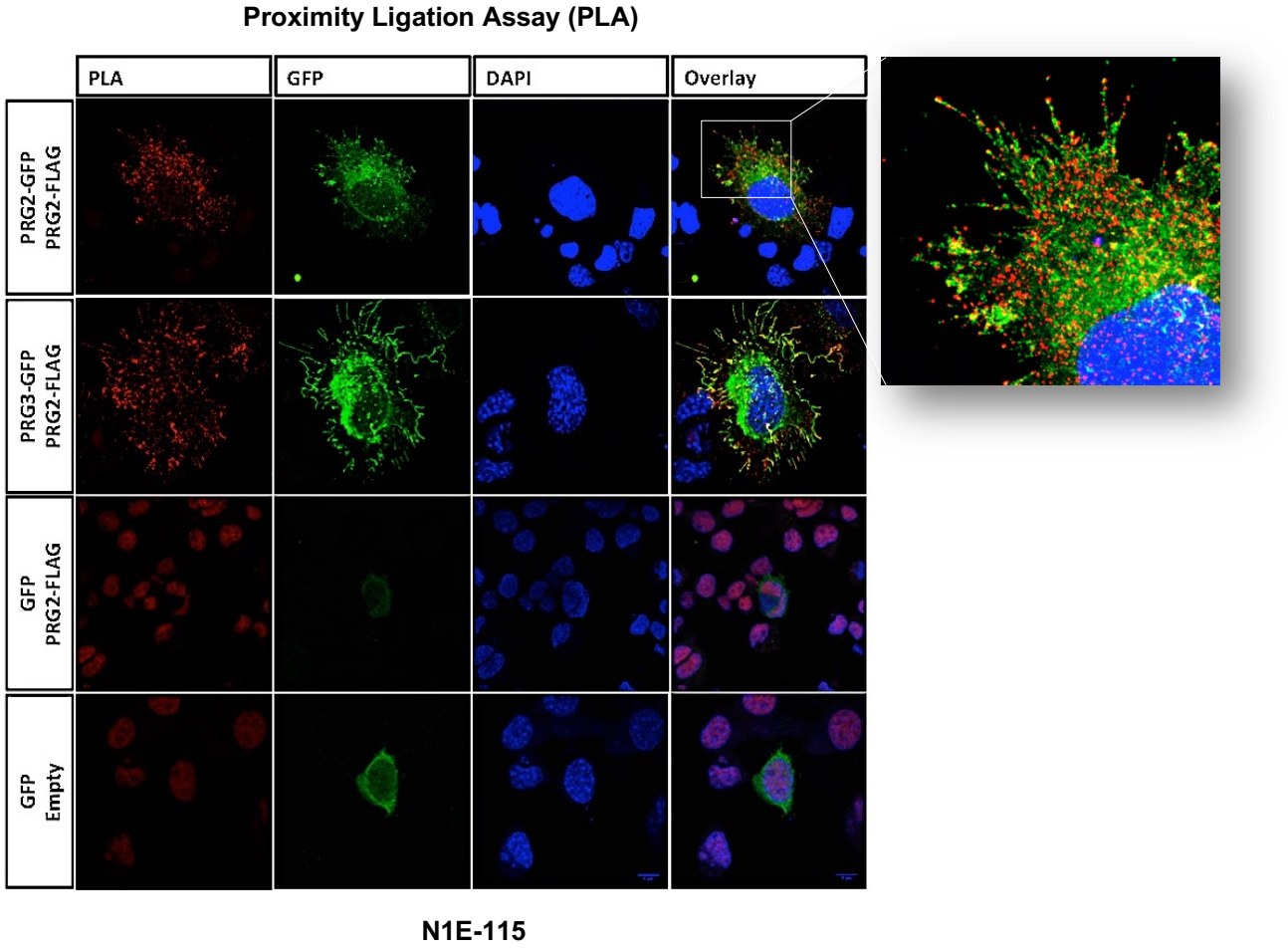


Figure 21: Proximity ligation assay study on PRG2 homo, and PRG2 & PRG3 heteromerization.
Maximum intensity projection of confocal images of Proximity ligation assay study on PRG2 homo, and PRG2 & PRG3 heteromerization in co-transfected N1E-115 cells. PLA signal indicates PRG2-PRG2, and PRG2-PRG3 co-localization in a close vicinity along particularly at the plasma membrane and filopodia.

3.2 Intracellular domain of PRG2 (PRG2ΔN)

3.2.1 Expression of intracellular domain of PRG2 in bacteria

In structural biology, if the protein of interest is an important target, it is commonly considered for attempting to purify domains or fragments. Though this may not address the global structural entity, it may still demonstrate insightful information in understanding some aspects of the protein. Therefore, in this section, I focus on the establishment of a pipeline to overexpress, purify, and characterize water-soluble mouse PRG2 C-terminal intracellular domain (PRG2ΔN) in bacteria cells under near-native condition. We used E. coli BL21 (DE3) strain, which is one of the most widely used hosts for high production level of heterologous proteins [124].

Our main rationale behind this strategy is to test whether PRG2ΔN makes homomultimeric complex. If the answer is yes, I can investigate the stoichiometry of homomultimerization. Since there was a range of homomultimeric complexes with full PRG2 purification, hereby I would like to answer if we can minimize multiple classes of homomultimers or separate an individual class of a multimeric complex by purifying water-soluble intracellular domain of PRG2 protein. In case of homomultimer purified in a single class, for example, a distinct class of purified homodimer or homotetramer can be the ideal candidate to subject in ingle-particle cryo-EM for structural investigations.

As highlighted in previous sections, the difficulties of using detergent to solubilize membrane protein, which is a harsh treatment that has to be carefully optimized to avoid protein aggregation and inactivation. One of the main advantages of working with the water-soluble domain of PRG2 is that the majority of available techniques for protein purifying, handling, and analyzing them were optimized for water-soluble proteins in an aqueous environment. Handling and studying membrane proteins required delicate treatment to disperse them in an aqueous solution, which is accomplished by adding a detergent to solubilize the biological membrane and form a soluble complex with the lipids and membrane proteins. Whereas, focusing purification of water-soluble intracellular domain does not require detergent for neither extraction nor stabilization of native complex.

To establish this pipeline, I also switched into a histidine tag pull-down assay, which is the standard option for water-soluble proteins overexpression and purification in E. coli. As it does not require antibody and elution peptide, it is highly scalable in terms of the financial budget. Histidine tagged proteins have an affinity for Ni2+ and several other metal ions that can be immobilized on chromatographic media using chelating ligands. Consequently, a protein containing a histidine tag will be selectively bound to metal-ion-charged media such as Ni Sepharose High Performance (HP) and Ni Sepharose 6 Fast Flow (FF) while most other cellular proteins will not bind or bind weakly. Elution is achieved by increasing the concentration of imidazole. This chromatographic technique is often termed immobilized metal ion affinity chromatography (IMAC) [125]. Hexahistidine tag (6xHis) has been chosen as longer histidine tags (with 8 or 10 histidine residues) are often used for membrane proteins to increase the binding strength and purity due to the ability to use more stringent washing steps [126]. Downsides of longer histidine tags are that expression levels have been reported to be decreased in some cases and that a higher imidazole concentration is required for elution which is destabilizing protein and trigger aggregation [125].

In this study, the intracellular domain of mouse PRG2 (PRG2ΔN) (residues 285-716) containing a 6xHistidine tag at C-terminal subcloned into prokaryotic vector pPAL. The recombinant plasmid was transformed into E. coli BL21 (DE3) strain for protein expression. The transformed E. coli strain was cultured, then induced with IPTG. Harvested cells were resuspended in 25 ml of 20 mM HEPES, 500 mM NaCl at pH 7.5 buffer supplemented with protease inhibitor, and cells were sonicated. The lysate was centrifuged, supernatants were collected for Nickel pull-down assay (HisTrap). Elution made by imidazole by using ÄKTA FPLC. Fractions were loaded onto a 12% SDS PAGE to investigate fractions which carries eluted His tagged PRG2ΔN (**Figure 22A**). Selected fractions (A8 to B15) were pooled, and the eluate was concentrated.

Subsequently, the eluate was loaded on the SEC to eliminate imidazole. Since imidazole leads to protein precipitation, this step has to be done as soon as possible. It is also a standard step to eliminate nonspecific proteins which binds on the Nickel column as we sperate all pull-downed proteins based on their molecular size and globular shape via

SEC. After SEC, all fractions were loaded onto a 12% SDS PAGE to find out which fractions were enriched with PRG2ΔN (**Figure 22B**).

PRG2ΔN positive fractions (A11 to B7) were pooled and loaded on anion exchange chromatography (IEC) (HiTrap) column. This is a further purification step to eliminate contaminant proteins that pull-downed with HisTrap due to intrinsic affinity to Nickel via metal-binding properties together with a similar molecular size with native PRG2ΔN complex, as they fragmented in the same fractions. Elution was performed by a gradient of 100 mM to 1 M NaCl concentration by using ÄKTA analytical FPLC. After IEC, fractions were loaded onto a 12% SDS PAGE to find out which fractions were enriched with PRG2ΔN (**Figure 22C**).

The most intense bands were observed from fractions A8 and A9, therefore both fractions were pooled and concentrated. This sample with a final concentration of 10 mg/ml was used for the crystallization screening. Rest was subjected to gel filtration SEC by ÄKTA micro FPLC on a Superose 12 3.2/300 column pre-equilibrated with 20 mM HEPES, 150 mM NaCl at pH 7.4. The final gel filtration step is a golden standard to ensure protein purity and complex stability. Fractions were loaded onto a 12% SDS PAGE. As we observed a defined band at 55 kDa weight on Coomassie-stained gel, we decided to pick B1 and B7 fractions for negative stain EM studies (**Figure 22D**). We observed a single band in B1 fraction with the size of PRG2ΔN, whereas in B7 fraction we observed an additional well-defined band at 70 kDa weigh with 1:1 stoichiometry.

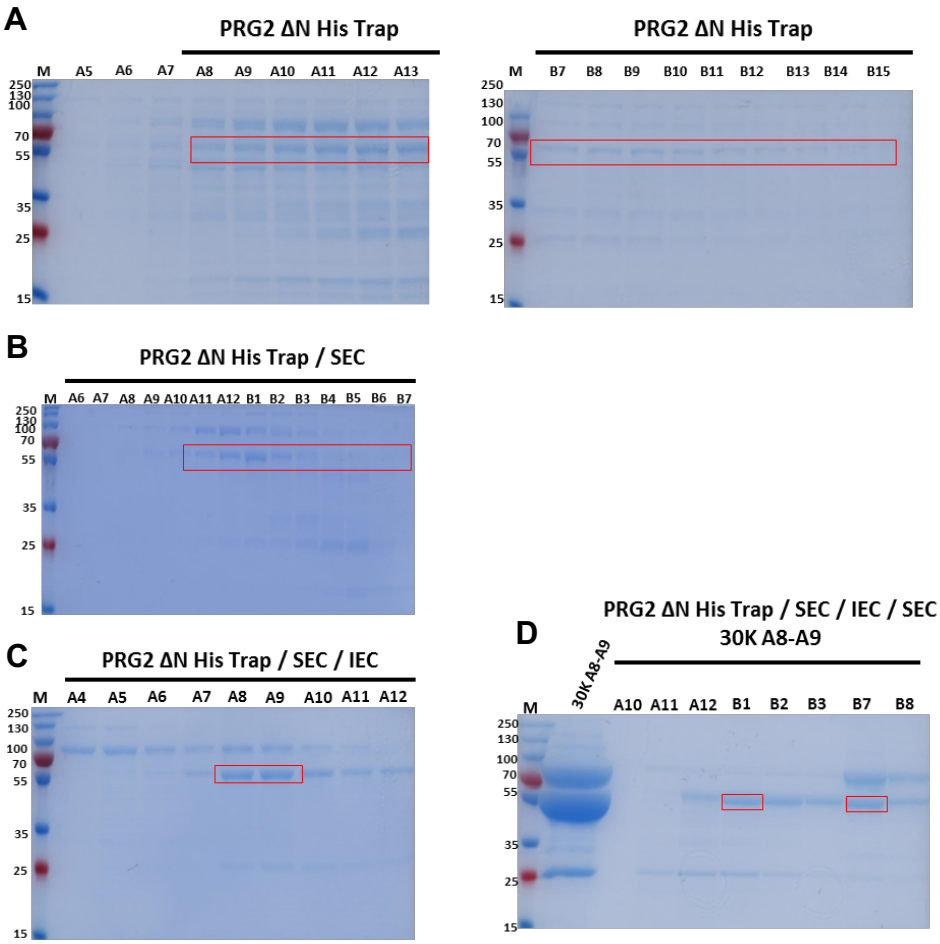


Figure 22: Expression of intracellular domain of PRG2 in E. coli.
Coomassie stained SDS PAGE images of **A** Nickel column pull down, **B** size exclusion, **C** anion exchange, **D** gel filtration of PRG2ΔN. Red boxes indicate selected fractions for downstream purification steps. M, protein molecular weight marker. SEC, size exclusion chromatography. IEC, anion exchange chromatography. 30K, 30 kDa cut off concentrate.

3.2.2 Mass Spectrometry analysis of PRG2ΔN

We observed multiple bands in purified native PRG2ΔN complexes in the SDS-PAGE experiment before and after the final SEC as seen in section 3.2.1 (**Figure 22D**). Therefore, together with Dr. Benno Kuropka, we performed LC-ESI-MS analysis on concentrated A8 and A9 fractions of purified native PRG2ΔN protein complex.

Six most prominent bands were excised from Coomassie-stained denaturing NuPAGE, digested with trypsin, peptides were extracted and measured by Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry (MALDI-MS). MS analysis reviled that, five bands were identified as prokaryotic proteins, which are PutA (UniProt Reference P09546), Odo1 (UniProt Reference P0AFG3), ArnA (UniProt Reference P77398), YihW

(UniProt Reference P32144), and SlyD (UniProt Reference P0A9K9) from top to bottom (**Figure 23**). As expected, band number four is identified as a cytoplasmic domain of mouse PRG2.

His-tag affinity purification of proteins expressed in E. coli in the presence of naturally histidine-rich host proteins, resulting in co-purification of the two most common E. coli contaminants are ArnA, a bifunctional enzyme involved in the modification of lipid A phosphates with aminoarabinose, and SlyD, a peptidyl-prolyl cis/trans isomerase. ArnA has several non-consecutive histidine residues, which are surface exposed and form clusters within the hexameric ArnA structure [127-129]. In contrast, SlyD is characterized by a 48-amino acid unstructured C-terminal tail containing 15 histidine [130,131]. Because the Ni-binding mechanism of ArnA and SlyD mimics that of the His-tag, both proteins are co-purified along with the target protein in His-tag affinity purifications. When protein expression is low, host proteins, especially ArnA and SlyD, have a similar abundance and compete for binding on nickel resins [132]. As a result, ArnA and SlyD are purified in equal amounts when compared to the target protein. Since we performed multiple steps of purification methods based on metal binding affinity, size, and ionic charge, we cannot eliminate a hypothesis that native PRG2ΔN complex forms physical interaction with these prokaryotic proteins in E. coli. Consequently, this result led us to find out an alternative solution to eliminate or minimize ArnA, and SlyD contaminations by using an alternative E. coli strain.

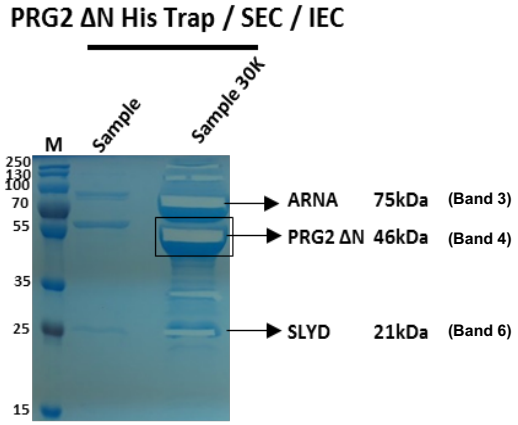


Figure 23: Mass Spectrometry analysis.
Coomassie stained SDS PAGE image of PRG2ΔN eluate shows two additional bands which were identified as prokaryotic metal binding proteins ArnA, and SlyD. Most intense protein band was depicted in black box which is identified as PRG2ΔN. M, protein molecular weight marker. SEC, size exclusion chromatography. IEC, anion exchange chromatography. 30K, 30 kDa cut off concentrate.

3.2.3 Negative-stain transmission electron microscopy of the PRG2ΔN

Although there were additional bands from purified native PRG2ΔN complex fractions B1 and B7, I proceed on testing these fractions on negative stan EM to gain knowledge about sample quality, particle density, homogeneity, and clusters. Both B1 and B7 fractions were the peak of two distinct curves on the final SEC chromatogram. As we identify that the B1 sample mostly enriched with PRG2ΔN at 55 kDa molecular weight range with minor contamination with prokaryotic protein SlyD at 25 kDa weight (Figure 22D), we were expected to obtain more information in B1 fraction laded EM grid. Besides, we also performed negative stain EM on B7 fraction although it is equally contaminated with prokaryotic protein ArnA at 70 kDa molecular weight (Figure 22D).

Representative micrograph shows that B7 fraction (**Figure 24B**) has a more defined particle profile comparing to the B1 micrograph (**Figure 24A**). Therefore, we proceed with particle picking on B7 fraction's micrographs. There were 99000 particles to generate 99 class averages (**Figure 24C**). As we know the sample is contaminated with ArnA protein (PDB #4wkg), we build up in-silico 2D projection views (**Figure 24D**) to eliminate ArnA particles from our class averages. We encountered that all well-defined particles are matching with the known native structure of the trimeric form of ArnA homodimer, and remaining particles are inhomogeneous particles. Therefore, we could not proceed further on image processing with this dataset, and repeat sample preparation by using E. coli BL21 (DE3) Lobstr strain, which is ArnA protein deficient, to test for minimization of ArnA contamination in section 3.2.4.

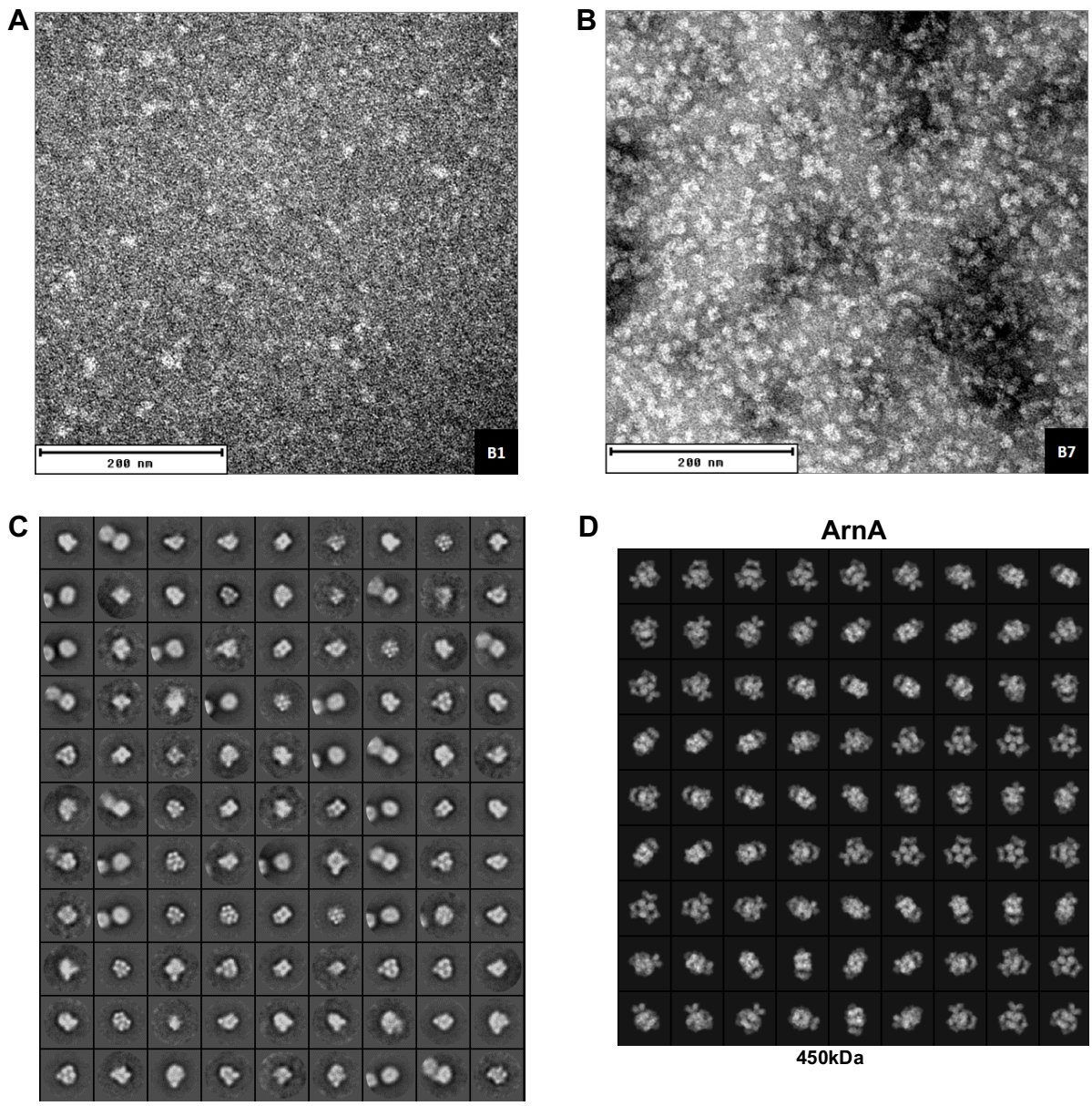


Figure 24: Single particle Negative-stain transmission electron microscopy of the PRG2ΔN.
An example of a negative stain EM micrograph **A** B1, **B** B7 fraction for particle picking. **C** An example of picked particles class averages from B7 fraction. **D** 2D projection view of native homo ArnA complex.

