Characterizing reactive glutamines in fibrinogen and elucidating factor XIII substrate specificity.

Kelly Njine Mouapi
University of Louisville

Follow this and additional works at: https://ir.library.louisville.edu/etd
Part of the Biochemistry Commons, and the Laboratory and Basic Science Research Commons

Recommended Citation
https://doi.org/10.18297/etd/2770
CHARACTERIZING REACTIVE GLUTAMINES IN FIBRINOGEN AND
ELUCIDATING FACTOR XIII SUBSTRATE SPECIFICITY

By
Kelly Njine Mouapi
B.A., Berea College, 2012
M.S., University of Louisville, 2015

A Dissertation
Submitted to the Faculty of the
College of Arts and Sciences of the University of Louisville
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy in Chemistry

Department of Chemistry
University of Louisville
Louisville, Kentucky

August 2017
CHARACTERIZING REACTIVE GLUTAMINES IN FIBRINOGEN AND ELUCIDATING FACTOR XIII SUBSTRATE SPECIFICITY

By

Kelly Njine Mouapi
B.A., Berea College, 2012
M.Sc., University of Louisville, 2015

A Dissertation Approved on

June 28th, 2017

by the following Dissertation Committee:

______________________________________________________________________________
Dissertation Chair
Dr. Muriel Maurer

______________________________________________________________________________
Dr. Michael Merchant

______________________________________________________________________________
Dr. Eugene Mueller

______________________________________________________________________________
Dr. Richard Wittebort
DEDICATION

This dissertation is dedicated to my grandmother Mrs Agnes Wubnyonga Fokong who nurtured my love for science at a tender age, sacrificed a lot so that I could obtain an education and excel as a student, and encouraged me to pursue my dreams.

To
My deceased aunt Ms Mercy Fokong Nukanjam whose death from a sudden heart attack in 2004 fueled my passion for cardiovascular disease research.

To
My family especially my parents Mr Hilaire Tsozock and Mrs Hope Nunyonga Tsozock, my aunt Flora Bidgemia, Ms Kathy Carr and Mr David Moore who equipped me with valuable life skills.

To

and

To my love and life partner Mr. Jonathan Davis. Words cannot describe the impact you have had in my personal and professional growth. You believed in me. You cheered me on all along the way. Thank you!
ACKNOWLEDGMENTS

As an old adage goes “It takes a village to raise a child” ~ African Proverb. I would like to thank the village that supported me through this journey with words of encouragement, resources, feedback and ideas that led to this dissertation.

I acknowledge my Ph.D. advisor and mentor Dr. Muriel Maurer for making this journey worthwhile. I am very grateful for your guidance through my graduate career and for mentoring me each step of the way. Your knowledge and expertise in the field of blood coagulation sparked my interests in this research. Your commitment to this research topic encouraged me to be just as committed, and your mentorship fostered my growth as a young investigator. You encouraged and helped me attend international conferences at an early stage during my career, which helped me to become a part of the scientific community in thrombosis. Thank you for supporting me, my passion for this research, and for sharing existing opportunities for professional growth with me.

A special thank you to my dissertation committee members for all your guidance, advice, training, and help. Thank you Dr. Eugene Mueller for making Enzymology so intriguing through your lectures and discussions. I learned a lot from your teaching style and aspire to impart other students like you did for me. Thank you Dr. Richard Wittebort for sharing your expertise in Protein NMR and for always being available to help. Thank you Dr. Michael Merchant for your insight throughout this research and especially for contributing to the progress of this dissertation with the LC-MS approach.

I would like to acknowledge a multitude of people who have trained me voluntarily throughout my time as a graduate student. I would like to thank Mr. Daniel Wilkey for his time, commitment to my experiments, and patience in teaching me how to use the Skyline software. Dr. Nazimuddin Nazimuddin for his contributions and training with plasmid design, protein expression, and also his valuable friendship. Thank you Dr. T Michael Sabo for your help with IPAP experiments and the 800 MHz NMR and Dr. Neal Stolowich for your help working with the 700 MHz NMR. I would like to acknowledge Mr. Jason Blumenkemp (JBI scientific) for his training on MALDI TOF hardware and his time. I would like to especially thank Mr. Steve Riley. I learned a lot from you and your commitment to making sure things worked.

