Development and Optimization of a Human Collagen-Targeted Protein-based MRI Contrast Agent for Early Detection of Chronic Human Diseases

Oluwatosin Ibhagui

Follow this and additional works at: https://scholarworks.gsu.edu/chemistry_diss

Recommended Citation
doi: https://doi.org/10.57709/24161345

This Dissertation is brought to you for free and open access by the Department of Chemistry at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Chemistry Dissertations by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.
Development and Optimization of a Human Collagen-Targeted Protein-based MRI Contrast Agent for Early Detection of Chronic Human Diseases

by

Oluwatosin Ibhagui

Under the Direction of Jenny J. Yang, Ph.D.

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in the College of Arts and Sciences

Georgia State University

2021
ABSTRACT
Chronic diseases are leading cause of death in US and globally. About 60% of adults in the US have at least one form of chronic disease while 40% suffer from 2 or more chronic diseases. Noninvasive, early detection, quantitative diagnosis, and staging of these diseases with high-resolution imaging remain a pressing unmet medical need. Type I collagen, a major constituent of extracellular matrix, is highly expressed in the microenvironment of various chronic and acute human diseases such as hepatocellular carcinoma and liver and lung fibrosis. Its expression level and spatial crosslink pattern are stage-dependent and is an attractive diagnostic and therapeutic target for many chronic diseases via pathological analysis. This dissertation reports the design and optimization of a human collagen-targeted protein-based MRI contrast agent (hProCA32.collagen) to extend MRI analysis for diagnosis and staging of liver and lung diseases using detection of overexpressed collagen biomarker using several mouse models. hProCA32.collagen exhibits strong collagen type I binding affinity and specificity over types III and IV. hProCA32.collagen has $10^4$ to $10^{11}$ metal selectivity for Gd$^{3+}$ over Ca$^{2+}$ and Zn$^{2+}$, respectively. It exhibits much stronger resistance against transmetallation than linear contrast agents. Importantly, hProCA32.collagen exhibits the high relaxivity values for both $r_1 (32 \pm 0.3 \text{ mM}^{-1}.\text{s}^{-1})$ and $r_2 (51 \pm 0.2 \text{ mM}^{-1}.\text{s}^{-1})$ per Gd$^{3+}$ at 1.4 T and $r_1 (18.5 \pm 1.5 \text{ mM}^{-1}.\text{s}^{-1})$ and $r_2 (105.6 \pm 2 \text{ mM}^{-1}.\text{s}^{-1})$ at 7.0 T that is 8 to 11-fold greater than clinically approved contrast agents Eovist and Gadovist. hProCA32.collagen enabled the early detection of idiopathic lung fibrosis (Ashcroft 2 of 8) as well as heterogeneously expressed UIP patterns validated by histology analysis and correlation. In addition, collagen mapping in a nicotine-induced COPD mouse model was achieved with hProCA32.collagen. hProCA32.collagen enabled the progressive and heterogeneous detection of liver diseases from an early-stage diabetes mice
model (Ishak 1 of 6) and late-stage (Ishak 5 of 6) mouse model to hepatocellular carcinoma.

The developed hProCA32.collagen is expected to overcome the major clinical barriers in early disease detection and staging with strong translational potential.

INDEX WORDS: MRI Contrast Agents, Fibrosis, Chronic Diseases, Fibrosis, Collagen Type I, Idiopathic Pulmonary Fibrosis
Development and Optimization of a Human Collagen-Targeted Protein-based MRI Contrast Agent for Early Detection of Chronic Human Diseases

by

Oluwatosin Ibhagui

Committee Chair: Jenny J. Yang

Committee: Ming Luo
Zhi-ren Liu.
Hans E. Grossniklaus

Electronic Version Approved:

Office of Graduate Services
College of Arts and Sciences
Georgia State University
August 2021
DEDICATION

This work is dedicated to God, my exceeding joy, the author and the finisher of my faith for proving to me beyond any reasonable doubt that He is my source and in charge of making ALL things beautiful in His time. To my father late Senior Evangelist Samuel Akinwunmi Odubade for instilling in me the “can do” spirit and that any door can be walked into. I hope you are proud of your little girl. To my beautiful mother Deaconess Juliana Odubade for her prayers, support and encouragement every step along the way; thank you for leaving everything to come help out when needed. To my husband, Eimienwanlan Ibhagui, for always giving me a voice, a listening ear, reading through my dissertation and for all his sacrifices just to make this happen. To my boys Jonathan and Joel Ibhagui for being the most beautiful part of my days and for always giving me a reason to try again no matter what. To my wonderful siblings and siblings-in-laws for trusting me, for their constant cheering, prayers, love and encouragement. To my mother-in-law for accepting me totally and coming over to help and the constant prayers. To my friend turned brother; Dr. Abiodun Anifowose for his support, advice and help in this journey. To my sweet friend, Opeyemi Aladesaye for her love, encouragement and support. To the “Magnolia men” for giving me a soft landing in the USA. To the Redeemed Christian Church of God, Family Praise Chapel for being my family and a home away from home, cheering me on and praying for me. And lastly, to my friends for always being so supportive and ready to pray with me. You all are a true definition of a support system.
ACKNOWLEDGEMENTS

Foremost, I would like to express my most sincere and heartfelt appreciation to Dr. Jenny J. Yang, my Ph.D. and life advisor! She has indeed been a blessing to me; she is absolutely the best advisor any graduate student could ever have. I am indeed grateful for the opportunity to join her lab in 2015 with no background whatsoever in biochemistry. I truly appreciate her experience and patience in grooming me, providing an environment where I could challenge myself and be better every single time. Her zeal and multitasking capability were a real source of inspiration for me to balance family life with research. Her ability to see the potentials and capability of everyone under her mentorship and nurture that potential to maturity is incredible. Her knowledge, experience, fighting spirit and passion for science is unparalleled. She not only played an essential role in developing my career and making me a detailed and well-grounded scientist, she also helped me become a better communicator with troubleshooting capability. I am so grateful for her guidance and advice during my research and for providing both moral and financial support for moving my research forward.

I would also like to extend my gratitude to previous members and mentors that helped me in the lab over the years. Corrie Purser who was my first mentor when I joined Dr. Yang’s lab and helped me navigate transitioning into a new field. Corrie was part of the MRI group and I learned some of the basic techniques such as expression and purification of proteins. I am indeed grateful to Dr. Mani Salarian for his detailed and incredible mentorship while teaching me all the important techniques and skills required for conducting research in MRI projects. His continued help, guidance, and support during my research has been valuable for my project advancement. Dr. Jingjuan Qiao has also been one of my most approachable mentors providing detailed suggestions both in *in vitro* and *in vivo* skills during all these years. Their mentorship was pivotal to the
completion of this work. I thank Dr. Shenghui Xue for his detailed guidance especially with lung fibrosis project, showing me how to see the big picture in the data processing stage. I am also very thankful to Dr. Shanshan Tan for her help during my research. My deepest gratitude goes to Dr. Bin Dong and Yiting Xu for helping with image processing and statistical analysis. And to all current and previous members of MRI group Dr. Florence Reddish for our friendship and our scientific/insightful discussions on several formulation and purification processes and for her great help with my dissertation, Zongxiang Gui, Dongjun Li, and Brenda-Ruth Nissan Mimba for their help during my research. I would also like to appreciate all our collaborators’ help in these multi-disciplinary projects; Drs. Ravi Chakra Turaga, Hua Yang, Malvika Sharma, Hongwei Han, Guangda Peng and Ganesh Satyanarayana as well as Falguni Mishra, for histology analysis and animal models’ development; Drs. Khan Hekmatyar, Yuguang Meng and Dr. Philip Sun for MR image acquisition and processing while also providing helpful suggestions; Drs. Alton Brad Farris and Gabriel Sica and Hua Yang for their pathology analysis and helpful suggestions towards the liver fibrosis/HCC and lung fibrosis project respectively. I really appreciate Dr. Hang Shi, Dr. Bingzhong Xue and Shirong Wang for helping to generate diabetes nonalcoholic fatty liver disease animal models. I also thank Dr. Rao Mikkavilli for his help in analyzing pharmacokinetics/pharmacodynamics data. I would also like to really appreciate Dr. Xiangming Ji and Maureen Beebe for their tremendous help in generating the nicotine-induced COPD mice model and Dr. Zuben Sauna for immunogenicity studies. Finally, I would like to express my heartfelt gratitude to all my committee members, Drs. Ming Luo, Hans E. Grossniklaus and Zhi-Ren Liu for accepting my invitation to be part of my dissertation committee without hesitation and for their helpful suggestions, detailed analysis, and assistance for supporting my research and help in completing this work. Thank you all so much for making this possible.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS V

LIST OF TABLES XVII

LIST OF FIGURES XIX

LIST OF ABBREVIATIONS XXXVII

1 INTRODUCTION 1

1.1 Chronic Human Diseases 1

1.2 Magnetic Resonance Imaging and its advantage as an imaging technique 5

1.2.1 Atomic Nuclei 6

1.2.2 Magnetic field 6

1.2.3 Radiofrequency Energy 7

1.2.4 MRI Pulse Sequence 8

1.2.5 MR image quality 9

1.3 MRI contrast Agents 10

1.3.1 Current MRI contrast agents and their limitations 15

1.3.2 MRI Relaxivity Theory 19

1.3.3 T1 and T2 Relaxation and factors contribute to relaxivity of contrast agents 22

1.3.4 Protein-based MRI contrast agents 29

1.4 Protein MRI contrast agent by design 32

1.4.2 High Relaxation Properties of ProCAs 35
1.5 Molecular Magnetic resonance Imaging by biomarker-targeted contrast agents

1.5.1 Role of Collagen as disease diagnosis biomarker

1.6 Current advances in the development of collagen-targeting peptides

1.7 Previous work done by Yang lab for collagen-Targeted ProCAs (ProCA32.collagen)

1.8 Motivation and Questions to be addressed

1.9 Overview of this Dissertation

2 MATERIALS AND METHODS

2.1 Cloning, plasmid generation for hProCA32.cys and hProCA32.collagen

2.2 Transformation and Expression of hProCA32.cys and hProCA32.collagen

2.3 Optimized transformation and Expression Protocol for hProCA32.collagen

2.3.1 Clone selection/Screening

2.3.2 Bacterial expression

2.3.3 Glycerol stock generation

2.3.4 Expression protocol for hProCA32.collagen from glycerol stock

2.4 Purification of hProCA32.cys and hProCA32.collagen

2.4.1 Laboratory Scale (shake flask)

2.4.2 Laboratory Scale purification using ammonium sulfate precipitation

2.5 PEGylation of Protein contrast agents

2.5.1 Lysine PEGylation
2.5.2 Optimized lysine PEGylation Procedure 65

2.5.3 Cysteine PEGylation 66

2.5.4 N-terminal PEGylation of hProCA32.collagen 66

2.6 Protein Concentration Determination 67

2.7 Relaxivity Studies 67

2.8 Gd\textsuperscript{3+}, Tb\textsuperscript{3+}, Ca\textsuperscript{2+}, and Zn\textsuperscript{2+} -Binding Affinity Determination 68

2.8.1 Calcium Titration 68

2.8.2 Terbium Titration 70

2.8.3 Zinc Titration 71

2.8.4 Gadolinium Titration 72

2.8.5 Determination of Free Gadolinium by xylenol orange method 72

2.8.6 Circular dichroism spectroscopy 73

2.8.7 Residual Calcium Measurement in Working Buffers 73

2.8.8 Transmetallation Studies 74

2.8.9 Gd\textsuperscript{3+} Bio distribution studies in different organs 76

2.8.10 Serum Stability Studies 76

2.8.11 Pharmacokinetics and Pharmacodynamic studies of hProCA32.collagen 76

2.9 MRI Scan 77

2.9.1 3-D ultrashort echo time (3D- UTE) 77

2.9.2 Rapid Acquisition with Refocused Echo (RARE) 77
2.9.3 Multislice spin Multi echo

3 DEVELOPMENT OF HUMAN PROTEIN-BASED MRI CONTRAST AGENT FOR LUNG CANCER IMAGING

3.1 Introduction

3.1.1 Lung cancer detection by noninvasive imaging modality

3.1.2 Current diagnostic methods of lung cancer and their limitation

3.1.3 Criteria and current status of development of protein-based MRI contrast agent for lung MR Imaging

3.1.4 Objective and overview of this chapter

3.2 Expression, Purification, and application of human protein contrast agents for MR Imaging

3.3 Expression and Purification of hProCA32.cys

3.4 Purification of hProCA32.cys

3.5 Cysteine PEGylation of hProCA32.cys

3.6 FPLC Characterization of PEGylated hProCA32.cys

3.6.1 Separation of PEGylated and UnPEGylated hProCA32.cys using the mono Q-column

3.6.2 Further Optimization PEGylation

3.7 Mass Spectrometric Analysis of purified hProCA32.cys

3.8 Relaxivity Measurement of PEGylated hProCA32.cys

3.8.1 Confirmation of hProCA32.cys binding ratio.
3.8.2 Characterization of hProCA32.cys using Circular dichroism (CD) 106

3.8.3 Estimation of residual calcium in purification buffer of hProCA32.cys 107

3.9 Metal binding studies of hProCA32.cys via fluorescence spectroscopy 109

3.10 Optimization of Purification protocol of hProCA32.cys 111

3.11 T-cell activation and Stability Studies of hProCA32.Cys 121

3.12 Screening and Solvent accessibility studies of FFQMVGLK 125

3.13 Transmetallation studies of hProCA32.cys 129

3.14 Gd$^{3+}$ Biodistribution in Different Organs 131

3.15 Development of KP mouse model for Lung Cancer for lung imaging 134

3.16 Magnetic Resonance Imaging of lung cancer using hProCA32.cys 134

3.17 Summary 138

3.18 Limitations of this work 139

4 DESIGN AND DEVELOPMENT OF HUMAN COLLAGEN-TARGETED PROTEIN CONTRAST AGENT (HPROCA32.COLLAGEN) FOR MR IMAGING 140

4.1.1 Introduction 140

4.2 Results and discussion 143

4.2.1 Rationale for designing variants for collagen targeted contrast agents 143

4.2.2 Expression, Purification and characterization of hProCA32c.collagen1 144

4.2.3 Expression, Purification and characterization of hProCA32.collagen1C 148

4.2.4 Expression, Purification and characterization of hProCA32.collagen1 151
4.2.5 Transformation and expression of newly developed collagen-targeting peptides 153

4.3 ELISA Assay for the developed hProCA32.collagen variants 156

4.3.1 Summary, conclusion and limitations in the selection of hProCA32.collagen1 as the new human collagen-targeted contrast agent 158

4.3.2 Future studies 159

5 FURTHER PRODUCTION OPTIMIZATION OF HPROCA32.COLLAGEN 159

5.1 Introduction 159

5.2 Materials and Methods 162

5.2.1 Collagen binding studies of hProCA32.collagen 162

5.3 Results and discussion 164

5.3.1 Glycerol stock generation 164

5.3.2 Laboratory Scale (shake flask) purification of hProCA32.collagen 170

5.3.3 Relaxivity Studies 172

5.3.4 Purification of hProCA32.collagen using Dasgip Parallel Bioreactor Systems 173

5.4 Optimization expression and purification with rhamannose induced expression vector 178

5.5 Comparison of collagen binding affinity of hProCA32.collagen purified from expression with old and new vectors 178

5.6 Optimization of Purification with fermentation facility (BFF) 180

5.7 Optimization of hProCA32.collagen purification 186
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.7.1</td>
<td>Probing the cause of aggregation in hProCA32.collagen</td>
<td>194</td>
</tr>
<tr>
<td>5.8</td>
<td>Probing the conformational requirement for collagen binding of the disulfide bond in the collagen binding peptide of hProCA32.collagen</td>
<td>215</td>
</tr>
<tr>
<td>5.8.1</td>
<td>Formulation of hProCA32.collagen by controlling oligomerization</td>
<td>222</td>
</tr>
<tr>
<td>5.8.2</td>
<td>Relaxivity and collagen binding studies of purified protein</td>
<td>234</td>
</tr>
<tr>
<td>5.8.3</td>
<td>N-terminal PEGylation of hProCA32.collagen</td>
<td>237</td>
</tr>
<tr>
<td>5.8.4</td>
<td>Tissue biodistribution and toxicity Test</td>
<td>240</td>
</tr>
<tr>
<td>5.8.5</td>
<td>Tricks and tips for Loading Gadolinium to unPEGylated hProCA32.collagen</td>
<td>243</td>
</tr>
<tr>
<td>5.9</td>
<td>Major conclusions, Challenges, and future work</td>
<td>247</td>
</tr>
<tr>
<td>6</td>
<td>APPLICATION OF HPROCA32.COLLAGEN IN EARLY THE DIAGNOSIS OF CHRONIC LUNG DISEASES</td>
<td>248</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>248</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Idiopathic Pulmonary Fibrosis (IPF)</td>
<td>250</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Animal model for IPF</td>
<td>252</td>
</tr>
<tr>
<td>6.1.3</td>
<td>Application and limitation of MRI in detection of IPF</td>
<td>257</td>
</tr>
<tr>
<td>6.1.4</td>
<td>Role of hyperpolarized MRI in IPF</td>
<td>258</td>
</tr>
<tr>
<td>6.1.5</td>
<td>Overview of this chapter</td>
<td>259</td>
</tr>
<tr>
<td>6.2</td>
<td>Results and discussion</td>
<td>260</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Intraperitoneal (IP) injection of bleomycin recapitulates the clinical features of IPF</td>
<td>262</td>
</tr>
</tbody>
</table>
6.2.2 Molecular imaging of IP injected bleomycin for idiopathic pulmonary fibrosis (IPF) imaging 263

6.2.3 Analysis of normal and bleomycin (BM)-induced IPF mice by conventional T1 and T2 sequence. 264

6.2.4 MR Imaging of normal and BM-induced IPF mice by T1-weighted rapid acquisition with refocused echo (RARE) and 3D ultra-short echo time (3D UTE). 268

6.2.5 3D-UTE analysis of normal and bleomycin (BM)-induced IPF mice 272

6.2.6 Detection of Usual interstitial pneumonia (UIP) pattern by hProCA32.collagen 280

6.2.7 MR Imaging of nicotine-induced lung fibrosis in mice 284

6.3 Comparison of imaging findings for IPF and COPD 290

6.4 Comparison with lung and liver enhancement, half life 291

6.4.1 Discussion of challenges in lung imaging 293

6.5 Major conclusion, limitations, and future direction 295

7 APPLICATION OF HPROCA32.COLLAGEN IN EARLY AND PROGRESSIVE DIAGNOSIS OF LIVER FIBROSIS 298

7.1 Introduction 298

7.1.1 The Human liver and liver disease 298

7.1.2 Fatty liver diseases 299

7.1.3 Role of obesity in NAFLD development 301
7.1.4 Organ Fibrosis

7.1.5 Mice Models for the development of NAFLD

7.1.6 Molecular Imaging of NAFLD and their limitations

7.1.7 Overview of this Chapter

7.2 METHODS

7.2.1 Development of Mouse model for NASH study using class I histone deacetylases (HDACs)

7.2.2 Development of Mouse model for early (minimal fibrosis) and Late stage (Cirrhosis) liver fibrosis

7.2.3 Development of Mouse model for Advance Liver fibrosis

7.3 RESULTS

7.4 MR Imaging of CGI-58 mouse model

7.4.1 MR Imaging of HDAC1 NASH/liver fibrosis model

7.5 Comparison of disease state and MRI observation of HDAC1 vs CGI-58 model

7.5.1 The effect of PEGylation on liver fibrosis imaging

7.6 Organ Harvesting and Analysis

7.7 Conclusion

8 APPLICATION OF HPROCA32.COLLAGEN IN PRECISION MRI (PMRI)

DETECTION OF HEPATOCELLULAR CARCINOMA (HCC)

8.1 Introduction
8.1.1 The role of cirrhosis in HCC development 338
8.1.2 The role of hepatitis B virus in HCC development 339
8.1.3 The role of hepatitis C virus in HCC development 339
8.1.4 Cancer-associated fibroblasts in hepatocellular carcinoma 339
8.1.5 HCC classification 340
8.1.6 Challenges Associated with HCC Diagnosis 341
8.2 Limitation associated with current clinical contrast agents 343
8.2.1 Overview of this chapter 344
8.3 Mice model 345
8.3.1 MR imaging of DEN-induced HCC after 12 weeks using rProCA32.collagen 346
8.3.2 MR imaging of DEN-induced HCC after 44 weeks. 352
8.3.3 Development of Diet/chemical-induced HCC 355
8.3.4 MR imaging of Diet/chemical-induced HCC 356
8.4 Conclusion 360
9 CONCLUSIONS, SIGNIFICANCE AND FUTURE DIRECTION 361
REFERENCES 369
LIST OF TABLES

Table 1-1 clinically approved gadolinium-based MRI contrast agents. ........................................ 18

Table 1-2: The collagen superfamily.[201, 206, 211-214]. (FACIT fibril-associated collagens with interrupted triple helices, MACIT membrane-associated collagens with interrupted triple helices, MULTIPLEXINs multiple triple-helix domains and interruptions) .......... 43

Table 1-3: Collagen-binding peptides and their applications. .............................................................. 53

Table 3-1: First Stability Constant Logarithms of selected ligands to metal ions [292] ...... 97

Table 3-2: Comparison of relaxivity values of hProCA32.cys with rProCA32 and clinical contrast agents .................................................................................................................................. 98

Table 3-3 Percentage composition of free gadolinium after separation by G-25 sephadex desalting column. ........................................................................................................................................ 105

Table 3-4: Solvent accessible surface area of residues with high promiscuity in hProCA32.cys ........................................................................................................................................ 129

Table 4-1: Summary of biophysical characteristics of hProCA32.collagen1 variants ......... 157

Table 5-1: Summary of the weight of pellet and other expression condition ....................... 166

Table 5-2: Effect of Reducing Agent on dimer band at different purification steps by under different conditions using BME and DTT ........................................................................................................ 226

Table 5-3: Comparison of Longitudinal and transverse relaxation rates of some commercially available gadolinium-based contrast agents with hProCA32.collagen and ProCA32.collagen at 37°C. ......................................................................................................................... 246

Table 5-4: Comparison of stability constants of some commercially available gadolinium-based contrast agents with hProCA32.collagen and ProCA32.collagen at 25°C. ...................... 246
LIST OF FIGURES

Figure 1-1: World Health Organization (WHO) estimate of the top ten global leading cause of death. Modified after [7].................................................................................................................................................. 2

Figure 1-2: Electromagnetic Spectrum with different range of electromagnetic radiation .......... 8

Figure 1-3: Electronic configuration Gd$^{3+}$ Metal depicting the presence of half-filled f-orbital responsible for the paramagnetic properties of gadolinium metal................................................. 12

Figure 1-4: Coordination polyhedron of Gd$^{3+}$ showing A. Molecular stick models of Gd(H2O)$_8^{3+}$ and B. Gd(H$_2$O)DTPA$^2$ modified after [65]. C. Schematic diagram of factors contributing to the relaxivity of GBCAs................................................................. 13

Figure 1-5: Chemical structure of the clinically approved gadolinium-based contrast agents with their chemical and commercial names. Modified after [74]......................................................... 16

Figure 1-6: Magnetic properties of atomic nuclei............................................................................. 21

Figure 1-7: Mechanism of T1 and T2 relaxation........................................................................... 23

Figure 1-8 EF-Hand Ca$^{2+}$ binding motif with the characteristics preferred residue[169] In addition to preferential oxygen binding, a preference is also made for negatively charged residues Aspartate and glutamate. Parvalbumin contains three paired helix-loop-helix EF-hand motifs AB motifs houses a two amino-acid deletion in the loop region......... 34

Figure 1-9: Component of the extracellular matrix. ....................................................................... 41

Figure 1-10 Pathogenesis of fibrosis in different organs................................................................. 45

Figure 1-11: Interaction of fibrillar collagen with core decorin protein, modified after [242] .... 49

Figure 2-1 Transmetallation experiment of ProCAs using Zn$^{2+}$ competition in phosphate buffer solution................................................................................................................................. 75

Figure 3-1 Expression of hProCA32.cys from Pet22b vector ............................................................. 84
Figure 3-2: SDS-PAGE characterization of step-by-step purification of hProCA32.cys............ 85
Figure 3-3 FPLC chromatogram and characterization of hProCA32-Cys................................. 86
Figure 3-4 SDS-PAGE characterization of hProCA32.cys under reduced conditions.............. 87
Figure 3-5 FPLC characterization of PEGylated and unPEGylated hProCA32.cys ................. 89

Figure 3-6 FPLC Characterization of PEGylated and UnPEGylated hProCA32.cys using the
mono Q-column......................................................................................................................... 90

Figure 3-7: Separation of PEGylated and unPEGylated hProCA32.collagen by FPLC.......... 91

Figure 3-8 Characterization of PEGylated hProCA32.Cys. .................................................... 93

Figure 3-9: Matrix-Assisted Laser Desorption/Ionization (MALDI) spectrum of hProCA32.Cys.
.................................................................................................................................................. 94

Figure 3-10: Relaxivity plot of hProCA32.cys......................................................................... 95

Figure 3-11 EGTA screening for calcium removal from hProCA32.cys................................. 96

Figure 3-12: Equilibrium dialysis system set up of hProCA32.cys in 20% chelex-100. ........... 98

Figure 3-13 R1 and R2 relaxivity plot of hProCA32.cys......................................................... 99

Figure 3-14 Relaxivity studies of hProCA32.cys................................................................. 100

Figure 3-15 Job plot of Relaxivity values for hProCA32.cys.................................................. 101

Figure 3-16 Spectrometric estimation of free [Gd^{3+}] in a 2:1 Gd^{3+} loaded hProCA32.cys..... 102

Figure 3-17 Fluorometric method of estimating free Gd^{3+} with Rhod-5N dye using different
binding ratio.................................................................................................................................. 103

Figure 3-18: Chemical structure of some calcium indicators.................................................. 104

Figure 3-19: FPLC Characterization and binding ration of hProCA32.cys with different ratios of
Gadolinium bound with a desalting column via gel filtration. .............................................. 105

Figure 3-20: Circular Dichroism spectra of hProCA32.cys prepared in 10mM HEPES buffer. 107
Figure 3-21: Characterization and relaxivity studies of calcium-loaded and calcium-free hProCA32.cys. ................................................................. 108

Figure 3-22 Scheme of Trp-Tb$^{3+}$ FRET and Gd$^{3+}$ Competition Assay................................. 109

Figure 3-23: Tb$^{3+}$, Zn$^{2+}$, Ca$^{2+}$ and Gd$^{3+}$ Titration of hProCA32.Cys................................. 110

Figure 3-24: Hi-Trap™ HIC Selection kit for rapid screening of hProCA32.cys purification. . 112

Figure 3-25: FPLC characterization of hProCA32.cys after different HIC column purification
    showing lack of binding.......................................................................................... 113

Figure 3-26: Addition of ammonium sulfate to protein for HIC purification shows precipitation
due to exposure of hydrophobic surface of the protein.............................................. 114

Figure 3-27: HIC purification of hProCA32.cys with modified protocol. ......................... 115

Figure 3-28: Screening of HIC column after pH adjustment............................................. 116

Figure 3-29: Ammonium sulfate purification and hydrophobic interaction chromatographic
    studies of hProCA32.cys...................................................................................... 119

Figure 3-30: βME reduced SDS-PAGE of hProCA32.cys after HIC chromatography......... 120

Figure 3-31: Absorbance spectra of different fractions obtained during the purification of
    hProCA32.cys........................................................................................................ 120

Figure 3-32: Cell proliferation monitoring by cell flow cytometry..................................... 121

Figure 3-33: Heat map created in collaboration with Dr Zuben Sauna at FDA. ................. 123

Figure 3-34: Immunogenic assessment of hProCA32.cys..................................................... 124

Figure 3-35: Solvent accessibility visualization of mutation of residue phenylalanine at position
    30 to cysteine ......................................................................................................... 126

Figure 3-36: Solvent accessibility visualization of mutation of residue phenylalanine at position
    31 to cysteine ......................................................................................................... 126
Figure 3-37: Solvent accessibility visualization of mutation of residue glutamine at position 32 to cysteine. ........................................................................................................................................................................ 127

Figure 3-38 Solvent accessibility visualization of mutation of residue methionine at position 33 to cysteine. ........................................................................................................................................................................................................ 127

Figure 3-39: Solvent accessibility visualization of mutation of residue valine at position 34 to cysteine. ........................................................................................................................................................................................................ 127

Figure 3-40: Solvent accessibility visualization of mutation of residue glycine at position 35 to cysteine. ........................................................................................................................................................................................................ 128

Figure 3-41: Solvent accessibility visualization of mutation of residue lysine at position 36 to cysteine. ........................................................................................................................................................................................................ 128

Figure 3-42: Solvent accessibility visualization of mutation of residue lysine at position 37 to cysteine. ........................................................................................................................................................................................................ 128

Figure 3-43: Gadolinium stability of lysine PEGylated hProCA32.cys (1 hProCA32.cys :1 Gd^{3+}) via transmetallation in phosphate buffer to mimic physiological condition and serum stability studies. Evolution of R1 Relaxation rate at increasing time points with respect to initial R1 value. ......................................................................................................................... 131

Figure 3-44: Clinical toxicity studies of hPoCA32.cys showing no significant toxicity resulting from hProCA32.cys injection into healthy mice compared to PBS injected mice. ........ 132

Figure 3-45: Bio distribution of Gadolinium of different tissues. Long-term distribution of Gd^{3+} in mice after injection of hProCA32.cys ........................................................................................................ 133

Figure 3-46: T1 and T2- weighted lung MR images of KP mice models before and after injection of hProCA32.cys .............................................................................................................................. 136
Figure 3-47 T1 and T2-weighted lung MR SNR of KP mice models before and after injection of hProCA32.cys ................................................................. 136

Figure 3-48 T2-W corona and axial MR images of KP mice M5 with lesions observed post injection correlated with hematoxylin and eosin (H&E) staining showing cell proliferation in observed tumor compared to normal lung tissue. ......................... 137

Figure 3-49 M3 KP mice shows selective MRI enhancement correlating with H&E staining. The areas with high MR intensity were observed to have abnormal lung morphology in H&E. ........................................................................................................... 137

Figure 4-1 Development of hProCA32.collagen for molecular imaging. ......................... 142

Figure 4-2: Expression of hProCA32c.collagen1. ........................................................................ 145

Figure 4-3: FPLC chromatogram of hProCA32c.collagen. ......................................................... 146

Figure 4-4: Characterization of Purified hProCA32c.collagen .................................................. 147

Figure 4-5: Estimation of the relaxation rates of hProCA32c.collagen1 in 10mM HEPES buffer pH8 ........................................................................................................................................ 148

Figure 4-6: Expression of hProCA32.collagen1C. ......................................................................... 149

Figure 4-7: Characterization of purified hProCA32.collagen1C .................................................. 150

Figure 4-8: Estimation of the relaxation rates of hProCA32.collagen1C. 1/T1 plot of hProCA32.collagen1C at 0mM, 0.02mM, 0.04mM, 0.08mM and 0.16 mM concentrations of Gd³⁺ to obtain the R1 relaxation rate B. 1/T2 plot of hProCA32.collagen1C at 0mM, 0.02mM, 0.04mM, 0.08mM and 0.16 mM concentrations of Gd³⁺ to obtain the R2 relaxation rate. The relaxivity reaction was done in 10mM HEPES pH 7.2 at 37°C. .... 150

Figure 4-9: Expression of hProCA32.collagen1 ........................................................................... 151

Figure 4-10: Characterization of purified hProCA32.collagen1C ................................................ 152
Figure 4-11: Relaxivity plots of hProCA32.collagen1 ................................................................. 153

Figure 4-12 Transformation and expression of developed LKK and LRR variants of hProCA32.collagen................................................................. 154

Figure 4-13: FPLC Chromatogram and UV/VIS characterization of hProCA32.collagenLRR 155

Figure 4-14: R1 and R2 relaxation rates of hProCA32.collagen1LRR obtained in 10mM HEPES showed high relaxivity values and potentially high MRI sensitivity ......................... 155

Figure 4-15: Relaxivity and calcium binding studies of developed LKK/LRR. ............... 156

Figure 4-16: ELISA Assay results of developed hProCA32.collagen variants. ............... 157

Figure 5-1: Examples of biologically relevant oxidative states and reaction pathways of cysteine [314]...................................................................................................................... 160

Figure 5-2: Types of ELISA methods. A., Direct ELISA, B., Indirect ELISA, C., optimized indirect ELISA, D., Sandwich ELISA, E., Competitive ELISA ................................. 164

Figure 5-3: Clone screening and selection .................................................................................. 165

Figure 5-4: Clone Selection and optimization in terrific broth medium................................. 167

Figure 5-5: Expression of hProCA32.collagen (clone 5) by BFF at 4 hours and 21 hours post induction with ampicillin showed no difference before and after IPTG induction. ...... 168

Figure 5-6: Expression of hProCA32.collagen (clone 5) by BFF at 4 hours and 21 hours post induction with ampicillin and chloramphenicol antibody shows no obvious expression after IPTG induction. Marker bands from top to bottom (250, 150, 100, 75, 50, 37, 25, 20, 15, and 10 kDa). ........................................................................................................ 168

Figure 5-7: Expression and Purification of hProCA32.collagen (clone 5) by BFF at 4 hours and ~18 hours post induction with ampicillin antibody only. ............................. 169
Figure 5-8: SDS-PAGE characterization of the purification stages of hProCA32.collagen from lysis step to dialysis stage preceding the FPLC separation................................. 171

Figure 5-9: Characterization of lab-purified hProCA32.collagen................................................. 172

Figure 5-10: R1 and R2 Relaxivity measurement of PEGylated hProCA32.collagen at 1.4T in 10mM HEPES pH 7.2.......................................................... 173

Figure 5-11: Expression of hProCA32.collagen (clone 5) by BFF at 4 hours and harvest (21 hours) after induction. Marker bands from top to bottom (250, 150, 100, 75, 50, 37, 25, 20, 15, and 10 kDa)........................................................... 174

Figure 5-12: Reduced expression gels of hProCA32.collagen (clone 5) by BFF at 4 hours and 21 hours post induction with ampicillin and chloramphenicol antibody. Minimal expression of hProCA32.collagen is present. Marker bands from top to bottom (250, 150, 100, 75, 50, 37, 25, 20, 15, and 10 kDa).......................................................... 175

Figure 5-13: Expression and Purification of hProCA32.collagen (clone 5) by BFF at 4 hours and ~18 hours post induction with ampicillin antibiotic only. A and C) Purification of hProCA32.collagen using Q Sepharose column. B) Reduced expression gel of hProCA32.collagen. Marker bands from top to bottom (250, 150, 100, 75, 50, 37, 25, 20, 15, and 10 kDa). D and E) Reduced purification gels of hProCA32.collagen samples from purification done in A and C. Marker bands from top to bottom (250, 150, 100, 75, 50, 37, 25, 20, 15, and 10 kDa).......................................................... 176

Figure 5-14: Confirmation of the purity level of obtained hProCA32.collagen from BFF........ 177

Figure 5-18: ELISA assay for hProCA32. collagen purified in GSU, UGA from old vector and new vector from GSU flask shake purification with and without reducing agent, BME. ........................................................................................................... 179
Figure 5-19: Physical changes of hProCA32.collagen during Purification at BFF, UGA at room temperature. ................................................................. 182

Figure 5-20: SDS-PAGE and FPLC chromatogram of purification steps of hProCA32.collagen1 .............................................................................................................. 183

Figure 5-21: Purification of hProCA32.collagen with close observation of physical changes. A., Clarification of sample into supernatant (SN) and pellet after heating, B. The effect of EGTA addition was observed, no precipitation was observed for samples without EGTA (1), sample with 40mM EGTA (2) and sample with 10mM EGTA (3). C. Buffer exchange and dialysis shows significant amount of precipitation. D., SDS-PAGE characterization of each step of the purification process. Marker bands are 68, 26, 14.3, and 6.5 kDa from top to bottom....................................................................................... 187

Figure 5-22: FPLC Chromatogram of EGTA loaded hProCA32.collagen with silver stained SDS-PAGE characterization ........................................................................................................ 189

Figure 5-23: Characterization of Purified hProCA32.collagen by UV spectroscopy and SDS-PAGE .................................................................................................................................. 190

Figure 5-24: Folding and refolding experiment of hProCA32.collagen ................................................................. 191

Figure 5-25: Calibration Chromatogram of 30ml sephadex size exclusion column with the standardized molecular weight ........................................................................................................ 192

Figure 5-26: SEC analysis of refolded hProCA32.collagen1 post urea-BME treatment ........... 192

Figure 5-27: SEC analysis of denatured hProCA32.collagen1 post second urea treatment..... 193

Figure 5-28: Expression of hProCA32.collagen ........................................................................................................ 194

Figure 5-29: Flowchart Representation of the new conditions and protocol tried for purification optimization for hProCA32.collagen using metal ions, reducing agent and arginine. ... 195
Figure 5-30: Different Precipitation level and supernatant appearance of the samples post heating................................................................. 197

Figure 5-31: Observable sample solution and Pellet appearance post supernatant analyzed by SDS-PAGE and pellet and supernatant estimation......................................................... 198

Figure 5-32: FPLC Chromatogram Representation condition one with higher buffer concentration.................................................................................................................. 199

Figure 5-33: FPLC Chromatogram Representation of conditions 2-5 listed above with 3x lower bound peak. .................................................................................................................. 200

Figure 5-34: Size Exclusion Chromatogram of the peak three fractions.............................................. 201

Figure 5-35: A 6-condition flow chart using 20mM with several conditions......................................... 202

Figure 5-36: hProCA32.collagen purification process. After the heat up step, all the protein yielded precipitate in contrast to previous observation................................................. 204

Figure 5-37: Addition of streptomycin to C1, C2, C4 and C5 solutions, C3 and C6 during arginine and BME incubation, respectively................................................................. 205

Figure 5-38: Addition of streptomycin to C1, C2, C4 and C5 solutions by next before and after centrifugation .......................................................................................................... 206

Figure 5-39: Analysis of pellet obtained for different conditions after dialysis step...................... 206

Figure 5-40: Agarose gel DNA analysis of the presence or absence of DNA at each step of the purification ....................................................................................................................... 207

Figure 5-41: Agarose gel DNA analysis of pellets and supernatant from the purification steps to track DNA shows a reduced DNA band after addition of streptomycin in conditions 1, 2 and 5 ....................................................................................................................... 208
Figure 5-42: FPLC chromatogram of the newly analyzed conditions showing a lower peak for the flow through compared to other peaks ................................................................. 208

Figure 5-43: SDS-PAGE and native gel analysis of purified hProCA32.collagen during purification and post Q-sepharose injection. .................................................................................. 209

Figure 5-44: SDS-PAGE Analysis of pellet and supernatant from the various purification steps. .................................................................................................................................................. 210

Figure 5-45: Process flow chart for purification optimization with DEAE column......... 211

Figure 5-46: SDS-PAGE and agarose gel for DEAE column purified hProCA32.collagen ..... 212

Figure 5-47: Different DNA precipitation methods and conditions using PEI ............... 213

Figure 5-48: Scheme for DNA Precipitation by Polyethyleneimine at low and high ionic strengths ........................................................................................................................................ 214

Figure 5-49: DNA Agarose gel for different DNA precipitation conditions explored using both low ionic and high ionic strengths. ........................................................................................................ 214

Figure 5-50: Collagen type 1-hProCA32.collagen dissociation curve of suspected monomer purified protein vs dimer containing protein. ........................................................................................................ 215

Figure 5-51: ELISA of monomer purified protein with and without BME......................... 217

Figure 5-52: ELISA assay of different peaks obtained from FPLC characterization in Fig. xx. 218

Figure 5-53: Mass spectrometry and concentration Estimation of hProCA32.collagen by SDS-PAGE quantification showed no protein. Marker band molecular weights are 68, 26, 14.3 and 6.5 kDa from top to bottom in A, B, and C........................................................................................................ 220

Figure 5-54: ELISA estimation of hProCA32.collagen purified by different methods with their corresponding SDS-PAGE.................................................................................................................. 221
Figure 5-55: Diagrammatic representation of workflow for cysteine addition for oligomer prevention. ................................................................. 223

Figure 5-56: SDS-PAGE of different steps in the cysteine-BME experiment for prevention of oligomerization of hProCA32.collagen. Mark band molecular weights are 68, 26, 14.3 and 6.5 kDa from top to bottom in both gels. ................................................................. 224

Figure 5-57: Use of higher concentration (2.5mM and 5mM) of cysteine during cell lysis. Last lane is the protein marker with bands corresponding to 68, 26, 14.3 and 6.5 kDa. .......................... 224

Figure 5-58: Cysteine and BME addition to hProCA32.collagen for prevention of oligomerization. ......................................................................................................................... 225

Figure 5-59: SDS-PAGE of different purification steps with DTT and BME as reducing agent .................................................................................................................. 226

Figure 5-60: Estimation of the role of pH in cysteine/BME reducing capability .................. 227

Figure 5-61: SDS-PAGE characterization of hProCA32.collagen with pH optimization protocol. Mark band molecular weights are 68, 26, 14.3, and 6.5 kDa if unlabeled. ............... 228

Figure 5-62: Agarose gel analysis for DNA removal of hProCA32.collagen ....................... 229

Figure 5-63: Optimization of the dialysis condition of hProCA32.collagen purification ....... 230

Figure 5-64: SDS-PAGE of the Optimization of the dialysis condition for reduction/prevention of oligomer formation. .......................................................................................... 231

*Figure 5-65: Dynamic light scattering studies of hProCA32.collagen.* ............................. 232

Figure 5-66: Relaxativity (r1 and r2) measurement of hProCA32.collagen........................... 234

Figure 5-67: Collagen binding dissociation curve of hProCA32.collagen versus hProCA32 shows a saturated binding curve compared to the hProCA32 curve which showed no-binding. ........................................................................................................ 236
Figure 5-68: Collagen binding specificity of hProCA32.collagen .......................................................... 237
Figure 5-69: N-terminal PEGylation of hProCA32.collagen ................................................................. 238
Figure 5-70: SDS-PAGE analysis of the mPEG-hProCA32.collagen ......................................................... 239
Figure 5-71: Optimized Gd$^{3+}$ loading method with dropwise addition shown in tube A with clear solution compared to bolus addition in tube B .......................................................... 245

Figure 6-1: Examples of chronic lung diseases affecting humans ............................................................... 249
Figure 6-2: Causes of Pulmonary fibrosis ................................................................................................. 250
Figure 6-3: Mechanism of gaseous exchange at the alveoli of normal lung compared to fibrotic lungs ........................................................................................................................................ 252
Figure 6-4: Development of bleomycin-induced idiopathic pulmonary fibrosis mice model by Drs. Ganesh and Guangda in Dr. Zhiren Liu’s lab ........................................................................ 261
Figure 6-5: Human features of idiopathic pulmonary fibrosis observed in our intraperitoneal (IP) administration of bleomycin by H&E staining ........................................................................ 261
Figure 6-6: Characterization of bleomycin-induced IPF mice model ...................................................... 263
Figure 6-7: Representative axial T1-W lung images of normal non-diseases mice showing no observable presence of fibrosis or abnormalities in the lung of two mice (M1 and M5). ........................................................................................................................................ 265
Figure 6-8: Representative axial T2-W lung images of normal non-diseases mice showing no observable presence of fibrosis or abnormalities in the lung of two mice (M1 and M5). ........................................................................................................................................ 266
Figure 6-9: Average signal-to-noise ratio (SNR) of the normal (healthy) mice with RARE acquisition ................................................................................................................................. 266
Figure 6-10: Representative axial T1-W and T2-W lung images ................................................................ 267
Figure 6-11: Representative axial T1-W and T2-W lung images ................................................................. 267

Figure 6-12: Representative axial T1-W and T2-W lung images of 3 weeks bleomycin-induced mice......................................................................................................................................................... 268

Figure 6-13: Summary of corona orientation of T1-W representative of RARE mice with injection of different MRI contrast agents. .................................................................................................................. 270

Figure 6-14: Signal to noise ratio estimation for the various MRI contrast agents tested in conventional MRI T1-W pulse sequence............................................................................................................. 271

Figure 6-15: 3D UTE MR Images of normal mice shows no significant enhancement in the lungs ............................................................................................................................................................... 273

Figure 6-16: 3D UTE MR SNR of normal mice shows no significant enhancement in the lungs across all the time points ....................................................................................................................................... 274

Figure 6-17: 3D UTE MR Images of 6-weeks bleomycin treated IPF pre, 1 h, 3h and 24 h after injection of non-targeted hProCA32 .............................................................................................................................................. 275

Figure 6-18: 3D UTE MR SNR of 6-weeks bleomycin treated IPF mice ...................................................... 275

Figure 6-19: 3D UTE MR Images of 6-weeks bleomycin treated IPF pre, 30 minutes, 3h and 24 h after injection of clinical contrast agent, Gadovist® .................................................................................................................. 276

Figure 6-20: 3D UTE MR SNR of 6-weeks bleomycin treated IPF ............................................................... 277

Figure 6-21: 3D UTE MR Images of 3-weeks bleomycin treated IPF pre, 1 h, 3h and 24 h after injection of hProCA32.collagen ............................................................................................................................................. 278

Figure 6-22: 3D UTE MR SNR of 3-weeks bleomycin treated IPF pre, 1 h, 3h and 24 h after injection of targeted hProCA32.collagen shows an increase in enhancement at 1 h time point in the lungs, this enhancement was increased at 3 h time point and reduced slightly at 24 h time point ............................................................................................................................................. 278
Figure 6-23: 3D UTE MR Images of 6-weeks bleomycin treated IPF pre, 1 h, 3h and 24 h after injection of hProCA32.collagen................................. 279

Figure 6-24: 3D UTE MR SNR values of 6-weeks bleomycin treated IPF pre, 1 h, 3h and 24 h after injection of hProCA32.collagen shows an increase in enhancement peaking at 1 h time point in the lungs, this enhancement was maintained at 3 h time point and 24 h time points.............................................................. 279

Figure 6-25: Detection of early and late stage IPF .................................................. 281

Figure 6-26: Histogram pixel analysis of lung after MR imaging............................. 282

Figure 6-27: Correlational studies of hProCA32.collagen MRI and histological findings in collaboration with Dr. Hua Yang....................................................... 283

Figure 6-28: M2 macrophage estimation in nicotine-induced COPD mice model. .................. 285

Figure 6-29: Spin echo multi-slice (T1-W) images of e-cigarette smoked mice for assessment of COPD (airway fibrosis) induction..................................................... 285

Figure 6-30: 3D UTE MR Images of nicotine-induced lung fibrosis pre, 1 h, 3 h and 24 h after injection of hProCA32.collagen........................................................................ 286

Figure 6-31: 3D UTE SNR values of nicotine-induced lung fibrosis pre, 1 h, 3 h and 24 h after injection of hProCA32.collagen........................................................................ 287

Figure 6-32: Hematoxylin and Eosin (H&E) staining with Sirius red stain of adjacent slices of nicotine-induced COPD Mice model shows collagen accumulation in the airways (bronchi) of the mice................................................................. 288

Figure 6-33: Detection of airway fibrosis in e-cigarette smoke transgenic mice model. BM-induced idiopathic pulmonary fibrosis in mice............................................. 289
Figure 6-34: Significance of specificity of hProCA32.collagen uptake in alveoli and bronchioles of IPF and nicotine-induced lung fibrosis................................................................. 290

Figure 6-35: In vivo pharmacodynamics and pharmacokinetics studies of hProCA32.collagen in different organs of 6 weeks bleomycin-treated mice......................................................... 292

Figure 6-36: Comparison of the lung and liver signal enhancement pattern............................... 293

Figure 6-37: Background selection troubleshooting of 3D-UTE lung data........................................ 295

Figure 7-1 A. The human liver and the several vital functions it performs in the human body (B). ........................................................................................................................................ 299

Figure 7-2 Nonalcoholic fatty liver disease risk factors, Progression, and prevalence. .......... 301

Figure 7-3 Experimental design and development of HDAC1 gene expressed NASH mouse model.................................................................................................................................. 308

Figure 7-4: Timeline of liver fibrosis model development and progression induced by western diet/CCl4.............................................................................................................................................. 309

Figure 7-5: Characterization of HDAC1 gene expressed NASH mouse model......................... 317

Figure 7-6: R1 mapping of HDAC induced fibrotic mice............................................................ 319

Figure 7-7: R2 mapping of HDAC induced fibrotic mice. A., Representative R1 maps of normal (control) mice (top row), HDAC1 KO mice (middle row), and HDAC1 gene mice (bottom row) before, 3h, 24 h and 48 h after injection of hProCA32.collagen contrast agent, B., change in R1 values of normal, HDAC1 wild type and HDAC1 knock out mice 3h, 24 h and 48 h after injection of hProCA32.collagen C., R1 values of normal, HDAC1 wild type and HDAC1 knock out mice before, 3h, 24 h and 48 h after injection of hProCA32.collagen.............................................................................................................. 320

Figure 7-8: Histological characterization of HDAC1 mice model .............................................. 321
Figure 7-9: TI-W MR images of mice 1 with HDAC1 gene .................................................. 322
Figure 7-10: TI-W MR images of mice 2 with HDAC1 gene ............................................. 323
Figure 7-11: TI-W MR images of mice 1 with HDAC1 knock out gene .............................. 324
Figure 7-12: TI-W MR images of mice 2 with HDAC1 knock out gene ............................. 325
Figure 7-13: Liver weight and Morphology post Euthanizing. .......................................... 311
Figure 7-14: RI and R2 mapping of CGI-58 induced fibrotic mice. .................................... 312
Figure 7-15: T1-W Images of human comparative gene knock out mice mode (CGI-58 LivKO)
A. representative T1-W images before and after injection of hProCA32.collagen and
hProCA32 B, C., SNR and CNR plot of CGI-58 mice models shows 24 h time point as
the targeting time point for hProCA32.collagen in both CGI-58 KO and WT mice
compared to hProCA32 ........................................................................................................... 313
Figure 7-16: Heavy T2-W images of early-stage liver fibrotic CGI-58 mice model. .......... 314
Figure 7-17: Heavy T2-W images of late stage liver fibrotic CGI-58 KO mice model with
estimation of SNR and CNR before, 3h, 24 h and 48 h after injection of
hProCA32.collagen ................................................................................................................ 315
Figure 7-18: Histological characterization of LivCGI-58 mice model ................................. 316
Figure 7-19: Comparison of histological results of HDAC1 vs CGI-58 mice model. ........ 326
Figure 7-20: Summary of RI maps and R1 values ................................................................. 327
Figure 7-21: MRI and histological analysis of liver fibrosis. ................................................. 329
Figure 7-22: Representative T1 inversion recovery images of western diet/CCL4 induced NASH
mice at different repetition times (TR) across 8 slices of MR images ............................. 330
Figure 7-23: Whole slides T2 images of western diet/CCL4 induced NASH mice at 12 different
echo times (TE) across 8 slices of MR images ................................................................. 331
Figure 7-24: Repetition time dependent image analysis.......................................................... 332
Figure 7-25: TE dependent image analysis............................................................................. 333
Figure 7-26: Comparison of T1W MR images all three mice models..................................... 334
Figure 7-27: Summary and comparison of liver enhancement using rat vs human
    ProCA32.collagen........................................................................................................... 335
Figure 7-28: Summary and comparison of liver morphology and weight of different mice model
    developed for liver fibrosis.............................................................................................. 336
Figure 8-1 Risk factors associated with development of hepatocellular carcinoma (HCC)...... 338
Figure 8-2: HCC nodules classification.................................................................................. 341
Figure 8-3 Diagnostic guidelines for hepatocellular carcinoma.......................................... 343
Figure 8-4: Gross liver images of mice and MR Imaging and euthanasia.............................. 346
Figure 8-5: Signal to noise ratio calculation of normal healthy mice...................................... 347
Figure 8-6: T2-W MR images of 12-weeks HCC mice before, 3 h and 24 h after
    hProCA32.collagen injection shows increase liver enhancement at 3 h time point and
    more tumor....................................................................................................................... 348
Figure 8-7: T2-W MR images of early-stage DEN-induced HCC ........................................... 349
Figure 8-8: T1-W MR images of 12-weeks HCC mice.......................................................... 350
Figure 8-9: T1-W Early stage HCC den-induced mice model shows a liver enhancement at 3 h
    time point but not 24 h time point.................................................................................. 351
Figure 8-10: Histological correlation of early HCC with MRI lesions shows a linear correction
    with an R value of over 97%............................................................................................ 352
Figure 8-11: MRI slices of the entire liver.............................................................................. 353
Figure 8-12: Analysis of TI-W MRI slices with similar anatomic properties........................... 353
Figure 8-13: Signal to noise ratio of tumor vs liver of HCC/Cirrhosis mice model showing .... 354

Figure 8-14: Correlational analysis of MR images with gross and histological Sirius red stain with collagen accumulation. ................................................................. 355

Figure 8-15: Left to right, T1-Weighted MR image slices of HCC mice liver ...................... 356

Figure 8-16: Characterization of diet-induced HCC by MRI and histology. ......................... 357

Figure 8-17: MR imaging and characterization of HCC liver versus tumor lesions. ............. 358

Figure 8-18: T1-W MRI and quantification of HCC with clinical contrast agent Eovist........ 358

Figure 8-19: MRI correlation of Rim-enhanced tumor nodule with HCC and collagen-specific staining ........................................................................................................ 359
LIST OF ABBREVIATIONS

3D: three-dimensional
3DGE: 3D gradient echo
AIF: arterial input function
ALP: alkaline phosphatase
ALT: alanine transaminase
AST: aspartate transaminase
ATGL: Adipose triglyceride lipase
BSA: bovine serum albumin
CAFs: Cancer Associated fibroblast
CD2: cluster of differentiation
CGI- Comparative gene identification
CNR: contrast-to-noise ratio
CNS: central nervous system
COPD: Chronic obstructive pulmonary diseases
CLD: Chronic liver diseases
CLD: Chronic lung diseases
CRD: Chronic respiratory diseases
CT: computed tomography
CV: column volume
Cys: cysteine
DD: dipole-dipole
DEC-MRI: dynamic contrast enhanced-MRI
DTPA: diethylene-triamine-pentaacetic acid
EDTA: ethylenediaminetetraacetic acid
EGFR: epidermal growth factor receptor
EGTA: ethylene glycol tetraacetic acid
FDA: U S Food and Drug Administration
FOV: field of view
FPLC: fast protein liquid chromatography
FSE: fast spin echo
FSEMS: fast spin echo multi-slices
Gd$^{3+}$: Gadolinium
Gd-DTPA: gadolinium-diethylene-triamine-pentaacetic acid
GMP: good manufacturing practice
GRF: glomerular filtration rate
GRP: gastrin-releasing peptide
GRPR: gastric-releasing peptide receptor
HCC: Hepatocellular carcinoma
HER2 or HER2/Neu: human epidermal growth factor receptor 2
HSCs: Hepatic stellate cells
ICP-OES: inductively coupled plasma atomic emission spectroscopy
IHC: immunohistochemistry
IPF: Idiopathic pulmonary fibrosis
IPTG: isopropyl β-D-1-thiogalactopyranoside
Kd: dissociation constant
LD50: median lethal dose

LDH: lactate dehydrogenase

LRET: luminescence resonance energy transfer

MEMS: multi-echo-multi-slices

Mn-DPDP: manganese dipyridoxal diphosphate

MRI: magnetic resonance imaging

MVD: mean vascular density

NAFLD: Non-alcoholic fatty liver disease

NASH: Non-alcoholic steatohepatitis

NHS: N-hydroxysuccinimide

NIR: near infrared

NMRD: Nuclear Magnetic Relaxation Dispersion

NSCLC: Non-small-cell lung cancer

NSF: nephrogenic systemic fibrosis

NTA: nitrilotriacetic acid

OD: optical density

PCR: polymerase chain reaction

PEG: polyethylene glycol

PET: positron emission tomography

PMSF: phenylmethylsulfonyl fluoride

ProCA: protein-based MRI contrast agent

ProCA3.bomb or ProCA3.bombesin: ProCA3 linked with bombesin peptide

PV: parvalbumin
R1: longitudinal relaxation rate
r1: longitudinal relaxivity
R2: transverse relaxation rate
r2: transverse relaxivity
RARE: Rapid acquisition refocusing echo
SBM: Solomon-Bloembergen-Morgan
SC: scalar or contact
SCLC: Small cell lung cancer
SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis
SE: spin echo
SPECT: single-photon emission computed tomography
T1: longitudinal relaxation time
T2: transverse relaxation time
TBST: Tris-buffered saline with Tween 20
TE: echo time
TG: Triglyceride
TR: repetition time
UTE: Ultra short echo
Vc: initial volume distribution
Vdss: steady state phase
VEGF: vascular endothelial growth factor
WT: wild type
ZFS: zero field splitting
Δ$K_{\text{obs}}$: the difference of the decay rates of Tb$3^+$ in the presence of H2O and D2O

τ: lifetime

Γ: the radiative decay
1 INTRODUCTION

1.1 Chronic Human Diseases

According to U.S. National Center for Health Statistics, chronic diseases are generally defined as conditions lasting three months or longer[1] requiring ongoing medical attention or limit activities of daily living or both. These diseases including cancer, lung diseases, obesity, arthritis, Alzheimer’s disease, stroke, kidney diseases, and cardiovascular disease are leading cause of death in the US and the world at large. About 60% of adults in the US have at least one form of chronic disease, while 40% suffer from two or more chronic diseases. The lung is a vital organ that is susceptible to airborne infections and injuries[2]. Lifestyle choices such as cigarette smoking, alcohol consumption, lack of physical exercise and unhealthy diets are the major risk factors associated with these diseases[3]. 90% of the annual health care cost is attributed to the support of people with two or more chronic conditions. So many efforts have been channeled into possible prevention of chronic disease as well as early detection and intervention strategies to improve the lives of patients with one or more chronic diseases [4, 5]. The diseases associated with the four-leading cause of death are cardiovascular diseases, cancer, chronic respiratory diseases, and diabetes [2, 6, 7]. These diseases (Figure 1-1) have also been identified as a major risk factor for increasing mortality associated with the SARS-COV-2 pandemic [8-11].
Figure 1-1: World Health Organization (WHO) estimate of the top ten global leading cause of death. Modified after [7].
In 2020, approximately 3.4 million patients died in the United States, 378,000 of which is associated with the SARS-COV2 infection [11, 12]. A major contributing factor to these figures is the increase in the surge of chronic respiratory diseases (CRDs), a group of diseases affecting the airways and other structures of the lungs. They account for the second highest cause of death globally. Lung diseases such as interstitial lung diseases, asthma, chronic obstructive pulmonary disease (COPD), cystic fibrosis, infection, and cancer are life-threatening conditions that slowly deteriorate quality of life. Asthma is a respiratory disease common in children and young adults. About 235 million people suffer from asthma with 3 million annual death, making it the third leading cause of death worldwide [13]. Active and passive cigarette smoking is a major risk factor of CRDs, while other factors include air pollution, age, sex, exposure to radiation etc. Despite advances in investigation of the possible etiologies of these diseases, the pathology and mechanism are yet to be fully understood [14]. Lung disease diagnosis is challenging for several reasons, high resolution computed tomography (HRCT) is the current gold standard for diagnosis of several CLDs [15] however, its ability to monitor disease progression is limited due to the use of ionizing radiation. In addition, it lacks the ability to give molecular information about the pathology of the diseases of the lungs and may require lung biopsy to establish accurate diagnosis [16]. Cancer in various organs represents another chronic disease of major concern. Cancer affecting various organs such as lungs and liver is the second leading cause of death worldwide [17]. In 2018, an estimated 9.6 million deaths resulted from cancer, and it is projected that about 1.9 million people will be diagnosed with cancer in 2021 with over 600,000 deaths [18-20]. A major strategy to the reduction of cancer related death will be to enable early detection of cancer for better therapeutic outcomes and precise and accurate cancer screening in at-risk patients. In
addition to effective treatment, the ability to separate aggressive from nonaggressive cancers, monitor recurrence and response to therapy is also an unmet medical need [21, 22]. Advances in early diagnosis has significantly improved the 5-year relative survival rate for some cancers further stressing the need for more improvement of cancer diagnostic tools to better increase survival chances and curative intent [23-26]. A primary liver cancer, hepatocellular carcinoma (HCC), has emerged as the second leading cause of cancer-related deaths globally [27]. HCC predominantly occurs in the setting of liver cirrhosis and is a typical example of inflammation-induced cancer. The causes of chronic liver disease promote the development of transformed or premalignant hepatocytes and leads to the development of HCC. Approximately 80% - 90% of HCC patients have liver cirrhosis with the rest having moderate to advanced liver fibrosis [28, 29]. Despite increasing knowledge on the causal factors of cirrhosis and progress in diagnosing and managing risk factors, the incidence rates for HCC are increasing. This increase has been linked to lack of accurate detection and screening methods for liver fibrosis, cirrhosis, and HCC. Ultrasound (US) screening for chronic liver disease is gaining popularity because of its cost and noninvasive surveillance method without any risk or radiation exposure for the patient. However, US imaging of HCC in a cirrhotic liver is limited by several factors such as abnormal liver texture, obesity, and low resolution of US. Generally, biopsy is the gold standard for diagnosis of liver disease. For instance in liver cancer, biopsy can only be performed after the tumor mass is detectable making it susceptible to high sampling error [30]. There is also the tendency of underdiagnoses and over diagnoses of diseases [31]. Up to 33% of patients have associated sampling error during biopsy sampling and analysis [32, 33]. An advancement in tissue biopsy is the introduction of liquid biopsy which provides an avenue for earlier detection through analysis of bodily fluids. The results obtained from this procedure
however are less reliable compared to tissue biopsy[30]. High resolution computed
tomography (HRCT) and magnetic resonance imaging (MRI) are a few imaging modalities
that are employed as biopsy guided imaging technique in cancer screening [34, 35]. Of these
two modalities, MR imaging is the most potentially beneficial due to its unique properties
(discussed in the next section). Additionally, the use of molecular imaging with target specific
biomarker in tandem with noninvasive MR imaging of chronic disease and tumor development
will be beneficial in elucidating the mechanism of disease progression and monitoring response
to therapy.