3.2.4 Expression of intracellular domain of PRG2 in E. coli Lobstr strain

In section 3.2.1, pipeline to overexpress and purify native PRG2ΔN complex has been established. I encountered that prokaryotic proteins ArnA and SlyD were co-purified with PRG2ΔN due to their metal-binding properties. In order to tackle with this problem, I find a commercial E. coli strain BL21 (DE3) Lobstr (Low background strain).

In this section, same pipeline was performed as in section 3.2.1 by using E. coli BL21 (DE3) Lobstr strain. Knocking out of both arnA and slyD genes lead to growth defects, therefore these strains are not viable options for recombinant protein expression [133,134]. To address this problem, the new strain has been genetically modified in arnA and slyD genes to manipulate their surface properties. Lobstr strain maintains normal cell growth but significantly reduces the Ni-binding affinities of both host proteins. Hence Lobstr strain drastically reduces ArnA and SlyD background contamination, thus enabling the purification of challenging low-expressing protein targets [132].

The recombinant plasmid was transformed into E. coli BL21 (DE3) Lobstr strain for protein expression. Cell growth, IPTG induction, harvest, and lysis methods has been done in the same way in section 3.2.1. The lysate was centrifuged, supernatants were collected for Nickel pull-down assay (HisTrap). Elution made by imidazole by using ÄKTA FPLC. Fractions were loaded onto a 12% SDS PAGE to investigate fractions which carries eluted His tagged PRG2ΔN (**Figure 25A**). Selected fractions (A6 to B11) were pooled, and the eluate was concentrated. Concentrated sample was subjected to SEC to eliminate imidazole. After SEC, all fractions were loaded onto a 12% SDS PAGE to find out which fractions were enriched with PRG2ΔN. PRG2ΔN positive fractions (A8 to B2) were pooled and loaded on anion exchange chromatography (IEC) (HiTrap) column (**Figure 25B**). The most intense bands were observed from fraction A5, therefore this fraction was concentrated for the final gel filtration purification step (**Figure 25C**). This sample with a final concentration of 10 mg/ml was used for the crystallization screen. Rest was subjected to gel filtration SEC by ÄKTA micro FPLC on a Superose 12 3.2/300 column pre-equilibrated with 20 mM HEPES, 150 mM NaCl at pH 7.4. Fractions were loaded onto a 12% SDS PAGE. Fraction A12 has the most intense band at 55 kDa weight on Coomassie-stained gel, therefore we decided to pick A12 fraction for negative stain EM studies (**Figure 25D**).

As it is expected, we enriched purified PRG2ΔN as seen at 55 kDa molecular weight, minimize ArnA protein band at above 70 kDa, and eliminate SlyD contaminations by using E. coli BL21 (DE3) Lobstr strain (**Figure 25D**) when we compare to E. coli BL21 (DE3) strain (**Figure 22D**).

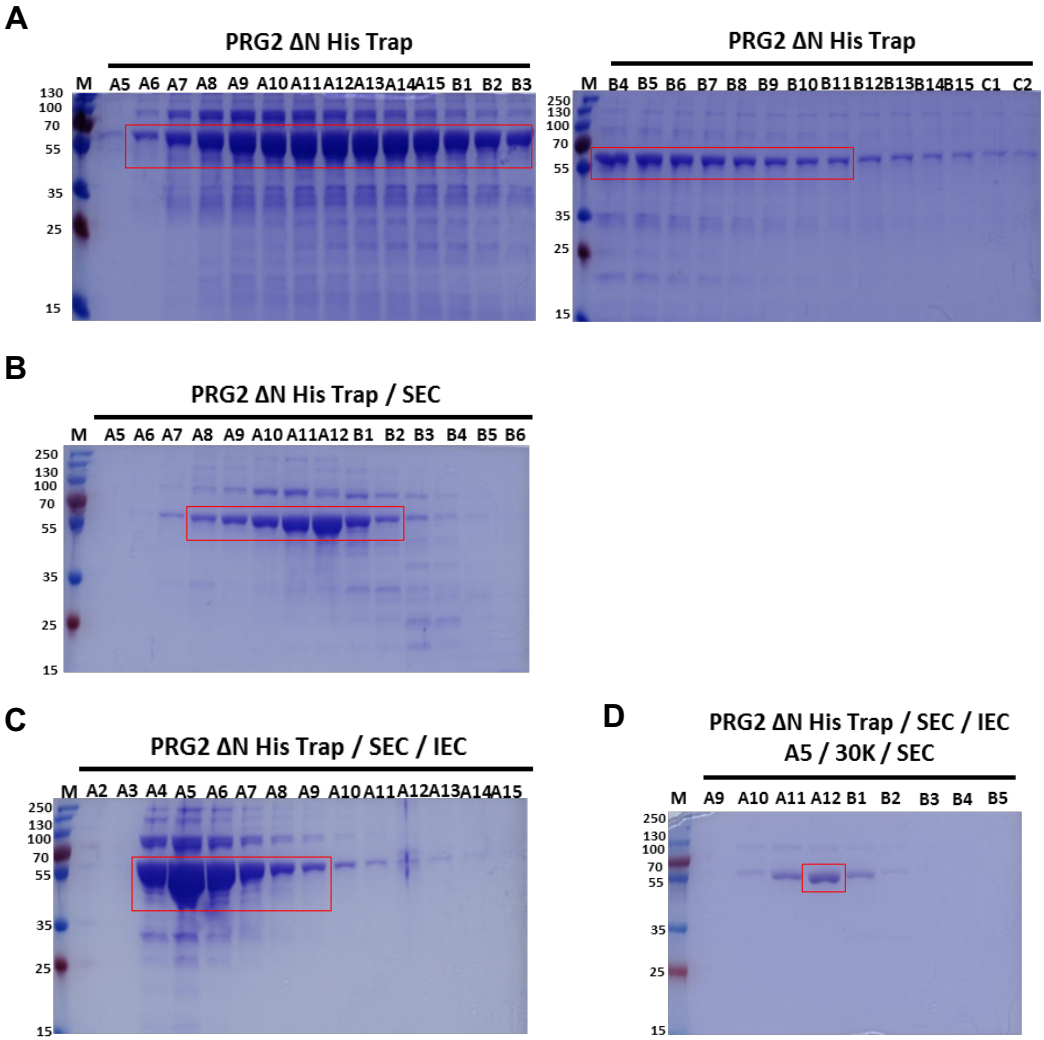


Figure 25: Expression of intracellular domain of PRG2 in prokaryote.
Coomassie stained SDS PAGE images of **A** Nickel column pull down, **B** size exclusion, **C** anion exchange, **D** gel filtration of PRG2ΔN. Red boxes indicate selected fractions for downstream purification steps. M, protein molecular weight marker. SEC, size exclusion chromatography. IEC, anion exchange chromatography. 30K, 30 kDa cut off concentrate.

3.2.5 Stability assay of PRG2ΔN with NanoDSF and SDS PAGE

In the previous section, I could use our establish pipeline to overexpress and purify native PRG2ΔN protein complex in E. coli BL21 (DE3) Lobstr strain with the elimination of prokaryotic metal-binding protein SlyD, and substantially reduce ArnA contaminations. Before embarking on a negative stain EM analysis with the final A12 fraction (**Figure 25D**), performing a set of biophysical protein sample quality assessment is vital. Purified protein quality control is the final and critical checkpoint of structural biology to ensure essential properties of protein sample purity, integrity, homogeneity, and activity. Approaches are to optimize the homogeneity, time-stability and storage

conditions of purified protein preparations, as well as methods to rapidly evaluate their reproducibility and robustness of established purification pipeline.

In this section, the reader can find an investigation of the storage conditions effect on protein stability. Flash freezing and storage at -20°C and thawing purified protein sample is one of the critical steps for protein structure. Unstable proteins can be exposed to significant denaturation, aggregation, and precipitation upon freezing/thawing [135]. These sensitive proteins may become unstable and lose their biological activity through a variety of physical or chemical mechanisms [136-138]. Also, purification and down-stream structural biology applications usually takes days, and during all these procedures sample kept on ice and 4°C temperature. Therefore, testing flash freezing and thawing, and storing purified protein at 4°C for more than a day are key methods to evaluate protein stability.

Hereby, stability assay has been performed by using nano differential scanning fluorimetry (Nano DSF) Prometheus NT.48 (Nano-Temper Technologies) to characterize the effect of freezing and thawing on thermal unfolding, chemical denaturation and aggregation properties of purified PRG2ΔN.

In this assay, 25 µl 1,2 mg/ml of A5 fraction (**Figure 25C**) of PRG2ΔN from anion exchange chromatography was mixed with 25 µl low salt tris buffer (50 mM Tris, 100 mM NaCl at pH 7.5) with 1:1 (v/v) ratio. The sample was either used directly or treated with one cycle of flash-freeze and thaw process and loaded into standard capillaries and subjected in Nano DSF for thermostability assay. Both untreated fresh, and once frozen and thaw samples produced an exact similar melting curve profile, and we observed 43,3°C Tm value in both cases (**Figure 26A**).

Furthermore, in order to test the stability of purified protein at long term 4°C temperature storage, two SDS PAGE run at day 0 and day 2 after incubation at 4°C. Fractions of PRG2ΔN from anion exchange chromatography freshly right after FPLC without any incubation period, and after two days of incubation at 4°C have been loaded onto 12% SDS PAGE. Coomassie-stained gels indicate that both samples have identical band

profiles (**Figure 26B**). These results suggest that freezing and thawing, and two days incubation at 4°C have no dramatic effect on PRG2ΔN sample stability and solubility.

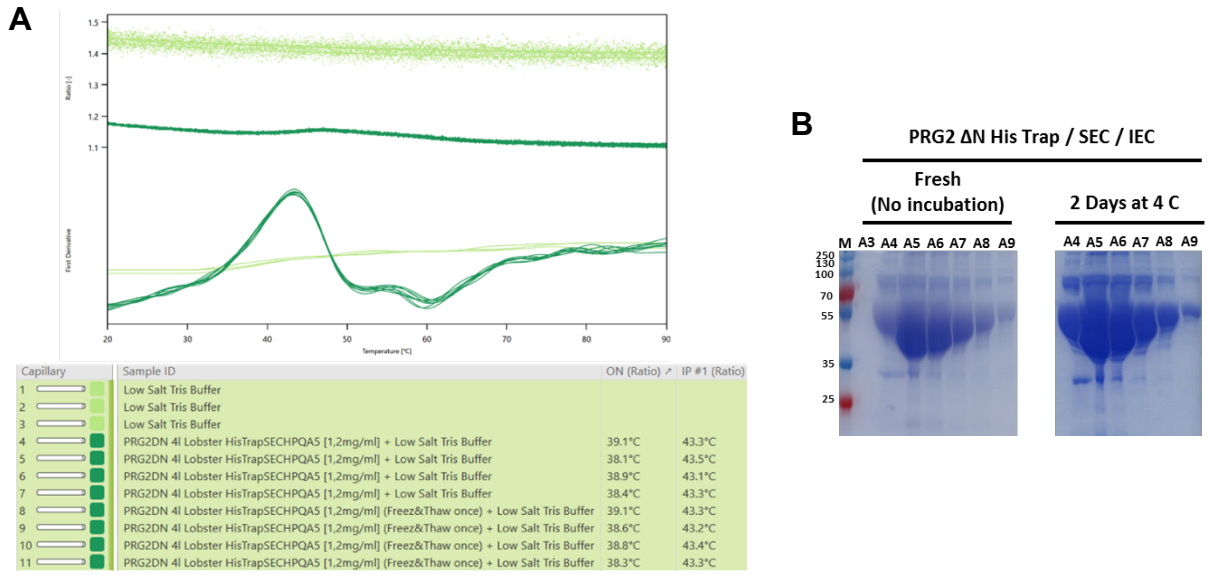


Figure 26: Stability assay of PRG2ΔN with NanoDSF and SDS PAGE.
A NanoDSF analysis of fresh versus frozen and thaw PRG2ΔN have 43,3°C Tm value. Four technical replicates were used. Light green represents Tris buffer control, dark green indicates PRG2ΔN sample. **B** Coomassie stained SDS PAGE images of fresh versus two days of incubation at 4°C PRG2ΔN fractions. M, protein molecular weight marker. SEC, size exclusion chromatography. IEC, anion exchange chromatography. 30K, 30 kDa cut off concentrate.

3.2.6 Filtration and Additive Screening of Purified PRG2ΔN with NanoDSF

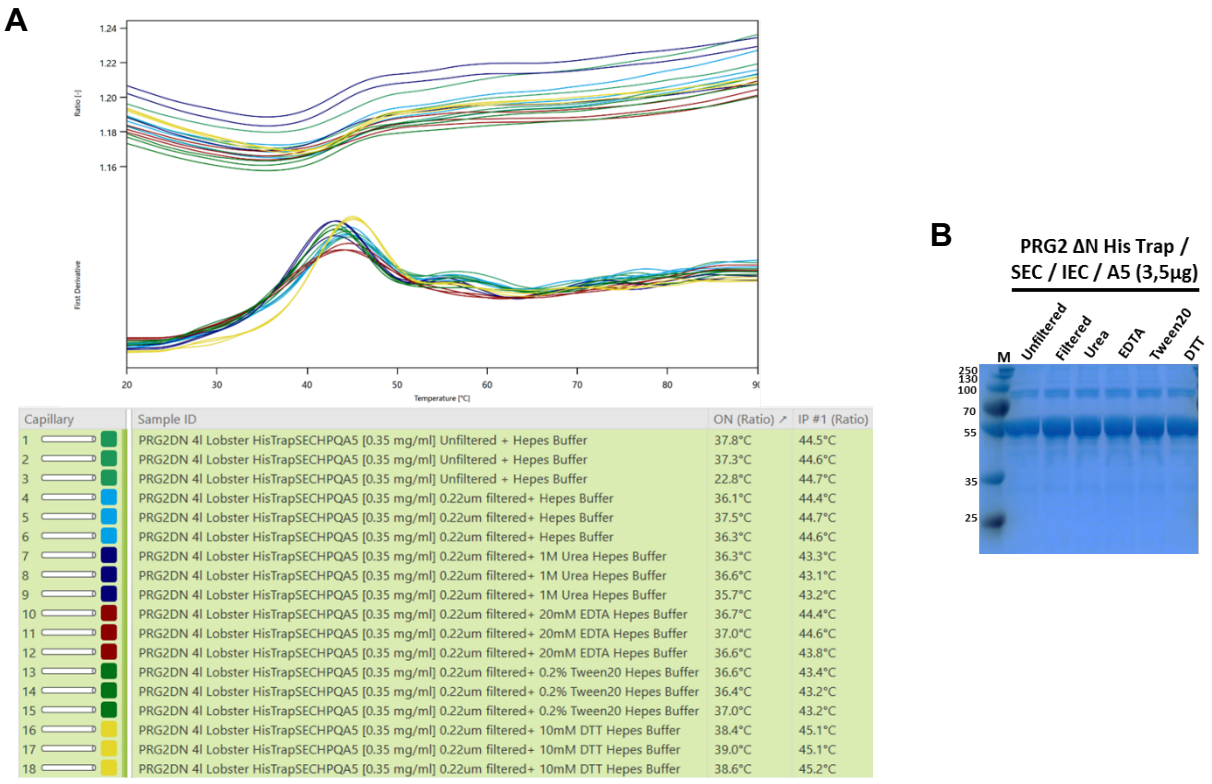
In structural biology, the filtration of all aquatic buffers is a necessity in order to eliminate microorganisms and insoluble aggregates. Also, purified sample filtration is a well-known step in protein crystallization to minimize sample heterogeneity and eliminate possible aggregates [139]. Hence, in this section, 0,22 µm pore size, hydrophilic PVDF centrifugal filtration has been tested on purified PRG2ΔN sample to see if sample purity and homogeneity could improve.

Hence, sample homogeneity before and after filtration has been investigated by using Nano DSF. In this assay, 20 µl 0,35 mg/ml of A5 fraction of purified PRG2ΔN from anion exchange chromatography was mixed with 20 µl HEPES buffer (20 mM HEPES, 150 mM NaCl, pH 7,4) with 1:1 (v/v) ratio. Each sample loaded into standard capillaries and subjected to Nano DSF for thermostability assay. Both samples before and after filtration produced an exact similar melting curve profile with observed 44,6°C Tm value (**Figure**

27A). This result suggests us, filtration step does not lead to sample loss. However, I did not observe any subtle improvement on the first derivative peak profile of purified PRG2ΔN sample before and after filtration in thermostability assay.

Moreover, the use of small organic molecules in the buffer system can have profound effects on protein stability, structure, and function. In an example, the stability of a protein is strongly dependent on the presence of hydrogen bond networks between solvent molecules. 2 to 8 M high concentration of urea exert its effect directly, by binding to the protein, or indirectly, by altering the solvent environment via disrupting the hydrogen bonding network between water molecules and reduce the stability of the native state of proteins [140,141]. However, 1M or lower concentration of urea as a chaotropic agent can eliminate protein aggregation and enhance protein solubilization [142]. Another example is Ethylenediaminetetraacetic acid (EDTA) used as a chelating agent in protein purification, which eliminates contaminating divalent cations and to inhibit protease activity. In addition to that EDTA is a well-known destabilizing reagent for metal-binding proteins [143]. Hence, EDTA has been tested as an additive to see some destabilization effects on metal-binding protein ArnA contaminants. The third example is Non-ionic surfactants Tween 20, which serves as the prevention of surface-induced denaturation and aggregation of recombinant proteins. It exerts detergent function with amphiphilic nature, stabilize purified protein hydrophobic domains by establishing protein-detergent complex at the hydrophilic surface, and prevent self-association of these hydrophobic domains. And the final example is the most popular reducing agent for proteins is dithiothreitol (DTT). The presence of a reducing agent in experimental solutions is often necessary for some proteins to maintain activity, stability, and prevent from aggregation by breaking interprotein disulphate bonds. 5 to 10 mM concentration of DTT is widely used in structural biology in the final gel filtration step in purification to minimize protein aggregation. Also, it has been used in protein refolding during the renaturation process [144]. Consequently, we also tested possible effects of 1 M chaotropic agent urea, 20 mM chelating agent Ethylenediaminetetraacetic acid (EDTA), 0,2% (v/v) Tween20, 10 mM reducing agent DTT additives on filtered PRG2ΔN sample by using Nano DSF to see possible effect on sample purity, homogeneity, and stability. We observed 43,2°C, 44,1°C, 44,3°C, and 45,1°C Tm values on respecting agents (**Figure 27A**). This result shows that the only addition of 10mM DTT has a slight

increase in the melting temperature of PRG2ΔN. In addition, each sample was loaded onto a 12% SDS PAGE, and Coomassie-stained image shows a similar band profile (**Figure 27B**). This gel image suggests us filtration process is compatible with the purified PRG2ΔN sample as there is no sample loss before and after filtration. Nevertheless, I could not see any decrease in the other contaminant bands which are below or above 55 kDa molecular weight.



Capillary	Sample ID	ON (Ratio) ↗	IP #1 (Ratio)
1	PRG2DN 4l Lobster HisTrapSECHPQA5 [0.35 mg/ml] Unfiltered + Hepes Buffer	37.8°C	44.5°C
2	PRG2DN 4l Lobster HisTrapSECHPQA5 [0.35 mg/ml] Unfiltered + Hepes Buffer	37.3°C	44.6°C
3	PRG2DN 4l Lobster HisTrapSECHPQA5 [0.35 mg/ml] Unfiltered + Hepes Buffer	22.8°C	44.7°C
4	PRG2DN 4l Lobster HisTrapSECHPQA5 [0.35 mg/ml] 0.22um filtered+ Hepes Buffer	36.1°C	44.4°C
5	PRG2DN 4l Lobster HisTrapSECHPQA5 [0.35 mg/ml] 0.22um filtered+ Hepes Buffer	37.5°C	44.7°C
6	PRG2DN 4l Lobster HisTrapSECHPQA5 [0.35 mg/ml] 0.22um filtered+ Hepes Buffer	36.3°C	44.6°C
7	PRG2DN 4l Lobster HisTrapSECHPQA5 [0.35 mg/ml] 0.22um filtered+ 1M Urea Hepes Buffer	36.3°C	43.3°C
8	PRG2DN 4l Lobster HisTrapSECHPQA5 [0.35 mg/ml] 0.22um filtered+ 1M Urea Hepes Buffer	36.6°C	43.1°C
9	PRG2DN 4l Lobster HisTrapSECHPQA5 [0.35 mg/ml] 0.22um filtered+ 1M Urea Hepes Buffer	35.7°C	43.2°C
10	PRG2DN 4l Lobster HisTrapSECHPQA5 [0.35 mg/ml] 0.22um filtered+ 20mM EDTA Hepes Buffer	36.7°C	44.4°C
11	PRG2DN 4l Lobster HisTrapSECHPQA5 [0.35 mg/ml] 0.22um filtered+ 20mM EDTA Hepes Buffer	37.0°C	44.6°C
12	PRG2DN 4l Lobster HisTrapSECHPQA5 [0.35 mg/ml] 0.22um filtered+ 20mM EDTA Hepes Buffer	36.6°C	43.8°C
13	PRG2DN 4l Lobster HisTrapSECHPQA5 [0.35 mg/ml] 0.22um filtered+ 0.2% Tween20 Hepes Buffer	36.6°C	43.4°C
14	PRG2DN 4l Lobster HisTrapSECHPQA5 [0.35 mg/ml] 0.22um filtered+ 0.2% Tween20 Hepes Buffer	36.4°C	43.2°C
15	PRG2DN 4l Lobster HisTrapSECHPQA5 [0.35 mg/ml] 0.22um filtered+ 0.2% Tween20 Hepes Buffer	37.0°C	43.2°C
16	PRG2DN 4l Lobster HisTrapSECHPQA5 [0.35 mg/ml] 0.22um filtered+ 10mM DTT Hepes Buffer	38.4°C	45.1°C
17	PRG2DN 4l Lobster HisTrapSECHPQA5 [0.35 mg/ml] 0.22um filtered+ 10mM DTT Hepes Buffer	39.0°C	45.1°C
18	PRG2DN 4l Lobster HisTrapSECHPQA5 [0.35 mg/ml] 0.22um filtered+ 10mM DTT Hepes Buffer	38.6°C	45.2°C

Figure 27: Filtration and Additive Screen with NanoDSF.
A NanoDSF analysis of 0,22 µm filtered versus unfiltered PRG2ΔN in 1:1 (v/v) HEPES buffer. In addition, 1M urea, 20mM EDTA, 0,2% (v/v) Tween20, 10 mM DTT additives in HEPES buffer were tested on filtered PRG2ΔN. Three technical replicates were used. There is no Tm difference with or without filtration. Only addition of DTT leads to increase thermostability of PRG2ΔN. There is a clear shift of yellow peak to the right-hand side comparing to the rest. **B** Coomassie stained SDS PAGE images of PRG2ΔN with or without additives. M, protein molecular weight marker. SEC, size exclusion chromatography. IEC, anion exchange chromatography. 30K, 30 kDa cut off concentrate.

