

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Genetic Incorporation of a Metal-chelating Unnatural Amino Acid and NMR Studies of the Viroporin p7 from Hepatitis C Virus

Permalink

<https://escholarship.org/uc/item/6sj1p43k>

Author

Wang, Vivian

Publication Date

2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Genetic Incorporation of a Metal-chelating Unnatural Amino Acid and
NMR Studies of the Viroprotein p7 from Hepatitis C Virus

A Thesis submitted in partial satisfaction of the requirements for the degree of
Master of Science

in

Chemistry

by

Vivian S. Wang

Committee in charge:

Stanley J. Opella, Chair
Patricia A. Jennings
Robert S. Pomeroy

2015

Copyright

Vivian S. Wang, 2015

All rights reserved.

This Thesis of Vivian S. Wang is approved, and it is acceptable in the quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2015

DEDICATION

To my mother, Amy, for her unending love,
selfless care, words of wisdom, and
for always believing in me.

EPIGRAPH

"He has made everything beautiful
in its own time; also He has put
eternity in their heart,
yet so that man does not find out
what God has done
from the beginning to the end."

Ecclesiastes 3:11

TABLE OF CONTENTS

Signature Page.....	iii
Dedication	iv
Epigraph.....	v
Table of Contents.....	vi
List of Abbreviations.....	viii
List of Figures	xi
Acknowledgements	xiii
Vita.....	xvii
Abstract of the Thesis.....	xviii
Chapter 1: Introduction to the study of p7 from HCV by NMR.....	1
1.1 Introduction to membrane proteins	1
1.2 NMR applications to membrane proteins.....	2
1.3 The viroporin p7 from HCV	5
Chapter 2: Expression and purification of hepatitis C genotype 1b J4 p7	8
2.1 Abstract.....	8
2.2 Introduction	9
2.3 Materials and methods	11
2.3.1 Cloning of p7 into a pET-31b(+) plasmid	11
2.3.2 Expression and growth of p7	13
2.3.3 Purification of p7.....	15
2.4 Results and discussion.....	17
2.5 Conclusion	25
Chapter 3: Paramagnetic relaxation enhancement of p7 by incorporation of the metal-chelating unnatural amino acid 2-amino-3-(8- hydroxyquinolin-3-yl)propanoic acid (HQA).	26
3.1 Abstract.....	26
3.2 Introduction	27
3.3. Paramagnetic protein NMR.....	30
3.4 Methods	37
3.4.1 Synthesis of HQA.....	37

3.4.2. <i>In vivo</i> incorporation of HQA into p7	39
3.4.3 Preparation of p7 for solution NMR and fluorescence studies.....	42
3.5 Results and discussion.....	43
3.5.1 Metal-chelating unnatural amino acid	43
3.5.2 Genetic incorporation of HQA into membrane proteins	46
3.5.3 Fluorescence induced by HQA-metal chelate	49
3.5.4 Intramolecular PREs by HQA-metal chelate	51
3.6 Conclusion	53
Chapter 4: NMR studies of p7 in various lipid membranes.....	55
4.1 Abstract	55
4.2 Introduction	55
4.3 Materials and methods	57
4.3.1 Reconstitution of p7 into DHPC and 6-O-PC micelles.....	57
4.3.2 Reconstitution of p7 into q=0.1 isotropic bicelles	57
4.3.3 Reconstitution of p7 into nanodiscs	58
4.3.3.1 Expression and purification of tobacco etch virus (TEV)	58
4.3.3.2 Expression and purification of membrane scaffolding proteins	59
4.3.3.3 Assembly and purification of nanodiscs.....	61
4.3.4 Reconstitution of p7 in DMPC proteoliposomes	62
4.4 Results and discussion.....	63
4.4.1 p7 in micelles and q=0.1 bicelles.....	63
4.4.2 p7 in nanodiscs.....	67
4.4.3 p7 in lipid bilayers	76
4.5 Conclusion	78
References.....	79

LIST OF ABBREVIATIONS

^1H	Proton
^2H	Deuterium
^{13}C	Carbon-13
^{15}N	Nitrogen-15
^{31}P	Phosphorus-31
6-O-PC	1,2-di-O-hexyl-sn-glycero-3-phosphocholine
ACN	Acetonitrile
AIBN	Azobisisobutyronitrile
AMS	Ammonium sulfate
CNBr	Cyanogen bromide
CP	Cross polarization
CSA	Chemical shift anisotropy
Cv	Column volume
DC	Dipolar coupling
DCM	Dichloromethane
DHPC	1,2-dihexanoyl-sn-glycero-3-phosphocholine
DMF	Dimethylformamide
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DMPG	1,2-dimyristoyl-sn-glycero-3-phospho-(1'-rac-glycerol)
DPC	Dodecylphosphocholine
DTT	Dithiothreitol

EDTA	Ethylenediamine tetracetic acid
FPLC	Fast protein liquid chromatography
HCV	Hepatitis C virus
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPC	Hexadecylphosphocholine
HPLC	High performance liquid chromatography
HQA	2-amino-3-(8-hydroxyquinolin-3-yl)propanoic acid
HSQC	Heteronuclear single quantum coherence
INEPT	Insensitive nuclei enhanced polarization transfer
IPTG	Isopropyl β -D-1-thiogalactopyranoside
kD	Kilo-dalton
LB	Luria-Bertani broth
M9	Minimal media
MAS	Magic angle spinning
MSP	Membrane scaffolding protein
NBS	N-Bromosuccinimide
Ni-NTA	Nickelnitrotriacetic acid
NOE	Nuclear Overhauser effect
NMR	Nuclear magnetic resonance
PBS	Phosphate buffered saline
PCS	Pseudocontact shift
PDB	Protein data bank
PMSF	Phenylmethanesulfonylfluoride

PRE	Paramagnetic relaxation enhancement
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
ssNMR	Solid-state NMR
TCEP	Tris(2-carboxyethyl)phosphine
TFA	Trifluoroacetic acid
TM	Transmembrane
TX-100	Triton X-100
UAA	Unnatural amino acid

LIST OF FIGURES

Figure 2.1. p7 vectors and sequence	10
Figure 2.2. Purification scheme of wild-type p7	18
Figure 2.3. SDS-PAGE of the purification of deuterated U- ¹⁵ N p7	19
Figure 2.4. SDS-PAGE of p7 autoinduction growths	21
Figure 2.5. FPLC chromatogram of post-CNBr cleavage separation	24
Figure 3.1. Mass spectrometry chromatogram and structure of HQA	36
Figure 3.2. Vector maps for unnatural amino acid incorporation into proteins in <i>E. coli</i>	46
Figure 3.3. SDS-PAGE of the purification of W48HQA p7	47
Figure 3.4. Comparison of ¹ H- ¹⁵ N HSQC spectra of uniformly ¹⁵ N-labeled p7 in DHPC micelles	48
Figure 3.5. Fluorescence spectra of HQA-incorporated membrane proteins in detergent micelles.....	50
Figure 3.6. Expanded region of the ¹ H- ¹⁵ N HSQC spectra of uniformly ¹⁵ N- labeled p7 in DHPC micelles	52
Figure 4.1. Chemical structures of commonly used lipids	64
Figure 4.2. ¹ H- ¹⁵ N HSQC of p7 in q=0.1 bicelles and 6-O-PC micelles	65
Figure 4.3. ¹ H- ¹⁵ N HSQC of p7 in 6-O-PC micelles and DHPC micelles.....	67
Figure 4.4. Purification of p7 in MSP1D1ΔH5 nanodiscs.....	68
Figure 4.5. ¹⁵ N-edited ¹ H 1-D spectra of nanodiscs by solution NMR	70
Figure 4.6. ¹ H- ¹⁵ N HSQC overlay of p7 in MSP1D1ΔH5 nanodiscs over p7 in q=0.1 bicelles.....	72
Figure 4.7. Native-PAGE vs. SDS-PAGE of p7 in nanodiscs	73

Figure 4.8. ^{31}P NMR of p7 in MSP1D1E3 nanodiscs.....	74
Figure 4.9. ^1H - ^{13}C cross-polarization and ^{13}C - ^{13}C correlation of p7 in DMPC proteoliposomes	77

ACKNOWLEDGEMENTS

First, I would like to thank my thesis advisor, Dr. Stanley Opella, for giving me the opportunity to learn and conduct research in his laboratory. I came into the lab with little experience in NMR and biochemistry, and so I am very grateful to him for accepting me as his student. I have grown so much as a scientist in my last two years in his laboratory. Without his guidance and encouragement, this Master's thesis would not be possible. I would like to also thank the rest of my thesis committee—Dr. Robert Pomeroy and Dr. Patricia Jennings—for their time and support of my research.

I would like to give a big thanks to all my colleagues in the Opella lab. Everyone has been patient and helpful with me as I learned to navigate the lab and eventually conduct my own experiments. I'd especially like to thank Dr. Ye Tian, Dr. Sang Ho Park, Jasmina Radoicic, Sabrina Berkamp, Dr. Hua Zhang, and Dr. Lindsay Dawson for getting me started in the wet lab. Without Tian's recommendation and encouragement I would not have ended up going back to graduate school. Sang Ho and Jasmina really helped me in troubleshooting and getting this thesis project started. Hua taught me the basics in protein purification, and Lindsay shared with me what she knew about p7. Sabrina has helped me so much from processing NMR spectra to troubleshooting wet lab problems. Thank you all for being the best mentors I could ask for.

I'd also like to thank my officemates who have come and gone—Dr. Anna Pavlova and Dr. Andreea Balaceanu. Thanks for listening and helping me troubleshoot. Thank you to Mignon Chu and Zheng Long, the other graduate students in the lab who have helped me in the lab. Thanks for all the laughs and learning experiences. I'd also like to thank Dr. Albert Chin Wu, Dr. Anna De Angelis, and Dr. Ratan Rai for all their help in the Center for NMR Spectroscopy and Imaging of Proteins (aka “the bubble”). Thanks for keeping all the instruments running smoothly, and thank you, Ratan, for helping me set up my first solid-state experiments. Of course, a big thank you goes to our lab assistant, Elena Vitoshka-Tarasov, for always being on top of things and keeping the lab organized and running smoothly. I'd also like to thank everyone else who has gone in and out of the lab—Dr. George Jiaozhi Lu, Dr. Gabriel Cook, Dr. Christopher Grant, Dr. Eugene Lin, Mitchell Zhao, Jun Heo, and Dr. Bibhuti Das. Finally from the lab, I'd also like to thank my undergraduate assistant researcher, Christa Lam, for her friendship and time in helping me synthesize HQA. Without you, Pac Hall wouldn't have been so much fun.

I would also like to thank Professor Michael Burkart, Dr. James La Clair, Dr. Heriberto Rivera, Dr. Justin Hammons, Kara Finzel, and Matt Jaremko from the Burkart Lab for all their help when I used their lab space for the synthesis of HQA. I'd like to acknowledge Jim and Harry especially for their patience, teaching me tips and tricks in the synthesis process and making sure I wasn't getting myself into too much trouble.

I would also like to thank Dr. Anthony Mrse and Dr. Xuemei Huang not only for all their help with solution NMR, but also for their friendship, words of wisdom and advice in life, NMR, and teaching. Thanks for keeping Protein Biochemistry Lab sane. I'd also like to thank George Anderson and Dr. Robert Pomeroy for teaching me most of what I know about analytical chemistry and helping me become a better teacher. I genuinely enjoyed being the teaching assistant for Analytical Chemistry Lab all these quarters. A big thank you especially to George for all his wisdom and career advice. Thanks for listening to me and encouraging me.

I'd also like to thank Dr. Donatella Chianelli, my supervisor from my undergraduate internship at GNF, for her help and encouragement from my undergraduate years until now. She's helped me plan Italy trips, written multiple recommendations for me, and supported all my professional endeavors. Thank you for everything you've done for me. The pEVOL-HQA plasmid was graciously donated by the Peter G. Schultz lab. I would also like to acknowledge Dr. Guipeun Kang (Judy E. Kim lab) for her help in obtaining the fluorescence studies in Chapter 3 and Dr. Yongxuan Su (UCSD Molecular Mass Spectrometry Facility) for his help in obtaining MS data in the synthesis of HQA and for being a wonderful mentor when I was an undergraduate assistant in the MS facility.

Lastly, I'd like to thank all my family and friends for their unending support, love, and prayer for me all these years. My mom, Amy, has been my biggest supporter and has never stopped believing in me. My sister, Susan, has

helped me so much outside of lab so that I could get all that I needed to get done in lab. My husband and best friend, Fred, is my number one fan. Thank you for putting up with my craziness during stressful times in graduate school (especially in writing this thesis!), for bringing me food when my experiments ran late, and for always seeing the best in me and cheering me on. I love you all so much. Finally, I thank my God, Jesus Christ, because out from Him and through Him and to Him are all things.

Thank you, everyone, for everything these past two years. This thesis would not have been possible without you.

Chapter 3 is a partial reprint of the material as it appears in the Journal of Biomolecular NMR, "Paramagnetic relaxation enhancement of membrane proteins by incorporation of the metal-chelating unnatural amino acid 2-amino-3-(8-hydroxyquinolin-3-yl)propanoic acid (HQA)" by S. H. Park, V. S. Wang, J. Radoicic, A. A. DeAngelis, S. Berkamp, and S. J. Opella, 2015. The thesis author was the secondary author of the paper.

VITA

- 2011 Bachelor of Science, University of California, San Diego
- 2013-2015 Teaching Assistant, Department of Chemistry and Biochemistry,
University of California, San Diego
- 2015 Master of Science, University of California, San Diego

PUBLICATION

Park, S.H., V.S. Wang, J. Radoicic, A.A. DeAngelis, S. Berkamp, and S.J. Opella.
(2014) "Paramagnetic relaxation enhancement of membrane proteins by
incorporation of the metal-chelating unnatural amino acid 2-amino-3-(8-
hydroxyquinolin-3-yl)propanoic acid (HQA)." *J. Biomol. NMR.* 61: 185-196.

ABSTRACT OF THE THESIS

Genetic Incorporation of a Metal-chelating Unnatural Amino Acid and NMR
Studies of the Viroporin p7 from Hepatitis C Virus

by

Vivian S. Wang

Master of Science in Chemistry

University of California, San Diego, 2015

Professor Stanley J. Opella, Chair

Hepatitis C virus (HCV) proteins are targets for potential pharmaceutical drugs since there are currently no vaccines against the virus, and drug treatments are expensive and prone to viral resistance. The channel-forming behavior and role of the viroporin p7 from HCV in virus assembly and release make it a promising choice for structure-based drug design. The structures of viroporins have only been successfully solved using nuclear magnetic spectroscopy (NMR). In the case of p7 and other

membrane proteins, the lipid environment plays a crucial role in the folding of the protein. Micellar environments have been known to distort membrane protein structures whereas bilayers provide a more native-like membrane environment. The structure of p7 has been solved in two micelle environments, but the distance between the transmembrane helices differ greatly between the two structures. One way to obtain long distance structural restraints in NMR is by obtaining paramagnetic relaxation enhancement (PRE) measurements through the incorporation of a metal-chelating unnatural amino acid (UAA). Here, an UAA with an 8-hydroxyquinoline moiety was successfully incorporated into the transmembrane portion of p7, and intramolecular paramagnetic relaxation enhancements were observed by solution NMR. The next step is to solve the structure of p7 in a lipid bilayer by solid-state NMR. In this research, efforts were made to move from isotropic detergent environments to larger native-like bilayer ones, with the goal of obtaining PREs with the UAA-incorporated p7 by solid-state NMR.

Chapter 1: Introduction to the study of p7 from HCV by NMR

1.1 Introduction to membrane proteins

Although membrane proteins, proteins that interact with the hydrophobic lipid bilayers of cell membranes, are the target of over 50% pharmaceutical drugs and comprise roughly 20-30% of genes in an organism's genome, their structure determination has been a challenging but growing work¹⁻⁴. Membrane proteins hold an important place in medicine research due to their function in signal transduction, ion transport, and cell-to-cell recognition. In humans, mutations or misfolding of membrane proteins have been linked to cancer, heart disease, and depression to name a few. Thus there is a particular need to understand the structure of membrane proteins in order to better understand their function and to make strides in structure-based drug design⁵⁻⁷.

To date, membrane proteins account for less than 1% of known protein structures. The first high resolution membrane protein structure published was that of the protein subunits of the photosynthetic reaction center from *Rhodospseudomonas viridis* in 1985⁸. Since then there have been over 1000 structures of membrane deposited in the Protein Database Bank (PDB) but this number pales greatly in comparison to the thousands of protein structures that are deposited into PDB each year. The hurdles present in membrane structure determination include their hydrophobic nature, flexibility, and instability¹. In order to isolate and purify these proteins, there are also challenges from the

use of detergents and denaturing reagents, poor solubility, low recombinant expression yields⁹, and the need for refolding and reconstitution into native-like lipid environments.

Membrane protein structure determination has proven to be a challenge in both x-ray crystallography⁷ and nuclear magnetic resonance (NMR)^{3, 4, 10, 11}, the two main methods used in structural biology. While x-ray crystallography uses x-ray diffraction to look at the electron density in a crystal⁷, NMR observes signals based on nuclear spin interactions in an external magnetic field. The main disadvantages of x-ray crystallography include the need for mutant proteins in order to obtain crystals and the inability to observe the protein in motion. NMR, on the other hand, can observe native proteins and their dynamics in both solution and solid-state³. It is mostly limited by instrumentation and complexity of the data but these two obstacles are becoming more trivial with advances in technology and computational biology.

1.2 NMR applications to membrane proteins

NMR itself can be divided into solution- and solid-state. Solution NMR has been widely used to study the structure and dynamics of soluble proteins¹²⁻¹⁴. Membrane proteins, however, have to be reconstituted into a lipid environment, increasing the size of the protein complex. In solution-state NMR, molecules undergo Brownian motion or fast isotropic tumbling^{3, 10, 15}, which averages out the anisotropic interactions on the NMR time scale and

results in sharp transition peaks. Because of this, solution NMR is limited by the tumbling or correlation time. Smaller molecules are able to tumble faster than larger molecules, so larger proteins give rise to decreased relaxation time, significant broadening of line widths, and lower intensity. Furthermore, larger proteins lead to spectral crowding. Broad line widths and spectral crowding give poorly resolved spectra, so solution NMR is mostly limited to proteins below 70kDa^{10, 15}. To circumvent some of these issues to a certain extent, deuterated proteins can be used^{16, 17}. Bacterial expression in D₂O results in 70-80% deuteration of the amino acid side chains (the side chain amide protons are exchangeable), which increases the relaxation time and improves spectral sensitivity and resolution. The use of deuterated protein, however, is limited by the high cost of deuterated media and bacterial adaptation to D₂O environments¹⁸.

Solid-state NMR (ssNMR), unlike solution NMR, has no size limitation^{4, 10, 11, 15, 19-21}. For solid-state of membrane proteins, the proteins have to be reconstituted in a large phospholipid bilayer in order to be pelleted and packed into a rotor. Micelles and small isotropic bicelles are too small to be isolated in this manner. In addition, the proteins should not undergo isotropic tumbling. Instead their mobility is restricted enough to only undergo rotational diffusion about the bilayer normal²². Thus the nuclear spin systems in ssNMR depend on the anisotropic or directionally dependent interactions. The two main anisotropic interactions are chemical shift anisotropy (CSA) and internuclear dipolar coupling (DC). Since these interactions are directionally

dependent they normally result in broad spectral line. To counter this, the sample can be spun at the magic angle (54.74° with respect to the magnetic field) to average the DC to zero and the CSA to a non-zero value. The sample can also be oriented or aligned (either mechanically or magnetically) so that the DC and CSA can be measured.

For both solution and solid-state NMR, long distance restraints for structure refinement can be obtained through paramagnetic relaxation enhancement (PRE) measurements. The unpaired electron from the paramagnetic center enhances the nuclear spin relaxation time and broadens the NMR signals. The extent of broadening and disappearance of peaks give insight to the distances between the paramagnetic center and certain nuclei—the closer to the paramagnetic site, the greater the broadening. Various paramagnetic centers include nitroxide radicals, Cu (II), Mn (II), and Gd (III). Metalloproteins have an innate ability to chelate to metals, but diamagnetic proteins need to be engineered in order to contain a paramagnetic center. Large lanthanide tags are often used, but their use is limited by their mobility or dynamics and the need for cysteine residues within the protein. A better alternative is the use of metal-chelating unnatural amino acids, which can be substituted for a single residue anywhere in the protein sequence. Their small size and rigidity within the protein make them an ideal choice for observing PREs.

As mentioned before, NMR studies of membrane proteins are also dependent on an appropriate lipid environment. Micelles, isotropic bicelles,

and nanodiscs are small enough to be detected by solution NMR²³, but again this is limited to smaller proteins. Micelle and isotropic bicelle environments²⁴⁻²⁷ are also not ideal or stable membrane bilayer mimetics²⁸ since short-chain lipids may distort the shape of integral membrane proteins. Proteoliposomes, large bilayer spheres made of long-chain lipids, form the closest resemblance to cell membranes, but their use is only feasible for solid-state NMR^{3, 9}. Large bicelles^{11, 25, 29}, nanodiscs³⁰⁻³³, and macrodiscs³⁴ have also been observed by solid-state NMR. In addition NMR experiment optimization and method development, the type of lipid and lipid environment are crucial in obtaining structures of membrane proteins in their near-native state.

1.3 The viroporin p7 from HCV

The hepatitis C virus (HCV) infects roughly 4 million new people annually worldwide, and there are currently 170 million people chronically affected (World Health Organization 2015, www.who.int). Unlike hepatitis A and B, there is no vaccine for hepatitis C, which is typically transferred from person-to-person through syringe sharing from drug usage and blood transfusions. The virus attacks liver cells, leading to cirrhosis and eventually hepatocellular carcinoma and death within 20-30 years. Treatment for HCV is expensive and thus not widely available to those who need it: one pill costs about \$1000 with regimens requiring at least 12 pills. With no vaccine available and the high cost of treatment, there is a need to develop cheaper alternative drugs³⁵.

HCV replication and release is localized in the endoplasmic reticulum of hepatocytes³⁶⁻³⁹. HCV has a positive RNA genome that encodes for a 3000aa polyprotein that is cleaved into ten mature proteins. At the N terminus of the polyprotein lie three structural proteins—the core protein and the envelope glycoproteins E1 and E2. At the C terminus are the six nonstructural proteins NS2, NS3, NS4A, NS4B, NS5A, and NS5B. Between the structural and nonstructure proteins is the viroporin p7^{36, 38-40}. Viroporins are small hydrophobic integral membrane proteins that form oligomeric pores or ion channels that disrupt the physiological properties of the host cell^{37, 41, 42}. Their role in affecting various cellular functions make p7 an ideal candidate for structure-based drug design³⁷. Although the oligomeric state of p7 is debated, the monomeric protein contains 63 residues that form two alpha-helical transmembrane portions with a short basic loop and N- and C-termini that face the endoplasmic reticulum lumen. Although p7 is not needed for virus replication, it is crucial for virus capsid assembly and release of the complete virus vesicle^{38, 39, 41}. Studies have also found that p7 forms ion channels to regulate pH in order to protect viral glycoproteins from degradation^{35, 38, 39}.

