Structural-Based Study of Intact Human Particulate Guanylyl Cyclase Receptor A (pGC-A) and Biophysical Characterization

by

Shangji Zhang

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved June 2021 by the
Graduate Supervisory Committee:

Petra Fromme, Chair
Stephen Johnston
Yuval Mazor

ARIZONA STATE UNIVERSITY

August 2021

ABSTRACT

Particulate Guanylyl Cyclase Receptor A (pGC-A) is an atrial natriuretic peptide receptor, which plays a vital role in controlling cardiovascular, renal, and endocrine functions. The extracellular domain of pGC-A interacts with natriuretic peptides and triggers the intracellular guanylyl cyclase domain to convert GTP to cGMP. To effectively develop a method that can regulate pGC-A, structural information regarding its intact form is necessary. Currently, only the extracellular domain structure of rat pGC-A has been determined. However, structural data regarding the transmembrane domain, as well as functional intracellular domain regions, need to be elucidated.

This dissertation presents detailed information regarding pGC-A expression and optimization in the baculovirus expression vector system, along with the first purification method for purifying functional intact human pGC-A. The first *in vitro* evidence of a purified intact human pGC-A tetramer was detected in detergent micellar solution. Intact pGC-A is currently proposed to function as a homodimer. Upon analyzing my findings and acknowledging that dimer formation is required for pGC-A functionality, I proposed the first tetramer complex model composed of two functional subunits (homodimer). Forming tetramer complexes on the cell membrane increases pGC-A binding efficiency and ligand sensitivity.

Currently, a two-step mechanism has been proposed for ATP-dependent pGC-A signal transduction. Based on cGMP functional assay results, it can be suggested that the binding ligand also moderately activates pGC-A, and that ATP is not crucial for the activation of guanylyl cyclase. Instead, three modulators can regulate different activation levels in intact pGC-A.

Crystallization of purified intact pGC-A was performed to determine its structure. During the crystallization condition screening process, I successfully selected seven promising initial crystallization conditions for intact human pGC-A crystallization. One selected condition led to the formation of excellent needle-shaped crystals. During the serial crystallography diffraction experiment, five diffraction patterns were detected. The highest diffraction resolution spot reached 3 Å.

This work will allow the determination of the intact human pGC-A structure while also providing structural information on the protein signal transduction mechanism. Further structural knowledge may potentially lead to improved drug design. More precise mutation experiments could help verify the current pGC-A signal transduction and activation mechanism.

DEDICATION

I would like to dedicate this dissertation to my lovely family, without whose support and encouragement this endeavor would not have been possible. I am immensely grateful to my dad. Without his blessing, my life would not have had an abundance of opportunities; I am very thankful to my mom. She has supported my studies despite the sadness of keeping us apart. My parents are traditional Chinese parents but open-minded. They have done their best to support my choices and pave my future maximally, simultaneously ensuring I had a carefree environment. They unconditionally supported my dreams, even when it meant I would have to leave them and go abroad to an unfamiliar country. Moreover, I would like to thank my husband. Thank you for sharing the ups and downs of our journey. Without your encouragement and support, I would not have made it to the finish line. Lastly, I would like to thank my special family member, Denden, my four-year-old dog. He could have a much better doggy life with others, but he still stays with me. I hope to fulfill all my promises to you soon. Promises made and promise kept. Woof-woof!

ACKNOWLEDGMENTS

I would like to thank my advisor, Petra Fromme, for all the advice and mentorship over the last five years. Thank you for affording me the opportunity to participate in numerous research projects and collaborative beamtimes. These research experiences were tremendous and eye-opening. I would also like to thank my committee member Yuval Mazor. Yuval always keeps his door open for me, answers my countless questions, listens to my difficulties, and provides valuable mentorship. Furthermore, I would like to thank my committee member Stephen Johnston. Thank you for allowing me to use your lab space and for providing suggestions and information for my future career decisions. It is my honor to have all of you as my committee members. Lastly, I would like to thank Debbie Hansen. Thanks for believing in me and assisting me with my research paper and job application package editing. I am blessed to have met all of you and receive your unconditional support.

TABLE OF CONTENTS

Page

LIST OF TABLES xi

LIST OF FIGURES xii

CHAPTER

1 INTRODUCTION 1

1.1 Cardiovascular Disease 1

1.2 Cardiometabolic Disease 4

1.3 Natriuretic Peptides 6

Different Types of Natriuretic Peptides 7

Current Natriuretic Peptide-Based Medication 12

Innovative Engineered Natriuretic Peptides 12

1.4 Natriuretic Peptide Receptors 14

Natriuretic Peptide Receptor A 15

Natriuretic Peptide Receptor B 16

Natriuretic Peptide Receptor C 17

1.5 The Natriuretic Peptides/ Natriuretic Peptide Receptor A/cGMP Pathway 18

Topology of Natriuretic Peptide Receptor A 19

Natriuretic Peptide Activates Natriuretic Peptide Receptor A 20

Protein Kinase-like Domain of Natriuretic Peptide Receptor A 23

Guanylyl Cyclase Domain of Natriuretic Peptide Receptor A 24

cGMP Is a Second Messenger in Cells 25

1.6 Natriuretic Peptide Receptor A Roles in Pathophysiology Regulation 28

CHAPTER Page

Regulation of Blood Pressure and Cardiac Homeostasis 28

Regulation of Renal Homeostasis 29

Regulation of Fat Cell Metabolism 30

1.7 Summary and Conclusion 33

2 EXPRESSION AND PURIFICATION OF INTACT HUMAN NATRIURETIC PEPTIDE RECEPTOR A 35

2.1 Abstract 35

2.2 Introduction 37

2.2.1 Protein Expression System 37

Prokaryotic System 37

Eukaryotic System 39

Cell-Free System 42

2.2.2 Baculovirus Expression Vector System 45

Baculovirus Life Cycle Overview of Baculovirology 45

Overview of Baculovirus Expression System 48

Insect Cell Expression System 52

Bac-to-Bac System 54

2.2.3 Protein Purification 57

Overview of Protein Purification 57

Difficulty of Membrane Protein Expression and Purification.....59

Precipitation Purification Method 62

Popular Chromatography Purification Methods 63

CHAPTER Page

Recombinant Protein with Tag Assisted Purification Methods 66

2.3 Materials and Methods 74

Construction of Pfastbac1-pGC-A Expression Vector 74

Insect Cell Culture 75

Recombinant Baculovirus Generation and Selection 76

Recombinant Baculovirus Amplification 79

Verification Of Intact Human pGC-A Expression 81

Optimization Of Virus MOI and Cell Harvest Days 82

Large Scale Intact Human pGC-A Purification 82

2.4 Results 84

Isolating Recombinant Bacmid Via Blue/ White Colony Selection 84

PCR Verification of Recombinant Bacmid.....85

Verification of Intact Human pGC-A Expression.....87

Optimization of Protein Expression 88

Large Scale Intact Human pGC-A Purification 89

2.5 Discussion 91

2.6 Conclusion and Outlook 94

3 PROTEIN CHARACTERIZATION AND FUNCTIONAL STUDY OF INTACT HUMAN NATRIURETIC PEPTIDE RECEPTOR A 96

3.1 Abstract 96

3.2 Introduction 97

3.2.1 Oligomerization of Natriuretic Peptide Receptor A 97

CHAPTER Page

3.2.2 Enzyme-linked Immunosorbent Assay 99

3.2.3 ATP Plays an Allosteric Regulatory Role in Active pGC-A Guanylyl Cyclase 103

3.3 Methods 105

High Resolution Clear Native Polyacrylamide Gel Electrophoresis 105

Competitive Functional ELISA Assay Protocol for cGMP Yield Determination 106

Whole Cell Activity Assay for Intact pGC-A Expressed in Sf9 Cells 109

Purified Intact pGC-A *in vitro* Functional Activity Test 110

3.4 Results 112

Different pGC-A Oligomeric States Observed in Gel Filtration Chromatography 112

Characterizations of different oligomeric peak fractions eluted from Superose6 column 115

Intact pGC-A High Oligomeric State Confirmed as A Tetramer 118

Intact pGC-A Tetramer Was Not Oligomerized Through Disulfide Bonds 122

Recombinantly Expressed Intact pGC-A Was Functional in A Whole-Cell Activity Assay 124

Purified Intact pGC-A Tetramer Was Functionally active in *In vitro* Activity Assay 126

3.5 Discussion 130

CHAPTER Page

3.6 Conclusion 138

4 CRYSTALLIZATION AND INITIAL DIFFRACTION OF INTACT HUMAN NATRIURETIC PEPTIDE RECEPTOR A 140

4.1 Abstract 140

4.2 Introduction 141

4.2.1 Overview of Protein Crystallization 141

4.2.2 Molecular Interactions in Crystallization 144

4.2.3 Methods for Protein Crystallization and Optimization 148

4.2.4 Membrane Protein Crystallization 153

4.2.5 X-ray Sources 158

4.2.6 Protein Crystal Diffraction Methods 161

4.3 Materials and Methods 168

Initial Protein Crystallization Screening 168

Crystallization Optimization Screening in F8 Condition 168

Crystallization Optimization Screening in C4 Condition 170

Real-time Protein Crystallization Under Microscope 170

Confirmation Of pGC-A Protein Crystals Formed in the F8 Condition 171

F8 Condition Protein Crystal Diffraction and Data Collection 172

4.4 Results 174

Initial Intact pGC-A Crystallization Screening 174

Further Optimization of the F8 Crystallization Condition 178

Optimization of C4 Crystallization Condition 184

CHAPTER Page

Real-time Protein Crystallization Under the Microscope 189

Confirmation of pGC-A Protein Formed Protein Crystals in F8 Condition 192

Diffraction Patterns of pGC-A Crystals under the F8 Condition via Serial Crystallography 195

Liquid Protein Phase in Crystallization 197

4.5 Discussion 199

4.6 Conclusion 206

5 OUTLOOK AND CONCLUSION 207

5.1 Future Directions 208

5.2 Outlook and Conclusion 210

REFERENCES 214

APPENDIX

A PERMISSIONS 261

LIST OF TABLES

Table Page

2.2.1 A List of Chromatography Purification Tags Along with Their Characteristic Features 72

2.2.2 A List of Newly Introduced Tags and Their Features....73

4.4.1 Summary of Seven Selected Crystallization Conditions After Initial Crystallization Screening with the MemGold2 HT-96 Kit....175

LIST OF FIGURES

Figure		Page
1.1	The Different Members of the Natriuretic Peptide Family	7
1.2	The Natriuretic Peptide Receptor Family and Their Binding Peptides	15
1.3	The Natriuretic Peptides/ Natriuretic Peptide Receptor A/cGMP Pathway	19
1.4	Structure of Rat pGC-A Extracellular Domain (ECD) and Proposed Rotation Model Mechanism	22
1.5	Summary of Protein Sequence Alignment	25
1.6	Natriuretic Peptide Receptor A is Involved in Different Vasodilation Mechanisms in Vascular Smooth Muscle Cells	27
1.7	Scheme of Natriuretic Peptide Receptor A (NPRA/pGC-A) Involved Human Adipocytes.	32
2.1	Summary Scheme of Ideal Major Experiment Steps from Cloning to Structure Determination of Intact Human pGC-A	36
2.2.1	General Scheme of the General Protein Production Process of Cell-free System	44
2.2.2	Summary of the Baculovirus Life Cycle	47
2.2.3	Summary of Baculovirus Expression Vector System (BEVS) Applications	51
2.2.4	Scheme of Generation of Recombinant Baculovirus and Expression of Recombinant Protein in The Bac-to-Bac™ Baculovirus Expression System.	57
2.2.5	Molecular Interactions in Immobilized Metal Ion Affinity Chromatography (IMAC)	68

Figure Page

2.3.1 Schematic of the pGC-A Donor Vector and Protein Diagram 74

2.3.2 Complete Sequences for the Expression-Optimized Sequences Encoding Intact pGC-A Protein 75

2.4.1 Blue-White Colony Screening 84

2.4.2 PCR Verification of Recombinant Bacmid Production 86

2.4.3 Western Blot (Anti-His) of Virus Passages Expressing pGC -A Proteins 86

2.4.4 Optimization of pGC -A Expression in Sf9 Cells 88

2.4.5 Purification of Intact pGC -A Via Affinity and Gel Filtration Chromatography.90

3.1 Summary Scheme of Major Experiments for Protein Characterizations and Functional Study 96

3.2.1 Summary of Four Different Types of Enzyme-Linked Immunosorbent Assay (ELISA) 101

3.3.1 Summary of Major Experimental Steps in Competitive cGMP Elisa Assay 108

3.4.1 Summary of Different pGC -A Oligomeric States in Superose6 Chromatography from different experiments 113

3.4.2 Summary of Characterization of Different Oligomeric Peak Fractions from Superose6 Column 116

3.4.3 Silver Stain of High Resolution Clear-Native PAGE 117

3.4.4 Summary of Intact pGC-A Superose6 Column Elution Profiles from Four Different Detergent Concentrations 119

3.4.5 Silver Stain of Different Oligomeric Samples Treated with or Without Overnight DTT in Clear-Native PAGE 123

Figure Page

3.4.6 Determination of Expressed Intact pGC-A Functionality via Whole-Cell Activity Assay 124

3.4.7 Purified Intact pGC-A *In Vitro* Functional Activity Test 127

4.2.1 Summary Illustration of the Protein Crystallization Phase Diagram....146

4.2.2 Schematic of Membrane Protein Crystallization in Lipidic Cubic Mesophase .155

4.2.3 Summary of Two Types of Membrane Protein Crystal Packing 157

4.2.4 Principle Behind the Generation of Coherent X-rays for X-ray Free-Electron Lasers (XFELs)....161

4.2.5 Schematic of Bragg's law 162

4.2.6 Scheme of the Serial Femtosecond Crystallography Experimental Layout at XFEL Beamlines....165

4.3.1 Summary of Protein Crystal Collecting and Washing Steps 172

4.4.1 Summary of Picked Seven Crystallization Conditions from the Initial Crystallization Screening 176

4.4.2 X-ray Powder Diffraction of Protein Crystallites under the F8 Condition 177

4.4.3 Needle-Shaped Crystals under Two Different Diluted F8 Conditions 178

4.4.4 Summary of Broad PEG3350 Screening of the F8 Crystallization Condition with Two Protein Concentrations 181

4.4.5 Summary of Broad Ammonium Sulfate Screening in the F8 Crystallization Condition with Two Protein Concentrations 182

Figure Page

4.4.6 Summary of Both PEG3350 and Ammonium Sulfate Narrow Screenings in the F8 Crystallization Condition 183

4.4.7 Summary of the Initial C4 Condition PEG250DME and Ammonium Sulfate Screening Results 186

4.4.8 Summary of PEG250DME and Ammonium Sulfate Narrow Screening Results under C4 Crystallization Condition at Room Temperature 187

4.4.9 Summary of PEG250DME and Ammonium Sulfate Narrow Screening Results under C4 Crystallization Condition in a 15°C Incubator.....188

4.4.10 Super-Dense Rod-Shaped Microcrystals Observed in Purification Buffer-Based Crystallization Conditions 189

4.4.11 Rod-Shaped Microcrystals Grown in Real-Time 191

4.4.12 Summary of Crystal Shapes at Different Temperatures 192

4.4.13 Presence of pGC-A in Crystals was Confirmed via Western Blot 195

4.4.14 Three Indexable Diffraction Patterns from pGC-A Crystals Collected via Serial Crystallography in APS 197

4.4.15 Liquid Phase Separation Crystallization in the C4 Condition 198

CHAPTER 1

INTRODUCTION

1.1 Cardiovascular Disease

Cardiovascular diseases are one of the leading causes of death in both developing and developed countries. According to the World Health Organization, an estimated 31% of worldwide deaths are due to cardiovascular disease (Capewell et al. 2010). Additionally, the high treatment costs contribute significantly to the economic burden of individuals and nations. Unlike developed countries which have higher health standards, the rate of cardiovascular disease continues to rise in low and middle-income countries (Gaziano et al. 2010, Owolabi et al. 2016). Cardiovascular diseases are a group of disorders associated with the heart and blood vessels (Mendis et al. 2011). Cardiovascular diseases are usually long-term conditions in which many risk factors are involved. Deposition, circulation and accumulation of fatty plaques inside the arteries is a major risk factor for cardiovascular disease (Frostegard 2013). An unhealthy diet, hypertension, diabetes, and obesity are commonly observed among cardiovascular disease patients who were found to have atheromas, and these are considered risk factors (Kjeldsen 2018, Leon and Maddox 2015). Furthermore, the patient's personal profile including age, sex and genetics also contributes to progression of cardiovascular disease (Hobbs 2015).

Hypertension is one of the most important factors for progression of cardiovascular disease. It can develop when the blood vessels become calcified and lose elasticity, which leads to a decreased ability to contract and relax (Flora and Nayak 2019). Clinical studies showed that treating high blood pressure reduces the risk of stroke by 35-40%, and the risk of heart failure by 50% (Antonakoudis et al. 2007). However,

about 12-15% of patients with hypertension are sub-classified as having resistant hypertension, which has a poor medical response to multiple anti-hypertensive treatments (Carey et al. 2018). Resistant hypertension is more likely to occur in patients with obesity, diabetes or renal dysfunction (Wei et al. 2018).

Typically, cardiovascular disease includes coronary, cerebral, peripheral arterial disease, rheumatic heart disease, congenital heart disease, deep vein thrombosis, and pulmonary embolism (Thiriet 2018). In coronary artery disease, blood flow to heart muscle cells is reduced or obstructed. This may lead to heart attack and heart failure. In the United States, coronary artery disease caused about 1 in every 5 deaths in 2005 (Cassar et al. 2009). A projection of total medical care costs for heart failure is estimated to increase from 31 billion US dollars in 2012 to 70 billion US dollars in 2030 (Heidenreich et al. 2013).

The treatment of cardiovascular disease is vast and varies depending on the specific illness. Several common cardiovascular ailments will be discussed in this section. As mentioned above, coronary artery disease is a common cause of heart failure associated with metabolic disease such as diabetes or high cholesterol (Cassar et al. 2009). Treatment of atherosclerotic disease involves antiplatelet medication such as aspirin (Thobani et al. 2019). In addition, high intensity cholesterol lowering therapy is also recommended and are commonly referred to as statins (atorvastatin, rosuvastatin, etc.) (Thondapu et al. 2019). By combining antiplatelet and cholesterol lowering therapies, the risk of atherosclerosis is greatly decreased (Potpara and Lip 2020).

Heart failure is a chronic condition in which the heart muscle loses the ability to effectively pump sufficient blood through the cardiovascular system (Guder and Stork

2019). The heart's ejection fraction is the percentage of blood that is pumped out of the left ventricle during heart contraction. Based on the ejection fraction of the left ventricle, heart failure is classified into either systolic heart failure or diastolic heart failure (Tomasoni et al. 2019, Federmann and Hess 1994). Systolic heart failure is also referred to as heart failure with reduced ejection fraction (HFrEF). Diastolic heart failure is referred to as heart failure with preserved ejection fraction (HFpEF) (Taylor et al. 2014). The treatment of heart failure is complex and varies depending on the type and severity, and the underlying cause of the disease. In both HFrEF and HFpEF, diuretics to increase urine output are typically used to optimize the patient's volume status as heart failure can lead to fluid retention due to the decreased ability of the heart to mobilize blood (Mullens et al. 2019). HFpEF is mostly caused by uncontrolled hypertension, which leads to hypertrophy and thickening of the left ventricular myocardium (Tam et al. 2017). Patients with HFpEF have a normal ejection fraction, but because of the myocardial thickening, the left ventricle is only able to be filled with a small amount of blood, which leads to fluid accumulation in the lungs and venous system. Treatment is primarily focused on blood pressure control, optimization of volume status with diuretic medications, and lifestyle modifications such as decreasing sodium intake and increasing physical activity (Tam et al. 2017, Cilia et al. 2019). Conversely, the therapeutic goal for HFrEF is to improve the ejection fraction of the left ventricle. The initial treatment of HFrEF involves the combination of multiple medication classes comprised of beta-blockers, angiotensin converting enzyme inhibitors (ACE-inhibitor) or angiotensin receptor blockers (ARBs), aldosterone receptor antagonists, and neprilysin inhibitors (Chavey et al. 2017). ACE inhibitors, ARBs, and aldosterone receptor antagonists reduce the effects of the renin-

angiotensin-aldosterone system in the kidneys (Ma et al. 2010, Berliner and Bauersachs 2017). Beta-blockers decrease adrenergic stimulation of the heart but the mechanism of improving HFrEF is still under discussion (Kotecha et al. 2017). Neprilysin is an enzyme in humans that degrades natriuretic peptides. Natriuretic peptides are beneficial in heart failure by stimulating the excretion of sodium and water by the kidneys. A neprilysin inhibitor would increase the activity of natriuretic peptides (Jhund and McMurray 2016).

The mortality and morbidity caused by cardiovascular diseases are a global phenomenon. Improving diagnostic testing, early treatment of potential risk factors, finding new therapeutic strategies, and developing better treatments are key to defeating cardiovascular disease.

1.2 Cardiometabolic Disease

Cardiometabolic diseases are also known as cardiometabolic syndrome, which is a combination of metabolic dysfunctions such as hypertension, insulin resistance, impaired glucose tolerance, dyslipidemia and abdominal obesity (Srivastava 2012). Cardiometabolic disease shares similar risk factors and tends to develop toward cardiovascular disease and type II diabetes. In the United States, about 25% of the adult population is diagnosed with cardiometabolic disease (Falkner and Cossrow 2014). For example, patients diagnosed with metabolic syndrome are more likely to die from coronary heart disease and three times more likely to develop heart attack and a stroke in their lifetime (Ford 2005). Cardiometabolic disease is diagnosed when a patient presents with at least three out of five criteria: increased waist circumference, elevated blood pressure, elevated fasting blood sugar level, decreased fasting high density lipoprotein

cholesterol, and increased fasting triglyceride levels based on the national cholesterol education program adult treatment panel III (Huang 2009).

As mentioned before, hypertension is one of the main features not only in metabolic syndrome but also in cardiovascular disease. More than 85% of patients with cardiometabolic disease are diagnosed with hypertension (Mozaffarian et al. 2015, Franklin 2006). Among all metabolic syndrome characteristics, obesity plays an important role in the development of hypertension (Morse et al. 2005). Both hypertension and obesity lead to excessive increase in sodium and water retention, which may also cause chronic kidney disease as a complication (Franklin 2006, Rahmouni et al. 2005).

Dyslipidemia in cardiometabolic disease is characterized by the levels of three major components: triglyceride-rich lipoproteins, high-density lipoproteins and small dense low-density lipoprotein particles (Ginsberg, Zhang, and Hernandez-Ono 2006). Insulin is a key hormone and regulates glucose levels in the blood. When insulin resistance appears, lipid metabolism is altered, and its oxidation is suppressed. Insulin resistance may lead to development of dyslipidemia (Ormazabal et al. 2018). Different types of lipoprotein defects increase the risk of cardiovascular disease (Ruotolo and Howard 2002).

Currently, cardiometabolic diseases have become a global epidemic and the major goal of cardiometabolic disease treatment is to reduce the risk of developing cardiovascular disease and diabetes (Grundy et al. 2005). Lifestyle modification to reduce total caloric intake and fat consumption is important in the treatment of dyslipidemia (American Heart Association Nutrition et al. 2006). However, lifestyle intervention alone

is frequently not sufficient in managing all lipoprotein defects, and medication is often still required (Miller 2009).

1.3 Natriuretic Peptides

The natriuretic peptide family is composed of different polypeptide-based hormones, which are highly conserved across species. These are the atrial natriuretic peptide (ANP), the brain/ B-type natriuretic peptide (BNP), the C-type natriuretic peptide (CNP), urodilatin (URO), the dendroaspis natriuretic peptide (DNP) and the ventricular natriuretic peptide (VNP) (Figure 1.1). Natriuretic peptides have diverse physiological roles. Each peptide interacts with its specific natriuretic peptide receptors on tissue cells. In general, natriuretic peptides play a pivotal role in body fluid homeostasis and response to metabolic homeostasis (Levin, Gardner, and Samson 1998, Wang 2012). In recent cancer studies, natriuretic peptides have also been shown to interact with the immune system and act as anti-inflammatory and anti-proliferative agents (Zhang, Li, et al. 2015, Kong et al. 2008, Mezzasoma et al. 2017). The natriuretic peptide family has tremendous therapeutic potential and many synthetic peptides have been proposed for medication treatment purposes. All natriuretic peptides have a 17 amino acid ring structure linked via disulfide bonds (Figure 1.1). The disulfide linked ring structure is essential for their biological function. Except for the C-type natriuretic peptide, all other natriuretic peptides also have a carboxy-terminal tail (van Kimmenade and Januzzi 2009).

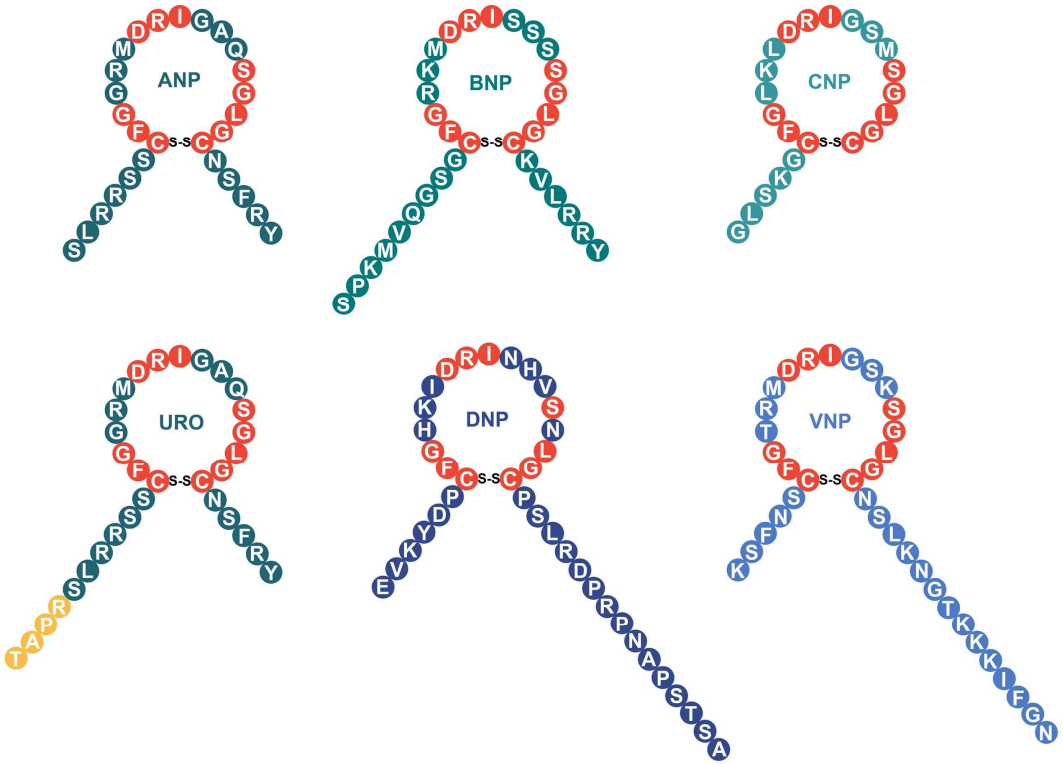

Figure 1.1 The different members of the natriuretic peptide family. Atrial natriuretic peptide (ANP), B-type natriuretic peptide (BNP), C-type natriuretic peptide (CNP), Urodilatin (URO), Dendroaspis natriuretic peptide (DNP) and Ventricular natriuretic peptide (VNP). Orange color indicates the conserved amino acids in the core structure.

Different Types of Natriuretic Peptide

Atrial Natriuretic Peptide (ANP)

ANP is mainly produced and secreted from cardiomyocytes (Macchia 1987). The atrial cardiomyocytes have electron-dense granules, which are similar to other endocrine secretory granules (Matsuo 2001). All natriuretic peptides are originally synthesized as preprohormones (Potter et al. 2009). After cleaving of the first 25 amino acids that represent the signal sequence, the remaining 126 amino acid peptide is named as proANP and is mainly stored in the atrial granules (Oikawa et al. 1984). proANP is then further

cleaved by a protease named corin to form the mature biologically active ANP (Yan et al. 2000). Matured ANP is a 28 amino acid peptide with a secondary ring structure formed by disulfide interactions (Lee and Burnett 2007). Alternatively, there is another unknown protease that cleaves proANP in the kidney to generate a 32 amino acid peptide, called urodilatin (Forssmann, Richter, and Meyer 1998). Urodilatin has been shown to play a role in regulation of salt intake and sodium and water homeostasis (Jorge et al. 2018). Released mature ANP is circulated to different targeted tissues. The concentration of ANP in circulation is in a pico-molar range and the average half-life of ANP is about two to five minutes in vivo (Weil et al. 1985, Sandefur and Jialal 2021). The short half-life in circulation is due to degradation by neutral endopeptidase (neprilysin) and the natriuretic peptide clearance receptor (Potter et al. 2009). Surprisingly, insulin-degradation enzyme also has the ability to cleave ANP, BNP and CNP peptides (Muller et al. 1992, Ralat et al. 2011). The binding affinity and degradation rate of ANP by insulin degradation enzyme are comparable to those of insulin peptide (Ralat et al. 2011, Manolopoulou et al. 2009).

The major physiological impact of ANP is to induce natriuresis, diuresis and vasodilation. Also, ANP inhibits renin and works as an antagonist in the renin-angiotensin-aldosterone system (Cannone et al. 2019). ANP is a peptide involved in intravascular fluid and blood pressure homeostasis (Matsuo 2001, Meems and Burnett 2016). Research has shown that all natriuretic peptides play a role in lipid mobilization and fat oxidation (Lafontan et al. 2008, Engeli et al. 2012). Specifically, intravenous injection of ANP increases the concentration of adiponectin (Tsukamoto et al. 2009). Adiponectin is another type of peptide secreted by adipocytes, which are involved in the

regulation of whole body energy and glucose homeostasis (Moschen, Wieser, and Tilg 2012). Adiponectin triggers the reduction of inflammation, atherogenesis and enhances insulin sensitivity (Sandefur and Jialal 2021, Cannone et al. 2019, Costello-Boerrigter and Burnett 2009). In addition, one study also indicated that ANP could hinder cancer metastasis by suppressing the inflammatory reaction of endothelial cells, thereby inhibiting cancer cell adhesion to vascular endothelial cells (Nojiri et al. 2015).

Brain/B-Type Natriuretic Peptide (BNP)

The brain/B-type natriuretic peptide (BNP) was named after its initial isolation from porcine brain tissue (Sudoh et al. 1988). Later, BNP was found to be abundant in cardiac ventricles and much less in the atrium where ANP is the majority secreted peptide (Matsuo 2001). In the BNP mRNA 3' untranslated region, there are highly repeated Adenylate-Uridylate-rich elements. The repeated AU-rich sequence tends to destabilize mRNA and regulates gene expression (Chen and Shyu 1995). Thus, BNP is considered more as a quick response agent expressed during a short time and is transcribed as needed. BNP mRNA codes the 134 amino acid preproBNP. After cleavage of the 26 amino acid signal sequence, proBNP is formed and contains 108 amino acids. Similar to the ANP formation, proBNP will be further cleaved to a matured 32 amino acid BNP before secretion (Yandle and Richards 2015).

ANP and BNP have similar physiological roles on stimulating natriuresis and vasodilatation. Vasodilation is the relaxation of blood vessels, which leads to a decrease in blood pressure. Both ANP and BNP function as vascular hormones to regulate blood pressure homeostasis (Matsuo 2001, Meems and Burnett 2016). However, BNP has a significantly longer half-life than ANP. In humans, the half-life of BNP is around 20

minutes (Smith et al. 2000). The degradation rate difference is due to differences in the peptide degradation mechanism and their affinity to natriuretic peptide clearance receptors. Unlike ANP, which is directly cleaved by neutral endopeptidase (neprilysin), BNP will be partially cleaved by a metalloprotease before further degradation by neutral endopeptidase (Pankow et al. 2007).

In healthy patients, the circulating levels of BNP are very low and are similar to the levels of ANP. However, BNP secretion increases dramatically and can be up to five times higher than ANP under conditions of cardiac stress (Matsuo 2001). The accuracy of BNP as a biomarker to diagnose acute heart failure is 83%, and elevated levels of BNP strongly predicted congestive heart failure in patients (Maisel et al. 2002).

C-Type Natriuretic Peptide (CNP)

The C-type natriuretic peptide was first isolated and purified from brain tissues, as was the case for BNP (Sudoh et al. 1990). CNP is highly expressed in brain tissue, vascular endothelial cells, and chondrocytes as a paracrine and autocrine factor (Moyes and Hobbs 2019). The human gene codes for a 126 amino acid preproCNP peptide. After cleaving of the signal sequence, the 103 amino acid proCNP is formed (Tawaragi et al. 1991) . There are two sizes of matured CNP: CNP-53 and CNP-22. CNP-53 is formed by cleaving proCNP with the intracellular serine endoprotease furin, while CNP-22 may be further modified by another extracellular enzyme based on CNP-53 (Wu et al. 2003). Although both sizes of CNP have biological activity, CNP-22 is the predominant biological peptide with 22 amino acids in the systemic circulation (Potter et al. 2009, Lumsden, Khambata, and Hobbs 2010).

CNP has a similarly short half-life as ANP does in human plasma. The circulated half-life is about 2.6 minutes (Hunt et al. 1994). Similar to ANP, CNP is inactivated either by neutral endopeptidase or internalized and degraded by the natriuretic peptide clearance receptor (Matsukawa et al. 1999). Unlike ANP and BNP, CNP has similar binding affinities to interact with natriuretic peptide receptor B and natriuretic peptide clearance receptor (He, Dukkipati, and Garcia 2006). Two different dramatic signal pathways expand the physiological functions of CNP. CNP not only modulates cardiac and vascular effects, but also controls angiogenesis, cell proliferation, anti-inflammation, and further functions (Moyes and Hobbs 2019). For example, the most recognized function of CNP in mammals is regulation of bone growth. CNP interacts with chondrocytes and stimulates endochondral ossification (Potter, Abbey-Hosch, and Dickey 2006, Peake et al. 2014).

Other Types of Natriuretic Peptides

There are other types of natriuretic peptides that were not originally discovered in mammals. Dendroaspis natriuretic peptide (DNP) is a 38-residue peptide originally isolated from the venom of the green Mamba snake (Schweitz et al. 1992). The biological function of DNP is still to be determined, but some studies indicated that DNP shares similar functions with other natriuretic peptides (Park et al. 2012). Ventricular natriuretic peptide (VNP) is another type of cardiac natriuretic peptide first discovered in eel cardiac ventricles (Takei, Ueki, and Nishizawa 1994). VNP may only exist in primitive ray-finned fishes and earlier-derived teleost species such as sturgeon, eel, and trout (Inoue et al. 2005). Studies in eel's natriuretic peptides showed that VNP has similar physiologic

effects on the cardiovascular system and osmoregulation when compared to ANP (Nobata et al. 2010, Miyanishi, Nobata, and Takei 2011).

Current Natriuretic Peptide-based Medication

ANP and BNP have been well studied and applied in synthetic peptide-based therapeutic development. In 1995, the recombinant formulation of human ANP carperitide was first approved in Japan for use in the treatment of acute heart failure (Saito 2010). However, large scale studies on carperitide were limited. More recently published carperitide studies showed that carperitide was significant associated with worse outcomes and increased in-hospital mortality (Nagai et al. 2019). A well designed large-scale randomized double-blinded study is needed to determine its clinical safety and efficacy, and it had been concluded that carperitide should be avoided as a first line treatment for acute heart failure (Matsue et al. 2015, Nagai et al. 2019). In the United States, the Food and Drug Administration (FDA) approved nesiritide as the human BNP-based treatment for patients with acute heart failure in 2001. Later, it was also available in Israel and Switzerland (Potter et al. 2009). However, newer studies showed that nesiritide may increase the risk of cardiovascular adverse effects and increased the risk of hypotension (Gong, Wu, and Li 2016, O'Connor et al. 2011). Currently, nesiritide is no longer recommended for routine use of treatment of heart failure due to a lack of significant improvement in clinical outcome or renal function (van Deursen et al. 2014).

Innovative Engineered Natriuretic Peptides

Peptide engineering opens a new field for natriuretic peptide-based therapeutics. Compared with endogenous peptides, engineering natriuretic peptides has an advantage

in efficacy and safety for cardiovascular treatment. Amino acid mutation is one of the approaches in peptide engineering. Switching key amino acids could enhance receptor activation and reduce peptide degradation speed. Also, identifying genetic alternations in natriuretic peptides and studying their effects on patients could select strong peptide candidates. For example, there are two unique peptides that were engineered based on the second approach: MANP (ZD100) and ASBNP.1 (ANX042) (Meems and Burnett 2016).

The engineered ANX042 peptide is based on a mutant that leads to an alternative splicing in BNP mRNA. The alternatively spliced BNP provides renal protection without introducing the undesired side effect hypotension which is caused by native BNP (Pan et al. 2009). The ANX042 peptide is a truncated version of the alternatively spliced BNP with 42 amino acids. ANX042 significantly boosts cGMP levels in renal tissue and enhances diuresis and natriuresis without increasing cGMP levels in vascular endothelial and smooth muscle cells (von Lueder et al. 2013). ANX042 is one of the promising treatments for heart failure. Currently, ANX042 has been in an interventional clinical trial since 2017. The estimated completion date is in September 2023 (ClinicalTrials.gov Identifier: NCT03019653).

MANP (mutant form of ANP) is another novel engineered peptide based on the discovery of a familial frame-shift mutation of the native ANP gene (Hodgson-Zingman et al. 2008). MANP contains an extra 12 amino acids at the C-terminal end of native ANP. Due to the extended C-terminal peptide tail, MANP has better protease resistance to neutral endopeptidase (neprilysin) (Dickey, Yoder, and Potter 2009). Previous *in vivo* canine model studies have shown that MANP has longer lasting and better effects on lowering blood pressure, natriuresis as well as aldosterone suppressing functions when

compared to native ANP (McKie et al. 2010). Hypertension is one of the leading conditions for causing heart failure. An epidemiology study demonstrated that there is a relative deficiency of natriuretic peptides in hypertension patients, and that patients with hypertension and metabolic disease have increased levels of aldosterone (Buglioni et al. 2015). Recent studies demonstrated that MANP decreases blood pressure, protects renal function, and increases adiponectin levels (Cannone et al. 2019, Chen et al. 2016). All of these studies have paved the way for new therapeutic applications of MANP in cardiovascular disease and metabolic syndrome. There is an ongoing clinical study to evaluate the effects of MANP in cardiovascular and metabolic diseases (Cannone et al. 2019). The first MANP interventional clinical Phase 1 trial was started in 2017 and the initial study completion date was in December 2019. The data may still be under analysis as there is currently no official update available for the study results (ClinicalTrials.gov Identifier: NCT03781739).

1.4 Natriuretic Peptide Receptors

So far, there are four subtypes of natriuretic peptide receptors identified. These are Natriuretic Peptide Receptor A (NPR-A), Natriuretic Peptide Receptor B (NPR-B), Natriuretic Peptide Receptor C (NPR-C), and Natriuretic Peptide Receptor D (NPR-D). All four peptide receptors have a relatively large peptide binding extracellular domain and small single transmembrane domain (Figure 1.2). NPR-A and NPR-B are also called particulate guanylyl cyclase receptor A (pGC-A) and B (pGC-B), respectively, due to an existing guanylyl cyclase receptor in the intracellular domain. NPR-C and NPR-D both lack the guanylyl cyclase activity and share a similar topology. NPR-D was revealed as a tetramer, but NPR-C appeared as a disulfide-linked homodimer (Hirose, Hagiwara, and

Takei 2001). NPR-D is a unique receptor and is only found in eel and pouched lamprey (Kashiwagi et al. 1995, Toop, Grozdanovski, and Potter 1998). Interestingly, NPR-D is mainly expressed in brain tissue (Kashiwagi et al. 1999), but the function of NPR-D remains to be clarified. NPR-A, NPR-B and NPR-C are known natriuretic peptide receptors in mammals. In the later paragraphs, all three mammalian receptors are introduced separately.

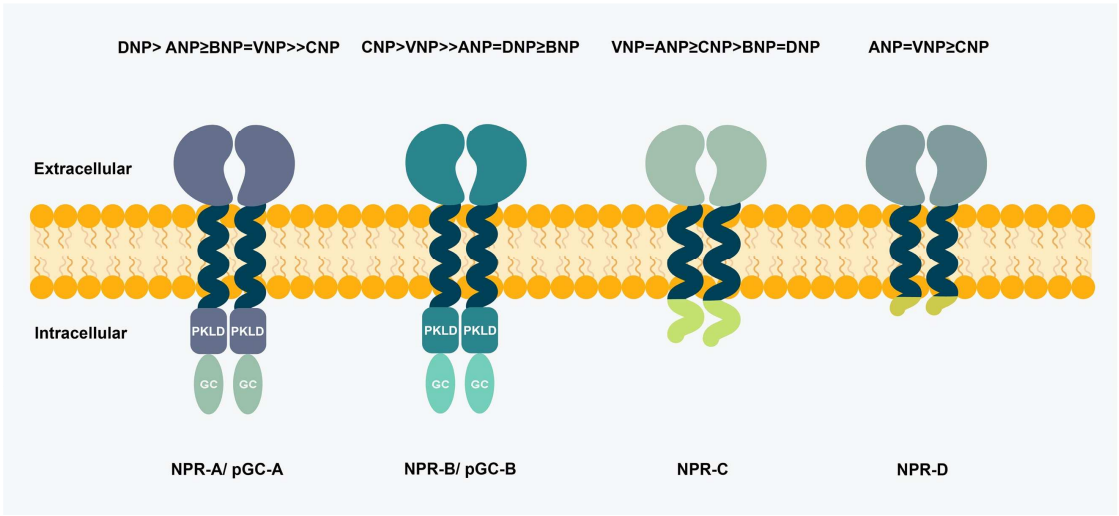

Figure 1.2 The natriuretic peptide receptor family and their binding peptides. NPR-A: natriuretic peptide receptor A; NPR-B: natriuretic peptide receptor B; NPR-C natriuretic peptide receptor C; NPR-D natriuretic peptide receptor D; PKLD: protein kinase like domain; GC: guanylyl cyclase domain; ANP: atrial natriuretic peptide; BNP: B-type natriuretic peptide; CNP: C-type natriuretic peptide; DNP: dendroaspis natriuretic peptide; VNP: ventricular natriuretic peptide.

Natriuretic Peptide Receptor A

The natriuretic peptide receptor A(NPRA) is also called particulate guanylyl cyclase receptor A (pGC-A). pGC-A is a 120kDa receptor with a single alpha helix in the transmembrane domain. Most natriuretic peptides activate pGC-A to initiate different biological reactions. pGC-A is the primary target for ANP and BNP in mammalian species. In a previous publication, it was suggested that pGC-A binds to the natriuretic

peptide with a stoichiometry of 2:1 (Rondeau et al. 1995). In mammals, the pGC-A peptide binding affinity ranks as ANP ⩾ BNP >> CNP (Suga et al. 1992). Binding natriuretic peptides at the extracellular domain (ECD) of pGC-A leads to activation of the intracellular guanylyl cyclase domain that converts the GTPs into cGMP. cGMP is an important second messenger that further initiates signaling cascades in cells by activating cGMP-dependent phosphodiesterases, protein kinases and ion-gated channels (Friebe, Sandner, and Schmidtko 2020). pGC-A is expressed in many organs and tissues including the heart, kidney and vascular endothelia. (Potter et al. 2009). There are controversial studies related to pGC-A internalization and degradation. One group suggested that most ANP-pGC-A complexes are internalized and degraded by the lysosome and small portions of intact pGC-A redistributed back to the cell membrane (Pandey 2002, 2015). However, other studies suggested that pGC-A from kidney or expressed in Chinese hamster ovary cells were not internalized or degraded in response to peptide binding. Instead, pGC-A is a membrane-resident protein and does not undergo ligand-induced endocytosis (Fan et al. 2005, Vieira et al. 2001, Koh et al. 1992).

Natriuretic Peptide Receptor B

Natriuretic peptide receptor B (NPRB) is also called particulate Guanylyl Cyclase receptor B (pGC-B). pGC-B is a 135 kDa receptor with a single alpha helix transmembrane domain. Unlike the pGC-A receptor, the principal peptide target for pGC-B is CNP. In mammals, pGC-A peptide binding affinity ranks as CNP>>ANP ⩾ BNP (Suga et al. 1992). pGC-B shares a similar structural topology and activation mechanism as PGC-A (Figure 1.2). Binding of CNP activates the intracellular guanylyl cyclase domain of pGC-B and converts GTP into cGMP. Even though pGC-A and pGC-B have

the same cGMP signaling pathway, pGC-B also has a unique biological function other than vascular and cardiac responses. pGC-B is expressed in many cell types including brain, heart, bone, chondrocyte, and ovarian tissue (Potter 2011). The CNP/pGC-B pathway is one of the most important regulators in bone growth and is involved in endochondral ossification (Vasques, Arnhold, and Jorge 2014). pGC-B is the main natriuretic receptor that is expressed in brain tissue. The CNP/pGC-B pathway might be involved in the central regulation of energy metabolism and food intake (Yamada-Goto et al. 2013). Also, the CNP/pGC-B pathway in granulosa cells inhibits oocyte maturation in mammals and prevents precocious meiotic maturation, which is critical for maturation and ovulation synchrony (Zhang, Yang, et al. 2015, Zhang et al. 2010). Additionally, one research report showed that providing CNP stimulates ovarian follicle development in infantile mice, and thus CNP could provide an alternative therapy for female infertility (Sato et al. 2012).

Natriuretic Peptide Receptor C

The natriuretic peptide receptor C (NPRC) is a disulfide- linked homodimer. Unlike pGC-A or pGC-B, NPRC lacks a protein kinase like domain and the guanylyl cyclase domain. Instead, NPRC has a short 37 amino acid intracellular tail, which has been shown to contain a G protein-activating sequence that involves a G protein-dependent signal cascade (Rose and Giles 2008). Due to the ability to activate G protein-dependent signal transduction while lacking seven transmembrane motifs that other G-Protein-Coupled-Receptors (GPCRs) have, NPRC may be considered as an atypical GPCR (Anand-Srivastava 2005, Egom 2019). In mammals, peptide binding affinity in NPRC ranks as ANP>>CNP $\geqslant$ BNP (Suga et al. 1992). Traditionally, NPRC was known as natriuretic

peptide clearance receptor and it is the most abundantly expressed natriuretic peptide receptor (Potter et al. 2009). The clearance mechanism has three main processes: internalization, peptide degradation, and receptor recycling (Nussenzveig, Lewicki, and Maack 1990). More studies indicate that NPRC plays more biological functions other than clearance of natriuretic peptides through different signal pathways. For example, NPRC is coupled to adenylyl cyclase/cAMP (cyclic adenosine monophosphate) signal transduction and inhibits adenylyl cyclase activity (Anand-Srivastava 2005). Also, NPRC has effects on the Gqα protein/mitogen-activated protein kinase (MAPK) and the PI3K-Akt pathway, which promotes cell proliferation in vascular smooth muscle cells and inhibits vascular hypertrophy (Li, Hashim, and Anand-Srivastava 2006).

1.5 The Natriuretic Peptides/ Natriuretic Peptide Receptor A/cGMP Pathway

Natriuretic peptide receptor A (NPRA) is classified as a member of the guanylyl cyclase family and is also called particulate guanylyl cyclase receptor A (pGC-A). pGC-A plays important biological and physiological roles in mammalian and vertebrate species (van Kimmenade and Januzzi 2009). The expression location of pGC-A in the body indicates some of the physiological roles of pGC-A (Figure 1.3). For example, pGC-A also acts as a neuroprotective modulator in the central nervous system (Mahinrad et al. 2016). In addition, pGC-A may be involved in the progression of cancer proliferation and angiogenesis (Qu et al. 2019, Mallela et al. 2013, Zhang, Li, et al. 2015, Zhang, Zhao, and Wang 2014). Interestingly, other studies showed that the ANP/pGC-A/cGMP pathway has anti-cancer effects, such as inhibiting tumor growth and cancer metastasis (Zhao et al. 2013, Serafino and Pierimarchi 2014, Nojiri et al. 2015). All the biological

actions are regulated by the natriuretic peptide/natriuretic peptide receptor A/cGMP signaling pathway.

Topology of Natriuretic Peptide Receptor A

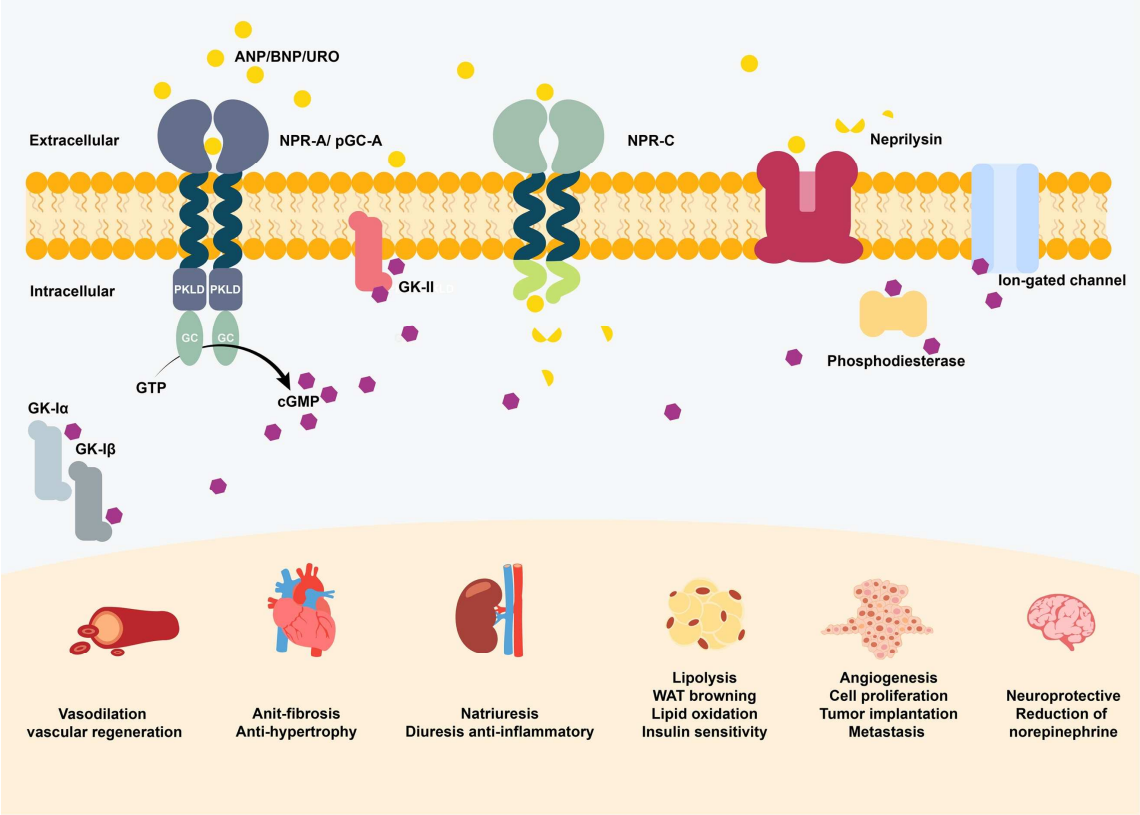

Figure 1.3 The natriuretic peptides/ natriuretic peptide receptor A/cGMP pathway. Different natriuretic peptides activate pGC-A to convert the guanosine triphosphate (GTP) into cGMP, which further interacts with protein kinase G (GK-Iα, GK-I β and GKII), ion- gated channels, and phosphodiesterases. The natriuretic peptide clearance receptor (NPR-C) and neprilysin degrade and inactivate peptides. The signal pathway activates different physiological actions in different tissues. PKLD: protein kinase like domain; GC: guanylyl cyclase domain; ANP: atrial natriuretic peptide; BNP: B-type natriuretic peptide; URO: urodilatin.

pGC-A shares a similar topological structure with other membrane-bound guanylyl cyclase proteins. pGC-A has three distinct regions: a peptide binding extracellular domain (ECD), a single helix transmembrane domain, and an intracellular

domain (ICD) that contains one protein kinase-like domain and a guanylyl cyclase catalytic domain (Figure 1.2 and Figure 1.3).

Atrial Natriuretic Peptide Activates Natriuretic Peptide Receptor A

Currently, only the extracellular domain (ECD) of rat pGC-A was solved and was accomplished with and without binding different substrate peptide ligands (Figure 1.4A, B) (Ogawa et al. 2004, Ogawa, Kodama, and Izumikawa 2020, van den Akker et al. 2000). The protein sequence of the human ECD pGC-A has an 85% similarity to the rat receptor. Based on the similarity, there may be three disulfide bonds formed in human pGC-A (Cys 92-118, Cys 196-245 and Cys 455-464). The glycosylation on the ECD of pGC-A is essential for proper protein folding and transport to the cell membrane. However, glycosylation in the ECD does not affect peptide binding (Heim, Singh, and Gerzer 1996, Koller, Lipari, and Goeddel 1993). The rat structure of the ECD of pGC-A shows that the potential glycosylation sites are not near the peptide binding site (Ogawa et al. 2004). One *in vivo* study suggested that ANP interacts with pGC-A under the chloride concentration dependent feedback-control mechanism (Misono 2000). The ECD of rat pGC-A also shows that there is a chloride atom near the l ligand binding pocket in each monomer of the ECD, and chloride interacts with the ECD via three hydrogen bonds (Ogawa et al. 2004).

ANP interacts with the ECD via three pairs of hydrogen bonds and a hydrophobic pocket in the ECD. Also, the ANP C-terminal segment partially forms a short parallel ß-sheet with the juxta-membrane site of the ECD (Ogawa et al. 2004, Misono et al. 2011). Based on the two structures of pGC-A ECD (with and without binding ANP), a rotation activation mechanism was proposed (Misono et al. 2005). They proposed that binding of

ANP induces a 24° rotation to each juxta-membrane signaling motif (Figure 1.4C). The juxta-membrane signaling motif is conserved among guanylyl cyclase receptors (Huo, Abe, and Misono 1999). Each monomer's juxta-membrane domain would rotate in the opposite direction. The hormone induced rotation transduces the rotation signaling further across the transmembrane helices and through the protein kinase-like domain toward the guanylyl cyclase catalytic domain. It was hypothesized that this rotation mechanism may bring the two guanylyl cyclase catalytic domains, which rotated in opposite directions, from a closed conformation into an active conformation (Figure 1.4D) (Ogawa et al. 2004, Misono et al. 2005).

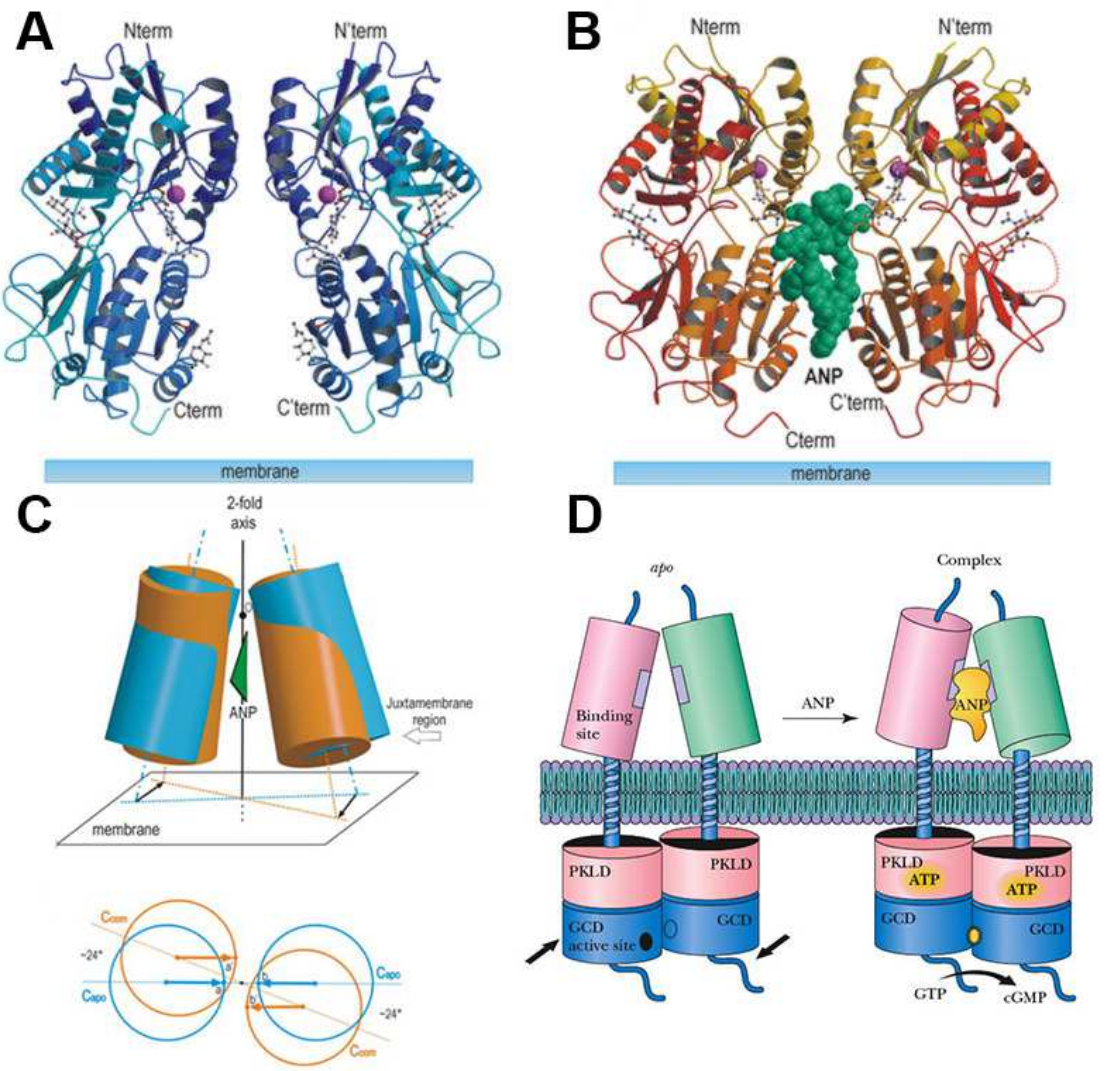

Figure 1.4 Structure of rat pGC-A extracellular domain (ECD) and proposed rotation model mechanism. A) Crystal structure of the rat apo-ECD dimer (PDB: 1DP4). B) Crystal structure of the rat ANP-pGC-A ECD complex (PDB: 1T34). ANP is shown in the space-filling model and in green. Two chloride atoms within the ECD are shown in ball shape and in purple. C) Illustrations of the rat ANP induced ECD dimer structure conformational changes. ANP induces a twisting motion on both ECD monomers. The apo position is shown in blue and the complex position is shown in orange. The simplified illustration circles represent the top view of the ECD dimer. The colored arrows indicate a parallel translocation of each ECD monomer before and after binding with ANP. Binding ANP induces a rotation mechanism and rotates each ECD domain by 24° counterclockwise. D) A proposed rotation mechanism for intact pGC-A. Abbreviations: GCD: guanylyl cyclase domain; PKLD, protein kinase like domain. Figures, caption and description are modified from (Misono et al. 2011) with permission from John Wiley and Sons.

Protein Kinase-like Domain of Natriuretic Peptide Receptor A

Protein kinase-like domain (PKLD) of human pGC-A contains 280 amino acids. Currently, the structural model of PKLD is based on a homology model of protein tyrosine kinases. The actual structure and regulatory functions remained to be discovered (Misono et al. 2011). There is a 43 residue-linker region of PKLD that connects to the guanylyl cyclase domain. This linker region may be necessary for oligomerization of pGC-A and regulatory function of the guanylyl cyclase domain (Wilson and Chinkers 1995, Lowe 1992). The PKLD is phosphorated on six residues (S497, T500, S502, S506, S510, and T513) near the putative ATP binding site. Site-directed mutagenesis of these residues has shown that phosphorylation of pGC-A is essential for peptide-dependent activation (Potter and Hunter 1998). However, the specific protein kinase that phosphorylates pGC-A has not yet been identified (Pandey 2015). There are many earlier studies which showed that ATP increases the activity of pGC-A and acts like an allosteric regulator. However, the mechanism is still under debate (Potter et al. 2009). There is a known conserved ATP binding motif GXGXXG that has been found in many protein kinases (Hanks, Quinn, and Hunter 1988). PKLD in pGC-A has a similar motif, GRGSNYG, near the phosphorylation sites and is suggested as a putative ATP binding site (Goraczniak, Duda, and Sharma 1992). One particular report suggested that PKLD mediates the regulatory reaction of ATP for signal transduction to activate the guanylyl cyclase and lower the pGC-A peptide binding affinity in order to release the ligand (Jewett et al. 1993). Controversially, one paper proposed ATP-independent activation of pGC-A, and they postulated that ATP may play a role in maintaining proper receptor phosphorylation status (Antos et al. 2005).

Guanylyl Cyclase Domain of Natriuretic Peptide Receptor A

The guanylyl cyclase catalytic domain (GCD) of human pGC-A contains 250 amino acids. So far, there are two solved GCD structures, which are both from photosynthetic organisms: CYG12, which is a soluble guanylyl cyclase catalytic domain from green algae, and Cya2, which is a membrane bound GCD from a prokaryotic cyanobacterium (Winger et al. 2008, Rauch et al. 2008). In these two structures, the guanylyl cyclase catalytic domain is represented in a dimer form and two GCD monomers are in twofold axis symmetry (Winger et al. 2008, Rauch et al. 2008). Based on the known GCD structures, at least two copies of pGC-A are required for guanylyl cyclase function. The protein sequence suggested that CYG12 may be more likely to show structural homology to the soluble guanylyl cyclase in mammals, and that Cya2 may be more similar in homology to membrane-bounded enzymes (Pandey 2011). When comparing the human pGC-A GCD sequence with CYG12 and Cya2, the sequence alignment showed that human pGC-A guanylyl cyclase is 51.5 % and 31.6 % identical to CYG12 and Cya2, respectively (Figure 1.5). Based on sequence homology, guanylyl cyclases and all known adenylyl cyclases share high sequence similarity and all belong to the class III nucleotide cyclase family (Linder and Schultz 2003). Many research studies suggested that intact pGC-A exists either as a homodimer (Qiu et al. 2004, Misono et al. 2005) or tetramer (Iwata et al. 1991, Jewett et al. 1993) in the native state, and the oligomerization of pGC-A is ligand independent (Chinkers and Wilson 1992, Lowe 1992).

```
Program: CLUSTALO
Alignment sequence: NPR1 and CYG12
Identical positions: 68
Identity: 51.515%
Similar positions: 33

   1   TIYFSDIVGFTALSAESTPMQVVTLLNDLYTCFDAVIDNF-DVYKVETIGDAYMVVSGLP        59
   1   TVLFSDIVGFTEIASRSSPLEVCSLLDELYQRFDAAIEEYPQLYKVETIGDAYMVVCNVT        60
       *: ******** :::.*:*::* :**::**  ***.*::: ::**************..:

  60   VRNGRLHACEVARMALALLDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGVVGLKMPRYCL       119
  61   VPCDDHADV-LLE--FALRMHEEASRVASSLGEPVRIRVGMHSGPVVAGVVGRKMPRFCL       117
       *  .      : .  :**    .: *:     * :*:*:*:*:*** ***** ****:**

 120   FGDTVNTASRME                                                       131
 118   FGDTVNTASRME                                                       129
       ************

Program: CLUSTALO
Alignment sequence: NPR1 and Cya2
Identical positions: 43
Identity: 31.618%
Similar positions: 34

   1   TIYFSDIVGFTALSAESTPMQVVTLLNDLYTCFDAVIDNFDVYKVETIGDAYMVVSGLPV        60
   1   TILTSDLRGFTSTSEGLNPEEVVKVLNIYFGKMADVITHHGGTIDEFMGDGILVLFGAPT        60
       **  **: ***: *   .* :**.:**  :  :  ** :..    * :**. :*: * *.

  61   RNGR----LHACEVARMALALLDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGVVGLK-MP       115
  61   SQQDDALRAVACGVE-MQLALREV--NQQVTGLGLQPLEMGIGINTGEVVVGNIGSEKRT       117
        :        ** *  * *** :.  . ::     : *.: ***:** * .* :* :

 116   RYCLFGDTVNTASRME                                                   131
 118   KYGVVGAQVNLTYRIE                                                   133
       :* :.*  ** : *:*
```

Figure 1.5 Summary of protein sequence alignment. protein sequence of human pGC-A guanylyl cyclase domain (NPR1, residues 876-1006) was compared to the guanylyl cyclase domain (CYG12, residues 477-605) from *Chlamydomonas reinhardtii* and guanylyl cyclase domain (Cya2, residues 443-7575) from *Synechocystis sp*. Dashes indicate mismatches or gaps, one dot indicates low similarity between residues, two dots indicate a higher similarity between residues, and pipes indicate identical residues.

cGMP Is a Second Messenger in Cells

Cyclic guanosine monophosphate (cGMP) was first isolated from rat urine (Ashman et al. 1963) and classified as a second messenger molecule in cells. Guanylyl cyclases catalyze guanosine-5'-triphosphate (GTP) to cGMP and pyrophosphate. So far, three different types of proteins were discovered as the intracellular targets of cGMP: cGMP-dependent protein kinases I and II (GKs), cGMP-dependent phosphodiesterases (PDEs) and cGMP-dependent ion-gated channels (CNGs) (Figure 1.3) (Friebe, Sandner,

and Schmidtko 2020). Production of cGMP is stimulated by various hormones and factors. Nitric oxide or natriuretic peptides target soluble guanylyl cyclases or membrane bound particulate guanylyl cyclases, respectively. GK-I is a homodimer with two isoforms: GK-Iα and GK-Iβ. GK-I plays a key role in vascular and cardiac functions in mammals (Park, Sandner, and Krieg 2018). One of the downstream effects of GK-I may involve increasing the Ca^{2+} concentration. For example, the cGMP dependent GK-I phosphorylates phospholamban, which a micro-peptide protein that modulates the sarcoplasmic reticulum (SR) Ca^{2+} ATPase (SERCA) to uptake Ca^{2+}. The SR is a membrane-bound intracellular organelle and involves in the storage and distribution of Ca^{2+} in cardiac cells (Figure 1.6). The SR participates in cardiac excitation-contraction coupling (Inserte et al. 2014, Frank and Kranias 2000). cGMP-dependent ion-gated channels are also called cyclic nucleotide-gated channels (CNCs). CNCs belong to the voltage-gated nonselective cation channel (Kaupp and Seifert 2002). CNCs are activated and open directly by binding of cGMP or cAMP (cyclic adenosine monophosphate) molecules. Binding cyclic nucleotides modulates the flow of sodium, potassium, and calcium into the cells and further triggers membrane depolarization or hyperpolarization events. CNCs play an important role in signal transduction in retinal photoreceptors and olfactory receptor neurons (Kaupp and Seifert 2002). Interestingly, CNCs behave as molecular amplifiers, and a lower concentration of cGMP causes large activity changes (Brown et al. 2006). CNCs contains four subunits, and each subunit has one cyclic nucleotide binding site. When four ligands fully bind to its subunits, the channel reaches its maximum conducting state (Ruiz and Karpen 1999). Phosphodiesterases (PDEs) tightly regulate degradation of cGMPs. There are eleven families of PDEs based on the

homology of the C-terminal catalytic domain (Baillie, Tejeda, and Kelly 2019). Among the different PDE proteins, PDEs 5, 6, and 9 have a higher affinity to hydrolyze cGMPs (Lee et al. 2015). Many cancer studies have shown decreased levels of cGMP with increased PDE activity (Baillie, Tejeda, and Kelly 2019, Yarla et al. 2019). There is a therapeutic strategy based on cGMP signaling pathway by either increasing cGMP concentration or reducing cGMP's catabolism with PDE inhibitors (Park, Sandner, and Krieg 2018).

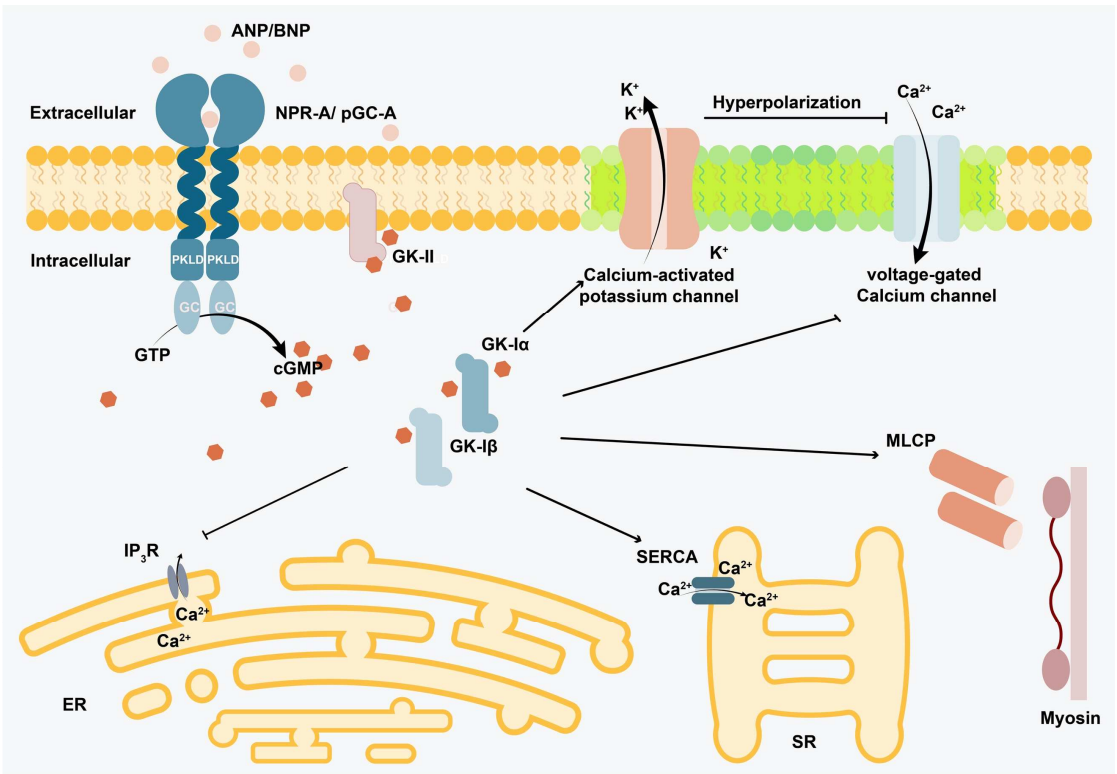

Figure 1.6 Natriuretic peptide receptor A is involved in different vasodilation mechanisms in vascular smooth muscle cells. pGC-A produced cGMPs activate cGMP-dependent protein kinase G proteins. In the vascular smooth muscle cells, the activated protein kinases further control the calcium ion levels and myosin filament-induced cell contraction. Abbreviations: ANP: atrial natriuretic peptide; BNP: B-type natriuretic peptide; ER: endoplasmic reticulum; GC: guanylyl cyclase domain; IP_3R: inositol trisphosphate receptor; MLCP: Myosin light chain phosphatase; NPR-A/pGC-A: natriuretic peptide receptor A; PKLD: protein kinase like domain; protein kinase G (GK-Iα, GK-I β and GKII), SERCA: Sarcoplasmic reticulum Ca2+ ATPase; SR: sarcoplasmic reticulum.

1.6 Natriuretic Peptide Receptor A Roles in Pathophysiology

The natriuretic peptide receptor A (pGC-A) is activated by ANP and BNP. Upon activation, pGC-A further activates different intracellular cGMP signaling pathways. The natriuretic peptides/natriuretic peptide receptor A/cGMP pathways play important roles in pathophysiology. Besides their well-known roles in blood pressure and cardiac homeostasis, as well as renal homeostasis, this pathway also regulates fat cell metabolism.

Regulation of Blood Pressure and Cardiac Homeostasis

When ANP interacts with pGC-A, it regulates blood pressure homeostasis through increasing water excretion (diuresis) and salt excretion (natriuresis) in the kidneys, and vasodilation of vascular smooth muscle cells to increase endothelium permeability (Potter et al. 2009). Vasodilation is achieved by cGMP dependent protein kinase GKI leading to a lower level of intracellular calcium (Figure 1.7) (Potter, Abbey-Hosch, and Dickey 2006). cGMP activated protein kinase GKIα targets and phosphorylates (activates) calcium-activated potassium channels (Alioua et al. 1998). Decreasing the potassium in the intracellular environment causes membrane hyperpolarization and further blocks calcium influx to the cell via voltage-gated calcium channels. The cGMP activated protein kinase GKI also stimulates the calcium/ATPase pump on the cell membrane to pump calcium out of the cell, but the exact cellular mechanism has yet to be determined (Potter, Abbey-Hosch, and Dickey 2006). Also, cGMP activated protein kinase GKIβ phosphorylates inhibit the inositol 1, 4, 5-trisphosphate receptor on the Endoplasmic Reticulum (ER) to pump calcium from storage vesicles to the cytoplasm (Schlossmann et al. 2000). In vascular smooth muscle cells, cGMP activated protein kinase GK-I indirectly triggers sarcoplasmic reticulum (SR) Ca^{2+} ATPase (SERCA) to increase uptake of Ca^{2+} into the

SR from the cytoplasm (Inserte et al. 2014, Frank and Kranias 2000). The SR is involved in storage and distribution of Ca^{2+} in cardiac cells. SR is a specialized membrane-bound intracellular organelle. It participates in cardiac excitation-contraction coupling (Inserte et al. 2014, Cornwell et al. 1991). It has also been shown that the cGMP activated protein kinase GKIα phosphorylates (activates) myosin light chain phosphatase to decrease the calcium sensitivity (Nakamura et al. 1999). Myosin light chain phosphatase plays roles in dephosphorylation of the regulatory light chain of the motor protein myosin-II, and modulates actomyosin-based contractility (Schaub 2007).