I am extremely grateful to Dr. Leonid Medved, Dr. Jeffrey Keillor, and Dr. Kerrie Smith for DNA samples and advice. I extend a special thanks to Dr. Helen Philippou and Dr. Robert Ariens for sharing their expertise with me. Thank you
Dr. Aleeta Powe for your mentoring and friendship. I would like to thank past members of the Maurer Lab: Dr. Marina Malovichko, Jacob Bell, Blake Lynch, Camela Riposo, Sara Zink, Chad Stephens and ENSCM (Montpellier) exchange student Lucille Wagner.

I would like to thank all my colleagues and new friends here at UofL. Dr. Elisha Otome and Dr. Sakatali Gori for guiding me. Ramya Billur, thank you for your friendship, advice, help, and support. Boris Anokhin thank you for all you do. I would like to acknowledge Mumiye Ogunwale for his friendship and being each other's keeper. To my friends and cohort members Wuyu Zhang, Rahul Jain, Andrew Haddad and Sara Biladeau. Thank you for your support. Many thanks to Ms Nalley Sherry and Ms Sabrina Haug for their hardwork with graduate students, and to the CGSA for organizing the Nobel Laureate Derby Lectures.

I would like to especially thank the Public Safety and escort service team at UofL for this ininitiative. I could work late nights in the lab knowing that I had a safe and free ride home. Thank you! To the Minority Association of Graduate Students (M.A.G.S), thank you. To some very special mentors Ms Prafula Sheth and Pastor Mr. & Mrs Akenroye, thank you all. To my church family at RCCG Agape house, you are appreciated.

To my family, especially my grandmother who helped me to discover my love for science, thank you for your words of encouragements. I acknowledge Berea College for amazing opportunities to explore scientific research as a career option. To all faculty and staff of the Department of Chemistry.

I would like to acknowledge my love and life partner Mr. Jonathan Davis for his support, advice, guidance, and love. Words are not enough to express my gratitude. Thank you!
ABSTRACT

CHARACTERIZING REACTIVE GLUTAMINES IN FIBRINOGEN AND
ELUCIDATING FACTOR XIII SUBSTRATE SPECIFICITY

Kelly Njine Mouapi
June 28, 2017

Fibrinogen is the most abundant protein involved in blood coagulation and has been associated with many pathological implications in cardiovascular disease. At the final stages of blood clot formation, the transglutaminase Factor XIIIa introduces γ-glutamyl-ε-lysyl covalent bonds between reactive glutamines and lysines in fibrin, which results in a tighter clot network that is resistant to fibrinolysis. Factor XIIIa crosslinks specific reactive glutamines on fibrinogen, selecting more reactive glutamines in the αC region of fibrinogen than any other chain. Although crosslinking pairs in the αC region have been identified, little is known about the extent of crosslinking and the role played by each reactive glutamine.

For the first time, we have ranked three reactive glutamines crosslinked under physiological conditions. A combination of MALDI-TOF mass spectrometry, LC-MS, and 2D $^{15}$N-$^1$H HSQC NMR studies was used on Fibrinogen αC(233-425) to understand Factor XIIIa’s substrate specificity. Factor
XIIIa’s specificity for each reactive glutamine in αC(233-425) was monitored by the enzyme's ability to cross-link a lysine mimic, glycine ethyl ester (GEE), to the reactive glutamines (Q237, Q328, Q366). We showed that Factor XIIIa crosslinks these three glutamines to GEE in the order Q237 >> Q366 ≈ Q328.

Fibrinogen αC(233-425) variants were generated in which each reactive glutamine (Q) was replaced with an inactive asparagine (N). We demonstrated that Factor XIIIa still crosslinks each reactive glutamine independently and is more selective in the absence of the most reactive glutamine Q237. To examine the role of the putative Factor XIII binding site on αC(233-425), we performed experiments in which an important amino acid E396 was replaced with an alanine (Fbg αCE396A). Our results showed that both plasma Factor XIII A2B2 and recombinant Factor XIII A2 can crosslink all three reactive glutamines in Fbg αCE396A to a similar extent as with the WT.