There are seven known genotypes of HCV with even more subtypes in each genotype. The amino acid sequences of p7 can vary between 10-50%^{5, 7, 43-45}. This genetic variability is accounted by the lack of a proof-reading mechanism of the viral RNA polymerase NS5A and fast replication rate of the virus. Despite having the same function, the structure of p7 has been surprisingly variable between genotypes^{6, 44}. The structure of p7 has been

solved in 1,2-dihexanoyl-sn-glycero-3-phosphocholine (DHPC) micelles³⁵, dodecylphosphocholine (DPC) micelles⁴⁶, methanol⁴⁷, and TFE/water⁴³, with some notable differences between them^{35, 38, 44, 46, 48}. In addition, the oligomeric states of p7 from different studies do not come to a consensus^{38, 43, 44, 46, 49}. Since none of the prior environments used for the study of p7 are ideal, the structure of p7 in a native-like environment—the lipid bilayer—needs to be examined in-depth using solid-state NMR. With accurate structures of the monomeric and oligomeric protein, structure-based drug design of a molecule disrupting the channel-forming ability of p7 can be developed to combat HCV infection.

Chapter 2: Expression and purification of hepatitis C genotype 1b J4 p7

2.1 Abstract

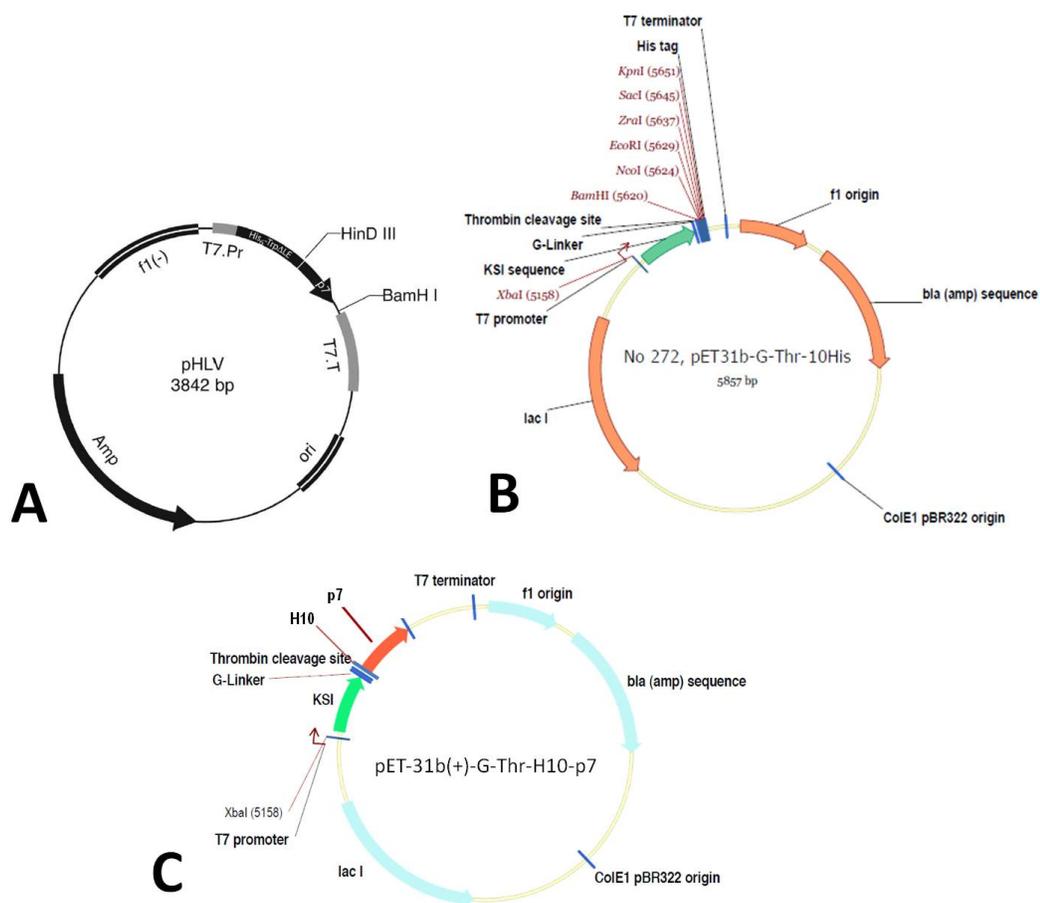
One method for paramagnetic relaxation enhancement studies of p7 utilizes the incorporation of a metal-chelating unnatural amino acid (UAA) as a single residue substitute in the protein sequence. *In vivo* incorporation of an UAA into a protein requires co-expression of two plasmids. Initial attempts in UAA incorporation failed due to plasmid incompatibility. Hepatitis C genotype 1b J4 p7 has been previously expressed in our lab using the expression vector pHLV-p7. Because this plasmid is not suitable for the co-expression with a second plasmid encoding the tRNA/tRNA synthetase orthogonal pair for an unnatural amino acid, this strain of p7 needed to be re-cloned into a vector that allowed for co-expression of a second plasmid. HCV p7 genotype 1b J4 p7 was successfully expressed and purified using a modified pET-31b(+) vector. Using this expression plasmid, p7 is driven into inclusion bodies using an N-terminal ketosteroid isomerase (KSI) fusion partner. A histidine tag was used to separate the p7 fusion from endogenous *E. coli* proteins. This vector contains one methionine flanking each end of the p7 region for cyanogen bromide chemical cleavage from the fusion partner. Alternatively, there is a thrombin cleavage site located between KSI and p7 for enzymatic cleavage of the two proteins. This vector is also compatible for co-expression of a second plasmid for UAA incorporation. The yield is 10 mg/L for pure wild-type

unlabeled p7, 6 mg/L for uniformly ^{15}N -labeled p7, and 3 mg/L for uniformly ^{13}C - ^{15}N -labeled p7.

2.2 Introduction

The hepatitis C virus displays high genetic diversity; it is divided into 6 genotypes with over 100 subtypes. Some genotypes vary over 33% in the viral genome^{39, 50}. Since there is currently no vaccine for HCV, there is a demand for drugs targeting HCV proteins to combat infection⁶. The search for an HCV-specific drug target is further complicated by the virus's high genetic variance. Genotype 1 is responsible for over 46% of cases, being the most prevalent genotype in the Americas, Europe, and East Asia. Genotype 3 accounts for about 30% cases primarily in India. Genotypes 2, 4, and 6 are responsible for about 22% of cases, and the rest are attributed to genotype 5³⁹.

Since genotype 1 is the most common variant, it is an ideal choice when looking at structure-based drug design. The three p7 genotype 1b strains (HCV-J, J4, and Con1) are slightly less common than p7 genotype 1a, but all of the p7 genotypes share over 80% sequence identity⁴³. Genotype 1b J4 was chosen for our studies due to its relative commonality and lower mutation deviation from the p7 consensus sequence. In addition, the wild-type J4 p7 sequence does not contain any methionines, making it suitable for cleavage reactions with cyanogen bromide to aid protein purification.



D 5' ALENLVVLNAA SVAGAHGILSFLVFFSAAWYIKGRLAPGAAYAFYGVWPLLLLLLLALPPRAYA 3'

Figure 2.1. p7 vectors and sequence. The original pHLV-p7 vector used a Trp Δ LE fusion protein (A). The modified pET-31b(+) #272 vector contains a five glycine linker in front of a thrombin cleavage site after KSI. EcoRI and SacI were the restriction enzyme sites chosen for ligation of the p7 sequence from the pHLV plasmid (B). After insertion of the p7 sequence, the leftover restriction enzyme residues were deleted and a ten histidine tag was inserted before the p7 sequence (C). The wild-type J4 p7 63-residue sequence contains a single C27S mutation (D).

Due to the hydrophobic nature and viroporin-like behavior of p7, its expression must be targeted into inclusion bodies in order to obtain moderate yields for NMR studies. Otherwise, buildup of p7 in the cells' cytoplasm is toxic to the cell. The bacterial plasmid pET-31b(+) contains a ketosteroid isomerase

(KSI) fusion protein under a T7 promoter (Figure 2.1). Expression with KSI has been known to drive target proteins into inclusion bodies¹¹. Previously, the expression of genotype 1b J4 p7 in our lab was done using the pHLV plasmid with a TrpLE 1413 polypeptide fusion protein (Figure 2.1)⁵¹. Although this construct also drives p7 into inclusion bodies, the pHLV plasmid has a high copy number in comparison with pET vectors. For co-expression with a low copy number plasmid, the high copy number pHLV vector becomes incompatible. The pET system is commercially available, has a medium copy number and has been used for the expression of membrane proteins, so it was the optimal candidate for re-cloning p7.

2.3 Materials and methods

2.3.1 Cloning of p7 into a pET-31b(+) plasmid

The wild-type J4 p7 strain contains the 63 residue amino acid sequence:

5'ALENLVVLNAASVAGAHGILSFLVFFCAAWYIKGRLAPGAAYAFYGVWPLLLLLLALPP RAYA3'. The pHLV-p7 plasmid made in our lab contained a cysteine 27 to serine mutation in order to aid purification, and this modified sequence was used for studies of wild-type p7. A modified pET-31b(+) vector⁵² containing a glycine linkage followed by a thrombin enzymatic cleavage site between the KSI fusion partner and the target protein and a C-terminal histidine tag was used. This modified pET-31b(+) plasmid (#272) had been previously used in our

lab (Figure 2.1). This plasmid containing the first transmembrane helix of CXCR1 was amplified using competent *E. coli* cells (NovaBlue Singles, Novagen) and purified using a plasmid miniprep kit (Qiagen).

The KOD Hot Start Polymerase kit (EMD Millipore) was used for cloning the KSI-p7 fusion protein. The following primers containing sequences for EcoRI and SacI restriction enzyme sites were obtained from Integrated DNA Technologies (IDT): 5'GGCGCAGAATTCATGGCATTGGAAAACCTGGTTGT3' and 5'GATATAGAGCTCTTATTAAGCGTAAGCACGCGGCGG3'. PCR protocol was followed using the KOD Hot Start protocol to amplify p7. The pET-31b(+)#272 plasmid and p7 PCR insert were cleaved with EcoRI and SacI enzymes (New England BioLabs) and the products were purified by running a 1% agarose gel and cutting out the appropriate DNA fragments. The cut plasmid and insert were ligated, transformed into NovaBlue Singles competent cells (Novagen), and sequenced by Eton Biosciences.

Extra residues leftover from the restriction enzymes and the C-terminal 10-histidine tag were removed (Figure 1C). A 10-histidine tag was inserted right after the thrombin cleavage site to aid in the purification process and increase yield for the wild-type protein. For unnatural amino acid incorporation, the C-terminal histidine tag was kept in order to separate incorporated and truncated protein.

2.3.2 Expression and growth of p7

All p7 constructs were grown in Luria Bertani (LB) media (10 g/L tryptone, 5 g/L yeast, 10 g/L sodium chloride, and 50 mg/L carbenecillin) for unlabeled protein. Minimal media (M9) (1 g/L ammonium sulfate (AMS), 10 g/L glucose, 1x MEM vitamin solution (Thermo Fisher Scientific), trace elements⁵³, 1 mM MgSO₄, 0.1 mM CaCl₂, 0.03 mM thiamine, 1x M9 salts (6 g/L Na₂HPO₄, 3 g/L KH₂PO₄, and 0.5 g/L NaCl), and 50 mg/L carbenecillin) or Bioexpress cell growth media (CIL) were used for isotopically labeled protein⁵⁴. For ¹⁵N-labeled protein, 1 g/L ¹⁵N AMS (¹⁵N₂, 99%, Cambridge Isotope Laboratories (CIL)) was used and for ¹³C-labeled protein, 2 g/L ¹³C glucose (U-¹³C₆, 99%, CIL) was used in the M9 media. Fresh transformations of p7 into C41(DE3) competent cells (Lucigen) on LB plates were used for each growth. A single, freshly transformed colony from an LB-agar plate was inoculated into 10 mL LB media and shaken at 37 °C and 200 rpm for 8 hours. The day growth was used to inoculate 100 mL LB overnight preculture shaken at 37 °C for 16 hours. In the following morning, 1-2% final concentration of the overnight preculture was inoculated into 500 mL of either LB, M9, or Bioexpress media in a 2.5 L baffled flask. The cells were induced at OD₆₀₀~0.5-0.6 with 1 mM final concentration of isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were grown for five hours at 37 °C and 200 rpm post-induction after which they were harvested by centrifuging at 6300 x g for 30 min.

Autoinduction media was also used for the expression of unlabeled and ¹⁵N-labeled p7⁵⁵. The protocol is similar to that of normal LB or M9 media;

however, autoinduction media also contained 4 mL/L glycerol, 0.5 g/L glucose, 2 g/L D-lactose, and 1x M9 salts in addition (or substitution) to the normal LB or M9 media components. For M9 autoinduction media, the amount of AMS used was also doubled. Rather than inducing with IPTG, autoinduction cultures were grown at 37 °C and 200 rpm shaking speed for 16 hours, harvested at 6300 x g, and stored at -80 °C.

In order to express deuterated ^{15}N -p7, *E. coli* C41 (DE3) competent cells (Lucigen) had to be adapted to D_2O ¹⁸. One colony from a freshly transformed LB-agar plate was restreaked on a 40% D_2O M9 agar plate and allowed to grow at 37 °C for 16 hours. From the 40% D_2O M9 plate, one colony was used to streak a 70% D_2O M9 plate. One colony from the 70% D_2O M9 plate was then streaked onto a 90% D_2O M9 plate and allowed to grow at 37 °C for 24 hours. One colony from the 90% D_2O M9 plate was then used to inoculate a day growth in 90% D_2O M9 media, which was grown for 8 hours at 37 °C and 200 rpm. One percent final concentration of the day growth was used to inoculate an overnight preculture containing 99% D_2O M9 that was shaken at 37 °C and 200 rpm for 24 hours. Two percent final concentration of the 99% D_2O M9 preculture was then used to inoculate either 99% D_2O M9 media (containing 1 g/L deuterated glucose (1,2,3,4,5,6,6-D7, 97-98%), 1 g/L ^{15}N ammonium sulfate (CIL), and 1x M9 salts re-dissolved in 99% D_2O) or deuterated ^{15}N -Bioexpress cell growth media (U-D, 98%; U- ^{15}N , 98%; CIL) diluted in 99% D_2O . The expression cultures were shaken at 37 °C and 200 rpm, induced at $\text{OD}_{600}\sim 0.6$ with 1 mM final concentration IPTG (dissolved in 99%

D₂O), and allowed to shake at 37 °C and 200 rpm for an additional 16 hours post induction. Cells were then harvested at 6300 x g and stored at -80 °C.

2.3.3 Purification of p7

Cell pellets from a 1 L growth were re-suspended in 60 mL cell lysis buffer (20 mM Tris-HCl, 500 mM NaCl, 15% glycerol, pH 8.0), lysed for 5 minutes using a Sonic Dismembrator 550 (Fisher Scientific) to obtain inclusion bodies and then centrifuged at 35k x g for 30 min at 4 °C. The supernatant was discarded. The inclusion bodies were re-suspended in 40 mL sodium dodecyl sulfate (SDS) binding buffer (1x phosphate buffered saline (from 20x PBS, Teknova), 1% SDS, 10 mM imidazole, 0.1% tris(2-carboxyethyl)phosphine (TCEP), pH 8.0), sonicated for 5 minutes and then centrifuged at 43.7k x g for 30 min at 15 °C. The supernatant was applied to two columns containing 10 mL column volume (cv) nickelnitrotriacetic acid (Ni-NTA) resin (Ni-NTA Superflow, Qiagen) and the protein was allowed to bind to the column for 1 hour at room temperature by agitation. For wild-type p7, the column was washed with 5 cv SDS binding buffer, 2 cv SDS wash buffer (20 mM HEPES, 1% SDS, 250 mM NaCl, 25 mM imidazole, pH 8.0), and eluted with 3 cv SDS elution buffer (20 mM HEPES, 1% SDS, 50 mM NaCl, 500 mM imidazole, pH 7.3).

After the Ni-purification, SDS was removed from the protein by dialysis against 4 L milli-Q H₂O over 4 days and 6 water changes. For the fourth water change, 0.5 g methyl-beta-cyclodextrin per liter of dialysis buffer was added.

The protein precipitated out of solution once all SDS was removed. After lyophilization, cyanogen bromide cleavage was used to remove any fusion protein and histidine tags. The lyophilized protein powder was dissolved in 70% formic acid (10 mL per 50 mg lyophilized powder) and cyanogen bromide crystals (99.995% trace metals basis, Sigma-Aldrich) were added (3 spatula tip scoops per 10 mL). The reaction was allowed to proceed in the dark at room temperature with gentle rocking for 5 hours. The reaction was neutralized using 1.5 equivalent volume of 1 N NaOH and dialyzed against milli-Q H₂O until the dialysis buffer was neutral in pH. Water was removed by lyophilization. The lyophilized protein powder was further purified using size-exclusion (SEC) fast protein liquid chromatography (FPLC).

For the FPLC purification, 1 mL of 2% SDS was added to 20 mg of lyophilized protein powder. The pH was increased by adding 50 μ L of 1 N NaOH to dissolve the protein powder. The volume was brought up to 5 mL using SDS FPLC buffer (1.2 mM NaH₂PO₄, 18.8 mM Na₂HPO₄, 4 mM SDS, 1 mM ethylenediamine tetracetic acid (EDTA), 1 mM NaN₃). The protein was purified on a Bio-Rad FPLC system using a Sephacryl S-200 HR column (HiPrep 26/60, GE Healthcare Life Sciences), with UV detection, a flow rate was 1.5 mL/min, and a run volume of 450 mL using SDS FPLC buffer. Fractions containing p7 were pooled together, and the SDS was removed via dialysis or using Bio-Beads SM-2 resin (Bio-Rad).

2.4 Results and discussion

Initial attempts to express p7 containing an unnatural amino acid were unsuccessful using the pHLV expression vector that was previously used for the expression of p7 in our lab. At the same time, a modified pET-31b(+) expression vector with a KSI fusion partner had been successfully used for *in vivo* incorporation of an unnatural amino acid into a protein. In order to make a valid comparison of the p7 protein with and without an unnatural amino acid, the expression and purification of p7 with and without UAA need to be consistent. Thus p7 needed to be cloned from the pHLV vector into a pET-31b(+) vector (Figure 2.1).

The p7 gene from the pHLV vector was successfully amplified and inserted between the EcoRI and SacI restriction enzymes sites of the modified pET-31b(+) #272 vector (Figure 2.1). Further modifications were made to the histidine tag and residues leftover from restriction enzyme ligation. Sequencing results (Eton Biosciences) showed an N-terminal KSI followed by a glycine linker, thrombin cleavage site, 10 histidine residues, and finally a methionine residue and p7 (Figure 2.1).

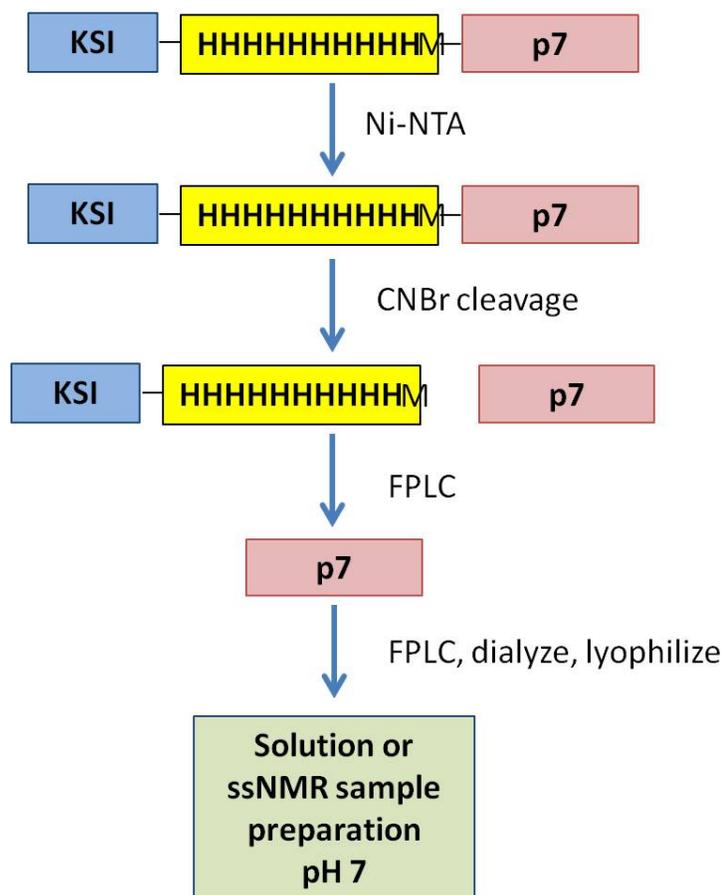


Figure 2.2. Purification scheme of wild-type p7. J4 p7 is expressed as a fusion with KSI. A 10-histidine tag is inserted between KSI and p7. A single methionine was added right before the p7 gene to allow for CNBr cleavage of KSI and p7. The Ni-NTA column yields pure fusion protein. The fusion is cleaved by CNBr. Uncleaved fusion protein, KSI, and p7 are separated by FPLC, and p7 is eluted out as a single peak for NMR studies.

In order to optimize expression of p7, competent cell line, induction time, and growth condition were examined. Various cell lines were tested for p7 expression including BL21(DE3) (NEB), C41(DE3) (Lucigen), C43(DE3) (Lucigen), Rosetta(DE3) (Novagen) competent cells of which the C41(DE3)⁵⁶ cells yielded the best growth and expression. The cells also gave the best yield when induced between OD₆₀₀~0.5-0.6. After post-induction, the cells were

either grown for five hours at 37 °C, 16 hours at 37 °C, or 16 hours at 20 °C. The yield was lower when grown for an extended time at 37 °C or at lower temperatures. Higher temperature promotes the formation of inclusion bodies and since the KSI-p7 fusion is directed into inclusion bodies, it made sense that protein yield was highest when the cells were grown at 37 °C post induction (Figure 2.2). Growth was slower in minimal media, especially if glucose was the limiting carbon source, but decent yields were obtained. In order to balance the cost of ^{13}C -glucose with optimal yield, it was found that using 2 g/L labeled glucose had more than two times the yield than using 1 g/L labeled glucose. The protein yield is 8 mg/L for unlabeled p7 (normal LB media), 6 mg/L for ^{15}N -labeled, and 3 mg/L for ^{13}C - ^{15}N labeled.

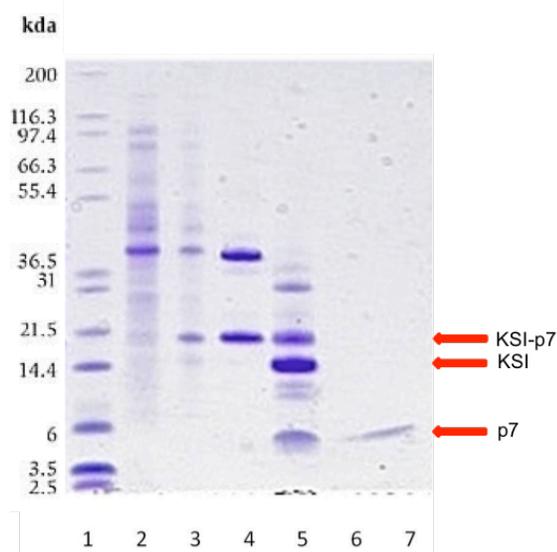


Figure 2.3. SDS-PAGE of the purification of deuterated U- ^{15}N p7. Bioexpress media was used. Lane 1 is the Mark12 ladder (Novex); lane 2 and 3 show pre- and post-induction, respectively; lane 4 shows the KSI-p7 fusion post Ni-NTA column; lane 5 shows the cleaved protein after cleavage with CNBr; and lane 6 and 7 show pure p7 after FPLC purification.