Partial overexpression of pGC-A in mice models clearly showed that pGC-A is directly involved in body salt sensitivity and blood pressure (Oliver et al. 1998). Genetic impairment of ANP in transgenic mice led to the development of hypertension (Melo et al. 1998). Additionally, the deficiency of pGC-A gene expression led to cardiac hypertrophy and atherosclerosis in mice models (Alexander et al. 2003). One study that compared the pathophysiologic differences in knockout pGC-A and wildtype mice models showed that the knockout mice have chronic hypertension and progressive cardiac changes. The knockout mice initially had cardiomyocyte hypertrophy. Then, they developed cardiac fibrosis with reduced cardiac contractility in the later stage (Kuhn et al. 2002). Genetic impairment of BNP in transgenic mice also caused extracellular matrix production, and the mice also developed cardiac fibrosis (Tamura et al. 2000). Decreasing the function through genetic alteration in either natriuretic peptides or pGC-A directly reduces the natriuretic peptides/natriuretic peptide receptor A/cGMP signaling pathway and adversely affects the blood pressure and cardiac homeostasis in mice.

Regulation of Renal Homeostasis

One knockout study of the natriuretic peptide receptor-A (pGC-A) in mice showed that ANP dependent diuresis and natriuresis are fully controlled via pGC-A (Kishimoto, Dubois, and Garbers 1996). In the kidney, ANP/pGC-A/cGMP pathway regulates diuresis and natriuresis by multiple mechanisms. The ANP/pGC-A interaction increases the glomerular permeability and filtration rate by increasing the pressure of glomerular capillaries (Potter, Abbey-Hosch, and Dickey 2006). cGMP-dependent signaling regulates glomerular afferent vasodilation and efferent vasoconstriction (Marin-Grez, Fleming, and Steinhausen 1986). In the proximal tubules and collecting duct of the kidney, the ANP/pGC-A complex leads to sodium and water resorption. ANP decreases proximal tubule resorption by inhibiting angiotensin-stimulated sodium and water transport (Harris, Thomas, and Morgan 1987). ANP triggered cGMP-dependent signaling inhibits epithelial sodium channels in the collecting ducts (Light, Corbin, and Stanton 1990). In addition, the ANP/pGC-A inhibits the renin-angiotensin-aldosterone system in kidney. Increased cGMP activates the cGMP-dependent protein kinase GKII to decrease renin secretion in the juxtaglomerular cells (Burnett, Granger, and Opgenorth 1984, Wagner et al. 1998). Other studies have shown that ANP decreases aldosterone secretion from the adrenal gland (Shi et al. 2001, Kudo and Baird 1984). ANP induced inhibition of aldosterone production may be involved with the cGMP-dependent phosphodiesterase PDE2. Activated PDE2 hydrolyses cAMP (cyclic adenosine monophosphate), which is another important secondary messenger in cells and participates in the aldosterone synthesis pathway (MacFarland, Zelus, and Beavo 1991).

Regulation of Fat Cell Metabolism

In the early 2000s, a new lipolysis pathway, which is involved in ANP/pGC-A/cGMP pathway, was for the first time demonstrated in human adipose tissue (Figure 1.7). *In vivo* natriuretic peptide infusion experiments have also shown that ANP has the most potency on lipolytic effects via the pGC-A receptor (Sengenes et al. 2000). Later studies have shown that the ANP/pGC-A/cGMP triggered lipolysis pathway in fat cells is exclusively found in primates. This could be due to the higher expression ratio of natriuretic peptide receptor C (NPRC) to pGC-A in non-primates (Sengenes et al. 2002). Interestingly, the ANP/pGC-A/cGMP pathway stimulated lipolysis pathway is not involve with phosphodiesterase PDE-3B, which is the main enzyme that hydrolyzes cAMP in the adipocyte and inhibits lipolytic signaling (DiPilato et al. 2015). The ANP/pGC-A/cGMP pathway activates cGMP-dependent protein kinase GKI to further activate perilipin A and hormone sensitive lipase (HSL) mediated triglyceride degradation (Schlueter et al. 2014). Perilipin A is a lipid droplet surface protein and will translocate HSL leading to triacylglycerol hydrolysis and release of fatty acids (Tansey et al. 2004). Both ANP and BNP activated pGC-A/cGMP pathways enhance lipid oxidation and burn released fatty acids by a more "energy efficient" β-oxidation pathway in adipose tissue, skeletal tissue, and liver (Schlueter et al. 2014). Moreover, natriuretic peptides turn on cGMP-dependent protein kinase GKI and activate AMP-activated protein kinase (AMPK) and p38-MAPK to enhance mitochondrial biogenesis and lead to browning of white adipose tissue (Schlueter et al. 2014, Bordicchia et al. 2012).

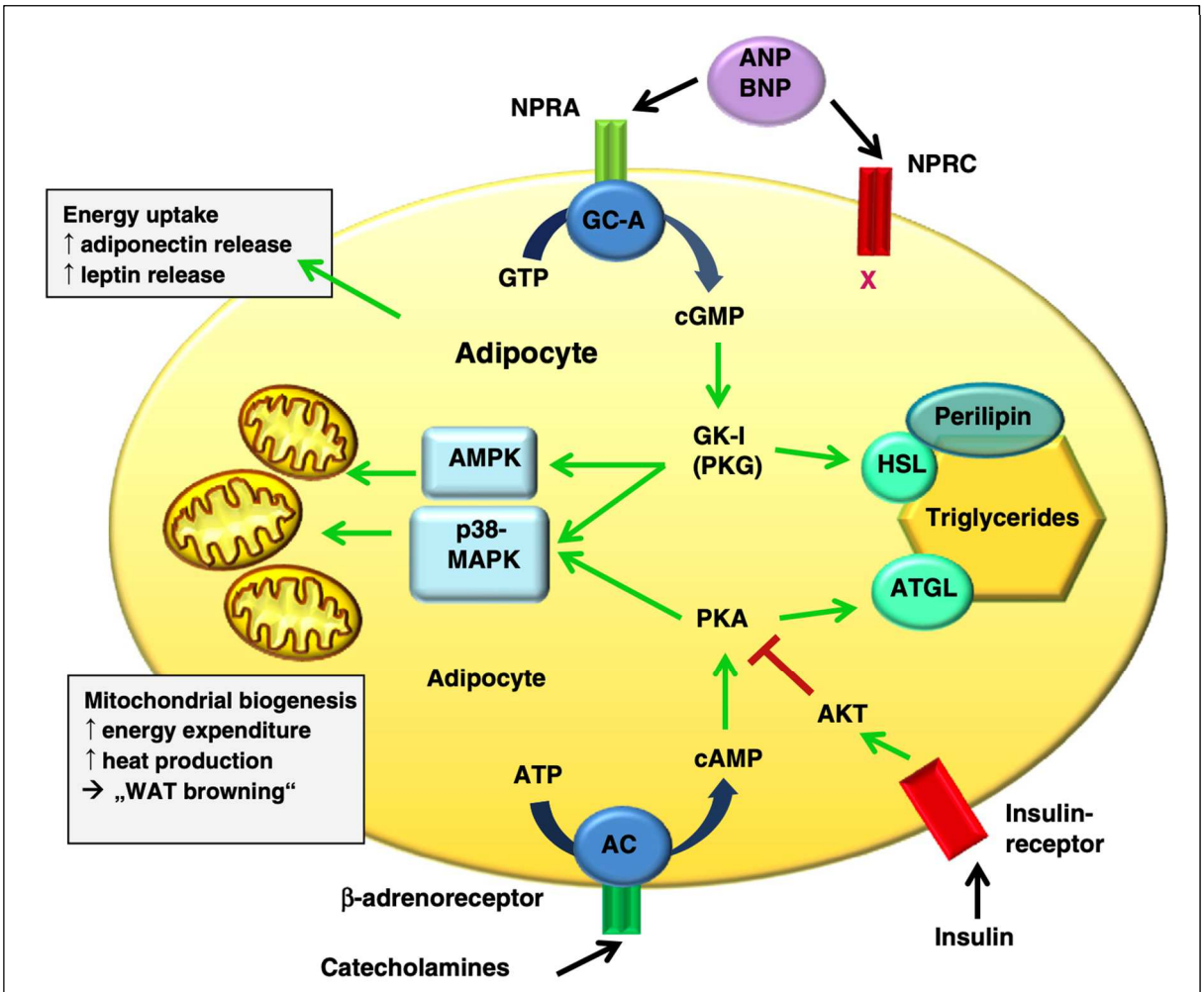

Figure 1.7 Scheme of natriuretic peptide receptor A (NPRA/pGC-A) involved human adipocytes. Binding of ANP and BNP with the extracellular domain of pGC-A induces a cGMP dependent pathway. Increased cGMPs further activate cGMP-dependent protein kinase G proteins and further activate lipolysis by phosphorylation of hormone sensitive lipases. In addition, activated protein kinase G stimulates mitochondrial biogenesis and browning of white adipose tissue by phosphorylation of AMPK and p38-MAPK. Abbreviation: AC: adenylyl cyclase; ANP: atrial natriuretic peptide; AMPK: AMP-activated protein kinase; ATGL: adipocyte triglyceride lipase; BNP: type-B natriuretic peptide; cGMP: cyclic guanosine monophosphate; HSL: hormone sensitive lipase; NPRA/pGC-A: natriuretic peptide receptor A; PKG: protein kinase G; PKA: protein kinase A; p38-MAPK: mitogen-activated protein kinase. Figure, caption, and description are modified from (Schlueter et al. 2014) with permission from Elsevier.

1.7 Summary and Conclusion

Natriuretic peptides and their receptors are well expressed in different organ systems in humans and have distinct physiologic outcomes. The natriuretic peptide receptor A (pGC-A) not only plays an important role in vascular and cardiac homeostasis but is also involved in lipid metabolism and cancer development. All these ongoing studies paved new therapeutic directions for ANP and the pGC-A receptor. ANP/pGC-A/cGMP pathways may have additional therapeutical potential beyond known body fluid and blood pressure homeostasis.

Currently, the natriuretic peptides ANP and BNP are used as biomarkers for acute cardiac diseases, and the approved endogenous peptide medications did not provide the therapeutic effects that we expected. There are many innovative engineered natriuretic peptides that may soon fill the therapeutic gap. There have been many studies that involve a genetic mutation approach to understand the pathophysiology of pGC-A in animal models over the last two decades, but many cellular mechanisms still need to be determined. There are some controversial study results related to natriuretic peptides and their receptors. Newly developed molecular biotechnology may help answer these questions.

With the detailed crystal structure information of the rat pGC-A ECD but lacking the structure information of the transmembrane domain and ICD of pGC-A, the actual signal transduction and activation mechanism of the ANP/pGC-A/cGMP pathway are still unclear. The knowledge of the full-length (intact) structure of the pGC-A receptor with and without the bound ANP is required to reveal the molecular basis of the function of the ANP/pGC-A/cGMP pathway. Importantly, pGC-A contains a GCD and is also

classified in the membrane-bound guanylyl cyclase family of a super-family of single-span transmembrane receptors (Drewett and Garbers 1994, Misono et al. 2005). Many single-span transmembrane receptors use a ligand-induced association activation mechanism in which the ligand triggers the receptor to bring the enzyme close to active site. The rotational mechanism initiating pGC-A signaling transduction and activating the GCD by rotation signaling is the first mechanism identified within single-span transmembrane receptors (Misono et al. 2011, Misono et al. 2005). Also, the membrane-bound guanylyl cyclase structure information is yet to be determined.

CHAPTER 2

EXPRESSION AND PURIFICATION OF INTACT HUMAN NATRIURETIC PEPTIDE RECEPTOR A

2.1 Abstract

The natriuretic peptide receptor A is also known as particulate guanylyl cyclase receptor A (pGC-A) owing to the GCD. pGC-A is an ANP receptor critical for controlling cardiovascular, renal, and endocrine homeostatic functions (Pandey 2011). To effectively develop a method for regulating pGC-A, structural information regarding its intact form needs to be determined. Currently, only the structure of the ECD of rat pGC-A has been determined, both with and without binding different substrate ligands (Ogawa et al. 2004, Ogawa, Kodama, and Izumikawa 2020, van den Akker et al. 2000). However, structural data regarding the transmembrane domain, as well as functional intracellular domain regions, remain unknown and need to be clarified. For expressing and purifying functional intact human pGC-A for structural purposes, *Spodoptera frugiperda* 9 (Sf9) insect cells were selected as the protein expression system. The insect cell expression system is one the most suitable for expressing proteins of eukaryotic origin. For example, many human GPCR structures were resolved using GPCRs expressed in the insect cell system (Schneider and Seifert 2010, Massotte 2003). The insect cell expression system provides several advantages in terms of protein expression, including protein post-translational modifications and a better protein folding system than bacterial expression systems (Massotte 2003). Compared with the traditional *Escherichia coli (E. coli)* expression system, the insect cell expression system allows for easier extraction of fully folded membrane proteins directly from the cell membrane, instead of the bacterial

expression of misfolded non-functional proteins or isolation and refolding of proteins from *E. coli* inclusion bodies (Baneyx and Mujacic 2004, Schmidt 2004). Figure 2.1 summarizes the ideal major experimental steps from plasmid design to pathways for structure determination of the intact human pGC-A. In this Chapter, I will focus on the expression and purification of the intact human pGC-A receptor.

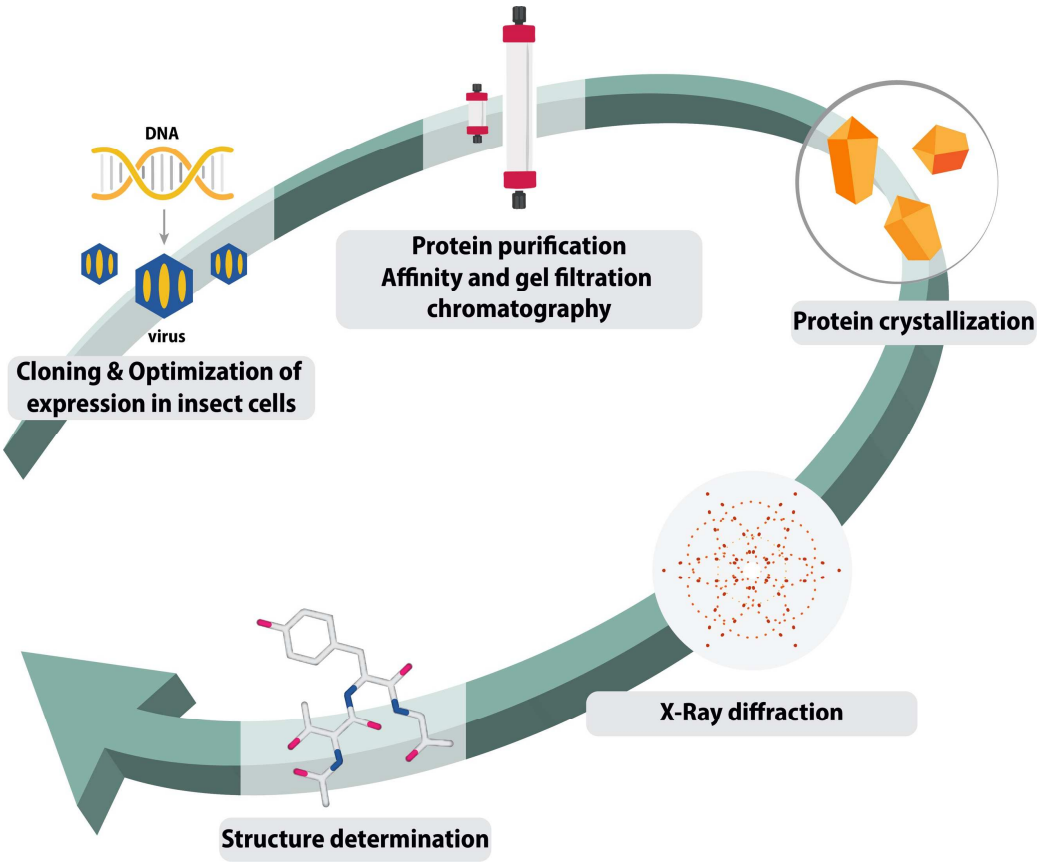

Figure 2.1 Summary scheme of ideal major experiment steps from cloning to structure determination of intact human pGC-A.

2.2 Introduction

2.2.1 Protein Expression System

Recombinant protein technology paved the foundation for advanced protein-based science. Especially in biopharmaceutical development and structural biology research, the first essential step is the precise expression of functional recombinant target proteins in different hosts. Recombinant technology achieved its first breakthrough in the early 1970s with the discovery of restriction enzymes (Celie, Parret, and Perrakis 2016). In 1976-1977, the first functional human hormone protein, somatostatin, was recombinantly expressed in *Escherichia coli* (Itakura et al. 1977). Within a few years, insulin, the first recombinant human protein, was expressed in *E. coli* and was approved by the United States FDA and marketed in the U.S. to treat diabetes (1982, Quianzon and Cheikh 2012). Currently, with the rapid development of biosciences and biotechnology, recombinant proteins can be expressed in different hosts. Expression systems can range from prokaryotic strains to eukaryotic mammalian cell lines. Additionally, recombinant proteins could be expressed in cell-free systems (Dondapati et al. 2020). Notably, that there is no perfect protein expression system, and several pros and cons need to be assessed based on research goals.

Prokaryotic System

Prokaryotic protein expression systems are the most widely employed both in academic research and bio-industry. These expression systems are also classified as bacterial expression systems and primarily based on *E. coli* expression (Porowinska et al. 2013). In general, bacterial expression offers high productivity with low cost. Within a relatively short time frame, bacterial systems can grow to high cell density and achieve

large-scale protein expression quantities with inexpensive chemicals. Moreover, the genetic information of *E. coli* is well-documented, and different regulatory promoters were well-characterized (Bervoets and Charlier 2019, Ishihama 2018). Numerous vectors and mutant *E. coli* strains are well-characterized, which can be selected based on specific research needs (Terpe 2006).

Typically, the *E. coli* protein expression system leads to fast-folding kinetics, which usually is a double-edged sword. For heterologous proteins with multiple domains, as well as for proteins that are toxic or contain hydrophobic domains such as membrane proteins, there is a high possibility for inclusion body formation in *E. coli* (Kaur, Kumar, and Kaur 2018). In addition, the *E. coli* expression system has minimal ability to include post-translational modifications. For example, as wildtype *E. coli* possesses a reducing cytoplasmic environment, the formation of intramolecular or intermolecular disulfide bonds is unlikely; these are crucial for rectifying the folding of several protein complexes. Due to the complexity of eukaryotic-based proteins, many eukaryotic proteins require proper folding mechanisms, essential lipids, or molecular chaperones to be functional and properly folded. Eukaryotic expression hosts have become the preferred choice for protein structure determination (Pandey et al. 2016). According to the Protein Data Bank (PDB) entries, about 8% of total protein structures were expressed in eukaryotic hosts by the end of 2017, and the use of eukaryotic hosts has steadily increased from 1992 to 2017 (Franke et al. 2018).

Up to 2017, the FDA-approved peptide and protein therapeutics database (HPdb) documented a total of 239 FDA-approved therapeutic peptides and proteins (Usmani et al. 2017). Specifically, there are more than 170 approved therapeutic protein-based drugs

from recombinant peptides or proteins (Puetz and Wurm 2019). From 1982 to January 2009, 151 recombinant proteins were licensed as pharmaceutical treatment by the FDA and European Medicines Agency (EMEA). Up to 2009, 29.8% of approved recombinant proteins were expressed in *E. coli*, 18.5% were expressed in yeast (*Saccharomyces cerevisiae*), 0.75% were expressed in insect cells, and 39% were expressed in mammalian cell lines (Ferrer-Miralles et al. 2009). Between 2000 to 2004, the ratio of mammalian and non-mammalian cells employed for expressing approved recombinant protein products was reversed. More FDA-approved recombinant protein products were expressed in mammalian cells (Ferrer-Miralles et al. 2009). After 2004, mammalian cell lines were predominantly used to express approved protein products. For example, about 60% of approved recombinant protein products were expressed in mammalian cells (Sanchez-Garcia et al. 2016). One reason is because mammalian cells can adequately express glycosylated proteins (Zhu 2012). Although the prokaryotic expression system is no longer considered the first choice for expressing recombinant pharmaceuticals, 30% of marketed pharmaceuticals continue to rely on the *E. coli* system owing to cost-effectiveness, unique physiological characteristics, and genetic manipulability (Sanchez-Garcia et al. 2016, Chen 2012).

Eukaryotic System

The yeast expression system is an intermediate between bacterial and the more complex insect or mammalian cell expression systems. Like prokaryotic systems, yeast is a unicellular organism but with a more complicated genome structure. The physiology, fermentation requirements, and genetic modifications are well-studied and established (Huang, Lowe, and Batt 2010). *S. cerevisiae* was the first eukaryote with a completely

sequenced eukaryotic genome (Bassett et al. 1996). Unlike the prokaryotic system, yeast possesses a eukaryotic expression system and thereby can perform critical post-translational modifications for targeted recombinant proteins. *S. cerevisiae* is considered a superior model to produce various pharmaceutical proteins compared to *E. coli*. Prior to 2010, all FDA- and EMEA-approved recombinant protein products exclusively derived from eukaryotic organisms were produced using *S. cerevisiae* (Huang, Lowe, and Batt 2010). Compared with insect and mammalian cell expression systems, yeast expression requires substantially lower nutrition demands and can be easily adapted to harsh large-scale industrial conditions for routine protein expression (Vieira Gomes et al. 2018). One main limitation of the yeast expression platform is hyper-mannosylation. Yeast tends to perform a high-mannose type N-glycosylation (Nielsen 2013). Hyper-mannosylated glycoproteins are more likely to lead to immunogenic reactions and thus reduce serum half-time and therapeutic efficiency (Baghban et al. 2019, Nielsen 2013), as well lead to allergic side effects in patients. Furthermore, yeast plasmids are unstable, and large-scale production requires a fermentation reactor (Baghban et al. 2019, Zhang, Moo-Young, and Chisti 1996).

Despite the cost and high-standard daily cell maintenance, the mammalian cell expression platform is the preferred system for mammalian-based proteins. More than 50% of approved therapeutic recombinant protein products are expressed in mammalian cell lines (Zhu 2012). The unique advantage of the mammalian cell system is that mammalian cells provide a natural and similar environment for protein synthesis. Mimicking the human protein expression environment offers almost native transcription and translation mechanisms, similar protein chaperone folding systems, and post-translational

modifications. For example, many eukaryotic proteins required native glycosylation for correct expression and folding (Chang et al. 2007). These advantages allow mammalian cells to produce near-natively formed functional proteins and become the best fit candidate for complicated eukaryotic protein expression (Dyson 2016).

The mammalian cell expression system can utilize either the transient transfection or stable transfection methods for protein expression. Under the transient transfection method, the recombinant virus with protein of interest (POI) DNA segment can be transfected into cells. The entire process is quick and can produce a high amount of target proteins. For example, the transiently transfected cell is generally ready to harvest from 24 to 96 h after cell transfection. However, transfected DNA material is not integrated into the host genome and is lost during cell passaging. The transient transfection method only allows limited effectiveness (McKenzie and Abbott 2018, Dukkipati et al. 2008). Under the stable transfection method, established stable transfection cells constantly produce the POI after integrating the DNA sequence into the host genome. The drawback of a stable cell line is that it is a time-consuming process. Multiple selective screenings are required to isolate one successful stable transfectant, which can constantly express the gene of interest. The low protein yield is another drawback of the stable transfection method, which can be attributed to the integration location on the host chromatin position or the low-copy number of transfected DNA successfully integrated into the host chromosome (Wurm 2004). Here, four different popular mammalian cell lines for protein structural purposes are listed: Chinese hamster ovary cells (CHO), human embryonic kidney cells (HEK293), baby hamster kidney cells (BHK-21), and monkey kidney fibroblast cells (COS-7) (Andrell and Tate 2013). CHO cells, BHK-21 cells, and murine myeloma cells (NS0 and Sp2/0) are

most frequently used for therapeutic production purposes (Dumont et al. 2016). In addition, there are also uniquely engineered, and "personalized" mammalian cell lines based on the targeted protein goals. For example, the cell line CHO Lec.3.2.8.1 is engineered to decrease the level and complexity of N-glycosylation on protein expression and can help protein crystallization for structural determination (Stanley 1989, Soler et al. 2018).

Cell-Free System

Both prokaryotic and eukaryotic protein expression platforms are based on utilizing cells as the fundamental expression host. In contrast, a cell-free protein synthesis system uses partially purified cellular components. The cell-free system operates the translation machinery, which is directly adapted from cells and avoids constraints of the cell membrane. Compared with traditional protein expression systems, a cell-free system provides high specificity and timesaving from initial protein synthesis to downstream protein applications (Dondapati et al. 2020). The cell-free expression system can be based on either prokaryotic or eukaryotic origins. In the cell-free method, the cell origin needs to be selected for the appropriate translation machinery, and dedicated preparatory steps are needed to lyse cells and isolate cell extracts without cell debris and genomic DNA (Gregorio, Levine, and Oza 2019). In principle, the cell extract contains all necessary components for protein transcription and translation (Figure 2.2.1). The isolated cell extract can be stored in a -80°C freezer for several years. This open expression system without cell membrane constraints allows the easy introduction of chemically modified amino acids into the cell-free system to express novel proteins (Bundy and Swartz 2010). Given this "beyond-the-cell" characteristic, the cell-free system is an appropriate tool to

study and express cytotoxic proteins and even some membrane proteins (Sobhanifar et al. 2010, Watanabe et al. 2010).

Over the last five years, cell-free expression methods have been rapidly developed. Different organisms were tested as the cell extract source, and numerous proteins were successfully expressed in a cell-free environment (Dondapati et al. 2020). For example, by employing the protein synthesis using recombinant elements (PURE) system, which is an optimized cell-free system, the membrane protein F_oF_1-ATP synthase was synthesized with detectable activity (Kuruma and Ueda 2016). However, significant challenges remain to be addressed. The cell-free system is considerably expensive and cannot appropriately produce proteins on a large scale, as the increased reaction volume (Zemella et al. 2015). Furthermore, isolated transcription and translation factors are less active than those in the original host cells (Dondapati et al. 2020, Botte, Deniaud, and Schaffitzel 2016). Certain absent chaperone proteins may affect protein folding and functionality. The cell-free expression platform has enormous potential. However, substantial work is required to optimize and decrease the cost of this expression system.

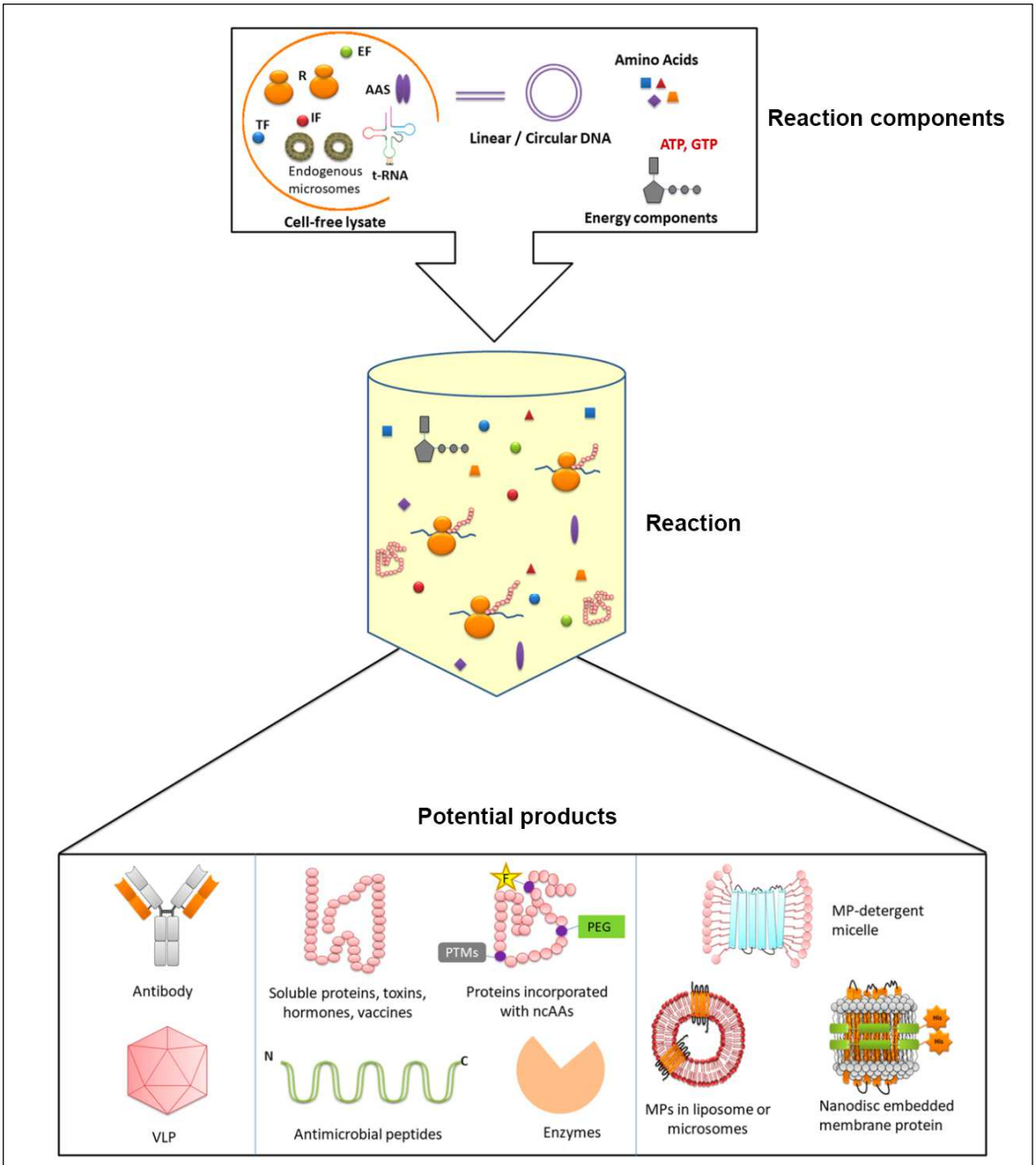

Figure 2.2.1 General scheme depicting the overall process of cell-free protein production. *aatRNA* aminoacyl-tRNA, *AAS* aminoacyl-tRNA synthetase, *ATP* adenosine triphosphate, *EF* elongation factor, *GSH* glutathione, *GSSG* glutathione-disulfide, *GTP* guanosine-5'-triphosphate, *IF* initiation factor, *IRES* internal ribosome entry site, *MP* membrane protein, *nCAA* non-canonical amino acid, *PDI* protein disulfide isomerase, *PEG* polyethylene glycol, *PTM* post-translational modification, *R* ribosomes, *t-RNA* transfer RNA, *TF* transcription fac- tor, *UTR* untranslated region, *VLP* virus like particle. Figure, caption, and description are used from (Dondapati et al. 2020) with permission from Springer Nature.

2.2.2 Baculovirus Expression Vector System

Baculovirus Life Cycle

Baculoviruses are rod-shaped insect viruses. The virus name is derived from the Latin word *baculum*, implying "stick," and is used to describe the shape of the viral nucleocapsid (Haase, Sciocco-Cap, and Romanowski 2015). Baculoviruses contain circular double-stranded DNA, and depending on the baculovirus species, the genome sizes range between 80 to 180 kbp (van Oers and Vlak 2007). There are two baculoviruses commonly used for recombinant protein expression: *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV) and *Bombyx mori* nuclear polyhedrosis virus (BmNPV) (van Oers 2011). Many baculovirus expression vector system (BEVS) are constructed and based on the AcMNPV genome to transfect *Spodoptera frugiperda* (Sf9) or BTI-Tn-5b1-4 (High Five) cells. AcMNPV can transfect different cell lines, but the BmNPV based expression system is limited to silkworm larvae (Wu et al. 2004, van Oers 2011).

Within the baculovirus life cycle, baculoviruses are classified into two virion types based on the envelope composition and virus function in the life cycle: budded virions (BVs) and occlusion derived virions (ODVs) (Clem and Passarelli 2013). Usually, the BV envelope only packs one nucleocapsid, and ODVs may carry multiple nucleocapsids. Subsequently, the viral nucleocapsid is transported into the host nucleus and uncoated for viral replication (Clem and Passarelli 2013). Virions are formed in the nucleus and budded out, with the envelope from the nuclear membrane forming BVs (Granados and Lawler 1981). Later, BVs are budded from the infected cell membrane, actively spreading the infection to other cells within the host; ODVs are embedded entirely in the nucleus of infected cells. Depending on the baculovirus genus, ODVs may pack single nucleocapsid

per virion (SNPV) or multiple nucleocapsids per virion (MNPV). When ODVs are embedded with a crystalline matrix, ODVs become occlusion bodies (OBs), which are the viral form and afford protection to the baculovirus to survive outside the host (van Oers 2011). Depending on the baculovirus genus, granulin protein-based granular OBs only carry one ODV and are classified as granulovirus; large polyhedral-based OBs can carry over hundreds of ODVs and are classified as nucleopolyhedrovirus (NPV) (Harrison et al. 2018).

The natural replication cycle of baculovirus occurs when the host consumes food contaminated with OBs (Figure 2.2.2). The alkaline environment in the host midgut triggers the dissolution of the polyhedrin matrix and releases the ODVs from the OBs. Released ODVs fuse, or through absorptive endocytosis, enter the midgut epithelial cells (Chambers et al. 2018). Then, viral nucleocapsids are released and transported to the host nucleus via actin filaments (Ohkawa and Welch 2018). The viral gene transcription has been divided into four phases: immediate-early, delayed-early, late, and very late phases (Hefferon and Miller 2002). In the immediate-early phase, host RNA polymerase II transcribes viral transregulators and early viral genes, which activate the delayed-early phase. In the delayed-early phase, some viral genes are transcribed to encode proteins required for viral DNA replication and late gene expression (Hefferon and Miller 2002). Simultaneously, viral proteins, such as apoptosis inhibitors, that prevent host defense mechanisms are expressed (Clem 2007). During the late phase, viral DNA is replicated, and nucleocapsids and viral envelope proteins readily assemble BVs for virus budding in the nucleus. In the very late phase, the production of BVs is reduced, and two proteins are expressed at high levels (Passarelli and Guarino 2007). Polyhedrin is an essential protein to

form the OB matrix. P10 protein plays a crucial role in releasing OBs from the host nucleus (Carpentier, Griffiths, and King 2008).

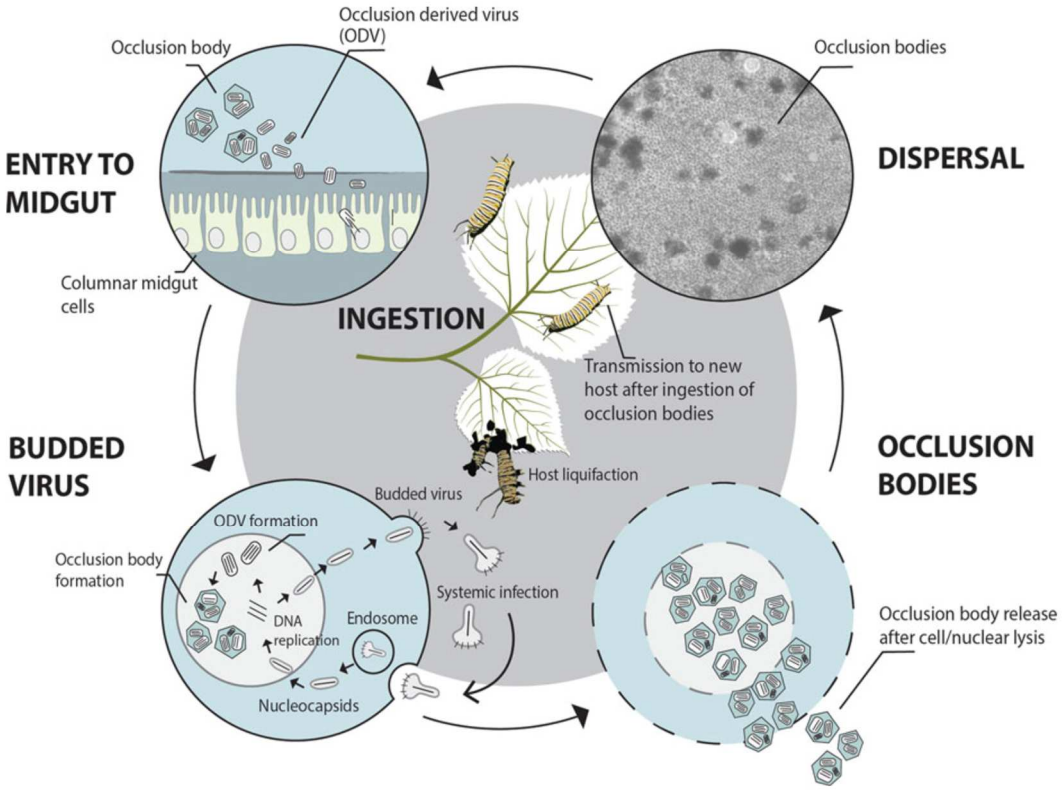

Figure 2.2.2 Summary of the baculovirus life cycle. **The ingestion step**: larvae ingest food containing baculovirus occlusion bodies (OBs). The OBs travel to the larva midgut. Due to the alkaline environment in the midgut, OBs dissolve and release the occlusion derived viruses (ODVs). **The entry to the midgut step**: The ODVs penetrate the columnar midgut cell membrane by a direct membrane fusion mechanism. The nucleocapsids are released from ODVs into the cytoplasm after the ODVs enter the cells. Then, the released nucleocapsids migrate into the nucleus and prepare for viral replication processes. The baculovirus is assembled inside the host's nucleus. **The budded virus (BV) step**: Newly generated BVs actively infect the larva host and enter the host cell via a receptor-mediated endocytosis. At the later endosome acidification step, the envelopes from BVs fuse with endosomal envelopes and release the viral nucleocapsids into the cytoplasm. **The occlusion bodies (OBs) step**: During the very late phase, nucleocapsids are accumulated in the nucleus to form ODVs. Several copies of formed ODVs are embedded within a polyhedrin crystalline matrix and generate the OBs. Newly formed OBs are released by natural host cell lysis, and the infection cycle will start again in a new host. Figure, caption and description are modified from (Chambers et al. 2018).

Overview of the Baculovirus Expression System

The BEVS is a powerful method to express recombinant proteins in various insect cell lines. The most popular BEVS is based on AcMNPV. AcMNPV is one of the most well-studied baculovirus species (Kool and Vlak 1993). In the last decades, the viral vectors were well-studied and manipulated to improve protein yield and quality. Simultaneously, different genetically engineered insect cell lines were generated for better protein folding mechanisms with post-translational modifications than traditional prokaryotic cells. Notably, the baculovirus has a large genome size. Directly cloning the foreign gene into the virus genome is impractical. One of the breakthroughs is the development of the bacmid technology (O'Reilly, Miller, and Luckow 1994). Bacmid is a recombinant baculovirus vector, which contains the baculovirus genome. In the engineered *E. coli* cell (DH10 β strain), the bacmid acts as a large plasmid with replication ability. After transforming the foreign gene (gene of interest) into the *E. coli* cell containing the bacmid via the standard heat shock method, the target gene contained transposition sites is spontaneously integrated into the baculovirus genome via site-specific transposition (Luckow et al. 1993).

Historically, the first generation of the recombinant baculovirus vector was generated by homologous recombination. However, the chance of homologous recombination is lower than 0.01%, and the recombinant virus purification is also time-consuming (van Oers 2011). Based on the original method, a linearized baculoviral genome (the second generation) was designed that enhanced the recombinant efficiency to up to 30% (Kitts, Ayres, and Possee 1990). Later, optimization of the 2nd generation linearized baculoviral genome was designed without the *Orf1629* gene. The second generation

baculovirus vector further improving the recombinant virus recovery to 90%. The *Orf1629* gene encodes a structural protein called phosphoprotein PP78/83, which is involved in nuclear actin filament formation during baculovirus infection and is essential to viral replication and infectivity (Slack and Arif 2007). Successful homologous recombination of the target gene restores and transfers the *Orf1629* gene back to the viral genome (Kitts and Possee 1993). As previously mentioned, the third generation bacmid technology involves a "bacterial artificial chromosome" vector with the baculovirus (AcMNPV) genome. Bacmid technology utilizes Tn7-recombinase to mediate transposition to insert a foreign gene into the bacmid instead of homologous recombination (Luckow et al. 1993). Verified recombinant bacmids can be directly inserted into insect cells using a chemical transfecting agent to generate high titer recombinant virus particles for future insect cell transfection and recombinant protein expression (van Oers 2011). The Invitrogen commercial Bac-to Bac™ system (used in this work) is based on the 3rd generation bacmid technology. Subsequently, a fourth generation of high-throughput recombinant method was developed based on the Bac-to Bac™ system. The new high-throughput recombinant method repairs the *Orf1629* gene back to the bacmid through homologous recombination (Possee et al. 2008). After homologous recombination was successfully achieved, the gene of interest inserts into the bacmid and restores the *Orf1629* gene simultaneously. The restoring *Orf1629* gene feature only allows the generated recombinant bacmid to replicate in insect cells after recovering the *Orf1629* gene (Radner et al. 2012, Hitchman, Possee, and King 2012). Unlike the Bac-to-Bac™ system, this new method avoids bacmid selection and verification steps. The commercial *flash*BAC™ from Oxford expression

technologies and BacMagic™ from Novagen are also based on the optimized bacmid concept with the Orf1629 gene.

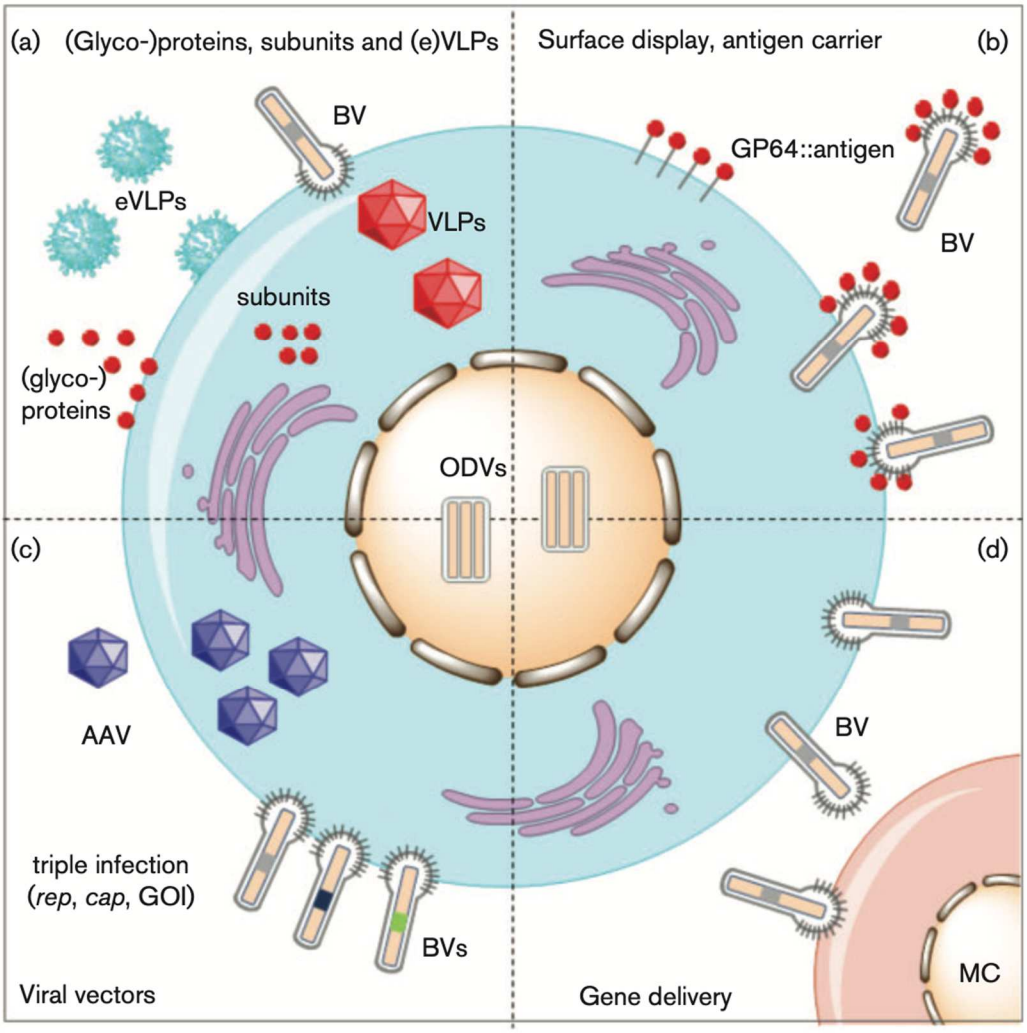

Figure 2.2.3 Summary of baculovirus expression vector system (BEVS) applications. A) Heterologous protein synthesis. Synthesized recombinant proteins could remain in the cytoplasm, translocated to cell membrane, or secreted from the cells. In addition, the recombinant proteins may form (enveloped) virus like particles (eVLPs/VLPs). Generated eVLPs may be further purified for vaccine production. B) Baculovirus surface display/antigen carrier. Target (antigen) protein is fused with the surface protein GP64 from baculovirus. GP64 serves as a carrier for the target protein and transports the target protein to the host cell membrane surface. Later, the target protein is integrated into the BV envelope upon budding. C) Viral vectors. BEVS is also used to generate adeno-associated virus (AAV) vectors for gene therapy. Different baculovirus vectors individually carry different partial AAV genes, such as replication gene segments, capsid gene segments, and the gene of interest. The recombinant AAVs can be generated inside the transfected insect cells. D) Gene delivery. BEVS is also used to deliver the gene of interest into mammalian cells. Harvested recombinant BVs transfect mammalian cells by endocytosis. Figure, caption and description are modified from (van Oers, Pijlman, and Vlak 2015) with permission from the Microbiology Society.

The application of the BEVS is not limited to the expression of recombinant proteins. The BEVS can also be utilized to synthesize baculovirus particles (virus-like particles) as antigen carriers for vaccine development. In addition, the BEVS can act as a gene delivery vector in mammalian cells for cancer vaccine treatment and express other viral vectors for gene therapy production (Figure 2.2.3) (van Oers, Pijlman, and Vlak 2015). For example, the first commercial vaccine produced by the BEVS is a veterinary vaccine for preventing classical swine fever in pigs (Moormann et al. 2000). By the end of 2015, there were nine BEVS-derived commercial products (Felberbaum 2015). Four of nine are human vaccines, including human papillomavirus and influenza vaccines, prostate cancer therapy, and lipoprotein lipase deficiency (Felberbaum 2015). AcMNPV baculovirus can enter mammalian cells by clathrin-mediated endocytosis and possibly micropinocytosis (Matilainen et al. 2005). Switching the baculoviral promoter with a mammalian viral promoter boosts gene expression level in mammalian cells, but the recombinant bacmid lacks replication ability in mammalian cells (Boyce and Bucher 1996, Thimiri Govinda Raj et al. 2020). The first concept approval for the BEVS gene delivery was the application in prostate cancer immunotherapy (Swift et al. 2013). The BEVS successfully expressed adeno-associated virus (AAV) vectors in Sf9 cells. The BEVS-produced recombinant AAV vector is the first AAV mediated gene therapy and was approved to treat a metabolic disorder (Haddley 2013).

Insect Cell Expression System

In the early 1980s, the first recombinant protein expressed via insect cell/BEVS was the human beta interferon (Smith, Summers, and Fraser 1983). There are some major steps for recombinant protein expression. In the first, insect cells need to reach an ideal log phase

(1.5-2.5 $\times 10^6$ cells/mL) concentration with 95% or higher cell viability. Second, the recombinant baculovirus, which contains the POI DNA, is added to the insect cell culture for transfection. The insect cell expression level is related to the characteristics of the target protein. In general, a lower expression level for functional glycoproteins is observed due to protein processing and trafficking steps (Contreras-Gomez et al. 2014).

Although the insect cell system provides the most of the post-translational modifications found in mammalian cells, there are some limitations. For example, the insect cell system limits N-linked glycosylation modification. In the N-linked glycosylation pathway, both insect and mammalian cells share the same high-mannose N-glycan precursor and intermediate. At the end of the glycosylation pathway, insect cells mainly produce a paucimannose N-glycan structure, but mammalian cells lead to a complex type with terminal sialics acids N-glycans (Contreras-Gomez et al. 2014, Shi and Jarvis 2007). This limitation can be improved by engineering the insect cell line with enzymes to produce the same mammalian glycosylation patterns (Hollister and Jarvis 2001, Aumiller et al. 2012).

More than 500 insect cell lines have been established from over 100 insect species (Lynn 2007). Cells derived from Lepidoptera insects have commonly been used for recombinant protein expression. Primarily, the insect cells isolated from undifferentiated ovarian and embryonic tissues have been employed as they allow the establishment of continuous diploid cell lines (Contreras-Gomez et al. 2014). The three most used insect cell lines in the BEVS include Sf9, Sf21, and High Five™. Sf9 (Summers and Smith 1987) and Sf21 (Vaughn et al. 1977) are both derived from the pupal ovarian tissue from *Spodoptera frugiperda* fall armyworm; High Five™ (Wickham et al. 1992) is derived from the

ovarian tissue from *Trichoplusia ni* cabbage looper. High Five ™ is commercialized by Invitrogen, and the actual cell line name is BTI-Tn-5b1-4 (Granados et al. 1994).

All three insect cell lines can be grown in both adherent culture, in which cells must be attached to a surface to grow, and suspension culture, in which cells grow in an agitated growth medium. Suspension culture provides cell culture size flexibility, and the serum-free media simplifies the downstream purification of secretory protein (McKenzie and Abbott 2018). Although High Five ™ tends to reach higher recombinant protein production on a per cell basis, Sf9 and Sf21 are more commonly used for large-scale protein expression due to higher cell density growth (Contreras-Gomez et al. 2014).

The BEVS may not be the ideal choice in the pharmaceutical industry, but it certainly demonstrates advantages in membrane protein structural biology. Based on the Protein Data Bank from 1995 to 2012, the insect cell expression system is the most common host choice among eukaryotic expression systems (Assenberg et al. 2013). More than 70% of known GPCR structures were solved using insect cells as the protein expression host (Abiko et al. 2021). This success is based on relatively cheaper and easier cell culture maintenance than mammalian cells and more advanced cellular protein synthesis mechanism and post-transitional modifications than yeast and *E. coli* (Saarenpaa, Jaakola, and Goldman 2015, Abiko et al. 2021). Surprisingly, reports showed that soluble proteins expressed in insect cells could be "auto-packed" and crystallized *in vivo* to form high diffraction quality protein crystals (Koopmann et al. 2012, Tang et al. 2020). However, the exact crystallization mechanism is yet to be determined.

Bac-to-Bac™ System

Based on the AcMNPV viral genome, a "bacterial artificial chromosome" (BAC) vector was used to carry the AcMNPV viral genome and known as a bacmid. The bacmid can replicate in the *E. coli* (DH10Bac strain). After transforming the DNA vector containing the gene of interest (GOI) into the DH10 β *E. coli,* a spontaneous genetic transposition occurs, and the GOI is swapped to create a recombinant bacmid (Luckow et al. 1993). The Bac-to-Bac™ system from Invitrogen became one of the most widely used BEVS in both bioindustry and academic research (Shang et al. 2017).

As mentioned earlier, the Bac-to-Bac™ system is based on site-specific transposition. The GOI is cloned on a special vector called pFastBac™. Depending on the specific pFastBac™ vector type, either the polyhedrin or p10 promoter from AcMNPV can be used to secure POI expression levels. Both sides of the pFastBac™ expression cassette are flanked with a Tn7 transposon (Tn7L and Tn7R) to achieve site-specific transposition (Figure 2.2.4). For improved accuracy, two antibiotic-resistant genes are present in the pFastBac™ vector. The ampicillin resistance gene is for general plasmid selection in *E. coli*; and the gentamicin resistance gene is within the mini-Tn7 element region and permits selection of the recombinant bacmid in the specialized DH10Bac *E. coli*.

The DH10Bac *E. coli* strain contains the bacmid and a helper plasmid. The bacmid has multiple features. First, the bacmid has a mini-*att*Tn7 target site. The spontaneous transposition occurs between the mini-Tn7 element on the pFastBac™ vector and the mini-*att*Tn7 target site on the bacmid. Second, the bacmid uses a mini-F replicon to allow low-copy viral genome (bacmid) replication in DH10Bac *E. coli*. Third, the mini-*att*Tn7 on the bacmid contains the LacZα peptide gene. Insertion of the mini-

Tn7 region (containing the GOI) into the mini-*att*Tn7 on the bacmid disrupts the expression of the LacZα peptide. Without the α peptide for the complementation process, the β-galactosidase synthesized in DH10Bac *E. coli* is inactive. Active β-galactosidase hydrolyzes X-gal from the environment to form 5-bromo-4-chloro-indoxyl, which spontaneously dimerizes to produce an insoluble blue pigment called 5,5'-dibromo-4,4'-dichloro-indigo (Julin 2018). Therefore, a successfully transposed recombinant bacmid can be selected via the blue-white colony method. Finally, the bacmid has a kanamycin resistance gene for selection. The helper plasmid codes an enzyme, transposase, which is essential for DNA transposition. Moreover, the helper plasmid contains an antibiotic-resistant gene for tetracycline (Invitrogen Life Technologies 2002). Collectively, the DH10Bac *E. coli* strain with a successfully transposed recombinant bacmid grows as a white colony on the X-gal LB plate, which also contains kanamycin, gentamicin, and tetracycline for selection. Once the recombinant colony is selected, recombinant bacmids are amplified and purified (by the same method as bacterial plasmid purification). Purified recombinant bacmids will be confirmed via PCR for the presence of the GOI. The bacmid viral DNA initializes the expression cascade, and high titer recombinant virus stock can be harvested from the culture supernatant for later protein expression (Figure 2.2.4) (Invitrogen Life Technologies 2002).

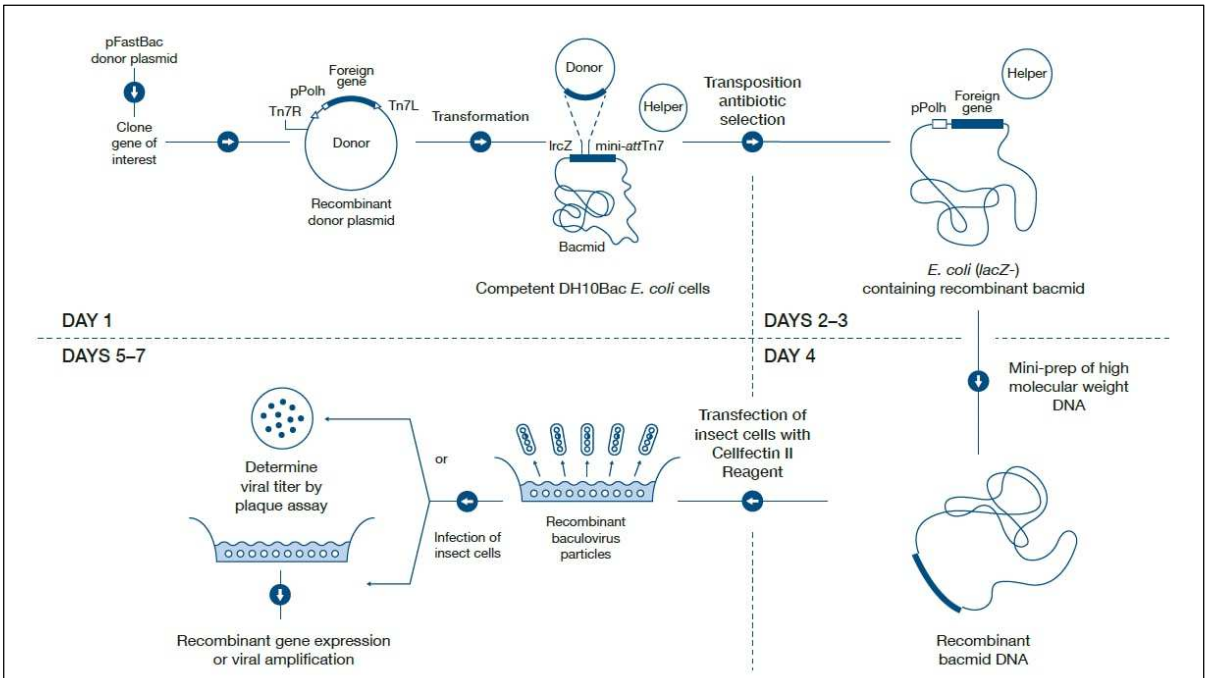

Figure 2.2.4 Scheme of generation of recombinant baculovirus and expression of recombinant protein in the Bac-to-Bac™ Baculovirus Expression System. Illustration was adapted from the Thermo Fisher Scientific website.

2.2.3 Protein Purification

Overview of Protein Purification

Protein-based biotechnology is associated with many industrial applications and fundamental research. All the industrial applications and experimental research require a high amount of purified POI with functionality. The first protein isolation attempt was reported in 1789 by Antoine Fourcroy. This was over two centuries ago; developing an effective and rapid protein purification method still has a high demand. Several milestones have been achieved in the protein purification methodology. During World War II, Edwin Joseph Cohn developed a protein purification technique to isolate the serum albumin and other proteins from blood plasma. The purification technique was substantially complicated and required changing buffer pH, temperature, ethanol concentration, and protein

concentration during each protein precipitation step (Scopes 2001). Cohn first introduced the ammonium sulfate precipitation method, and this precipitation method is still widely employed in current protein purification research (Wingfield 2001, Cohn 1922). In 1924, Theodor Svedberg built a revolutionary tool, the prototypical ultracentrifuge. The ultracentrifuge instrument paved the road toward modern protein purification and sedimentation analysis of proteins (Schuck 2003, Mandal et al. 2017). The next purification technique breakthrough is chromatography. There are two major types: liquid chromatography and gas chromatography. Liquid chromatography is commonly used to separate biomolecules, such as nucleotides and proteins (Ali et al. 2010). Liquid chromatography involves a column packed with a special resin to form the solid stationary phase. When the liquid mobile phase, which contains the POI, passes through the resin, POI interacts with the resin via different physicochemical interactions. These different interactions could be protein charge (ion-exchange chromatography), protein molecular size (size exclusion chromatography), specific (ligand) binding interaction (affinity chromatography), and hydrophobicity (hydrophobic interaction chromatography). Depending on the interaction method, the POI can be eluted at a specific point or after impurities passed through the chromatography (Labrou 2021, McCue 2009).

Lastly, recombinant DNA technology has played a game-changing role in protein science over the last decade. Recombinant DNA technology provides an opportunity to express the POI in other host expression systems for maximum protein yield and avoid directly utilizing/ purifying with the original infectious pathogen hosts. Genetic optimization or mutation can improve protein properties for downstream applications such as protein function studies or structural investigations. Additionally, the recombinant DNA

technology simplified protein purification steps by fusing an affinity tag to the POI (Wingfield 2015, Labrou 2021).

Untagged proteins isolated from natural sources typically require multiple purification steps. Despite tagged proteins undergoing simplified purification methods, additional purification steps may still be needed. In protein science, one fundamental goal is to achieve maximum protein yield and purity, both efficiently and economically. Generally, the sample preparation step mainly separates the POI from other cell debris by centrifugation after lysis of the protein source materials. Depending on the types of POIs, alternate and non-chromatography purification methods are employed to further isolate and purify the targeted POI during the capture step. The goal of the capture step removes significant contaminants from the POI. The intermediate purification step can be omitted but may be useful for untagged proteins purified directly from the native source. In the intermediate purification step, most contaminants (nucleic acids, other proteins) should be removed before undergoing the final polishing step. In the polishing step, the remaining impurities should be removed. Usually, a high-resolution gel filtration column is used in the final purification step. The gel filtration column can further separate aggregated and degraded POIs and change the protein into another buffer for downstream analysis.

Difficulty of Membrane Protein Expression and Purification

Membrane proteins play an important role in biological and physiological functions in both prokaryotic and eukaryotic organisms. In most organisms, 20-30% of the genome is dedicated to the expression of membrane proteins (Kermani 2020). Membrane proteins are amphipathic macromolecules with both hydrophobic regions and hydrophilic regions. The hydrophobic region is generally a large central transmembrane

domain that is embedded in the cell membrane bilayers. The bilayers are comprised of different amphipathic lipids (Sadaf et al. 2015).

Understanding the functionality of membrane proteins requires expression and purification, which is more difficult than for soluble proteins. The primary factor is low expression yield (Birch et al. 2018). In some bacterial-based expression host studies, overexpression of membrane proteins tended to cause growth inhibition and disruption of central metabolism in host cells (Gubellini et al. 2011, Jensen et al. 2017). Due to the "toxic nature" of membrane proteins, heterologous expression of membrane proteins leads to lower protein yield (Gubellini et al. 2011, Pandey et al. 2016). More importantly, some membrane proteins may require essential lipids, specific chaperones, and proper post-translational modifications for correct protein folding (Pandey et al. 2016).

Purification of membrane proteins is also very challenging. There are two major factors that affect membrane protein purification: poor solubilization and instability after extraction from its native membrane (Birch et al. 2018). Purification is performed in aqueous solutions, but membrane proteins are insoluble in aqueous solutions due to their exposed hydrophobic regions after isolation from lipid membranes. Additionally, the bilayer architecture of cell membranes provides a lateral pressure on membrane proteins that prevents protein aggregation and denaturation (Anglin and Conboy 2008, Samuli Ollila et al. 2007). Furthermore, many membrane proteins require large conformational changes to achieve cellular functionalities (Orellana 2019, Alenghat and Golan 2013). The highly dynamic structure of membrane proteins tends to lead to self-degradation after extraction from the membrane (Sadaf et al. 2015, Deller, Kong, and Rupp 2016).

Detergent molecules have historically been the most popular amphipathic agent for membrane protein solubilization (Stroud, Hall, and Dafforn 2018). In the early 1980s, the fabrication of detergent-based micelles led to a breakthrough in membrane protein solubilization (Caffrey 2003). Detergents have the same amphipathic characteristics as membrane lipids. Detergent molecules form micelles by a hydrophobic effect when the detergent concentration reaches a critical micelle concentration (White and Wimley 1999). When non-polar or hydrophobic molecules are added into aqueous solutions, the original intermolecular hydrogen interactions in water are disrupted. These non-polar or hydrophobic molecules cause water molecules to reorganize, which results in decreased entropy in the system (Bhairi 1997). To compensate for the loss of entropy, water molecules force the hydrophobic molecules into a small enclosed space (Bhairi 1997). Because of this hydrophobic effect, detergent-formed spherical micelles have a hydrophobic interior and a hydrophilic exterior. Detergent amphipathic molecules surround the hydrophobic segment of membrane proteins and mimic the native membrane lipidic system (Sadaf et al. 2015). Generally, detergents play two critical roles in membrane protein purification: 1) Firstly, detergents extract most of the membrane proteins out of the cell membranes and into an aqueous solution; 2) Secondly, detergent micelles keep the targeted membrane proteins soluble and stable in the later purification steps. There are at least 120 detergents for membrane protein studies (Sadaf et al. 2015). However, choosing the appropriate types and concentrations of detergent is crucial to prevent membrane protein aggregation and self-denaturation. Therefore, it is common for multiple detergent screenings to be performed during the purification of membrane proteins (Kotov et al. 2019).

Precipitation Purification Method

The precipitation-based protein purification method is strictly dependent on the characteristics of the POI. Precipitation is a crude purification method and allows the removal of up to 90% of cell proteins (Lojewska et al. 2016). Typically, the precipitation method is combined with protein dialysis and other chromatography methods to remove the remaining impurities and change the buffer system (Liu et al. 2020). One common precipitation method is based on the protein isoelectric point (pI). When the buffer pH reaches the target protein's pI value, the target protein also reaches the lowest solubility. Contaminant proteins also have pI values. Adjusting the buffer pH can separate the target protein from other proteins (contaminant proteins are precipitated out from the solution). One disadvantage of isoelectric precipitation is that extremely low or high pH conditions may impair target protein functionality (Liu et al. 2020). Nevertheless, isoelectric precipitation is widely used in the food industry due to its economic and convenience features (Fawole et al. 2018). Another popular precipitation method is salting out. The addition of salt to the protein solution removes water molecule competition between the added salt and protein molecules and mediates the hydrophobic interaction of the protein to reduce protein solubility (Liu et al. 2020). The goal of salting-out is to precipitate contaminant proteins and keep targeted protein in solution. In addition, salting-out is preferred prior to hydrophobic interaction chromatography (HIC) purification. The high salt solution also enhances hydrophobic reactions between HIC chromatography resins and proteins to achieve higher resolution (Tsumoto et al. 2007). For example, ammonium sulfate precipitation is the most popular choice for the salting-out method owing to its high solubility, low price, and effectiveness. After removing the protein contaminants, the HIC

may be used to further purify the protein sample, which is present in a high concentration of ammonium sulfate (Duong-Ly and Gabelli 2014b).

Popular Chromatography Purification Methods

HIC is based on hydrophobicity to separate different molecules. Separation in HIC is based on a reversible interaction between proteins and the hydrophobic bead surface in the HIC (McCue 2009). A high ionic-strength buffer enhances the binding interaction. This feature renders HIC a perfect intermediate purification stage after ammonium sulfate precipitation or ion chromatography without buffer exchange. Proteins bind to the HIC in a high ionic-strength buffer, and the binding buffer usually contains 1 to 2 M ammonium sulfate or 3 M sodium chloride (McCue 2009). The hydrophobic ligands in the HIC interact with protein by hydrophobic interaction. Targeted proteins are concentrated after elution from HIC in the column. As mentioned earlier, the interaction between HIC and soluble proteins is reversible. Elution can be performed by decreasing the buffer salt concentration. In addition, organic solvents such as glycols and alcohols can be added to the elution buffer to reduce the eluent polarity and elute proteins (McCue 2009). Notably, organic solvents, such as alcohol, may cause protein denaturing. Unfortunately, HIC may not be a good choice for membrane protein purification. Generally, membrane proteins are amphiphilic and embedded in detergent micelles. However, detergent molecules tend to strongly bind to the HIC column. Therefore, application of HIC to membrane protein samples may either interfere with the protein binding to the column or induce protein denaturation.

Ion exchange chromatography is a valuable purification method for protein isolated from natural sources. It utilizes the surface charge of molecules for separation. Usually, ion chromatography is used as an intermediate purification stage. Based on the

charge of the resin or cellulose inside the column, ion chromatography is classified as anion exchange chromatography (positively charged stationary phase) and cation exchange chromatography (negatively charged stationary phase). The charged resin or cellulose interacts with substances with opposite charges in the solution. Protein surface charge is determined by the amino acid composition that defines the pI value and the environmental pH value. In the most popular design, the protein binds and undergoes anion exchange chromatography when the solution pH is above target protein pI value, and binding to cation exchange resign material when the solution pH is below the pI value (Jungbauer and Hahn 2009). Typically, the buffer pH should be 0.5-1 pH unit above the pI for anion exchange and 0.5-1 pH unit below the pI for cation exchange chromatography. However, some proteins may require further buffer pH adjustments. A trial and error process is needed on a case-by-case basis (Ahamed et al. 2008). After impurities pass through ion chromatography, bonded proteins can be released from resins by either increasing the elution buffer salt concentration or adjusting the buffer pH to compete with and interrupt protein-resin interaction.

Size exclusion chromatography (SEC) is also called gel filtration chromatography. As the name implies, SEC can be used for biomolecule purification based on the molecule hydrodynamic radius, which is determined by both molecule size and shape. The slurry packed in the size exclusion column presents well-defined pore sizes. After injecting the sample into SEC, small molecules are diffused and retained in pores; medium molecules are partially diffused, and the large molecules do not enter the slurry beads. In other words, smaller protein sizes are carried further along the pathway, thus eluting out later; larger proteins will take a short path and will elute out first

(Irvine 2001, Duong-Ly and Gabelli 2014a). Unlike the affinity column, SEC does not directly bind and concentrate the protein. The sample region is expanded, and sample concentration is diluted after passing through the SEC column. For maximum resolution, the suggested sample loading volume is 0.5-2% of the total column volume. SEC is typically used in the last polishing stage to provide a homogenous protein sample. It can remove remainder impurities, including the aggregated and degraded target protein. Additionally, SEC is used as a desalting and buffer exchange column (Duong-Ly and Gabelli 2014a).

The affinity chromatography (AC) mechanism is based on the reversible interaction between the target protein and a specific ligand on the column matrix. This interaction can be bio-specific (e.g., hormone ligand) or non-bio-specific (e.g., metal chelates and triazine dyes). Commonly, the AC matrix is composed of agarose or polyacrylamide, and the specific ligand is covalently attached to the matrix. In rare cases, the ligand binds to the AC matrix via noncovalent interactions. For example, the streptavidin-coated AC was used to purify biotinylated proteins (Cheah and Yamada 2017). AC is widely used for purifying antibodies. AC can be considered an essential step during protein purification. The development of AC technology dramatically simplified the purification process due to high selectivity, repeatability, and rapidity (Ayyar et al. 2012). The AC method includes three steps: (1) sample loading on the AC column, followed by interaction of target proteins and binding with ligands on the matrix; (2) the AC column matrix is flushed with wash buffer, which interrupts non-specific contaminant protein binding; and (3) target protein is eluted via elution buffer, which has a robust competing ability with the target protein or can disrupt protein-ligand interaction (Mahmoudi Gomari et al. 2020). Typically, AC is used in

the early or first purification stage to remove most of the contaminants in the sample. In some cases, AC can be used as one-step purification because of the high selectivity. Based on the types of ligands on the AC matrix, AC can be classified into four groups: 1) immunoaffinity-based AC; 2) specific ligand bio-affinity; 3) general ligand bio-affinity; 4) aptamer-based AC (Mahmoudi Gomari et al. 2020). Group 1 and 2 are selectively used for protein purification; Group 3 and 4 are mainly used in a nonselective manner. For example, to purify DNA-binding proteins based on the aptamer-based AC, aptamers are attached to the matrix. In theory, all DNA-binding proteins would bind to the matrix. Proteins that do not interact with DNA will be eluted and lead to sufficient enrichment of the potential DNA-binding protein targets (Navani, Mok, and Yingfu 2009). Immunoaffinity-based AC provides high selectivity and specificity in the antibody-antigen interaction. However, immunoaffinity-based AC is expensive, with inadequate regeneration capacity. When working with a newly discovered protein, the production of a new monoclonal antibody for the new protein target is not a desirable choice (Mahmoudi Gomari et al. 2020). Merging recombinant DNA technology with AC development has propelled protein science to new levels. Fusing an immunological tag on the target protein avoids the limitations of traditional immunoaffinity-based AC (Brizzard 2008).

Recombinant Protein with Tag-Assisted Affinity Purifications

Genetic fusion of the POI with a unique sequence tag simplifies the protein purification process and enhances biomolecule imaging and localization *in vivo*. Different types of exogenous tag sequences offer bio-specific affinity properties to target proteins. Tag fusion is not only able to provide a specific selectivity to the POI with high throughput but also boosts protein stability and solubility if needed (Arnau et al. 2006,

Walls and Loughran 2011). Furthermore, fusion tags can be removed later to meet downstream purposes such as clinical applications (Arnau et al. 2006). Selecting the right tag to assist in protein yield improvement and purity is the first and essential step in current protein science. The most popular purification tags and the features list are presented in Table 2.2.1. Table 2.2.2 lists recently introduced tags and their features (Mahmoudi Gomari et al. 2020).

The FLAG tag is an octamer epitope tagging first designed in 1984 based on the antibody-antigen interaction (Munro and Pelham 1984). A FLAG tag contains the DYKDDDDK sequence motif with high affinity to monoclonal M1, M2, and M5 antibodies. Some FLAG variants are selected via direct ELISA to improve tag selective recognition properties (Slootstra et al. 1997). The polar environment of the aromatic residues influences the strength of the antibody-antigen interaction (Sela and Arnon 1960). The negatively charged aspartic acid at the N-terminus and the hydrophilic hexapeptide (KDDDDK) at the C-terminus ensure the FLAG tag provides a maximum antigenicity. Additionally, the FLAG tag C-terminus contains a cleavage site for enterokinase to recognize and offer an additional tag removable feature (Einhauer and Jungbauer 2001). With the same elution concept as AC, bounded FLAG-tagged proteins can be eluted off the AC by a high concentration of FLAG-tagged peptides, EDTA, or low pH buffer (Table 2.2.1). Due to the difficulty of generating antibodies for some proteins, epitope tags provide the possibility for protein immunocytochemical detection and immunoaffinity-based

purification. One iconic example is fusing the FLAG epitope tag in G-protein-coupled receptors (Milligan 1999, Hunter, Grimsey, and Glass 2016).