For structural characterization, αC(233-425) was expressed in 15N or 15N/13C-enriched media. 2D HSQC NMR experiments confirmed that the 15N-labeled αC(233-425) is mostly intrinsically disordered. Residues surrounding Q366 and Factor XIII's binding site αC(389-402) showed a higher propensity for order. 2D NMR experiments suggest the possibility of pre-structured motifs (PreSMOS) within Fibrinogen αC(233-425).
# TABLE OF CONTENTS

ACKNOWLEDGMENTS..............................................................................................................IV

ABSTRACT ..........................................................................................................................VI

LIST OF FIGURES................................................................................................................IX

CHAPTER

1. INTRODUCTION ............................................................................................................... 1

2. MATERIALS AND METHODS .......................................................................................... 25

3. RANKING REACTIVE GLUTAMINES IN THE FIBRINOGEN αC REGION THAT ARE
   TARGETED BY THE BLOOD COAGULANT FACTOR XIII .............................................. 50

4. INVESTIGATING THE ROLE OF KEY MUTATIONS IN FIBRINOGEN αC (233 - 425)
   IN CONTROLLING FACTOR XIII'S ABILITY TO CROSSLINK REACTIVE
   GLUTAMINES (Q237, Q328, Q366) .................................................................................. 70

5. STRUCTURAL CHARACTERIZATION OF FIBRINOGEN αC (233 - 425) USING NMR
   SPECTROSCOPY ............................................................................................................... 97

6. RESEARCH SUMMARY AND FUTURE DIRECTION ....................................................... 119

REFERENCES ......................................................................................................................133

APPENDIX ..........................................................................................................................146

CURRICULUM VITAE ...........................................................................................................157
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The Blood Coagulation Cascade</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>Fibrinogen, Factor XIII and Thrombin.</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>Crystal Structure of Fibrinogen</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>Fibrinogen αC region</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>Fibrin Polymerization</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>Crystal Structure of Factor XIII</td>
<td>15</td>
</tr>
<tr>
<td>7</td>
<td>Calcium activated Factor XIII crystal structure</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>Factor XIII Transglutaminase reaction mechanism</td>
<td>18</td>
</tr>
<tr>
<td>9</td>
<td>Transamidation reaction catalyzed by Factor XIII</td>
<td>19</td>
</tr>
<tr>
<td>10</td>
<td>Amino acid sequence of Fibrinogen αC region</td>
<td>22</td>
</tr>
<tr>
<td>11</td>
<td>Protein Expression with Akta Prime and GST-Trap columns</td>
<td>27</td>
</tr>
<tr>
<td>12</td>
<td>SDS PAGE Analysis of Fibrinogen αC (233-425)</td>
<td>28</td>
</tr>
<tr>
<td>13</td>
<td>Western Blot Analysis of Fibrinogen αC (233-425)</td>
<td>29</td>
</tr>
<tr>
<td>14</td>
<td>Chymotrypsin and GluC digest of Fibrinogen αC (233-425)</td>
<td>32</td>
</tr>
<tr>
<td>15</td>
<td>Scheme for MALDI-TOF MS Kinetic Assay</td>
<td>35</td>
</tr>
<tr>
<td>16</td>
<td>Representative data showing reactant and product peaks</td>
<td>36</td>
</tr>
<tr>
<td>17</td>
<td>MALDI-TOF MS Kinetic Assay in Fibrinogen γ(148-411)</td>
<td>38</td>
</tr>
<tr>
<td>18</td>
<td>Comparing MALDI-TOF MS and LC-MS assay ranking</td>
<td>41</td>
</tr>
<tr>
<td>19</td>
<td>Pulse sequence for typical HSQC experiment</td>
<td>43</td>
</tr>
<tr>
<td>20</td>
<td>2D HSQC NMR Kinetic Assay in Fibrinogen αC (233-425)</td>
<td>45</td>
</tr>
<tr>
<td>21</td>
<td>2D HSQC NMR Kinetic Assay in Fibrinogen γ(148-411)</td>
<td>46</td>
</tr>
<tr>
<td>22</td>
<td>Representative 2D HSQC NMR spectrum for $^{15}$N-labeled Fibrinogen αC (233-425)</td>
<td>48</td>
</tr>
<tr>
<td>23</td>
<td>Proteolytic digest of Fibrinogen αC (233-425)</td>
<td>58</td>
</tr>
<tr>
<td>24</td>
<td>Representative MALDI-TOF MS spectra with GluC and Chymotrypsin</td>
<td>59</td>
</tr>
<tr>
<td>25</td>
<td>Ranking Reactive Glutamines in Fibrinogen αC (233-425)</td>
<td>61</td>
</tr>
</tbody>
</table>
Figure 26. Complementary 2D HSQC NMR spectra for WT and Q→N variants