Autoinduction media can also be used (Figure 2.4)⁵⁵. In normal growths, IPTG is used to mimic allolactose and initiate transcription of the lac operon. Autoinduction uses lactose in the media rather than IPTG, which converts to allolactose to initiate transcription. Since glucose is the primary carbon source metabolized first by *E. coli*, balancing the amount of glucose and lactose in the media can optimize the induction time. After all the glucose is consumed, lactose is metabolized and at the same time induces transcription of the recombinant protein. Thus, the bacterial culture does not need to be monitored for optimal cell density for induction, and the rate of induction is slower, allowing the culture to reach higher cell densities especially in the expression of toxic proteins.

Although the cell density is much higher when using autoinduction media (2-3 times that of normal media), the protein yield is not always correlated to the higher cell density. The yield is higher (10 mg/L) for unlabeled p7, but is only comparable for ¹⁵N-labeled p7. Moreover, the ¹⁵N autoinduction minimal media had inconsistent results and it is not feasible to obtain ¹³C-labeling with autoinduction due to the high cost of ¹³C-labeled lactose and glycerol. These factors offset the convenience of using autoinduction media for NMR. Thus, regular minimal media was used to obtain isotopically labeled p7, but autoinduction was used to obtain unlabeled p7. As seen in figure 2.4, The LB autoinduction growth expression was very similar to the normal LB growth. In the case of the M9 growth, induction seems to occur sometime between OD₆₀₀~1 (lane 2) and 2 (lane 4).

Adjusting for the lower concentration loaded into lane 4, there is clearly also some expression in the M9 media, though at lower levels compared to LB autoinduction.

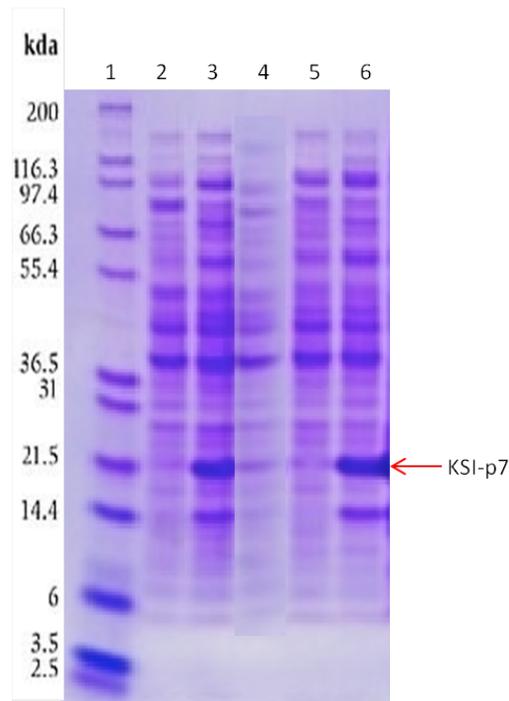


Figure 2.4. SDS-PAGE of p7 autoinduction growths. Lane 1 is the Mark12 Unstained standard (Novex), lane 2 is an M9 autoinduction growth taken at OD₆₀₀~1, lane 3 is an LB autoinduction growth, lane 4 is an M9 autoinduction growth taken at OD₆₀₀~2, lane 5 is a regular LB growth pre-induction, and lane 6 is the regular LB growth post-induction. Lane 4 came from the same gel but has been cropped over.

For large proteins, deuteration allows for better spectral quality^{16, 17, 57-59}.

The slower tumbling of large proteins give rise to faster relaxation times, peak broadening, and low intensity. Deuteration of the protein side chains (amide hydrogens are exchangeable) increases the relaxation time and increases the spectral resolution since it is not picked up by ¹H NMR. In order to make deuterated protein, the bacteria must be able to grow in D₂O. High

abundance of deuterium in cells is toxic. *E. coli*, however, has been known to be able to adapt to grow in heavy water. Once a colony is adapted to D₂O, it remains adapted. For this reason, the transformed cells were restreaked multiple times plates using increasing amounts of D₂O in order to select colonies that could metabolize D₂O better. Several batches of deuterated protein were made this way (Figure 2.2). Although all the colonies chosen were adapted to D₂O media, not all growths were consistent, suggesting that there is still high variation in the mutations that allow *E. coli* to survive in D₂O conditions. Bioexpress media was also favored over minimal media for expression of deuterated p7. The limiting factor for minimal media is the need for deuterated glucose in order to obtain over 90% ²H-labeled protein. Regular glucose can also be used if such high labeling is not required. Although deuterated protein is ideal for NMR since it increases the relaxation time and thereby increasing peak resolution, its use is limited to the high cost of D₂O and deuterated glucose.

Ni-affinity column chromatography allows for the efficient separation of the KSI-p7 fusion from all other proteins and debris found in the solubilized inclusion bodies due to the presence of a histidine tag in the p7 construct (Figure 2.2). The inclusion body solubilizes easily in SDS buffers, which can be easily removed by thorough dialysis. Once all SDS has been removed the protein precipitates as a white powder, which can then be lyophilized. The KSI-p7 sequence contains only one methionine residue, located at the immediate N-terminus of the p7 sequence, so cyanogens bromide allows for

facile cleavage of the two proteins. Once the reaction is quenched with the addition of sodium hydroxide, the proteins precipitate in solution. Dialysis and lyophilization removes the acid and water to leave the protein powder. SDS-PAGE analysis shows some uncleaved protein along with KSI and p7 as the darkest bands (Figure 2.3).

Both KSI and p7 solubilize easily in basic SDS buffers, making FPLC an optimal choice to separate the two proteins. The uncleaved protein is 21.6 kDa, KSI is 14.8 kDa, and p7 is 6.7 kDa. While the uncleaved protein and KSI elute closely together, pure p7 is eluted as a single isolated peak (Figure 2.5). Comparing the sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) with the FPLC chromatogram, cyanogen bromide cleavage typically achieves 90% cleavage. SDS-PAGE of the final FPLC peak shows the pure protein without any other impurities (Figure 2.3).

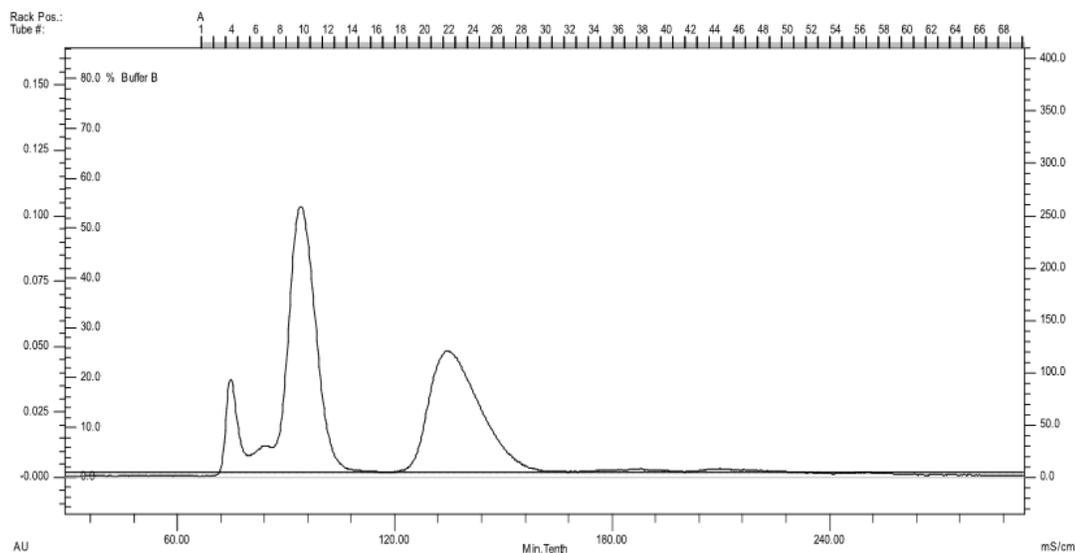


Figure 2.5. FPLC chromatogram of post-CNBr cleavage separation. Cyanogen bromide cleaves the fusion protein after the histidine tag. Approximately 90% of the protein is cleaved, leaving a mixture of uncleaved fusion protein, KSI, and p7. The aggregated protein (fractions 3-4) uncleaved fusion protein (fractions 5-6) and KSI (fractions 8-10) were separated from pure p7 (fractions 19-29) by FPLC using a SDS buffer and Sephacryl S-200 column.

From transformation of p7 into competent cells to the last FPLC run, the time to obtain pure p7 is roughly two weeks, with most of the time spent on dialysis. Cutting short the dialysis time results in residual detergent or buffer that interferes with succeeding steps of the purification. Despite the long purification process, p7 is obtained with high purity (Figure 2.3). The protein also has limited exposure to highly denaturing solutions (e.g. urea, guanidine, organic solvents), and can be refolded properly in various lipids for NMR characterization.

2.5 Conclusion

The J4 p7 sequence was successfully cloned into the modified pET-31b(+) #272 vector from the original pHLV vector. The new plasmid contains a thrombin cleavage site, a histidine tag between the thrombin cleavage site and p7, and a single methionine residue right before the p7 sequence. The fusion protein is cleaved using cyanogens bromide, cutting at the C-terminus of the methionine to give pure wild-type p7. FPLC separated the uncleaved fusion protein, KSI, and p7 proteins.

Expression yield of p7 in this new construct is comparable to that of p7 in pHLV. In addition, the purification process does not use any organic solvents or denaturants such as urea or guanidine hydrochloride. Guanidine hydrochloride and high performance liquid chromatography (HPLC) purification had been previously used with the pHLV construct. The low number of purification steps also helps retain p7 yield since protein loss through incomplete cleavage and columns is limited. Moreover, the pET-31b(+) vector can be used for co-expression with a second plasmid in order to incorporate an unnatural amino acid into p7, the goal for making a new construct.

Chapter 3: Paramagnetic relaxation enhancement of p7 by incorporation of the metal-chelating unnatural amino acid 2-amino-3-(8-hydroxyquinolin-3-yl)propanoic acid (HQA)

3.1 Abstract

The use of paramagnetic constraints in protein NMR is an active area of research because of the benefits of long-range distance measurements (>10 Å). One of the main issues in successful execution is the incorporation of a paramagnetic metal ion into diamagnetic proteins. The most common metal ion tags are relatively long aliphatic chains attached to the side chain of a selected cysteine residue with a chelating group at the end where it can undergo substantial internal motions, decreasing the accuracy of the method. An attractive alternative approach is to incorporate an unnatural amino acid that binds metal ions at a specific site on the protein using the methods of molecular biology. Here we describe the successful incorporation of the unnatural amino acid 2-amino-3-(8-hydroxyquinolin-3-yl)propanoic acid (HQA) into two different membrane proteins by heterologous expression in *E. coli*. Fluorescence and NMR experiments demonstrate complete replacement of the natural amino acid with HQA and stable metal chelation by the mutated proteins. Evidence of site-specific intra- and inter-molecular PREs by NMR in micelle solutions sets the stage for the use of HQA incorporation in solid-state NMR structure determinations of membrane proteins in phospholipid bilayers.

3.2 Introduction

Many analogies can be drawn between the earliest protein NMR studies and present day studies of proteins in biological supramolecular assemblies, such as membrane proteins, amyloid fibrils, chromatin, etc. Both were severely limited by resolution and sensitivity of the spectra, and the ability to interpret the data in terms of the three-dimensional structures of the proteins, which has always been the major goal of protein NMR spectroscopy. Starting with the initial NMR spectrum of a protein in solution⁶⁰, which consisted of four broad, overlapping signals, it was clear that additional steps were needed to extract the underlying spectroscopic and structural information. Some gains resulted from increasing the ¹H resonance frequency and the introduction of signal averaging of the continuous wave signals⁶¹. Nonetheless, the early demonstrations of protein NMR were limited to a few residues whose resonances could be resolved in the spectra of the most favorable globular proteins available in large quantities at the time, for example ribonuclease and lysozyme. The introduction of isotopic labeling of proteins was a major step forward in attaining both improved spectral resolution and resonance assignments¹⁷ but was restricted to selected examples, such as Staphylococcal nuclease because heterologous expression of proteins was not yet feasible. The only available spectral parameters were the isotropic chemical shift frequencies, which varied among amino acid residues because of the differences in environment resulting from protein folding, and resonance line widths that reflected protein

dynamics. The first spectra of membrane proteins were not reported for another ten years⁶²⁻⁶⁵.

In the early 1970s the addition of directly bound paramagnetic species made a substantial difference in the prospects for NMR of proteins. Two similar, parallel paths were introduced. McConnell et al.^{66, 67} exploited the stable paramagnetic center of 1-oxyl-2,2,6,6-tetramethyl-4-piperidinyI (TEMPO) to covalently label lysozyme and then measure the broadening effects on resonances of bound ligands. The goal then, as now, was to make direct distance measurements between the electron spin-label and nuclei at covalently bonded sites on the protein. However, this was not possible because the studies were limited by the experiments being performed at the relatively low field strength corresponding to a ^1H resonance frequency of 100 MHz and other technical issues. Around the same time, Campbell et al.¹² were able to convincingly demonstrate the ability of a lanthanide ion (Gd^{3+}) bound to lysozyme to selectively broaden resonances from residues proximate to the binding site, aided in large part by performing the experiments at the significantly higher ^1H resonance frequency of 270 MHz. These early experiments that exploited the broadening effects of paramagnetic species, whether TEMPO-containing spin labels or lanthanide ions bound to protein ligands, were key predecessors for the current activity in paramagnetic NMR of proteins.

Solution NMR of proteins has advanced significantly over the past 40 years based on improvements in instrumentation, implementation of new

spectroscopic techniques, and the application of sophisticated computational methods to both the processing of experimental data and structure calculations. Nonetheless, limitations remain for structure determination of several important classes of proteins, especially large proteins in complexes in aqueous solution and membrane proteins in various detergent/lipid environments. Both classes of proteins present difficulties for the resolution of individual resonances that result from both the number of overlapping resonances and the broad line widths of the resonances associated with slowly reorienting proteins. Even as these problems have been incrementally addressed, there remains the problem of resolving and assigning a sufficient number of $^1\text{H}/^1\text{H}$ NOEs for structure determination with conventional approaches that measure inter-proton distances of $<5 \text{ \AA}$. Membrane proteins with multiple trans-membrane helices have the additional problem of identifying the correct alignment and relative positioning of the helices, which is difficult to do with only measurements of short-range distances. Further progress in NMR spectroscopy of these classes of proteins would be greatly aided by the ability to measure relatively long-range ($>10 \text{ \AA}$) distances. Such long-range distance constraints have many benefits, including improving the resolution of protein structure determination, defining the overall folding topology, and identifying residues in binding sites. They are especially advantageous in offering a method for positioning multiple trans-membrane helices in membrane proteins^{68, 69}.

This Perspective focuses on membrane proteins, in particular the use of paramagnetic metals attached by tags to otherwise diamagnetic proteins. Our primary research interest is in structure determination of membrane proteins in their native environment of phospholipid bilayers^{3, 11}. However, on the path towards this goal, aspects of membrane protein sample preparation and, to some extent, experimental methods are first worked out with micelle⁷⁰, bicelle²⁵, or nanodisc³⁴ samples that are tractable for solution NMR. This is the situation for the examples described here, as we demonstrate the applicability of the genetically incorporated unnatural amino acid (UAA) 2-amino-3-(8-hydroxyquinolin-3-yl)propanoic acid (HQA) (Figure 3.1)⁷¹ as a metal-binding tag to enable the use of paramagnetic ions to provide long-range paramagnetic relaxation enhancements that serve as intra- and inter-molecular distance measurements in membrane proteins (Figure 3.1).

3.3. Paramagnetic protein NMR

The unpaired electron on a paramagnetic ion has spectroscopic effects that are several orders of magnitude larger than those of the spin $S = 1/2$ nuclei (^1H , ^{13}C , ^{15}N) that are commonly observed in NMR studies of proteins^{20, 68, 72-77}. Briefly, paramagnetic ions induce three effects on the diamagnetic spectra of proteins, only two of which are generally observed and incorporated into the experiments—paramagnetic relaxation enhancements (PREs) and pseudocontact shifts (PCSs). In addition, there are through-bond contact shifts; however, because they only occur in close

proximity to the metal ions, their effects are typically overwhelmed by strong PRE broadening of the resonances. In general, both PREs and PCSs are widely used in paramagnetic NMR studies of proteins. Their occurrence and properties can be controlled by the selection of the metal ions, the ligands, and other factors. The magnetic susceptibility tensor is a key parameter. The two metals used here, Mn^{2+} and Gd^{3+} , have isotropic magnetic susceptibility tensors and therefore induce only PRE effects. Nitroxide spin labels also only induce PRE effects⁷⁸. Many other metals, especially the lanthanides (with the exception of gadolinium) and Co^{2+} have highly anisotropic susceptibility tensors⁷⁹. This gives rise to PCSs, which can result in large changes in chemical shifts at distant sites.

There are other beneficial effects of adding paramagnetic metals to the samples in a controlled manner. They can reduce the longitudinal relaxation times so that data can be acquired much more quickly⁸⁰⁻⁸³, and they can weakly align proteins in solution^{84, 85}, providing an alternative to conventional alignment media for the measurement of residual dipolar couplings (RDCs).

Much of the groundwork for the use of paramagnetic ions in protein NMR was laid by Bertini et al. on metalloproteins, which contain a natural paramagnetic center^{74, 86}. For convenience, in appropriate cases, the metal ion can be exchanged for one with more favorable properties for the studies of interest. As an example, Bertini and Pintacuda have used metal ions in structural studies of superoxide dismutase⁸⁷. They were able to measure many

^{15}N and ^{13}C PREs based on the high-resolution two-dimensional heteronuclear correlation 'fingerprint' spectrum of the protein. These paramagnetic constraints significantly reduced the RMSD of the calculated protein structure.

However, most proteins, especially the membrane proteins of interest, are not metalloproteins. Because of the advantages of introducing a paramagnetic ion, there has been a great deal of activity in the design and implementation of tags that attach a paramagnetic ion to proteins. There are three main approaches to specifically attaching a paramagnetic metal to a protein. One is to attach residues corresponding to a natural or engineered metal binding site to the C- or N- terminus or a loop of the protein. The second is to attach a chemical linker to a reactive site, almost universally a surface cysteine side chain, which can bind a metal ion. The third is to incorporate an unnatural metal-binding amino acid into the sequence at a specific location. The incorporation of an unnatural metal-binding amino acid into membrane proteins is the principal subject of this Perspective.

In the first case, twelve amino acid residues corresponding to an "EF-hand" calcium-binding site were added to the N-terminus of the membrane protein Vpu from HIV-1⁸⁸. This provided a covalently attached lanthanide ion binding site. The protein itself was not altered by the added residues, as evidenced by a lack of perturbation of the chemical shifts, and it was possible to observe long-range paramagnetic effects in the spectra. In addition, the added lanthanide served to weakly align the protein for measurement of residual dipolar couplings. Imperiali and Schwalbe designed seventeen

residue lanthanide binding tags (LBTs) for proteins with improved properties over native calcium binding sites⁸⁹. Subsequently, they inserted LBTs into protein loops⁹⁰, which were shown to give complementary results. In a similar vein, Gaponenko et al.⁹¹ fused zinc fingers to the N- and C- termini of a protein and demonstrated that they could be substituted with paramagnetic cobalt and manganese.

Most paramagnetic NMR studies of proteins have placed the metal ion on the surface of the protein with a covalent tag⁹². Generally this has been done through a linkage to a selected cysteine side chain. In many cases this requires the removal of competitive reactive sites through mutation. It has also become a very active area of research with the development of linkers to two cysteine residues to reduce the local dynamics of the metal ion in order to increase the precision of the measurements. There have also been examples where other protein functional groups are involved in the chelation for the same reason. As a result of this activity, this area has been the subject of a number of reviews, including those by Hass and Ubbink⁹³ and Otting⁷³. The tables and figures of these reviews provide a thorough listing of the wide variety of tags that have been utilized to tag proteins with paramagnetic ions. Figure 3 of the review by Otting⁷³ is particularly helpful in understanding and planning paramagnetic NMR experiments. It shows the relative paramagnetism and asymmetry of the magnetic susceptibility tensors in a way that facilitates direct comparisons.

Using covalent paramagnetic tags, Otting and coworkers have utilized pseudocontact shifts in their studies⁹⁴. They discussed how the assignment problem caused by the large shifts could be addressed. One way is to take advantage of the shifts for two bonded nuclei, e.g., ^1H and ^{15}N , which shift along parallel frequencies from the same metal ion. Fast exchange between diamagnetic and paramagnetic species enables titration of the protein and the resulting paramagnetic ion-induced chemical shifts. Pseudocontact shifts can be used as constraints for calculations of protein structures. With the use of multiple tags, significant improvements in the precision of structures result from the measurement of PCSs.

Ubbink et al. have also focused on the use of PCSs in the structure determination of protein complexes⁷². Their results have been facilitated by the development of metal binding tags. The magnetic properties of lanthanides are relatively insensitive to variations in the coordination. However, since many tags are more flexible than a typical chemical coordination site, it is essential to restrict their dynamics. This has been achieved with bulky tags, including peptide tags as discussed above, two-point attached tags⁹⁵, and tags that interact with protein side chains⁹⁶. With dynamically restricted lanthanide tags, significant PCSs can be observed over very large distances, perhaps beyond 100 Å⁹³.

The groups of Jaroniec and Huber have used a variety of tagged mutants of the protein GB1 as a model polycrystalline protein^{21, 76, 97}. Jaroniec et al. observed substantial PRE effects in well resolved magic angle spinning

spectra by comparing samples with and without free electrons on the nitroxide moiety. Significantly, some of the strong signals in the control sample are missing in the corresponding spectrum of the spin-labeled protein. A key feature of their experimental samples is that the spin-labeled protein was diluted with unlabeled protein to enable the intramolecular effects of interest to be separated from any intermolecular effects of nearby protein molecules. Although clearly beneficial in certain applications, the use of nitroxides as relaxation agents suffers from a major draw-back, namely the presence of large transverse PREs, which leads to severely attenuated signal intensities for numerous residues; this precludes quantitative PRE and distance measurements. They overcame this problem by using tags containing a more rapidly relaxing paramagnetic center, namely Cu^{2+98} , that do not elicit significant paramagnetic shifts due to its effectively isotropic magnetic susceptibility tensor. The longitudinal PREs could be used to determine the global fold of GB1. They were able to calculate the backbone fold of GB1 based solely on PREs using only six mutants.

Paramagnetic NMR has been applied to a number of membrane proteins. Building on the structure of the Anabaena sensory rhodopsin (ASR)⁹⁹ Ladizhansky et al. employed nitroxide spin labels and PRE data to map the oligomerization interface of the receptor¹⁰⁰. Gottstein et al.¹⁰¹ have summarized the alpha helical membrane proteins whose structures have been determined with the aid of PREs. Other notable studies of membrane proteins using paramagnetic NMR include: protein-lipid interactions of outer

membrane protein X (OmpX) and DHPC using various nitroxide spin labels²⁶, improvements of the outer membrane protein A (OmpA) backbone structure via 'parallel spin labeling,'¹⁰² characterization of the Influenza M2 proton channel¹⁰³, topology determination in DPC micelles of SCO3063 and YbdK, two bacterial histidine kinase membrane proteins¹⁰⁴, DsbB¹⁹, structural studies of the *M. tuberculosis* RV1761C protein¹⁰⁵, and the characterization of an 80 residue region of the GPCR Ste2p using PREs in solution NMR¹⁰⁶.