3.2.7 Detection of native complex sizes of PRG2ΔN with Native PAGE

As it was performed in section 3.1.11, hereby I gain the advantage of Native PAGE analysis to investigate molecular weight, homogeneity, and complex stoichiometry of purified PRG2ΔN from E. coli BL21 (DE3), and E. coli BL21 (DE3) Lobstr strains.

30 kDa cut off filter concentrated anion exchange fractions A7 and A8 from BL21 (DE3) (**Figure 22C**), and A4 to A9 pool from Lobstr strain (**Figure 25C**) in 30 to 1,5 µg concentrations have been loaded onto 3-12% Native PAGE Bis-Tris Gel for Coomassie staining, and WB with His antibody. Single level of heptameric/octameric complex formation and monomeric form with purified PRG2ΔN has been observed (**Figure 28**), yet there were multiple orders of native protein complexes in full PRG2 in section 3.1.11 (**Figure 17**). Besides, there is no smeared band after gel filtration in the PRG2ΔN sample in comparison with gel filtrated full PRG2 protein (**Figure 18B**). This shows that the intracellular domain of PRG2 protein is more stable than full PRG2 protein. In addition, we observed two distinct classes of native PRG2ΔN complex as monomeric form below 66 kDa molecular weight marker and heptameric or octameric form below 480 kDa marker (**Figure 28**). The heptameric/octameric complex class is more defined than monomeric PRG2ΔN. This indicates that the intracellular domain of PRG2 also tends to form homomultimerization to maintain stability and lower the free energy level. This finding is a piece of novel evidence for the C-terminus of PRG2 protein can be sufficient for homomultimerization. When we compare two E. coli strain at the heptameric/octameric and monomeric band level, there is a dramatic difference at Coomassie-stained gel. purified PRG2ΔN from E. coli BL21 (DE3) Lobstr strain has crispy, sharp, and well-defined bands in comparison to purified PRG2ΔN from E. coli BL21 (DE3) strain with less background protein bands as we eliminate prokaryotic metal-binding protein contaminants (**Figure 17**).

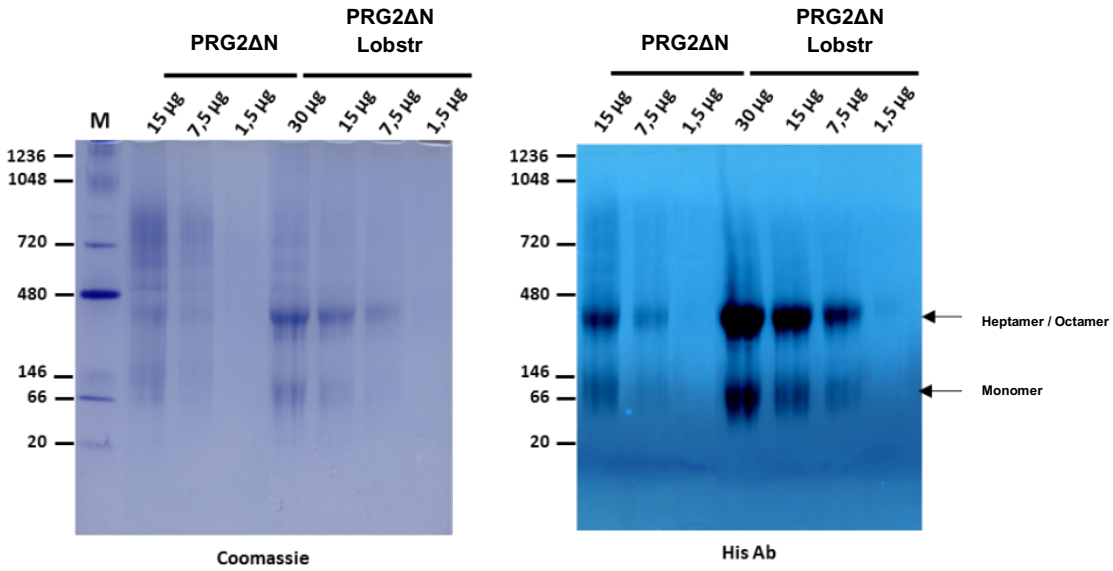


Figure 28: Detection of native complex sizes of PRG2ΔN with Native PAGE.
Native PAGE Coomassie and WB image of PRG2ΔN eluate from E. coli BL21 (DE3), and Lobstr strains show mono- and heptameric/octameric stages. M, protein molecular weight marker.

3.2.8 Detection of native complex sizes of PRG2ΔN with Dynamic Light Scattering

The extent of monodispersity versus polydispersity of a protein in solution can be measured with biochemical or biophysical techniques such as size exclusion chromatography (SEC) [145], Native PAGE [109-111], and Dynamic Light Scattering (DLS) [26]. DLS measures the diffusion coefficient of a native protein complex in buffer condition based on Brownian motion of small molecules move faster than large ones. Upon exposure of monochromatic light, light scattered by moving protein molecules display particle motion. The distances between particles are changing in aquatic environment, and scattered light fluctuates by time. This light scattering signal provides quantitative information about the diffusion coefficient behavior of a protein in solution. This data can be transformed to detect sample mono- or polydispersity, population of polydisperse particle, and theoretical diameter of each particle population. Hence, I performed (DLS) measurement as a complementary analysis to Native PAGE to identify sample homogeneity, size and molecular weight of purified PRG2ΔN.

For this measurement, purified PRG2ΔN eluates has been used from A4 to A9 fractions pool from E. coli BL21 (DE3) Lobstr strain with 0,1 mg/ml concentrations. At first, the sample was spun down to eliminate any aggregates, then filtered through 0,22 µm

nitrocellulose membrane. The sample was loaded on disposable cuvette and measured with DynaPro NanoStar.

As a result of this experiment, I observed polydispersity in the sample (**Figure 29A**), which is a parallel result with Native PAGE that was observed (**Figure 28**). 99% of all classes of heterogenic particles species are falling in 6,5 nm in radius (13nm in diameter assuming that particles are spherical), and 350 to 400 kDa in weight (**Figure 29B**) which forms heptameric/octameric complex in parallel with our Native PAGE analysis (**Figure 28**).

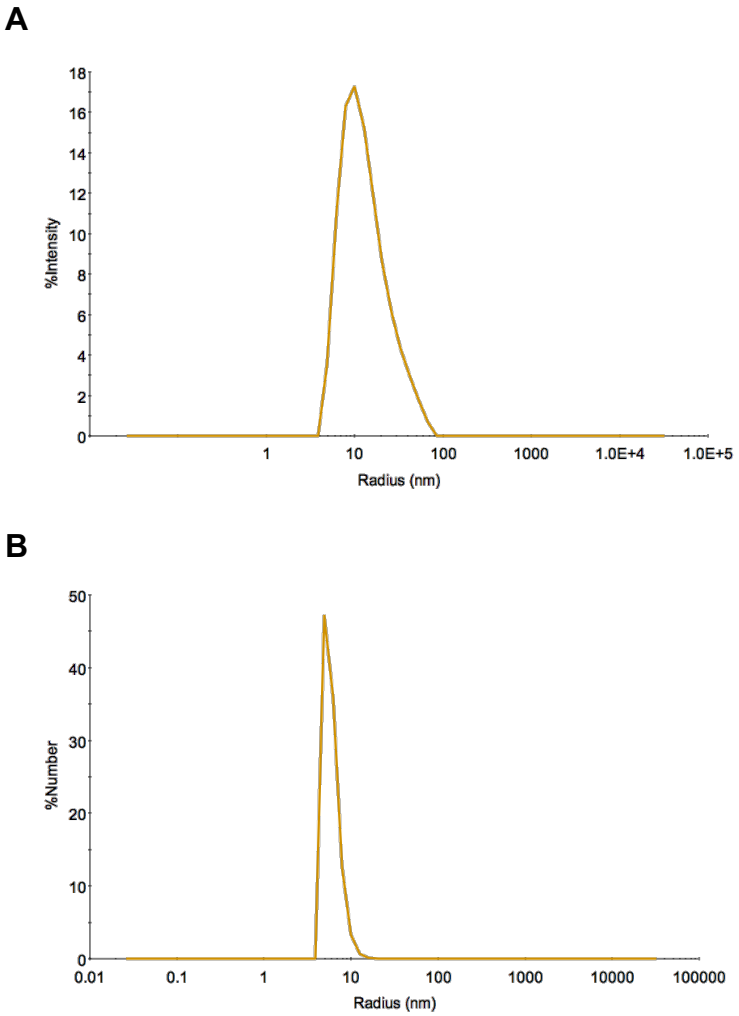


Figure 29: Detection of native complex sizes of PRG2ΔN with Dynamic Light Scattering. **A** DLS analysis of PRG2ΔN has been shown that sample is heterogenic with a range of 6.5 to 80 nm radius. **B** Main species (> = 99% mass) has an average radius of 6.5 nm with a sharp single peak. This main species corresponds to a protein of approximately 350 to 400 kDa if we assume a spherical shape.

3.2.9 Negative-stain transmission electron microscopy of the PRG2dN from E. coli Lobstr strain

Next, sets of protein quality controls have been performed to assess and improve protein stability by testing diverse storage conditions, eliminating insoluble protein aggregates with filtration, screening for buffer additives, and investigating homogeneity, homomultimerization, molecular weight and size of the purified PRG2ΔN complex.

After all these biophysical assessments for protein sample quality measurements, I received the green light to proceed with a negative stain EM investigation on purified PRG2ΔN from E. coli BL21 (DE3) strain. In this experiment, A12 fraction from the last gel filtration step of purified PRG2ΔN complex from E. coli BL21 (DE3) Lobstr strain was used (**Figure 25D**). This sample was snap freeze and stored at -20°C and thawed once. In order to eliminate any aggregates, sample was subjected in SEC one more time. Eventually, I observed a reproducible chromatogram profile in SEC separation and A11 fraction, which is the maxima peak, was used for negative stain EM investigation.

Representative micrograph shows that A11 fraction has a wide range of particle sizes and shapes heterogeneity (**Figure 30**). In the control case, we loaded HEPES buffer (20 mM HEPES, 500 mM NaCl at pH 7,5) only where we do not see any defined particles. I was confident about sample purity and oligomerization stage which is heptameric/octameric complex, we start to questioner if negative stain reagent could trigger complex disassociation and aggregation. Therefore, we proceed to test alternative staining reagents in the next section.

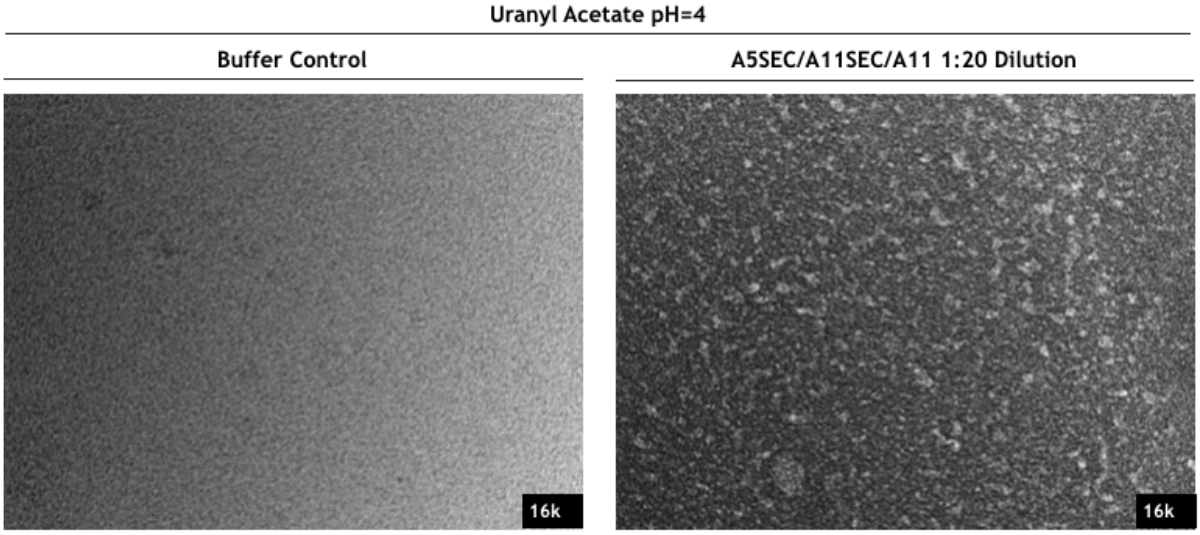


Figure 30: Negative-stain transmission electron microscopy of the PRG2dN from E. coli Lobstr strain.
An example of a negative stain EM micrograph of native PRG2ΔN complex, and HEPES buffer control. 16k: 16000x magnification.

3.2.10 Testing alternative agents for negative staining

The charges on the surface of proteins are mainly dependent on its isoelectric point and local pH level. Since the isoelectric point is constant for each protein, the surface chemistry of the molecules is dependent on ambient pH. Therefore, the pH of the stain solution takes a decisive role in the negative staining method [146]. As we know that full PRG2 has theoretical isoelectric point 5.7, and it shows weak thermostability at low pH levels equal to 4 and 5 (Table 8). The theoretical isoelectric point of PRG2ΔN is 5.2, and we expect a similar negative effect of low pH levels on protein stability.

The most commonly used negative staining reagents are uranyl acetate and uranyl formate with low pH levels 4 and 4.5 respectively. One of the advantages is to use these stains are they act as a fixative, preserving many protein-protein interactions on a millisecond time scale [147], although the low pH of the stain could be detrimental to some biological samples as it could lead to precipitate [146].

Alternative negative staining agents were tested with higher pH levels in order to test possible effects of low pH level in Uranyl Acetate negative stain method on protein complex integrity. For this test, I used 50 µl of 2mg/ml PRG2ΔN eluate of SEC A12

fraction from Lobstr strain was loaded into the SEC one more time to eliminate aggregates. A11 fraction was used for negative stain EM by using Uranyl Formate pH 4,5 (**Figure 31A**), Uranylless pH 4,5 (**Figure 31B**), Ammonium Molbidate pH 5 (**Figure 31C**), and Phosphotungstic Acid pH 7 (**Figure 31D**). Representative micrographs show that Ammonium Molbidate staining is the best among all tested stains without any staining artifact and most defined particle view. Nevertheless, the sample was still heterogeneous (**Figure 31C**). Therefore, we concluded to use standard Uranyl Acetate for further experiments.

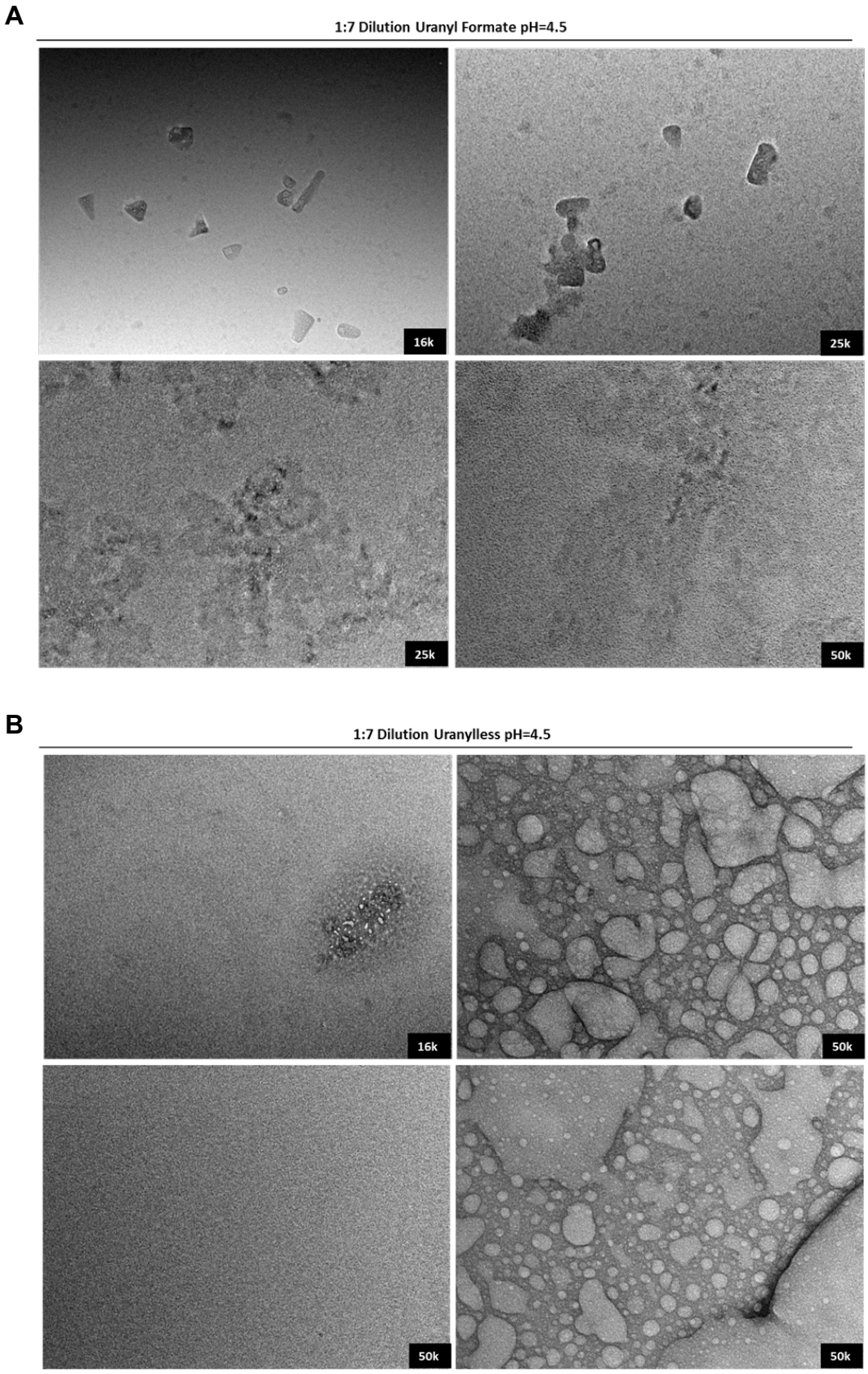


Figure 31: Testing alternative agents for negative staining.
An example of a negative stain EM micrograph of native PRG2ΔN complex at 16k, 25k, and 50k magnification by using **A** Uranyl Formate pH 4,5 , **B** Uranylless pH 4,5.

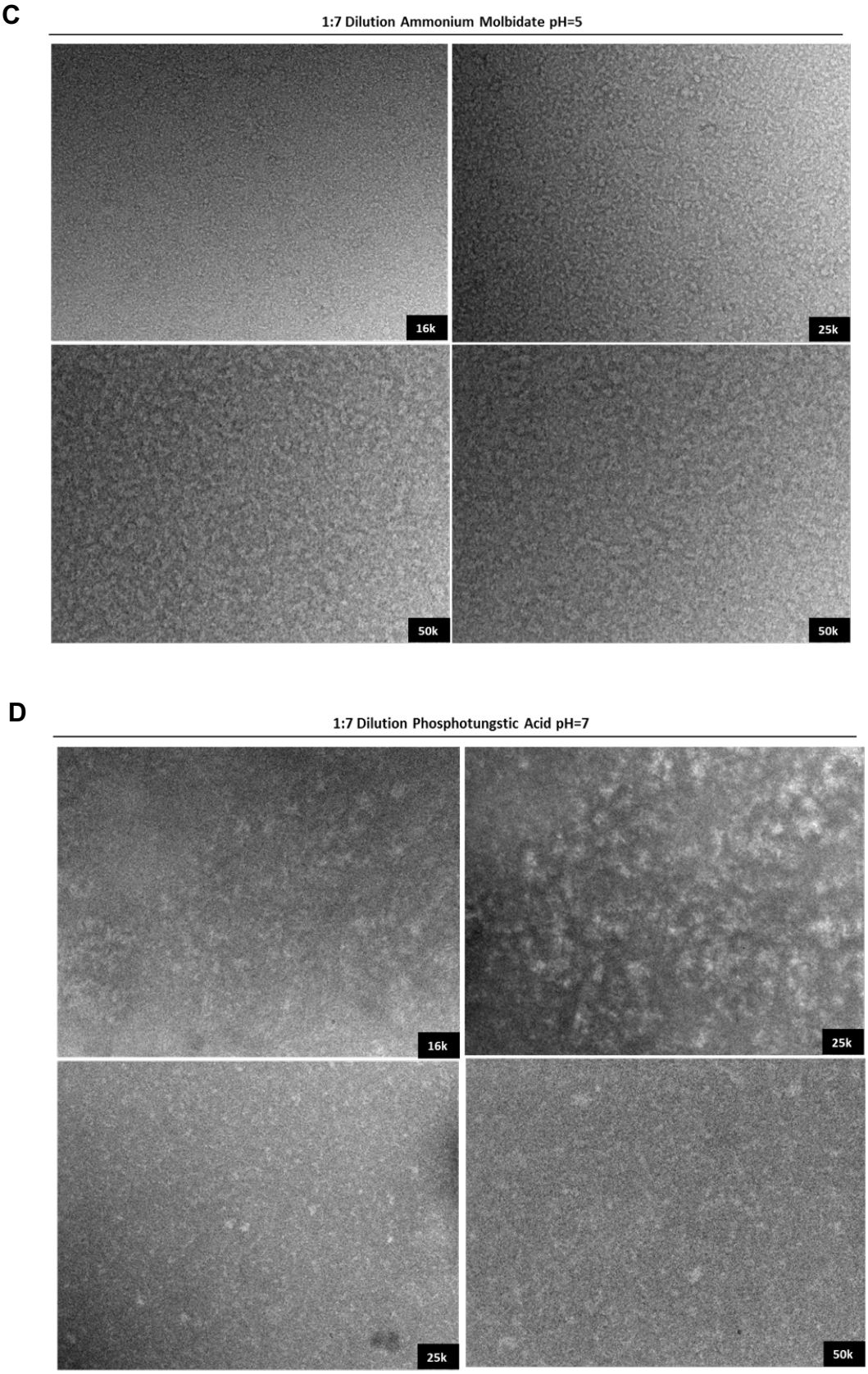


Figure 31: Testing alternative agents for negative staining.
An example of a negative stain EM micrograph of native PRG2ΔN complex at 16k, 25k, and 50k magnification by using **C** Ammonium Molbidate pH 5, **D** Phosphotungstic Acid pH 7.