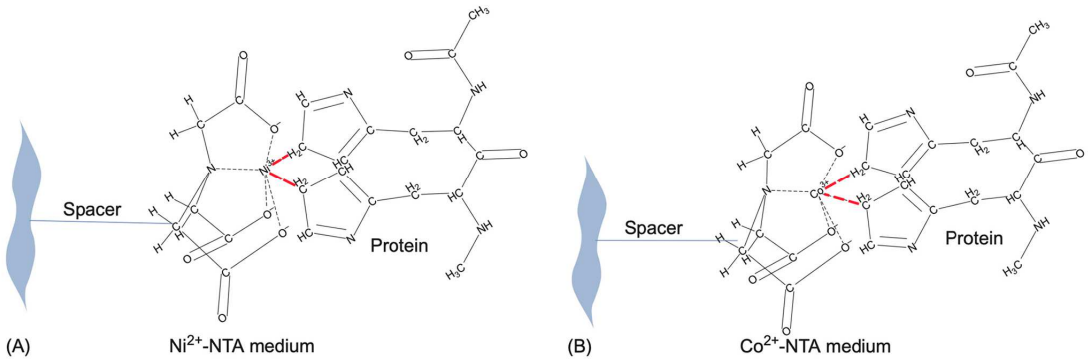

Figure 2.2.5 Molecular interactions in immobilized metal ion affinity chromatography (IMAC). (A) Nickel-charged affinity resin interaction with histidine residues during purification and (B) cobalt-charged affinity resin interaction with histidine residues during purification. The interaction between metal ion and side chain of histidines were marked with red dash lines. Figure and caption are modified from (Rai, Arya, and Kumar 2021) with permission from ELSEVIER.

Immobilized metal ion affinity chromatography (IMAC) is based on protein sequences, which contain electron donor atoms (nitrogen, sulfur, or oxygen) that interact with divalent metal ions (nickel, cobalt, copper, or zinc) to form strong coordinated covalent bonds. In IMAC, specific metal ions are immobilized on the column resin via chelator linkers (Mahmoudi Gomari et al. 2020). The His-tag system, which fuses multiple (6-10) histidine residues on either the N- or C- terminus of the POI, is the most common strategy employed in IMAC. Both divalent nickel and cobalt cations are frequently used in the His tag system. The side chain of histidine has a high affinity to the metal ions chelated on the column matrix. Two side chains from histidine coordinate to one metal ion by replacing two water molecules within the chelated structure (Figure 2.2.5). The His-tag is a small peptide sequence that does not generally interfere with native protein structure. More

than 60% of crystallography research has fused this poly-His tag to their POI (Mahmoudi Gomari et al. 2020). However, one major drawback of the His-tag system is unwanted binding by the IMAC to contaminant proteins with natural histidine-rich regions (Zhao, Li, and Liang 2013).

The glutathione S-transferase (GST) tag is an enzyme based affinity tag. GST is a cytosolic enzyme and binds to the substrate glutathione (GSH), and GSH linked resins bind recombinant proteins with the GST tag. The purified GST-tagged proteins can be eluted using an elution buffer with excessive GSH. GST was the first tag system used in protein purification and pull-down assays in 1988 (Mahmoudi Gomari et al. 2020). In addition, fusing the GST tag to the N-terminus of the POI enhances target protein solubility and expression efficacy (Malhotra 2009). However, the high expression of the GST-tagged protein may lead to improper protein folding and protein aggregation. Furthermore, there are alternative transcription initiation sites in the DNA sequence of the GST-tag. Given the different transcription initiation points, there is a high tendency to form truncated proteins (Schäfer et al. 2015, Bernier et al. 2018). Recently, one report has revealed the mutation of the GST coding sequence (299-301 nt) from "TTG" to "CTC" facilitated production of the GST-tagged full-length POI (Mahmoudi Gomari et al. 2020, Bernier et al. 2018).

Aggregating tag is another type of fusion tag. This tag promotes specific aggregation to targeted protein and modulates protein aggregation, both in *in vivo* or *in vitro* environments. Due to this unique feature, aggregating tags are considered an alternative non-chromatography (column-free) purification method (Lin et al. 2015). Based on the protein aggregation status, aggregating tags are broadly classified into three categories: tags that trigger the active protein to form aggregates *in vivo*; tags that trigger

inactive protein to form aggregates *in vivo*; tags that trigger *in vivo* soluble protein to form aggregates *in vitro* (Lin et al. 2015).

Certain aggregating tags trigger the formation of protein aggregates within the host (*in vivo* aggregation) and lead to sustained expression by reducing the concentration of the soluble form of the targeted protein in cells. It is especially beneficial to express toxic proteins in host cells because aggregated toxic proteins remain inaccessible to the host cells (Yang, Pistolozzi, and Lin 2018). For example, N^{pro} is a highly hydrophobic self-cleaving aggregating tag derived from swine fever virus nucleocapsid protein (Achmuller et al. 2007). The N^{pro} tag induces the formation of inactive protein aggregates *in vivo*, which can be applied to express toxic peptides. However, cleavage of the aggregation tag followed by a refolding process is still necessary for the downstream applications. Furthermore, the N^{pro} tag is incompatible with expressing large proteins. These two drawbacks limit the range of applications (Yang, Pistolozzi, and Lin 2018).

Elastin-like polypeptides (ELPs), Annexin B1, and repeats-in-toxin (RTX) aggregating tags can express the soluble protein *in vivo* and reversibly aggregate protein *in vitro*. The reversible aggregation mechanism in ELPs is based on temperature and salt concentration; Annexin B1 and RTX tags are based on calcium metal ion concentration. In recent reports, ELPs are popular aggregating tags for protein purification (Mahmoudi Gomari et al. 2020). After cell lysis, ELPs tagged soluble protein can turn into aggregated forms by increasing the temperature above the transition point. Then, reversible aggregated proteins can be easily separated and collected *in vitro* via centrifugation (Lin et al. 2015). Typically, ELPs are combined with a self-cleavable intein (small peptide) to a tag-free final protein product. For triggering intein self-cleavage activity, redissolved purified fusion

proteins are treated with buffer pH shift or a thiol-rich reagent (Yang, Pistolozzi, and Lin 2018). However, the most significant limitation of ELPs is the large molecular size of the aggregating tag. In addition, the aggregating tag requires a long-expression time at low temperatures (18-20°C), which is limited to the expression of non-thermosensitive proteins (Mahmoudi Gomari et al. 2020, Lin et al. 2015)

Table 2.2.1 A list of chromatography purification tags along with their characteristic features. The table is modified and used from (Mahmoudi Gomari et al. 2020) with permission from ELSEVIER.

Tag	Length and/or sequence	MW[a] (KDa)	Position	Column	Elution system	Purity	Rate of usage	Estimated cost[b]	Origin	Advantages	Disadvantages
Arginine-tag	RRRRR	0.8	C-term	Cation exchange resin sp-sephadex	NaCl linear gradient from 0 to 400 mM at alkaline pH>8.0	95%	NA[c]	€108/10 g (sigma)	Synthetic	Easy to uses	Negative effects on the tertiary structure of proteins
CBD	51	5.59	N/C term	Chitin	30–50 mM dithiothreitol, b-mercaptoethanol or cysteine	99%	NA	$94/20 ml (biolabs)	*Bacillus circulans*	High purity	Target Loses first methionine after the tag removal
CBP	26	2.96	N/C term	Calmodulin	EDTA or EGTA with 1M NaCl	NA	7.5%	€555/10 ml (sigma)	*H. sapiens*	High specificity and yield, and keeping native structure of fusion protein during mild elution conditions	Unsuitable for eukaryotic systems
FLAG-tag	DYKDDDDK	1.01	N/C term	M1 or M2 Antibodies coupled with agarose	pH 3.0 or 2–5 mM EDTA	95%	16%	€615/1 ml (sigma)	Synthetic	Fewer effects on target protein	Low capacity, which increases the costs
GST	211	26	N/C term	Glutathione	5–10 mM reduced glutathione	80%	NA	€286/5 ml (Sigma)	*Schistos majaponicum*	Increases solubility and has high specificity cleavage from fusion proteins	Decreases solubility, immunogenicity
HSV-tag	QPELAPED	0.9	C-term	mAb	Low pH buffers	>90%	NA	$299/0.1 ml	*herpes simplex virus*	High price	Low yield
HA-tag	YPYDVPDYA	1.0	N/C term or internal	12-CA5 antibody immobilized on agarose	Low pH or HA peptide	>80%	7.5%	€639/1 ml (sigma)	*Human influenza virus*	Useful for the detection and purification	Interference with protein folding
HaloTag7	312	33	N/C term	Chloroalkane Coupled with Agarose, HaloLinkTm resin	HaloLink buffer	95%	NA	NA	*Rhodococcus*	Increases expression, solubility and high purity	Negative effect on protein function
His-tag	HHHHHH	0.84	N/C term, or internal	Ni^{2+}-NTA*, Co^{2+}-CMA, Zn^{2+}, Cu^{2+}	Imidazole 20–250 mM or low pH	>90%	23%	$305/25 ml (qiagen)	Synthetic	Low interference with the protein function, and easy access	Interference with histidine containing proteins, dimerization of target protein
MBP	396	40	N/C term	Amylose	10 mM Maltose	75%	NA	$192/15 ml (biolabs)	*E. coli*	Increases solubility and does not leave extra amino acids	High price, low efficiency, immunogenicity
Myc	CEQKLISEEDL	1.3	N/C term or internal	9E10 antibody coupled with agarose	Low pH	>90%	*5%*	*€722/1 ml (sigma)*	*H. sapiens*	High purity	The resin is expensive
NusA	495	55	N/C term	NA	NA	NA	NA	NA	*E. coli*	Increases solubility	Large size and low yield
Protein-A	449	42	N/C term	IgG Sepharose	0.1 M glycine-HCl, pH 3.0	NA	NA	NA	*S. aureus.*	Improving protein solubility	Large size and Interference with protein folding
Protein-G	465	58	N/C term	IgG Sepharose	0.1 M glycine-HCl, pH 3.0	NA	NA	NA	*Streptococcus*	High purity	Large size and high price
Streptavidin binding peptide	38	4.31	C-term	Streptavidin	Biotin containing buffer	>90%	NA	NA	*Synthetic*	High stability	Interference with protein folding
Strep-II	WSHPQFEK	1.06	N/C term	Strep-Tactin	2.5 mM desthiobiotin	99%	6.5%	€592/10 ml (sigma)	*Bacteria*	Versatility	The resin is expensive
SUMO	100	18	N-term	Ni-NTA, NI-IMAC	NA	>90%	NA	NA	*H. sapiens*	Increases the solubility	Not useful to purification
Streptavidin	123	12.96	N/C term	Biotin	Biotin containing buffer	NA	NA	$365/5 mg	*Streptomyces avidinii*	Good expression	Disruption of protein folding in elution step
S-tag	KETAAAKFERQHMDS	1.75	N/C/ term or internal	S-fragment of RNase A	Acidic buffer	>90%	NA	NA	*RNase A*	High stability with good expression	Limited use
T7-tag	MASMTGGQQMG	1.1	Internal or N-term	mAb	low pH Acidic buffers (like acetic acid, pH 2.9)	>95%	NA	$72/1mg	Bacteriophage T7	Positive effect on protein expression	Acidic condition for elution may impact on protein
TrxA	109	12	N/C term	Phenylarsinine oxide	5-1000 mM β-mercaptoethanol	NA	NA	NA	*E. coli*	Suitable for crystallization, increases solubility	Not useful to purification
V5-tag	GKPIPNPLLGLDST	1.42	C-term	mAb	low pH elution	>95%	NA	€622/1 ml (sigma)	*Simion virus 5*	Small size	Low yield

Table 2.2.2 A list of newly introduced tags and their features. The table is modified and used from (Mahmoudi Gomari et al. 2020) with permission from ELSEVIER.

Tag	Length and/or sequence	MW[a] (KDa)	Resin	Purity	Origin	Advantage
AGIA-tag	AGIA	0.34	Ra48	NA[b]	Dopamine receptor D1	High specificity
ALFA-tag	SRLEEELRRRLTE	1.69	NbALFA	NA	Synthetic	Good expression
C-tag	APEE	0.45	NbSyn2	>70%	Synthetic	High yield
Car9	DSARGFKKPGKR	1.35	Silica Gel Columns	80%	TrxA	Affordability and good performance
C9R	RRRGRRRGRRRG	1.59	Ni2+-NTA	>90	Synthetic	Good expression and solubility enhancer
CL7	128 residues	14.7	IM7	97-100%	Colicin E7 DNase	High specificity and high efficiency
CRDSAT	156 residues	17.36	Lactose-Sepharose column	>95%	Galectin	low-cost and high stability
CP5	GQHVT	0.55	Ra62	71-89%	Dopamine receptor D1	High purity and activity
CspB50	50 residues	5.64	Acidic condition (pH 6.8)	>70%	Cell surface protein B	Cost benefit and suitable for unstable protein
CusF	110 residues	12.25	Cu2+ containing column	>95%	E. coli	Solubility enhancer and no interfering with protein folding
G196-epitope	DLVPR	0.6	G196 mAb	>70%	GST	High yield
GFP	238 residues	26.9	A4C7	94%	Aequorea victoria	Low toxicity and tolerance to pH changes
Glu6	EEEEEE	0.8	Bare iron oxide nanoparticles in pH 7	70%	Synthetic	High stability
HB-tag	MASKAQKAQAKQWKQAQKAQKAQAKQAKQAKQW	3.75	Heparin-Sepharose column	100%	Heparin	Affordability and reasonable speed
HHHNSWD	7	0.93	Ni2+ resin	NA	Synthetic	Low interference with protein function
LecB	108	11.28	Mannose agarose	NA	Pseudomonas aeruginosa	High yield
LSLa	150	17.2	Agarose-4B	NA	Mushroom Laetiporus	Good stability
MAP-tag	GDGMVPPGIEDK	1.09	Mab-1	NA	Podoplanin	High affinity for mab-1
mbT4L	121 residues	13.6	Ni2+-NTA	Higher than His-tag	T4 lysozyme	strong affinity for immobilized metal ions
MhPA14	193	20	Carbohydrates coupled resins	NA	MhLap	Low steric clashes and versatility
PagP	164 residues	19	Denaturing condition	>95% (with His-tag)	E. coli	High efficiency
Phosphorylatable tag	EDSDSSSEDE	1.09	Fe3+ containing column	NA	Casein kinase II	Good performance and straightforwardly
RAP	DMVNPGLEDRIE	1.4	PMab-2	NA	Podoplanin	Multipurpose use
Ribosomal protein L2 (E. coli)	71 residues	7.9	Diatoms silica	>90%	E. coli	High efficiency
SB7	RQSSRGR	0.84	Silica or Shirasu particles	90%	CotB1	Cost-benefit
S1v1	(AEAEAHAH)2	1.65	Ni2+-NTA	>90	Zuotin protein	Positive effects on protein expression and stability
ssTorA	MNNNDLFQASRRRFLAQL GGLTVAGMLGPSLLTPRR	3.98	Low speed centrifugation	NA	E. coli	Suitable difficult-to-express and toxic proteins
T5	QRVRELAV	0.97	5G10	NA	Bacteriophage	Fewer effects on target protein
1D4	TETSQVAPA	0.91	Rho1D4	NA	Bovine rhodopsin	Versatility
2B8	RDPLPAFPP	1.01	Anti-2B8	NA	Deinococcus radiodurans	High specificity
Zacid	58	6.6	IgG coupled column	NA	Protein-A	High yield

2.3 Materials and Methods

Construction of pFastBac1-pGC-A Expression Vector

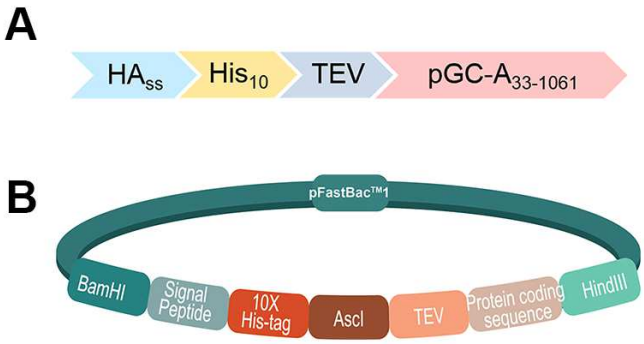

Figure 2.3.1 Schematic of the pGC-A donor vector and protein diagram. (A) Protein diagram. (B) The intact pGC-A donor vector is 7877 bp. Abbreviations: HA_{ss}, hemagglutinin signal sequence; TEV, tobacco etch virus protease cleavage site.

The Invitrogen Bac-to-Bac BEVS protocol was followed to express intact human pGC-A (NCBI Reference Sequence NP_000897.3). The plasmids construct pFastBac1-pGC-A was designed to allow pGC-A (amino acids 33-1061) expression in Sf9 insect cells (Figure 2.3.1 B). The gene encoding the pGC-A protein (GenBank accession number NP_000897.3) was optimized by GenScript for expression in insect cells using GenScript's OptimumGene algorithm. The pFastBac1-pGC-A donor vector was sequence verified and shipped from GenScript. The designed protein diagram is shown in Figure 2.3.1A, and the complete protein and DNA sequences are represented in Figure 2.3.2A and 2.3.2B, respectively. The native signal peptide sequence (amino acids 1-32) was replaced with a cleavable N-terminal influenza virus hemagglutinin signal peptide (Guan, Kobilka, and Kobilka 1992). A tobacco etch virus (TEV) protease cleavable His10-tag was added between the signal peptide and the GOI. Cloning of pGC-A constructs into the *Bam*HI and *Hin*dIII sites of the insect cell donor vector pFastBac1 was performed by GenScript. The designed pFastBac1 construct is shown in Figure 2.3.1B.

Figure 2.3.2 Complete sequences for the expression-optimized sequences encoding intact pGC-A protein. (A) Protein sequence. (B) Plasmid DNA sequence.

Insect Cell Culture

Spodoptera frugiperda (Sf9) cells were purchased (Thermo Fisher, Cat. A35243). Cells were seeded at a density of 0.6×10^6 viable cells/mL. Insect cells were grown in suspension flasks with Sf900-II SFM medium (Gibco, Cat. 12659017). Seeded cells were incubated at 27°C with shaking at 130 rpm on an orbital shaker and passaged when the cell density reached 2×10^6 viable cells/mL. Cell counting was performed using the

Thermo Countess II cell counter. During each passage, 10 μL of the suspension cell culture was mixed with 10 μL Trypan Blue stain (Thermo Fisher, Cat. C10288) to measure the cell density and viability. The cells were sub-cultured at the same seeding density until reaching the 30^{th} passage. Then, another vial of frozen cell stock was used for a new passage.

A frozen Sf9 cell line stock was generated and preserved in a liquid nitrogen tank storage box. For the best cell recovery, cell cultures in the 10^{th} to 15^{th} passage (lower passage) were used for preparing the frozen stock. In brief, 50 mL of the Sf9 cell culture (2.5 to 3.5 × 10^6 cells/mL; viability above 95%) was pelleted down and resuspended with fresh Sf900-II SFM medium (Gibco, Cat. 12659017) to achieve a 2 × 10^7 cells/mL concentration. An equal volume of the frozen buffer, containing 20% dimethyl sulfoxide (DMSO), 20% fetal bovine serum (FBS), and 60% fresh Sf900-II SFM medium, was added to the resuspended cell culture to achieve the final 1 ×10^7 cells/mL concentration. Diluted cell cultures were aliquoted (1.1 mL) into cryotubes (Wheaton, Cat. W985872100). The cryotubes were loaded to the Mr. Frosty™ Freezing Container (Thermo Scientific, Cat. 5100-0001) and chilled in a -80°C freezer overnight. The following day, frozen tubes were transferred to the liquid nitrogen storage box.

Recombinant Bacmid Generation and Selection

pFastBac1-pGC-A Plasmid Transformation and Purification. The DH5α competent *E. coli* cell strain was used to amplify and store the pFastBac1-pGC-A plasmid purchased from GenScript. Briefly, 100 μL of DH5α culture was mixed with 100 ng of pFastBac1 plasmids with a gentle mix and then placed on an ice bath for 30 min. The chilled cell-plasmid mixture was heat-shocked for 50 s at 42°C and immediately placed

back on the ice bath for 2 min. Then, 900 μL of filtered SOC complete buffer (2 % tryptone, 0.5 % yeast extract, 10 mM NaCl, 2.5 mM KCl, 5 mM $MgCl_2$, 5 mM $MgSO_4$, 0.1% glucose) was added to the cell-plasmid mixture, and the mixture was incubated at 37°C for 1 h with shaking at 225 rpm. Next, 100 μL of transformed cell culture was spread on the LB medium plate with 50 μg/mL carbenicillin (substituted for 100 μg/mL ampicillin), and the plate was incubated at 37°C overnight.

A single well-isolated white colony was used to inoculate 4 mL of LB medium with 50 μg/mL of carbenicillin. The inoculated culture was incubated at 37°C overnight with shaking at 225 rpm on an orbital shaker. pFastBac1-pGC-A plasmids were purified with a standard miniprep kit QIAprep Spin Miniprep kit (QIAGEN, Cat. 27106). The overnight cell culture was pelleted and resuspended with 250 μL Buffer P1 (100 μg/mL RNase A added) in a 1.5 mL micro-centrifuge tube. Then, 250 μL of Buffer P2 was added, and the tube was gently inverted 6 times. Next, 350 μL Buffer N3 was added, and the tube was thoroughly inverted 6 times. The tube was centrifuged for 10 min at 13K g-force. In the next step, 800 μL of the supernatant was transferred to the spin column and centrifuged for 1 min at 13K g-force. After discarding the flowthrough, 0.5 mL of Buffer PB and 0.75 mL of Buffer PE were separately added, with the same procedure performed as previously mentioned. Finally, 50 μL of Buffer EB was added, and the tube was centrifuged for 1 min at full speed to elute purified plasmids. All of the listed buffer solutions were directly from the commercial QIAprep Spin Miniprep Kits (QIAGEN, Cat. 27106).

Glycerol Stock for DH5α with pFastBac1-pGC-A Plasmid To generate the glycerol cell stock, one well-isolated single DH5α colony, which contained the pFastBac1-pGC-A plasmid, was inoculated into 4 mL LB medium with 50 μg/mL carbenicillin at 37°C overnight with shaking at 225 rpm on an orbital shaker. Next, 750 μL of the overnight culture was mixed with 750 μL frozen media (autoclaved LB medium with 30% glycerol) and stored in a -80°C freezer.

Recombinant Bacmid Generation To generate the pGC-A recombinant bacmids, the Invitrogen Bac-to-Bac BEVS protocol was followed (Figure 2.2.4). Purified 50 μL pFastBac1-pGC-A plasmids were transformed into *E. coli* DH10Bac competent cells (Invitrogen, Cat. 10361-012). For each transformation, one vial of commercial DH10Bac competent cells was thawed on ice and mixed with 1 ng of purified pFastBac1-pGC-A plasmids with a gentle mix. The mixed cells were cooled on an ice bath for 30 min. The chilled cell-plasmid mixture was heat-shocked for 45 s at 42°C and immediately placed back on the ice bath for 2 min. Then, 900 μL of filtered SOC complete buffer (room temperature) was added to the cell-plasmid mixture, and the mixture was incubated at 37°C for 4 h with shaking at 225 rpm. Then, 100 μL transformed *E. coli* DH10Bac cells were grown on the X-gal LB plate (containing 50 μg/mL kanamycin, 7 μg/mL gentamicin, 10 μg/mL tetracycline, 100 μg/mL Bluo-gal, and 40 μg/mL IPTG) for 48 h at 37°C. The single and well-isolated white colonies were picked for whole-cell colony Polymerase Chain Reaction (PCR).

Colony PCR Well-isolated white colonies were picked with a 20 μL pipet tip and transferred to a PCR tube with 20 μL molecular biology grade water (Fisher Scientific

Cat. MT46000CV). Well-mixed samples were heated for 10 min at 90°C and centrifuged at 13 K g-force for 5 min to collect the supernatant. Then, 10 µL of colony supernatant solution was mixed with 0.3 µL 10 mM each Deoxynucleotide (dNTP) Solution Mix (New England Biolabs Cat. N0447S), 0.6 µL 100 µM forward primer BACM13F (5'-CCCAGTCACGACGTTGTAAAACG-3'), 0.6 µL 100 µM reverse primer BACM13R (5'-AGCGGATAACAATTTCACACAGG-3'), 6 µL iProof High-Fidelity polymerase (Bio-rad, Cat. 1725301), and 23.28 µL molecular grade water. The initial denaturation time was 1 min at 98°C and followed with repeated cycle: 10 s denaturation time at 98°C; 45 s annealing time at 50°C; 6 min extension time at 72°C. The total PCR had 35 runs.

Recombinant Baculovirus Generation and Amplification

The PCR confirmed whole-cell colonies were inoculated in 4 mL LB medium (containing 50 µg/mL of kanamycin,7 µg/mL of gentamicin, and 10 µg/mL of tetracycline) and cultured overnight at 37°C with shaking at 225 rpm on an orbital shaker. The recombinant bacmid DNA was purified with the Invitrogen PureLink HiPure miniprep kit (Invitrogen, Cat. K210002). In summary, overnight cell culture was pelleted and resuspended with 400 µL resuspension buffer (50 mM Tirs-HCl, pH 8.0 and 10 mM EDTA) with 20mg/mL RNase A in a 1.5 mL micro-centrifuge tube. Then, 400 µL of lysis buffer (0.2 M NaOH and 1% (w/v) SDS) was added, and the tube was gently inverted 5 times. Next, 400 µL precipitation buffer (3.1 M potassium acetate, pH 5.5) was added, and the tube was thoroughly inverted until the mixture was homogenous. The tube was centrifuged for 10 min at 17 K g-force. Finally, 1 mL of the supernatant was transferred to an equilibrated column [equilibrated with 2 mL of equilibration buffer (0.1 M sodium acetate, pH 5.0; 0.6 M NaCl, and 0.15% (v/v) Triton® X-100)]. The column

was washed twice with 2.5 mL wash buffer (0.1 M sodium acetate, pH 5.0 and 825mM NaCl), and recombinant bacmid was eluted with 0.9 mL elution buffer (100 mM Tris-HCl, pH 8.5 and 1.25 M NaCl). Then, 0.63 mL isopropanol was added to the eluted solution. The inverted mixed tube was chilled on ice for 10 min. The sample was centrifuged at 17 K g-force for 20 min at 4°C. After discarding the supernatant, 1 mL of 70% ethanol was added to the tube. The inverted tube was further centrifuged at 17 K g-force for 5 min at 4°C. After discarding the supernatant, the bacmid pellet in the tube was air-dried for 10 min. Finally, 40 µL of TE buffer (10 mM Tris-HCl, pH 8.0 and 0.1 mM EDTA) was added to gently resuspend the purified bacmid. All of the listed buffer solutions were directly from the commercial Invitrogen PureLink HiPure miniprep kit (Invitrogen, Cat. K210002).

The purified 40 µL bacmid DNA (50-100 ng/µL) was mixed with 100 µL fresh Sf900-II SFM medium (Gibco, Cat. 12659017) via finger flick. Next, 8 µL Cellfectin II reagent (Thermo Scientific, Cat. 10362100) was mixed with insect cell fresh cell media via vortexing. Then, 108 µL transfecting reagent was added to the bacmid solution with finger tapping, followed by incubation for 30 min. The ratio of the transfection mixture is 1 µg DNA/8 µL Cellfectin II reagent. The transfected mixture was added to 1 mL (1×10^6 cell/mL) insect cell culture in a 24-well deep plate (Thomas Scientific, Cat. 504361) and incubated for 7 days at 27°C with shaking at 130 rpm to generate the Passage 1 virus stock. After 7 days of virus generation, Passage 1 culture was centrifuged for 5 min at 1K g-force. The supernatant, which contained the P1 recombinant virus, was stored and used to generate the Passage 2 virus stock. Then, 200 µL of P1 virus supernatant was added to 4 mL (0.5×10^6 cell/mL) insect cell culture in a 24-well plate and incubated for 3 days at

27°C with shaking at 130 rpm to generate the Passage 2 virus stock. After 3 days of virus generation, the Passage 2 culture was centrifuged for 5 min at 1 K g-force. The supernatant, which contained the P2 virus, was stored and used to generate the Passage 3 virus stock. Subsequently, 500 µL of P2 virus supernatant was added to 30 mL (3×10^6 cell/ml) insect cell culture and incubated for 4 days at 27°C with shaking at 130 rpm to generate the final Passage 3 virus stock. All cell pellets from three virus passages were saved for the Western blot protein expression test. The Passage 3 viral concentration was measured before storing in liquid nitrogen. The virus titer was determined based on viable transfected cell diameter size change and statistically analyzed and calculated by the method of moment (Janakiraman et al. 2006). The precise virus titer protocol was previously published (Janakiraman, Forrest, and Seshagiri 2006), and the exact steps were followed during the titer determination experiment. Briefly, 2 mL of different serial dilutions of virus stock and one mock media were prepared. Then, each virus dilution (2 mL) was added to 20 mL of 2×10^6 cells/mL Sf9 cells with 95% viability and incubated at 27°C with shaking 130 rpm on an orbital shaker for 24 h. Subsequently, 1 mL of each transected cell culture was used to measure the cell diameter via the Vi-CELL counter in duplicate mode. Then, the measured cell diameter data were analyzed, statically calculating the virus titer using the method of moment (Janakiraman, Forrest, and Seshagiri 2006).

Verification of pGC-A Expression

In order to verify the expression of pGC-A, previous serial virus amplification-generated cell pellets were collected and used for Western blotting. For gel analyses, 5 µL of cell pellet suspensions from different stages were mixed with 35 µL of 4X SDS-

PAGE loading buffer (Bio-Rad, Cat. 1610747) with β-mercaptoethanol. Samples were boiled at 90°C for 10 min and analyzed on 4-12% acrylamide Bis-Tris gels (Invitrogen, Cat. NP0321BOX). For western blots using nitrocellulose membranes, the primary antibody was 1:2000 anti-His5-tag monoclonal (Qiagen, Cat.34660), and the secondary antibody was 1:2000 HRP-conjugated anti-mouse-IgG (H+L) polyclonal (Invitrogen, Cat. 626520).

Optimization of Virus MOI and Cell Harvest Days

During the screening test, a multiplicity of infection (MOI) of 0, 3, 5, 10 was used to transfect 4 mL of Sf9 insect cell culture. MOI stands for the ratio of the number of the virus to the number of cells in cell culture. A 4 mL volume of culture was collected after post-transfection at 48, 72, and 96 h. A 100 μL volume of cell culture was pelleted and lysed with 50 μL of 1X SDS-sample loading buffer (Bio-Rad, Cat. 1610747). All samples were heated up at 96°C for 10 min before loading on SDS-PAGE. Samples were analyzed by SDS-PAGE on 4□12% acrylamide Bis-Tris gels (Invitrogen, Cat. NW00127BOX) in MOPS-SDS running buffer (Invitrogen, Cat. NP0001). For western blotting, the primary antibody, anti-pGC-A, was purchase from R&D Systems (Cat. MAB4860); the secondary antibody, HRP-conjugated anti-mouse IgG, was procured from Jackson ImmunoResearch (Cat. 515-035-062).

Large Scale pGC-A Purification

Insect cells were grown and scaled up in commercial Sf-900 III SFM serum-free media (Gibco, Cat. 12659017) in a 27°C incubator. A culture with 1.5 L of 2×10^6 Sf9 insect cells/mL was transfected with generated pGC-A virus with a MOI of 10. Seventy-two hours post-infection, the cells were harvested by centrifugation at 2 K g-force speed

for 10 mins at 4°C. Harvested cells were suspended in ice-chilled lysis buffer (50 mM HEPES-NaOH, pH 7.5, 10 mM $MgCl_2$, 20 mM KCl, 0.01 mg/mL deoxyribonuclease I [DNaseI, Sigma-Aldrich, Cat. 11284932001], and 10% glycerol) and lysed on an ice bath via sonication (1 s pulse on and 2 s pulse off for a total of 3.5 min at 40% power; sonicator model: Branson Sonifier sfx550). The cellular membrane fraction was collected using a 45-Ti rotor at 45 K rpm using a Beckman Coulter ultracentrifuge for 30 min at 4°C. Then, a 15 mL glass KIMBLE tissue homogenizer was used to resuspend the cellular membrane fraction with hypertonic buffer (50 mM HEPES-NAOH [pH 7.5], 10 mM $MgCl_2$, 20 mM KCl, 5 mM DTT, 1 M NaCl, and 10% glycerol) on ice; subsequently, the solubilization buffer (50 mM HEPES-NAOH [pH 7.5], 150 mM NaCl, 5 mM DTT, and 10% glycerol) was used to wash the membrane fraction with a tissue homogenizer on ice. The washed membrane fraction was pelleted down in a 70-Ti at 45 K rpm using a Beckman Coulter ultracentrifuge for 30 min at 4°C. The washed and pelleted membrane fraction was solubilized with 0.75% *n*-dodecyl-β-D-maltopyranoside (DDM; Glycon Biochemicals, Cat. D97002-C) and 0.15% cholesterol hemisuccinate (CHS; Sigma-Aldrich, Cat. C6512-5G) for 1 h on ice. The solubilized membrane solution was centrifuged in 70-Ti at 45K rpm using a Beckman Coulter ultracentrifuge for 30 min at 4°C. The solubilized membrane supernatant was incubated with pre-equilibrated 300-400 μL Ni^{2+}-charged immobilized metal AC resin (Bio-Rad, Cat. 7800801) for 30 mins at 4°C. The resin was washed with 30 column volumes (CVs) of buffer A (50 mM HEPES-NAOH [pH 7.5], 500 mM NaCl, 5 mM DTT, 50 mM imidazole-HCl, 10% glycerol, 0.05% DDM, and 0.01% CHS), 15 CVs of buffer B (50 mM HEPES-NAOH [pH 7.5], 500 mM NaCl, 5 mM DTT, 90 mM imidazole-HCl, 10% glycerol, 0.05%

DDM, and 0.01% CHS) and eluted with 10 CVs of buffer C (50 mM HEPES-NAOH [pH 7.5], 500 mM NaCl, 5 mM DTT, 500 mM imidazole-HCl, 10% glycerol, 0.05% DDM, and 0.01% CHS). The eluted sample was collected and concentrated in a 15 mL 100 kDa cutoff Amicon Ultracel-100 regenerated cellulose membrane concentrator (EMD Millipore, Cat. UFC910024) at 3 K g-force at 4°C. The concentrated elution fraction (300-500 µL volume) was injected into pre-equilibrated (50 mM HEPES-NAOH [pH 7.5], 150 mM NaCl, 5 mM DTT, 10% glycerol, 0.05% DDM, and 0.01% CHS) column Superose6 10/300 GL (GE, Cat. 29091596). The ÄKTA pure chromatography system used with a Superose6 column at a flow rate of 0.45 mL/min for a total of 1 CV (24 mL) run at 4°C. Both peak fractions were collected and concentrated to 1-2 mg/mL (depending on the crystallization condition) for crystallization.

2.4 Results

Isolating Recombinant Bacmid Via Blue/ White Colony Screen

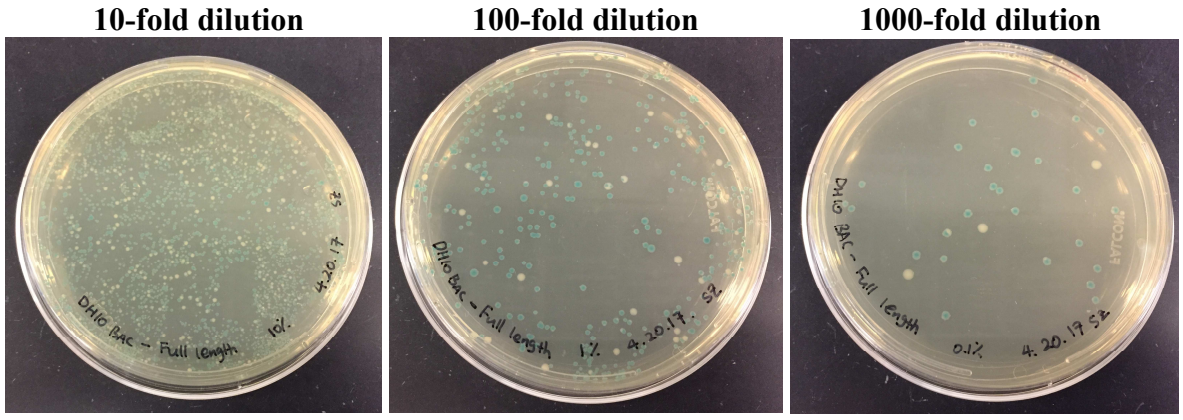

Figure 2.4.1 Blue-white colony screening. Transformed DH10Bac™ *E. coli* solution was serially diluted and sprayed on selected plates for screening for colonies with generated recombinant Bacmid. Plates were incubated for 72 h at 37°C.

Classic blue-white colony screening was performed to confirm if transformed pFastBac1-pGC-A successfully transposed the gene for the pGC-A receptor into the bacmid. After transforming the pFastBac1-pGC-A plasmids into the DH10Bac™ *E. coli* culture, 100 µL of transformed culture was applied with three serial dilutions. Diluted cultures were spread on the blue/white colony selection plate with antibiotics and IPTG. During the screening, pFastBac1-pGC-A successfully transformed into the DH10Bac™ *E. coli* as indicted by a white color on the X-gal plate (Figure 2.4.1). The GOI was integrated into the bacmid DNA by the site-specific transposition of the Tn7 transposon. Then, integrated of the pGC-A gene disrupted the lacZα gene in the bacmid. Without expressing the lacZα gene, the transformed colony was unable to produce active β-galactosidase and was white in color. Figure 2.4.1 shows that white colonies appeared in selected plates, indicating that recombinant bacmid was successfully generated after transforming pFastBac1-pGC-A plasmids into DH10Bac™ *E. coli*. In all serially diluted plates, the ratio of white colonies was low. These results indicated that the spontaneous transposition frequency was not high in DH10Bac™ *E. coli*.

PCR Verification of Recombinant Bacmid

To confirm if recombinant bacmid was truly formed in the picked white colonies, colony PCR was applied. Before PCR, one randomly chosen white colony was inoculated in LB media with kanamycin, gentamicin, and tetracycline antibiotics. Then, the Invitrogen PureLink HiPure miniprep Kit was used to purify the amplified bacmids. Purified amplified bacmids and isolated colonies were directly verified by PCR using iProof High-Fidelity DNA Polymerase and the primers BACM13F (CCCAGTCACGACGTTGTAAAACG) and BACM13R

(AGCGGATAACAATTTCACACAGG). The expected size of the bacmid containing the intact pGC-A gene is 5495 bp. The PCR analysis (Figure 2.4.2) showed DNA bands near 6000 bp size in both picked colonies and purified bacmids samples. The PCR results confirmed that the GOI was present in both selected white colonies (Figure 2.4.2, colonies #2-3) and purified bacmids derived from isolated colonies (Figure 2.4.2, P). Empty bacmid yields 300 bp, and bacmid containing pFastBac1 empty vector is 2300 bp (Figure 2.4.2, C). However, picked colonies (Figure 2.4.2, #1 and 4) had no signal on the DNA gel, probably due to failure during colony PCR steps.

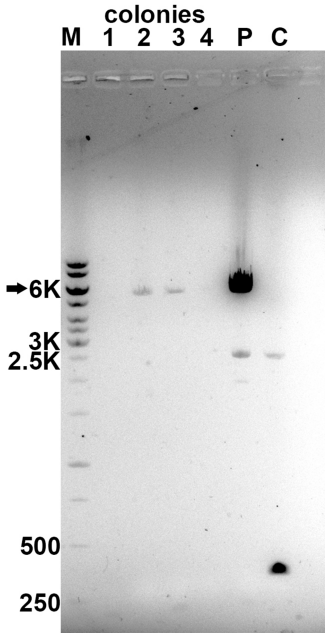

Figure 2.4.2 PCR verification of recombinant bacmid production in four isolated colonies and one amplified purified bacmid derived from the previously selected white colony. The expected size of the bacmid containing the intact pGC-A is 5495 bp. Empty bacmid yields 300 bp. Bacmid containing pFastBac1 empty vector is 2300bp. Abbreviations: M, molecular weight marker; colonies 1 to 4, individual white colonies; P, purified bacmid DNA; C, blue colony control.

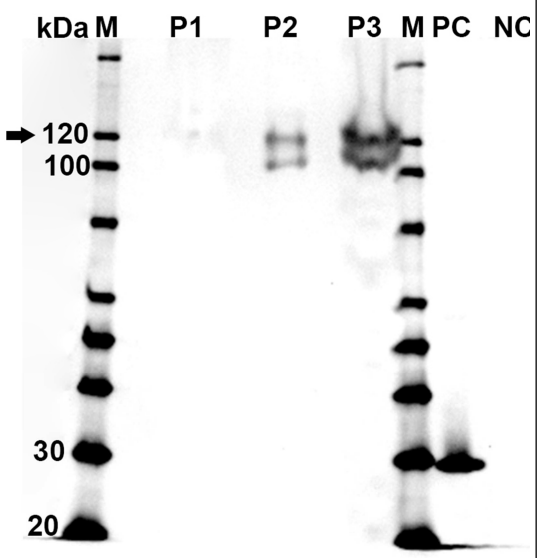

Figure 2.4.3 Western blot (anti-His) of virus passages expressing pGC-A proteins. The intact pGC-A is approximately 120 kDa. Abbreviations: M, molecular weight marker; P1P, P2P, P3P, intact insect cell pellets in virus passages 1, 2, and 3, respectively; PC, positive control (purified $GFPHis_6$, 27.7 kDa); NC, negative control (Sf9 cell pellet/supernatant without virus transfection).

Verification of pGC-A Expression

The following two steps were performed to evaluate if recombinant baculovirus was successfully generated after insect cell transfection with verified bacmids. First, the insect cell pellets (P1, P2, and P3) from three successive recombinant baculovirus virus amplifications were collected. Secondly, western blotting was performed to verify that the targeted protein was expressed in the collected Sf9 cell pellets (Figure 2.4.3). The results indicated the presence of intact pGC-A expression during virus amplification in

the insect cells. In the later collected cell pellets (P2 and P3), intact pGC-A was strongly expressed when compared with the negative control. In this experiment, the un-transfected Sf9 cells were used as the negative control and purified GFP-His_6 protein was used as the positive control. The western blot results verified that pGC-A was successfully expressed by generated recombinant baculovirus.

Optimization of Protein Expression

After successfully generating the recombinant baculovirus for intact human pGC-A expression, the next step was to optimize protein expression. For optimizing the pGC-A expression levels, Sf9 insect cells were tested in small-scale expressions as follows. Different MOI of 0, 3, 5, 10 were used to each transfect 4 mL of insect cell culture. Transfected cells were collected at 48, 72, and 96 h. Western blotting of total cell proteins revealed that the best expression parameter combination was MOI of 10 and harvest at 72 h (Figure 2.4.4, well 8). The higher the virus concentration (MOI 10) used, the better protein expression level 72 h post-transfection (Figure 2.4.4, well 8) was observed. The MOI of 3 and 5 showed similar protein expression levels within the same post-transfection time (Figure 2.4.4, compare wells 2 with 3, 6 with 7, and 10 with 11), with slightly larger yields at a MOI of 10 for transfection at 72 h and 96 h (lanes 8 and 12, respectively), but not for 48 h (lane 4). The pGC-A expression level generally decreased as the post-transfection hours increased (Figure 2.4.4, compare lanes 2-4 with 6-8 and with 10-12). This is probably due to the self-lysis of transfected cells during virus transfection. The longer the viruses were transfected in cells, the higher the likelihood of cell lysis. The different sized bands between 150kDa and 100kDa may be caused by inconsistent glycosylation levels and protease activity. These results indicate that pGC-A

was successfully expressed in Sf9 cells at all MOI's and from 48 to 96 h post-transfection.

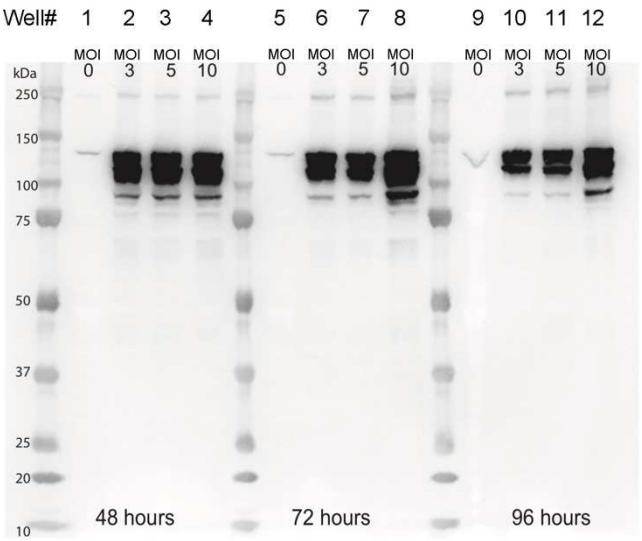

Figure 2.4.4 Optimization of pGC-A expression in Sf9 cells. Western blotting (anti-pGC-A) of total cell protein over different virus multiplicities of infection (MOI) and transfection times. MOI of 0 indicates no virus transfection. Intact pGC-A is approximately 120 kDa.

Large Scale pGC-A Purification

In order to purify intact pGC-A, a culture with 1.5 L of 2×10^6 cells/mL of transfected Sf9 insect cells was setup and cells were harvested 72 h post-infection. Harvested cells were lysed for cell plasma membrane isolation. The isolated membrane was solubilized with detergent to solubilize the membrane proteins in form of protein-detergent micelles. Solubilized intact pGC-A was purified by nickel-affinity chromatography (Ni-AC) and size exclusion chromatography using a Superose6 column. After membrane solubilization, samples from each step of the purification procedure were collected and analyzed by gel electrophoresis. Western immunoblotting results are

presented in Figure 2.4.5A, and Coomassie-stained SDS-PAGE results are presented in Figure 2.4.5B&C.

The western blotting (anti-pGC-A) results revealed that large quantities of intact pGC-A were still present in the membrane pellets (Figure 2.4.5 A, well# 2). The detergent-solubilized membrane supernatant underwent gravity nickel affinity chromatography. No pGC-A was visible in the column flowthrough (Figure 2.4.5, lanes 3 and 9). These results indicated that solubilized pGC-A was well-captured in the AC. After washing the AC with purification buffers containing 50 and 100 mM imidazole (Figure 2. 2.4.5, lane 4, 5, 10 and 11), pGC-A was eluted with the purification buffer containing 500 mM imidazole (Figure 2.4.5, lanes 6 and 12). The Coomassie blue-stained results showed that the majority of contaminant proteins were washed off with the purification buffer containing 50 mM imidazole (Figure 2. 2.4.5, lane 10). Although some pGC-A was present in the 100 mM imidazole wash (Figure 2.4.5, lane 5), this wash also contained many other contaminants (Figure 2.4.5, lane 11). Sample eluted from the buffer containing 500 mM imidazole showed a single protein band in the Coomassie blue-stained gel. The western blotting results indicated that the single band represented the intact pGC-A protein (Figure 2.4.5, lane 6 &12). This purified pGC-A was concentrated and further purified using the Superose6 gel filtration column. The peak fractions from Superose6 columns were combined and are shown in Figure 2.4.5, lane 14. Together, the intact pGC-A was successfully purified.

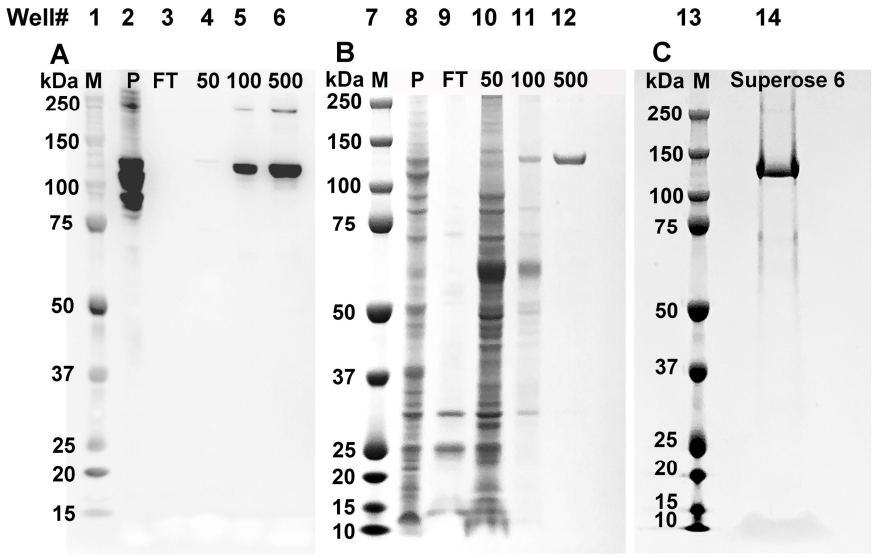

Figure 2.4.5 Purification of intact pGC-A via affinity and gel filtration chromatography. (**A**) Western blot (anti-pGC-A) of samples from cell lysis after the affinity chromatography. (**B**) Coomassie blue staining of cell lysis after the affinity chromatography. (**C**) Coomassie blue stain of the eluted fraction from Superose6 column used for protein crystallization. M, molecular weight marker; P, membrane pellet from ultracentrifuge after solubilization with DDM with CHS; FT, flowthrough from the affinity column; 50, 100, 500, imidazole (mM) at elution.

2.5 Discussion

In the studies presented in this chapter, I have established the intact pGC-A expression platform based on the Bac-to-Bac™ system and developed the large-scale protein purification pipeline for downstream protein structural and functional research. To achieve a high protein expression levels, I replaced the predicted signal peptide (amino acid residues 1-32) with the highly expressing hemagglutinin signal peptide (MKTIIALSYIFCLVFA). The hemagglutinin signal peptide was first reported as a cleavable sequence to enhance translocation of target membrane receptor into endoplasmic reticulum membrane and facilitate functional receptor expression (Guan, Kobilka, and Kobilka 1992). Substitution of native signal peptides with that from hemagglutinin was often used for GPCR studies and expression of other membrane

receptors (Guan, Kobilka, and Kobilka 1992, Dunham and Hall 2009, Al-Sabah et al. 2020). Enhancement of the expression of some receptors on the cell surface by switching to the hemagglutinin signal peptide was confirmed with confocal FRET imaging (Quitterer et al. 2011, Al-Sabah et al. 2020). Based on the SignalP 5.0 server, the native pGC-A signal peptide has 97.91% cleavage probability upon expression on the cell surface. Upon switching to the hemagglutinin signal peptide on pGC-A, the cleavage probability was 72.57%. The high cleavage probability may indicate that the mature expressed native pGC-A on the cell surface has no signal peptide.

To simplify protein purification, a decahistidine purification tag (HHHHHHHHHH) was added after the substituted signal peptide. As an alternate, I may have chosen FLAG epitope tag for immunoaffinity-based affinity purification. The FLAG affinity tag provides much higher specificity and selectivity. Notably, more histidine-rich contaminant proteins are present in eukaryotic cells than in the prokaryotic expression system (Gavin et al. 2002, Lee et al. 2012). Considering the cost-effectiveness, high availability, and accessibility of IMAC resin in the laboratory, a poly-histidine tag was used as an affinity tag. To increase the IMAC resin selectivity in the insect cell lysed solution, the traditional six-His tag was replaced with a ten-His tag to boost high binding affinity. Although most structural studies reveal that His-tags does not interfere with protein folding and crystallization, some reports show that His-tags could interfere with crystallization (Booth et al. 2018). Accordingly, a TEV cleavage site was added during the construct design to avoid potential problematic outcomes. In addition, insertion of a TEV cleavage site can allow a second reverse purification in IMAC after TEV protease treatment (Block et al. 2008).

One challenge encountered was the yield of recombinant bacmid. According to the Invitrogen Bac-to-Bac™ system manual, 2 mL culture inoculated with the selected single colony can achieve 500 ng/μL in 40 μL TE buffer. However, my average recombinant bacmid yield was 50-100 ng/μL with doubled (4 mL) culture size. Bacmid is a low-copy number "bacterial artificial chromosome" vector contained baculovirus genome. The additional large recombinant bacmid DNA size may have made the recombinant bacmid yield even worse.

In the large-scale pGC-A purification, high purity protein was achieved, but the protein yield is very low with an average protein yield of 0.05 mg pGC-A isolated from 1.5 L of insect cell culture. Initially, I added 100 mM imidazole in the second IMAC wash buffer, and pGC-A was partially washed out (Figure 2.4.5A). Subsequently, I switched to 90 mM imidazole, but it did not help the general low yield of this protein. The low yield may be attributed to the protein's own character. One similar research report has expressed rat intact pGC-A in baculovirus Sf9 insect cells, revealing similar modest protein production (Chinkers, Singh, and Garbers 1991). Protein sequence BLAST analysis shows that sequence alignments between intact pGC-A in humans (UniProtKB/Swiss-Prot: P16066.1) and rat (XP_032753748.1) has 91.38% identity. After searching pGC-A related reports, the low protein yield could be attributed to glycosylation. Previous reports have shown that glycosylation in the extracellular domain (ECD) of pGC-A does not affect peptide binding (Abe et al. 1993, Misono et al. 2011). However, the glycosylation on the ECD of pGC-A is essential for proper functional protein folding and transport to the cell membrane. (Heim, Singh, and Gerzer 1996, Koller, Lipari, and Goeddel 1993). The insect cell expression system is the most

preferred choice for expressing eukaryotic proteins for structural investigations (Assenberg et al. 2013). More than 70% of known GPCR structures were solved using insect cells as the host (Abiko et al. 2021). However, insect cells do not offer the same native mammalian protein synthesis and post-translational modifications. Some membrane proteins may require a near native environment for functional protein expression. The main drawbacks of insect cells for expressing certain membrane proteins include limited N-linked glycosylation modifications and the different lipid environment that contains higher amounts of cholesterol (Goehring et al. 2014). Both insect and mammalian cells share the same high-mannose N-glycan precursor and intermediate. At the end of the glycosylation pathway, insect cells mainly produce a paucimannose structure, but mammalian cells generate a complex type with terminal sialic acid N-glycans. The limitations of the insect cell system could be the reason why most of the pGC-A functional research is based on either the native animal resources or expression in mammalian cell lines (Shi et al. 2001, Koh et al. 1992, Lowe 1992). In my current purification protocol, CHS was added after membrane solubilization to mimic the native membrane environment and stabilize intact pGC-A. CHS is a detergent used as an additive to replace cholesterol in membrane protein crystallization (Kulig et al. 2014). CHS helped to stabilize the protein during the purification step, but the only concern that still persists is the large-scale expression of intact pGC-A in insect cells. Some glycoengineered insect cell lines are capable of producing glycoproteins with complex terminally sialylated N-glycans (Geisler and Jarvis 2012); however, they are yet to be commercialized.

2.6 Conclusion and Outlook

In summary, I successfully established the insect cell expression platform for functional intact human pGC-A via the Bac-to-Bac™ system. After generating high titer recombinant virus, protein expression optimization was performed by testing different virus concentrations and cell harvest days. To address the protein yield issue, I would contact research labs studying glycoengineered insect cells. It could be helpful to verify whether glycosylation is the cause for the low protein yield and keep glycoengineered cell lines for other glycoprotein-involved projects. Additionally, a mammalian cell line should be explored for pGC-A expression. In last five years, the cost difference between mammalian and insect cell lines has narrowed. Using a mammalian cell line may be the most optimal condition for studying the pGC-A protein. Furthermore, I established the first purification method for the structural determination of intact human pGC-A. On average, 0.05mg of pGC-A can be purified from 1.5 L of transfected insect cell culture. This purification method mainly includes a high salt membrane wash, membrane solubilization, and His-tag affinity and gel filtration chromatography. The successfully established intact human pGC-A expression and purification methods pave the road for future protein structure determination and functional studies.

CHAPTER 3

PROTEIN CHARACTERIZATION AND FUNCTIONAL STUDY OF INTACT HUMAN NATRIURETIC PEPTIDE RECEPTOR A

3.1 Abstract

Natriuretic peptide receptor A is also known as particulate guanylyl cyclase receptor A (pGC-A) due to the guanylyl cyclase domain (GCD). pGC-A is an atrial natriuretic peptide (ANP) receptor, critical for controlling cardiovascular, renal, and endocrine homeostasis functions (Pandey 2011). Currently, only the extracellular domain (ECD) of the rat pGC-A structure is available, with and without binding different substrate ligands (Ogawa et al. 2004, Ogawa, Kodama, and Izumikawa 2020, van den Akker et al. 2000). The previous crystal structure of rat pGC-A ECD, with and without ANP substrate, proposed that pGC-A may exist as a homodimer with a rotation mechanism. Each ECD domain rotates by 24° counterclockwise after ligand binding. (Misono et al. 2005, Qiu et al. 2004). Conversely, earlier studies of heterologously expressed intact pGC-A in mammalian cells and partially purified native pGC-A from bovine adrenal cortex membranes, suggested that intact pGC-A may exists as tetramer (Iwata et al. 1991, Jewett et al. 1993) in the native state, and the oligomerization of pGC-A is ligand-independent (Lowe 1992). In Chapter 3, the intact human pGC-A oligomerization was characterized. In addition, a functional assay was performed to measure the level of cGMP synthesis by incubating the human ANP and GTP molecules with purified pGC-A in solution or expressed in cells directly. Figure 3.1 summarizes the main protein characterization steps in Chapter 3.

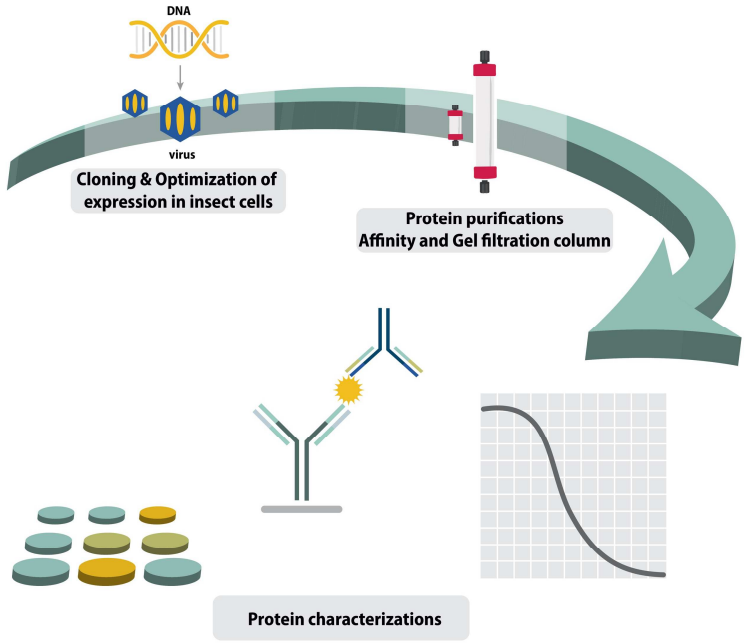

Figure 3.1 Summary scheme of major experiments for protein characterization and functional study.

3.2 Introduction

3.2.1 Oligomerization of Natriuretic Peptide Receptor A

Previous reports have suggested that natriuretic peptide receptor A (pGC-A) is activated by a hormone-induced dimerization mechanism, analogous to growth hormone and cytokine receptors (Labrecque et al. 1999, Koesling, Bohme, and Schultz 1991). However, dimer, trimer, and tetramer oligomeric states of pGC-A have been reported without introducing the binding ligand, ANP (Wilson and Chinkers 1995, Misono et al. 2005). One study revealed that human intact pGC-A is self-associated before ANP binding (Lowe 1992). Additionally, purification of the ECD of rat pGC-A has shown that the ECD undergoes spontaneous dimerization in the absence of ANP binding (Misono et al. 1999). Collectively, oligomerization of intact pGC-A is hormone-independent. However, one

question that needs to be resolved is which oligomeric state is the native state of intact pGC-A (Potter et al. 2009).

Based on the crystal structure information of rat pGC-A ECD with and without ANP, a rotation activation mechanism was proposed for homodimer intact pGC-A (Misono et al. 2005, Misono et al. 2011). Binding of ANP induces a 24° rotation to each juxta-membrane signaling motif (Figure 1.4 C&D). The juxta-membrane signaling motif is a conserved ECD region among guanylyl cyclase receptors (Huo, Abe, and Misono 1999). Each monomer's juxta-membrane domain rotates in the opposite direction. This rotation mechanism may further transduce the rotation signal toward the intracellular domain (ICD) and bring the two guanylyl cyclase domains (GCDs), which are proposed to rotate in opposite directions, from a closed conformation into an active conformation (Figure 1.4 C&D) (Misono et al. 2005).

Conversely, the partially purified intact pGC-A from bovine adrenal cortex tissue indicates that pGC-A exists as a disulfide-linked tetramer in its native state (Iwata et al. 1991). A higher oligomer structure may improve the local concentration of the receptor to increase the ligand-binding sensitivity (Iwata et al. 1991). Additionally, crosslinking of human intact pGC-A (expressed in the human HEK293 cell line) with disuccinimidyl suberate crosslinker revealed a possible tetrameric complex in an immunoblot detection experiment (Lowe 1992). Based on these protein characterization results, a tetramer model was proposed, such that the pre-association of pGC-A occurs at the ICD (Lowe 1992). In this proposed tetramer model, ligand-binding on the ECD triggers the ATP molecule to bind to the protein kinase like domain (PKLD). The binding of ATP to the PKLD eliminates the PKLD-negative regulation on the GCD. While the GCD is active, ATP

binding on the PKLD also induces a low-affinity state on the ECD to release the ligand (Jewett et al. 1993).

3.2.2 Enzyme-linked Immunosorbent Assay (ELISA)

Enzyme-linked immunosorbent assay (ELISA) is a widely used diagnostic and analytical technique in biomedical research. ELISA is often used to detect and quantify a targeted antigen or antibody in a given sample (Gan and Patel 2013). The immunology concept underlying ELISA is simple: utilizing enzyme-labeled antigen or antibody to detect/bind targeted biological molecules present at minimal concentrations. The most common conjugated labeling enzymes are alkaline phosphatase (AP) and horseradish peroxidase (HRP). Both conjugated enzymes can convert the substrate into a colored or light-emitting product (Beyzavi et al. 1987). AP is a natural enzyme and can be isolated from the bovine intestines. AP has a more linear reaction rate than HRP, increasing the sensitivity by extending the incubation time. The most common substrate for AP is *p*-nitrophenyl phosphate, which can be converted to a yellow product with light absorbance at 405 nm (Danikowski and Cheng 2019). HRP is a natural enzyme commonly found in the horseradish root. HRP catalyzes the oxidation of different organic substrates by hydrogen peroxide and has a higher turnover rate than AP. The most common substrate for HRP is tetramethylbenzidine (TMB), which generates a blue color product after oxidization (Khanmohammadi et al. 2018).

Generally, ELISA is designed as a 96-well solid-phase technique. The immunoreactant (antigen or antibody) is immobilized on the microtiter plate by noncovalent interactions. The sample solution containing the molecules of interest is incubated on the plate. After washing the plate, the secondary enzyme-conjugated antibody

is added to bind with the molecules of interest. A qualitative or quantitative measure can be achieved by detecting the colorimetric or chemiluminescence signal (Reen 1994). Depending on the specific reaction optimizations, ELISA can be sub-classified as direct ELISA, indirect ELISA, sandwich ELISA, and competitive ELISA (Figure 3.2.1).

Direct ELISA

Direct ELISA is the simplest ELISA with fewer steps than other ELISAs. The antigen (sample solution) is immobilized in the 96-well plate, and an enzyme-labeled antibody directly detects the antigen without involving a secondary antibody. However, directly coating the antigen on the plate may cause high background noise. Furthermore, as the primary antibody is directly conjugated with the enzyme, each targeted antigen (protein) requires a specific antibody for detection.

Indirect ELISA

The indirect ELISA method is the most popular method for diagnostic measurements. In this method, a secondary antibody incubation is added to the direct ELISA method. Indirect ELISA has an identical immuno-concept to that of the Western blot technique. After antigen (sample solution) is coated on the plate, the primary antibody is added to bind with the specific antigen. Then, the enzyme-labeled secondary antibody is added to recognize and bind to the primary antibody. The indirect ELISA method has high sensitivity as multiple secondary antibodies could bind to the primary one and increase the binding sensitivity. However, the secondary antibody may non-specifically bind to other contaminants (Boguszewska et al. 2019).

Sandwich ELISA

The key immuno-concept of sandwich ELISA is antibody paring. Instead of coating the targeted antigen (sample solution) on the 96-well plate, the antibody that specifically recognizes the targeted antigen is coated and called capture antibody. Then, the sample containing the targeted antigen is incubated with the capture antibody. After washing the plate, the second antibody with conjugated enzyme, called detection antibody, is added to recognize the target antigen. This sandwich method provides high sensitivity and specificity (Boguszewska et al. 2019). Unlike direct or indirect ELISA, sample antigen purification is not required in the sandwich method. This method is helpful to analyze complex samples such as blood samples (Gan and Patel 2013).

Competitive ELISA

Competitive ELISA is based on the competition between the targeted sample antigen and standard reference antigen bound to the primary (capture) antibody on the 96-well plate. Essentially, the sample antigen competes with the standard reference antigen to bind to the limited capture antibody. The higher the concentration of the sample antigen in the solution, the lower the possibility that the standard reference antigen binds to the primary antibody (weaker ELISA output signal). The signal output is inverse to the target sample antigen concentration (Reen 1994). Competitive ELISA has predominantly been used as a quantitative measurement, which quantifies the concentration of the targeted antigen in the sample. First, the target sample antigen was incubated with a limited amount of primary capture antibodies. Then, the antibody-antigen complexes are added to the plate coated with standard reference antigen. The higher the targeted antigens concentration is in the sample; the fewer primary capture antibodies are available to bind with the reference antigens. The secondary detection antibody binds to the reference antigen bonded primary

antibody. Accordingly, the generated signal is proportional to the concentration of the reference antigen and inverse to the sample antigen concentration (Gan and Patel 2013).

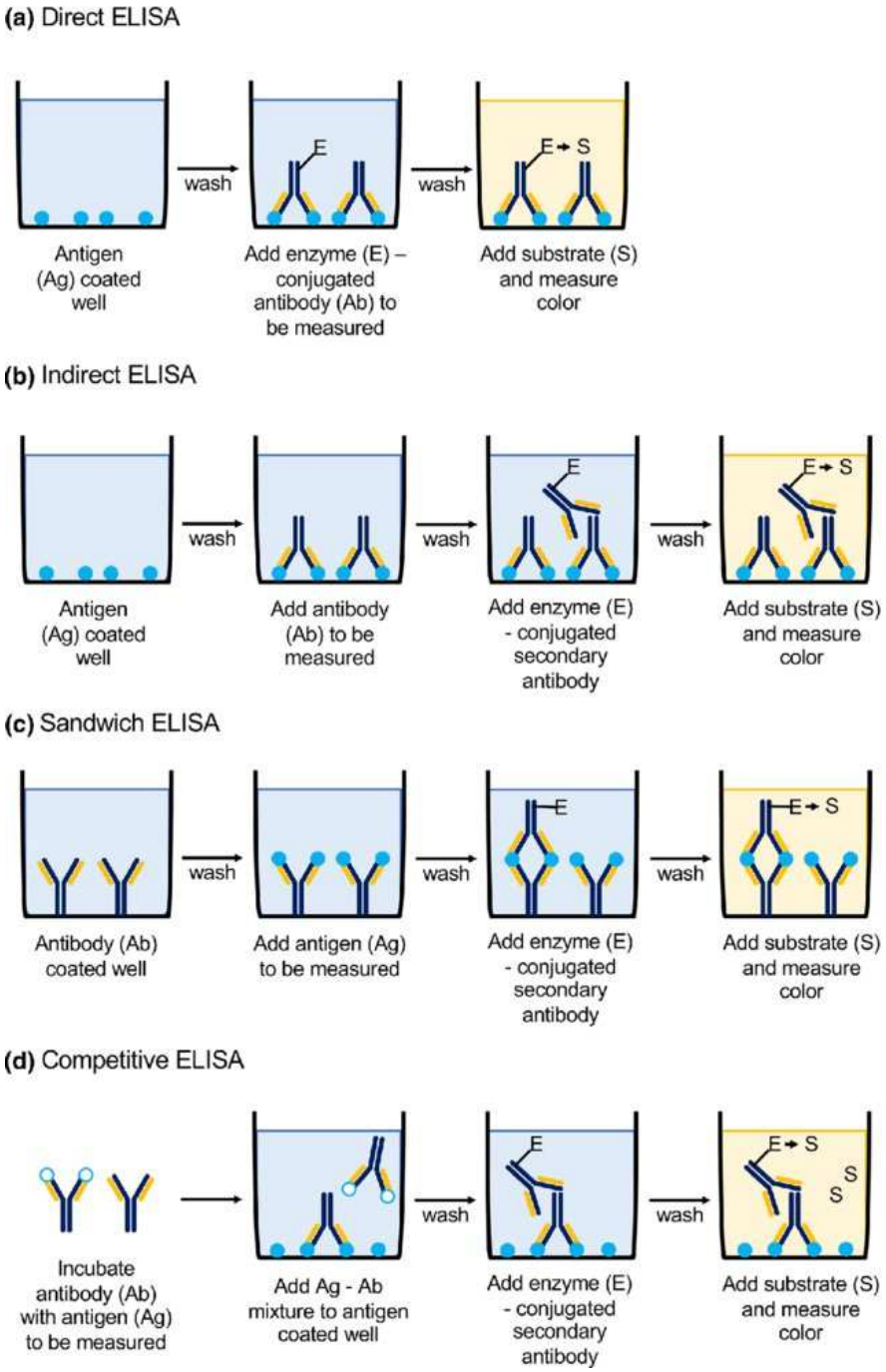

Figure 3.2.1 Summary of four different types of enzyme-linked immunosorbent assay (ELISA): A) direct ELISA method, B): indirect ELISA method; C): sandwich ELISA method; and D): competitive ELISA method. Figure, caption, and description are modified from (Boguszewska et al. 2019) with permission from Springer Nature.

3.2.3 ATP Plays an Allosteric Regulatory Role in Active pGC-A Guanylyl Cyclase

Assessing the ATP function in regulating the pGC-A signal transduction machinery is intricate and controversial. Earlier studies have suggested that ATP molecules accelerate the activity of the pGC-A guanylyl cyclase domain (GCD) (Kurose, Inagami, and Ui 1987, Chang et al. 1990). Subsequent research has suggested that ATP is crucial for signal transduction and GCD activity. In other words, ATP may be involved in the allosteric regulation of GCD activity (Marala, Sitaramayya, and Sharma 1991, Chinkers, Singh, and Garbers 1991). Originally, the ATP regulation domain was proposed on the protein kinase like domain (PKLD) (Chinkers and Garbers 1989, Goraczniak, Duda, and Sharma 1992). Later, a glycine-rich cluster sequence motif (Gly^{503}-Arg-Gly^{505}-Ser^{506}-Asn-Tyr-Gly^{509}) starting 40 amino acids after the transmembrane domain of pGC-A in the PKLD was determined to be essential in both ATP binding and signaling transduction to guanylyl cyclase. In order to emphasize the functionality of the glycine-rich motif, it was also termed as ATP-regulated module (ARM) (Goraczniak, Duda, and Sharma 1992). Studies evaluating the natriuretic peptide receptor C (NPRC) have revealed that ATP is essential for NPRC signal transduction activity (Duda et al. 1993). Unexpectedly, replacing the natriuretic peptide receptor A (NPRA) ARM sequence motif with an NPRC motif has shown identical functionality. The ATP-dependent transduction activity was nearly identical in pGC-A and NPRC (Duda, Goraczniak, and Sharma 1993a). Furthermore, a point mutation experiment on each glycine residue in the ARM sequence motif revealed that the Gly505 is critical to maintain the integrity of the ATP-binding pocket and controls the guanylyl cyclase activity (Duda, Goraczniak, and Sharma 1993b). Moreover, the ARM (glycine-rich

motif) is not part of the direct ATP-binding pocket but plays vital roles in the integrity of the tertiary structure of the ATP-binding pocket (Sharma 2010). These discoveries support earlier research that pGC-A is a modular protein. The binding signal generated from the interaction of atrial natriuretic peptide (ANP) with the ECD of pGC-A is passed to the transmembrane domain. The ARM sequence motif directly below the transmembrane region recruits the ATP molecule to bind and further induces allosteric effects on both ECD and GCD: ATP binding lowers the ECD affinity to bind with the peptide and triggers peptide dissociation; simultaneously, ATP binding activates guanylyl cyclase and accelerates cGMP production (Sharma 2002, 2010).

In addition to the direct ATP allosteric modification of the pGC-A signal transduction in the ARM, another study proposed that signal transduction modification involves an indirect process via phosphorylation (Potter and Hunter 1999a, b). A few research groups have shown that phosphorylation of the pGC-A ATP-binding domain (AAs Ser^{497} to Thr^{513}) via a hypothetical protein kinase indirectly affects guanylyl cyclase activation. In total, six residues were marked as critical phosphorylation sites (Potter and Hunter 1999a, b). Sharma et al. speculated whether pGC-A phosphorylation was exclusively mediated via ATP activation of pGC-A signal transduction. Later, the "3D ARM ATP transduction model" analysis and experimental validation revealed that the phosphorylation of Ser^{506} in the ATP-binding domain depended on the ATP allosteric step (Duda et al. 2005). Possibly, phosphorylation of the other five putative phosphorylation amino acids on the PKLD (Ser^{497}, Thr^{500}, Ser^{502}, Ser^{510}, Thr^{513}) is also dependent on the ATP allosteric step (Potter and Hunter 1999b) (Sharma 2010).