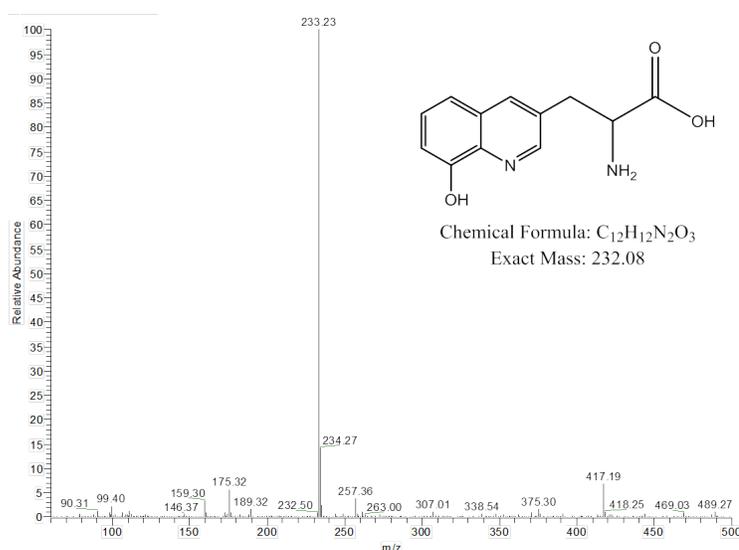


Figure 3.1. Mass spectrometry chromatogram and structure of HQA. The unnatural amino acid HQA contains the 8-hydroxyquinoline moiety. The mass spectrometry spectrum was taken using electrospray ionization, positive mode on a Thermo LCQdeca-MS.

Perhaps the most powerful approach to introducing tags into proteins is unnatural amino acid incorporation^{107, 108}. This very promising method for paramagnetic NMR will be discussed in more depth in the following section along with a sampling of recent studies performed in our laboratory using membrane proteins and unnatural amino acid incorporation of a novel

paramagnetic tag. The best currently available metal-binding unnatural amino acid, HQA, is shown in Fig. 3.1, which we used in our studies.

3.4 Methods

3.4.1 Synthesis of HQA

The synthesis of HQA has been previously reported⁷¹.

Synthesis of 8-methoxy-3-methylquinoline. O-anisidine (5.3 mL, 47 mmol) was added to 70% sulfuric acid (18 mL) at 90 °C. Sodium iodide (140 mg, 0.93 mmol) was added to the solution at 110 °C. Methacrolein (5 mL, 61 mmol) was then added to the reaction mixture at 110 °C dropwise over one hour. Reaction was stirred for one hour at 110-120 °C. After cooling to room temperature, the reaction mixture was transferred into 300 mL 1 M sodium carbonate and extracted into dichloromethane, dried over magnesium sulfate and concentrated in vacuo. The crude product was purified using flash silica column chromatography (1:1:1 dichloromethane (DCM)/hexanes/EtOAc). Mixed fractions were purified again on silica flash column chromatography (100% hexanes to 5% MeOH/30% EtOAc/65% hexanes). 2.7 g (33%) of pure product was obtained as a brown, hard solid.

Synthesis of 3-methylquinolin-8-ol. 8-methoxy-3-methylquinoline (2.7 g, 15.6 mmol) was refluxed in 46-48% hydrobromic acid (27 mL) for 30 hrs. After cooling to room temperature, 3 M sodium hydroxide was added until pH~8. The mixture was diluted with dichloromethane and washed with 1 M sodium

carbonate, dried over magnesium sulfate, and concentrated in vacuo. The crude product was purified using flash silica column chromatography (100% hexanes to 10% EtOAc/90% hexanes). Solvent was removed in vacuo to give 1.75 g (71%) of the product as a pale yellow solid.

Synthesis of 3-methylquinolin-8-yl acetate. 3-methylquinolin-8-ol (1.75 g, 11 mmol) in acetic anhydride (14.5 mL) was stirred at 130 °C for 30 min. The reaction was cooled to room temperature and dried under air. The residue was diluted with ethyl acetate and washed with saturated sodium bicarbonate, dried over magnesium sulfate, and concentrated in vacuo to give 2.1 g (95%) of the product as a tan solid.

Synthesis of diethyl 2-acetamido-2-((8-acetoxyquinolin-3-yl)methyl)malonate. All the reactions in this step were done under argon and in dry flasks. Azobisisobutyronitrile (AIBN) (171 mg, 1.04 mmol) was added to 3-methylquinolin-8-yl acetate (2.1 g, 10.4 mmol) and N-Bromosuccinimide (NBS) (1.86 g, 10.4 mmol) in CCl₄ (75 mL). The reaction mixture was stirred at 70 °C for 4.5 hours. CCl₄ was removed in vacuo. Residue was diluted with dichloromethane, washed with saturated sodium bicarbonate, dried over magnesium sulfate, and concentrated in vacuo to give the crude brominated product (~59% conversion by NMR) as an orange solid. Diethyl acetamidomalonate (6.5 g, 30 mmol) in dry dimethylformamide (DMF) (20 mL) was added dropwise to NaH (1.2 g, 30 mmol, 60% in mineral oil) in dry DMF (20 mL) at 0 °C. The mixture was stirred for 30 min at 0 °C and then 8 mL of the mixture was added to the crude brominated product dissolved in dry DMF (12

mL) dropwise at 0 °C. The reaction was then allowed to stir at room temperature for 45 min. The reaction mixture was diluted with ethyl acetate, washed twice with 10% sodium thiosulfate (2 x 100 mL), dried over magnesium sulfate, and concentrated in vacuo. The crude product was purified by silica flash column chromatography (30% EtOAc/70% hexane s to 75% EtOAc/25% hexanes) to give 2.5 g (58%) of the pure product as a light yellow solid.

Synthesis of 2-amino-3-(8-hydroxyquinolin-3-yl)propanoic acid dihydrochloride. Diethyl 2-acetamido-2-((8-acetoxyquinolin-3-yl) methyl) malonate (1.8 g, 4.3 mmol) in 36-38% HCl (10mL) was refluxed at 110 °C for 5 hours while stirring. Volatiles were removed under air overnight to give the product (0.9 g, 60%) as a yellow solid in the HCl salt form.

3.4.2. *In vivo* incorporation of HQA into p7

The plasmid pEVOL-HQA¹⁰⁹ was obtained from Peter Schultz's lab at The Scripps Research Institute. The pEVOL-HQA plasmid contained the gene for the tRNA and tRNA-synthetase pair specific for HQA and the TAG stop codon. The pET-31b(+) plasmid containing wild-type p7⁵¹ and the C-terminal histidine tags was modified to contain p7 W48TAG. Quikchange Lightning kit and the following primers were used for to make the mutant protein from wild-type p7: 5'GCTTACGCTTTCTACGGTGGTTAGCCGCTGCTGCTGCTGCTGCTG3' and 5'CAGCAGCAGCAGCAGCAGCGGCTAAACACCGTAGAAAGCGTAAGC3'.

The pEVOL-HQA and p7 plasmids were co-transformed into C41(DE3) cells. The growth was similar to that of wild-type p7 for unlabeled W48HQA p7

except for in the incorporation of an unnatural amino acid, 0.02% final concentration L-arabinose and 1 mM final concentration UAA at $OD_{600} \sim 0.4$ were added. These cultures were then induced at $OD_{600} \sim 0.6$ with an additional 0.02% L-arabinose and 1 mM final concentration IPTG. For U- ^{15}N labeled protein, W48HQA p7 was prepared using a media exchange method^{54, 57}: 500 mL Luria-Bertani (LB) media with chloramphenicol/carbenicillin was inoculated with 10 % overnight LB starter culture and grown in a shaker/incubator at 37 °C. At $OD_{600} \sim 0.4$ the cells were induced with a final concentration of 0.02 % L-arabinose and grown for 2 hrs at 37 °C. They were then spun down by centrifugation at 1,350 x g. The cell pellets were gently re-suspended in 500 mL minimal media with ^{15}N -labeled ammonium sulfate as the sole nitrogen source, induced with 0.02 % final concentration of L-arabinose and grown at 37 °C for one more hour following which 1 mM HQA and 100 μ M IPTG were added. Cells were grown for 4 h, after which they were harvested via centrifugation at 6,200 x g and stored at -80 °C prior to further purification.

Cell pellets were lysed using 30mL lysis buffer (20 mM Tris-HCl, 500 mM NaCl, 15% glycerol, pH 8) for 500mL of growth. Lysozyme and DNase I were added to the resuspended pellet. The resuspension was lysed for 5 minutes using a Sonic Dismembrator 550 (Fisher Scientific) and centrifuged for 30 min at 35k x g and 4 °C. The supernatant was discarded and the inclusion body was resuspended using 25mL SDS binding buffer (1x PBS, 1% SDS, 10 mM imidazole, 0.1% TCEP, pH 8). The resuspension was sonicated for 5 minutes and

then centrifuged for 30 min at 43.7k x g and 15°C. The supernatant was applied to a 10 mL bed volume of a Ni-NTA (Ni-NTA Superflow, Qiagen) affinity column. The protein was allowed to bind to the column over 1 hour. The column was washed with 5 cv SDS binding buffer and then the column buffer was exchanged by washing with 20 cv thrombin cleavage buffer (20 mM Tris-HCl, 300 mM NaCl, 0.1% hexadecylphosphocholine (HPC), pH 8). The fusion protein was cleaved using 4000U/L thrombin dissolved in thrombin cleavage buffer. Cleavage was allowed to occur at room temperature for 16 hours by gentle agitation of the column. The column was then washed with 10 cv SDS wash buffer (20 mM HEPES, 1% SDS, 250 mM NaCl, 25 mM imidazole, pH 8) and protein was eluted with 3 cv SDS elution buffer (20 mM HEPES, 1% SDS, 50 mM NaCl, 500 mM imidazole, pH 7.3).

After the Ni-NTA column, SDS was removed from the protein by dialysis against milli-Q H₂O over 4 days and 6 water changes. For the third and fourth water changes, 0.5 g methyl-beta-cyclodextrin was added per liter of the dialysis buffer. The protein precipitated out of solution once all SDS was removed and was lyophilized. Cyanogen bromide cleavage was used to remove any fusion protein and histidine tags. The lyophilized protein powder was dissolved in 70% formic acid (10 mL per 50 mg lyophilized powder) and cyanogens bromide crystals were added (3 spatula tip scoops per 10 mL). The reaction was allowed to proceed in the dark at room temperature with gentle rocking for 5 hours. The reaction was neutralized using 1.5 equivalent volume of 1 N NaOH and dialyzed against milli-Q H₂O until the dialysis buffer was

neutral in pH. Water was removed by lyophilization. The lyophilized protein powder was used for further purification by HPLC.

For HPLC purification, 20 mg of lyophilized protein powder was dissolved using 20 drops of glacial acetic acid and 4 mL 49.9% ACN/ 50% H₂O/ 0.1% TFA and sonicating in a Branson 1510 sonicator for 30 minutes. The protein was purified on a Waters 600 HPLC using a diphenyl column (Pursuit XRs, 100 x 21.2 mm, 5 μ m, Agilent Technologies) using a flow rate of 10 mL/min. The gradient started from 90% A (99.9% H₂O, 0.1% TFA) and 10% B (90% ACN, 9.9% H₂O, 0.1% TFA) to 10% A and 90% B over 45 minutes. The fractions containing p7 were dried under nitrogen gas and then lyophilized to give the pure protein powder.

3.4.3 Preparation of p7 for solution NMR and fluorescence studies

NMR samples were prepared by dissolving the lyophilized protein in a solution containing 150 mM DHPC, 20 mM HEPES, 10% D₂O, pH 7.0 at a final protein concentration of 50-100 μ M. For the PRE experiments, aliquots of stock solutions of 10 or 100 mM MnCl₂ were added to the protein samples at final Mn²⁺ concentrations of 0.5 mM. The NMR experiments were performed on a Bruker Avance 600 MHz spectrometer equipped with a 5-mm triple-resonance cryoprobe with z-axis gradient. One-dimensional ¹⁵N-edited ¹H NMR and ¹H-¹⁵N HSQC¹¹⁰ NMR spectra were obtained at 50 °C.

Fluorescence experiments were performed on a FluoroMax-4 spectrofluorometer (Horiba Scientific, New Jersey, NY) at room temperature

with excitation at 400 nm. Fluorescence of the wild-type and HQA-incorporated proteins at 20-50 μM (200 μL) was measured in 300 mM SDS or 150 mM DPC detergent micelles, 20 mM HEPES, pH 7.0, and different concentrations of ZnSO_4 (0, 50, 100, 150, 200, and 300 μM). Fluorescence quenching experiments were performed on Zn^{2+} -bound p7 W48HQA by titrating EDTA to final concentrations of 0, 200, 400, 600, 900, and 1800 μM .

3.5 Results and discussion

3.5.1 Metal-chelating unnatural amino acid

Although the use of fusions with cysteines and metal-binding peptides for site-specific paramagnetic labeling of proteins has helped protein structure determination, analysis of protein–ligand binding, and descriptions of molecular dynamics, their use is limited by their bulk, local dynamics, and potential sample heterogeneity. Large conformational spaces for the paramagnetic center lead to observed PRE measurements whose properties may be influenced by the flexibility of the probe itself. Rigid paramagnetic probes have been introduced into proteins to counter this problem^{111, 112}, but the location of anchoring cysteine residues in the protein limit their placement and their often-bulky structure can also affect native protein structure. The incorporation of unnatural amino acids (UAAs) addresses both the issues of bulk and local dynamics and allows for homogeneous protein samples. They can be placed anywhere in the protein without being limited by disulfide

bonds, cysteine residues, and attachments to the protein termini⁷³. In addition, a paramagnetic metal ion can be introduced anywhere in the protein with minimal protein structure perturbation¹¹³ due to a single amino acid substitution with a UAA with high metal affinity in the protein sequence.

Site-specific genetic incorporation of HQA involves an orthogonal aminoacyl-tRNA/tRNA synthetase pair specific to the HQA that reads the amber TAG codon⁷¹. Under normal circumstances, *E. coli* would recognize TAG as a stop codon. With the addition of the orthogonal pair, however, HQA is added to the growing protein chain. The orthogonal tRNA synthetase is not recognized by endogenous tRNA/amino acid and vice versa¹⁰⁷. Of the more than 100 unnatural amino acids incorporated successfully, a small fraction is useful for NMR. These include isotopically labeled *p*-methoxy-phenylalanine and its fluorinated analogs^{114, 115}, photocaged unnatural amino acids for site-specific labeling¹¹³, metal-chelating unnatural amino acids^{71, 111, 116}, and an amino acid that ligates with a lanthanide tag¹¹⁵. Incorporation of isotopically labeled amino acids at specific residues have helped identify ligand binding sites and conformational changes of large proteins. Aside from these, to date, only six unnatural amino acids that have been successfully incorporated into proteins can be used as paramagnetic probes. Of these, four are metal chelating and have been used for NMR studies¹¹³. It is also possible to use unnatural amino acids¹¹⁷ to incorporate electron spin-labels in specific sites^{118, 119}.

The use of lanthanides for PRE studies are of particular interest because of their varying magnetic properties; different lanthanides can be incorporated into the same metal-chelation site. Previous lanthanide tags could only be non-covalently bound to the N- or C- termini of a protein, resulting in flexibility of the tag and poor measurements. While the lanthanide-p-azido-l-phenylalanine (AzF) fusion allowed for placement of the tag independent of the location of other residues in the protein, there are limitations that hinder its widespread application¹¹¹. The lanthanide tag is also limited by the need for known structural data for tether optimization. Using metal-chelating unnatural amino acids, on the other hand, provides a straightforward way to introduce a paramagnetic ion for PRE measurements¹¹⁶. For example, the UAA bipyridylalanine (BpyAla) was successfully incorporated into the West Nile virus NS2B-NS3 protease and bound to cobalt (II) to obtain PCS measurements¹¹⁶. Of the three metal-chelating unnatural amino acids, only HQA chelates to lanthanides^{71, 113}, making it the best candidate for PRE studies. Moreover, HQA is the smallest of the three UAA, making it a favorable replacement for the similarly sized aromatic canonical amino acids, in particular tryptophan. It is also fluorescently active, allowing for easy detection of metal-bound protein.

Successful efforts to obtain NMR spectra of HQA-incorporated protein have not been reported previously¹¹³. Here we show by solution NMR the successful incorporation of HQA into the second transmembrane helix of viroporin p7 from the hepatitis C virus. We were also able to successfully obtain

NMR data showing specific binding of a paramagnetic ion to the HQA-incorporated proteins, offering powerful insight to protein–ligand binding and protein structure.

3.5.2 Genetic incorporation of HQA into membrane proteins

Genetic incorporation of the unnatural amino acid 2-amino-3-(8-hydroxyquinolin-3-yl)propanoic acid (HQA) was done via amber codon suppression as demonstrated by Schultz et al⁷¹. For the p7 construct, the tryptophan at residue 48 in the sequence was chosen as the site of mutation due to its location in the second transmembrane helix, with the primary goal to obtain intramolecular PRE data.

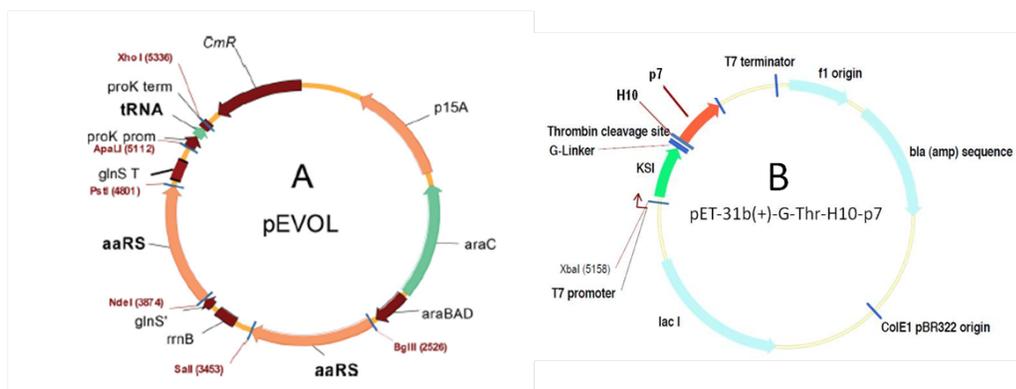


Figure 3.2. Vector maps for unnatural amino acid incorporation into proteins in *E. coli*. A. The pEVOL vector optimized for efficient expression of orthogonal pairs of tRNA/aminoacyl-tRNA synthetase. B. Modified pET-31b(+) vector for HQA-incorporated membrane protein expression as a KSI-fusion protein containing a thrombin cleavage site between KSI and the target membrane protein.

The pEVOL vector encoding the orthogonal tRNA/aminoacyl-tRNA synthetase pair for HQA¹⁰⁹ and the modified pET31b(+) vector encoding our target mutant (Fig. 3.2) were co-expressed in Bioexpress media and we were

able to obtain high yields of mutant, HQA incorporated, membrane protein for the p7 construct. Growth and incorporation protocols for both the orthogonal pair and target genes were optimized and included testing various growth temperatures, induction times and concentrations, as well as optimizing the growth media itself (data not shown). Following growth and induction both the truncated and incorporated forms of the p7 protein were produced, as can be seen in Fig. 3.3, lane 1. Due to the presence of a C-terminal histidine tag, which is only present in the full-length mutant and not in the truncated form (Fig. 3, top), the two could be successfully separated via Ni-immobilized metal affinity chromatography (IMAC). Further purification by HPLC resulted in high yields of pure protein (Fig. 3.3).

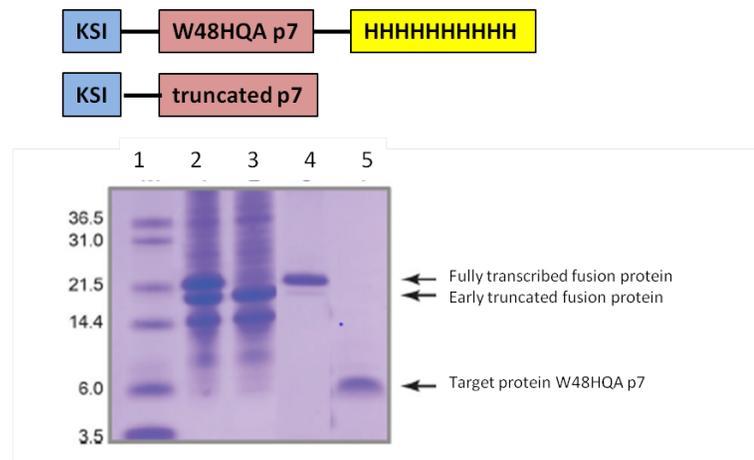


Figure 3.3. SDS-PAGE of the purification of W48HQA p7. Schematic drawings of the proteins prepared by using two expression vectors shown in Fig. 3.2 (top) and a SDS-PAGE showing the purification of a HQA-incorporated membrane protein p7 (bottom): lane 1, Mark12 ladder; lane 2 inclusion bodies solubilized in detergents; lane 3, flow through of the Ni-affinity chromatography; lane 4, elution from the Ni-affinity chromatography; lane 5, HPLC pure W48HQA p7 after CNBr cleavage. The His-tag attached to the C-terminus of p7 facilitates the separation of the fully transcribed proteins from the early-terminated proteins using Ni-affinity chromatography.

In the ^{15}N HSQC spectra obtained at 600 MHz W48HQA p7 in DHPC micelles at pH 7 (Fig. 3.4) it was observed that the resonances from the tryptophan indole nitrogen sites, which were mutated at positions 10 and 48, respectively, were noticeably absent in the spectra of the mutant proteins, indicating that the unnatural amino acid was successfully incorporated into the proteins and at the correct positions. Furthermore, the overlap between the spectra of the wild-type and mutant proteins is very significant for both mutants (Fig. 3.4), indicating that incorporation of the unnatural amino acid did not cause significant chemical shift perturbations in the spectra, and that the native structures and conformations are retained.

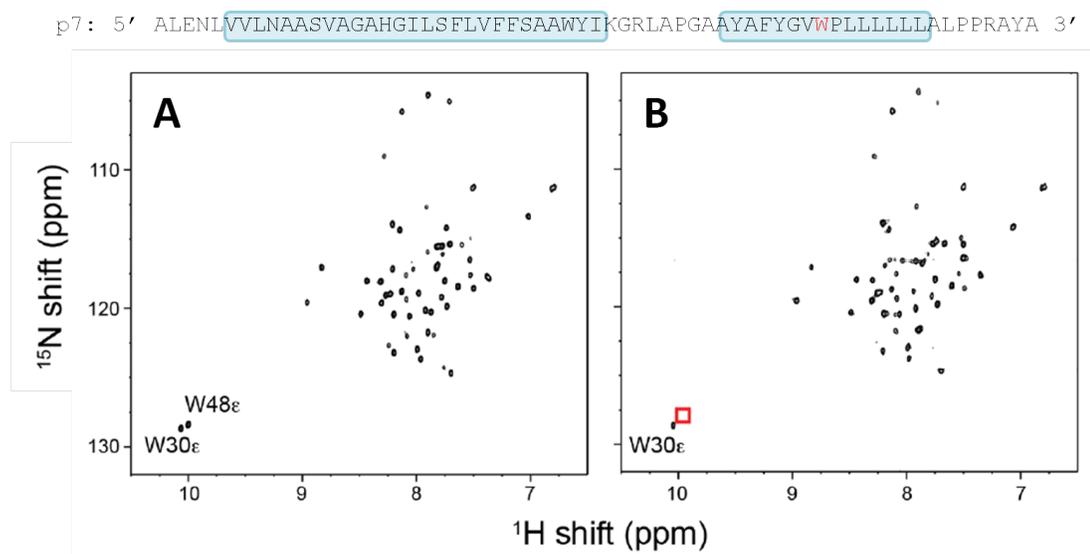


Figure 3.4. Comparison of ^1H - ^{15}N HSQC spectra of uniformly ^{15}N -labeled p7 in DHPC micelles. A. Wild-type p7. B. W48HQA p7. Absence of Trp indole $\text{NH}\epsilon$ signal shown in a red box indicates complete incorporation of unnatural amino acid HQA into residue Trp 48 in p7. The mutation site is colored in red in the amino acid sequences and schematic drawing of p7. Note that the amide chemical shifts for wild-type and HQA-incorporated p7 are identical except in the vicinity of the mutated site.