1.2 Magnetic Resonance Imaging and its advantage as an imaging technique

A major recurring theme in the diagnosis of human diseases is associated limitation in all the
imaging modality currently available for patient screening and disease monitoring despite their
advantages. MRI has emerged as a leading imaging modality for monitoring human diseases
due to its capability for high-resolution 3D images, non-invasiveness, whole body imaging
without the use of ionizing radiation and superior soft tissue contrast [36-39]. MRI takes
advantage of the approximately 70% water [40] in the body. Hydrogen nuclei from resident
body water molecule are the most useful atoms to use in imaging. Another advantage of
employing hydrogen in imaging is due to its odd number of protons which is a basic
requirement for MR imaging. MR imaging uses the unpaired proton in an atom to provide the
magnetic moment due to the intrinsic “spin” property the unpaired proton possesses. By
utilizing the magnetic properties of water protons, MRI provide high resolution three
dimensional (3D) images of the body, it is non-invasive, it is capable of deep tissue penetration
and allows for whole body imaging. These advantages sets it aside from other imaging
modality (Table1) and affords its use in diagnosis and monitoring of various types of cancers
and other human diseases [38]. The use of MRI has opened a world of opportunities and great advancement in both science and medical imaging at large but it is limited by sensitivity. The basic principle of nuclear magnetic resonance (NMR) and MRI focuses on the interaction of atomic nuclei, radio frequency energy and a magnetic field. An understanding of the underlying principles of MRI and relaxation theory is imperative for exploring, optimizing and advancing the use of MRI as an indispensable tool in non-invasive clinical imaging. Each of the terms highlighted above will be discussed in detail below.

1.2.1 Atomic Nuclei

Every atom possesses a small region of high density at its center containing the positively charged proton and the neutron with no charge. This positively charged nucleus is surrounded by negatively charged electrons orbiting around it. The hydrogen atom employed for use in NMR/MRI consist of a single proton in the nucleus with a circulating electron orbiting around it. Since the electron is of little use in NMR/MRI, the term ‘nuclear’ stems from the use of the proton occupying the nucleus.

1.2.2 Magnetic field

MRI exploits the magnetic properties of specific atomic nuclei such as hydrogen to give detailed characterization of the nuclei of interest. The hydrogen proton rotates on its axis like the planet’s orbit around the sun. This moving proton charges produces Magnetic field since Magnetic fields exert forces on charged particles in motion. The magnitude of the magnetic force \( F \) on a charge \( q \) moving at a velocity \( v \) in a magnetic field of strength \( B \) is given by the equation:
\[ F = qvB(\sin \theta) \]  

(1.1)

Where \( \theta \) is the angle between the directions of \( \mathbf{v} \) and \( \mathbf{B} \).

The magnitude of the magnetic field strength is expressed in units of tesla (T) in honor of Nikola Tesla for his contributions to the field of magnetic fields and their practical applications. The direction of the magnetic force \( \mathbf{F} \) is perpendicular to the plane formed by \( \mathbf{v} \) and \( \mathbf{B} \) as determined by the right hand rule which states that, to determine the direction of the magnetic force on a positive moving charge, curl the fingers of your right hand in the direction of the current flow through the coil, and your outstretched arm will point in the direction of the magnetic field.[41-43]

1.2.3 Radiofrequency Energy

Radiofrequency energy is part of the electromagnetic spectrum (a range of various types of electromagnetic radiation). Electromagnetic radiation can be described as a stream of photons, each traveling at the speed of light in a wave-like pattern. It ranges from radiofrequency at the far left to X-rays at the far right and visible light positioned in the middle. Each of these radiations have a wavelength, Energy and frequency function associated to them. The **wavelength** is the distance between peaks of the wave, the **frequency** is the number of waves passing a fixed position or completing a cycle per second. The amplitude of the wave is the height or power of the wave Figure 1-2. The phase and frequency of these waves are utilized in MRI/NMR.[44]
Figure 1-2: Electromagnetic Spectrum with different range of electromagnetic radiation

1.2.4 MRI Pulse Sequence

Pulse sequence describes a predefined set of changing magnetic gradients (radiofrequency and gradient pulses), usually repeated many times during a scan, wherein the time interval between pulses and the amplitude and shape of the gradient waveforms will control NMR signal reception and affect the characteristics of the MR images. It sets the specific number, strength, and timing of the RF and gradient pulses. The MRI pulse sequences are dependent on several parameters such as repetition time (TR) which is the time between consecutive 90-degree RF (the end of one RF excitation and the beginning of the next RF pulse). Echo time (TE) is the time between the application of the radiofrequency excitation pulse and the peak of the signal induced in the coil. Both TR and TE are measured mostly in milliseconds. The flip angle Flip angle (The flip angle is the rotation of the net magnetization vector by a radiofrequency
pulse relative to the main magnetic field), **Inversion time** (the time between the 180° inverting pulse and the 90°-pulse is called the inversion time) etc. Some of the commonly used pulse sequence include **Spin Echo, Gradient Echo, Inversion Recovery, Saturation Recovery** and **Diffusion Weighted** Sequences. In MRI, differences observed between different pulse sequences in various tissues are responsible for the contrast observed [45-47].

### 1.2.5 MR image quality

A point on this screen is called pixel which is a concatenation of the word **picture and element**. MR images are made up of a series of voxels (volume element). Each square on the MR image corresponds to a volume of tissue on the body. The MRI machine is designed to measure the NMR signal from each of this small volume, localize them in 3D space and plot them on a matrix to generate a visible picture. The resolution (the number of pixels in a given field of view (FOV)) of the image generated can be **in-plane, temporal or spatial**. The temporal resolution describes the time lag between consecutive images of time and is defined mathematically as

\[ VPS \times TR \]  

Where VPS is view per segment and TR is the repetition time. The temporal resolution of an image acquired with a TR of 8ms on 4 segments will have a temporal resolution of 32ms. A high in-plane resolution results in better morphologic patterns visualization in MR imaging [48]. In-plane resolution is directly proportional to pixel size, and it is expressed mathematically as FOV/matrix. Lastly, spatial resolution predicts how well morphological structures are separated and distinguished from one another. It is defined by voxel size of the
acquired images. The size of the voxel and by extension, the image resolution depends on the FOV, slice thickness and the matrix size. The higher the resolution of the image, the better the quality of the image obtained. A balance between signal to noise (SNR) and resolution must be established because a resolution higher than the acceptable range will produce a grainy image with low SNR and a blurry image with high SNR if the resolution is reduced below the acceptable range [49, 50]. Despite the advantages of MR imaging, a major limitation to the use of MRI relative to other imaging techniques such as Positron emission tomography (PET), HRCT, and Single-photon emission computed tomography (SPECT) is low sensitivity necessitating the need for exogenous contrast agents in up to 35% of clinical MR imaging [51].

1.3 MRI contrast Agents

MRI contrast agent gives an observable difference between a pathological and a normal tissues/organ in diagnostic imaging thereby increasing the information available for diagnostic purposes [52]. A major distinctive characteristics between MRI contrast agents and other frequently used imaging probes in other modality is the indirect detection of abnormalities with MRI probes in MR imaging as opposed to other probes such as probes for PET which relies on direct visualization of radioactive substances (radiotracers) to quantify change in metabolic processes, and other physiological activities in the human body [53]. The basis of using a contrast agent in MRI is to accelerate the relaxation of water protons in the surrounding tissue. There are different types of MRI contrast agents such as those based on iron, zinc, manganese, Gadolinium, nanoparticles etc. These are broadly classified as paramagnetic, ferromagnetic, or super paramagnetic metal ion-based contrast agents [54, 55]. All MRI contrast agents acts by decreasing both longitudinal (T1) and transverse (T2) relaxation times of available water protons albeit to different extents. Nearly all MR images enhance tissue contrast by utilizing
T1, T2 and proton density characteristics of tissues under observation. In T1-weighted imaging, tissues with short T1-values such as the adipose tissues will appear bright while tissues with low T1 values such as cerebrospinal fluid will appear dark [56, 57]. T1-W MRI contrast agents have gained more acceptance due to their ability to produce images with better contrast. Trivalent Gadolinium (Gd\(^{3+}\)) contains seven unpaired electrons in its 4f shells making it the most highly paramagnetic single cation known (Figure 1-3). These electrons remain ‘available’ when bound and thereby maintaining their paramagnetic status even in molecular bonding. This results from dipolar interactions between water nuclei (in tissue) and electron spins at the metallic center, a phenomenon known as paramagnetic relaxation. Usually, body protons from water, lipid and protein undergo Brownian motion [58] generating relatively small fluctuating fields as compared to that generated by electrons which is about 658x greater than proton [59]. As the Gd\(^{3+}\) complex diffuses in solution, a fluctuating magnetic field is created. This fluctuation results in water relaxation if it is close enough to the hydrogen Larmor frequency as electron is transferred from the gadolinium to water [60]. As a paramagnetic ion with long electron spin relaxation time and high magnetic moment, Gd\(^{3+}\) is one of the most widely used ions in T1-weighted MRI contrast agents.
Gadolinium (Gd$^{3+}$) is a lanthanide metal with long electron spin relaxation time, and high magnetic moment. It shortens T1 to a greater extent than T2 and has emerged as one of the most widely used ions in T1-weighted MRI imaging [61-64]. These set of MR imaging probes are referred to as gadolinium-based contrast agents (GBCAs). GBCAs use have been remarkably successful due to their ability to noninvasively provide essential diagnostic information unlike other elements such as Iron oxide nanoparticles and manganese (II) complex which even though approved for imaging the liver were not commercially successful. GBCAs are used in at least 35% of all magnetic resonance imaging (MRI) exams and approximately 60% of neuro MRI exams are conducted with a GBCA. Thus, representing a global GBCA administration of about 40 million [65].
Figure 1-4: Coordination polyhedron of Gd$^{3+}$ showing A. Molecular stick models of Gd(H$_2$O)$_8^{3+}$ and B. Gd(H$_2$O)DTPA$^2-$ modified after [66]. C. Schematic diagram of factors contributing to the relaxivity of GBCAs

Furthermore, the relaxivity of these contrast agents are described by a combination of the Zimmerman-Brittin and Solomon-Bloembergen-Morgan (SBM) models which has been simplified to provide formulation of the relaxation effects of gadolinium in tissue mathematically as a combination of equation 1.3 as follows [67];

$$\frac{1}{T_{i,obs}} = \frac{1}{T_{i,b}} + r_i[C_A]; \ i = 1,2 \quad (1.3)$$

From equation 1.5, it can be deduced that the percent change in tissue relaxation rate due to administered GBCAs linearly correlates with the concentration and the relaxivity of that contrast agent but inversely proportional to the baseline relaxation rate of the tissue of interest before contrast. As a result of the intrinsic relaxation rates, a high relaxivity or a high dose will be required of contrast agents to overcome this baseline relaxation to obtain an observable contrast with GBCAs up to at least 10% higher relativity than the baseline or buffer relaxivity
From equation 1.5, it can be deduced that the percent change in tissue relaxation rate due to administered GBCAs linearly correlates with the concentration and the relaxivity of that contrast agent but inversely proportional to the baseline relaxation rate of the tissue of interest before contrast. As a result of the intrinsic relaxation rates, a high relaxivity or a high dose will be required of contrast agents to overcome this baseline relaxation to obtain an observable contrast with GBCAs up to at least 10% higher relativity than the baseline or buffer relaxivity [53]. For most commercial GBCAs, they possess low relaxivity and would require high dose which will inherently lead to the propensity for Gd\(^{3+}\) toxicity. The decision on the type of GBCAs to be employed for a particular image and the dosage requirement has several implications some of which can be fatal. It is therefore important to design contrast agents which allows for the lowest dosage possible by designing contrast agents with improved relaxivities. Several factors contributes to the relaxivity of GBCAs and no one factor can be optimized effectively enough to achieved the best relaxivities [68]. Some of the factors include number of water molecules present in the inner coordination sphere of Gd\(^{3}\), temperature, magnetic field strength, mean residency time of the inner sphere water molecules, water exchange rate, number of exchangeable water molecules or hydrogen atoms in the second, third and subsequent coordination spheres, water residence time, temperature. The inherent tissue T1 relaxation time (Ti\(_b\)) as well as the relaxivity of GBCA both depends on field strength but to different level. The sum of relaxation time of both the tissue and GBCA is inversely proportional to field strength, with a fractionally greater effect at high field than at low field since both \(r_1\) of Gd\(^{3+}\) and \(r_1\) of baseline tissue increase nonlinearly and disparately as the
resonance frequency decreases.[67, 69, 70]. The interaction between Gd$^{3+}$ and water proton which affects its relaxation happens through a dipolar mechanism depending on $1/r^6$ reflecting the distance between the ion and the nucleus. It is therefore critical to have at least one exchangeable water in the inner coordination sphere of Gd$^{3+}$ binding site. The number of water in the inner coordination sphere is represented by $q$ and is called hydration number, however if this hydration number is bigger than one then the complex should be stable since the water can be displaced by ligands inside the inner coordination sphere[68, 71]. Water molecules in the inner coordination sphere, second sphere and outer sphere will have decreasing water exchange rate due to their distance from the gadolinium ion Figure 1-4.

1.3.1 Current MRI contrast agents and their limitations

GBCAs are by far the most prominent MRI contrast agents used in approximately 40 % of all MRI scans and in about 60% of neuro MRI exams Figure 1-5 [72, 73]. They are also called positive contrast agents because they increase the MR signal in regions where they distribute. The contrast agent creates a brighter or darker image specifically in the area of diseased tissue. GBCAs acts as a catalyst to speed up the rate of relaxation of water proton atoms with a rate termed relaxivity [53]. The first clinically approved contrast agent was Gd- Magnevist® (gadopentetate dimeglumine,) in 1988 and contrast-enhanced MRI has since gained universal applicability in various human diseases such as in liver, kidney, and gall bladder [74].
Figure 1-5: Chemical structure of the clinically approved gadolinium-based contrast agents with their chemical and commercial names. Modified after [75]. Multihance® and eovist® are hepatobiliary agents [76].

The first generation of contrast agents produced were specific for different organs such as Gadoxetate disodium (eovist) which is taken up by the hepatocyte for liver specific MR imaging and are mostly extracellular fluid (ECF) agents as they extravasate into the
extracellular space and are relatively quickly excreted with each pass through the kidneys. In preclinical stage, there has been continuous efforts to create gadolinium-based contrast agents in molecular imaging. Gadofosveset trisodium (also known as Ablavar™ or Vasovist™) was an intravenously administered blood pool GBCA used in magnetic resonance imaging until it was discontinued in 2017. Its application was in MR angiography (MRA) because of its strong binding affinity to plasma proteins [77]. Most GBCAs are intravenously administered, while some in addition to intravenous (IV) injection can also be administered via intraarticular (IA) route. The route in which the contrast agent is being administered plays an indispensable role on biodistribution and safety. The majority of GBCAs are approved at a dose of ~ 0.1 mmol Gd/kg; however, for MRA, at least twice this dose is required. Gadolinium is a heavy metal which is highly toxic in biological systems. It is therefore imperative that GBCA be eliminated from the body in its bound form to avoid de-chelation and subsequent metal toxicity from gadolinium retention and brain in several areas of the body including the brain [62, 78-85]. The lower the injection dosage, the less toxicity associated with such contrast agent since GBCAs have been found to cause nephrogenic systemic fibrosis (NSF) - fibrosis of the skin and internal organs which can be fatal. NSF was found to have resulted from gadolinium retention in patient with renal impairment [74, 86-88] and more recently gadolinium retention was observed in individuals with healthy kidney and renal function[89]. Since the first approval of clinical MRI contrast agent, gadopentetate dimeglumine (Magnevist), in 1987, over 450 million GBCAs have successfully administered worldwide for an array of different tissues, diagnostics, characterization of tumors and other imaging functions. In different cases of diseases, the change impermeability of blood vessels is crucial for the function of these MRI contrast agents. For instance, stroke, liver fibrosis, neuron degenerative diseases and so on all
have changes in permeability of blood vessels. These clinical contrast agents can function as probes for organs such as liver, kidney, and gall bladder.\[90, 91\]. GBCAs are made up of a gadolinium ion and a chelator bound in a linear or macrocyclic (Table 1-1 and Figure 1-5).

**Table 1-1 clinically approved gadolinium-based MRI contrast agents modified after [74, 92].**

<table>
<thead>
<tr>
<th>Brand name</th>
<th>Generic name</th>
<th>Kinetic stability</th>
<th>Structure</th>
<th>Current status (EMA)</th>
<th>Usage (recommended dosage in mmol/kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prohance</td>
<td>Gadoteridol</td>
<td>High</td>
<td>Macrocyclic/nonionic</td>
<td>Maintain usage</td>
<td>CNS: Extracranial/ extraspinal (0.1)</td>
</tr>
<tr>
<td>Ablavar/</td>
<td>Gadofosveset</td>
<td>Intermediate</td>
<td>Linear ionic</td>
<td>Discontinued due to poor sales</td>
<td>MRA (0.03)</td>
</tr>
<tr>
<td>Vasovist</td>
<td>trisodium</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eovist</td>
<td>Gadoxetate</td>
<td>Intermediate</td>
<td>Linear ionic</td>
<td>Maintain usage</td>
<td>Liver (0.025)</td>
</tr>
<tr>
<td></td>
<td>disodium</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Magnevist</td>
<td>Gadopentetate</td>
<td>Low</td>
<td>Linear ionic</td>
<td>Suspend administration through I.V injection</td>
<td>CNS 0.1 Extracranial/ extraspinal 0.1</td>
</tr>
<tr>
<td></td>
<td>dimeglumine</td>
<td></td>
<td></td>
<td></td>
<td>Body 0.1</td>
</tr>
<tr>
<td>Omniscan</td>
<td>Gadodiamide</td>
<td>Low</td>
<td>Linear nonionic</td>
<td>Suspended</td>
<td>Body/CNS/Intrathoracic, intra-abdominal/ pelvic (0.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Kidney 0.05</td>
</tr>
<tr>
<td>Multihance</td>
<td>Gadobenic</td>
<td>Intermediate</td>
<td>Linear ionic</td>
<td>Maintain usage</td>
<td>CNS/MRA (0.1)</td>
</tr>
<tr>
<td>Gadovist</td>
<td>acid</td>
<td></td>
<td></td>
<td></td>
<td>CNS/Liver/Kidney/MRA/ Whole body* (0.1)</td>
</tr>
<tr>
<td>Dotarem</td>
<td>Gadobutrol</td>
<td>High</td>
<td>Macrocyclic/nonionic</td>
<td>Maintain usage</td>
<td>CNS: Extracranial/ Extrapinal/Body (0.1)</td>
</tr>
<tr>
<td>Optimark</td>
<td>Gadoversetamide</td>
<td>Low</td>
<td>Linear nonionic</td>
<td>Suspended</td>
<td>CNS/Liver (0.1)</td>
</tr>
</tbody>
</table>

*Usage for whole body imaging restricted to the EU*

In July 2017, the European’s medicines agency (EMA) decided to either suspend or restrict the use of some of the clinically approved contrast agents. Multihance, widely used in the central nervous system (CNS) and in MRA, was restricted to use in the liver which together with eovist can be applied to MR imaging of the liver \[93\]. Furthermore, magnevist was restricted to use in scans associated with the joints due to low dosage associated with joint imaging. All macrocyclic contrast agents were maintained but the lowest dosage for the best image is advised \[93\]. Structurally speaking, all currently available commercial Gd\(^{3+}\) based MRI contrast agents are all formed by coordination of one gadolinium ion to organic chelators.
such as 1,4,7,10-tetraacetic acid (DOTA), diethylenetriaminepentaacetic acid (DTPA) and their derivatives. Gd$^{3+}$ has a total of 9 coordination sites and thus is coordinated to 8 oxygen and or Nitrogen molecule Figure 1-5 and one water molecule through Gd-O interactions [71]. Despite the significant progress made in the development of MRI contrast agents, the clinically available contrast agents possess low relaxivity and cannot detect diseases at early stages especially in cases such as in organ fibrosis where biomarker expression play a pivotal role in disease progression. It has now become an unmet medical need to develop targeted and sensitive MRI contrast agents with improved specificity, biomarker targeting capability, and better selectivity. This will in no small way help enable early detection of different chronic pathologies and improve patient life.

1.3.2 **MRI Relaxivity Theory**

The relaxation and relaxivity theory are the dogma of MRI and have been extensively employed as the basis upon which a wide range of NMR and MRI application is founded. Application of a magnetic field induces movement of electrons and thus creates an electric current. In MRI, the detection of the change in energy state of the body proton (hydrogen) when subjected to a strong static magnetic field (B) which polarizes the nuclear spin of the proton partially is what is converted to an observable image. Following the introduction of a rightly tuned radio frequency pulse, exciting the spins back into their higher energy state and separates them into weak (those aligned with the static magnetic field) and strong signals (those oriented against the static magnetic field) as spin relaxation from a higher energy state to a lower energy state occurs (A proton possesses two quanta mechanical state) [37, 94]. Atomic nuclei with an odd atomic mass or atomic number possess a quantum mechanical property referred to as “spin” which it utilizes to create its own magnetic field that can be used to interact
with external nuclei with spin properties or an external magnetic field such as that used in MRI experiment. In contrast, atomic nuclei with an even atomic mass/number are static, lacking the spin property required for magnetic interactions. Odd nuclei have a spin quantum number of $\frac{1}{2}$ with two allowed spin states $+\frac{1}{2}$ and $-\frac{1}{2}$ which possess the same energy in the absence of an external magnetic field. These nuclei can absorb and emit electromagnetic radiation. MRI takes advantage of the high prevalence of hydrogen in the body and the magnetic properties of the proton in a hydrogen atom both from abundance of water and/or from energy storing molecules such as fat and carbohydrates both of which contains an abundance of hydrogen (glucose possesses 12 hydrogens on a carbon oxygen hexagonal backbone and the free fatty acid contains 2 to 3 H attached to its carbon backbone). Specifically, the human body constitutes ~70% water, or ~90 M of hydrogen atoms provided by these water molecules. MRI depends on the relaxation of the single proton nuclei for an image signal output. Hydrogen atoms have nuclei with multiple spin states; a hydrogen atom ($^1$H) contains one unpaired proton, an electron and no neutron thereby having a net spin of $\frac{1}{2}$ conferring on it an intrinsic spin or magnetic property.
Atoms in a static magnetic field are subjected to a second oscillating electromagnetic field in the form of radio frequency radiation, which causes the nucleus to resonate[95]. In the absence of an external magnetic field, proton magnets or spins are oriented randomly cancelling each other out with no significant net magnetization. In the presence of an external magnetic field however, all the spins of proton magnet line up mostly in the direction of the applied magnetic field. On application of an external magnetic field ($B_0$), the spin is split into $+\frac{1}{2}$ and $-\frac{1}{2}$ magnetic moments. The $+\frac{1}{2}$ state possesses a lower energy and is aligned in the direction of the external magnetic field, while that of the higher energy $-\frac{1}{2}$ spin state is aligned in opposing direction to the external magnetic field (Figure 1-6). The difference in these spin states is dependent on the strength of the external magnetic field applied and is mostly small compared to what is needed to generate an MR image. The magnitude of the change in energy ($\Delta E$) is directly proportional to the applied magnetic field strength. In a typical MRI experiment, the
sample is placed in a homogenous magnetic field environment (Bo). As expected, majority of the spin will exist in the lower energy $+\frac{1}{2}$ state, aligning with the external magnetic field. On application of a radiofrequency (RF) pulse, this spin is converted to the higher energy $-\frac{1}{2}$ state and upon removal of the RF pulse, the spin returns to its original lower energy $-\frac{1}{2}$ state (Figure 1-6). The term given to the process of moving from a spin state of $+\frac{1}{2}$ to $-\frac{1}{2}$ is "resonance" because of the dependence of the radiofrequency and magnetic fields on the vibrational frequency. MRI signals are obtained from the water protons, provided by the body when exposed to a primary magnetic field and radiofrequency pulse. During this process, the nuclei’s vector is at either spin-up or spin-down direction in alignment with the magnetic field (Bo). The applied radio-frequency pulse then generates a rotational torque ($\tau_r$) on the nuclei’s vector. This process then causes precession of the nuclei spin, that is a spiral motion along the applied rf angle, between the spin up and spin down directions Figure 1-6 [96-100]. The parameters obtained from this process is analyzed and transformed to produce a readable MR image. This process is dependent on the magnetic field strength ($B_0$), gyromagnetic property of the nuclei ($\gamma$), the Larmour frequency ($\omega$) and is represented mathematically in Equation 1.4.

$$\omega = \gamma B_0$$ (1.4)

1.3.3 T1 and T2 Relaxation and factors contribute to relaxivity of contrast agents

The MRI signal loss which is the relaxation of water is defined as either spin-lattice relaxation time (T1) or spin-spin relaxation time (T2). When a proton is placed in a magnetic field, the overall magnetization will be on the z axis. When an RF pulse is given, then the net magnetization will be oriented in different directions if the RF pulse is in place. Immediately after that RF pulse is ended, the protons will relax to their original position.
All the MR contrast agents currently available for clinical use in the United States are small molecule Gd (III) complexes. Mn-DPDP, a manganese-based molecule, and a superparamagnetic iron-oxide nanoparticle (SPION) formulation (ferumoxide) were approved for liver imaging, and an oral SPION formulation was approved for gastrointestinal imaging. The observed relaxivity of gadolinium-based contrast agents has contributions from the water molecule(s) that bind directly to the gadolinium ion (inner-sphere water), long lived water molecules and exchangeable protons that make up the second sphere of coordination, and water molecules that diffuse near the contrast agent (outer-sphere).

Figure 1-7: Mechanism of T1 and T2 relaxation.
A. T1 curve plotting the recovery of longitudinal magnetization over time after the removal of a RF pulse, the T1 time specify the relaxation time in terms of the time required for the magnetization to reach 63% of its maximum. B. T2 curve plotting of the decay of transverse magnetization over during the relaxation process. The T2 time specify the relaxation time in terms of the time at which 37% of initial magnetization is present.
The T1 relaxivity values of commercially available GBCAs are relatively between 3.6 to 6.3 mM\(^{-1}\)s\(^{-1}\) which is small compared to what is theoretically possible [101]. Due to the similarities in the \(r_1\) and \(r_2\) values of majority of current GBCAs, it can be easy to speculate that GBCAs have the same shortening effect on both T1 and T2 relaxation times for tissues. However, this is not true for \textit{in vivo} studies [102]. Most biological tissues possess an intrinsic disparity in their native relaxation times with the native T1 5 to 20 x longer than T2 in the absence of application of contrast agent. They will inherently contribute to the unequal effects of gadolinium on observed relaxation rates resulting in a more observable effect of the GBCA on tissue T1 [67, 69, 103, 104]. Being either predominantly T1 or T2 is dependent on the background tissue relaxation time pre-contrast. For example, the proton density in the lung is at least 10-fold lower than other tissues thereby resulting in significantly lower lung MRI signal compared to other tissues such as liver, brain etc. [105]. The amount of contrast agent accumulating in the tissue of interest and can be explained using the relaxation rates (S\(^{-1}\)) of the contrast agent after administration described as the inverse of the relaxation times, \(\frac{1}{T_{i\text{obs}}}\). Accordingly, the observed T1 and T2 relaxation rates post contrast administration (1/T1obs and 1/T2obs) can be derived by summing up the rates contributed by the native tissue and the paramagnetic contrast agent, as shown in equation 1.5 and 1.6.

\[
\frac{1}{T_{i\text{obs}}} = \frac{1}{T_{i\text{b}}} + \frac{1}{T_{i\text{c}}}; \quad i = 1,2
\]  

(1.5)

\[
R_{i\text{ob}} = R_{i\text{b}} + R_{i\text{c}}; \quad i = 1,2
\]

(1.6)

Where 1/T_{i\text{obs}} and R_{i\text{obs}} are the observed relaxation rate post contrast administration with \(i = 1, 2\).

1/T_{i\text{b}} and R_{i\text{b}} are the buffer relaxation rate contributed by the tissue before contrast agent and
1/T_{ic} and R_{ic} are the relaxation rate contributed by the administered contrast agent.

The relationship between contrast agent concentration \([C]\) and its relaxivity is expressed mathematically as,

\[
\frac{1}{T_{ic}} = ri[C]; \ i = 1,2
\]

(1.7)

For relaxation to occur, changing the magnetic properties of water require the water proton to encounter a fluctuating magnetic field. In the absence of a contrast agent, this is usually achieved by encounter with other protons such as those in proteins, fat, or water. This intrinsic water proton undergoes Brownian motion while in circulation resulting in small fluctuating magnetic field which acts on other water protons to achieve relaxation. Dr. Xue in Yang lab [102] has shown by simulation using the Solomon-Bloembergen-Morgan (SBM) theories at \(\tau_R\) values ranging from 100 ps to 100 ns and clinical field strengths (1.5 T and 3T). He observed that clinical contrast agents such as magnevist® with \(\tau_R \sim 100\) ns have low \(r_2\) value (5 mM\(^{-1}\cdot\text{s}^{-1}\)) rendering them unable to result in any significant \textit{in vivo} \(T_2\) enhancement at physiologically permitted injection doses and feasible echo times. Thus, only \(r_1\) properties are practically application for in vivo studies.

Electrons generate better fluctuating magnetic field than protons and with Gd\(^{3+}\) possessing seven unpaired electrons, several factors act together to generate a significant fluctuating magnetic field to produce high relaxation rate as the gadolinium complex diffuses in solution. Correlation time (\(\tau_c\)) is the time required to characterize the fluctuating magnetic dipoles responsible for spin transitions and spin relaxation. Its reciprocal, \(1/\tau_c\), is the average rate constant for these fluctuating dipoles caused by electronic relaxation (1/T_{ie}) of Gd\(^{3+}\), rotational diffusion (1/\(\tau_R\)) of the H\(_2\)O-Gd\(^{3+}\) complex, water exchange in and out of the first coordination
sphere (1/τ_m) or the second coordination sphere (1/τ_m') also creates a fluctuating magnetic field for the proton nucleus. The shortest time constant of all the above determines the observed relaxation of the water proton and invariably the relaxivity of that contrast agent [53, 68, 71, 106, 107]. Gd^{3+} can bound and release over a million water protons per second. The rate in which the water is exchanged in and out of the inner coordination sphere of Gd^{3+} should be fast enough for maximum exchange with bulk water in the solvent surrounding of the Gd^{3+} metal.

As shown in Figure 1-4, protons from the inner, second and outer sphere water molecules contribute to the relaxation observed for GBCAs and can be represented by the equation below:

\[ r_{i_{obs}} = r_{i IS} + r_{i 2nd} + r_{i OS} \quad i = 1, 2 \]  

(1.8)

Where \( r_{i IS} \), \( r_{i 2nd} \) and \( r_{i OS} \) are relaxation rates corresponding to inner sphere, second sphere and outer sphere water. For small chelates, \( r_{i 2nd} \) and \( r_{i OS} \) contributions often were neglected. However, Xue et al from Yang lab has shown that they played important roles in relaxivity [102, 108, 109].

The inner and second sphere relaxivity can be defined by parameters such as the water number (q), relaxation time of water bound to the Gd^{3+} metal (T_{1m}), how long that water stays bound (mean residence time \( \tau_m \)) and the mole fraction of the Gd^{3+} per solvent (P_M) as shown in equation 1.9:

\[ \frac{1}{T_1} = \frac{P_M q}{T_{1m} + \tau_m} \]  

(1.9)

The mean residence time of bound water to Gd^{3+} is governed by several factors summarized in equation 1.1.
\[
\frac{1}{\tau_{1m}} = \frac{1}{T_{1DD}^{DD}} + \frac{1}{T_{1SC}^{SC}}
\]
(1.10)

Where \( T_{1DD} \) and \( T_{1SC} \) are relaxation times resulting from dipole-dipole interactions and scalar contact interactions, respectively. Both interactions also get contributions from several factors expressed mathematically as equations 1.11 to 1.12

\[
\frac{1}{T_{1DD}^{DD}} = \frac{2\gamma_I g^2 \mu_B^2}{15 r_{Gd-H}^6} S(S + 1) \left( \frac{\mu_0}{4\pi} \right)^2 \left( \frac{7\tau_{c2}}{1 + \omega_s^2 \tau_{c2}^2} + \frac{3\tau_{c1}}{1 + \omega_s^2 \tau_{c1}^2} \right)
\]
(1.11)

\[
\frac{1}{T_{1SC}^{SC}} = \frac{2S(S + 1)}{3} \left( \frac{A}{\hbar} \right)^2 \left( \frac{\tau_{e2}}{1 + \omega_s^2 \tau_{e2}^2} \right)
\]
(1.12)

Where \( \gamma_I \) is the nuclear gyromagnetic ratio, \( g \) is the electron g-factor, \( \mu_B \) is Bohr magneton, \( \mu_0 \) is permeability of vacuum, \( S \) is the spin quantum number of Gd\(^{3+} \) metal ion, \( r_{Gd-H} \) is the distance between Gd\(^{3+} \) and a water proton, \( \tau_c \) is the rotational correlation time, \( A/\hbar \) is the hyperfine coupling constant of electron-nuclear, \( \omega_I \) is the nuclear spin of water proton, and \( \omega_s \) is the electron spin resonance frequency[107, 110-112]. The influence of local magnetic field fluctuation contributed by electronic relaxation time of metal ion, rotational correlation time of metal water complex, and residence time of the bound water on rotational correlation time is given by

\[
\frac{1}{\tau_{ci}} = \frac{1}{\tau_R} + \frac{1}{\tau_m} + \frac{1}{T_{ie}} \quad i = 1,2
\]
(1.13)

From equation 1.19, \( \tau_{ie} \) representing the Zero Field Splitting (ZFS) of electron spin level resulting from collision between ions with spin greater than half and solvent molecules. \( \tau_{ie} \) is dependent on factors described by equations 1.14 and 1.15 below:

\[
\frac{1}{T_{1e}^{1e}} = 2C \left( \frac{1}{1 + \omega_s^2 \tau_{e}^2} + \frac{4}{1 + 4\omega_s^2 \tau_{e}^2} \right)
\]
(1.14)

\[
\frac{1}{T_{2e}^{2e}} = C \left( \frac{5}{1 + \omega_s^2 \tau_{e}^2} + \frac{2}{1 + 4\omega_s^2 \tau_{e}^2} + 3 \right)
\]
(1.15)
\[ C = \frac{1}{50} \Delta^2 \tau_v [4S(S + 1) - 3] \]  

(1.16)

Where \( \tau_v \) is the correlation time for ZFS distortion, \( \Delta \) is the mean-square zero-field-splitting energy.

Based on these equations, we know that \( \tau_e \) is dependent on \( T_{1e} \) which is turn is dependent on the magnitude and duration of ZFS. At low magnetic fields, the fluctuations are minimal and by extension, the \( 1/T_{1e} \) value will therefore be high enough to dominate \( \tau_c \). At high field, the opposite effect is experienced and the \( 1/T_{1e} \) becomes increasingly negligible as magnetic field is increased, high \( \tau_r \) then becomes an important factor in improving relaxivity. To generate an effective image, a contrast agent concentration of 0.1-0.6 mM of clinical approved contrast agents is needed [68, 113, 114]. The required high contrast agent concentration, i.e. low dose efficiency, hence, leads to high toxicity from gadolinium deposition and inability for molecular imaging. New techniques for contrast agent detection have emerged including the use of chemical exchange saturation transfer (CEST) technique, by saturating exchangeable hydrogen resonances which then chemically transfer this saturation to bulk water resulting in slightly decreased water signal [115-118].

Due to the low relaxivity and sensitivity of small chelator based GBCAs, several strategies have been employed to improve the observed \textit{in vivo} signal of MRI contrast agents. Some of the strategies employed includes encapsulation of Gd\(^{3+} \) ion into caged structures and micelles [119], Gadolinium chelation to polymers [120, 121], proteins and particles [109, 122-124], each with significantly improved relaxivity, sensitivity and blood circulation and retention time. A good example is the albumin-targeted MRI contrast agent (MS-325) which binds to blood serum albumin with significantly improved relaxivity and circulates with the blood
through veins and arteries (blood pool agent). MS-325 was approved for angiographic (blood vessel) imaging [125-128]. An alternative approach to the development of high relaxivity MRI contrast agent is the use of native and engineered proteins as MRI probes since about 30% of proteins already require metal cofactors such as calcium, copper, magnesium, zinc, iron, and manganese for proper folding, stability, and function [129-131]. To obtain high relaxivity contrast agents parameters governing the T1 relaxivity such as the hydration number (q), the rotational correlation time (τR), and the water residency time (τm) can be optimized [53, 101, 132, 133]. An increase in the rotational correlation time and hydration number q is increased (q = 1 in approved contrast agents), as well as a decrease in the coordinated water residency time τm to 1–30 ns (τm = 150–1000 ns for approved contrast agents) will significantly improve the relation rates of contrast agents [101, 109, 134]. New techniques for contrast agent detection have emerged including the use of chemical exchange saturation transfer (CEST) technique, by saturating exchangeable hydrogen resonances which then chemically transfer this saturation to bulk water resulting in slightly decreased water signal [115-118].

1.3.1  **Protein-based MRI contrast agents**

As shown in Equation 1.13, the longitudinal relaxation rate (r1) for small molecule-based MRI contrast agents is limited by their fast rotational motion τc at picoseconds. By slowing down the rotational motion of the Gd-L complex, the relaxivity of the complex can be significantly increased. This can be partially achieved by increasing the molecular weight of gadolinium complexes using polymers, dendrimers, proteins, polymerized liposomes, nanoparticle
emulsions, or viruses for complexation [135-140]. Hemoglobin, and engineered proteins like the heme domain of flavocytochrome P450-BM3 (BM3h) have been used to function as contrast agent to monitor activities in human brains as well as animals [141-144]. Furthermore, studies have shown that metalloproteins can be engineered to trap different metals such as iron in ferritin with T2 property making it a great target as T2-w MRI agents in in vivo applications [145, 146] with darkening effect. Their clinical applications are yet to be demonstrated.

Moreover, there are certain proteins with well folded structures and the right sizes, allowing for good circulation and tissue penetration. Globular proteins with molecular weight between 10–14 kDa correspond to a size of 2–3 nm and would result in easy filtration through the glomerulus [147]. Therefore, selecting a protein scaffold for Gd³⁺ MRI contrast agents with improved metal binding and relaxation properties and metal-binding properties will enable several modifications for molecular imaging such as addition of a biomarker targeting moiety for targeted MR imaging to target various disease biomarkers by protein modification.

Generally, there are two approaches to the development of protein-based Gd³⁺ contrast agents [133]. In the first approach, the Gd³⁺ chelate is conjugated to protein via covalent or non-covalent bonds. Gd³⁺ chelates were designed to interact non-covalently with proteins or directly via covalent bond as seen with MS-325, and because of the increased molecular weight of the complex, relaxivity is improved [125, 132, 148]. For example, in a study by Ogan et.al [149], Gd-DTPA chelates were covalently attached to a series of poly-L-lysine chains, with molecular weight ranging from 3kDa up to 102kDa. The longitudinal relaxivity, of these paramagnetically labeled albumins increased with increasing (DTPA-Gd³⁺)/BSA ratio which
is attributed to the larger molecular weight of the contrast agent and hence a higher rotational correlation time. Several other studies using human serum albumin (HSA) or bovine serum albumin (BSA) have been done, the results showed that the use of macromolecular compounds possesses a $r_1$ value that is 3x higher than that of the monomeric chelates [136, 149-151]. One of the major limitations of these contrast agents is the tendency for these large molecules chelated to Gd$^{3+}$ to experience internal motion. The rotational correlation time of a contrast agent at any given magnetic field have contributions from the slow overall motion of the large molecule and from faster local motions because of internal flexibility which directly influences the water residency time and hence the relaxivity of the contrast agent [152-154]. In addition, large molecular size also limits to their in vivo diffusion and resulted in prolonged retention time.

In the second approach, Gd$^{3+}$ binding pockets are designed into the proteins or the protein acts as a chelator for Gd$^{3+}$ metal binding [155]. Over the years, our lab has successfully designed protein based engineered proteins-based MRI contrast agents with high relaxivities that are 20-times higher than that of Gd-DTPA and are stable to de-chelation of Gd$^{3+}$ [102, 108, 109, 155-161]. By rational protein design, we developed three generations of protein-based MRI contrast agents utilizing well-folded proteins with appropriate size enabling good circulation time and tissue penetration as well as ease of excretion by staying within the acceptable excretion size ($\leq$70kDa) for the glomerular filtration [162]. The first generation of protein-based contrast agents was achieved by creating high-coordination Gd$^{3+}$ binding sites in a stable host protein using amino acid residues and water molecules as the metal coordinating ligands. Gd$^{3+}$ binding site was designed into domain 1 of cell adhesion protein 2 (CD2). This protein contrast agent tagged ProCA1 possessed the appropriate rotational correlation time (10–30 ns) for optimum
relaxivity, the right size (about 2.5 nm) and good circulation and tissue penetration [108, 161, 163]. Despite the several potential of our designed ProCA1 MRI contrast agents, it has a relatively low Gd$^{3+}$ binding affinity ($K_d = 8.7 \times 10^{-13}$ M), compared with clinical MRI contrast agents [$1.9 \times 10^{-21}$ M for DTPA and $1.4 \times 10^{-17}$ M for DTPA–bis(methylamide) (DTPA-BMA)], even though its metal selectivity for Gd$^{3+}$ over Mg$^{2+}$/Ca$^{2+}$/Zn$^{2+}$ are better than DTPA-BMA and Gd-DTPA [155].

1.4 Protein MRI contrast agent by design

Over the years, our lab has successfully designed protein based engineered Protein Contrast Agents (termed ProCAs) with high relaxivities. Gd$^{3+}$ binding pockets are designed into the proteins or the protein acts as a chelator for Gd$^{3+}$ metal bindinith 20-times higher than that of Gd-DTPA and are stable to dechelation of Gd$^{3+}$ [102, 108, 109, 156-158, 160, 164]. By rational protein design, we developed three generations of protein-based MRI contrast agents utilizing well-folded proteins with appropriate size enabling good circulation time and tissue penetration as well as ease of excretion by staying within the acceptable excretion size ($\leq 70kDa$) for the glomerular filtration [162]. The first generation of protein-based contrast agents was achieved by creating high-coordination Gd$^{3+}$ binding sites in a stable host protein using amino acid residues and water molecules as the metal coordinating ligands. Gd$^{3+}$ binding site was designed into domain 1 of cell adhesion protein 2 (CD2) [161]. This protein contrast agent tagged ProCA1 possessed the appropriate rotational correlation time (10–30 ns) for optimum relaxivity, the right size (about 2.5 nm) and good circulation and tissue penetration [108, 161,
Compared with clinical MRI contrast agents, ProCA1 exhibits stronger metal selectivity for Gd$^{3+}$ over Mg$^{2+}$/Ca$^{2+}$/Zn$^{2+}$ are better than DTPA-BMA and Gd-DTPA [155]. However, ProCA1 has a relatively lower Gd$^{3+}$ binding affinity ($K_d = 8.7 \times 10^{-13}$ M) than DTPA [1.9 $\times$ 10$^{-21}$ M] and DTPA–bis(methylamide) (DTPA-BMA)] (Omniscan) (1.4 $\times$ 10$^{-17}$ M). Drs. Li and Jiang from Yang lab has shown that PEGylation improves the relaxivity and biocompatibility ProCAs while Dr. Qiao et. al further demonstrated that using ProCA1 with engineered biomarker binding moiety can target tumors with high intratumoral distribution (Please see section 1.5 for molecular imaging).

1.4.1.1 Choice of natural metal binding proteins as scaffold protein for Protein MRI contrast agents

In the second generation of our lab designed protein contrast agent (ProCA2), scaffold protein with intrinsic metal binding properties such as EF-hand motifs was utilized. Calmodulin (CaM) was selected as the scaffold protein. This protein-based MRI contrast agent (ProCA2) was generated by modifying metal binding sites to optimize metal binding affinity of Gd$^{3+}$ which improved the metal binding affinity of ProCA2 compared to ProCA1.

The third generation (ProCA3 series) was developed to address the limitations of ProCA1 and ProCA2 for further metal binding stability. For this contrast agent, we selected alpha parvalbumin (PA) as the scaffold protein for chelation of Gd$^{3+}$metal during MR imaging [102, 164]. Parvalbumin (PA) is a classical small, mostly cytosolic, vertebrate specific Ca$^{2+}$ binding protein belonging to the EF-hand superfamily whose expression is usually tissue and cell-specific. They are classified as calcium buffers, calcium transporter or calcium shuttle protein [165, 166]. Parvalbumin is an acidic protein with molecular weight ~ 12kDa and isoelectric point (pI) of about 5. Both the rat and the human counterpart contains a total of 110 amino
acids with just about 8% difference in primary structure. In mammals, parvalbumin is found in fast muscle, brain, and several endocrine glands, where it is thought to constitute an important calcium buffer. They exist either as the α-parvalbumin isoform containing 109-110 residues or the β-isofom with 108 residues. Parvalbumin contains two functional EF-hand calcium binding motifs [167].

Figure 1-8 EF-Hand $Ca^{2+}$ binding motif with the characteristics preferred residue[168] In addition to preferential oxygen binding, a preference is also made for negatively charged residues Aspartate and glutamate. Parvalbumin contains three paired helix-loop-helix EF-hand motifs AB motifs houses a two amino-acid deletion in the loop region.

As shown in Figure 1-8, the EF-hand motif is composed primarily of a helix-loop-helix structural component. The “loop” is the calcium binding site of the protein indicated with red colored balls. α-parvalbumin has three EF hand motifs and is structurally related to CaM and troponin C.

The choice of alpha parvalbumin as scaffold for our contrast agent is based on its high thermal stability, high solubility, high resistance to the enzyme cleavage, and high tolerance mutations [102, 164, 168]. The calcium binding affinity (dissociation constant) $K_d$ of parvalbumin is
estimated as $8.35 \pm 0.29 \times 10^{-9}$ M. The resulting protein contrast agent (ProCA3) was developed with several mutations and upon characterization and detailed analysis, ProCA32 was selected for use in MR imaging [102, 108, 109, 164]. Rat ProCA32 was developed by Dr. Xue in Yang lab by creating an additional negatively charged ligand in the metal binding site of rat alpha parvalbumin by optimizing hydration number while achieving high metal binding affinity and specificity. This was achieved with a S56D mutation to enhance Gd$^{3+}$ binding affinity and selectivity. Furthermore, F103W mutation was made to introduce luminescence resonance energy transfer (LRET) property to the protein to determine its metal binding affinity and water number by Tb$^{3+}$ luminescence lifetime decay. The resulting contrast agents are termed ProCAs (Protein Contrast Agents). We have shown that rProCA32 has q=0.5 shared by the two binding sites of Gd$^{3+}$ sufficient for increased ProCA (both rat-based and human-based) relaxivity [102, 108, 109, 159, 160, 164]. One of major goal of this dissertation is to further develop ProCA32 for molecular imaging and clinical applications.

1.4.2 High relaxation properties of ProCAs

The relaxivity of a contrast agent is linearly correlated with its sensitivity [69, 90]. All clinically approved MRI contrast agents (e.g. Gd-DTPA, Gd-BOPTA and Gd-EOB-DTPA) have relaxivities between 3.6 to 6.3 mM$^{-1}$s$^{-1}$ at 1.5 and 3T, these are relatively low values since the estimated local concentrations of Gd$^{3+}$ must be around 125 µM to robustly observe contrast differences in tissue [133]. The first parameter contributing to the high relaxivity value of our developed ProCAs is its optimized hydration number. As shown in previous equations, the relaxivity of a contrast agent has a linear relationship with water number (q). An increase in the number of water molecules coordinated to the Gd$^{3+}$ ion may directly increase the relaxivity of a contrast agent. In a study to develop Egad, an enzyme responsive contrast agent, water
access to Gd$^{3+}$ ($q = 0$) was disabled due to the addition of a galactopyranose moiety to a Gd-DO3A chelate. However, upon galactosidase activation, galactopyranose sugar was cleaved from Gd$^{3+}$, thereby allowing access to water molecules, resulting in increased water number from 0.7 to 1.2 [169] and a subsequent increase in the $r1$ value from 0.90 mM$^{-1}$ s$^{-1}$ to 2.72 mM$^{-1}$ s$^{-1}$ [170]. On the other hand, the stable $q = 2$ complex of Gd-DO3A is capable of forming ternary complexes with phosphate or other endogenous coordinating anions such as bicarbonate resulting in the displacement of the coordinating water ligands, and a subsequent reduction in $T_1$ relaxivity [171]. It is therefore important to appropriately tune the hydration number without compromising relaxivity. We have shown that ProCA32 has $q$=0.5 shared by the two binding sites of Gd$^{3+}$ sufficient for increased ProCA (both rat based and human based) relaxivity [102, 108, 109, 159, 160, 164]. Moreover, larger surface of proteins can also contribute to high relaxivity through second sphere relaxivity that could reach as high as 3.3 and 8.8 mM$^{-1}$ s$^{-1}$ at 20 and 60 MHz, respectively ($\tau R = 10$ ns, $\tau m = 10$ ns, $q = 4$, $r_{GdH} = 5$ Å$^\circ$). According to all the parameters discussed above, it is highly likely that hProCA32.collagen high relaxivity benefits from coordination water from the inner, second, and outer spheres of which cause the increase of relaxivity.

1.5 Molecular Magnetic resonance Imaging by biomarker-targeted contrast agents

In the current clinical setting, one of the major limitations of imaging modalities is the lack of precise and accurate detection, characterization, and quantification of relevant pathobiological processes involved in these diseases at molecular and cellular levels [172, 173]. For instance,
the use of biopsy has associated invasiveness, high costs and sampling error, post biopsy trauma that may lead to hospitalization in up to 6.9% of biopsy cases [174], intra/interobserver variability [175, 176], mortality in 0.01–0.1% of cases [177-179], lack of accuracy due to substantial sampling error (for instance up to 33% is observed in liver cirrhosis), cost associated morbidity, pain and bruising at the biopsy site, prolonged bleeding, infection of the biopsy site, and pneumothorax and hemothorax [32, 180, 181].

In addition to developing contrast agent with improved relaxivity for lower dosage, another approach to overcome the safety and sensitivity limitations associated with the current MRI contrast agents is to develop MRI contrast agents with targeting capability to certain disease molecular biomarkers for specific biomarker detection. Typically, a biomarker is a biological feature that can be easily identified and quantified as an indicator of a normal or pathological biological condition [182]. MRI contrast agents biochemically targeted are designed to specifically adhere to molecular biomarkers expressed in different pathological conditions thus providing a retained image enhancement beyond the capability of a non-targeted agent.

Molecular imaging describes the visualization and quantitative characterization of biological processes in living organism both at the cellular and molecular levels [173]. Most widely used molecular imaging techniques include computed tomography, single photon emission computed tomography, optical imaging, ultrasound imaging, magnetic resonance imaging, and magnetic resonance spectroscopy. They typically require the use of either endogenous or exogenous probes/contrast agent for better image quality by putting the imaging cellular events, such as gene expression, and their downstream modalities, diseases and organ sites into perspective [173, 183]. It aids medical diagnosis by directly measuring the therapeutic effect of drugs in addition to providing visualization of the presence, severity and location of
diseases. This capability provides a tool for image-guided therapies and treatment decisions non-invasively[184]. Molecular imaging affords direct quantitation of molecular targets and biomarkers such as collagen, human epidermal growth factor (EGFR) human epidermal growth factor 2 (HER2), elastin, fibrin etc.[173, 184, 185]. Molecular imaging aims to address the diagnostic gaps currently observed in clinical practices that conventional tests have failed to address to improve patient care. There have been several diagnostic gaps observed including limitations associated with early detection of several diseases such as organ fibrosis, image-guided identification of different tumor subtypes, quantitative monitoring of disease progression and drug treatment effect, especially for targeted therapy in patients classified as high risks. The ability of molecular imaging to enable specific targeting of different biomarkers or proteins in several organs and tissues provides an avenue for obtaining information on key pathophysiologic processes which will in no small way enable early detection of several pathologies, and management and ultimately improving prognosis and patient survival. [186][184]. Molecular MRI possesses a unique capacity for reproducible and non-invasive assessment of biological and pathological processes in various organ. Molecular MRI (mMRI) of disease biomarkers can be an indispensable and ideal diagnostic tool to elucidate the underlying mechanism of diseases, monitoring disease progression, drug treatment and activity during preclinical and clinical application. However, one of the major limitations of mMRI in the investigation of specific disease biomarkers for diagnosis and drug treatment monitoring is the lack of sensitive contrast agents with high relaxivity, targeting capability/specificity, and tissue permeability.
Molecular probes employed in molecular imaging are small molecules, antibodies, nucleic acids, or peptides tagged to a chemical label to enable specific identification and visualization of receptors, proteins, or biological processes [187, 188]. They can also be referred to as contrast agents, molecular beacon, tracers. Generally, they consist of a chemical agent which has the required chemical properties for the principles surrounding the imaging modality for molecular read out (for instance, a paramagnetic property is required for a metal to be usable in MRI), a targeting moiety which possesses high affinity for the target molecular biomarker or receptor and a flexible linker connecting both. Targeting moiety is expected to specifically bind to a target such as a biomarker, receptor, DNA sequence or protein. The majority of disease biomarkers have very low expression levels such as collagen type I with expression level around 1-20µM [187, 189] and receptors usually at nanomolar or sub micromolar range. Since current clinically approved contrast agent requires to have a local concentration of Gd$^{3+}$ = 125 µM (which corresponds to a relaxation rate change of about 0.5 s$^{-1}$) to observe contrast differences in tissue [133, 190]. Targeted MR contrast agents with similar T$_1$ relaxivity values using existing chelator contrast agents, will not be able to image molecular biomarkers.

Development of targeted contrast agents with higher relaxivity will potentially provide better tissue enhancement and by extension, allow for disease signature mapping, detection of smaller lesions, or better contrast with lower dose relative to current approved compounds. Majority of disease biomarkers in different organs have very low expression levels and targeted contrast agents, using high relaxivity compounds would allow the detection of these lower concentration targets [191, 192]. To mitigate this, use of molecular probes that specifically targets biomarkers associated with different diseases have been described [188, 193, 194].
Efforts have been put into the development of targeted mMRI contrast agents with improved sensitivity [195, 196]. These MRI contrast agents target biomarkers expressed on the surface of blood vessels and not so many have tried to target the biomarkers outside of the blood vessels. To improve the dose efficiency of MRI contrast agents, a peptide fragments were developed with high affinity to biomarkers, applied by conjugation but have slow diffusion when targeting tumors in vivo. In addition, antibodies does not have the proper distribution and penetration in tumor blood vessel [157]. On the other hand, peptides have a size small enough for good tumor penetration [197, 198]. Despite these great achievements in targeted contrast agent development, peptide-based MRI agents need further improvement in targeting capability and resistance to protease cleavage. There is therefore an unmet medical need for the development of MRI contrast agents with higher relaxivity. Early detection of diseases usually corresponds to lower overexpression of biomarkers and is crucial to patient survival and better prognosis outcome. Therefore, development of sensitive MRI contrast agents with high sensitivity, metal binding affinity, selectivity, and biomarker-targeting capability is needed to overcome the limitations associated with molecular MRI early detection of different chronic diseases such as liver fibrosis, monitoring progression and possible disease reversal [199-201].

1.5.1 **Role of Collagen as disease diagnosis biomarker**

A common denominator crucial for normal tissue development and function is the extracellular matrix [202] which consists of proteins and carbohydrates that supports different biological structures and processes such as tissue development, elasticity, and preservation of organ structure Figure 1-9 [187, 203]. Diseases involving inflammation, organ fibrosis, tumor invasion, and injury are all attributed to the transition of the extracellular matrix (ECM) from
homeostasis to remodeling, which can significantly change the biochemical and biomechanical features of ECM components.

Figure 1-9: Component of the extracellular matrix. Extracellular matrix undergoes remodeling, a sequence of quantitative and qualitative changes in the ECM composition during disease development and progression. In fibrosis, cancer, and many pathological conditions, this composition becomes dysregulated, stiff, and overexpressed[187].
Collagens are proteins present in multicellular eukaryotic organisms. They are the most abundant group of proteins in connective tissues and the extracellular matrix. The extracellular matrix (ECM) is the noncellular component of tissues and organs providing essential mechanical support while performing several regulatory functions for cells. The main components of the ECM are glycoproteins, proteoglycans, elastin, microfibrillar proteins and collagens. They control important cellular events such as cell migration, differentiation, proliferation, adhesion, and survival. An equilibrium in physiological processes is important for normal function of tissue. A disruption in this equilibrium results in development and progression of a series of pathological conditions such as observed in osteoarthritis, cancer, pulmonary fibrosis, liver fibrosis, chronic kidney diseases, and atherosclerosis. [204, 205].

Collagen is the most abundant and diverse ECM molecule making up 25% to 35% of human body protein [206]. The collagen superfamily is made up of 28 members (Table 1-2). A structural feature common to all 28 types is the presence of a triple helix. Collagen type I, a fibrillar collagen is the most abundant collagen making up about 90% of collagen types [207]. The fibrous, structural protein is responsible for the tensile strength of the ECM and contains three polypeptide α-chains, displaying a polyproline-II conformation, a right-handed supercoil and a glycine residue in the center for flexibility and close parking between adjacent chains. [208-210]. It is the main component of the organic part of bone, skin, tendon, organs, and vascular ligature. The ECM protein collagen is absent in normal tissue or present at very low levels and as the most abundant collagen in organs, abnormal and uncontrolled overexpression of collagen in organs such as lungs, liver, and vasculature can lead to collagen-induced arthritis, fibrosis and several collagen vascular diseases (Table 1-2). It has therefore been found to be of particular importance as diagnostics and therapeutic biomarker of these disorders by
quantitation of collagen overexpression or degradation. Another importance factor is that collagen is easily accessible as noncellular component. Furthermore, MRI contrast agent capable of estimating collagen levels with addition capability to bind cross-linked collagen would revolutionize detection of early stages of diseases such as fibrosis, cirrhosis and further improve drug treatment monitoring. The current research-based collagen-targeted contrast agents for detection of fibrosis, myocardial infarction or pancreatic cancer are based on small molecules which may limit the contrast agent in attaining the required sensitivity and specificity for early detection of diseases [90, 195, 211].