Subsequently, another group has suggested the order in which the two processes occur in signal transduction: based on their hypothesis, pGC-A phosphorylation is essential prior to activation of the ATP-dependent allosteric step (Foster and Garbers 1998). However, experiments with ATP non-hydrolyzable analogs and phosphatase inhibitor (microcystin) were used to measure the guanylyl cyclase activity level. Accordingly, it was determined that ATP-dependent allosteric modification is not affected by phosphorylation. An opposite order of the two processes was proposed in pGC-A signal transduction (Duda et al. 2000, Sharma 2002). Based on this hypothesis, first, activation of the ATP-dependent occurs as an allosteric step, which partially activates the GCD and triggers a conformational change; second, the phosphorylation of the ATP-binding domain fully activates guanylyl cyclase (Sharma 2010, Duda et al. 2000).

Notably, one report challenged the role of ATP in pGC-A signal transduction and postulated that activation of the guanylyl cyclase is ATP independent (Antos et al. 2005). Later, Sharma and colleagues analyzed issues listed in this article (Burczynska, Duda, and Sharma 2007). Sharma et al. highlighted that although ATP addition failed to improve the pGC-A guanylyl cyclase activity during the first 15 s, it robustly accelerated the guanylyl cyclase activity from 5 to 60 s (Antos et al. 2005). In addition, there were a couple of technical flaws in this controversial article. Such as the cell system in this study may not sensitive to the ANP. (Burczynska, Duda, and Sharma 2007, Sharma 2010).

3.3 Methods

High resolution Clear-Native PAGE

The High Resolution Clear Native Polyacrylamide Gel Electrophoresis (hrCN-PAGE) method was performed as described by Wittig et al. (Wittig, Karas, and Schagger 2007). Briefly, 12 µL of concentrated protein was mixed with 12 µL of 2X native PAGE sample loading buffer (100 mM sodium chloride,100 mM imidazole-HCl, 4 mM 6-aminohexanoic acid, 10% glycerol, and 2 mM EDTA [pH 7.0]) immediately before loading of the sample to the native polyacrylamide gel. During the running buffer preparation, 0.05% N-dodecyl β D-maltoside (DDM) was added to the cathode buffer (50 mM Tricine-NaOH and 7.5 mM imidazole-HCl [pH 7.0]). The cathode buffer and the anode buffer (25 mM imidazole-HCl [pH 7.0]) were freshly prepared for single usage. The 4-16% native polyacrylamide gel was commercially available from Thermo Scientific (Invitrogen, Cat. BN1002BOX). Native gel electrophoresis was performed in an ice-chilled basket, and hrCN-PAGE was run at a constant 150 V for a total of 2 h.

Competitive functional ELISA assay protocol for cGMP yield determination

The detailed competitive assay steps were performed as described in the cyclic GMP complete enzyme immunoassay kit (Enzo Life Science Inc, Cat. 901-164). The precise, detailed procedure for quantifying the cGMP yield in the provided sample solution was performed as follows: The pGC-A sample solution with unknown cGMP content was diluted 10-fold (e.g., 110 µL of the original pGC-A sample solution was mixed with 990 µL sample buffer solution). Then, the diluted sample solution was further serial diluted 4-fold using the dilution factor ratio 1:5 (e.g., 100 µL sample solution is diluted in 400 µL sample buffer solution). To calculate the unknown cGMP sample concentration, a standard reference cGMP (5000 pmol/mL) was serially diluted 5 times (500 pmol/mL, 100 pmol/mL, 20 pmol/mL, 4 pmol/mL, and 0.8 pmol/mL). In the ELISA

plate, all standard, sample, and control wells were run in duplicate. To prepare the ELISA plate, 150 µL of sample buffer solution was added to the NSB (non-specific binding) control well, and 100 µL sample buffer solution was added to the B0 (maximum binding) well. Both diluted standard cGMP solutions and unknown samples were individually pipetted (100 µL) into each well. After adding unknown and standard samples to the plate, 50 µL of alkaline phosphatase-conjugated cGMP (blue color) was added to each well, except for the TA (total activity) and blank wells. Then, 50 µL of primary detection cGMP antibody (yellow color) was added to each well, except for the TA, blank, and NSB wells. Each well should be green in color, except for NSB (blue color), TA, and blank (empty, no color). The plate was sealed with a ELISA plate sealer and incubated on a plate shaker (500 rpm) at room temperature for 2 h. The 10× stock wash buffer was diluted to 1× working concentration. After emptying the plate contents, 400 µL wash solution was added to each well and then emptied. A total of three washes were performed. After the final wash, the wells were emptied by inverting the plate and firmly tapping the plate on a lint-free paper towel to remove any remaining wash buffer. Then, 5 µL of the alkaline phosphatase-conjugated cGMP (blue color) was pipetted to the TA wells, followed by the addition of 200 µL of the *p*-nitrophenyl phosphate substrate solution into each well. The plate was sealed with a plate sealer and incubated at room temperature for 1 h without shaking. Later, 50 µL of stop solution was pipetted into each well. After the plate reader was blanked using substrate blank wells, the optical density (OD) was measured at 405 nm. The major competitive ELISA assay steps are shown in Figure 3.3.1.

For calculating the unknown concentration, the average net OD for each standard and sample was calculated by subtracting the average NSB OD from the average OD for each standard and sample. The percent bound in each well was calculated by dividing the average net OD by the net B0 (maximum binding) OD.

$$Percent\ Bound = \frac{Net\ OD}{Net\ B0\ OD} * 100$$

Then, the percent bound (B/Bo) versus cGMP concentration for standards was plotted with a straight line passing through the points. The cGMP concentration can be determined by interpolation. For calculating the cGMP levels in the sample accurately, the 4-parameter logistic curve fitting program was used for ELISA data analysis. GraphPad Prism9 (GraphPad Software, Inc., La Jolla, CA, USA) was used for in statistical evaluation. Calculator software was found online (https://www.arigobio.com/elisa-analysis; https://www.aatbio.com/tools/four-parameter-logistic-4pl-curve-regression-online-calculator).

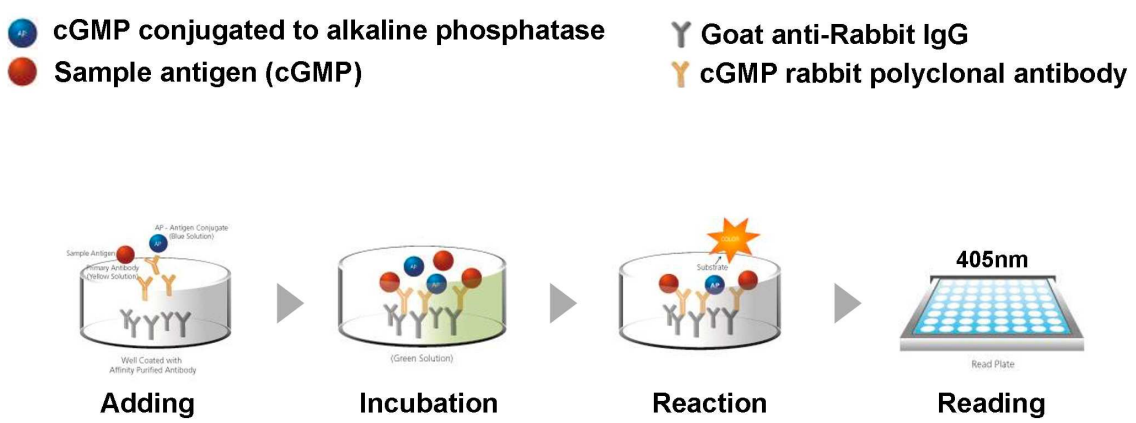

Figure 3.3.1 Summary of major experimental steps in competitive cGMP ELISA assay. 96-well ELISA plate was coated with goat anti-rabbit IgG (gray antibody). Primary detection cGMP antibody (yellow antibody) was incubated with sample solution and reference cGMP conjugated with alkaline phosphatase (blue). After incubation (green color solution) and washing steps, pNpp substrate was added to the plate. The pNpp substrate was hydrolyzed to a yellow color product and the plate was read at 405 nm. The amount of signal is indirectly proportional to the amount of cGMP in the sample. The picture was adapted from Enzo Life Science cyclic GMP kit manual with modifications.

Whole-cell activity assay for intact pGC-A expressed in Sf9 Cells

The cyclic GMP complete enzyme immunoassay kit (Enzo Life Science Inc, Cat. 901-164) was employed to measure the whole-cell activity assay. The exact cell preparation and ligand incubation process have been described previously (Chen et al. 2019, Chen et al. 2020). Both transfected and wildtype Sf9 insect cells were grown for 3 days in the incubator. Later, insect cells were pelleted via centrifugation and suspended in fresh Sf-900 III SFM medium (Gibco, Cat. 12659017). Various concentrations of ligand MANP (mutant ANP) (Bachem, customized peptide) (1×10^{-5} M, 1×10^{-6} M, 1×10^{-7} M, 1×10^{-8}) with nonspecific phosphodiesterase inhibitor isobutylmethylxanthine (IBMX) (5×10^{-3} M) (Sigma, Cat. I5879-250MG) were added and incubated with insect cells for 20 min. Then, cells were pelleted and lysed with lysis buffer (sodium acetate buffer containing proteins and sodium azide) provided in the assay kits. The exact ELISA step-

by-step procedure was described above. Briefly, the lysed samples, standards, and controls were added to a 96-well microtiter plate coated with goat anti-rabbit IgG antibody (Enzo Life Science Inc, Cat. 80-0060). The solution of cGMP conjugated to ALP was then added, followed by a solution of rabbit polyclonal antibody to cGMP. After incubation for 2 h at room temperature with shaking at 500 rpm, the plate was washed three times with wash buffer. P-nitrophenyl phosphate substrate solution was added to the plate. Then, the plate was incubated at room temperature for 1 h without shaking. The stop solution was added, and the plate was read at 405 nm, with a blank against the substrate.

Purified Intact pGC-A in vitro Functional Activity Test

The cyclic GMP complete enzyme immunoassay kit instructions (Enzo Life Science Inc, Cat. 901-164) was followed to measure the purified pGC-A activity with a few modifications. Purified pGC-A protein peak fractions (tetramer and monomer size) were collected separately. The same amount of concentrated peak fractions was mixed with human ANP (hANP) (Sigma-Aldrich, Cat. A1663-.5MG) and GTP (Sigma-Aldrich, Cat. G8877-25MG) and incubated overnight on an ice bath. Depending on the experimental setting, ATP (Sigma-Aldrich, Cat. A1852-1VL) was also added during the incubation. To maintain hANP stability and integrity, 5 mM DTT was removed in the last Superose6 gel filtration, and no reducing reagent was added to the protein sample buffer.

In this specific experiment, both tetramer and monomer size fraction samples were concentrated to 0.36 mg/mL. To prepare the non-ATP incubated samples, 35 µL of each protein sample was mixed with 5 µL of hANP (5 mg/mL) and 35 µL of GTP (13.7 mg/mL). The total volume of the non-ATP incubated sample was 75 µL. For the non-

ATP incubated control sample, 35 µL of GTP (13.7 mg/mL) was mixed with 40 µL of protein purification buffer (50 mM HEPES-NAOH [pH 7.5], 150 mM NaCl, 10% glycerol, 0.05% DDM, and 0.01% CHS). For preparing the ATP incubated samples, 35 µL of each protein sample was mixed with 5 µL of hANP (5 mg/mL) and 35 µL of GTP (13.7 mg/mL) and 35 µL of ATP (26.4 mg/mL). The total volume of the ATP incubated sample was 110 µL. For the ATP incubated control sample, 35 µL of GTP (13.7 mg/mL) and 35 µL of ATP (26.4 mg/mL) were mixed with 40 µL of protein purification buffer. All prepared samples were incubated on an ice bath in the cold room and subjected to overnight incubation.

The exact ELISA procedure was described above. Briefly, after overnight incubation, all ELISA standards, controls, and samples were prepared and diluted with protein purification buffer (50 mM HEPES [pH 7.5], 150 mM NaCl, 5 mM DTT, 10% glycerol, 0.05% DDM, and 0.01% CHS). Then, 100 µL of the standards, controls, and samples were added to a 96-well microtiter plate coated with goat anti-rabbit IgG antibody (Enzo Life Science Inc, Cat. 80-0060). Next, a cGMP solution conjugated to ALP was then added, followed by the reagent of rabbit polyclonal antibody to cGMP. After 2 h of incubation at room temperature with shaking at 500 rpm, the ELISA plate was washed three times using the wash buffer (Tris buffered saline containing detergents). P-nitrophenyl phosphate substrate solution was added to the ELISA plate. The plate was incubated at room temperature for 1 h without shaking. Then, a stop solution (trisodium phosphate solution) was added, and the ELISA plate was read at 405 nm, with a blank against the substrate.

3.4 Results

Different pGC-A oligomeric states observed in gel filtration chromatography

In each intact human pGC-A purification experiment, 20 to 25 g of transfected insect cells was lysed for protein purification. The Superose6 gel filtration column was extensively used to purify pGC-A before crystallization. Interestingly, different oligomeric states were observed under the same purification protocol. To determine the relative molecular size of the different oligomers, 5 column molecular weight standard proteins were injected into the Superose6 column to generate the column performance elution profile (Figure 3.4.1 A). In the standard elution profile, standard protein thyroglobulin (669 kDa) was eluted at 14.19 mL, ferritin (440 kDa) was eluted at 15.96 mL, and aldolase (158 kDa) was eluted at 17.58 mL, ovalbumin (44 kDa) was eluted at 18.49 mL, and aprotinin (6.5 kDa) was eluted at 21.60 mL (Figure 3.4.1 A). On comparing the Superose6 column standard chromatography profile (Figure 3.4.1 A) with the pGC-A oligomer elution profiles (Figure 3.4.1 B, C& D), the results suggested that the pGC-A oligomers were either tetramer (480 kDa) or monomer (120 kDa).

The chromatography profiles in Figure 3.4.1 B, C& D summarize three common intact pGC-A oligomeric states during the last Superose6 purification. Interestingly, the ratio of the relative tetramer size to relative monomer size differed in each purification, although the purification protocol was identical without any optimization. In Figure 3.4.1B, the major peak was eluted around 17.3 mL, and no other oligomeric state was visible. The elution volume was remarkably close to the standard aldolase protein (158 kDa) elution profile at 17.58 mL, which indicated that the majority peak in the Figure 3.4.1B represents the monomer size of intact pGC-A. As shown in Figure 3.4.1B, the

majority peak was eluted around 15.0 mL, with a minor shoulder peak around 17.1 mL (monomer). The majority peak elution volume was extremely close to the standard ferritin protein (440 kDa) elution profile at 15.96 mL, indicating that the majority peak in Figure 3.4.1B represents the tetramer size of intact pGC-A. In Figure 3.4.1 C, two prominent elution volume peaks can be observed at approximately 14.76 mL and 17.29 mL, representing the tetramer and monomer size, respectively. In addition, a minor peak was eluted at 16.62 mL in Figure 3.4.1B. It may represent the dimer state of intact pGC-A. However, it is difficult to define the size due to the Superose6 resolution limitation. In addition, the possible dimer peak was rarely observed during the performed purifications. Most peaks represent either the tetramer or monomer.

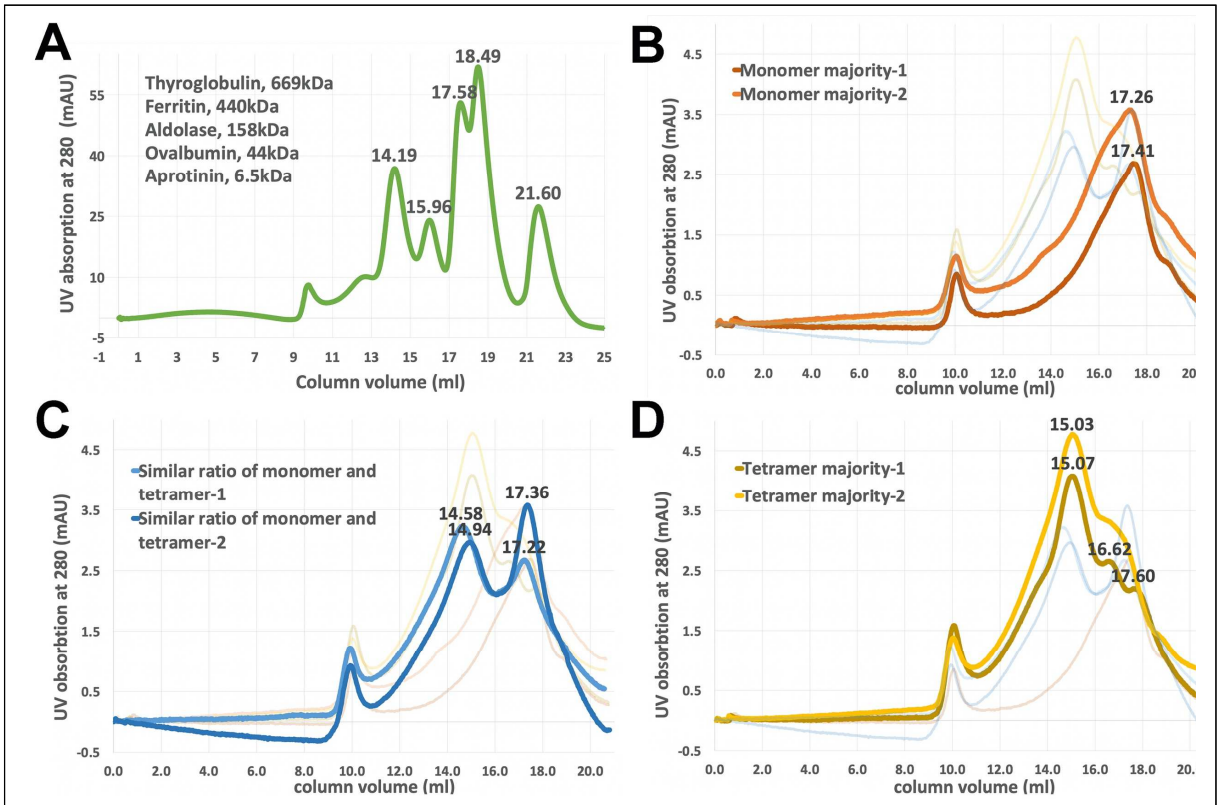

Figure 3.4.1 Summary of different pGC-A oligomeric states in Superose6 chromatography from different experiments. A dynamic ratio of tetramer and monomer can be observed in each Superose6 purification step under same purification protocol. The dynamic ratio between tetramer and monomer is affected by protein concentration. A) The Superose6 10/300GL column performance profile. Five standard proteins were used to generate relative molecule elution points based on different molecular sizes. The 5 standard proteins were thyroglobulin (669 kDa) eluted at 14.19 ml, ferritin (440 kDa) eluted at 15.96 ml, aldolase (158 kDa) eluted at 17.58 ml, ovalbumin (44 kDa) eluted at 18.49 ml, and aprotinin (6.5 kDa) eluted at 21.60 ml; B) Peak intensity at 17.3 ml corresponds to pGC-A monomeric state. pGC-A monomer is the major peak determined by chromatography. Other ratios were faded out in the background and served as supplemental comparison; C) Peak intensity at 14.76 ml and 17.29 ml corresponds to pGC-A tetrameric and monomeric state respectively. pGC-A tetramer and monomer present similar ratios in the chromatographic separation. Other ratios were faded out in the background and served as supplemental comparison; D) Peak intensity at 15.05 ml corresponds to pGC-A tetrameric state. pGC-A tetramer is the major peak determined by chromatography. Other ratios were faded out in the background and served as a supplemental comparison.

Collectively, the Superose6 elution chromatography profiles of intact human pGC-A could be classified as the tetramer majority group (Figure 3.4.1D), monomer

majority group (Figure 3.4.1B), and a separation that contains similar ratio of tetramer and monomer group (Figure 3.4.1C). This unique and distinctive protein characteristic may be dependent on the protein concentration. On carefully comparing the peak sizes and UV absorption levels between Figure 3.4.1 B, C, and D, it was observed that the higher the protein concentration in the injected pGC-A sample, the larger the size of the earlier eluting peak, which represents the tetrameric in size. Similar results were also indirectly observed during injection of the concentrated elution sample of the affinity chromatography into the Superose6 column. The smaller sample volume (approximately 300 μL) and more concentrated the protein was before injection, the larger the tetramer peak size observed; the larger the sample volume (approximately 450 μL), and the lower the protein concentration, the smaller the peak size was observed for the tetramer. Interestingly, this concentration-dependent protein oligomerization phenomenon was observed beyond the intact pGC-A protein. Several of research groups have reported that purification of the ECD of pGC-A shows different oligomeric states during gel filtration. The oligomerization is concentration-dependent, and the dimerization affinity can be further enhanced upon hormone binding (Misono et al. 1999, van den Akker et al. 2000, Ogawa et al. 2009). In addition, self-oligomer formation has been found to occur during the purification of the ECD of natriuretic peptide receptor C (He et al. 2001).

Characterizations of different oligomeric peak fractions eluted from Superose6 column

As mentioned in the previous section, most of the intact pGC-A oligomeric states observed in the Superose6 column best match either a tetramer or a monomer size. The potential dimer peak was rarely observed. In one specific purification experiment, the P1 peak was labeled for the column void volume peak, and the potential tetramer (P2), dimer

(P3), and monomer (P4) size peaks were observed and labeled P1 to P4 in Figure 3.4.2 A. To confirm that all these oligomer size peaks were pGC-A, both western blot (anti-pGC-A) and Coomassie blue gel electrophoresis analyses were performed (Figure 3.4.2 B). Both gel electrophoresis analysis results revealed that P2, P3, and P4 were pGC-A proteins. Not surprisingly, the P2 and P3 peaks, which were suspected to represent the tetramer and dimer oligomers based on the elution volume, showed higher oligomer bands at 250 kDa in the western blot. All fraction samples were incubated with β-mercaptoethanol added SDS sample buffer and heated at 95ºC for 10 min. Heating protein samples with reducing reagent added SDS sample buffer can reduce disulfide bonds and linearize protein structures. The oligomerization of intact pGC-A may not be solely dependent on potential disulfide bond formation. The remaining P2, P3 and P4 fractions were fully concentrated and analyzed by dynamic light scattering (DLS) to verify whether the protein concentration altered the intact pGC-A oligomeric state (Figure 3.4.2 C). All peaks demonstrated a uniform and straight pattern in the heat map. Furthermore, all three concentrated peak samples had the same radius reading at 11 nm, suggesting a markedly similar protein size among the concentrated samples after concentrating the fractions of the monomer, dimer, and tetramer peaks. The molecular weight estimated from the size of 11nm (110nm) corresponded to 800kDa. This means that the protein eluting in P2, P3, and P4 as trammer, dimer and monomer, all form a higher oligomeric state during the concentration step. Collectively, the DLS results suggested that the intact pGC-A oligomerization was protein concentration-dependent (Figure 3.4.2 C).

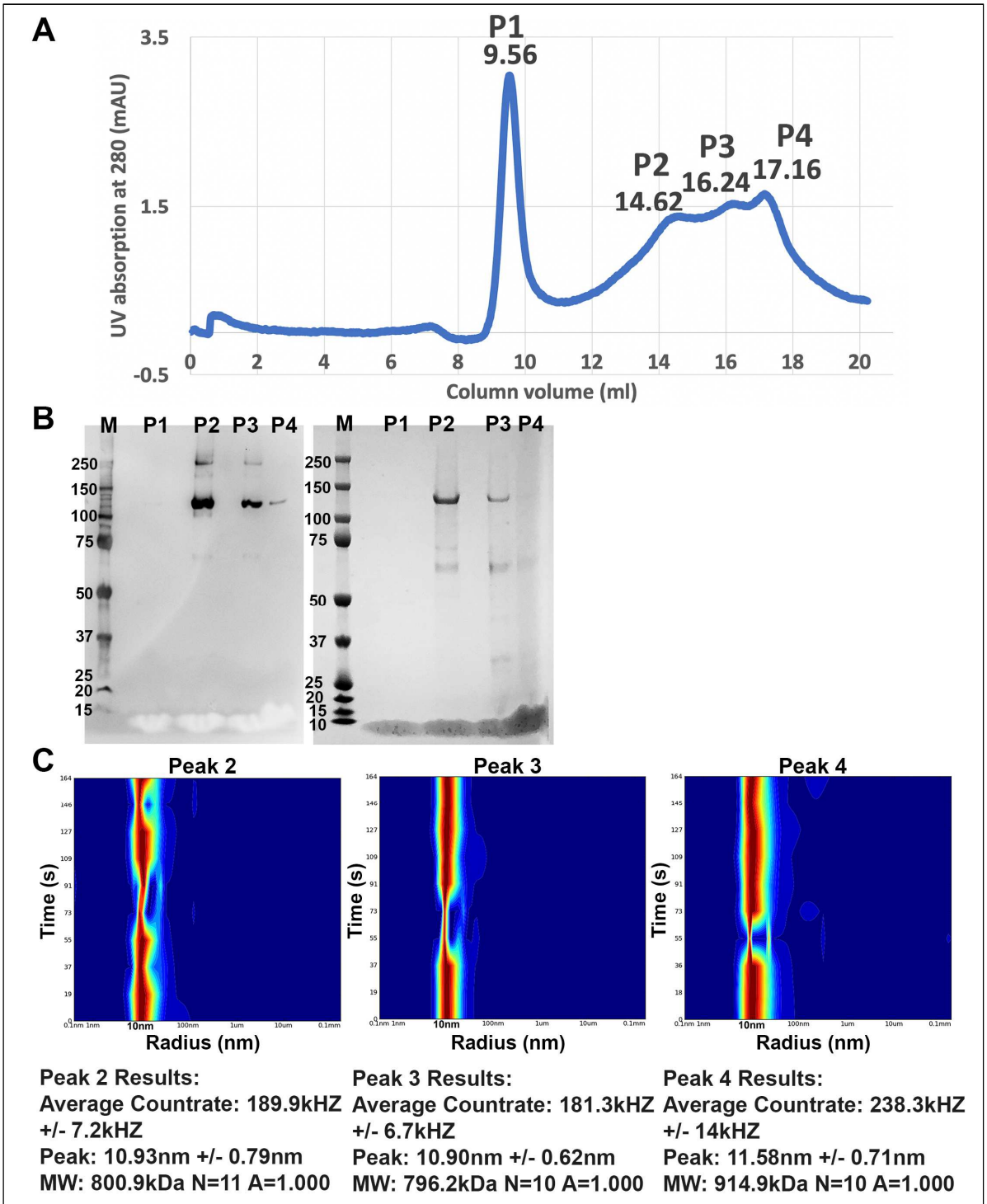

Figure 3.4.2. Summary of characterization of different oligomeric peak fractions from Superose6 column. A) Different oligomeric peaks are shown in the Superose 6 elution profile; B) Western blotting (left) (anti-pGC-A) and Coomassie Blue staining (right) were used to analyze different peak fractions; C) Dynamic light scattering was used to analyze the fully concentrated protein samples from Peak 2, 3, and 4.

Intact pGC-A high oligomeric state confirmed as a tetramer

In the recent literature, pGC-A is proposed to be a homodimeric receptor in cells based on the dimeric structure of the ECD of rat pGC-A, solved with and without ligands (Misono et al. 2005, Rukavina Mikusic et al. 2018). In earlier studies, intact pGC-A expressed in mammalian cells or partially purified from bovine revealed tetrameric oligomerization in its native state of intact pGC-A (Iwata et al. 1991, Jewett et al. 1993). In the current study, the intact pGC-A elution profile from Superose6 chromatography corroborated the earlier tetramer theory. The major peak eluted from the Superose6 column was either relative to tetramer or monomer size (Figure 3.4.1). In addition, the DLS results confirmed that protein concentration affected intact pGC-A oligomerization (Figure 3.4.2).

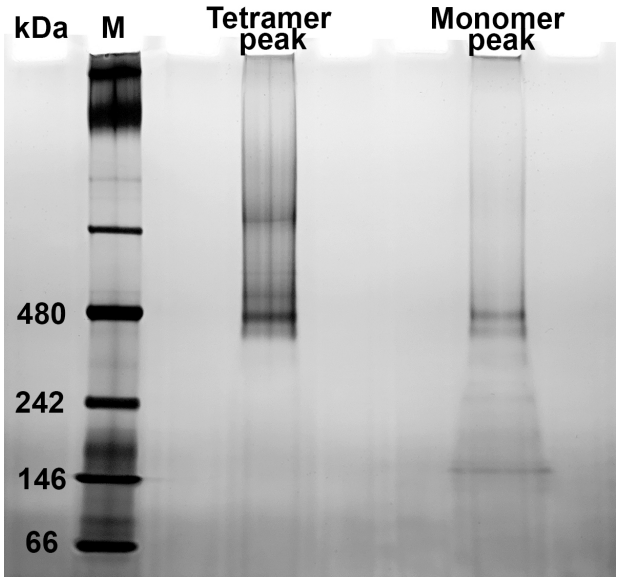

Figure 3.4.3 Silver staining of high resolution clear-native PAGE. Superose6 column eluted tetramer and monomer size peaks were concentrated and analyzed in 4-16% native gel. The intact pGC-A monomer is 120 kDa size. PAGE, Polyacrylamide gel electrophoresis.

Indeed, the gel filtration chromatography is based on separation by molecular size. However, the separation process is not strictly processed by molecular mass. The molecular shape also affects migration speed in the column (O'Fagain, Cummins, and O'Connor 2011). Next, I attempted to verify whether the intact pGC-A high oligomer state eluted from the Superose6 column represents the true tetramer (Figure 3.4.1) and that the protein concentration truly converted the monomer into higher oligomers as indicated by DLS (Figure 3.4.2 C). Accordingly, two peak fractions eluted from the Superose6 column, representing the tetramer size and monomer sizes; these were collected and concentrated separately for high resolution clear native polyacrylamide gel electrophoresis (hrCN-PAGE) analysis. The prepared tetramer and monomer peak samples were loaded on a 4-16% clear native gel, followed by silver staining (Figure 3.4.3). The silver-stained hrCN-PAGE results revealed that the tetramer peak fraction sample featured a strong signal band at 480 kDa, representing a tetramer. Interestingly, the sample from the concentrated monomer size peak also showed a strong tetramer size band signal. These results matched the previous DLS characterization results (Figure 3.4.2 C) and revealed that protein concentration converted the monomer into a higher oligomer state. However, no significant homodimer (240 kDa) band signal was observed in neither the tetramer nor monomer peak samples. A possible minor band signal in the monomer peak indicated homodimer size; however, even with silver staining, visualization remained challenging (Figure 3.4.3). In conclusion, the higher oligomeric state peak sample from the Superose6 column chromatography was confirmed as a tetramer. Moreover, the protein concentration contributed to converting monomer into the higher oligomer state (tetramer state).

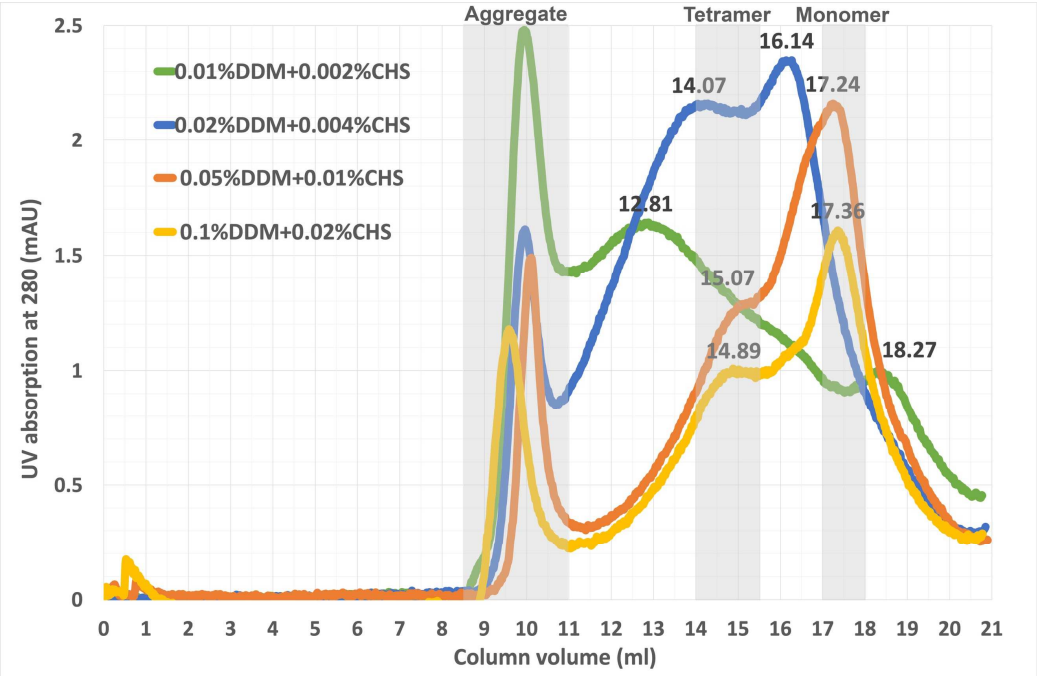

Figure 3.4.4 Summary of intact pGC-A Superose6 column elution profiles from four different detergent concentrations. Four different detergent concentrations were tested during the protein purification step, including 0.01% DDM+0.002% CHS, 0.02% DDM+0.004% CHS, 0.05% DDM+0.01% CHS, and 0. 1% DDM+0.02% CHS. DDM, *n*-dodecyl-β-D-maltopyranoside; CHS, cholesterol hemisuccinate.

Following purification using hrCN-PAGE, the purified intact pGC-A was confirmed as a tetramer. There is a possibility that the detergent during protein purification may prompt to change the oligomeric state. For example, insufficient detergent concentration may force the formation of a detergent micelles containing two copies of the dimer or four copies of monomers, which may appear as a tetramer. To address the potential detergent issue, different concentrations of detergent were used for purification after cell membrane solubilization. In this experiment, 4 different detergent concentrations were tested: 0.01% DDM+0.002% CHS, 0.02% DDM+0.004% CHS, 0.05% DDM+0.01% CHS and 0. 1% DDM+0.02% CHS. As DDM has a CMC of

approx. 0.01% the DDM concentrations corresponded to 1x CMC, 2 x CMC, 5X CMC and 10 x CMC. The Suprose6 elution profiles for these 4 different detergent concentrations are shown in Figure 3.4.4.

In Figure 3.4.4, the results showed a general trend that a higher percentage of detergent tended to shift the protein toward monomer size (lower molecular weight). The critical micelle concentration (CMC) for DDM is 0.0087%. The formed DDM micelle has a relatively large size (65-70 kDa) (Chae et al. 2010, Stetsenko and Guskov 2017). For the membrane protein purification, at detergent concentration has to be higher than the CMC and a twofold the CMC detergent concentrations is optimal for maintenance of the protein detergent micelles (Lalaurie et al. 2018). Furthermore, the detergent:protein weight to weight ratio should exceed 4:1 (Anatrace, Detergents and Their Uses in Membrane Protein Science).

As expected, the sample purified in the 0.01% DDM+0.002% CHS purified sample showed a much higher aggregation peak at the void volume (9.6 mL) than other detergent concentrations, with an additional degradation peak at 18.27 mL. The Superose6 column standard protein elution profile is presented in Figure 3.4.1 A. In the 0.02% DDM+0.004% CHS concentration elution profile, although a potential dimer peak (16.14 mL) occurred in the elution profile, there was no monomer peak, with an even higher 669 kDa oligomer peak (14.07 mL) formed before the potential dimer peak (Figure 3.4.4). The results suggested that the 0.02% DDM+0.004% CHS concentration may be insufficient for intact pGC-A purification and tend to form aggregation. In contrast, the detergent concentration at 0.05% DDM+0.01% CHS and 0.1% DDM+0.02% CHS showed similar elution profiles. Most of the peak is of monomer size

(17.3 mL), with a minor tetramer size peak (15.0 mL) (Figure 3.4.4). The same elution profiles suggested that 0.05% DDM+0.01% CHS was sufficient detergent concentration for intact pGC-A purification. In addition, there was no dimer peak when the detergent concentration increased. The 0.05% DDM+0.01% CHS detergent concentration was the only concentration used during intact pGC-A purification. Together, the "missing" dimer peak and observed tetramer were independent of the detergent concentration.

There is another possibility that 0.05% DDM+0.01% CHS is an excessive detergent concentration that breaks the intact tetramer into four monomers during pGC-A purification. While the 0.05% DDM+0.01% CHS detergent concentration was used during most of the pGC-A purifications, this may need to be revisited. The 0.02% to 0.05% detergent concentration may be more suitable for isolation of the intact tetramer.

Intact pGC-A tetramer was not oligomerized through disulfide bonds

In an earlier study, the partially purified bifunctional pGC-A from bovine adrenal cortex tissue was proposed to exist as a disulfide-linked tetramer in its native state (Iwata et al. 1991). The number of cysteine residues in pGC-A varies between species. For instance, there are a total of 17 cysteine residues in intact human pGC-A. It is unclear if all cysteine residues are involved in the formation of intramolecular or intermolecular disulfide bonds. In my previous gel electrophoresis analysis, purified intact pGC-A was incubated with a reducing agent, with SDS sample buffer added and heated at 95ºC for 10 min. However, there was a higher pGC-A oligomer signal in the western blot (Figure 3.4.2 B). To test if the formation of intermolecular disulfide bonds caused the tetrameric oligomeric state, no reducing reagent was added in this specific purification experiment.

Based on the Superose6 column standard elution profile (Figure 3.4.1 A), three peak fractions from each peak, which represented the tetramer (480 kDa), dimer (240 kDa), and monomer (120 kDa) size, were collected separately. Each concentrated peak fraction was further split into two portions. One portion contained the sample only, and another portion was mixed with 1 M dithiothreitol (DTT). All samples were incubated on ice overnight. Then, the incubated samples were loaded on the native gel. The silver-stained hrCN-PAGE revealed no difference in all three peak samples, both with and without DTT (Figure 3.4.5). Accordingly, the tetrameric oligomer is not caused by a disulfide bond. Furthermore, in the dimer size peak sample, a potent protein signal was observed in form of tetramer and monomer size bands, with only a minor dimer peak (Figure 3.4.5). These results indicated that the dimer is rarely formed when compared with tetramer and monomer. Collectively, these experimental results showed that the pGC-A tetramer is not oligomerized through disulfide bonds. Furthermore, the rarely observed dimer in the Superose6 column elution profile was not due to the column resolution limit. The results imply that the dimer is not a favored oligomeric state.

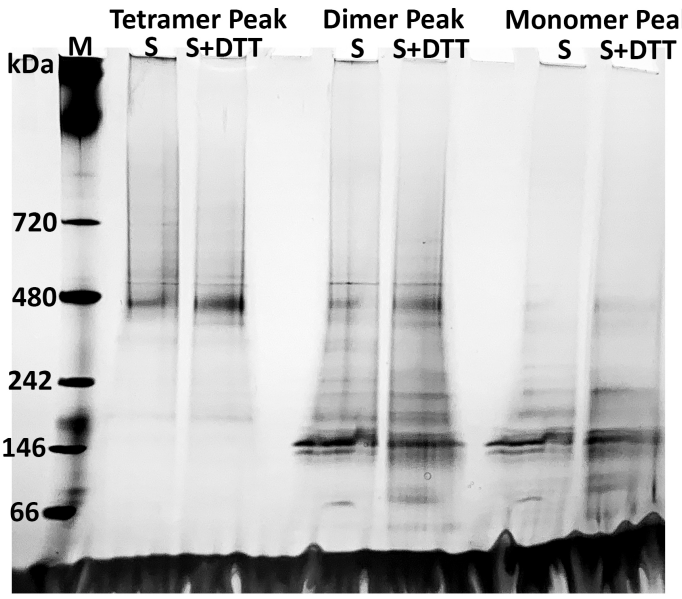

Figure 3.4.5 Silver staining of different oligomeric samples treated with or without overnight DTT in clear-native PAGE. Three peak samples, which represent tetramer, dimer, and monomer size of intact pGC-A, were concentrated and split in half. S, concentrated sample only; S+DTT, concentrated sample incubated with 1 M DTT overnight. DTT, Dithiothreitol.

Recombinantly expressed intact pGC-A was functional in a whole-cell activity assay

The binding of ANP to the ECD of intact pGC-A triggers signal transduction toward the intracellular GCD. Then, the activated GCD converts GTP into cGMP for further physiological modulations. To test if the expressed intact human pGC-A was functional in Sf9 cells, the whole-cell functional activity was based on measuring the produced cGMP concentration via a competitive ELISA assay (Figure 3.4.6). Transfected cells expressing pGC-A (Sf9-pGC-A), or non-transfected cells (Sf9), were incubated with different concentrations of MANPs. MANP is a novel natriuretic peptide developed based on the cardiac hormone ANP that contains a C-terminal extension that increases binding to the receptor and decreased binding to its counterplayer neprilysin which degrades ANP (Chen et al. 2020). To prevent degradation of produced cGMPs in cells,

isobutylmethylxanthine (IBMX), a non-specific inhibitor of cAMP and cGMP phosphodiesterases, was also added during the ligand incubation.

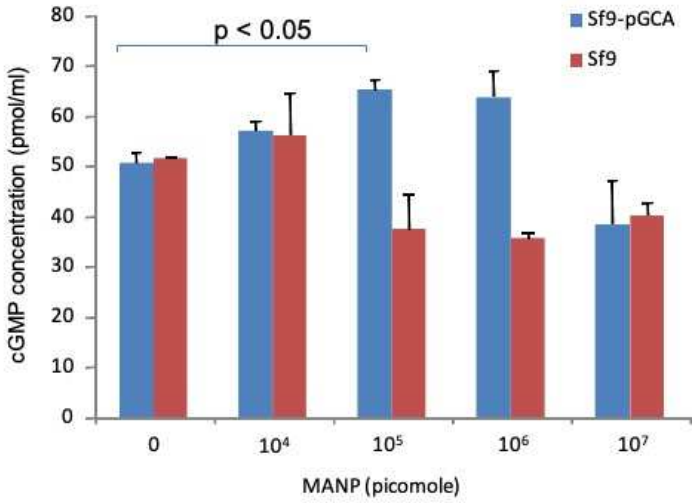

Figure 3.4.6. Determination of expressed intact pGC-A functionality via whole-cell activity assay. The competitive ELISA assay measured the cGMP yield level in Sf9 cells expressing intact pGC-A versus control Sf9 cellst. Both cell types were incubated with different concentrations of MANP ligand (0 to 10^7 pmol). Incubation with 0 pmol MANP served as a negative control.

Figure 3.4.6 shows that without adding MANP (0 picomole), Sf9 cells expressing intact pGC-A, and the control Sf9 cells had almost the same basal cGMP level. In the presence of MANP, cells that expressed active pGC-A generated a higher cGMP level when compared with non-transfected Sf9 cells. There was a statistically significant difference between cells expressing intact pGC-A and non-transfected cells when cells were incubated with 10^5 and 10^6 picomole MANP. However, on increasing the MANP concentration to 10^7 picomole, the level of cGMP in both cells was lower than in the control group (basal cGMP level). This may be caused by toxicity induced by high-dose MANP, which could become toxic to the cells. In conclusion, the whole-cell activity assay showed that the expressed intact pGC-A was functionally active in the Sf9 cell membrane.

Purified intact pGC-A tetramer was functionally active in in vitro activity assay

Intact pGC-A is a membrane protein receptor. The ECD of pGC-A interacts with ANP and triggers the intracellular GCD to convert GTP to cGMP. In earlier studies, results have shown that ATP binds to the intracellular PKLD region and plays a regulatory role in the activation pGC-A GCD (Kurose, Inagami, and Ui 1987, Jewett et al. 1993, Chinkers, Singh, and Garbers 1991). Subsequently, more detailed studies of the regulation of pGC-A by ATP have proposed that the ATP dependent ANP-pGC-A-cGMP signal transduction mechanism may be a two-step ordered process (Duda et al. 2000): First, activation of the ATP-dependent allosteric step, partially activates the GCD through binding of ATP to the ATP regulatory module domain; second, the phosphorylation of the cyclase fully activates the guanylyl cyclase (Sharma 2010, Duda et al. 2000). However, this hypothesis is still under discussion as another group suggested a two processes signal transduction model in the opposite order: they proposed that the phosphorylation of the ATP-binding domain would occur and must take place before activating the ATP-dependent allosteric step (Foster and Garbers 1998).

To determine if purified different intact pGC-A oligomers were functional and ATP was obligated for GCD activation, two Superose6 peak fractions, which represented tetramer and monomer sizes, were collected separately and slightly concentrated for competitive cGMP ELISA detection. The same protein concentration of tetramer and monomer samples was established, and the sample were each split, creating two monomer and two tetramer samples that were incubated with (Figure 3.4.7 A) and without ATP (Figure 3.4.7 B). The ELISA data were analyzed by GraphPad Prism9 software, and the 4-parameter logistic curve fitting program (VØlund 1978) was used to

generate the standard cGMP concentration curve based on the cGMP percent bound (B/B0) (Figure 3.4.7 C). Based on the standard cGMP curve, unknown cGMP concentrations from different samples were calculated. All raw data points were first analyzed via the ROUT method (Q=1%) to remove significantly impact outlier values, which substantially duffer from the rest of data in nonlinear regression analysis. The Q value is the maximum desired setting for the dales discovery rate. Q=1% represents no more than 1% of the identified outliers will be false (Motulsky and Brown 2006). In this experiment, one outlier value (9.4×10^4 pmol/mL) from the tetramer ATP incubation group was removed before analyzing pGC-A cGMP functional assay data (Figure 3.4.7 A, B &D).

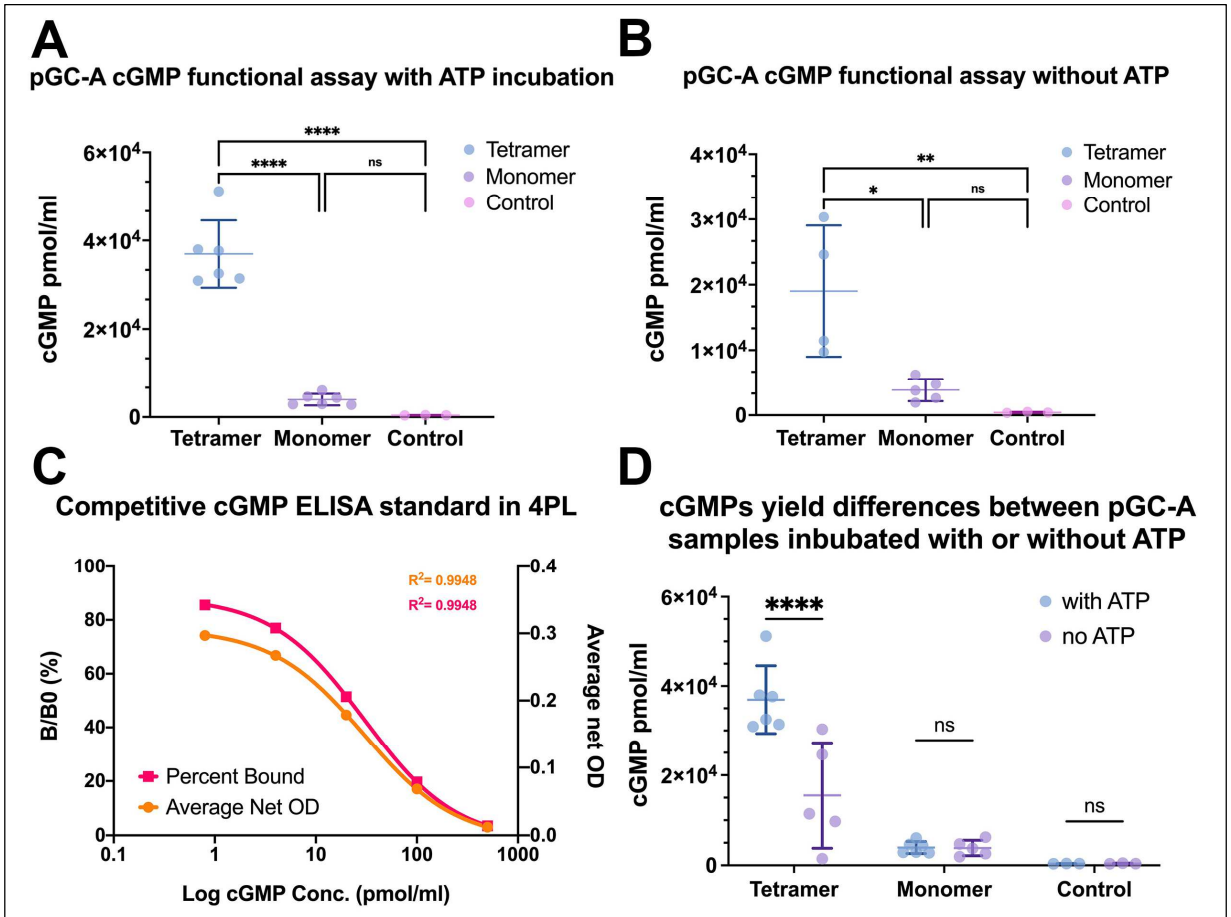

Figure 3.4.7 Purified intact pGC-A *in vitro* functional activity test. A) Analyzing functional activity for two pGC-A oligomeric states with ATP incubation. The control group was GTP and ATP in sample buffer. B) Analyzing functional activity for two pGC-A oligomeric states without ATP incubation. The control group was GTP only in the sample buffer. C) Competitive cGMP ELISA standard fit in four parameter logistic (4PL) curve. The left Y-axis is the B/B0 (%) value and represents the percentage of bound cGMP. The right Y-axis represents the average net optical density (OD) reading at 405 nm. Both standard curves were generated with a 95% confidence interval. D) The cGMP yield differences between pGC-A oligomer samples incubated with or without ATP were analyzed. All raw data points were analyzed via the ROUT method (Q=1%) to remove significantly impossible outlier values before data analysis. One-way ANOVA was used to determine the statistical significance between samples and control in graphs A and B. Two-way ANOVA was used to determine the statistical significance in graph D. Each dot represented to the sample point and plotted as mean ± standard deviation (SD). * $P \leq 0.05$, ** $P \leq 0.01$, **** $P \leq 0.0001$ and ns $P \geq 0.05$, Not significant.

In Figure 3.4.7 A, the competitive cGMP ELISA assay was used to analyze the cGMP yield from purified pGC-A samples with ATP incubation. The results revealed that the tetramer sample had a significant cGMP higher yield when compared with the control without protein addition and the monomer sample. Additionally, there was no significant cGMP yield difference between the monomer sample and the control group. In Figure 3.4.7 B, the competitive cGMP ELISA assay analyzed the cGMP yields from purified pGC-A samples without pre-ATP incubation. The results in Figure 3.4.7 B results are very interesting as the show that the isolated pGC-A receptor can convert GTP into cGMP without prior activation by ATP. The relative activity of tetramer and monomer is similar to the results with ATP activation depicted in Figure 3.4.7 A. The major difference was that the cGMP yield in the tetramer sample was two times higher with ATP activation compared to the data without ATP activation. (Compare Figure 3.4.7A and B).

Next, to determine whether ATP-regulated pGC-A guanylyl cyclase activity was still compatible with purified intact pGC-A in micelle solution, as previously reported (Duda et al. 2000), the cGMP yield differences between pGC-A oligomer samples incubated with or without ATP were further analyzed (Figure 3.4.7 D). The two-way ANOVA results showed a significant cGMP yield difference between tetramer size samples with or without ATP incubation, consistent with previous reports (Duda et al. 2000). Additionally, there were no significant differences in monomer sample and the control group with or without ATP incubation. These significant findings corroborated the previous one-way ANOVA analysis, which indicated that the monomer sample had only minimal activity. My conclusion is that the very low activity detected may be caused

by small amounts of tetramer may have re-formed in the sample. Thereby the likely conclusion is that the monomer is not functional (Figure 3.4.7 A& B). In conclusion, the functional assay showed for the first time that purified intact pGC-A tetramer was functional and confirmed that ATP plays a vital role in activating detergent purified pGC-A *in vitro*.

3.5 Discussion

According to the most current hypothesis pGC-A was assumed to be a homodimer (Rukavina Mikusic et al. 2018), which was based on the crystal structures of rat pGC-A ECD solved with and without ANP substrate. ANP binds to two ECDs by forming hydrogen bonds, interacting with a hydrophobic pocket in the ECD and partially forming a short parallel ß-sheet with the juxta-membrane site of one ECD at the C-terminal segment of ANP (Ogawa et al. 2004, Misono et al. 2011). Subsequently, a rotation mechanism homodimer model was proposed for the function of the intact full-length receptor (Misono et al. 2005, Misono et al. 2011). Indeed, the known structures of two heterologously expressed GCD were crystallized in dimer form (Winger et al. 2008, Rauch et al. 2008). Each GCD monomer has twofold symmetry. There are two symmetric guanylyl cyclase active sites in the central cavity (Misono et al. 2011). The proposed intact pGC-A homodimer state secures and meets the requirements for ligand binding and guanylyl cyclase functionality (Figure 1.4 C&D).

My results suggest that intact pGC-A is a tetramer in its native state, which is supported by early results from research groups who also worked with the intact pGC-A that had shown different oligomeric states without ligand presence (Wilson and Chinkers 1995). One truncation experiment that expressed intact human pGC-A in HEK-293 cells

has shown that the ICD of pGC-A was necessary for self-association, and in their studies the higher oligomer formation was ligand-independent (Lowe 1992). In another study, the partially purified intact pGC-A from bovine adrenal cortex tissue proposed that the pGC-A may exist as a disulfide-linked tetramer in its native state (Iwata et al. 1991). However, my research findings did not corroborate the disulfide-linked concept. Different pGC-A oligomeric state peak fractions were incubated with or without 1 M DTT (final 167 mM DTT in the samples) overnight in my study. There was no molecular weight shift on the native gel when comparing the protein samples with and without DTT incubation (Figure 3.4.5). In their partial pGC-A purification protocol, 1 mM DTT was added to prevent formation of artificial disulfide bonds (Iwata et al. 1991). In my routine intact pGC-A purification, 5 mM DTT was freshly added to each step.

For the analysis by SDS electrophoresis, purified pGC-A was mixed with reducing agent and after overnight incubation the SDS sample buffer was added, and the samples were heated at 95ºC for 10 min before gel electrophoresis analysis. Higher oligomer state signals were shown in anti-pGC-A western blots (Figure 2.4.4 A and Figure 3.4.4 B) with and without the reducing agent DTT.

Furthermore, my results suggested that high detergent concentration partially disassemble the tetramer into monomers, but that the higher protein concentration promoted pGC-A monomers to be re-oligomerized into the functional active tetramer. (Figure 3.4.2 C and Figure 3.4.3).

Overall, the findings in the present study suggested they oligomerization into the functional active tetramer is concentration dependent and that very likely noncovalent interactions (like hydrophobic or electrostatic interactions as well as H-binds) facilitate the

tetramer formation, The disulfide interaction (if they exist) may not be important for tetramer formation. Additionally, one previous literature study shared a similar outcome.

The previous cross-linked immunoblot detection study showed that not all the available NPR-A was covalently aggregated (Lowe 1992). The formation of disulfide links was inefficient on pGC-A oligomerization, and most of the intact pGC-A was in a noncovalently associated state (Lowe 1992). Furthermore, another previous report stated that their FPLC chromatography showed intact rat pGC-A without ligand eluted at the tetrameric size, but overnight density gradient centrifugation revealed all migrated pGC-A as monomers (Chinkers and Wilson 1992). The tetramer to monomer dissociation observed was consistent with the idea that pGC-A monomers may noncovalently oligomerize to tetramers.

The possibility of an artificial disulfide bond would form a higher pGC-A oligomer was ruled out. Is it possible that heterologous overexpression of intact pGC-A caused ligand-independent oligomerization? Herein, the intact pGC-A was expressed in insect cells. The protein expression level was extremely low. One liter of transfected Sf9 cells only yielded less than 0.05 mg purified pGC-A. In a similar previous report, rat intact pGC-A in baculovirus Sf9 insect cell showed similar modest protein production (Chinkers, Singh, and Garbers 1991). Another research study has expressed intact rat pGC-A on COS cells (fibroblast-like cells derived from monkey kidney tissue), and they did not favor the concept of overexpression-induced oligomerization in the cell membrane (Chinkers and Wilson 1992). In addition, a previously mentioned study that partially purified intact pGC-A from bovine adrenal cortex tissue has shown that the pGC-A exists as a tetramer (Iwata et al. 1991). Collectively, the conclusion can be drawn that

tetramer configuration is the active and native form of pGC-A and is not an artificial product from heterologous expression.

Is it also possible that my intact pGC-A purification process artificially triggered the tetramer by dimerization of homodimers? My earlier results suggested that high protein concentration triggers tetramer formation (Figure 3.4.2 C and Figure 3.4.3). If the dimer would be the native state, dimer peaks should be as frequent as monomers observed in gel filtration. During purification, intact pGC-A was isolated from the cell membrane and embedded in detergent micelles. Next, to rule out the possibility that detergent caused oligomerization, different concentrations of DDM with CHS were screened (Figure 3.4.4). In the 0.02% DDM+0.004% CHS concentration elution profile, a very small potential dimer peak (16.14 mL) was observed, but an even much higher tetramer oligomer size peak (14.07 mL) was formed. No monomer was formed under these conditions. The results suggested that 0.02% DDM+0.004% CHS concentration may be insufficient for intact pGC-A purification and may tend to aggregation. Conversely, the detergent concentration at 0.05% DDM+0.01% CHS and 0.1% DDM+0.02% CHS showed a remarkably similar elution profile. A majority monomer size (17.3 mL) peak was observed, with a minor tetramer size shoulder (15.0 mL) (Figure 3.4.4). The similar elution profiles suggested that 0.05% DDM+0.01% CHS was sufficient detergent concentration for intact pGC-A purification.

On the other hand, higher detergent concentrations that may causing partial disassembly of the tetramers into monomers. The detergent concentration at 0.05% DDM+0.01% CHS and 0.1% DDM+0.02% CHS showed elution profiles, with the majority monomer size (17.3 mL) peak, and only a minor tetramer size shoulder (15.0

mL) (Figure 3.4.4). The mainly monomeric elution profiles suggested that 0.05% DDM+0.01% CHS may be excessive and partially disassembled the intact tetramer into monomers.

Originally, I was postulated that a column resolution limit might have caused the "missing" dimer peak. The dimer peak may be covered by overlapped tetramer and monomer peaks. However, after analyzing protein elution peaks based on oligomer sizes in native gel, most protein signals were still from tetramer and monomer peaks, with no obvious dimer signal (Figure 3.4.5). In conclusion, a dimer oligomer state may not be typically favored. While I used the high detergent concentration of 0.05% for most of my isolations, the DLS results showed that the tetramer re-forms when the protein is concentrated, therefore I can safely assume that the crystallization experiments described in the next chapter contain the active, tetrameric PGC-A receptor. In future studies, it may be advantageous to reduce the detergent concentration and test different detergent concentrations between 0.02% to 0.05% DDM.

I next addressed how pGC-A may form a tetramer in its native state among different cell sources. There was a possibility that the intracellular domain (ICD) of pGC-A may interact with each other to form a higher oligomer. One study has reported that the ICDs played an essential role in pGC-A self-association (Lowe 1992). Based on the previous discussion, the ICD may trigger oligomerization through noncovalent interactions such as hydrogen bonds and hydrophobic interactions. Protein sequence BLAST comparison showed that human (UniProtKB/Swiss-Prot: P16066.1) and bovine (NP_001179680.1) intact pGC-A are 92.93% identical. There are 16 and 17 cysteine residues in the bovine and intact pGC-A respectively. The tetramer in bovine could

possibly be disulfide-linked, but disulfide bonds may not be the primary molecular force in the human pGC-A oligomerization.

Generally, higher oligomerization is common in membrane-protein receptors. For example, many GPCRs have been shown to form a dimer or higher oligomers in the cell membrane. Fluorescence labeled single-molecule analysis suggested that GPCRs on the cell surface constantly self-associate and dissociate in a dynamic equilibrium manner (Calebiro et al. 2013). It is reasonable to assume that the tetramer state of pGC-A represents the native state of pGC-A.

I acknowledge that at least dimer formation is required for pGC-A functionality. Intact pGC-A may functionally work as a homodimer, but the homodimer could be just a partial of the tetramer complex. In other words, one tetramer complex may be composed of two functional units (homodimer). The report on partially purified pGC-A from bovine made an interesting point: A higher oligomer structure may improve the local concentration of the receptor to increase the ligand-binding sensitivity (Iwata et al. 1991). I agree that the advantage of the tetramer complex could increase the local pGC-A concentration. In addition, the crystal structure of ANP-rat pGC-A ECD complex showed binding of ANP is asymmetric between ECD monomer1 and ECD monomer2 (Ogawa et al. 2004). These results suggested that one ligand has two binding possibilities in one potential functional unit (homodimer). Each of the four copies of identical pGC-A monomers provides a total of four random combination possibilities to form a tetramer complex composed of 2 homodimers. This assortment provides different possibilities of ligand binding sites. Furthermore, the released ligand from one functional subunit (homodimer) may recycle to the second pair in the tetramer complex. There could also be

cooperativity between the ligand binding sites in the ECDs of and the GDCs of the tetramer. Cooperativity and recycling of ligands may further increase pGC-A binding efficiency and ligand sensitivity. All three natriuretic peptides are degraded either via clearance natriuretic peptide receptor C or by neprilysin enzymatic degradation (Potter, Abbey-Hosch, and Dickey 2006, Meems and Burnett 2016). Neprilysin is a membrane protein expressed with various tissue distributions. The ANP ligand has a much higher degradation rate than other natriuretic peptides (Potter, Abbey-Hosch, and Dickey 2006). ANP has a fast clearance rate in humans and the half-time ranges between 2 to 2.5 min (Pandey 2011). Furthermore, the concentration of ANP in plasma is in picomolar range (Schiffrin 1990). Forming a tetramer complex in the native state may become essential to sense the ligand in the body ensuring for maximal protein functionality and signal transduction.

The cGMP functional assay results revealed that the purified tetramer in detergent micelle solution was functionally active (Figure 3.4.7 A&B). It also indirectly confirmed that the tetramer formation was not caused by artificial protein aggregation during protein purification but is the native functional state of the receptor. Although the monomer samples showed slightly higher cGMP yield than the control group, the cGMP yield difference between monomer and control group was not statistically significant (Figure 3.4.7 A&B). Based on the cGMP yield results from the monomer, it can be indirectly suggested that the tetramerization of pGC-A is essential for functional activity and is a ligand-independent event. This result was consistent with previous research, showing that pGC-A oligomerization of pGC-A occurred in a ligand-independent fashion (Chinkers and Wilson 1992). The two-way ANOVA analysis showed a significant cGMP yield

difference between the tetramer sample with or without ATP incubation, which is in line with earlier reports that ATP plays an important regulatory role in activating pGC-A (Duda et al. 2000, Burczynska, Duda, and Sharma 2007).

The current pGC-A signal transduction mechanism proposes an ATP-dependent two-step ordered process. First, ATP binds to the ATP-binding pocket to activate the ATP-dependent allosteric step, which partially activates the GCD and triggers a conformational change; second, the phosphorylation of the potential ATP-binding domain fully activates guanylyl cyclase (Sharma 2010, Duda et al. 2000). However, in my cGMP functional assay, purified tetramer size protein still showed significant cGMP yield without ATP incubation (Figure 3.4.7 B), but the cGMP yield was increased dramatically with ATP incubation (Figure 3.4.7 A). My findings suggested that ATP may not be critical for initiating pGC-A activation, as suggested by previous reports (Chinkers, Singh, and Garbers 1991). Binding the ligand will moderately activate guanylyl cyclase function. One research article has revealed that activation of the guanylyl cyclase was ATP-independent (Antos et al. 2005). My functional assay results were consistent with previous research and showed that ATP accelerated the guanylyl cyclase activity (Burczynska, Duda, and Sharma 2007). However, my functional assay results (Figure 3.4.7 B) also indicated that the ligand itself will activate pGC-A even without ATP (Antos et al. 2005). Collectively, my study suggested that three modulators modulate different activation levels in intact pGC-A. The first level is ligand binding to the ECD, which moderately activates the pGC-A; the second level is binding ATP to the ATP-binding pocket at the ICD, which acts as an augment modulator and boosts protein

activity; the third level is pGC-A phosphorylation, which fully activates the guanylyl cyclase functionality.

3.6 Conclusion

In summary, I analyzed different pGC-A oligomeric states observed in Superose6 chromatography. The results showed that the higher oligomeric state represented the tetramer. As my finding conflicted with the current understanding that pGC-A is a homodimer, different oligomeric peak fractions eluted from the Superose6 column were carefully characterized. Surprisingly, my results suggested protein concentration played a stimulation role in converting pGC-A monomers into a tetramer. In order to rule out the possibility that the tetramer was an artificial product due to detergent concentration, different detergent concentration screenings were performed. The results revealed that a standard 0.05% DDM with 0.01% CHS detergent concentration may or may not destabilize the tetramer, however it is re-formed when the protein is concentrated. The DTT reducing reagent incubation experiment showed that the tetramer formation was not disulfide bond associated. The results suggested that other noncovalent interactions could possibly oligomerize monomers to tetramers. In the high resolution clear native gel electrophoresis analysis study, different oligomeric peak samples were analyzed, where the dimer rarely appeared. The results suggested that the dimer is not a favored oligomerization state and the "missing" dimer peak was not due to column resolution limit or overlapped tetramer and monomer peaks. Lastly, the cGMP functional assay showed that the pGC-A tetramer was functionally active, which indirectly rejected the theory that the tetramer protein aggregation. Tetramer and monomer samples were incubated with or without ATP during the cGMP functional assay. The monomer sample results showed insignificant activity,

whereas the tetramer is highly active which indicates that the tetramerization of pGC-A is a ligand-independent event. In addition, the tetramer sample results optimized the current two-step ordered mechanism for ATP-dependent pGC-A signal transduction. My study suggests that three modulators regulate different activation levels in intact pGC-A. The first level is the binding of the ligand to the ECD, which moderately activates the pGC-A; the second level is ATP binding to the ATP-binding pocket at the ICD, which acts as an augment modulator and boosts protein activity; the third level may involve pGC-A phosphorylation, which could then fully activate the guanylyl cyclase functionality. More studies including mass spectroscopy has to be performed in the future to detect phosphorylation and follow its time course and influence on the catalytic activity.

CHAPTER 4

CRYSTALLIZATION AND INITIAL DIFFRACTION OF INTACT HUMAN NATRIURETIC PEPTIDE RECEPTOR A

4.1 Abstract

Natriuretic peptide receptor A is also known as particulate guanylyl cyclase receptor A (pGC-A) because of guanylyl cyclase domain. pGC-A is an ANP receptor critical to control cardiovascular, renal, and endocrine homeostatic functions (Pandey 2011). To effectively develop a method to regulate pGC-A, structural information regarding its intact form is required. Currently, only an ECD of ligand-bound and ligand-free rat pGC-A structure is available (Ogawa et al. 2004, Ogawa, Kodama, and Izumikawa 2020, van den Akker et al. 2000). However, the structure of the transmembrane domain and functional intracellular domain regions need to be determined to understand the function and activation mechanisms of the receptor. In Chapter 2, I described how the intact human pGC-A was successfully expressed and purified from the insect host cell. Regarding the structure of pGC-A, Chapter 4 primarily focuses on screening purified intact pGC-A with commercial crystallization kits to select potential crystallization conditions and optimizations. Furthermore, verification of crystals formed by the target protein is an essential step before analyzing crystals for diffraction and structural determination. Figure 2.1 illustrates the schematic for the ideal major experimental steps from plasmid design to structure determination of intact human pGC-A. This Chapter focuses on protein crystallization and crystal diffraction of the intact human pGC-A receptor.

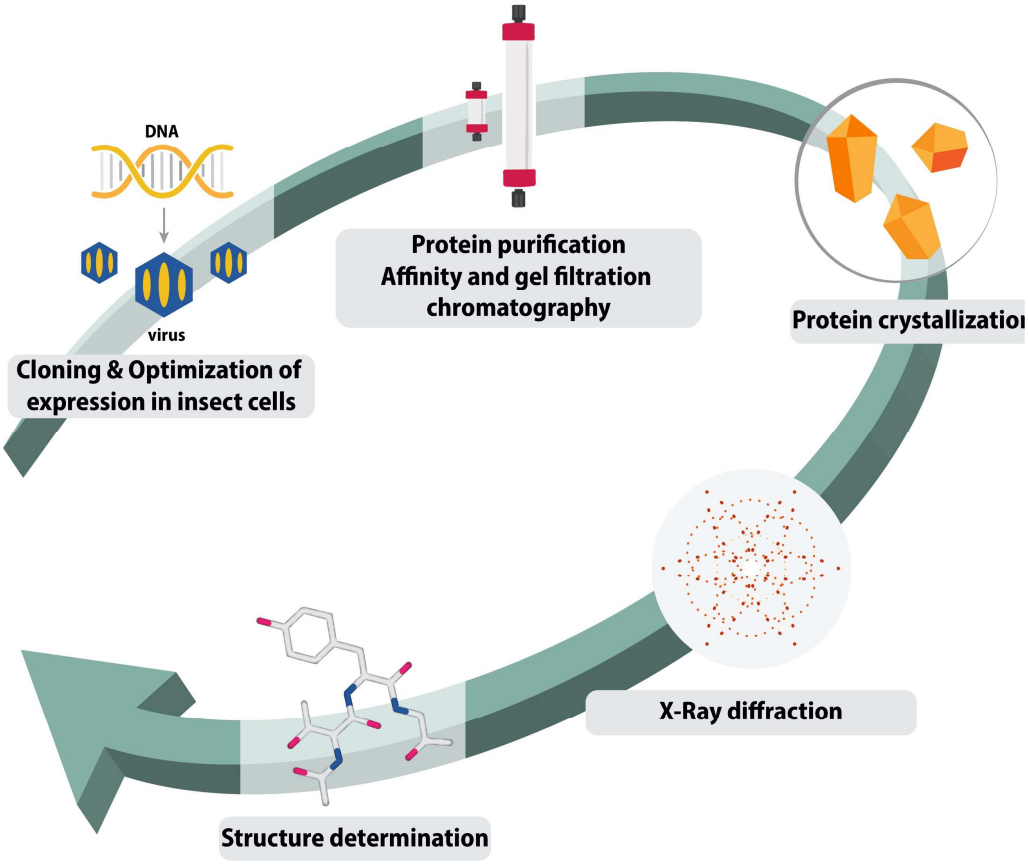

Figure 2.1 Summary scheme of ideal major experiment steps from cloning to structure determination of intact human pGC-A.

4.2 Introduction

4.2.1 Overview of Protein Crystallization

The first protein crystallization was reported in 1840. Friedrich Ludwig Hünefeld accidentally discovered that hemoglobin formed small plate-like crystalline material in a blood sample (Hünefeld 1840). Years later, more diverse proteins were crystallized, often serendipitously during other experiments. From the 19th century to the first half of the 20th century, crystallization served as a practical purification step or characterization method in many protocols. Furthermore, important factors, such as salt concentration, organic solvent,

solution pH, and temperature, were noted to affect protein crystallization (Giege 2013). The application of crystallization expanded to structural biology after the first pepsin protein crystal diffraction pattern was revealed in 1934 (Bernal and Crowfoot 1934). Combined with the development of improved protein crystallization methodologies and advanced X-ray crystallography, the first protein structure was reported in 1958 for myoglobin (Kendrew et al. 1958). Subsequently, numerous crystallization techniques have been developed, and the principles of crystallization studied. Protein crystallization may appear to be a straightforward technique; however, bottlenecks exist in most cases, especially when well-ordered crystals are needed for structural determination by X-ray crystallography. Before cryo-EM technology made a large leap in improving resolution (Earl et al. 2017), macromolecule crystallization was a crucial requirement to obtain three-dimensional structural information. In addition to X-ray crystallography and cryo-EM technology, nuclear magnetic resonance (NMR) is also used for high resolution protein structure determination. NMR allows for targeted biomolecules to stay in solution at physiological temperature, which preserves the native state of the biomolecules. Also, NMR has the unique ability to observe and record molecular dynamic motions within the structure (Takeuchi, Baskaran, and Arthanari 2019). Unfortunately, the biomolecule size limitation is still the major challenge in this field and it is difficult to determine a protein structure larger than 50 kDa in size (Gauto et al. 2019).

In theory, the key to macromolecule crystallization is to obtain a protein solution in a supersaturated state. Then, individual factors that impact solubility, precipitation, and crystal formation are systematically examined and adjusted. The goal of initial screening is to determine the conditions that form small crystallites or nano/ microcrystalline

precipitates, or that lead to phase separation, with the latter being a state from which membrane protein crystals can grow. Then, based on the initial conditions, each crystallization factor can be optimized to obtain crystals during extensive screening (McPherson 2009). However, each macromolecule has a complicated physical and chemical system (phase diagram), and altering any factor affects the function of the environment (i.e., other factors). For example, solubility as a function of pH may also be influenced by a change in ionic strength, temperature, and the concentration and nature of the precipitate. The protein may also become denatured under certain conditions. Generally, macromolecules, such as proteins, are difficult to purify, especially to obtain 100% purity. Even with 100% purity, the protein could contain various conformational states, which could hinder the formation of ordered crystals. Protein size, shape, and self-complexity also affect the ability to form intermolecular contacts. Compared to small molecules, protein crystals have fewer protein-protein contact points and weak binding energy. As mentioned earlier, diverse factors can influence protein crystallization. For example, slightly altered buffer pH, salt concentration, and saturation level can modulate intermolecular contact energy. Notably, crystallization conditions are unique to each protein (Durbin and Feher 1996). Screening for potential crystallization conditions is a complicated and time-consuming process. Obtaining sufficient protein crystal size with high-quality diffraction is another bottleneck in X-ray crystallography (Schonherr, Rudolph, and Redecke 2018). Among protein species, membrane proteins play an essential role in the biology and physiology of all organisms. Approximately 20–30% of coding sequences of genomes code for transmembrane proteins (Krogh et al. 2001). However, it is difficult to obtain membrane protein crystal with high-quality diffraction for X-ray

crystallography because of low-protein yield, hydrophobic properties, and instability following membrane extraction. Furthermore, they contain lipids and/or detergent micelles that make it difficult to tightly pack membrane proteins into highly ordered crystals (Kermani 2020, Birch et al. 2018).