3.5.3 Fluorescence induced by HQA-metal chelate

8-hydroxyquinoline forms chelate compounds with many metal ions including lanthanides, and Cd^{2+} , Mg^{2+} , and Zn^{2+} form a strong fluorescent complex with 8-hydroxyquinoline¹²⁰. In order to investigate site-specific fluorescence from binding of metal ions to HQA-incorporated membrane proteins in detergent micelles, W48HQA p7 protein was titrated with Zn^{2+} ions and fluorescence was measured (Fig. 3.5). No fluorescence of the HQA-incorporated protein was observed in the absence of Zn^{2+} , but the fluorescence increased with increasing concentrations of Zn^{2+} , and was saturated at about 2:1 (Zn^{2+} :HQA) molar ratio for p7. As a control, the fluorescence of the wild-type protein was examined; however they showed no fluorescence in the presence of 200 μM Zn^{2+} (data not shown).

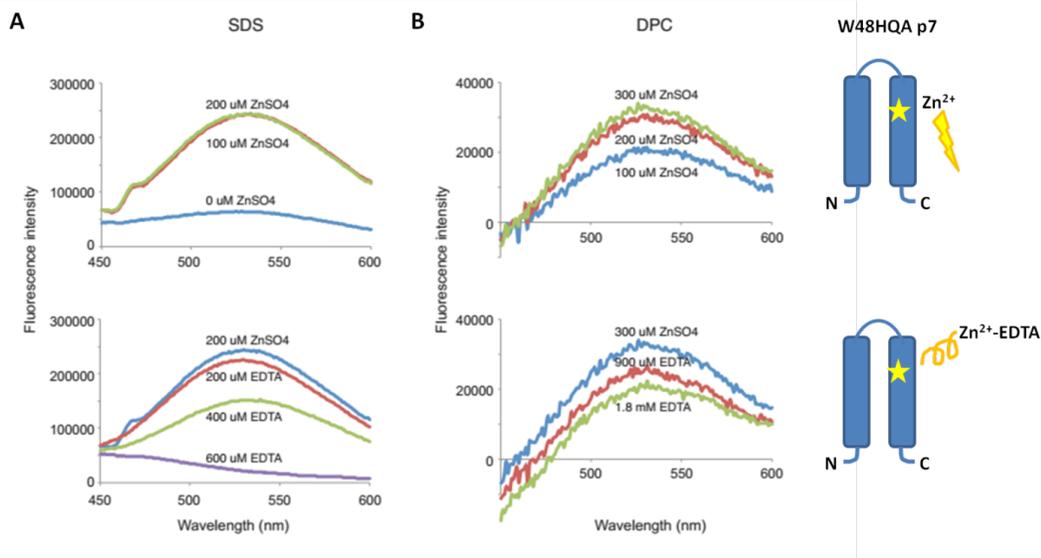


Figure 3.5. Fluorescence spectra of HQA-incorporated membrane proteins in detergent micelles. A. W48HQA p7 complexed with varying concentrations of ZnSO₄ in SDS micelles (top) and W48HQA p7 in SDS complexed with 200 μM ZnSO₄ and with varying concentrations of EDTA (bottom). B. W48HQA p7 complexed with varying concentrations of ZnSO₄ in DPC micelles (top) and W48HQA p7 in DPC complexed with 300 μM ZnSO₄ and with varying concentrations of EDTA (bottom). Protein concentration was approximately 50 μM.

Complete quenching of fluorescence of Zn²⁺-HQA chelates was observed at a 3:1 (EDTA:HQA-Zn²⁺) molar ratio for W48HQA p7 in SDS micelles (Fig. 3.5A), suggesting that Zn²⁺ forms stable chelates with HQA in the transmembrane region of p7 in micelles. However, the type of detergent may affect binding saturation. In DPC micelles, W48HQA p7 requires a higher concentration of ZnSO₄ for fluorescence and a higher concentration of EDTA to quench fluorescence although the overall fluorescence is much lower than that of the SDS data. This suggests that the choice of detergent has a great effect in the folding of the membrane protein. It appears that p7 is folded in

DPC but denatured in SDS. Regardless, the fluorescence data shows that HQA is incorporated into p7 and it is available for metal chelation. These results prove that genetic incorporation of the unnatural amino acid HQA into membrane proteins can serve as a powerful site-specific biophysical probe for studies of the structures and dynamics of membrane proteins in membrane environments.

3.5.4 Intramolecular PREs by HQA-metal chelate

In Fig. 3.6, the comparison of HSQC spectra of p7 constructs demonstrates the different PRE effects on the wild-type and HQA-incorporated p7 samples. All of the p7 resonances, except for the first three N-terminal residues, were observed and assigned in DHPC micelles at pH 7 (Fig. 3.6a). When paramagnetic Mn^{2+} ions were titrated to the solution containing wild-type p7, several resonances (residues 4, 17, 18, and 63) were significantly broadened, indicating these residues are located in the solvent accessible regions of the molecular surface of p7 in detergent micelles (Fig. 3.6b).

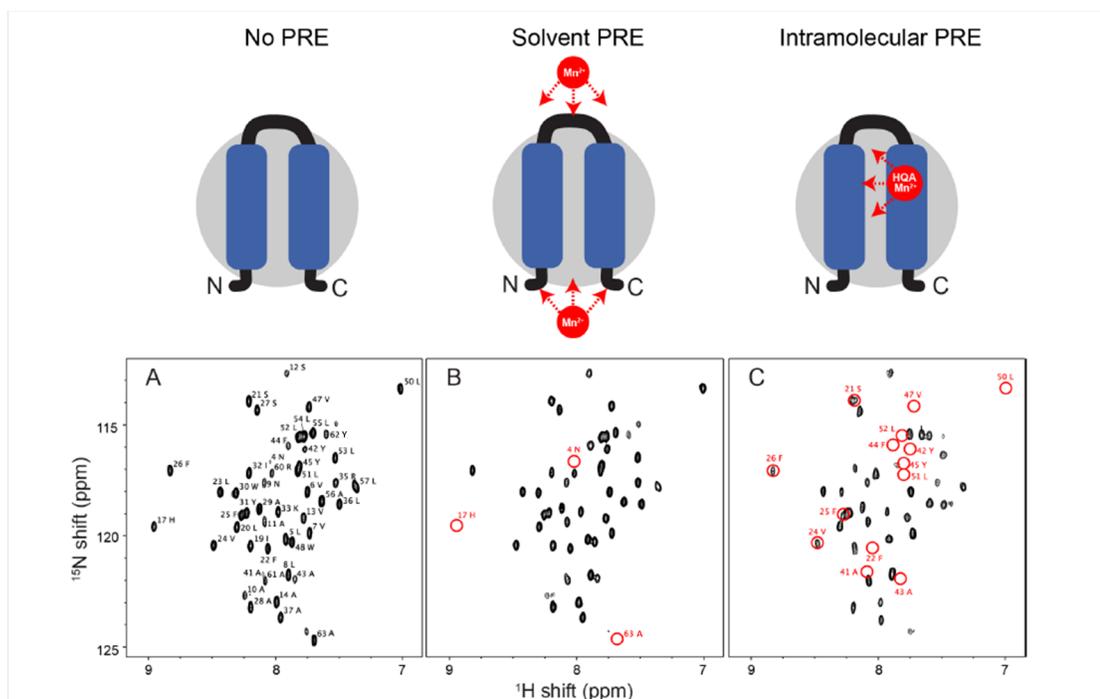


Figure 3.6. Expanded region of the ^1H - ^{15}N HSQC spectra of uniformly ^{15}N -labeled p7 in DHPC micelles. A. Wild-type p7 alone. B. Wild-type p7 in the presence of 0.5 mM MnCl_2 . C. W48HQA p7 in the presence of 0.5 mM MnCl_2 . The resonances broadened by solvent PREs or intramolecular PREs are indicated by red circles. Note that glycine residues 15, 18, 34, and 46 do not lie in the expanded region.

When Mn^{2+} ions were titrated in the W48HQA p7 sample at an approximately 5:1 (Mn^{2+} :W48HQA p7) molar ratio, many more signals were broadened or disappeared compared to the wild-type p7 (Fig. 3.6c). This result suggests that Mn^{2+} can specifically bind to the HQA located in the second transmembrane helix of p7 in detergent micelles, and therefore yields intramolecular PREs. Many of the missing resonances (residues 41–47 and residues 50–52) are near the mutated position of the residue 48 as expected. However, it is noteworthy that the residues 21, 22, 24–26 were also significantly broadened or disappeared, indicating that these residues, which are located

in the first transmembrane helix, are in close proximity to residue 48 which is located in the second transmembrane helix of p7. These intramolecular PREs are very helpful in structure determination of membrane proteins, since they provide long-range distance restraints that are very challenging to obtain by conventional NOE experiments.

Although successful efforts to obtain NMR spectra of HQA-incorporated protein have not been reported previously due to metal mediated protein oligomerization¹¹³, we do not observe any evidence of oligomerization or aggregation of HQA-incorporated membrane proteins in detergent micelles. Optimization of experimental conditions was straightforward.

3.6 Conclusion

The research described here is a prelude to the use of paramagnetic tags on membrane proteins in phospholipid bilayers. This involves the use of solid-state NMR because the membrane proteins are immobilized by their interactions with the phospholipids in the bilayer environment. Thus, significant background comes from prior solid-state NMR studies of paramagnetically tagged proteins.

The *in vivo* incorporation of unnatural amino acids into proteins is a well-established technique requiring an orthogonal tRNA/aminoacyl-tRNA synthetase pair specific for the unnatural amino acid that is incorporated at a position encoded by a TAG amber codon. Recently developed metal-chelating unnatural amino acid 2-amino-3-(8-hydroxyquinolin-3-yl)propanoic

acid (HQA) forms highly stable complexes with various transition metal ions and lanthanides, serving as an excellent probe for paramagnetic relaxation enhancement NMR experiments. Optimization of the expression of orthogonal aminoacyl-tRNA synthetases/suppressor tRNA pairs from the pEVOL vector and target proteins from the KSI-fusion expression system provided high yields of HQA-incorporated proteins, essentially equivalent to those for wild-type proteins. NMR spectral comparisons of the wild-type and HQA-incorporated mutants demonstrate complete incorporation of HQA into the membrane proteins. Zn^{2+} -HQA induced fluorescence confirms a stable metal chelation.

Chapter 3 is a partial reprint of the material as it appears in the Journal of Biomolecular NMR, "Paramagnetic relaxation enhancement of membrane proteins by incorporation of the metal-chelating unnatural amino acid 2-amino-3-(8-hydroxyquinolin-3-yl)propanoic acid (HQA)" by S. H. Park, V. S. Wang, J. Radoicic, A. A. DeAngelis, S. Berkamp, and S. J. Opella, 2015. The thesis author was the secondary author of the paper.

Chapter 4: NMR studies of p7 in various lipid membranes

4.1 Abstract

Viroporin p7 from HCV, recombinantly expressed in *E. coli*, was successfully reconstituted into various membrane mimetics. For solution-state NMR, p7 in DHPC and 1,2-di-O-hexyl-sn-glycero-3-phosphocholine (6-O-PC) micelles as well as in $q=0.1$ DHPC/DMPC (1,2-dimyristoyl-sn-glycero-3-phosphocholine) bicelles (where 'q' is the molar ratio of long-chain lipid to detergent) showed good resolution and that the monomeric protein was well-folded. Reconstituted p7 in nanodiscs were examined using both solution- and solid-state NMR. The size of the nanodisc appears to be a challenge to both areas in the study of integral membrane proteins as the isotropic tumbling is too slow for highly resolved spectra using solution NMR, but it is also not large enough to sediment fully and therefore contains too much motion for solid-state NMR (ssNMR). Finally, p7 was also reconstituted into DMPC proteoliposomes for ssNMR studies. In proteoliposomes, p7 appears to undergo rotational diffusion initially but quickly loses rotational diffusion after 24 hours.

4.2 Introduction

Understanding protein function and developing new pharmaceutical drugs rely heavily on first elucidating the structure of a protein. For membrane proteins, the puzzle is complicated due to the innate need for the protein to

be correctly folded in its native environment^{9, 15, 23, 34, 121}. In other words, the lipid environment affects the folding of a membrane protein¹²². For solution NMR, only micelles²³, small isotropic bicelles^{24, 25}, and nanodiscs³² are small enough to achieve isotropic tumbling and resolved spectra. Unfortunately, cell membranes are not like any of these small lipid mimetics; rather, they are large, asymmetric, liquid crystalline phospholipid bilayers. Thus solid-state NMR spectra of membrane proteins reconstituted in large bicelles²⁹ and nanodiscs, macrodiscs^{29, 34}, and proteoliposomes⁹ give more realistic structures of the protein in its native environment^{3, 4}.

Currently, there are no crystal structures of a full-length viroporin, so NMR spectroscopy has been the primary method used to obtain structural information for this class of proteins^{123, 124}. In the case of p7 from HCV, only solution NMR studies have been published^{35, 36, 43, 45, 46}, and the results vary greatly from one another due to different sample environment and different p7 genotypes. So far, there are two structures solved in organic solvents (TFE/water⁴³ and methanol⁴⁷) and two in detergent micelles (DHPC³⁵ and DPC⁴⁶) as the sample environment. The structures differ in the length of the transmembrane helices, the location of kinks, and the distance between the two helices. In addition, p7 has seven main genotypes with over a hundred subtypes, and the amino acid sequences between two genotypes can differ as much as 50%³⁹. Interestingly, the two structures published of p7 in two different detergent micelles show high discrepancy in the distance between the two transmembrane helices. This begs the question: although two different

genotypes and two different detergent micelles were used, can the same protein from two different genotypes form two different monomeric structures but have the same viroporin function or does the detergent choice affect the structure regardless of genotypic differences? In this study I examine p7 reconstituted into detergent micelles, small isotropic bicelles²³, nanodiscs³², and phospholipid bilayers³.

4.3 Materials and methods

4.3.1 Reconstitution of p7 into DHPC and 6-O-PC micelles

NMR samples were prepared by dissolving lyophilized uniformly ¹⁵N labeled p7 purified using HPLC in a solution containing 150 mM DHPC (Avanti Polar Lipids) or 6-O-PC (D44, 98%, CIL) micelles, 20mM HEPES, 10% D₂O, pH 7.0 at a final protein concentration of 50-100 μM. The NMR experiments were performed on a Bruker Avance 600 MHz spectrometer equipped with a 5-mm triple-resonance cryoprobe with z-axis gradient. One-dimensional ¹⁵N-edited ¹H NMR and ¹H-¹⁵N heteronuclear single quantum coherence (HSQC) spectra were obtained at 50 °C.

4.3.2 Reconstitution of p7 into q=0.1 isotropic bicelles

NMR samples were prepared by dissolving lyophilized ¹⁵N p7 purified using HPLC in a solution containing 150 mM final detergent concentration with 1 DMPC (Anatrace) to 10 DHPC molar ratio, 20mM HEPES, 10% D₂O, pH 7.0 at

a final protein concentration of 50-100 μM . The NMR experiments were performed on a Bruker Avance 600 MHz spectrometer equipped with a 5-mm triple-resonance cryoprobe with z-axis gradient. One-dimensional ^{15}N -edited ^1H NMR and ^1H - ^{15}N HSQC spectra were obtained at 50 $^\circ\text{C}$.

4.3.3 Reconstitution of p7 into nanodiscs

4.3.3.1 Expression and purification of tobacco etch virus (TEV)

The TEV protease expression vector pRK793 was transformed into Rosetta(DE3) competent cells¹²⁵. One colony was selected and grown in 10 mL LB containing carbenecillin (50 mg/L) and chloramphenicol (34 mg/L) for 8 hours at 37 $^\circ\text{C}$. One percent of the day growth was used to inoculate 100 mL of an overnight LB (plus 2 g/L glucose) preculture containing the same amount of antibiotics. Two and a half percent of the overnight preculture was used to inoculate 500 mL of LB containing carbenecillin (50 mg/L) and chloramphenicol (34 mg/L), glucose (2 g/L), and M9 salts (1.5 g/L Na_2HPO_4 , 0.75 g/L KH_2PO_4 , and 125 mg/L NaCl). The culture was grown at 37 $^\circ\text{C}$ and 245 rpm shaking speed and was induced with 1 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) when OD_{600} reached 1. After induction, the cells were grown at 30 $^\circ\text{C}$ for 5 hours. The cells were harvested and the pellet was stored at -80 $^\circ\text{C}$.

The pellet from 500 mL growth was resuspended in 27 mL of iced lysis buffer (50 mM sodium phosphate dibasic, 200 mM NaCl, 10% glycerol, 25 mM

imidazole, pH 8). Cells were sonicated on ice for 60s using a Sonic Dismembrator 550 (Fisher Scientific). Five percent poly(ethyleneimine) was added to a final concentration of 0.1% and the sonicated cells were centrifuged at 35k x g and 4 °C for 30 min. One half liter of growth was purified over one 10 mL bed volume Ni-NTA beads (Ni-NTA Superflow, Qiagen). The column was equilibrated with iced lysis buffer. The supernatant was applied to the column and allowed to flow through the column slowly twice. The column was then washed with 7 column volumes (cv) of chilled lysis buffer. The protein was washed stepwise using 10 mL each of chilled lysis buffer containing 70 mM, 115 mM, 160 mM, and 205 mM imidazole. The protein was eluted with 20 mL of chilled elution buffer (50 mM sodium phosphate dibasic, 200 mM NaCl, 10% glycerol, 250 mM imidazole, pH 8). Fractions containing pure TEV were concentrated to 2 mg/mL and stored at -20 °C with 5 mM dithiothreitol (DTT) and 1 mM EDTA added.

4.3.3.2 Expression and purification of membrane scaffolding proteins

The membrane scaffolding protein (MSP) MSP1D1ΔH5 and MSP1D1E3 genes in the expression vector pET28-His-MSP1 were obtained from the Francesca Marassi lab^{31, 32}. A fresh transformation of the plasmid containing MSP1D1ΔH5 into BL21(DE3) cells were used for the expression of MSP1D1ΔH5, and a glycerol stock was used for the expression of MSP1D1E3. A 100 mL preculture containing LB and kanamycin (30 mg/L) was grown for 16 hours at 37 °C. Two percent final concentration of the preculture was used to

inoculate 1 L of TB containing kanamycin (30 mg/L). The cells were induced at $OD_{600} \sim 1$ with 1 mM IPTG (final concentration), grown for 4 hours at 37 °C, and harvested. The cells were resuspended in MSP lysis buffer (20 mM PBS, pH 7.4), 2 μ L DNase I, 1% Triton X-100 (TX-100), and 1 mM phenylmethanesulfonylfluoride (PMSF). The cells were lysed by sonicating for 5 min using a Sonic Dismembrator 550 (Fisher Scientific) and centrifuged at 30K x g. One liter of growth was purified over 10 cv of Ni-NTA. The supernatant was applied twice to a column equilibrated with MSP lysis buffer. The column was washed with 20 cv MSP wash buffer 1 (40 mM Tris-HCl, 300 mM NaCl, 1% TX-100, pH 8), 10 cv MSP wash buffer 2 (40 mM Tris-HCl, 300 mM NaCl, 20 mM imidazole, 50 mM sodium cholate, pH 8), and 10 cv MSP wash buffer 3 (40 mM Tris-HCl, 300 mM NaCl, pH 8). The protein was eluted stepwise using 1 cv of MSP wash buffer 3 with increasing amounts of imidazole (50 mM, 150 mM, and 250 mM) and then with 2 cv of MSP elution buffer (40 mM Tris-HCl, 400 mM NaCl, 400 mM imidazole, pH 8). The protein was dialyzed 24 hr against MSP dialysis buffer (20 mM Tris-HCl, 0.5 mM EDTA, 100 mM NaCl, pH 7.4). The MSP protein was cleaved with TEV protease (2 mg TEV per 100 mg MSP) and 1 mM DTT for 16 hours, dialyzing against MSP dialysis buffer containing 1 mM DTT. The protein was then dialyzed against MSP dialysis buffer with no DTT for 4 hours. The protein was applied twice to a second Ni-NTA column equilibrated with MSP wash buffer 3 containing 16 mM imidazole. The column was washed with 3 cv MSP wash buffer 3 containing 16 mM imidazole to obtain the cleaved MSP1D1

protein. The purified protein was dialyzed twice against MSP dialysis buffer containing 1 mM EDTA over 24 hours, concentrated, and stored at -20 °C.

4.3.3.3 Assembly and purification of nanodiscs

A molar ratio of 1 MSP1D1ΔH5 to 50 lipids to 0.25 p7 was used for the assembly of the smaller MSP1D1ΔH5 nanodiscs¹²². A molar ratio of 1 MSP1D1E3 to 150 lipids to 1 p7 was used for the assembly of the larger MSP1D1E3 nanodiscs. The lipids used for these nanodiscs were DMPC (Avanti Polar Lipids) and 1,2-dimyristoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (DMPG) (Anatrace) in a molar ratio of 1 DMPG to 3 DMPC.

FPLC-purified p7 was concentrated to 0.5 mg/mL. Lipids were added to the protein and agitated for one hour until the lipids dissolved into solution. The MSP protein was then added and the mixture was rocked gently for one hour at room temperature. Bio-Beads SM-2 Adsorbent (Bio-Rad Laboratories, 1 g per 5 mL solution) was added to the mixture to remove SDS and rocked overnight. The Biobeads were filtered off and a SDS detection kit (G-Biosciences) was used to check for residual SDS. More Biobeads were added to remove SDS if necessary. The assembled nanodiscs were filtered through a 0.45 micron filter and purified by FPLC using a Superdex 200 prep grade size-exclusion column (HiLoad 16/60, GE Healthcare Life Sciences) on a Bio-Logic DuoFlow FPLC (Bio-Rad). The flow rate was 0.8 mL/min over 180 min using nanodisc FPLC buffer (20 mM Tris-HCl, 50 mM NaCl, pH 7.4) The main peak contained a mixture of full and empty nanodiscs. The fractions were

concentrated to 250 μ L using a 15 mL Amicon centrifuge concentrator and used for solution NMR.

Solution NMR experiments were performed on a Bruker Avance 600 MHz spectrometer equipped with a 5-mm triple-resonance cryoprobe with z-axis gradient. One-dimensional ^{15}N -edited ^1H NMR and ^1H - ^{15}N HSQC spectra were obtained at 50 $^{\circ}\text{C}$.

For solid-state experiments, the nanodiscs were sedimented using a Beckman Coulter Airfuge Air-Driven Ultracentrifuge equipped with an A-100/30 rotor at 90k rpm. The bottom 50 μ L of the sedimented sample was packed into a 3.2mm Varian rotor. Solid-state spectra were obtained on a Magnex Scientific 700 MHz Bruker Avance solid-state spectrometer equipped with a home-built 3.2 mm triple-resonance MAS probe. Phosphorus spectra were obtained on a Magnex Scientific 500 MHz Varian Inova solid-state spectrometer equipped with a home-built static ^{31}P - ^1H probe.