Table 1-2: The collagen superfamily. [202, 207, 212-215]. (FACIT fibril-associated collagens with interrupted triple helices, MACIT membrane-associated collagens with interrupted triple helices, MULTIPLEXINs multiple triple-helix domains and interruptions)

<table>
<thead>
<tr>
<th>Type of collagen</th>
<th>Class</th>
<th>Location</th>
<th>Associated diseases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collagen I</td>
<td>Fibril</td>
<td>Non-cartilaginous connective tissue: dermis, bone, tendon, ligament</td>
<td>Osteogenesis imperfecta, Ehlers–Danlos syndrome, Infantile cortical hyperostosis</td>
</tr>
<tr>
<td>Collagen II</td>
<td>Fibril</td>
<td>Cartilage, vitreous</td>
<td>Ehlers–Danlos syndrome, Dupuytren’s contracture</td>
</tr>
<tr>
<td>Collagen III</td>
<td>Fibril</td>
<td>Co-distribution with collagen I: skin, blood vessels, intestine</td>
<td>Classical Ehlers–Danlos syndrome</td>
</tr>
<tr>
<td>Collagen IV</td>
<td>Network</td>
<td>Basement membrane</td>
<td>Alport syndrome, Goodpasture’s syndrome</td>
</tr>
<tr>
<td>Collagen V</td>
<td>Fibril</td>
<td>Co-distribution with collagen I: bone, dermis, cornea, placenta</td>
<td>Ulrich myopathy, Bethlem myopathy, Atopic dermatitis</td>
</tr>
<tr>
<td>Collagen VI</td>
<td>Network</td>
<td>Muscle, bone, cartilage, cornea, dermis</td>
<td>Epidermolysis bullosa dystrophica</td>
</tr>
<tr>
<td>Collagen VII</td>
<td>FACIT</td>
<td>Dermis, bladder</td>
<td>Posterior polymorphous corneal dystrophy 2</td>
</tr>
<tr>
<td>Collagen VIII</td>
<td>Network</td>
<td>Dermis, brain, heart, kidney</td>
<td>Multiple Epiphyseal Dysplasia Type 2 (EDM2) and EDM3</td>
</tr>
<tr>
<td>Collagen IX</td>
<td>FACIT</td>
<td>Co-distribution with collagen II: cartilage, cornea, vitreous</td>
<td>Schmid metaphyseal dysplasia</td>
</tr>
<tr>
<td>Collagen X</td>
<td>Network</td>
<td>Hypertrophic cartilage</td>
<td>Collagenopathies</td>
</tr>
<tr>
<td>Collagen XI</td>
<td>Fibril</td>
<td>Co-distribution with collagen II: cartilage, intervertebral disc</td>
<td>-</td>
</tr>
<tr>
<td>Collagen XII</td>
<td>FACIT</td>
<td>Co-distribution with collagen I: dermis, tendon</td>
<td>-</td>
</tr>
<tr>
<td>Collagen XIII</td>
<td>MACIT</td>
<td>Endothelial cells, dermis, eye, heart</td>
<td>-</td>
</tr>
<tr>
<td>Collagen XIV</td>
<td>FACIT</td>
<td>Co-distribution with collagen I: bone, dermis, cartilage</td>
<td>-</td>
</tr>
<tr>
<td>Collagen XV</td>
<td>Multiplexin</td>
<td>Located between collagen fibrils that are close to BM, capillaries, testis, kidney, heart.</td>
<td>-</td>
</tr>
<tr>
<td>Collagen XVI</td>
<td>FACIT</td>
<td>Integrated into collagen fibrils and fibrillin-1 microfibrils, dermis, kidney</td>
<td>-</td>
</tr>
<tr>
<td>Collagen XVII</td>
<td>MACIT</td>
<td>Hemidesmosomes in epithelia</td>
<td>Bullous pemphigoid and certain forms of junctional epidermolysis bullosa</td>
</tr>
<tr>
<td>Collagen XVIII</td>
<td>Multiplexin</td>
<td>Basement membrane, liver</td>
<td>Knobloch syndrome</td>
</tr>
<tr>
<td>Collagen XIX</td>
<td>FACIT</td>
<td>Rare, basement membrane</td>
<td>-</td>
</tr>
<tr>
<td>Collagen XX</td>
<td>FACIT</td>
<td>Cornea (chick)</td>
<td>-</td>
</tr>
<tr>
<td>Collagen XXI</td>
<td>FACIT</td>
<td>Stomach, kidney</td>
<td>-</td>
</tr>
<tr>
<td>Collagen XXII</td>
<td>FACIT</td>
<td>Tissue junctions</td>
<td>-</td>
</tr>
<tr>
<td>Collagen XXIII</td>
<td>MACIT</td>
<td>Heart, retina</td>
<td>-</td>
</tr>
<tr>
<td>Collagen XXIV</td>
<td>Fibril</td>
<td>Bone, cornea</td>
<td>-</td>
</tr>
<tr>
<td>Collagen XXV</td>
<td>MACIT</td>
<td>Brain, heart and testis</td>
<td>-</td>
</tr>
</tbody>
</table>
A major hallmark of fibrosis is uncontrolled accumulation of collagen. Fibrosis also described as tissue scarring is marked by overexpression of proteins such as collagen in the ECM in response to chronic tissue injury leading to uncontrolled accumulation of scar tissue and eventually to organ failure and death. 45% of deaths in the United States is as a result of fibrotic disorders[216]. Organ fibrosis typically stems from inadequate wound healing responses to ongoing tissue injury resulting in deposition of ECM and tissue remodeling[217]. Fibrosis is a major disease for an increasing number of individuals and is a common pathological condition of many persistent inflammatory diseases, such as idiopathic pulmonary fibrosis (IPF), progressive kidney disease, liver cirrhosis and chronic obstructive pulmonary fibrosis (COPD) (Figure 1-10). One of the major sources of ECM in hepatic pathologies are activated hepatic stellate cells (HSCs)[218]. Recent studies have highlighted the role of the ECM and shown the importance of deregulated ECM dynamics in molecular pathway and disease development mechanism. Uncontrolled ECM production or a reduction in ECM turnover are contributing factors to the development of several organ fibrosis [219, 220]. Collagen type I is an appealing target for molecular imaging of a variety of diseases such as liver fibrosis, liver metastasis, breast cancer, pancreatic cancer and aortic aneurysms. In organ fibrosis and different cancers, collagen accumulation as well as its concentration increases by the disease progression. For example, pulmonary fibrosis marked by excessive collagen accumulation occurs in conditions such as chronic obstructive pulmonary disease [221] and interstitial lung diseases [222]. For all these pathologies it would be useful to non-invasively detect, assess, and monitor fibrosis. Treatment decisions hinge on both the identity and severity of the disease. For instance, non-alcoholic steatohepatitis (NASH), which afflicts 1–2% of the U.S. population, progresses to
cirrhosis of the liver in about 20% of cases. Early detection and accurate characterization of these diseases can improve patient outcomes. Another attractive feature of collagen I is its accessibility as an ECM protein further affirming its clinical translation potential.

Figure 1-10 Pathogenesis of fibrosis in different organs

Developments in MRI probes targeting ECM proteins such as collagen, fibronectin, and elastin via conjugation of existing contrast agents to targeting moieties and their applications to various diseases is an active area of research. There are many different conjugation strategies reported for development of these targeted probes, however the translation of these contrast agents into clinic is a challenging process which relies on the thermodynamic stability and kinetic inertness of the Gd$^{3+}$ chelating unit, its conjugation to a targeting moiety with strong affinity and specificity towards the desired target, improved relaxivity; superior pharmacokinetic properties and fast clearance. In addition, the targeted probe requires good acquisition time or imaging window with no in vivo toxicity. Better understanding of the structure of the ligands along with choice of a good target, facilitates the rational design of targeted probes. At the end, in a complex environment such as ECM, addressing specificity,
heterogeneity and various patterns of targets as well as detection of ECM changes at early stages of the disease is still a major unmet medical challenge. In this work, collagen type I was identified as the appropriate biomarker for different disease diagnosis.

1.6 Current advances in the development of collagen-targeting peptides

Diseases involving fibrosis, inflammation, tumor invasion, and injury are all associated with the transition of ECM from homeostasis to remodeling. With collagen as the most abundant protein, its presence in the ECM of most tissues is required as part of normal tissue development and balance maintenance. They also aid in providing physical support to tissues by occupying the intercellular space, acting as a flexible, dynamic and mobile contributor to defining cellular behaviors and tissue function [223]. Thus, abnormal or overexpression of collagen can provide a quantitative parameter useful for monitoring progression and improvement of several diseases including organ fibrosis, cancer, arthritis, and other collagen vascular diseases [224, 225]. Several works directed at targeting either overexpressed or degraded collagen for diagnostic or therapeutic interventions will be reviewed in this section. Many ECM-binding proteins contain specific domains that have affinity to collagen. Takagi et. al reported a collagen-binding decapeptide derived from bovine propolypeptide of the von Willebrand factor. The collagen binding domain (CBD) \textbf{WREPSFCALS} sequence was reported to bind to the 21.5/21-kDa fragment to collagen/gelatin. Using solid phase adsorption assay, the dissociation constant \( K_d = 100 \, \mu M \) for collagen type I was reported [226]. Dai and colleagues described their development of a collagen binding peptide (CBP), \textbf{TKKTLRT} from collagenase. They observed that the sequence inhibited the diffusion of various growth factors while increasing its tissue retention capability thereby leading to improve wound repair.
Because this protein was originally designed for inhibiting collagenase activity, an apparent inhibition constant of 0.3 µM was reported for anti-TKKTLRT [227]. TKKTLRT was conjugated with vascular endothelial growth factor (VEGF) to obtain CBD-VEGF, a fusion protein which was reported to specifically bind type I collagen in vitro. In this study, the TKKTLRT collagen binding peptide was employed as a delivery system for VEGF for improvement of cardiac function and angiogenesis. The reported kd of CBD-VEGF by ELISA assay was 0.43 µM [228]. Narayana et.al reported a ‘collagen-hug’ model (designed to allow multidomain collagen binding proteins to bind their extended rope-like ligand) for the binding of the sub segment of the collagen adhesion (CNA) protein isolated from the cell surface of Staphylococcus aureus [229]. The intrinsic collagen-binding domain, referred to as CNA35 has two soluble domains, N1 (residues 31 – 140) and N2 (residues 141 – 344) connected by a long linker (residue 164 to 173). The reported Kd value for CNA-35 is 0.2µM, although a range of values up to 30µM has been reported depending on binding assay employed and the CNA conjugating partner. For example, CNA was used in the fluorescent imaging of atherosclerosis upon conjugation with Oregon green to obtain CNA35-OG488 with a Kd value of 0.5µM [230-233]. A major limitation of the use of this probe is its inability to cross the endothelial barrier [233]. Decorin, an ECM protein has been extensively utilized in the development of collagen-targeting peptides because of its proven role in wound healing and fibrinogenesis in addition to its ability to bind collagen, especially type I, II, III, VI and XIV [234-238]. It has an intrinsic collagen I binding dissociation constant of $3.5 \times 10^{-9}$ M [239]. Decorin belongs to the small leucine-rich proteoglycan (SLRPs) family, it participates in ECM assembly and regulates the bioactivities of cell growth factors. Several studies have shown that decorin acts as a ligand of several cytokines and growth factors by directly or indirectly
interacting with the corresponding signaling molecules involved in cell growth, differentiation, proliferation, adhesion, and metastasis [240-242]. Decorin consists of a collagen-binding protein core containing 12 tandem leucine-rich repeats (LRR) attached to a dermatan sulfate (DS) glycosaminoglycan (GAG) side chain (Figure 1-11). Therefore, by mimicking the binding epitopes of collagen-decorin interactions, small peptides can be developed as biomimetic systems capable of targeting collagen. Researchers therefore explores the structural information from both decorin and collagen to obtain a suitable binding approach for complex formation of both molecules.
Kresse et al. reported the importance of the sixth leucine-rich repeat (Met\textsuperscript{176} to Lys\textsuperscript{201}) in decorin and more importantly, the critical role of glutamate 180 in collagen binding [244]. Using the RGD (arginine-glycine-aspartic acid) cell-attachment sequence and two distinct glutamic acid (glu)-rich regions present in mammalian bone sialoprotein (BSP) in conjugation with the Met\textsuperscript{176} to Lys\textsuperscript{201} collagen binding domain (MIVI\_ELGTNPLKSSGIENGAFQGMKK),
Goldberg et al. reported a binding affinity of 0.394 \( \mu \text{M} \) for this binding domain by solid state adsorption method [245, 246]. Stuart et al. reported a synthetic collagen-binding peptidoglycan (DS-SILY) with the capability to inhibit matrix metalloproteinase-1 and 13 (MMP-1 and MMP-13) -mediated collagen degradation by targeting collagen fibril organization to restore elements of healthy tissue. The peptide sequence: \textbf{RRANAALKAGELYKSILYG C} was derived from the collagen I platelet receptor following mutation of cysteine residue to serine, a modification which allowed dermatan sulfate (DS) to be conjugated to the terminus of the peptide. DS-SILY has a dissociation constant of 0.86 \( \mu \text{M} \) which was discovered to have been majorly produced from the GELYKSILY amino acid portion of the residues [247]. According to the docking studies and homology modeling experiment conducted by Lozzo et al, collagen has a high binding specificity for the inner surface of decorin [248]. More specifically, the major decorin-binding site in collagen was reported as the \( \text{c1} \) band of the collagen fibril, near one of the major intermolecular cross-linking sites. Additionally, using circular dichroism, Pearson and coworkers [249] observed that interaction with collagen is independent of the secondary structure of decorin but instead relies on the presence and preservation of the disulfide bridges. McQuillan et al further reported that mutations in the inner surface of decorin interfered with collagen binding [250] and there is a need for proper folding of the protein core of decorin to effectively bind to collagen [251]. Based on these findings, Neffe et al. reported residues \textbf{LHERHLNNN} belonging to the inner surface of decorin as the peptide with the highest collagen I binding amongst a total of 8 screened peptides. In their studies, the first set of peptides consists of six inner surface and two outer surface peptides respectively. A general LRR consensus sequences for proteins with highly repetitive structures (\textbf{LxxLxLxxN}\( \gamma \)L) was developed enabling the identification of specific binding peptides with
relevance as binding epitopes for protein–protein interaction. Using Surface Plasmon resonance (SPR) detection technique, they observed that all but one of the peptides based on the inner surface consensus showed significant binding potential with collagen I with Kd values ranging from 0.17 µM to 220 µM, while the two sequences corresponding to the outer surface of decorin showed no binding. No further use of this peptide has been reported. Lee et al. constructed a phage display experiment to express both Collagen mimetic peptide (CMP) and streptavidin binding peptide using recombinant DNA techniques by taking advantage of the conserved amino acid sequence naturally existing in collagen referred to as CMP. The CMP motif contains Gly-Pro-Pro repeats; 7GPP (Gly-Pro-Pro)\textsubscript{7} on the minor coat proteins. The construct, M\textsubscript{13}-7GPP\textsubscript{III}, was able to selectively bind several types of collagens using five copies of the CMP motif with a dissociation constant of $1.5 \times 10^{10}$ phages/mL [252]. Next, Yang et al utilized pro-MMP-2, the precursor of Matrix metalloproteinase-2 (MMP-2), one of the enzymes that interact with collagen type I to develop a collagen-targeted probe CBP1495 with sequence CPKESCNLFVLKD after it was accidentally discovered that CBP1495 can bind to both the CMP (GPO)\textsubscript{9} and type I collagen with Kd values of 0.63 µM and 0.86 µM respectively using ELISA method. This was employed as a collagen-targeted SPECT/CT imaging probe for pulmonary and hepatic fibrosis by chelation to $^{99}$Tc to obtain $^{99m}$Tc-CBP1495 prepared using Gly-(D)-Ala-Gly-Gly as a chelator [253]. Jandrot-Perrus’ lab [254] developed the RVMHGLHLDDE (Collagelin) collagen binding sequence by taking advantage of the collagen binding specificity of Glycoprotein VI (GPVI) in the main platelet receptor of types I and III collagen. An anti-GPVI antibody that neutralizes collagen-binding was used to screen a bacterial random peptide library, identifying a cyclic motif. Using surface plasmon resonance (SPR) technique, a binding affinity for collagelin to type-I collagen of 0.1
μM was reported. Collagen binding and imaging capability was tested in myocardial infarction rat models histologically and with $^{99m}$Tc labelled collagen SPECT imaging [254]. It is worthy of note that most of these peptides have not been used in in vivo studies especially for MRI diagnostic purposes.

Caravan et al. reported a cyclic, disulphide-containing polypeptide by phage display as a collagen specific substrate after systemic substitution of amino acids nonessential for collagen binding by alanine [132]. The collagen-targeting probes termed EP-3533 consisting of a peptide of 16 amino acids (GKWHCTTKFPHHYCLY) and 3 Gd-DTPA chelators connected to the N-terminal of the peptide and lysine ε-side chains via thiourea linkage. EP-3533 is formed because of three amino acids flanking a cyclic peptide of 10 amino acids that is formed through a disulfide bond. Using the percentage of bound compound versus collagen ratio where “bound” represents the amount of EP-3533 bound to the incubated collagen, the dissociation constant of 1.8 ± 1.0 μM with N>7 equivalent binding sites was reported [195]. EP-3533 was used in MRI for staging and treatment response monitoring in liver fibrosis [255, 256], detection and quantification of myocardial infarction [257], pancreatic cancer [258] and pulmonary fibrosis animal models [189]. EP 3533 possesses a relaxivity higher than Gd-DTPA but about 3x lower than protein-based MS-325 blood pool agent. In addition to this, the use of a linear gadopentetate dimeglumine (Gd-DTPA) chelate resulted in Gd$^{3+}$ retention in the kidneys and bone of mice 24 hours after injection indicating the presence of free unchelated Gd$^{3+}$ and a high risk for nephrogenic systemic fibrosis [188, 195, 259, 260]. To address the toxicity associated with EP 3533 probe, an improved version, CM-101 was developed by conjugating the 16-amino acid peptide to three t-butyl–protected macrocyclic chelator DOTA. While this probe possesses negligible Gd$^{3+}$ deposition compared to EP-3533, the relaxivity
values dropped by 31% and 62.5% for r₁ and r₂. Since relaxation rates (r₁ and r₂) are linearly correlated to the sensitivity of the contrast agent in molecular MR imaging, this may potentially result in less sensitivity and a high dose requirement for CM-101. Furthermore, the Kd value for CM-101 doubled implicating that the collagen binding affinity decreased by 50% [211, 259]. A summary of these collagen binding peptides and their binding properties are shown in Table 1-3. Several limitations associated to these peptides have necessitated the need for the development of a collagen specific MR imaging probe with high relaxivity, biomarker targeting binding and specificity while maintaining low or no toxicity.

**Table 1-3: Collagen-binding peptides and their applications.**

<table>
<thead>
<tr>
<th>Targeting Peptide</th>
<th>Target</th>
<th>Agent</th>
<th>Application</th>
<th>Binding Affinity (Kd)</th>
<th>Binding studies used</th>
</tr>
</thead>
<tbody>
<tr>
<td>WREPSFCALS (von Willebrand factor)</td>
<td>Collagen Type I</td>
<td>V-bFGF</td>
<td>Promote vascularization and cellularization</td>
<td>100 µM</td>
<td>Solid phase</td>
</tr>
<tr>
<td>MIVELGTNPLOCKSGENGAFQGMKK (Leucine-rich repeats (LRR))</td>
<td>Collagen Type I</td>
<td>P2S Collagen Binding Peptide</td>
<td>Induce collagen mineralization in bone defects</td>
<td>0.394 µM</td>
<td>Solid phase</td>
</tr>
<tr>
<td>N1 and N2 domains of a collagen adhesion protein from Staphylococcus aureus bacteria.</td>
<td>Collagen type I, III and IV</td>
<td>CNA35</td>
<td></td>
<td>0.2 µM to 30 µM</td>
<td>Solid phase</td>
</tr>
<tr>
<td>N1 and N2 domains of a collagen adhesion protein from Staphylococcus aureus bacteria.</td>
<td>Collagen type II, III and IV</td>
<td>CNA35-OG488</td>
<td></td>
<td>0.5 µM</td>
<td>Solid phase</td>
</tr>
<tr>
<td>TKKTLRT (Collagenase)</td>
<td>Collagen type I</td>
<td>C-bFGF</td>
<td>Promote vascularization and cellularization</td>
<td>≤100 µM</td>
<td>ELISA</td>
</tr>
<tr>
<td>CBD-VEGF</td>
<td>Collagen type I</td>
<td>CBD-VEGF</td>
<td>Induce angiogenesis and improve cardiac function</td>
<td>0.43 µM</td>
<td>ELISA</td>
</tr>
<tr>
<td>GPPGPgpGPPGPPGPPGPPGPP (CMP)</td>
<td>Collagen I, III, IV, V</td>
<td>M13-7GPP</td>
<td>Fluorescence Imaging</td>
<td>1.5 x 10^10 phages/mL</td>
<td>ELISA</td>
</tr>
<tr>
<td>CPKESCNLFLVKD (Collagen I platelet receptor)</td>
<td>Collagen Type I</td>
<td>CBP1495</td>
<td>GAG</td>
<td>0.86 µM</td>
<td>ELISA</td>
</tr>
<tr>
<td>RRANAALKAGELYKSIKL (Collagen I platelet receptor-original sequence)</td>
<td>Collagen Type I</td>
<td>DS-SILY</td>
<td>Modulate collagen fibrillogenesis</td>
<td>10 nM</td>
<td>Microplate fluorescence binding assay</td>
</tr>
<tr>
<td>LRELHLNNN (Inner surface of LRR 10 of Decorin)</td>
<td>Collagen Type I</td>
<td>Decorin-Based Peptides</td>
<td>Promote controlled release of molecules ≥2 kDa</td>
<td>0.17 µM</td>
<td>Surface Plasmon Resonance</td>
</tr>
<tr>
<td>LSELRLHAN</td>
<td>Collagen Type I</td>
<td>Decorin-Based Peptides</td>
<td>Promote controlled release of molecules ≥2 kDa</td>
<td>19 µM</td>
<td>Surface Plasmon Resonance</td>
</tr>
<tr>
<td>LSELRLHEN</td>
<td>Collagen Type I</td>
<td>Decorin-Based Peptides</td>
<td>Promote controlled release of molecules ≥2 kDa</td>
<td>34 µM</td>
<td>Surface Plasmon Resonance</td>
</tr>
<tr>
<td>LTELHLDNN</td>
<td>Collagen Type I</td>
<td>Decorin-Based Peptides</td>
<td>Promote controlled release of molecules ≥2 kDa</td>
<td>160 µM</td>
<td>Surface Plasmon Resonance</td>
</tr>
<tr>
<td>LSELRLHNN</td>
<td>Collagen Type I</td>
<td>Decorin-Based Peptides</td>
<td>Promote controlled release of molecules ≥2 kDa</td>
<td>220 µM</td>
<td>Surface Plasmon Resonance</td>
</tr>
<tr>
<td>GKWHTKTFPHHYCLY-(DOTA)_3</td>
<td>Collagen Type I</td>
<td>CM-101</td>
<td></td>
<td>3.6 µM</td>
<td>Bound/collagen binding model</td>
</tr>
<tr>
<td>RVMHGLHLGDDE (GPVI)</td>
<td>Collagen Type I</td>
<td>Collagelin</td>
<td>Scintigraphy, myocardial autoradiography</td>
<td>0.11 µM</td>
<td>Surface Plasmon Resonance</td>
</tr>
<tr>
<td></td>
<td>Collagen Type III</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A summary of these collagen binding peptides and their binding properties are shown in Table 1-3. Several limitations associated to these peptides have necessitated the need for the development of a collagen specific MR imaging probe with high relaxivity, biomarker targeting binding and specificity while maintaining low or no toxicity.
1.7 Previous work done by Yang lab for collagen-Targeted ProCAs (ProCA32.collagen)

ProCA32.collagen1 was designed by engineering a collagen type I targeting peptide moiety to the C-terminal of rat ProCA32 through a flexible linker to maximize collagen-targeting capacity and relaxivity without compromising the metal binding capability (denoted as rProCA32.collagen). rProCA32.collagen exhibits the highest relaxivity values for r1 (34 ± 0.12 mM⁻¹.s⁻¹) and r2 (50 ± 0.16 mM⁻¹.s⁻¹) per Gd³⁺ at 1.4 T and r1 (21.3 ± 0.5 mM⁻¹.s⁻¹) and r2 (108.5 ± 1.2 mM⁻¹.s⁻¹) at 7.0 T. rProCA32.collagen enabled detection of both early (Ishak 3 of 6) and late stage mouse liver fibrosis as well as early stage nonalcoholic steatohepatitis (Ishak 1 of 6) in different models with strong metal binding affinity and selectivity. The targeted contrast agent is also capable of detecting disease heterogeneity with high collagen type I binding affinity with dissociation constant of Kd = 1.42 ± 0.2 μM. In addition, rProCA32.collagen facilitated the sensitive detection and staging of hepatic micrometastasis as small as 0.144 mm² and two different tumor growth patterns in a uveal melanoma animal models using multiple imaging methodologies owing to its dual relaxation property [261]. rProCA32.collagen has largely reduced dose and strong resistance against transmetallation (10⁴ -10¹² -fold higher metal selectivity for Gd³⁺ over Ca²⁺ and Zn²⁺) compared to existing contrast agents [159, 187]. rProCA32.collagen has been shown to have the capability for detection of different diseases including thoracic aortic aneurysm and hepatocellular carcinoma at early stages with high confidence and response to treatment in different mice models [158]. These exciting achievements motivated us to further optimize several key aspects including using human form of rProCA32 for development of hProCA32.collagen.
Therefore, an optimization of ProCA32.collagen to hProCA32.collagen is required due to the potential of immunogenic response from the rat scaffold of parvalbumin used in its design. This dissertation seeks to address this limitation by designing a human protein-based collagen-targeted MRI contrast agent (see detailed question in 1.9 below).

1.8 Motivation and Questions to be addressed

Noninvasive early detection and stage of several human diseases remains a challenge for both the research and clinical stages. Early detection is crucial in improving prognosis, treatment efficiency and strategies or eventual reversal of diseases. The growth of the field of contrast-enhanced MRI has been tremendous. However, clinical MRI contrast agents have faced major limitations in terms of the capability of these contrast agents to detect diseases at an early stage due to their insufficient sensitivity and targeting capability. This dissertation aims to further develop our pioneered novel Gd$^{3+}$ protein-based MRI contrast agents (ProCA) to overcome the challenges associated with current MRI contrast agents in clinical setting. I will specifically address following major questions:

First, can a new MRI contrast agent be developed based on human protein scaffold to mitigate immunogenic response of the host upon administration of the contrast agent since there are 9 mutation difference between rat and human α-parvalbumin?

Secondly, will this human analogue maintain the improved sensitivity to target a specific biomarker for early-stage diagnosis of chronic diseases observed with ProCA32.collagen1 compared to the previously reported rat analogue?

Third, can the safety of an MRI contrast agent be improved by reducing its dose and toxicity and by ensuring PEGylation and collagen biomarker specificity?
Fourth, can a new MRI contrast agent capable of monitoring diseases progression and treatment be developed?

Fifth, can contrast agents be obtained as a single form with little or no adducts such as dimer or oligomers?

Sixth, can the binding and specificity of our developed collagen-targeted contrast agent to collagen type I be optimized while evaluating its susceptibility to binding to other types of collagen types?

Seventh, can chronic diseases including lung diseases be accurately staged with the use of a noninvasive MRI with our developed contrast agent and

Lastly, can the accuracy and specificity of chronic disease diagnosis be improved by taking advantage of a T1-T2 dual mode contrast agent combined with imaging methodologies?

The lungs as a major organ responsible for the respiratory system contains significantly lower proton density than majority of the organs in the body and has additional limitation for MR imaging. We hypothesize that our developed MRI contrast agent, hProCA32.collagen, has the capability to extend MRI application to achieve early diagnosis and staging of chronic lung diseases such as lung fibrosis and chronic obstructive pulmonary diseases (COPD). hProCA32.collagen can specifically bind collagen expression in heterogeneous cystic clustering honeycomb of IPF bleomycin mouse model as well was collagen expression associated with nicotine-induced airway fibrosis with heterogeneous and collagen specific pattern. In addition, hProCA32.collagen is able to detect both early and late stage (Ishak score of 1, 3 and 5) of mouse liver fibrosis in multiple mouse models as well as hepatocellular carcinoma (HCC) tumor nodules as small as 0.1mm.
1.9 Overview of this Dissertation

In this Ph.D dissertation, I report the development of a human collagen type I targeted contrast agent. This improved version addresses the clinical translational needs of our previously developed targeted contrast agent without loss of biophysical characteristics observed with the rat analog ProCA32.collagen. hProCA32.collagen has improved sensitivity and targeting capability compared to other previously developed non-targeted contrast agent in our lab and current clinical contrast agents.

Chapter 2 details experimental methods and protocols used for the work described in this dissertation.

Chapter 3 describes the expression, purification, and biophysical properties of hProCA32.cys such as its metal binding affinities, relaxivities, pharmacokinetic, transmetallation, and collagen type I affinity, serum stability and toxicity studies and application of hProCA32.cys in the detection of non-small cell lung cancer (NSCLC).

Chapter 4 details the process of selection of the protein scaffold that offers the least immunogenic response as well as the selection of the best collagen targeting sequence for collagen targeting. Furthermore, it describes the rationales for design of a novel collagen type I targeted protein contrast agent (hProCA32.collagen) and its expression, purification and biophysical properties of hProCA32.collagen such as its metal binding affinities, relaxivities, pharmacokinetic, transmetallation, collagen type I affinity, serum stability and toxicity studies.

Chapter 5 elucidates our effort in formulation and optimization of the purification of our developed contrast agent.
Chapter 6 discusses the successful application of hProCA32.collagen in bleomycin induced as well as nicotine induced lung fibrosis. hProCA32.collagen can map out the collagen pattern as well as cystic honeycombing in both pathologies.

Chapter 7 details the results of the application of hProCA32.collagen in early detection of Nonalcoholic steatohepatitis (NASH) in Nonalcoholic fatty liver disease (NAFLD) in mouse models on a NASH diet with HDAC1 gene and liver fibrosis in CGI-58 knock out mice model using both r1 and r2 properties.

Chapter 8 reports the studies of hepatocellular carcinoma (HCC) tumor detection in liver in DEN-induced mouse model using hProCA32.collagen as well as NASH+CCl4 induced HCC. hProCA32.collagen can degenerative nodules from HCC tumors.

Chapter 9 describes and summarizes the major findings of this dissertation and future direction.
2 MATERIALS AND METHODS

2.1 Cloning, plasmid generation for hProCA32.cys and hProCA32.collagen

hProCA32.cys a non-targeted contrast agent, a human version of the previously reported non-targeted rat analogue ProCA32 based on alpha parvalbumin scaffold\[108, 159, 164\] are described in this work. The human analog of ProCA32.collagen (hProCA32.collagen) was created using the non-targeted construct (hProCA32) as the main component of contrast agent. hProCA32.collagen was designed by linking type I human collagen targeting peptide to human protein-based MRI contrast agent (hProCA32) using a 3-glycine (GGG) flexible hinge. Several variants were generated, sequenced, and characterized (Chapter 4). hProCA32.collagen was sub cloned into a Pet22b vector using NdeI and Xho1 restriction enzymes. The resultant plasmid was confirmed by Sanger sequencing (Genewiz).

2.2 Transformation and Expression of hProCA32.cys and hProCA32.collagen

Both hProCA32.cys and hProCA32.collagen were expressed in Escherichia coli BL21(DE3)pLyss strain using modified purification protocol reported previously\[102, 159, 262\]. For this procedure, 50 μL of the BL21(DE3)pLyss competent cell and 0.5 μL of hProCA32.collagen plasmid (150 μg) was mixed together in a 1.5 mL Eppendorf tube and allowed to sit on ice for exactly 30 minutes. The mixture was heat shocked in a 42°C water bath for exactly 60 seconds and then placed back on ice for thermal recovery for exactly 2 minutes. 50 μL of Luria broth (LB) media was added and mixed with the DNA/competent cell mixture, this was incubated at 37°C for 30 minutes to aid bacterial growth. 50 μL of this solution was streaked onto an agar plate containing ampicillin antibiotic under aseptic conditions. The plates were then placed upside down (to avoid contamination) in the 37°C incubator overnight. The
next day, a healthy colony (single standing at location of application of transformation culture which is typically the center of the 10 cm LB agar plate) of bacteria was carefully selected for expression. While operating next to a lighted Bunsen burner. The colony was added to 250 mL of autoclaved LB and 10mL autoclaved LB media with 250 μL of ampicillin antibiotics. This grew in the shaker at 37°C overnight at 220 rpm. 35 mL of the grown medium was added per 1 L each of LB media for a total of 4L large scale expression. The large-scale media were placed in an aseptic shaker at 37°C, 220 rpm until the optical density at 600 nm reached about 0.6 AU. 0.25 mM IPTG was added for inducing the bacteria growth after which the temperature was reduced to 25°C for optimum and controlled growth in the shaker overnight. The next day, the bacteria cells (pellet) were collected via centrifugation at 7,000 rpm, 4°C for 30 minutes. About 6 g/L of the pellet was obtained and thereafter stored at -20°C or suspended in lysis buffer depending on intended time of use.

2.3 Optimized transformation and Expression Protocol for hProCA32.collagen

2.3.1 Clone selection/Screening

As will be discussed in Chapter 5, clone selection and screening were performed for hProCA32.collagen for reproducibility of results and ensuring the same clone is expressed and purify every time. To achieve this, a single clone was selected for transformation and purification. 0.5 μL hProCA32.collagen plasmid was transformed into 50 μL BL21 (DE3)pLyss competent cell, heat shocked for 60 seconds at 42°C, mixed with 1.5mL LB media (for controlled colonial growth) and plated on 10 cm LB agar plate. This was incubated at 37°C overnight. A total of 14 colonies were observed on the plate. Each of the colonies were
inoculated in 5 mL LB media (containing 5 μL ampicillin) at 37°C, 220 rpm overnight. After running SDS gel to determine the clone with the best expression outlook based on size and SDS-PAGE run charge, 1mL of the inoculum resultant culture media of the selected clones were spun down at 1300 rpm for 1 min in an Eppendorf tabletop micro centrifuge and the pellet was saved for SDS-PAGE and the other 4 mL fraction were stored at -80°C in 1 mL aliquots.

2.3.2 **Bacterial expression**

A streak of the saved culture aliquots observed by SDS-PAGE to possess the gene of interest at the desired 14.3 kDa marker were taken using an inoculation loop and transferred into two 100 mL LB media per colony containing 100 μL kanamycin antibiotic and allowed to express until OD600 reached 0.6. At this point, 0.25 mM IPTG was added and allowed to grow for either 4 hours or 18 hours each (2 flasks per colony) to decide on the optimum condition for gene expression using a ratio of the soluble fraction to the insoluble fraction as benchmark.

2.3.3 **Glycerol stock generation**

The ‘winning’ clone from section 2.3.2 was recovered by serial expression into 5 ml, 100 ml and 250 ml LB respectively. 25 mL of the grown media was inoculated into 0.5 L LB Medium with ampicillin antibiotic and allowed to grow to OD of 0.6. The culture medium was spun down, and the pellet was resuspended in 100 mL LB plus 20% final glycerol concentration then stored immediately at -80°C.

2.3.4 **Expression protocol for hProCA32.collagen from glycerol stock**

A major advantage of the use of glycerol stock is its reproducibility and long-term stability. To recover this bacterial media, a streak of the glycerol stock was incubated in 5 ml LB medium using an inoculation tube, allowed to grow overnight at 37°C, 220 rpm. The inoculum was stored at 4°C for about 8 hours to stunt growth and keep stabile until it was transferred into 100
mL ampicillin LB and subsequently grown overnight at 37°C and 220 rpm. The resultant inoculum was transferred into 250 mL LB media and grow overnight as previously described; this medium was used as the small-scale medium for transfer into large scale (1 L). 35 mL of the small-scale media was transferred per liter, optical density was monitored and upon IPTG induction at OD$_{600nm}$, the temperature was dropped to 25°C and left to grow overnight. The cell pellet was obtained by centrifugation at 7000 rpm for 40 minutes at 4°C

2.4 Purification of hProCA32.cys and hProCA32.collagen

2.4.1 Laboratory Scale (shake flask)

The collected pellet was thawed on ice until defrosted, to this, 30 mL per 6 g pellet (1 L) of 10 mM HEPES, pH 7.2 was added with 1 µL benzo nuclease and vortexed to mix. The ultrasonication for cell breakage was done ten times (10x) with 10 minutes’ break after each cycle. The homogenized mixture obtained was centrifuged at 17000 rpm for 40 minutes at 4°C. The supernatant expected to contain the cytosol concentrated solution of the protein of interest was collected and heated at about 80°C-90°C for 20 minutes. The solution was cooled then centrifuged at 8000 rpm for 20 minutes at 4°C. To the obtained supernatant, a final concentration of 3% streptomycin was added for precipitation of nuclei acids, the solution was left overnight for optimal reaction. The precipitate was separated from the solution by centrifugation at 8000 rpm for 20 minutes at 4°C. The supernatant was collected and heated in an 80° -85°C water bath for 20 minutes exactly and thereafter centrifuged at 8000 rpm for 20 minutes at 4°C. The resultant supernatant was dialyzed in 2 L 10 mM HEPES buffer, pH 8. The buffer system was changed at least three times with one-time overnight dialysis. The dialysate and retentate was separated during the process. The retentate was centrifuged at 8000 rpm for 20 minutes at 4°C. 40 mM EGTA was added to selectively chelate calcium out and the
solution reacted for about 48 hours. The solution was then filtered with a 0.45 um Whatman syringe filter. The filtrate was then injected into the AKTA FPLC column 10 ml per injection and allowed to run and the fraction collected until all the solution had been injected and collected. SDS-PAGE of each of the peak observed on the chromatogram was analyzed and the peak corresponding to the protein of interest (~12.42 kDa) was concentrated using an amicon concentrator with a 3 kDa membrane.

2.4.2 Laboratory Scale purification using ammonium sulfate precipitation

To 10 g of the Bacterial pellet, 200 mL 50 mM Tris, pH 7.4 was added to the pellet and vortex to mix, the mixture was sonicated and centrifuged at 8000 rpm, 15°C, for 30 mins. SDS page for both supernatant and pellet were obtained, the pH and concentration of ammonium sulfate of the supernatant was adjusted to 7 using 50 mM Tris, 3 M Ammonium sulfate, pH 7.0 stock solution to 1.2 M of final concentration of Ammonium sulfate in protein solution, the supernatant was centrifuged again at 8000 rpm, 15°C, for 30 mins, SDS PAGE for both pellet and supernatant were The resulting supernatant was loaded to phenyl Sepharose high performance substitution column (Phe HS column) after the column had been brought to equilibrium with 50 mM Tris, 1.2M Ammonium sulfate pH 7.0. The flow-through was collected since it was expected that hProCA32.cys will not bind to the column at this pH, the loaded column was then washed with 50 mM Tris, 0.2M Ammonium sulfate pH 7.0. Next, the column was washed with 20 mM Tris pH 8.0. The flow through pH was adjusted to pH 6.0 using 3M Ammonium sulfate and 250 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$ pH 5.7 stock solution until the final concentration of Ammonium sulfate is 1.5 M and final concentration of Na$_2$HPO$_4$/NaH$_2$PO$_4$ is 50 mM. The solution was observed for precipitate and cloudiness and
in the absence of any, it was loaded unto the Phe HS column. 50 mM Na₂HPO₄/NaH₂PO₄, 1.5 M Ammonium sulfate pH 6.0 solution was used to equilibrate the column with hProCA32.cys binding to the column at the reduced pH. 50 mM Na₂HPO₄/NaH₂PO₄, 0.5M Ammonium sulfate pH 6.0 was used to rid the column of impurities after which 20mM Tris pH 8.0 was used to elute out the protein. The column was regenerated using filtered double distilled water.

2.5 PEGylation of Protein contrast agents

Protein PEGylation is a process of covalent attachment of a Polyethylene glycol (PEG) group to a protein tertiary structure. It improves solubility due to the hydrophilicity of polyethylene glycol. The main purpose is to increase protein’s half-life in the vascular circulation while maintaining a large therapeutic index. It is an important strategy for enhancing the pharmacokinetic properties of protein therapeutics. For a successful PEGylation, the Protein structure, Total charge, Surface charge, isoelectric point (pI) and Binding sites of PEGylation reagent was considered. hProCA32.cys is a 12.24 kDa protein. There are several available pathways for PEGylation such as O-GlycoPEGylation, Transglutaminase mediated PEGylation, disulfide bridging PEGylation, thiol group PEGylation, amine PEGylation, carboxyl PEGylation, N-terminal PEGylation and non-covalent PEGylation [263]. Advantages of PEGylation includes the ability to optimize the protein solubility, blood circulation time, reduce immunogenicity and avoid possible enzymatic degradation. hProCA32.cys possess 16 lysine residues with no cysteine residue in its primary sequence. Carrying out lysine PEGylation will therefore result in nonspecific and uncontrolled PEGylation.
2.5.1 *Lysine PEGylation*

PEGylation can increase the solubility, decrease immunogenicity of the protein drugs. Polyethylene glycol (PEG) reagent was mixed with ProCA in a 5:1 ratio. The mixture was incubated at room temperature for 2 hours with constant shaking. The reaction was stopped by addition of 10% v/v 100mM Tris/HCl buffer pH 8. The excess PEG reagent was removed by dialysis using the Amicon concentrator with 3kDa membrane. Both commassie brilliant blue and iodine staining of SDS-PAGE gel were used to confirm PEGylation success. Commassie brilliant blue were used to stain the protein bands and iodine were used to stain PEG moiety.

2.5.2 *Optimized lysine PEGylation Procedure*

To minimize cost of production and maximize PEGylation efficiency, a new protocol was introduced with the use of the G-25 fine Sephadex column. HCl solution with a concentration of 2 mM was made from a stock concentration of 12 M. The 2 mM HCl was used to dissolve PEG powder and make the concentration to be 20% in this case, we dissolved 200 µg of PEG in 800 µl of 2 mM HCl solution. Protein preparation was done by using DTT (5 mM, the stock is 1 M) to reduce the protein at room temperature. This was allowed to incubate for about 5 min, either ultra-concentrator or spin filter was used to concentrate and remove the DTT. The PEG in HCl was added to the reduced protein immediately. The ration of protein and PEG could be 1:1, 1:1.5 and 1:2. The reaction happens within 1 min. The mixture was shaken at room temperature for 1 hr to increase the reaction yield a little. A separation of free PEG reagent from PEGylated protein as well as separation of monoPEGylated protein and unPEGylated protein was done by FPLC purification by G-25 fine sephadex column and mono Q column, respectively.
2.5.3  

**Cysteine PEGylation**

hProCA32.cys possesses a cysteine residue at the c-terminal which is utilized for site specific cysteine PEGylation since hProCA32.cys has 16 lysine residues with no cysteine residue in its primary sequence. Carrying out lysine PEGylation will therefore result in nonspecific and uncontrolled PEGylation. To make up for this, site mutagenesis of the modified parvalbumin scaffold was done to introduce a cysteine residue at the C-Terminal of the protein. To achieve this, 2 ml 0.9mM hProCA32.cys was mixed with 200µL tris (2-carboxyethyl) phosphine (TCEP), a reducing agent to break any disulfide bonds present in the reaction mixture. The solution was mixed with a tube rotator and incubated at room temperature for 30 minutes after which the unreacted/residual TCEP was dialyzed out through a 3 kDa membrane. Next 0.011 g of PEG reagent (2 kDa) solute was added in a 1:5 (Protein: PEG) concentration ratio and allowed to react for two hours at room temperature. The amount of PEG added was calculated as shown below.

\[
900 \times 10^{-6} M \times 3 mL \times 2000 g \times 5 = 0.027 g \text{ PEG reagent}
\]

The unreacted PEG reagent was removed using a 3kDa membrane amicon® concentrator.

2.5.4  

**N-terminal PEGylation of hProCA32.collegen**

After several attempt, it was discovered that the presence of 2 cysteine in the collagen targeting moiety plays an inhibitory role in cysteine PEGylation of hProCA32.collegen. While still aiming at achieving site specific PEGylation, N-terminal PEGylation was done. The mPEG-propionaldehyde derivative (Mw 5kDa) and protein (molar ratio= 1:1 and 2:1) were dissolved in sodium acetate buffer (50 mmol/L, pH 5.5) in the presence of sodium borohydride (2.5 mmol/L) as a reducing agent. The PEGylation reaction was conducted for 24 h at room temperature and the excess reducing agent was removed by repeated ultrafiltration using a
centrifugal filter with MWCO 10kDa. The reaction was terminated with 5mM Tris pH 8 and analyzed.

2.6 Protein Concentration Determination

Protein concentration of stock solution were determined by measuring the absorbance of the solution at 280nm using the UV/VIS Spectrophotometer at an extinction coefficient of 7200 M$^{-1}$ cm$^{-1}$ for hProCA32.Cys as reported and 21890 M$^{-1}$ cm$^{-1}$ obtained from extinction coefficient calculation using tryptophan, phenylalanine, tyrosine residues. Lambert Beer’s equation was used to solve for the concentration.

$$A = \log \frac{I_o}{I} = \varepsilon cl$$

Equation 2.1

Where $A$ is the absorbance (value obtained from the spectrophotometer)

$I_o$ is the incident Intensity of light

$I$ is the resulting light intensity

$L$ is the path length and

$C$ is the concentration of protein

2.7 Relaxivity Studies

The relaxation times (T1 and T2) were measured by Bruker relaxometer (1.47 T). The contrast agents with different concentration were prepared in 50 mM HEPES, pH 7.4. The relaxivity was calculated using equation 1.5. The relaxivity of ProCA3 variants were measured in the buffer containing 50 mM HEPES, 150 mM NaCl, and pH 7.2 at 37 °C. The protein possesses
two calcium binding sites that were modified to gadolinium binding sites, the samples were prepared in a 250µL HEPES buffer a 2:1 gadolinium to protein ratio respectively from 20 µM to 160 µM concentration of Gadolinium. We then calculate the per Gd$^{3+}$ relaxivity of each contrast agent concentration using the equation above. An alternate method used was to keep either the Gd$^{3+}$ or the protein concentration constant with varying concentration of the other using the buffer system.

2.8 Gd$^{3+}$, Tb$^{3+}$, Ca$^{2+}$, and Zn$^{2+}$ -Binding Affinity Determination

2.8.1 Calcium Titration

Fluorescence spectra were recorded using a QM1 fluorescence spectrophotometer (PTI) with a xenon short arc lamp at ambient temperature. Tryptophan fluorescence was monitored using excitation at 280 nm and emission between 310 and 315 nm with 0.35-0.5mm excitation and 0.7-1mm emission slit width, respectively. The fluorescence response of the ProCAs (hProCA32.cys and hProCA32.collagen) to calcium was analyzed using a Photon Technology International spectrophotometer. The spectra were collected using the Felix32 fluorescence analysis software, the slit width was set at 0.4mm and 0.8mm for excitation at 280nm and emission 300-400nm, respectively. A 10µM concentration of protein in the buffer system was used Different concentration of calcium chloride were added into the protein-containing buffer system and the free calcium was obtained using a foreknown Ca-EGTA $K_d$ 1.51 x $10^{-7}$M in Tsein derived equation below:

$$[Ca^{2+}]_{free} = K_{dca-EGTA} \times \frac{[Ca-EGTA]}{[EGTA]_{free}} \quad \text{Equation 2.2}$$

$$f = \frac{[Ca^{2+}]_{free}^n}{K_{dCa-ProCA}^n + [Ca^{2+}]_{free}^n} \quad \text{Equation 2.3}$$
Where $[\text{Ca}^{2+}]_{\text{free}}$ is the free $\text{Ca}^{2+}$ concentration at each titration point,

Kd $\text{Ca-EGTA}$ is the dissociation constant of EGTA for $\text{Ca}^{2+}$,

$[\text{Ca-EGTA}]$ denotes the concentration of total Ca-EGTA complex at each titration point,

$f$ represents fractional change of fluorescent signal, and

$[\text{EGTA}]_{\text{free}}$ is the concentration of EGTA at each titration point.

The modifications made to parvalbumin includes a F103W mutation where the tryptophan residue is extensively utilized for absorbance and fluorescence. The mutation is also close to the binding site of the protein hence its use as an indicator to monitor calcium binding. The Hill equation commonly used to study the kinetics of reactions that exhibit a sigmoidal behavior was used to fit the curve obtained.

$$V = \frac{V_{\text{max}}[S]^n}{(K_{0.5})^n + [S]^n} \quad \text{Equation 2.4}$$

$V$ is the reaction velocity (rate of reaction progression per unit time)

$V_{\text{max}}$ is the maximum velocity of the reaction. It has the same units as the reaction velocity ($V$).

It is the highest reaction rate that can be achieved at saturating substrate concentrations.

$[S]$ is the substrate concentration

$K_{0.5}$ is the half-maximal concentration constant. It may also be referred to as $K_{\text{half}}$. It is the substrate concentration that gives rise to a reaction velocity that is 50% of $V_{\text{max}}$

$n$ is the Hill coefficient
If $n > 1$, the reaction/process is thought to exhibit positive cooperativity with respect to substrate binding to the protein. A value of larger than 1 for the Hill coefficient ($n > 1$) also suggests that there are more than one substrate binding sites in the protein under study.

### 2.8.2 **Terbium Titration**

$\text{Tb}^{3+}$ binding affinity of hProCA32.cys was obtained by fluorescence spectroscopic method. $\text{Tb}^{3+}$ fluorescence at 545nm was monitored using tryptophan excitation at 280 nm and terbium emission between 500 and 600 nm with 0.35mm excitation and 0.65mm emission slit width respectively. DTPA solution was used as chelator to chelate out residue terbium in solution. The fluorescence observations were done both for the DTA-buffer system with no protein and for Protein containing solution. A final concentration of 30 $\mu$M hProCA32.cys and 5 mM DTPA solution was prepares in a 150 mM NaCl, 10 mM HEPES buffer system, pH 7.2. To this different concentration of $\text{Tb}^{3+}$ was added until a final concentration of 5 mM was reaches at which point DTPA is expected to have been fully competed out. The fluorescence response of the hProCA32.cys to $\text{Tb}^{3+}$ was analyzed using a Photon Technology International spectrophotometer. The spectra were collected using the Felix32 fluorescence analysis software, the apparent $K_d$ ($K_{app}$) was obtained by one-to-one binding Equation and the dissociation constants of hProCA32.cys were fitted by the equation 2.5 below;

$$[\text{ Tb}^{3+}]_{\text{free}} = K_{dTb,DTPA} \times \frac{[\text{Tb}^{3+} - \text{DTPA}]}{[\text{DTPA}]_{\text{free}}} \quad \text{Equation 2.5}$$

$$f = \frac{[\text{Tb}^{3+}]^n_{\text{free}}}{K_{d,ProCA} + [\text{Tb}^{3+}]^n_{\text{free}}} \quad \text{Equation 2.6}$$

Where $[\text{Tb}^{3+}]_{\text{free}}$ is the free $\text{Tb}^{3+}$ concentration calculated from buffer system, $K_{dTb,DTPA}$ is the dissociation constant between $\text{Tb}^{3+}$ and DTPA,
[Tb-DTPA] is the concentration of Tb-DTPA complex,

[DTPA]_{free} is the free DTPA concentration in the buffer,

f is the fractional change,

$K_{dTb,ProCA}$ is the dissociation constant between Tb$^{3+}$ and protein

### 2.8.3 Zinc Titration

The fluorescence spectrum for the calculation of zinc binding to ProCA was obtained as described previously. 1 µM Fluozin-1 mixed with 2 µM ZnCl$_2$ prepared in 50 mM HEPES, 100 mM NaCl buffer, pH 7.2 and was excited at 495 nm with emission spectra between 500 and 600 nm subsequently collected. This was recorded as the blank sample, after which pre-calculated volume of protein ranging from concentration 0-100µM was added or until saturation was established as observed by absence of further increase in fluorescence intensity. The apparent dissociation constant was obtained by fitting this data into a 1:1 competition equation 2.7 shown below. The fluorescence signal changes were recorded in an equilibrium state after each titration point.

$$f = \frac{([P]_T+[M]_T+K_{dapp})-\sqrt{([P]_T+[M]_T+K_{dapp})^2-4[P]_T[M]_T}}{2[P]_T}$$

Equation 2.7

$$K_{dZn,ProCA} = \frac{K_{dZn,fluozin-1}}{K_{dapp}-K_{dTb,fluozin-1}}$$

Equation 2.8

Where $K_{dZn,ProCA}$ is the dissociation constant between Zn$^{2+}$ and ProCA protein,

$K_{dapp}$ is the apparent $K_d$ calculated by Equation 2.7 above,

$K_{dTb,fluozin-1}$ is the dissociation constant between Tb$^{3+}$ and Fluozin-1.

$K_{dTb,fluozin-1}$ is measured by already established literatures
2.8.4 **Gadolinium Titration**

The dissociation constant obtained for Tb$^{3+}$ above was used in the calculation of binding constant for Gd$^{3+}$. FRET competition method was applied to measure the Gd$^{3+}$ binding affinity to ProCA. 10µM of ProCA and 20 µM Tb$^{3+}$ were incubated with a series of 0 to 1000 µM of GdCl$_3$ at room temperature overnight. The Tb$^{3+}$ FRET changes were measured by the emission of Tb$^{3+}$ at 545 nm. The apparent $K_d$ were fitted by 1:1 binding Equation 2.7 and $K_d$ of Gd$^{3+}$ was calculated with equation 2.9

$$K_{dGd,\text{ProCA}} = K_{d\text{app}} \times \frac{K_{dTb,\text{ProCA}}}{K_{dTb,\text{ProCA}} + [Tb^{3+}]_T} \quad \text{Equation 2.9}$$

Where $K_{dGd,\text{ProCA}}$ is the dissociation constant between Gd$^{3+}$ and ProCA,

$K_{d\text{app}}$ is the apparent Kd for Gd$^{3+}$,

$K_{dTb,\text{ProCA}}$ is the dissociation constant between Tb$^{3+}$ and ProCA previously obtained section 2.7.3 and

$[Tb^{3+}]_T$ is the total concentration of Tb$^{3+}$ in the experiments.

2.8.5 **Determination of Free Gadolinium by xylenol orange method**

50mM acetate buffer solution was prepared by adding 287µL Acetic acid (density = 1.05g/ml) to 80mL Deionized water. The pH was thereafter adjusted with 3M NaOH to 5.8, the volume was made up to 100mL. After washing and drying all cuvette, 1µM xylenol orange solution was prepared by dissolving 3 mg xylenol orange in 250mL acetate buffer, pH 5.8. The absorbance spectrum of the unused solution was obtained for future comparison at the time of thawing. The peaks at 433nm and 573nm was thereafter used in the evaluation of free Gd$^{3+}$ per the following equation.
\[ Gd^{free} \propto \frac{Abs^{573}}{Abs^{433}} \] 
Equation 2.10

\[ Gd^{free} = A + B \frac{Abs^{573}}{Abs^{433}} \] 
Equation 2.11

A standard calibration curve of the ratios of the absorbance was plotted against increasing known concentration of free Gadolinium. The concentration used were 0 µM, 2 µM, 5 µM, 7 µM, 10 µM, 15 µM, 20 µM, 25 µM, 30 µM, 40 µM, 45 µM, 50 µM. For the first point, 800µL of the xylenol orange solution was measured as blank. Then gradual addition of the concentrations of free gadolinium (high purity standard) was done with the absorbance value recorded. The amount of free Gadolinium in the protein was estimated, the absorbance value recorded was used in the pre-obtained calibration curve to calculate the free Gd\(^{3+}\) concentration.

2.8.6 **Circular dichroism spectroscopy**

The CD spectra of samples were recorded in a Jasco-810 spectropolarimeter at room temperature. The far-UV CD spectrum of different ratios of hProCA32.cys:Gd\(^{3+}\) were measured in a 1-mm path length cell in 10 mM HEPES (pH 7.4) and the near-UV CD spectrum. All spectra were obtained as the average of ten scans with a scan rate of 100 nm/min. The ellipticity was measured from 190 to 260 nm (far-UV) and 250 to 340 nm (near-UV). This was converted to mean residue molar ellipticity after subtracting the spectrum of baseline buffer sample as blank.

2.8.7 **Residual Calcium Measurement in Working Buffers**

The UV-visible cuvettes were thoroughly rinsed with deionized water several times. The cuvette was then filled with 100 mM EDTA and allowed to stand for 3-5 minutes to remove residual calcium contamination. The cuvette was then rinsed with deionized water (or
investigated buffer) several times to remove the EDTA. Next, the buffer was added to the cuvette. The dye (in principle any dye with strong calcium-binding affinity, $K_a$ less than 1 µM, can be used) was added. 10 µM of calcium green will result in the absorbance out of the linear range. In other case, the dye concentration lower than 3 µM may not enough to measure some buffers since the calcium background up to 2-3 µM is normal. Suggested dye concentrations are 3-8 µM for using calcium green to measure the background level of 0.5-2.5 µM calcium). In this case, a final concentration of 7µM Rhod-5N was used in a 1mL solution of 100mM Tris, 100mM KCl, PBS or HEPES buffers were used. The absorbance range covering the positions with the greatest changes was scanned and recorded as Point $A_{\text{buffer-only}}$. 5 µL 100 mM EGTA (or EDTA) was added to the sample and mixed by inversion. The same absorbance range was scanned and recorded as. Point 2: $A_{\text{EDTA}}$. 5 µL 1 M calcium was added to the sample and scanned again and recorded as Point 3: $A_{\text{Ca}}$. The absorbance values were taken (the dilution effects of the last two samples can be normalized by simply multiplying by 1.005 and 1.01, respectively. However, if it is ignored, it will not cause any obvious difference).

The background calcium concentration was calculated using the equation.

$$C = C_{\text{dye}} \times (A_{\text{buffer-only}} - A_{\text{EDTA}})/(A_{\text{Ca}} - A_{\text{EDTA}})$$

Where $C_{\text{dye}}$ is the dye concentration.

### 2.8.8 Transmetallation Studies

1.2mM phosphate and 30mM phosphate buffer were prepared for ProCA and clinical contrast agents respectively. The amount of each reagent required to make 2.5mM concentration of clinical contrast agent and 110µM (1:2) ProCA was estimated and for other binding rations such as 1:1, 1:2, 1:1.8, the calculations were adjusted accordingly. The clinical contrast agents
have already been formulated with gadolinium in a 1:1 ratio while gadolinium was added to ProCAs after the appropriate calculation done, 100µM Zinc was added to all the solutions with a final volume of 250 µL. The solutions were incubated at 37°C and the longitudinal relation rates (R1) at different time points were measured for at least 3 days while keeping the temperature at 37°C. A general description of the reaction is shown diagrammatically below.

\[
\text{Gd} - L^{(n)} \leftrightarrow L^{(n+3)} + \text{Gd}^{3+} \\
L^{(n+3)} + \text{Zn}^{2+} \leftrightarrow \text{Zn} - L^{(n+1)} \\
\text{Gd}^{3+} + \text{PO}_4^{3+} \rightarrow \text{GdPO}_4 \text{ (precipitate)}
\]

Figure 2-1 Transmetallation experiment of ProCAs using Zn\(^{2+}\) competition in phosphate buffer solution
2.8.9 **Gd³⁺ Bio distribution studies in different organs**

7- to 10-wk-old female CD1 mice were injected with 0.28 mmol/kg of ProCA (containing 0.56 mmol/kg of Gd³⁺) \( (n = 3/\text{contrast agent}) \). Blood was collected from the heart (cardiac puncture) 5 days after injection. Tissue samples from the mice were collected after euthanasia. The collected tissues were digested with 70% (v/vol) HNO₃ at 130 °C overnight while the serum samples were sent for mice clinical pathological studies with saline injection for the control mice. Furthermore, the HNO₃ solution containing tissue digestions were and diluted with 2% (v/vol) HNO₃ to 10 mL. The samples were filtered with the 0.45µm filter and subsequently analyzed by ICP-OES at gadolinium wavelength of 342.246 nm. YCl₃ at 2 ppm was used as internal standard.

2.8.10 **Serum Stability Studies**

150 µl of 500 µM of protein sample was incubated with 150 µl human serum. A protein only sample and serum only sample were included as controls. The sample was left sitting at 37 °C with SDS-PAGE sample collected daily. 20 µl of sample was mixed with 20µl SDS sample buffer during time points 0, 2 hr, 4 hr, 6 hr, 2 days, 3 days, 4 days, 6 days, 8 days, 10 days, 12 days, and 14 days. Upon SDS-PAGE sample collection, the samples were boiled and stored at -20°C.

2.8.11 **Pharmacokinetics and Pharmacodynamic studies of hProCA32.collagen**

Single-dose administration of ProCAs via tail vein injection was carried out in mice divided into groups of four consisting four mice each to reduce the frequency of blood collection in individual mice at a specific point in time. The back of the hind leg of the mice was shaved while being restrained with an approved animal restrainer. The hind leg was immobilized, and
a slight pressure was applied gently above the joint of the knee to aid blood flow and limit animal movement. Ointment was applied on the shaved site to reveal the sephaneous vein; the vein was then punctured carefully with a 20G needle and blood collected with an aspirating pipette. Pressure was applied to the punctured site to stop the bleeding and data (blood samples) was collected at various time points in all groups of mice over a period of 14 days after which the animals were sacrificed and their tissues collected for ICP-OES.

2.9 MRI Scan

2.9.1 3-D ultrashort echo time (3D-UTE)

Lung Fibrosis was imaged with T1-weighted 3D ultra-short TE (3D-UTE) MRI. Ultra-short TE was 0.16 ms. TR = 12 ms. Data were acquired with radial sampling. Isotropic FOV was 48 mm, isotropic matrix size was 192 and isotropic spatial resolution was 250 mm. The acquisition time was 19 minutes.

2.9.2 Rapid Acquisition with Refocused Echo (RARE)

Two-dimensional spin echo multi-slice (T1-weighted images) were collected with the following acquisition parameters: TR = 500 ms, TE = 15 ms, Segment/ETL = 32/8, FOV = 40 × 40 mm, matrix = 256 × 256, total image slices = 8, slice thickness = 1.0 mm without gap the total acquisition time was 2 min.

2.9.3 Multislice spin Multi echo

Fast spin echo multi slice (T2-weighted images) were obtained with the following parameters: TR = 2000 ms, FOV = 40 × 40 mm, effective TE = 40 ms, matrix = 256 × 256, total image slices = 8, slice thickness = 1.0 mm. Both T1-W and T2-W images were taken without gap.
3 DEVELOPMENT OF HUMAN PROTEIN-BASED MRI CONTRAST AGENT FOR LUNG CANCER IMAGING

3.1 Introduction

3.1.1 Lung cancer detection by noninvasive imaging modality

Lung cancer is the second most common cancer and the leading cause of cancer death. It is estimated that over 230,000 cases are expected to be recorded in 2021 with over 55% of these cases leading to fatality in the US. It is more prevalent in individuals 65 or older although people <45 have sometimes been diagnosed with lung cancer [264, 265]. Although only 10% of smokers develop lung cancer, smoking is the most common risk factor for lung cancer, with approximately 90% to 95% of new lung cancers arising from active smoking[266, 267] and passive smoking accounting for 25% of lung cancer cases [268] other risk factors include occupational exposures, chronic lung diseases, environmental exposures and lifestyle choices[269]. The onset of lung cancer is premalignant changes in the epithelium, developing from squamous metaplasia through carcinoma in situ and finally invasive cancer [270]. Lung cancer is divided into two main types: small cell and non-small cell lung cancers. Small cell lung cancer (SCLC) has a higher propensity for an early spread with a corresponding fast tumor doubling time, it accounts approximately for 15% of lung cancer with 70% of patient having metastasized cancer at the time of diagnosis[271]. Non-small cell lung cancers (NSCLC) accounts for about 85% of lung cancers, it is made up of adenocarcinoma (the most common lung cancer type observed in previous smokers as well as non-smokers) which is predominantly observed in the pleural and sub pleural parts of the lung and has a slower spread
rate compared to SCLC. A second type of NSCLC is squamous cell carcinoma originating in the lung squamous cell and is associated with smoking history, this type of cell is pathologically prevalent in the bronchus. Third class is the large cell carcinoma with no restriction to any area of the lungs. Like SCLC, it tends to grow and spread quickly, making therapeutic options difficult. Other less common subtypes are adenosquamous carcinoma and sarcomatoid carcinoma. Altogether, NSCLC have similar treatment and prognostic options[272]. The survival of the lung disease is greatly hampered by late-stage diagnosis as most of the diagnostic finding occurs at advance stage or after lung cancer metastasis has occurred despite that there’s been considerable improvement in the outcomes of lung cancer due to a well establish etiology [273-276]. This chapter describes our effort to develop of protein MRI contrast agent for the possibility of using MRI in lung imaging of NSCLC.