4.2.2 Molecular Interactions in Crystallization

The protein crystallization process can be illustrated by a crystallization phase diagram (Figure 4.2.1). This diagram shows that crystallization is a process where a supersaturated environment is gradually created while protein solubility declines to a point where the protein molecules spontaneously form nuclei in the nucleation zone. Once stable nuclei are formed, the surrounding protein molecules are continuedly and spontaneously bound to the surface of the developing crystal. This is a dynamic process where more molecules bind than dissociate from the growing crystal. The protein crystallization phase diagram contains four protein solubility zones: 1) precipitation zone: here the protein solution is in a highly supersaturated state, leading to non-ordered protein precipitation; 2) nucleation zone: in this zone, supersaturation is moderate and triggers protein crystal nucleation; 3) metastable zone: this zone features a slightly lower supersaturation level than the nucleation zone; supersaturation is not high enough to form stable nuclei but seed crystals larger than the critical radius grow into a single crystal and this facilitates protein crystal growth but no new nuclei are formed; 4) undersaturation zone: in this zone the protein solution is non-saturated, i.e., the protein is soluble, and no crystallization occurs. The phase diagram is an essential tool to optimize the initial screening of crystallization conditions, and it can be “personalized” for the POI (Chayen and Saridakis 2008). Generally, crystals formed by macromolecules are composed of an average of 50% solvent,

although this number may vary between 25 and 90%, depending on the arrangement of the protein molecules in the crystal (Matthews 1968, McPherson 1999). There are far fewer intermolecular contact points in macromolecular crystals than in small molecules. In crystallization, these contacts establish the crystal lattice interaction and maintain crystal integrity. Because of this property, protein crystals are typically limited in size and have a soft texture. Large solvent cavities in the crystal can also limit the resolution of crystal diffraction because of flexibility in the packing caused by large spaces between adjacent molecules, slightly different molecular positions, and weak lattice forces (McPherson 2009).

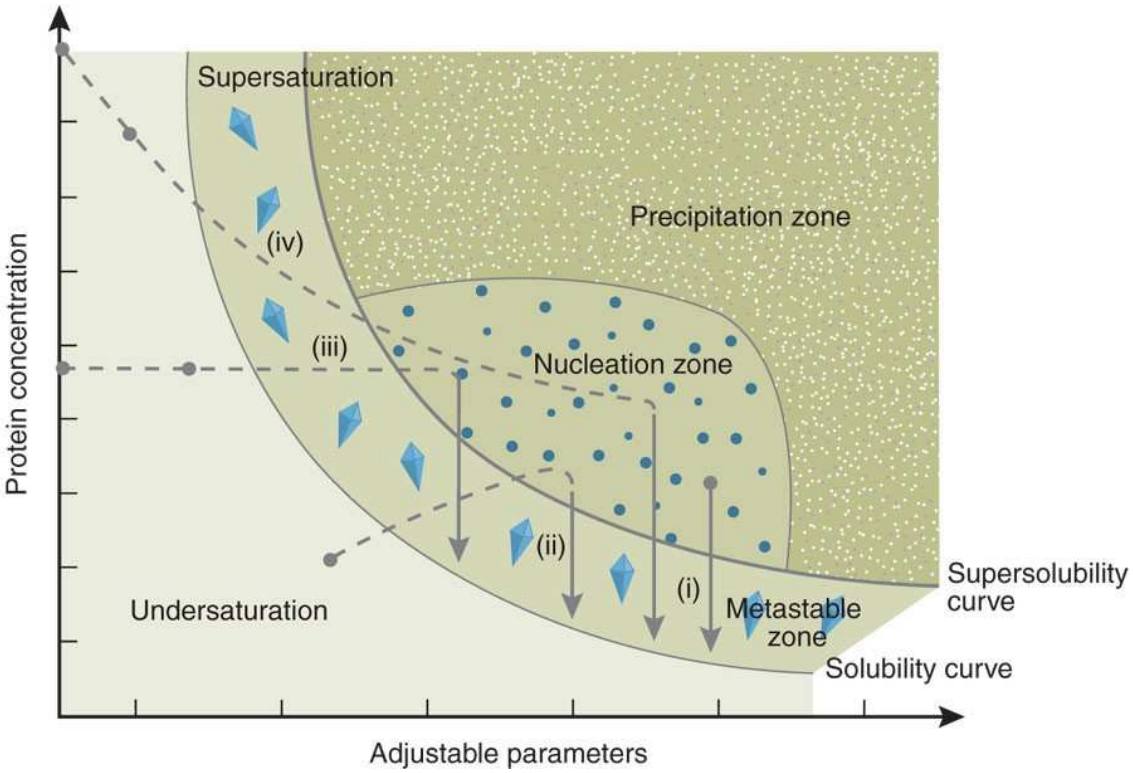

Figure 4.2.1 Summary illustration of the protein crystallization phase diagram. The x-axis represents the adjustable parameters, including different concentrations of the crystallization precipitants. The y-axis represents the protein concentration. There are four different crystallization methods represented in this crystallization phase diagram: (i) microbatch method, (ii) vapor diffusion method, (iii) dialysis method, and (iv) free-interface diffusion method. Each crystallization method has a different path to the nucleation and metastable zones. The exact pathway is shown with either straight or dashed gray lines. The solubility curve represents the concentration of the protein in the solute that is in equilibrium with crystals. An undersaturated solute will dissolve the formed protein crystals to reach protein concentration equilibrium. The supersolubility curve represents the boundary between the metastable zone (protein crystal growth) and spontaneous nucleation zone (protein crystal nucleation) or precipitation zone. Figure, caption, and description are modified from (Chayen and Saridakis 2008) with permission from Springer Nature.

The crystallization process is driven by the thermodynamics of protein-protein interactions under specific solvent conditions. Similar to the interactions that drive the folding of proteins, protein-protein interactions are driven by the difference in free energy between molecules in contact with each other in the crystal and molecules in solution separated by their hydration shell and the surrounding solution (Durbin and Feher 1996).

Compared with the interactions in protein-protein geometric arrangement (e.g., degree of specificity and strength), the non-specific intermolecular forces (e.g., dipole-dipole, Van der Waals dispersion) are less important (Durbin and Feher 1996, Janin 1995). The number and geometry of contacts are essential to crystal formation. If two binding sites are available in the unit cell, then only one-dimensional crystals are formed (circular, helical, or straight). If three binding sites are present in the unit cell, then two-dimensional crystals form (planar, cylindrical, or spherical). In the case of four protein contact sites in the unit cell, three-dimensional crystals form. In each unit cell, most three-dimensional crystals have more than four contacts (Durbin and Feher 1996). The greater the number of contacts and more rigid the packing, the greater is the potential order in the crystal. X-ray resolution may be improved by reducing the mobility of the unit cell (Sheriff et al. 1985, Phillips Jr 1990). There are specific interactions between any two molecules that affect crystal packing. For example, direct and indirect hydrogen bonds and electrostatic interactions raised by divalent metal bonds are common interactions at protein contact sites in crystals. Solvent-derived interactions could occur between any two molecules and may also play an essential role in crystallization. To adjust the solvent-derived interactions, the concentrations of salt and the precipitant (for example PEG) molecule are generally modified. Ions raise electrostatic interactions and high salt concentrations (i.e., high ionic strength) enhance hydrophobic interactions (Melander and Horváth 1977) by competing with the protein for the hydrating water/solvent molecules. PEG acts as a space-filling background molecule in macromolecular crystallization and introduces the excluded volume to stimulate attraction between two molecules, favoring protein association (Middaugh et al. 1979). In summary, the molecular interactions in crystallization are

influenced by a delicate balance of interactions between protein molecules, the solvent, and the solutes in the solvent. Crystallization is highly dependent on the protein molecular self-characterization and specific solvent composition.

4.2.3 Methods for Protein Crystallization and Optimization

Proteins form crystals when they are exposed to an environment that gradually lowers their solubility. In most common crystallization methods, salts, PEG, and organic solvents are added to the protein solution and they act as precipitation agents. Precipitates reduce protein solubility and create a supersaturated solution. The initiation of crystallization occurs when a given radius threshold (called the critical radius of a nucleus) is attained. Then, proteins interact with neighboring protein molecules to form stable nuclei. When the change in supersaturation is too rapid, the proteins do not have time to pack in an ordered form, and they precipitate out of the solution in the form of an amorphous precipitate. Ideally, the process in which supersaturation is reached is slower and controlled, and an intermediate level of supersaturation allows the proteins to form in ordered arrays, thereby growing into crystals based on the first formed crystalline nucleus (Gabrielsen et al. 2009). Despite several efforts to grow protein crystals with high throughput, many of the first crystals detected are typically substandard for data collection. Optimization of crystal size and quality is a crucial step in structure determination by standard X-ray crystallography. Slight modifications in crystal growth conditions, including pH, protein concentration, and temperature, often lead to improvements in crystal growth. It should be noted that new X-ray diffraction techniques with X-ray Free Electron lasers allow data collection from a stream of nano or micron-sized crystals and therefore do not require the growth of large macroscopic single crystals

(Chapman et al. 2011, Boutet et al. 2012, Gisriel et al. 2020). The packing of the molecules in the unit cell is still critical for serial fs crystallography.

Vapor diffusion is the most commonly used method for the crystallization of proteins. It can be sub-classified into hanging drop and sitting drop methods depending on the crystallization droplet location (Ducruix and Giegé 1999). The plate reservoir wells are filled with a large volume of crystallization precipitate solution (200–1000 μL). Then, a small amount of protein solution is mixed with the precipitate solution to form a crystallization droplet (1–10 μL for manual plate setup, 10 nL to 100 nL for a setup using a crystallization robot) and placed either in the central well (sitting drop) or on a siliconized coverslip (hanging drop). The droplet and reservoir are in a sealed system. The crystallization plate is incubated at a stable temperature. During the incubation time, the water concentration in the crystallization droplet and reservoir slowly reaches an equilibrium via vapor diffusion. The reservoir solution contains a higher concentration of the precipitant than the crystallization droplet. The osmolarity of the crystallization droplet asymptotically reaches the same osmolarity as the reservoir solution because water evaporates from the droplet into the reservoir. Consequently, the concentrations of protein, solutes, and precipitants simultaneously increase in the drop. Ideally, the crystallization drop will reach the supersaturation state and initiate protein crystallization (McPherson 2017). To adjust the speed at which equilibrium is reached, the incubation temperature can be changed, or silicon oil can be added to the surface of the reservoir well (Gabrielsen et al. 2009). The advantage of the hanging drop is that the formed crystal may remain at the bottom of the crystallization droplets because of gravity. However, the sitting drop method was popularized because it can be set up via crystallization robots (Hui and Edwards 2003,

Zhu et al. 2014). Vapor diffusion is typically performed in a 24-well plate for manual setup and a 96-well format for automated setups. This allows for slightly modified conditions to be compared and for parallel tests in both initial and optimization crystallization screening to be conducted. However, the vapor diffusion method also has some drawbacks. Crystal size depends on the size of the drops; thus, only very small crystals can be grown in automated crystallization experiments with 10–50 nL drop size. Large crystals can be grown in manual experiments with hanging and sitting drops. The main limitation of vapor diffusion is unspecified crystallization conditions because the protein and solutes are concentrated simultaneously. Consequently, it is impossible to exactly map supersaturation, where the first nuclei occur, and under which conditions they grow in the metastable zone. Therefore, the exact crystallization compositions are not controllable, and this may affect the reproducibility of protein crystallization (Chayen and Saridakis 2008, Rayment 2002b).

Microbatch experiments utilize low-density paraffin oil as the crystallization incubation environment. The mixed crystallization droplets are directly dispensed into the oil-based environment. Given the difference in density, the crystallization droplets sink to the bottom, and the top layer of oil prevents droplet evaporation. Unlike in the vapor diffusion method, the protein sample is directly mixed at the ideal precipitant concentration to achieve the final supersaturation state. No decrease in volume occurs during the incubation period, and the formed crystals do not suffer from undesired effects, such as where crystals first transition into an aggregated phase or dissolution, which can occur in vapor diffusion experiments. However, precipitant solutions that contain small volatile organic solvent molecules that could diffuse through the oil layer may not be suitable for use in the microbatch method (Chayen 1998).

In vivo crystallization was first documented in 1853. Charcot-Leyden crystals (CLCs) were first observed in a leukemia patient. The CLC protein is dominant in human eosinophils and basophils. *In vivo* crystallization of CLCs has been considered the hallmark of eosinophil-associated allergic inflammation (Leonidas et al. 1995). Subsequently, naturally occurring *in vivo* crystallization has been observed in several species, with no preference for a cellular compartment. However, the precise physiological functions of self-crystallization in cells remain unclear (Schonherr, Rudolph, and Redecke 2018). This poorly understood natural crystallization mechanism provides an opportunity to crystallize heterologous POI *in vivo.* This new approach avoids the conventional protein purification and crystallization pathway (Koopmann et al. 2012, Redecke et al. 2013).

Free interface diffusion allows protein crystals to form in a capillary or tube. The protein solution is loaded into the capillary via the capillary effect. Protein in the loaded capillary is overlaid with the precipitant solution. A wide range of precipitant gradients can be established in a single trial by slowly diffusing one solution into the other. Theoretically, this is extremely useful for fine-tuning screening of optimization conditions. Additionally, this methodology is particularly applied in microgravity crystallization studies (McPherson 2017). Free interface diffusion has several variations, such as counter-diffusion; in this variation, a liquid-gel interface between the two solutions can slow the equilibration process, thereby preventing rapid diffusion compared to the traditional liquid-liquid interface (Garcia-Ruiz and Ng 2006).

Crystal Seeding is another optimization method in which protein crystals ("seeds") are transferred into the new protein solution and they act as nuclei to form larger crystals. The added crystal seeds are the basis for the new crystallization droplets.

The seeding method can be used to grow larger and higher-quality crystals without the need to alter growth conditions. Typically, seeding can be sub-classified into micro-seeding and macro-seeding. During micro-seeding, the crystallization nucleation zone is avoided. Small, previously formed crystals are mixed with the precipitant solution and may be crushed into many nanosized or very small micron-sized crystals. These nano/microcrystal suspensions are added to the new crystallization droplets and serve as crystal nuclei. The disadvantage is the inability to control the exact number of seeds added to each new crystallization droplet. Normally, different dilutions of a seed stock suspension are tested to determine the best microseeding conditions. Unlike the micro-seeding method, macro-seeding requires high quality, 5–50 μm sized crystal. The single chosen crystal is washed several times in the precipitant solution to remove impurities before transfer to the new crystallization droplet. Macro-seeding is advantageous because it increases crystal size (Bergfors 2003, McPherson and Gavira 2014).

Second-order non-linear imaging of chiral crystals (SONICC) is an imaging system to distinguish chiral crystals from non-chiral crystals (e.g., salt crystals) and amorphous protein precipitates. There are two imaging methods combined in the SONICC system: the second harmonic generation (SHG) method that detects crystallinity (SONICC) and the ultraviolet two-photon excited fluorescence (UV-TPEF) method that detects proteinaceous samples by Trp fluorescence. SONICC can detect microcrystals as small as 100 nm and distinguish crystal showers from amorphous aggregates (Haupert

and Simpson 2011). In the SHG method, samples are exposed to two infrared fs laser pulses (wavelength 1064 nm). Chiral crystals exhibit a frequency-doubled response and generate an SHG signal at a wavelength of 532 nm. However, in protein crystals with a high symmetry space group, the SONICC signal is diminished or canceled, and these proteins will be "SONICC inactive." Furthermore, some small molecules and certain salt ions (organic buffers) may form non-central symmetric crystals, i.e., chiral crystals UV-TPEF that can be detected with fluorescence from Trp, thereby ruling out potential false-positive results. UV-TPEF is analogous to UV fluorescence and excites amino acids, such as tryptophan. The green light (wavelength 532 nm) has been employed to image samples. The frequency-doubled wavelength of green light is 266 nm, which excites tryptophan in the sample, and the excited fluorescence (350–400 nm) is recorded to create a fluorescence image (Haupert and Simpson 2011).

4.2.4 Membrane Protein Crystallization

In structural biology research, the rate-limiting step in the formation of structured protein from purified protein is often called the crystallization step. If obtaining a sufficient size and well-diffracted protein crystal is the major bottleneck, then acquiring membrane protein crystals with the same quality as the soluble proteins may be impractical. Although detergent-based micelles mimic the lipid environment and solubilize membrane proteins, they make it difficult to manipulate the amphiphilic region in the solution. Frequently, membrane proteins tend to aggregate after leaving the native lipid environment (Caffrey 2003). Vapor diffusion is one of the most wildly used methods in protein crystallization. However, mixed crystallization droplets often form phase separation, where the detergent micelles are excluded from the aqueous phase and form detergent-rich oily drops in

solution, which are often visible on the surface of the crystallization cover slides (Gabrielsen et al. 2009). Crystals of membrane protein often grow inside the phase-separated drops or at the droplet–water interface. The addition of specific amphiphilic molecules can boost membrane protein crystallization under the phase separation condition or in the detergent-rich phase (Garavito and Rosenbusch 1980, Yurkova, Demin, and Abdulaev 1990). In some cases, the detergent micelle structure cannot fully replace the lipid environment, and certain lipid compositions are required to interact and stabilize the target membrane proteins. Required lipids can be added to the vapor diffusion droplets as additives, creating a protein-detergent-lipid complex during crystallization (Zhang et al. 2003). There are several commercially available screening kits for membrane protein crystallization. These screening kits are designed based on the most successful crystallization conditions that form high-resolution crystal structures for both α-helical and β-barrel transmembrane proteins. The purpose of using commercial kits is to initially screen potential crystallization conditions. Based on the chosen conditions, multi-optimization steps may lead to the formation of well-diffracting crystals (Kermani 2020).

Lipid mesophase crystallization has been developed to overcome the difficulty in membrane protein crystallization; monoolein is the most commonly used monoacylglycerol (Caffrey et al. 2009). Lipids have the inherent ability to form a highly ordered liquid crystalline state called the mesophase. The amphiphilic nature of lipid molecules induces spontaneous self-assembly and forms the mesophase (Caffrey 2003). When water is mixed with lipids, the self-assembly and phase separation of the lipids result in the lipidic cubic phase (LCP). The cubic phase contains three-dimensional lipid bilayers and interconnecting water channels (Figure 4.2.2). To reconstitute membrane proteins into the lipidic cubic

phase, a 40% protein solution is mixed with 60% monoolein lipid using two gas-tight Hamilton syringes with a coupler (Landau and Rosenbusch 1996, Cheng et al. 1998). The introduction of the precipitant agents into the lipidic cubic phase further transforms the phase into a transient lamellar phase (Caffrey 2008). In the lamellar phase, the membrane proteins are enriched simultaneously and form protein crystals. Nucleation and crystallization subsequently occur in this phase (Caffrey 2015). LCP crystallization method provides a native-like membrane environment and can stabilize the protein at a wide range of temperature and pH values. It has been the "default" method to crystallize membrane proteins like G-protein-coupled receptors (Kermani 2020).

Figure 4.2.2 Schematic of membrane protein crystallization in lipidic cubic mesophase. Proteins are reconstituted into the lipid bilayer cubic phase (Left). After introducing the precipitant solution, phase separation occurred in the protein-lipid complex. Membrane proteins slowly diffused into a lamellar sheetlike domain and locally concentrated to form protein crystals (Middle). The magnified picture shows protein-protein interactions and proteins are orderly stacking along the lipid lamellar domain (Right). Figure, caption, and description are modified from (Li, Shah, and Caffrey 2013) with permission from American Chemical Society.

For 3D crystals formed by membrane proteins, two types of crystal packing occur (Figure 4.2.3). In type I, crystal packing is achieved by hydrophobic protein–protein and polar lipid–protein interactions. Type I packing layers of the membrane protein are formed by either lipids or detergent molecules, and contact sites between the layers are mediated

by hydrophilic interactions, including H bonds and electrostatic interactions. Type I crystals are often formed in lipid meso systems, but membrane proteins crystallized in detergent can also form type I crystals. An example is Photosystem I, which is crystallized in detergent at a low ionic strength. Here, detergent molecules form a lamellar phase of detergent molecules in the a/b plane of the crystals with a very high solvent content of 78% (Jordan et al. 2001, Gisriel et al. 2019). Type II packing is achieved by polar protein–protein interactions, with the hydrophobic region covered by the detergent micelle. Type II packing is mainly formed in crystallization experiments where crystals are grown in detergent micelles (Birch et al. 2018, Qutub et al. 2004).

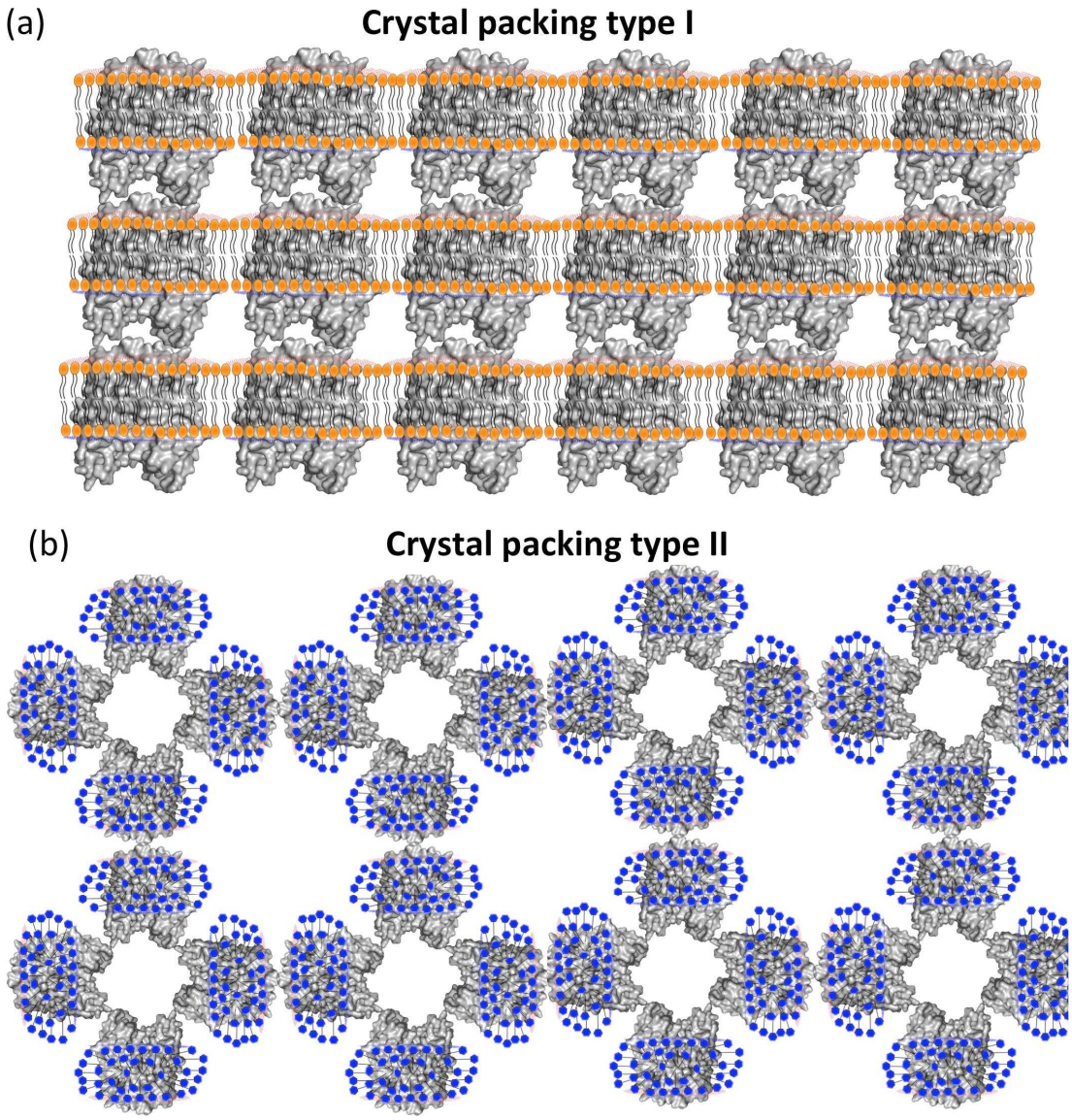

Figure 4.2.3 Summary of two types of membrane protein crystal packing. A) Crystal packing type I. Membrane proteins (gray-space filled) are crystallized within a lipid environment. The lipid molecules mimic the native biological membrane. Lipid–protein and protein–protein interactions contribute to the crystal packing. The orange dots represent the lipid molecules. B) Crystal packing type II. Membrane proteins are crystallized from a detergent micelle solution. Unlike in the type I crystal packing mechanism, the hydrophobic region of the membrane protein in the detergent solution is covered by detergent micelles. The polar protein–protein interactions contribute to the crystal packing. Generally, type II crystal packing forms larger solvent channels within the crystals. The blue dots represent the detergent molecules. Figure, caption, and description are modified from (Birch et al. 2018) with permission from ELSEVIER.

4.2.5 X-ray Sources

In 1985, Wilhelm Röntgen discovered and systematically studied electromagnetic energy waves with a short wavelength in the angstrom region; these waves were later named X-rays. Invisible X-ray light has an extremely high energy and is emitted via electron radiation. In the light spectrum chart, X-rays have a wavelength ranging from 0.01 to 10 nm. Based on the X-ray energy level, X-rays are sub-classified as soft or hard X-rays. X-rays with a high photon energy are called hard X-rays and have a wavelength below 0.3 nm. X-rays with a low photon energy are called soft X-rays and have a wavelength between 0.3 nm and 10 nm (Attwood and Sakdinawat 2017). The development of new X-ray sources continues to benefit many fields. For example, X-rays are an excellent resource in the medical field because they can penetrate soft tissues. They can easily penetrate flesh to determine if the bone is fractured. Furthermore, because of the similarity in size between hard X-ray wavelengths and interatomic distances, hard X-rays are critical in determining atomic-resolution structures using X-ray crystallography (Garman and Schneider 1997).

X-rays can be produced by hitting a material comprising a metal with sufficiently high energy-charged particles. The X-ray tube is the most commonly used source for their generation. The X-ray tube is a vacuum-sealed tube and has anode and cathode sides. At the anode side, it contains a copper stem with a metal target. At the cathode side, there is a tungsten filament operating at a high voltage. A circuit heats the filament to emit electrons, which are accelerated by the potential difference in voltage at both sides, toward the metal target at the anode. The accelerated electrons cause inner electron shell vacancy (“knocking out” electrons) of the target metal atoms. The electrons in the

outer orbitals of the target metal atoms “drop down” to fill the shell vacancy and release energy in the form of X-rays. The released energy or emitted X-rays are usually generated perpendicular to the electron beam. However, the X-ray tube is not efficient. Only 1% of the electric energy is converted to X-rays, with the remainder of the energy being converted to heat. The copper stem transduces the heat into a water cooling system (Zink 1997). Synchrotron radiation is a specialized X-ray source. Emitted electrons travel and accelerate via bending magnets in a curved storage ring path. Under the strong magnetic field, emitted electrons are accelerated to near light speed (Kim 1986). Within the storage ring, there are several types of magnets aligned straight or curved in an alternating order. In the straight section, there are focusing magnets and undulators. As the name implies, focusing magnets ensure that all electrons follow the ideal orbital path before reaching the undulator. The undulator contains magnets with alternating polarization to force the electrons to a bending path where they emit X-rays. The electrons are bent back and forth along the horizontal plane; in each constructive bend, the electrons emit radiation. The bending magnets force the electrons to adjust from a straight motion to a curving orbital path after the undulator. The electrons emit synchrotron light at the tangent to the orbital path under the bending magnets. The emitted light covers a broad continuous spectrum, i.e., from microwaves to X-rays (Bakshi 2006, Bharti and Goyal 2019). Synchrotron radiation sources can meet different light spectrum needs. Typically, a monochromator is used to select the wavelength of interest.

With increasing demands for improving brightness and achieving a coherent beam and short electron pulse duration, hard X-ray free-electron lasers (XFEL) became the

fourth X-ray generation source, providing coherent sub-picosecond pulses (Emma et al. 2010). The original idea for designing an XFEL is attributed to John Madey at Stanford in the 1970s (Winick 1997). Using XFEL as the light source, atomic and molecular level resolution can be achieved with much more coherence X-ray. Additionally, single-molecule and nano-scale ultrafast dynamic processes can be observed on a femtosecond scale before X-ray damaging the protein crystals. Unlike traditional synchrotrons, XFELs do not have a circular storage ring. Instead, they consist of a linear particle accelerator and straight undulators. In the first part of the European XFEL, electrons pass through a near 2 km long accelerator with superconducting resonators before "racing" through the undulators. Every electron has a velocity near the speed of light. In the second part, a long, periodically arranged undulator of alternating magnets forces the electrons in a wiggling path, where they emit X-rays. The electrons then interact with the X-rays; this interaction accelerates the slow electrons, decelerates the fast ones, and ultimately leads them into bunches. When all the electrons in a bunch wiggle in phase (i.e., are coherent), very bright and ultrashort X-ray pulses are emitted from each bunch. Thus, radiation emitted from every single electron will be amplified exponentially (Pellegrini and Stöhr 2003). This bunching of electrons in an XFEL is called coherence, and it amplifies the power of the emitted radiation. During the stimulated emission process, each electron is stimulated via an electromagnetic field emitted by other electrons. Consequently, the radiation is intensified (Figure 4.2.4). For example, if there were a total of N electrons in one electron bunch, the total power of radiation would be N^2. After this interaction, each electron bunch emits extremely short and intense X-ray pulses (Shi 2000).

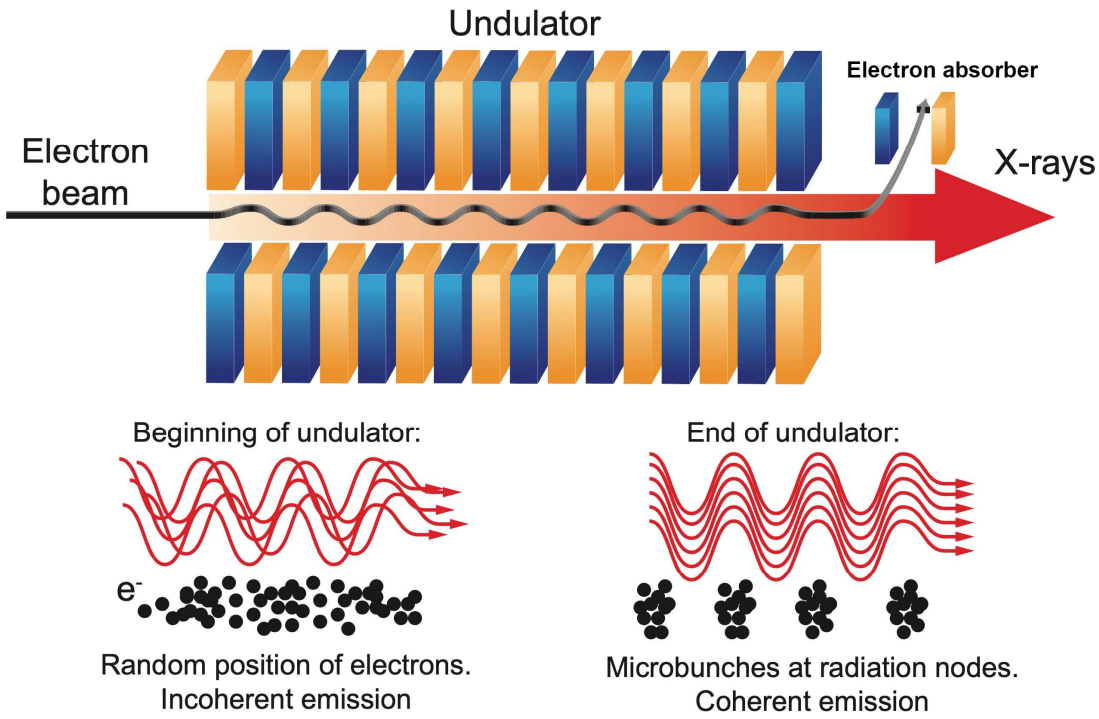

Figure 4.2.4 Principle behind the generation of coherent X-rays for X-ray free-electron lasers (XFELs). The generated electron beam passes through several magnet undulators at near light speed. The undulators force the traveling electrons to move in a sinusoidal motion and emit coherent X-ray photons. In the beginning, electrons are randomly distributed and emit incoherent photons. As the randomly distributed electrons pass through more undulators, the generated electromagnetic field forces the electrons to form micro bunches at the radiation nodes of the electromagnetic wave. At the end of the undulators, all electrons move in the same phase and emit coherent X-ray beams. Figure, caption, and description are modified from (Nilsson et al. 2011) with permission from AIP Publishing.

4.2.6 Protein Crystal Diffraction Methods

It is well established that a small object can be directly observed under a microscope when it is larger than the wavelength of light (1 μm). The scattered light from the object can be collected and focused via the objective lens of the microscope. Electromagnetic radiation with a wavelength close to the atomic distance of 1Å is required to determine protein structures at an atomic resolution. The carbon–carbon bond is

approximately 1.5 Å. However, no lens can collect and bend the focus of the scattered X-rays. The atomic-level structure must, therefore, be reconstructed via diffraction by measuring the experimental intensities of the diffracted waves using indirect methods to determine the phases. When the X-ray beam overlaps with the electron cloud of an atom, elastic scattering can occur and a secondary wave with the same wavelength but in a different direction can be generated (Bragg's law, Figure 4.2.5). A protein crystal is composed of repeating unit cells. Secondary waves from planes in the crystal interfere constructively, leading to reflections on the detector. The diffraction from the scattered waves is located in the reciprocal space (Ewald Sphere). A crystal must be rotated in the beam until all planes meet the Bragg's condition and all reflections are recorded. From the intensities of the reflections, the amplitudes of the scattered waves are determined; however, the phase is lost and has to be determined indirectly (phase problem). There are three parameters, namely wavelength (determined by the X-ray source) and amplitude and phase (determined by the electron density map). These are calculated and contain information regarding the location of the atom in the unit cell of the crystal (Parker 2003).

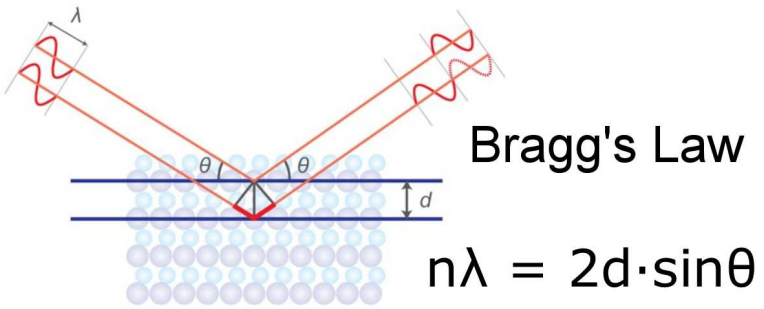

Figure 4.2.5 Schematic of Bragg's law. In Bragg's law, the n represents an integer; λ is the wavelength of the X-rays diffracted from the crystal sample; d is the spacing between layers of atoms; and θ is the angle between the incident X-ray beam and the surface of the crystal sample. Illustration was adapted and modified from Rigaku® website with modifications.

In a traditional synchrotron protein crystal diffraction setup, suitable size protein crystals can be directly mounted in a capillary tube at room temperature. Alternatively, frozen crystals can be mounted and frozen in a loop by maintaining a -177°C cooling condition with a stream of liquid nitrogen; this allows data collection under cryogenic conditions. A protein crystal is mounted on the goniometer head, which can be manually adjusted to ensure the upcoming X-ray beam and crystal align ideally. There is a motorized detector system behind the mounted crystal to allow the crystal-to-detector distance to be altered for observing all reflections. In general, the initial detector location allows the collection of X-ray diffraction from 80 Å to a maximum 1.5–3 Å resolution range. In a diffraction pattern, the resolution of spots collected on the detector increases as the diffracted angle increases. A photon-counting hybrid pixel array detector (HPAD) was recently used to determine the diffraction intensity at individual photon levels with a wide dynamic range (Brönnimann and Trüb 2020). There is a beam stop between the crystal and detector to prevent the intense X-ray beam from damaging the detector or causing overexposure in the central region of the detector (Smyth and Martin 2000).

Unlike the traditional crystal mounting method, the serial femtosecond crystallography (SFX) approach is applied in XFEL data collection. This is because ultrabright XFEL pulses destroy contacted crystals with a single pulse. In SFX, thousands of microcrystals are delivered in a liquid jet and interact with the upcoming XFEL beam in random orientations. Therefore, single-shot still diffraction patterns are collected from thousands of crystals at a random orientation that are assembled and integrated. The short XFEL pulse avoids radiation damage because the crystals diffract before they are

destructed by the X-ray pulses (Neutze et al. 2000). In the SFX approach, sub-micrometer to micrometer protein crystals can be delivered at room temperature for structural determination. Furthermore, protein conformational dynamic transitions can be observed and studied with a sub-picosecond time resolution (Kupitz et al. 2014, Pande et al. 2016). The experimental setup in SFX is similar to the traditional synchrotron (Figure 4.2.6). Because of the extreme beam intensity in XFEL and the short pulse duration, the previously mentioned synchrotron HPAD detector is unsuitable. Other detectors have been developed for XFEL, including the Cornell-SLAC pixel array detector (CSPAD), adaptive gain integrating pixel detector (AGIPD), and adjusting gain detector for the aramis user station (JUNGFRAU) (Mishin et al. 2019). To reduce the background scattering signal, the sample chamber is either under vacuum or filled with helium. Helium-filled sample chambers provide greater experimental simplicity and flexibility than vacuum-sealed chambers. Many sample delivery methods can be adapted to the helium chamber, and sample environments can be adjusted according to needs. However, moderate noise occurs as a tread-off. Many sample delivery methods can be adapted to the helium chamber and sample environments can be adjusted by needs. However, moderate noise occurs as a tread-off.

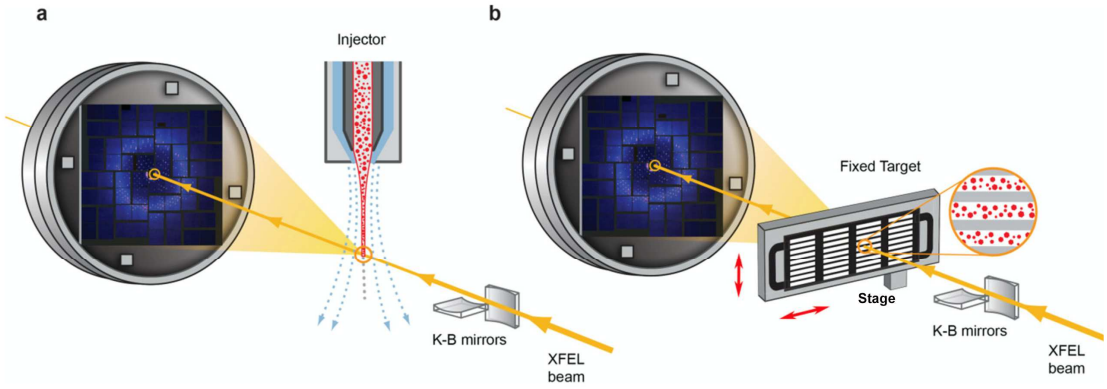

Figure 4.2.6 Scheme of the serial femtosecond crystallography experimental layout at XFEL beamlines. A) Injector-based crystal delivery system layout. In the injector-based crystal delivery system, the microcrystals are constantly flowing against the X-rays. B) Fixed target crystal delivery system layout. Microcrystals are mounted on a solid support. The fast-moving stage controls the diffraction locations. In this XFEL beamline, the coherent X-rays are focused by K-B mirror optics before diffracting the targeted objects. Generally, both delivery methods are performed in a vacuum or a helium environment to reduce scattering from the direct beam. However, the vacuumed or helium chamber is not necessary for all XFEL experiments. Figure, caption, and description are modified from (Mishin et al. 2019) with permission from Taylor & Francis.

Preparation and delivery of the crystals are the most critical steps in XFEL experiments. Several sample delivery methods have been designed and can be separated into two main categories: 1) injector-based methods and 2) fixed target methods. In the injector-based methods, stream of the crystals is constant across the XFEL beam; the crystals are embedded in special carrier matrixes with different levels of viscosities (Nam 2019). In the fixed target method, the protein crystals are either deposited or directly grown on a solid support chip; the XFEL beam can quickly raster the chip (Mishin et al. 2019, Pearson and Mehrabi 2020). XFEL techniques have been increasingly developed for sample delivery, data collection, and data processing during the last few decades (Sauter 2015, Nelson et al. 2016, White 2019, Martiel, Muller-Werkmeister, and Cohen 2019, Schulz et al. 2019). All these efforts allow for room temperature crystallography,

facilitating the successful determination of protein structure with limited crystal sizes. Additionally, time-resolved studies provide the key to observing the conformation transition moment in the targeted macromolecules; reactions can be triggered "on the fly" by rapid mixing or light excitation.

In addition to X-ray protein crystal diffraction, electron intersection with matter can achieve atomic structure determination via cryo-electron microscopy (Cryo-EM), as well as from single-particle scattering or diffraction from 2D or small 3D crystals. Apart from the popular 2D electron crystallography, micro-electron diffraction (MicroED) can also diffract 3D microcrystals with continuous rotation (Nannenga et al. 2014). In the MicroED method, protein microcrystals are diffracted under cryogenic conditions; data processing with this crystal diffraction method is similar to that with traditional X-ray crystallography.

Electron diffraction is a powerful tool to determine macromolecule structure and has many advantages over X-ray crystallography. First, electrons have much stronger interactions with the object than X-rays and expose less energy to the object during diffraction (less radiation damage). Second, because of the inherently strong interaction with molecules, electron diffraction can achieve high resolution with a small protein crystal size (Henderson 1995). Lastly, diffracted and scattered electrons can be focused by the (electron) microscope, which allows crystal imaging and accurate phase information recovery (Nannenga and Gonen 2016). The MicroED technique was proposed in 2013 and was developed based on collecting multiple diffraction patterns from a single crystal and sampling with a large enough region of reciprocal space to perform traditional indexing and diffraction data processing (Shi et al. 2013). Later, adapting the continuous

rotation data collection method in MicroED improved the diffraction dataset quality and allowed processing of datasets via conventional X-ray crystallography software (Hattne et al. 2015). The MicroED experimental setup is a hybrid method, including elements of traditional X-ray crystallography and single-particle cryo-EM. Targeted proteins are normally crystallized via the vapor diffusion method, and crystals are later transferred to a glow-discharged cryo-grid by pipetting. The sample is placed on a grid, excessive solution on the grid is blotted, and the grid is flash-frozen or vitrified using a vitrification device (e.g., a vitrobot). Vitrified samples are stored in liquid nitrogen and can be loaded into high-resolution cryo-EM instruments. Both MicroED and 2D electron crystallography require a fast-moving sample stage in the cryo-EM instrument. In cryo-electron diffraction (Cryo-ED), once a suitable size crystal is identified, the crystal is exposed to the electron beam with continuous stage rotation motions (Shi et al. 2016). Currently, the main advantage of MicroED over SFX is the availability of instruments and the requirement of a small amount of crystal sample. However, many methodological optimizations are required to integrate the diffraction data and more accurate electron scattering factors are required to improve the data refinement step (Nannenga and Gonen 2016). In addition, the growth of the nanocrystals is very challenging because the thickness of the crystal needs to be < 200 nm to avoid multiple scattering of the diffracted electrons. Recently, a more comprehensive study shows protein crystals that were less than 1 μm thick can be applied for MicroED methods (Martynowycz et al. 2017). Lastly, radiation damage is massive and data collection must take place under cryogenic conditions (ideally in liquid helium), which hinders the time-resolved study.

4.3 Materials and Methods

Initial Protein Crystallization Screening

Purified pGC-A protein from the Superose6 chromatography was concentrated to 2 mg/mL using a 100 kDa cutoff amicon concentrator (EMDmillipore, Cat. UFC510096) by bench centrifuging (12 K g-force speed) at 4°C. To reduce the possibility of the protein directly reaching the precipitation zone where an amorphous precipitate is formed, during the initial crystallization screening process, 80% concentration of the original MemGold2 HT-96 kit (Molecular dimensions, Cat. MD1-64) was manually screened in 24-well plates by the hanging drop (vapor diffusion) method. Next, to avoid protein degradation, plates were set up and incubated in a cold room at 4°C. First, 1 μL of concentrated protein sample was mixed with 1 μL of 20% diluted MemGold2 HT-96 screening precipitant solution on a glass cover slide. Then, glycerol was added to the precipitant solution in each well (a final total of 10% glycerol concentration) to compensate for the glycerol brought into the crystallization drop by the protein solution; the glass cover slide was flipped and sealed back onto the well. The plates were checked 3 and 5 days after setup. During the initial crystallization screening, promising crystallization conditions were identified based on the formation of crystalline structures that showed a signal in the SONICC imaging system. Images of drops in the 24-well plates were captured under visible light, SONICC Hi-Res, and UV TPEF Hi-Res modes.

Crystallization Optimization Screening in F8 Condition

Based on the original F8 condition (Table 4.4.1), two diluted F8 conditions were prepared: the 60% original F8 condition and 80% original F8 condition (note that the 80% condition was the same as the initial screening condition). The same hanging drop

method was used to set up 12 wells in the plate with a 2 mg/mL protein concentration. After comparing the two diluted conditions, the 80% original F8 condition (0.32 M of ammonium sulfate, 0.08 M of MES buffer [pH 6.5], and 8% of PEG3350) was chosen as the starting condition for screening optimization. The following stock chemical solutions were used for fine screening under the F8 condition: 1 M MES buffer, [pH 6.5] (Molecular Dimensions, Cat. MD2-013-PH), 3.5 M ammonium sulfate (Molecular Dimensions, Cat. MD2-100-35), and 50% PEG3350 (Molecular Dimensions, Cat. MD2-100-9).

To further optimize the 80% original F8 condition, both the salt (ammonium sulfate) and the precipitant (PEG3350) were applied to create broad step gradient screenings in a 24-well plate. In the AB rows, the percentage of PEG3350 was varied from 5% to 13%, with 0.08 M of MES buffer [pH 6.5] and 0.32 M of ammonium sulfate not altered. In the CD rows, the concentration of ammonium sulfate varied from 0 M to 0.6 M, with 0.08 M of MES buffer [pH 6.5] and 8% of PEG3350 unaltered. Additionally, two protein concentrations: 1.18 mg/mL and 2.25 mg/mL were used during the board screenings.

Based on the previous broad screening results, both the salt (ammonium sulfate) and the precipitant (PEG3350) were used to create narrow step gradient screenings in a 24-well plate. The protein concentration was 1.25 mg/mL. In the AB rows, the concentration of PEG3350 was varied from 5% to 10% with 0.08 M of MES buffer [pH 6.5] and 0.32 M of ammonium sulfate (not changed). In the CD rows, the concentration of ammonium sulfate was varied from 0 M to 0.4 M with 0.08 M of MES buffer [pH 6.5] and 8% of PEG3350 (not changed).

Crystallization Optimization Screening in C4 Condition

Unlike the F8 condition, the board C4 condition screening was performed in a 15°C incubator with a protein concentration of 1 mg/mL. Next, to keep the 24-well plate cool, the plate was placed in an ice bath and all crystallization solutions were chilled on ice before use. The same vapor diffusion (hanging drop, 1 µL + 1 µL) method was used to optimize the C4 condition (0.1 M of ammonium sulfate, 0.05 M of Tris-buffer [pH 7.0], and 22% of PEG250DME). During the board screening, the initial PEG250DME concentration was screened in the range from 15% to 30% with 100 mM ammonium sulfate and 50 mM Tris-buffer at pH 7.0 (not changed). The initial ammonium sulfate concentration was screened in the range of 0–0.6 M with 100 mM ammonium sulfate and 50 mM Tris-buffer at pH 7.0 (not changed). Stock chemical solutions for screening of the C4 condition were 1 M Tris buffer (pH 7.0; Molecular Dimensions, Cat. MD2-028-PH), 3.5 M ammonium sulfate (Molecular Dimensions, Cat. MD2-100-35), and 40% PEG250DME (Molecular Dimensions, Cat. MD2-100-180).

To further optimize the C4 condition based on the initial screening, the PEG250DME concentration range was narrowed from 17% to 29% with 100 mM ammonium sulfate and 50 mM Tris-buffer at pH 7.0 (not changed). The ammonium sulfate concentration was narrowed from 0 M to 0.33 M with 22% PEG250DME and 50 mM Tris-buffer at pH 7.0 (not changed). Both narrow screenings were performed in a 15°C incubator and at room temperature with 1 mg/mL purified pGC-A.

Real-time Protein Crystallization under Microscope

To rationally test whether intact pGC-A crystallized in the purification buffer (50 mM HEPES [pH 7.5] and 150 mM NaCl) by adding a defined amount of PEG3350, both

PEG3350 (6%–12% with 150 mM NaCl) and sodium chloride concentration screenings (0–0.32 mM with 8% PEG3350) were performed as described above. The crystallization plate was set up with 1 mg/mL protein and incubated in the cold room at 4°C. The plate was brought from the cold room temperature to room temperature for regular imaging. The heat from the microscope promoted microcrystal nucleation and formed a crystal shower in real-time under the microscope. To screen the exact temperature that stimulated rod shape microcrystal formation, four stable temperatures were tested. The plates were separately incubated in the 4°C cold room, 10°C incubator, 15°C incubator, and at room temperature.

Confirmation of pGC-A Protein Crystals Formed in the F8 Condition

To confirm if the needle-shaped crystals that grew in the F8 condition were composed of the pGC-A protein, 3-week-old crystallization drops were collected via pipetting the drop content into a PCR tube. In this experiment, the content of five crystallization drops was collected into a PCR tube. Collected crystals were pelleted down by centrifugation and washed with crystallization precipitant solution. Figure 4.3.1 summarizes the scheme of the crystal washing and collection processes. The protein crystals were pelleted at a maximum 17K g-force for 10 minutes in 4°C. The crystallization drop mixture in the PCR tube, a sample of the centrifugation supernatant, washed crystal pellets, and a sample of the washed crystal pellet supernatant, as well as the content of one intact crystallization drop were each mixed with SDS sample buffer in 1: 1 volume ratio. To partially preserve the multimers in the protein crystals, the prepared SDS samples were directly loaded into the SDS-PAGE without heating. Western blot analysis was used to detect the pGC-A signal. For the western blot, the primary antibody,

anti-pGC-A, was ordered from R&D Systems, (Cat. MAB4860) and the secondary antibody, HRP-conjugated anti-mouse IgG, was ordered from Jackson ImmunoResearch (Cat. 515-035-062).

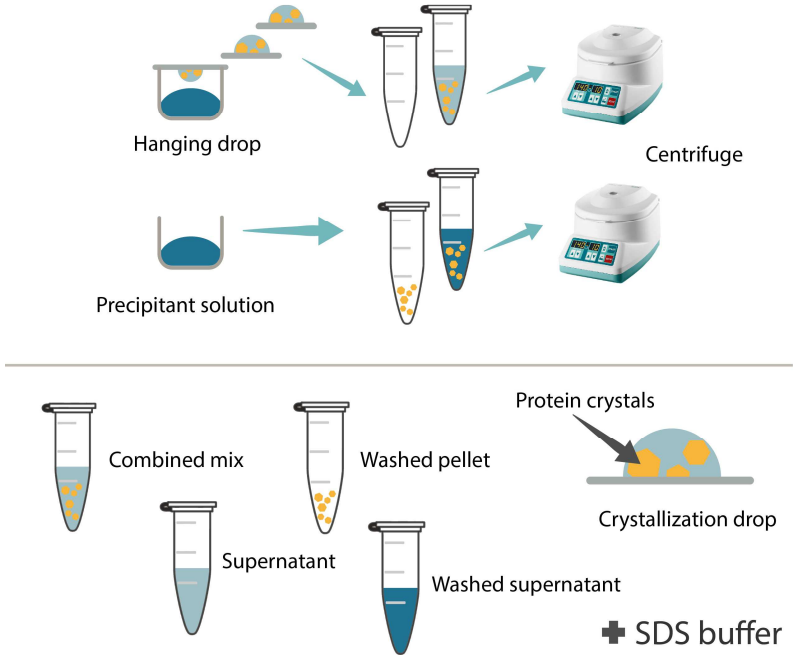

Figure 4.3.1 Summary of protein crystal collection and washing steps. This diagram shows both protein crystal collection and the washing processes with crystallization precipitant solution. After washing, each collected sample was mixed with SDS sample buffer for anti-pGC-A western blot.

F8 Condition Protein Crystal Diffraction and Data Collection

As mentioned before, 1 mg/mL of intact human pGC-A was crystallized under the 80% F8 condition (0.32 M ammonium sulfate, 0.08 M MES [pH 6.5], 7–12% PEG3350). Needle-shaped crystals from 96 crystallization drops (four plates) were harvested and placed into 0.65 mL Eppendorf tubes and shipped to the Advanced Photon Source (APS, Argonne National Laboratory, Chicago, IL) within a temperature-controlled 4°C container. Sample preparation was performed as described previously (Martin-Garcia et al. 2017). Briefly, 16 µL lipid mixture was prepared by mixing the crystallization

precipitant solution (25%) with molten 9.9 MAG (monoolein, 75%) (Anatrace, Cat. LCP18 1 GM) in a glass dual-syringe lipid mixer. Then, the crystal slurry was gently centrifuged at 100 rcf for 20 min at 4°C to pellet the crystals. After removing the supernatant, pelleted crystals (12 µL) were gently loaded to a 100 µL glass syringe and mixed with previously prepared lipid/precipitant mixture (the LCP) (16 µL) in a glass dual-syringe lipid mixer. Later, another 8 µL of 9.9 MAG was added and mixed with the total 36 µL lipid crystal mixture. Finally, 2 µL of 7.9 MAG monomyristolein lipid (Anatrace, Cat. 7Z-LCP14 250 MG) and 2 µL of 9.9 MAG were added to titrate the lipid crystal mixture at a final volume of 40 µL. Once the homogeneous transparent LCP was formed, the LCP embedded crystal sample was ready for sample delivery. Finally, 40 µL of LCP was prepared with 40% crystal solution and 60% lipids. The LCP embedded crystal sample was loaded into an LCP injector reservoir with a 50 µm glass capillary nozzle (Weierstall et al. 2014). The sample flow rate was varied between 15 nL/min to 300 nL/min during data collection, and the LCP-embedded crystal sample was delivered to the synchrotron X-ray beam at room temperature.

Serial data collection was performed on the GM/CA 23-ID-D beamline at APS using the same experimental setup previously described (Martin-Garcia et al. 2017). Data collection was performed in a shutterless mode at a repetition rate of 100 Hz (10 ms exposure time), which was achieved using the continuous readout of the PILATUS3 6M detector. X-ray diffraction data were collected at a photon energy of 11.5 keV (1.07 Å) and a beam size of 16 × 10 µm (H × V). The sample-to-detector distance was fixed at 300 mm. A total of 270,000 frames were recorded on the PILATUS3 6M detector, five of

which were identified as hits. Diffraction patterns were indexed by the CrystFEL program (White et al. 2016).

4.4 Results

Initial Intact pGC-A Crystallization Screening

In this study, the conditions under which purified intact human pGC-A formed protein crystals were determined. Purified pGC-A protein from the Superose6 column was concentrated to 2 mg/mL and initially screened with the commercial MemGold2 HT-96 kit. To decrease the possibility that protein directly precipitated during the initial crystallization screening process, an 80% concentration of the original commercial kit was used as the precipitant. Diluted crystallization conditions were manually screened in 24-well plates by the hanging drop (vapor diffusion) method in a cold room (4°C). After manual screening using the MemGold2 HT-96 kit in the cold room (4°C), seven crystallization conditions were selected based on the formation of crystalline protein nano/micro crystals, as identified by positive signals in the SONICC imaging system (Figure 4.4.1). The exact crystallization conditions from selected wells are listed in Table 4.4.1. Additionally, other conditions appeared promising. To narrow down the total number of crystallization conditions for further optimization, seven of 96 conditions were selected after the initial screening.

Table 4.4.1 Summary of seven selected crystallization conditions after initial crystallization screening with the MemGold2 HT-96 kit

Well #	Conc.	Salt name	Conc.	Buffer name	Conc. v/v (w/v)	Precipitant
G8	-	-	0.05M	Tris, pH 7.5	17%	PEG 350 MME
F8	0.4M	Ammonium sulfate	0.1M	MES, 6.5	10%	PEG3350
B1	0.1M	Sodium chloride	0.01M	MES, 6.5	19%	PEG1000
C4	0.09M	Ammonium sulfate	0.05M	Tris, pH7.0	22%	PEG 250 DME
G12	-	-	0.1M	HEPES, pH7.5	28%	PEG600
F11	0.4M	Ammonium thiocyanate	0.1M	Sodium acetate, pH 4.5	15%	PEG 4000
F12	0.4M	Sodium thiocyanate	0.1M	Sodium acetate, pH 4.0	16%	PEG 4000

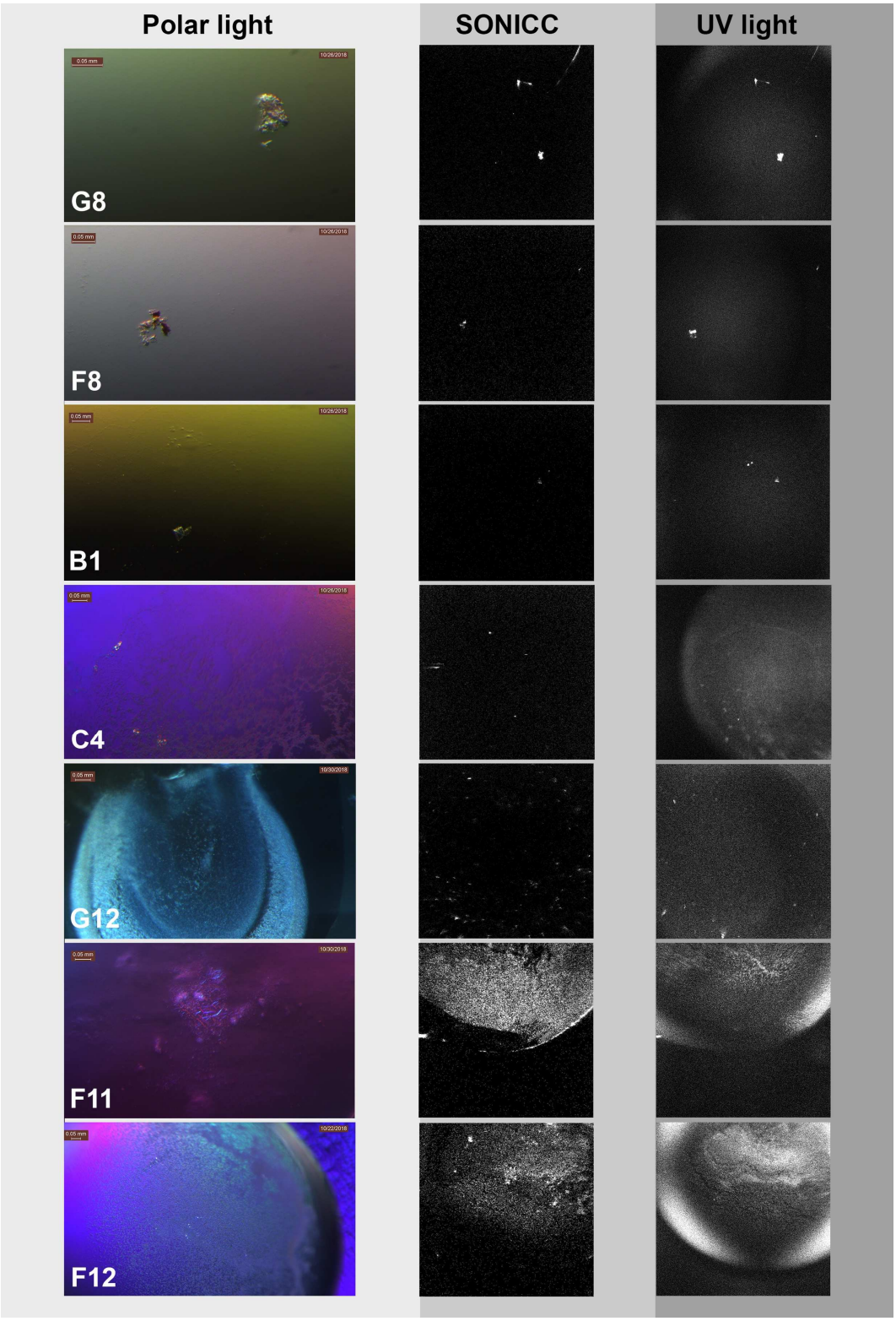

Figure 4.4.1 Summary of seven selected crystallization conditions from the initial crystallization screening. Manually screened plates were imaged by light microscopy, UV Trp fluorescence, and SONICC to identify promising crystal hits.

After setting up new crystallization plates based on the seven selected crystallization conditions, one 3D crystallite with 3 μm width and 5 μm height was formed under the F8 condition (Figure 4.4.2). To assess whether this was a protein crystallite, the F8 crystallite was mounted in a loop, and X-ray diffraction was tested at a X-ray home source (Figure 4.4.2). In our home-source X-ray facility, the Rigaku microfocus rotating anode X-ray generator (MicroMax-007HF) was used to provide the X-ray beam. Because of the extremely small crystallite size, the crystallite was transferred to a CrystalQuick™ 96-well plate, which allowed *in situ* observation of X-ray diffraction (Watanabe, Murai, and Tanaka 2002). After exposing the crystal in the plate to X-rays at room temperature, a classic protein crystal powder diffraction pattern was observed (Figure 4.4.2B). The powder diffraction results suggested that the 3D crystallite formed under the F8 condition was formed by protein (possibly the targeted intact pGC-A protein). The diffraction concentric ring indicated the crystallites might have the highest resolution powder diffraction rings between 3.24 to 6.06 Å (Figure 4.2.4 C).

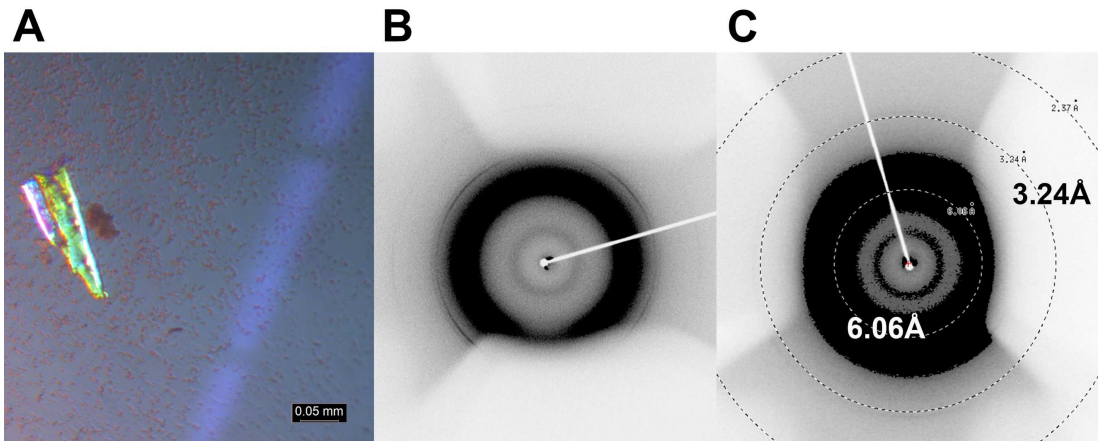

Figure 4.4.2 X-ray powder diffraction of protein crystallites under the F8 condition. A) Micro-crystallites crystallized under the F8 condition imaged with polarized light; B) Multiple powder diffraction rings generated by diffraction at the ASU X-ray home-source; C) Powder diffraction results with diffraction concentric ring.

Further Optimization of F8 Crystallization Condition

Based on the powder diffraction results (Figure 4.4.2), the F8 condition was chosen for further optimization. Two further diluted F8 conditions were screened in a 24-well plate format. Among 12 crystallization wells, some crystallization drops showed needle-shaped crystals at both 60% and 80% concentrations of the original F8 condition (Figure 4.4.3). On comparing the morphology presented by these needle-shaped crystals, the 80% concentration of the original F8 condition was chosen for further optimization. The 80% concentration of the F8 condition contained 0.32 M of ammonium sulfate, 0.08 M of MES [pH 6.5], and 8% PEG3350.

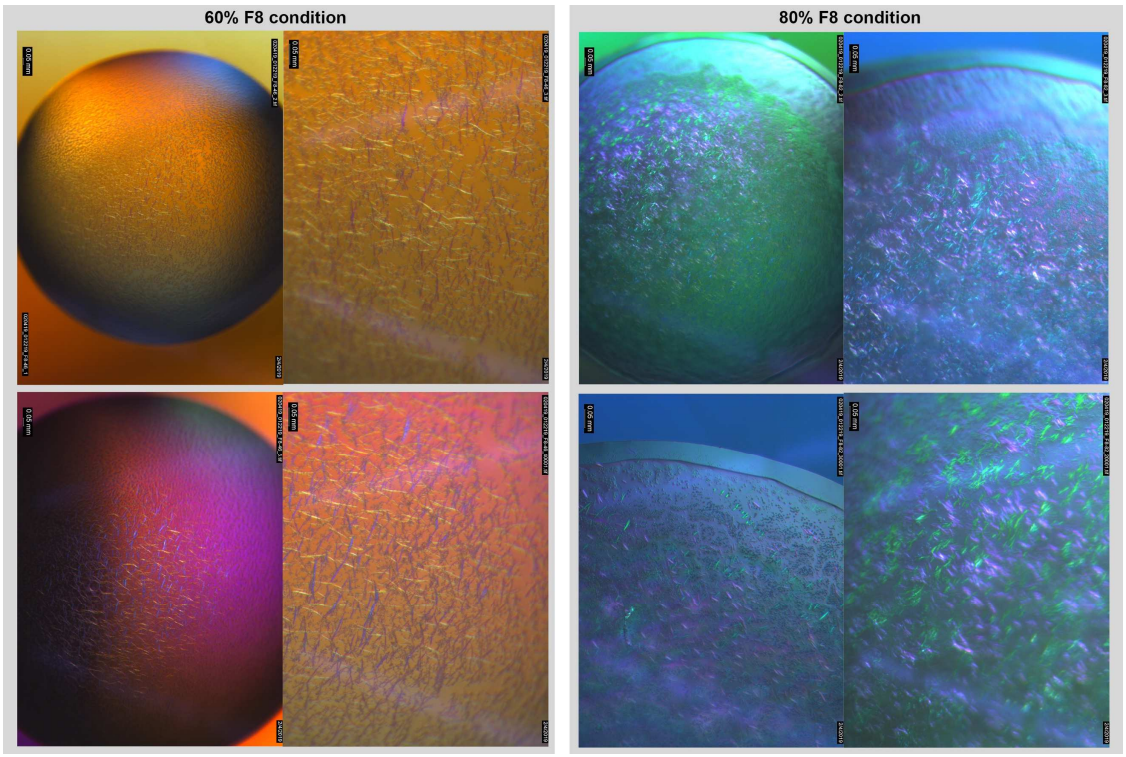

Figure 4.4.3 Needle-shaped crystals under two different diluted F8 conditions. The gray panel on the left shows the crystallization drop equilibrated with a precipitant solution corresponding to 60% of the original F8 condition. The gray panel on the right shows the crystallization drop corresponding to 80% of the original F8 condition. Crystallization was screened with 2 mg/mL purified protein concentration.

To further optimize the 80% concentration of the F8 original crystallization condition, broad gradient screening was applied for both salt and precipitant concentrations. Additionally, the broad screening was performed with two protein concentrations. The board PEG3350 screening results with different protein concentrations are shown in Figure 4.4.4, and the results of the board ammonium sulfate screening with different protein concentrations are shown in Figure 4.4.5. In both PEG3350 and ammonium sulfate optimization screens, the crystals grown under the lower protein concentration (1.18 mg/mL) exhibited a better morphology and featured a smaller precipitation zone size in the center of the drops. To further optimize the crystallization conditions and reduce the protein sample size needed for further fine screening, the 1.18 mg/mL protein concentration was chosen for further optimization. The results from the broad PEG3350 screening showed that the crystallization drops started to form tiny needle-shaped crystals in 5% PEG precipitant. The high viscosity of high PEG concentrations prevented protein crystals from growing in the 10.84% to 13% PEG 3350 concentrations (Figure 4.4.4). The comparison of the screening results from the ammonium sulfate gradient screening showed that ammonium sulfate was necessary to form needle-shaped crystals. However, a higher ammonium sulfate concentration also led to precipitation in the crystallization drops, with fewer protein crystals (Figure 4.4.5).

Both narrow gradient screenings were applied to further optimize the PEG3350 and ammonium sulfate concentration at a lower protein concentration (1.25 mg/mL). The narrow PEG3350 screening was set from 5% to 10% (Figure 4.4.6 A), and the narrow ammonium sulfate screening was set from 0 M to 0.4 M (Figure 4.4.6 B). In the narrow PEG3350 screening, the PEG range of 5%–9.05% grew better needle-shaped crystals. A

trend where higher PEG concentration induced smaller needle-shaped crystals to form was noted. The narrow ammonium sulfate screening indicated that ammonium sulfate salt was necessary to form needle-shaped crystals (Figure 4.4.6 B, 0M salt yielded no crystals). There were many clusters of needle-shaped crystals formed at lower ammonium sulfate concentrations. It can be postulated that ammonium sulfate facilitated protein crystal formation in the drops after phase separation occurred in the crystallization drops (Figure 4.4.6 B, 36.4 mM to 145.6 mM). The ammonium sulfate concentrations from 218 mM to 291 mM produced better crystals in the drop without forming clusters or precipitation zones in the center of the drops. Comparing the higher protein concentration with the lower one (Figure 4.4.5 B and Figure 4.4.6 B), the center precipitation zones were much smaller with the lower protein concentration condition and the high ammonium sulfate concentrations. In conclusion, to crystalize the intact pGC-A with a lower protein concentration, the optimal PEG 3350 concentration ranged from 5% to 10%, and the optimal ammonium sulfate concentration ranged from 218 mM to 400 mM.

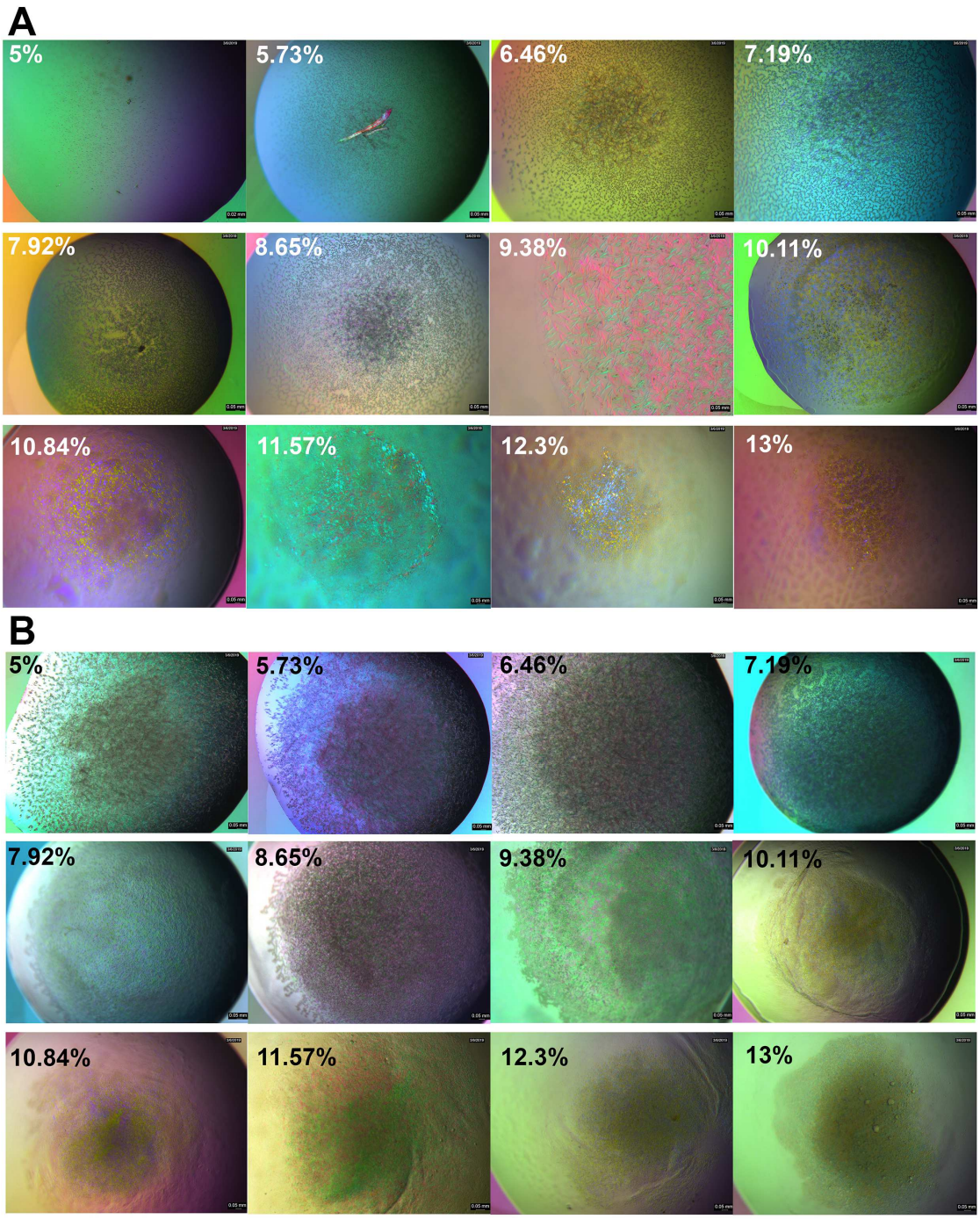

Figure 4.4.4 Summary of broad PEG 3350 screening of the F8 crystallization condition with two protein concentrations. A) 5–13% PEG 3350 screening at a protein concentration of 1.18 mg/mL; B) 5–13% PEG 3350 screening at a protein concentration of 2.25 mg/mL.