3.2.11 Limited Proteolysis Assay

As it has been observed homomultimerization with purified PRG2ΔN complex, I gain the advantage of limited proteolysis assay to find out which domains are situated at the inner core region of the complex. In order to retrieve structural information through proteolytic digestion, it is necessary to limit the reaction of proteolysis [148] with a relatively low concentration of a protease, which cuts at recognition sites throughout the protein's outermost accessible hydration shell [149]. Hence, the core of the complex can maintain undigested as restriction recognition sites are not available for limited activity of the proteolytic enzyme. Therefore, knowing this inner domain could enlighten us about binding pockets that could be in charge of the homomultimerization of PRG2ΔN.

For this experiment, PRG2ΔN eluate of SEC A12 fraction from Lobstr strain has been incubated with a serial dilution of chymotrypsin which has theoretical 13 cleavage sites (**Figure 32A**). Reactions were quenched to limit proteolysis. Following digestion with a protease, protein samples are analyzed by Coomassie-stained SDS-PAGE to identify cleavage products. The appearance of lower molecular weight bands represents the digestion of the protein, and intensity and appearance/disappearance of bands closely monitored. In comparison with the undigested sample, differential band at 0,012 mg/ml chymotrypsin digested sample has been observed (**Figure 32B**). This band excised for LC-MS analysis. MS analysis reviled 14 peptides cover the entire PRG2ΔN sequence (**Figure 32C**). Since I observed all theoretical cleavage peptide products of PRG2ΔN from this partially digested band, it is not possible to predict any core domain of native PRG2ΔN complex by using this data obtained from this protein band.

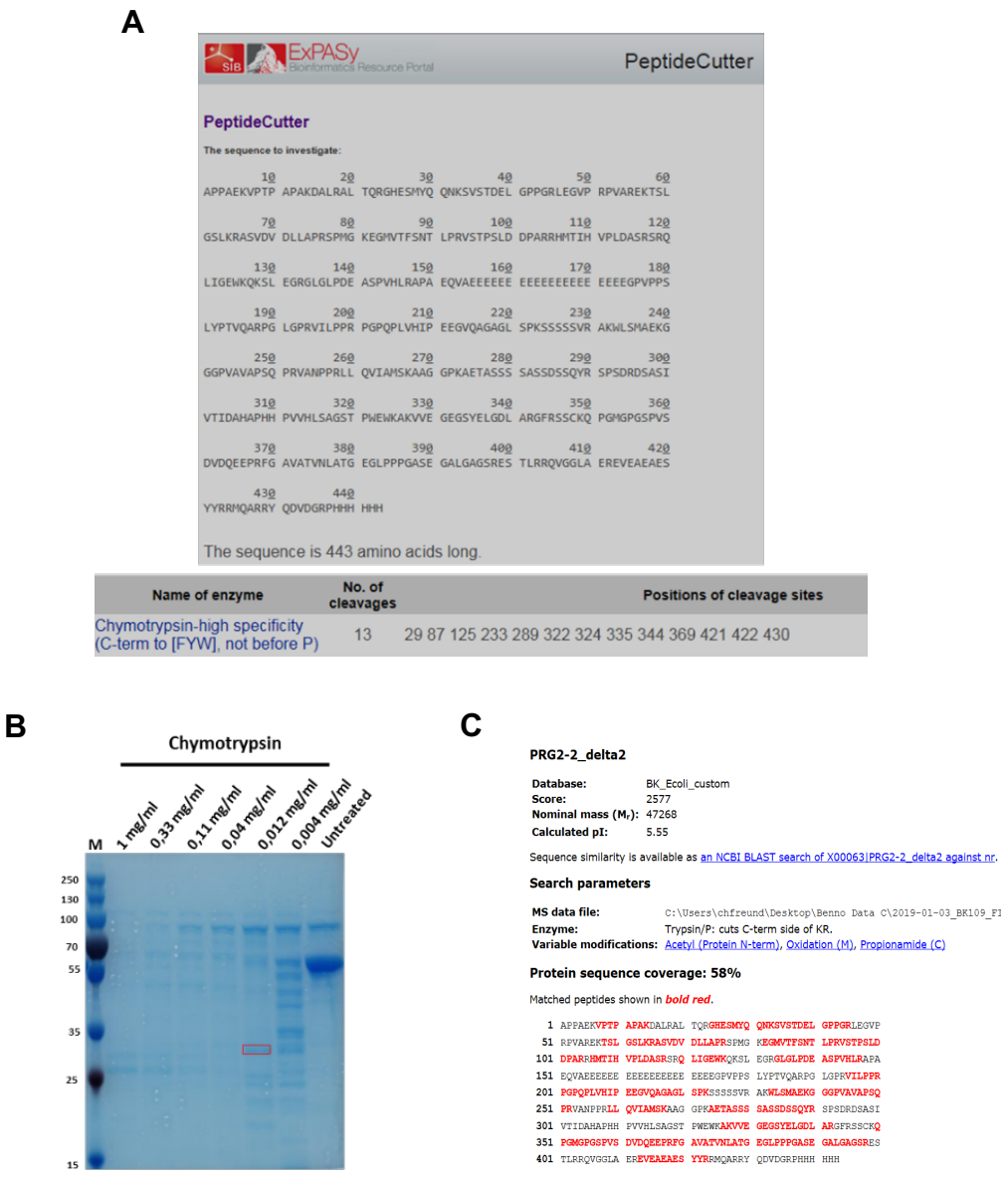


Figure 32: Limited Proteolysis Assay.
A Theoretical cleavage sites and number of PRG2ΔN with chymotrypsin (https://web.expasy.org/peptide_cutter/). **B** Coomassie stained SDS PAGE image of digested PRG2ΔN. Red box indicates excised band for LC-MS analysis. **C** Result of MS analysis. Red letters indicate identified amino acids in digested peptides. M, protein molecular weight marker.

3.2.12 Expression of intracellular domain of PRG2 in eukaryote

It is first a summary. In section 3.1, pipeline has been established to overexpress and purify full PRG2 protein in near-native conditions to maintain a biologically active complex for structural investigation. When sets of biophysical and biochemical assay in combination with negative stain EM investigation have been performed, I encountered that sample purification was achieved successfully, yet purified complex forms diverse

classes of homooligomeric complexes. The majority of particle classes fall in the category of eight-fold symmetric bagel shape structure with 12 nm size in diameter. As the single-particle EM image analysis method requires homogenous particle distribution, and we could not achieve to separate distinct homomultimeric complexes, we decided to establish a new overexpression and purification pipeline with water-soluble intracellular domain of PRG2 C-term (PRG2ΔN) in E. coli.

In section 3.2.1, PRG2ΔN in E. coli BL21 (DE3) strain overexpressed and purified, yet I ended up with co-purification of prokaryotic metal-binding proteins ArnA and SlyD contaminants. Hence, in section 3.2.4, I tested a mutagenic E. coli BL21 (DE3) Lobstr strain which is slyD gene-deficient, and arnA gene mutated to implement our established purification pipeline. As I expected, contamination of SlyD completely eliminated, and ArnA contamination minimized, additionally PRG2ΔN protein yield increased. After biophysical and biochemical tests to investigate purified protein sample quality, stability, size and molecular weight of the complex measures, this sample has been subjected in negative stain EM investigation. Unpredictably, we observed sub-optimal sample heterodispersity with multiple homomultimerization states.

Consequently, in this section, we decided to switch back into the eukaryotic overexpression system to overexpress and purify native PRG2ΔN complex. As explained earlier, using mammalian cell lines to overexpress recombinant protein is to introduce accurate protein folding and maintain post-translational modifications. Therefore, overexpression and purification of PRG2ΔN in HEK293S GnTI- cells has been tested to investigate the improvement of purified protein complex particles in terms of stability, density, definition, homogeneity, and monodispersity in negative stain EM analysis.

In this experiment, intracellular domain of mouse PRG2 (PRG2ΔN) (residues 285-716) containing a 6xHistidine tag at C-terminal has been subcloned into eukaryotic overexpression plasmid pTM4. The recombinant plasmid was transiently transfected in HEK293S GnTI- cells for protein expression. Cells were harvested one day after transfection, lysed with sonication, and soluble protein fraction in the supernatant was collected after centrifugation for Nickel pull-down assay (HisTrap). Elution made by

imidazole. Subsequently, the eluate was filtered via a 0.22 µm pore size column to eliminate aggregates and concentrated with 30 kDa cut off centrifugal filter before loaded into the SEC on a Superose 12 3.2/300 column as a final purification step (**Figure 33A**). In this final SEC step, we also eliminate imidazole which leads to protein aggregation.

It has been seen that B3 fraction is the maxima of the gel filtrated purified PRG2ΔN complex. Next, pull down fractions were loaded onto a 12% SDS PAGE and stained with Coomassie, and blotted for His Ab (**Figure 33B**). Gel images show us pull-down assay worked effectively. Final gel filtration step eliminates some impurities which are visible on Coomassie-stained gel in between 100 to 130 kDa and above 70 kDa molecular weight markers at concentrated and filtered eluate (EL) sample comparing to B3 fraction. Eventually, I decided to use B3 fraction for both negative stain EM investigation, and mass spectrometric analysis.

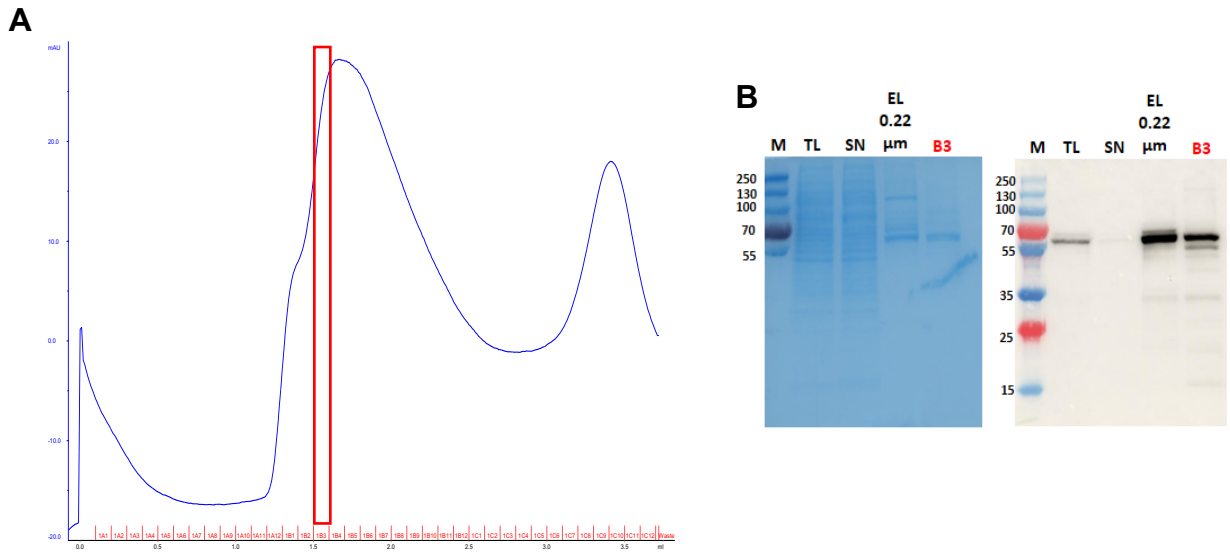


Figure 33: Expression and purification of intracellular domain of PRG2 in eukaryote.
A SEC chromatogram shows 280 nm wavelength of absorption for PRG2ΔN eluate. B3 fraction is depicted in red box, which is the initiation of PRG2 peak maxima. **B** SDS PAGE Coomassie and WB images show enrichment of PRG2ΔN after pull down assay, and SEC B3 fraction shows retention of PRG2ΔN after SEC. M, protein molecular weight marker. TL, total cell lysate. SN, unbounded proteins after Nickel pull down assay. EL, eluate.

3.2.13 Mass Spectrometry analysis and single particle negative stain EM study

Before subjecting the sample on a grid for negative stain EM investigation, we performed mass spectrometry analysis on B3 fraction to evaluate sample purity and quality. 30 µg concentrations of purified protein loaded onto 12% SDS-PAGE, and it run till all proteins enter in resolving layer of the gel for Coomassie staining. A thick section of the protein band was excised and digested with trypsin. Peptides were extracted and measured by Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry (MALDI-MS). MS analysis reviled that mouse PRG2ΔN protein peptides are identified as the top hit. The rest of the list consists of chaperone and heat shock protein family members, which are expected proteins in the exogenous protein overexpression approach (**Table 9**). This result assures us our purification pipeline was successful, and B3 fraction is pure enough to proceed on EM investigation.

Table 9: List of identified proteins via MS analysis.
B3 fraction of PRG2ΔN has been used for LC-MS/MS. Identified proteins are sorted based on identified peptide match numbers.

Description	Type	Mass	Num. of significant matches	Sequence coverage
PRG2-2_delta2	PRG2	47325	72	0,7
TCPB_HUMAN T-complex protein 1 subunit beta	Chaperone	57794	38	0,61
GRP75_HUMAN Stress-70 protein	Heat shock protein	73920	37	0,48
GRP78_HUMAN 78 kDa glucose-regulated protein	Chaperone	72402	36	0,44
TCPE_HUMAN T-complex protein 1 subunit epsilon	Chaperone	60089	35	0,49
TCPQ_HUMAN T-complex protein 1 subunit theta	Chaperone	60153	34	0,47
TCPG_HUMAN T-complex protein 1 subunit gamma	Chaperone	61066	31	0,37
TCPZ_HUMAN T-complex protein 1 subunit zeta	Chaperone	58444	30	0,43
TCPA_HUMAN T-complex protein 1 subunit alpha	Chaperone	60819	29	0,5
CH60_HUMAN 60 kDa heat shock protein	Heat shock protein	61187	27	0,51
TCPH_HUMAN T-complex protein 1 subunit eta	Chaperone	59842	27	0,41

Accordingly, B3 fraction from the last gel filtration step of purified PRG2ΔN complex from HEK293S GnTI- cells (**Figure 33B**) was loaded on a carbon-coated film EM grid, which was electrostatically discharged. Stained with 2% heavy metal salts uranyl acetate negative staining solution and subjected to EM image acquisition. Representative negative stain EM micrograph shows well-defined particles with two sub-populations of big particles with a shape of the ring structure, and small particle with a circular or triangular ring pattern (**Figure 34**). As particle density, signal to noise level, homogeneity

of individual particle classes in micrographs are satisfactory, single-particle image processing and analysis need to be done in the future.

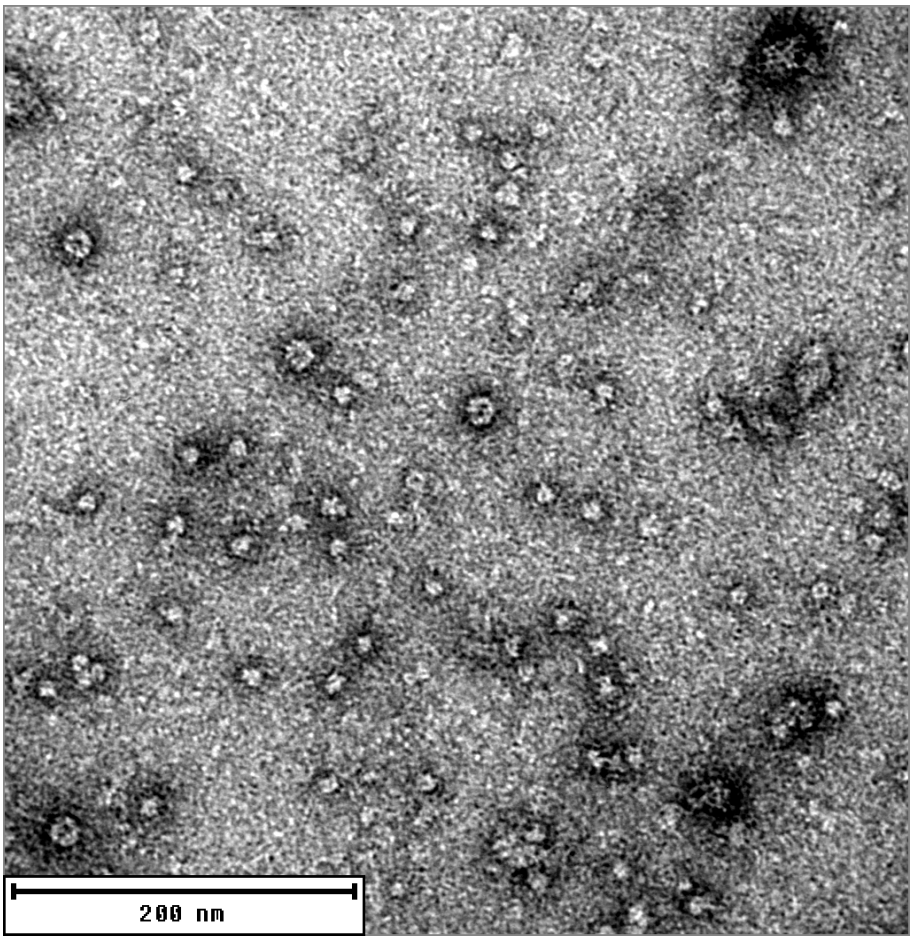


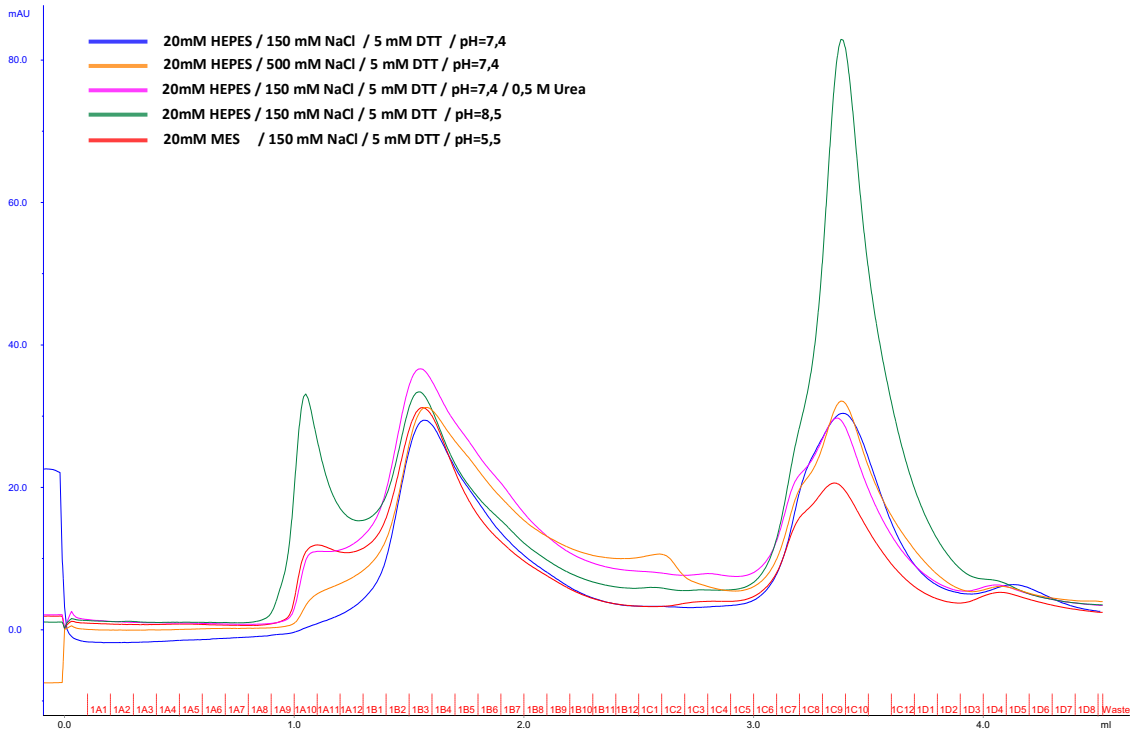
Figure 34: Single particle negative stain EM study.
An example of a negative stain EM micrograph of B3 fraction of PRG2ΔN. Scale bar is 200 nm.