3.1.2 **Current diagnostic methods of lung cancer and their limitation**

A key factor contributing to the high mortality rate for lung cancers is the late diagnosis of lung cancer. Early, accurate and effective cancer diagnosis is critical for improved prognosis and guiding treatment options. Lung cancer is usually suspected in symptomatic patients such as chronic smokers with respiratory symptoms such as cough, dyspnea, hemoptysis, and chest pain. The diagnosis and staging are usually done simultaneously. This usually commences with a chest X-Ray, then computer tomographic (CT) scan which provides a first outlook and classification of the suspected disease into either SCLC or NSCLC. Large lymphadenopathy and direct mediastinal invasion are characteristic of SCLC. Another feature present in about 78% of SCLC patients is a mass adjacent to or in the hilum of the lung.[277, 278]. An absence of these SCLC distinguishing features results in the diagnosis of NSCLC. Depending on the location and size of the lesion, other approaches include sputum cytology bronchoscopy and
trasthoracic needle aspiration (TTNA). The choice of the most applicable diagnostic method for the suspected type and location is made during a multidisciplinary discussion. Early detection of SCLC is difficult due to high propensity for an early spread with a corresponding fast tumor doubling time as well as lack of specific symptoms. This has hampered the current screening approaches ineffective in diagnosing patients at early disease stages. Chest X-ray of lung cancer does not have the required sensitivity for cancer detection especially with small lung cancer and currently, the gold standard for lung cancer imaging is the use of computed tomography (CT). Certain limitations associated with CT scans include the use of ionizing radiation, and the need for biopsy in over 50% of patient with at least one noncalcified nodule.[279] There is therefore an unmet medical need for a non-invasive and precise imaging modality for early detection of lung cancer for improving patient prognosis, and image guided therapeutic options. Due to the extremely low proton density and susceptibility difference between air and tissue, MRI of the lung using conventional T1 and T2 weighted pulse sequence has several limitations and remains a major challenge for MR imaging of the lung. MRI has been extensively examined as a potential imaging modality for early detection and monitoring of lung disease progression [280-282] due to its advantage of not using harmful ionized radiation, superior ability to characterize tissue properties and dynamic image acquisition capacities. Lung imaging by MRI especially in small animals presents some unique challenges [283, 284] [285]. First is the relatively low tissue density (10x lower) due to high oxygen (from air) and low water proton within the lungs, severely limiting signal-to-noise. In addition, variations in magnetic “susceptibility,” associated with the many air–tissue interfaces of the alveoli and bronchioles, create local magnetic field in homogeneities (field gradients) that can
lead to shortening of some relaxation times. Further, respiratory and cardiac motion can lead to significant image blurring [286].

3.1.3 **Criteria and status of development of protein-based MRI contrast agent for lung MR Imaging**

The current clinically available MRI contrast agents lacks the required sensitivity to meet patients’ needs in diagnosis and treatment monitoring of several diseases including lung cancer. In this dissertation chapter, we discuss our strategy for lung cancer imaging by developing a contrast agent using protein as the chelating ligand for Gd$^{3+}$ with high sensitivity, metal binding capability and in vivo properties. Several important criteria were considered; first, the desired protein-based MRI contrast agent should be thermostable and resistant to degradation and denaturation factors such as protease usually present in the body. In addition to protein stability, the protein-Gd$^{3+}$ complex should be stable throughout its residence time in the body to avoid dechelation due to the toxicity of Gd$^{3+}$ with LD$_{50}$ reported as 0.2mmol/kg [55] and the propensity of Gd$^{3+}$ to interfere with the physiological activity of specific, particularly proteins susceptible to Gd$^{3+}$ binding, either by forming a complex with functional side chain groups or by displacing essential metal ions such as iron, zinc, copper, manganese in metalloproteins. It may also interfere with the routine signal transduction pathways modulated by other metals such as Ca$^{2+}$. Another important criterion is for the contrast agent to possess high relaxivity for high sensitivity to enable dose reduction and dose efficiency. If relaxivity is higher, then the required injection dosage of contrast agent will be decreased thereby resulting in reduced toxicity caused by the release of free Gd$^{3+}$. A third criterion considered for our protein contrast agent is the possession of the right size for optimized bio
distribution in the organs and tissue while possessing the capability for glomerulus excretion. Furthermore, developed protein MRI contrast agent should have the capability to specifically target molecular biomarkers. Lastly, the developed protein contrast agent should possess low immunogenic susceptibility for clinical translation and applications.

Over the past years, our lab has designed several protein-based MRI contrast agents based on different proteins such as cell adhesion 2, calmodulin and α-parvalbumin all of which possesses high relaxivity, proper blood retention and bio distribution properties as well as low toxicity [102, 108, 109, 157, 159, 160, 164, 261, 287, 288]. The rat version with lysine PEGylation (ProCA32.P40) was compared with Eovist in the detection of metastatic liver metastasis [102]. ProCA32.P40 has been shown to have a significantly higher $r_1$ and $r_2$ per-Gd relaxivities (about 10-fold greater) than clinical contrast agent, Eovist at 1.4 T, allowing acquisition of both positively and negatively enhanced liver images at physiologically achievable concentrations in the liver using $T_1$- and $T_2$-weighted pulse sequences. In addition, by combining the contrasts in $T_1$- and $T_2$-weighted images obtained by using the ratio, ProA32-P40 enabled a 100-fold increase in its detection limit for identifying lesions compared with eovist. Due to its high sensitivity and reduced dosage, ProCA32.P40 was reported as a potential candidate for clinical use [102].

3.1.4 **Objective and overview of this chapter**

A major limitation of ProCA32.P40 however, is the potential of ProCA32.P40 to elicit immunogenic response upon injection in humans due to the use of rat-based protein scaffold. Another limitation is the non-specificity in the PEGylation of lysine residues, lysine accounts for at least 6% of human proteins [289] and PEGylation of such protein will be considerably nonspecific.
This chapter focuses on the development of an improved protein-based contrast agent using human parvalbumin as scaffold. To this effect, a human-based protein scaffold was selected to address the immunogenic concern while site-specific PEGylation was done by genetically encoding a single cysteine residue at the C-terminus of the human version of (hProCA32). This will subsequently be referred to as hProCA32.cys. Site-specific PEGylation of cysteine was decided upon because cysteine residue is present as less than 1% of the total amino acid content of many protein biomolecules and may be present as a cysteine by forming a disulfide bonds rendering them inert to many thiol-specific reagents [290, 291]. We will then report the development of animal model for lung cancer and preliminary application of hProCA32.cys for lung cancer detection in a mouse model developed by Dr. Zhiren Liu’s lab.

3.2 Expression, Purification, and application of human protein contrast agents for MR Imaging

3.3 Expression and Purification of hProCA32.cys

hProCA32.cys was developed as an optimized analogue of ProCA32 previously reported based on rat α-parvalbumin scaffold. It is the backbone upon which a targeted contrast agent hProCA32.collagen was designed for use as MRI contrast agent with enhanced biomarker-targeting specificity and relaxivity without compromising the need for reduced toxicity. This protein was expressed and purified. The transformation process is described in section 2.1. The bacterial growth curves and SDS-PAGE gels for expression are shown in Figure 3-1 Lane 3, 4 and 5 of the SDS gel shows that hProCA32.cys was successfully expressed in bacteria cells as indicated by increased 14.3kDa band on the SDS-PAGE compared to the same position before
IPTG induction. The grown medium was centrifuged, and the cells were harvested in the pellet as observed in lane 5 of the SDS gel.

![Image of IPTG induction process]

Figure 3-1 Expression of hProCA32.cys from Pet22b vector. A. General Expression scheme of hProCA32.cys, B., SDS-PAGE of expression culture before and after IPTG induction shows no expression of the hProCA32.cys gene before IPTG induction (lane 2) and appearance of the expressed gene at band around the expected 12kDa band after 1 h (lane 3) and at harvest and upon centrifugation. C., Bacterial growth curve of hProCA32.cys from start to harvest. The plot shows a lag phase and after IPTG induction, a log phase was observed.

3.4 Purification of hProCA32.cys

hProCA32.cys was purified as described in section 2.4.1 As shown in Figure 3-2, the supernatant obtained after sonication (cell lysis) stage had 95% of hProCA32.cys lYZed into
the supernatant compared to about 5% in the pellet. The supernatant obtained from dialysis was incubated with EGTA and subsequently injected in batches (9ml per batch) for separation by FPLC. A representation of the chromatogram obtained is shown in Figure 3-3.

![Figure 3-2: SDS-PAGE characterization of step-by-step purification of hProCA32.cys. The purification of hProCA32.cys following cell lysis shows the appearance of less impurity band with each step until addition of EGTA (lane 13) with majority of the protein in the supernatant.](image)

Fractions eluted for each peak were analyzed by SDS PAGE. Lane 2 to 4 in the top gel in Figure 3-3 shows a little to no protein present which could correspond to unbound proteins as nucleic acids have been suspected to be abundant in solutions at physiological pH. Peak 2 containing the highest amount of hProCA32.cys as observed on SDS-PAGE was concentrated to a concentration of 730.2µM. This results in the production of 28mg of hProCA32.cys protein.
Figure 3-3 FPLC chromatogram and characterization of hProCA32-Cys. SDS-PAGE for fraction confirmation (Lane 5 and 6) contained the hProCA32.cys protein at the expected molecular weight and upon concentration, the SDS-PAGE from the concentrate showed purified hProCA32.cys (arrow)

3.5 Cysteine PEGylation of hProCA32.cys

The linkage of polyethylene glycol (PEG) has been extensively employed as a post-translational modification methodology to prolong the circulating half-lives of protein as well as improving the biomedical efficacy and physicochemical properties of protein biologics [292]. The cysteine PEGylation reaction was carried out as described in 2.5.3. hProCA32.cys was mixed in a 10% v/v ratio with tris(2-carboxyethyl) phosphine (TCEP), a reducing agent to break any disulfide bonds possible present in the reaction mixture. The solution was mixed with shaking at room temperature for 30 minutes after which the unreacted/residual TCEP was dialyzed out through a 3kDa membrane. Next PEG reagent (2kDa) solute was added in a 1:5
concentration ratio and allowed to react for two hours at room temperature. The unreacted PEG reagent was removed using a 3kDa membrane.

Figure 3-4 SDS-PAGE characterization of hProCA32.cys under reduced conditions. Iodine and Coomassie brilliant blue staining showed the successful PEGylation of hProCA32.cys. On the iodine staining on the right, the 14.3 kDa band showed no protein while there’s a higher band showing at about 24kDa due to slower migration rate of the PEGylated protein. Iodine staining are more sensitive to PEGylated product. Staining of unPEGylated protein products depends on the concentration of sample loaded [293, 294]. On the left, the normalized Coomassie staining showed both the unPEGylated and PEGylated portion.

To confirm the PEGylation yield, both iodine and Coomassie brilliant blue staining were carried out since protein staining alone (i.e., Coomassie, silver, etc.) cannot accurately distinguish between a PEGylated protein product from a high-molecular-weight protein-protein aggregate [294]. From Figure 3-4, the intensity of the SDS-PAGE bands were calculated based on ImageJ analysis of both the Coomassie and iodine stained gels [295] where
it was estimated that PEGylation was achieved with about 70% efficiency while about 30% unPEGylated protein still present (Coomassie blue stain). The molecular weight of the PEGylated band observed at about 24kDa is inconsistent with the expected 16.3kDa band since the PEG reagent is 2kDa and only one cysteine is present in the protein. The discrepancy however could be because of slow migration of the PEGylated fraction as a result of charge. Although the bands have similar molecular weight expected for the dimer formation, the SDS-PAGE was carried out under reduced conditions (5% βME) for both unPEGylated and PEGylated fractions and is unlikely to have dimer formation upon PEGylation. As described by M. Kurfurst, the mobility of the PEG molecule carried out at 8 – 25 % polyacrylamide gel (15% polyacrylamide was used in this study) is linear to the log of its molecular weight and this linearity is parallel to those of standard protein markers [294]. This implies that as the size of PEG and ratio of PEG:Protein increases, the mobility of the complex reduces. As shown in Figure 3-4, attaching 5 PEG molecules to hProCA32.cys significantly reduced the mobility rate relative to standard protein markers which showed up at lower band intensity due to the higher specificity and sensitivity of iodine staining towards PEG products over protein unbound to PEG.

3.6 FPLC Characterization of PEGylated hProCA32.cys

To achieve a good separation between the PEGylated and unPEGylated hProCA32.cys, the sample went through a second round of FPLC purification using the ion exchange technique using both anion and cation exchange columns. For anion exchange Q-column, 10mM HEPES and 1M NaCl/10mM HEPES pH 8 were used as binding and elution buffer respectively while
cation exchange S-column used 10mM HEPES and 1M NaCl/10mM HEPES pH 6 as binding and elution buffer respectively. The Q-column could not achieve well resolved separation as both PEGylated and unPEGylated were under one broad peak (Figure 3-5A) although the SDS-PAGE displayed different PEGylation states of hProCA32.cys. A plausible explanation for higher molecular weight bands observed for fraction 17 (F17) could be a result of sample storage or incomplete quenching of the PEGylation reaction hence continued PEGylation. F18, showed no unPEGylated band but presence of mono (red) and multiPEGylated peaks. As the fractions increased from F18 to F26, the higher molecular weight band reduced with less high molecular weight bands compared to F17 and the higher molecular weight band continued to decrease until F26 where majority of unPEGylated hProCA32 exists. Using S-column, (Figure 3-5B), a lower pH ≤ 4, the pH at which hProCA32.cys is expected to possess a positive charge also resulted in no binding.

Figure 3-5 FPLC characterization of PEGylated and unPEGylated hProCA32.cys
Cation and anion exchange column with no separation observed in A. Q-column and no binding in B. S-column. Even though the PEGylated product had binding to the Q-column, an observable separation could not be achieved due.
3.6.1 Separation of PEGylated and UnPEGylated hProCA32.cys using the mono Q-column

To improve the FPLC chromatogram resolution for better separation efficiency of PEGylated from unPEGylated fractions, mono-Q column was used. Since monoBeads media are based on highly rigid, monodisperse porous beads for high resolution purification and high capacity, it is suitable for achieving high level purity as well as separation of proteins and other biomolecules per charge or isoelectric point with excellent reproducibility and durability it can be used at high flow rates, without compromising its function and stability. We utilized it to achieve a well resolved separation between PEGylated and unPEGylated protein. Figure 3-6 shows a better separation compared to that achieved by the Q- column shown in Figure 3-5A using the same buffer conditions (buffer A: 10mM HEPES and buffer B: 1M NaCl/10mM HEPES pH 8). The PEGylated fraction had a SDS-PAGE band at approximately 18kDa possibly due to slow migration due to PEG reagent as previously discussed in section 3.6.

Figure 3-6 FPLC Characterization of PEGylated and UnPEGylated hProCA32.cys using the mono Q-column.
As can be observed, two peaks (red arrows) with different UV values were observed, the lower UV value for the unPEGylated fraction (P1) is due to absence of protein and can be suspected as the unreacted PEG reagent/unPEGylated hProCA32.cys while P2 is the PEGylated hProCA32.cys. This was confirmed in the iodine staining (notice that the band is around 18kDa, a value closer to the expected 14.2kDa PEGylated band).

3.6.2 Further Optimization PEGylation

It was observed that using a 1:5 protein to PEG reagent ratio resulted in majority of the protein being multiPEGylated (Figure 3-7). A 1:1 and 1:2 ratio was used to prevent the formation of multiPEGylated products,

*Figure 3-7: Separation of PEGylated and unPEGylated hProCA32.collagen by FPLC.*
The unbound contaminants were eluted first, the first major peak (P1) corresponds to the PEGylated fraction.

In addition, a prolonged buffer A binding phase was used to extend the binding phase in the program method to facilitate binging of the components of the PEGylated products. Using the “hold” function, the gradient of the binding phase was kept constant until peak 1 (P1) was completely eluted. This resulted in a better separation of the mixture. As can be observed in the iodine-stained SDS-PAGE in Figure 3-7, P1 is the multiPEGylated peak) while P2 is the monoPEGylated peak, well separated from each other (Figure 3-7).

The use of Hi-Prep G-25 fine sephadex desalting column was introduced to help provide reliable separation of the unreacted PEG reagent (if any) from the PEGylated protein to aid termination of the PEG reaction. As shown in Figure 3-8, the G-25 column was able the separate unreacted PEG (peak 2) which had no visible band on SDS-PAGE. In addition to this, the desalting column also removed bands (arrows) around the 26kDa molecular weight. The major peak (peak 1) corresponding to the PEGylated fraction in the chromatogram showed the expected PEGylation band on SDS. Also, no significant difference in PEGylation efficiency was observed between a 1:1 and 1:2 binding ratio for optimal PEGylation.
Figure 3-8 Characterization of PEGylated hProCA32.Cys.
A. FPLC of G-25 Superdex column separated the residual PEG reagents (little peak) from the protein of interest (hProCA32.Cys). The SDS-PAGE on the right (B) also showed the PEGylated protein (lane 2-6) and lane having no band on the SDS-PAGE gel. Both silver stain and Coomassie staining showed no significant difference between 1:1 binding and 1:2 binding making the newly developed protocol more cost effective while retaining observable PEGylation efficiency. Using the G-25 column helped achieve a reduction in the bands observed around the 26kDa molecular weight.

3.7 Mass Spectrometric Analysis of purified hProCA32.cys
To determine and confirm the accurate mass of the eluents from the PEGylation step, the samples were analyzed by mass spectrometric studies as shown in Figure 3-9 below. The sample was observed to contain a significant number of multiple peaks. The predicted mass for unPEGylated and PEGylated are 12242Da and 14242Da respectively. Figure 4.7 shows matrix-assisted laser desorption/ionization (MALDI) analysis of the protein with a 14.36kDa peak corresponding to 12,242 + 2000 (PEG reagent mass) + 54 (3 H₂O adduct) + 66 (3Na adduct) = 14364Da.
Figure 3-9: Matrix-Assisted Laser Desorption/Ionization (MALDI) spectrum of hProCA32.Cys. The expected peak (blue circle). The additional 120Da peak were suspected to be contributions from three molecules of water and 3 sodium atoms.

3.8 Relaxivity Measurement of PEGylated hProCA32.cys

The relaxivity r1 and r2 values of hProCA32.cys was measured as outlined in section 2.7, the relaxivity of hProCA32.cys was measured in the buffer containing 10 mM HEPES, pH 7.2 at 37 °C. A 2:1 molar ratio of Gd³⁺- hProCA32.cys was used with Gd³⁺ concentrations of 20 mM, 40 mM, 80 mM and 160 mM. The per-Gd³⁺ relaxivity was thereafter calculated. As shown in Figure 3-13, the relaxivity value obtained for hProCA32.cys post PEGylation was 18 mmol⁻¹.S⁻¹ and 27 mmol⁻¹.S⁻¹ for longitudinal (T1) and transverse (T2) relaxation respectively. The values are 46% and 39% lower than the values reported for rProCA32.P40, the PEGylated rat
analogue of hProCA32.cys possibly due to the presence of Ca$^{2+}$ in the binding pocket of hProCA32.cys thereby preventing complete Gd$^{3+}$ loading of the protein.

To ensure the removal of endogenous metal ions such as Ca$^{2+}$ during the purification of hProCA32.cys, EGTA was screened at the following conditions, 30 mM EGTA pH 4.5, 40 mM EGTA pH 4.5, 30 mM EGTA pH 5, 40 mM EGTA pH 5, 50 mM EGTA pH 5, 30 mM EGTA pH 6, 40 mM EGTA pH 6, 50 mM EGTA pH 6) as shown from left to right in Figure 3-11. It was observed that hProCA32.cys had the most stability with no precipitation after pH adjustment to pH 6 compared to the other conditions evaluated. The effect of EGTA on hProCA32.cys was analyzed by SDS-PAGE at different time points as shown in Figure 3-11 with no significant difference between all of the time. The samples were further characterized by UV/Vis spectrometry Figure 3-11 C. A small shoulder peak at the ~ 292nm wavelength (red

---

**Figure 3-10:** Relaxivity plot of hProCA32.cys.  
A. T1-longitudinal relaxation of hProCA32.cys, B. T2-Transverse relaxation of hProCA32.cys

\[
\frac{1}{T_1} = r_1 [\text{Gd}^{3+}] \\
\frac{1}{T_2} = r_2 [\text{Gd}^{3+}]
\]

**Figure Details:**
- **A:** T1 relaxation graph showing a linear relationship with the equation \( \frac{1}{T_1} = 18 \text{ mmol}^{-1}\text{s}^{-1} [\text{Gd}^{3+}] \).
- **B:** T2 relaxation graph showing a linear relationship with the equation \( \frac{1}{T_2} = 27 \text{ mmol}^{-1}\text{s}^{-1} [\text{Gd}^{3+}] \).
arrow) often was used to reflect calcium binding form. The capability of EGTA to eliminate the Ca\(^{2+}\) at 292 nm was observed. It is therefore important to chelate out the residual calcium to make the binding site available for loading of Gd\(^{3+}\).

Figure 3-11 EGTA screening for calcium removal from hProCA32.cys. A solubility studies of hProCA32.cys after addition of different concentration of EGTA with pH alteration, B., SDS-PAGE for optimum incubation time for EGTA reaction of hProCA32.cys, C., UV/Vis, spectrometric detection of calcium in EGTA-loaded versus non EGTA-loaded hProCA32.cys.

20mM EGTA was previously reported [159]. To ensure separation of holo (metal loaded hProCA32.cys) and apo form of hProCA32.cys, two methods were employed. In the first method, the concentration of EGTA was gradually increased to obtain the optimum concentration for Ca\(^{2+}\). In Figure 3-11A, various concentrations, and pH conditions for optimal chelation with EGTA were evaluated. EGTA solubility was more favorable at pH 6 compared to the lower pH tested. Giving that the isoelectric point (pI) of hProCA32.cys is around 5 and the Q-column employed for FPLC purification works better at high pH and to ensure a charged state for hProCA32.cys, pH 7.4 was utilized as the working pH for this experiment. Furthermore, EGTA chelation has a fast kinetics and does not necessarily have to be incubated for longer time as shown in Figure 3-11B, EGTA resulted in a single band compared to the double band observed in Figure 3-3 (arrow) with 40mM EGTA and no noticeable difference with increasing incubation time. EGTA concentration of 40mM at pH 7.4 improved the R1
and R2 values by 26% and 11% respectively. In the second method, hProCA32.cys was dialyze against Chelex 100 resin as depicted in the scheme in Figure 3-12.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Mg^{2+}</th>
<th>Ca^{2+}</th>
<th>Zn^{2+}</th>
<th>Cd^{2+}</th>
<th>Gd^{3+}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetate</td>
<td>0.5</td>
<td>0.5</td>
<td>1.1</td>
<td>1.6</td>
<td>2.0</td>
</tr>
<tr>
<td>Lactate</td>
<td>0.9</td>
<td>1.1</td>
<td>1.9</td>
<td>1.5</td>
<td>2.9</td>
</tr>
<tr>
<td>Maleate</td>
<td>N/A</td>
<td>1.1</td>
<td>2.0</td>
<td>2.4</td>
<td>3.8</td>
</tr>
<tr>
<td>Malonate</td>
<td>2.1</td>
<td>1.5</td>
<td>3.0</td>
<td>2.6</td>
<td>4.3</td>
</tr>
<tr>
<td>D- Tartrate</td>
<td>1.4</td>
<td>1.8</td>
<td>2.7</td>
<td>N/A</td>
<td>4.2</td>
</tr>
<tr>
<td>Malate</td>
<td>1.7</td>
<td>2.0</td>
<td>2.9</td>
<td>2.4</td>
<td>4.8</td>
</tr>
<tr>
<td>Citrate</td>
<td>3.4</td>
<td>3.5</td>
<td>5.0</td>
<td>3.8</td>
<td>7.8</td>
</tr>
<tr>
<td>Daunorubicin</td>
<td>3.7</td>
<td>3.3</td>
<td>4.5</td>
<td>N/A</td>
<td>7.2</td>
</tr>
<tr>
<td>Oxydiacetate</td>
<td>1.8</td>
<td>3.4</td>
<td>3.6</td>
<td>3.2</td>
<td>5.4</td>
</tr>
<tr>
<td>Iminodiacetate</td>
<td>3.0</td>
<td>2.6</td>
<td>7.2</td>
<td>5.7</td>
<td>6.7</td>
</tr>
<tr>
<td>2,6- Dipicolinate</td>
<td>2.3</td>
<td>4.6</td>
<td>6.3</td>
<td>6.7</td>
<td>8.7</td>
</tr>
<tr>
<td>Alanine</td>
<td>2.0</td>
<td>1.2</td>
<td>4.6</td>
<td>3.8</td>
<td>N/A</td>
</tr>
<tr>
<td>Aspartate</td>
<td>2.4</td>
<td>1.6</td>
<td>5.8</td>
<td>4.4</td>
<td>5.7</td>
</tr>
<tr>
<td>Nitrioltriacetate</td>
<td>5.5</td>
<td>6.4</td>
<td>10.7</td>
<td>9.8</td>
<td>11.4</td>
</tr>
<tr>
<td>EGTA</td>
<td>5.3</td>
<td>10.9</td>
<td>12.6</td>
<td>16.5</td>
<td>17.5</td>
</tr>
<tr>
<td>EDTA</td>
<td>8.8</td>
<td>10.6</td>
<td>16.4</td>
<td>16.4</td>
<td>17.4</td>
</tr>
</tbody>
</table>

Chelex-100, a styrene divinylbenzenzene copolymer which contains iminodiacetate ion (Table 3-1) with a high selectivity for divalent ions such as Ca^{2+} was used to chelate possible residual calcium in the hProCA32.cys purification system.
Figure 3-12: Equilibrium dialysis system set up of hProCA32.cys in 20% chelex-100. At equilibrium, it is expected that the chelex-100 buffer with a relatively high selectivity for Ca\textsuperscript{2+} divalent ions will bind to the calcium present in the protein supernatant (dark blue circle represents calcium ions).

0.05g/mL of chelex-100 was added to 10mM HEPES buffer pH 8 (addition directly to hProCA32.cys increases the pH of the protein) and hPoCA32.cys was dialyzed against this buffer overnight at 4\degree C. Using chelex-100 as chelating ligand, \( r_1 \) and \( r_2 \) relaxivity values were increased by 13\% and 6.7 \% respectively. Upon PEGylation of the EGTA incubated samples, the \( r_1 \) and \( r_2 \) relaxivity values were increased from \( r_1 \) and \( r_2 \) values of 18 mmol\textsuperscript{-1}s\textsuperscript{-1} and 27 mmol\textsuperscript{-1}s\textsuperscript{-1} pre-optimizations to about 24 mmol\textsuperscript{-1}s\textsuperscript{-1} and 36 mmol\textsuperscript{-1}s\textsuperscript{-1} respectively post optimization and PEGylation (Figure 3-13 A-H). This represents a 32\% and 35\% increase in the \( r_1 \), \( r_2 \) values respectively. This remarkable increase also reduces the relaxivity difference previously observed with published data for rProCA32 by 27\% and 19\% respectively.

Table 3-2: Comparison of relaxivity values of hProCA32.cys with rProCA32 and clinical contrast agents.

<table>
<thead>
<tr>
<th>Contrast Agent</th>
<th>( r_1 ) at 1.4 T mmol\textsuperscript{-1}s\textsuperscript{-1}</th>
<th>( r_1 ) at 7.0 T mmol\textsuperscript{-1}s\textsuperscript{-1}</th>
<th>( r_2 ) at 1.4 T mmol\textsuperscript{-1}s\textsuperscript{-1}</th>
<th>( r_2 ) at 7.0 T mmol\textsuperscript{-1}s\textsuperscript{-1}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Magnevist (Gd-DTPA)</td>
<td>3.3</td>
<td>5.1</td>
<td>3.9</td>
<td>9.4 ± 1.3</td>
</tr>
<tr>
<td>Eovist (EOB-DTPA)</td>
<td>5.38 ± 0.02</td>
<td>5.37</td>
<td>6.54 ± 0.06</td>
<td>7.01</td>
</tr>
<tr>
<td>Dotarem (DOTA)</td>
<td>3.9 ± 0.2</td>
<td>N/A</td>
<td>3.2 ± 0.7</td>
<td>N/A</td>
</tr>
<tr>
<td>MultiHance (BOPTA)</td>
<td>6.20</td>
<td>N/A</td>
<td>8.7</td>
<td>N/A</td>
</tr>
<tr>
<td>ProHance (HP-DO3A)</td>
<td>4.39</td>
<td>N/A</td>
<td>5.0</td>
<td>N/A</td>
</tr>
<tr>
<td>PEG.ProCA32</td>
<td>33.14 ± 0.32</td>
<td>18.9</td>
<td>44.61 ± 0.12</td>
<td>48.6 ± 0.1</td>
</tr>
</tbody>
</table>
Table:

<table>
<thead>
<tr>
<th>Protein</th>
<th>Value</th>
<th>Error</th>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEG.hProCA32.cys</td>
<td>24 ± 0.41</td>
<td>N/A</td>
<td>36 ± 0.51</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Figure 3-13 R1 and R2 relaxivity plot of hProCA32.cys. The relaxivity values obtained before protocol optimization (A-B) increases with dialysis with chelex-100 (C-D) while incubating with 40mM EGTA increases the value by about 26% for R1 and 11% for R2 from pre-optimization (E,F). PEGylation of hProCA32.cys incubated with 40mM EGTA increases the R1 and R2 values further by 32% and 35% respectively compared to 10mM EGTA.

Next, the relaxivity of hProCA32.cys was obtained by using two different experimental set up in order to probe the real binding ratio between hProCA32.cys and Gd\(^{3+}\) metal. For the first method, as previously described; 2:1 Gd\(^{3+}\)- hProCA32.cys ratio concentration was used. The concentrations of hProCA32.cys were 10µM, 20µM, 40µM, 80µM. The Gd\(^{3+}\) relaxivity was thereafter calculated. For the second method, one of the components was fixed per time while the other is varied (50µM hProCA32.cys with varying Gd\(^{3+}\) concentration and vice versa). As shown in Figure 3-14, either hProCA32.cys or Gd\(^{3+}\) were fixed at 50µM in each of the experiments.
Figure 3-14 Relaxivity studies of hProCA32.cys. T1, T2 time plot of fixed hProCA32.cys (A, D) and Gd$^{3+}$ (B, E). C, F. 1:2 formulation mixture of hProCA32.cys:Gd$^{3+}$ for both PEGylated and unPEGylated protein obtained after FPLC separation using 10mM HEPES pH 8 and 10 mM HEPES pH 8 + 1M NaCl as binding and elution buffers respectively.

The data was also plotted as Job plots to provide both qualitative and quantitative information about the stoichiometry of Gd$^{3+}$ binding to hProCA32.cys[297]. These are shown in Figure 3-15 A, B, D and E. r1 and r2 relaxivity values of 31.77 mmol$^{-1}$.s$^{-1}$ and 45.6 mmol$^{-1}$.s$^{-1}$ were obtained for 1:1 ratio with varying hProCA32.cys concentration and fixed Gd$^{3+}$ concentration. While 33.92 mmol$^{-1}$.s$^{-1}$ and 48.38 mmol$^{-1}$.s$^{-1}$ were obtained for fixed hProCA32.cys. These values are higher than those obtained for the 2:1 binding point in the experiments. The job plot in Figure 3-15 also favored a 1:1 stoichiometric since the peak mostly peaked at the midpoint.
Figure 3-15 Job plot of Relaxivity values for hProCA32.cys. T1, T2 Relaxivity plot of fixed hProCA32.cys (A,B) and Gd^{3+} (D,E). C, F. 1:2 formulation mixture of hProCA32.cys:Gd^{3+} for both PEGylated and unPEGylated protein obtained after FPLC separation.

Due to the proposed two gadolinium binding sites on the protein scaffold. It was expected that the relaxivity value will peak at the 2:1 binding ratio i.e., 100uM Gd^{3+}: 50uM hProCA32.cys in Figure 3-15 for both fixed Gd^{3+} metal and hProCA32.cys. The discrepancy could be insufficient binding due to presence of competing metals or slow Gd^{3+}-binding kinetics [298].

3.8.1 Confirmation of hProCA32.cys binding ratio.

To further probe the low relaxation rates observed for the 2:1 binding stoichiometry, the concentration of free Gd^{3+} after chelation to hProCA32.cys was estimated as described by Barge et. al [299]. Gd^{3+} was added to hProCA32.cys, Gd^{3+} loading was confirmed by UV/VIS spectrometry by increased ~ 292nm wavelength peak compared to pure hProCA32.cys (Figure 3-16A). The free Gadolinium content was then measured using 1µM xylenol orange prepared in 50mM acetate buffer, pH 5.8. The color observed for Xylenol Orange depends in buffer solution is pH dependent. At neutral or acidic pH (acetate buffer solution), xylenol orange
solution is yellow, addition of Gadolinium metal changes the solution pH, with pH values tending towards more basic values due to the deprotonation of the phenolic hydroxyl group on the xylenol orange compound which then leads to an extended delocalization of electrons in solution and the color moves toward the violet end of the electromagnetic spectrum[299]. The estimation of free Gd\(^{3+}\) in this method is based on the differences between the visible spectra of free xylenol orange compared to complexed xylenol orange dye. The first spectrum with no Gd\(^{3+}\) possesses two absorption maxima at 433 and 573 nm. As Gd\(^{3+}\) is added (from 0 to 50 \(\mu\)M), the 433nm maximum decreases while the 573nm maximum increases until saturation is reached. To obtain the concentration of free Gd\(^{3+}\), \([\text{Gd}^{3+}]_{\text{free}}\) the absorbance values of the protein sample is extrapolated from the obtained calibration curve. As shown in Figure 3-16, the xylenol orange absorption increased and decreased as expected at the two maxima. The sample (green arrow) was estimated to contain 0.8\(\mu\)M free Gd\(^{3+}\) by estimation from the calibration curve in Figure 3-16C.

Figure 3-16 Spectrometric estimation of free \([\text{Gd}^{3+}]\) in a 2:1 Gd\(^{3+}\) loaded hProCA32.cys. A. The absorption spectra of 20\(\mu\)M hProCA32.cys with and without Gd\(^{3+}\). B. UV spectrometric determination of Gd\(^{3+}\) complexed by Xylenol Orange. Gd\(^{3+}\) concentration increase (0, 1, 1.5, 3, 7, 15, 20, 25, 30, 35, 40, 45 and 50 \(\mu\)M) results in a decrease of the band intensity at 433 nm and a simultaneous increase of the 563nm band. The green arrow and curve show the absorbance profile of Gd-hProCA32.cys complex with a slight overlap with the 1.5 \(\mu\)M point and a shift from the isosbestic due to a change in pH of the sample containing hProCA32.cys.
Standard curve resulting from absorption values of Xylenol Orange with increasing Gd$^{3+}$ concentration. Points after 25µM have been masked to obtain a better fitting.

The xylenol orange estimation method have been reported to lack accuracy especially in the presence of weakly coordinated metal ions outside the chelation sphere[299]. Xylenol orange was therefore replaced with Rhod-5N, a fluorescent probe with Ca$^{2+}$ binding affinity of 320µM (Invitrogen catalog), a value considerably lower than most calcium indicators. Due to the size similarity between Ca$^{2+}$ and Gd$^{3+}$, Rhod-5N was utilized in the estimation of the free Gd$^{3+}$ concentration of Gd.hProCA32.cys prepared with the following hProCA32.cys:Gd$^{3+}$ ratios; 1:1, 1:1.5, 1:1.8, 1:2, 2:1 and 1:4 (Figure 3-17).

Figure 3-17 Fluorometric method of estimating free Gd$^{3+}$ with Rhod-5N dye using different binding ratio.
A. 1:1 binding of hProCA32.cys to Gd$^{3+}$ B. 1:1.5 binding of hProCA32.cys to Gd$^{3+}$ C. 1:1.8 binding of hProCA32.cys to Gd$^{3+}$ D. 1:2 binding of hProCA32.cys to Gd$^{3+}$ E. 1:4 binding of hProCA32.cys to Gd$^{3+}$ F. 2:1 binding of hProCA32.cys to Gd$^{3+}$ G. Fluorescence intensity plot for binding of Gd$^{3+}$ to Rhod-5N H. Normalized intensity of binding of Gd$^{3+}$ to Rhod-5N shows the binding affinity to be 17 µM. I. Estimated free Gd$^{3+}$ concentration for all the ratio tested.

In Figure 3-18 below, a list of several fluorescent dyes used for several metal binding assays is shown and Rhod-5N and xylenol orange highlighted.

![Chemical structure of some calcium indicators](image)

**Figure 3-18:** Chemical structure of some calcium indicators. A. Fura-2 B. Fura-2-6F, C. Mag-Fura-2, D. Fluo-5N, E. Rhod-5N, F. Xylenol orange. Depending on the binding affinity of the metal of interest to these dyes, each of them can be successfully employed in metal titrations.

The use of Rhod-5N will be a better quantification method because lanthanides generally have a strong fluorescent signal for Ca$^{2+}$ indicators [300]. The gadolinium binding affinity to Rhod-5N was estimated as by previous Yanglab member, Dr. Jin as ~14.5 µM (unpublished data). This value was used in the calculation of free gadolinium concentration upon binding. To do this, three different Gd$^{3+}$: hProCA32.Cys concentrations were incubated and allowed to bind overnight at 4°C for slow and optimized binding. Next, removal of unbound free Gd$^{3+}$ was
achieved using G-25 desalting column. In Figure 3-19, only ratios 1:2 and 1:4 had a second peak (red circles) in their chromatograms. The estimated free Gd$^{3+}$ concentration for 1:2 hProCA32.cys:Gd$^{3+}$ was 3nM (Table 3-3), a value at least 200 times lower than the value obtained for xylenol orange for the same 1:2 ratio. This result highlights that the 1:1 binding ratio with highest relaxivity observed previously may be due to presence of free metal in solution and thus requires removal by desalting as the proposed 1:2 protein: gadolinium ratio had the lowest associated free gadolinium only followed by 1:1 binding.

![Figure 3-19: FPLC Characterization and binding ratio of hProCA32.cys with different ratios of Gadolinium bound with a desalting column via gel filtration. The protein peak with a higher molecular weight is expected to be eluted first while the free gadolinium peak is indicated with red ring. 2:1 and 1:4 binding others had no visible metal peak possible due to just enough or insufficient gadolinium for complex formation.](image)

<table>
<thead>
<tr>
<th>Binding Ration</th>
<th>[Gd$^{3+}$]$_{free}$ (µM)</th>
<th>%Composition</th>
<th>pH After Gd$^{3+}$</th>
<th>pH After EGTA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:1 hProCA32.Cys: Gd$^{3+}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:1.5 hProCA32.Cys: Gd$^{3+}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:1.8 hProCA32.Cys: Gd$^{3+}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:2 hProCA32.Cys: Gd$^{3+}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:4 hProCA32.Cys: Gd$^{3+}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2:1 hProCA32.Cys: Gd$^{3+}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.8.2  **Characterization of hProCA32.cys using Circular dichroism (CD)**

To probe the secondary structure, folding and binding properties of hProCA32.cys, spectra of hProCA32.cys were recorded in a Jasco-810 spectropolarimeter at room temperature. The far-UV CD spectrum of the different ratios of hProCA32.cys:Gd$^{3+}$ were measured in a 1-mm path length cell with 10 mM HEPES (pH 7.4) and the far-UV CD spectrum were obtained. All spectra were obtained as the average of ten scans with a scan rate of 100 nm/min. The ellipticity was measured from 190 to 260 nm (far-UV). This was converted to mean residue molar ellipticity after subtracting the baseline spectrum (buffer sample as blank). To estimate the free Gd$^{3+}$ content, hProCA32.Cys was incubated with gadolinium in different ratio; 1:1, 1:1.5, 1:1.8, 1:2, 2:1 and 1:4 hProCA32.Cys: Gd$^{3+}$ respectively at 4°C for 48hrs. Circular dichroism was introduced to understand the extent of conformational change with metal binding, the 1:2 loading ratio had the closest conformation to the control hProCA32.cys sample (Figure 3-20).

<table>
<thead>
<tr>
<th>Ratio</th>
<th>hProCA32.cys</th>
<th>Gd$^{3+}$</th>
<th>7.47</th>
<th>7.39</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:1</td>
<td>0.025</td>
<td>0.253</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:1.5</td>
<td>0.055</td>
<td>0.552</td>
<td>7.44</td>
<td>7.40</td>
</tr>
<tr>
<td>1:1.8</td>
<td>0.040</td>
<td>0.405</td>
<td>7.42</td>
<td>7.42</td>
</tr>
<tr>
<td>1:2</td>
<td>0.003</td>
<td>0.031</td>
<td>7.42</td>
<td>7.42</td>
</tr>
<tr>
<td>1:4</td>
<td>0.082</td>
<td>0.827</td>
<td>7.43</td>
<td>7.42</td>
</tr>
<tr>
<td>2:1</td>
<td>0.012</td>
<td>0.120</td>
<td>7.42</td>
<td>7.42</td>
</tr>
</tbody>
</table>
3.8.3  Estimation of residual calcium in purification buffer of hProCA32.cys

Since the scaffold protein for hProCA32.cys is a calcium-binding protein, another plausible reason for lower relaxivity for the 1:2 binding ratio compared to 1:1 could be unavailability of the second Gd$^{3+}$ binding pocket on hProCA32.cys due to occupation of the binding pocket by Ca$^{2+}$ ions possibly resulting from presence of Ca$^{2+}$ in lysis/purification buffer. To confirm this hypothesis, FPLC peaks corresponding to calcium loaded (holo) and calcium-free (apo) were collected as separate fractions as shown in Figure 3-21A, a single definite band was observed for fractions in peak 5 compared to double bands in peak 2, this was further confirmed EGTA addition studies in Figure 3-21B. The lanes representing hProCA32.cys with EGTA addition had a single defined band compared to those without EGTA. Furthermore, addition of β-mercaptoethanol (BME) removed the bands belonging to dimer formation (role and effect of dimerization is discussed for hProCA32.collagen in chapter 5). The relaxivity values for apo
and holo hProCA32.cys also showed that the apo conformation of hProCA32.cys is the best condition of Gd$^{3+}$ binding Figure 3-21 C-F.

**Figure 3-21:** Characterization and relaxivity studies of calcium-loaded and calcium-free hProCA32.cys.
A. FPLC chromatogram and SDS-PAGE of hProCA32.cys after Q-column purification B. SDS-PAGE analysis of EGTA and BME addition to peak 2 fraction of hProCA32.cys FPLC in A reveals that EGTA addition removes the lower band in the SDS-PAGE lane 1 and 5 compared the lanes with no EGTA 2 and 7. C-D. $r_1$ and $r_2$ relaxivity values of peak 5 (apo/calcium-free fraction), E-F. $r_1$ and $r_2$ relaxivity values of peak 2 (holo/calcium-loaded fraction) with lower relaxivity further confirming binding site occupation by Ca$^{2+}$ over Gd$^{3+}$.
3.9 Metal binding studies of hProCA32.cys via fluorescence spectroscopy

These experiments were carried out using the Fluorescence spectroscopic technique.

**Trp-Tb$^{3+}$ FRET and Gd$^{3+}$ Competition Assay**

![Diagram of Trp-Tb$^{3+}$ FRET and Gd$^{3+}$ Competition Assay](image)

*Figure 3-22 Scheme of Trp-Tb$^{3+}$ FRET and Gd$^{3+}$ Competition Assay.*

This technique relies on the absorption of light by tryptophan at 315nm followed by the subsequent emission of fluorescence light by hProCA32.cys+Tb$^{3+}$ at 545nm. There is a competition of binding between Gd$^{3+}$ and Tb$^{3+}$ and the fluorescence intensity decreases with increasing Tb$^{3+}$ concentration since Tb$^{3+}$ possesses fluorescent properties compared to Gd$^{3+}$.

The calcium binding affinity of the hProCA32.Cys was estimated using the intrinsic properties of the tryptophan residue incorporated into the protein sequence by site directed mutagenesis F103W [102, 108, 164]. Binding affinities for Zn$^{2+}$, Gd$^{3+}$ and Tb$^{3+}$ were determined, and the general binding scheme is outlined in Figure 3-22. A metal-buffer system using
150mM NaCl was employed[164]. Figure 3-23 shows the hill binding curves for hProCA32.cys to metals yielding $k_d$ values of 0.44nM, 7.2µM, $8.2 \times 10^{-22}$M and $2 \times 10^{-22}$M for $\text{Ca}^{2+}$, $\text{Zn}^{2+}$, $\text{Gd}^{3+}$ and $\text{Tb}^{3+}$ respectively.

**Figure 3-23: $\text{Tb}^{3+}$, $\text{Zn}^{2+}$, $\text{Ca}^{2+}$ and $\text{Gd}^{3+}$ Titration of hProCA32.Cys**

A. The DTPA-buffer system shows an increase in fluorescence intensity with increasing concentration of $\text{Tb}^{3+}$ added up to 5mM (top left) this was also the case for the ProCA32.Cys-$\text{Tb}^{3+}$ plot (bottom left). B. Zinc titration fluorescence intensity and binding plot showing gradual decrease in fluozin-1 fluorescent intensity with increasing protein concentration. C. Calcium binding curve, the EGTA plot shows a decrease as increasing concentration of EGTA was added up to 5mM where there’s been complete chelation of residual calcium in the buffer.
system (top left) while the calcium plot showed a steady increase as more calcium is added (bottom left). The 5mM EGTA added to the buffer was a condition optimized for complete removal of residual calcium in the buffer before more calcium is added to avoid false results. Calcium concentration up to 2mM was then added gradually until saturation was observed. The curve was fit on a 1:1 hill binding equation to obtain the dissociation constant ($K_d$) of the free Ca$^{2+}$ metal. (D) A Tb$^{3+}$ - Gd$^{3+}$ competition assay of hProCA32.Cys, the free Gd$^{3+}$ concentration is calculated from the competition of Gd$^{3+}$ by Tb$^{3+}$ an amino acid residue with fluorescent capabilities. A buffer system was used to monitor possible fluctuations and as a calculation parameter, a total concentration of 20µM Tb$^{3+}$ was added to various concentrations of Gd$^{3+}$ in 10µM. The buffer used buffer contains 50mM HEPES, 100mM NaCl pH 7.2.

3.10 Optimization of Purification protocol of hProCA32.cys

The previously developed and utilized laboratory purification method of ProCAs has yielded relatively good results. However, it possesses two heating steps that may lead to uncontrolled denaturation of the protein of interest. Hydrophobic Interaction Chromatography (HIC) is a widely used protein purification technique that separates proteins according to differences in their surface hydrophobicity by utilizing a reversible interaction between these proteins and the hydrophobic surface of a HIC medium. It complements other techniques that separate according to charge, size, or bio specific recognition. The technique is an ideal next step when samples have been subjected to ammonium sulfate precipitation (frequently used for initial sample concentration and clean-up) or after separation by ion exchange chromatography. In both situations, the sample contains a high salt concentration and can be applied directly to the HIC column with little or no additional preparation. The elevated salt level enhances the interaction between the hydrophobic components of the sample and the chromatography medium. In contrast to other chromatographic methods for purifying proteins (e.g. gel filtration, affinity, and ion exchange), hydrophobic interaction chromatography (HIC) commonly requires experimental determination (referred to as screening or "scouting") to
select the most suitable chromatographic medium for protein purification such as hProCA32.cys. For our protocol, we desired a step that can considerably reduce cost, save time and provide mild treatment condition for the protein. This means that the two-step heating, addition of streptomycin and dialysis step can be taken off the protocol especially because addition of streptomycin has been observed to affect the binding of hProCA32.cys with smeared appearance as can be seen in Figure 3-2. First, several columns with varying degree of hydrophobicity was used to screen for the best condition (Figure 3-24).

![Figure 3-24: Hi-Trap™ HIC Selection kit for rapid screening of hProCA32.cys purification. The different hydrophobic interaction chromatographic columns screened for optimized purification of hProCA32.cys. The high performance (A and G) columns provide better resolution compared to fast flow columns.](image)

The most used chaotropic salts for hydrophobic interaction are (NH$_4$)$_2$SO$_4$, Na$_2$SO$_4$, NaCl, KCl and CH$_3$COONH$_4$. In the first trial, 1.5M NaCl a milder chaotropic salt was used to ensure
stability of the protein post addition of salt. In practice sodium, potassium or ammonium sulfates can effectively promote ligand-protein interactions in HIC and have a stabilizing influence on protein structure[301]. The binding buffer (buffer A) was 50mM phosphate buffer, while the elution buffer (buffer B) was 1.5M NaCl, 50mM phosphate buffer, both were adjusted to pH 7. The results obtained after several trials showed a similar trend for all column tried with no protein binding at any instance using Figure 3-25 as a representative result of the chromatograms obtained.

Figure 3-25: FPLC characterization of hProCA32.cys after different HIC column purification showing lack of binding. A., Octyl fast flow hydrophobic column shows no binding with the protein coming out as flow through during the binding stage B. Butyl fast flow hydrophobic column also showed no binding with the protein coming out as flow through during the binding stage C. Phenyl fast flow hydrophobic column showed no binding to the column since the protein eluted out during
the binding phase D., Phenyl high performance hydrophobic column shows no binding with the protein coming out as flow through during the binding stage. E. SDS-PAGE characterization of the elution fraction from each of the column tried.

Figure 3-26: Addition of ammonium sulfate to protein for HIC purification shows precipitation due to exposure of hydrophobic surface of the protein.

The loading condition was changed such that the protein was loaded unto the program before start of the automated program for each run Ammonium sulfate was also used in place of NaCl in the buffer B but a similar result with no binding was obtained (result not shown. Next, Ammonium sulfate was directly added to the hProCA32.cys solution leading to precipitation that cleared out upon filtration (Figure 3-26). Upon loading to Butyl column, a drop in the UV absorbance (which is usually indicative of protein concentration) was observed. A second peak possibly resulting from binding and elution of pH 7 binding protein (red arrow, Figure 3-27B) was also observed, indicative of a possibility for optimized binding condition. The process was repeated in Figure 3-27C and the same observation was found. Upon using Phenyl High
performance, high substitution, a more hydrophobic column due to aromaticity, two peaks were observed (Figure 3-27D) and SDS-PAGE characterization of the peaks in Figure 3-27E.

Figure 3-27: HIC purification of hProCA32.cys with modified protocol. A loading of hProCA32.cys showing a high and saturated peak at the beginning of the chromatogram indicating a lack of binding, B. Reinjected of unbound hProCA32.cys shows possible binding despite low peak resolution, C. Using butyl (C) and phenyl (D) HIC column.
• Due to the high amount of charged residues present at the binding sites (site 1: DKDKDGFIEEDE; site 2: DKDGDGKIGVDE), the protein pH was adjusted to pH 6, a pH close to the protein pI (5).
• Ammonium sulfate was made as a stock solution and added as a liquid as opposed to the previous solid addition that yielded precipitates.
• The sample was loaded without filtration to ensure hydrophobic patch exposure.
• Lastly, the sample was completely loaded on the column first before running the FPLC program.

Figure 3-28: Screening of HIC column after pH adjustment.
A., Sample loading without pH adjustment. B., Phenyl High performance, high substitution column with pH adjustment. C., Butyl High performance, high substitution column with pH adjustment. E., Precipitation of hproCA32.collagen at pH 5.5 resulting from charged residues and F., 20% Biorad quick colorimetric assay for protein detection showed protein present in the waste.
The primary sequence of the hProCA32.cys contains a considerably high number of charged residues and could be a plausible reason there had been no binding previously keeping in mind that the pH of the protein sample and buffers were adjusted to 5.5, a value relatively close to its pI. The protein precipitated at this pH (Figure 3-11). The result obtained (Figure 3-28 A and C), was not comprehensive due to the unfolding and inability of the protein to bind to the hydrophobic column but a slight binding was observed for phenyl column (Figure 3-28 B). Figure 3-28 F showed that the non-binding waste contained significant amount of unbound hProCA32.cys using 20% Biorad as a quick easy to read method. In addition to optimizing the pH condition increased to pH 6 to effect binding as well as rid the cell lysate of impurities. The protocol was optimized as follows; To 10 g of the Bacterial pellet, 200 mL 50 mM Tris, pH 7.4 was added to the pellet and vortex to mix, the mixture was sonicated and centrifuged at 8000 rpm, 15°C, for 30 mins. SDS page for both supernatant and pellet were obtained, the pH and concentration of ammonium sulfate of the supernatant was adjusted to 7 using 50 mM Tris, 3M Ammonium sulfate, pH 7.0 stock solution to 1.2 M of final concentration of Ammonium sulfate in protein solution, the supernatant was centrifuged again at 8000 rpm, 15°C, for 30 minutes, SDS PAGE for both pellet and supernatant were The resulting supernatant was loaded to phenyl sepharose high performance substitution column (Phe HS column) after the column had been brought to equilibrium with 50 mM Tris, 1.2M Ammonium sulfate pH 7.0. The flow-through was collected since it was expected that hProCA32.cys will not bind to the column at this pH, the loaded column was then washed with 50 mM Tris, 0.2M Ammonium sulfate pH 7.0. Next, the column was washed with 20 mM Tris pH 8.0. The flow through pH was adjusted to pH 6.0 using 3M Ammonium sulfate and 250 mM Na₃H₂PO₄/NaH₂PO₄ pH 5.7 stock solution until the final concentration of Ammonium sulfate is 1.5M and final concentration of
Na$_2$HPO$_4$/NaH$_2$PO$_4$ is 50 mM. The solution was observed for precipitate and cloudiness and in the absence of any, it was loaded unto the Phe HS column. 50 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$, 1.5M Ammonium sulfate pH 6.0 solution was used to equilibrate the column with hProCA32.cys binding to the column at the reduced pH. 50 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$, 0.5M Ammonium sulfate pH 6.0 was used to get rid of the impurities after eluting out the protein with 20mM Tris pH 8.0. The column was regenerated using filtered double distilled water. Figure 3-29A-F described the step-by-step purification of hProCA32.cys at with SDS-PAGE (silver stain) characterization. The UV spectrum also indicated DNA removal with A$_{260}$/A$_{280}$ ratio less than 1 (Figure 3-31B and E) The resulting fraction suspected to be hProCA32.cys is shown as a single clean peak in the chromatogram Figure 3-29 E as well as the UV spectrum in Figure 3-31 B and E. The yield obtained (~19mg/ml) is low relative to lab scale method (80 ± 30 mg/ml). Further analysis is suggested to achieve a much better yield.
Figure 3-29: Ammonium sulfate purification and hydrophobic interaction chromatographic studies of hProCA32.cys.

A. Protein loading to phenyl Sepharose HP column at non-binding pH 7
B. Significant binding of impurities with pI close to 7 on the HIC column with hProCA33.cys collected in the waste. The elution corresponding to undesired proteins showed a high UV absorbance for the protein impurity present in hProCA32.cys lysate
C. Loading of hProCA32.cys to the column had no significant UV peak, D, E. Elution of hProCA32.cys at pH 6 resulted in a relatively pure with predominately high UV peak for several trials of loading for reproducibility. F. Column regeneration after hProCA32.cys load and elution had no significant protein peak.
Figure 3-30: βME (5%) reduced SDS-PAGE of hProCA32.cys after HIC chromatography. The use of this method resulted in a significant reduction of the impurity bands obtained for the lysed protein in lane 1 compared to lane 4 where most high molecular weight bands have been separated. Lane 2 had some of the hProCA32.cys lost with the non-bound proteins while the column regenerating fractions had no significant amount of hProCA32.cys.

Figure 3-31: Absorbance spectra of different fractions obtained during the purification of hProCA32.cys. A. Fraction obtained from the first elution corresponding to presence of protein impurities with pI close to 7 the binding pH. B, E hProCA32.cys fraction. C,D Wash and loading waste
after hProCA32.cys loading. F., overlay of concentrated and diluted fractions. The $A_{260}/A_{280}$ ratio greater than 1 indicates the presence of high DNA component removal in waste and first elution step compared to a 0.68 $A_{260}/A_{280}$ for hProCA32.cys further confirming its relative purity.

### 3.11 T-cell activation and Stability Studies of hProCA32.Cys

The ability of hProCA32.cys to induce a humoral and/or cell mediated immune responses was evaluated in vitro using T-cell activation although it is not possible to measure immunogenicity directly and accurately without injection into human. This measurement is important for evaluation of hProCA32.cys’s susceptibility to provoke immune response for its translational potential as a protein-based MRI contrast agent. In addition, the stability of hProCA32.cys under the conditions for which it will be used are necessary information to submit to regulatory authorities before it can be considered for use in clinical studies. As shown in Figure 3-32C1 and C2, hProCA32 displayed considerably lower immunogenicity compared to positive control (Figure 3-32 B1 and B2) and rat ProCA32 (rProCA32) Figure 3-32 D1 and D2 respectively.

**Figure 3-32:** Cell proliferation monitoring by cell flow cytometry.

Furthermore, the potential for hProCA32.Cys to elicit immunogenic response was evaluated by screening of its amino acid residues, to identifying regions that can be classified as foreign and likely to induce immune response. This was done in collaboration with Dr Zuben Sauna from FDA. The intensity or clarity of the heat map was utilized as a tool to evaluate the binding affinity of the ‘foreign protein’ to the Major histocompatibility complex (MHC) class II protein of humans. Since MHC class II proteins are polymorphic, the intensity distribution of the heat map is proportional to the likelihood of the epitope at each point on the map to be immunogenic. Generally, a clear heat-map indicates a protein that has limited potential immunogenicity and vice versa (Figure 3-33). If some HLA alleles do light up, the weighted promiscuity scores (example shown in Figure 3-34B) will further indicate the relative risk to the population. For the heat map in Figure 3-34A, the X-axis is ProCA sequences, Y-axis is HLA (human leukocyte antigen) [302]. This was used a basis for development of hProCA32.collagen discussed further in Chapter 4.
Figure 3-33: Heat map created in collaboration with Dr Zuben Sauna at FDA. Depicting the affinities of individual MHC-II molecules (y axis) to wild-type peptides (x axis) from regions of factor VIII spanning the three highly recurrent hemophilia A–causing missense mutations analyzed [303, 304].

hProCA32.cys was investigated by computer simulation to calculate its propensity for immunogenicity. Figure 3-34A describes an *in silico* immunogenic assessment method developed to estimate the capability of one or more amino acid in hProCA332.cys to bind to human leukocyte antigen (HLA) component of its major histocompatibility complex (MHC).
alleles [302, 305] and subsequently resulting in immunogenic response. hProCA32.cys residues at positions 30 to 37; FFQMVGLK was found to bind to HLA class II alleles of individuals by computer simulations (Figure 3-34B). These immunogenic residues were optimized in the selection of the right scaffold for targeted contrast agent discussed in chapter 4.

Figure 3-34: Immunogenic assessment of hProCA32.cys.
A. The affinity of individual MHC-II (y-axis) to hProCA32.cys amino acid residues (x-axis) to immunogenicity. hProCA32.Cys displayed a considerably low immunogenicity on the heat map showing only an 8-residue susceptibility. B. The promiscuity scores of the potential immunogenic residues have been weighted for the MHC-II allele frequency of the African, Caucasian, North American, North and South east Asian and the general world populations.
C. 3D structure of hProCA32.cys with immunogenic residues “FFQMVGLK” shown as sticks.
D 3D stick structure of hProCA32cys with “FFQMVGLK” residues shown as spheres.

3D visualization of hProCA32.cys is shown with residues FFQMVGLK as stick (Figure 3-34C) and spheres (Figure 3-34).
3.12 **Screening and Solvent accessibility studies of FFQMVGLK**

The eight residues identified as potentially capable of inducing immunogenicity were individually visualized for possible cysteine mutation to create a cysteine molecule directly in the protein scaffold by site-directed mutagenesis for PEGylation. The solvent accessible surface area (SASA) of hProCA32.cys was estimated on molecular visualization software, the rotamers with the least steric hindrance and highest probability were selected upon mutagenesis In Pymol software (Table 3-4). Each of these residues were visualized as shown in Figure 3-35 to Figure 3-42. Both phenylalanine and valine at positions 30 and 34 were completely buried while residues Q32, M33, and K37 had the highest accessibility/exposure before mutagenesis by simulation with SASA values of 102.17, 79.15, and 84.435, respectively and after computational mutagenesis, the SASA values were similar for all 8 residues. Although glutamine at position 32 had the highest SASA value representing the residue with the highest solvent accessibility and exposure, only one glutamine residue exists in hProCA32.cys and its mutation may have adverse effect since glutamine is considered a biologically important molecule responsible for specific metabolic function [306] and its replacement may further enhance the immunogenic properties of hProCA32.cys. Meanwhile, the K37C mutation having the second highest SASA value of 84.4Å contains 16 lysine residue in hProCA32.cys and a mutation of lysine to cysteine in previous rat ProCA32 had no significant effect on its function (Shenghui Xue unpublished data).
Figure 3-35: Solvent accessibility visualization of mutation of residue phenylalanine at position 30 to cysteine.

Figure 3-36: Solvent accessibility visualization of mutation of residue phenylalanine at position 31 to cysteine.
Figure 3-37: Solvent accessibility visualization of mutation of residue glutamine at position 32 to cysteine.

Figure 3-38: Solvent accessibility visualization of mutation of residue methionine at position 33 to cysteine.

Figure 3-39: Solvent accessibility visualization of mutation of residue valine at position 34 to cysteine.
Figure 3-40: Solvent accessibility visualization of mutation of residue glycin at position 35 to cysteine.

Figure 3-41: Solvent accessibility visualization of mutation of residue lysine at position 36 to cysteine.