4.3.4 Reconstitution of p7 in DMPC proteoliposomes

Right after FPLC purification, 2 mg of uniformly ^{13}C - ^{15}N labeled p7 were concentrated using a 15 mL Amicon centrifuge concentrator to a final volume of 5 mL. 12 mg of DMPC (Anatrace) was added to the solution and allowed to dissolve by gently rocking the mixture for 1 hour at room temperature. The mixture was dialyzed against 20 mM HEPES pH 7.3 over several water changes until the dialysis showed formation of the proteoliposomes^{52, 126} as the solution turned from clear to slightly opaque. The

mixture was then ultracentrifuged for 4 hours at 40k and 15°C using a Beckman 45 Ti rotor in an Optima L-90K Ultracentrifuge (Beckman Coulter). The opaque pellet was resuspended in 9 mL 20 mM HEPES pH 7.3 and ultracentrifuged for 6 hours at 60k and 15 °C using a Beckman 70.1 Ti rotor. To prepare solid-state samples, the pellet was resuspended again in 4 mL 20 mM HEPES pH 7.3 and ultracentrifuged for 16 hours at 90k and 15 °C using a Beckman NVT 90 rotor. The pellet was packed into a Varian 3.2 mm rotor. Solid-state spectra were obtained on a Magnex Scientific 700 MHz solid-state spectrometer equipped with a home-built MAS triple-resonance probe.

All solution-state NMR spectra were processed with NMRPipe (NIH) and visualized with Sparky (UCSF). ^1H - ^{15}N correlation experiments were performed using the fast-HSQC phase-sensitive $g\text{-}2\text{D}$ ^1H - ^{15}N HSQC pulse sequence with WATERGATE, with the acquisition 1024 t_2 points for 108 t_1 increments. HSQC spectra were processed by applying a sin bell function, solvent suppression, zero filling to 4k point, and auto Fourier transform in both the direct and indirect dimensions. All solid-state spectra were processed using Bruker Topspin 1.3 (www.bruker.com).

4.4 Results and discussion

4.4.1 p7 in micelles and q=0.1 bicelles

Phospholipids are composed of a polar phosphate head group exposed to the solvent and a hydrophobic tail composed of varying lengths

of hydrocarbon chains (Figure 4.1). The polar head group and the hydrophobic tail spontaneously arrange themselves with the lipid tails facing each other and the polar heads facing the solvent. Short chain lipids will form small micelles while long chain lipids will form large bilayers²³. A mixture of long and short chain lipids can be used to form bicelles of various sizes. The main phosphate head groups of cell membranes include phosphocholines and phosphoglycerols. Since integral membrane proteins span the phospholipid bilayers, the lipid tails as well as the head group will interact with the protein. Depending on the thickness of the lipid membrane (determined by hydrocarbon chain length) and the properties of the phosphate head group, the same membrane protein can potentially have different structures.

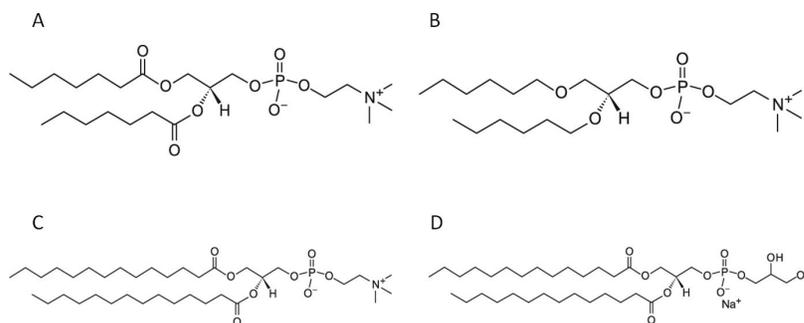


Figure 4.1. Chemical structures of commonly used lipids. DHPC contains two chains with six hydrocarbons linked to the phosphate head by ester bonds (A). 6-O-PC is very similar to DHPC except the hydrophobic tails are linked to the phosphate head by an ether bond (B). DMPC has the same phosphocholine head group like DHPC, but the two chains have twelve hydrocarbons (C). DMPG is similar to DMPC but has a phosphoglycerol head group rather than a phosphocholine head group (D). This gives DMPC a neutral overall charge but DMPG a net negative charge.

Pure p7 protein lyophilized powder after purification by HPLC solubilizes readily in detergent for solution NMR. The pH is adjusted to 7.0 using 1 N NaOH,

and most of the protein remains soluble at physiological pH. Previously, a histidine tag was left at the C-terminus of p7, but the extra residues decreases the solubility of p7 in detergent micelles.

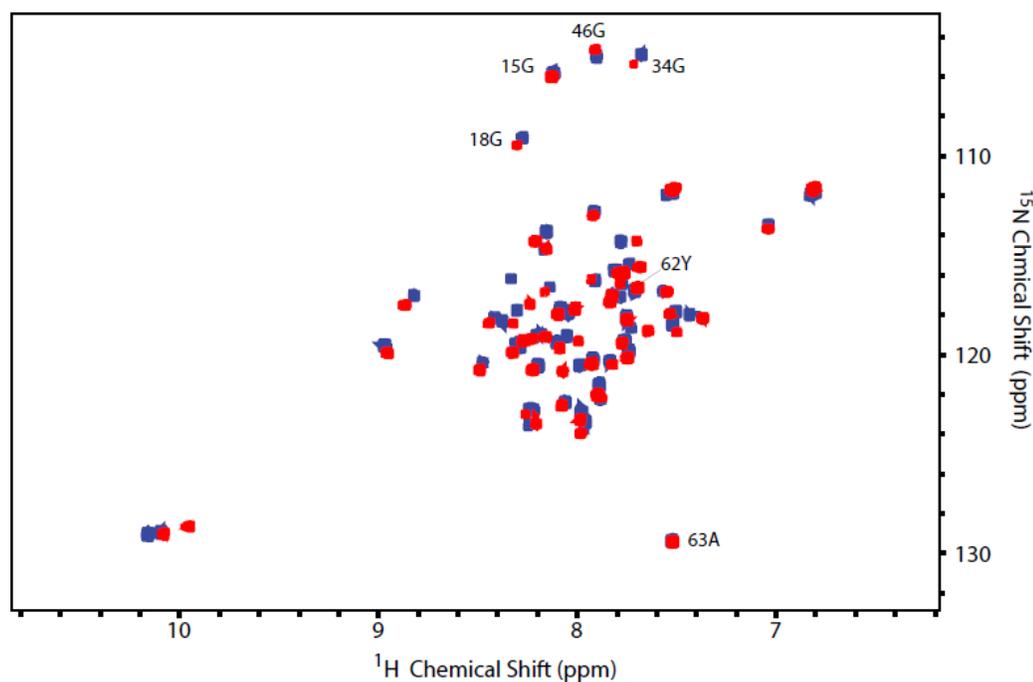


Figure 4.2. ^1H - ^{15}N HSQC of p7 in $q=0.1$ bicelles and 6-O-PC micelles. HSQC of p7 in $q=0.1$ bicelles (red) is overlaid HSQC of p7 in 6-O-PC micelles (blue). Sample contained 50-100 μM p7 at pH 7.0. Both spectra were taken on a Bruker Avance 600 MHz magnet at 50 $^\circ\text{C}$ with 256 scans.

The spectra of p7 in micelles and $q=0.1$ bicelles showed good resolution and overlap with one another and with the published data of p7 in DHPC micelles. Comparing the three HSQCs together, there was greater overlap between the $q=0.1$ bicelles and the 6-O-PC micelles spectra (Figure 4.2) than the DHPC micelles and the 6-O-PC micelles (Figure 4.3). Deuterated p7 was used for the 6-O-PC and $q=0.1$ bicelles spectra, so the peaks are more

resolved than that of DHPC. Furthermore, deuterated 6-O-PC lipids were used, resulting in the best peak resolution. At pH 7.0, four out of the five glycines in p7 can be seen are well overlaid in all three spectra. The most dramatic shifts were alanine 63 and tyrosine 62, the last residues of the p7 sequence. The residues in this region lie in the terminal region of p7, showing that the interaction with the glycerol backbone region of the phospholipid may have minor effects on the soluble regions of p7 while the transmembrane regions have the same structure (Figure 4.3). This is expected since the sizes of the small $q=0.1$ bicelles and DHPC and 6-O-PC micelles are approximately the same, causing the two transmembrane helices of p7 to be folded in the same way. The different linkages between DHPC and 6-O-PC are responsible for the shifts of residues that lie near the membrane surface. While the shifts primarily affect the solvent-exposed regions of p7, the slight differences in the polar head group may have a greater effect for transmembrane proteins that interact more the phospholipid surface.

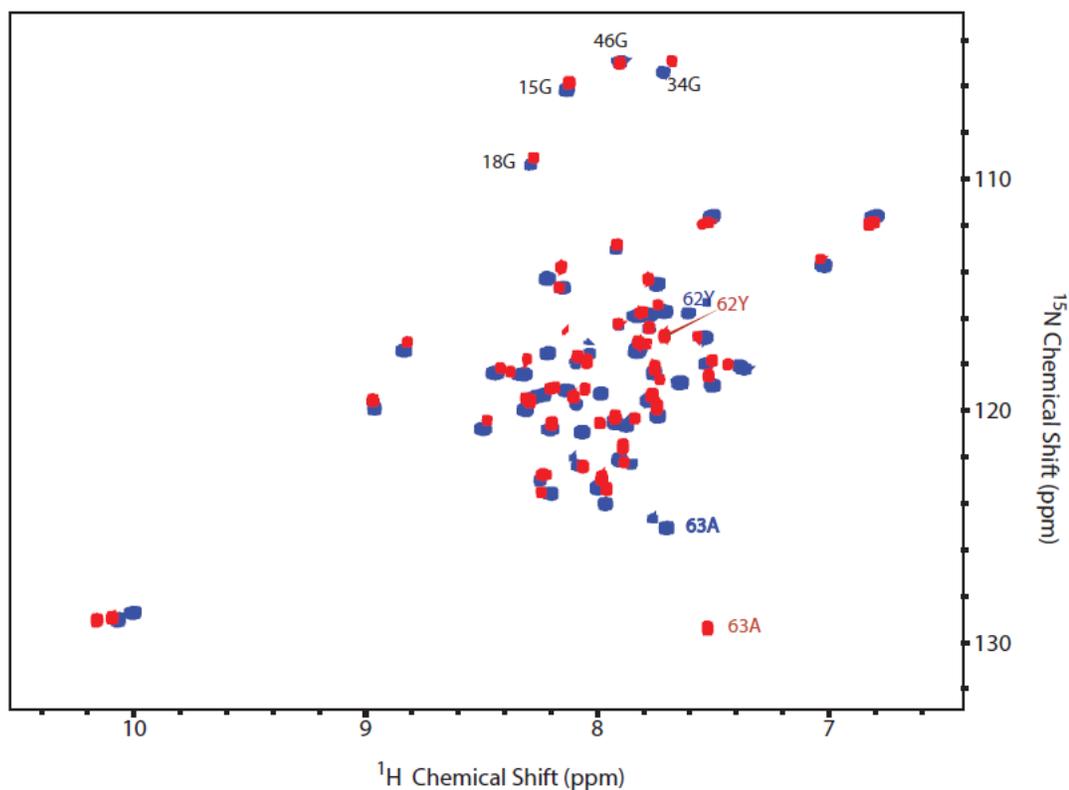


Figure 4.3. ^1H - ^{15}N HSQC of p7 in 6-O-PC micelles and DHPC micelles. HSQC of p7 in 6-O-PC micelles (red) is overlaid HSQC of p7 in DHPC micelles (blue). Sample contained 50-100 μM p7 at pH 7.0. Both spectra were taken on a Bruker Avance 600 MHz magnet at 50 $^\circ\text{C}$ with 256 scans.

4.4.2 p7 in nanodiscs

In the formation of nanodiscs, the ratio of protein to MSP to phospholipid is key^{30, 32, 127, 128}. Too much MSP or phospholipid can result in aggregation of the nanodiscs. Ideally, two membrane scaffold proteins form one nanodisc with one target membrane protein embedded in the center³¹. In reality, there is usually a mixture between full and empty nanodiscs. If the target protein has a histidine tag, then the full and empty nanodiscs can be separated using a Ni-NTA column. In the case of p7, the histidine tag

decreases protein solubility, so the full and empty nanodiscs cannot be separated. Using size-exclusion FPLC, the assembled nanodiscs elute as one peak with aggregates eluting slightly earlier and MSP eluting out later (Figure 4.4).

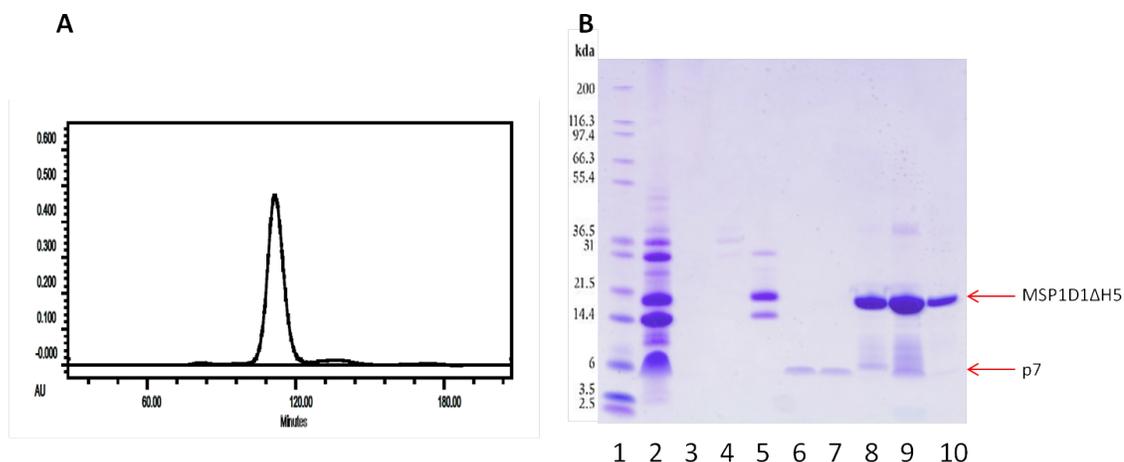


Figure 4.4. Purification of p7 in MSP1D1 Δ H5 nanodiscs. A. The main peak eluting at around 110 min contains a mixture of assembled full and empty nanodiscs. The small peak in front of the main peak contains aggregates. The small peak behind the main peak is the excess MSP. The molar ratio is 0.5 p7 to 2 MSP1D1 Δ H5 to 100 phospholipids (3 DMPC to 1 DMPG). B. The SDS-PAGE of the nanodisc assembly shows Mark12 ladder in lane 1, 15 N p7 post CNBr cleavage in lane 2, FPLC fractions post CNBr cleavage in lanes 3-7, the purified MSP1D1 Δ H5 in lane 8, the assembled nanodiscs using the pure p7 from lanes 6-7 before FPLC in lane 9, and the main elution peak post FPLC nanodisc purification. The p7 band is faint due to the low concentration of p7 compared to MSP (from both full and empty nanodiscs).

The optimized ratio for MSP1D1 Δ H5 nanodiscs is 0.5 p7 to 2 MSP1D1 Δ H5 to 100 phospholipids. At first, the phospholipids used were all DMPC, but these nanodiscs did not give an NMR spectrum with good resolution. Instead 3:1 DMPC:DMPG molar ratio was used. The molar ratio of p7 to MSP1D1 Δ H5 could be increased to 1:2 to lower the number of empty nanodiscs, but using less p7 gave the most uniform peak with the lowest aggregate peak based on the

FPLC chromatogram (Figure 4.4). In addition, there is also the possibility of p7 oligomerizing at higher molar ratios. For MSP1D1E3 nanodiscs, the ratios were not optimized since this was only a preliminary trial with the view to eventually making solid-state nanodisc samples. The molar ratio used for the MSP1D1E3 nanodiscs was 1 p7 to 1 MSP1D1E3 to 150 lipids (3:1 DMPC:DMPG).

After concentrating the assembled nanodisc peak, the samples were studied using solution NMR. Running the nanodisc sample at pH 7 and a temperature of 323K gave the best spectra (Figure 4.5). The sample started precipitating out at pH 4. The signal resolution and intensity also started decreasing as the temperature was lowered to room temperature. Comparing the nanodiscs with $q=0.1$ bicelles, the general peak shape was present for both MSP1D1E3 and MSP1D1 Δ H5 nanodiscs. However, the peaks were broadened for MSP1D1 Δ H5 nanodiscs, and the signal intensity was very low for MSP1D1E3 nanodiscs (Figure 4.5). While it is expected that a higher temperature gives better intensity due to faster tumbling, it was expected that a lower pH would give better signal due to faster exchange. Unfortunately, the sample had started to precipitate within one hour of lowering the pH. The broadened signal for the nanodiscs is expected since nanodiscs are much larger than $q=0.1$ bicelles and tumble much slower. However, the low intensity for the MSP1D1E3 nanodisc could indicate either a low concentration or that the nanodisc is tumbling too slowly to be detected by solution NMR.

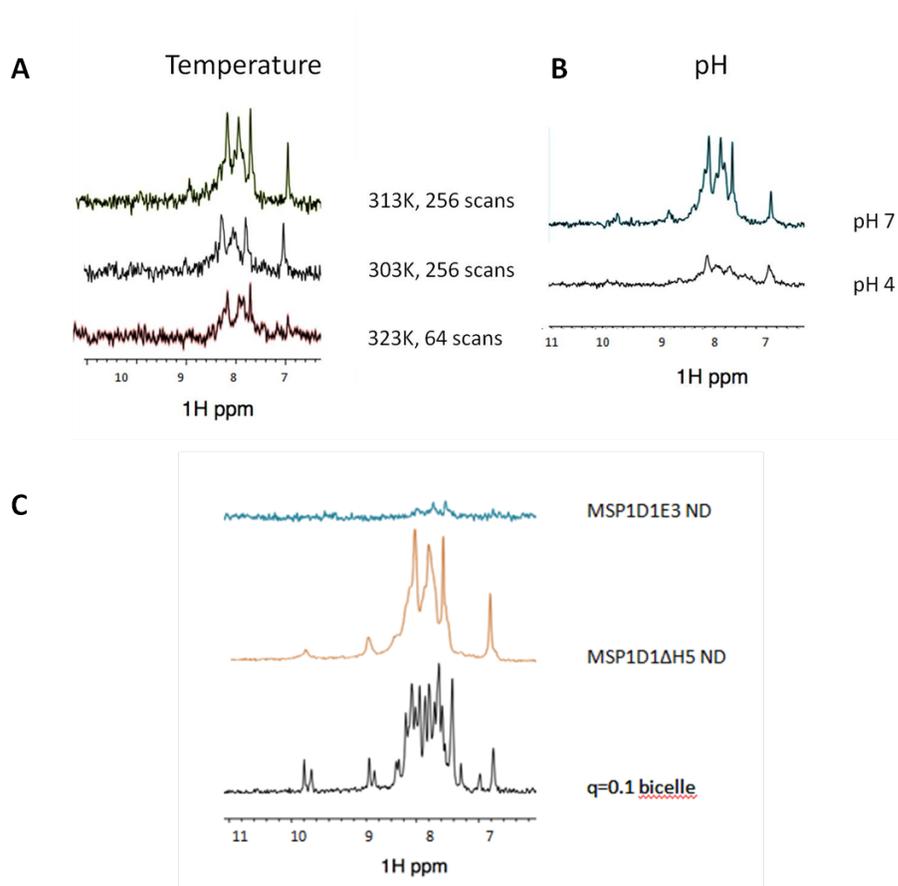


Figure 4.5. ^{15}N -edited ^1H 1-D spectra of MSP1D1 Δ H5 nanodiscs by solution NMR. The results were recorded on a Bruker Avance 600 MHz magnet. The peak intensity at pH 7 decreases with temperature. 323K was the optimum temperature (A). The peak resolution and intensity decreased going from pH 7 to pH 4, 323K, 1024 scans (B). The peaks are broadened for MSP1D1 Δ H5 nanodiscs and almost disappear for MSP1D1E3 nanodiscs when compared to q=0.1 bicelles, 323K, pH 7, 1024 scans (C).

A ^1H - ^{15}N HSQC was obtained for p7 in MSP1D1 Δ H5 nanodiscs (Figure 4.6). Since nanodiscs are much larger in size (diameters of ~ 8.2 nm for MSP1D1 Δ H5 and ~ 13 nm for MSP1D1E3)^{30, 34, 127, 129} compared to q=0.1 bicelles (diameter of 2-4 nm)²⁴, the peaks are greatly broadened. In addition, only about 30 peaks could be picked up despite using a more concentrated sample. In solution NMR, fast molecular tumbling averages out the chemical shift anisotropy tensors to zero and gives an isotropic peak. The smaller a

molecule is, the faster the tumbling and the sharper the peaks are. Highly resolved, single peaks can be seen in Figure 4.6 for p7 in $q=0.1$ bicelles, indicating the protein is small enough to undergo fast molecular tumbling. Larger proteins tumble slowly, giving poorly resolved peaks. This appears to be the case for p7 in nanodiscs, which are roughly four times the size of $q=0.1$ bicelles. Many of the peaks have large broadening while other cannot be detected.

As seen in figure 4.6, some peaks are much more intense than others. One glycine (G15) has high intensity while the others barely appear after increasing the contour levels. As expected the solvent-exposed terminal alanine 63 shows up with high intensity, but what is surprising is that a larger number of transmembrane peaks also appear with high intensity. For example, histidine 17, valine 47, and leucine 50 are all within the transmembrane regions and overlay well with the bicelle spectrum. About half of the remaining peaks also overlay well but obtaining a good spectrum is hindered by the poor resolution or low signal intensity. This suggests that the structure of p7 in nanodiscs is similar to the isotropic bicelle and micelle structures but is more different from any one of those structures than the isotropic bicelles and micelles are from each other. It would be interesting to do further studies using deuterated p7^{16, 28, 57, 58, 130} and deuterated lipid to obtain better resolution and study the structural differences in nanodiscs.