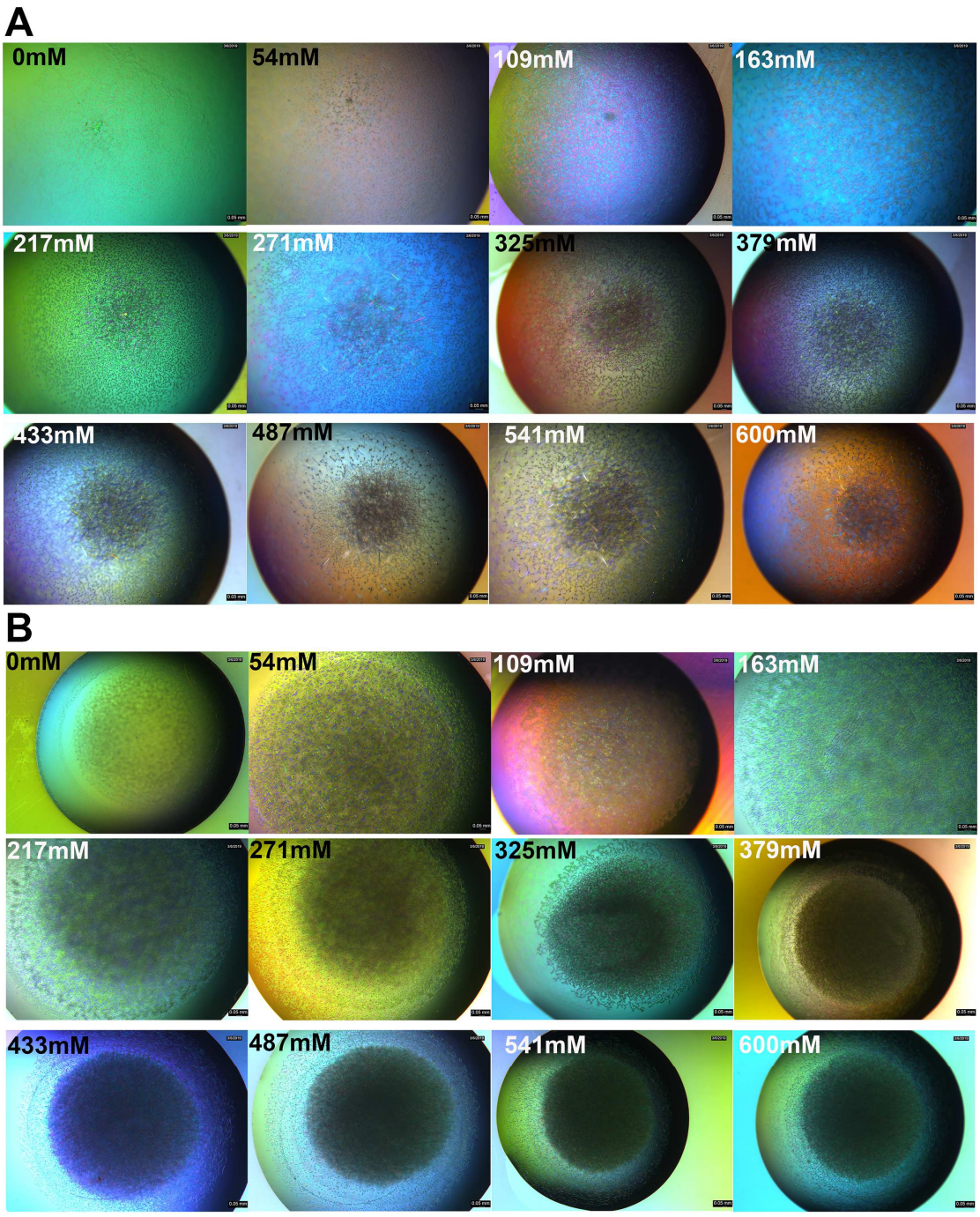

Figure 4.4.5 Summary of broad ammonium sulfate screening in F8 crystallization condition with two protein concentrations. A) 0–0.6M ammonium sulfate screening at a protein concentration of 1.18 mg/mL; B) 0–0.6M ammonium sulfate screening at a protein concentration of 2.25 mg/mL.

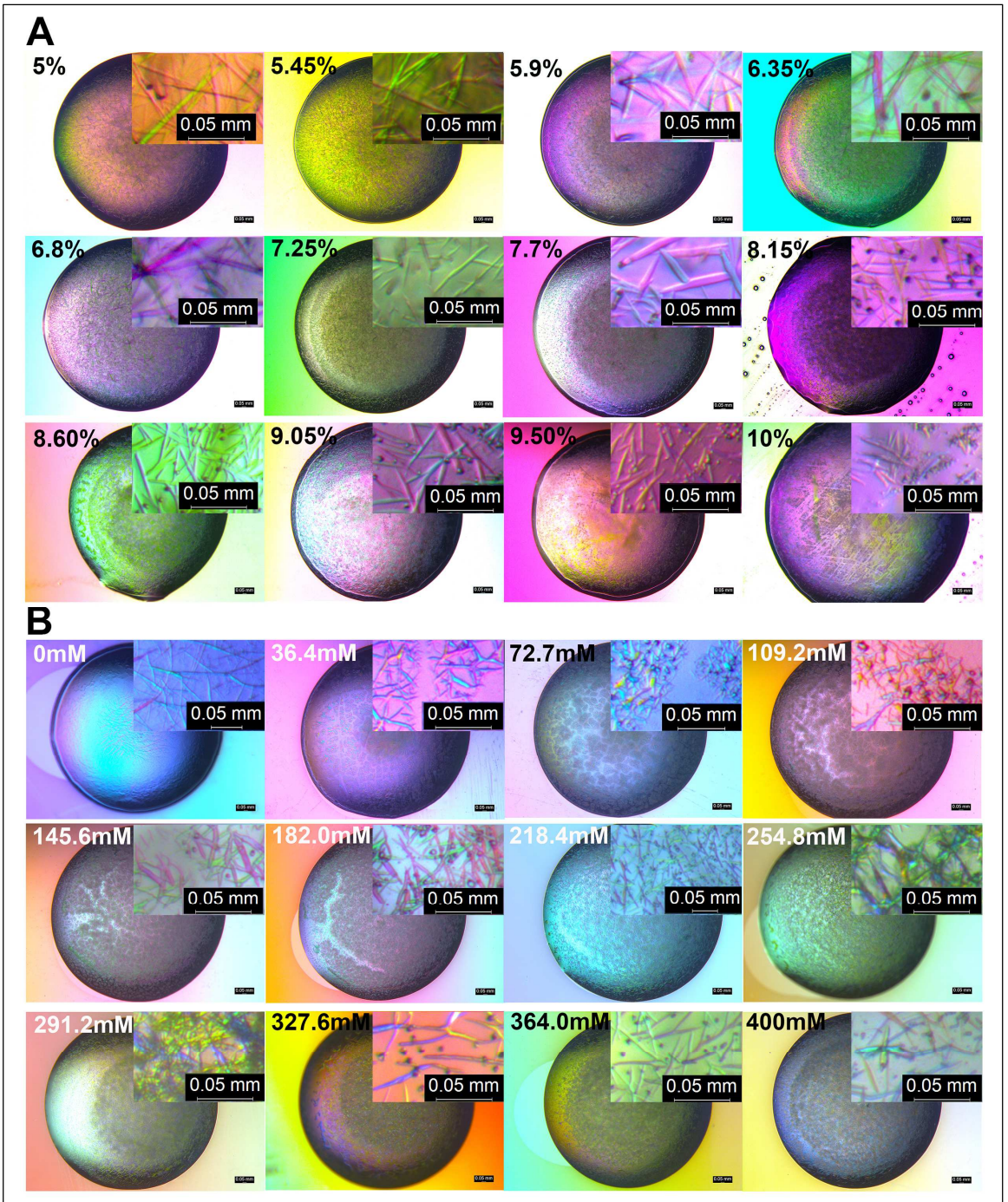

Figure 4.4.6 Summary of both PEG3350 and ammonium sulfate narrow screenings derived from the F8 crystallization condition. A) 5–10% PEG3350 screening at a protein concentration of 1.25 mg/mL; B) 0–0.4M ammonium sulfate screening at a protein concentration of 1.25 mg/mL.

Optimization of C4 Crystallization Condition

Comparing the previously optimized F8 crystallization condition with the other six initial selected conditions, the C4 crystallization condition was highly similar to the F8 condition but with distinct differences. The C4 condition had the same salt (ammonium sulfate) with different concentrations; however, the buffer and pH were different (Tris pH7.0 vs. MES pH 6.5), as were the PEG precipitants (PEG250DME VS. PEG3350) (Table 4.4.1).

To explore the C4 condition, both initial PEG250DME and ammonium sulfate screenings were performed with 1 mg/mL purified pGC-A. The initial PEG250DME concentration was screened in the range from 15% to 30% with a constant 0.1 M ammonium sulfate concentration, and the initial ammonium sulfate concentration was screened in the range of 0–0.6 M with constant PEG250DME at 22% concentration. Excellent short needle-shaped crystals were formed in both higher PEG and ammonium sulfate concentration gradients after 5 d incubation at 15°C, where the first images of crystals were obtained 10 d after setup of the crystallization plates (Figure 4.4.7 A). The images in Figure 4.4.7 B are obtained after 14 d. Based on the initial results, the promising PEG250DME range was 17%–29%, and the promising ammonium sulfate range was 0–0.33 M (Figure 4.4.7 B). Interestingly, 6 d after the first imaging, many super-long needle-shaped crystals were formed by transformation into bigger crystals by drawing protein in equilibrium from solution and nearby 5 to 10 μm small needle-shaped crystals, which disappeared (Figure 4.4.7 B&C). There were Ostwald ripening events that occurred in some crystallization drops. The smaller crystals were depleted and provided "building materials" for larger growing ones, with a "halo" effect created around the

larger growing crystals. Spontaneous formation of larger crystals could be attributed to thermodynamic favoring. Small crystal formation is kinetically favored, but molecules on the surface of the small crystals are less energetically stable because of the larger surface area to volume ratio (Boistelle and Astier 1988). However, the formation of many small nuclei of crystals decreased the supersaturation and consequently favored crystal growth over nucleation (Figure 4.2.1). The protein crystals in the screened C4 conditions continued to form larger sizes gradually over 14 d. However, the sharp features of the crystals were lost; an uneven surface noticeably appeared on the bigger crystals. For the C4 condition, early crystal harvesting is vital for harvesting higher-quality crystals.

Based on the previous initial screening, narrowed PEG250DME screening (17%–29%) and ammonium sulfate (0–0.33 M) screening were performed at two different temperatures: room temperature (Figure 4.4.8) and 15ºC (Figure 4.4.9). All images were taken 11 days after the setup of the crystallization screening. In the PEG250DME gradient screening, both temperature crystallization drops showed that higher PEG concentrations potentially promoted longer and larger needle-shaped crystals with the same incubation time. Additionally, the temperature setting affected the speed of crystallization. At the same PEG concentration, crystals grown at room temperature had bigger sizes than crystals grown in the 15°C incubator. In the ammonium sulfate screening, ammonium sulfate showed a similar effect as PEG250DME. The higher the ammonium sulfate concentration in the crystallization drops, the longer and larger the needle-shaped crystals grew. Notably, when higher concentration of ammonium sulfate was applied to the drops, the high concentration (higher than 119 mM) inhibited crystal formation at 15°C and room temperature. Additionally, a high concentration of

PEG250DME (higher than 25%–29%) and ammonium sulfate (179 mM–327 mM) at room temperature formed spiky salt crystals. In conclusion, to obtain better quality needle-shaped crystals from the C4 condition, the protein crystals should grow at a lower temperature (15°C). The optimal PEG250DME concentration ranged from 18% to 26%, and the optimal ammonium sulfate concentration was from 30 mM to 150 mM.

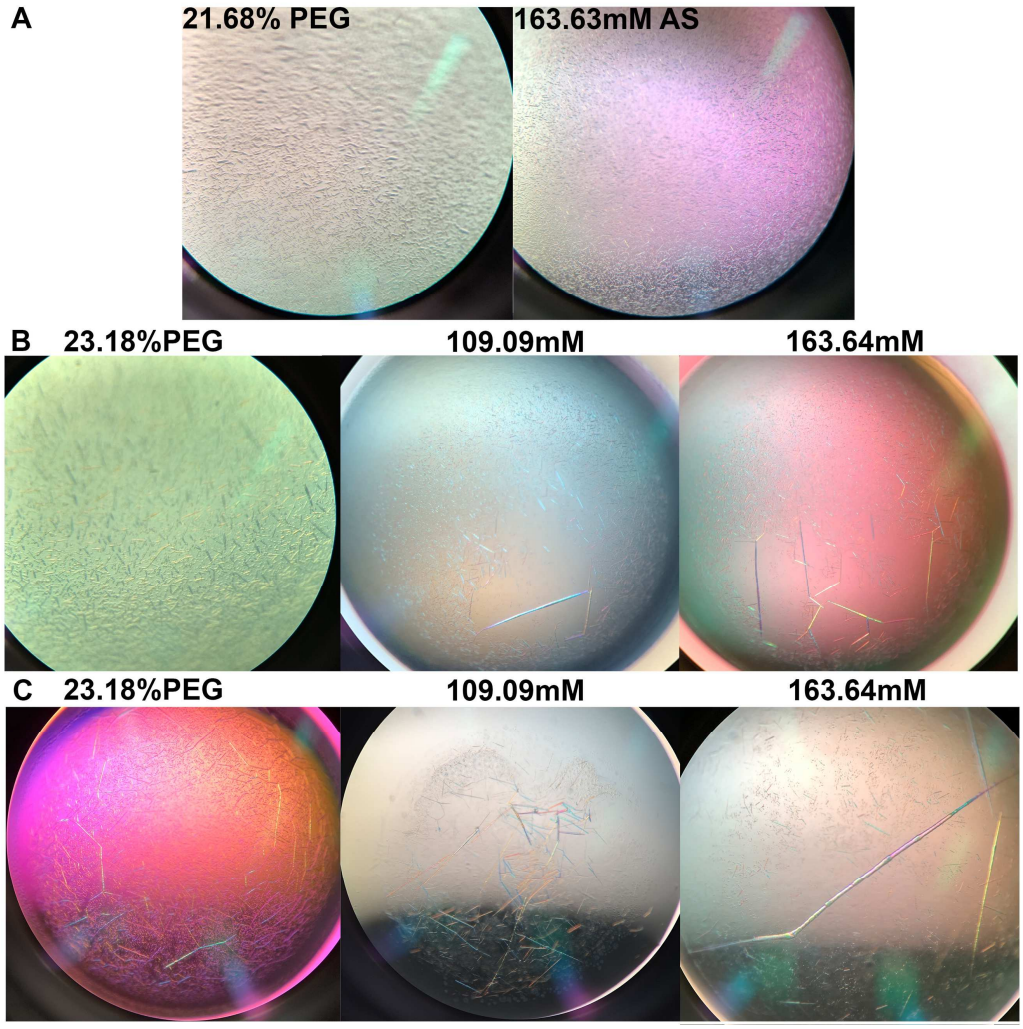

Figure 4.4.7 Summary of the initial C4 condition PEG250DME and ammonium sulfate screening results. Selected PEG and ammonium sulfate screening images taken A) 10 d after crystallization setup; B) 14 d after crystallization setup; C) 20 d after crystallization setup. All images were obtained using a cellular phone camera through the microscope eyepiece.

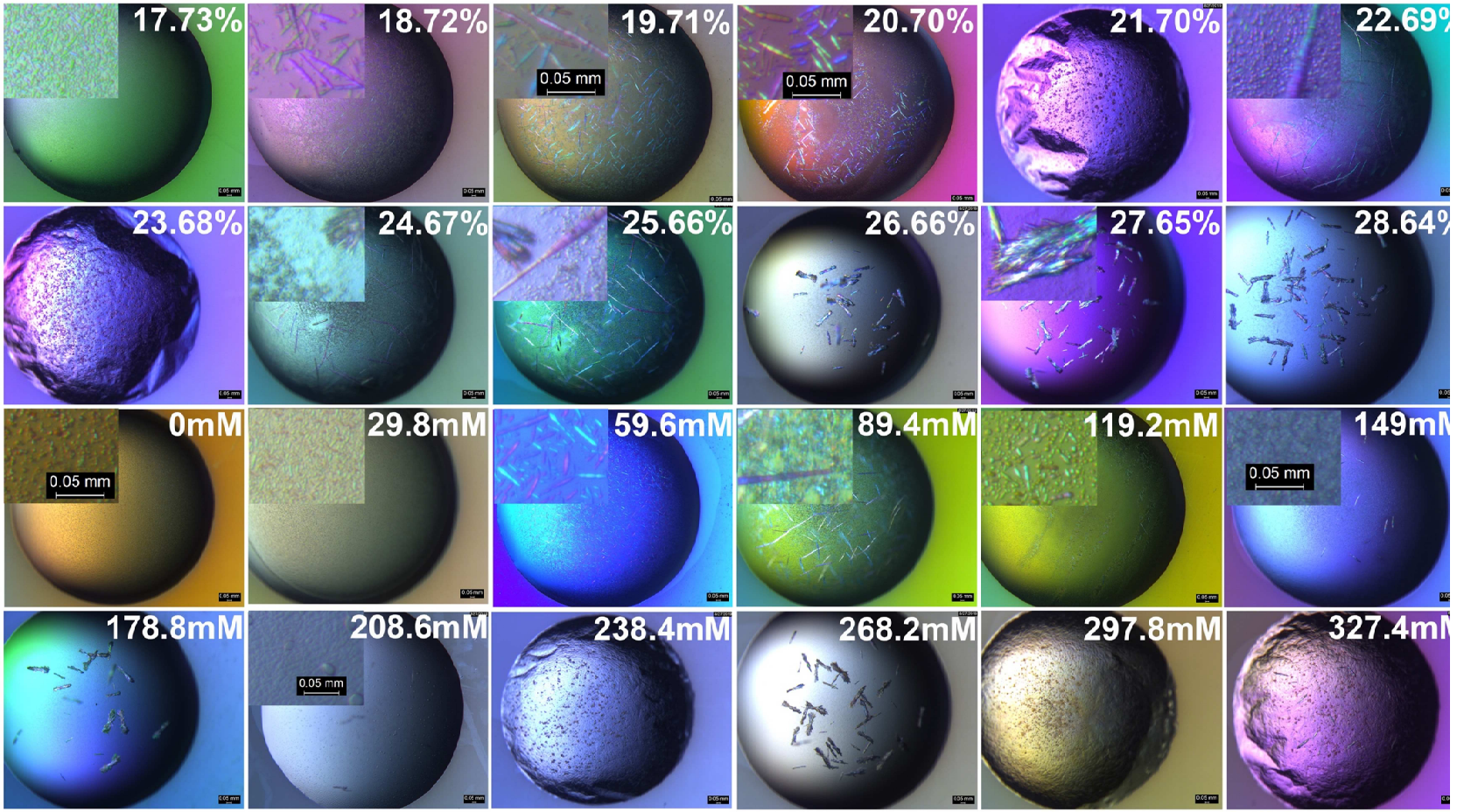

Figure 4.4.8 Summary of PEG250DME and ammonium sulfate narrow screening results under C4 crystallization condition at room temperature. The top two rows are PEG250DME screening at 17–29%; the bottom two rows are ammonium sulfate screening at 0–0.33 M. Images taken 11 d after setup of the plates.

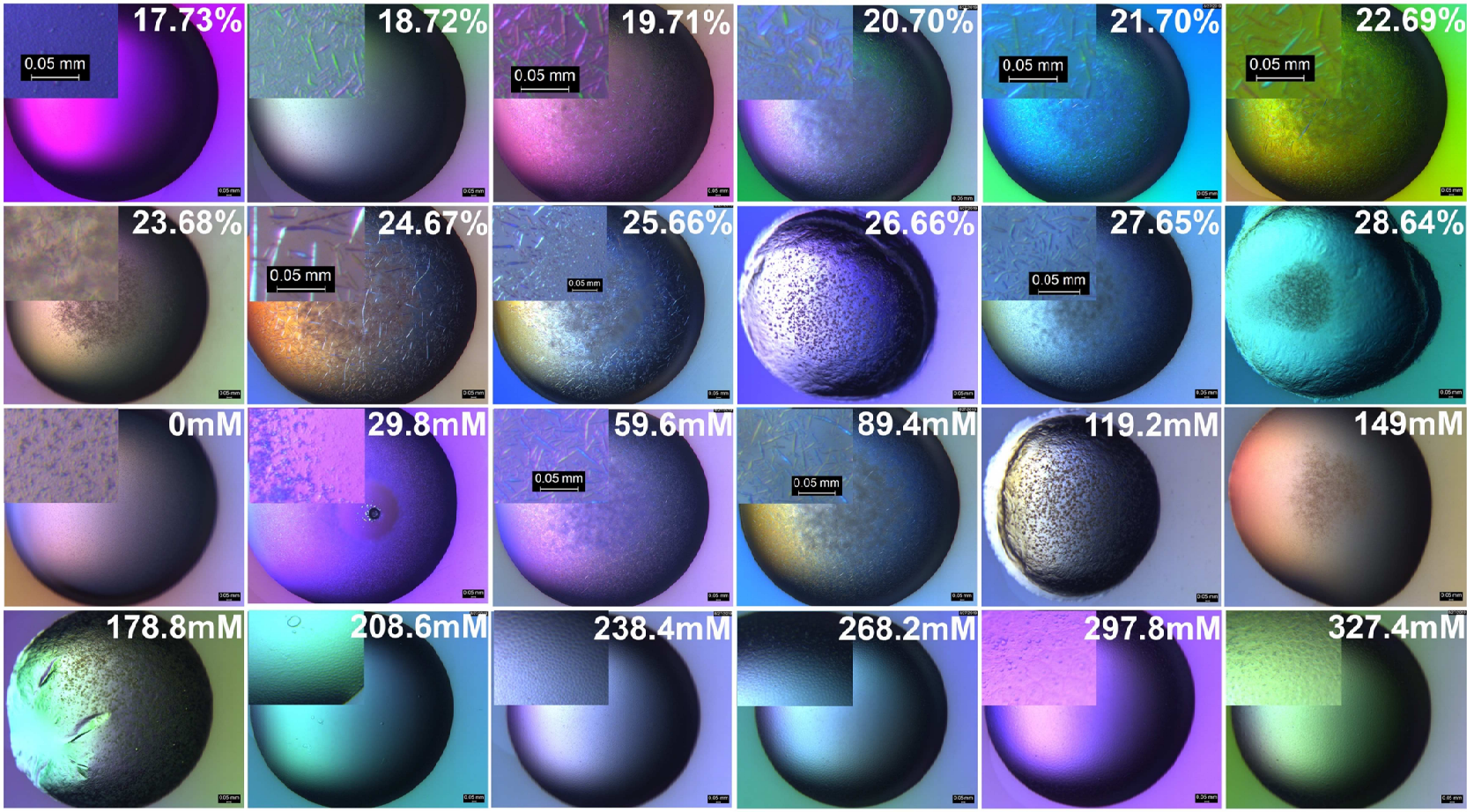

Figure 4.4.9 Summary of PEG250DME and ammonium sulfate narrow screening results under C4 crystallization condition in a 15°C incubator. The top two rows are PEG250DME screening at 17–29%; the bottom two rows are ammonium sulfate screening at 0–0.33 M. Images taken 11 d after setup of the plates.

Real-time Protein Crystallization Under Microscope

Next, whether intact pGC-A could be crystallized in the purification buffer by adding PEG3350 was explored. Both PEG3350 and sodium chloride concentration screenings were performed. The purification buffer-based crystallization screening was set up with 1 mg/mL protein concentration and incubated in the cold room at 4°C. Super-dense microcrystals grew in the crystallization drop with high PEG3350 concentrations (PEG 8.75% and PEG 9.3%) (Figure 4.4.10). Plates were placed in the cold room after imaging. During the second imaging time, many previously imaged drops became clear drops and reformed microcrystals in real-time.

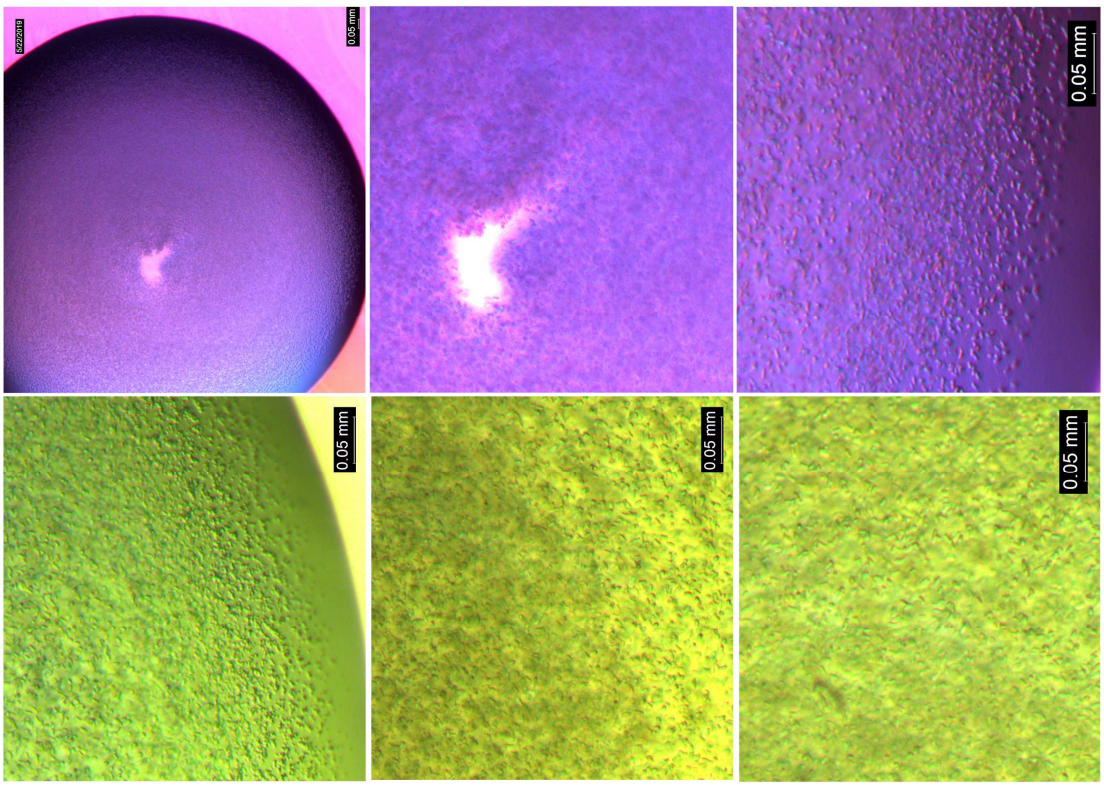

Figure 4.4.10 Super-dense rod-shape microcrystals observed in purification buffer-based crystallization condition.

This transition process was recorded by obtaining pictures and video recordings (Figure 4.4.11). Many rod-shaped microcrystals formed during the temperature transition

from 4°C to room temperatures. However, formed microcrystals began to dissolve after direct heating under the microscope for 45 min. To determine the exact temperature that triggered protein crystal nucleation, four different temperature conditions (4°C, 10°C, 15°C, and room temperature) were assessed to determine the growth of rod-shaped crystals under crystallization condition in the the purification buffer (Figure 4.4.12). However, only the dot-shaped microcrystals grew under these four temperatures. Under the room temperature conditions, the dot-shaped microcrystals were more packed in the center of the drop. Each room temperature grown dot featured strong birefringence, but the dots were less round (irregular shapes) than crystals grown at the three lower temperatures. In conclusion, the rod-shaped microcrystals grew during the temperature transition from 4°C to a higher temperature. The constant increase in temperature may stimulate thermodynamically favorable events by changing the solvent entropy and therefore kinetically improved formation of dot-shaped and rod-shaped crystals (Vekilov et al. 2002).

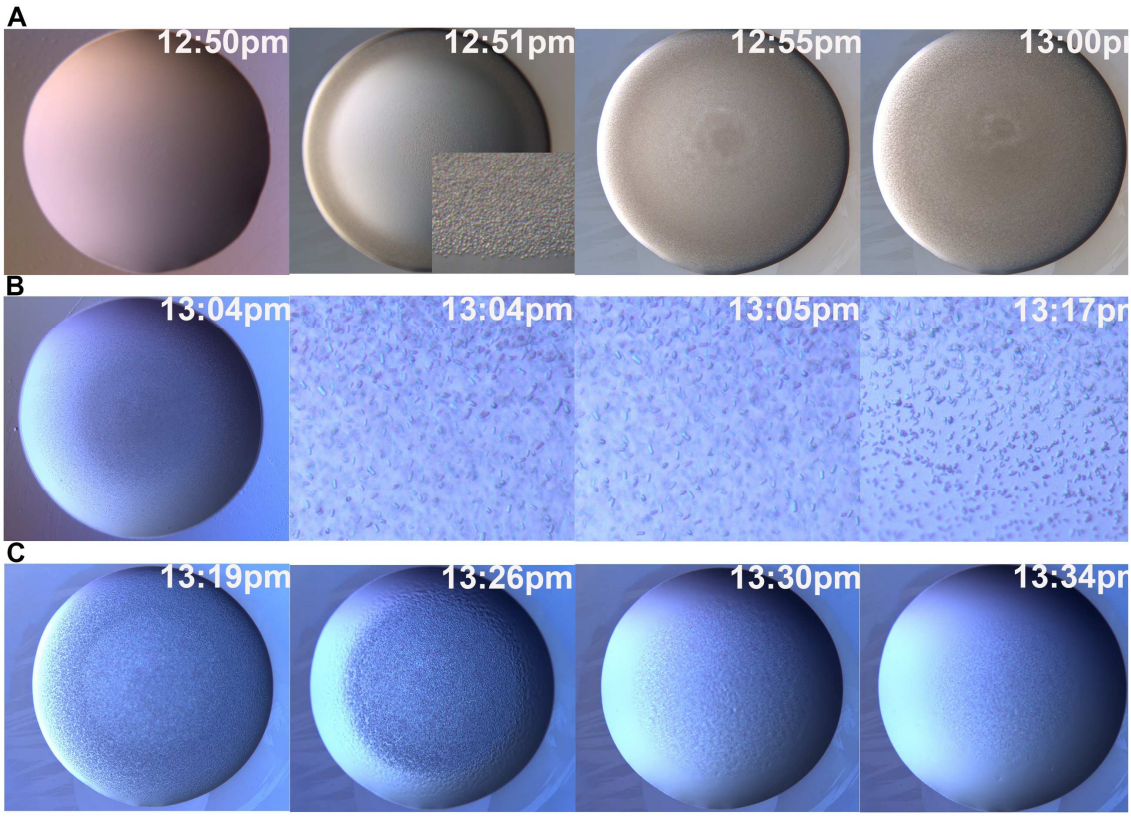

Figure 4.4.11 Rod-shaped microcrystals grown in real-time. The rod-shaped microcrystals were grown from clear drops during the temperature transition under the microscope. The plate was set up at 4°C and warmed up to room temperature under the microscope. However, when continuous heating of the plate to over 30°C occurred, the formed microcrystals dissolved in the drop. A) Crystals grow from a clear drop; B) many uniform rod-shaped microcrystals are formed in the drop; C) formed crystals start to dissolve.

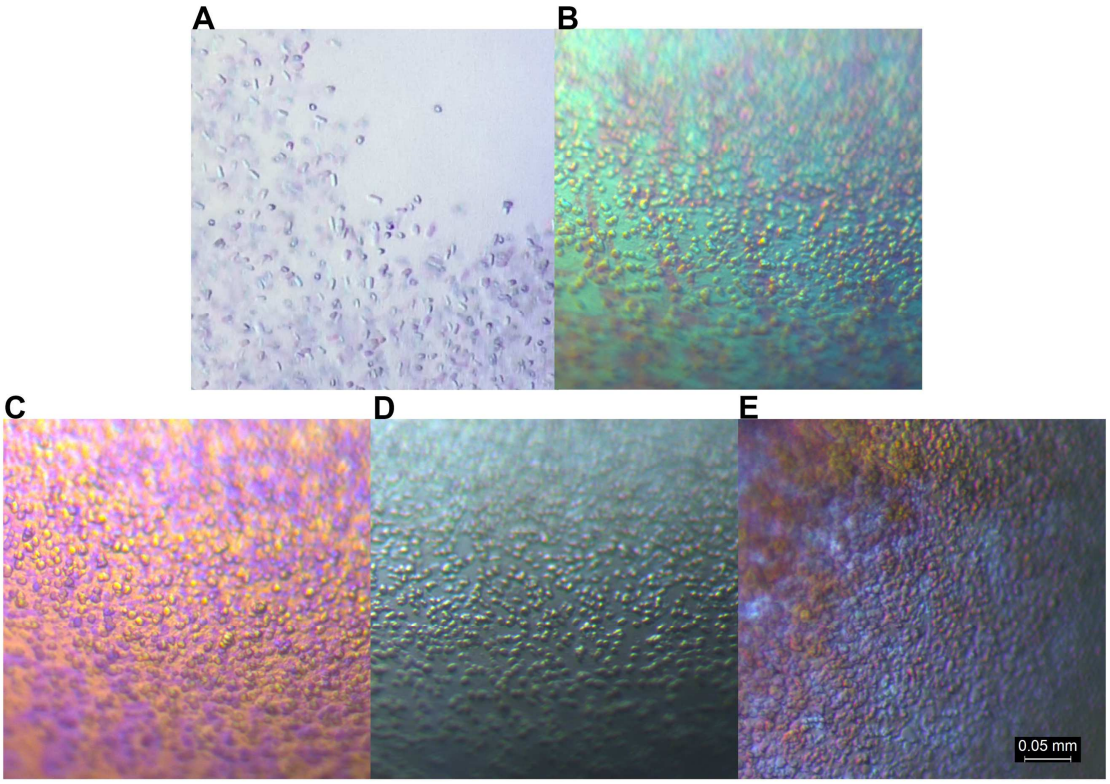

Figure 4.4.12 Summary of crystal shapes at different temperatures. The purification buffer-based crystallization condition was set up at five different temperatures. A) transition temperature from cold room (4°C) temperature to 23°C. Rod-shaped crystals formed; B) cold room (4°C) temperature. Dot-shaped crystals formed; C) 10°C incubator. Dot-shaped crystals formed; D) 15°C incubator. Dot-shaped crystals formed; E) room temperature (23°C). Super-dense dot-shaped crystals formed. All plates were crystallized with 1 mg/mL purified pGC-A.

Confirmation of pGC-A Protein Formed Crystals in F8 Condition

To confirm the needle-shaped shape crystals formed under the F8 condition were composed of purified intact pGC-A protein, the crystals were isolated from combined crystallization drops via centrifugation. The crystal pellet was washed with the precipitant solution and pelleted down again for later western blot signal determination (Figure 4.3.1). During the crystal collection and washing process, the sample from each step was mixed with SDS sample buffer to detect the pGC-A protein via western blotting. Figure

4.4.13 shows the anti-pGC-A western blot results and confirms that the protein crystals were formed by pGC-A. In Figure 4.4.13, well# 1, 2, and 4 had weaker western blot signals than Figure 4.4.13, well# 3, representing the pelleted crystals. The weaker pGC-A signal was caused by fewer pGC-A proteins in the sample. In Figure 4.4.13, well# 3 contained the strongest pGC-A signal during the crystal collection and washing process compared to well# 1 (the combined crystallization drops), # 2 (the supernatant collected from the centrifuged combined mix), and # 4 (the supernatant from washed crystal pellet), which indicated that most of the pGC-A was in the pelleted crystals (Figure 4.4.13, well# 3). The weak western blot signals in the supernatant samples (Figure 4.4.13, well# 2 and 4) were likely from the un-crystallized pGC-A protein in solution and shear force partially dissolved needle-shaped crystals from pipetting.

Unlike salt crystals, protein crystals are fragile and tend to break from pressure (Deller and Rupp 2014). The weak western blot signals in Figure 4.4.13, well# 4 were probably caused by resuspending the crystals with the precipitant solution by pipetting. Many microcrystals were stuck on the cover glass during the crystal drop collection/harvesting process, making transfer into the PCR tube a challenge. To test the pGC-A signal from the entire drop, SDS sample buffer was directly added to one crystallization drop (Figure 4.4.13, well# 5). The directly mixed crystallization drop (Figure 4.4.13, well# 5) had a much higher pGC-A western blot signal than the washed crystal pellet (Figure 4.4.13, well# 4). This was caused by formed crystals sticking on the hanging drop silicon coverslip, which were difficult to transfer into the PCR tube. In Figure 4.4.13, well# 6 represents purified pGC-A protein before crystallization. It served as a positive control in this experiment. The supernatant (Figure 4.4.13, well# 2) would

have a strong signal similar to the crystallization drops (Figure 4.4.13, well# 5) if pGC-A protein did not form crystals and stayed in solution. Additionally, in Figure 4.4.13, well# 3 and 5 showed strong pGC-A signals above the 250 kDa molecular weight in the western blot. However, the main protein signal band in sample of the purified protein before the crystallization (Figure 4.4.13, well# 6) was 120 kDa, consistent with the monomer size of intact pGC-A. These higher molecular weight bands represented higher oligomeric states of pGC-A. Indeed, natural dimerization of pGC-A was observed during pGC-A purification (Figure 2.4.4) and is also seen in the positive control (Figure 4.4.13, well# 6). Interestingly, in Figure 4.4.13, both well# 3 and 5 showed much higher oligomeric states in the crystal samples. Possibly, these higher molecular weight signals may represent the higher oligomeric states that dominated the protein crystals.

Notably, samples loaded on the SDS-PAGE were incubated in SDS without the heating process. The native protein oligomeric state may be stabilized against SDS denaturation in the crystals. Additionally, the crystallization plate was set up 3 weeks prior and stored in the cold room. Intact pGC-A is easily degraded in solution, but there was no pGC-A degradation signal (western blot signal below 120 kDa) detected in the western blot. Collectively, the western blot results indicated that pGC-A formed needle-shaped crystals under the F8 condition and that the protein is stabilized in its native oligomeric form in the crystals.

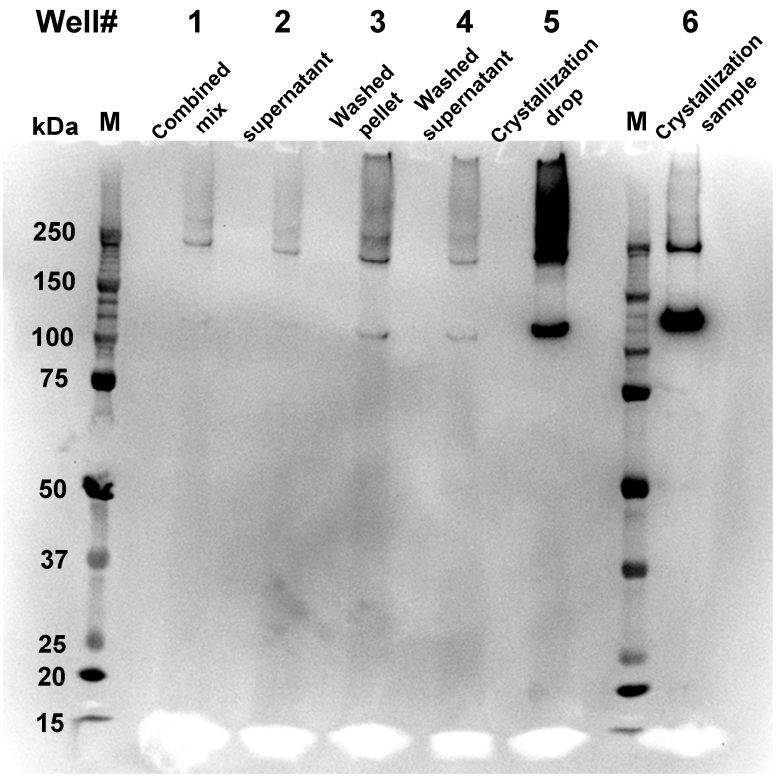

Figure 4.4.13 Presence of pGC-A in crystals was confirmed via western blot. The anti-pGC-A antibody was used in the western blot. The pGC-A monomer pGC-A is 120 kDa. Well #1: Combined mix: combined crystallization drops collected in the PCR tube. Well #2: Supernatant: the supernatant collected from the centrifuged combined mix; Well #3: washed pellet: the crystal pellet was washed with the precipitant solution and collected again via centrifugation; Well #4: washed supernatant: the supernatant from washed crystal pellet; Well #5: crystallization drop: one crystallization hanging drop directly mixed with SDS sample buffer; Well #6: crystallization sample: purified pGC-A before crystallization, serving as a positive control.

Diffraction Patterns of pGC-A Crystals under F8 Condition via Serial Crystallography

Harvested needle-shaped crystals grown under the F8 condition were shipped to APS at Argonne National Laboratory in a temperature-controlled container. Unlike other membrane proteins directly crystallized in the LCP, harvested pGC-A crystals must be embedded into LCP before delivering to the X-ray source. The final delivery LCP was prepared by mixing the crystal solution (40%) with molten monoolein lipid (60%) in a glass dual-syringe lipid mixer until a homogeneous transparent LCP was formed. The

detailed mixing steps are described in the methods section. A specialized LCP injector was used to deliver the LCP-crystal mixture in a serial crystallography experiment (Weierstall et al. 2014, Martin-Garcia et al. 2017). During the serial crystallography experiment at APS, LCP-embedded crystals were constantly jetting perpendicular to the X-ray beam. Unlike traditional single-crystal diffraction, in the serial crystallography method single crystal snapshots in random orientation are collected from a stream of microcrystals and data sets consist of diffraction patters from thousands of individual microcrystals. Each diffraction pattern is thereby obtained from a different crystal. During this specific experiment, LCP-embedded pGC-A crystals were jetted across the X-ray source for approximately 2 h. A total of 270,000 frames were recorded on the detector. Five diffraction patterns were detected for the first time for intact pGC-A crystals. Accordingly, three out of five recorded diffraction patterns could be indexed by the CrystFEL software (Figure 4.4.14) (White et al. 2016). Based on the three indexable diffraction patterns, the CrystFEL indexing results showed that the needle-shaped crystals likely represent crystals in the P1 space group. The crystal unit cell showed the cell length: a=67.6 Å, b=119.0 Å, and c=153.1 Å and the cell angles: alpha=72.8°, beta=80.6°, gamma=80.1°. The Matthew correlation coefficient predicted four copies of the monomer pGC-A (or a tetramer) in the unit cell with 48% solvent content. The highest diffraction spot reached 3 Å, which can be observed in the bottom image of Figure 4.4.14 B.

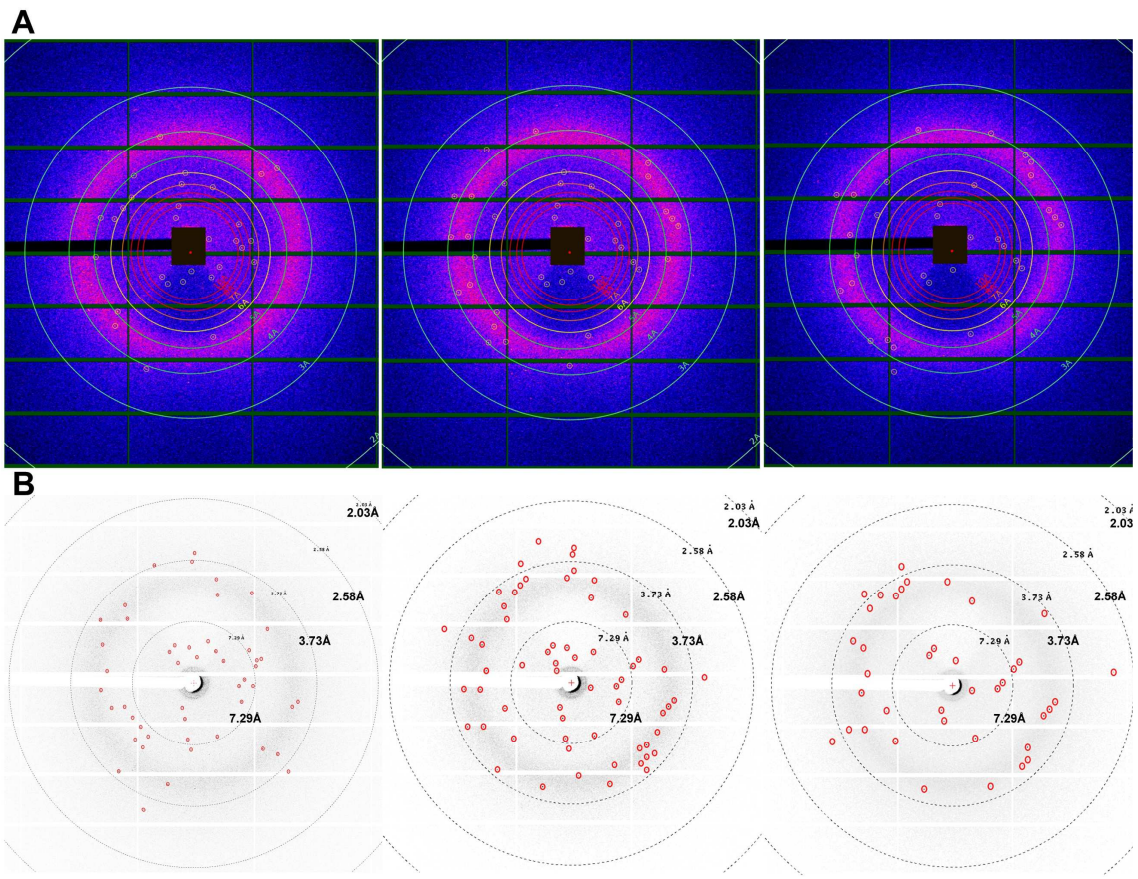

Figure 4.4.14 Three indexable diffraction patterns from pGC-A crystals collected via serial crystallography in APS. Three indexable diffraction patterns with the highest resolution of 3 Å. A) Blue/pink diffraction graphs show diffraction patterns analyzed using CrystFEL software. B) The original diffraction patterns in the white background images. All diffraction dots were manually circled in red for better visualization.

Liquid Protein Phase in Crystallization

During the crystallization optimization, an interesting phenomenon was observed; many of the needle-shaped crystals grew after phase separation. The crystallization drops contained two layers, and some formed a second oil-like phase. This is often observed in membrane protein crystallization containing a detergent. The high detergent/protein environment extrudes from the aqueous phase and creates "oil drops," which have a different density than the aqueous phase (Gabrielsen et al. 2009). The high protein

concentrated droplets subsequently shrank during the equilibration process and the protein crystals disappeared. The liquid phase crystallization was imaged during the C4 condition optimization process. Figure 4.4.15 shows that needle-shaped crystals were growing out from the "oil drops," and crystal growth directions were guided by the merged "oil drops" shapes. The more the "oil drops" merged into one big shape, the bigger were the needle-shaped crystals that formed. When the "oil drops" were well-isolated from each other, microcrystal clusters were formed (Figure 4.4.6 B).

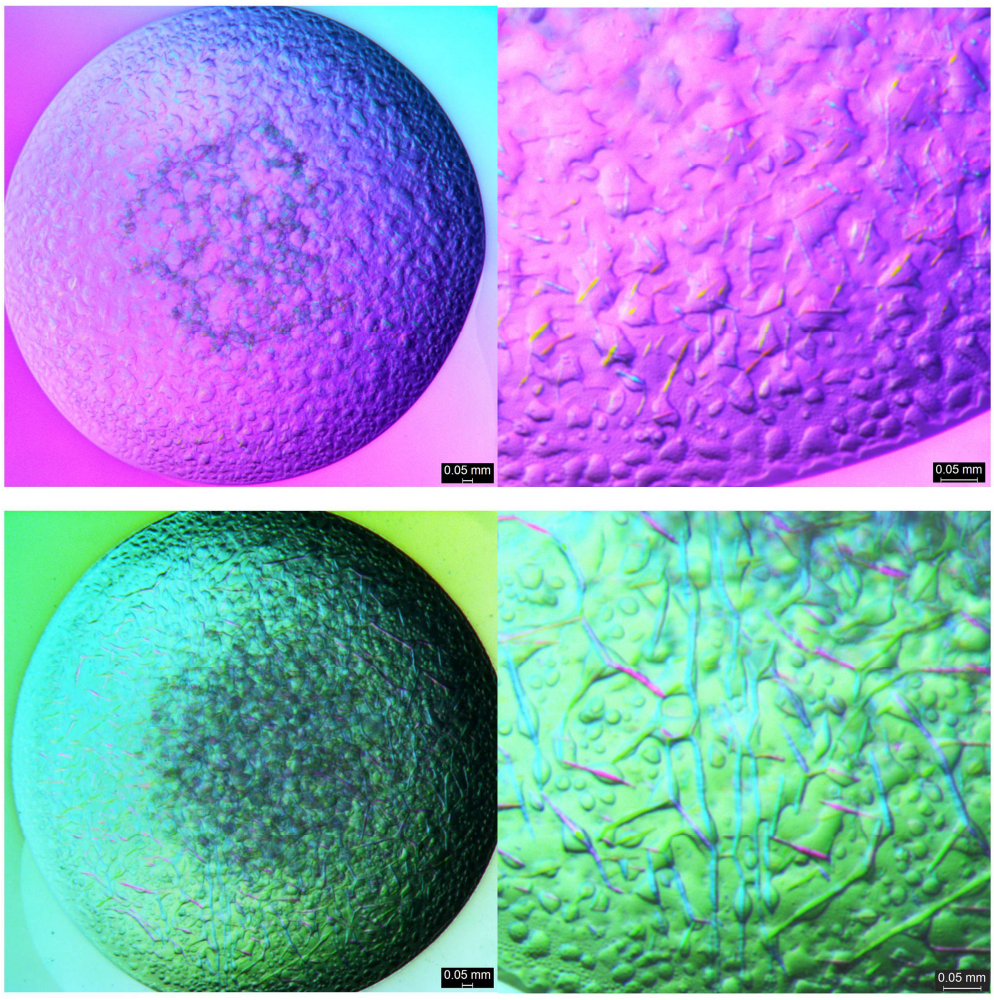

Figure 4.4.15 Liquid phase separation crystallization under the C4 condition. Needle crystals of intact pGC-A grown in contact with the surfaces of merged/closed oil drop inclusions.

4.5 Discussion

In the beginning, crystallization of intact human pGC-A was screened with the commercial membrane protein crystallization kit, MemGold2HT. Commercial screening kits are designed based on known crystallization conditions of multiple different proteins, which successfully yielded high-resolution protein structures deposited in the protein data bank (PDB). Rational screening of the target protein under a broad matrix screen could maximize the possibility of determining first suitable crystallization conditions. Depending on the transmembrane domain types, the membrane screening kits are designed specifically for either α-helix or β-barrel protein domains. Most pharmaceutical drug targets are α-helical membrane proteins, which play important physiological roles in humans (Parker and Newstead 2012). The β-barrel membrane proteins are mainly found in the bacterial outer membrane, which modulate bacterial pathogenicity and human metabolism (Newstead et al. 2008). There are limited commercial screening kits for membrane proteins. Understanding the concept of the sparse matrix design and choosing the correct screening kits are the first key steps in protein crystallization. The MemGold2 kits were designed based on the crystallization conditions of 254 α-helical membrane proteins for which structures are deposited in the PDB (2008–2012). Unlike the original MemGold kits, the updated MemGold2 kits contained more crystallization conditions derived from successful crystallization of channels, transporters, and structures from the family of GPCRs. (Parker and Newstead 2012). Intact pGC-A is a transmembrane domain peptide receptor, which contains one transmembrane helix/monomeric unit. Therefore, the MemGold2HT kits were the best starting point for initial crystallization screening.

Another rational approach in membrane protein crystallization screening is detergent screening. Up to 2012, the class of detergents called alkyl maltopyranosides were the most successful class of detergents used in transporter and channel membrane protein crystallization. The most commonly used alkyl maltopyranoside detergents are *n*-dodecyl-β-d-maltopyranoside (DDM) and *n*-decyl-β-d-maltopyranoside (DM) (Parker and Newstead 2012). With their long alkyl chains, they can completely cover the membrane integral region of the membrane protein thereby optimally maintaining the structural integrity and function of the membrane proteins. However, shorter chain detergents from smaller micelles and therefore have a higher probability of forming high diffraction crystals by facilitating the crystal packing process (Sonoda et al. 2010). Because of resolution considerations, the alkyl glucopyranoside detergent, *n*-octyl-β-d-and glucopyranoside (OG) is considered as one of the first choice for screening α-helix membrane proteins in the vapor diffusion method (Parker and Newstead 2012). In addition, the amine oxide, *n*-lauryl dimethylamine-*N*-oxide (LDAO), contains both positive and negative charged groups with an overall zero net charge in the polar head. LDAO is a nonionic detergent and provides intermediate harsh level between ionic and nonionic detergent (Seddon, Curnow, and Booth 2004). LDAO is one of the most successful detergents for both membrane protein purification and crystallization (Stetsenko and Guskov 2017). Intact pGC-A was successfully isolated and purified in DDM detergent. However, further crystallization detergent screening has not yet been performed and may lead to crystals diffracting to higher resolution in the future.

Ideally, the most efficient way to achieve initial crystallization screening is using an automatic screening robot. However, I chose to focus on initial manual screening after

initially attempting the robot screening technique for the following reasons. First, there is a reproducibility issue when screening for the crystallization conditions in small drops. The results from robot small volume screening are not fully reproducible when the drop size is increased in the vapor diffusion method (Newman, Xu, and Willis 2007). Second, robot drop mixing is not accurate. Some crystallization precipitate may not align or fully mix with the protein drops. Additionally, the crystallization solutions from commercial kits often contain chemicals that increase the viscosity which make automated droplet dispensing more challenging. My protein target pGC-A had a relatively large molecular size. There is a high chance that most of the initial screening drops precipitate after direct mixing. To avoid precipitation during the initial screening, I diluted the commercial kit solutions to 80% concentration before mixing it with protein. It is difficult to prepare the "personalized" diluted screening solution for robot screening. Lastly, intact pGC-A tends to degrade in solution; I switched to manual screening of the MemGold2 kits in a 24-well plate in a cold room and kept the protein on ice at all times.

As mentioned in Chapter 2, the purified pGC-A yield was low. A concentrated 2 mg/mL pGC-A sample from 25 g transfected insect cells was only sufficient to set up one 24-well plate. Generally, the protein concentration used for crystallization depends on the natural solubility of the protein. Soluble proteins are usually crystallized at much higher protein concentrations. Typically, the protein concentration varies from 5 to 50 mg/mL, and a higher concentration usually yields better crystallization results (Dessau and Modis 2011). However, over-concentrating the protein also increases instability and aggregation. Besides the protein solubility, the molecular size of the protein also affects the protein solubility in crystallization. One article provides a general suggestion that

smaller proteins below 10 kDa may require 20–50 mg/mL for crystallization, whereas larger proteins above 30 kDa may require only 2–5 mg/mL concentration (Kermani 2020). This suggestion is a general guide since each protein is unique and many trial and error attempts are still required for targeted protein. For example, microcrystals of the bovine rhodopsin photoreceptor were crystallized in a wide protein concentration range from 0.4 to 15 mg/mL (Yurkova, Demin, and Abdulaev 1990).

Intact pGC-A was concentrated to 2 mg/mL for the initial screening as aggregation was observed at higher protein concentrations. Later, the optimization screening showed that the protein at 1 mg/mL concentration also produced the same needle-shaped crystal shower with less precipitate zone in the center of the drop (Figure 4.4.4 and 4.4.5) under the F8 condition. To retain the protein sample for additional optimization screening, I lowered the protein concentration to 1 mg/mL for further screening. This shows that the solubility of pGC-A is extremely low. Interestingly, the pGC-A protein can also form crystals when the protein concentration was lowered to 0.6 mg/mL under C4 conditions, purification conditions, and F11 conditions. As mentioned earlier, although the bovine rhodopsin was crystallized at a concentration from 0.4 to 15 mg/mL, the best crystals with high diffraction were obtained at a 2.5 mg/mL protein concentration (Yurkova, Demin, and Abdulaev 1990). In future work, the intact pGC-A should be regularly screened at 1–1.5 mg/mL to minimize protein sample consumption and have a higher chance of forming better-diffracted crystals.

During the initial screening, more than 10 conditions were identified to be promising for further optimization. I only listed the top seven in table 3.4.1. Interestingly, these seven conditions are similar to each other. Depending on the components, they can

be sub-classified into two groups. Based on the buffer pH, these conditions were either at pH 7.0 with +/- 0.5 range or pH 4.5 with +/- 0.5 range. Based on the precipitant PEG, the majority conditions contained a lower percentage of high molecular PEG molecules (PEG3350 or 4000) or a higher percentage of low molecular PEG molecules (PEG250 or PEG600). Based on the salt, the ammonium cation and sodium cation were frequently added monovalent ions in these conditions. Indeed, ammonium sulfate facilitated protein crystallization under both F8 (Figure 4.4.5) and C4 (Figure 4.4.9) conditions. Adding a suitable amount of ammonium sulfate stimulated protein crystals to form after phase separation. However, a higher concentration of ammonium sulfate eventually prevented crystal growth. The ammonium sulfate concentration followed the shape of the normal distribution Bell curve in statistics. At both ends of the curve, too low (0–50 mM) or too high (300–600 mM) concentrations of ammonium sulfate hampered protein crystallization. Precipitants are critical elements in protein crystallization. Precipitant chemicals reduce protein solubility by interaction with water molecules. The classical precipitants in crystallization are salt and PEG (McPherson 2017). Varying the concentrations of salt and PEG during crystallization optimization is crucial in protein crystallization.

During the crystallization optimization screenings, many microcrystals favored growth on the glass cover slides; this was mainly observed with lower concentrations of PEG (Figure 4.4.18; 17.73%) or salt (Figure 4.4.5; 163 mM). Microcrystals grown on the glass surface were uniformly sized with a high crystal density. However, it was challenging to collect these crystals during the crystal harvesting step. These densely grown microcrystals were tightly bound to the crystallization surface (cover glass). In

theory, protein crystals binding to the contact hardware surface is more likely to occur when using the sitting drop method. The hanging drop method is supposed to reduce the crystal binding to the surface as crystals would settle to the bottom of the drop where they are not in contact with the glass cover surface.

There are several explanations for why the crystals in this experiment favored growth on the cover glass surface. First, the cover glass serves as nucleation site for the formation of protein crystal nuclei, which then continue to grow along the surface. Second, phase separation is the first step before protein crystals grow from the "oil drops" during crystallization. The majority of "oil drops" are created on the contact surface. In other words, the location of the "oil drops" determines the location of crystal growth. The liquid phase separation is caused by the difference in density between the aqueous precipitant solution and the detergent enriched low -density phase separated droplet (oil drop) (Kuznetsov, Malkin, and McPherson 2001). Membrane proteins often concentration in the detergent rich phase. Cleary, needle-shaped crystals formed under the C4 condition showed that needle-shaped crystals were grown in contact with the surface or even within the detergent-rich droplets (Figure 4.4.15). Additionally, intact pGC-A is a membrane protein that inherently prefers a hydrophobic environment/surface. The cover glass I used was silicon coated to provide a hydrophobic surface. The siliconized slides assisted the hanging drop to from a small well defined droplet shape and to not flatten on the glass. However, the hydrophobic surface may trigger high detergent droplets with my protein enriched in it to cluster together on the contact surface.

During the crystal diffraction experiments, there were only five diffraction patterns recorded from a total of 270,000 frames collected in a 2-hour serial crystallography diffraction experiment at the APS synchrotron. The hit rate for this experiment was very low with 0.00185%. There are some reasons that could explain the low hit rate. Each F8 condition crystallization drop contained hundreds of needle-shaped crystals, but most tended to bind to the cover glass during crystal harvesting. Thereby, the crystal harvesting process was not efficient and left many crystals on the cover glass. Furthermore, the harvested crystals were not delivered for X-ray diffraction immediately. Instead, the harvested needle-shaped microcrystals were shipped to APS beamline first and then embedded within LCP delivery media. As exact temperature stability cannot be maintained during shipping the majority of the crystals may have been disintegrated during shipment.

Crystals grown from vapor diffusion undergo a "brutal" harvesting process where the solvent can evaporate and thereby are more prone to damage during harvesting compared to crystals grown from larger volume from batch methods. The subtle environmental change may dissolve crystals once the cover slides with crystal drops are removed from the precipitant solutions (Rayment 2002a). Obtained crystals were checked under the microscope before delivery media mixing. The morphology of harvested crystals appeared poor compared the crystals before harvesting.

Furthermore, the delivery medium could also have led to disintegration of most of the crystals. LCP would be a great delivery media for a membrane protein naturally crystallized in LCP. However, embedding of pre-grown protein crystals in detergent micelles required multi-step mixing in a glass dual-syringe lipid mixer. A strong shear

force is created whenever the sample passes through the coupler, which connects the dual-syringe mixer to the other side of the syringe. In addition to the challenge of crystal embedding LCP is a lipidic membrane-mimetic matrix environment. As mentioned earlier in the Introduction section of Chapter 4, membrane protein crystal packing is different between LCP and detergent micelle solution (Figure 4.2.3). Briefly, LCP mimics the biological membrane assembly. The crystal packing forces are lipid–protein and protein–protein interactions. However, crystal grown in form of a protein-detergent micelle are mainly formed by polar protein–protein interactions. Detergent molecules cover the hydrophobic surface of the membrane proteins. Mixing the lipid molecules with grown crystals to form embedded LCP media for crystal delivery might induce the risk of lipids in the LCP slowly replacing the protein detergent molecules, eventually damaging the crystal structure. Finally, membrane crystals grown from detergent solutions often have larger solvent channels than crystals formed in LCP (Birch et al. 2018). The crystal packing order and high solvent content may affect crystal diffraction quality.

4.6 Conclusion

In summary, I successfully selected seven promising initial crystallization conditions for intact human pGC-A crystallization. Among these, two conditions were further optimized by variation of salt, PEG, and buffer pH. The concentrations of salt and PEG are crucial to protein crystallization. However, the pH plays a less important role as crystals were observed over a large pH range from pH 4.5 to pH 8. At lower protein concentration similar microcrystals were observed during the crystallization optimization process compared to a higher protein concentration sample. To conserve protein samples for further screenings, the protein concentration was adjusted from 2 mg/mL to 1 mg/mL.

Under the C4 condition, the lowest (0.6 mg/mL) protein concentration enabled the formation of protein crystals under higher PEG concentration conditions. During rational crystallization screening, uniformly sized rod-shaped microcrystals grew in the purification buffer with PEG3350 as precipitant during a transition temperature range from 4°C to room temperature. However, only dot-shaped crystals formed under stable temperature conditions. Needle-shaped crystals from the F8 condition were harvested and used to confirm that these microcrystals were composed of intact pGC-A protein by anti-PGC-A western blotting. This is the first study that successfully crystallized intact human pGC-A for future structure determination. After confirming that the needle-shaped crystals were formed by my pGC-A protein, four crystal plates from the F8 condition were harvested. Crystals were embedded within LCP delivery media. During the 2 h serial crystallography diffraction experiment, five diffraction patterns were detected, and three of the five patterns were indexable. The highest diffraction resolution spot reached 3 Å. These diffraction patterns showed a promising sign that this crystallization condition could lead to the determination of the intact human pGC-A structure in the future. Although the crystallization conditions were not perfect, further optimizations including, additive screening to grow higher-quality crystals could be undertaken.

CHAPTER 5

OUTLOOK AND CONCLUSION

5.1 Future Directions

The successfully established intact human pGC-A expression and purification methods have paved the road for future protein structure determination and functional studies. Promising pGC-A protein crystallization conditions were selected. Confirmed pGC-A crystals from F8 conditions showed diffraction up to 3 Å. These results may form the basis for the structure determination of intact human pGC-A by X-ray crystallography in the future. Based on the known diffraction data analysis, the crystal solvent content was assumed to be 48%. This valuable information provides an opportunity to diffuse the peptide igands into the crystals. Direct ligand diffusing or co-crystallization methods may provide a shortcut for acquiring the active ANP-pGC-A complex structure. In addition to resolving structures by X-ray crystallography, other experiments could be undertaken in the future including cryo-EM studies.

Expression of Intact pGC-A in Mammalian Cell Lines

Although I optimized protein expression from construct design to insect cell culture, the intact human pGC-A expression level remained low. As I discussed in Chapter 2, glycosylation may be the underlying cause for this issue. To address the protein yield issue, I propose to contact research labs studying glycoengineered insect cells. It would be helpful to verify if glycosylation is the underlying cause for the low protein yield. Additionally, switching to a mammalian cell line may boost pGC-A expression. In previous intact pGC-A cGMP functional studies, almost all studies expressed intact pGC-A in mammalian cells. To corroborate my findings with other

studies, I suggest reproducing my experiments using *in vitro* functional assays with protein purified from mammalian cells. Furthermore, reproduction of the previous functional studies and delicately quantifying the guanylyl cyclase activity in different modulators (e.g., ligand, ATP, and protein kinase enzyme) would be informative. First, I propose measuring the pGC-A activity level *in vivo* with or without ligand. To determine how phosphorylation affects pGC-A activity assays, microcystin, the phosphatase inhibitor, could also be tested. The previous *in vivo* experimental setup should be duplicated and tested with and without a microcystin inhibitor. Second, I suggest measuring the detergent-purified pGC-A activity level *in vitro* with ligand, ligand+ATP, ligand+AMP-PNP, and ligand+ATPgS. To determine whether phosphorylation occurs before or after ATP binding, the previous *in vitro* experimental setup should be duplicated and tested with and without protein kinase enzyme incubation, which should also include mass spectroscopy studies to identify the phosphorylation site and quantify the phosphorylation level.

Further Optimization Protein Crystallization

In the initial screening, there were seven promising candidate conditions; however, I only fully optimized two of them. There is a possibility that other crystallization conditions could grow better diffracting protein crystals. Additionally, the protein detergent composition could be optimized. Other short-chain detergents could be tested during the final gel filtration or in the crystallization step to partially replace the DDM for potential better crystal packing. Furthermore, additive screening could help to grow higher-quality crystals. Other than optimizing the protein crystal condition itself, screening beamtime at the synchrotron is also essential. Different cryoprotectant

solutions could be screened during such a screening beamtime. I observed that CryoProtX™ (Molecular Dimensions, MD1-61) is a good cryoprotectant kit for cryo screening.

5.2 Outlook and Conclusion

During my 5 years of research, I have established the Bac-to-Bac™ insect cell expression platform in my laboratory and successfully expressed intact human pGC-A in Sf9 cells. After generating a high-titer recombinant virus, protein expression optimization was performed by testing different virus concentrations and optimized the time for cell harvest. Additionally, this dissertation presents the first purification method for structure determination of intact human pGC-A. This purification method mainly includes a high salt membrane wash, membrane solubilization, His-tag affinity column, and gel filtration chromatography. The successfully established intact human pGC-A expression and purification methods have paved the road for future protein structure determination and functional studies.

Next, I analyzed different pGC-A oligomeric states observed during Suprose6 size exclusion chromatography. The results revealed that the higher oligomeric state of pGC-A represented the tetramer. Given the conflicting results when compared with the current proposal in the literature that pGC-A is a homodimer, different oligomeric peak fractions eluted from the Superose6 column were carefully characterized. Surprisingly, my results suggested that protein concentration played a stimulatory role in converting pGC-A monomers into a tetramer. Different detergent concentrations were screened to rule out the possibility that the tetramer was an artificial product. The detergent concentration screenings results revealed that a standard 0.05% DDM with 0.01% CHS

detergent concentration was sufficient for stabilizing the protein. The high resolution clear native gel electrophoresis analysis studies suggested that the dimer was not favored during oligomerization, and the "missing" dimer peak could not be attributed to the column resolution limit or overlapped tetramer and monomer peaks. Subsequently, the DTT reducing reagent incubation experiment suggested that other noncovalent interactions could oligomerize monomers into tetramers.

Based on my study and knowledge that dimer formation is required for pGC-A functionality, I proposed the first tetramer complex model composed of two functional subunits (homodimer). Intact pGC-A may functionally work as a homodimer, but the homodimer could be part of the tetramer complex. Inside the tetramer complex, four copies of identical pGC-A monomers provide more ligand binding positions possibilities for one ligand. Additionally, the released ligand from one functional subunit (homodimer) may recycle to another pair in the tetramer complex. The combinations and potential cooperativity of ligand binding could further increase pGC-A binding efficiency and ligand sensitivity.

This dissertation presents the first *in vitro* proof that purified intact human pGC-A tetramer is active in a detergent micelle solution, which indirectly rejected the tetramer as a protein aggregation theory. To verify if an ATP molecule played an allosteric regulatory role in the pGC-A activity, tetramer and monomer samples were incubated with or without ATP during the cGMP functional assay. There was no significant cGMP yield in the monomer sample with or without ATP. The results indirectly suggested that the oligomerization of pGC-A is a ligand-independent event.

Currently, ATP-dependent pGC-A signal transduction is a two-step ordered mechanism. This dissertation also presents an optimized mechanism. Herein, the cGMP functional assay results suggested that three modulators regulate different activation levels in intact pGC-A. The first level is the binding of the ligand to the ECD, which moderately activates the pGC-A. The second level is the binding of ATP to the ATP-binding pocket at the ICD, which acts as an allosteric modulator and boosts protein activity. The third level is phosphorylation of pGC-A, which fully activates the guanylyl cyclase functionality.

Next, to bridge the structure information gap and verify the signal transduction mechanism, purified intact pGC-A was crystallized for structure determination. During the crystallization condition screening process, I successfully selected seven suitable initial crystallization conditions for intact human pGC-A crystallization. The further crystallization modification results revealed that intact human pGC-A was successfully crystallized for the first time towards future structure determination. After confirming that the needle-shaped crystals were formed by the targeted protein, harvested crystals were embedded with LCP delivery media for X-ray diffraction. During the 2 h serial crystallography diffraction experiment, five diffraction patterns were detected, of which three were indexable. The highest diffraction resolution spot reached 3 Å. This crystallization condition could lead to the determination of the intact human pGC-A structure.

This work will lead in the future to the structure determination of intact human pGC-A and thereby can provide structural information on the protein signal transduction mechanism. Furthermore, precise amino acid point mutation experiments could be

designed to study the signal transduction mechanism. Intact pGC-A presents a fine-tuned activity mechanism. Different modulators may initiate different strength activity levels of pGC-A. With the inactive pGC-A and active ANP-pGC-A complex structure information, more details will be discovered about the function and the activation mechanisms of pGC-A in the near future.

REFERENCES

Abe, T., K. Nishiyama, R. Snajdar, X. He, and K. S. Misono. 1993. "Aortic smooth muscle contains guanylate-cyclase-coupled 130-kDa atrial natriuretic factor receptor as predominant receptor form. Spontaneous switching to 60-kDa C-receptor upon cell culturing." *Eur J Biochem* 217 (1):295-304. doi: 10.1111/j.1432-1033.1993.tb18246.x.

Abiko, L. A., M. Rogowski, A. Gautier, G. Schertler, and S. Grzesiek. 2021. "Efficient production of a functional G protein-coupled receptor in E. coli for structural studies." *J Biomol NMR* 75 (1):25-38. doi: 10.1007/s10858-020-00354-6.

Achmuller, C., W. Kaar, K. Ahrer, P. Wechner, R. Hahn, F. Werther, H. Schmidinger, M. Cserjan-Puschmann, F. Clementschitsch, G. Striedner, K. Bayer, A. Jungbauer, and B. Auer. 2007. "N(pro) fusion technology to produce proteins with authentic N termini in E. coli." *Nat Methods* 4 (12):1037-43. doi: 10.1038/nmeth1116.

Ahamed, T., S. Chilamkurthi, B. K. Nfor, P. D. Verhaert, G. W. van Dedem, L. A. van der Wielen, M. H. Eppink, E. J. van de Sandt, and M. Ottens. 2008. "Selection of pH-related parameters in ion-exchange chromatography using pH-gradient operations." *J Chromatogr A* 1194 (1):22-9. doi: 10.1016/j.chroma.2007.11.111.