Figure 3-42: Solvent accessibility visualization of mutation of residue lysine at position 37 to cysteine.
Table 3-4: Solvent accessible surface area of residues with high promiscuity in hProCA32.cys

<table>
<thead>
<tr>
<th>Residue no</th>
<th>Mutation</th>
<th>Solvent accessible surface area (Å²)</th>
<th>Amino acid composition</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>F30C</td>
<td>0.000/262.194</td>
<td>8</td>
<td>Buried</td>
</tr>
<tr>
<td>31</td>
<td>F31C</td>
<td>0.256/263.801</td>
<td>8</td>
<td>Slightly exposed</td>
</tr>
<tr>
<td>32</td>
<td>Q32C</td>
<td>102.173/264.822</td>
<td>1</td>
<td>Exposed</td>
</tr>
<tr>
<td>33</td>
<td>M33C</td>
<td>79.148/263.494</td>
<td>6</td>
<td>Exposed</td>
</tr>
<tr>
<td>34</td>
<td>V34C</td>
<td>0.000/263.578</td>
<td>6</td>
<td>Buried</td>
</tr>
<tr>
<td>35</td>
<td>G35C</td>
<td>0.255/263.128</td>
<td>9</td>
<td>Slightly exposed</td>
</tr>
<tr>
<td>36</td>
<td>L36C</td>
<td>0.216/262.576</td>
<td>9</td>
<td>Slightly exposed</td>
</tr>
<tr>
<td>37</td>
<td>K37C</td>
<td>84.435/263.139</td>
<td>16</td>
<td>Exposed</td>
</tr>
</tbody>
</table>

Since all the promiscuous residues identified are at positions away from the functional metal binding loops, they can be considered for cysteine mutation for site specific PEGylation or as a scaffold for development of human protein based targeted MRI contrast agents without significantly affecting the function and conformational stability of the protein [167]. F31, E32, M33, G35, L36 and K37 have residues exposed and accessible but the lysine residue at position 37 was mutated to cysteine due to its SASA value and abundance of lysine residues relative to others.

3.13 Transmetallation studies of hProCA32.cys

The transfer or dechelation of Gd³⁺ from the ligand (hProCA32.cys) and chelation of Zn²⁺ to hProCA32.cys was estimated. There are several physiological metals that could be probable to replace Gd³⁺ chelated to the ligand (hProCA32.cys) but the most probable is Zn²⁺, because its concentration in the body is relatively high (~2g) [307]. Zinc is an important divalent cation in biological systems, influencing DNA synthesis, microtubule polymerization, gene expression, apoptosis, immune system function, the activity of enzymes such as carbonic anhydrase and matrix metalloproteinases (MMP), an anti-cancer defense mechanism [308]. Zn²⁺ is also functionally active in synaptic transmission and is a contributory factor in neurological...
disorders including epilepsy and Alzheimer’s disease. Transmetallation can be assessed through the evolution of the paramagnetic longitudinal relaxation rate of water protons at 37°C in phosphate buffer solutions[309]. When the transmetallation of a Gd-complex by diamagnetic Zn\(^{2+}\) ions occur in such a buffered solution, the released Gd\(^{3+}\) ions react with PO\(_4\)^{3-} ions to form GdPO\(_4\) and a decrease in the proton relaxation rate (R1) is thus observed for a transmetallated complex. Seven agents; Gadovist®, Prohance®, hProCA32.Cys, Eovist®, multihance®, magnevist® and omniscan® were analyzed. As seen in Figure 3-43 the PEGylated hProCA32.cys with 1:1 ratio of Gd\(^{3+}\) maintain its longitudinal relaxation values up to over 4000 minutes with ~99% retention depicting that our protein contrast agent is both kinetically and thermodynamically stable, similar to clinical contrast agents of macrocyclic compounds. hProCA32.Cys also remained unexchangeable for over 4000s while linear contrast agents such as magnevist and omniscan are not stable with decreased relaxivity as a function of time.
131

Figure 3-43: Gadolinium stability of lysine PEGylated hProCA32.cys (1 hProCA32.cys :1 Gd³⁺) via transmetallation in phosphate buffer to mimic physiological condition and serum stability studies. Evolution of R1 Relaxation rate at increasing time points with respect to initial R1 value.

### 3.14 Gd³⁺ Biodistribution in Different Organs

7 to 10-wk-old female CD1 mice were injected with 0.28 mmol/kg of hProCA32-Cys (containing 0.56 mmol/kg of Gd³⁺) (n = 3). Blood was collected from the heart (cardiac puncture) 5 days after injection. Tissue samples from the mice were collected after euthanasia. The collected tissues were digested with 70% (wt/vol) HNO₃ at 42°C overnight while the serum samples were sent for mice clinical chemistry studies with PBS injection for the control mice. The results for the clinical chemistry studies are shown in Figure 3-44.
Figure 3-44: Clinical toxicity studies of hPoCA32.cys showing no significant toxicity resulting from hProCA32.cys injection into healthy mice compared to PBS injected mice. A, C. Body protein concentration estimation of hProCA32.cys injected mice showed considerable levels compared to PBS-injected control mice. B. Ratio of protein to ions in mice injected with both hProCA32.cys and PBS had no significant difference in the values D. Enzyme levels including alanine aminotransferase (ALT), alkaline phosphatase (ALP) and aspartate aminotransferase (AST) showed a significant increase in AST value of mice injected with hProCA32.cys indicative of possible injury in liver, brain, heart, or kidney. E, F component level of ions showed no significant difference between PBS and hProCA32.cys mice. The discrepancy observed for glucose level (PBS) and creatine kinase (hProCA32.cys) are due to blood sample hemolysis.
Furthermore, the HNO\textsubscript{3} solution containing tissue digestions were collected the next day and diluted with 2\% (wt/vol) HNO\textsubscript{3} to 10 mL. The samples were filtered with the 0.45\textmu m filter and subsequently analyzed by ICP-OES at gadolinium wavelength of 342.246 nm. YCl\textsubscript{3} at 2 ppm was used as internal standard. The biodistribution of hProCA32.cys was obtained by direct estimation of Gd\textsuperscript{3+} concentration as shown in Figure 3-45.

![Figure 3-45: Bio distribution of Gadolinium of different tissues. Long-term distribution of Gd\textsuperscript{3+} in mice after injection of hProCA32.cys. After injection of hProCA32-Cys (n = 3), for 5 days, Gd\textsuperscript{3+} retention in various tissues was measured by ICP-OES. Data are expressed as mean ± SD. ProCA32-Cys shows considerably low in all organs and the high retention observed in the liver and spleen can be attributed to the very high dosage injected.](image)
3.15 Development of KP mouse model for Lung Cancer for lung imaging

Common mutations in human NSCLC are activating mutations in K-RAS (10–30%) and loss of function point mutations in p53 (50–70%). It is therefore crucial to develop mouse models that will give a wealth of information about the disease progression and possible treatment opportunities. Recent advances in the genomic analysis of lung cancer has proven the importance of certain genes such as the MET, KRAS, LKB1, P53, EGFR, BRAF, PIK3CA, ALK and ROSI gene in the development of lung cancer [310]. Development of KP mouse models was developed by the lab of Dr Zhiren Liu of biology Department, Georgia State University. K-ras: p53 (KP) NSCLC models induce tumors similar to human diseases’ histopathological findings. Activation of an oncogenic allele of K-ras can initiate the tumorigenesis process, additional point mutation of p53 will significantly enhances tumor progression, leading to a more rapid development of advanced lung adenocarcinomas[311]. The resulting KP mice was injected with hProCA32.cys for lung imaging.

In this chapter, the high R1 and R2 relaxation properties of hProCA32.cys made it an attractive candidate for use in the detection of lung cancer in a Kras/P53 (KP) mouse model which contains a conditional point mutation in the transformation related protein 53 gene and a point mutation in the KRAS gene both of which generate non-functional proteins.

3.16 Magnetic Resonance Imaging of lung cancer using hProCA32.cys

KP mice model of about 20-30g were selected for MR Imaging on a 7T Varian MRI scanner at the University of Georgia, Athens. 0.025mmol/kg hProCA32.cys was injected into mice. T1-weighted MR images were collected before and after injection of contrast agents using spin echo sequence with TR = 500 ms, TE = 15 ms, FOV = 4 × 4 cm, matrix = 512 × 512, thickness = 1 mm. T2-weighted pulse sequence was also utilized, fast spin echo MRI of mice were
collected before and after injection of MRI contrast agents with TR = 2 s, exposure = 10ms, effective TE = 40 ms FOV = 4 × 4 cm, matrix = 512 × 512, thickness = 1 mm. The images were analyzed by Image J and MRI Cron software. The intensities of lungs and tumor were plotted over time with corresponding SNR and CNR values estimated. In lungs, with significantly lower proton density, a notable enhancement can be observed as seen in both the grayscale and color MR shown in Figure 3-46. T2-W images of both mice (M3 and M5) had a significant increase in the lung signal to noise ratio upon injection of hProCA32.cys in the tumor area compared to normal lung area. While M3 showed a noticeably diseases lung with the entire right lobe damaged, M5 only had a metastatic lesion (arrows) both of which were taken as region of interest for tumor area calculation in each mouse resulting in a similar value obtained since the overall intensity in both mice were similar (Figure 3-46). In Figure 3-47, a 102% and 182% increase SNR at 3 h and 24 h time points were observed for M5 while M3 had a 63% increase for T2W image SNR at 3h time point (more time points were not obtained due to mouse death). Additionally, a 71% SNR increased was obtained for M3 on T1-W images at 3 h post hProCA32.cys injection time point. Figure 3-48 shows an attempt to correlate the lesion observed on MRI with histology. This needs to be further verified although the size was consistent over the time points from 3h to 48 h time point. Figure 3-49 confirms that the MRI enhancement obtained for M3 was specifically observed in the cranial, accessory, and caudal lobes. This was consistent with hematoxylin and eosin (H&E) staining results.
Figure 3-46: T1 and T2- weighted lung MR images of KP mice models before and after injection of hProCA32.cys.
In the first mice (M3) had a significant increase in selected areas of the lung with the right lobe having no proton density light up (no enhancement) possibly due to normal lung function on T2W images (Top row), his enhancement was further increased in T1W images (middle row). In the second mice however (M5), the lung had no significant enhancement (labelled “L”). At the end of the lung and closer to the liver (LV), an abnormal mass can be observed with increased intensity.

Figure 3-47 T1 and T2- weighted lung MR SNR of KP mice models before and after injection of hProCA32.cys
Figure 3-48 T2-W corona and axial MR images of KP mice M5 with lesions observed post injection correlated with hematoxylin and eosin (H&E) staining showing cell proliferation in observed tumor compared to normal lung tissue.

Figure 3-49 M3 KP mice shows selective MRI enhancement correlating with H&E staining. The areas with high MR intensity were observed to have abnormal lung morphology in H&E.
3.17 Summary

We have developed a human protein-based MRI contrast agent (hProCA32.Cys) using human parvalbumin. A Cys residue was introduced to achieve site-specific PEGylation. The relaxivity value obtained for PEGylated hProCA32.cys was 24 mmol\(^{-1}\).S\(^{-1}\) and 36 mmol\(^{-1}\).S\(^{-1}\) for longitudinal \(r_1\) and transverse \(r_2\) relaxation respectively. The values are 27% and 19% lower than the values reported for rProCA32.P40, the PEGylated rat analogue but at least 5 to 8 fold high than clinically approved contrast agent. The metal binding affinity of hProCA32.cys was obtained to be 0.44nM, 7.2\(\mu\)M, 8.2 \(\times\) 10\(^{-22}\)M and 2 \(\times\) 10\(^{-22}\)M for Ca\(^{2+}\), Zn\(^{2+}\), Gd\(^{3+}\) and Tb\(^{3+}\) respectively. These values are comparable with those obtained for rProCA32.P40 [102] showing high stability for Gd\(^{3+}\) and a 10\(^{11}\)-fold greater selectivity for Gd\(^{3+}\) over Zn\(^{2+}\) compared with existing contrast agents. Preliminary T-Cell Proliferation assay data have been able to show that with hProCA32.Cys developed from human protein scaffold, no significant activation of T-cell was observed. This may give an indication of possibly negligible immunogenicity. Further experiments with direct injection of human subjections are required to confirm its low immunogenicity. Computer simulation using MHC class II alleles of diverse origin by Dr. Zuben also showed that this protein is like to have immunogenicity that is limited to a specific portion of the protein sequence such as residues 30 to 37 in the protein scaffold.

The study of the \textit{in vitro} transmetallation process of Gd-complexes carried out in this work clearly confirms hPoCA32.cys possess a high kinetic and thermodynamic inertness up to at least 4320 minutes retaining its longitudinal relaxation up to ~99%, similar to the stable macrocyclic agents such as Prohance® and Gadovist®. Further optimization with different ratio and formulation conditions are needed to ensure highest metal binding affinity stability, specificity, and thermodynamic inertness.
The developed hProCA32.cys facilitated the detection of NSCLC in a KP generated mouse model for both tumor lesion <0.2mm and advanced lung cancer affecting the entire right lobe without significantly affecting the middle lobe. This was specifically enhanced by hProCA32.cys with up to 71% verified increase in signal to noise ratio. MRI Images shows a promising contrast enhancement post administration of hProCA32.cys contrast and need to be further explored for statistical relevance (for the current report N=1). As a future study, more animal (N≥ 3) should be used with additional control (healthy mice) and histological verification are needed.

3.18 Limitations of this work

In this work, the immunogenicity measurement may not be sufficient for clinical decision since the data were carried out in vitro. A true test of the immunogenicity will require direct injection into humans being the end users. For the MRI experiment, one mouse was used each for the studies, it is therefore important to use more mice for reproducibility and for statistical significance, and with the development of the collagen-targeted contrast agent, the role of cancer associated fibroblast in the tumor microenvironment with collagen deposition can be utilized to target the collagen biomarker overexpressed in the lung cancer patient/mice model. Further studies and correlation between MRI images and histology results are needed. Using a targeted agent such as hProCA32.collagen could potentially help to map out the heterogeneity associated with NSCLC with high sensitivity and specificity.
4 DESIGN AND DEVELOPMENT OF HUMAN COLLAGEN-TARGETED PROTEIN CONTRAST AGENT (HPROCA32.COLLAGEN) FOR MR IMAGING

4.1.1 Introduction

As discussed in Error! Reference source not found., Collagen type I is one of the major diagnostic biomarker and therapeutic target for many chronic diseases including chronic lung disease, liver diseases (e.g., liver fibrosis and cirrhosis), chronic obstructive pulmonary diseases (COPD), different types of cancers and metastasis, heart and kidney failure, nephrogenic systemic and pulmonary fibrosis [225]. There is therefore unmet need to develop noninvasive imaging methodologies and contrast agents to detect early stages of lung fibrosis, and stage fibrosis severity for the progression of different diseases [312]. The sensitivity and specificity of clinically approved MRI contrast agents have been a challenge as high doses of such contrast media are required, and they are non-biomarker targeted [101].

The use of biomarker-targeted MRI contrast agent can therefore be a powerful tool for the diagnosis, monitoring and treatment monitoring human pathologies. Yang lab has pioneered in developing protein MRI contrast agent with significantly improved relaxivity and metal binding. Previous members such as Drs. Xue and Salarian have demonstrated that by using rat parvalbumin scaffold protein, protein contrast agents’ rProCA32 [102], and collagen-targeted rat ProCA32.collagen can be created based on knowledge of collagen binding peptides and related structural biology [159] (discussed in section 1.6).

To enhance the translational capability of our previously reported rat analogue rProCA32.collagen [159] such as overcoming potential immunogenicity potential etc.
future patient application, one of the major effort in this dissertation study is to use a human based protein scaffold hProCA32 (Chapter 3) to develop a human collagen-targeted protein-based MRI contrast agent hProCA32.collagen.

We hypothesized that if hProCA32.collagen maintains the metal binding, collagen binding, safety and relaxivity properties of rProCA32.collagen, hProCA32.collagen will possess similar or better in vivo properties in MR imaging. Combining the clinical relevance, reduced immunogenicity and better tolerance of previously described non-targeted human Protein contrast agent (hProCA32.cys) with the targeting power of an additional collagen binding peptide to hProCA32, we anticipate that we can achieve a precise mapping of collagen in human diseases especially where collagen overexpression and arrangement aids early disease detection and definition of molecular subtypes for treatment.

To test this hypothesis, five developed collagen-binding variants will be characterized using relaxivity, and collagen binding affinity measurement and compared with our previously reported rProCA32.collagen, clinically approved contrast agents, and other collagen targeting agents [195, 211, 226-232].

In this chapter, efforts for the design and characterization of relaxivity and collagen binding capability of several human collagen-targeted protein-based MRI contrast agent (hProCA32.collagen) were first discussed.
Finally, the developed hProCA32.collagen will be used in MR imaging of different chronic mice disease models to address unmet medical needs for early detection of lung fibrosis (Chapter 5), liver fibrosis (Chapter 6) and HCC by MRI (Chapter 7).

Figure 4-1 Development of hProCA32.collagen for molecular imaging.
A. hProCA32 with the mutations for Gd$^{3+}$ binding (yellow balls), B., hProCA32 possesses 16 lysine residues (purple sticks). C., hProCA32.cys with a single cysteine residue at the C-terminus for PEGylation. D., Addition of collagen targeting moiety through a “GGG” flexible linker results in hProCA32.collagen. This targeted variant was overlaid on the previously developed rat analogue with a 96% match after sequence alignment.
**4.2 Results and discussion**

**4.2.1 Rationale for designing variants for collagen targeted contrast agents**

We developed five potential collagen type 1 binding moiety by either using the unaltered hProCA32 backbone, hProCA32 with cysteine added at the c-terminus (hProCA32.cys) or hProCA32 with K37C mutation based on the simulation of HLA alleles as shown in Figure 3-34, section 3.12, Table 4-1. All variants investigated in this chapter were expressed and purified as outlined in section 2.2 and section 2.4.1 respectively.

The sequence of the development of hProCA32.collagen is shown in Figure 4-1. As discussed in 3.1.3, hProCA32 (Figure 4-1A) contains sixteen lysine residues that can be utilized for PEGylation (Figure 4-1B) thereby resulting in non-specific PEGylation of the protein. hProCA32 with (hProCA32.cys) and without cysteine attachment (hProCA32) were utilized for the development of the five potential variants for development of a collagen biomarker-targeted human protein-based MRI contrast agent (hProCA32.collagen).

The first variant termed **hProCA32.collagen1** was designed from the hProCA32 scaffold with no additional mutation. A flexible “GGG” linker was introduced to covalently attach a 17-amino acid residue collagen targeting moiety; KKWHCYTFPHHYCVYG. In the second variant, **hProCA32C.collagen1** hProCA32 scaffold with a K37C mutation at position 37 of the amino acid sequence was linked to the collagen targeting moiety through a flexible GGG linker and a 17-amino acid residue collagen targeting moiety; KKWHCYTFPHHYCVYG. hProCA32c.collagen1 is made up of 130 amino acids with a molecular weight of 14.5kDa and isoelectric point of 5.48 from hProCA32.cys scaffold attached through the GGG linker to the 17-amino acid collagen targeting moiety.
Two additional variants hProCA32.collagen1\textit{LRR} and hProCA32.collagen1\textit{LKK} with modification at the collagen targeting moiety region were developed. The LKK and LRR terms represents the first three residues in the targeting moiety. hProCA32.collagen1LKK contains the hProCA32 scaffold, a K37C, a GGG linker and a newly designed the peptide “\textit{LKKLHLNNNELFPHHY}” as the collagen-targeting peptide while hProCA32.collagen1LRR contains the hProCA32 scaffold, a K37C, a GGG linker and a designed collagen targeting moiety; \textit{LRRLHLNNNRLFPHHY}. The difference between them lies in the fact that hProCA32.collagen1LRR possesses arginine residues in place of lysine in hProCA32.collagen1LKK. Another important difference is the presence of glutamate at the position 123 in the LKK construct compared to arginine is in LRR construct. The glutamate residue has been previously described to be of critical importance for the successful interaction off collagen bonding peptides with type I collagen [244]. Other consideration for amino acid residue selection were charge, hydrophobicity and flexibility [132, 195, 244, 313, 314].

4.2.2 \textit{Expression, Purification and characterization of hProCA32c.collagen1}

As shown in Figure 4-24-34-4C, transformation on an ampicillin coated plate resulted in colony growth indicative of successful transformation of hProCA32c.collagen. The bacterial growth curve as well as the SDS-PAGE characterization in Figure 4-24-34-4B-C also confirms that hProCA32c.collagen1 was indeed expressed at the right molecular weight (red box). The pellet obtained was with a $5.57 \pm 0.30 \text{g/L}$ cell pellet. The yield obtained was about 33.1mg/L with r1 and r2 values of 34.2 mmol$^{-1}.s^{-1}$ and 50.8 mmol$^{-1}.s^{-1}$ respectively. This value is similar...
to the value obtained for rProCA32.collagen1. Peak 3 was further concentrated and characterized.

Figure 4-24-34-4: Expression of hProCA32c.collagen1.
A. Colonies of E.coli PET22b vector after transformation. B. Bacterial growth curve of hProCA32c.collagen1 indicating a lag bacterial expression before the IPTG induction point (orange box) and log phase after IPTG induction. C. SDS-PAGE characterization of expressed hProCA32c.collagen1 before (lane 1) and after IPTG induction (lane 3). The 14.3 band in the protein marker (lane 2) indicates expression of hProCA32c.collagen1 gene relative to the absence of the 14.3kDa band in lane 1 before IPTG induction.
Figure 4-5: FPLC chromatogram of hProCA32c.collagen. Shows a presence of hProCA32c.collagen in the different peaks, including peak 1 eluted during the loading phase where proteins and macromolecules with inappropriate charge (positive) are expected to elute. Peak 3 has similar UV absorption value to peak 1 (~800 mAU). The SDS-PAGE shows that the peaks 1 to 5 contain hProCA32C.collagen in all the peaks. Peak 6 with a UV absorbance of ~2000 mAU had no protein band which could be due to significantly high prevalence of DNA fraction.
Figure 4-6: Characterization of Purified hProCA32c.collagen.
A. UV/Vis spectrum after concentration of peak 3  B. BME reduced SDS-PAGE showed pure hProCA32c.collagen final product.
Figure 4-7: Estimation of the relaxation rates of hProCA32c.collagen1 in 10mM HEPES buffer pH 8.
A., 1/T1 plot of hProCA32c.collagen1 at 0mM, 0.02mM, 0.04mM, 0.08mM and 0.16 mM concentrations of Gd$^{3+}$ to obtain the R1 relaxation rate. B. 1/T2 plot of hProCA32c.collagen1 at 0mM, 0.02mM, 0.04mM, 0.08mM and 0.16 mM concentrations of Gd$^{3+}$ to obtain the R2 relaxation rate. The relaxivity reaction was done in 10mM HEPES pH 7.2 at 37°C.

4.2.3  Expression, Purification and characterization of hProCA32.collagen1C

hProCA32.collagen1C contains 131 amino acids with a pI of 5.7 and a molecular weight of 14.63kDa. As shown in Figure 4-8, the transformation of hProCA32.collagen1C had 2.5x more colonies than hProCA32c.collagen1, the OD$_{600nm}$ obtained on bacteria growth curve at 22 h is 2.5 compared to 2.1 value at harvest for hProCA32c.collagen. The cell mass was $6.00 \pm 0.19g/L$. 
Figure 4-8: Expression of hProCA32.collagen1C.
A. Colonies of E.coli PET22b vector after transformation. B. Bacterial growth curve of hProCA32.collagen1C indicating a lag bacterial expression before the IPTG induction point (arrow) and log phase after IPTG induction. C. SDS-PAGE characterization of expressed hProCA32.collagen1C before (lane 1) and after IPTG induction (lane 3).

Upon purification, the unresolved peak (Figure 4-9A, orange circle) analyzed by SDS-PAGE (not shown) showed similar characteristics and the fractions were therefore pooled together, concentrated and characterized. In Figure 4-9B, the reduced SDS-PAGE showed a pure protein product with $A_{260}/A_{280}$ of 0.59 depicting a higher ratio of protein compared to DNA contamination Figure 4-9C. The yield obtained was 40.2mg/L and the r1 and r2 relaxivity were 43.9 mmol$^{-1}.s^{-1}$ and 65.9 mmol$^{-1}.s^{-1}$ respectively which is significantly higher than those obtained for hProCA32c.collagen1 and rProCA32.collagen1.
Figure 4-9: Characterization of purified hProCA32.collagen1C. 
A. FPLC chromatogram showing a higher absorbance at the 20 – 30% buffer B elution peak corresponding to the peak containing hProCA32.collagen1C. The broad peak was pooled together and concentrated B. The SDS-PAGE of the concentrated peak which contains more protein than DNA (C)

Figure 4-10: Estimation of the relaxation rates of hProCA32.collagen1C. 1/T1 plot of hProCA32.collagen1C at 0mM, 0.02mM, 0.04mM, 0.08mM and 0.16 mM concentrations of Gd^{3+} to obtain the R1 relaxation rate B. 1/T2 plot of hProCA32.collagen1C at 0mM, 0.02mM,
0.04 mM, 0.08 mM and 0.16 mM concentrations of Gd^{3+} to obtain the R2 relaxation rate. The relaxivity reaction was done in 10 mM HEPES pH 7.2 at 37°C.

4.2.4 **Expression, Purification and characterization of hProCA32.collagen1**

Next, hProCA32.collagen1 with hProCA32 scaffold with neither an additional cysteine residue at the C-terminal nor a K37C mutation was evaluated. hProCA32.collagen1 contains a GGG linker and the collagen targeting moiety; **KKWHCYTYFPHHYCVYG**. The protein consists of 130 amino acid residues with a pI of 5.7 and a molecular weight of 14.5 kDa. The transformation showed presence of healthy colonies comparable to those observed for hProCA32c.collagen1 but 2x lower than hProCA32.collagen1C. The SDS-PAGE for the expression also showed no gene expression before IPTG induction (Figure 4-11B, C red box). The expression yield obtained was 6.32 g/L.

![Figure 4-11: Expression of hProCA32.collagen1.](image)

**A.** Colonies of E.coli PET22b vector after transformation. **B.** Bacterial growth curve of hProCA32.collagen1 indicating a lag bacterial expression before the IPTG induction point (red box) and log phase after IPTG induction. **C.** SDS-PAGE characterization of expressed hProCA32.collagen1C before (lane 1) and after IPTG induction (lane 2). The 14.3 kDa band in the protein marker (lane 2) indicates expression of hProCA32.collagen1C compared to no noticeable band at 14.3 kDa band (red box).
Figure 4-12: Characterization of purified hProCA32.collagen1C.
A. FPLC chromatogram showing the highest UV absorbance at the 30% buffer B elution peak corresponding to the peak containing hProCA32.collagen1 (red arrow). The fractions were pooled together and concentrated B. The SDS-PAGE of the concentrated peak which contains more protein than DNA C. reduced SDS-PAGE characterization showed purified protein at the expected molecular weight.

As shown in Figure 4-12, the protein peak (red arrow) eluted at the 30% buffer B elution stage with the highest protein UV absorbance. The protein was pooled together and characterized using UV/VIS as well as SDS-PAGE. The resulting protein had a significantly higher protein/DNA peak with $A_{260}/A_{280}$ value of 0.49. The $r1$ and $r2$ relaxation rates were 30.97 mmol$^{-1}$ and 49.64 mmol$^{-1}$ respectively which is comparable to the value obtained for the rat analogue.
Transformation and expression of newly developed collagen-targeting peptides

hProCA32.collagen1.LKK and hProCA32.collagen1.LRR was cloned and developed after which the sequence was confirmed to be correct by sequencing. Both were expressed, purified and characterized. The results are shown in Figure 4-14. The expression yield for both LRR and LKK were about 5.3g/L and 6g/L respectively with similar bacteria growth curves. The SDS-PAGE for the purification steps showed a successful purification of hProCA32.collagen1.LKK and hProCA32.collagen1LRR (Figure 4-14D lane 9 and 13 respectively). The hProCA32.collagen1LKK and hProCA32.collagen1LRR result was compared with the other three variants based on the previously used collagen-targeting moiety as shown in Table 4-1, Figure 4-18.
Figure 4-14 Transformation and expression of developed LKK and LRR variants of hProCA32 collagen.

A. Transformation of developed hProCA32.LKK/LRR on culture plate, B. weight of pellet obtained post expression of the developed contrast agents, C. optical density growth curve at 600nm (OD\textsubscript{600}) curve showing lag and log phase of the expression D. SDS-PAGE of the expression and purification of hProCA32.LKK (left) and LRR (right).
Figure 4-15: FPLC Chromatogram and UV/VIS characterization of hProCA32.collagenLRR

Figure 4-16: R1 and R2 relaxation rates of hProCA32.collagen1LRR obtained in 10mM HEPES showed high relaxivity values and potentially high MRI sensitivity.

Figure 4-17 shows the results of the relaxivity and calcium binding studies of hProCA32.collagen1.LKK. The reported r1 and r2 values obtained for hProCA32.collagen1.LKK are 30mmol⁻¹ S⁻¹ and 44mmol⁻¹ S⁻¹ respectively while LRR possesses R1 and R2 relaxivity values of 29.79 mmol⁻¹s⁻¹ and 49.76 mmol⁻¹s⁻¹ respectively.
The calcium dissociation constant ($K_d$) value of $8.5 \times 10^{-8}$ M was obtained for hProCA32.collagenLKK compared to rat ProCA32.Collagen with $4.93 \times 10^{-8}$ M but the binding profile showed a low dynamic range.

![Figure 4-17: Relaxivity and calcium binding studies of developed LKK/LRR.](image)

Relaxivity plot $r1$ and $r2$ of hProCA32.collagen1.LKK as well as the EGTA/calcium binding plots (A and B) of hProCA32.collagen1.LKK plotted using Adair (C) and Hill equations (D).

### 4.3 ELISA Assay for the developed hProCA32.collagen variants

Relaxivity values for all the hProCA32.collagen variants showed a considerable high relaxivity (Table 4-1) comparable to the previously developed rProCA32.collagen. As shown in Figure 4-18, collagen binds the protein variants with increasing protein concentration and eventually saturating at high concentrations. The binding dissociation constants for rProCA32.collagen and hProCA32.collagen1 were $1.42 \, \mu$M and $1.85 \, \mu$M respectively. For the other variants, the binding affinities were at least 10 x lower than rat analogue.
Figure 4-18: ELISA Assay results of developed hProCA32.collagen variants.
A. ELISA binding plot for rProCA32.collagen1 shows a high absorbance range with a sigmoidal binding profile with saturation at high concentration. B. ELISA binding plot for unmodified hProCA32.collagen1 also shows an absorbance range closely matching the rat analogue with a sigmoidal binding profile. C-F. ELISA binding plots for the other four variants shows at least 3x lower absorbance values and lower binding affinity (Table 4-1).

Table 4-1: Summary of biophysical characteristics of hProCA32.collagen1 variants; (ND: Not determined)

<table>
<thead>
<tr>
<th>Variant</th>
<th>Pellet(g/L)</th>
<th>Relaxivity R1, R2</th>
<th>Yield (mg/L)</th>
<th>Collagen binding (Kd), µM</th>
<th>Log (Ko)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rProCA32.collagen1</td>
<td>6.10 ± 0.20</td>
<td>34mmol⁻¹S⁻¹, 50mmol⁻¹S⁻¹</td>
<td>37</td>
<td>1.42</td>
<td>8.7</td>
</tr>
<tr>
<td>hProCA32.collagen1</td>
<td>6.32 ± 0.12</td>
<td>31mmol⁻¹S⁻¹, 50mmol⁻¹S⁻¹</td>
<td>35.2</td>
<td>1.85</td>
<td>ND</td>
</tr>
<tr>
<td>hProCA32.collagen1C</td>
<td>5.57 ± 0.30</td>
<td>43mmol⁻¹S⁻¹, 66mmol⁻¹S⁻¹</td>
<td>40.2</td>
<td>15</td>
<td>ND</td>
</tr>
<tr>
<td>hProCA32.collagen1LKK</td>
<td>6.15 ± 0.10</td>
<td>30mmol⁻¹S⁻¹, 44mmol⁻¹S⁻¹</td>
<td>51</td>
<td>15</td>
<td>7.7</td>
</tr>
<tr>
<td>hProCA32.collagen1LRR</td>
<td>5.41 ± 0.21</td>
<td>31mmol⁻¹S⁻¹, 50mmol⁻¹S⁻¹</td>
<td>50</td>
<td>23.5</td>
<td>ND</td>
</tr>
</tbody>
</table>
4.3.1 **Summary, conclusion and limitations in the selection of hProCA32.collagen1**

*as the new human collagen-targeted contrast agent*

To develop a human collagen targeted hProCA32.collagen, 5 variants based on previously designed hProCA32 scaffold with or without cysteine amino acid residue were carefully selected. hProCA32.collagen1, designed with no additional cysteine residue had the highest collagen binding affinity (1.85 µM) and a $r_1$ and $r_2$ relaxation rates of 31 mmol·S⁻¹ and 50 mmol·S⁻¹ respectively comparable to the previously reported values for rProCA32.collagen1. The protein yield obtained was 35.2 mg/L. Conversely, hProCA32.collagen1C displayed a higher protein yield (40.2 mg/L), 38.7% and 32% increase in $r_1$ and $r_2$ relaxation rates compared to hProCA32.collagen1. However, the collagen binding affinity is 8.1 and 10.6 times lower than the values obtained for hProCA32.collagen1 and rProCA32.collagen1 respectively. For the hProCA32.collagen1.LKK and hProCA32.collagen1LRR variants with a different peptide composition in the collagen targeting moiety, a protein yield at least 24.4% higher than the KKWHCYTYFPHHYCVYG-based hProCA32.collagen1 was observed. Due to this increased yield, hProCA32.collagen1LKK with >50% higher collagen binding affinity than the hProCA32.collagen1LRR analogue was considered for further optimization and a possible improved collagen-binding property. Due to lack of improved collagen binding affinity and low dynamic range observed for the calcium binding plot, hProCA32.collagen1LKK was not push further for development of the human collagen contrast agent. Since the long-term goal in the use of the developed contrast agent is dependent on both relaxivity for highly sensitive MR imaging and collagen affinity for biomarker specificity, hProCA32.collagen1 (hereinafter
referred to as hProCA32.collagen) was selected for further optimization, characterization and application in various diseases discussed in chapters 5 to 8.

4.3.2 Future studies

The collagen binding peptide of hProCA32.collagen may be further utilized with hProCA scaffold with additional cysteine mutation both at the c-terminus and at any other positions in the scaffold protein since all the variants with additional cysteine in this study also displayed reduced binding affinity possibly caused by the presence of cysteine in the protein scaffold.

5 FURTHER PRODUCTION OPTIMIZATION OF HPROCA32.COLLAGEN

5.1 Introduction

In Chapter 4, we discussed the steps leading to the design and selection of human collagen protein-based MRI contrast agent hProCA32.collagen. This protein had the highest collagen binding affinity with relaxivity values similar to the previously reported rProCA32.collagen1. A key steps in the development of hProCA32.collagen as a contrast agent in MR imaging of human diseases is the formulation of the protein drug into a dosage form with the ability to achieve the minimum 18-month shelf life and stability during long-term storage[315]. PEGylation has been shown to help with stability and solubility, but it is important to identify factors that may result in protein degradation, aggregation, misfolding which might lead to adverse effects during and after drug delivery including anaphylactic shock. Some of these factors include the ionic strength and concentration of the buffer employed, pH, surface active agent and excipient where applicable [316, 317]. It is therefore important to start the formulation optimization by using a more stable vector with high expression yield and long-
term stability. Cysteine, although found in relatively low abundance on protein is one of the critical amino acids present in functionality sites of proteins [318]. The sulfur atom present in cysteine can exist in a range of oxidation states from -2 to +6 making cysteine’s thiol suitable for use as a reducing agent or nucleophile while its disulfide form (cystine) can act as an oxidizing agent or an electrophile (Figure 5-1) [319, 320].

![Diagram of cysteine oxidation states and reaction pathways](image)

**Figure 5-1:** Examples of biologically relevant oxidative states and reaction pathways of cysteine [320]

In this chapter, several efforts in the optimization of expression, purification and characterization of hProCA32.collagen will be discussed. Several optimization steps was employed from the expression through purification step to ensure the production of pure hProCA32.collagen final product with high relaxivity, sensitivity and biomarker sensitivity.
First, due to the absence of cysteine residue in this protein scaffold and with the assumption that the two cysteine residues in the targeting moiety are in cyclic conjugation with each other, N-terminal PEGylation was done for PEGylation specificity in place of cysteine PEGylation. Furthermore, with the presence of two cysteine molecules in hProCA32.collagen, it is pertinent to understand the plausible mechanism of collagen binding to determine the effect and extent of involvement of the individual or both cysteine molecules present in the targeting moiety of the protein. A major question will then be; is it better to the cysteine formed between a monomer molecules as intra-disulfide bond or in a multimer such as dimer as inter disulfide bond? To answer this question, a protocol designed to produce a high level of monomer was developed with the used of varying ratio of Glutathione disulfide (GSSG) and Glutathione (GSH) (The details of the protocol in not discussed report). Additionally, Improper folding and protein misfolding are the major causes of protein aggregation [321]. Aggregation is an umbrella term for the interactions leading to self-association of protein macromolecules into clusters different from the native quaternary structure of the protein[322]. Under normal physiological conditions, aggregation of protein is prevented but several factors may contribute to protein aggregation in vivo such as pH, temperature, ionic strength, agitation, the presence of metal ions and the amino acid composition of the protein may make the protein susceptible to destabilization reactions during manufacturing or storage[323, 324]. Failure to control and prevent protein aggregation can significantly limit the successful development and commercialization of protein biologics such as hProCA32.collagen [317, 325, 326]. It is difficult to associate protein aggregate with a particular stage of protein production because it can occur at any point of the production process ranging from the gene expression stage to different steps during purification, formulation and storage [327-331]. Therefore, several
factors were optimized in the purification of hProCA32.collagen to reduce and possibly eradicate the formation of aggregates during and after protein production. Further, the selectivity of hProCA32.collagen towards different types of collagen; I, III and IV was evaluated showing the high specificity of hProCA32.collagen to type I collagen. Lastly, the biodistribution and in vivo stability were examined.

5.2 Materials and Methods

5.2.1 Collagen binding studies of hProCA32.collagen

Enzyme-Linked Immunosorbent Assay (ELISA) is a highly specific, sensitive and easy-to-perform technique designed for detecting and quantifying substances such as peptides, proteins, antibodies and hormones. In this experiment, the antigen (collagen type 1) is first immobilized on a solid surface (96 well plate) and then complexed with an antibody that is linked to an enzyme (hProCA32.collagen added to the antigen after blocking with Bovine Serum Albumin (BSA). Detection is accomplished by assessing the conjugated enzyme activity via incubation with a substrate to produce a measurable product as and enzymes on the antibody elicit a chromogenic or fluorescent signal. The specificity of this method requires an exact antigen to be recognized by corresponding antibody with a substrate with an effectiveness of detection and quantification of the enzyme reaction. Several types of ELISA methods have been successfully utilized and are described in Figure 5-2.

1. **Direct ELISA**: In direct ELISA, a target protein or antibody is coated on the surface of a microplate such as a 96-well plate and then incubated with an enzyme-linked antibody. The activity of the enzyme is next evaluated as absorbance at 450nm. With the aid of a substrate designed for that specific task. This is the simplest form of ELISA assay.
2. **Indirect ELISA**: This ELISA method like direct ELISA starts with coating the target protein on the microplate after which it is incubated with a primary antibody. Next, an enzyme-linked secondary antibody against the primary is introduced. Substrate is introduced and the activity of the enzyme is measured as absorbance at 450nm.

3. **Optimized ELISA (used in this report)**: This is basically indirect ELISA with slight modifications. The microplate is coated with the target antigen (collagen type I) which is then incubated with the protein of interest (enzyme). A primary antibody is introduced followed by a secondary antibody and substrate for quantification of enzyme activity as absorbance at 450nm.

4. **Sandwich ELISA**: The microplate is first coated with an antibody against the target protein then incubated with the target protein. This is then incubated with another target protein-specific enzyme-linked antibody. Substrate is added and detection is carried out by quantifying absorbance at 450nm. The antibody that was coated on the plate as well as the enzyme-linked antibody should have the capacity to recognize different epitopes of the protein of interest. This method is highly accurate and specific

5. **Competitive ELISA**: An antibody specific for a target protein is immobilized on the surface of microplate wells and incubated with samples containing the target protein and a known amount of enzyme-labeled target protein. After the reaction, the activity of the microplate well-bound enzyme is measured as absorbance at 450nm. When the antigen level in the sample is high, the level of antibody-bound enzyme-labeled antigen is lower and the color is lighter. Conversely, when it is low, the level of antibody-bound enzyme-labeled antigen is higher and the color, darker. The kind of binding profile obtained for this method is usually opposite the other methods.
The method employed in this study is described in Figure 5-2C and has been employed for several batch and optimized conditions for purified protein.

5.3 Results and discussion

5.3.1 Glycerol stock generation

To ensure a reproducible protein production technique, it is imperative to establish the right clones that carries the hProCA32,collagen backbone. The successful clone(s) with the best expression yield as well as the right molecular weight (14.5kDa) on SDS-PAGE was selected as described in 2.3.1. As shown in Figure 5-3 A, 14 colonies were selected from the ampicillin-coated agar plate out of which clones 1, 5, 7 and 9 were selected. Next, the expression yield of each of the clones were tested in both Luria broth (Zongxiang Gui unpublished data) and
terrific broth media. The soluble to insoluble ratio for the selected clone were estimated by imageJ analysis as shown in Figure 5-4 and Table 5-1. From Table 5-1, it can be observed that a similar expression yield was obtained for soluble and insoluble fractions at the 4h time point (Figure 5-4A). Conversely, at the 18h time point (harvest), an expression band with relatively higher intensity band was obtained for the soluble fractions compared to the insoluble fractions Figure 5-4B. In Figure 5-4C, whole cell, containing both soluble and insoluble portion of the protein were analyzed by imageJ of the SDS-PAGE before, 4 h and 18 h post induction time points of all selected clones. The expression yield for whole cell normalized to 100% at 4 h time point showed a decrease in expression to 40%, 50%, 40% and 45% for clones 1,5,7, and 9 respectively at the harvest time point. The soluble to insoluble ratio at 4 h post induction time point were 50/50, 48/52, 51/49 and 55/45 while the ratio at harvest were 68/32, 75/25, 65/35 and 51/49 for clones 1,5,7 and 9 respectively. To confirm if the protein expressed was the desired/target protein, western blot Figure 5-4D was done and as can be observed, no band was observed before induction compared to after induction time points, the soluble portion also had the highest band intensity despite the overexposure and high loading concentration of the western blot.

![Figure 5-3: Clone screening and selection for hProCA32.collagen.](image-url)
A., LB agar plate with labelled colonies transformed and expressed to obtain the best clone. B., SDS-PAGE of the transformed bacterial colonies 1 to 14 (lanes 1-14) with the expected protein band at 14.5kDa identified (green box).

Table 5-1: Summary of the weight of pellet and other expression condition

<table>
<thead>
<tr>
<th></th>
<th>Clone 1</th>
<th>Clone 5</th>
<th>Clone 7</th>
<th>Clone 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight of pellet in 40mls @ 4 hrs</td>
<td>0.74</td>
<td>0.84</td>
<td>0.76</td>
<td>0.76</td>
</tr>
<tr>
<td>Vol. of HEPES added @ 4 hrs</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Weight of pellet in 40mls @ 18 hrs</td>
<td>0.99</td>
<td>1.03</td>
<td>0.98</td>
<td>1.02</td>
</tr>
<tr>
<td>Vol. of HEPES added @ 18 hrs</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Weight of pellet in ~470mls (Harvest)</td>
<td>8.38 (*11.75)</td>
<td>9.4(*11.75)</td>
<td>8.91(*11.75)</td>
<td>9.17(*11.75)</td>
</tr>
<tr>
<td>% of soluble/insoluble protein@ 4 hr</td>
<td>50/50</td>
<td>48/52</td>
<td>51/49</td>
<td>55/45</td>
</tr>
<tr>
<td>% of soluble/insoluble protein@ harvest</td>
<td>68/32</td>
<td>75/25</td>
<td>65/35</td>
<td>51/49</td>
</tr>
<tr>
<td>% expression over time</td>
<td>12/100/40</td>
<td>22/100/50</td>
<td>14/100/40</td>
<td>15/100/45</td>
</tr>
<tr>
<td>Estimated g soluble protein expressed @ 4 hr per 0.5L</td>
<td>3.45/4.60</td>
<td>3.78/5.04</td>
<td>3.64/4.86</td>
<td>3.92/5.23</td>
</tr>
<tr>
<td>Estimated g soluble protein expressed @ harvest per 0.5L</td>
<td>6.31/8.42</td>
<td>7.24/9.66</td>
<td>5.97/7.96</td>
<td>4.88/6.50</td>
</tr>
<tr>
<td>Estimated g soluble protein expressed @ harvest per 0.5L based on decreased band intensity SDS</td>
<td>3.79/5.05</td>
<td>3.62/4.83</td>
<td>3.58/4.77</td>
<td>2.69/3.58</td>
</tr>
</tbody>
</table>

*Expected based on ideal condition calculation

This increase in expression level at 4 h post induction time point (Figure 5-4A) however simultaneously occurs with at least a 40% decrease in expression of gene of interest at the harvest time point indicated by decrease intensity on SDS-PAGE in Figure 5-4B. Overall, clones 5 had the best expression condition in both LB medium (done by Zongxiang Gui) and TB (shown in this report) and was subsequently used in glycerol stock generation described in 2.3.3.
Figure 5-4: Clone Selection and optimization in terrific broth medium.

A., SDS-PAGE of soluble and insoluble portions of clones 1, 5, 7 and 9, at 4 h time point B., SDS-PAGE of soluble and insoluble portions of clones 1,5,7 and 9, at harvest, C. SDS-PAGE of whole cell in clones 1,5,7 and 9, before, 4 h and 18 h after IPTG induction, D., western blot of clone 5 at different time points

5.3.1.1 Expression protocol for hProCA32.collagen from glycerol stock of clone 5

This batch of expression was carried out at the bioexpression and fermentation facility (BFF) of the university of Georgia. In the first run (Figure 5-5A-B), both pre-induced and post induction sample contains showed similar expression of hProCA32.collagen gene as soluble and insoluble fractions obtained by lysis of cell pellet in B-per buffer. To further probe the origin of the “leaky expression”, chloramphenicol was used in addition to ampicillin antibiotics. However, the use of chloramphenicol antibody shown in Figure 5-6Figure 5-7 appears to have inhibited the expression of the gene of interest.
Figure 5-5: Expression of hProCA32.collagen (clone 5) by BFF at 4 hours and 21 hours post induction with ampicillin showed no difference before and after IPTG induction.

A. First trial using four different media such as LB, and TB-LG: Terrific broth with 0.4% glycerol (low glycerol) and TB-HG 2% glycerol (high glycerol).

B. Second trial using the same condition as A. During this experiment, the TB-HG condition had an 85% soluble content (red arrow).

Figure 5-6: Expression of hProCA32.collagen (clone 5) by BFF at 4 hours and 21 hours post induction with ampicillin and chloramphenicol antibody shows no obvious expression after
IPTG induction. Marker bands from top to bottom (250, 150, 100, 75, 50, 37, 25, 20, 15, and 10 kDa).

In the second trial of the ampicillin only expression, the result (Figure 5-5B) showed a better expression for the TB media with 2% glycerol at 4 h post injection time point with an 85% to 15% soluble to insoluble ratio compared to the chloramphenicol/ampicillin mix shown in Figure 5-6. The weight of pellet obtained from 20L expression of the TB-HG condition was about 500g (25g/L) representing a 5 x increase in expression yield obtained during laboratory scale expression. This obtained pellet was subsequently purified as described in 2.4. The purified pool of fraction from the FPLC in Figure 5-7A, C showed most of the protein to be either in the 5% pool or 20% pool while higher gradients (40 and 60%) showed multiple bands equivalent to less presence of impurities. Therefore, the fractions from lower gradients 5-20% were pooled and exchanged into the pH 7.2 HEPES buffer. 5g protein equivalent to a yield of 250mg/L was obtained for this purification process.

Figure 5-7: Expression and Purification of hProCA32.collagen (clone 5) by BFF at 4 hours and ~18 hours post induction with ampicillin antibody only.
A,C. FPLC chromatogram of hProCA32.collagen at BFF B. SDS-PAGE characterization of the different purification stages of hProCA32.collagen D,E. SDS-PAGE characterization of the different fraction from the FPLC characterization of hProCA32.collagen. Buffer A is 10mM HEPES, Buffer B: 10mM HEPES and 1M NaCl.

5.3.2 **Laboratory Scale (shake flask) purification of hProCA32.collagen**

At the laboratory scale level in GSU, a simultaneous expression and purification of clone 5 of hProCA32.collagen was done as described in 2.4.1. As shown in Figure 5-8, the bands not consistent with the molecular weight of hProCA32.collagen gradually reduced as the purification steps proceeded. Peaks 2 and 3 from the FPLC chromatogram in Figure 5-9A was pooled, concentrated and analyzed by SDS-PAGE with significantly high monomer ratio (85%) compared to dimer fraction (about 15 %) Figure 5-9B.
Figure 5-8: SDS-PAGE characterization of the purification stages of hProCA32 collagen from lysis step to dialysis stage preceding the FPLC separation.
Figure 5-9: Characterization of lab-purified hProCA32.collagen. A. FPLC chromatogram of showing higher protein peak (peak 2 and 3) than DNA peak (peak 4). B. The protein was concentrated and analyzed on a non-reduced SDS-PAGE showing pure protein product with about 15% dimer portion.

5.3.3 Relaxivity Studies

The relaxation times (T1 and T2) were measured by Bruker relaxometer (1.47 T). The contrast agents with different concentration were prepared in 50 mM HEPES, pH 7.4. The relaxivity was calculated using equation 1.5. The relaxivity of ProCA3 variants were measured in the buffer containing 50 mM HEPES, 150 mM NaCl, and pH 7.2 at 37 °C. The relaxivity values obtained were 34mmol·s⁻¹ and 51 mmol·s⁻¹ for R1 and R2 respectively as shown in Figure 5-10.
Figure 5-10: R1 and R2 Relaxivity measurement of PEGylated hProCA32.collagen at 1.4T in 10mM HEPES pH 7.2.

5.3.4 Purification of hProCA32.collagen using Dasgip Parallel Bioreactor Systems

To test the best condition for the selected clone 5, a small-scale condition selection was tried where four conditions; LB, TB, HDCC (developed by BFF) and auto induction condition were observed in a DASGIP bioreactor system to ensure control and targeted optimization of the expression parameters pH and temperature. The first run (Figure 5-11) was inconclusive as a great percentage of expression was shown in the pre-induced samples both soluble and insoluble fractions. To avoid this, chloramphenicol antibody together with ampicillin was in the expression to avoid leaky expression that could have been the reason for the observed expression. The new result is shown in Figure 5-12, despite the absence of “leaky expression”, it appears as though the presence of chloramphenicol hindered the expression of the gene of
interest since the normalized gel (OD<sub>600</sub> of 1) showed no significant expression on the SDS-PAGE.

**Figure 5.11:** Expression of hProCA32.collagen (clone 5) by BFF at 4 hours and harvest (21 hours) after induction. *Marker bands from top to bottom (250, 150, 100, 75, 50, 37, 25, 20, 15, and 10 kDa)*
Next, ampicillin only with a more tightly controlled fermenter was done for 20L expression. The result, shown in Figure 5-13 showed a better expression and no significant protein expression before IPTG induction compared to previous result obtained. At the 4-hour post induction time point, more expression at the expected 14.5kDa band is observed (red arrow) compared to what is seen at the 18h harvest time point (yellow arrow). This clearly indicates that the expression system is not stable to maintain expression up to 18 hrs. The weight of pellet obtained from 20L was about 500g which was subsequently purified using the established protocol. The purified pool of fraction from the FPLC in Figure 5-13C revealed that most of the protein with high purity are either in the 5% pool (purest, red arrow) or 20% pool while higher gradients (40 and 60%) showed multiple bands equivalent to less pure fractions and oligomerization. Therefore, the fractions from lower gradients 5-20% were
pooled and exchanged into the pH 7.2 HEPES buffer. From this pool, a total of 5g of protein was collected which is equivalent to a yield of 250mg/L calculated based on converting the amount recovered using the molecular weight of hProCA32.collagen and then dividing by the total grams expressed per liter.

Figure 5-13: Expression and Purification of hProCA32.collagen (clone 5) by BFF at 4 hours and ~18 hours post induction with ampicillin antibiotic only. A and C) Purification of hProCA32.collagen using Q Sepharose column. B) Reduced expression gel of hProCA32.collagen. Marker bands from top to bottom (250, 150, 100, 75, 50, 37, 25, 20, 15, and 10 kDa). D and E) Reduced purification gels of hProCA32.collagen samples from purification done in A and C. Marker bands from top to bottom (250, 150, 100, 75, 50, 37, 25, 20, 15, and 10 kDa).

5.3.4.1 Characterization of Samples Received from Bio expression Fermentation Facility (BFF), UGA.

The final product from the purified hProCA32.collagen expression run was analyzed by SDS-PAGE as shown in Figure 5-14. In lane 8 of the 5% BME reduced SDS-PAGE in Figure 5-14, the purified sample even though it was a pool of fractions with mostly monomer, the final
sample contains impure fractions with molecular up to 40kDa in addition to the desired 14.5kDa band. The lane corresponding to column wash (Figure 5-13D) also showed a high level of purity compared to the high gradient fractions. This pool of wash was subsequently re-dialyzed and re-concentrated for enhanced purity. Another SDS-PAGE with reduced conditions of the crude wash, final sample, wash protein purified (in GSU) was ran. The result (Figure 5-14B) showed mostly pure band with similarities to the purified wash sample (lane 2 and 4 respectively). The wash which initially contained a significant amount of impurities was repurified successfully. Next, all FPLC samples that were saved for further characterization during the FPLC process were re-analyzed on SDS-PAGE (Figure 5-14C-D). The first six lanes in Figure 5-13C were the wash fractions pooled together as crude wash and shown in lane 3 of Figure 5-14B re-dialyzed and concentrated for enhanced purity as observed in Figure 5-14B lane 4.

Figure 5-14: Confirmation of the purity level of obtained hProCA32.collagen from BFF.
A. Final protein product after purification steps, B. reanalysis of BFF final product, crude wash and protein obtained in the wash step repurified at GSU shows similarity in purity between the further purified wash and the final protein sample compared to the original sample sent as the washout protein with high impurity bands. C., SDS-PAGE of representative fractions and wash from all six injections. D., SDS-PAGE characterization of fractions. Marker bands from top to bottom in C and D (250, 150, 100, 75, 50, 37, 25, 20, 15, and 10 kDa).

5.4 Optimization expression and purification with rhamannose induced expression vector

To further facilitate large scale protein production in E.coli for clinical application, we devoted effort to develop L-rhamnose-inducible system with rhaBAD promoter for better tightly regulated and tunable system [332-334]. A new vector called pD861-SR vector containing a rhamnose-inducible promoter rhaBAD (P_rhaBAD) capable of high-level protein expression in the presence of L-rhamnose for tight regulation by glucose in the absence of rhamnose was developed by Dr. Qiao [335].

Working with Dr. Liangwei Li and Zhongxiang Gui in Yang lab the best rhammose concentration for optimal expression condition via clone selection were screened. Its expression and purification were performed in a similar way to the old vector pet22b described in section 2.3 and 2.4.

5.5 Comparison of collagen binding affinity of hProCA32.collagen purified from expression with old and new vectors

Purified products from both old (pet22b) and new (pD861-SR) were further characterized by ELISA to compare the efficiency of new vector cloned hProCA32.collagen to collagen type I binding. Old vector purified protein was purified in two batches; one using fermentation method (20L) and the other using small scale shake flask (lab)
Figure 5-15: ELISA assay for hProCA32. Collagen purified in GSU, UGA from old vector and new vector from GSU flask shake purification with and without reducing agent, BME.

A. ELISA plot of shake flask purified hProCA32.collagen (old vector) with an without reducing agent, B. ELISA plot of shake flask purified hProCA32.collagen (new vector) with an without reducing agent, C. ELISA plot of large scale purified hProCA32.collagen (old vector) with an without reducing agent, D. Overlay of all the three protein showing the highest binding affinity for hProCA32.collagen purified from new vector with and without B reducing agent (blue). hProCA32.collagen purified from old vector by shake flask had a dissociation constant of 2µM before addition of βME but on addition of the reducing agent, the binding constant reduced to 115µM while the old vector hProCA32.collagen purified from large scale fermentation process had no significant collagen 1 binding, the binding points did not reach saturation due to incomplete binding. E. Color change of all the three protein products characterized.

As shown in Figure 5-15, the new vector based purified protein showed the highest reaction with collagen 1 binding dissociation constant of 0.91 ± 0.1µM and absorbance (color change) while pet22b based vector (old vector) purified protein purified both in GSU shake flask and BFF had a lower binding reaction. Furthermore, β-mercapto ethanol (βME) had a greater effect on protein purified from Pet22b based vector compared to the pD-861-SR vector. The binding affinity for the new vector protein was 0.91± 0.1µM. There was no significant shift in the binding profile of the new vector purified hProCA32.collagen. In contrast, the old vector purified protein showed a significant loss of binding (2 ± 0.12 µM reduced to 115 ± 10 µM in GSU purified protein and BFF purified protein did not reach saturation indicating an
incomplete binding and a need for more points, the calculated $k_d$ was estimated at $120.4 \pm 10$). These changes can both be seen in the binding profile (Figure 5-15 A-D) as well as the colorimetric change (Figure 5-15E).

There are several possible causes for the differences in collagen binding affinity for the two vectors. One possible cause is the difference in purification temperature that can impact stability. The FPLC purification step for the new vector at BFF was performed at room temperature instead of in a cold box at 4°C (old vector) which may compromise the stability and aggregation status of the hProCA32.collagen purified. Second, the old vector of Pet22b has at least 5x lower copy number compared to the PD-86SR vector which has a higher replication system and expresses much higher desired protein with a considerably higher yield. Such higher expression level leads to higher protein concentration that is likely to have higher oligomerization. Third, the largest cause for the variation in collagen binding may be due to protein concentration estimated using absorbance. The high DNA content at 260 nm may obscure the protein concentration calculated using the 280 nm absorbance peak.

5.6 Optimization of Purification with fermentation facility (BFF)

To fully understand protein purification using fermentation, next we worked with the bio expression and fermentation facility in University of Georgia (BFF), to perform step-by-step analysis of each purification step using designed new vector expression system at the BFF. For this experiment, kanamycin antibody and Rhamnose was used with a start temperature of 37°C which was subsequently reduced to 24°C after induction in 20 L fermenter using LB medium. A 155g/20L cell mass was harvested. The expression yield quantified by imageJ
normalized to 100% at harvest yielded 25% soluble fraction (50% soluble to 50% insoluble) at 4 h (green arrow) compared to 63% soluble fraction at harvest. This corresponds to a yield \( \sim 98 \text{g} \) of soluble protein in cell pellet (4.9g of soluble protein/L).

*Figure 5-16: SDS-PAGE characterization of new vector hProCA32.collagen. The expression obtained at harvest (red arrow) for soluble portion was at >2x higher than obtained at the 4 h time point (green arrow).*

Cell pellets were mixed with buffer by blending for dissolution, homogenization, and subsequently lysed using a cell disruptor. The pictorial representation of the purification steps is shown in Figure 5-17. During the lysis step, the temperature of the cell disruptor increased up to 67°C despite the in-built chiller in the cell disruptor which may have exposed the protein to aggregation/oxidation tendencies. As shown in lane 1 Figure 5-18A, the cell lysate contained a significant amount of higher oligomer bands.
Figure 5-17: Physical changes of hProCA32.collagen during Purification at BFF, UGA at room temperature.
A. Direct heating of the lyzed cells results in high temperature and possible denaturation of hProCA32.collagen. B. Buffer exchange and pH change of hProCA32.collagen with tangential flow filtration process. C. Precipitation of EGTA-loaded hProCA32.collagen compared to 3x diluted sample and EGTA-free hProCA32.collagen. D. pH adjustment of hProCA32.collagen after EGTA addition compared to EGTA-free aliquot. E. EGTA-free (left) and EGTA-loaded (right) hProCA32.collagen after filtration by 0.45µm filter. F. Precipitation observed in EGTA-loaded hProCA32.collagen at room temperature, G. FPLC process at room temperature.
Next, the lysate was centrifuged at 10000 rpm for 30 mins at 4°C and samples from both pellet and supernatant were collected for characterization. At BFF, cell supernatant was directly placed on heater without proper controlling of temperature using water bath at 80°C (Figure 5-17A). The thermometer readings went as high as 100°C. There is a significant increase in
dimer band formation Figure 5-18A (blue arrow), likely due to overheating caused denaturation, misfolding, and other complications. According to the protocol, the supernatant is expected to be heat up at 80°C but because the cell lysate was directly placed on heat (Figure 5-17A), the thermometer readings went as high as 100°C which could have led to denaturation, fragmentation, misfolding, formation of molten globule and other complications. Upon analysis of the supernatant post heating step, there is a significant increase in dimer band formation Figure 5-18A (blue arrow). The resulting sample with precipitation was centrifuged at 10000rpm for 30 mins at 4°C and SDS samples collected. To the resulting 800mL supernatant, 24 g (3% final concentration) of streptomycin was added and stirred at room temperature for 2 hours to remove DNA. The resulting solution was clarified by centrifugation at 10000rpm for 30 mins, to which 1μg/mL of DNase 1 was added and allowed to react overnight to further digest DNA. Tangential flow filtration (TFF) was employed to dialyze via concentration then dilution, buffer-exchange and concentrate the solution to a 10 X lower volume for FPLC purification at room temperature.