3.2.14 Optimizing buffer condition to enhance intracellular domain of PRG2 sample homogeneity

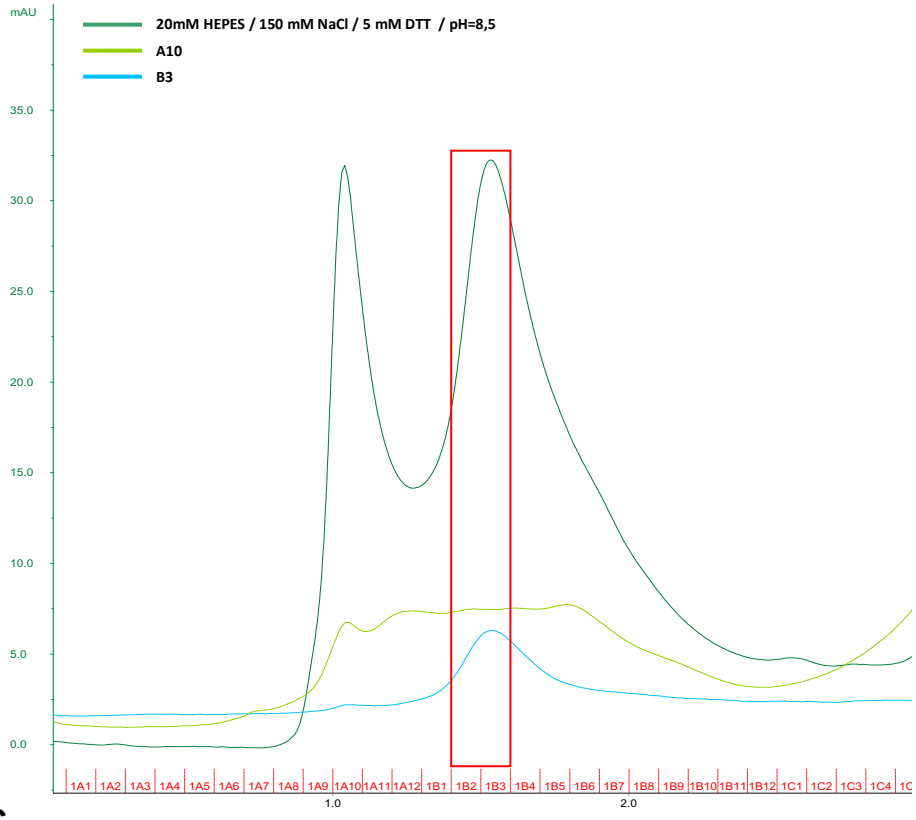
As I have seen a fusion of two peaks in SEC chromatograph with maxima at B1 and B4 fractions from purified PRG2ΔN in HEK293S GnTI- cells (**Figure 33A**), I tested a variety of buffer conditions to improve separation of these two peaks. This step could help us to eliminate possible sample heterogeneity due to a mixture of two possible distinct particle classes. To do so, we tested four additional buffer conditions. Initial 20 mM HEPES, 150 mM NaCl, 5 mM DTT at pH 7,4 buffer condition kept as a control indicated as a blue line in the chromatogram (**Figure 35A**) to compare with modified buffers with 500 mM high salt as an orange line, the addition of 0,5 M Urea as a pink line, increasing

pH to 8,5 as a green line, and testing low pH at 5,5 by switching HEPES to 20 mM MES buffer as a red line (**Figure 35A**). The SEC result shows that HEPES buffer at pH 8,5 leads to two separation of two distinct peaks at A10 and B3 fractions (**Figure 35A**) green line. We loaded each fraction onto 12% SDS PAGE for Coomassie staining and WB with His Ab. Gel results indicate a well-defined band near 55 kDa molecular weight marker range in B2, B3, B4 fractions (**Figure 35C**). Whereas, in A10 fraction there is a vague band at the same molecular weight (**Figure 35C**). Furthermore, in order to see sample stability, B3 fraction was reloaded into SEC, and we observed a normal distribution curve with a peak maxima at B3 fraction (**Figure 35B**). Taken together, B3 fraction of purified PRG2ΔN in HEK293S GnTI- cells under 20 mM HEPES, 150 mM NaCl, 5 mM DTT at pH 8,5 buffer condition consists of the stable protein complex. To test sample homogeneity, B3 fraction has been loaded on a carbon grid for negative stain. Negative stain EM image acquisition, single-particle image processing, and image analysis needs to be done in the future.

A


mAU
80.0
60.0
40.0
20.0
0.0
20mM HEPES / 150 mM NaCl / 5 mM DTT / pH=7,4
20mM HEPES / 500 mM NaCl / 5 mM DTT / pH=7,4
20mM HEPES / 150 mM NaCl / 5 mM DTT / pH=7,4 / 0,5 M Urea
20mM HEPES / 150 mM NaCl / 5 mM DTT / pH=8,5
20mM MES / 150 mM NaCl / 5 mM DTT / pH=5,5
1A1 1A2 1A3 1A4 1A5 1A6 1A7 1A8 1A9 1A10 1A11 1A12 1B1 1B2 1B3 1B4 1B5 1B6 1B7 1B8 1B9 1B10 1B11 1B12 1C1 1C2 1C3 1C4 1C5 1C6 1C7 1C8 1C9 1C10 1C12 1D1 1D2 1D3 1D4 1D5 1D6 1D7 1D8 Waste
0.0
1.0
2.0
3.0
4.0
ml

B


mAU
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
20mM HEPES / 150 mM NaCl / 5 mM DTT / pH=8,5
A10
B3
1A1 1A2 1A3 1A4 1A5 1A6 1A7 1A8 1A9 1A10 1A11 1A12 1B1 1B2 1B3 1B4 1B5 1B6 1B7 1B8 1B9 1B10 1B11 1B12 1C1 1C2 1C3 1C4 1C
1.0
2.0

C

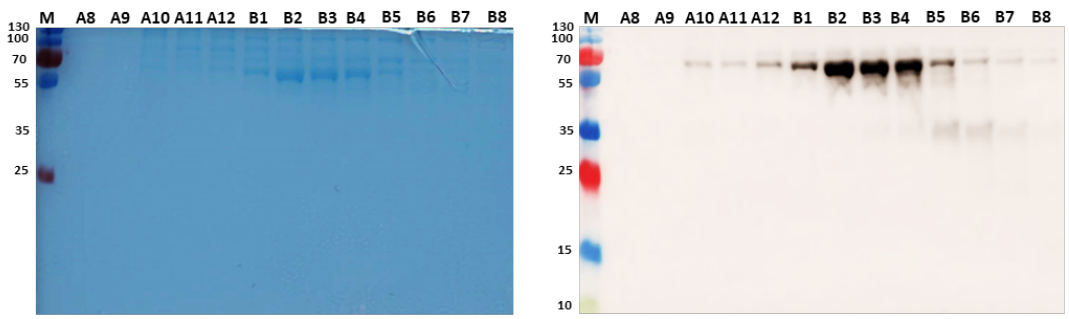

M A8 A9 A10 A11 A12 B1 B2 B3 B4 B5 B6 B7 B8
130
100
70
55
35
25
M A8 A9 A10 A11 A12 B1 B2 B3 B4 B5 B6 B7 B8
130
100
70
55
35
25
15
10

Figure 35: Optimizing buffer condition to enhance intracellular domain of PRG2 sample homogeneity.
A SEC chromatogram shows 280 nm wavelength of absorption for PRG2ΔN eluate in different buffer conditions. Changing pH level to 8,5 leads to separation of two distinct peaks at A10 and B3 fractions which is represented by green chromatogram. **B** Reloading B3 and A10 fractions on SEC to test sample stability. Distinct chromatogram obtained with a peak maxima at B3 fraction with reloaded B3 fraction. **C** SDS PAGE Coomassie and WB images show retention of purified PRG2ΔN in the peak of SEC fraction B1 to B5 with a maxima in B2 and B3 fractions also represent the most intense bands in gels near 55 kDa molecular weight marker. M, protein molecular weight marker.

3.2.15 Microscale Thermophoresis (MST)

3.2.15.1 PRG2, PRG2ΔCa, PRG2ΔCc, and PRG2ΔN interaction with PTEN

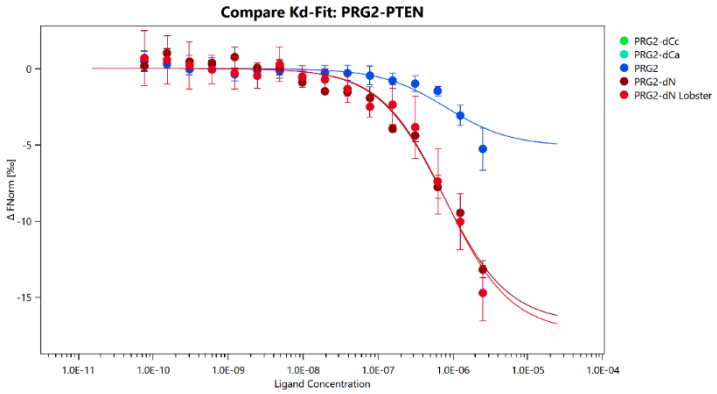
As discussed earlier, the purification method has to be stringent enough to selectively maximize the yield of protein of interest and minimize contamination of all other proteins, and mild enough to maintain native protein structure and direct interprotein interactions. All biophysical structural investigation experiments rely on the assumption that purified protein samples are pure and homogeneous, concentration is assessed precisely, and all of the protein is solubilized and in a natively active state. Therefore, a purified protein complex has to be evaluated for the final quality control to assess sequentially the essential properties of any protein sample: purity and integrity, homogeneity and activity [150].

In previous sections, sample purity, stability, and homogeneity assessments have been performed on purified full PRG2 and PRG2ΔN protein complexes. Hereby, it can find further analysis of functional assay to investigate the biological activity of purified protein complexes. MicroScale Thermophoresis (MST) method has been performed, which is a standard biomolecular interaction studies, to measure binding affinities of purified protein complexes with known physical interactors to investigate conservation of physiologically relevant biophysical structure and biochemical activities [151]. In recent years, MST was successfully employed to analyze a large variety of bio-molecular interactions including protein-protein [152-154], and protein-lipid interactions [155,156]. Hence, this method can be used to validate the biomolecular function of purified protein complexes.

In this section, the MST method has been performed to measure the binding affinity of purified native PRG2 complex to well-known interacting protein PTEN [58,61,62] to evaluate the conservation of physiologically active tertiary and quaternary protein structure. In this assay, I measured purified PRG2 (full protein 1-716), PRG2ΔCa (residues 1-438), PRG2ΔCc (residues 1-284), and PRG2ΔN (285-716) complexes with PTEN by using Monolith NT.115 (Nano-Temper Technologies). In this study, 20 μM purified PRG2, PRG2ΔCa, PRG2ΔCc, and PRG2ΔN complexes were fluorescently labeled with Cy5. Access dye was eliminated with gel filtration on a Superose 12 3.2/300 column (GE Healthcare) pre-equilibrated with 20 mM HEPES pH 7.4, 150 mM NaCl. The peak protein fraction was collected to perform MST measurements. MST was measured individually with 20 nM fluorescently labeled PRG2, PRG2ΔCa, PRG2ΔCc, and PRG2ΔN in 20 mM HEPES pH 7.4, 150 mM NaCl, 0.05 % (v/v) Tween 20 buffer with sixteen-fold serial dilution of 2.5 μM PTEN (Origene) ligand in 25 mM Tris-HCl pH 7.3, 100 mM glycine, 10 % (v/v) glycerol. The samples were loaded into standard capillaries, and MST was measured. Measurement was repeated seven times for PRG2, four times for PRG2ΔCc with four independent samples preparade, four times for PRG2ΔCa with three independent samples preparade, and three times with PRG2ΔN from E. coli BL21 sample, and six times with PRG2ΔN from E.coli Lobstr sample. Data analyzes were performed using the Nano-Temper analysis software MO.Affinity Analysis (Nano-Temper Technologies). The Kd constants between a protein and its ligand can be calculated using the standard fitting model derived from the law of mass action.

The apparent dissociation constant for PRG2-PTEN is 789 nM, and PRG2ΔN-PTEN is 732 and 791 for E. coli BL21 and E. coli Lobstr samples respectively (**Figure 36**). Whereas, there were no binding detected for PRG2ΔCa and PRG2ΔCc with PTEN. Taken together all these results, purified PRG2, as well as PRG2ΔN complex, makes direct physical interaction with PTEN protein. This is an indication for the purified complex maintains a physiologically active structure. Besides, this investigation indicates that the PTEN binding domain of PRG2 protein is in between 439 to 716 amino acid residues as there is no PTEN binding detected with PRG2ΔCa. In future studies, purified PRG2ΔN complex in HEK293S GnTI- cells MST analysis with PTEN protein

needs to be done to confirm PRG2 protein C-terminal end, downstream of Poly-E box, is in charge to make physical interaction with PTEN.



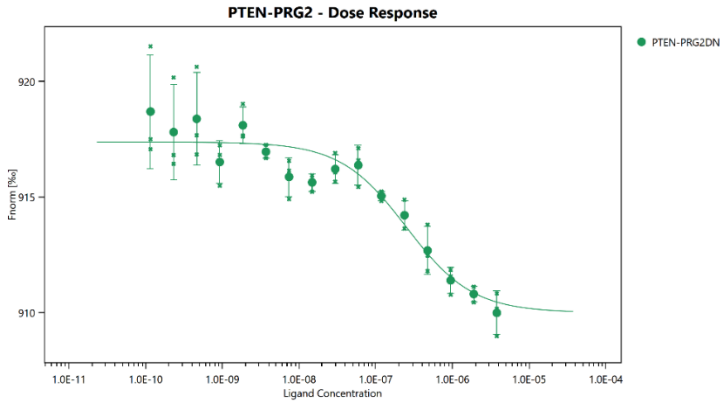
	Host	Kd (nM)	Std. Dev.	n
PRG2-PTEN	HEK293 GnTI-	798	±189	7
ΔCa-PTEN	HEK293 GnTI-	n.d.		4
ΔCc-PTEN	HEK293 GnTI-	n.d.		7
ΔN-PTEN	E.coli BL21	732	±73	3
ΔN-PTEN	E.coli Lobster	791	±75	6

Figure 36: PRG2-PTEN Microscale Thermophoresis.
Detection of nanomolar K_d of PRG2-PTEN and PRG2ΔN-PTEN interactions using the Monolith NT.115. PTEN at concentrations from 2.5 µM down to 798 nM was titrated against 20 nM of NT647-labeled PRG2 in A8-35 from HEK293S GnTI-, PRG2ΔN from E. coli BL21 and E. coli Lobstr. Plotting of the change in thermophoresis and concomitant fitting of the data yielded a K_d of 798 ± 189 nM for PRG2, 732 ± 73 nM for PRG2ΔN from E. coli BL21, and 791 ± 75 nM for PRG2ΔN from E. coli Lobstr.

To confirm the interaction between native PRG2ΔN complex and PTEN protein, I performed MST by chemically labeling 2 µM PTEN with the same Cy5 dye contains NHS-ester chemistry. Since PTEN is in incompatible Tris buffer for amine-reactive labeling, we exchanged buffer Tris buffer with 20 mM HEPES pH 7.4, 150 mM NaCl, 0.05 % (v/v) Tween 20 via SEC. After chemical labeling, access dye was eliminated and protein purified by gel filtration on a Superose 12 3.2/300 column (GE Healthcare) pre-equilibrated with 20 mM HEPES pH 7.4, 150 mM NaCl, 0.05 % (v/v) Tween 20. The peak protein fraction was collected to perform MST measurements.

MST was measured with 20 nM fluorescently labelled PTEN in 20 mM HEPES pH 7.4, 150 mM NaCl, 0.05 % (v/v) Tween 20. The samples were added to capillaries, and MST was measured. Sixteen-fold serial dilution of 7.6 µM PRG2ΔN from E. coli Lobstr ligand in 20 mM HEPES pH 7.4, 150 mM NaCl, 0.05 % (v/v) Tween 20 was prepared, mixed with a constant concentration of labeled PTEN. Molecules loaded into capillaries and analyzed in the instrument by subsequent scanning of each capillary. Measurements were repeated three times. The apparent dissociation constant for PTEN-PRG2ΔN is 247 nM (**Figure 37**). This result is a proof of concept for PTEN protein and purified

PRG2ΔN complex are direct physical interactors, and our purification pipeline maintains PRG2 PTEN binding property.



	Host	Kd (nM)	Std. Dev.	n
PTEN - ΔN	E.coli Lobster	247	±61	3

Figure 37: PTEN-PRG2ΔN Microscale Thermophoresis.
Detection of nanomolar K_d of PTEN-PRG2 interaction using the Monolith NT.115. PRG2ΔN at concentrations from 7.6 µM down to 0,232 nM was titrated against 20 nM of NT647-labeled PTEN. Plotting of the change in thermophoresis and concomitant fitting of the data yielded a K_d of 247 ± 61 nM.

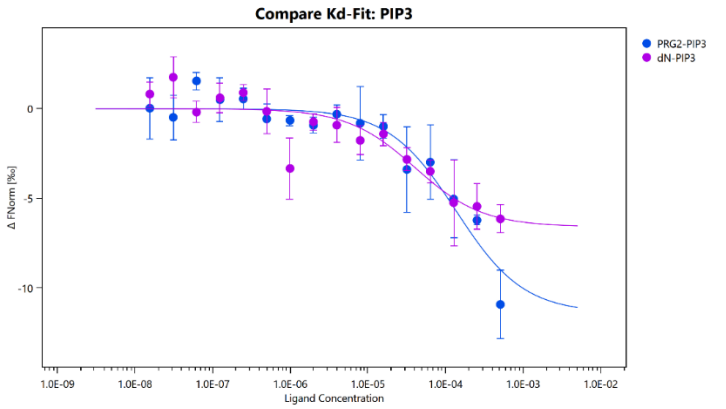
3.2.15.2 PRG2, PRG2ΔCa, PRG2ΔCc, and PRG2ΔN interaction with PI(4,5)P2 and PI(3,4,5)P2

After successfully applying the MST assay with purified PRG2 complexes, I performed further quantitative analysis. I further investigate PRG2-PIPs interaction to understand if PRG2 has any direct physical interaction with PIPs, if yes, which domain takes a role in this interaction. This in vitro investigation is out of the scope of this thesis, yet results will be novel and help scientists to better understand the molecular interaction of PRG2-PTEN-PIPs. MST measurement of PRG2, PRG2ΔCa, PRG2ΔCc, and PRG2ΔN with PI(4,5)P2 and PI(3,4,5)P2 was performed in the same way except standard capillaries replaced by hydrophobic capillaries to prevent surface adsorption of phosphatidylinositols. Fluorescently labeled PRG2, PRG2ΔCa, PRG2ΔCc, and PRG2ΔN concentrations were fixed to 20 nM, and the highest PI(4,5)P2 and PI(3,4,5)P2 concentrations were 584 to 510 µM respectively for sixteen-fold serial dilution in 20 mM HEPES pH 7.4, 150 mM NaCl, 0.05 % (v/v) Tween 20. Measurement was repeated four times for PRG2, six times for PRG2ΔCa, three times for PRG2ΔCc, and three times for PRG2ΔN from E. coli BL21. Measured dissociation constant for PRG2-PI(4,5)P2 is 239 µM, and PRG2-PI(3,4,5)P2 is 137 µM, and PRG2ΔN-PI(4,5)P2 is 17 µM, and PRG2ΔN-PI(3,4,5)P2 is 40 µM (**Figure 38**). There was no binding detected for PRG2ΔCa and

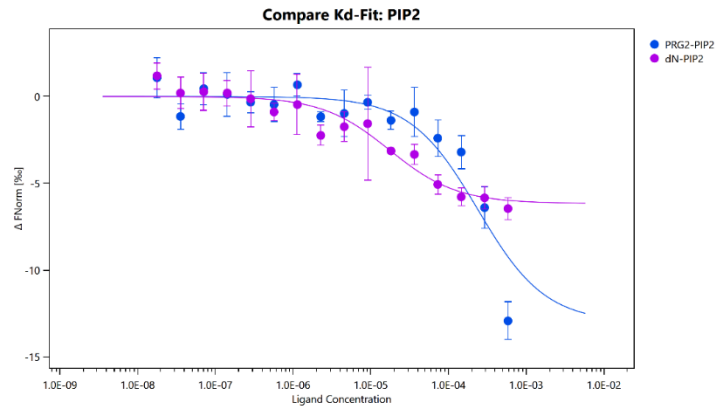
PRG2ΔCc with PI(4,5)P2 and PI(3,4,5)P2. This result indicates that the intracellular domain of PRG2 protein is in charge to make a physical interaction with PIPs in the inner leaflet of the plasma membrane.

To rule out unspecific binding affinity of PI(4,5)P2 and PI(3,4,5)P2 on A8-35, purified opsin protein in A8-35 was tested for MST measurement. Opsin was fluorescently labeled, and concentration was fixed to 20 nM. Buffer conditions and phosphatidylinositols concentrations were kept the same as the PRG2 experiment. Measurements were repeated two times for each experiment. There was no binding detected between opsin and phosphatidylinositols (data has not been shown).

Besides, to rule out electrostatic interaction of negatively charged PI(4,5)P2 and PI(3,4,5)P2 with positively charged His-tag, we tested MST measurement with His tagged phytochrome AGP2 protein with PI(4,5)P2 and PI(3,4,5)P2 . AGP2 was fluorescently labeled and concentration was fixed to 20 nM. Buffer conditions and phosphatidylinositols concentrations were kept the same as the PRG2 experiment. Measurements were repeated two times for each experiment. There was no binding detected between AGP2 and phosphatidylinositols (data has not been shown).