Al-Sabah, S., L. Adi, M. Bunemann, and C. Krasel. 2020. "The Effect of Cell Surface Expression and Linker Sequence on the Recruitment of Arrestin to the GIP Receptor." *Front Pharmacol* 11:1271. doi: 10.3389/fphar.2020.01271.

Alenghat, F. J., and D. E. Golan. 2013. "Membrane protein dynamics and functional implications in mammalian cells." *Curr Top Membr* 72:89-120. doi: 10.1016/B978-0-12-417027-8.00003-9.

Alexander, M. R., J. W. Knowles, T. Nishikimi, and N. Maeda. 2003. "Increased atherosclerosis and smooth muscle cell hypertrophy in natriuretic peptide receptor A-/-apolipoprotein E-/- mice." *Arterioscler Thromb Vasc Biol* 23 (6):1077-82. doi: 10.1161/01.ATV.0000071702.45741.2E.

Ali, I., H. Y. Aboul-Enein, P. Singh, R. Singh, and B. Sharma. 2010. "Separation of biological proteins by liquid chromatography." *Saudi Pharm J* 18 (2):59-73. doi: 10.1016/j.jsps.2010.02.001.

Alioua, A., Y. Tanaka, M. Wallner, F. Hofmann, P. Ruth, P. Meera, and L. Toro. 1998. "The large conductance, voltage-dependent, and calcium-sensitive K+ channel, Hslo, is a target of cGMP-dependent protein kinase phosphorylation in vivo." *J Biol Chem* 273 (49):32950-6. doi: 10.1074/jbc.273.49.32950.

American Heart Association Nutrition, Committee, A. H. Lichtenstein, L. J. Appel, M. Brands, M. Carnethon, S. Daniels, H. A. Franch, B. Franklin, P. Kris-Etherton, W. S. Harris, B. Howard, N. Karanja, M. Lefevre, L. Rudel, F. Sacks, L. Van

Horn, M. Winston, and J. Wylie-Rosett. 2006. "Diet and lifestyle recommendations revision 2006: a scientific statement from the American Heart Association Nutrition Committee." *Circulation* 114 (1):82-96. doi: 10.1161/CIRCULATIONAHA.106.176158.

Anand-Srivastava, M. B. 2005. "Natriuretic peptide receptor-C signaling and regulation." *Peptides* 26 (6):1044-59. doi: 10.1016/j.peptides.2004.09.023.

Andrell, J., and C. G. Tate. 2013. "Overexpression of membrane proteins in mammalian cells for structural studies." *Mol Membr Biol* 30 (1):52-63. doi: 10.3109/09687688.2012.703703.

Anglin, T. C., and J. C. Conboy. 2008. "Lateral pressure dependence of the phospholipid transmembrane diffusion rate in planar-supported lipid bilayers." *Biophys J* 95 (1):186-93. doi: 10.1529/biophysj.107.118976.

Antonakoudis, G., L. Poulimenos, K. Kifnidis, C. Zouras, and H. Antonakoudis. 2007. "Blood pressure control and cardiovascular risk reduction." *Hippokratia* 11 (3):114-9.

Antos, L. K., S. E. Abbey-Hosch, D. R. Flora, and L. R. Potter. 2005. "ATP-independent activation of natriuretic peptide receptors." *J Biol Chem* 280 (29):26928-32. doi: 10.1074/jbc.M505648200.

Arnau, J., C. Lauritzen, G. E. Petersen, and J. Pedersen. 2006. "Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins." *Protein Expr Purif* 48 (1):1-13. doi: 10.1016/j.pep.2005.12.002.

Ashman, D. F., R. Lipton, M. M. Melicow, and T. D. Price. 1963. "Isolation of adenosine 3', 5'-monophosphate and guanosine 3', 5'-monophosphate from rat urine." *Biochem Biophys Res Commun* 11:330-4. doi: 10.1016/0006-291x(63)90566-7.

Assenberg, R., P. T. Wan, S. Geisse, and L. M. Mayr. 2013. "Advances in recombinant protein expression for use in pharmaceutical research." *Curr Opin Struct Biol* 23 (3):393-402. doi: 10.1016/j.sbi.2013.03.008.

Attwood, David, and Anne Sakdinawat. 2017. *X-rays and extreme ultraviolet radiation: principles and applications*: Cambridge university press.

Aumiller, J. J., H. Mabashi-Asazuma, A. Hillar, X. Shi, and D. L. Jarvis. 2012. "A new glycoengineered insect cell line with an inducibly mammalianized protein N-glycosylation pathway." *Glycobiology* 22 (3):417-28. doi: 10.1093/glycob/cwr160.

Ayyar, B. V., S. Arora, C. Murphy, and R. O'Kennedy. 2012. "Affinity chromatography as a tool for antibody purification." *Methods* 56 (2):116-29. doi: 10.1016/j.ymeth.2011.10.007.

Baghban, R., S. Farajnia, M. Rajabibazl, Y. Ghasemi, A. Mafi, R. Hoseinpoor, L. Rahbarnia, and M. Aria. 2019. "Yeast Expression Systems: Overview and Recent Advances." *Mol Biotechnol* 61 (5):365-384. doi: 10.1007/s12033-019-00164-8.

Baillie, G. S., G. S. Tejeda, and M. P. Kelly. 2019. "Therapeutic targeting of 3',5'-cyclic nucleotide phosphodiesterases: inhibition and beyond." *Nat Rev Drug Discov* 18 (10):770-796. doi: 10.1038/s41573-019-0033-4.

Bakshi, Vivek. 2006. *EUV sources for lithography*. Vol. 149: SPIE press.

Baneyx, F., and M. Mujacic. 2004. "Recombinant protein folding and misfolding in Escherichia coli." *Nat Biotechnol* 22 (11):1399-408. doi: 10.1038/nbt1029.

Bassett, D. E., Jr., M. A. Basrai, C. Connelly, K. M. Hyland, K. Kitagawa, M. L. Mayer, D. M. Morrow, A. M. Page, V. A. Resto, R. V. Skibbens, and P. Hieter. 1996. "Exploiting the complete yeast genome sequence." *Curr Opin Genet Dev* 6 (6):763-6. doi: 10.1016/s0959-437x(96)80033-5.

Bergfors, T. 2003. "Seeds to crystals." *J Struct Biol* 142 (1):66-76. doi: 10.1016/s1047-8477(03)00039-x.

Berliner, D., and J. Bauersachs. 2017. "Current Drug Therapy in Chronic Heart Failure: the New Guidelines of the European Society of Cardiology (ESC)." *Korean Circ J* 47 (5):543-554. doi: 10.4070/kcj.2017.0030.

Bernal, John D, and Dorothy Crowfoot. 1934. "X-ray photographs of crystalline pepsin." *Nature* 133 (3369):794-795.

Bernier, S. C., L. P. Morency, R. Najmanovich, and C. Salesse. 2018. "Identification of an alternative translation initiation site in the sequence of the commonly used Glutathione S-Transferase tag." *J Biotechnol* 286:14-16. doi: 10.1016/j.jbiotec.2018.09.003.

Bervoets, I., and D. Charlier. 2019. "Diversity, versatility and complexity of bacterial gene regulation mechanisms: opportunities and drawbacks for applications in synthetic biology." *FEMS Microbiol Rev* 43 (3):304-339. doi: 10.1093/femsre/fuz001.

Beyzavi, K., S. Hampton, P. Kwasowski, S. Fickling, V. Marks, and R. Clift. 1987. "Comparison of horseradish peroxidase and alkaline phosphatase-labelled antibodies in enzyme immunoassays." *Ann Clin Biochem* 24 (Pt 2):145-52. doi: 10.1177/000456328702400204.

Bhairi, Srirama M. 1997. *Detergents: a guide to the properties and uses of detergents in biological systems*: Calbiochem-Novabiochem International.

Bharti, Amardeep, and Navdeep Goyal. 2019. "Fundamental of Synchrotron Radiations." In *Synchrotron Radiation-Useful and Interesting Applications*. IntechOpen.

Birch, J., D. Axford, J. Foadi, A. Meyer, A. Eckhardt, Y. Thielmann, and I. Moraes. 2018. "The fine art of integral membrane protein crystallisation." *Methods* 147:150-162. doi: 10.1016/j.ymeth.2018.05.014.

Block, H., J. Kubicek, J. Labahn, U. Roth, and F. Schafer. 2008. "Production and comprehensive quality control of recombinant human Interleukin-1beta: a case study for a process development strategy." *Protein Expr Purif* 57 (2):244-54. doi: 10.1016/j.pep.2007.09.019.

Boguszewska, K., M. Szewczuk, S. Urbaniak, and B. T. Karwowski. 2019. "Review: immunoassays in DNA damage and instability detection." *Cell Mol Life Sci* 76 (23):4689-4704. doi: 10.1007/s00018-019-03239-6.

Boistelle, R., and J. P. Astier. 1988. "Crystallization mechanisms in solution." *Journal of Crystal Growth* 90 (1):14-30. doi: https://doi.org/10.1016/0022-0248(88)90294-1.

Booth, W. T., C. R. Schlachter, S. Pote, N. Ussin, N. J. Mank, V. Klapper, L. R. Offermann, C. Tang, B. K. Hurlburt, and M. Chruszcz. 2018. "Impact of an N-terminal Polyhistidine Tag on Protein Thermal Stability." *ACS Omega* 3 (1):760-768. doi: 10.1021/acsomega.7b01598.

Bordicchia, M., D. Liu, E. Z. Amri, G. Ailhaud, P. Dessi-Fulgheri, C. Zhang, N. Takahashi, R. Sarzani, and S. Collins. 2012. "Cardiac natriuretic peptides act via p38 MAPK to induce the brown fat thermogenic program in mouse and human adipocytes." *J Clin Invest* 122 (3):1022-36. doi: 10.1172/JCI59701.

Botte, M., A. Deniaud, and C. Schaffitzel. 2016. "Cell-Free Synthesis of Macromolecular Complexes." *Adv Exp Med Biol* 896:79-95. doi: 10.1007/978-3-319-27216-0_6.

Boutet, S., L. Lomb, G. J. Williams, T. R. Barends, A. Aquila, R. B. Doak, U. Weierstall, D. P. DePonte, J. Steinbrener, R. L. Shoeman, M. Messerschmidt, A. Barty, T. A. White, S. Kassemeyer, R. A. Kirian, M. M. Seibert, P. A. Montanez, C. Kenney, R. Herbst, P. Hart, J. Pines, G. Haller, S. M. Gruner, H. T. Philipp, M. W. Tate, M. Hromalik, L. J. Koerner, N. van Bakel, J. Morse, W. Ghonsalves, D. Arnlund, M. J. Bogan, C. Caleman, R. Fromme, C. Y. Hampton, M. S. Hunter, L. C. Johansson, G. Katona, C. Kupitz, M. Liang, A. V. Martin, K. Nass, L. Redecke, F. Stellato, N. Timneanu, D. Wang, N. A. Zatsepin, D. Schafer, J. Defever, R. Neutze, P. Fromme, J. C. Spence, H. N. Chapman, and I. Schlichting. 2012. "High-resolution protein structure determination by serial femtosecond crystallography." *Science* 337 (6092):362-4. doi: 10.1126/science.1217737.

Boyce, F. M., and N. L. Bucher. 1996. "Baculovirus-mediated gene transfer into mammalian cells." *Proc Natl Acad Sci U S A* 93 (6):2348-52. doi: 10.1073/pnas.93.6.2348.

Brizzard, B. 2008. "Epitope tagging." *Biotechniques* 44 (5):693-5. doi: 10.2144/000112841.

Brönnimann, Christian, and Peter Trüb. 2020. "Hybrid pixel photon counting X-ray detectors for synchrotron radiation." *Synchrotron Light Sources and Free-Electron Lasers: Accelerator Physics, Instrumentation and Science Applications*:1191-1223.

Brown, R. L., T. Strassmaier, J. D. Brady, and J. W. Karpen. 2006. "The pharmacology of cyclic nucleotide-gated channels: emerging from the darkness." *Curr Pharm Des* 12 (28):3597-613. doi: 10.2174/138161206778522100.

Buglioni, A., V. Cannone, A. Cataliotti, S. J. Sangaralingham, D. M. Heublein, C. G. Scott, K. R. Bailey, R. J. Rodeheffer, P. Dessi-Fulgheri, R. Sarzani, and J. C. Burnett, Jr. 2015. "Circulating aldosterone and natriuretic peptides in the general community: relationship to cardiorenal and metabolic disease." *Hypertension* 65 (1):45-53. doi: 10.1161/HYPERTENSIONAHA.114.03936.

Bundy, B. C., and J. R. Swartz. 2010. "Site-specific incorporation of p-propargyloxyphenylalanine in a cell-free environment for direct protein-protein click conjugation." *Bioconjug Chem* 21 (2):255-63. doi: 10.1021/bc9002844.

Burczynska, B., T. Duda, and R. K. Sharma. 2007. "ATP signaling site in the ARM domain of atrial natriuretic factor receptor guanylate cyclase." *Mol Cell Biochem* 301 (1-2):93-107. doi: 10.1007/s11010-006-9400-7.

Burnett, J. C., Jr., J. P. Granger, and T. J. Opgenorth. 1984. "Effects of synthetic atrial natriuretic factor on renal function and renin release." *Am J Physiol* 247 (5 Pt 2):F863-6. doi: 10.1152/ajprenal.1984.247.5.F863.

Caffrey, M. 2003. "Membrane protein crystallization." *J Struct Biol* 142 (1):108-32. doi: 10.1016/s1047-8477(03)00043-1.

Caffrey, M. 2015. "A comprehensive review of the lipid cubic phase or in meso method for crystallizing membrane and soluble proteins and complexes." *Acta Crystallogr F Struct Biol Commun* 71 (Pt 1):3-18. doi: 10.1107/S2053230X14026843.

Caffrey, M., J. Lyons, T. Smyth, and D. J. Hart. 2009. "Chapter 4 Monoacylglycerols: The Workhorse Lipids for Crystallizing Membrane Proteins in Mesophases." In *Current Topics in Membranes*, 83-108. Academic Press.

Caffrey, Martin. 2008. "On the mechanism of membrane protein crystallization in lipidic mesophases." *Crystal Growth and Design* 8 (12):4244-4254.

Calebiro, D., F. Rieken, J. Wagner, T. Sungkaworn, U. Zabel, A. Borzi, E. Cocucci, A. Zurn, and M. J. Lohse. 2013. "Single-molecule analysis of fluorescently labeled

G-protein-coupled receptors reveals complexes with distinct dynamics and organization." *Proc Natl Acad Sci U S A* 110 (2):743-8. doi: 10.1073/pnas.1205798110.

Cannone, V., A. Cabassi, R. Volpi, J. C. Burnett, and Jr. 2019. "Atrial Natriuretic Peptide: A Molecular Target of Novel Therapeutic Approaches to Cardio-Metabolic Disease." *Int J Mol Sci* 20 (13). doi: 10.3390/ijms20133265.

Capewell, S., E. S. Ford, J. B. Croft, J. A. Critchley, K. J. Greenlund, and D. R. Labarthe. 2010. "Cardiovascular risk factor trends and potential for reducing coronary heart disease mortality in the United States of America." *Bull World Health Organ* 88 (2):120-30. doi: 10.2471/BLT.08.057885.

Carey, R. M., D. A. Calhoun, G. L. Bakris, R. D. Brook, S. L. Daugherty, C. R. Dennison-Himmelfarb, B. M. Egan, J. M. Flack, S. S. Gidding, E. Judd, D. T. Lackland, C. L. Laffer, C. Newton-Cheh, S. M. Smith, S. J. Taler, S. C. Textor, T. N. Turan, W. B. White, Education American Heart Association Professional/Public, Hypertension Publications Committee of the Council on, Cardiovascular Council on, Nursing Stroke, Cardiology Council on Clinical, Genomic Council on, Medicine Precision, Disease Council on Peripheral Vascular, Care Council on Quality of, Research Outcomes, and Council Stroke. 2018. "Resistant Hypertension: Detection, Evaluation, and Management: A Scientific Statement From the American Heart Association." *Hypertension* 72 (5):e53-e90. doi: 10.1161/HYP.0000000000000084.

Carpentier, D. C., C. M. Griffiths, and L. A. King. 2008. "The baculovirus P10 protein of Autographa californica nucleopolyhedrovirus forms two distinct cytoskeletal-like structures and associates with polyhedral occlusion bodies during infection." *Virology* 371 (2):278-91. doi: 10.1016/j.virol.2007.09.043.

Cassar, A., D. R. Holmes, Jr., C. S. Rihal, and B. J. Gersh. 2009. "Chronic coronary artery disease: diagnosis and management." *Mayo Clin Proc* 84 (12):1130-46. doi: 10.4065/mcp.2009.0391.

Celie, P. H., A. H. Parret, and A. Perrakis. 2016. "Recombinant cloning strategies for protein expression." *Curr Opin Struct Biol* 38:145-54. doi: 10.1016/j.sbi.2016.06.010.

Chae, P. S., S. G. Rasmussen, R. R. Rana, K. Gotfryd, R. Chandra, M. A. Goren, A. C. Kruse, S. Nurva, C. J. Loland, Y. Pierre, D. Drew, J. L. Popot, D. Picot, B. G. Fox, L. Guan, U. Gether, B. Byrne, B. Kobilka, and S. H. Gellman. 2010. "Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins." *Nat Methods* 7 (12):1003-8. doi: 10.1038/nmeth.1526.

Chambers, A. C., M. Aksular, L. P. Graves, S. L. Irons, R. D. Possee, and L. A. King. 2018. "Overview of the Baculovirus Expression System." *Curr Protoc Protein Sci* 91:5 4 1-5 4 6. doi: 10.1002/cpps.47.

Chang, C. H., K. P. Kohse, B. Chang, M. Hirata, B. Jiang, J. E. Douglas, and F. Murad. 1990. "Characterization of ATP-stimulated guanylate cyclase activation in rat lung membranes." *Biochim Biophys Acta* 1052 (1):159-65. doi: 10.1016/0167-4889(90)90071-k.

Chang, V. T., M. Crispin, A. R. Aricescu, D. J. Harvey, J. E. Nettleship, J. A. Fennelly, C. Yu, K. S. Boles, E. J. Evans, D. I. Stuart, R. A. Dwek, E. Y. Jones, R. J. Owens, and S. J. Davis. 2007. "Glycoprotein structural genomics: solving the glycosylation problem." *Structure* 15 (3):267-73. doi: 10.1016/j.str.2007.01.011.

Chapman, H. N., P. Fromme, A. Barty, T. A. White, R. A. Kirian, A. Aquila, M. S. Hunter, J. Schulz, D. P. DePonte, U. Weierstall, R. B. Doak, F. R. Maia, A. V. Martin, I. Schlichting, L. Lomb, N. Coppola, R. L. Shoeman, S. W. Epp, R. Hartmann, D. Rolles, A. Rudenko, L. Foucar, N. Kimmel, G. Weidenspointner, P. Holl, M. Liang, M. Barthelmess, C. Caleman, S. Boutet, M. J. Bogan, J. Krzywinski, C. Bostedt, S. Bajt, L. Gumprecht, B. Rudek, B. Erk, C. Schmidt, A. Homke, C. Reich, D. Pietschner, L. Struder, G. Hauser, H. Gorke, J. Ullrich, S. Herrmann, G. Schaller, F. Schopper, H. Soltau, K. U. Kuhnel, M. Messerschmidt, J. D. Bozek, S. P. Hau-Riege, M. Frank, C. Y. Hampton, R. G. Sierra, D. Starodub, G. J. Williams, J. Hajdu, N. Timneanu, M. M. Seibert, J. Andreasson, A. Rocker, O. Jonsson, M. Svenda, S. Stern, K. Nass, R. Andritschke, C. D. Schroter, F. Krasniqi, M. Bott, K. E. Schmidt, X. Wang, I. Grotjohann, J. M. Holton, T. R. Barends, R. Neutze, S. Marchesini, R. Fromme, S. Schorb, D. Rupp, M. Adolph, T. Gorkhover, I. Andersson, H. Hirsemann, G. Potdevin, H. Graafsma, B. Nilsson, and J. C. Spence. 2011. "Femtosecond X-ray protein nanocrystallography." *Nature* 470 (7332):73-7. doi: 10.1038/nature09750.

Chavey, W. E., R. V. Hogikyan, R. Van Harrison, and J. M. Nicklas. 2017. "Heart Failure Due to Reduced Ejection Fraction: Medical Management." *Am Fam Physician* 95 (1):13-20.

Chayen, N. E. 1998. "Comparative studies of protein crystallization by vapour-diffusion and microbatch techniques." *Acta Crystallogr D Biol Crystallogr* 54 (Pt 1):8-15. doi: 10.1107/s0907444997005374.

Chayen, N. E., and E. Saridakis. 2008. "Protein crystallization: from purified protein to diffraction-quality crystal." *Nat Methods* 5 (2):147-53. doi: 10.1038/nmeth.f.203.

Cheah, J. S., and S. Yamada. 2017. "A simple elution strategy for biotinylated proteins bound to streptavidin conjugated beads using excess biotin and heat." *Biochem Biophys Res Commun* 493 (4):1522-1527. doi: 10.1016/j.bbrc.2017.09.168.

Chen, C. Y., and A. B. Shyu. 1995. "AU-rich elements: characterization and importance in mRNA degradation." *Trends Biochem Sci* 20 (11):465-70. doi: 10.1016/s0968-0004(00)89102-1.

Chen, H., J. Neutel, D. H. Smith, D. Heublein, and J. Burnett. 2016. "A FIRST-IN-HUMAN TRIAL OF A NOVEL DESIGNER NATRIURETIC PEPTIDE ZD100 IN HUMAN HYPERTENSION." *Journal of the American College of Cardiology* 67:1946.

Chen, R. 2012. "Bacterial expression systems for recombinant protein production: E. coli and beyond." *Biotechnol Adv* 30 (5):1102-7. doi: 10.1016/j.biotechadv.2011.09.013.

Chen, Y., J. J. Schaefer, S. R. Iyer, G. E. Harders, S. Pan, S. J. Sangaralingham, H. H. Chen, M. M. Redfield, and J. C. Burnett, Jr. 2020. "Long-term blood pressure lowering and cGMP-activating actions of the novel ANP analog MANP." *Am J Physiol Regul Integr Comp Physiol* 318 (4):R669-R676. doi: 10.1152/ajpregu.00354.2019.

Chen, Y., Y. Zheng, S. R. Iyer, G. E. Harders, S. Pan, H. H. Chen, T. Ichiki, J. C. Burnett, Jr., and S. J. Sangaralingham. 2019. "C53: A novel particulate guanylyl cyclase B receptor activator that has sustained activity in vivo with anti-fibrotic actions in human cardiac and renal fibroblasts." *J Mol Cell Cardiol* 130:140-150. doi: 10.1016/j.yjmcc.2019.03.024.

Cheng, A., B. Hummel, H. Qiu, and M. Caffrey. 1998. "A simple mechanical mixer for small viscous lipid-containing samples." *Chem Phys Lipids* 95 (1):11-21. doi: 10.1016/s0009-3084(98)00060-7.

Chinkers, M., and D. L. Garbers. 1989. "The protein kinase domain of the ANP receptor is required for signaling." *Science* 245 (4924):1392-4. doi: 10.1126/science.2571188.

Chinkers, M., S. Singh, and D. L. Garbers. 1991. "Adenine nucleotides are required for activation of rat atrial natriuretic peptide receptor/guanylyl cyclase expressed in a baculovirus system." *J Biol Chem* 266 (7):4088-93.

Chinkers, M., and E. M. Wilson. 1992. "Ligand-independent oligomerization of natriuretic peptide receptors. Identification of heteromeric receptors and a dominant negative mutant." *J Biol Chem* 267 (26):18589-97.

Cilia, L., A. Saeed, H. V. Ganga, and W. C. Wu. 2019. "Heart Failure With Preserved Ejection Fraction: Prevention and Management." *Am J Lifestyle Med* 13 (2):182-189. doi: 10.1177/1559827617695219.

Clem, R. J. 2007. "Baculoviruses and apoptosis: a diversity of genes and responses." *Curr Drug Targets* 8 (10):1069-74. doi: 10.2174/138945007782151405.

Clem, R. J., and A. L. Passarelli. 2013. "Baculoviruses: sophisticated pathogens of insects." *PLoS Pathog* 9 (11):e1003729. doi: 10.1371/journal.ppat.1003729.

Cohn, Edwin Joseph. 1922. "STUDIES IN THE PHYSICAL CHEMISTRY OF THE PROTEINS: I. THE SOLUBILITY OF CERTAIN PROTEINS AT THEIR ISOELECTRIC POINTS." *The Journal of general physiology* 4 (6):697.

Contreras-Gomez, A., A. Sanchez-Miron, F. Garcia-Camacho, E. Molina-Grima, and Y. Chisti. 2014. "Protein production using the baculovirus-insect cell expression system." *Biotechnol Prog* 30 (1):1-18. doi: 10.1002/btpr.1842.

Cornwell, T. L., K. B. Pryzwansky, T. A. Wyatt, and T. M. Lincoln. 1991. "Regulation of sarcoplasmic reticulum protein phosphorylation by localized cyclic GMP-dependent protein kinase in vascular smooth muscle cells." *Mol Pharmacol* 40 (6):923-31.

Costello-Boerrigter, L. C., and J. C. Burnett, Jr. 2009. "A new role for the natriuretic peptides: metabolic regulators of the adipocyte." *J Am Coll Cardiol* 53 (22):2078-9. doi: 10.1016/j.jacc.2009.02.037.

Danikowski, K. M., and T. Cheng. 2019. "Colorimetric Analysis of Alkaline Phosphatase Activity in S. aureus Biofilm." *J Vis Exp* (146). doi: 10.3791/59285.

Deller, M. C., L. Kong, and B. Rupp. 2016. "Protein stability: a crystallographer's perspective." *Acta Crystallogr F Struct Biol Commun* 72 (Pt 2):72-95. doi: 10.1107/S2053230X15024619.

Deller, M. C., and B. Rupp. 2014. "Approaches to automated protein crystal harvesting." *Acta Crystallogr F Struct Biol Commun* 70 (Pt 2):133-55. doi: 10.1107/S2053230X14000387.

Dessau, M. A., and Y. Modis. 2011. "Protein crystallization for X-ray crystallography." *J Vis Exp* (47). doi: 10.3791/2285.

Dickey, D. M., A. R. Yoder, and L. R. Potter. 2009. "A familial mutation renders atrial natriuretic Peptide resistant to proteolytic degradation." *J Biol Chem* 284 (29):19196-202. doi: 10.1074/jbc.M109.010777.

DiPilato, L. M., F. Ahmad, M. Harms, P. Seale, V. Manganiello, and M. J. Birnbaum. 2015. "The Role of PDE3B Phosphorylation in the Inhibition of Lipolysis by Insulin." *Mol Cell Biol* 35 (16):2752-60. doi: 10.1128/MCB.00422-15.

Dondapati, S. K., M. Stech, A. Zemella, and S. Kubick. 2020. "Cell-Free Protein Synthesis: A Promising Option for Future Drug Development." *BioDrugs* 34 (3):327-348. doi: 10.1007/s40259-020-00417-y.

Drewett, J. G., and D. L. Garbers. 1994. "The family of guanylyl cyclase receptors and their ligands." *Endocr Rev* 15 (2):135-62. doi: 10.1210/edrv-15-2-135.

Ducruix, Arnaud, and Richard Giegé. 1999. *Crystallization of nucleic acids and proteins: a practical approach*: Practical Approach Series.

Duda, T., R. M. Goraczniak, and R. K. Sharma. 1993a. "Core sequence of ATP regulatory module in receptor guanylate cyclases." *FEBS Lett* 315 (2):143-8. doi: 10.1016/0014-5793(93)81151-o.

Duda, T., R. M. Goraczniak, and R. K. Sharma. 1993b. "The glycine residue of ATP regulatory module in receptor guanylate cyclases that is essential in natriuretic factor signaling." *FEBS Lett* 335 (3):309-14. doi: 10.1016/0014-5793(93)80408-m.

Duda, T., R. M. Goraczniak, A. Sitaramayya, and R. K. Sharma. 1993. "Cloning and expression of an ATP-regulated human retina C-type natriuretic factor receptor guanylate cyclase." *Biochemistry* 32 (6):1391-5. doi: 10.1021/bi00057a001.

Duda, T., V. Venkataraman, S. Ravichandran, and R. K. Sharma. 2005. "ATP-regulated module (ARM) of the atrial natriuretic factor receptor guanylate cyclase." *Peptides* 26 (6):969-84. doi: 10.1016/j.peptides.2004.08.032.

Duda, T., P. Yadav, A. Jankowska, V. Venkataraman, and R. K. Sharma. 2000. "Three dimensional atomic model and experimental validation for the ATP-Regulated Module (ARM) of the atrial natriuretic factor receptor guanylate cyclase." *Mol Cell Biochem* 214 (1-2):7-14. doi: 10.1023/a:1007144328682.

Dukkipati, A., H. H. Park, D. Waghray, S. Fischer, and K. C. Garcia. 2008. "BacMam system for high-level expression of recombinant soluble and membrane glycoproteins for structural studies." *Protein Expr Purif* 62 (2):160-70. doi: 10.1016/j.pep.2008.08.004.

Dumont, J., D. Euwart, B. Mei, S. Estes, and R. Kshirsagar. 2016. "Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives." *Crit Rev Biotechnol* 36 (6):1110-1122. doi: 10.3109/07388551.2015.1084266.

Dunham, J. H., and R. A. Hall. 2009. "Enhancement of the surface expression of G protein-coupled receptors." *Trends Biotechnol* 27 (9):541-5. doi: 10.1016/j.tibtech.2009.06.005.

Duong-Ly, K. C., and S. B. Gabelli. 2014a. "Gel filtration chromatography (size exclusion chromatography) of proteins." *Methods Enzymol* 541:105-14. doi: 10.1016/B978-0-12-420119-4.00009-4.

Duong-Ly, K. C., and S. B. Gabelli. 2014b. "Salting out of proteins using ammonium sulfate precipitation." *Methods Enzymol* 541:85-94. doi: 10.1016/B978-0-12-420119-4.00007-0.

Durbin, S. D., and G. Feher. 1996. "Protein crystallization." *Annu Rev Phys Chem* 47:171-204. doi: 10.1146/annurev.physchem.47.1.171.

Dyson, M. R. 2016. "Fundamentals of Expression in Mammalian Cells." *Adv Exp Med Biol* 896:217-24. doi: 10.1007/978-3-319-27216-0_14.

Earl, L. A., V. Falconieri, J. L. Milne, and S. Subramaniam. 2017. "Cryo-EM: beyond the microscope." *Curr Opin Struct Biol* 46:71-78. doi: 10.1016/j.sbi.2017.06.002.

Egom, E. E. 2019. "Pulmonary Arterial Hypertension Due to NPR-C Mutation: A Novel Paradigm for Normal and Pathologic Remodeling?" *Int J Mol Sci* 20 (12). doi: 10.3390/ijms20123063.

Einhauer, A., and A. Jungbauer. 2001. "The FLAG peptide, a versatile fusion tag for the purification of recombinant proteins." *J Biochem Biophys Methods* 49 (1-3):455-65. doi: 10.1016/s0165-022x(01)00213-5.

Emma, P., R. Akre, J. Arthur, R. Bionta, C. Bostedt, J. Bozek, A. Brachmann, P. Bucksbaum, R. Coffee, F. J. Decker, Y. Ding, D. Dowell, S. Edstrom, A. Fisher, J. Frisch, S. Gilevich, J. Hastings, G. Hays, Ph Hering, Z. Huang, R. Iverson, H. Loos, M. Messerschmidt, A. Miahnahri, S. Moeller, H. D. Nuhn, G. Pile, D. Ratner, J. Rzepiela, D. Schultz, T. Smith, P. Stefan, H. Tompkins, J. Turner, J. Welch, W. White, J. Wu, G. Yocky, and J. Galayda. 2010. "First lasing and operation of an ångstrom-wavelength free-electron laser." *Nature Photonics* 4 (9):641-647. doi: 10.1038/nphoton.2010.176.

Engeli, S., A. L. Birkenfeld, P. M. Badin, V. Bourlier, K. Louche, N. Viguerie, C. Thalamas, E. Montastier, D. Larrouy, I. Harant, I. de Glisezinski, S. Lieske, J. Reinke, B. Beckmann, D. Langin, J. Jordan, and C. Moro. 2012. "Natriuretic peptides enhance the oxidative capacity of human skeletal muscle." *J Clin Invest* 122 (12):4675-9. doi: 10.1172/JCI64526.

Falkner, B., and N. D. Cossrow. 2014. "Prevalence of metabolic syndrome and obesity-associated hypertension in the racial ethnic minorities of the United States." *Curr Hypertens Rep* 16 (7):449. doi: 10.1007/s11906-014-0449-5.

Fan, D., P. M. Bryan, L. K. Antos, R. J. Potthast, and L. R. Potter. 2005. "Down-regulation does not mediate natriuretic peptide-dependent desensitization of natriuretic peptide receptor (NPR)-A or NPR-B: guanylyl cyclase-linked natriuretic peptide receptors do not internalize." *Mol Pharmacol* 67 (1):174-83. doi: 10.1124/mol.104.002436.

Fawole, FJ, NP Sahu, N Shamna, V Phulia, BO Emikpe, AA Adeoye, AZ Aderolu, and OM Popoola. 2018. "Effects of detoxified Jatropha curcas protein isolate on growth performance, nutrient digestibility and physio□metabolic response of Labeo rohita fingerlings." *Aquaculture Nutrition* 24 (4):1223-1233.

Federmann, M., and O. M. Hess. 1994. "Differentiation between systolic and diastolic dysfunction." *Eur Heart J* 15 Suppl D:2-6. doi: 10.1093/eurheartj/15.suppl_d.2.

Felberbaum, R. S. 2015. "The baculovirus expression vector system: A commercial manufacturing platform for viral vaccines and gene therapy vectors." *Biotechnol J* 10 (5):702-14. doi: 10.1002/biot.201400438.

Ferrer-Miralles, N., J. Domingo-Espin, J. L. Corchero, E. Vazquez, and A. Villaverde. 2009. "Microbial factories for recombinant pharmaceuticals." *Microb Cell Fact* 8:17. doi: 10.1186/1475-2859-8-17.

Flora, G. D., and M. K. Nayak. 2019. "A Brief Review of Cardiovascular Diseases, Associated Risk Factors and Current Treatment Regimes." *Curr Pharm Des* 25 (38):4063-4084. doi: 10.2174/1381612825666190925163827.

Ford, E. S. 2005. "Risks for all-cause mortality, cardiovascular disease, and diabetes associated with the metabolic syndrome: a summary of the evidence." *Diabetes Care* 28 (7):1769-78. doi: 10.2337/diacare.28.7.1769.

Forssmann, W. G., R. Richter, and M. Meyer. 1998. "The endocrine heart and natriuretic peptides: histochemistry, cell biology, and functional aspects of the renal urodilatin system." *Histochem Cell Biol* 110 (4):335-57. doi: 10.1007/s004180050295.

Foster, D. C., and D. L. Garbers. 1998. "Dual role for adenine nucleotides in the regulation of the atrial natriuretic peptide receptor, guanylyl cyclase-A." *J Biol Chem* 273 (26):16311-8. doi: 10.1074/jbc.273.26.16311.

Frank, K., and E. G. Kranias. 2000. "Phospholamban and cardiac contractility." *Ann Med* 32 (8):572-8. doi: 10.3109/07853890008998837.

Franke, B., C. Opitz, S. Isogai, A. Grahl, L. Delgado, A. D. Gossert, and S. Grzesiek. 2018. "Production of isotope-labeled proteins in insect cells for NMR." *J Biomol NMR* 71 (3):173-184. doi: 10.1007/s10858-018-0172-7.

Franklin, S. S. 2006. "Hypertension in the metabolic syndrome." *Metab Syndr Relat Disord* 4 (4):287-98. doi: 10.1089/met.2006.4.287.

Friebe, A., P. Sandner, and A. Schmidtko. 2020. "cGMP: a unique 2nd messenger molecule - recent developments in cGMP research and development." *Naunyn Schmiedebergs Arch Pharmacol* 393 (2):287-302. doi: 10.1007/s00210-019-01779-z.

Frostegard, J. 2013. "Immunity, atherosclerosis and cardiovascular disease." *BMC Med* 11:117. doi: 10.1186/1741-7015-11-117.

Gabrielsen, Mads, Alastair T. Gardiner, Petra Fromme, and Richard J. Cogdell. 2009. "Chapter 6 Membrane Protein Crystallization: Approaching the Problem and Understanding the Solutions." In *Current Topics in Membranes*, 127-149. Academic Press.

Gan, S. D., and K. R. Patel. 2013. "Enzyme immunoassay and enzyme-linked immunosorbent assay." *J Invest Dermatol* 133 (9):e12. doi: 10.1038/jid.2013.287.

Garavito, R. M., and J. P. Rosenbusch. 1980. "Three-dimensional crystals of an integral membrane protein: an initial x-ray analysis." *J Cell Biol* 86 (1):327-9. doi: 10.1083/jcb.86.1.327.

Garcia-Ruiz, JM, and JD Ng. 2006. "Counterdiffusion capillary crystallization for high-throughput applications." *Protein Crystallization Strategies for Structural Genomics*:111-126.

Garman, Elspeth F, and Thomas R Schneider. 1997. "Macromolecular cryocrystallography." *Journal of Applied Crystallography* 30 (3):211-237.

Gauto, D. F., L. F. Estrozi, C. D. Schwieters, G. Effantin, P. Macek, R. Sounier, A. C. Sivertsen, E. Schmidt, R. Kerfah, G. Mas, J. P. Colletier, P. Guntert, A. Favier, G. Schoehn, P. Schanda, and J. Boisbouvier. 2019. "Integrated NMR and cryo-EM atomic-resolution structure determination of a half-megadalton enzyme complex." *Nat Commun* 10 (1):2697. doi: 10.1038/s41467-019-10490-9.

Gavin, A. C., M. Bosche, R. Krause, P. Grandi, M. Marzioch, A. Bauer, J. Schultz, J. M. Rick, A. M. Michon, C. M. Cruciat, M. Remor, C. Hofert, M. Schelder, M. Brajenovic, H. Ruffner, A. Merino, K. Klein, M. Hudak, D. Dickson, T. Rudi, V. Gnau, A. Bauch, S. Bastuck, B. Huhse, C. Leutwein, M. A. Heurtier, R. R. Copley, A. Edelmann, E. Querfurth, V. Rybin, G. Drewes, M. Raida, T. Bouwmeester, P. Bork, B. Seraphin, B. Kuster, G. Neubauer, and G. Superti-Furga. 2002. "Functional organization of the yeast proteome by systematic analysis of protein complexes." *Nature* 415 (6868):141-7. doi: 10.1038/415141a.

Gaziano, T. A., A. Bitton, S. Anand, S. Abrahams-Gessel, and A. Murphy. 2010. "Growing epidemic of coronary heart disease in low- and middle-income countries." *Curr Probl Cardiol* 35 (2):72-115. doi: 10.1016/j.cpcardiol.2009.10.002.

Geisler, C., and D. L. Jarvis. 2012. "Innovative use of a bacterial enzyme involved in sialic acid degradation to initiate sialic acid biosynthesis in glycoengineered insect cells." *Metab Eng* 14 (6):642-52. doi: 10.1016/j.ymben.2012.08.005.

Giege, R. 2013. "A historical perspective on protein crystallization from 1840 to the present day." *FEBS J* 280 (24):6456-97. doi: 10.1111/febs.12580.

Ginsberg, H. N., Y. L. Zhang, and A. Hernandez-Ono. 2006. "Metabolic syndrome: focus on dyslipidemia." *Obesity (Silver Spring)* 14 Suppl 1:41S-49S. doi: 10.1038/oby.2006.281.

Gisriel, C., J. Coe, R. Letrun, O. M. Yefanov, C. Luna-Chavez, N. E. Stander, S. Lisova, V. Mariani, M. Kuhn, S. Aplin, T. D. Grant, K. Dorner, T. Sato, A. Echelmeier, J. Cruz Villarreal, M. S. Hunter, M. O. Wiedorn, J. Knoska, V. Mazalova, S. Roy-Chowdhury, J. H. Yang, A. Jones, R. Bean, J. Bielecki, Y. Kim, G. Mills, B. Weinhausen, J. D. Meza, N. Al-Qudami, S. Bajt, G. Brehm, S. Botha, D. Boukhelef, S. Brockhauser, B. D. Bruce, M. A. Coleman, C. Danilevski, E. Discianno, Z. Dobson, H. Fangohr, J. M. Martin-Garcia, Y. Gevorkov, S. Hauf, A. Hosseinizadeh, F. Januschek, G. K. Ketawala, C. Kupitz, L. Maia, M. Manetti, M. Messerschmidt, T. Michelat, J. Mondal, A. Ourmazd, G. Previtali, I. Sarrou, S. Schon, P. Schwander, M. L. Shelby, A. Silenzi, J. Sztuk-Dambietz, J. Szuba, M. Turcato, T. A. White, K. Wrona, C. Xu, M. H. Abdellatif, J. D. Zook, J. C. H. Spence, H. N. Chapman, A. Barty, R. A. Kirian, M. Frank, A. Ros, M. Schmidt, R. Fromme, A. P. Mancuso, P. Fromme, and N. A. Zatsepin. 2019. "Membrane protein megahertz crystallography at the European XFEL." *Nat Commun* 10 (1):5021. doi: 10.1038/s41467-019-12955-3.

Gisriel, C., G. Shen, V. Kurashov, M. Y. Ho, S. Zhang, D. Williams, J. H. Golbeck, P. Fromme, and D. A. Bryant. 2020. "The structure of Photosystem I acclimated to far-red light illuminates an ecologically important acclimation process in photosynthesis." *Sci Adv* 6 (6):eaay6415. doi: 10.1126/sciadv.aay6415.

Goehring, A., C. H. Lee, K. H. Wang, J. C. Michel, D. P. Claxton, I. Baconguis, T. Althoff, S. Fischer, K. C. Garcia, and E. Gouaux. 2014. "Screening and large-scale expression of membrane proteins in mammalian cells for structural studies." *Nat Protoc* 9 (11):2574-85. doi: 10.1038/nprot.2014.173.

Gong, B., Z. Wu, and Z. Li. 2016. "Efficacy and safety of nesiritide in patients with decompensated heart failure: a meta-analysis of randomised trials." *BMJ Open* 6 (1):e008545. doi: 10.1136/bmjopen-2015-008545.

Goraczniak, R. M., T. Duda, and R. K. Sharma. 1992. "A structural motif that defines the ATP-regulatory module of guanylate cyclase in atrial natriuretic factor signalling." *Biochem J* 282 (Pt 2):533-7. doi: 10.1042/bj2820533.

Granados, R. R., and K. A. Lawler. 1981. "In vivo pathway of Autographa californica baculovirus invasion and infection." *Virology* 108 (2):297-308. doi: 10.1016/0042-6822(81)90438-4.

Granados, Robert R., Li Guoxun, Anja C. G. Derksen, and Kevin A. McKenna. 1994. "A new insect cell line from Trichoplusia ni (BTI-Tn-5B1-4) susceptible to

Trichoplusia ni single enveloped nuclear polyhedrosis virus." *Journal of Invertebrate Pathology* 64 (3):260-266. doi: https://doi.org/10.1016/S0022-2011(94)90400-6.

Gregorio, N. E., M. Z. Levine, and J. P. Oza. 2019. "A User's Guide to Cell-Free Protein Synthesis." *Methods Protoc* 2 (1). doi: 10.3390/mps2010024.

Grundy, S. M., J. I. Cleeman, S. R. Daniels, K. A. Donato, R. H. Eckel, B. A. Franklin, D. J. Gordon, R. M. Krauss, P. J. Savage, S. C. Smith, Jr., J. A. Spertus, F. Costa, Association American Heart, Lung National Heart, and Institute Blood. 2005. "Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement." *Circulation* 112 (17):2735-52. doi: 10.1161/CIRCULATIONAHA.105.169404.

Guan, X. M., T. S. Kobilka, and B. K. Kobilka. 1992. "Enhancement of membrane insertion and function in a type IIIb membrane protein following introduction of a cleavable signal peptide." *J Biol Chem* 267 (31):21995-8.

Gubellini, F., G. Verdon, N. K. Karpowich, J. D. Luff, G. Boel, N. Gauthier, S. K. Handelman, S. E. Ades, and J. F. Hunt. 2011. "Physiological response to membrane protein overexpression in E. coli." *Mol Cell Proteomics* 10 (10):M111 007930. doi: 10.1074/mcp.M111.007930.

Guder, G., and S. Stork. 2019. "COPD and heart failure: differential diagnosis and comorbidity." *Herz* 44 (6):502-508. doi: 10.1007/s00059-019-4814-7.

Haase, S., A. Sciocco-Cap, and V. Romanowski. 2015. "Baculovirus insecticides in Latin America: historical overview, current status and future perspectives." *Viruses* 7 (5):2230-67. doi: 10.3390/v7052230.

Haddley, K. 2013. "Alipogene tiparvovec for the treatment of lipoprotein lipase deficiency." *Drugs Today (Barc)* 49 (3):161-70. doi: 10.1358/dot.2013.49.3.1937398.

Hanks, S. K., A. M. Quinn, and T. Hunter. 1988. "The protein kinase family: conserved features and deduced phylogeny of the catalytic domains." *Science* 241 (4861):42-52. doi: 10.1126/science.3291115.

Harris, P. J., D. Thomas, and T. O. Morgan. 1987. "Atrial natriuretic peptide inhibits angiotensin-stimulated proximal tubular sodium and water reabsorption." *Nature* 326 (6114):697-8. doi: 10.1038/326697a0.

Harrison, R. L., E. A. Herniou, J. A. Jehle, D. A. Theilmann, J. P. Burand, J. J. Becnel, P. J. Krell, M. M. van Oers, J. D. Mowery, G. R. Bauchan, and Consortium Ictv Report. 2018. "ICTV Virus Taxonomy Profile: Baculoviridae." *J Gen Virol* 99 (9):1185-1186. doi: 10.1099/jgv.0.001107.

Hattne, J., F. E. Reyes, B. L. Nannenga, D. Shi, M. J. de la Cruz, A. G. Leslie, and T. Gonen. 2015. "MicroED data collection and processing." *Acta Crystallogr A Found Adv* 71 (Pt 4):353-60. doi: 10.1107/S2053273315010669.

Haupert, Levi M., and Garth J. Simpson. 2011. "Screening of protein crystallization trials by second order nonlinear optical imaging of chiral crystals (SONICC)." *Methods* 55 (4):379-386. doi: https://doi.org/10.1016/j.ymeth.2011.11.003.

He, X. L., A. Dukkipati, and K. C. Garcia. 2006. "Structural determinants of natriuretic peptide receptor specificity and degeneracy." *J Mol Biol* 361 (4):698-714. doi: 10.1016/j.jmb.2006.06.060.

He, Xl, Dc Chow, M. M. Martick, and K. C. Garcia. 2001. "Allosteric activation of a spring-loaded natriuretic peptide receptor dimer by hormone." *Science* 293 (5535):1657-62. doi: 10.1126/science.1062246.

Hefferon, K. L., and L. K. Miller. 2002. "Reconstructing the replication complex of AcMNPV." *Eur J Biochem* 269 (24):6233-40. doi: 10.1046/j.1432-1033.2002.03342.x.

Heidenreich, P. A., N. M. Albert, L. A. Allen, D. A. Bluemke, J. Butler, G. C. Fonarow, J. S. Ikonomidis, O. Khavjou, M. A. Konstam, T. M. Maddox, G. Nichol, M. Pham, I. L. Pina, J. G. Trogdon, Committee American Heart Association Advocacy Coordinating, Thrombosis Council on Arteriosclerosis, Biology Vascular, Radiology Council on Cardiovascular, Intervention, Cardiology Council on Clinical, Epidemiology Council on, Prevention, and Council Stroke. 2013. "Forecasting the impact of heart failure in the United States: a policy statement from the American Heart Association." *Circ Heart Fail* 6 (3):606-19. doi: 10.1161/HHF.0b013e318291329a.

Heim, J. M., S. Singh, and R. Gerzer. 1996. "Effect of glycosylation on cloned ANF-sensitive guanylyl cyclase." *Life Sci* 59 (4):PL61-8. doi: 10.1016/0024-3205(96)00306-2.

Henderson, R. 1995. "The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules." *Q Rev Biophys* 28 (2):171-93. doi: 10.1017/s003358350000305x.

Hirose, S., H. Hagiwara, and Y. Takei. 2001. "Comparative molecular biology of natriuretic peptide receptors." *Can J Physiol Pharmacol* 79 (8):665-72.

Hitchman, R. B., R. D. Possee, and L. A. King. 2012. "High-throughput baculovirus expression in insect cells." *Methods Mol Biol* 824:609-27. doi: 10.1007/978-1-61779-433-9_33.

Hobbs, F. D. 2015. "Prevention of cardiovascular diseases." *BMC Med* 13:261. doi: 10.1186/s12916-015-0507-0.

Hodgson-Zingman, D. M., M. L. Karst, L. V. Zingman, D. M. Heublein, D. Darbar, K. J. Herron, J. D. Ballew, M. de Andrade, J. C. Burnett, Jr., and T. M. Olson. 2008. "Atrial natriuretic peptide frameshift mutation in familial atrial fibrillation." *N Engl J Med* 359 (2):158-65. doi: 10.1056/NEJMoa0706300.

Hollister, J. R., and D. L. Jarvis. 2001. "Engineering lepidopteran insect cells for sialoglycoprotein production by genetic transformation with mammalian beta 1,4-galactosyltransferase and alpha 2,6-sialyltransferase genes." *Glycobiology* 11 (1):1-9. doi: 10.1093/glycob/11.1.1.

Huang, C. J., A. J. Lowe, and C. A. Batt. 2010. "Recombinant immunotherapeutics: current state and perspectives regarding the feasibility and market." *Appl Microbiol Biotechnol* 87 (2):401-10. doi: 10.1007/s00253-010-2590-7.

Huang, P. L. 2009. "A comprehensive definition for metabolic syndrome." *Dis Model Mech* 2 (5-6):231-7. doi: 10.1242/dmm.001180.

Hui, R., and A. Edwards. 2003. "High-throughput protein crystallization." *J Struct Biol* 142 (1):154-61. doi: 10.1016/s1047-8477(03)00046-7.

Hünefeld, Friedrich Ludwig. 1840. *Der Chemismus in der thierischen Organisation: Physiologisch-chemische Untersuchungen der materiellen Veränderungen, oder des Bildungslebens im thierischen Organismus; insbesondere des Blutbildungsprocesses, der Natur der Blut körperchen und ihrer Kernchen. Ein Beitrag zur Physiologie und Heilmittellehre*: Brockhaus.

Hunt, P. J., A. M. Richards, E. A. Espiner, M. G. Nicholls, and T. G. Yandle. 1994. "Bioactivity and metabolism of C-type natriuretic peptide in normal man." *J Clin Endocrinol Metab* 78 (6):1428-35. doi: 10.1210/jcem.78.6.8200946.

Hunter, M. R., N. L. Grimsey, and M. Glass. 2016. "Sulfation of the FLAG epitope is affected by co-expression of G protein-coupled receptors in a mammalian cell model." *Sci Rep* 6:27316. doi: 10.1038/srep27316.

Huo, X., T. Abe, and K. S. Misono. 1999. "Ligand binding-dependent limited proteolysis of the atrial natriuretic peptide receptor: juxtamembrane hinge structure essential for transmembrane signal transduction." *Biochemistry* 38 (51):16941-51. doi: 10.1021/bi9919448.

Inoue, K., T. Sakamoto, S. Yuge, H. Iwatani, S. Yamagami, M. Tsutsumi, H. Hori, M. C. Cerra, B. Tota, N. Suzuki, N. Okamoto, and Y. Takei. 2005. "Structural and functional evolution of three cardiac natriuretic peptides." *Mol Biol Evol* 22 (12):2428-34. doi: 10.1093/molbev/msi243.

Inserte, J., V. Hernando, M. Ruiz-Meana, M. Poncelas-Nozal, C. Fernandez, L. Agullo, C. Sartorio, U. Vilardosa, and D. Garcia-Dorado. 2014. "Delayed phospholamban phosphorylation in post-conditioned heart favours Ca2+ normalization and

contributes to protection." *Cardiovasc Res* 103 (4):542-53. doi: 10.1093/cvr/cvu163.

Invitrogen Life Technologies. 2002. "Bac-to-Bac® Baculovirus Expression Systems." *Invitrogen Life Technologies*.

Irvine, G. B. 2001. "Determination of molecular size by size-exclusion chromatography (gel filtration)." *Curr Protoc Cell Biol* Chapter 5:Unit 5 5. doi: 10.1002/0471143030.cb0505s06.

Ishihama, A. 2018. "Building a complete image of genome regulation in the model organism Escherichia coli." *J Gen Appl Microbiol* 63 (6):311-324. doi: 10.2323/jgam.2017.01.002.

Itakura, K., T. Hirose, R. Crea, A. D. Riggs, H. L. Heyneker, F. Bolivar, and H. W. Boyer. 1977. "Expression in Escherichia coli of a chemically synthesized gene for the hormone somatostatin." *Science* 198 (4321):1056-63. doi: 10.1126/science.412251.

Iwata, T., K. Uchida-Mizuno, T. Katafuchi, T. Ito, H. Hagiwara, and S. Hirose. 1991. "Bifunctional atrial natriuretic peptide receptor (type A) exists as a disulfide-linked tetramer in plasma membranes of bovine adrenal cortex." *J Biochem* 110 (1):35-9. doi: 10.1093/oxfordjournals.jbchem.a123539.

Janakiraman, V., W. F. Forrest, B. Chow, and S. Seshagiri. 2006. "A rapid method for estimation of baculovirus titer based on viable cell size." *J Virol Methods* 132 (1-2):48-58. doi: 10.1016/j.jviromet.2005.08.021.

Janakiraman, V., W. F. Forrest, and S. Seshagiri. 2006. "Estimation of baculovirus titer based on viable cell size." *Nat Protoc* 1 (5):2271-6. doi: 10.1038/nprot.2006.387.

Janin, Joel. 1995. "Elusive affinities." *Proteins: Structure, Function, and Bioinformatics* 21 (1):30-39.

Jensen, H. M., T. Eng, V. Chubukov, R. A. Herbert, and A. Mukhopadhyay. 2017. "Improving membrane protein expression and function using genomic edits." *Sci Rep* 7 (1):13030. doi: 10.1038/s41598-017-12901-7.

Jewett, J. R., K. J. Koller, D. V. Goeddel, and D. G. Lowe. 1993. "Hormonal induction of low affinity receptor guanylyl cyclase." *EMBO J* 12 (2):769-77.

Jhund, P. S., and J. J. McMurray. 2016. "The neprilysin pathway in heart failure: a review and guide on the use of sacubitril/valsartan." *Heart* 102 (17):1342-7. doi: 10.1136/heartjnl-2014-306775.

Jordan, P., P. Fromme, H. T. Witt, O. Klukas, W. Saenger, and N. Krauss. 2001. "Three-dimensional structure of cyanobacterial photosystem I at 2.5 A resolution." *Nature* 411 (6840):909-17. doi: 10.1038/35082000.

Jorge, A. R. C., P. H. S. Costa, H. S. A. Monteiro, and M. C. Fonteles. 2018. "High Salt Intake Promotes Different Responses to Urodilatin and Uroguanylin in the Isolated Rat Kidney." *Horm Metab Res* 50 (2):152-159. doi: 10.1055/s-0043-120669.

Julin, Douglas A. 2018. "Blue/White Selection." In *Molecular Life Sciences: An Encyclopedic Reference*, edited by Robert D. Wells, Judith S. Bond, Judith Klinman and Bettie Sue Siler Masters, 72-73. New York, NY: Springer New York.

Jungbauer, A., and R. Hahn. 2009. "Ion-exchange chromatography." *Methods Enzymol* 463:349-71. doi: 10.1016/S0076-6879(09)63022-6.

Kashiwagi, M., T. Katafuchi, A. Kato, H. Inuyama, T. Ito, H. Hagiwara, Y. Takei, and S. Hirose. 1995. "Cloning and properties of a novel natriuretic peptide receptor, NPR-D." *Eur J Biochem* 233 (1):102-9. doi: 10.1111/j.1432-1033.1995.102_1.x.

Kashiwagi, M., K. Miyamoto, Y. Takei, and S. Hirose. 1999. "Cloning, properties and tissue distribution of natriuretic peptide receptor-A of euryhaline eel, Anguilla japonica." *Eur J Biochem* 259 (1-2):204-11. doi: 10.1046/j.1432-1327.1999.00023.x.

Kaupp, U. B., and R. Seifert. 2002. "Cyclic nucleotide-gated ion channels." *Physiol Rev* 82 (3):769-824. doi: 10.1152/physrev.00008.2002.

Kaur, J., A. Kumar, and J. Kaur. 2018. "Strategies for optimization of heterologous protein expression in E. coli: Roadblocks and reinforcements." *Int J Biol Macromol* 106:803-822. doi: 10.1016/j.ijbiomac.2017.08.080.

Kendrew, J. C., G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, and D. C. Phillips. 1958. "A three-dimensional model of the myoglobin molecule obtained by x-ray analysis." *Nature* 181 (4610):662-6. doi: 10.1038/181662a0.

Kermani, A. A. 2020. "A guide to membrane protein X-ray crystallography." *FEBS J.* doi: 10.1111/febs.15676.

Khanmohammadi, M., M. B. Dastjerdi, A. Ai, A. Ahmadi, A. Godarzi, A. Rahimi, and J. Ai. 2018. "Horseradish peroxidase-catalyzed hydrogelation for biomedical applications." *Biomater Sci* 6 (6):1286-1298. doi: 10.1039/c8bm00056e.

Kim, Kwang-Je. 1986. "Brightness, coherence and propagation characteristics of synchrotron radiation." *Nuclear Instruments and Methods in Physics Research*

Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 246 (1-3):71-76.

Kishimoto, I., S. K. Dubois, and D. L. Garbers. 1996. "The heart communicates with the kidney exclusively through the guanylyl cyclase-A receptor: acute handling of sodium and water in response to volume expansion." *Proc Natl Acad Sci U S A* 93 (12):6215-9. doi: 10.1073/pnas.93.12.6215.

Kitts, P. A., M. D. Ayres, and R. D. Possee. 1990. "Linearization of baculovirus DNA enhances the recovery of recombinant virus expression vectors." *Nucleic Acids Res* 18 (19):5667-72. doi: 10.1093/nar/18.19.5667.

Kitts, P. A., and R. D. Possee. 1993. "A method for producing recombinant baculovirus expression vectors at high frequency." *Biotechniques* 14 (5):810-7.

Kjeldsen, S. E. 2018. "Hypertension and cardiovascular risk: General aspects." *Pharmacol Res* 129:95-99. doi: 10.1016/j.phrs.2017.11.003.

Koesling, D., E. Bohme, and G. Schultz. 1991. "Guanylyl cyclases, a growing family of signal-transducing enzymes." *FASEB J* 5 (13):2785-91. doi: 10.1096/fasebj.5.13.1680765.

Koh, G. Y., D. R. Nussenzveig, J. Okolicany, D. A. Price, and T. Maack. 1992. "Dynamics of atrial natriuretic factor-guanylate cyclase receptors and receptor-ligand complexes in cultured glomerular mesangial and renomedullary interstitial cells." *J Biol Chem* 267 (17):11987-94.

Koller, K. J., M. T. Lipari, and D. V. Goeddel. 1993. "Proper glycosylation and phosphorylation of the type A natriuretic peptide receptor are required for hormone-stimulated guanylyl cyclase activity." *J Biol Chem* 268 (8):5997-6003.

Kong, X., X. Wang, W. Xu, S. Behera, G. Hellermann, A. Kumar, R. F. Lockey, S. Mohapatra, and S. S. Mohapatra. 2008. "Natriuretic peptide receptor a as a novel anticancer target." *Cancer Res* 68 (1):249-56. doi: 10.1158/0008-5472.CAN-07-3086.

Kool, M., and J. M. Vlak. 1993. "The structural and functional organization of the Autographa californica nuclear polyhedrosis virus genome." *Arch Virol* 130 (1-2):1-16. doi: 10.1007/BF01318992.

Koopmann, R., K. Cupelli, L. Redecke, K. Nass, D. P. Deponte, T. A. White, F. Stellato, D. Rehders, M. Liang, J. Andreasson, A. Aquila, S. Bajt, M. Barthelmess, A. Barty, M. J. Bogan, C. Bostedt, S. Boutet, J. D. Bozek, C. Caleman, N. Coppola, J. Davidsson, R. B. Doak, T. Ekeberg, S. W. Epp, B. Erk, H. Fleckenstein, L. Foucar, H. Graafsma, L. Gumprecht, J. Hajdu, C. Y. Hampton, A. Hartmann, R. Hartmann, G. Hauser, H. Hirsemann, P. Holl, M. S. Hunter, S. Kassemeyer, R. A. Kirian, L. Lomb, F. R. Maia, N. Kimmel, A. V. Martin, M. Messerschmidt, C.

Reich, D. Rolles, B. Rudek, A. Rudenko, I. Schlichting, J. Schulz, M. M. Seibert, R. L. Shoeman, R. G. Sierra, H. Soltau, S. Stern, L. Struder, N. Timneanu, J. Ullrich, X. Wang, G. Weidenspointner, U. Weierstall, G. J. Williams, C. B. Wunderer, P. Fromme, J. C. Spence, T. Stehle, H. N. Chapman, C. Betzel, and M. Duszenko. 2012. "In vivo protein crystallization opens new routes in structural biology." *Nat Methods* 9 (3):259-62. doi: 10.1038/nmeth.1859.

Kotecha, D., M. D. Flather, D. G. Altman, J. Holmes, G. Rosano, J. Wikstrand, M. Packer, A. J. S. Coats, L. Manzano, M. Bohm, D. J. van Veldhuisen, B. Andersson, H. Wedel, T. G. von Lueder, A. S. Rigby, A. Hjalmarson, J. Kjekshus, J. G. F. Cleland, and Group Beta-Blockers in Heart Failure Collaborative. 2017. "Heart Rate and Rhythm and the Benefit of Beta-Blockers in Patients With Heart Failure." *J Am Coll Cardiol* 69 (24):2885-2896. doi: 10.1016/j.jacc.2017.04.001.

Kotov, V., K. Bartels, K. Veith, I. Josts, U. K. T. Subhramanyam, C. Gunther, J. Labahn, T. C. Marlovits, I. Moraes, H. Tidow, C. Low, and M. M. Garcia-Alai. 2019. "High-throughput stability screening for detergent-solubilized membrane proteins." *Sci Rep* 9 (1):10379. doi: 10.1038/s41598-019-46686-8.

Krogh, A., B. Larsson, G. von Heijne, and E. L. Sonnhammer. 2001. "Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes." *J Mol Biol* 305 (3):567-80. doi: 10.1006/jmbi.2000.4315.

Kudo, T., and A. Baird. 1984. "Inhibition of aldosterone production in the adrenal glomerulosa by atrial natriuretic factor." *Nature* 312 (5996):756-7. doi: 10.1038/312756a0.

Kuhn, M., R. Holtwick, H. A. Baba, J. C. Perriard, W. Schmitz, and E. Ehler. 2002. "Progressive cardiac hypertrophy and dysfunction in atrial natriuretic peptide receptor (GC-A) deficient mice." *Heart* 87 (4):368-74. doi: 10.1136/heart.87.4.368.

Kulig, W., J. Tynkkynen, M. Javanainen, M. Manna, T. Rog, I. Vattulainen, and P. Jungwirth. 2014. "How well does cholesteryl hemisuccinate mimic cholesterol in saturated phospholipid bilayers?" *J Mol Model* 20 (2):2121. doi: 10.1007/s00894-014-2121-z.

Kupitz, C., S. Basu, I. Grotjohann, R. Fromme, N. A. Zatsepin, K. N. Rendek, M. S. Hunter, R. L. Shoeman, T. A. White, D. Wang, D. James, J. H. Yang, D. E. Cobb, B. Reeder, R. G. Sierra, H. Liu, A. Barty, A. L. Aquila, D. Deponte, R. A. Kirian, S. Bari, J. J. Bergkamp, K. R. Beyerlein, M. J. Bogan, C. Caleman, T. C. Chao, C. E. Conrad, K. M. Davis, H. Fleckenstein, L. Galli, S. P. Hau-Riege, S. Kassemeyer, H. Laksmono, M. Liang, L. Lomb, S. Marchesini, A. V. Martin, M. Messerschmidt, D. Milathianaki, K. Nass, A. Ros, S. Roy-Chowdhury, K. Schmidt, M. Seibert, J. Steinbrener, F. Stellato, L. Yan, C. Yoon, T. A. Moore, A. L. Moore, Y. Pushkar, G. J. Williams, S. Boutet, R. B. Doak, U. Weierstall, M.

Frank, H. N. Chapman, J. C. Spence, and P. Fromme. 2014. "Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser." *Nature* 513 (7517):261-5. doi: 10.1038/nature13453.

Kurose, H., T. Inagami, and M. Ui. 1987. "Participation of adenosine 5'-triphosphate in the activation of membrane-bound guanylate cyclase by the atrial natriuretic factor." *FEBS Lett* 219 (2):375-9. doi: 10.1016/0014-5793(87)80256-9.

Kuruma, Y., and T. Ueda. 2016. "Corrigendum: The PURE system for the cell-free synthesis of membrane proteins." *Nat Protoc* 11 (3):616. doi: 10.1038/nprot0316.616d.

Kuznetsov, Yurii G, Alexander J Malkin, and Alexander McPherson. 2001. "The liquid protein phase in crystallization: a case study—intact immunoglobulins." *Journal of crystal growth* 232 (1-4):30-39.

Labrecque, J., N. Mc Nicoll, M. Marquis, and A. De Lean. 1999. "A disulfide-bridged mutant of natriuretic peptide receptor-A displays constitutive activity. Role of receptor dimerization in signal transduction." *J Biol Chem* 274 (14):9752-9. doi: 10.1074/jbc.274.14.9752.

Labrou, N. E. 2021. "Protein Purification Technologies." *Methods Mol Biol* 2178:3-10. doi: 10.1007/978-1-0716-0775-6_1.

Lafontan, M., C. Moro, M. Berlan, F. Crampes, C. Sengenes, and J. Galitzky. 2008. "Control of lipolysis by natriuretic peptides and cyclic GMP." *Trends Endocrinol Metab* 19 (4):130-7. doi: 10.1016/j.tem.2007.11.006.

Lalaurie, C. J., V. Dufour, A. Meletiou, S. Ratcliffe, A. Harland, O. Wilson, C. Vamasiri, D. K. Shoemark, C. Williams, C. J. Arthur, R. B. Sessions, M. P. Crump, J. L. R. Anderson, and P. Curnow. 2018. "The de novo design of a biocompatible and functional integral membrane protein using minimal sequence complexity." *Sci Rep* 8 (1):14564. doi: 10.1038/s41598-018-31964-8.

Landau, E. M., and J. P. Rosenbusch. 1996. "Lipidic cubic phases: a novel concept for the crystallization of membrane proteins." *Proc Natl Acad Sci U S A* 93 (25):14532-5. doi: 10.1073/pnas.93.25.14532.

Lee, C. Y., and J. C. Burnett, Jr. 2007. "Natriuretic peptides and therapeutic applications." *Heart Fail Rev* 12 (2):131-42. doi: 10.1007/s10741-007-9016-3.

Lee, D. I., G. Zhu, T. Sasaki, G. S. Cho, N. Hamdani, R. Holewinski, S. H. Jo, T. Danner, M. Zhang, P. P. Rainer, D. Bedja, J. A. Kirk, M. J. Ranek, W. R. Dostmann, C. Kwon, K. B. Margulies, J. E. Van Eyk, W. J. Paulus, E. Takimoto, and D. A. Kass. 2015. "Phosphodiesterase 9A controls nitric-oxide-independent cGMP and hypertrophic heart disease." *Nature* 519 (7544):472-6. doi: 10.1038/nature14332.

Lee, D. W., M. Peggie, M. Deak, R. Toth, Z. O. Gage, N. Wood, C. Schilde, T. Kurz, and A. Knebel. 2012. "The Dac-tag, an affinity tag based on penicillin-binding protein 5." *Anal Biochem* 428 (1):64-72. doi: 10.1016/j.ab.2012.06.007.

Leon, B. M., and T. M. Maddox. 2015. "Diabetes and cardiovascular disease: Epidemiology, biological mechanisms, treatment recommendations and future research." *World J Diabetes* 6 (13):1246-58. doi: 10.4239/wjd.v6.i13.1246.

Leonidas, Demetrios D., Boris L. Elbert, Zeqi Zhou, Hakon Leffler, Steven J. Ackerman, and K. Ravi Acharya. 1995. "Crystal structure of human Charcot–Leyden crystal protein, an eosinophil lysophospholipase, identifies it as a new member of the carbohydrate-binding family of galectins." *Structure* 3 (12):1379-1393. doi: https://doi.org/10.1016/S0969-2126(01)00275-1.

Levin, E. R., D. G. Gardner, and W. K. Samson. 1998. "Natriuretic peptides." *N Engl J Med* 339 (5):321-8. doi: 10.1056/NEJM199807303390507.

Li, D., S. T. Shah, and M. Caffrey. 2013. "Host Lipid and Temperature as Important Screening Variables for Crystallizing Integral Membrane Proteins in Lipidic Mesophases. Trials with Diacylglycerol Kinase." *Cryst Growth Des* 13 (7):2846-2857. doi: 10.1021/cg400254v.

Li, Y., S. Hashim, and M. B. Anand-Srivastava. 2006. "Intracellular peptides of natriuretic peptide receptor-C inhibit vascular hypertrophy via Gqalpha/MAP kinase signaling pathways." *Cardiovasc Res* 72 (3):464-72. doi: 10.1016/j.cardiores.2006.08.012.

Light, D. B., J. D. Corbin, and B. A. Stanton. 1990. "Dual ion-channel regulation by cyclic GMP and cyclic GMP-dependent protein kinase." *Nature* 344 (6264):336-9. doi: 10.1038/344336a0.

Lin, Z., Q. Zhao, L. Xing, B. Zhou, and X. Wang. 2015. "Aggregating tags for column-free protein purification." *Biotechnol J* 10 (12):1877-86. doi: 10.1002/biot.201500299.

Linder, J. U., and J. E. Schultz. 2003. "The class III adenylyl cyclases: multi-purpose signalling modules." *Cell Signal* 15 (12):1081-9. doi: 10.1016/s0898-6568(03)00130-x.

Liu, S., Z. Li, B. Yu, S. Wang, Y. Shen, and H. Cong. 2020. "Recent advances on protein separation and purification methods." *Adv Colloid Interface Sci* 284:102254. doi: 10.1016/j.cis.2020.102254.

Lojewska, E., T. Kowalczyk, S. Olejniczak, and T. Sakowicz. 2016. "Extraction and purification methods in downstream processing of plant-based recombinant proteins." *Protein Expr Purif* 120:110-7. doi: 10.1016/j.pep.2015.12.018.

Lowe, D. G. 1992. "Human natriuretic peptide receptor-A guanylyl cyclase is self-associated prior to hormone binding." *Biochemistry* 31 (43):10421-5. doi: 10.1021/bi00158a001.

Luckow, V. A., S. C. Lee, G. F. Barry, and P. O. Olins. 1993. "Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in Escherichia coli." *J Virol* 67 (8):4566-79. doi: 10.1128/JVI.67.8.4566-4579.1993.

Lumsden, N. G., R. S. Khambata, and A. J. Hobbs. 2010. "C-type natriuretic peptide (CNP): cardiovascular roles and potential as a therapeutic target." *Curr Pharm Des* 16 (37):4080-8. doi: 10.2174/138161210794519237.

Lynn, D. E. 2007. "Available lepidopteran insect cell lines." *Methods Mol Biol* 388:117-38. doi: 10.1007/978-1-59745-457-5_6.

Ma, T. K., K. K. Kam, B. P. Yan, and Y. Y. Lam. 2010. "Renin-angiotensin-aldosterone system blockade for cardiovascular diseases: current status." *Br J Pharmacol* 160 (6):1273-92. doi: 10.1111/j.1476-5381.2010.00750.x.

Macchia, D. D. 1987. "Atrial natriuretic factor: a hormone secreted by the heart." *Pharm Weekbl Sci* 9 (6):305-14. doi: 10.1007/BF01956510.

MacFarland, R. T., B. D. Zelus, and J. A. Beavo. 1991. "High concentrations of a cGMP-stimulated phosphodiesterase mediate ANP-induced decreases in cAMP and steroidogenesis in adrenal glomerulosa cells." *J Biol Chem* 266 (1):136-42.

Mahinrad, S., A. J. M. de Craen, S. Yasar, D. van Heemst, and B. Sabayan. 2016. "Natriuretic peptides in the central nervous system: Novel targets for cognitive impairment." *Neurosci Biobehav Rev* 68:148-156. doi: 10.1016/j.neubiorev.2016.05.022.

Mahmoudi Gomari, M., N. Saraygord-Afshari, M. Farsimadan, N. Rostami, S. Aghamiri, and M. M. Farajollahi. 2020. "Opportunities and challenges of the tag-assisted protein purification techniques: Applications in the pharmaceutical industry." *Biotechnol Adv* 45:107653. doi: 10.1016/j.biotechadv.2020.107653.