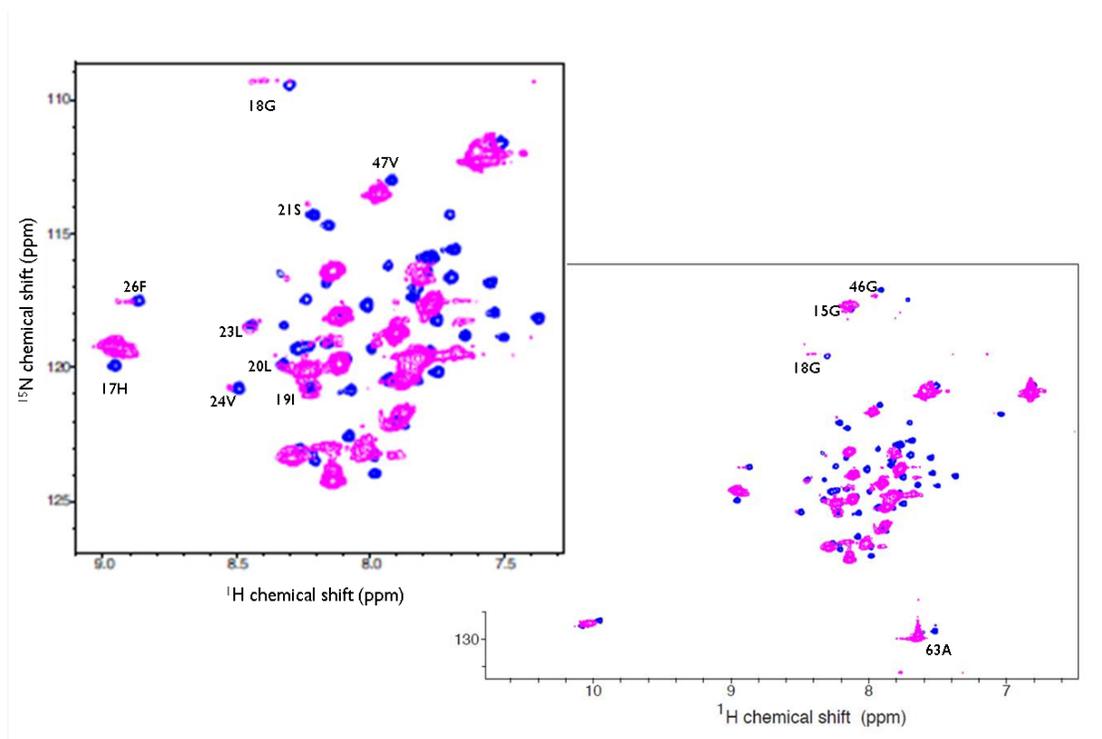


Figure 4.6. ^1H - ^{15}N HSQC overlay of p7 in MSP1D1 Δ H5 nanodiscs over p7 in q=0.1 bicelles. The q=0.1 bicelle spectrum is in blue, and the nanodisc spectrum is in magenta. Samples contained 150 μM p7 for the nanodisc sample and 50-100 μM p7 for the bicelles. Experiment was conducted on a Bruker Avance 600 MHz magnet at 323K, pH 7, and 256 scans.

Since nanodiscs have been used to obtain oligomeric structures of proteins¹²⁹ and p7 is known to oligomerize, a Native-PAGE was run to see if p7 oligomerizes in the MSP1D1 Δ H5 nanodiscs (Figure 4.7). Wild-type and truncated interleukin 8 (IL-8) were used as a control since wild-type IL-8 dimerizes¹³¹ but truncated IL-8 is not able to. The MSP1D1 Δ H5 belt protein, empty nanodisc, and p7 were also run separately. Samples of the p7 nanodiscs at pH 4 and pH 7 were also compared to see if the change in pH promoted protein oligomerization.

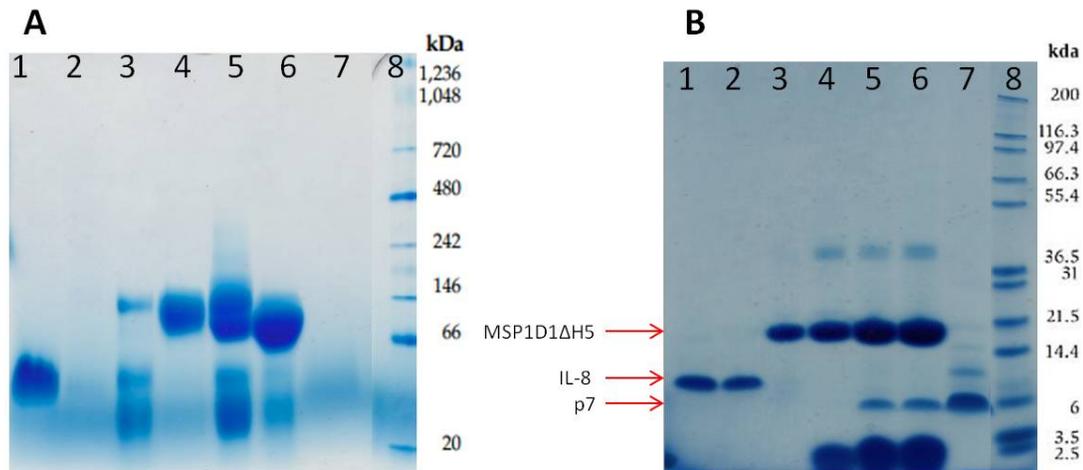


Figure 4.7. Native-PAGE vs. SDS-PAGE of p7 in nanodiscs. Native-PAGE (A) and SDS-PAGE (B) were run to check for oligomeric p7. The samples were the same for both gels. Truncated IL-8 is in lane 1, wild-type IL-8 is in lane 2, MSP1D1 Δ H5 is in lane 3, empty nanodiscs are in lane 4, p7 in MSP1D1 Δ H5 nanodiscs at pH 4 are in lane 5, p7 in MSP1D1 Δ H5 nanodiscs at pH 7 are in lane 6, p7 in SDS is in lane 7, and either NativeMark (A) or Mark12 (B) ladders are in lane 8.

The results were inconclusive since it appeared that the truncated IL-8 should run around 8 kDa and not over 20 kDa as shown, and the wild-type IL-8 did not appear to oligomerize. Also, MSP1D1 Δ H5 oligomerizes by itself and the empty nanodisc ran at a higher lower molecular weight compared to the full nanodiscs at pH 7. Given that an empty nanodisc is approximately 72kDa, it may be hard to tell based on the Native-PAGE separation and band size if p7 forms oligomers. However, at pH 4, it appears as if the nanodisc is partially dissembled due to appearance of the MSP1D1 Δ H5 bands, accounting for the precipitation. At pH 7, the nanodisc is still intact. This shows that nanodiscs stability is pH-dependent, dissembling at lower pH but stable at physiological pH. Considering the low p7 to belt protein ratio, it is not likely that p7 oligomerized in these small nanodiscs. For larger nanodiscs, however, the

amount of phospholipids can be altered to select for oligomerized protein. This is worth examining further in the case of p7 since its oligomeric state is still debated.

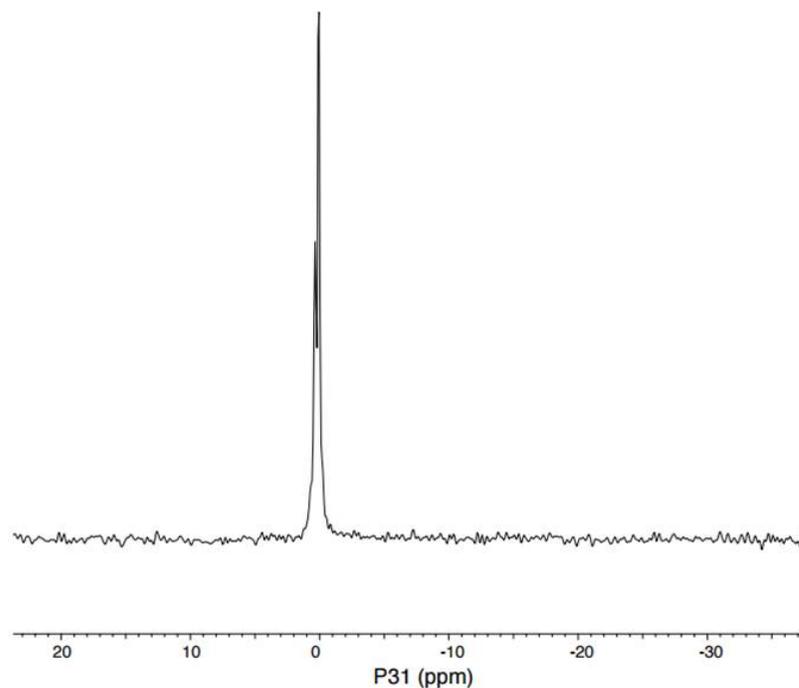


Figure 4.8. ^{31}P NMR of p7 in MSP1D1E3 nanodiscs. The two isotropic peaks correspond to the phospholipids DMPC and DMPG in a 3 to 1 ratio due to the 3:1 DMPC:DMPG ratio used in the assembly of the nanodiscs.

For solid-state NMR, the MSP1D1 Δ H5 nanodiscs were sedimented. While there was a clear separation between a very fluid-like top layer and a denser lower layer, it was difficult extracting the bottom portion after ultracentrifuging. Small disturbances would easily mix the two layers together. Nonetheless, the bottommost portion of the sedimentation was packed into a Varian 3.2 mm MAS rotor. Both ^1H - ^{15}N cross polarization (CP) and ^1H - ^{15}N insensitive nuclei enhanced polarization transfer (INEPT) experiments did not

show any signal, probably due to inefficient sedimentation of the small nanodisc. In a second attempt, larger MSP1D1E3 nanodiscs were also packed into an MAS rotor for ssNMR. Instead of sedimenting the sample, the FPLC fraction was concentrated to 200 μ L and water was removed under nitrogen gas until the sample was 50 μ L in volume. The concentrated sample of the MSP1D1E3 nanodiscs also did not show any signal from ^1H - ^{15}N CP or ^1H - ^{15}N INEPT (data not shown), suggesting that there was not enough p7 in the sample or that motion was not restricted enough. Unlike solution NMR which depends on fast molecular tumbling, solid-state NMR requires that the protein be motionally restricted. The protein should still undergo rotational diffusion in the lipid bilayer but be rigid enough to align to the magnetic field. Typically, a solid-state sample for a membrane protein in lipid bilayers is gel-like with 100-200% hydration. In my case, the concentrated nanodisc sample was more fluid-like than gel-like, which can explain the lack of signal due to the protein not being motionally restricted.

The MSP1D1E3 solid-state sample was examined by ^{31}P NMR to check if the sample was stable and if the lipids had deteriorated (Figure 4.8). The spectrum showed two isotropic peaks, matching the two different phospholipids used for the nanodisc. For a solid-state sample a powder pattern would be expected. Instead, isotropic peaks indicate that the nanodiscs were not aligned with the field or motionally restricted enough for solid-state NMR.

With more optimization, nanodiscs are feasible for solid-state NMR. Small nanodiscs can tumble fast enough to obtain solution NMR spectra but may not 'solidify' enough for solid-state NMR. Larger nanodiscs are not ideal for solution NMR due to its slow tumbling but may be able to sediment better for solid-state NMR. The protein, MSP, and lipid ratios for assembling p7 in larger nanodisc still need to be optimized.

4.4.3 p7 in lipid bilayers

Proteoliposome samples of p7 in DMPC (1 mg p7 to 6 mg DMPC) were successfully prepared with 300% hydration. As seen in Figure 4.9, the freshly prepared sample underwent rotational diffusion but did not continue to do so after spinning at 5 kHz overnight at 25 °C. The ^{13}C - ^{13}C correlation was obtained at 10 kHz and 8 °C when the protein was not undergoing rotational diffusion (Figure 4.9). Although the quality of the sample may have been compromised, the spectrum showed relatively good resolution and reproducibility. Two samples of p7 in DMPC were made and although both of them stopped undergoing rotational diffusion within 24 hours, both ^{13}C - ^{13}C correlation spectra looked identical.

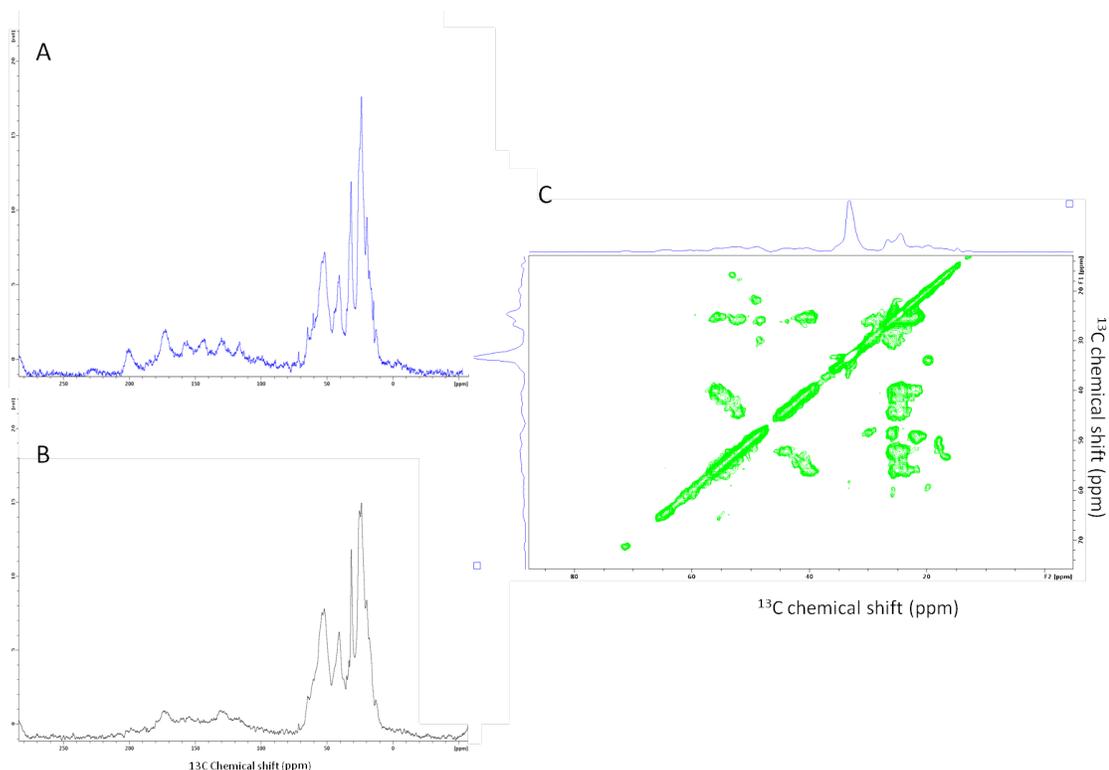


Figure 4.9. ^1H - ^{13}C cross-polarization and ^{13}C - ^{13}C correlation of p7 in DMPC proteoliposomes. The sample contained 1.5 mg U- ^{13}C - ^{15}N p7 in DMPC with 200-300% hydration. CP experiments were taken at 298K, 5 kHz spin rate, with 256 scans. A. After 1 day, the sample was not undergoing rotational diffusion since the side bands are seen in the carbonyl region (150-200 ppm). B. The freshly prepared sample did undergo rotational diffusion since the side bands are not seen in the carbonyl region. C. The ^{13}C - ^{13}C correlation was taken at 278K, 10 kHz spin rate, and 384 scans. The acquisition time was 20 ms in the indirect dimension and 9 ms in the direct dimension and the decoupling power was 92.6 kHz.

Despite the small size of viroporin p7, its complex properties make it a challenge to study by ssNMR. Further efforts need to be made to ensure the protein undergoes rotational diffusion over the period of time needed to obtain ssNMR data. In the future reconstitution methods, protein-lipid ratios, and even types of lipids can be varied in order to optimize reconstitution of p7 into proteoliposomes. Since ssNMR is not limited by the size of the molecule, it can also be used to examine the oligomeric state of p7 in its native-state.

4.5 Conclusion

The structure of p7 appears to alter slightly with small changes in the lipid environment. This is seen comparing the HSQC spectra of p7 in $q=0.1$ bicelles, DHPC micelles, and 6-O-PC micelles. Small variations in the glycerol head group and hydrocarbon tail length affect the amino acid residues that are closest to the membrane surface. Comparing p7 in small MSP1D1 Δ H5 nanodiscs with $q=0.1$ bicelles, it appears that there is a greater chemical shift in about half the residues. Although the HSQC data does not indicate a dramatically different structure, it may be worth investigating further to help refine the current micelle structure. The use of nanodiscs for solution NMR of transmembrane proteins is hindered by their relatively large size, so deuterated protein and lipids would need to be used for further studies. Additionally, nanodiscs have been used to study oligomeric states of proteins, and this would be an ideal vehicle to study the oligomeric state of p7. Additionally, nanodiscs have been used for solid-state NMR, but the protocol for these still needs to be optimized. The structure of p7 has also been studied by ssNMR in our lab; however, the behavior of p7 in lipid bilayers appears to be challenging. While the sample may undergo rotational diffusion after packing, it stops doing so after a few hours. There still needs to be further studies to optimize a proteoliposome sample that undergoes rotational diffusion for a period long enough to obtain solid-state data.

REFERENCES

- [1] Carpenter, E. P., Beis, K., Cameron, A. D., and Iwata, S. (2008) Overcoming the challenges of membrane protein crystallography, *Curr Opin Struct Biol* 18, 581-586.
- [2] Moraes, I., Evans, G., Sanchez-Weatherby, J., Newstead, S., and Stewart, P. D. (2014) Membrane protein structure determination - the next generation, *Biochim Biophys Acta* 1838, 78-87.
- [3] Radoicic, J., Lu, G. J., and Opella, S. J. (2014) NMR structures of membrane proteins in phospholipid bilayers, *Q Rev Biophys* 47, 249-283.
- [4] Opella, S. J. (2015) Solid-state NMR and membrane proteins, *J Magn Reson* 253, 129-137.
- [5] Zein, N. N. (2000) Clinical significance of hepatitis C virus genotypes, *Clin Microbiol Rev* 13, 223-235.
- [6] Mihm, U., Grigorian, N., Welsch, C., Herrmann, E., Kronenberger, B., Teuber, G., von Wagner, M., Hofmann, W. P., Albrecht, M., Lengauer, T., Zeuzem, S., and Sarrazin, C. (2006) Amino acid variations in hepatitis C virus p7 and sensitivity to antiviral combination therapy with amantadine in chronic hepatitis C, *Antivir Ther* 11, 507-519.
- [7] Messina, J. P., Humphreys, I., Flaxman, A., Brown, A., Cooke, G. S., Pybus, O. G., and Barnes, E. (2015) Global distribution and prevalence of hepatitis C virus genotypes, *Hepatology* 61, 77-87.
- [8] Deisenhofer, J., Epp, O., Miki, K., Huber, R., and Michel, H. (1985) Structure of the protein subunits in the photosynthetic reaction centre of *Rhodospseudomonas viridis* at 3A resolution, *Nature* 318, 618-624.
- [9] Yao, Y., Ding, Y., Tian, Y., Opella, S. J., and Marassi, F. M. (2013) Membrane protein structure determination: back to the membrane, *Methods Mol Biol* 1063, 145-158.

- [10] Opella, S. J. (2014) The development of solid-state NMR of membrane proteins, *Biomed Spectrosc Imaging* 3, 81-105.
- [11] Opella, S. J. (2013) Structure determination of membrane proteins in their native phospholipid bilayer environment by rotationally aligned solid-state NMR spectroscopy, *Accounts of chemical research* 46, 2145-2153.
- [12] Campbell, I. D., Dobson, C. M., Williams, R. J., and Xavier, A. V. (1973) The determination of the structure of proteins in solution: lysozyme, *Annals of the New York Academy of Sciences* 222, 163-174.
- [13] Wuthrich, K. (1990) Protein structure determination in solution by NMR spectroscopy, *J Biol Chem* 265, 22059-22062.
- [14] Wagner, G., and Wuthrich, K. (1978) Dynamic model of globular protein conformations based on NMR studies in solution, *Nature* 275, 247-248.
- [15] Cook, G. A., and Opella, S. J. (2010) NMR studies of membrane proteins, *Methods Mol Biol* 637, 263-275.
- [16] Tang, M., Comellas, G., Mueller, L. J., and Rienstra, C. M. (2010) High resolution $(1)(3)\text{C}$ -detected solid-state NMR spectroscopy of a deuterated protein, *J Biomol NMR* 48, 103-111.
- [17] Markley, J. L., Putter, I., and Jardetzky, O. (1968) High-resolution nuclear magnetic resonance spectra of selectively deuterated staphylococcal nuclease, *Science* 161, 1249-1251.
- [18] Paliy, O., Bloor, D., Brockwell, D., Gilbert, P., and Barber, J. (2003) Improved methods of cultivation and production of deuterated proteins from *E. coli* strains grown on fully deuterated minimal medium, *J Appl Microbiol* 94, 580-586.
- [19] Tang, M., Berthold, D. A., and Rienstra, C. M. (2011) Solid-State NMR of a Large Membrane Protein by Paramagnetic Relaxation Enhancement, *The journal of physical chemistry letters* 2, 1836-1841.

- [20] Knight, M. J., Felli, I. C., Pierattelli, R., Emsley, L., and Pintacuda, G. (2013) Magic angle spinning NMR of paramagnetic proteins, *Accounts of chemical research* 46, 2108-2116.
- [21] Jaroniec, C. P. (2012) Solid-state nuclear magnetic resonance structural studies of proteins using paramagnetic probes, *Solid state nuclear magnetic resonance* 43-44, 1-13.
- [22] Jardetzky, O., and Wade-Jardetzky, N. G. (1980) Comparison of protein structures by high resolution solid state and solution NMR, *FEBS Lett* 110, 133-135.
- [23] Mus-Veteau, I. (2014) *Membrane proteins production for structural analysis*, Springer, New York.
- [24] Wu, H., Su, K., Guan, X., Sublette, M. E., and Stark, R. E. (2010) Assessing the size, stability, and utility of isotropically tumbling bicelle systems for structural biology, *Biochim Biophys Acta* 1798, 482-488.
- [25] Son, W. S., Park, S. H., Nothnagel, H. J., Lu, G. J., Wang, Y., Zhang, H., Cook, G. A., Howell, S. C., and Opella, S. J. (2012) 'q-Titration' of long-chain and short-chain lipids differentiates between structured and mobile residues of membrane proteins studied in bicelles by solution NMR spectroscopy, *J Magn Reson* 214, 111-118.
- [26] Hilty, C., Wider, G., Fernandez, C., and Wuthrich, K. (2004) Membrane protein-lipid interactions in mixed micelles studied by NMR spectroscopy with the use of paramagnetic reagents, *Chembiochem* 5, 467-473.
- [27] Cook, G. A., Zhang, H., Park, S. H., Wang, Y., and Opella, S. J. (2011) Comparative NMR studies demonstrate profound differences between two viroporins: p7 of HCV and Vpu of HIV-1, *Biochim Biophys Acta* 1808, 554-560.
- [28] Triba, M. N., Devaux, P. F., and Warschawski, D. E. (2006) Effects of lipid chain length and unsaturation on bicelles stability. A phosphorus NMR study, *Biophys J* 91, 1357-1367.

- [29] Raschle, T., Hiller, S., Eitzkorn, M., and Wagner, G. (2010) Nonmicellar systems for solution NMR spectroscopy of membrane proteins, *Curr Opin Struct Biol* 20, 471-479.
- [30] Ritchie, T. K., Grinkova, Y. V., Bayburt, T. H., Denisov, I. G., Zolnerciks, J. K., Atkins, W. M., and Sligar, S. G. (2009) Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs, *Methods in enzymology* 464, 211-231.
- [31] Ding, Y., Fujimoto, L. M., Yao, Y., and Marassi, F. M. (2015) Solid-state NMR of the *Yersinia pestis* outer membrane protein Ail in lipid bilayer nanodiscs sedimented by ultracentrifugation, *J Biomol NMR* 61, 275-286.
- [32] Bayburt, T. H., and Sligar, S. G. (2010) Membrane protein assembly into Nanodiscs, *FEBS Lett* 584, 1721-1727.
- [33] Bayburt, T. H., Grinkova, Y. V., and Sligar, S. G. (2006) Assembly of single bacteriorhodopsin trimers in bilayer nanodiscs, *Arch Biochem Biophys* 450, 215-222.
- [34] Park, S. H., Berkamp, S., Cook, G. A., Chan, M. K., Viadiu, H., and Opella, S. J. (2011) Nanodiscs versus macrodiscs for NMR of membrane proteins, *Biochemistry* 50, 8983-8985.
- [35] Cook, G. A., Dawson, L. A., Tian, Y., and Opella, S. J. (2013) Three-dimensional structure and interaction studies of hepatitis C virus p7 in 1,2-dihexanoyl-sn-glycero-3-phosphocholine by solution nuclear magnetic resonance, *Biochemistry* 52, 5295-5303.
- [36] Saint, N., Montserret, R., Chipot, C., and Penin, F. (2009) Structural and functional analysis of the HCV p7 protein, *Methods Mol Biol* 510, 125-143.
- [37] OuYang, B., and Chou, J. J. (2014) The minimalist architectures of viroporins and their therapeutic implications, *Biochim Biophys Acta* 1838, 1058-1067.