From Figure 5-18G, a major observation from the FPLC showed a reduced shoulder (blue arrow) in the calcium free EGTA loaded sample, the shoulder can be assumed to be the calcium-loaded form (holo form) of the protein which has been reduced with the addition of EGTA. The non-reducing SDS-PAGE results in Figure 5-18A-F clearly shows that the protein exists in more than the monomer form in different gradient steps of the FPLC run for both EGTA loaded and EGTA free samples. This indicates that several factors at different steps need to be optimized such as the heating temperature, heat duration and pH. During this stage, the sample became greenish cloudy as soon as the buffer was exchanged from pH 7 to pH 8 (Figure 5-17B).
Upon dilution, a color change was observed reducing the greenish color to greenish yellow, however, the cloudiness prevailed (smaller bottle in Figure 5-17D). The solution was divided into two parts of the calcium loaded form pre FPLC and apo-form with addition of EGTA. The direct addition of EGTA to the solution caused a considerable amount of precipitate with significant altered pH (pH of solution reduced to 3.58). After pH was adjusted back to 8, the solution became clear. Both samples were stored at 4°C during the FPLC system equilibration with three column volumes of buffer A (10mM HEPES, pH 8) at room temperature. Due to the precipitation observed for the EGTA-loaded sample, the sample was clarified and filtered to be applied unto the Q-sepharose column. The supernatant obtained was completely clear (Figure 5-17E). Next the column was washed until the UV returned to baseline to remove unbound protein or junk fractions followed by step gradient elution; 5%, 20%, 40% and 60% respectively. As shown in Figure 5-17G, upon injection of hProCA32.collagen at room temperature, high back pressure was observed for the EGTA loaded sample during the loading step.

To test the effect of temperature, concentration for the possible high back pressure in column purification, we prepared three samples (Figure 5-17C top row) using this batch of protein stored at 4°C in cold room. The sample was stored and placed at room temperature and became turbid after 5 minutes at room temperature. The original EGTA-loaded stock without dilution, its 3x dilution, and the EGTA-free stock were left at room temperature where it was observed that the diluted sample was clear with no precipitate even when left overnight at room temperature. In contrast, the undiluted EGTA-loaded sample formed precipitate at room temperature (Figure 5-17C, Figure 5-17F). Based on the observation, we hypothesize that apo-form of the protein in the presence of EGTA could be contributing to the
precipitation/aggregation due to less folded protein allowing oligomers to form more readily at room temperature. To minimize oligomerization, the EGTA loaded samples were diluted 3x before applying to the column.

5.7 Optimization of hProCA32.collagen purification

As discussed previously, factors observed at the BFF in UGA necessitated a close look at some of the purification conditions to improve monomer formation while reducing dimer ratio. In addition, the ELISA result shown in Figure 5-15, the two protein products purified in our lab performed better during collagen binding before addition of βME. To this end, a 4L purification was done comparing our observation with what was observed at each step of the purification process during at BFF.

➢ **Sonication Step:** 7mls representing a 43% increase of HEPES-NaOH pH 7.2 was added per gram of pellet. This was then vortexed rapidly for dissolution and 1μL per falcon tube (~35mls solution) of benzonase is added then sonicated on ice for 2 minutes each, 10 cycles and 10 minutes break after each cycle. The lysate is thereafter centrifuged at 17000rpm for 30 minutes at 4°C. During this batch of experiment, the solution temperature was monitored and just as observed in UGA, the temperature rose to about 55°C before cooling back, this heat-cool cycle could contribute to denaturing and incorrect folder opening the protein up for aggregation possibilities. To optimize this step for future purification wet ice will be used during subsequent sonication cycles to better cool the solution.
Boiling Step: During this step, a beaker half-filled with water was placed on the hot plate, this was brought to boiling and the hot plate temperature was reduced. Next, loosely capped falcon tubes containing supernatant from sonication step was placed into the beaker with a thermometer placed in to monitor the temperature (Figure 5-19A). The beaker with the protein supernatant was allowed to stay on the hot plate for 20 minutes (with the falcon tube placed as described). The temperature was monitored here as well and it was observed that the solution temperature rose to as high as 90°C.

Figure 5-19: Purification of hProCA32.collagen with close observation of physical changes. A., Clarification of sample into supernatant (SN) and pellet after heating, B. The effect of EGTA addition was observed, no precipitation was observed for samples without EGTA (1), sample with 40mM EGTA (2) and sample with 10mM EGTA (3). C. Buffer exchange and dialysis shows significant amount of precipitation. D., SDS-PAGE characterization of each step of the purification process.
Upon centrifugation, 2g of pellet was obtained per 35ml of supernatant (5g of pellet) and a clear solution was obtained for the supernatant.

- **Addition of Streptomycin:** 3% final concentration of streptomycin sulfate was added to the resulting supernatant for precipitation of nucleic acid and left to react overnight at 4°C. The solution is clarified. No major observation or change was observed for this step. A decrease in the amount of streptomycin to 1% may however be sufficient for precipitation since the excess streptomycin interferes with the SDS-PAGE running. As shown in Figure 5-19D, lane 10 and 11 corresponding to streptomycin precipitation step did not run properly due to interference from streptomycin.

- **Dialysis and buffer exchange:** To get the solution pH to pH 8, used for the FPLC step, the supernatant was dialyzed in 50 volumes of 10mM HEPES-NaOH pH 8. As shown in Figure 5-19C, the solution became cloudy after dialysis similar to what was observed at BFF (Figure 5-17B) which had been previously attributed to salts and protein fragments less than 3kDa (the pore size of the dialysis bag used). This further indicated that the protein is not stable at pH 8. The solution was clarified via centrifugation, the supernatant and pellet were collected (Figure 5-19C) and analyzed and as seen non-reduced SDS-PAGE in Figure 5-19D.

- **Calcium Removal by EGTA:** 40mM final concentration from a 200mM stock was added to the clarified sample from the dialysis step described above. As soon as EGTA was added, precipitation occurred (number 2 Figure 5-19B) even on dilution of EGTA concentration to 10mM (number 3 Figure 5-19B) compared to the sample with no EGTA addition (number 1, Figure 5-19B). Both 40mM and 10mM EGTA samples has the same effect on the sample showing a significant amount of the protein in the pellet although most junk observed up till
the dialysis step was removed into the pellet thereby improving the purity of the protein with the major band appearing as a monomer band.

Figure 5-20: FPLC Chromatogram of EGTA loaded hProCA32.collagen with silver stained SDS-PAGE characterization
Showing presence of hProCA32.collagen in the FPLC fractions as well as the concentrated fractions (blue arrow)

Both EGTA loaded and EGTA free samples were injected into the FPLC system for complete separation of the purified hProCA32.collagen from DNA and other proteins with similar characteristic as our desired protein during purification (Figure 5-20). Each of the peaks were analyzed on a UV Spectrophotometer as shown in Figure 5-21 A-D. Next the EGTA loaded sample was concentrated to a higher concentration to observe the extent of aggregation possible at higher concentrations. Figure 5-19C and F shows a significantly monomer band (83.8 %) for the purified protein. The calculated protein yield for the concentrated purified protein is ~285 mg/L
Figure 5-21: Characterization of Purified hProCA32.collagen by UV spectroscopy and SDS-PAGE.
A. Summary of UV/Vis spectra of the different peaks observed in the FPLC chromatogram of purified hProCA32.collagen, B., UV/Vis spectra of samples in the flow through (peak 1) during the column binding phase of both EGTA-loaded, diluted EGTA sample and EGTA-free sample. C-D. Spectra of EGTA-loaded and EGTA-free fractions E., Coomassie brilliant blue and F., silver stain SDS-PAGE of concentrated EGTA-loaded sample. Marker band molecular weights are the same as listed in C.

To determine if the observation from the SDS gel in Figure 5-21 is a true representation of the condition of the protein, or could have been altered by the action of SDS, a chaotropic environment using 8M urea was created wherein the protein was unfolded and refolded as shown in the flow chart in Figure 5-22.
Briefly put, 24ml of 1mg/ml hProCA32.collagen1 was prepared in 8M urea and 1% BME. Next, this solution was added dropwise into 2L 20mM Tris pH 8.5 with stirring. The solution was incubated for a total of 144 hours although protein concentration to achieve refolding started after 48 hours incubation. To analyze the native state of the refolded sample, a pre-calibrated (size exclusion chromatogram) SEC analytic column (Figure 5-23) was used. The resulting chromatogram (Figure 5-24) showed a predominant peak at molecular weight of 64kDa strongly suggesting aggregation or oligomer formation. This could be attributed to the long hours used for protein concentration where prolonged stirring and incubation was done. Also the protein became cloudy after concentration. The solution was centrifuged and analyzed, however, the SDS gel shown in Figure 5-24 lane 7 to 10 does not give much information due to inappropriate silver staining.
Figure 5.23: Calibration Chromatogram of 30ml sephadex size exclusion column with the standardized molecular weight

Figure 5.24: SEC analysis of refolded hProCA32.collagen1 post urea-BME treatment
The clarified sample obtained after refolding was denatured a second time with direct addition of 8M urea and once again analyzed using size exclusion chromatography. In Figure 5-25, the sample, incubated for 15 minutes with 8M urea was injected for analysis and similar to what was previously obtained, predominantly oligomer peaks were observed. This indicates that the aggregates already formed before the denaturing-refolding experiment and may not be reversible by urea addition.

![SEC analysis of denatured hProCA32.collagen1 post second urea treatment.](image)

Based all the aforementioned observations, it is important to pinpoint exactly what stage of the purification contributes significantly to aggregate formation.
5.7.1 Probing the cause of aggregation in hProCA32.collagen

Figure 5-26: Expression of hProCA32.collagen.
A. Bacterial growth curve of hProCA32.collagen shows log phase expression indicating a successful expression process. B. SDS-PAGE characterization of hProCA32.collagen before and after induction of hProA32.collagen showing no significant expression at the 14.5kDa before induction (green arrow) compared to after gene expression further confirming the successful expression of hProCA32.collagen (red arrow).

Starting with the expression step, it can be observed in Figure 5-26, a smeared band spanning different molecular weights can be observed before induction. The intensity of the bands for
the higher than 14.5kDa molecular weights bands had no noticeable difference after rhamnose induction despite that the protein was successfully expressed at the 14.5kDa molecular weight. This indicated that the higher oligomer bands were already present before the hProCA32.collagen gene was expressed. With this in mind, a few conditions were tried to remove and prevent further formation of aggregates.

Figure 5-27: Flowchart Representation of the new conditions and protocol tried for purification optimization for hProCA32.collagen using metal ions, reducing agent and arginine.

Factors that were considered for the conditions includes the ionic strength of the lysis buffer, the buffer environment. In order to slightly alter the protein-protein interaction [336, 337], the solution pH was increased slightly from 7.2 to 7.4. The various conditions tried shown in Figure 5-27 are summarized below;

**Condition one:** The first condition was with a trial of a higher buffer concentration to enhance the buffering capability, 1mM calcium added to help with protein folding since the scaffold of
hProCA32.collagen is the calcium buffering protein parvalbumin with two EF-hand calcium binding motif and 10mM magnesium added to effectively assist benzonuclease in cleaving nucleic acids [338]. In addition to Ca$^{2+}$ and Mg$^{2+}$ metal ions, 5mM βME was added to keep the protein reduced throughout most of the purification process.

**Condition two:** This condition is the reference (control) condition that is exactly what has been previously used during the purification of hProCA32.collagen in the absence of metal ions and reducing agent.

**Condition three:** Represents condition two, the control condition with addition of metals for folding and nuclei acid removal.

**Condition four:** Replicates condition one but with a lower buffer concentration on 10mM i.e the old protocol with metals and reducing agents.

**Condition five:** Represents the previous condition with metals and addition of arginine which helps to prevent protein aggregation.

From the breakdown of conditions highlighted above, it can be observed that apart from the first condition which explores a higher buffer concentration, the condition seeks to improve on the previous condition already established. It should be noted a pH adjustment from the previously used pH 7.2 to pH 7.4 first to monitor the effect of pH on protein folding and prevention of aggregation as well as to minimize pH shock during dialysis and FPLC. Therefore, the pH adjustment step before the FPLC will no longer be needed since the pH adopted for purification is sufficient for binding during FPLC.
Figure 5-28: Different Precipitation level and supernatant appearance of the samples post heating.

Figure 5-28 represents various precipitation outlook of the sample after 80°C heating of the supernatant from the lysis step. 15ml of buffer per gram of pellet was added for each condition with addition of 1μL benzonuclease, this was vortexed to mix after which the cells were lysed by 10x sonication with 10 mins break after each cycle. The resulting lysate was sonicated at 17000rpm for 30 minutes at 4°C. The supernatants were then subjected to heat at 80°C for 5 minutes, cooled to 4°C and subsequently centrifuged at 8000rpm for 15 minutes. As shown in Figure 5-28 (top), heating is expected to induce precipitation of thermally less stable proteins and it can be seen that most of the conditions had some form of precipitation except the sample loaded with 0.5M arginine and the control sample. One plausible explanation for condition five

<table>
<thead>
<tr>
<th>Weight of pellet</th>
<th>Condition 1: 1.78 g</th>
<th>2: <strong>0.6 g</strong></th>
<th>3: 1.96 g</th>
<th>4: 1.97 g</th>
<th>5: <strong>0.36 g</strong></th>
</tr>
</thead>
</table>

*Considerably low pellet weight
could stem from the ability of arginine to keep protein in solution with no selectivity or specificity, preventing precipitation of thermally less stable protein which in our case is undesired. 1.5% v/v final concentration of streptomycin sulfate from a stock concentration of 30% w/v was added and allowed to incubate overnight at 4°C.

![Figure 5-29: Observable sample solution and Pellet appearance post supernatant analyzed by SDS-PAGE and pellet and supernatant estimation.](image)

From Figure 5-29, it can be observed that condition five to which arginine was added still did not generate precipitate as shown in others and sample two, the control condition has the most precipitate (2.16g) which could a combination of the expected precipitate from boiling step and streptomycin step. Conditions 2, 3 and 4 had relatively the same weight of pellet generated.

The samples ought to be dialyzed in the buffer in which each was lysed at 4°C with three buffer changes then left overnight at same temperature but since there would be no need for buffer exchange, that step was skipped. The samples were centrifuged and analyzed. From the SDS-
PAGE in Figure 5-29, condition one seem to have pure protein formed at the boiling stage while condition 3 and 4 also have almost pure protein that the same boiling stage.

![Image of FPLC Chromatogram and SDS-PAGE](image)

**Figure 5-30: FPLC Chromatogram Representation condition one with higher buffer concentration**

However, for conditions two and five it seems to have so much junk that wasn’t rid of even after boiling. It also supports the theory that arginine kept all proteins both junk and desired in solution thereby resulting in a smeared SDS-PAGE. On injection of the samples, majority of the sample was observed in the flow through (peak one), Figure 5-30 and Figure 5-31 were selected as representation of the five conditions, only condition one has a considerably high peak although not as high as the unbound peak. This could be due to several factors such as; presence of the calcium and magnesium metals competing on the column, pH or streptomycin
and other salts. One way to solve this problem would be to include the dialysis step in the purification process.

Figure 5-31: FPLC Chromatogram Representation of conditions 2-5 listed above with 3x lower bound peak.

Condition one peak 3 was analyzed by size exclusion chromatography described previously and a broad peak with molecular weights ranging from 56kDa to 7.5kDa was observed representing different form of the protein (Figure 5-32). Due to the low binding observed, a new set of conditions were tried.
Figure 5-32: Size Exclusion Chromatogram of the peak three fractions. From condition one, the UV absorbance was considerably low which could be due to loss of protein. The calibration standard (red chromatogram) showed majority of the peaks in the higher molecular weight region.

For these new sets of conditions shown in Figure 5-33, several considerations were put in place to optimize the previous conditions shown in Figure 5-27.
Figure 5-33: A 6-condition flow chart using 20mM with several conditions.

All conditions in this set of experiment has the same buffer concentration of 20mM HEPES pH 7.4, the pellet dilution ratio was reduced to a 1:10 pellet: buffer ratio. Also, the heating temperature was reduced to 65°C to prevent denaturation that may lead to fragmentation and improper folding. Each of the conditions were also handled differently at different stages with different compositions

**Condition one:** The first condition was with 1mM calcium added to help with protein folding since the scaffold of hProCA32.collagen is the calcium buffering protein parvalbumin with two EF-hand calcium binding motif and 10mM magnesium added to effective assist benzonuclease in cleaving nucleic acids. Also, the buffer used to lyse the cells was prepared with 0.2M arginine which is 2.5x lower than the previous concentration to enhance a milder condition for ensuring protein solubility. After cell lysis via sonication, the lysate was centrifuge, the resulting supernatant heated at a 65°C and centrifuged at 8000rpm for 15mins. 2% final concentration of streptomycin was added and allowed to incubate overnight, after
which dialysis in 20mM HEPES, pH 8 carried out with three buffer changes before being filtered and injected to the FPLC system.

**Condition two:** This condition is the control condition that is close to what has been done during purification in the lab employing no metals or reducing agent during the purification process except for the higher concentration of HEPES used. The proceeding process is same as aforementioned for condition one.

**Condition three:** Represents condition two, the control condition with addition of metals for folding and nuclei acid removal. In addition, 0.2M arginine was added after the sonication step, incubated at 4°C overnight. Next day, the solution was centrifuged at 8000rpm for 15 mins at 4°C after which it is heated at 65°C for 5 mins and purified in the steps highlighted above post boiling.

**Condition four:** Replicates condition three at lysate step but follows the steps for condition two up to the streptomycin step at which point 0.2M arginine was added, incubated at 4°C overnight before being dialyzed then injected into the Q-sepharose column.

**Condition five:** Was the previously designated condition one and since it had a relatively better binding than the other conditions, it is a positive control which will include dialysis step to observe for an improved binding.

**Condition six:** Patterned after the control protocol with addition of 5mM BME after the boiling step and allowed to react at 4°C overnight. Then reaction with 2% streptomycin, dialysis, and injection into the Q-sepharose column.
Figure 5-34: hProCA32.collagen purification process. After the heat up step, all the protein yielded precipitate in contrast to previous observation.

The first observation different from the past experiment was the precipitation observed for arginine containing sample (condition one) up to ~2g which corresponded to the average weight of pellet observed in the first set of experiments with samples without arginine. Condition three which also contains arginine (red box) followed similar trend (Figure 5-34).
Figure 5-35: Addition of streptomycin to C1, C2, C4 and C5 solutions, C3 and C6 during arginine and BME incubation, respectively.

From Figure 5-35 and Figure 5-36, streptomycin precipitation expected to lead to the precipitation of DNA [339] resulted in no precipitation for the arginine-containing sample (condition one) while conditions 2, 4 and 5 with no arginine addition had DNA precipitation which was separated by centrifugation. For both sample C4 and C6, addition of either BME or arginine post boiling did not generate precipitate either.
Figure 5-36: Addition of streptomycin to C1, C2, C4 and C5 solutions by next before and after centrifugation.

Figure 5-37: Analysis of pellet obtained for different conditions after dialysis step.

As explained previously, dialysis at pH 8 had been suspected to lead to protein precipitation (Figure 5-19), therefore by reducing the pH to 7.4, a significantly lower precipitation was
observed compared to previous data Figure 5-37.

**Figure 5-38:** Agarose gel DNA analysis of the presence or absence of DNA at each step of the purification.

Next, we estimated the rate of DNA removal from our purified hProCA32.collagen since arginine addition prevented precipitation upon streptomycin addition. From Figure 5-38, a close look at the agarose gel for all samples containing arginine (C1, C3 after boiling and C5 after arginine addition; the green bar) all showed no peak. Arginine is a positively charged amino acid, therefore, to confirm if it was not interfering in the DNA running, agarose gel for the pellet samples were also analyzed (Figure 5-39) which showed the presence of DNA in the pellet.
Figure 5-39: Agarose gel DNA analysis of pellets and supernatant from the purification steps to track DNA shows a reduced DNA band after addition of streptomycin in conditions 1, 2 and 5.

Figure 5-40: FPLC chromatogram of the newly analyzed conditions showing a lower peak for the flow through compared to other peaks.
The FPLC chromatogram in Figure 5-40 showed a considerably reduced peak height for the flow through compared to the peaks in the binding region. Also the last peak, suspected to be the DNA peak is lowered. The fractions were analyzed as shown in Figure 5-41 and Figure 5-42. As can be observed in Figure 5-41 is the presence of predominantly monomer band at the shoulder peaks just after the flow through. The native gel result is inconclusive.

Figure 5-41: SDS-PAGE and native gel analysis of purified hProCA32.collagenduring purification and post Q-sepharose injection. A., SDS-PAGE of conditions 1, 2 and 5 supernatant after dialysis as well as FPLC fractions B-D., Native gel of each purification step for all conditions.
Figure 5-42: SDS-PAGE Analysis of pellet and supernatant from the various purification steps. 
A. SDS-PAGE analysis of supernatant obtained for conditions 1-6 after cell lysis by sonication. 
B. SDS-PAGE analysis of pellet obtained for conditions 1-6 after cell lysis by sonication and 
pellet for conditions 1 to 3. C., Analysis of different fractions and conditions after FPLC 
purification. D-F. Pellet, supernatant and FPLC fraction analysis by SDS-PAGE.
Figure 5-43: Process flow chart for purification optimization with DEAE column

In Figure 5-42A, the gel shows an m-shaped protein band for all the samples after streptomycin treatment which was not be removed by dialysis until the binding phase of FPLC where it was eluted as Q column flow through. This could also suggest that addition of streptomycin could be a major contributor to low binding to the ion exchange column during FPLC. To this end, a falcon tube replicating the steps preceding streptomycin addition was set as the control for a second experimental set up employing the use of Diethylaminoethyl (DEAE), a positively charged resin used in ion-exchange chromatography, for the separation and purification of proteins and nucleic acids (Figure 5-43). The DEAE beads were pre activated at room temperature in the binding buffer for 2 days below been washed severally and mixed with the supernatant obtained from the boiling step and allowed to rotate overnight at 4°C. Both the flow through from the DEAE column and the supernatant from streptomycin addition were dialyzed in 20mM HEPES pH 7.4 for 36 hours. The results obtained from SDS-PAGE and agarose gel reveals a pure protein obtained albeit with double band at the boiling step which became single.
but with about 5% dimer band for streptomycin compared to single band at the desired 14.5kDa band with multiple bands above 29kDa. On the agarose gel, both the DEAE and the streptomycin were able to reduce the DNA composition of the protein, but the streptomycin was more effective for DNA removal. A plausible explanation for the lower efficiency of the DEAE beads could be due to reduced activity stemming from broken seal and prolonged storage.

Figure 5-44: SDS-PAGE and agarose gel for DEAE column purified hProCA32.collagen

The result obtained (Figure 5-44) show that dialysis for longer hours could rid the protein of the streptomycin attachment that mostly inhibits binding to the FPLC column and DEAE is a potential column that can be employed to effectively remove DNA. Due to the pI of the protein, using the DEAE column was used extensively to probe the possibility of separation of impurities during the FPLC stage but the binding of hProCA32.collagen to the column was significantly low with most of the desired protein washing out during the binding phase (experiment done by another lab member). To this end, polyethleneimine (PEI) was explored
for DNA precipitation. Several conditions were explored as shown in Figure 5-45 and Figure 5-46 as described by Burgess [340].

Figure 5-45: Different DNA precipitation methods and conditions using PEI.

Three sets of reactions were carried out as described in the flow diagram (Figure 5-46), as can be observed in Figure 5-47, the DNA agarose showed the lowest DNA content for the samples treated with 0.3% DNA in the absence of NaCl.
Figure 5-46: Scheme for DNA Precipitation by Polyethyleneimine at low and high ionic strengths

Figure 5-47: DNA Agarose gel for different DNA precipitation conditions explored using both low ionic and high ionic strengths. The absence of NaCl aided the most DNA precipitation with the least DNA showing up as DNA band (red box and arrow).
5.8 Probing the conformational requirement for collagen binding of the disulfide bond in the collagen binding peptide of hProCA32.collagen

The resulting purified protein shown in Figure 5-48 generated a purified protein with about 95% monomer portion (SDS-PAGE not shown). A control protein purified with the protocol described in section 2.5 generated purified protein with 40% dimer portion (Figure 5-48F) was also estimated with ELISA. As observed previously, the addition of BME compromised the binding affinity of hProCA32.collagen (Fig 3.3A-B).

![Image](image_url)

**Figure 5-48:** Collagen type 1-hProCA32.collagen dissociation curve of suspected monomer purified protein vs dimer containing protein. A. Purified protein with 40% dimer portion, B. Purified protein with 40% dimer portion plus 5% BME and C. Spectra overlay of both reduced and non-reduced protein with 40% dimer D. color photographs of Purified protein with 40% dimer portion, BME reduced Purified protein with 40% dimer portion and Purified protein with 95% monomer E,F. SDS-PAGE of the purified proteins; lane 1: Protein marker, lane 2: reanalyzed 95% monomer protein with aggregate but no dimer, lane 3: reduced Purified protein with 40% dimer portion, lane 4: non reduced Purified protein with 40% dimer portion.
From the spectra overlay in Figure 5-48, the protein shown to possess 95% monomer portion had an almost flat plot (red solid circles) with corresponding lack of color development post ELISA experiment Figure 5-48E. In contrast, both BME treated and untreated protein with 40% dimer showed a positive binding profile, reduced on addition of BME (Figure 5-48A-B). The color change observed as well is more visible in these two proteins (Figure 5-48D). Post ELISA experiment, all samples (monomer, dimer containing and reduced dimer-containing protein) were reanalyzed by SDS-PAGE. The dimer containing protein retained its conformation while the majorly monomer protein formed a considerably high aggregate. The lack of reaction during ELISA as well as aggregate formation could have resulted from unspecific interaction of GSSG/GSH with the protein post purification. The binding pocket for collagen type I may have been compromised resulting to lack of binding. Another possibility could be incorrect protein concentration estimation. For the ELISA experiment, the protein concentration was measured using the Lambert Beer’s UV spectrometric method. To confirm the blocking of GSSG/GSH and the monomer binding affinity of hProCA32.collagen, the protein purification was repeated and FPLC fractions distributed over peaks 2 to 8 was observed same as previous (Figure 5-49A). The ELISA had absolutely no color change and therefore could not be plotted (Figure 5-49C-D).
Figure 5.49: ELISA of monomer purified protein with and without BME.
A. FPLC Chromatogram of hProCA32.collagen distributed over 10 peaks. B. SDS-PAGE analysis of the purification stages of hProCA32.collagen and each FPLC peak. C-D. ELISA microplates after completion of ELISA experiment, no color change was observed.

In order to rule out the possibility of incorrect concentration estimation, the protein concentration was SDS-PAGE image j analysis and the concentrations was significantly lower than observed for UV measurement (Figure 5.50)
Figure 5-50: ELISA assay of different peaks obtained from FPLC characterization under reducing and oxidizing conditions.

A-B., Peak 4 with and without redox buffer shows incomplete binding, hence low collagen affinity. C-D., Peak 8 with and without redox buffer shows a relatively improved binding profile for the sample without buffer while the sample with redox buffer had a similar binding profile as that obtained for peak 4. E. Random sampling of fractions with and without dimer composition. F. Mass spectrometry analysis of peak 2 showing presence of dimer and monomer bands after sample storage (data provided by Dongjun Li)

In figure 3.6, hProCA32.collagen was purified using GSSG/GSH. The difference between this purification and the previous method is the introduction of dialysis and slight concentration
steps to get rid of the unbound GSSG and GSH after which the ELISA experiment was carried out with and addition of redox buffer (GSSG: GSH 10:1). Of the peaks, peak 4 (mostly monomer) and peak 8 (with about 30% dimer) were selected for analysis with and without redox buffer (Figure 5-51A-D). In addition to this, a control experiment using raw samples from different peaks (peak 2, 3, 5 and 6 with more monomer as well as peak 9 containing dimer) was done simultaneously (Figure 5-51E) and plotted linearly. The binding profile for the monomer protein with and without redox buffer showed significantly less binding compared to the dimer containing fraction (peak 8) without redox buffer. Addition of redox buffer however resulted on loss of binding on peak 8 fraction. This could confirm that redox buffer interferes with the binding of hProCA32.collagen to collagen type I. Peaks 2, 3, 4, 5 and 6 didn’t show a significant color change/absorbance values compared to peak 8 and 9 with dimer fractions.
Figure 5-51: Mass spectrometry and concentration Estimation of hProCA32.collagen by SDS-PAGE quantification showed no protein. Marker band molecular weights are 68, 26, 14.3 and 6.5 kDa from top to bottom in A, B, and C.
Figure 5-52: ELISA estimation of hProCA32.collagen purified by different methods with their corresponding SDS-PAGE.
A. New vector purified unPEGylated hProCA32.collagen by shake flask method (protein purified and by Zongxiang Gui) B., Lysine PEGylated hProCA32.collagen purified by BFF and stored for 21 weeks. C. BFF purified unPEGylated hProCA32.collagen (new vector)

Figure 5-52 shows our most recent attempt at estimating the binding affinity of our contrast agent to collagen. hProCA32.collagen was freshly purified, while a previously purified protein in UGA BFF facility stored at -20°C for about 8 months and 5 months respectively were simultaneously analyzed. All protein contained at least 40% dimer portion as shown in the SDS-PAGE. The $K_d$ for the freshly prepared protein was consistent and within the 1µM range while PEGylation and oxidation seemed to improve the binding of UGA old vector protein (compared to that reported in Figure 5-15). The new vector purified protein from UGA had a better binding than reported for unPEGylated old vector purified in UGA. This could be due to vector or oxidation experienced over the months of storage.

Further characterizations and specifically monomer (with no tendency for aggregation) and dimer (100%) need to be done to confirm most of the assertions in this report.
5.8.1 Formulation of hProCA32.collagen by controlling oligomerization

The first step in the formulation is an overall optimization of the expression condition for hProCA32.collagen with the use of an expression vector capable of production efficiency, stability and high induction response[333]. To obtain hProCA32.collagen in the monomeric it is important to push to keep the two cysteine residues in the collagen binding moiety in conjugation with each other since we have previously established in Figure 5-15 that addition of reducing agent reduces the collagen binding affinity of the protein. In addition, for clinical use of hProCA32.collagen, it is important to obtain the purified protein in the monomeric form without other forms. We therefore added cysteine to selectively prevent oligomerization. The workflow shown in Figure 5-53, using a range of cysteine concentration, we tried to decide on the right amount of cysteine required for oligomer prevention. Cysteine concentration as low as 0.05mM and as 0.1 mM was used. All pellets were lyzed in a 1:10 pellet: buffer ratio of 20mM HEPES buffer, pH 7.4 with 1mM Ca$^{2+}$ and 10mM Mg$^{2+}$. Conditions 1 and 2 had 0.05mM cysteine added while 3 and 4 had 0.1mM cysteine. In addition to cysteine, conditions 2 and 4 had 1% BME added. The use of cysteine as an oxidizing agent has been well established but, in this experiment, it was used to provide excess thiol group and possibly an oxidizing agent with BME as a reducing agent. As shown in Figure 5-54A, the presence of DNA resulted in the presence of higher oligomer in the hProCA32.collagen (ROI). In Figure 5-54B, addition of polyethylenimine (PEI) for DNA precipitation also helped oligomer reduction. Another interesting observation is the presence of dimer and tetramer bands (blue and green arrows) for conditions with no BME addition.
Furthermore in conditions 1 and 3, oligomerization was completely eliminated with 0.1mM cysteine compared to 0.05mM (Figure 5-54B, lane 8, red arrow). This was further confirmed in condition 4 where only dimer and trimer bands were observed with no oligomer bands. This observation indicates that 0.1mM cysteine is sufficient for oligomer prevention while BME can remove multimer formation.
Figure 5-54: SDS-PAGE of different steps in the cysteine-BME experiment for prevention of oligomerization of hProCA32.collegen.
A. Supernatant (SN) of samples 1 to 4 (S1-S4) from lysis step to heating step up to post dialysis step (B). (Standard marker (M) bands are same as in A)

However, upon dialysis, some of the purified protein especially for condition 2 and 4 went back in the oligomer state. To improve this, we increased the cysteine concentration to 2.5 mM and 5mM.

Figure 5-55: Experimental set up for use of higher concentration (2.5mM and 5mM) of cysteine during cell lysis with corresponding SDS-PAGE for each steps.
As shown in Figure 5-55, no major difference was observed between 2.5mM and 5mM cysteine concentration with no oligomer reduction all through the purification steps unlike what was observed with 0.1 mM cysteine. A major observation however is the reduction of oligomer after dialysis which is a sharp contrast from the use of 0.1mM cysteine concentration previously used.

**Figure 5-56: Cysteine and BME addition to hProCA32.collagen for prevention of oligomerization.**

BME is a very compound and can evaporate easily from the solution which may explain the reason for the oligomerization observed after dialysis since the disulfide bond formation reaction is a dynamic process hence the need for a more stable reducing agent such as DTT.
which is less volatile and more stable in solution.

Figure 5-57: SDS-PAGE of different purification steps with DTT and BME as reducing agent

As shown in Figure 5-57, conditions 4 and 6 with 1% BME and 5mM DTT respectively had the best purity after PEI addition. Both conditions also had the lowest concentration of oligomer, 0% of dimer was observed after PEI addition for the condition with 1% BME and 0.1 mM cysteine (Table 5-2) although oligomers were observed after dialysis for all the conditions, similar to what was previously observed in Figure 5-54.

Table 5-2: Effect of Reducing Agent on dimer band at different purification steps by under different conditions using BME and DTT

<table>
<thead>
<tr>
<th>Sample condition</th>
<th>Reducing agent</th>
<th>Percentage of dimer (%)</th>
<th>Percentage of dimer (%)</th>
<th>Percentage of dimer (%)</th>
<th>Percentage of dimer (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Cell lysis</td>
<td>Heating</td>
<td>DNA removal</td>
<td>Dialysis</td>
</tr>
<tr>
<td>1</td>
<td>-</td>
<td>21</td>
<td>21.50</td>
<td>15</td>
<td>25</td>
</tr>
<tr>
<td>2</td>
<td>0.1% BME</td>
<td>15.8</td>
<td>18.1</td>
<td>15.1</td>
<td>23.5</td>
</tr>
<tr>
<td>3</td>
<td>0.5% BME</td>
<td>10.05</td>
<td>9.14</td>
<td>2.01</td>
<td>24.2</td>
</tr>
</tbody>
</table>
Since the oligomers returned each time after dialysis. The role of pH was elucidated to determine the optimum pH for best result.

**Figure 5-58: Estimation of the role of pH in cysteine/BME reducing capability**
As shown in Figure 5-58, condition 1 was the control condition with no cysteine or BME while others had either 0.1mM cysteine or 2.5 mM cysteine with or without BME at different pH strength. All the condition used all previous lysis conditions with heating at 55°C.

![Figure 5-59: SDS-PAGE characterization of hProCA32.collagen with pH optimization protocol. Mark band molecular weights are 68, 26, 14.3, and 6.5 kDa if unlabeled.](image)

Figure 5-59 shows a significant reduction in higher oligomer bands occurring with 0.1mM cysteine +1% BME at both pH 7 and 7.4 (orange arrow) post heating with a more profound reduction at pH 7 (red arrow) which also proved to be a better pH for 2.5mM cysteine (yellow arrow) compared to the result obtained at pH 7.4 (green arrow). After dialysis however, the effect was eliminated giving multimer bands like that obtained post lysis).
Figure 5-60: Agarose gel analysis for DNA removal of hProCA32.collagen

In addition, PEI addition removed DNA significantly as shown in Figure 5-60 which stayed removed after dialysis indicated that DNA contamination is not responsible for the oligomer formation. Therefore pH 7.1 condition with 0.1mM cysteine was further optimized. This therefore necessitated a look into the dialysis condition which has proven to be the step where oligomers are observed mostly.
Figure 5-61: Optimization of the dialysis condition of hProCA32.collagen purification

Figure 5-61 shows our effort in optimizing the dialysis condition for hProCA32.collagen purification to prevent oligomer formation after dialysis. For this experiment, 0.1mM cysteine, 1% BME and 5mM DTT which had shown positive result for oligomer removal as shown above. Furthermore, Gd$^{3+}$, BME, cysteine and DTT were used in the dialysis buffer. The combination are diagrammatical represented in Figure 5-61. All of these were carried out at pH 7.
Figure 5-62: SDS-PAGE of the Optimization of the dialysis condition for reduction/prevention of oligomer formation.

As can be observed in Figure 5-62A, using pH 7, the protein sample with 1% BME had the purest band on SDS-PAGE after heating (lane 2 and 3, red box) while the samples with DTT even though purer relative to control condition 1, they had multimeric components present. This observation persisted over different steps until dialysis step where both dimer and tetramer bands were observed but not oligomers.
Figure 5-63: Dynamic light scattering studies of hProCA32.collagen.
A., Sample previously purified as monomer and allowed to aggregate by storing in 4°C for 30 days showed a peak with diameter ~7000nm corresponding to suspected aggregate peak and an additional peak of ~240 nm. B., Control sample with 1mM Ca^{2+} and 10mM Mg^{2+} in 20mM HEPES lysis buffer showed a peak with diameter ~11nm corresponding to suspected monomer peak and an additional peak of ~100 nm. C., Sample with 1% BME, 0.1mM cysteine, 1mM Ca^{2+} and 10mM Mg^{2+} in 20mM HEPES lysis buffer plus BME and cysteine in dialysis buffer showed monomer peak at ~10nm and additional peak at ~300nm. D., Sample with 1% BME, 0.1mM cysteine, 1mM Ca^{2+} and 10mM Mg^{2+} in 20mM HEPES lysis buffer plus Gd^{3+}, BME and cysteine in dialysis buffer showed mostly 142nm peak. E., Sample with 5mM DTT, 0.1mM cysteine, 1mM Ca^{2+} and 10mM Mg^{2+} in 20mM HEPES lysis buffer plus Gd^{3+}, BME and cysteine in dialysis buffer showed predominant peaks at ~145nm. F. Sample with 5mM DTT, 0.1mM cysteine, 1mM Ca^{2+} and 10mM Mg^{2+} in 20mM HEPES lysis buffer plus BME and cysteine in dialysis buffer showed predominant peaks at ~148nm. F. Sample with 5mM DTT, 0.1mM cysteine, 1mM Ca^{2+} and 10mM Mg^{2+} in 20mM HEPES lysis buffer plus DTT and cysteine in dialysis buffer showed predominant peaks at ~150nm. G. Sample with 5mM DTT, 0.1mM cysteine, 1mM Ca^{2+} and 10mM Mg^{2+} in 20mM HEPES lysis buffer plus Gd^{3+}, BME and cysteine in dialysis buffer showed predominant peaks at ~145nm.H. Overlay of the DLS spectrum

In Figure 5-63, each of the conditions were characterized using dynamic light scattering, which is a non-invasive technique (sample is exactly as is after measurement) for measuring the size
of particles and molecules in suspension [341, 342]. Based on a previous standardization experiment (result not shown), hProCA32.collagen with a molecular weight of 14.5 kDa is expected to have a particle size at about 3.8 nm and a polydispersity index (PDI) ≤ 1.0. Most of the samples had PDI values within the recommended limit and are below 0.7. In Figure 5-63A, a positive control sample which was stored at 4°C for 2 months to promote aggregates formation showed a PDI of 0.36 with two distinct peaks at ~240 nm and ~7000 nm respectively. This is over 60x higher than the expected 3.8 nm particle size suspected to be due to the presence of high oligomers as shown in its SDS-PAGE characterization (arrow). In Figure 5-63B, the bacteria cells were lysed with only 1 mM Ca\(^{2+}\) and 10 mM Mg\(^{2+}\) added to the 20 mM HEPES buffer. We observed that the particle size corresponding to ~7000 nm was absent while the dominant peak was a dispersed peak around 150 nm. This can be attributed to the presence of monomer, dimer and trimer bands with no significant amount of oligomer as shown its corresponding SDS-PAGE. This is also reflected in the 0.56 PDI value obtained for this sample indicative of a more dispersed sample population compared to A. Condition 2 in Figure 5-63C had the highest PDI (1.00) and two prominent peaks at 10.1 nm and 341 nm. This sample also had a relatively pure SDS-PAGE with mostly monomer population and no observable oligomers. This DLS characterization reflects the presence of smaller population relative to others. As can be observed in Figure 5-63 C-G, addition of either DTT, Gd\(^{3+}\) or BME in the dialysis buffer resulted in similar peaks with the more predominant peak at ~150 nm corresponding to a mixture of monomer, dimer, and trimer forms of hProCA32.collagen, similar characteristics was observed in SDS-PAGE. The protein yield obtained for this batch of purification was ~320 mg/L.
5.8.2 **Relaxivity and collagen binding studies of purified protein**

The relaxivity of purified hProCA32.collagen was estimated as well as its collagen binding affinity to ensure that the protein retained functionality after addition of cysteine and BME in the purification protocol. A higher relaxation rate will translate to higher sensitivity and lower dose requirement. As shown in Figure 5-64, hProCA32.collagen has r1 and r2 relaxivity values of 30 mmol\(^{-1}\)s\(^{-1}\) and 52 mmol\(^{-1}\)s\(^{-1}\) respectively. This value is consistent with that reported for the targeted rat analogues of ProCAs such as ProCA32.collagen1 and ProCA32.CXCR4 [159, 160, 187]. In comparison to clinically approved contrast agent, hProCA32.collagen has at least 8-fold higher relaxivity.

![Figure 5-64: Relaxativity (r1 and r2) measurement of hProCA32.collagen.](image)

A. r1 and r2 relaxivity plot of hProCA32.collagen at clinical field strength (1.5 T) in comparison to clinically approved gadolinium-based MRI contrast agents B. r1 and r2 relaxivity plot of hProCA32.collagen at 7T.

Furthermore, we quantified the collagen binding affinity of hProCA32.collagen. Collagen solution purified from rat tail was diluted 10 x in 0.1 M Sodium Carbonate-Bicarbonate, pH 9.6. This was incubated overnight at 4°C. The next day, the plate was washed with 350 μL 1X TBST 3 times at 5 minutes each at room temperature while shaking at 120 rpm, collagen was blocked with 350 μL/well blocking buffer (5% BSA dissolved in TBST) for 2 hours at room
temperature with shaking at 120 rpm. The plate was washed 3 x after which serial dilutions of hProCA32.collagen was prepared from 0.05 to 40 µM in 1X TBS was added to the plate. The protein was incubated 4°C overnight. The next day, the plate was washed 3x as previously described, 100µL of primary antibody (dilution ratio of 1:100) was added to the wells and allowed to incubate for 2 h at room temperature. The wells were washed 3 x with 350 µL/well 1X TBST. Dilution of secondary antibody (1:1000) was done with 1X TBS and added at 100 µL/well to the plate, this was incubated for 45 minutes. The plate was washed with 350 µL/well 1X TBST 3 times for 5 minutes each at room temperature after which 100 μL/well TMB and incubated for 10 minutes at room temperature exactly for all wells. The reaction was stopped with 100 µL/well 2N Sulfuric Acid and the absorbance value was read. As shown in Figure 5-65, hProCA32 shows no binding to collagen type I compared to hProCA32.collagen with a dissociation constant K_d of 0.18 µM. To confirm the translational potential and specificity of hProCA32.collagen to type I collagen compared to different types of collagen, we measure the collagen binding affinity of hProCA32.collagen to types I, III and IV. As shown in Figure 5-66, hProCA32.collagen has a dissociation constant of 0.9 µM which is 7 and 13 x higher than that of collagen types III and IV.
Figure 5-65: Collagen binding dissociation curve of hProCA32.collagen versus hProCA32 shows a saturated binding curve compared to the hProCA32 curve which showed no-binding.
Figure 5-66: Collagen binding specificity of hProCA32.collagen.  
A. collagen binding of hProCA32.collagen to human collagen type I, B., collagen binding of hProCA32.collagen to human collagen type III, C. collagen binding of hProCA32.collagen to human collagen type IV, D., overlay of the binding affinity plots of hProCA32.collagen to type I, II and IV collagen.

5.8.3  **N-terminal PEGylation of hProCA32.collagen**

The amine group at the N-terminus of proteins is one of the accessible amine group that can be utilized in PEGylation. Lysine is the most abundant amino acid residue in proteins [343]. By taking advantage of the fact that primary amine residues in protein have different pKa (7.8) for the N-terminal α-amino group and 10.1 for the amino group in the ε position of lysine
residues [290, 344]. As shown in Figure 5-67, the reaction was not successful as no observable PEGylation was seen.

![Figure 5-67: N-terminal PEGylation of hProCA32.collagen.](image)

A., SDS-PAGE with Coomassie brilliant blue staining with about 90% of the protein still in the unPEGylated state B., iodine staining showing inconsistent trend with no obvious PEGylation observed.

This reaction was optimized by adjusting the pH and reducing conditions. The modified procedure was carried out as described in section 2.5.4.
Figure 5-68: SDS-PAGE analysis of the mPEG-hProCA32.collagen
SDS-PAGE were stained with iodine, Coomassie Blue and silver staining. MonoPEGylated hProCA32.collagen was observed at the lanes corresponding to 1:1 binding ratio with a 65% PEGylation yield while 35% of the protein remained unPEGylated. By contrast, the reactions with both 5 and 10 x PEG reagent both showed multiPEGylated bands indicated saturation of reagent and possible PEGylation of amine groups in lysine.

In Figure 5-68, a PEGylation efficiency of 65%, 86% and ~94% were observed for 1:1, 1:5 and 1:10 Protein: PEG ratio respectively. While a higher PEGylation reaction completion was observed with high PEG concentration, specificity was observed for 1:1 binding ration. Addition of a higher concentration of PEG is likely to have led to the reaction of primary amines in hProCA32.collagen and can be improved by shortening the reaction time.
5.8.4 **Tissue biodistribution and toxicity Test**

Tissue biodistribution and toxicity test was carried out for hProCA32.collagen as described in section **Error! Reference source not found.**. After injection of either PBS (control) or hProCA32.collagen, the physical morphology and the weight of the mice tissues were analyzed. As shown in Figure 5-69, there was significant difference between the morphology and tissue weight of PBS injected mice versus hProCA32.collagen injected mice. The weight of the PBS injected mice’ brain, spleen, lung, kidney and liver tissues were 0.39 ± 0.03, 0.135 ± 0.06, 0.24 ± 0.04, 0.495 ± 0.06, and 1.69 ± 0.3 respectively while the values obtained for hProCA32.collagen cohorts were 0.395 ± 0.02, 0.126 ± 0.05, 0.31 ± 0.04, 0.54 ± 0.06, and 1.365 ± 0.16 respectively.

![Figure 5-69: Tissue toxicity test of hProCA32.collagen.](image)

A., Gross tissue specimen of healthy CD-1 mice injected with either PBS or hProCA32.collagen. There was no significant difference between the appearance of the tissue specimen for PBS injected cohort compared to hProCA32.collagen injected mice. B., the brain, spleen, lung, kidney and liver of both groups of mice were evaluated. No significant difference was observed between the weights of brain, spleen, lung, kidney and liver of mice injected with PBS or hProCA32.collagen. (P values for brain, spleen, lung, kidney and liver are 0.8952, 0.3721, 0.3415, 0.6615 and 0.4474 respectively for N=4 per group)
Figure 5-70 shows that hProCA32.collagen exhibited no tissue toxicity in all kinds of mice tissues and organs studied, including liver, kidney, muscle, spleen, brain, and heart.

Figure 5-70: Tissue toxicity test of hProCA32.collagen.
0.025 mmol/kg ProCA32-P40 or 100µL PBS (control) was i.v. injected in CD-1 mice. H&E staining was used to evaluate the toxicity of hProCA32.collagen to these tissues after injection of either hProCA32.collagen or PBS control. An experienced pathologist was blinded to the groups of H&E staining to evaluate the organ toxicity of hProCA32.collagen. hProCA32.collagen shows no toxicity in all kinds of mice tissues and organs studied, including liver, kidney, muscle, spleen, brain, and heart. (N = 4 per group).
hProCA32.collagen demonstrated predominantly hepatobiliary biodistribution at 1 day and a combination of renal, hepatobiliary, and splenic elimination in CD-1 mice at 14 days.

Biodistribution studies (Figure 5-71) showed low tissue gadolinium levels at 1 day and 14 days after hProCA32.collagen injection in bone and brain. Bone uptake have been associated with “free” unchelated Gd\(^{3+}\) [85].

Figure 5-71: Long-term distribution of Gd\(^{3+}\) in mice after injection of hProCA32.collagen for 1 day (N = 4) and 14 days (N= 4). hProCA32.collagen shows no Gd\(^{3+}\) either bone or brain.

Furthermore, the acute toxicity, measured with liver enzymes (ALT, ALP), urea nitrogen, bilirubin, and total protein from CD-1 mice 48 hours post tail veil injection of hProCA32.collagen (0.025 mmol/kg) and PBS were analyzed. The bloods of the mice were collected 14 days post injection. The blood from CD1 mice with injection of PBS were used as control. Based on the results from IDEXX from 3 mice in each group, all the tests including creatine, ALT, ALP, cholesterol, triglycerol had similar values for between the control group and the group injected with hProCA32.collagen. In addition, the values were within the normal range reported for these tests. This indicates that the function of the organs including kidney
and liver have not been altered by the injection of hProCA32.collagen. (Figure 5-72). All together, these toxicity tests confirmed that hProCA32.collagen does not cause chronic tissue toxicity.

Figure 5-72: Acute toxicity test of hProCA32.collagen.
The mice clinical pathology profiles were obtained by analyzing mice serum samples collected 14 days after injection of PBS (n = 3), or 0.025 mmol/kg hProCA32.collagen (n = 3).

5.8.5 Tricks and tips for Loading Gadolinium to unPEGylated hProCA32.collagen

A major advantage of protein PEGylation is its ability to enhance protein stability and solubility. Therefore, addition of Gd$^{3+}$ to hProCA32.collagen showed no precipitation of
turbidity. The rate of complex formation between Gd\(^{3+}\) and hProCA32.collagen depends on factors such as reaction temperature, pH, concentration of hProCA32.collagen and binding environment. For the binding environment, a fifth carboxylate group (S56D) was introduced into the CD EF-hand binding motif of \(\alpha\)-Parvalbumin to enhance the Gd\(^{3+}\) binding affinity [102, 164, 345] introducing an extra “hard acid” for enhanced Gd\(^{3+}\) binding. The reaction pH can be controlled by adjusting the Gd\(^{3+}\) pH to 7-8 before addition to hProCA32.collagen. The concentration of hProCA32.collagen typically used for MR imaging is 5mM but with the non-PEGylated hProCA32.collagen, addition of Gd\(^{3+}\) to protein solution \(\geq\) 3.5mM resulted in precipitation. To improve cost and reduce the circulation half-life of hProCA32.collagen, it may be important to formulate it in the non-PEGylated form without compromising kinetic and thermodynamic stability. As shown in Figure 5-73, addition of pH adjusted Gd\(^{3+}\) solution in a dropwise fashion resulted in a clear solution compared to adding the Gd\(^{3+}\) solution at once.
Figure 5-73: Optimized $\text{Gd}^{3+}$ loading method with dropwise addition shown in tube A with clear solution compared to bolus addition in tube B.
Table 5-3: Comparison of Longitudinal and transverse relaxation rates of some commercially available gadolinium-based contrast agents with hProCA32.collagen and ProCA32.collagen at 37°C.

<table>
<thead>
<tr>
<th>Contrast Agent</th>
<th>$^{a}r_1$ at 1.4 T</th>
<th>$^{a}r_1$ at 7.0 T</th>
<th>$^{b}r_2$ at 1.4 T</th>
<th>$^{b}r_2$ at 7.0 T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Magnevist®</td>
<td>3.3</td>
<td>5.1</td>
<td>3.9</td>
<td>9.4 ± 1.3</td>
</tr>
<tr>
<td>Eovist®</td>
<td>5.38 ± 0.02</td>
<td>5.37</td>
<td>6.54 ± 0.06</td>
<td>7.01</td>
</tr>
<tr>
<td>Dotarem®</td>
<td>3.9 ± 0.2</td>
<td>N/A</td>
<td>3.2 ± 0.7</td>
<td>N/A</td>
</tr>
<tr>
<td>MultiHance®</td>
<td>6.20</td>
<td>N/A</td>
<td>8.7</td>
<td>N/A</td>
</tr>
<tr>
<td>ProHance®</td>
<td>4.39</td>
<td>N/A</td>
<td>5.0</td>
<td>N/A</td>
</tr>
<tr>
<td>ProCA32-P40</td>
<td>33.14 ± 0.32</td>
<td>18.9</td>
<td>44.61 ± 0.12</td>
<td>48.6 ± 0.1</td>
</tr>
<tr>
<td>ProCA32.collagen</td>
<td>34 ± 0.12</td>
<td>21.3 ± 0.5</td>
<td>50 ± 0.16</td>
<td>108.5 ± 1.2</td>
</tr>
<tr>
<td>hProCA32.collagen</td>
<td>34.5 ± 0.5</td>
<td>18.5 ± 1.2</td>
<td>50.6 ± 0.61</td>
<td>105.6 ± 2.0</td>
</tr>
</tbody>
</table>

*Longitudinal relaxation rates measured at 37°C.
*Transverse relaxation rates measured at 37°C.
*N/A: not available.

Table 5-4: Comparison of stability constants of some commercially available gadolinium-based contrast agents with hProCA32.collagen and ProCA32.collagen at 25°C.

<table>
<thead>
<tr>
<th>Contrast Agent</th>
<th>Log ($K_{11}$)$^{a}$</th>
<th>Injection Dosage (mmol/kg)</th>
<th>Log ($K_{12}$)$^{a}$</th>
<th>Log ($K_{22}$)$^{a}$</th>
<th>Log ($K_{23}$)$^{a}$</th>
<th>Log ($K_{31}$)$^{b}$</th>
<th>Log ($K_{32}$)$^{b}$</th>
<th>Log ($K_{33}$)$^{b}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Magnevist®</td>
<td>22</td>
<td>0.1</td>
<td>22.46</td>
<td>10.75</td>
<td>18.6</td>
<td>12.24</td>
<td>4.13</td>
<td></td>
</tr>
<tr>
<td>Eovist®</td>
<td>N/A$^{d}$</td>
<td>0.025</td>
<td>23.6</td>
<td>11.82</td>
<td>18.78</td>
<td>12.22</td>
<td>5.18</td>
<td></td>
</tr>
<tr>
<td>Dotarem®</td>
<td>N/A$^{d}$</td>
<td>0.1</td>
<td>24.7</td>
<td>17.23</td>
<td>21.05</td>
<td>7.46</td>
<td>3.65</td>
<td></td>
</tr>
<tr>
<td>MultiHance®</td>
<td>N/A$^{d}$</td>
<td>0.05</td>
<td>21.91</td>
<td>N/A</td>
<td>17.04</td>
<td>N/A</td>
<td>4.87</td>
<td></td>
</tr>
<tr>
<td>ProHance®</td>
<td>N/A$^{d}$</td>
<td>0.1-0.3</td>
<td>23.8</td>
<td>14.83</td>
<td>19.37</td>
<td>10.07</td>
<td>4.37</td>
<td></td>
</tr>
<tr>
<td>ProCA32-P40</td>
<td>21.08$^{c}$</td>
<td>0.0013</td>
<td>22.44$^{c}$</td>
<td>9.55$^{c}$</td>
<td>8.77$^{c}$</td>
<td>13.1</td>
<td>14.3</td>
<td></td>
</tr>
<tr>
<td>ProCA32.collagen1</td>
<td>22.53$^{c}$</td>
<td>0.0013</td>
<td>22.54$^{c}$</td>
<td>8.71$^{c}$</td>
<td>6.36$^{c}$</td>
<td>14.7</td>
<td>16.2</td>
<td></td>
</tr>
<tr>
<td>hProCA32.collagen1</td>
<td>22.03$^{c}$</td>
<td>0.0013</td>
<td>22.20$^{c}$</td>
<td>8.11$^{c}$</td>
<td>6.8$^{c}$</td>
<td>14.1</td>
<td>15.4</td>
<td></td>
</tr>
</tbody>
</table>

*Stability constants measured by potentiometric method.
*Stability constant values estimated by subtraction of stability constants of contrast agents.
*Stability constant values measured by the fluorescence resonance energy transfer between tryptophan and metal in protein contrast agents.
*N/A: not available.

The use of the human scaffold did not have a negative effect on the bio characteristics of hProCA32.collagen compared to ProCA32.collagen.
5.9 Major conclusions, Challenges, and future work

hProCA32.collagen purified in Pet22b vector had a low expression yield at harvest compared to 4 h post induction time point. hProCA32.collagen was subsequently sub cloned into the new vector by Dr. Liangwei Li. A major focus of this chapter was to obtain the purified hProCA32.collagen in its monomeric state due to the observed presence of dimerized hProCA32.collagen as well as higher oligomer bands. Using a combination of cysteine and 1% βME at pH 7.1 to 7.4, a significant reduction of dimer and higher oligomer bands were observed. The new vector possessed a higher copy number and a higher protein yield ~350 ± 53 mg/L of protein about 10 times higher than hProCA32.collagen in Pet22b vector. N-terminal PEGylation of hProCA32.collagen in a 1:1 ratio resulted in 65% PEGylation yield and needs to be further optimized. The improved hProCA32.collagen exhibits 8- to 10-fold improvement in both $r_1$ and $r_2$ compared to all the current clinically used contrast agents at both 1.4 T ($r_1 = 30 ± 0.6 \text{ mM}^{-1}\text{s}^{-1}$ per Gd$^{3+}$ (60 ± 1.2 mM$^{-1}$s$^{-1}$ per participle) and $r_2 52 ± 0.9$ mM$^{-1}$s$^{-1}$ per Gd$^{3+}$ (104 ± 1.8 mM$^{-1}$s$^{-1}$ per particle) and 7.0 T ($r_1 = 18.5 ± 1.2 \text{ mM}^{-1}\text{s}^{-1}$ per Gd$^{3+}$ (37.0 ± 2.4 mM$^{-1}$s$^{-1}$ per participle) and $r_2 105.6 ± 2.0 \text{ mM}^{-1}\text{s}^{-1}$ per Gd$^{3+}$ (211.2 ± 4.0 mM$^{-1}$s$^{-1}$ per particle) at 37°C. hProCA32.collagen exhibits $10^8$ –$10^{15}$-fold greater metal selectivity for Gd$^{3+}$ over physiological metals such as Zn$^{2+}$ and Ca$^{2+}$ compared to Dotarem, a clinically approved contrast agent. Clinical toxicity test of hProCA32.collagen showed values in the normal range of each test evaluated. Furthermore, injection of hProCA32.collagen exhibited no noticeable tissue toxicity.

hProCA32.collagen displayed a high specificity for collagen type 1 over types 3 and 4. This is exciting in two phases, on one hand, collagen type one being the most abundant protein can be easily detected with high specificity making it a potential biomarker for disease monitoring in
collagen vascular diseases. On the other hand, collagen type 3 and 4 can be detected in diseases with overexpression of these types of collagens such as autosomal dominant polycystic kidney disease (ADPKD) [346]. hProCA32.collagen like our previously reported rProCA32.collagen enabled early detection of NAFLD. Using a diabetes/obesity model, (discussed in detail in chapter 7) hProCA32.collagen detected NASH stage F1 as well as late-stage cirrhosis with SNR values directly indicative of collagen accumulation as estimated by collagen proportionate area (CPA) analysis. Most importantly, the use of a human protein in place of rat analogue used for the development of rProCA32.collagen [102, 108, 155, 157, 159, 160, 187, 210] significantly reduces the immunogenic potential of our developed contrast agent while maintaining sensitivity and specificity. As a future direction, it will be beneficial to explore different excipients that will keep the protein in its formulation form throughout its shelf life. Other conditions that can further be optimized are pH and storage temperature.

6 APPLICATION OF HPROCA32.COLLAGEN IN EARLY THE DIAGNOSIS OF CHRONIC LUNG DISEASES

6.1 Introduction

Chronic lung diseases (CLD) are progressive lung disorders that typically affect the lungs and other aspects of the respiratory system, some of which includes Asthma, pulmonary pneumonia, chronic obstructive pulmonary disease, virus infections, interstitial lung diseases (lung fibrosis) etc. (Figure 6-1). Active and passive cigarette smoking is a major risk factor of CLDs while other factors include air pollution, age, sex, exposure to radiation etc. Despite
advances in investigation of the possible etiologies of these diseases, the pathology and mechanism are yet to be fully understood.[14]

Figure 6-1 Examples of chronic lung diseases affecting humans.

Development of chronic lung fibrosis is a hallmark of interstitial lung diseases (ILD) which is a broad term used for a group of over 200 diffuse parenchymal lung disorders characterized by progressive shortness of breath, cough, restrictive physiology, and impaired gas exchange[347]. There are several risk factors associated with ILDs, some are caused by known factors like drugs, radiation, autoimmune conditions, exposure to environmental toxins such as asbestos, and they can as well be granulomatous or rare non categorized diseases, such as chronic eosinophilic pneumonia. The last category of ILDs is caused by unknown factors, these
sets of diseases are referred to as non-specific interstitial pneumonia and idiopathic interstitial pneumonia Figure 6-2.

6.1.1  **Idiopathic Pulmonary Fibrosis (IPF)**

Pulmonary fibrosis is a specific form of chronic interstitial lung disease of unknown etiology more characteristically associated with older adults’ age > 50 years with usual interstitial pneumonia (UIP) pattern. It can develop because of occupational factors or diseases associated with specific exposure to agents known to damage the lungs. These include medications such as bleomycin, occupational exposures to particles such as asbestos, tobacco smoke, or other agents in the environment that can cause an immune reaction (hypersensitivity pneumonitis), it can also arise from genetic insufficiency [348], or after trauma and acute lung injury leading to fibroproliferative acute respiratory distress syndrome. When lung fibrosis develops from unknown etiology, it is referred to as Idiopathic Pulmonary fibrosis (IPF).