	Host	Kd (µM)	Std. Dev.	n
PRG2-PIP3	HEK293 GnTI-	137	±25	4
ΔCa-PIP3	HEK293 GnTI-	n.d.		6
ΔCc-PIP3	HEK293 GnTI-	n.d.		3
ΔN-PIP3	E.coli BL21	40	±15	3



	Host	Kd (µM)	Std. Dev.	n
PRG2-PIP2	HEK293 GnTI-	239	±57	4
ΔCa-PIP2	HEK293 GnTI-	n.d.		6
ΔCc-PIP2	HEK293 GnTI-	n.d.		3
ΔN-PIP2	E.coli BL21	17	±4	3

Figure 38: PRG2-Phosphatidylinositols Microscale Thermophoresis.
Detection of micromolar Kd of PRG2-Phosphatidylinositols interaction using the Monolith NT.115. PI(4,5)P2 and PI(3,4,5)P2 individually at concentrations from 584 µM down to 17,8 nM, and 510 µM down to 15,6 nM respectively, was titrated against 20 nM of NT647-labeled PRG2, PRG2ΔCa, PRG2ΔCc in A8-35, and PRG2ΔN. Plotting of the change in thermophoresis and concomitant fitting of the data yielded Kd of 239 ± 57.1 µM for PRG2-PI(4,5)P2 , 137 ± 25.7 µM for PRG2-PI(3,4,5)P2 , 17 ± 4 µM for PRG2ΔN-PI(4,5)P2 , and 40 ± 15 µM for PRG2ΔN-PI(3,4,5)P2 .

3.2.15.3 Investigation of PRG2-PTEN binding domain

In order to confirm our findings of PRG2-PTEN interaction, and pinpoint where PTEN is binding on PRG2, I performed a dot blot analysis. According to our MST findings, PTEN binds to the cytoplasmic domain of PRG2. In our initial attempt, I tested this result by spotting purified PRG2, PRG2ΔN, and PRG2ΔCa on the nitrocellulose membrane. Then the membrane was blocked and incubated with PTEN protein. After washing, blot coupled with PTEN rabbit antibody, then rabbit HRP antibody for ECL detection. The result indicates that PTEN binds on full PRG2, and PRG2ΔN, but not on PRG2ΔCa (**Figure 39**), which is a parallel finding with MST results in section 3.2.15.1.

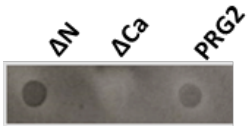


Figure 39: Investigation of PRG2-PTEN binding domain via dot blot.
Dot blot image of purified PRG2, PRG2ΔC, and PRG2ΔCa spots on nitrocellulose membrane coupled with PTEN protein, and detection made by PTEN antibody to identify PTEN binding partners. Depicted blot shows that both PRG2 and PRG2ΔC interact with PTEN.

In the next step, designed 8 short polypeptides with 34 amino acid long covers the cytoplasmic domain of PRG2 (residues 435 to 672) (**Table 10**) have been used to perform dot blot analysis with PTEN protein to narrow down PRG2-PTEN binding pocket. Each peptide was spotted on nitrocellulose membrane. As a positive control, PTEN protein was also spotted on the membrane. Then, the assay was performed with PTEN protein coupled with rabbit PTEN antibody and detected by rabbit HRP. The result indicates a faint signal on peptide number 26 that covers the poly-E box (**Figure 40**). In section 3.2.15.1, we find out that the PTEN binding domain of PRG2 is in between 439 to 716 amino acids. This result confirms this finding as peptide 24 covers 435 to 472 amino acid sequences of PRG2 protein.

Table 10: List of polypeptides which covers downstream of PRG2 intracellular domain from poly-E box.

Peptide number	Peptide amino acid sequence
26	EQVAEEEEEEEEEEEEEEEEEEEGPVPPSLYPTVQA
27	RPGLGPRVILPPRPGPQPLVHIPEEGVQAGA
28	GLSPKSSSSSVRAKWLSMAEKGGGPVAVAPSQPR
29	VANPPRLLQVIAMSKAAGGPKAETASSSSASSDS
30	SQYRSPSDRDSASIVTIDAHAPHHPVVHLSAGST
31	PWEWKAKVVEGEGSYELGDLARGFRSSCKQPGMG
32	PGSPVSDVDQEEPRFGAVATVNLATGEGLPPPGA
33	SEGALGAGSRESTLRRQVGGLAEREVEAEAESYY

MLAMKEKNKTPKDSMTLLPCFYFVELPIVASSIVSLYFLELTDLFKPAKVGFQCYDRALSMPYVETNEEL
IPLLMLLSLAFAAPAASIMVGEGMVYCLQSRLWGRGPGGVEGSINAGGCNFNSFLRRTVRFVGVHVFGLC
ATALVTDVIQLATGYHTPFFLTVCKPNYTLLGTSCESNPYITQDICSGHDTHAILSARKTFPSQHATLSA
FAAVYVSMYFNAVISDTTKLLKPILVFAFAIAAGVCGLTQITQYRSHPVDVYAGFLIGAGIAAYLACHAV
GNFQAPPAEKVPTPAPAKDALRALTQRGHESMYQQNKSVSTDELGPPGRLEGVPRPVAREKTSLGSLKRA
SVDVDLLAPRSPMGKEGMVTFSNTLPRVSTPSLDDPARRHMTIHVPLDASRSRQLIGEWKQKSLEGRGLG
LPDEASPVHLRAPA EQVAEEEEEEEEEEEEEEEEEGPVPPSLYPTVQA RPGLGPRVILPPRPGPQ
#26 #27
PLVHIPEEGVQAGA GLSPKSSSSSVRAKWLSMAEKGGGPVAVAPSQPR VANPPRLLQVIAMSKAAGGP
#28 #29
KAETASSSSASSDS SQYRSPSDRDSASIVTIDAHAPHHPVVHLSAGST PWEWKAKVVEGEGSYELGDL
#30 #31
ARGFRSSCKQPGMG PGSPVSDVDQEEPRFGAVATVNLATGEGLPPPGA SEGALGAGSRESTLRRQVGG
#32 #33
LAEREVEAEAESYY RRMQARRYQD

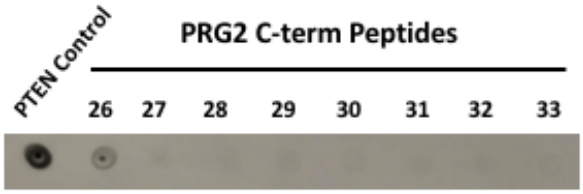


Figure 40: Narrowing down PRG2-PTEN binding domain via dot blot.
Dot blot image of peptides, which covers poly-E box downstream of intra cellular domain of PRG2, spots on nitrocellulose membrane coupled with PTEN protein, and detection made by PTEN antibody to identify PTEN binding domain. Blot exhibits positive signal on PTEN control, and faint signal from peptide number 26.

4. Discussion

Membrane proteins account for approximately one-third of the entire proteome in mammalian organisms. Additionally, they play crucial roles in many fundamental cell processes. Nevertheless, merely a minor fraction of membrane proteins' 3D structures have been identified. For example, as of September 2020, 49307 human protein structures deposited in the protein data bank (PDB), and only 514 are transmembrane protein. The main reason for this gap and unmet need is due to technical challenges associated with difficulties of membrane protein extraction, solubilization, and purification under native conditions. One of the main hurdles is the low abundance of IMP in cells, as presence is limited with the occupancy of the physical capacity of the cell membrane. Other challenges can be listed as; low expression of IMPs, truncations, misfolding, aggregation, missing posttranslational modifications, and mislocalization. Moreover, overexpression of membrane proteins can be toxic to the cell.

In this thesis, I aim to investigate the molecular structure of an integral membrane protein PRG2 in a near-native complex. Hence, I could contribute to filling the aforementioned gap by using in-house state-of-the-art biophysical and biochemicals methods with subject matter experts and key opinion leaders in the field. To begin with, I embarked on establishing a pipeline for overexpression and purification of the native PRG2 protein complex. This journey includes empirical multistep tests, systematic screening, optimizations, and quality control checks.

The expression level of native PRG2 protein in rodents' brain is insufficiently abundant to isolate material for biochemical and structural studies, therefore, it has to be overexpressed in heterologous systems by using recombinant DNA technology. Since there is no universal expression system for all membrane proteins, it requires a trial and error approach. Thus, I tested several mammalian cell-lines to overexpress PRG2. The main reason for testing mammalian cell expression systems is to maintain accurate protein folding with molecular chaperones and post-translational modifications (PTMs), which are vital for protein structure and function [70, 71]. The highest PRG2 expression level was observed with the HEK293S GnTI$^-$ cell line, which is lacking N-acetylglucosaminyltransferase I (GnTI) (**Figure 8**), hence missing complex N-glycans post-translational modification [25]. The main reason to establish this cell line is to

minimize the large diversity of glycosylation, which interferes with the purification as well as the crystallization process. Moreover, this cell line also increases the yield of protein expression and leads to preferential IMPs localization in the plasma membrane. It has been successfully employed to overexpress a wide variety of mammalian membrane proteins [77–80]. The transient transfection method has been chosen over stable cell-line creation to mitigate the tangible risk of unexpected mutagenesis, the silence of gene expression, alteration of genomic integrity by being in culture for long passages can be critical for protein structure. PEI as a chemical transient transfection reagent was prepared in-house and performed efficiently compared to commercial transfection reagents (**Figure 9**). As the PEI transient transfection method is cost-effective, it has been used for large-scale PRG2 overexpression in adherent cells.

Initial and the most critical steps in IMP purification in the native condition is the solubilization of the membrane to extract the protein of interest in an active form, maintaining physiological structure for subsequent purification steps, structural and functional analysis [97]. Detergents can keep IMP in solution, and avoid precipitation, however, they are also denaturant. Therefore, using detergents that are mild non-denaturing, yet stringent enough to extract hydrophobic TM domains from the plasma membrane is an essential key to success. These detergents maintain protein-protein interactions while braking lipid–lipid interactions and lipid-protein interactions without affecting the protein's structural features, such that it can be isolated in its biologically active form [98]. N-dodecyl-β-d-maltoside (DDM) is the most commonly used detergent in membrane protein purification [99]. Therefore, we used DDM to establish our purification pipeline optimization. However, when I load each fraction of immunoprecipitated PRG2 protein, the vast majority was observed in the pellet fraction (**Figure 14**). After screening commercially available non-ionic and zwitterionic mild detergents, Fos-Choline-14 has a fifty-fold higher solubilization rate than DDM (Figure 15). Therefore, Fos-Choline-14 was used in the initial solubilization step to maximize the PRG2 extraction yield. After IP, detergent has been exchanged with the amphipathic polymer, amphipol A8-35 to maintain solubility and eliminate aggregation in SEC (**Figure 16B**).

After establishing the PRG2 purification pipeline in the native complex, the next step was to investigate molecular weight, homogeneity, and stoichiometry of the complex. Native polyacrylamide gel electrophoresis (Native PAGE) analysis indicates multiple degrees of complex formation in dimeric, tetrameric, and octameric stages of PRG2 multimer in addition to the most abundant monomeric form of PRG2 (**Figure 17**). Multimerization of integral transmembrane proteins are common phenomena, such as oligomerization of 7-transmembrane G protein-coupled receptors (GPCRs) regulates the pharmacological characteristics of the receptors and influence their coupling to G proteins [111,112], and transient receptor potential (TRP) proteins, which are 6-transmembrane ion channels assemble in homo- and hetero-tetrameric complexes to enhance ion permeation property [113,114]. As seen in the examples, IMPs with multiple TM domains multimerize to perform functional implications in the cell membrane. This result also aligned with our extraction and purification pipeline is strong enough to extract PRG2 complex from the cellular membrane, and soft enough to maintain native complex integrity.

In order to utilize the single-particle EM method, the purified sample has to ensure homogeneity. The above established native PRG2 protein purification pipeline yields multiple degrees of multimeric PRG2 protein classes. Ideally, these multiple classes of native PRG2 complexes have to be reduced in one class to perform single-particle EM analysis. Therefore, an additional analytical SEC separation method was applied. After the SEC, the native PRG2 protein complex has been observed as a smear in denaturing SDS PAGE WB analysis in the range of dimmer to tetramer (**Figure 18B**). At this stage, loading one of the PRG2 complex containing individual fraction in the SEC as a tandem lading fashion results in the loss of chromatogram peak. This result is a clear indication of the destabilization of the complex. Additionally, photo crosslinking with UV light and chemical crosslinking with formaldehyde and glutaraldehyde have been tested to investigate if crosslinking could enhance complex stability and help us to distinguish individual native PRG2 complex classes by using SEC. However, all these methods result in protein aggregation. GraFix (Gradient Fixation), which is an intramolecular chemical crosslinking with a density gradient ultracentrifugation method [118], has also been tested for crosslinking PRG2 proteins in their native complex. This method also failed to separate individual native PRG2 complexes.

There were multiple classes of purified native PRG2 complexes before and after the analytical SEC. In order to confirm that all identified PRG2 complex classes consist of pure PRG2 protein, mass spectrometry analysis has been performed. All identified protein bands in Coomassie gel are also identified in WB gel (**Figure 20A**). This is a clear indication of sample purity as all bands' profiles are having the same fingerprint. Excision of seven explicitly visible bands for LC-ESI-MS analysis indicates that PRG2 peptides are obtained as the highest hit for all bands. This result is a piece of solid evidence as proof of the success of an established purification pipeline and confirms native PRG2 establishes a various degree of homomultimerization. Also, the B7 fraction (**Figure 18A**) has been analyzed for initial negative stain EM investigation to pave the way in molecular morphology, organization, and heterogeneity of the sample. Negative stained EM chromatograms show a high-quality image of successful negative stain with adequate contrast between particle to grid background and optimum particle density (**Figure 19**). Most particles fall in eight-fold symmetric bagel shaped particles with 12 nm size in diameter. These particles could be tetramer of PRG2 homomultimer with a top angle view (**Figure 19C**). All 138 particle classes were used to build up the initial 3D model generation. However, since there is a wide range of particle heterogeneity between class averages, an initial low-resolution 3D model generation attempt could not be achieved. Besides, the same sample was prepared with direct cryogenic freezing of purified native PRG2 complex to eliminate any possible artifacts that may arise due to negative staining. Unfortunately, these grids were not subjected to cryo-EM image acquisition as we did not achieve an ideal sample homogeneity in our established purification pipeline. Nevertheless, these grids are kept in cryo storage and can be subjected to single-particle cryo-EM analysis in the future.

Proximity ligation assay has been performed to extend our understanding of PRG2 homomerization and pinpoint the cellular location of these multimers. In Figure 21, the confocal image shows red puncta signal to indicate PRG2 homomultimerization which has been observed on the plasma membrane. In this assay, PRG2-PRG3 interaction has also been observed similarly to PRG2 homomultimerization. Functional cooperation between PRG protein family members could exert a specific role in the cell.

In conclusion, an established pipeline is powerful enough to extracts, solubilize, and purify full PRG2 protein under the native complex. Nevertheless, the aforementioned findings indicate that the native PRG2 complex consists of various degrees of homomultimers. As further physical separation nor crosslinking stabilization methods did not achieve to obtain an individual class of multimers, the single-particle EM method could not provide us a piece of insightful structural information. Besides, purification yield was in the sub-microgram level, therefore, protein crystallization could not be an alternative option for this endeavor. Hence, changing the strategy to purify the water-soluble intracellular domain of PRG2 protein under native condition was a foreseeable approach. On this strategy, E. coli has been used to overexpress, purify, and characterize water-soluble mouse PRG2 C-terminal intracellular domain (PRG2ΔN) under near-native condition. Switching to bacteria allows us to increase the yield of purified protein together with the elimination of complex detergent solubilization steps. However, this approach is missing eukaryotic molecular chaperons, complex glocalization patterns, and posttranslational modifications which are crucial for both protein structure and function. Therefore, I call purified complex as “near-native” since protein expression conditions are not the same as using the mammalian protein expression system. The main aim of this approach is to purify the homomultimeric complex for single-particle cryo-EM and X-ray crystallography to investigate the 3D molecular structure of the cytoplasmic domain of PRG2 protein.

To establish a new pipeline, initially, PRG2ΔN has been overexpressed in standard E. coli BL-21 strain followed by protein extraction and multi-step purification including Nickel pull-down assay, size exclusion chromatography, anion exchange chromatography, and gel filtration. Eventually, a 10 mg/ml concentration of protein purified. At a Coomassie-stained gel, there was a defined band at 55 kDa weight (**Figure 22D**). I observed a single band in B1 fraction with the size of PRG2ΔN, whereas in B7 fraction we observed an additional well-defined band at 70 kDa weigh with 1:1 stoichiometry. Both fractions were subjected to negative stain EM studies for testing sample quality, particle density, homogeneity, and clusters. Both B1 and B7 fractions were the peaks of two distinct curves on the final SEC chromatogram. In parallel, Mass Spectrometry analysis has been done to reveal the purified protein profile. The most prominent bands were investigated by Matrix-Assisted Laser Desorption/Ionization

Mass Spectrometry (MALDI-MS). MS analysis reviled that, purified protein complex contains prokaryotic proteins, which are PutA (UniProt Reference P09546), Odo1 (UniProt Reference P0AFG3), ArnA (UniProt Reference P77398), YihW (UniProt Reference P32144), and SlyD (UniProt Reference P0A9K9) (**Figure 23**) in addition to the cytoplasmic domain of mouse PRG2. Two of these prokaryotic proteins, ArnA, and SlyD are well-known metal-binding proteins, therefore both protein able to bind on the Nickel column and co-purified with His-tagged protein of interest [132].

Going back to negative stain EM analysis, the micrograph shows that the B7 fraction (**Figure 24B**) has a more defined particle profile comparing to the B1 micrograph (**Figure 24A**). As we know the sample is contaminated with ArnA protein (PDB #4wkg), we build up in-silico 2D projection views (**Figure 24D**) to eliminate ArnA particles from our class averages. We encountered that all well-defined particles are matching with the known native structure of the trimeric form of ArnA homodimer, and the remaining particles are inhomogeneous.

This purification pipeline includes multiple steps purification methods includes metal binding affinity assay, size exclusion, and ionic charge chromatography. As biophysical characteristics of both SlyD and ArnA contaminants are diverse than PRG2ΔN in molecular weight and ionic charge, it would be expected that both contaminant proteins were eliminated from PRG2ΔN. As this was not the case, ArnA and SlyD could also be physical interactors of the PRG2ΔN protein complex. There are a couple of strategies to overcome these prokaryotic metal-binding contaminants; one of them is using 8xHis-tag instead 6xHis to increase the metal binding affinity of over-expressed protein and increase imidazole concentration to wash out contaminants including prokaryotic metal-binding proteins. The next strategy is using a stronger promoter plasmid to increase the over-expression yield of protein of interest to eliminate a similar abundance of a protein of interest with ArnA and SlyD leading competition on nickel column binding. Another alternative solution is to eliminate or minimize ArnA, and SlyD contaminations by using an alternative E. coli strain. In this strategy, the established pipeline has been tested by using commercial E. coli strain BL21 (DE3) Lobstr (Low background strain) which is knocked out for slyD, and deficient for arnA genes [132-134].

Utilizing this strategy resulted in increased purification yield of PRG2ΔN, minimize ArnA protein, and eliminate SlyD contaminations (**Figure 25D**) when we compare to E. coli BL21 (DE3) strain (**Figure 22D**). Purified PRG2ΔN molecular weight, homogeneity, and complex stoichiometry investigated by Native PAGE analysis and Dynamic Light Scattering. Native Page experiment indicates heptameric/octameric complex formation and monomeric form with purified PRG2ΔN (**Figure 28**). To confirm this finding, the sample was also investigated by DSL as a complementary analysis. The result highlights polydispersity in the sample (**Figure 29A**), which is a parallel result with Native PAGE that was observed. 99% of all classes of heterogenic particle species are falling at 350 to 400 kDa molecular weight (**Figure 29B**) represents heptameric/octameric complex form similar to Native PAGE analysis (Figure 28). In conclusion, the intracellular domain of PRG2 also tends to form homomultimerization to maintain stability and lower the free energy level, and the intracellular domain of PRG2 protein is able to form homomultimerization. Purified PRG2ΔN from E. coli BL21 (DE3) strain investigated by negative stain EM resulted in a wide range of particle sizes and shapes heterogeneity (**Figure 30**). This result highlighted the fact that PRG2ΔN establishes a range of homomultimers. Native PAGE and DLS results indicate that majority of purified protein is heptameric/octameric complex size and monomer, negative stain EM micrograph contains a higher degree of heterogeneity. To eliminate possible negative effects of low pH level (4) negative staining reagent Uranyl Acetate, alternative staining reagents Uranyl Formate pH 4,5, Uranylless pH 4,5, Ammonium Molbidate pH 5, and Phosphotungstic Acid pH 7 were tested. EM micrographs indicate no improvements in image quality and particle homogeneity (**Figure 31A-D**).

Consequently, PRG2ΔN overexpressed in HEK293S GnTI- cells to maintain relevant protein folding and post-translational modifications to test the improvement of purified protein complex particles in terms of stability, density, definition, homogeneity, and monodispersity in negative stain EM analysis. However, as a future prospect, it is worth testing the standard HEK293T cell line to perform an established purification pipeline to maintain native glycosylation. Purified complex analyzed by both negative stain EM, and mass spectrometer. As expected, peptides are identified as PRG2ΔN by MALDI-MS analysis including chaperone and heat-shock proteins, which are commonly found contaminants in the exogenous protein over-expressions (**Table 9**). This result assures

an established purification pipeline can yield purified PRG2ΔN complex. Negative stain EM micrograph indicates two well-defined but distinct populations of large particles, one of which is the ring structure as similar to full PRG2 results, and the other is a rather small particle with a circular or triangular ring pattern (**Figure 34**). Since particle density, signal to noise level, homogeneity of individual particle classes in micrographs are adequate, single-particle image processing and analysis need to be done in the future.

In addition to that, there was a fusion of two peaks in SEC chromatograph from purified PRG2ΔN in HEK293S GnTI- cells (**Figure 33A**). In order to improve purification by separating these two peaks, alternative buffer conditions have been tested. HEPES buffer at pH 8,5 improves the separation of these two peaks, which may contribute to eliminating possible sample heterogeneity due to a mixture of two possible distinct particle classes. To test sample homogeneity, the best fraction (**Figure 35**) has been loaded on a carbon grid. Negative stain EM image acquisition, single-particle image processing, and image analysis need to be done in the future.