Maisel, A. S., P. Krishnaswamy, R. M. Nowak, J. McCord, J. E. Hollander, P. Duc, T. Omland, A. B. Storrow, W. T. Abraham, A. H. Wu, P. Clopton, P. G. Steg, A. Westheim, C. W. Knudsen, A. Perez, R. Kazanegra, H. C. Herrmann, P. A. McCullough, and Investigators Breathing Not Properly Multinational Study. 2002. "Rapid measurement of B-type natriuretic peptide in the emergency diagnosis of heart failure." *N Engl J Med* 347 (3):161-7. doi: 10.1056/NEJMoa020233.

Malhotra, Arun. 2009. "Chapter 16 Tagging for Protein Expression." In *Methods in Enzymology*, edited by Richard R. Burgess and Murray P. Deutscher, 239-258. Academic Press.

Mallela, J., S. Ravi, F. Jean Louis, B. Mulaney, M. Cheung, U. Sree Garapati, V. Chinnasamy, C. Wang, S. Nagaraj, S. S. Mohapatra, and S. Mohapatra. 2013. "Natriuretic peptide receptor A signaling regulates stem cell recruitment and angiogenesis: a model to study linkage between inflammation and tumorigenesis." *Stem Cells* 31 (7):1321-9. doi: 10.1002/stem.1376.

Mandal, A., J. C. Boatz, T. B. Wheeler, and P. C. van der Wel. 2017. "On the use of ultracentrifugal devices for routine sample preparation in biomolecular magic-angle-spinning NMR." *J Biomol NMR* 67 (3):165-178. doi: 10.1007/s10858-017-0089-6.

Manolopoulou, M., Q. Guo, E. Malito, A. B. Schilling, and W. J. Tang. 2009. "Molecular basis of catalytic chamber-assisted unfolding and cleavage of human insulin by human insulin-degrading enzyme." *J Biol Chem* 284 (21):14177-88. doi: 10.1074/jbc.M900068200.

Marala, R. B., A. Sitaramayya, and R. K. Sharma. 1991. "Dual regulation of atrial natriuretic factor-dependent guanylate cyclase activity by ATP." *FEBS Lett* 281 (1-2):73-6. doi: 10.1016/0014-5793(91)80361-6.

Marin-Grez, M., J. T. Fleming, and M. Steinhausen. 1986. "Atrial natriuretic peptide causes pre-glomerular vasodilatation and post-glomerular vasoconstriction in rat kidney." *Nature* 324 (6096):473-6. doi: 10.1038/324473a0.

Martiel, I., H. M. Muller-Werkmeister, and A. E. Cohen. 2019. "Strategies for sample delivery for femtosecond crystallography." *Acta Crystallogr D Struct Biol* 75 (Pt 2):160-177. doi: 10.1107/S2059798318017953.

Martin-Garcia, J. M., C. E. Conrad, G. Nelson, N. Stander, N. A. Zatsepin, J. Zook, L. Zhu, J. Geiger, E. Chun, D. Kissick, M. C. Hilgart, C. Ogata, A. Ishchenko, N. Nagaratnam, S. Roy-Chowdhury, J. Coe, G. Subramanian, A. Schaffer, D. James, G. Ketwala, N. Venugopalan, S. Xu, S. Corcoran, D. Ferguson, U. Weierstall, J. C. H. Spence, V. Cherezov, P. Fromme, R. F. Fischetti, and W. Liu. 2017. "Serial millisecond crystallography of membrane and soluble protein microcrystals using synchrotron radiation." *IUCrJ* 4 (Pt 4):439-454. doi: 10.1107/S205225251700570X.

Martynowycz, Michael W., Calina Glynn, Jennifer Miao, M. Jason de la Cruz, Johan Hattne, Dan Shi, Duilio Cascio, Jose Rodriguez, and Tamir Gonen. 2017. "MicroED Structures from Micrometer Thick Protein Crystals." *bioRxiv*:152504. doi: 10.1101/152504.

Massotte, D. 2003. "G protein-coupled receptor overexpression with the baculovirus-insect cell system: a tool for structural and functional studies." *Biochim Biophys Acta* 1610 (1):77-89. doi: 10.1016/s0005-2736(02)00720-4.

Matilainen, H., J. Rinne, L. Gilbert, V. Marjomaki, H. Reunanen, and C. Oker-Blom. 2005. "Baculovirus entry into human hepatoma cells." *J Virol* 79 (24):15452-9. doi: 10.1128/JVI.79.24.15452-15459.2005.

Matsue, Y., N. Kagiyama, K. Yoshida, T. Kume, H. Okura, M. Suzuki, A. Matsumura, K. Yoshida, and Y. Hashimoto. 2015. "Carperitide Is Associated With Increased In-Hospital Mortality in Acute Heart Failure: A Propensity Score-Matched Analysis." *J Card Fail* 21 (11):859-64. doi: 10.1016/j.cardfail.2015.05.007.

Matsukawa, N., W. J. Grzesik, N. Takahashi, K. N. Pandey, S. Pang, M. Yamauchi, and O. Smithies. 1999. "The natriuretic peptide clearance receptor locally modulates the physiological effects of the natriuretic peptide system." *Proc Natl Acad Sci U S A* 96 (13):7403-8. doi: 10.1073/pnas.96.13.7403.

Matsuo, H. 2001. "Discovery of a natriuretic peptide family and their clinical application." *Can J Physiol Pharmacol* 79 (8):736-40.

Matthews, B. W. 1968. "Solvent content of protein crystals." *J Mol Biol* 33 (2):491-7. doi: 10.1016/0022-2836(68)90205-2.

McCue, J. T. 2009. "Theory and use of hydrophobic interaction chromatography in protein purification applications." *Methods Enzymol* 463:405-14. doi: 10.1016/S0076-6879(09)63025-1.

McKenzie, E. A., and W. M. Abbott. 2018. "Expression of recombinant proteins in insect and mammalian cells." *Methods* 147:40-49. doi: 10.1016/j.ymeth.2018.05.013.

McKie, P. M., A. Cataliotti, G. Boerrigter, H. H. Chen, S. J. Sangaralingham, F. L. Martin, T. Ichiki, and J. C. Burnett, Jr. 2010. "A novel atrial natriuretic peptide based therapeutic in experimental angiotensin II mediated acute hypertension." *Hypertension* 56 (6):1152-9. doi: 10.1161/HYPERTENSIONAHA.110.159210.

McPherson, A. 2017. "Protein Crystallization." *Methods Mol Biol* 1607:17-50. doi: 10.1007/978-1-4939-7000-1_2.

McPherson, A., and J. A. Gavira. 2014. "Introduction to protein crystallization." *Acta Crystallogr F Struct Biol Commun* 70 (Pt 1):2-20. doi: 10.1107/S2053230X13033141.

McPherson, Alex. 2009. "Introduction to the crystallization of biological macromolecules." *Current Topics in membranes* 63:5-23.

McPherson, Alexander. 1999. *Crystallization of biological macromolecules*.

Meems, L. M. G., and J. C. Burnett, Jr. 2016. "Innovative Therapeutics: Designer Natriuretic Peptides." *JACC Basic Transl Sci* 1 (7):557-567. doi: 10.1016/j.jacbts.2016.10.001.

Melander, Wayne, and Csaba Horváth. 1977. "Salt effects on hydrophobic interactions in precipitation and chromatography of proteins: an interpretation of the lyotropic series." *Archives of biochemistry and biophysics* 183 (1):200-215.

Melo, L. G., A. T. Veress, C. K. Chong, S. C. Pang, T. G. Flynn, and H. Sonnenberg. 1998. "Salt-sensitive hypertension in ANP knockout mice: potential role of abnormal plasma renin activity." *Am J Physiol* 274 (1):R255-61. doi: 10.1152/ajpregu.1998.274.1.R255.

Mendis, Shanthi, Pekka Puska, Bo Norrving, World Health Organization., World Heart Federation., and World Stroke Organization. 2011. *Global atlas on cardiovascular disease prevention and control*. Geneva: World Health Organization in collaboration with the World Heart Federation and the World Stroke Organization.

Mezzasoma, L., M. J. Peirce, A. Minelli, and I. Bellezza. 2017. "Natriuretic Peptides: The Case of Prostate Cancer." *Molecules* 22 (10). doi: 10.3390/molecules22101680.

Middaugh, C Russel, William A Tisel, Robert N Haire, and Andreas Rosenberg. 1979. "Determination of the apparent thermodynamic activities of saturated protein solutions." *Journal of Biological Chemistry* 254 (2):367-370.

Miller, M. 2009. "Dyslipidemia and cardiovascular risk: the importance of early prevention." *QJM* 102 (9):657-67. doi: 10.1093/qjmed/hcp065.

Milligan, G. 1999. "Exploring the dynamics of regulation of G protein-coupled receptors using green fluorescent protein." *Br J Pharmacol* 128 (3):501-10. doi: 10.1038/sj.bjp.0702824.

Mishin, A., A. Gusach, A. Luginina, E. Marin, V. Borshchevskiy, and V. Cherezov. 2019. "An outlook on using serial femtosecond crystallography in drug discovery." *Expert Opin Drug Discov* 14 (9):933-945. doi: 10.1080/17460441.2019.1626822.

Misono, K. S. 2000. "Atrial natriuretic factor binding to its receptor is dependent on chloride concentration: A possible feedback-control mechanism in renal salt regulation." *Circ Res* 86 (11):1135-9. doi: 10.1161/01.res.86.11.1135.

Misono, K. S., H. Ogawa, Y. Qiu, and C. M. Ogata. 2005. "Structural studies of the natriuretic peptide receptor: a novel hormone-induced rotation mechanism for transmembrane signal transduction." *Peptides* 26 (6):957-68. doi: 10.1016/j.peptides.2004.12.021.

Misono, K. S., J. S. Philo, T. Arakawa, C. M. Ogata, Y. Qiu, H. Ogawa, and H. S. Young. 2011. "Structure, signaling mechanism and regulation of the natriuretic peptide receptor guanylate cyclase." *FEBS J* 278 (11):1818-29. doi: 10.1111/j.1742-4658.2011.08083.x.

Misono, K. S., N. Sivasubramanian, K. Berkner, and X. Zhang. 1999. "Expression and purification of the extracellular ligand-binding domain of the atrial natriuretic peptide (ANP) receptor: monovalent binding with ANP induces 2:2 complexes." *Biochemistry* 38 (2):516-23. doi: 10.1021/bi982127v.

Miyanishi, H., S. Nobata, and Y. Takei. 2011. "Relative antidipsogenic potencies of six homologous natriuretic peptides in eels." *Zoolog Sci* 28 (10):719-26. doi: 10.2108/zsj.28.719.

Moormann, R. J., A. Bouma, J. A. Kramps, C. Terpstra, and H. J. De Smit. 2000. "Development of a classical swine fever subunit marker vaccine and companion diagnostic test." *Vet Microbiol* 73 (2-3):209-19. doi: 10.1016/s0378-1135(00)00146-2.

Morse, S. A., R. Zhang, V. Thakur, and E. Reisin. 2005. "Hypertension and the metabolic syndrome." *Am J Med Sci* 330 (6):303-10. doi: 10.1097/00000441-200512000-00008.

Moschen, A. R., V. Wieser, and H. Tilg. 2012. "Adiponectin: key player in the adipose tissue-liver crosstalk." *Curr Med Chem* 19 (32):5467-73. doi: 10.2174/092986712803833254.

Motulsky, H. J., and R. E. Brown. 2006. "Detecting outliers when fitting data with nonlinear regression - a new method based on robust nonlinear regression and the false discovery rate." *BMC Bioinformatics* 7:123. doi: 10.1186/1471-2105-7-123.

Moyes, A. J., and A. J. Hobbs. 2019. "C-type Natriuretic Peptide: A Multifaceted Paracrine Regulator in the Heart and Vasculature." *Int J Mol Sci* 20 (9). doi: 10.3390/ijms20092281.

Mozaffarian, D., E. J. Benjamin, A. S. Go, D. K. Arnett, M. J. Blaha, M. Cushman, S. de Ferranti, J. P. Despres, H. J. Fullerton, V. J. Howard, M. D. Huffman, S. E. Judd, B. M. Kissela, D. T. Lackland, J. H. Lichtman, L. D. Lisabeth, S. Liu, R. H. Mackey, D. B. Matchar, D. K. McGuire, E. R. Mohler, 3rd, C. S. Moy, P. Muntner, M. E. Mussolino, K. Nasir, R. W. Neumar, G. Nichol, L. Palaniappan, D. K. Pandey, M. J. Reeves, C. J. Rodriguez, P. D. Sorlie, J. Stein, A. Towfighi, T. N. Turan, S. S. Virani, J. Z. Willey, D. Woo, R. W. Yeh, M. B. Turner, Committee American Heart Association Statistics, and Subcommittee Stroke Statistics. 2015. "Heart disease and stroke statistics--2015 update: a report from the American Heart Association." *Circulation* 131 (4):e29-322. doi: 10.1161/CIR.0000000000000152.

Mullens, W., K. Damman, V. P. Harjola, A. Mebazaa, H. P. Brunner-La Rocca, P. Martens, J. M. Testani, W. H. W. Tang, F. Orso, P. Rossignol, M. Metra, G. Filippatos, P. M. Seferovic, F. Ruschitzka, and A. J. Coats. 2019. "The use of diuretics in heart failure with congestion - a position statement from the Heart Failure Association of the European Society of Cardiology." *Eur J Heart Fail* 21 (2):137-155. doi: 10.1002/ejhf.1369.

Muller, D., C. Schulze, H. Baumeister, F. Buck, and D. Richter. 1992. "Rat insulin-degrading enzyme: cleavage pattern of the natriuretic peptide hormones ANP, BNP, and CNP revealed by HPLC and mass spectrometry." *Biochemistry* 31 (45):11138-43. doi: 10.1021/bi00160a026.

Munro, S., and H. R. Pelham. 1984. "Use of peptide tagging to detect proteins expressed from cloned genes: deletion mapping functional domains of Drosophila hsp 70." *EMBO J* 3 (13):3087-93.

Nagai, T., N. Iwakami, M. Nakai, K. Nishimura, Y. Sumita, A. Mizuno, H. Tsutsui, H. Ogawa, T. Anzai, and Jroad-Dpc investigators. 2019. "Effect of intravenous carperitide versus nitrates as first-line vasodilators on in-hospital outcomes in hospitalized patients with acute heart failure: Insight from a nationwide claim-based database." *Int J Cardiol* 280:104-109. doi: 10.1016/j.ijcard.2019.01.049.

Nakamura, M., K. Ichikawa, M. Ito, B. Yamamori, T. Okinaka, N. Isaka, Y. Yoshida, S. Fujita, and T. Nakano. 1999. "Effects of the phosphorylation of myosin phosphatase by cyclic GMP-dependent protein kinase." *Cell Signal* 11 (9):671-6. doi: 10.1016/s0898-6568(99)00036-4.

Nam, K. H. 2019. "Sample Delivery Media for Serial Crystallography." *Int J Mol Sci* 20 (5). doi: 10.3390/ijms20051094.

Nannenga, B. L., and T. Gonen. 2016. "MicroED opens a new era for biological structure determination." *Curr Opin Struct Biol* 40:128-135. doi: 10.1016/j.sbi.2016.09.007.

Nannenga, B. L., D. Shi, A. G. W. Leslie, and T. Gonen. 2014. "High-resolution structure determination by continuous-rotation data collection in MicroED." *Nat Methods* 11 (9):927-930. doi: 10.1038/nmeth.3043.

Navani, N. K., W. K. Mok, and L. Yingfu. 2009. "In vitro selection of protein-binding DNA aptamers as ligands for biosensing applications." *Methods Mol Biol* 504:399-415. doi: 10.1007/978-1-60327-569-9_22.

Nelson, G., R. A. Kirian, U. Weierstall, N. A. Zatsepin, T. Farago, T. Baumbach, F. Wilde, F. B. Niesler, B. Zimmer, I. Ishigami, M. Hikita, S. Bajt, S. R. Yeh, D. L. Rousseau, H. N. Chapman, J. C. Spence, and M. Heymann. 2016. "Three-dimensional-printed gas dynamic virtual nozzles for x-ray laser sample delivery." *Opt Express* 24 (11):11515-30. doi: 10.1364/OE.24.011515.

Neutze, R., R. Wouts, D. van der Spoel, E. Weckert, and J. Hajdu. 2000. "Potential for biomolecular imaging with femtosecond X-ray pulses." *Nature* 406 (6797):752-7. doi: 10.1038/35021099.

Newman, J., J. Xu, and M. C. Willis. 2007. "Initial evaluations of the reproducibility of vapor-diffusion crystallization." *Acta Crystallogr D Biol Crystallogr* 63 (Pt 7):826-32. doi: 10.1107/S0907444907025784.

Newstead, S., J. Hobbs, D. Jordan, E. P. Carpenter, and S. Iwata. 2008. "Insights into outer membrane protein crystallization." *Mol Membr Biol* 25 (8):631-8. doi: 10.1080/09687680802526574.

Nielsen, J. 2013. "Production of biopharmaceutical proteins by yeast: advances through metabolic engineering." *Bioengineered* 4 (4):207-11. doi: 10.4161/bioe.22856.

Nilsson, D., A. Holmberg, H. Sinn, and U. Vogt. 2011. "Zone Plates for Hard X□Ray FEL Radiation." *AIP Conference Proceedings* 1365 (1):120-123. doi: 10.1063/1.3625319.

Nobata, S., A. Ventura, H. Kaiya, and Y. Takei. 2010. "Diversified cardiovascular actions of six homologous natriuretic peptides (ANP, BNP, VNP, CNP1, CNP3, and CNP4) in conscious eels." *Am J Physiol Regul Integr Comp Physiol* 298 (6):R1549-59. doi: 10.1152/ajpregu.00789.2009.

Nojiri, T., H. Hosoda, T. Tokudome, K. Miura, S. Ishikane, K. Otani, I. Kishimoto, Y. Shintani, M. Inoue, T. Kimura, N. Sawabata, M. Minami, T. Nakagiri, S. Funaki, Y. Takeuchi, H. Maeda, H. Kidoya, H. Kiyonari, G. Shioi, Y. Arai, T. Hasegawa, N. Takakura, M. Hori, Y. Ohno, M. Miyazato, N. Mochizuki, M. Okumura, and K. Kangawa. 2015. "Atrial natriuretic peptide prevents cancer metastasis through vascular endothelial cells." *Proc Natl Acad Sci U S A* 112 (13):4086-91. doi: 10.1073/pnas.1417273112.

Nussenzveig, D. R., J. A. Lewicki, and T. Maack. 1990. "Cellular mechanisms of the clearance function of type C receptors of atrial natriuretic factor." *J Biol Chem* 265 (34):20952-8.

O'Connor, C. M., R. C. Starling, A. F. Hernandez, P. W. Armstrong, K. Dickstein, V. Hasselblad, G. M. Heizer, M. Komajda, B. M. Massie, J. J. McMurray, M. S. Nieminen, C. J. Reist, J. L. Rouleau, K. Swedberg, K. F. Adams, Jr., S. D. Anker, D. Atar, A. Battler, R. Botero, N. R. Bohidar, J. Butler, N. Clausell, R. Corbalan, M. R. Costanzo, U. Dahlstrom, L. I. Deckelbaum, R. Diaz, M. E. Dunlap, J. A. Ezekowitz, D. Feldman, G. M. Felker, G. C. Fonarow, D. Gennevois, S. S. Gottlieb, J. A. Hill, J. E. Hollander, J. G. Howlett, M. P. Hudson, R. D. Kociol, H. Krum, A. Laucevicius, W. C. Levy, G. F. Mendez, M. Metra, S. Mittal, B. H. Oh, N. L. Pereira, P. Ponikowski, W. H. Tang, S. Tanomsup, J. R. Teerlink, F. Triposkiadis, R. W. Troughton, A. A. Voors, D. J. Whellan, F. Zannad, and R. M.

Califf. 2011. "Effect of nesiritide in patients with acute decompensated heart failure." *N Engl J Med* 365 (1):32-43. doi: 10.1056/NEJMoa1100171.

O'Fagain, C., P. M. Cummins, and B. F. O'Connor. 2011. "Gel-filtration chromatography." *Methods Mol Biol* 681:25-33. doi: 10.1007/978-1-60761-913-0_2.

O'Reilly, David R, Lois K Miller, and Verne A Luckow. 1994. *Baculovirus expression vectors: a laboratory manual*: Oxford University Press on Demand.

Ogawa, H., Y. Qiu, L. Huang, S. W. Tam-Chang, H. S. Young, and K. S. Misono. 2009. "Structure of the atrial natriuretic peptide receptor extracellular domain in the unbound and hormone-bound states by single-particle electron microscopy." *FEBS J* 276 (5):1347-55. doi: 10.1111/j.1742-4658.2009.06870.x.

Ogawa, H., Y. Qiu, C. M. Ogata, and K. S. Misono. 2004. "Crystal structure of hormone-bound atrial natriuretic peptide receptor extracellular domain: rotation mechanism for transmembrane signal transduction." *J Biol Chem* 279 (27):28625-31. doi: 10.1074/jbc.M313222200.

Ogawa, Haruo, Masami Kodama, and Kei Izumikawa. 2020. "Structural insight into hormone recognition and transmembrane signaling by the atrial natriuretic peptide receptor." *bioRxiv*:2020.07.09.196394. doi: 10.1101/2020.07.09.196394.

Ohkawa, T., and M. D. Welch. 2018. "Baculovirus Actin-Based Motility Drives Nuclear Envelope Disruption and Nuclear Egress." *Curr Biol* 28 (13):2153-2159 e4. doi: 10.1016/j.cub.2018.05.027.

Oikawa, S., M. Imai, A. Ueno, S. Tanaka, T. Noguchi, H. Nakazato, K. Kangawa, A. Fukuda, and H. Matsuo. 1984. "Cloning and sequence analysis of cDNA encoding a precursor for human atrial natriuretic polypeptide." *Nature* 309 (5970):724-6. doi: 10.1038/309724a0.

Oliver, P. M., S. W. John, K. E. Purdy, R. Kim, N. Maeda, M. F. Goy, and O. Smithies. 1998. "Natriuretic peptide receptor 1 expression influences blood pressures of mice in a dose-dependent manner." *Proc Natl Acad Sci U S A* 95 (5):2547-51. doi: 10.1073/pnas.95.5.2547.

Orellana, L. 2019. "Large-Scale Conformational Changes and Protein Function: Breaking the in silico Barrier." *Front Mol Biosci* 6:117. doi: 10.3389/fmolb.2019.00117.

Ormazabal, V., S. Nair, O. Elfeky, C. Aguayo, C. Salomon, and F. A. Zuniga. 2018. "Association between insulin resistance and the development of cardiovascular disease." *Cardiovasc Diabetol* 17 (1):122. doi: 10.1186/s12933-018-0762-4.

Owolabi, Mayowa, Jaime J Miranda, Joseph Yaria, and Bruce Ovbiagele. 2016. "Controlling cardiovascular diseases in low and middle income countries by

placing proof in pragmatism." *BMJ Global Health* 1 (3):e000105. doi: 10.1136/bmjgh-2016-000105.

Pan, S., H. H. Chen, D. M. Dickey, G. Boerrigter, C. Lee, L. S. Kleppe, J. L. Hall, A. Lerman, M. M. Redfield, L. R. Potter, J. C. Burnett, Jr., and R. D. Simari. 2009. "Biodesign of a renal-protective peptide based on alternative splicing of B-type natriuretic peptide." *Proc Natl Acad Sci U S A* 106 (27):11282-7. doi: 10.1073/pnas.0811851106.

Pande, K., C. D. Hutchison, G. Groenhof, A. Aquila, J. S. Robinson, J. Tenboer, S. Basu, S. Boutet, D. P. DePonte, M. Liang, T. A. White, N. A. Zatsepin, O. Yefanov, D. Morozov, D. Oberthuer, C. Gati, G. Subramanian, D. James, Y. Zhao, J. Koralek, J. Brayshaw, C. Kupitz, C. Conrad, S. Roy-Chowdhury, J. D. Coe, M. Metz, P. L. Xavier, T. D. Grant, J. E. Koglin, G. Ketawala, R. Fromme, V. Srajer, R. Henning, J. C. Spence, A. Ourmazd, P. Schwander, U. Weierstall, M. Frank, P. Fromme, A. Barty, H. N. Chapman, K. Moffat, J. J. van Thor, and M. Schmidt. 2016. "Femtosecond structural dynamics drives the trans/cis isomerization in photoactive yellow protein." *Science* 352 (6286):725-9. doi: 10.1126/science.aad5081.

Pandey, A., K. Shin, R. E. Patterson, X. Q. Liu, and J. K. Rainey. 2016. "Current strategies for protein production and purification enabling membrane protein structural biology." *Biochem Cell Biol* 94 (6):507-527. doi: 10.1139/bcb-2015-0143.

Pandey, K. N. 2002. "Intracellular trafficking and metabolic turnover of ligand-bound guanylyl cyclase/atrial natriuretic peptide receptor-A into subcellular compartments." *Mol Cell Biochem* 230 (1-2):61-72.

Pandey, K. N. 2011. "Guanylyl cyclase / atrial natriuretic peptide receptor-A: role in the pathophysiology of cardiovascular regulation." *Can J Physiol Pharmacol* 89 (8):557-73. doi: 10.1139/y11-054.

Pandey, K. N. 2015. "Endocytosis and Trafficking of Natriuretic Peptide Receptor-A: Potential Role of Short Sequence Motifs." *Membranes (Basel)* 5 (3):253-87. doi: 10.3390/membranes5030253.

Pankow, K., Y. Wang, F. Gembardt, E. Krause, X. Sun, G. Krause, H. P. Schultheiss, W. E. Siems, and T. Walther. 2007. "Successive action of meprin A and neprilysin catabolizes B-type natriuretic peptide." *Circ Res* 101 (9):875-82. doi: 10.1161/CIRCRESAHA.107.153585.

Park, M., P. Sandner, and T. Krieg. 2018. "cGMP at the centre of attention: emerging strategies for activating the cardioprotective PKG pathway." *Basic Res Cardiol* 113 (4):24. doi: 10.1007/s00395-018-0679-9.

Park, S. A., T. G. Kim, M. K. Han, K. C. Ha, S. Z. Kim, and Y. G. Kwak. 2012. "Dendroaspis natriuretic peptide regulates the cardiac L-type Ca2+ channel activity by the phosphorylation of alpha1c proteins." *Exp Mol Med* 44 (6):363-8. doi: 10.3858/emm.2012.44.6.041.

Parker, J. L., and S. Newstead. 2012. "Current trends in alpha-helical membrane protein crystallization: an update." *Protein Sci* 21 (9):1358-65. doi: 10.1002/pro.2122.

Parker, M. W. 2003. "Protein structure from x-ray diffraction." *J Biol Phys* 29 (4):341-62. doi: 10.1023/A:1027310719146.

Passarelli, A. L., and L. A. Guarino. 2007. "Baculovirus late and very late gene regulation." *Curr Drug Targets* 8 (10):1103-15. doi: 10.2174/138945007782151324.

Peake, N. J., A. J. Hobbs, B. Pingguan-Murphy, D. M. Salter, F. Berenbaum, and T. T. Chowdhury. 2014. "Role of C-type natriuretic peptide signalling in maintaining cartilage and bone function." *Osteoarthritis Cartilage* 22 (11):1800-7. doi: 10.1016/j.joca.2014.07.018.

Pearson, A. R., and P. Mehrabi. 2020. "Serial synchrotron crystallography for time-resolved structural biology." *Curr Opin Struct Biol* 65:168-174. doi: 10.1016/j.sbi.2020.06.019.

Pellegrini, Claudio, and Joachim Stöhr. 2003. "X-ray free-electron lasers—principles, properties and applications." *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 500 (1-3):33-40.

Phillips Jr, George N. 1990. "Comparison of the dynamics of myoglobin in different crystal forms." *Biophysical journal* 57 (2):381-383.

Porowinska, D., M. Wujak, K. Roszek, and M. Komoszynski. 2013. "[Prokaryotic expression systems]." *Postepy Hig Med Dosw (Online)* 67:119-29. doi: 10.5604/17322693.1038351.

Possee, R. D., R. B. Hitchman, K. S. Richards, S. G. Mann, E. Siaterli, C. P. Nixon, H. Irving, R. Assenberg, D. Alderton, R. J. Owens, and L. A. King. 2008. "Generation of baculovirus vectors for the high-throughput production of proteins in insect cells." *Biotechnol Bioeng* 101 (6):1115-22. doi: 10.1002/bit.22002.

Potpara, T. S., and G. Y. H. Lip. 2020. "Combining Anticoagulant and Antiplatelet Therapies for Chronic Atherosclerotic Disease: A Focus on Diabetes Mellitus as a High-Risk Patient Group." *Circulation* 141 (23):1855-1858. doi: 10.1161/CIRCULATIONAHA.120.046905.

Potter, L. R. 2011. "Guanylyl cyclase structure, function and regulation." *Cell Signal* 23 (12):1921-6. doi: 10.1016/j.cellsig.2011.09.001.

Potter, L. R., S. Abbey-Hosch, and D. M. Dickey. 2006. "Natriuretic peptides, their receptors, and cyclic guanosine monophosphate-dependent signaling functions." *Endocr Rev* 27 (1):47-72. doi: 10.1210/er.2005-0014.

Potter, L. R., and T. Hunter. 1998. "Phosphorylation of the kinase homology domain is essential for activation of the A-type natriuretic peptide receptor." *Mol Cell Biol* 18 (4):2164-72. doi: 10.1128/mcb.18.4.2164.

Potter, L. R., and T. Hunter. 1999a. "A constitutively "phosphorylated" guanylyl cyclase-linked atrial natriuretic peptide receptor mutant is resistant to desensitization." *Mol Biol Cell* 10 (6):1811-20. doi: 10.1091/mbc.10.6.1811.

Potter, L. R., and T. Hunter. 1999b. "Identification and characterization of the phosphorylation sites of the guanylyl cyclase-linked natriuretic peptide receptors A and B." *Methods* 19 (4):506-20. doi: 10.1006/meth.1999.0893.

Potter, L. R., A. R. Yoder, D. R. Flora, L. K. Antos, and D. M. Dickey. 2009. "Natriuretic peptides: their structures, receptors, physiologic functions and therapeutic applications." *Handb Exp Pharmacol* (191):341-66. doi: 10.1007/978-3-540-68964-5_15.

Puetz, John, and Florian M Wurm. 2019. "Recombinant proteins for industrial versus pharmaceutical purposes: a review of process and pricing." *Processes* 7 (8):476.

Qiu, Y., H. Ogawa, M. Miyagi, and K. S. Misono. 2004. "Constitutive activation and uncoupling of the atrial natriuretic peptide receptor by mutations at the dimer interface. Role of the dimer structure in signalling." *J Biol Chem* 279 (7):6115-23. doi: 10.1074/jbc.M310225200.

Qu, J., X. Zhao, X. Liu, Y. Sun, J. Wang, L. Liu, J. Wang, and J. Zhang. 2019. "Natriuretic peptide receptor a promotes breast cancer development by upregulating MMP9." *Am J Cancer Res* 9 (7):1415-1428.

Quianzon, C. C., and I. Cheikh. 2012. "History of insulin." *J Community Hosp Intern Med Perspect* 2 (2). doi: 10.3402/jchimp.v2i2.18701.

Quitterer, U., A. Pohl, A. Langer, S. Koller, and S. Abdalla. 2011. "A cleavable signal peptide enhances cell surface delivery and heterodimerization of Cerulean-tagged angiotensin II AT1 and bradykinin B2 receptor." *Biochem Biophys Res Commun* 409 (3):544-9. doi: 10.1016/j.bbrc.2011.05.041.

Qutub, Yasser, Ilya Reviakine, Carrie Maxwell, Javier Navarro, Ehud M. Landau, and Peter G. Vekilov. 2004. "Crystallization of Transmembrane Proteins in cubo:

Mechanisms of Crystal Growth and Defect Formation." *Journal of Molecular Biology* 343 (5):1243-1254. doi: https://doi.org/10.1016/j.jmb.2004.09.022.

Radner, S., P. H. Celie, K. Fuchs, W. Sieghart, T. K. Sixma, and M. Stornaiuolo. 2012. "Transient transfection coupled to baculovirus infection for rapid protein expression screening in insect cells." *J Struct Biol* 179 (1):46-55. doi: 10.1016/j.jsb.2012.04.013.

Rahmouni, K., M. L. Correia, W. G. Haynes, and A. L. Mark. 2005. "Obesity-associated hypertension: new insights into mechanisms." *Hypertension* 45 (1):9-14. doi: 10.1161/01.HYP.0000151325.83008.b4.

Rai, Praveen, Hemant Arya, and Diwakar Kumar. 2021. "Chapter 13 - Protein purification and desalting." In *The Design & Development of Novel Drugs and Vaccines*, edited by Tarun Kumar Bhatt and Surendra Nimesh, 181-201. Academic Press.

Ralat, L. A., Q. Guo, M. Ren, T. Funke, D. M. Dickey, L. R. Potter, and W. J. Tang. 2011. "Insulin-degrading enzyme modulates the natriuretic peptide-mediated signaling response." *J Biol Chem* 286 (6):4670-9. doi: 10.1074/jbc.M110.173252.

Rauch, A., M. Leipelt, M. Russwurm, and C. Steegborn. 2008. "Crystal structure of the guanylyl cyclase Cya2." *Proc Natl Acad Sci U S A* 105 (41):15720-5. doi: 10.1073/pnas.0808473105.

Rayment, I. 2002a. "Small-scale batch crystallization of proteins revisited: an underutilized way to grow large protein crystals." *Structure* 10 (2):147-51. doi: 10.1016/s0969-2126(02)00711-6.

Rayment, Ivan. 2002b. "Small-Scale Batch Crystallization of Proteins Revisited: An Underutilized Way to Grow Large Protein Crystals." *Structure* 10 (2):147-151. doi: https://doi.org/10.1016/S0969-2126(02)00711-6.

Redecke, L., K. Nass, D. P. DePonte, T. A. White, D. Rehders, A. Barty, F. Stellato, M. Liang, T. R. M. Barends, S. Boutet, G. J. Williams, M. Messerschmidt, M. M. Seibert, A. Aquila, D. Arnlund, S. Bajt, T. Barth, M. J. Bogan, C. Caleman, T. C. Chao, R. B. Doak, H. Fleckenstein, M. Frank, R. Fromme, L. Galli, I. Grotjohann, M. S. Hunter, L. C. Johansson, S. Kassemeyer, G. Katona, R. A. Kirian, R. Koopmann, C. Kupitz, L. Lomb, A. V. Martin, S. Mogk, R. Neutze, R. L. Shoeman, J. Steinbrener, N. Timneanu, D. Wang, U. Weierstall, N. A. Zatsepin, J. C. H. Spence, P. Fromme, I. Schlichting, M. Duszenko, C. Betzel, and H. N. Chapman. 2013. "Natively inhibited Trypanosoma brucei cathepsin B structure determined by using an X-ray laser." *Science* 339 (6116):227-230. doi: 10.1126/science.1229663.

Reen, D. J. 1994. "Enzyme-linked immunosorbent assay (ELISA)." *Methods Mol Biol* 32:461-6. doi: 10.1385/0-89603-268-X:461.

Rondeau, J. J., N. McNicoll, J. Gagnon, N. Bouchard, H. Ong, and A. De Lean. 1995. "Stoichiometry of the atrial natriuretic factor-R1 receptor complex in the bovine zona glomerulosa." *Biochemistry* 34 (7):2130-6. doi: 10.1021/bi00007a005.

Rose, R. A., and W. R. Giles. 2008. "Natriuretic peptide C receptor signalling in the heart and vasculature." *J Physiol* 586 (2):353-66. doi: 10.1113/jphysiol.2007.144253.

Ruiz, M., and J. W. Karpen. 1999. "Opening mechanism of a cyclic nucleotide-gated channel based on analysis of single channels locked in each liganded state." *J Gen Physiol* 113 (6):873-95. doi: 10.1085/jgp.113.6.873.

Rukavina Mikusic, Natalia L., María I. Rosón, Nicolás M. Kouyoumdzian, Silvana M. Cantú, Belisario E. Fernández, and Marcelo R. Choi. 2018. "Natriuretic Peptide Receptor Type A (NPRA)." In *Encyclopedia of Signaling Molecules*, edited by Sangdun Choi, 3344-3351. Cham: Springer International Publishing.

Ruotolo, G., and B. V. Howard. 2002. "Dyslipidemia of the metabolic syndrome." *Curr Cardiol Rep* 4 (6):494-500. doi: 10.1007/s11886-002-0113-6.

Saarenpaa, T., V. P. Jaakola, and A. Goldman. 2015. "Baculovirus-mediated expression of GPCRs in insect cells." *Methods Enzymol* 556:185-218. doi: 10.1016/bs.mie.2014.12.033.

Sadaf, Aiman, Kyung Ho Cho, Bernadette Byrne, and Pil Seok Chae. 2015. "Chapter Four - Amphipathic Agents for Membrane Protein Study." In *Methods in Enzymology*, edited by Arun K. Shukla, 57-94. Academic Press.

Saito, Y. 2010. "Roles of atrial natriuretic peptide and its therapeutic use." *J Cardiol* 56 (3):262-70. doi: 10.1016/j.jjcc.2010.08.001.

Samuli Ollila, O. H., Tomasz Róg, Mikko Karttunen, and Ilpo Vattulainen. 2007. "Role of sterol type on lateral pressure profiles of lipid membranes affecting membrane protein functionality: Comparison between cholesterol, desmosterol, 7-dehydrocholesterol and ketosterol." *Journal of Structural Biology* 159 (2):311-323. doi: https://doi.org/10.1016/j.jsb.2007.01.012.

Sanchez-Garcia, L., L. Martin, R. Mangues, N. Ferrer-Miralles, E. Vazquez, and A. Villaverde. 2016. "Recombinant pharmaceuticals from microbial cells: a 2015 update." *Microb Cell Fact* 15:33. doi: 10.1186/s12934-016-0437-3.

Sandefur, C. C., and I. Jialal. 2021. "Atrial Natriuretic Peptide." In *StatPearls*. Treasure Island (FL).

Sato, Y., Y. Cheng, K. Kawamura, S. Takae, and A. J. Hsueh. 2012. "C-type natriuretic peptide stimulates ovarian follicle development." *Mol Endocrinol* 26 (7):1158-66. doi: 10.1210/me.2012-1027.

Sauter, N. K. 2015. "XFEL diffraction: developing processing methods to optimize data quality." *J Synchrotron Radiat* 22 (2):239-48. doi: 10.1107/S1600577514028203.

Schäfer, Frank, Nicole Seip, Barbara Maertens, Helena Block, and Jan Kubicek. 2015. "Chapter Nine - Purification of GST-Tagged Proteins." In *Methods in Enzymology*, edited by Jon R. Lorsch, 127-139. Academic Press.

Schaub, Marcus C. 2007. "Myosin Light Chain Phosphatase." In *xPharm: The Comprehensive Pharmacology Reference*, edited by S. J. Enna and David B. Bylund, 1-3. New York: Elsevier.

Schiffrin, E. L. 1990. "Vascular receptors for atrial natriuretic peptide in hypertension." *Int J Rad Appl Instrum B* 17 (7):673-6. doi: 10.1016/0883-2897(90)90083-d.

Schlossmann, J., A. Ammendola, K. Ashman, X. Zong, A. Huber, G. Neubauer, G. X. Wang, H. D. Allescher, M. Korth, M. Wilm, F. Hofmann, and P. Ruth. 2000. "Regulation of intracellular calcium by a signalling complex of IRAG, IP3 receptor and cGMP kinase Ibeta." *Nature* 404 (6774):197-201. doi: 10.1038/35004606.

Schlueter, N., A. de Sterke, D. M. Willmes, J. Spranger, J. Jordan, and A. L. Birkenfeld. 2014. "Metabolic actions of natriuretic peptides and therapeutic potential in the metabolic syndrome." *Pharmacol Ther* 144 (1):12-27. doi: 10.1016/j.pharmthera.2014.04.007.

Schmidt, F. R. 2004. "Recombinant expression systems in the pharmaceutical industry." *Appl Microbiol Biotechnol* 65 (4):363-72. doi: 10.1007/s00253-004-1656-9.

Schneider, E. H., and R. Seifert. 2010. "Sf9 cells: a versatile model system to investigate the pharmacological properties of G protein-coupled receptors." *Pharmacol Ther* 128 (3):387-418. doi: 10.1016/j.pharmthera.2010.07.005.

Schonherr, R., J. M. Rudolph, and L. Redecke. 2018. "Protein crystallization in living cells." *Biol Chem* 399 (7):751-772. doi: 10.1515/hsz-2018-0158.

Schuck, P. 2003. "On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation." *Anal Biochem* 320 (1):104-24. doi: 10.1016/s0003-2697(03)00289-6.

Schulz, J., J. Bielecki, R. B. Doak, K. Dorner, R. Graceffa, R. L. Shoeman, M. Sikorski, P. Thute, D. Westphal, and A. P. Mancuso. 2019. "A versatile liquid-jet setup for the European XFEL." *J Synchrotron Radiat* 26 (Pt 2):339-345. doi: 10.1107/S1600577519000894.

Schweitz, H., P. Vigne, D. Moinier, C. Frelin, and M. Lazdunski. 1992. "A new member of the natriuretic peptide family is present in the venom of the green mamba (Dendroaspis angusticeps)." *J Biol Chem* 267 (20):13928-32.

Scopes, R. K. 2001. "Overview of protein purification and characterization." *Curr Protoc Protein Sci* Chapter 1:Unit 1 1. doi: 10.1002/0471140864.ps0101s00.

Seddon, A. M., P. Curnow, and P. J. Booth. 2004. "Membrane proteins, lipids and detergents: not just a soap opera." *Biochim Biophys Acta* 1666 (1-2):105-17. doi: 10.1016/j.bbamem.2004.04.011.

Sela, M., and R. Arnon. 1960. "Studies on the chemical basis of the antigenicity of proteins. 1. Antigenicity of polypeptidyl gelatins." *Biochem J* 75:91-102. doi: 10.1042/bj0750091.

Sengenes, C., M. Berlan, I. De Glisezinski, M. Lafontan, and J. Galitzky. 2000. "Natriuretic peptides: a new lipolytic pathway in human adipocytes." *FASEB J* 14 (10):1345-51.

Sengenes, C., A. Zakaroff-Girard, A. Moulin, M. Berlan, A. Bouloumie, M. Lafontan, and J. Galitzky. 2002. "Natriuretic peptide-dependent lipolysis in fat cells is a primate specificity." *Am J Physiol Regul Integr Comp Physiol* 283 (1):R257-65. doi: 10.1152/ajpregu.00453.2001.

Serafino, A., and P. Pierimarchi. 2014. "Atrial natriuretic peptide: a magic bullet for cancer therapy targeting Wnt signaling and cellular pH regulators." *Curr Med Chem* 21 (21):2401-9. doi: 10.2174/09298673216661402051401 52.

Shang, H., T. A. Garretson, C. M. S. Kumar, R. F. Dieter, and X. W. Cheng. 2017. "Improved pFastBac donor plasmid vectors for higher protein production using the Bac-to-Bac(R) baculovirus expression vector system." *J Biotechnol* 255:37-46. doi: 10.1016/j.jbiotec.2017.06.397.

Sharma, R. K. 2002. "Evolution of the membrane guanylate cyclase transduction system." *Mol Cell Biochem* 230 (1-2):3-30.

Sharma, R. K. 2010. "Membrane guanylate cyclase is a beautiful signal transduction machine: overview." *Mol Cell Biochem* 334 (1-2):3-36. doi: 10.1007/s11010-009-0336-6.

Sheriff, Steven, Wayne A Hendrickson, Ronald E Stenkamp, Larry C Sieker, and Lyle H Jensen. 1985. "Influence of solvent accessibility and intermolecular contacts on atomic mobilities in hemerythrins." *Proceedings of the National Academy of Sciences* 82 (4):1104-1107.

Shi, D., B. L. Nannenga, M. J. de la Cruz, J. Liu, S. Sawtelle, G. Calero, F. E. Reyes, J. Hattne, and T. Gonen. 2016. "The collection of MicroED data for macromolecular crystallography." *Nat Protoc* 11 (5):895-904. doi: 10.1038/nprot.2016.046.

Shi, D., B. L. Nannenga, M. G. Iadanza, and T. Gonen. 2013. "Three-dimensional electron crystallography of protein microcrystals." *Elife* 2:e01345. doi: 10.7554/eLife.01345.

Shi, S. J., H. T. Nguyen, G. D. Sharma, L. G. Navar, and K. N. Pandey. 2001. "Genetic disruption of atrial natriuretic peptide receptor-A alters renin and angiotensin II levels." *Am J Physiol Renal Physiol* 281 (4):F665-73. doi: 10.1152/ajprenal.2001.281.4.F665.

Shi, W. 2000. "SASE X-Ray free electron laser in DESY." *Journal of the Society of Chinese Physists* 6:5-16.

Shi, X., and D. L. Jarvis. 2007. "Protein N-glycosylation in the baculovirus-insect cell system." *Curr Drug Targets* 8 (10):1116-25. doi: 10.2174/138945007782151360.

Slack, J., and B. M. Arif. 2007. "The baculoviruses occlusion-derived virus: virion structure and function." *Adv Virus Res* 69:99-165. doi: 10.1016/S0065-3527(06)69003-9.

Slootstra, J. W., D. Kuperus, A. Pluckthun, and R. H. Meloen. 1997. "Identification of new tag sequences with differential and selective recognition properties for the anti-FLAG monoclonal antibodies M1, M2 and M5." *Mol Divers* 2 (3):156-64. doi: 10.1007/BF01682203.

Smith, G. E., M. D. Summers, and M. J. Fraser. 1983. "Production of human beta interferon in insect cells infected with a baculovirus expression vector." *Mol Cell Biol* 3 (12):2156-65. doi: 10.1128/mcb.3.12.2156.

Smith, M. W., E. A. Espiner, T. G. Yandle, C. J. Charles, and A. M. Richards. 2000. "Delayed metabolism of human brain natriuretic peptide reflects resistance to neutral endopeptidase." *J Endocrinol* 167 (2):239-46. doi: 10.1677/joe.0.1670239.

Smyth, M. S., and J. H. Martin. 2000. "x ray crystallography." *Mol Pathol* 53 (1):8-14. doi: 10.1136/mp.53.1.8.

Sobhanifar, S., S. Reckel, F. Junge, D. Schwarz, L. Kai, M. Karbyshev, F. Lohr, F. Bernhard, and V. Dotsch. 2010. "Cell-free expression and stable isotope labelling strategies for membrane proteins." *J Biomol NMR* 46 (1):33-43. doi: 10.1007/s10858-009-9364-5.

Soler, D. C., A. E. Young, A. Vahedi-Faridi, and T. S. McCormick. 2018. "Generation of Flp-in(tm)-ready DG44 and Lec 3.2.8.1 CHO cell lines for quick and easy constitutive protein expression." *Biotechniques* 65 (1):41-46. doi: 10.2144/btn-2018-0075.

Sonoda, Y., A. Cameron, S. Newstead, H. Omote, Y. Moriyama, M. Kasahara, S. Iwata, and D. Drew. 2010. "Tricks of the trade used to accelerate high-resolution structure determination of membrane proteins." *FEBS Lett* 584 (12):2539-47. doi: 10.1016/j.febslet.2010.04.015.

Srivastava, A. K. 2012. "Challenges in the treatment of cardiometabolic syndrome." *Indian J Pharmacol* 44 (2):155-6. doi: 10.4103/0253-7613.93579.

Stanley, P. 1989. "Chinese hamster ovary cell mutants with multiple glycosylation defects for production of glycoproteins with minimal carbohydrate heterogeneity." *Mol Cell Biol* 9 (2):377-83. doi: 10.1128/mcb.9.2.377.

Stetsenko, Artem, and Albert Guskov. 2017. "An Overview of the Top Ten Detergents Used for Membrane Protein Crystallization." *Crystals* 7 (7):197.

Stroud, Z., S. C. L. Hall, and T. R. Dafforn. 2018. "Purification of membrane proteins free from conventional detergents: SMA, new polymers, new opportunities and new insights." *Methods* 147:106-117. doi: 10.1016/j.ymeth.2018.03.011.

Sudoh, T., K. Kangawa, N. Minamino, and H. Matsuo. 1988. "A new natriuretic peptide in porcine brain." *Nature* 332 (6159):78-81. doi: 10.1038/332078a0.

Sudoh, T., N. Minamino, K. Kangawa, and H. Matsuo. 1990. "C-type natriuretic peptide (CNP): a new member of natriuretic peptide family identified in porcine brain." *Biochem Biophys Res Commun* 168 (2):863-70. doi: 10.1016/0006-291x(90)92401-k.

Suga, S., K. Nakao, K. Hosoda, M. Mukoyama, Y. Ogawa, G. Shirakami, H. Arai, Y. Saito, Y. Kambayashi, K. Inouye, and et al. 1992. "Receptor selectivity of natriuretic peptide family, atrial natriuretic peptide, brain natriuretic peptide, and C-type natriuretic peptide." *Endocrinology* 130 (1):229-39. doi: 10.1210/endo.130.1.1309330.

Summers, Max D, and Gale E Smith. 1987. "A manual of methods for baculovirus vectors and insect cell culture procedures."

Swift, S. L., G. C. Rivera, V. Dussupt, R. M. Leadley, L. C. Hudson, C. Ma de Ridder, R. Kraaij, J. E. Burns, N. J. Maitland, and L. J. Georgopoulos. 2013. "Evaluating baculovirus as a vector for human prostate cancer gene therapy." *PLoS One* 8 (6):e65557. doi: 10.1371/journal.pone.0065557.

Takei, Y., M. Ueki, and T. Nishizawa. 1994. "Eel ventricular natriuretic peptide: cDNA cloning and mRNA expression." *J Mol Endocrinol* 13 (3):339-45. doi: 10.1677/jme.0.0130339.

Takeuchi, K., K. Baskaran, and H. Arthanari. 2019. "Structure determination using solution NMR: Is it worth the effort?" *J Magn Reson* 306:195-201. doi: 10.1016/j.jmr.2019.07.045.

Tam, M. C., R. Lee, T. M. Cascino, M. C. Konerman, and S. L. Hummel. 2017. "Current Perspectives on Systemic Hypertension in Heart Failure with Preserved Ejection Fraction." *Curr Hypertens Rep* 19 (2):12. doi: 10.1007/s11906-017-0709-2.

Tamura, N., Y. Ogawa, H. Chusho, K. Nakamura, K. Nakao, M. Suda, M. Kasahara, R. Hashimoto, G. Katsuura, M. Mukoyama, H. Itoh, Y. Saito, I. Tanaka, H. Otani, and M. Katsuki. 2000. "Cardiac fibrosis in mice lacking brain natriuretic peptide." *Proc Natl Acad Sci U S A* 97 (8):4239-44. doi: 10.1073/pnas.070371497.

Tang, Y., J. Saul, N. Nagaratnam, J. M. Martin-Garcia, P. Fromme, J. Qiu, and J. LaBaer. 2020. "Construction of gateway-compatible baculovirus expression vectors for high-throughput protein expression and in vivo microcrystal screening." *Sci Rep* 10 (1):13323. doi: 10.1038/s41598-020-70163-2.

Tansey, J. T., C. Sztalryd, E. M. Hlavin, A. R. Kimmel, and C. Londos. 2004. "The central role of perilipin a in lipid metabolism and adipocyte lipolysis." *IUBMB Life* 56 (7):379-85. doi: 10.1080/15216540400009968.

Tawaragi, Y., K. Fuchimura, S. Tanaka, N. Minamino, K. Kangawa, and H. Matsuo. 1991. "Gene and precursor structures of human C-type natriuretic peptide." *Biochem Biophys Res Commun* 175 (2):645-51. doi: 10.1016/0006-291x(91)91614-i.

Taylor, R. S., V. A. Sagar, E. J. Davies, S. Briscoe, A. J. Coats, H. Dalal, F. Lough, K. Rees, and S. Singh. 2014. "Exercise-based rehabilitation for heart failure." *Cochrane Database Syst Rev* (4):CD003331. doi: 10.1002/14651858.CD003331.pub4.

Terpe, K. 2006. "Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems." *Appl Microbiol Biotechnol* 72 (2):211-22. doi: 10.1007/s00253-006-0465-8.

Thimiri Govinda Raj, D. B., N. A. Khan, S. Venkatachalam, and S. Arumugam. 2020. "BacMam System for Rapid Recombinant Protein Expression in Mammalian Cells." *Methods Mol Biol* 2125:205-208. doi: 10.1007/7651_2019_249.

Thiriet, Marc. 2018. "Cardiovascular Disease: An Introduction." In *Vasculopathies: Behavioral, Chemical, Environmental, and Genetic Factors*, 1-90. Cham: Springer International Publishing.

Thobani, A., D. S. Dhindsa, B. D. DeMoss, M. Raad, P. B. Sandesara, L. S. Sperling, and J. T. Baer. 2019. "Usefulness of Aspirin for Primary Prevention of Atherosclerotic

Cardiovascular Disease." *Am J Cardiol* 124 (11):1785-1789. doi: 10.1016/j.amjcard.2019.08.040.

Thondapu, V., O. Kurihara, T. Yonetsu, M. Russo, H. O. Kim, H. Lee, T. Soeda, Y. Minami, and I. K. Jang. 2019. "Comparison of Rosuvastatin Versus Atorvastatin for Coronary Plaque Stabilization." *Am J Cardiol* 123 (10):1565-1571. doi: 10.1016/j.amjcard.2019.02.019.

Tomasoni, D., M. Adamo, C. M. Lombardi, and M. Metra. 2019. "Highlights in heart failure." *ESC Heart Fail* 6 (6):1105-1127. doi: 10.1002/ehf2.12555.

Toop, T., D. Grozdanovski, and I. C. Potter. 1998. "Natriuretic peptide binding sites in the gills of the pouched lamprey Geotria australis." *J Exp Biol* 201 (Pt 11):1799-808.

Tsukamoto, O., M. Fujita, M. Kato, S. Yamazaki, Y. Asano, A. Ogai, H. Okazaki, M. Asai, Y. Nagamachi, N. Maeda, Y. Shintani, T. Minamino, M. Asakura, I. Kishimoto, T. Funahashi, H. Tomoike, and M. Kitakaze. 2009. "Natriuretic peptides enhance the production of adiponectin in human adipocytes and in patients with chronic heart failure." *J Am Coll Cardiol* 53 (22):2070-7. doi: 10.1016/j.jacc.2009.02.038.

Tsumoto, K., D. Ejima, A. M. Senczuk, Y. Kita, and T. Arakawa. 2007. "Effects of salts on protein-surface interactions: applications for column chromatography." *J Pharm Sci* 96 (7):1677-90. doi: 10.1002/jps.20821.

U.S. Food and Drug Administration. 1982. "Human insulin receives FDA approval." *FDA Drug Bull* 12 (3):18-9.

Usmani, S. S., G. Bedi, J. S. Samuel, S. Singh, S. Kalra, P. Kumar, A. A. Ahuja, M. Sharma, A. Gautam, and G. P. S. Raghava. 2017. "THPdb: Database of FDA-approved peptide and protein therapeutics." *PLoS One* 12 (7):e0181748. doi: 10.1371/journal.pone.0181748.

van den Akker, F., X. Zhang, M. Miyagi, X. Huo, K. S. Misono, and V. C. Yee. 2000. "Structure of the dimerized hormone-binding domain of a guanylyl-cyclase-coupled receptor." *Nature* 406 (6791):101-4. doi: 10.1038/35017602.

van Deursen, V. M., A. F. Hernandez, A. Stebbins, V. Hasselblad, J. A. Ezekowitz, R. M. Califf, S. S. Gottlieb, C. M. O'Connor, R. C. Starling, W. H. Tang, J. J. McMurray, K. Dickstein, and A. A. Voors. 2014. "Nesiritide, renal function, and associated outcomes during hospitalization for acute decompensated heart failure: results from the Acute Study of Clinical Effectiveness of Nesiritide and Decompensated Heart Failure (ASCEND-HF)." *Circulation* 130 (12):958-65. doi: 10.1161/CIRCULATIONAHA.113.003046.

van Kimmenade, R. R., and J. L. Januzzi, Jr. 2009. "The evolution of the natriuretic peptides - Current applications in human and animal medicine." *J Vet Cardiol* 11 Suppl 1:S9-21. doi: 10.1016/j.jvc.2009.01.001.

van Oers, M. M. 2011. "Opportunities and challenges for the baculovirus expression system." *J Invertebr Pathol* 107 Suppl:S3-15. doi: 10.1016/j.jip.2011.05.001.

van Oers, M. M., G. P. Pijlman, and J. M. Vlak. 2015. "Thirty years of baculovirus-insect cell protein expression: from dark horse to mainstream technology." *J Gen Virol* 96 (Pt 1):6-23. doi: 10.1099/vir.0.067108-0.

van Oers, M. M., and J. M. Vlak. 2007. "Baculovirus genomics." *Curr Drug Targets* 8 (10):1051-68. doi: 10.2174/138945007782151333.

Vasques, G. A., I. J. Arnhold, and A. A. Jorge. 2014. "Role of the natriuretic peptide system in normal growth and growth disorders." *Horm Res Paediatr* 82 (4):222-9. doi: 10.1159/000365049.

Vaughn, J. L., R. H. Goodwin, G. J. Tompkins, and P. McCawley. 1977. "The establishment of two cell lines from the insect Spodoptera frugiperda (Lepidoptera; Noctuidae)." *In Vitro* 13 (4):213-7. doi: 10.1007/BF02615077.

Vekilov, P. G., A. R. Feeling-Taylor, S. T. Yau, and D. Petsev. 2002. "Solvent entropy contribution to the free energy of protein crystallization." *Acta Crystallogr D Biol Crystallogr* 58 (Pt 10 Pt 1):1611-6. doi: 10.1107/s0907444902014312.

Vieira Gomes, A. M., T. Souza Carmo, L. Silva Carvalho, F. Mendonca Bahia, and N. S. Parachin. 2018. "Comparison of Yeasts as Hosts for Recombinant Protein Production." *Microorganisms* 6 (2). doi: 10.3390/microorganisms6020038.

Vieira, M. A., M. Gao, L. N. Nikonova, and T. Maack. 2001. "Molecular and cellular physiology of the dissociation of atrial natriuretic peptide from guanylyl cyclase a receptors." *J Biol Chem* 276 (39):36438-45. doi: 10.1074/jbc.M102208200.

VØlund, Aage. 1978. "Application of the four-parameter logistic model to bioassay: comparison with slope ratio and parallel line models." *Biometrics*:357-365.

von Lueder, T. G., S. J. Sangaralingham, B. H. Wang, A. R. Kompa, D. Atar, J. C. Burnett, Jr., and H. Krum. 2013. "Renin-angiotensin blockade combined with natriuretic peptide system augmentation: novel therapeutic concepts to combat heart failure." *Circ Heart Fail* 6 (3):594-605. doi: 10.1161/CIRCHEARTFAILURE.112.000289.

Wagner, C., A. Pfeifer, P. Ruth, F. Hofmann, and A. Kurtz. 1998. "Role of cGMP-kinase II in the control of renin secretion and renin expression." *J Clin Invest* 102 (8):1576-82. doi: 10.1172/JCI4044.

Walls, D., and S. T. Loughran. 2011. "Tagging recombinant proteins to enhance solubility and aid purification." *Methods Mol Biol* 681:151-75. doi: 10.1007/978-1-60761-913-0_9.

Wang, T. J. 2012. "The natriuretic peptides and fat metabolism." *N Engl J Med* 367 (4):377-8. doi: 10.1056/NEJMcibr1204796.

Watanabe, M., K. Miyazono, M. Tanokura, T. Sawasaki, Y. Endo, and I. Kobayashi. 2010. "Cell-free protein synthesis for structure determination by X-ray crystallography." *Methods Mol Biol* 607:149-60. doi: 10.1007/978-1-60327-331-2_13.

Watanabe, N., H. Murai, and I. Tanaka. 2002. "Semi-automatic protein crystallization system that allows in situ observation of X-ray diffraction from crystals in the drop." *Acta Crystallogr D Biol Crystallogr* 58 (Pt 10 Pt 1):1527-30. doi: 10.1107/s0907444902014154.

Wei, F. F., Z. Y. Zhang, Q. F. Huang, W. Y. Yang, and J. A. Staessen. 2018. "Resistant hypertension." *Kardiol Pol* 76 (7):1031-1042. doi: 10.5603/KP.a2018.0129.

Weierstall, U., D. James, C. Wang, T. A. White, D. Wang, W. Liu, J. C. Spence, R. Bruce Doak, G. Nelson, P. Fromme, R. Fromme, I. Grotjohann, C. Kupitz, N. A. Zatsepin, H. Liu, S. Basu, D. Wacker, G. W. Han, V. Katritch, S. Boutet, M. Messerschmidt, G. J. Williams, J. E. Koglin, M. Marvin Seibert, M. Klinker, C. Gati, R. L. Shoeman, A. Barty, H. N. Chapman, R. A. Kirian, K. R. Beyerlein, R. C. Stevens, D. Li, S. T. Shah, N. Howe, M. Caffrey, and V. Cherezov. 2014. "Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography." *Nat Commun* 5:3309. doi: 10.1038/ncomms4309.

Weil, J., R. E. Lang, H. Suttmann, U. Rampf, F. Bidlingmaier, and R. Gerzer. 1985. "Concomitant increase in plasma atrial natriuretic peptide and cyclic GMP during volume loading." *Klin Wochenschr* 63 (24):1265-8. doi: 10.1007/BF01738451.

White, S. H., and W. C. Wimley. 1999. "Membrane protein folding and stability: physical principles." *Annu Rev Biophys Biomol Struct* 28:319-65. doi: 10.1146/annurev.biophys.28.1.319.

White, T. A. 2019. "Processing serial crystallography data with CrystFEL: a step-by-step guide." *Acta Crystallogr D Struct Biol* 75 (Pt 2):219-233. doi: 10.1107/S205979831801238X.

White, T. A., V. Mariani, W. Brehm, O. Yefanov, A. Barty, K. R. Beyerlein, F. Chervinskii, L. Galli, C. Gati, T. Nakane, A. Tolstikova, K. Yamashita, C. H. Yoon, K. Diederichs, and H. N. Chapman. 2016. "Recent developments in CrystFEL." *J Appl Crystallogr* 49 (Pt 2):680-689. doi: 10.1107/S1600576716004751.

Wickham, T. J., T. Davis, R. R. Granados, M. L. Shuler, and H. A. Wood. 1992. "Screening of insect cell lines for the production of recombinant proteins and infectious virus in the baculovirus expression system." *Biotechnol Prog* 8 (5):391-6. doi: 10.1021/bp00017a003.

Wilson, E. M., and M. Chinkers. 1995. "Identification of sequences mediating guanylyl cyclase dimerization." *Biochemistry* 34 (14):4696-701. doi: 10.1021/bi00014a025.

Winger, J. A., E. R. Derbyshire, M. H. Lamers, M. A. Marletta, and J. Kuriyan. 2008. "The crystal structure of the catalytic domain of a eukaryotic guanylate cyclase." *BMC Struct Biol* 8:42. doi: 10.1186/1472-6807-8-42.

Wingfield, P. 2001. "Protein precipitation using ammonium sulfate." *Curr Protoc Protein Sci* Appendix 3:Appendix 3F. doi: 10.1002/0471140864.psa03fs13.

Wingfield, P. T. 2015. "Overview of the purification of recombinant proteins." *Curr Protoc Protein Sci* 80:6 1 1-6 1 35. doi: 10.1002/0471140864.ps0601s80.

Winick, Herman. 1997. "Fourth generation light sources." Proceedings of the 1997 Particle Accelerator Conference (Cat. No. 97CH36167).

Wittig, I., M. Karas, and H. Schagger. 2007. "High resolution clear native electrophoresis for in-gel functional assays and fluorescence studies of membrane protein complexes." *Mol Cell Proteomics* 6 (7):1215-25. doi: 10.1074/mcp.M700076-MCP200.

Wu, C., F. Wu, J. Pan, J. Morser, and Q. Wu. 2003. "Furin-mediated processing of Pro-C-type natriuretic peptide." *J Biol Chem* 278 (28):25847-52. doi: 10.1074/jbc.M301223200.

Wu, X., C. Cao, Y. Xu, and X. Lu. 2004. "Construction of a host range-expanded hybrid baculovirus of BmNPV and AcNPV, and knockout of cysteinase gene for more efficient expression." *Sci China C Life Sci* 47 (5):406-15. doi: 10.1360/03yc0128.

Wurm, F. M. 2004. "Production of recombinant protein therapeutics in cultivated mammalian cells." *Nat Biotechnol* 22 (11):1393-8. doi: 10.1038/nbt1026.

Yamada-Goto, N., G. Katsuura, K. Ebihara, M. Inuzuka, Y. Ochi, Y. Yamashita, T. Kusakabe, A. Yasoda, N. Satoh-Asahara, H. Ariyasu, K. Hosoda, and K. Nakao. 2013. "Intracerebroventricular administration of C-type natriuretic peptide suppresses food intake via activation of the melanocortin system in mice." *Diabetes* 62 (5):1500-4. doi: 10.2337/db12-0718.

Yan, W., F. Wu, J. Morser, and Q. Wu. 2000. "Corin, a transmembrane cardiac serine protease, acts as a pro-atrial natriuretic peptide-converting enzyme." *Proc Natl Acad Sci U S A* 97 (15):8525-9. doi: 10.1073/pnas.150149097.

Yandle, T. G., and A. M. Richards. 2015. "B-type Natriuretic Peptide circulating forms: Analytical and bioactivity issues." *Clin Chim Acta* 448:195-205. doi: 10.1016/j.cca.2015.07.004.

Yang, X., M. Pistolozzi, and Z. Lin. 2018. "New trends in aggregating tags for therapeutic protein purification." *Biotechnol Lett* 40 (5):745-753. doi: 10.1007/s10529-018-2543-2.

Yarla, N. S., H. Gali, G. Pathuri, S. Smriti, M. Farooqui, J. Panneerselvam, G. Kumar, V. Madka, and C. V. Rao. 2019. "Targeting the paracrine hormone-dependent guanylate cyclase/cGMP/phosphodiesterases signaling pathway for colorectal cancer prevention." *Semin Cancer Biol* 56:168-174. doi: 10.1016/j.semcancer.2018.08.011.

Yurkova, E. V., V. V. Demin, and N. G. Abdulaev. 1990. "Crystallization of membrane proteins: bovine rhodopsin." *Biomed Sci* 1 (6):585-90.

Zemella, A., L. Thoring, C. Hoffmeister, and S. Kubick. 2015. "Cell-Free Protein Synthesis: Pros and Cons of Prokaryotic and Eukaryotic Systems." *Chembiochem* 16 (17):2420-31. doi: 10.1002/cbic.201500340.

Zhang, H., G. Kurisu, J. L. Smith, and W. A. Cramer. 2003. "A defined protein-detergent-lipid complex for crystallization of integral membrane proteins: The cytochrome b6f complex of oxygenic photosynthesis." *Proc Natl Acad Sci U S A* 100 (9):5160-3. doi: 10.1073/pnas.0931431100.

Zhang, J., M. Li, Y. Yang, Y. Yan, J. Li, J. Qu, and J. Wang. 2015. "NPR-A: A Therapeutic Target in Inflammation and Cancer." *Crit Rev Eukaryot Gene Expr* 25 (1):41-6. doi: 10.1615/critreveukaryotgeneexpr.2015012447.

Zhang, J., Z. Zhao, and J. Wang. 2014. "Natriuretic peptide receptor A as a novel target for cancer." *World J Surg Oncol* 12:174. doi: 10.1186/1477-7819-12-174.

Zhang, M., Y. Q. Su, K. Sugiura, G. Xia, and J. J. Eppig. 2010. "Granulosa cell ligand NPPC and its receptor NPR2 maintain meiotic arrest in mouse oocytes." *Science* 330 (6002):366-9. doi: 10.1126/science.1193573.

Zhang, W., Y. Yang, W. Liu, Q. Chen, H. Wang, X. Wang, Y. Zhang, M. Zhang, and G. Xia. 2015. "Brain natriuretic peptide and C-type natriuretic peptide maintain porcine oocyte meiotic arrest." *J Cell Physiol* 230 (1):71-81. doi: 10.1002/jcp.24682.

Zhang, Z., M. Moo-Young, and Y. Chisti. 1996. "Plasmid stability in recombinant Saccharomyces cerevisiae." *Biotechnol Adv* 14 (4):401-35. doi: 10.1016/s0734-9750(96)00033-x.

Zhao, X., G. Li, and S. Liang. 2013. "Several affinity tags commonly used in chromatographic purification." *J Anal Methods Chem* 2013:581093. doi: 10.1155/2013/581093.

Zhao, Z., J. Zhang, M. Li, Y. Yang, K. Sun, and J. Wang. 2013. "ANP-NPRA signaling pathway--a potential therapeutic target for the treatment of malignancy." *Crit Rev Eukaryot Gene Expr* 23 (2):93-101. doi: 10.1615/critreveukargeneexpr.2013006641.

Zhu, J. 2012. "Mammalian cell protein expression for biopharmaceutical production." *Biotechnol Adv* 30 (5):1158-70. doi: 10.1016/j.biotechadv.2011.08.022.

Zhu, Y., L. N. Zhu, R. Guo, H. J. Cui, S. Ye, and Q. Fang. 2014. "Nanoliter-scale protein crystallization and screening with a microfluidic droplet robot." *Sci Rep* 4:5046. doi: 10.1038/srep05046.

Zink, F. E. 1997. "X-ray tubes." *Radiographics* 17 (5):1259-68. doi: 10.1148/radiographics.17.5.9308113.

APPENDIX A

PERMISSIONS

Chapter 1:

Figure 1.4 modified and reprinted from Misono, K. S., Philo, J. S., Arakawa, T., Ogata, C. M., Qiu, Y., Ogawa, H., & Young, H. S. (2011). Structure, signaling mechanism and regulation of the natriuretic peptide receptor guanylate cyclase. *The FEBS journal*, *278*(11), 1818-1829. With permission from FEBS journal.

Figure 1.7 reprinted from Schlueter, N., A. de Sterke, D. M. Willmes, J. Spranger, J. Jordan, and A. L. Birkenfeld. 2014. "Metabolic actions of natriuretic peptides and therapeutic potential in the metabolic syndrome." *Pharmacol Ther* 144 (1):12-27. doi: 10.1016/j.pharmthera.2014.04.007. With permission from ELSEVIER.

Chapter 2:

Figure 2.2.1 modified and reprinted from Dondapati, S. K., M. Stech, A. Zemella, and S. Kubick. 2020. "Cell-Free Protein Synthesis: A Promising Option for Future Drug Development." *BioDrugs* 34 (3):327-348. doi: 10.1007/s40259-020-00417-y.With permission from Springer Nature.

Figure 2.2.3 reprinted from van Oers, M. M., G. P. Pijlman, and J. M. Vlak. 2015. "Thirty years of baculovirus-insect cell protein expression: from dark horse to mainstream technology." *J Gen Virol* 96 (Pt 1):6-23. doi: 10.1099/vir.0.067108-0. With permission from the Microbiology Society.

Figure 2.2.5 modified and reprinted from Rai, Praveen, Hemant Arya, and Diwakar Kumar. 2021. "Chapter 13 - Protein purification and desalting." In *The Design & Development of Novel Drugs and Vaccines*, edited by Tarun Kumar Bhatt and Surendra Nimesh, 181-201. Academic Press. With permission from ELSEVIER.

Chapter 3:

Figure 3.2.1 reprinted from Boguszewska, K., M. Szewczuk, S. Urbaniak, and B. T. Karwowski. 2019. "Review: immunoassays in DNA damage and instability detection." *Cell Mol Life Sci* 76 (23):4689-4704. doi: 10.1007/s00018-019-03239-6. With permission from Springer Nature.

Chapter 4:

Figure 4.2.1 reprinted from Chayen, N. E., and E. Saridakis. 2008. "Protein crystallization: from purified protein to diffraction-quality crystal." *Nat Methods* 5 (2):147-53. doi: 10.1038/nmeth.f.203. With permission from Springer Nature.

Figure 4.2.2 reprinted from Li, D., S. T. Shah, and M. Caffrey. 2013. "Host Lipid and Temperature as Important Screening Variables for Crystallizing Integral Membrane Proteins in Lipidic Mesophases. Trials with Diacylglycerol Kinase." *Cryst Growth Des*

13 (7):2846-2857. doi: 10.1021/cg400254v. With permission from American Chemical Society.

Figure 4.2.3 reprinted from Birch, J., D. Axford, J. Foadi, A. Meyer, A. Eckhardt, Y. Thielmann, and I. Moraes. 2018. "The fine art of integral membrane protein crystallisation." *Methods* 147:150-162. doi: 10.1016/j.ymeth.2018.05.014. With permission from ELSEVIER.

Figure 4.2.4 modified and reprinted from Nilsson, D., A. Holmberg, H. Sinn, and U. Vogt. 2011. "Zone Plates for Hard X□Ray FEL Radiation." *AIP Conference Proceedings* 1365 (1):120-123. doi: 10.1063/1.3625319. With permission from AIP Publishing.

Figure 4.2.6 modified and reprinted from Mishin, A., A. Gusach, A. Luginina, E. Marin, V. Borshchevskiy, and V. Cherezov. 2019. "An outlook on using serial femtosecond crystallography in drug discovery." *Expert Opin Drug Discov* 14 (9):933-945. doi: 10.1080/17460441.2019.1626822. With permission from Taylor & Francis.