![Diagram of Pulmonary Fibrosis Causes](image)

**Figure 6-2 Causes of Pulmonary fibrosis**
IPF has been reported to occur throughout the world in many different racial and ethnic groups. Studies in the United States have suggested that Caucasians are more likely to be diagnosed with IPF and have higher mortality rates from IPF than African Americans. Genetics play a role in the development of the familial cases of IPF. About 8 percent of familial cases can be attributed to a single set of genes [349-351]. IPF typically develops in a gradual, subtle way, with a gradual onset of dyspnea and dry cough. The symptoms often progress to become severe and lethal [189]. IPF is classified as a rare disease despite having a similar frequency to cancer of the brain, stomach, and testicular cancer [352]. The median survival remains unacceptably poor, at approximately 3-5 years [353]. The recently approved therapies, nintedanib and pirfenidone, have provided new hope for patients diagnosed with IPF with their proven ability to improve progression-free survival by slowing down the decline of pulmonary function [354-356]. With the introduction and proven positive effects of these therapies, there exists a need for precise and early diagnosis of IPF with clear distinguishing pattern from other ILDs or chronic pulmonary disorders. Currently, there are no effective therapies capable of targeting these underlying fibrotic changes in the lung. Additionally, there are no diagnostic tools to detect early-stage fibrosis and quantify lung fibrosis. Progress for selecting patients for lung transplantation and other treatments has been constrained by the lack of non-invasive diagnostic methods and methods for monitoring disease progression. There is no effective treatment for IPF, and lung transplantation is the only treatment that had been proven to improve prognosis [357]. The process characterized by alveolar epithelial cell injury and hyperplasia, inflammatory cell accumulation, fibroblast hyperplasia, deposition of extracellular matrix, likely resulting in chronic inflammatory infiltrates, myofibroblast hyperplasia, and disordered collagen deposition and scar formation which typically ends in
loss of lung elasticity and highly dense alveoli thereby impairing the process of O₂ – CO₂ exchange a critical function of the lungs[358]. The alveoli (air sac) responsible for gaseous exchange consist of an alveoli fibrotic membrane which becomes impermeable in fibrotic conditions (Figure 6-3). The extent of pulmonary Injury and individual susceptibility to pulmonary damage is dependent on age and genetic predisposition, with subsequent innate immune system and fibroblast activation[359]

![Figure 6-3 Mechanism of gaseous exchange at the alveoli of normal lung compared to fibrotic lungs](image)

6.1.2 **Animal model for IPF**

The role of animal models in the understanding and development of potential treatment therapies for diseases cannot be overemphasized. Chronic diseases are not easy to model especially for a disease like IPF with unclear etiology and natural history; no one trigger is known to be capable of inducing IPF in animal [360].
There are no natural models for IPF induction[361]. To this effect, several studies have applied different techniques to developing mice models of experimentally induced pulmonary fibrosis some of which are elucidated below. In fluorescein isothiocyanate mice model, a single intra tracheal administration of Fluorescein isothiocyanate (FITC) leads to a reproducible lung injury, evidenced by increased in lung collagen content with evidence of specific immunity to the fluorescein hapten [359]. The use of a fluorescein enables visualization of the areas of the lung where deposition occurs via immunofluorescence imaging for the characteristic green color of the FITC since fibrosis has been observed to the primarily present in areas of FITC accumulation. The fibrotic response is developed both in balb-c as well as C57BL/6 mice[362]. Disadvantages of this animal model include absence of representative UIP findings and predominant inflammatory infiltrates that precede fibrosis. A second model involves the use of silica for pulmonary fibrosis development. Intratracheal administration of silica fibers in mice results in the development of fibrotic lesions closely resembling those seen in individuals with certain occupations, including those in the textile industry, mining and clay manufacturing [363]. The strain of mice utilized is a contributing factor to the extent and mechanism of fibrotic response with C57BL/6 mice displaying a higher degree of susceptibility. Just as described for the FITC model, fibrosis occurs primarily in areas of silica deposits[364]. The proposed fibrotic pathogenesis is associated with limited inflammation and enhanced fibrotic lesions mediated by increased pro-fibrotic growth factors and cytokines. Several routes of administration exist for this model with intratracheal route being the easiest and cheapest delivery route resulting in fibrosis between 2 to 4 weeks. A major disadvantage of this model is its lack of reproducibility, high variability in fibrotic pattern amongst several cohorts, absence of representative histopathological UIP findings lesions such as fibroblastic foci,
regional heterogeneity, and hyperplastic epithelium is another major shortcoming of this animal model [364-366]. Another model is the radiation-induced mice models, radiation injury may occur in any organ within the radiation field such as the lungs triggering inflammation and ultimately activation of myofibroblasts from differentiated fibroblasts which in turn leads to overexpression of collagen and other extracellular matrix. Radiation fibrosis syndrome (RFS) in mice model is typically generated by whole thorax radiation exposure (18 Gray; dose rate 0.7 Gray/minute) using a Gamma cell Cesium-137 unit to elicit lung damage. Lung fibrosis in this model is dependent on inflammation and free-radical-mediated DNA damage and less on TGF-β with better success in C57BL/6 mice compared to other strains. The lung injury is usually local if surrounding organs are adequately shielded. However, the fibrosis development is relatively slow, observed after 24 weeks depending on the dose of radiation, fraction size, and treated volume [367-370]

6.1.2.1 The bleomycin mice model

Bleomycin is a chemotherapeutic antibiotic belonging to a family of compounds produced by Streptomyces verticillis. They possess highly effective apoptotic properties against tumor cells and are frequently used in cancer chemotherapy with pulmonary toxicity as a major side effect[371]. It possesses activity against gram-negative bacteria. It has been used in mice, hamsters, rats, guinea pigs, dogs, rabbits, and primates even though mice and rats are the most common specie in which the drug is used[372]. Additionally, several routes of delivery such as intratracheal, intravenous, intraperitoneal, subcutaneous, and inhalational have been described with the intratracheal route being the most utilized due to its low cost and ability. Bleomycin administration results in lung injuries by inducing DNA strand breakage and oxidant injury[373]. In this study, we utilized the intraperitoneal injection route of
bleomycin (BLM) to recapitulate histopathology consistent with human idiopathic pulmonary fibrosis [374]. 25mg of bleomycin was given to each mice twice weekly for 3 weeks to induce early stage IPF, 4.5 weeks for moderate IPF and 6 weeks for advanced stage IPF.

6.1.2.2 Role of Nicotine in fibrosis

Chronic airway inflammation is a cardinal feature of chronic obstructive pulmonary disease (COPD). Nicotine is a highly addictive chemical present in ~0.6–3.0% of the dry weight of tobacco including electronic cigarette, prolonged exposure is associated with the development of chronic lung diseases such as chronic obstructive pulmonary disease [375, 376]. In many organ systems, nicotine inhalation leads to the recruitment of inflammatory cells which modulates fibrosis by altering the functions of fibroblasts promoting tissue remodeling, extracellular matrix deposition such as collagen and alterations in lung structure and function [377-379]. Nicotine acts by binding to nicotinic acetylcholine receptors (nAChRs), which are pentameric acetylcholine-triggered channel proteins that form homomeric and heteromeric α and β chain pentameters. nAChRs are nonselective ligand-gated ion channels that allow several different positively charged ions, such as potassium, sodium, and calcium, to cross the cell membrane. The use of various subunits gives rise to nicotine's differential effects on the activation of signaling mechanisms. The binding of nicotine to the α7-nAChR receptor triggers an influx of calcium across the plasma membrane, thereby activating a downstream signaling mechanism that contributes to dysregulated growth, angiogenesis, release of growth factors, and modulation of the microenvironment [380-382]. Nicotine has been shown to stimulate lung fibroblasts to express collagen type I expression without alteration to the overall lung architecture. The lack of effect on the lung architecture may mask subtle injury until it results in excessive tissue damage after injury [378].
6.1.2.3  Current Limitation in detection of IPF

An accurate diagnosis of IPF relies on exclusion of other known causes of interstitial lung disease (ILDs). Chest X-ray shows the scar tissue typical of pulmonary fibrosis which may be important for disease and treatment monitoring but has the tendencies for inconclusive diagnosis; requiring other tests before diagnosis can be made. Pulmonary Function Test (PFT) such as spirometry, maximal respiratory pressure, arterial blood gases etc. generally give a general overview of how well the lungs carry out their respiratory function hinting of a possible impairment. These measurements no doubt has been extremely useful in identifying lungs with impairment, however, it is not a standard method of diagnosis and does not present restrictions in lung function[383, 384]. The current golden standard for the detection of IPF is the high-resolution computed tomography (HRCT) [357, 385]. Using HRCT for diagnosis of IPF requires the presence of usual interstitial pneumonia (UIP) patterns. High resolution computed tomography (HRCT) scanning has proven to be invaluable in the diagnosis of the disease, and currently represents a gold standard for detection, but this method has drawbacks including exposure to ionizing radiation, limited tissue characterization, and limited ability to assess dynamic contrast enhancement[386]. At least 30% of patients do not have clear radiologic features of fibrosis[387], such as honeycombing (the most important feature essential to IPF diagnosis), basilar predominance of reticulation in the sub pleural region and in majority of the cases, traction bronchiectasis [388-390], requiring the use of additional surgical lung biopsy for further diagnosis and confirmation[391]. Further, interpretation of HRCT data can vary significantly during diagnosis, since some patients may present features consistent with basilar predominance of reticulation in the sub pleural region without honeycombing in place of usual interstitial pneumonia (UIP) patterns[384, 387, 392]. Additionally, surgical lung
biopsy in patients with suspected IPF often have high sampling error and comorbidities [390, 393, 394] leaving this set of patients at high risk for post-biopsy ventilator dependence and in certain cases, air leakage. Diagnosis can be difficult for clinicians as it requires ruling out other contributing factors for the interstitial pneumonia such as occupational hazards, connective tissue diseases, or drug toxicity. Further tracking of disease progression and monitoring clinical trials is problematic, and physiological changes are often only noted with acute exacerbations.

6.1.3 Application and limitation of MRI in detection of IPF

Compared to HRCT, MRI has no associated radiation exposure and can be used for pediatric and pregnant patients, it allows for multiple and repeated measurements and can be used to monitor disease progression and treatment effects. However it has major limitations when applied in lung imaging[395]. The lung being the site for oxygen exchange possesses certain clinical features such as proton density (0.1g/cm³, a value ten times lower than liver tissue proton density) and susceptibility difference in the air tissue interface that are significantly different from most tissues such as liver or brain [105]. This has created a major limitation of MR imaging of the lungs[395]. This implies that the paramagnetic nature of air oxygen versus the diamagnetic tissue oxygen will give rise to bulk magnetic susceptibility difference ($\Delta \chi = 8$ ppm) at the lung–air interfaces which forms a static local field gradient. As more airways and alveoli in the lungs are encountered, multiple microscopic surfaces will yield local magnetic field gradients inhomogeneity smaller than the size of a typical imaging voxel which leads to a fast and continuous dephasing in gradient echo imaging pulse sequence; magnetic field inhomogeneity described by an apparent transverse relaxation time $T2^*$ becomes extremely short making gradient echo MRI of lung parenchyma highly challenging and requires the use of ultra-short echo time MRI[105, 396, 397].
6.1.4 Role of hyperpolarized MRI in IPF

Due to the low proton density and air-tissue interface susceptibility, the introduction of gaseous MRI contrast agents based without the use of proton has become important in MR imaging of the lungs. Most notably is the use of hyperpolarized gas such as $^{129}$Xe and $^3$He are elements in group 18 of the periodic table. These are electrical conductors, fluorescent, stable, noble gases with filled outer shell and nuclear spin $\frac{1}{2}$. By employing dynamic nuclear polarization (DNP) technique, the nuclear spin alignment of the hyperpolarized gases can be increased by ~100,000-fold compared to thermal equilibrium possible with conventional MRI [398, 399]. They image the lung airspaces and airways as opposed to the surrounding tissues imaged by proton MRI and by extension, the lower density of gases decreases the obtainable signal intensity in resulting MR images. To overcome this limitation, hyperpolarization process was introduced [400, 401] to enhance the magnetization of gases by about 5 orders of magnitude. In hyperpolarization, the nuclear spins inside a sample are aligned to enable the imaging of gases without the limitation associated with intrinsic low spin density [285]. In clinical settings, patients are positioned inside the bore and asked to inhale a specified gas such as $^3$He and $^{129}$Xe and the measurement is taken immediately at breath-hold. $^{129}$Xe despite being the only noble gas with anesthetic properties under normobaric conditions was the first noble gas tried in MRI studies [402-404]. At the time, the advancement of $^{129}$Xe hyperpolarized MR imaging was hampered by its low polarization leading to low MRI signal intensity and thereby causing a shift to the use of $^3$He. $^3$He offered a larger gyrometric ratio almost 3x that of $^{129}$Xe, $^3$He also overcame the low signal intensity limitation observed with $^{129}$Xe, does not function as an anesthesia[404]. However, sustainability of $^{129}$Xe MRI has been crippled by its limited availability and exorbitant pricing[405]. This has led to the reevaluation of the safety and
tolerability of multiple inhalation of hyperpolarized $^{129}$Xe for hyperpolarized MRI\[406-408]\ since inhalation of $^{129}$Xe in large enough alveolar concentrations (>70%) \[404]\ possesses anesthetic properties\[285]. Despite the advantages of this technique, multi-center clinical studies and technical validation are major limitations\[409].

In summary, both the IPF and COPD require a major improvement for precise and accurate detection of these diseases. Significant progress has been made at the clinical level in noninvasive imaging of lung fibrosis. However, the techniques and protocols required for lung imaging require either the use of invasive methods or exposure to ionizing radiation. These modalities also have associated detection error and may not be capable for accurate fibrosis detection, staging and treatment monitoring. At the preclinical level, molecular imaging of liver fibrosis via MRI, PET and SPECT is gradually shifting towards more specific detection of changes in biomarker levels that directly correlate with disease progression and/or treatment responses. Lung fibrosis is characterized by activated fibroblasts accumulation and excessive deposition of fibrotic extracellular matrix proteins, especially type I collagen. Therefore, targeting collagen type I with our developed collagen-targeted contrast agent hProCA32.collagen is expected to be clinically relevant in the early noninvasive detection of and staging of lung fibrosis. This possibly reversing lung fibrosis with treatment at an early stage and will have strong application toward major clinical consequences including cirrhosis, impaired liver function, and hepatocellular carcinoma.

6.1.5 **Overview of this chapter**

In this Chapter, we describe our progress in the application of hProCA32.collagen in MR detection of lung fibrosis in bleomycin induced mice model for idiopathic pulmonary fibrosis will be discussed using both conventional T1 and T2 weighted MRI pulse sequences and
optimized 3D T1-W MRI with very short TE (3D UTE). To prove the capability of our contrast agent to specifically map out collagen accumulation regardless of the location, a second animal model utilizing electronic cigarette (e-cigarette) in a nicotine-induced lung fibrosis for chronic obstructive pulmonary diseases transgenic mice model for airway fibrosis was evaluated by hProCA32.collagen since collagen accumulation is expected at the bronchial structures of the lungs in contrast to the lung interstitial space for IPF.

6.2 Results and discussion

To evaluate the application of hProCA32.collagen in detection of collagen accumulation in the interstitial space of the lungs, hProCA32.collagen was used in MR imaging of bleomycin-induced lung fibrosis. For the development of bleomycin in mice, twice 8-week-old male C57BL/6 mice were given intraperitoneal injection of 25mg/kg bleomycin sulfate, twice weekly. To induce different stages of lung fibrosis, bleomycin induction lasted for 3, 4.5 and 6 weeks for early, moderate, and late stage respectively (Figure 6-4). One week after stopping bleomycin treatment for each stage, the mice were taken for MRI analysis. Then the lungs were harvested and weighted out and then saved for the following hydroxyproline assay and histology examinations.
Figure 6-4: Development of bleomycin-induced idiopathic pulmonary fibrosis mice model by Drs. Ganesh and Guangda in Dr. Zhiren Liu’s lab

As shown in Figure 6-4 and Figure 6-5, I.P injection of bleomycin progressively induces the development of IPF in the lung alveoli. The subpleural region of the lungs exhibits cystic honeycombing features with patchy collagen network.

Figure 6-5: Human features of idiopathic pulmonary fibrosis such as cystic honeycombing (black stars), sub pleural abnormality (blue arrow), epithelial hypercellularity (black arrow),...
airway hyperplasia (orange arrow) and fibroblast foci (green arrow) observed in our intraperitoneal (IP) administration of bleomycin by H&E staining.

6.2.1 Intraperitoneal (IP) injection of bleomycin recapitulates the clinical features of IPF

As shown in Figure 6-6, lung fibrosis development and progression was characterized by heterogeneous patchwork appearance with alternating areas of collagen-rich (orange arrow) and normal lung (blue arrow), fibroblastic foci (arrow head) with overall architectural distortion prominent at the subpleural region (black arrow) alveoli in the bleomycin mice model was characterized by collagen area, α-SMA quantitation of histological tissue sections and colorimetric analysis of hydroxyproline content. The hydroxyproline quantification reflected a progressive increase with disease advancement from 53 ± 12 μg/lung for normal mice, 80 ± 20 μg/lung for 3 weeks BM-treated mice to 130 ± 25 μg/lung for 6 weeks BM-treated mice (Figure 6-6C). The α-SMA levels also increased with disease progression from 0.51 ± 0.2% , 4.2 ± 0.1 % for 3 weeks BM-treated mice to 8.51 ± 0.35% (Figure 6-6D). The collagen area increased progressively at different stages from 1.01 ± 0.065% for normal mice, 8.9 ± 0.2 % for 3 weeks BM-treated mice, to 13.22±1.01 % for 6 weeks BM treated mice (Figure 6-6E). The Ashcroft score quantified as described[410] increased from 0 for normal mice, to 2 ± 0.5 for 3 weeks BM-treated mice and 5 ± 1.2 for 6 weeks BM-treated mice
Figure 6-6: Characterization of bleomycin-induced IPF mice model. A. representative images of bleomycin-induced mouse stained with hematoxylin and eosin (H&E), B. Sirius red and C. α smooth muscle actin (α-SMA), D. Ashcroft scoring of normal and bleomycin-induced mice E-G. Quantification of collagen area, hydroxyproline and α-SMA in normal and bleomycin-induced IPF mice models. Blue arrows represent areas with heterogeneous cystic honeycombing, yellow arrow shows hyper cellularity similar to human features. (All scale bars = 100µm)

6.2.2 Molecular imaging of IP injected bleomycin for idiopathic pulmonary fibrosis (IPF) imaging

In order to test our contrast agent capability for in lung imaging by MRI, bleomycin sulfate, a cytotoxic chemotherapy drug was injected via intraperitoneal route into C57/BL6 mice twice
weekly for 3 weeks, 4.5 weeks and 6 weeks respectively inducing early, moderate and late stage IPF respectively. Then, the PEGylated hProCA32.collagen contrast agent (100 µL, 5mM) was injected into the mice via the tail vein. For all the mice, T1-weighted rapid acquisition refocused echo (RARE) and three-dimensional ultra-short echo (3D UTE) MR images of the lungs were collected using either a Bruker or Varian 7.0 T MRI scanner. The first cohort of bleomycin treated mice were selected for confirmation of IPF development and were imaged on a 7-T Agilent MRI scanner at the Biomedical imaging research center (BIRC) of the University of Georgia. Animals were anesthetized by isoflurane inhalation, their respiration rate as well as temperature were monitored with a rodent physiological monitoring system while maintaining Anesthesia within 65 to 70 breaths per minute (BPM). The second cohort of bleomycin-treated mice and nicotine-induced fibrotic mice were imaged on a 7.0 T Bruker scanner at Emory University Yerkes National Primate Research center as described in section 2.9.

6.2.3 Analysis of normal and bleomycin (BM)-induced IPF mice by conventional T1 and T2 sequence.

The first cohort of bleomycin treated mice were selected for confirmation of IPF development and were imaged on a 7-T Agilent MRI scanner at the Biomedical imaging research center (BIRC) of the University of Georgia. Animals were anesthetized by isoflurane inhalation, their respiration rate as well as temperature were monitored with a rodent physiological monitoring system while maintaining Anesthesia within 65 to 70 breathes per minute (BPM).
Figure 6-7: Representative axial T1-W lung images of normal non-diseases mice showing no observable presence of fibrosis or abnormalities in the lung of two mice (M1 and M5).

Both T1 weighted (T1-W) (Figure 6-7) and T2-W (Figure 6-8) images of the normal non-diseased lungs displayed an absence of lesions or abnormality in the lungs of a healthy mice. The calculated signal to noise ratio (SNR) as well showed no observable difference in SNR from pre through 24 h time point.
Figure 6-8 Representative axial T2-W lung images of normal non-diseases mice showing no observable presence of fibrosis or abnormalities in the lung of two mice (M1 and M5).

Figure 6-9 Average signal-to-noise ratio (SNR) of the normal (healthy) mice with RARE acquisition. There was no significant difference in the T1-W and T2-W SNR over the various time points from pre, 1 h, 3 h and 24 h time points.
Figure 6-10: Representative axial T1-W and T2-W lung images.
3 weeks bleomycin-induced mice showed minimal fibrotic structure (orange arrow) with increased signal to noise ratio after injection of hProCA32.collagen. The fibrotic structure was more obvious at the 1 h and 3 h time point compared to the pre time point. Similar patterns were observed in all animals studied injected with bleomycin for 3 weeks, indicating that fibrosis even though present at 3 weeks, it is minimal.

Figure 6-11: Representative axial T1-W and T2-W lung images.
3 weeks bleomycin induced mice showed minimal fibrotic structure (orange arrow) with increased signal to noise ratio after injection of hProCA32.collagen. The fibrotic structure was more obvious at the 1 h and 3 h time point compared to the pre time point. Similar patterns were observed in all animals studied injected with bleomycin for 3 weeks, indicating that fibrosis even though present at 3 weeks, it is minimal.

Figure 6-12: Representative axial T1-W and T2-W lung images of 3 weeks bleomycin-induced mice. The image showed minimal fibrotic structure (orange arrow) with increased signal to noise ratio after injection of hProCA32.collagen. The fibrotic structure was more obvious at the 1 h and 3 h time point compared to the pre time point. Similar patterns were observed in all animals studied injected with bleomycin for 3 weeks, indicating that fibrosis even though present at 3 weeks, it is minimal.

6.2.4 MR Imaging of normal and BM-induced IPF mice by T1-weigted rapid acquisition with refocused echo (RARE) and 3D ultra-short echo time (3D UTE).

The second cohort of bleomycin treated mice were imaged on a 7.0 T Bruker scanner at Emory University Yerkes National Primate Research center. Figure 6-13 summarizes the T1-W RARE images before and after injection of contrast agents. After injection of hProCA32.collagen, both 3 weeks bleomycin-treated (early stage) and 6- weeks treated (late stage) mice had a significant enhancement with fibrotic network at 3 h time point. At the 1 h
time point both the non-targeted hProCA32 and targeted hProCA32.collagen had significant enhancement with the highest observed in late-stage BM-treated mice. Clinical contrast agent, gadovist showed no significance across all the time points. This can be seen confirmed in the SNR values in Figure 6-14 estimated at the 3 h time point. The average percentage SNR increase for convention T1-W images on both the Bruker and Varian MRI scanner was 40% and 120% respectively for early and late-stage IPF mice respectively (Figure 6-14), while the T2-W pulse sequence showed a 25.2 and 61.5% increase in early and late-stage IPF SNR values, respectively 3 hours after injection of hProCA32.collagen.
Figure 6-13: Summary of corona orientation of T1-W representative of RARE mice with injection of different MRI contrast agents.

A. hProCA32.collagen injection of normal mice at pre, 1 h and 3 h post injection time points
B. hProCA32.collagen injection of 3-weeks BM-treated mice at pre, 1 h and 3 h post injection
time points, C., hProCA32.collagen injection of 6-weeks BM-treated mice at pre, 1 h and 3 h post injection time points, D., non-targeted hProCA32 injection of 6-weeks BM-treated mice at pre, 1 h and 3 h post injection time points and E., Injection of commercial contrast agent; Gadovist® to 6-weeks BM-treated mice at pre, 1 h and 3 h time points.

Furthermore, as demonstrated by the histogram plot (Figure 6-15), the fibrosis pattern in normal mice injected with hProCA32.collagen as well as Gadovist injected late stage IPF are relatively homogenous while both early stage and late-stage mice exhibits heterogeneity in the pixel distribution.
Figure 6-15: Histogram Pixel-by-pixel analysis of RARE MRI pulse sequence for A., Normal healthy mice, B., Late-stage IPF mice with clinical contrast agent Gadovist injection, C., hProCA32.collagen-injected early stage IPF mice and D., Late-stage IPF mice with hProCA32.collagen injection.

6.2.5 3D-UTE analysis of normal and bleomycin (BM)-induced IPF mice

hProCA32.collagen was used in comparison with non-targeted hProCA32 and gadovist in 3D UTE MRI pulse sequence to determine whether the collagen-targeted ability could specifically detect pulmonary fibrosis with the clinical hallmarks of IPF. All 3D UTE images were performed in corona orientation. As shown in Figure 6-16 and Figure 6-17, no significant increase was observed before and after injection of hProCA32.collagen in normal mice (n=4) with no bleomycin treatment.
Figure 6-16: 3D UTE MR Images of normal mice shows no significant enhancement in the lungs.

The SNR value increased from 9.18 ± 0.42 for the pre time point to 9.49 ± 0.29, 9.52 ± 0.35 and 9.43 ± 0.24 at 1h, 3h and 24 h time points respectively. This represents an average of a 3% increase after injection of hProCA32.collagen.
Another set of mice (n=4) were treated with bleomycin sulfate for 6 weeks to induce late stage IPF after which non-targeted hProCA32 was injected. As shown in Figure 6-18, a signal increase can be seen at the 1 h time point compared to pre, 3h and 24 h time points which had no significant increase.
Figure 6-18: 3D UTE MR Images of 6-weeks bleomycin treated IPF pre, 1 h, 3h and 24 h after injection of non-targeted hProCA32.

Figure 6-19: 3D UTE MR SNR of 6-weeks bleomycin treated IPF mice. Images captured before, 1 h, 3h and 24 h after injection of non-targeted hProCA32 shows an increase in enhancement at 1 h time point in the lungs, this enhancement was reduced at 3 h time point and 24 h time point.
This increase was quantitatively quantified as shown in Figure 6-19. The SNR value obtained for the pre time point (9.77 ± 0.26) increased to 14.65 ± 0.42 at 1h time, 11.38 ± 0.41 for 3 h time point and 9.58 ± 0.3 for the 24 h time point. This corresponds to a 49.9% and 20% increase at the 1 h and 3 h time points respectively.

![Figure 6-20: 3D UTE MR Images of 6-weeks bleomycin treated IPF pre, 30 minutes, 3h and 24 h after injection of clinical contrast agent, Gadovist®](image)

The next set of mice (n=4) treated with bleomycin sulfate for 6 weeks to induce late stage IPF were injected with clinically approved contrast agent, gadovist. As can be seen in Figure 6-20, it is apparent that the signal increase in the lungs mice after injection of gadovist across the time points is comparable to that observed for the lungs of normal mice.
Figure 6-21: 3D UTE MR SNR of 6-weeks bleomycin treated IPF. Images pre, 1 h, 3h and 24 h after injection of non-targeted Gadovist shows an increase in enhancement at 1 h time point in the lungs, this enhancement was reduced at 3 h time point and 24 h time point.

The SNR values obtained for these mice was not significantly different across all time points. In contrast, injection of hProCA32.collagen in early stage IPF mice model (n=5) resulted in a significant enhancement at 1 h time point and a peak enhancement at the 3 h time point (Figure 6-22).
Figure 6-22: 3D UTE MR Images of 3-weeks bleomycin treated IPF pre, 1 h, 3h and 24 h after injection of hProCA32.collagen.

Figure 6-23: 3D UTE MR SNR of 3-weeks bleomycin treated IPF pre, 1 h, 3h and 24 h after injection of targeted hProCA32.collagen shows an increase in enhancement at 1 h time point in the lungs, this enhancement was increased at 3 h time point and reduced slightly at 24 h time point.
This Lung SNR value before hProCA32.collagen injection was 10.51 ± 0.41. This value increased to 13.78 ± 0.34 at 1 h time representing a 31.6%. More importantly, a 60% increase with SNR value of 16.79 ± 0.35 was observed after 3 h of hProCA32.collagen injection.

*Figure 6-24: 3D UTE MR Images of 6-weeks bleomycin treated IPF pre, 1 h, 3h and 24 h after injection of hProCA32.collagen.*

*Figure 6-25: 3D UTE MR SNR values of 6-weeks bleomycin treated IPF pre, 1 h, 3h and 24 h after injection of hProCA32.collagen shows an increase in enhancement peaking at 1 h*
time point in the lungs, this enhancement was maintained at 3 h time point and 24 h time points.

Interestingly, mice treated with bleomycin for 6 weeks to generate late stage IPF and injected with hProCA32.collagen (n=6) had a 139% increase 1 h after injection of hProCA32.collagen, the enhancement was maintained up to 80.5% at 3 h time point.

6.2.6 Detection of Usual interstitial pneumonia (UIP) pattern by hProCA32.collagen

We previously reported the existence of portal hypertension associated with late-stage fibrosis in mouse liver which resulted in the slow washout of ProCA32.collagen1, the rat analogue from the liver [159]. In this same vein, the slow wash out and collagen targeting associated with hProCA32.collagen at the 3 h time point could be due to pulmonary hypertension which has been identified as a co-morbidity to lung fibrosis [411, 412], characterized by fibrotic destruction of the lung parenchyma, resulting in vascular bed density loss and hypoxic vasoconstriction in the lungs [413-415].

The UIP pattern of IPF usually identified on HRCT by a predominance of sub pleural reticulations and honeycombing at the basal sub pleural region of the lungs up to 40% of which can be missed or diagnosed as inconsistent with UIP patterns and usually diagnosed at the severe stage[416]. Due to the progressive nature of IPF [417], it is important to be able detect IPF with possible UIP pattern at an early stage.
Figure 6-26: Detection of early and late stage IPF.
A. Axial Ultrashort echo time (UTE)-MRI of the lungs of normal, early stage and late stage bleomycin (BM) treated mice. Coronal T1-weighted Ultrashort echo time (UTE)-MR images of mice from various groups of mice (A) Healthy mice before, 1 h, 3 h and 24 h after injection of hProCA32.collagen; (B) mice treated with bleomycin for 3 weeks (early stage) pre, 1h, 3h and 24 h after injection of hProCA32.collagen with heterogeneous enhancement of collagenous honeycombing at the sub pleural and paraseptal areas of the lung (black arrows) (C) 6 weeks BM treatment (late stage) pre, 1h, 3h and 24 h after injection of hProCA32.collagen with heterogeneous enhancement of collagenous honeycombing at the sub pleural and paraseptal areas of the lung (black arrows) with traction bronchiectasis (arrow heads) (D) 6 weeks BM treatment (late stage) pre, 1h, 3h and 24 h after injection of non-targeted hProCA32 with homogenous and rapid enhancement of collagenous honeycombing at the sub pleural and paraseptal areas of the lung at the 1 h time point (black arrows), at the 3 h time point, a noticeable wash out of hProCA32.collagen was observed. (E) 6 weeks BM treatment (late stage) pre, 30 mins, 3h and 24 h after injection of Gadovist® with no significant enhancement in the lung. F. Lung SNR plot of healthy, early stage and late stage IPF mice before and after injection of the MRI contrast agents.

As shown in Figure 6-26 A-E, hProCA32.collagen had no significant effect on normal mice with preserved lung architecture compared to early and late stage BM-induced lung fibrosis. The sub pleural region of the mice (black arrows) showed honeycombing that was detected in
both early and late stage IPF with hProCA32.collagen compared to non-targeted hProCA32 which showed a significant uptake at the 1 h time point without detecting honeycombing and gadovist with no significant enhancement after injection. In addition, hProCA32.collagen delineate traction bronchiectasis with “rim” enhancement (arrow heads).

![Figure 6-27: Histogram pixel analysis of lung after MR imaging.](image)

A., 3weeks BM-treated mice before and at peak enhancement (3h) after injection of hProCA32.collagen with heterogeneous lung SNR distribution compared to the pre time point B., 6 weeks BM-treated mice before and at peak enhancement (30 mins) after injection of gadovist C., normal mice before and at peak enhancement (1 h) after injection of hProCA32.collagen. D., 6 weeks BM-treated mice before and at peak enhancement (1h) after injection of hProCA32.collagen with heterogeneous lung SNR distribution compared to the pre time point with a more widely distributed SNR compared to 3-weeks treated.

In Figure 6-27 A-D, a dramatic increase in SNR value with heterogeneous pixel distribution is observed after injection of hProCA32.collagen contrast agent in both early stage and late stage
bleomycin-induced lung fibrosis mice model as depicted by the SNR histogram plot compared to normal, healthy mice and late stage IPF mice with injection of clinical contrast agent; gadovist. Figure 6-27 B and C both showed a homogenous distribution of SNR values. Patchy collagen network with heterogeneous cystic honeycombing is a major histopathological finding of IPF, this was detected in the bleomycin mice models injected with hProCA32.collagen at 3 weeks and 6 weeks after bleomycin injection.

Figure 6-28: Correlational studies of hProCA32.collagen MRI and histological findings in collaboration with Dr. Hua Yang.
A. Corona MR imaging of 6 weeks bleomycin-induced IPF mouse model shows relatively low MRI signal in the lungs before injection of hProCA32.collagen and high hProCA32.collagen MRI signal predominantly in sub pleural and basilar regions of the lungs in MRI which
corresponds directly with the dense area and red collagen area in H&E and Sirius red staining respectively. B. Pulmonary nodules

As shown in Figure 6-28, increased MRI signal was observed in the sub pleural and basilar regions of the lungs corresponding to areas with high lung injury and collagen accumulation in both H&E and Sirius red staining. This further confirms the capability of hProCA32.collagen in IPF staging and detection.

6.2.7 MR Imaging of nicotine-induced lung fibrosis in mice

To further demonstrate the specificity of hProCA32.collagen to collagen accumulation in lung disease, a second cohort of mice were tested for lung fibrosis imaging in collaboration with Dr Ji and Maureen Meister of the nutrition department of Georgia State University. In this mice model developed by Dr. Ji, C57BL/6J-βENaC-Tg mouse model with overexpression of the sodium channel nonvoltage-gated 1 beta subunit, (Scnn1b) was exposed to e-cigarette smoke for one hour, 2x daily, for 10 days. As shown in Figure 6-29, tumor infiltrating macrophages (M2) macrophages which are major contributors to extracellular matrix formation, activated by exposure to cytokines thereby producing proline to induce collagen production was significantly higher in e-cigarette smoke mice further confirming the presence of ECM deposition and collagen accumulation. To further confirm lung fibrosis development in the mice model, the mice were imaged on a As shown in Figure 6-30, representative e-cigarette-induced COPD mice (COPD 1 and COPD 2 mice) showed an increased signal and lung abnormality in the central region corresponding to areas of bronchi location in the lung with significantly increased SNR after injection of hProCA32.collagen.
Figure 6-29: M2 macrophage estimation in nicotine-induced COPD mice model. A. selected out M2 macrophages from second gated cell population by CD206 M2 macrophage marker and F4/80 macrophage markers in control and B. e-cigarette. C. Percentage of total macrophage in control and e-cigarette smoked mice. (Data provided by Dr. Guangda Peng)

Figure 6-30: Spin echo multi-slice (T1-W) images of e-cigarette smoked mice for assessment of COPD (airway fibrosis) induction.
Next as described for IPF mice above, we performed T1-W RARE and 3D UTE MR Images on a 7T Bruker MRI scanner. The 3D UTE result shown in Figure 6-31 depicts the same observation as Figure 6-30. More importantly, a significant SNR increase was observed both at the 1 h time point and the 3 h time point of the nicotine-induced mice model with maximum accumulation of hProCA32.collagen in the airways. Whole scan histology with zoomed-in bronchi region shows architectural destruction with increased collagen accumulation.

![3 D UTE COPD Mice Pre 1 h 3 h 24 h](image)

*Figure 6-31: 3D UTE MR Images of nicotine-induced lung fibrosis pre, 1 h, 3 h and 24 h after injection of hProCA32.collagen.*

To demonstrate the specificity of hProCA32.collagen, its uptake into IPF lung was compared to its uptake in nicotine-induce lung fibrosis in a COPD mice model. hProCA32.collagen accumulation was specifically observed in the bronchi in contrast to alveolar accumulation observed with IPF mice model with peak enhancement at the 1 h time point (Figure 6-31). This was seen both in 3D-UTE images as well as RARE images (Figure 6-34). Comparison of pathophysiology, collagen accumulation and MRI signal can be seen in Figure 6-35 with both MRI and histology data showing at the bronchi airways.
Figure 6-32: 3D UTE SNR values of nicotine-induced lung fibrosis pre, 1 h, 3 h and 24 h after injection of hProCA32.collagen.
Figure 6-33: Hematoxylin and Eosin (H&E) staining with Sirius red stain of adjacent slices of nicotine-induced COPD Mice model shows collagen accumulation in the airways (bronchi) of the mice (zoomed in box area)
Figure 6-34: Detection of airway fibrosis in e-cigarette smoke transgenic mice model. BM-induced idiopathic pulmonary fibrosis in mice. 
A. RARE (T1-weighted) MRI of normal (healthy) mice, transgenic epithelial sodium channel (ENac) mice, and transgenic epithelial sodium channel (ENac) mice with e-cigarette smoke before, and 3h after injection of hProCA32.collagenB. Lung signal-to-noise ratio (SNR) increase after injection OF hProCA32.collagen. E-cigarette smoked mice had a significantly higher SNR increase compared to healthy mice with enhancement in the lung airways (blue arrow). C. T1-weighted 3D UTE MRI of normal (healthy) mice, transgenic epithelial sodium channel (ENac) mice, and transgenic epithelial sodium channel (ENac) mice with e-cigarette smoke before, and 3h after injection of hProCA32.collagen D-G quantification of lung fibrosis in e-cigarette smoke mice. D. Lung signal-to-noise ratio (SNR) increase after injection of hProCA32.collagen E. AUC_1–24 h boxplot analysis demonstrating the ability of hProCA32.collagen to distinguish normal from early stage airway fibrosis and late-stage fibrosis in e-cigarette smoked mice model (P < 0.001, unpaired two-tailed student’s t test; the midline is the median of the data, with the upper and lower limits of the box being the third and first quartile, 75th and 25th percentile, respectively. The whiskers extends up to 1.5 times
the interquartile range and show the minimum and maximum as they are all within that distance). F. Histogram pixel analysis of normal mice vs e-cigarette mice model before and at peak enhancement (1 h) after injection of hProCA32.collagen G. Histological analysis of e-cigarette smoked mice versus normal mice showed increased collagen accumulation in the lung airways of e-cigarette smoke mice. All data are shown as mean ± SD, n = 4 biologically independent animals in each group, p <0.001. Histology scale bar =100 µm.

6.3 Comparison of imaging findings for IPF and COPD

A difference in hProCA32.collagen uptake was observed in BM-induced lung fibrosis at the alveolar region versus nicotine-induced COPD fibrotic model in the lung bronchi. The histology collagen area calculation also depicted an increased collagen overexpression in bronchi of nicotine-induced mouse model and in alveoli of BM-Induced IPF mouse model. This increase is observed as an increase in hProCA32.collagen-MRI SNR value for both mouse models as shown in Figure 6-35.

Figure 6-35: Significance of specificity of hProCA32.collagen uptake in alveoli and bronchioles of IPF and COPD and nicotine-induced lung fibrosis.
A-B. Collagen accumulation and binding in IPF (top) and COPD (bottom) with overexpression of collagen 1 (red color) at alveolar and the bronchial airways respectively compared to

---

The text describes the results of a study comparing the effects of nicotine exposure on lung fibrosis in mice, using histology and imaging techniques. The study found increased collagen accumulation in the lung airways of e-cigarette smoke mice compared to normal mice. Additionally, it observed differences in hProCA32.collagen uptake in BM-induced lung fibrosis and nicotine-induced COPD fibrotic models, with increased collagen overexpression in bronchi of nicotine-induced mouse models and in alveoli of BM-Induced IPF mouse models. The increase in hProCA32.collagen-MRI SNR value for both mouse models is significant. The text is accompanied by images and graphs illustrating the findings.
normal healthy mice (middle panel), human collagen-targeted protein MRI contrast agent (hProCA32.collagen) detects collagen overexpression in IPF and COPD with alveolar (top) and bronchial airways (bottom) in 3D UTE MRI. C-D., Quantification of percentage collagen area in nicotine-induced mice model and D. bleomycin-induced mice model, E. Comparison of lung signal enhancement (%) in normal mice, nicotine-induced lung fibrosis model for COPD and bleomycin-induced mice model.

6.4 Comparison with lung and liver enhancement, half life

To elucidate the in vivo organ biodistribution of hProCA32.collagen, a time course biodistribution in mice was estimated in different organs by drawing a region of interest in these organs. Our findings (Figure 6-36 and Figure 6-37A) indicates that hProCA32.collagen had the highest signal increase at 1 h time point providing blood pool enhancement and a 139% signal increase in the lung, and a 10% increase in the liver for 6 weeks bleomycin treated mice, this was accompanied with no significant enhancement observed for all other organs evaluated. At the 3 h time points, the lung enhancement reduced to ~80% while a 30% and 12% increase in SNR was observed in the liver and kidney respectively. At 24 h time point, all the organs had SNR values like baseline while kidney had the most significant (68%) increase. In contrast, no SNR increase was observed at any point in the muscle. In addition, hProCA32.collagen displayed a difference in pattern of enhancement between late-stage IPF and early-stage IPF. More importantly, while the peak enhancement for late stage of lung fibrosis is observed at 1 h time point, and a delayed enhancement at 3 h time point, the hProCA32.collagen did not have a significant uptake in the liver until the 3 h time point. The increase in SNR value for late stage IPF at 1 h time may be indicative of pulmonary hypertension [418] while the 3 h time point represent the targeting time point.
Figure 6-36: In vivo pharmacodynamics and pharmacokinetics studies of hProCA32.collagen in different organs of 6 weeks bleomycin-treated mice.
A. 3D-UTE T1-W whole body MR image of mouse showing different organs and enhancement pattern
B. The SNR quantification of muscle, kidney (C) and liver (D) shows that hProCA32.collagen is present in the liver at 1 h post injection time point, with significant liver enhancement at the 3 h post injection time point and a return to baseline at 24 h post injection time point. Conversely, no significant increase in muscle enhancement was observed at any point post injection while the kidney increases progressively from 1 h post injection time point to 24 h time point.

To further understand this observation, a comparison to previously published data for liver fibrosis mouse model [159] using the rat analogue, rProCA32.collagen was made (Figure 6-37B), a clear difference in the peak enhancement time points for lung fibrosis and liver fibrosis. For example, for both early and late-stage liver fibrosis using R₁ values. The serial change of R₁ values can be seen in Figure 6-37B where all R₁ values in both early stage of the disease and normal liver increased to 2.0–2.4 s⁻¹, 3 h post injection of rProCA32.collagen1. However, at 24 h post injection, R₁ values of normal liver decreased but fibrotic and livers
with NASH showed an increase to 2.4 and 2.8 s\(^{-1}\), respectively. These values decreased for both livers as the contrast agent was washed out of the liver 48 h post injection. However, at 3 h post injection of rProCA32.collagen, \(R_1\) of late-stage fibrotic livers for both animal models increased to >3.5–4.1 s\(^{-1}\) with \(\Delta R_1 \geq 2.0\) s\(^{-1}\). The Increase of \(R_1\) at 3 h post-injection for late-stage fibrosis likely suggests the existence of intrahepatic angiogenesis.

![Graph A](image1.png)

**Figure 6-37**: Comparison of the lung and liver signal enhancement pattern. *Serial Molecular Imaging of hProCA32.collagen in IPF versus rProCA32.collagen in liver fibrosis [159] shows the capability of both contrast agents to distinguish early from late-stage lung and liver fibrosis in a time-dependent manner.*

6.4.1 **Discussion of challenges in lung imaging**

Several challenges observed for lung MRI image process
1) Low voxel size and proton density for convention sequences such as RARE, FSEM and SEM, and

2) High background signal in some images due to magnetic inhomogeneity.

To overcome these challenges, percentage increase was utilized for conventional sequence images since hProCA32.collagen could detect fibrotic network in IPF models compared to normal non-diseased mice. In addition, the pattern of enhancement observed for hProCA32.collagen injected mice was significantly different from non-targeted hProCA32 or gadovist injected mice (Figure 6-26) with no visible pattern observed. The lung enhancement (%) correlated linearly with those obtained with the high voxel 3D UTE sequence. Next, due to high magnetic inhomogeneity, some of the total mice imaged had high background noise which led to SNR values deviating significantly from the other sets of data. As shown in Figure 6-38, early-stage bleomycin induced IPF was used as an example. In these mice, the pre time point had an SNR value of 15 while the 1 h and 3 h time points had values of 8 and 15.4 respectively. While this may have been easily ruled as an outlier, a closer look at the background values in the “pre” time point presented a major disparity as different slices are examined. (Figure 6-38B). To bring this into context, SNR calculation using background value obtained in slice 108 will result in a much lower SNR calculation compared to slice 106. If it has been established that the background area selected is not from surrounding tissues of image artifacts (Figure 6-38), then it will be critical to select background values that best represents the image acquired. Multiple areas should be circled as regions of interest in the background across several slices. The values should be compared with each other, and outlier mean intensities should be eliminated and an average of the other mean intensities should be used.
This method was utilized for SNR calculation and the values were better depicted and were observed to follow the same trend seen with the other early-stage mice.

![Image of Figure 6-38: Background selection troubleshooting of 3D-UTE lung data. A., 3D-UTE lung MRI obtained at different time points showing different level of background noise across the time points. B., Raw background mean and standard deviations for different slices in the same mice. C., 3D visualization of image shows mouse hind leg in position circled in A as area considered to be image background. D., Calculated SNR value before background correction and E., after background correction.]

### Major conclusion, limitations, and future direction

The use of conventional MRI of the lung faces major challenge for years due to the low proton density in the lung and the fast signal decay due to susceptibility artefacts at air-tissue interfaces. In addition, development of animal model mimic human disease characteristics is also required.
In this chapter, we demonstrate following major findings.

First, in collaboration with Dr. Liu’s lab, an IPF model was developed by intraperitoneal injection of Bleomycin twice weekly to progressively develop IPF with features closely resembling human IPF features at sub pleural and paraseptal regions in addition to cystic honeycombing and traction bronchiectasis.

Secondly, using RARE T1 and T2 weighted pulse sequences, we show that hProCA32.collagen enabled conventional T1 and T2 weighted MRI was successful in the specific detection of IPF at the 3 h targeting time point in both early and late-stage diseases with notable fibrotic network. Additionally, the use of the optimized 3D-UTE pulse sequence, hProCA32.collagen was able to delineate the histopathological features of human IPF such as traction bronchiectasis, cystic honeycombing with specific presence at the apico-basal region (sub pleural) of the lung with associated heterogeneity. The SNR-AUC values obtained by MRI increased with increasing disease progression with linear correlation by histological characterization of several markers such as collagen area and α-SMA. The specificity with which hProCA32.collagen binds to collagen expressed in IPF lung matches with the areas with collagen accumulation and pulmonary nodules in Sirius red staining.

Third, pixel by pixel histogram analysis showed heterogeneous hProCA32.collagen uptake in areas corresponding to high heterogenous collagen distribution associated with human IPF compared to normal mice and late-stage mice injected with clinical contrast agent Gadovist. This observation was seen by dual MR pulse sequence (RARE and 3D UTE).

Fourth, collagen overexpression in the lung bronchi was detected by hProCA32.collagen in a nicotine-induced COPD mice model with similar accumulation pattern predominantly in the airways and bronchioles of the mice lungs as observed for histological collagen staining.
Fifth, the distribution dynamics and kinetics of hProcA32.collagen in different organ shows its ability to detect collagen in specific organs based on distribution times. For example, the peak enhancement for lung fibrotic mice with high collagen content was 1 h compared to 3 h observed in liver of the same mice where no significant damage was done to the liver. Conversely, in liver fibrosis imaging by rProCA32.collagen and hProCA32.collagen (chapter 7) the peak enhancement for late-stage liver fibrosis model with high collagen content is at 3 h time point while the targeting time point for both early stage with minimal collagen accumulation and late stage was 24 h due to the biodistribution pattern of the collagen-targeted contrast agent. This is important in clinical translation enabling the use of hProCA32.collagen in time-dependent detection and staging of diseases in various organs.

More importantly, ECM plays a critical role in carcinogenesis and cancer progression. Collagen is the major component of the tumor microenvironment and participates in cancer fibrosis with type I collagen secreted by fibroblasts during tumor growth and epithelial-mesenchymal transition [419-421]. Both patients with COPD and IPF have been associated with increased risk of lung cancer [422-427]. Non-small cell lung cancer accounting for >80% of all lung cancer has been reported to secrete type I collagen [428] and the use of collagen-targeted hProCA32.collagen MRI can significantly improve image guided treated of lung cancer by precise detection of collagen secreted by the lung cancer cells and the tumor microenvironment. The specificity in the uptake of hProCA32.collagen to collagen accumulation depicted in two different mice models reported in this chapter further enforces its clinical relevance for separating different disease pattern and early detection of lung diseases with non-invasive MRI.
7 APPLICATION OF HPROCA32.COLLAGEN IN EARLY AND PROGRESSIVE DIAGNOSIS OF LIVER FIBROSIS

7.1 Introduction

7.1.1 The Human liver and liver disease

The liver is the largest organ inside the human body and one of the organs most susceptible to injury. It is responsible for several regulatory activities of chemicals in the body. The liver is responsible for over 500 vital functions some of which include, bile secretion/removal, fat, protein and carbohydrate metabolism, enzyme activation etc. (Figure 7-1). It plays a central role in lipid metabolism, importing serum free fatty acids and manufacturing, storing and exporting lipids and lipoproteins which makes it susceptible to injury or diseases caused by viruses such as in hepatitis A, B or C, Ingestion of drug, excessive alcohol etc. nonalcoholic fatty liver diseases leading to fibrosis, cirrhosis, cancer, etc. [429]. Despite the propensity of the liver to injury, it possesses an incredible regeneration capability with peak occurring when two-thirds of the liver is removed. A less than two thirds removal of the liver mass results in slower liver regeneration while an amount greater than two thirds also impacts mitotic activity and DNA synthesis.[430]
Figure 7-1 A. The human liver and the several vital functions it performs in the human body (B).

7.1.2 Fatty liver diseases

A fatty liver also known as hepatic steatosis is the result of the excess hepatic fatty accumulation in liver cells. Just as the name implies, a fatty liver contains a lot of fat especially at an amount too much for an individual’s body to handle, typically fatty components greater than 5% of the liver [431]. This can be either an aftermath of excessive alcohol consumption known as alcoholic fatty liver disease or non-alcohol induced usually described as nonalcoholic fatty liver disease (NAFLD). Non-alcoholic fatty liver disease (NAFLD) is divided into non-alcoholic fatty liver (NAFL) and non-alcoholic steatohepatitis (NASH). NAFL refers to hepatic lipid accumulation (steatosis) with no obvious cell damage or inflammation while NASH similar to NAFL, has steatosis in addition to significant inflammation, cellular damage and varying degree of overexpression of extracellular matrix component which can progress to cirrhosis and hepatocellular carcinoma (HCC) that may require liver transplant [432, 433]. Lipid accumulation (as triglyceride), obesity, insulin resistance, high triglyceride, dyslipidemia (high plasma TG and/or low plasma HDL-
cholesterol concentrations), and hypertension are some of the metabolic disorders that have been implicated in the pathogenesis of NAFLD although a lot is yet to be understood about the underlying mechanisms of the disease Figure 7-2 [434]. NAFLD is a silent epidemic [435] affecting about 30–40% of the US adults and about 10% (40% for obese pediatric population) of the children population translating to over a hundred adults and 10 million children, respectively. Of these population, about 20% will progress from NAFL to NASH and about 15-25% of the NASH patients will develop cirrhotic liver from which a subset will progress to HCC [436-438] making it the most common chronic liver disease in the world [439]. NAFLD has the propensity for chronic cardio-metabolic abnormalities, including type 2 diabetes mellitus (T2DM), cardiovascular disease and the metabolic syndrome [440]. There are several stages of fibrosis according to the Metavir scoring system including stage F0 with no fibrosis in the liver, stage F1 involves fibrous portal expansion with mild fibrosis, stage F2 has few bridges or septa with moderate fibrosis, stage F3 depicts numerous bridges or septa with associated advanced fibrosis while the stage F4 is the cirrhosis stage (Figure 7-2).
Obesity is a complex human disease with associated chronic health problems such as high blood pressure, type 2 diabetes, fatty liver diseases, cardiovascular disease, sleep apnea, stroke, osteoarthritis etc. Obesity results from excess energy storage resulting from an imbalance between food intake and energy expenditure. Adipose tissue lipolysis, de novo lipogenesis (DNL) and dietary intake are three causal factors of excess hepatic fatty acids.[437] Accumulation of lipid droplets as triglycerides is a feature of NAFLD, excessive accumulation lowers the ability of the liver to control the use, storage and export of intrahepatic triglycerides results in uncontrolled DNL and subsequently, lipotoxicity.[441] Lipotoxicity and high TG content is common in patient with insulin resistance which contributes to impaired glucose metabolism and insulin sensitivity in muscle and in the liver.[442] In obesity-related NAFLD, there is an increase of free fatty acid (FFA) delivery to the liver especially during the fed state,
due to adipose tissue insulin resistance[443-445]. Compensatory increase in very low-density lipoprotein (VLDL) secretion is not sufficient to overcome the excess formation of triglycerides. [440, 446]

7.1.4 Organ Fibrosis

Fibrosis is the scarring and thickening of tissues in response to an injury. It is a common feature in several pathologies, characterized by the accumulation of extracellular matrix (ECM) components such collagen and is associated with up to 45% of deaths in industrialized countries[351]. In general, fibrosis mechanisms include exacerbated injury-related responses that are dysregulated versions of conventional tissue repair processes[447]. Core elements in disease progression can be acute, recurrent or persistent epithelial and endothelial injuries that can both initiate and sustain fibrosis[448]. These pathogenic pathways further involve changes in numerous inflammatory, endothelial and mesenchymal cells, resulting in increased production of fibronectin, osteopontin, hyaluronan, proteoglycans, laminins and a variety of collagen subtypes[449]. The abnormal ECM deposition which is typical of fibrosis leads to a disruption and distortion of the tissue architecture and function[450]. Although the pathophysiology of fibrosis is similar in various tissues, the regenerative capacity and the ability to reverse advanced fibrosis vary tremendously. Resolution of fibrosis is for example possible in the liver when the underlying cause is removed; this is for instance the case where complete suppression of hepatitis B virus replication and curative treatment of hepatitis C [451]. Fibrosis is a common endpoint for most tissues, with diverse causal factors such as acute injury, chronic inflammation, autoimmune reactions and genetic alterations. In some cases, the
exact etiology may even remain unknown as in the case of idiopathic pulmonary fibrosis (IPF) and NAFLD figure.

### 7.1.5 Mice Models for the development of NAFLD

Several animal models including dietary such as high fat, high cholesterol, western diet, methionine and choline deficient etc., genetically induced, chemically induced and a combination of both have been developed.[452] First, we developed a diabetes/obesity mice model utilizing expression of family 1 histone deacetylases (HDAC1), a class of enzymes that remove acetyl groups from an amino acid on a histone. HDACs are classified into four classes, class I to IV. These classes are further divided into families based on their sequence similarities. Class 1 is divided into HDAC1 to IV etc. [453]. HDAC proteins are vital regulators of fundamental cellular events, such as cell cycle progression, differentiation, and tumorigenesis. Abnormal HDACs can contribute to many different human diseases including cancer, neurodegenerative disorders, cardiac hypertrophy, and pulmonary diseases[454]. Histone acetylation and deacetylation are regulated by the balanced action of histone acetyltransferases and histone deacetylases (HDACs). HDAC1 is essential in energy regulation in brown adipocytes and glucose homeostasis [455, 456]. Next, we utilized the human comparative gene identification mice model. The gene targeting the CGI-58 construct was assembled from two DNA segments of the mouse CGI-58 gene amplified by PCR from the genomic DNA of R1 mouse embryonic stem (ES) cells. The gene-targeting construct upon linearization and purification was introduced into R1 mouse ES cells. Positive and negative selections was done, and the ES cell clones were screened for homologous recombination by PCR and Southern blotting. The selected clones were injected into C57BL/6 blastocysts. Chimeric male mice were allowed to mate with C57BL/6 female mice and the germ-line
transmission was assessed by PCR genotyping. Mice with the right allele were mated with 129S4/Sv-Tg-Gt (ROSA)26Sortm1(FLP1) dym mice (Jackson Laboratory, stock #003946). The progeny was subsequently crossed with C57BL/6 mice to isolate the Flp transgene leading to the generation of heterozygous CGI-58 floxed (CGI-58 f/+ ) mice harboring an allele in which exon 3 (the largest exon in the mouse CGI-58 gene) is flanked by two LoxP sites. These CGI-58 f/+ mice were then crossed with Albumin-cre mice of C57BL/6J mice to generate CGI-58 f/+ /Alb-cre mice and CGI-58 f/+ mice, which had 93.75% C57BL/6 background. Further cross breeding of these two genotypes generated homozygous CGI-58 f/+ mice with Alb-cre transgene (LivKO) mice and homozygous CGI-58 f/+ mice without Alb-cre transgene (wild type) mice [457]. The HDAC1 mouse model is based on the epigenetics and brown fat thermogenesis in the development of diabetes since HDAC1 is a negative regulator of the brown adipocyte thermogenic program. This model mimicked the natural occurrence in human and is less aggressive, progressing from steatosis to early NASH (stage F1). In contrast, the CGI-58 knock out mouse model with western diet and fructose water treatment, the process is artificial, extreme, and aggressive model capable of developing late-stage liver fibrosis/cirrhosis (F4), by feeding the mice with western diet and fructose water, we also observed moderate fibrosis (F2) in mice. In this model, the focus was more on the role of adipose triglyceride lipase (ATGL) in triglyceride (TG) hydrolysis. By knocking out the liver CGI-58 gene, the hepatic TG hydrolase activity was reduced resulting in the development of NAFLD. Lastly, the western diet/cholesterol/CCl4 model takes advantage of the activation hepatic stellate cells (HSCs) which are critical source of activated myofibroblast that can drive the development of fibrosis in NASH and cirrhosis. This model represents advanced-stage liver
fibrosis (F3). All mice models were characterized and confirmed by histology and MR imaging analysis.

7.1.6 Molecular Imaging of NAFLD and their limitations

Currently, there is lack of simple, reproducible, precise, and noninvasive diagnostic tool for disease staging in patients with NAFLD. Liver biopsy is the gold standard to distinguish patients with NASH from those without NASH or bland steatosis, therefore, it is imperative to accurately stage fibrosis for better disease prognosis. Establishing an accurate diagnosis of NASH is of major clinical importance. A histologic diagnosis of NASH is associated with cardiovascular disease and more rapid progression of liver disease. To accurately distinguish NASH from NAFL, liver biopsy is required. NAFL is defined as bland steatosis with minimal or no inflammation, whereas NASH is characterized by macro vesicular steatosis, hepatocyte injury, and mixed lobular inflammation with or without zone 3 perisinusoidal fibrosis.[458-460]. Recently, an array of clinical biomarkers has been utilized for predicting NASH and fibrosis. For example, there has been a development of the NAFLD fibrosis score, a scoring system used that utilizes age, body mass index, albumin level, Aspartate aminotransferase (AST) - Alanine aminotransferase (ALT) ratio to predict fibrosis in NAFLD. A major limitation of most of the biomarkers and scoring systems is their inability to achieve high sensitivity for the diagnosis of early and moderate stage fibrosis. Another important factor that could result in high false positive is the upregulation in bilirubin, decrease in haptoglobin, Gilbert syndrome, and cholestasis. Noninvasive imaging modalities are better alternatives in the diagnosis of NAFLD. Early diagnosis and risk stratification are essential for effective management. Current imaging methods such as ultrasound, CT, and MRI have demonstrated their values to serve as noninvasive imaging biomarkers to evaluate NAFLD progression, but
they are still relatively limited in the detection of inflammation and collagen accumulation [461-463]. For instance, MRI is preferred as the most definitive imaging modality to diagnose hepatic steatosis qualitatively and quantitatively since both fat and water contain protons that contribute to the overall signal observed in the liver MRI especially with different precession rate for fat protons and water proton allowing accurate lipid quantitation. This difference in precession rate is utilized in chemical-shift-encoded MRI (Dixon method) one of three MRI imaging techniques whereby fat and water undergo phase difference interference at predictable regular intervals. Fat and water signals cancel at out-of-phase (OP) and add at in-phase (IP) echo times enabling the assessment of hepatic fat by comparison of sequential OP (indicated as signal loss) to IP images [463-466]. Transient elastography (TE) has been routinely used as a tool for non-invasive assessment of liver fibrosis especially in chronic viral hepatitis, it has questionable accuracy since its measurement is based on the shear wave speed through the liver reflecting the stiffness of the liver and not actual fibrosis. Several factors such as inflammation and congestion, extrahepatic cholestasis etc [467]. There is therefore a need to detect, stage and monitor treatment of liver fibrosis noninvasively and accurately.

7.1.7 Overview of this Chapter

Here, our efforts towards application of hProCA32.collagen1 in early detection of NASH and cirrhosis using different mice models including an obesity induced and a genetically induced mouse model is reported. Based on histology analysis, hProCA32.collagen1 can detect fibrosis at early stage in HDAC1 mice model as well as CGI-58 KO mice models. The time dependent staging and heterogeneous uptake of hProCA32.collagen is also reported. We will address several questions; First, we will discuss different mice models for the development of liver fibrosis in mice using various pathways. Next, we ascertain the presence of liver injury
morphological changes by characterization. Our efforts towards application of hProCA32.collagen1 in early detection of NASH (inflammation and fibrosis) using an obesity induced and a genetically induced mice model is then reported. Based on histology analysis, hProCA32.collagen1 can detect fibrosis at early stage in HDAC1 mice model as well as CGI-58 KO mice models. The time dependent staging and heterogeneous uptake of hProCA32.collagen is also reported.