To go for the extra mile, purified PRG2 proteins, both full PRG2, intracellular domain PRG2ΔC, and cytoplasmic domain PRG2ΔN have been subjected to a biophysical test to analyze the functionality of these purified proteins. Ideal purification method expected to be stringent enough to purify proteins from other cellular components and proteome, yet mild enough to maintain native complex to conserve biologically active 3D structure. Purified protein complexes sample purity, stability, and homogeneity assessments have been performed via biochemical and biophysical assays. Next, the Micro Scale Thermophoresis (MST) method has been performed as a functional assay to investigate the biological activity of purified protein complexes. In this assay, we could measure binding affinities of purified PRG2, PRG2ΔC and PRG2ΔN protein complexes with known physical interactors PTEN to investigate the conservation of physiologically relevant biophysical structure and biochemical activities [61, 62]. Additionally, we investigate lipid binding affinity by testing PIPs (PI(4,5)P2 and PI(3,4,5)P2). PTEN functions predominantly via transient localization to the plasma membrane, by dephosphorylating PI(3,4,5)P3 to PI(4,5)P2 [63-65], and PIPs gradient has been postulated to regulate cell polarity and migration [68]. In vivo studies in rodent and cell culture studies have been shown that local inhibition of PTEN activity with PRG2, could

lead inhibition of PTEN in a close vicinity of PRG2 accommodating nanodomains of the plasma membrane to initiate PI(3,4,5)P3 accumulations in these nanodomains [62]

MST assay indicates parallel result with literature, PRG2 and PTEN are physical interactors, additionally, PRG2ΔN is able to interact with PTEN with a similar extend of affinity in the nanomolar range, which is a strong molecular interaction range (**Figure 36**). While there is no binding detected for the integral membrane domain of purified PRG2 (PRG2ΔCa and PRG2ΔCc) with PTEN (**Figure 36**). As a proof of concept study, dot blot assay has been conducted and results indicate parallel findings with MST where PRG2, PRG2ΔN can bind on PTEN but PRG2ΔCa can't (**Figure 39**). Both results aligned with PTEN binding domain of PRG2 protein could be in between 439 to 716 amino acid residues which locates in the cytoplasm. When we performed dot blot analysis with short peptides cover covers the cytoplasmic domain of PRG2 (residues 435 to 672) as indicated in Table 10, peptide covers the poly-E box shows PTEN binding (**Figure 40**). Taken together, these results claim that purified complex is not only pure, but also structurally well preserved, and able to bind a known physical interactor PTEN. Moreover, the direct PTEN binding domain could be localized in the cytoplasmic domain. Another novelty of this finding indicates precise disassociation constant of PRG2 and PTEN protein is 200 to 800 nM range (**Figure 36-37**).

Besides, similar to PTEN results, full PRG2 and cytoplasmic domain (PRG2ΔN) have interacted with PIPs (PI(4,5)P2 and PI(3,4,5)P2) in the micromolar range (15 to 250 μM), and there was no binding detected with purified integral membrane domain (PRG2ΔC). This result indicates that the intracellular domain of PRG2 protein could be in charge to make a physical interaction with PIPs in the inner leaflet of the plasma membrane where potentially can create steric hindrance to restrict their accessibility to PTEN. Therefore, may lead to local nanodomains of PI(3,4,5)P3 accumulations along the axonal plasma membrane and facilitating the formation of axonal branches. This mechanism could be an additional paradigm to hypothesized inhibition of local PTEN dephosphorylation activity of PIPs via direct interaction with PRG2 which facilitates axonal filopodia protrusion. Eventually, either one or both mechanisms could take a part in orchestration for the formation of axonal filopodia at these nanodomains.

Acknowledgment

This dissertation was developed during my work as a research assistant at the Institute of Biochemistry and the Institute of Medical Physics and Biophysics of the Charité - Universitätsmedizin Berlin. During the entire study time, I was comprehensively supervised by Dr. George Leondaritis and Dr. Michal Szczepek, and signs of progress were discussed with my thesis advisory committee members Prof. Dr. Britta Eickholt, Prof. Dr. Christian Spahn, Dr. Patrick Scheerer. Additionally, I gained the advantage of EM image analysis expertise of Dr. Justus Loerke, EM grid preparation expertise from Jörg Bürger, technical assistance from Brian Bauer and Anja Koch. Thank you all, and I wish you all the best.

My sincere thanks go to my dear friends Dr. Joel Kaufmann, and Birgit Schroeer, it was such a great pleasure to share the office space with you, I will never forget our daily talks about life, politics, sports, art, and science. Last but not least, thank you Dr. Dennis Kwiatkowski to have a joyful popular science talk and discuss alternative career goals and opportunities outside of the academy.

References

1. Almén, M.S., Nordström, K.J., Fredriksson, R. and Schiöth, H.B., 2009. Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC biology*, *7*(1), p.50.

2. Wallin, E. and Heijne, G.V., 1998. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Science*, *7*(4), pp.1029-1038.

3. Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W. and Funke, R., 2001. Consortium IHGS. *Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, et al: Initial sequencing and analysis of the human genome. Nature*, *409*, pp.860-921.

4. Rawlings, A.E., 2016. Membrane proteins: always an insoluble problem? *Biochemical Society Transactions*, *44*(3), pp.790-795.

5. Kim, M.S., Pinto, S.M., Getnet, D., Nirujogi, R.S., Manda, S.S., Chaerkady, R., Madugundu, A.K., Kelkar, D.S., Isserlin, R., Jain, S. and Thomas, J.K., 2014. A draft map of the human proteome. *Nature*, *509*(7502), p.575.

6. Bill, R.M., Henderson, P.J., Iwata, S., Kunji, E.R., Michel, H., Neutze, R., Newstead, S., Poolman, B., Tate, C.G. and Vogel, H., 2011. Overcoming barriers to membrane protein structure determination. *Nature biotechnology*, *29*(4), p.335.

7. Spiegel, A.M. and Weinstein, L.S., 2004. Inherited diseases involving G proteins and G protein–coupled receptors. *Annu. Rev. Med.*, *55*, pp.27-39.

8. Spiegelberg, B.D. and Hamm, H.E., 2007. Roles of G-protein-coupled receptor signaling in cancer biology and gene transcription. *Current opinion in genetics & development*, *17*(1), pp.40-44.

9. Ng, D.P., Poulsen, B.E. and Deber, C.M., 2012. Membrane protein misassembly in disease. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, *1818*(4), pp.1115-1122.

10. Overington, J.P., Al-Lazikani, B. and Hopkins, A.L., 2006. How many drug targets are there?. *Nature reviews Drug discovery*, *5*(12), p.993.

11. Deisenhofer, J., Epp, O., Miki, K., Huber, R. and Michel, H., 1985. Structure of the protein subunits in the photosynthetic reaction centre of Rhodopseudomonas viridis at 3Å resolution. *Nature*, *318*(6047), p.618.

12. Moraes, I., Evans, G., Sanchez-Weatherby, J., Newstead, S. and Stewart, P.D.S., 2014. Membrane protein structure determination—the next generation. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, *1838*(1), pp.78-87.

13. Gutmann, D.A., Mizohata, E., Newstead, S., Ferrandon, S., Henderson, P.J., Van Veen, H.W. and Byrne, B., 2007. A high-throughput method for membrane protein solubility screening: The ultracentrifugation dispersity sedimentation assay. *Protein science*, *16*(7), pp.1422-1428.

14. Niesen, F.H., Berglund, H. and Vedadi, M., 2007. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature protocols*, *2*(9), p.2212.

15. Postis, V.L., Deacon, S.E., Roach, P.C., Wright, G.S., Xia, X., Wright, G.S., Xia, X., Ingram, J.C., Hadden, J.M., Henderson, P.J. and Phillips, S.E., 2008. A high-throughput assay of membrane protein stability. *Molecular membrane biology*, *25*(8), pp.617-624.

16. Vergis, J.M., Purdy, M.D. and Wiener, M.C., 2010. A high-throughput differential filtration assay to screen and select detergents for membrane proteins. *Analytical biochemistry*, *407*(1), pp.1-11.

17. le Maire, M., Champeil, P. and Møller, J.V., 2000. Interaction of membrane proteins and lipids with solubilizing detergents. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, *1508*(1-2), pp.86-111.

18. Helenius, A. and Simons, K.A.I., 1975. Solubilization of membranes by detergents. *Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes*, *415*(1), pp.29-79.

19. Garavito, R.M. and Ferguson-Miller, S., 2001. Detergents as tools in membrane biochemistry. *Journal of Biological Chemistry*, *276*(35), pp.32403-32406.

20. Tribet, C., Audebert, R. and Popot, J.L., 1996. Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. *Proceedings of the National Academy of Sciences*, *93*(26), pp.15047-15050.

21. Whiles, J.A., Deems, R., Vold, R.R. and Dennis, E.A., 2002. Bicelles in structure–function studies of membrane-associated proteins. *Bioorganic chemistry*, *30*(6), pp.431-442.

22. Inagaki, S., Ghirlando, R. and Grisshammer, R., 2013. Biophysical characterization of membrane proteins in nanodiscs. *Methods*, *59*(3), pp.287-300.

23. Perez-Aguilar, J.M. and Saven, J.G., 2012. Computational design of membrane proteins. *Structure*, *20*(1), pp.5-14.

24. Popot, J.L., 2010. Amphipols, nanodiscs, and fluorinated surfactants: three nonconventional approaches to studying membrane proteins in aqueous solutions. *Annual review of biochemistry*, *79*, pp.737-775.

25. Reeves, P.J., Callewaert, N., Contreras, R. and Khorana, H.G., 2002. Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proceedings of the National Academy of Sciences*, *99*(21), pp.13419-13424.

26. Stetefeld, J., McKenna, S.A. and Patel, T.R., 2016. Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophysical reviews*, *8*(4), pp.409-427.

27. Asmari, M., Ratih, R., Alhazmi, H.A. and El Deeb, S., 2018. Thermophoresis for characterizing biomolecular interaction. *Methods*, *146*, pp.107-119.

28. Mueller, A.M., Breitsprecher, D., Duhr, S., Baaske, P., Schubert, T. and Längst, G., 2017. Microscale thermophoresis: a rapid and precise method to quantify protein–nucleic acid interactions in solution. In *Functional Genomics* (pp. 151-164). Humana Press, New York, NY.

29. Filarsky, M., Zillner, K., Araya, I., Villar-Garea, A., Merkl, R., Längst, G. and Németh, A., 2015. The extended AT-hook is a novel RNA binding motif. *RNA biology*, *12*(8), pp.864-876.

30. Seidel, S.A., Dijkman, P.M., Lea, W.A., van den Bogaart, G., Jerabek-Willemsen, M., Lazic, A., Joseph, J.S., Srinivasan, P., Baaske, P., Simeonov, A. and Katritch, I., 2013. Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions. *Methods*, *59*(3), pp.301-315.

31. Jerabek-Willemsen, M., Wienken, C.J., Braun, D., Baaske, P. and Duhr, S., 2011. Molecular interaction studies using microscale thermophoresis. *Assay and drug development technologies*, *9*(4), pp.342-353.

32. Niesen, F.H., Berglund, H. and Vedadi, M., 2007. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature protocols*, *2*(9), p.2212.

33. Lo, M.C., Aulabaugh, A., Jin, G., Cowling, R., Bard, J., Malamas, M. and Ellestad, G., 2004. Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Analytical biochemistry*, *332*(1), pp.153-159.

34. Alexandrov, A.I., Mileni, M., Chien, E.Y., Hanson, M.A. and Stevens, R.C., 2008. Microscale fluorescent thermal stability assay for membrane proteins. *Structure*, *16*(3), pp.351-359.

35. Gill, S.C. and Von Hippel, P.H., 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Analytical biochemistry*, *182*(2), pp.319-326.

36. Poole C. F., Poole S. K., 1991. Chromatography Today. Elsevier Science.

37. GE Healthcare, 2007. Ion Exchange Chromatography and Chromatofocusing: Principles and Methods, GE Healthcare.

38. Ayyar, B.V., Arora, S., Murphy, C. and O'Kennedy, R., 2012. Affinity chromatography as a tool for antibody purification. *Methods*, *56*(2), pp.116-129.

39. Walters, R.R., 1985. Affinity chromatography. *Analytical chemistry*, *57*(11), pp.1099A-1114A.

40. Hage, D.S., Anguizola, J.A., Bi, C., Li, R., Matsuda, R., Papastavros, E., Pfaunmiller, E., Vargas, J. and Zheng, X., 2012. Pharmaceutical and biomedical applications of affinity chromatography: recent trends and developments. *Journal of pharmaceutical and biomedical analysis*, *69*, pp.93-105.

41. Porath, J., Carlsson, J.A.N., Olsson, I. and Belfrage, G., 1975. Metal chelate affinity chromatography, a new approach to protein fractionation. Nature, 258(5536), pp.598-599.

42. Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R. and Stüber, D., 1988. Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. Bio/technology, 6(11), pp.1321-1325.

43. Ohi, M., Li, Y., Cheng, Y. and Walz, T., 2004. Negative staining and image classification—powerful tools in modern electron microscopy. *Biological procedures online*, *6*(1), p.23.

44. Dubochet, J., Adrian, M., Chang, J.J., Homo, J.C., Lepault, J., McDowall, A.W. and Schultz, P., 1988. Cryo-electron microscopy of vitrified specimens. *Quarterly reviews of biophysics*, *21*(2), pp.129-228.

45. Al-Amoudi, A., Chang, J.J., Leforestier, A., McDowall, A., Salamin, L.M., Norlén, L.P., Richter, K., Blanc, N.S., Studer, D. and Dubochet, J., 2004. Cryo-electron microscopy of vitreous sections. *The EMBO journal*, *23*(18), pp.3583-3588.

46. Chen, L.F., Blanc, E., Chapman, M.S. and Taylor, K.A., 2001. Real space refinement of acto-myosin structures from sectioned muscle. *Journal of structural biology*, *133*(2-3), pp.221-232.

47. Gabashvili, I.S., Agrawal, R.K., Spahn, C.M., Grassucci, R.A., Svergun, D.I., Frank, J. and Penczek, P., 2000. Solution structure of the E. coli 70S ribosome at 11.5 Å resolution. *Cell*, *100*(5), pp.537-549.

48. Kato T., Namba K., The 1.54 A resolution structure of apoferritin by CRYOARM300 with Cold-FEG. To Be Published. https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-9865

49. Saibil, H.R., 2000. Macromolecular structure determination by cryo-electron microscopy. *Acta Crystallographica Section D: Biological Crystallography*, *56*(10), pp.1215-1222.

50. Harris, J.R. and Scheffler, D., 2002. Routine preparation of air-dried negatively stained and unstained specimens on holey carbon support films: a review of applications. *Micron*, *33*(5), pp.461-480.

51. Bräuer, A.U., Savaskan, N.E., Kühn, H., Prehn, S., Ninnemann, O. and Nitsch, R., 2003. A new phospholipid phosphatase, PRG-1, is involved in axon growth and regenerative sprouting. *Nature neuroscience*, *6*(6), p.572.

52. Savaskan, N.E., Bräuer, A.U. and Nitsch, R., 2004. Molecular cloning and expression regulation of PRG-3, a new member of the plasticity-related gene family. *European Journal of Neuroscience*, *19*(1), pp.212-220.

53. Broggini, T., Nitsch, R. and Savaskan, N.E., 2010. Plasticity-related Gene 5 (PRG5) Induces Filopodia and Neurite Growth and Impedes Lysophosphatidic Acid–and Nogo-A–mediated Axonal Retraction. *Molecular biology of the cell*, *21*(4), pp.521-537.

54. Bräuer, A.U. and Nitsch, R., 2008. Plasticity-related genes (PRGs/LRPs): a brain-specific class of lysophospholipid-modifying proteins. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids*, *1781*(9), pp.595-600.

55. Strauss, U. and Bräuer, A.U., 2013. Current views on regulation and function of plasticity-related genes (PRGs/LPPRs) in the brain. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids*, *1831*(1), pp.133-138.

56. Sigal, Y.J., McDERMOTT, M.I. and Morris, A.J., 2005. Integral membrane lipid phosphatases/phosphotransferases: common structure and diverse functions. *Biochemical Journal*, *387*(2), pp.281-293.

57. Yue, F., Cheng, Y., Breschi, A., Vierstra, J., Wu, W., Ryba, T., Sandstrom, R., Ma, Z., Davis, C., Pope, B.D. and Shen, Y., 2014. A comparative encyclopedia of DNA elements in the mouse genome. *Nature*, *515*(7527), p.355.

58. Brosig A., 2018. Analysis of a PTEN-associated protein scaffold containing the transmembrane protein PRG2. http://dx.doi.org/10.17169/refubium-623

59. Yu, P., Agbaegbu, C., Malide, D.A., Wu, X., Katagiri, Y., Hammer, J.A. and Geller, H.M., 2015. Cooperative interactions of LPPR family members in membrane localization and alteration of cellular morphology. *J Cell Sci*, *128*(17), pp.3210-3222.

60. Velmans, T., Battefeld, A., Geist, B., Farrés, A.S., Strauss, U. and Bräuer, A.U., 2013. Plasticity-related gene 3 promotes neurite shaft protrusion. *BMC neuroscience*, *14*(1), p.36.

61. Van Diepen, M.T., Parsons, M., Downes, C.P., Leslie, N.R., Hindges, R. and Eickholt, B.J., 2009. MyosinV controls PTEN function and neuronal cell size. Nature cell biology, 11(10), pp.1191-1196.

62. Brosig, A., Fuchs, J., Ipek, F., Kroon, C., Schrötter, S., Vadhvani, M., Polyzou, A., Ledderose, J., van Diepen, M., Holzhütter, H.G. and Trimbuch, T., 2019. The Axonal Membrane Protein PRG2 Inhibits PTEN and Directs Growth to Branches. Cell reports, 29(7), pp.2028-2040.

63. Li, J., Yen, C., Liaw, D., Podsypanina, K., Bose, S., Wang, S.I., Puc, J., Miliaresis, C., Rodgers, L., McCombie, R. and Bigner, S.H., 1997. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. *Science*, *275*(5308), pp.1943-1947.

64. Steck, P.A., Pershouse, M.A., Jasser, S.A., Yung, W.A., Lin, H., Ligon, A.H., Langford, L.A., Baumgard, M.L., Hattier, T., Davis, T. and Frye, C., 1997. Identification of a candidate tumour suppressor gene, MMAC1, at chromosome 10q23. 3 that is mutated in multiple advanced cancers. *Nature genetics*, *15*(4), p.356.

65. Li, D.M. and Sun, H., 1997. TEP1, encoded by a candidate tumor suppressor locus, is a novel protein tyrosine phosphatase regulated by transforming growth factor β. *Cancer research*, *57*(11), pp.2124-2129.

66. Lee, Y.R., Chen, M. and Pandolfi, P.P., 2018. The functions and regulation of the PTEN tumour suppressor: new modes and prospects. *Nature Reviews Molecular Cell Biology*, *19*(9), p.547.

67. Liliental, J., Moon, S.Y., Lesche, R., Mamillapalli, R., Li, D., Zheng, Y., Sun, H. and Wu, H., 2000. Genetic deletion of the Pten tumor suppressor gene promotes cell motility by activation of Rac1 and Cdc42 GTPases. *Current Biology*, *10*(7), pp.401-404.

68. Iijima, M. and Devreotes, P., 2002. Tumor suppressor PTEN mediates sensing of chemoattractant gradients. *Cell*, *109*(5), pp.599-610.

69. Kreis, P., Leondaritis, G., Lieberam, I. and Eickholt, B.J., 2014. Subcellular targeting and dynamic regulation of PTEN: implications for neuronal cells and neurological disorders. *Frontiers in molecular neuroscience*, *7*, p.23.

70. Chaudhary, S., Pak, J.E., Gruswitz, F., Sharma, V. and Stroud, R.M., 2012. Overexpressing human membrane proteins in stably transfected and clonal human embryonic kidney 293S cells. *Nature protocols*, *7*(3), p.453.

71. Ulloa-Aguirre, A. and Conn, P.M., 2014. Intracellular trafficking of G protein-coupled receptors to the plasma membrane in health and disease. In *Cellular Endocrinology in Health and Disease* (pp. 341-364). Academic Press.

72. Estes, S. and Melville, M., 2013. Mammalian cell line developments in speed and efficiency. In *Mammalian cell cultures for biologics manufacturing* (pp. 11-33). Springer, Berlin, Heidelberg.

73. Aruffo, A., 1998. Transient expression of proteins using COS cells. *Current protocols in neuroscience*, *2*(1), pp.4-7.

74. Swiech, K., Picanço-Castro, V. and Covas, D.T., 2012. Human cells: new platform for recombinant therapeutic protein production. *Protein expression and purification*, *84*(1), pp.147-153.

75. Gluzman, Y., 1981. SV40-transformed simian cells support the replication of early SV40 mutants. *Cell*, *23*(1), pp.175-182.

76. Graham, F.L., Smiley, J., Russell, W.C. and Nairn, R., 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *Journal of general virology*, *36*(1), pp.59-72.

77. Chelikani, P., Reeves, P.J., Rajbhandary, U.L. and Khorana, H.G., 2006. The synthesis and high-level expression of a β2-adrenergic receptor gene in a tetracycline-inducible stable mammalian cell line. *Protein Science*, *15*(6), pp.1433-1440.

78. Reeves, P.J., Thurmond, R.L. and Khorana, H.G., 1996. Structure and function in rhodopsin: high level expression of a synthetic bovine opsin gene and its mutants in stable mammalian cell lines. *Proceedings of the National Academy of Sciences*, *93*(21), pp.11487-11492.

79. Takayama, H., Chelikani, P., Reeves, P.J., Zhang, S. and Khorana, H.G., 2008. High-level expression, single-step immunoaffinity purification and characterization of human tetraspanin membrane protein CD81. *PLoS One*, *3*(6).