- [38] Madan, V., and Bartenschlager, R. (2015) Structural and Functional Properties of the Hepatitis C Virus p7 Viroporin, *Viruses* 7, 4461-4481.
- [39] Gentzsch, J., Brohm, C., Steinmann, E., Friesland, M., Menzel, N., Vieyres, G., Perin, P. M., Frentzen, A., Kaderali, L., and Pietschmann, T. (2013) Hepatitis c Virus p7 is critical for capsid assembly and envelopment, *PLoS Pathog* 9, e1003355.
- [40] Gonzalez, M. E., and Carrasco, L. (2003) Viroporins, *FEBS Lett* 552, 28-34.
- [41] Opella, S. J. (2015) Relating structure and function of viral membrane-spanning miniproteins, *Curr Opin Virol* 12, 121-125.
- [42] Nieva, J. L., Madan, V., and Carrasco, L. (2012) Viroporins: structure and biological functions, *Nat Rev Microbiol* 10, 563-574.
- [43] Montserret, R., Saint, N., Vanbelle, C., Salvay, A. G., Simorre, J. P., Ebel, C., Sapay, N., Renisio, J. G., Bockmann, A., Steinmann, E., Pietschmann, T., Dubuisson, J., Chipot, C., and Penin, F. (2010) NMR structure and ion channel activity of the p7 protein from hepatitis C virus, *J Biol Chem* 285, 31446-31461.
- [44] Kalita, M. M., Griffin, S., Chou, J. J., and Fischer, W. B. (2015) Genotype-specific differences in structural features of hepatitis C virus (HCV) p7 membrane protein, *Biochim Biophys Acta* 1848, 1383-1392.
- [45] Bentham, M. J., Foster, T. L., McCormick, C., and Griffin, S. (2013) Mutations in hepatitis C virus p7 reduce both the egress and infectivity of assembled particles via impaired proton channel function, *J Gen Virol* 94, 2236-2248.
- [46] OuYang, B., Xie, S., Berardi, M. J., Zhao, X., Dev, J., Yu, W., Sun, B., and Chou, J. J. (2013) Unusual architecture of the p7 channel from hepatitis C virus, *Nature* 498, 521-525.
- [47] Foster, T. L., Thompson, G. S., Kalverda, A. P., Kankanala, J., Bentham, M., Wetherill, L. F., Thompson, J., Barker, A. M., Clarke, D., Noerenberg, M., Pearson, A. R., Rowlands, D. J., Homans, S. W., Harris, M., Foster, R., and

- Griffin, S. (2014) Structure-guided design affirms inhibitors of hepatitis C virus p7 as a viable class of antivirals targeting virion release, *Hepatology* 59, 408-422.
- [48] Dev, J., Bruschweiler, S., Ouyang, B., and Chou, J. J. (2015) Transverse relaxation dispersion of the p7 membrane channel from hepatitis C virus reveals conformational breathing, *J Biomol NMR* 61, 369-378.
- [49] Luik, P., Chew, C., Aittoniemi, J., Chang, J., Wentworth, P., Jr., Dwek, R. A., Biggin, P. C., Venien-Bryan, C., and Zitzmann, N. (2009) The 3-dimensional structure of a hepatitis C virus p7 ion channel by electron microscopy, *Proc Natl Acad Sci U S A* 106, 12712-12716.
- [50] Gottwein, J. M., Scheel, T. K., Jensen, T. B., Lademann, J. B., Prentoe, J. C., Knudsen, M. L., Hoegh, A. M., and Bukh, J. (2009) Development and characterization of hepatitis C virus genotype 1-7 cell culture systems: role of CD81 and scavenger receptor class B type I and effect of antiviral drugs, *Hepatology* 49, 364-377.
- [51] Cook, G. A., Stefer, S., and Opella, S. J. (2011) Expression and purification of the membrane protein p7 from hepatitis C virus, *Biopolymers* 96, 32-40.
- [52] Park, S. H., Das, B. B., Casagrande, F., Tian, Y., Nothnagel, H. J., Chu, M., Kiefer, H., Maier, K., De Angelis, A. A., Marassi, F. M., and Opella, S. J. (2012) Structure of the chemokine receptor CXCR1 in phospholipid bilayers, *Nature* 491, 779-783.
- [53] Shiloach, J., and Fass, R. (2005) Growing E. coli to high cell density--a historical perspective on method development, *Biotechnol Adv* 23, 345-357.
- [54] Marley, J., Lu, M., and Bracken, C. (2001) A method for efficient isotopic labeling of recombinant proteins, *J Biomol NMR* 20, 71-75.
- [55] Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures, *Protein Expr Purif* 41, 207-234.

- [56] Dumon-Seignovert, L., Cariot, G., and Vuillard, L. (2004) The toxicity of recombinant proteins in *Escherichia coli*: a comparison of overexpression in BL21 (DE3), C41 (DE3), and C43 (DE3), *Protein Expr Purif* 37, 203-206.
- [57] Venditti, V., Fawzi, N. L., and Clore, G. M. (2012) An efficient protocol for incorporation of an unnatural amino acid in perdeuterated recombinant proteins using glucose-based media, *Journal of Biomolecular Nmr* 52, 191-195.
- [58] Meissner, A., Briand, J., and Sphirensen, O. W. (1998) Editing of multidimensional NMR spectra of partially deuterated proteins. Measurement of amide deuterium isotope effects on the chemical shifts of protein backbone nuclei, *J Biomol NMR* 12, 339-343.
- [59] Cook, G. A., and Opella, S. J. (2011) Secondary structure, dynamics, and architecture of the p7 membrane protein from hepatitis C virus by NMR spectroscopy, *Biochim Biophys Acta* 1808, 1448-1453.
- [60] Saunders, M., Wishnia, A., and Kirkwood, J. G. (1957) The Nuclear Magnetic Resonance Spectrum of Ribonuclease, *Journal of the American Chemical Society* 79, 3289-3290.
- [61] Meadows, D. H., Markley, J. L., Cohen, J. S., and Jardetzky, O. (1967) Nuclear magnetic resonance studies of the structure and binding sites of enzymes. I. Histidine residues, *Proc Natl Acad Sci U S A* 58, 1307-1313.
- [62] Cross, T. A., and Opella, S. J. (1979) NMR of fd coat protein, *Journal of supramolecular structure* 11, 139-145.
- [63] Hagen, D. S., Weiner, J. H., and Sykes, B. D. (1978) Fluorotyrosine M13 coat protein: fluorine-19 nuclear magnetic resonance study of the motional properties of an integral membrane protein in phospholipid vesicles, *Biochemistry* 17, 3860-3866.
- [64] Hagen, D. S., Weiner, J. H., and Sykes, B. D. (1979) Investigation of solvent accessibility of the fluorotyrosyl residues of M13 coat protein in deoxycholate micelles and phospholipid vesicles, *Biochemistry* 18, 2007-2012.

- [65] Cross, T. A., and Opella, S. J. (1980) Structural properties of fd coat protein in sodium dodecyl sulfate micelles, *Biochemical and biophysical research communications* 92, 478-484.
- [66] Morrisett, J. D., Wien, R. W., and McConnell, H. M. (1973) The use of spin labels for measuring distances in biological systems, *Annals of the New York Academy of Sciences* 222, 149-162.
- [67] Wien, R. W., McConnel.Hm, and Morriset.Jd. (1972) Spin-Label-Induced Nuclear Relaxation - Distances between Bound Saccharides, Histidine-15, and Tryptophan-123 on Lysozyme in Solution, *Biochemistry* 11, 3707-&.
- [68] Ganguly, S., Weiner, B. E., and Meiler, J. (2011) Membrane protein structure determination using paramagnetic tags, *Structure* 19, 441-443.
- [69] Chen, H., Ji, F., Olman, V., Mobley, C. K., Liu, Y., Zhou, Y., Bushweller, J. H., Prestegard, J. H., and Xu, Y. (2011) Optimal mutation sites for PRE data collection and membrane protein structure prediction, *Structure* 19, 484-495.
- [70] Mesleh, M. F., Lee, S., Veglia, G., Thiriot, D. S., Marassi, F. M., and Opella, S. J. (2003) Dipolar waves map the structure and topology of helices in membrane proteins, *J Am Chem Soc* 125, 8928-8935.
- [71] Lee, H. S., Spraggon, G., Schultz, P. G., and Wang, F. (2009) Genetic incorporation of a metal-ion chelating amino acid into proteins as a biophysical probe, *J Am Chem Soc* 131, 2481-2483.
- [72] Liu, W. M., Overhand, M., and Ubbink, M. (2014) The application of paramagnetic lanthanoid ions in NMR spectroscopy on proteins, *Coordination Chemistry Reviews* 273, 2-12.
- [73] Otting, G. (2010) Protein NMR using paramagnetic ions, *Annual review of biophysics* 39, 387-405.
- [74] Bertini, I., Luchinat, C., Parigi, G., and Pierattelli, R. (2008) Perspectives in paramagnetic NMR of metalloproteins, *Dalton transactions*, 3782-3790.

- [75] Clore, G. M. (2013) Seeing the invisible by paramagnetic and diamagnetic NMR, *Biochem Soc Trans* 41, 1343-1354.
- [76] Sengupta, I., Nadaud, P. S., and Jaroniec, C. P. (2013) Protein structure determination with paramagnetic solid-state NMR spectroscopy, *Accounts of chemical research* 46, 2117-2126.
- [77] Zhuang, T. D., Lee, H. S., Imperiali, B., and Prestegard, J. H. (2008) Structure determination of a Galectin-3-carbohydrate complex using paramagnetism-based NMR constraints, *Protein Science* 17, 1220-1231.
- [78] McConnell, H. M., and McFarland, B. G. (1970) Physics and chemistry of spin labels, *Q Rev Biophys* 3, 91-136.
- [79] Abragam, A., and Bleaney, B. (2012) *Electron Paramagnetic Resonance of Transition Ions*, Oxford University Press, Oxford.
- [80] Inubushi, T., and Becker, E. D. (1983) Efficient Detection of Paramagnetically Shifted Nmr Resonances by Optimizing the Weft Pulse Sequence, *Journal of Magnetic Resonance* 51, 128-133.
- [81] Ward, M. E., Wang, S., Krishnamurthy, S., Hutchins, H., Fey, M., Brown, L. S., and Ladizhansky, V. (2014) High-resolution paramagnetically enhanced solid-state NMR spectroscopy of membrane proteins at fast magic angle spinning, *J Biomol NMR* 58, 37-47.
- [82] Parthasarathy, S., Nishiyama, Y., and Ishii, Y. (2013) Sensitivity and resolution enhanced solid-state NMR for paramagnetic systems and biomolecules under very fast magic angle spinning, *Accounts of chemical research* 46, 2127-2135.
- [83] Ullrich, S. J., Holper, S., and Glaubitz, C. (2014) Paramagnetic doping of a 7TM membrane protein in lipid bilayers by Gd(3)(+)-complexes for solid-state NMR spectroscopy, *J Biomol NMR* 58, 27-35.
- [84] Prestegard, J. H., al-Hashimi, H. M., and Tolman, J. R. (2000) NMR structures of biomolecules using field oriented media and residual dipolar couplings, *Q Rev Biophys* 33, 371-424.

- [85] Tolman, J. R., Flanagan, J., Kennedy, M., and Prestegard, J. H. (1995) Nuclear magnetic dipole interactions in field-oriented proteins: Information for structure determination in solution, *Proc Natl Acad Sci U S A* 92, 9279-9283.
- [86] Balayssac, S., Bertini, I., Lelli, M., Luchinat, C., and Maletta, M. (2007) Paramagnetic ions provide structural restraints in solid-state NMR of proteins, *J Am Chem Soc* 129, 2218-2219.
- [87] Knight, M. J., Pell, A. J., Bertini, I., Felli, I. C., Gonnelli, L., Pierattelli, R., Herrmann, T., Emsley, L., and Pintacuda, G. (2012) Structure and backbone dynamics of a microcrystalline metalloprotein by solid-state NMR, *Proc Natl Acad Sci U S A* 109, 11095-11100.
- [88] Ma, C., and Opella, S. J. (2000) Lanthanide ions bind specifically to an added "EF-hand" and orient a membrane protein in micelles for solution NMR spectroscopy, *J Magn Reson* 146, 381-384.
- [89] Wohnert, J., Franz, K. J., Nitz, M., Imperiali, B., and Schwalbe, H. (2003) Protein alignment by a coexpressed lanthanide-binding tag for the measurement of residual dipolar couplings, *J Am Chem Soc* 125, 13338-13339.
- [90] Barthelmes, K., Reynolds, A. M., Peisach, E., Jonker, H. R., DeNunzio, N. J., Allen, K. N., Imperiali, B., and Schwalbe, H. (2011) Engineering encodable lanthanide-binding tags into loop regions of proteins, *J Am Chem Soc* 133, 808-819.
- [91] Gaponenko, V., Dvoretzky, A., Walsby, C., Hoffman, B. M., and Rosevear, P. R. (2000) Calculation of z-coordinates and orientational restraints using a metal binding tag, *Biochemistry* 39, 15217-15224.
- [92] Keizers, P. H., and Ubbink, M. (2011) Paramagnetic tagging for protein structure and dynamics analysis, *Progress in nuclear magnetic resonance spectroscopy* 58, 88-96.
- [93] Hass, M. A., and Ubbink, M. (2014) Structure determination of protein-protein complexes with long-range anisotropic paramagnetic NMR restraints, *Curr Opin Struct Biol* 24, 45-53.

- [94] Yagi, H., Pilla, K. B., Maleckis, A., Graham, B., Huber, T., and Otting, G. (2013) Three-dimensional protein fold determination from backbone amide pseudocontact shifts generated by lanthanide tags at multiple sites, *Structure* 21, 883-890.
- [95] Liu, W. M., Skinner, S. P., Timmer, M., Blok, A., Hass, M. A., Filippov, D. V., Overhand, M., and Ubbink, M. (2014) A two-armed lanthanoid-chelating paramagnetic NMR probe linked to proteins via thioether linkages, *Chemistry* 20, 6256-6258.
- [96] Su, X. C., Man, B., Beeren, S., Liang, H., Simonsen, S., Schmitz, C., Huber, T., Messerle, B. A., and Otting, G. (2008) A dipicolinic acid tag for rigid lanthanide tagging of proteins and paramagnetic NMR spectroscopy, *J Am Chem Soc* 130, 10486-10487.
- [97] Li, J., Pilla, K. B., Li, Q., Zhang, Z., Su, X., Huber, T., and Yang, J. (2013) Magic angle spinning NMR structure determination of proteins from pseudocontact shifts, *J Am Chem Soc* 135, 8294-8303.
- [98] Nadaud, P. S., Helmus, J. J., Kall, S. L., and Jaroniec, C. P. (2009) Paramagnetic ions enable tuning of nuclear relaxation rates and provide long-range structural restraints in solid-state NMR of proteins, *J Am Chem Soc* 131, 8108-8120.
- [99] Wang, S., Munro, R. A., Shi, L., Kawamura, I., Okitsu, T., Wada, A., Kim, S. Y., Jung, K. H., Brown, L. S., and Ladizhansky, V. (2013) Solid-state NMR spectroscopy structure determination of a lipid-embedded heptahelical membrane protein, *Nature methods* 10, 1007-1012.
- [100] Wang, S., Munro, R. A., Kim, S. Y., Jung, K. H., Brown, L. S., and Ladizhansky, V. (2012) Paramagnetic relaxation enhancement reveals oligomerization interface of a membrane protein, *J Am Chem Soc* 134, 16995-16998.
- [101] Gottstein, D., Reckel, S., Dotsch, V., and Guntert, P. (2012) Requirements on paramagnetic relaxation enhancement data for membrane protein structure determination by NMR, *Structure* 20, 1019-1027.

- [102] Liang, B., Bushweller, J. H., and Tamm, L. K. (2006) Site-directed parallel spin-labeling and paramagnetic relaxation enhancement in structure determination of membrane proteins by solution NMR spectroscopy, *J Am Chem Soc* 128, 4389-4397.
- [103] Su, Y., Hu, F., and Hong, M. (2012) Paramagnetic Cu(II) for probing membrane protein structure and function: inhibition mechanism of the influenza M2 proton channel, *J Am Chem Soc* 134, 8693-8702.
- [104] Yeo, K. J., Kim, H. Y., Kim, Y. P., Hwang, E., Kim, M. H., Cheong, C., Choe, S., and Jeon, Y. H. (2010) Rapid exploration of the folding topology of helical membrane proteins using paramagnetic perturbation, *Protein Sci* 19, 2409-2417.
- [105] Page, R. C., Lee, S., Moore, J. D., Opella, S. J., and Cross, T. A. (2009) Backbone structure of a small helical integral membrane protein: A unique structural characterization, *Protein Sci* 18, 134-146.
- [106] Cohen, L. S., Arshava, B., Neumoin, A., Becker, J. M., Guntert, P., Zerbe, O., and Naider, F. (2011) Comparative NMR analysis of an 80-residue G protein-coupled receptor fragment in two membrane mimetic environments, *Biochim Biophys Acta* 1808, 2674-2684.
- [107] Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C., and Schultz, P. G. (1989) A general method for site-specific incorporation of unnatural amino acids into proteins, *Science* 244, 182-188.
- [108] Liu, C. C., and Schultz, P. G. (2010) Adding new chemistries to the genetic code, *Annu Rev Biochem* 79, 413-444.
- [109] Young, T. S., Ahmad, I., Yin, J. A., and Schultz, P. G. (2010) An enhanced system for unnatural amino acid mutagenesis in *E. coli*, *J Mol Biol* 395, 361-374.
- [110] Bodenhausen, G., and Ruben, D. J. (1980) Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy, *Chemical Physics Letters* 69, 185-189.

- [111] Loh, C. T., Ozawa, K., Tuck, K. L., Barlow, N., Huber, T., Otting, G., and Graham, B. (2013) Lanthanide tags for site-specific ligation to an unnatural amino acid and generation of pseudocontact shifts in proteins, *Bioconjug Chem* 24, 260-268.
- [112] Liu, Z., Gong, Z., Guo, D. C., Zhang, W. P., and Tang, C. (2014) Subtle dynamics of holo glutamine binding protein revealed with a rigid paramagnetic probe, *Biochemistry* 53, 1403-1409.
- [113] Jones, D. H., Cellitti, S. E., Hao, X., Zhang, Q., Jahnz, M., Summerer, D., Schultz, P. G., Uno, T., and Geierstanger, B. H. (2010) Site-specific labeling of proteins with NMR-active unnatural amino acids, *J Biomol NMR* 46, 89-100.
- [114] Cellitti, S. E., Jones, D. H., Lagpacan, L., Hao, X., Zhang, Q., Hu, H., Brittain, S. M., Brinker, A., Caldwell, J., Bursulaya, B., Spraggon, G., Brock, A., Ryu, Y., Uno, T., Schultz, P. G., and Geierstanger, B. H. (2008) In vivo incorporation of unnatural amino acids to probe structure, dynamics, and ligand binding in a large protein by nuclear magnetic resonance spectroscopy, *J Am Chem Soc* 130, 9268-9281.
- [115] Deiters, A., Geierstanger, B. H., and Schultz, P. G. (2005) Site-specific in vivo labeling of proteins for NMR studies, *Chembiochem* 6, 55-58.
- [116] Nguyen, T. H., Ozawa, K., Stanton-Cook, M., Barrow, R., Huber, T., and Otting, G. (2011) Generation of pseudocontact shifts in protein NMR spectra with a genetically encoded cobalt(II)-binding amino acid, *Angew Chem Int Ed Engl* 50, 692-694.
- [117] Zhang, W. H., Otting, G., and Jackson, C. J. (2013) Protein engineering with unnatural amino acids, *Curr Opin Struct Biol* 23, 581-587.
- [118] Fleissner, M. R., Brustad, E. M., Kalai, T., Altenbach, C., Cascio, D., Peters, F. B., Hideg, K., Peucker, S., Schultz, P. G., and Hubbell, W. L. (2009) Site-directed spin labeling of a genetically encoded unnatural amino acid, *Proc Natl Acad Sci U S A* 106, 21637-21642.

- [119] Schmidt, M. J., Borbas, J., Drescher, M., and Summerer, D. (2014) A genetically encoded spin label for electron paramagnetic resonance distance measurements, *J Am Chem Soc* 136, 1238-1241.
- [120] Soroka, K., Vithanage, R. S., Phillips, D. A., Walker, B., and Dasgupta, P. K. (1987) Fluorescence Properties of Metal-Complexes of 8-Hydroxyquinoline-5-Sulfonic Acid and Chromatographic Applications, *Analytical Chemistry* 59, 629-636.
- [121] Cross, T. A., Sharma, M., Yi, M., and Zhou, H. X. (2011) Influence of solubilizing environments on membrane protein structures, *Trends Biochem Sci* 36, 117-125.
- [122] Ding, Y., Fujimoto, L. M., Yao, Y., Plano, G. V., and Marassi, F. M. (2015) Influence of the lipid membrane environment on structure and activity of the outer membrane protein Ail from *Yersinia pestis*, *Biochim Biophys Acta* 1848, 712-720.
- [123] Acharya, R., Carnevale, V., Fiorin, G., Levine, B. G., Polishchuk, A. L., Balannik, V., Samish, I., Lamb, R. A., Pinto, L. H., DeGrado, W. F., and Klein, M. L. (2010) Structure and mechanism of proton transport through the transmembrane tetrameric M2 protein bundle of the influenza A virus, *Proc Natl Acad Sci U S A* 107, 15075-15080.
- [124] DiMaio, D. (2014) Viral miniproteins, *Annu Rev Microbiol* 68, 21-43.
- [125] Tropea, J. E., Cherry, S., and Waugh, D. S. (2009) Expression and purification of soluble His(6)-tagged TEV protease, *Methods Mol Biol* 498, 297-307.
- [126] Casagrande, F., Maier, K., Kiefer, H., Opella, S. J., and Park, S. H. (2011) Expression and Purification of G-Protein-Coupled Receptors for Nuclear Magnetic Resonance Structural Studies, In *Production of Membrane Proteins: Strategies for Expression and Isolation* (Robinson, A. S., Ed.), pp 297-316, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany.
- [127] Denisov, I. G., Grinkova, Y. V., Lazarides, A. A., and Sligar, S. G. (2004) Directed self-assembly of monodisperse phospholipid bilayer Nanodiscs with controlled size, *J Am Chem Soc* 126, 3477-3487.

- [128] Hagn, F., Etzkorn, M., Raschle, T., and Wagner, G. (2013) Optimized phospholipid bilayer nanodiscs facilitate high-resolution structure determination of membrane proteins, *J Am Chem Soc* 135, 1919-1925.
- [129] Grinkova, Y. V., Denisov, I. G., and Sligar, S. G. (2010) Engineering extended membrane scaffold proteins for self-assembly of soluble nanoscale lipid bilayers, *Protein Eng Des Sel* 23, 843-848.
- [130] Triba, M. N., Warschawski, D. E., and Devaux, P. F. (2005) Reinvestigation by phosphorus NMR of lipid distribution in bicelles, *Biophys J* 88, 1887-1901.
- [131] Rajarathnam, K., Kay, C. M., Clark-Lewis, I., and Sykes, B. D. (1997) Characterization of quaternary structure of interleukin-8 and functional implications, *Methods in enzymology* 287, 89-105.