7.2 METHODS

7.2.1 Development of Mouse model for NASH study using class I histone deacetylases (HDACs)

It is especially important to develop reliable noninvasive methods to monitor patients at risk for alcoholic liver diseases (ALD) and NASH. We tested whether molecular MR imaging with hProCA32.collagen will allow accurate detection of fibrosis at early stages, monitoring of liver fibrosis progression, and response to treatment. When identified in an early or moderate stage, ALD can be controlled by cessation of alcohol consumption. The contrast agent can later be applied to many types of fibrotic diseases. Early-stage fibrosis was induced in 6-8 weeks C57BL/6 mice (Jackson Laboratories, Bar Harbor, ME) by daily feeding with high fat diet with 4.2% fructose in drinking water. As shown in Figure 7-3, HDAC1 gene induced NASH mouse model was developed in a Cre/loxP system into ROSA26 locus. The mice were housed with a 12/12-h light/dark cycle in temperature- and humidity-controlled rooms with unrestricted access to water and food.
A more aggressive NASH and liver fibrosis model was used using CGI-58 gene knockout mice. Two different groups of mice (n=4 in each group) on NASH diet were chosen for the studies. One group had a liver CGI-58 knocked out while the other group had no CGI-58 knocked out and are therefore designated the wild type (WT) mice. Both groups were simultaneously fed with non-sterile western diet (D12079B) and 42g/L fructose in drinking water starting from 6 weeks old. Intraperitoneal (i.p.) injection was done twice a week. Fructose in drinking water was autoclaved and changed twice a week. As control, normal mice were purchased with no treatment and placed on regular chow diet.
7.2.3 Development of Mouse model for Advance Liver fibrosis

For these set of mice, 9-week-old mice were fed with western diet containing 21.1% fat, 41% Sucrose, and 1.25% Cholesterol by weight (Teklad diets, TD. 120528) and a high sugar solution 23.1 g/L d-fructose (Sigma-Aldrich, F0127) and 18.9 g/L d-glucose (Sigma-Aldrich, G8270). CCl₄ (Sigma-Aldrich, 289116-100ML) was administered at the dose of 0.2 µl (0.32 µg)/g of body weight will be injected via intra-peritoneal route once per week, which is much lower than the dose that is usually given for fibrosis induction with CCl₄ alone. Mice were euthanized (n=2) weekly to assess disease progression until 18 weeks. Mice were imaged and euthanized by exsanguination after ketamine and xylazine anesthesia (Figure 7-4). Liver and serum samples were collected and processed for histological, serological and gene expression analysis.

![Figure 7-4: Timeline of liver fibrosis model development and progression induced by western diet/CCl₄.](image)

In this model, a western diet (WD), which is high-fat, high-fructose and high-cholesterol was used to mimics fast food style diets which have been implicated in the pathogenesis in
nonalcoholic steatohepatitis (NASH) in humans[468]. By combination of WD to induce obesity and insulin resistance with carbon tetrachloride (CCl₄) which has reported to induce liver injury and fibrosis in mice[469] by triggering oxidative stress resulting in inflammation and apoptosis [470, 471].

### 7.3 RESULTS

#### 7.4 MR Imaging of CGI-58 mouse model

First, to ascertain the in vivo capability of hProCA32.collagen, the CGI-58 mouse model, that had been previously reported to be successfully detected by rProCA32.collagen was utilized in hProCA32.collagen MR imaging. T1 weighted, quantitative T1(t1), and quantitative T2 map were collected for all groups of mice and quantitative analysis was performed on the liver signal as well as histological data. 100 μL, 5mM concentration (0.02 mmol/kg) of hProCA32.collagen and hProCA32 were injected to CGI-58 KO mice. The r1 and r2 mapping is shown in Figure 7-6 At 3 h post-injection of hProCA32.collagen1, an increase in R1 to 3.51 s⁻¹ was observed (ΔR1~ 1.58 s⁻¹, an approximate 55% increase in the liver area) in late-stage fibrotic liver, compared to pre-injection of hProCA32.collagen1. The R1 increase at 24 h post-injection was specific for the early-stage mice. In contrast, the targeting agent did not result in a significant change in normal liver (ΔR1 ~0.3 s⁻¹) at either 3 or 24 h. Consistent with the R1 map, the R2 map and CNR and SNR results also exhibited the same pattern at 24 h for both early and late stages of fibrosis, with higher changes for the late-stage fibrotic liver (Figure 7-7). These values and enhancement patter are similar to the observed change with rProCA32.collagen1 [159]. In contrast, hProCA32 without collagen binding capability did demonstrate any significant enhancement observed in R1 or R2 maps at these same time points.
Both collagen presence and liver injury were confirmed by histological staining and ALT test (Figure 7-10).

Figure 7-5: Liver weight and Morphology post Euthanizing.
A. Gross picture of mice liver after euthanasia showing smooth morphology and architecture for normal mice compared to both CGI-58 WT and CGI-58 KO mice. B-D. Quantification of bio characteristics of the euthanized mice livers, B. ALT liver injury assessment of the mice showed a significant difference between both group relative to normal mice, the increased ALT values observed for CGI-58 wild type mice can be attributed to early stage NASH with injury and collagen accumulation, this was be confirmed by histological analysis. C. mice body weight was significantly increased for CGI-58 WT mice compared to normal and CGI-58 KO groups of mice D., Liver weight for each group of mice indicates that mice with CGI-58 KO gene had the highest increase in liver weight compared to normal mice and CGI-58 WT mice despite the increase observed in body weight. This was due to increase in fat accumulation and subsequently liver injury.

The liver morphological characteristics is shown in Figure 7-5, majority of the mice with CGI-58 KO gene showed no significant difference in overall body weight when compared to the normal mice while CGI-58 wild type with no gene knockout had a significantly higher body
weight due to fat accumulation. On comparing the alanine transaminase (ALT) level indicative of liver injury, both the CGI-58 KO mice had an increased ALT value of 170 ± 31 U/L and 93 ± 15 U/L for CGI-58 KO and CGI-58 wild type respectively compared to 30 ± 3 U/L for normal mice (Figure 7-11 B). In addition, the liver weight for both CGI-58 KO mice and wild type CGI-58 were 6.92 ± 0.78 g and 2.5 ± 0.28 g, respectively compared to normal mice liver weight of 1.01 ± 0.1 g.

Figure 7-6: RI and R2 mapping of CGI-58 induced fibrotic mice.
A., Representative maps and corresponding change in RI values of normal (control) mice, CGI-58 WT mice, and CGI-58 gene KO mice before, and 3 h after injection of hProCA32.collagen and hProCA32 contrast agents, B., Representative maps and corresponding change in RI values of normal (control) mice, CGI-58 WT mice, and CGI-58 gene KO mice before, and 24 h after injection of hProCA32.collagen and hProCA32 contrast agents. C. Representative R2 maps of normal (control) mice, CGI-58 WT mice, and CGI-58 gene KO mice before, and 3 h after injection of hProCA32.collagen and hProCA32 contrast agents. D. Representative R2 maps of normal (control) mice (top row), CGI-58 WT mice, and CGI-58 gene KO mice before, and 3 h after injection of hProCA32.collagen contrast agent. E. RI values of normal, CGI-58 wild type and CGI-58 knock out mice before, 3h, 24 h and 48 h after injection of targeted and non-targeted contrast agents c., r1 values of normal, cgi-58 wild type and cgi-58 knock out mice before, 3h, 24 h and 48 h after injection of targeted and non-targeted contrast agents.
Figure 7-7: T1-W Images of human comparative gene knock out mice model (CGI-58 LivKO) A. representative T1-W images before and after injection of hProCA32.collagen and hProCA32 B. C., SNR and CNR plot of CGI-58 mice models shows 24 h time point as the targeting time point for hProCA32.collagen in both CGI-58 KO and WT mice compared to hProCA32.

TI-W SNR values for hProCA32.collagen1 injected in CGI-58 KO were 29.5 ± 0.1.2, 22 ± 0.98 and 15.4 ± 0.59 for 3 h, 24 h and 48 h time points, respectively while the SNR value for wild type mice were 18.3 ± 0.14, 21.03± 0.48 and 14.46 for 3 h, 24 h and 48 h time points, respectively. The highest SNR value for the late-stage liver fibrosis mice (CGI-58 KO) was observed at 3 h compared to 24 h time point observed for early-stage CGI-58 wild type mice. Both mice showed significant increase in SNR at 24 h time point with a 20% higher SNR for late-stage liver fibrosis compared to early stage.
Heavy T2-W images has SNR values of 21.2 ± 2.51, 24.7 ± 0.62, 26.5 ± 1.2 and 20.4 ± 3.21 for early-stage liver fibrosis at pre, 3h, 24 h and 48 h time points respectively. The CNR values for the same sets of mice were 0.81 ± 0.4, 1.54 ± 0.65, 2.11 ± 0.19 and 0.98 ± 0.32, respectively (Figure 7-8).

Figure 7-8: Heavy T2-W images of early-stage liver fibrotic CGI-58 mice model.
The heavy T2 W images SNR values had similar increase at both 3 h and 24 h time point while the CNR values increase from $0.81 \pm 0.36$ before hProCA32.collagen injection to $1.82 \pm 0.41$ at 3 h post contrast and $2.05 \pm 0.15$ at 24 h post contrast. The CNR value decreased significantly at 48 h post contrast (Figure 7-9). The heavy T2 SNR values could therefore not clearly distinguish early from late-stage fibrosis compared to the T1-W SNR values. This was compensated for in the CNR calculation.
Histological analysis of the CGI-58 mice models (Figure 7-10) revealed consistent data with the $r_1$ and $r_2$ values as well as the CNR and SNR results. For both livers with and without CGI-58 Knockout, lipid droplets accumulation was observed translating to fatty liver, while collagen expression was significantly higher in liver CGI-58 KO mice with a collagen proportionate area (CPA) of $19.8 \pm 2.5$ compared to $8.2 \pm 1.21$ for wild type early-stage liver fibrosis mice with Ishak scoring of $2.3 \pm 0.65$ and $4.92 \pm 0.3$ respectively.

*Figure 7-10: Histological characterization of LivCGI-58 mice model. A., Hematoxylin and Eosin staining for normal, WT and LivCGI-58 KO mice with corresponding Sirius red stain (B). C., quantification of the collagen proportionate area and Ishak Scoring (D) for normal, WT and LivCGI-58 KO mice. (Scale bar = 100µm for n=4 mice per group)
7.4.1  *MR Imaging of HDAC1 NASH/liver fibrosis model*

To further test the detection ability of hProCA32 collagen for even earlier stage liver fibrosis, a diabetes mouse model with steatosis, inflammation and very early-stage fibrosis was evaluated. MRI imaging was obtained as described in 2.9.2 and 2.9.3. As shown in Figure 7-11B, the mice with HDAC1 gene showed no significant difference in overall body weight when compared to the HDAC1 KO mice and normal healthy mice. However, measurement of alanine transaminase (ALT) levels as a marker of liver injury, showed an ALT value of $132 \pm 22$ U/L in HDAC1 mice compared to $58 \pm 25$ U/L and $30 \pm 3$ U/L in HDAC1 KO and normal mice, respectively (Figure 7-11C). In addition, the liver weight for HDAC1 KO mice had no significant difference from normal mice while HDAC1 mice had a liver weight of $2.2 \pm 0.2$ g which is double the size observed for normal mice ($1.01 \pm 0.1$ g).

*Figure 7-11: Characterization of HDAC1 gene expressed NASH mouse model.*
A., Gross picture of mice liver after euthanasia showing similar morphology and smooth architecture for normal and HDAC1 KO mice compared to HDAC wt. B-D. Quantification of bio characteristics of the euthanized mice livers, B. mice body weight with no significant difference for all groups of mice, C. ALT liver injury assessment of the mice showed a significant difference between both group relative to normal mice, the increased ALT values observed for HDAC1 KO mice can be attributed to the high fat diet (steatosis) and not NASH.
or scarring from collagen accumulation, this can be confirmed by histological analysis. D., Liver weight for each group of mice indicates that mice with HDAC1 gene had the highest increase in liver weight compared to normal mice and HDAC1 KO mice.

Steatosis was induced in HDAC1 gene induced NASH mouse model developed in a Cre/loxP system into ROSA26 locus, in 6-8 weeks old C57BL/6 mice (Jackson Laboratories, Bar Harbor, ME) fed with high fat diet and 4.2% fructose in water daily. Since collagen type I has been implicated in the progression of liver fibrosis, we evaluated the capability of molecular MR imaging with hProCA32.collagen in quantitative detection of both early- and late-stage non-alcoholic fatty liver diseases and liver fibrosis in vivo in different mouse models, taking advantage of the high $r_1$ and $r_2$ values of the contrast agent. Figure 7-12A shows R1 map of fibrotic liver in a diabetes mouse model, hProCA32.collagen1 is capable of detecting early-stage NASH in nonalcoholic fatty liver in family 1 histone deacetylase (HDAC 1) knock out and wild type mouse model using both R1 and R2 maps. This model is a natural, progressive and non-aggressive mice model, it also better mimics the progression of NAFLD diseases in humans. HDAC 1 knock out prevents the development of NAFLD in NASH diet-fed mice. At 24 h post-injection of hProCA32.collagen1, early-stage fibrotic liver (Ishak 1 of 6) exhibited $\Delta R_1$ of 1.12 s$^{-1}$, which is 4.5-fold higher than $\Delta R_1$ of 0.25 s$^{-1}$ for normal liver (Figure 7-12BFigure 7-13). Consistently, $\Delta R_2$ determined by T2 map is 10.7 s$^{-1}$ compared to 4.2 s$^{-1}$ for normal liver at the same time point (Figure 7-13B). These results correlate well with body weight and liver weight that indicated similarities between the HDAC KO mice and normal non-fed mice (Figure 7-11 C and D). The ALT level however was elevated in HDAC KO mice (Figure 7-11 B).
Figure 7-12: RI mapping of HDAC induced fibrotic mice.
A., Representative R1 maps of normal (control) mice (top row), HDAC1 KO mice (middle row), and HDAC1 gene mice (bottom row) before, 3h, 24 h and 48 h after injection of hProCA32.collagen contrast agent, B., change in R1 values of normal, HDAC1 wild type and HDAC1 knock out mice 3h, 24 h and 48 h after injection of hProCA32.collagen C., R1 values of normal, HDAC1 wild type and HDAC1 knock out mice before, 3h, 24 h and 48 h after injection of hProCA32.collagen.
Figure 7-13: R2 mapping of HDAC induced fibrotic mice. A., Representative R1 maps of normal (control) mice (top row), HDAC1 KO mice (middle row), and HDAC1 gene mice (bottom row) before, 3h, 24 h and 48 h after injection of hProCA32.collagen contrast agent, B., change in R1 values of normal, HDAC1 wild type and HDAC1 knock out mice 3h, 24 h and 48 h after injection of hProCA32.collagen C., R1 values of normal, HDAC1 wild type and HDAC1 knock out mice before, 3h, 24 h and 48 h after injection of hProCA32.collagen.

Liver injury and presence of fibrosis were verified by Sirius red collagen staining. Figure 7-14A represents histological staining of HDAC1 KO and HDAC wild type mice in comparison to normal healthy mice. The hematoxylin and eosin staining of the healthy control group and the HDAC1 had no significant amount of lipid droplets compared to mice with the HDAC1 gene with high accumulation of lipid droplets and collagen staining evident from the Sirius red staining. The collagen proportionate area quantification (Figure 7-14) of control mice was estimated to be 0.2 ± 0.01 while HDAC KO and wild type were 0.65 ± 0.05 and 4.89 ± 0.23 respectively. The HDAC wild type were also given an average score of 1 (Figure 7-14 C-D) depicting very early stage of NASH.
Figure 7-14: Histological characterization of HDAC1 mice model. A. Hematoxylin and eosin stain and B. Sirius red staining of normal (control) mice, HDAC1 KO mice, and HDAC1 gene mice. C., Estimation of collagen proportionate area and D., Ishak scoring of normal (control) mice HDAC1 KO mice and HDAC1 gene mice.
An increase in T1-W SNR values of $(8.14 \pm 0.35)$ obtained for the pre time point to $8.98 \pm 0.16$ for 3 h time point, $14.3 \pm 2.92$ at 24 h time point and $12.2 \pm 1.37$ at the 48 h time point representing a 10%, 75% and 50% SNR increase for 3h, 24 h and 48 h time points respectively was observed for the mice with HDAC gene. In addition, the CNR values were $0.97 \pm 0.05$, $3.29 \pm 0.15$, $5.62 \pm 0.21$ and $4.28 \pm 0.5$ for pre, 3 h, 24 h and 48 h time points, respectively. The highest enhancement and contrast was observed at 24 h time point as depicted in the SNR and CNR values, respectively for representative images of two different mice with HDAC1 gene (Figure 7-15 and Figure 7-16).

Figure 7-15: TI-W MR images of mice 1 with HDAC1 gene
A., before, 3h, 24 h and 48 h after injection of hProCA32.collagen contrast agent B., signal-to-noise ratio quantification C., contrast-to-noise ratio. (Error bar was generated based on SNR and CNR calculation from different slices)
In contrast, no significant increase across time points was observed for mice with HDAC1 gene knockout. The T1-W SNR values pre contrast (9.18 ± 0.15) increased to 9.78 ± 0.25 for 3 h time point, 10.3 ± 0.41 at 24 h time point and 9.2 ± 0.10 at the 48 h time point respectively, representing a 6%, 12% and 0.2% SNR increase for 3h, 24 h and 48 h time points respectively. In addition, the CNR values were 0.97 ± 0.05, 3.29 ± 0.15, 5.62 ± 0.21 and 4.28 ± 0.5 for pre, 3 h, 24 h and 48 h time points respectively. (Figure 7-17 and Figure 7-18).
Figure 7-17: TI-W MR images of mice 1 with HDAC1 knock out gene
A., before, 3h, 24 h and 48 h after injection of hProCA32.collagen contrast agent B., signal-to-noise ratio estimation C., contrast-to-noise ratio. (Error bar was generated based on SNR and CNR calculation from different slices)
Figure 7-18: TI-W MR images of mice with HDAC1 knock out gene A., before, 3h, 24 h and 48 h after injection of hProCA32.collagen contrast agent B., signal-to-noise ratio estimation C., and contrast-to-noise ratio. (Error bar was generated based on SNR and CNR calculation from different slices)

7.5 Comparison of disease state and MRI observation of HDAC1 vs CGI-58 model

The HDAC1 model corresponding to steatosis and early-stage NASH was compared to the more aggressive CGI-58 mice model to determine the capability of hProCA32.collagen to progressively detect steatosis in HDAC1 gene KO mice, early NASH/liver fibrosis in HDAC1 gene/ CGI-58 wild type mice and late-stage liver fibrosis in CGI-58 KO mice.
Figure 7-19: Comparison of histological results of HDAC1 vs CGI-58 mice model. A Hematoxylin and eosin (H&E) staining (top row) and Sirius red staining (bottom row) of the two mice models, B., collagen proportionate area of HDAC1 mice model and CHI-58 mice model.

As shown in Figure 7-19, the collagen proportionate area increased with disease progression from normal to late stage liver fibrosis. A similar trend was observed for the Ishak’s scoring.
Figure 7-20: Summary of RI maps and R1 values
Comparison of normal, HDAC1 mice model and CGI-58 mice models shows an increasing R1 values from healthy mice to late-stage fibrotic mice.

The trend in histology is similar to that observed for MRI data as represented by the overlay of both animal models in Figure 7-20. The targeting time point is the 24 h time point although late-stage liver fibrosis model showed the highest enhancement at 1 h, likely due to intrahepatic angiogenesis and portal hypertension [159, 472]. This progressive increase in R1 value at 24 h time point shows that hProCA32.collagen can progressively detect and stage liver steatosis, non-alcoholic steatohepatitis and late-stage liver fibrosis.
7.5.1 The effect of PEGylation on liver fibrosis imaging

As described in chapter 5, the use of non-PEGylated hProCA32.collagen may be beneficial for reducing the circulation half-life of the targeted contrast agent and reduce cost of production if stable and soluble protein can be obtained without the use of PEG reagent. To this end, the imaging capability of non-PEGylated hProCA32.collagen was compared with its PEGylated analogue. As shown in Figure 7-21, the T1-W SNR value increase from 27.4 ± 0.89 obtained at pre time point to 45.6 ± 3.1 for 3h and 40.01 ± 2.0 at 24 h time point. This is a 66% and a 45% increase at 3 h and 24 h time points respectively. In addition, the T1-W CNR values increased from 10.47 ± 0.53 at the pre time point to 24.89 ± 1.2 at the 3 h time point and 23.8 ± 2.4 at the 24 h time point. Conversely, the T2-W SNR values decreased significantly from 18.3 ± 2.1 to 11.88 ± 1.03 and 10.99 ± 0.5 at 3h and 24 time points respectively.
Figure 7-21: MRI and histological analysis of liver fibrosis.
A. TIW and heavy T2W MR Images of western diet + CCl₄ fed mice model pre, 3 h and 24 h after injection of hProCA32.collagen B. Sirius red (top) and hematoxylin and eosin (H&E) staining (bottom) of western diet + CCl₄ fed mice model C-D., TI-W and T2-W SNR quantification of hProCA32.collagen enabled MR images E. Comparison of T-W SNR and CNR values.

Due to the discrepancy observed for the T2-W pulse sequence, each slide and sequence parameter of the T1 and T2 maps were individually analyzed as shown in Figure 7-22. This was compared to the T1 map. The TR values used for saturation recovery image acquisition were 298, 351, 532, 752, 1035, 1431, 2000 and 2500ms with a total number of slices = 8.
Figure 7-22: Representative T1 inversion recovery images of western diet/CCL4 induced NASH mice at different repetition times (TR) across 8 slices of MR images.

A similar analysis was done by analyzing each TE values in the T2 map Figure 7-23 with TE values of 6.5, 13, 19.5, 26, 32.5, 39, 45.5, 52, 58.5, 65, 71.5, and 78 ms respectively.
Figure 7-23: Whole slides T2 images of western diet/CCL4 induced NASH mice at 12 different echo times (TE) across 8 slices of MR images.

As can be deduced from Figure 7-24 and Figure 7-25, the use of individual TE and TR values yielded a similar trend in both T1 and T2 pulse sequence with the highest enhancement at 3 h time point and 24 h as the targeting time point.
Figure 7-24: Repetition time dependent image analysis.
A. T1 images of NASH mice liver at TR of 1.035s. B. Signal to noise ratio analysis results in a 2x increase in SNR values obtained at 3 h time point and a 3x increase in SNR at 24 h time point (due to time limitation, no 48 h time points were collected for these sets of mice).
Figure 7-25: TE dependent image analysis.
A. T2 images of NASH mice liver at echo time of 0.0195s. This image demonstrates a 2.5x increase in SNR and CNR values obtained at 3 h time point and a 2x increase in SNR at 24 h time point (due to time limitation, no 48 h time points were collected for these sets of mice).
Figure 7-26: Comparison of T1W MR images all three mice models. A. normal mice, B. HDAC1 gene mice with early-stage NASH, C. CGI-58 wild type for early-stage fibrosis D. Western diet/CCl4 mice for moderate stage fibrosis and E. CGI-58 KO mice model for late-stage fibrosis.
Figure 7-27: Summary and comparison of liver enhancement using rat vs human ProCA32.collagen.

A. SNR of mice liver before, 3h, 24h and 48 h after injection of hProCA32.collagen in all mice models.

B. R1 relaxivity values of mice liver before, 3h, 24h and 48 h after injection of rProCA32.collagen in all mice models.

7.6 Organ Harvesting and Analysis

After MR imaging, the mice were sacrificed, and their organ was harvested for histology analysis and further studies. Figure 7-28 not only showed a significantly increased liver weight with disease progression but also points to a difference in liver architecture as disease progress from steatosis to early stage to late-stage liver fibrosis mice in 3 different animal models for liver fibrosis.
7.7 Conclusion

hProCA32.collagen was successfully utilized in detection of liver fibrosis. Based on our results, the contrast agent can detect and distinguish different stages of the disease from early stage 1 (Ishak system) to late stage 5 liver fibrosis with high sensitivity. This sensitivity can be well demonstrated by increase in SNR of liver with fibrosis at 3 and 24 hours after injection of hProCA32.collagen using T1 and T2W MRI pulse sequences with no significant increase in SNR was observed for normal liver without fibrosis. The stage of fibrosis was confirmed with histology analysis. The development of collagen targeting contrast agent is expected to have broad applications in detection and staging of fibrosis in chronic diseases and liver metastasis from various types of cancer and probing heterogeneous microenvironment changes upon disease progression and treatment. Furthermore, less injection dosage can be used to avoid toxicity as our results demonstrate. hProCA32.collagen combined with MRI can be used as a quantification and promising method for diagnosing stage specific chronic liver diseases. MR imaging with the collagen-targeted contrast agent of hProCA32.collagen provided robust
detection and staging of liver fibrosis in three different mice models. All these models represent different mechanisms in the development of liver fibrosis in humans based on the cause of liver injury, morphology, and disease stage. hProCA32.collagen-enabled MRI progressively detected and staged non-alcoholic liver fibrosis from the earliest stage possible (F1) to the cancer stage (hepatocellular carcinoma) discussed in chapter 8.

8 APPLICATION OF HPROCA32.COLLAGEN IN PRECISION MRI (PMRI)
DETECTION OF HEPATOCELLULAR CARCINOMA (HCC)

8.1 Introduction
Hepatocellular carcinoma (HCC) is the most common primary liver cancer. Liver cancer is the second most common cause of death worldwide with HCC being the fifth most common cancer in men, the ninth most common cancer in women [27] accounting for about 85% of all liver cancers[473]. Over half a million people are diagnosed with HCC annually worldwide [474]. The major risk factors for HCC are chronic hepatitis B (HBV) and C (HCV) which currently affects over 250 million and 71 million people worldwide, respectively. Other factors include smoking, chronic alcohol consumption, lifestyle choices, diabetes, obesity Figure 8-1 [475-479]. HCC is the most common cause of death in patients with liver cirrhosis [126, 480]. Approximately, 800,000 new cases occur annually [481, 482] with accompanying 700,000 deaths from the disease [473, 483]. The stage at which the disease is diagnosed is a major contributor to the 5-year survival rate of the patient. Majority of HCC cases occur alongside chronic liver injury such as hepatitis or cirrhosis [484, 485]. It is extremely important to achieve precise, accurate and early detection and characterization of liver nodules since there are curative options available to patients with early HCC which includes liver transplant, resection
and ablation therapy [486]. The management of HCC in patients is directly proportional to whether it is benign or malignant and at what stage it was detected, these factors determine the prognosis of the disease [92].

Figure 8-1 Risk factors associated with development of hepatocellular carcinoma (HCC)

8.1.1 The role of cirrhosis in HCC development

Liver cirrhosis represents the main indication for patient screening and the highest factor for HCC-related fatality [487, 488]. Regardless of the etiology of cirrhosis, up to 90% of cirrhotic livers possess the tendency to progress to HCC [480] and hence can be considered a premalignant condition for HCC [29, 126]. Based on demographics, HCC is associated with high occurrence in eastern and southern Asia, central Africa and the pacific islands, moderately high occurrence in East Africa, West Africa, southern Europe, and the Caribbean. HCC occurs the lowest in South Africa, eastern and Western Europe, western Asia, North Africa, South and North America, Australia, South central Asia, Northern Europe and Central America respectively [29, 489-491]. During a study [492] spanning a period of 14 years involving 112 patients with cirrhotic liver and ultrasound-confirmed HCC, it was observed that tumor progression was the cause of death for 63%, 69%, and 83% of the patients who died in the
first 4 years, second 4 years and third 4 years of disease monitoring respectively. An increase in HCC mortality rate [493-495] has also been recorded over time.

8.1.2  **The role of hepatitis B virus in HCC development**

The development of HCC in patients with chronic hepatitis B virus (HBV) is promoted by oncogenic viral proteins and through the integration of its genome into the genome of the host followed by the activation of oncogenes and tumor suppressor genes repression[482, 496]. Approximately 50% of the global HCC incidence is associated with chronic HBV infection [479, 497]. HCC-development in HBV patients is often dependent on sex (more prominent in male than female), age, heredity, prior exposure to aflatoxin, tobacco use an alcohol consumption, coinfection with HCV or hepatitis delta virus, high levels of HBV DNA, or infection with HBV with genotype C[479].

8.1.3  **The role of hepatitis C virus in HCC development**

Hepatitis C virus (HCV) is implicated in several liver diseases progressively from steatohepatitis, to fibrosis, cirrhosis, and hepatocellular carcinoma (HCC) [498]. With prevalence unique to different countries, an estimated 2-3 % of the world population is living with chronic HBV infection [499, 500]. HCV can promote HCC through hepatic inflammation and subsequently, fibrosis which in turn contribute to cell carcinogenesis and apoptosis resulting in cell proliferation of transformed hepatocytes. A second mechanism where HCV viral protein expression promotes mutations and malignant transformation of infected cells has also been proposed[501].

8.1.4  **Cancer-associated fibroblasts in hepatocellular carcinoma**

The surrounding environment of tumors, typically referred to as tumor microenvironment (TME) play a crucial role in tumor development, growth, and progression [502-505]. CAFs
are the most important component of the TME, promoting the carcinogenesis, proliferation and invasion of cancer cells thereby secreting various factors that are related to extracellular matrix (ECM) synthesis or remodeling, such as collagens, fibronectins and ECM-degrading proteases[506, 507], they also secrete cytokines such as C-X-C Motif Chemokine Ligand 12 (CXCL12) [508], interleukin-6 (IL-6) and vascular endothelial growth factor A (VEGFA)[509] [510]. Fibroblasts found in cancer tissues have been found to have similarity in morphology with the myofibroblasts activated during wound healing process[511]. A cross talk between cancer cells and CAFs has been reported to exist in different cancer models [512]. Hepatic stellate cells (HSCs) are major source of ECM in parenchymal liver disease including NASH[218] and HCC tumor microenvironment. In hepatocellular carcinoma, activated HSCs transform into myofibroblast-like cells to induce fibrosis in response to hepatic injury or chronic inflammation that leads to cirrhosis and HCC [513, 514].

8.1.5  **HCC classification**

Hepatocellular carcinomas can be classified as focal/massive, nodular/multifocal, and diffuse/infiltrative (Figure 8-2). Enhancement of tumor rim on delayed post-contrast images leading to a capsule-appearance is considered relative specific indication of HCC [515]. Regenerative nodules are usually a result of chronic liver disease marked by hepatic cell proliferation surrounded by fibrous tissue[516]. Regenerative nodules have the tendency to progress to premalignant and malignant lesions [517, 518] by first developing into a low grade dysplastic nodule (LGDN) then high grade dysplastic nodule (HGDN), well-differentiated HCC, and lastly moderately/poorly differentiated HCC[519]. During these processes, neoangiogenesis occurs where the normal hepatic parenchymal portal venous inflow is replaced with recruited, abnormal, arterial vessels [520-522]. Well differentiated tumors are
classified as tumors ≤ 2 cm in diameter [523] and poorly and moderately differentiated tumors are tumors >2 cm which are more susceptible to tumor neovascularity and increased arterial flow [524].

\[ \text{Figure 8-2: HCC nodules classification} \]

### 8.1.6 Challenges Associated with HCC Diagnosis

Hepatocellular carcinoma (HCC) exists as a rapidly growing tumor with poor survival rate and prognosis. Liver transplantation is the most efficient treatment for HCC. Unfortunately, due to several factors, only selected patients qualify for the liver transplant procedure. It is therefore important to achieve accurate selection of patients. Currently, selection protocols in clinical setting involve the use of ultrasonography (US), Computed tomography (CT), MRI and less frequently biopsy. It is recommended that patients with cirrhosis get evaluated for possible development of hepatocellular carcinoma using abdominal ultrasound (US) every 6 months. A
major limitation to this approach is the suboptimal sensitivity of ultrasound to HCC detection at the early stage (<2 cm) in high-risk patients with cirrhosis and obesity [525]. Less than 30% of patients with HCC benefit from early-stage diagnosis, when potentially curative treatments with favorable survival outcomes such as i.e., resection, liver transplantation, and local ablation are applicable [526]. Although CT has a better spatial resolution, MRI is superior to CT for characterizing liver masses (contrast resolution). Non-contrast MRI are divided into T1-in phase, opposed-phase which can be used to identify intracellular fat [527], T2-W images compare intensity of liver lesions to spleen intensity [528], diffusion weighted images (DWI) and apparent diffusion coefficient (ADC) images which detects restricted diffusion observed in abscess and malignancy[529]. These pulse sequences have limited detection capability and are used as supportive diagnostic modality [530-532]. For better characterization, MRI with liver-specific contrast agents such as gadoxetate disodium (eovist) and gadopenate dimeglumine (MultiHance) has been employed in HCC diagnosis [519], these agents have multiple vascular phases divided into pre-contrast, late arterial (25 seconds), portal venous (about 65 s post injection), and delayed venous (3 minutes post injection) phases [533, 534]. Despite the significant advancement in MR imaging of HCC nodules, up to 60% of small malignant nodules < 1 cm in size in the background of cirrhotic liver, are missed during MR imaging [535].
8.2 Limitation associated with current clinical contrast agents

Magnetic resonance imaging (MRI) is the imaging modality of choice in HCC diagnosis[536-538] with the capability of HCC diagnosis without need for surgical biopsy[539]. The most frequently used contrast agents for liver magnetic resonance imaging have been classified into five categories including extracellular agents, reticuloendothelial agents, blood pool agents, hepatobiliary agents and combined agents [540]. For example, Gadoxetic acid, eovist is a Gadolinium-based hepatobiliary contrast agent used for routine dynamic multiphasic studies, offering both the conventional arterial and venous phases. In addition, it has a 20-min delayed hepatobiliary phase where the hepatocyte is expected to have taken up 50 % of the administered dose. An implication of this hepatobiliary hypointensity is that nontumorous lesions and
lesions with non-functional hepatocytes, such as most HCCs adenoma, may not accumulate contrast agent leading to low signal intensity of focal lesions compared to the surrounding liver in the hepatobiliary phase and hence misdiagnosis [541-545]. In addition, early HCC has characteristic well-preserved hepatocyte function and may facilitate the uncontrolled uptake of eovist in the liver leading to high false negative results [545]. The injection rate at which eovist is administered has been found to affect the behavior of gadoxetic acid in detection of liver lesions, therefore, the injection rate may hinder the correct recognition of the contrast uptake in the arterial phase thereby limiting the established criteria based on the recognition of arterial uptake followed by wash-out[545-549]. Furthermore, wash-out recognition in such nodules may be impaired, as uptake of GA in the biliary phase may overlap with the delayed venous one. Put together, the use of clinical contrast agents in HCC has been associated with great success but still has certain limitations [550, 551]. These suggest that current clinical MRI contrast agents have associated weaknesses for the detection of tumor lesions in patients with suspected or proven HCC exposing an unmet medical need for research into more sensitive and specific MRI contrast agents for HCC detection, staging and treatment monitoring. HCC has a major impact on public health, and the detection of HCC at an early stage improves outcomes [546]. hProCA32.collagen has been developed to meet these unmet medical needs and our effort in its application in MR detection of HCC is reported in this chapter.

8.2.1 Overview of this chapter

In this chapter, we first report the detection of HCC using our previously developed rProCA32.collagen in a DEN-induced mice model at both early (12 weeks) and late-stage treated for 44 weeks (this work was done with Dr. Salarian). Next, we describe our result using the newly developed hProCA32.collagen in a western diet and carbon tetrachloride (CCL4)
mice model. hProCA32.collagen could successfully map out areas with collagen accumulation observed as Mosaic pattern around the edges of the tumors.

### 8.3 Mice model

In generation of the mice model both female and male mice with C57BL/6 genetic background were used. All animal experiments described were performed in accordance with National Institutes of Health (NIH) guide for the care and use of laboratory animals and were approved by the institutional animal care and use committee (IACUC) of Georgia State University, Emory University and the University of Georgia. The mice were housed with a 12/12-h light/dark cycle in temperature- and humidity-controlled rooms with unrestricted access to water and food. Two different mouse models were developed to induce HCC development. In the first model, 14-day-old mice were treated with a single dose of DEN (Sigma–Aldrich # N0756) dissolved in saline at a dose of 25 mg/kg body weight by intraperitoneal injection route. Mice were randomly selected and 3 weeks intervals to monitor the rate of HCC development by histological and biochemical analyses. Mice were randomly distributed in various groups having equal males and females. At 12 weeks, the early-stage HCC mice were imaged by MRI and late stage were imaged at 44 weeks. Following euthanizing, mice livers were isolated (Figure 8-4), characterized and the numbers of visible tumors on the liver surface were identified and processed for histological analysis by formalin fixation and paraffin embedding. Sections were stained and examined microscopically.
Figure 8-4: Gross liver images of mice and MR Imaging and euthanasia. Images show normal, smooth liver outlook for healthy, control mice and development of small and large tumor nodules at 12 weeks and 44 weeks respectively.

8.3.1 **MR imaging of DEN-induced HCC after 12 weeks using rProCA32.collagen**

For this section, rProCA32.collagen was first evaluated for HCC detection in a DEN-induced mice model. At 12 weeks post DEN injection, mice were injected with rProCA32.collagen through the tail-vein. T2-weighted, T1-weighted, T1 map inversion recovery and T2 maps were collected pre, 3 h and 24 h after injection of rProCA32.collagen. As shown in Figure 8-5 and Figure 8-6, after injection of rProCA32.collagen at 3 h time point, there is enhancement in both the normal mice and HCC mice liver regions. At 24 h time point however, the normal mice had its enhancement return to baseline while the DEN-injected mice still had a significant enhancement at 24 h time point possibly due to early-stage fibrosis.
After injection of rProCA32.collagen shown an uptake of rProCA32.collagen in mice liver after injection, this enhancement was reduced back to baseline (pre) at 24 h time point.

Figure 8-6 further shows the uptake of rProCA32.collagen to separate liver imaging from tumor uptake. Compared to 2 tumor lesions detected overall across all MR slices, rProCA32.collagen at the 24 h time point detected up to 6 lesions (observed by eye) in one slice (Figure 8-7) using T2-W images. The SNR value for the tumor at 24 h time point also increased by 80% from 6.48 ± 0.34 to 11.66 ± 0.56.
Figure 8-6: T2-W MR images of 12-weeks HCC mice before, 3 h and 24 h after rProCA32.collagen injection shows increase liver enhancement at 3 h time point and more tumor
**Figure 8-7: T2-W MR images of early-stage DEN-induced HCC**

A. Before, 3 h and 24 h after rProCA32.collagen1 injection showing uptake at 3 h in mostly liver region and in tumor region at 24 h time point. B, signal to noise ratio estimation before, 3 h and 24 h after injection of rProCA32.collagen1, C., The number of tumors observed in displayed slices of MR images.

In T1-W images however, no major distinction can be observed between the tumor signal and liver signal at the 24 h time point. The same observation exists however for both T1 and T2 w images at 3 h time point (Figure 8-8 and Figure 8-9).
Figure 8-8: T1-W MR images of 12-weeks HCC mice
Before, 3 h and 24 h after rProCA32.collagen injection shows increase liver enhancement at 3 h time point and some tumor enhancement at 24 h time point.
Figure 8-9: T1-W images of HCC DEN-induced mice model.
A., Before, 3 h and 24 h after rProCA32.collagen injection shows increased liver enhancement at 3 h time point but not 24 h time point.

On histological assessment, the tumor observed correlated linearly with the tumor enhanced on the T2-W MR images as shown in Figure 8-10
Figure 8-10: Histological correlation of early HCC with MRI lesions shows a linear correction with an R value of over 97%. Each tumor observed on MR images matches with the location and size of tumor in histology (tumors are color coded).

8.3.2 MR imaging of DEN-induced HCC after 44 weeks.

At 44 weeks post DEN injection, mice were injected with rProCA32.collagen through the tail-vein. T2- weighted, T1-weighted, T1 map inversion recovery and T2 maps were collected pre, 3 h and 24 h after injection of rProCA32.collagen. As shown in Figure 8-11 and Figure 8-12, after injection of rProCA32.collagen at both 3 h and 24 h time points, there is enhancement in both the tumor region and cirrhotic regions across several slices.
Figure 8-11: MRI slices of the entire liver
Shows significant uptake of rProCA32.collagen at 3 h post injection of rProCA32.collagen across all 16 slices.

Furthermore, by analyzing several slices, it can be seen in slice 1 that both cirrhosis and HCC tumor exist with no significant distinction. In slice 5, majority of the cirrhotic regions are observed at 3 h time point compared to HCC at 24 h time point. While slice 8 mostly shows cirrhosis Figure 8-12.

Figure 8-12: Analysis of T1-W MRI slices with similar anatomic properties
Before, 3 h and 24 h after injection of rProCA32.collagen. The intensity increased after injection of rProCA32.collagen with peak enhancement at the 3 h time point. By using the grayscale images, hProCA32.collagen can show HCC regions on cirrhotic liver (red circle).

Consistent with what was observed with 12-weeks treated mice, T1-W MRI had a significant increase at 3 h time point and 24 h time point. The area designated at tumor area was not distinguishable from cirrhotic region.

Figure 8-13: Signal to noise ratio of tumor (circles) vs liver of HCC/Cirrhosis mice model showing increased enhancement at 3 h time point for both liver and tumor quantified by signal to noise ratio in the bottom figure.
On T2-W image however, the tumor was distinguishable at the 24 h time point compared 3 h time point. The tumor regions (H and red circle) were correlated with Sirius stain with collagen enhancement (Figure 8-14).

**Figure 8-14:** Correlational analysis of MR images with gross and histological Sirius red stain with collagen accumulation. The enhancement pattern of collagen observed in MRI matched with the area in gross image (labelled H) and the areas of collagen accumulation in Sirius red with corresponding tumor measurements.

8.3.3 **Development of Diet /chemical -induced HCC**

In this model, 9-week-old mice were fed with western diet (WD) containing 21.1% fat, 41% Sucrose, and 1.25% Cholesterol by weight (Teklad diets, TD. 120528) and a high sugar solution 23.1 g/L d-fructose (Sigma-Aldrich, F0127) and 18.9 g/L d-glucose (Sigma-Aldrich, G8270). CCl₄ (Sigma-Aldrich, 289116-100ML) at the dose of 0.2 µL (0.32 µg)/g of body weight and injected intra-peritoneally once per week, which is much lower than the dose that is usually given for fibrosis induction with CCl₄ alone. These mice were taken for MR imaging after 32 weeks and euthanized by exsanguination after ketamine and xylazine anesthesia.
Liver and serum samples were collected and processed for histological, serological and gene expression analysis.

8.3.4 **MR imaging of Diet/chemical-induced HCC**

T1-Weighted, T2-Weighted, T1-map and T2-mapping sequences were collected for mice with diet/chemical-induced HCC. As shown in Figure 8-15 and Figure 8-16, T2-Weighted MRI showed hypo intensity of the liver at 24 h (post) injection of hProCA32.collagen.

*Figure 8-15: Left to right, T1-Weighted MR image slices of HCC mice liver*

More lesions across slices were detected after proCA32.collagen injection (72 lesions) similar to ex vivo (79) imaging on T1-W images comparable to pre time point (15).
Figure 8-16: Characterization of diet-induced HCC by MRI and histology.
A. T1-T2-W MR images of diet-induced HCC B. ex vivo imaging of HCC mice model with histology correlation C. T2-W signal to noise ratio quantification shows a decrease in liver signal at 3 h and 24 h time points compared to pre time point, D. In contrast, the tumor SNR on T1-W images increased significantly at both 3 h and 24 h post injection time points. E. Quantification of tumor detected by non-contrast enhanced MRI vs hProCA32.collagen injected and ex vivo imaging.

The SNR value for tumors on T1-W MRI increased from 16.95 ± 2.75 pre contrast to 25.57 ± 2.92, 46.55 ± 3.72, and 20.08 ± 3.32 for 3h, 24 h and 48 h time points, respectively (Figure 8-17).
Figure 8-17: MR imaging and characterization of HCC liver versus tumor lesions. A. Grayscale and color images of diet-induced HCC mice model, B. signal to noise ratio quantification of mice liver and C. mice nodules. D. ex vivo MRI.

Figure 8-18: T1-W MRI and quantification of HCC with clinical contrast agent Eovist. A., T1-W images of Eovist-injected MR images before, 1 h, 3 h and 24 h post injection. B. signal to noise ratio of MR images of liver and C. tumor nodules.
As shown in Figure 8-18, injection of Eovist lack the capability for detection of lesions relative to the pre time point. At the 3 h time point, there is tumor enhancement with SNR increase from 17.11 ± 0.91 to 18.15 ± 0.95, 20.6 ± 0.35, and 15 ± 0.50 for 1 h, 3 h and 24 h respectively.

Figure 8-19: MRI correlation of Rim-enhanced tumor nodule with HCC and collagen-specific staining

Tumor-associated fibrosis in HCC has highly heterogeneous patterns which include fibrous capsules [552]. Collagen capsule around HCC nodule is a classic observation in HCC nodules, as shown in Figure 8-19, the HCC tumor nodules displayed an increased intensity at 3 h post hProCA32.collagen injection. At the 3 h time point, hepatocytes uptake of hProCA32.collagen is observed while 24 h time point showed retained enhancement at the edge of the lesion (blue arrows) The presence of nodule was confirmed by 70kDa Heat shock protein (HSP-70) and Glypican-3 staining specific for HCC. A positive staining using Sirius red is an indication of collagen type overexpression. Since hProCA32.collagen has higher specificity for type I collagen, it is important to specifically determine which collagen type was stained on Sirius
red. To do this, Masson trichrome staining was done for specificity for type I while reticulin staining was stained for type III collagen (Figure 8-19B).

8.4 Conclusion

rProCA32.collagen1 was successful in the detection of DEN-induced rProCA32.collagen1 enhanced both cirrhotic and tumor areas in the model as confirmed by several pulse sequences. Due to both high $r_1$ and $r_2$ property, T1- and T2- weighted and T1 and T2 maps can be used to increase the level confidence for the detection and avoid observing artifacts in the MRI. To ascertain hProCA32.collagen capability in HCC detection, a western diet/CCl$_4$-induced hepatocellular carcinoma (HCC) mouse model reported to have collagen accumulation in the HCC nodules was utilized. hProCA32.collagen differentiated areas of underlying cirrhosis, from normal liver and tumor using its high $r_1$ and $r_2$ properties for T1 and T2 weighted images. Collagen arrangement pattern with specific collagen enhancement at the tumor surrounding was detected by hProCA32.collagen forming a rim at the tumor boundary at 24 h time point compared to total enhancement at 3 h time point post injection in WD/CCl$_4$ induced HCC mouse model. This result matches histological results using collagen specific Masson Trichrome, Sirius red and reticulin staining as well as HCC specific staining such as Glypican-3 and HSP-70. More importantly, the delayed collagen enhancement at the 24 h time point corresponds directly with the collagen enhancement observed for Masson trichrome staining. This capability to specifically map out collagen patterns and arrangement in HCC tumor can
lead to a major shift in the clinical imaging making hProCA32.collagen a very attractive imaging contrast agent in disease treatment and monitoring.

9 CONCLUSIONS, SIGNIFICANCE AND FUTURE DIRECTION

The use of molecular imaging combined with the use of a sensitive, specific, and precise MRI contrast agent such as hProCA32.collagen will result in accurate staging and detection of different diseases. As discussed in Chapter 1, there is a major gap between molecular imaging and the availability of sensitive and specific MRI contrast agents with high relaxivity and metal binding selectivity.

In chapter 3, we developed a human protein-based MRI contrast agent (hProCA32.Cys) using human parvalbumin. A Cys residue was introduced to achieve site-specific PEGylation. The relaxivity value obtained for PEGylated hProCA32.cys was 24 mmol^{-1}.S^{-1} and 36 mmol^{-1}.S^{-1} for longitudinal ($r_1$) and transverse ($r_2$) relaxation respectively. The values are 27% and 19% lower than the values reported for rProCA32.P40, the PEGylated rat analogue but at least 5 to 8-fold high than clinically approved contrast agent. The metal binding affinity of hProCA32.cys was obtained to be 0.44nM, 7.2µM, $8.2 \times 10^{-22}$M and $2 \times 10^{-22}$M for Ca$^{2+}$, Zn$^{2+}$, Gd$^{3+}$ and Tb$^{3+}$ respectively. These values are comparable with those obtained for
rProCA32.P40 [102] showing high stability for Gd$^{3+}$ and a $10^{11}$-fold greater selectivity for Gd$^{3+}$ over Zn$^{2+}$ compared with existing contrast agents. Preliminary T-Cell Proliferation assay data have been able to show that with hProCA32.Cys developed from human protein scaffold, no significant activation of T-cell was observed. This may give an indication of possibly negligible immunogenicity. Further experiments with direct injection of human subjections are required to confirm its low immunogenicity. Computer simulation using MHC class II alleles of diverse origin by Dr. Zuben also showed that this protein is like to have immunogenicity that is limited to a specific portion of the protein sequence such as residues 30 to 37 in the protein scaffold.

The study of the *in vitro* transmetallation process of Gd-complexes carried out in this work clearly confirms hPoCA32.cys possess a high kinetic and thermodynamic inertness up to at least 4320 minutes retaining its longitudinal relaxation up to ~99%, similar to the stable macrocyclic agents such as Prohance® and Gadovist®. Further optimization with different ratio and formulation conditions are needed to ensure highest metal binding affinity stability, specificity, and thermodynamic inertness.

The developed hProCA32.cys facilitated the detection of NSCLC in a KP generated mouse model for both tumor lesion <0.2mm and advanced lung cancer affecting the entire right lobe without significantly affecting the middle lobe. This was specifically enhanced by hProCA32.cys with up to 71% verified increase in signal to noise ratio. MRI Images shows a promising contrast enhancement post administration of hProCA32.cys contrast and need to be further explored for statistical relevance (for the current report N=1).
In chapter 4, five potential hProCA43.collagen variants were created based on previously designed hProCA32 scaffold with or without cysteine amino acid residue were carefully selected. hProCA32.collagen1, designed with no additional cysteine residue had the highest collagen binding affinity (1.85 µM) and a $r_1$ and $r_2$ relaxation rates of 31 mmol$^{-1}$ S$^{-1}$ and 50 mmol$^{-1}$ S$^{-1}$ respectively comparable to the previously reported values for rProCA32.collagen1. The protein yield obtained was 35.2 mg/L. Conversely, hProCA32.collagen1C displayed a higher protein yield (40.2 mg/L), 38.7% and 32% increase in $r_1$ and $r_2$ relaxation rates compared to hProCA32.collagen1. However, the collagen binding affinity is 8.1 and 10.6 times lower than the values obtained for hProCA32.collagen1 and rProCA32.collagen1 respectively. For the hProCA32.collagen1.LKK and hProCA32.collagen1LRR variants with a different peptide composition in the collagen targeting moiety, a protein yield at least 24.4% higher than the KKWHCYTYFPHHYCVYG-based hProCA32.collagen1 was observed. Due to this increased yield, hProCA32.collagen1LKK with >50% higher collagen binding affinity than the hProCA32.collagen1LRR analogue was considered for further optimization and a possible improved collagen-binding property. Due to lack of improved collagen binding affinity and low dynamic range observed for the calcium binding plot, hProCA32.collagen1LKK was not push further for development of the human collagen contrast agent. Since the long-term goal in the use of the developed contrast agent is dependent on both relaxivity for highly sensitive MR imaging and collagen affinity for biomarker specificity, hProCA32.collagen1 (hereinafter referred to as hProCA32.collagen) was selected for further optimization, characterization and application in various diseases discussed in chapters 5 to 8.
As discussed in chapter 5, clone selection and optimization of purification conditions for hProCA32.collagen was carried out to generate a single component (monomer) final product. To obtain the purified hProCA32.collagen in its monomeric state due to the observed presence of dimerized hProCA32.collagen as well as higher oligomer bands. Using a combination of cysteine and 1% βME at pH 7.1 to 7.4, a significant reduction of dimer and higher oligomer bands were observed. The new vector possessed a higher copy number and a higher protein yield ~350 ± 53 mg/L of protein about 10 times higher than hProCA32.collagen in Pet22b vector. N-terminal PEGylation of hProCA32.collagen in a 1:1 ratio resulted in 65% PEGylation yield and needs to be further optimized. N-terminal PEGylation resulted in a 65% yield for 1:1 PEGylation ratio and can be farther improved. The improved hProCA32.collagen exhibits 8- to 10-fold improvement in both r₁ and r₂ compared to all the current clinically used contrast agents at both 1.4 T (r₁ = 30 ± 0.6 mM⁻¹s⁻¹ per Gd³⁺ (60 ± 1.2 mM⁻¹s⁻¹ per participle) and r₂ 52 ± 0.9 mM⁻¹s⁻¹ per Gd³⁺ (104 ± 1.8 mM⁻¹s⁻¹ per particle) and 7.0 T (r₁ = 18.5 ± 1.2 mM⁻¹s⁻¹ per Gd³⁺ (37.0 ± 2.4 mM⁻¹s⁻¹ per participle) and r₂ 105.6 ± 2.0 mM⁻¹s⁻¹ per Gd³⁺ (211.2 ± 4.0 mM⁻¹s⁻¹ per particle) at 37°C. hProCA32.collagen exhibits 10⁸−10¹⁵-fold greater metal selectivity for Gd³⁺ over physiological metals such as Zn²⁺ and Ca²⁺ compared to Dotarem, a clinically approved contrast agent. Clinical toxicity test of hProCA32.collagen showed values in the normal range of each test evaluated. Furthermore, injection of hProCA32.collagen exhibited no noticeable tissue toxicity.

hProCA32.collagen displayed a high specificity for collagen type 1 over types 3 and 4. This is exciting in two phases, on one hand, collagen type one being the most abundant protein can be easily detected with high specificity making it a potential biomarker for disease monitoring in collagen vascular diseases. On the other hand, collagen type 3 and 4 can be detected in diseases
with overexpression of these types of collagens such as autosomal dominant polycystic kidney disease (ADPKD) [346]. hProCA32.collagen like our previously reported rProCA32.collagen enabled early detection of NAFLD. Using a diabetes/obesity model, (discussed in detail in chapter 7) hProCA32.collagen detected NASH stage F1 as well as late-stage cirrhosis with SNR values directly indicative of collagen accumulation as estimated by collagen proportionate area (CPA) analysis. Most importantly, the use of a human protein in place of rat analogue used for the development of rProCA32.collagen [102, 108, 155, 157, 159, 160, 187, 210] significantly reduces the immunogenic potential of our developed contrast agent while maintaining sensitivity and specificity. As a future direction, it will be beneficial to explore different excipients that will keep the protein in its formulation form throughout its shelf life. Other conditions that can further be optimized are pH and storage temperature.

In Chapter 6, hProCA32.collagen was successfully utilized for first time to achieve robust early detection of idiopathic pulmonary fibrosis (IPF). First, the IPF model was developed by intraperitoneal injection of Bleomycin twice weekly to progressively develop IPF with features closely resembling human IPF features at sub pleural and paraseptal regions in addition to cystic honeycombing and traction bronchiectasis. Secondly, using RARE T1 and T2 weighted pulse sequences, we show that hProCA32.collagen enabled conventional T1 and T2 weighted MRI was successful in the specific detection of IPF at the 3 h targeting time point in both early and late-stage diseases with notable fibrotic network. Additionally, the use of the optimized 3D-UTE pulse sequence, hProCA32.collagen was able to delineate the histopathological features of human IPF such as traction bronchiectasis, cystic honeycombing with specific presence at the apico-basal region (sub pleural) of the lung with associated heterogeneity. The SNR-AUC values obtained by MRI increased with increasing disease progression with linear
correlation by histological characterization of several markers such as collagen area and α-SMA. The specificity with which hProCA32.collagen binds to collagen expressed in IPF lung matches with the areas with collagen accumulation and pulmonary nodules in Sirius red staining.

Third, pixel by pixel histogram analysis showed heterogeneous hProCA32.collagen uptake in areas corresponding to high heterogenous collagen distribution associated with human IPF compared to normal mice and late-stage mice injected with clinical contrast agent Gadovist. This observation was seen by dual MR pulse sequence (RARE and 3D UTE).

Fourth, collagen overexpression in the lung bronchi was detected by hProCA32.collagen in a nicotine-induced COPD mice model with similar accumulation pattern predominantly in the airways and bronchioles of the mice lungs as observed for histological collagen staining.

Fifth, the distribution dynamics and kinetics of hProcA32.collagen in different organ shows its ability to detect collagen in specific organs based on distribution times. For example, the peak enhancement for lung fibrotic mice with high collagen content was 1 h compared to 3 h observed in liver of the same mice where no significant damage was done to the liver. Conversely, in liver fibrosis imaging by rProCA32.collagen and hProCA32.collagen (chapter 7) the peak enhancement for late-stage liver fibrosis model with high collagen content is at 3 h time point while the targeting time point for both early stage with minimal collagen accumulation and late stage was 24 h due to the biodistribution pattern of the collagen-targeted contrast agent. This is important in clinical translation enabling the use of hProCA32.collagen in time-dependent detection and staging of diseases in various organs.

More importantly, ECM plays a critical role in carcinogenesis and cancer progression. Collagen is the major component of the tumor microenvironment and participates in cancer
fibrosis with type I collagen secreted by fibroblasts during tumor growth and epithelial-mesenchymal transition \([419-421]\). Both patients with COPD and IPF have been associated with increased risk of lung cancer \([422-427]\). Non-small cell lung cancer accounting for >80% of all lung cancer has been reported to secrete type I collagen \([428]\) and the use of collagen-targeted hProCA32.collagen MRI can significantly improve image guided treated of lung cancer by precise detection of collagen secreted by the lung cancer cells and the tumor microenvironment. The specificity in the uptake of hProCA32.collagen to collagen accumulation depicted in two different mice models reported in this chapter further enforces its clinical relevance for separating different disease pattern and early detection of lung diseases with non-invasive MRI.

In Chapter 7, we applied hProCA32.collagen in early detection of nonalcoholic steatohepatitis using both its \(r_1\) and \(r_2\) properties. We show that hProCA32.collagen has strong capability in detection of fibrosis using liver specific comparative gene identification-58 knock out (CGI58-KO) mice model fed with western diet and 4.2% fructose water for 42 weeks as the most advanced stage (Ishak score 6/F4) and histone deacetylase complex (HDAC 1) expression diabetes mice model fed with NASH diet and 4.2% fructose water for at least 24 weeks for early stage (Ishak 1/F1). The peak enhancement for early and moderate stage HDAC1 mice model and wild type CGI-58 model (CGI-58 WT) was observed at 24 hours post hProCA32.collagen injection time point while western diet/chemical induced fibrotic mice model and CGI-58 KO mice models with advanced and late-stage fibrosis peaked at 3 h time point with SNR and \(R_1\) values at 24 h significantly higher than early stage 24 h time point. This result is consistent with our previously reported finding rProCA32.collagen1 with ability
to distinguish early stage (peak enhancement at 24 h) versus late stage (peak enhancement at 3 h and prolonged enhancement up to 24 h).

hProCA32.collagen possesses features making it suitable for molecular imaging with tendency for translation and early disease staging, treatment monitoring and interventions compared to other clinical contrast agents in precision and early detection of human diseases.

In Chapter 9, we first tried rProCA32.collagen1 in the detection of DEN-induced rProCA32.collagen1 to enhanced both cirrhotic and tumor areas in the model as confirmed by several pulse sequences. Due to both high r1 and r2 property, T1- and T2- weighted and T1 and T2 maps can be used to increase the level confidence for the detection and avoid observing artifacts in the MRI. To ascertain hProCA32.collagen capability in HCC detection, a western diet/CCl4-induced hepatocellular carcinoma (HCC) mouse model reported to have collagen accumulation in the HCC nodules was utilized. hProCA32.collagen differentiated areas of underlying cirrhosis, from normal liver and tumor using its high r1 and r2 properties for T1 and T2 weighted images. Collagen arrangement pattern with specific collagen enhancement at the tumor surrounding was detected by hProCA32.collagen forming a rim at the tumor boundary at 24 h time point compared to total enhancement at 3 h time point post injection in WD/CCl4 induced HCC mouse model. This result matches histological results using collagen specific Masson Trichrome, Sirius red and reticulin staining as well as HCC specific staining such as Glypican-3 and HSP-70. More importantly, the delayed collagen enhancement at the 24 h time point corresponds directly with the collagen enhancement observed for Masson trichrome staining. This capability to specifically map out collagen patterns and arrangement in HCC
tumor can lead to a major shift in the clinical imaging making hProCA32.collagen a very attractive imaging contrast agent in disease treatment and monitoring.

The capability of hProCA32.collagen to specifically map out collagen patterns and arrangement in different disease models will lead to a major shift in the clinical imaging making hProCA32.collagen a very attractive imaging contrast agent in disease treatment and monitoring.

The development of hProCA32.collagen is projected to bridge this gap by providing targeted accurate and biomarker specific MR imaging for disease staging and detection.

REFERENCES


57. Drew, A.M.Z., *T1 values (1.5T)*. Radiopedia.


158. Salarian, M., Development of Novel Collagen-targeted Protein based MRI Contrast Agent for Imaging of Chronic Liver and Heart Diseases, in Chemistry. 2018, Georgia State University: Scholarworks, Georgia State University. p. 361.


164. Xue, S., Design of Novel Protein-based MRI Contrast Agernets with High
Relaxivity and Stability for Biomedical Imaging in biology. 2013, Georgia State University: Scholar works.