80. Reeves, P.J., Callewaert, N., Contreras, R. and Khorana, H.G., 2002. Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase l-negative HEK293S stable mammalian cell line. *Proceedings of the National Academy of Sciences*, *99*(21), pp.13419-13424.

81. Wurm, F.M., 2004. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nature biotechnology*, *22*(11), pp.1393-1398.

82. Recillas-Targa, F., 2006. Multiple strategies for gene transfer, expression, knockdown, and chromatin influence in mammalian cell lines and transgenic animals. *Molecular biotechnology*, *34*(3), pp.337-354.

83. Glover, D.J., Lipps, H.J. and Jans, D.A., 2005. Towards safe, non-viral therapeutic gene expression in humans. *Nature Reviews Genetics*, *6*(4), pp.299-310.

84. Büssow, K., 2015. Stable mammalian producer cell lines for structural biology. *Current opinion in structural biology*, *32*, pp.81-90.

85. Schenborn, E.T. and Goiffon, V., 2000. DEAE-dextran transfection of mammalian cultured cells. In *Transcription Factor Protocols* (pp. 147-153). Humana Press.

86. Holmen, S.L., Vanbrocklin, M.W., Eversole, R.R., Stapleton, S.R. and Ginsberg, L.C., 1995. Efficient lipid-mediated transfection of DNA into primary rat hepatocytes. *In Vitro Cellular & Developmental Biology-Animal*, *31*(5), pp.347-351.

87. Washbourne, P. and McAllister, A.K., 2002. Techniques for gene transfer into neurons. *Current opinion in neurobiology*, *12*(5), pp.566-573.

88. Kim, T.K. and Eberwine, J.H., 2010. Mammalian cell transfection: the present and the future. *Analytical and bioanalytical chemistry*, *397*(8), pp.3173-3178.

89. Longo, P.A., Kavran, J.M., Kim, M.S. and Leahy, D.J., 2013. Transient mammalian cell transfection with polyethylenimine (PEI). In *Methods in enzymology* (Vol. 529, pp. 227-240). Academic Press.

90. Muller, N., Girard, P., Hacker, D.L., Jordan, M. and Wurm, F.M., 2005. Orbital shaker technology for the cultivation of mammalian cells in suspension. *Biotechnology and bioengineering*, *89*(4), pp.400-406.

91. Meissner, P., Pick, H., Kulangara, A., Chatellard, P., Friedrich, K. and Wurm, F.M., 2001. Transient gene expression: recombinant protein production with suspension-adapted HEK293-EBNA cells. *Biotechnology and bioengineering*, *75*(2), pp.197-203.

92. Tharmalingam, T., Ghebeh, H., Wuerz, T. and Butler, M., 2008. Pluronic enhances the robustness and reduces the cell attachment of mammalian cells. *Molecular biotechnology*, *39*(2), pp.167-177.

93. McGowen, J.C. and Mellors, A., 1979. Relationship between the solubility of amino acids. *Journal of Applied Biochemistry*, *1*, pp.423-428.

94. Perrin D.D., Dempsey B. (1974) Applications of pH Buffers. In: Buffers for pH and Metal Ion Control. Springer, Dordrecht

95. Arakawa, T. and Timasheff, S. N. (1984). Mechanism of protein salting in and salting out by divalent cation salts: balance between hydration and salt binding. *Biochemistry*, 23(25):5912–5923.

96. Zayas, J.F., 1997. Solubility of proteins. In *Functionality of proteins in food* (pp. 6-75). Springer, Berlin, Heidelberg.

97. Duquesne, K. and Sturgis, J.N., 2010. Membrane protein solubilization. In *Heterologous Expression of Membrane Proteins* (pp. 205-217). Humana Press.

98. Seddon, A.M., Curnow, P. and Booth, P.J., 2004. Membrane proteins, lipids and detergents: not just a soap opera. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, *1666*(1-2), pp.105-117.

99. Parker, J.L. and Newstead, S., 2016. Membrane protein crystallisation: current trends and future perspectives. In *The Next Generation in Membrane Protein Structure Determination* (pp. 61-72). Springer, Cham.

100. Mandon, E.D., Agez, M., Pellegrin, R., Igonet, S. and Jawhari, A., 2017. Novel systematic detergent screening method for membrane proteins solubilization. *Analytical biochemistry*, *517*, pp.40-49.

101. King, M.S., Crichton, P.G., Ruprecht, J.J. and Kunji, E.R., 2018. Publisher Correction: Concerns with yeast mitochondrial ADP/ATP carrier's integrity in DPC. *Nature structural & molecular biology*, *25*(10), pp.988-988.

102. Yang, Q., Brüschweiler, S., Zhao, L. and Chou, J.J., 2018. Reply to 'Concerns with yeast mitochondrial ADP/ATP carrier's integrity in DPC'and 'Dynamics and interactions of AAC3 in DPC are not functionally relevant'. *Nature structural & molecular biology*, *25*(9), pp.749-750.

103. Kotov, V., Bartels, K., Veith, K., Josts, I., Subhramanyam, U.K.T., Günther, C., Labahn, J., Marlovits, T.C., Moraes, I., Tidow, H. and Löw, C., 2019. High-throughput stability screening for detergent-solubilized membrane proteins. *Scientific reports*, *9*(1), pp.1-19.

104. Liao, M., Cao, E., Julius, D. and Cheng, Y., 2013. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature*, *504*(7478), pp.107-112.

105. Chen, Y., Clarke, O.B., Kim, J., Stowe, S., Kim, Y.K., Assur, Z., Cavalier, M., Godoy-Ruiz, R., Desiree, C., Manzini, C. and Blaner, W.S., 2016. Structure of the STRA6 receptor for retinol uptake. *Science*, *353*(6302), p.aad8266.

106. Zubcevic, L., Herzik, M.A., Chung, B.C., Liu, Z., Lander, G.C. and Lee, S.Y., 2016. Cryo-electron microscopy structure of the TRPV2 ion channel. *Nature structural & molecular biology*, *23*(2), pp.180-186.

107. Bai, X.C., Rajendra, E., Yang, G., Shi, Y. and Scheres, S.H., 2015. Sampling the conformational space of the catalytic subunit of human γ-secretase. *Elife*, *4*, p.e11182.

108. Tribet, C., Audebert, R. and Popot, J.L., 1996. Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. *Proceedings of the National Academy of Sciences*, *93*(26), pp.15047-15050.

109. Schägger, H. and von Jagow, G., 1991. Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Analytical biochemistry*, *199*(2), pp.223-231.

110. Wittig, I., Braun, H.P. and Schägger, H., 2006. Blue native PAGE. *Nature protocols*, *1*(1), p.418.

111. Schagger, H., Cramer, W.A. and Vonjagow, G., 1994. Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Analytical biochemistry*, *217*(2), pp.220-230.

112. Schamel, W.W., Arechaga, I., Risueño, R.M., van Santen, H.M., Cabezas, P., Risco, C., Valpuesta, J.M. and Alarcón, B., 2005. Coexistence of multivalent and monovalent TCRs explains high sensitivity and wide range of response. *The Journal of experimental medicine*, *202*(4), pp.493-503.

111. Rivero-Müller, A., Jonas, K.C., Hanyaloglu, A.C. and Huhtaniemi, I., 2013. Di/oligomerization of GPCRs—mechanisms and functional significance. In *Progress in molecular biology and translational science* (Vol. 117, pp. 163-185). Academic Press.

112. Gahbauer, S. and Böckmann, R.A., 2016. Membrane-mediated oligomerization of G protein coupled receptors and its implications for GPCR function. *Frontiers in physiology*, *7*, p.494.

113. Hellwig, N., Albrecht, N., Harteneck, C., Schultz, G. and Schaefer, M., 2005. Homo- and heteromeric assembly of TRPV channel subunits. *Journal of cell science*, *118*(5), pp.917-928.

114. Schindl, R. and Romanin, C., 2007. Assembly domains in TRP channels. *Biochemical Society transactions*, *35*(Pt 1), p.84.

115. Ogawa, T. and Hirokawa, N., 2018. Multiple analyses of protein dynamics in solution. *Biophysical reviews*, *10*(2), pp.299-306.

116. Laue, T.M., 1996. Analytical ultracentrifugation. *Current protocols in protein science*, *4*(1), pp.7-5.

117. Schuck, P., 2013. Analytical ultracentrifugation as a tool for studying protein interactions. *Biophysical reviews*, *5*(2), pp.159-171.

118. Stark, H., 2010. GraFix: stabilization of fragile macromolecular complexes for single particle cryo-EM. In *Methods in enzymology* (Vol. 481, pp. 109-126). Academic Press.

119. Bianchi, M., Turner, H.L., Nogal, B., Cottrell, C.A., Oyen, D., Pauthner, M., Bastidas, R., Nedellec, R., McCoy, L.E., Wilson, I.A. and Burton, D.R., 2018. Electron-microscopy-based epitope mapping defines specificities of polyclonal antibodies elicited during HIV-1 BG505 envelope trimer immunization. *Immunity*, *49*(2), pp.288-300.

120. Sander, B. and Golas, M.M., 2011. Visualization of bionanostructures using transmission electron microscopical techniques. *Microscopy research and technique*, *74*(7), pp.642-663.

121. Melchior, V., Hollingshead, C.J. and Cahoon, M.E., 1980. Stacking in lipid vesicle-tubulin mixtures is an artifact of negative staining. *The Journal of cell biology*, *86*(3), pp.881-884.

122. De Carlo, S. and Harris, J.R., 2011. Negative staining and cryo-negative staining of macromolecules and viruses for TEM. *Micron*, *42*(2), pp.117-131.

123. Söderberg, O., Gullberg, M., Jarvius, M., Ridderstråle, K., Leuchowius, K.J., Jarvius, J., Wester, K., Hydbring, P., Bahram, F., Larsson, L.G. and Landegren, U., 2006. Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nature methods*, *3*(12), pp.995-1000.

124. Fathi-Roudsari, M., Akhavian-Tehrani, A. and Maghsoudi, N., 2016. Comparison of three Escherichia coli strains in recombinant production of reteplase. *Avicenna journal of medical biotechnology*, *8*(1), p.16.

125. Bornhorst, J.A. and Falke, J.J., 2000. [16] Purification of proteins using polyhistidine affinity tags. In *Methods in enzymology* (Vol. 326, pp. 245-254). Academic Press.

126. Grisshammer, R. and Tucker, J., 1997. Quantitative evaluation of neurotensin receptor purification by immobilized metal affinity chromatography. *Protein expression and purification*, *11*(1), pp.53-60.

127. Wülfing, C., Lombardero, J. and Plückthun, A., 1994. An Escherichia coli protein consisting of a domain homologous to FK506-binding proteins (FKBP) and a new metal binding motif. *Journal of Biological Chemistry*, *269*(4), pp.2895-2901.

128. Bolanos-Garcia, V.M. and Davies, O.R., 2006. Structural analysis and classification of native proteins from E. coli commonly co-purified by immobilised metal affinity chromatography. *Biochimica et Biophysica Acta (BBA)-General Subjects*, *1760*(9), pp.1304-1313.

129. Gatzeva-Topalova, P.Z., May, A.P. and Sousa, M.C., 2005. Structure and mechanism of ArnA: conformational change implies ordered dehydrogenase mechanism in key enzyme for polymyxin resistance. *Structure*, *13*(6), pp.929-942.

130. Williams, G.J., Breazeale, S.D., Raetz, C.R. and Naismith, J.H., 2005. Structure and function of both domains of ArnA, a dual function decarboxylase and a formyltransferase, involved in 4-amino-4-deoxy-L-arabinose biosynthesis. *Journal of Biological Chemistry*, *280*(24), pp.23000-23008.

131. Weininger, U., Haupt, C., Schweimer, K., Graubner, W., Kovermann, M., Brüser, T., Scholz, C., Schaarschmidt, P., Zoldak, G., Schmid, F.X. and Balbach, J., 2009. NMR solution structure of SlyD from Escherichia coli: spatial separation of prolyl isomerase and chaperone function. *Journal of molecular biology*, *387*(2), pp.295-305.

132. Andersen, K.R., Leksa, N.C. and Schwartz, T.U., 2013. Optimized E. coli expression strain LOBSTR eliminates common contaminants from His-tag purification. *Proteins: Structure, Function, and Bioinformatics*, *81*(11), pp.1857-1861.

133. Roof, W.D., Fang, H.Q., Young, K.D., Sun, J. and Young, R., 1997. Mutational analysis of slyD, an Escherichia coli gene encoding a protein of the FKBP immunophilin family. *Molecular microbiology*, *25*(6), pp.1031-1046.

134. Robichon, C., Luo, J., Causey, T.B., Benner, J.S. and Samuelson, J.C., 2011. Engineering Escherichia coli BL21 (DE3) derivative strains to minimize E. coli protein contamination after purification by immobilized metal affinity chromatography. *Appl. Environ. Microbiol.*, *77*(13), pp.4634-4646.

135. Cao, E., Chen, Y., Cui, Z. and Foster, P.R., 2003. Effect of freezing and thawing rates on denaturation of proteins in aqueous solutions. *Biotechnology and bioengineering*, *82*(6), pp.684-690.

136. Carpenter, J.F., Manning, M.C. and Randolph, T.W., 2002. Long-Term Storage of Proteins. *Current protocols in protein science*, *27*(1), pp.4-6.

137. Patro, S.Y., Freund, E. and Chang, B.S., 2002. Protein formulation and fill-finish operations. *Biotechnology annual review*, *8*, pp.55-84.

138. Izutsu, K.I., 2014. Stabilization of therapeutic proteins in aqueous solutions and freeze-dried solids: an overview. In *Protein Downstream Processing* (pp. 435-441). Humana Press, Totowa, NJ.

139. Nadeau, J.L., 2017. *Introduction to experimental biophysics: biological methods for physical scientists*. CRC Press.

140. Salvi, G., De Los Rios, P. and Vendruscolo, M., 2005. Effective interactions between chaotropic agents and proteins. *Proteins: Structure, Function, and Bioinformatics*, *61*(3), pp.492-499.

141. Bennion, B.J. and Daggett, V., 2003. The molecular basis for the chemical denaturation of proteins by urea. *Proceedings of the National Academy of Sciences*, *100*(9), pp.5142-5147.

142. Bhuyan, A.K., 2002. Protein stabilization by urea and guanidine hydrochloride. *Biochemistry*, *41*(45), pp.13386-13394.

143. Janecki, D.J. and Reilly, J.P., 2005. Denaturation of metalloproteins with EDTA to facilitate enzymatic digestion and mass fingerprinting. *Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry*, *19*(10), pp.1268-1272.

144. Mirnajd, G.S., Farajnia, S., Mahboudi, F. and Babaei, H., 2011. Optimizing refolding condition for recombinant tissue plasminogen activator.

145. Casey, J.R. and Reithmeier, R.A., 1991. Analysis of the oligomeric state of Band 3, the anion transport protein of the human erythrocyte membrane, by size exclusion high performance liquid chromatography. Oligomeric stability and origin of heterogeneity. *Journal of Biological Chemistry*, *266*(24), pp.15726-15737.

146. Cao, B., Xu, H. and Mao, C., 2011. Transmission electron microscopy as a tool to image bioinorganic nanohybrids: The case of phage-gold nanocomposites. *Microscopy research and technique*, *74*(7), pp.627-635.

147. Zhao, F.Q. and Craig, R., 2003. Capturing time-resolved changes in molecular structure by negative staining. *Journal of structural biology*, *141*(1), pp.43-52.

148. Hubbard, S.J., 1998. The structural aspects of limited proteolysis of native proteins. *Biochimica et biophysica acta*, *1382*(2), pp.191-206.

149. Quevillon-Cheruel, S., Collinet, B., Trésaugues, L., Minard, P., Henckes, G., Aufrère, R., Blondeau, K., Zhou, C.Z., Liger, D., Bettache, N. and Poupon, A., 2007. Cloning, production, and purification of proteins for a medium-scale structural genomics project. In *Macromolecular Crystallography Protocols* (pp. 21-37). Humana Press.

150. Daviter, T. and Fronzes, R., 2013. Protein sample characterization. In *Protein-Ligand Interactions* (pp. 35-62). Humana Press, Totowa, NJ.

151. Jerabek-Willemsen, M., André, T., Wanner, R., Roth, H.M., Duhr, S., Baaske, P. and Breitsprecher, D., 2014. MicroScale Thermophoresis: Interaction analysis and beyond. *Journal of Molecular Structure*, *1077*, pp.101-113.

152. Lin, C.C., Melo, F.A., Ghosh, R., Suen, K.M., Stagg, L.J., Kirkpatrick, J., Arold, S.T., Ahmed, Z. and Ladbury, J.E., 2012. Inhibition of basal FGF receptor signaling by dimeric Grb2. *Cell*, *149*(7), pp.1514-1524.

153. Keren-Kaplan, T., Attali, I., Estrin, M., Kuo, L.S., Farkash, E., Jerabek-Willemsen, M., Blutraich, N., Artzi, S., Peri, A., Freed, E.O. and Wolfson, H.J., 2013. Structure-based in silico identification of ubiquitin-binding domains provides insights into the ALIX-V: ubiquitin complex and retrovirus budding. *The EMBO journal*, *32*(4), pp.538-551.

154. Arbel, N., Ben-Hail, D. and Shoshan-Barmatz, V., 2012. Mediation of the antiapoptotic activity of Bcl-xL protein upon interaction with VDAC1 protein. *Journal of Biological Chemistry*, *287*(27), pp.23152-23161.

155. Van Den Bogaart, G., Meyenberg, K., Diederichsen, U. and Jahn, R., 2012. Phosphatidylinositol 4, 5-bisphosphate increases Ca2+ affinity of synaptotagmin-1 by 40-fold. *Journal of Biological Chemistry*, *287*(20), pp.16447-16453.

156. Van Den Bogaart, G., Thutupalli, S., Risselada, J.H., Meyenberg, K., Holt, M., Riedel, D., Diederichsen, U., Herminghaus, S., Grubmüller, H. and Jahn, R., 2011. Synaptotagmin-1 may be a distance regulator acting upstream of SNARE nucleation. *Nature structural & molecular biology*, *18*(7), p.805.

Statutory Declaration

"I, Fatih Mehmet Ipek , by personally signing this document in lieu of an oath, hereby affirm that I prepared the submitted dissertation on the topic Structure investigations of plasticity-related gene 2 (PRG2), independently and without the support of third parties, and that I used no other sources and aids than those stated.

All parts which are based on the publications or presentations of other authors, either in letter or in spirit, are specified as such in accordance with the citing guidelines. The sections on methodology (in particular regarding practical work, laboratory regulations, statistical processing) and results (in particular regarding figures, charts and tables) are exclusively my responsibility.

Furthermore, I declare that I have correctly marked all of the data, the analyses, and the conclusions generated from data obtained in collaboration with other persons, and that I have correctly marked my own contribution and the contributions of other persons (cf. declaration of contribution). I have correctly marked all texts or parts of texts that were generated in collaboration with other persons.

My contributions to any publications to this dissertation correspond to those stated in the below joint declaration made together with the supervisor. All publications created within the scope of the dissertation comply with the guidelines of the ICMJE (International Committee of Medical Journal Editors; www.icmje.org) on authorship. In addition, I declare that I shall comply with the regulations of Charité – Universitätsmedizin Berlin on ensuring good scientific practice.

I declare that I have not yet submitted this dissertation in identical or similar form to another Faculty.

The significance of this statutory declaration and the consequences of a false statutory declaration under criminal law (Sections 156, 161 of the German Criminal Code) are known to me."

Date _____ Signature _____

Declaration of your own contribution to any publications

Fatih Ipek contributed the following to the below listed publications:

Brosig, A., Fuchs, J., Ipek, F., Kroon, C., Schrötter, S., Vadhvani, M., Polyzou, A., Ledderose, J., van Diepen, M., Holzhütter, H.G. and Trimbuch, T., 2019. The axonal membrane protein PRG2 inhibits PTEN and directs growth to branches. Cell reports, 29(7), pp.2028-2040.
Contribution: Figure1D, Figure 1E, Figure 1F.

_____
Signature, date and stamp of first supervising university professor / lecturer

_____
Signature of doctoral candidate

Curriculum vitae

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

Complete list of publications

1. Brosig, A., Fuchs, J., Ipek, F., Kroon, C., Schrötter, S., Vadhvani, M., Polyzou, A., Ledderose, J., van Diepen, M., Holzhütter, H.G. and Trimbuch, T., 2019. The axonal membrane protein PRG2 inhibits PTEN and directs growth to branches. Cell reports, 29(7), pp.2028-2040.
2. Fatih, M.I., Janssens, K. and Timmerman, V., 2012, December. Modelling a sensory neuropathy caused by mutations in SPTLC2 in Drosophila. In JOURNAL OF NEUROGENETICS (Vol. 26, pp. 62-62).
3. Sogut, M.S., Ozturk, N.C., Ozturk, H.U., Pinar, O., Ipek, F.M., Cayir, E., Gokce, A., Karasu, D., Ozturk, D.C., Ozturk, Y. and Denizci, A.A., 2012. Production of industrial enzymes for feed industry. New Biotechnology, (29), p.S108.
4. Pannier, L., Sweeney, T., Hamill, R.M., Ipek, F., Stapleton, P.C. and Mullen, A.M., 2009. Lack of an association between single nucleotide polymorphisms in the bovine leptin gene and intramuscular fat in Bos taurus cattle. Meat science, 81(4), pp.731-737.
5. Salih, B.A. and Ipek, F.M., 2007. Global research on Helicobacter pylori. Digestive diseases and sciences, 52(1), p.147.
6. Saribasak, H., Saribasak, N.N., Ipek, F.M., Ellwart, J.W., Arakawa, H. and Buerstedde, J.M., 2006. Uracil DNA glycosylase disruption blocks Ig gene conversion and induces transition mutations. The Journal of Immunology, 176(1), pp.365-371.