Figure 27. 2D HSQC NMR spectra for Fibrinogen αC (233-425) E396A

Figure 28. Reactive Glutamine Q237 crosslinked by Factor XIIIa in Fibrinogen αC (233-425) WT vs E396A

Figure 29. Comparing all reactive glutamines crosslinked in Fibrinogen αC (233-425) WT vs E396A

Figure 30. Monitoring Q237 crosslinked by Factor XIII A2B2 in Fbg αC (233-425) WT vs E396A

Figure 31. Comparing all three reactive glutamines crosslinked by Factor XIII A2B2 in Fbg αC (233-425) WT vs E396A

Figure 32. LC-MS kinetic assay comparing reactive glutamines crosslinked in Fibrinogen αC(233-425) WT and Q→N variants

Figure 33. Q328 crosslinking extent compared in Fibrinogen αC(233-425) WT, Q237N, and Q237N/Q366N

Figure 34. Q366 crosslinking extent compared in Fibrinogen αC(233-425) WT, Q237N, and Q237N/Q328N

Figure 35. PONDR and PrDOS Order-disorder predictor for Fibrinogen αC(233-425)

Figure 36. 2D HSQC for 15N-labeled Fibrinogen αC(233-425)

Figure 37. Scheme for IPAP HSQC Experiments

Figure 38. 2D HSQC for 15N-labeled Fibrinogen αC(233-425) with Pf1 Phage

Figure 39. Representative In-Phase/Antiphase IPAP HSQC Spectra

Figure 40. Representative Sum and Difference IPAP HSQC Spectra

Figure 41. Summary of RDCs calculated from Fibrinogen αC (233-425)

Figure 42. Comparison of 1D 13C HSQC in 15N/13C versus 15N-labeled Fibrinogen αC (233-425)

Figure 43. 3D HNCO for 15N/13C Fibrinogen αC (233-425)

Figure 44. Representative strip plot from 3D HNCO
CHAPTER 1

INTRODUCTION

The formation of a blood clot is triggered by the body’s need to prevent excessive bleeding, eliminate foreign bodies and prevent infection especially following an injury. Blood clots create a plug at the site of injury preventing excessive bleeding and provide a barrier against infection, a process known as haemostasis. Thrombosis is the formation of a blood clot within a blood vessel following injury. The major components in a clot include; platelets, red blood cells, white blood cells, and the fibrin(ogen) network. Not to our surprise, the mode of recruiting, activating, sustaining and subsequently, terminating the clotting process are very interconnected and tightly controlled by both cellular and molecular components. These processes leading to the formation of a stable blood clot can be divided into four major categories: initiation, propagation, amplification, and stabilization. Several proteins and enzymes are involved in these processes summarized in the blood coagulation cascade described below.
The Blood Coagulation Cascade

Upon injury, the extracellular matrix layer is exposed which triggers the initiation phase in blood coagulation.\textsuperscript{5} One of the major components recruited during the initiation phase are platelets.\textsuperscript{2} Platelets are produced in the liver and circulate freely in the plasma. The exposure of extracellular matrix during an injury triggers a series of events that recruits platelets to the site.\textsuperscript{2} The membrane surface of platelets contains glycoproteins that bind to specific ligands in the vessel wall.\textsuperscript{6} This interaction promotes platelet adhesion and aggregation to form a platelet plug at the site of injury.\textsuperscript{7} Platelets are therefore known to provide primary haemostasis since they make up the first mechanism of defense at the site of injury.

Secondary haemostasis is provided by plasma coagulation factors triggered by Tissue Factor (TF) – another major component involved during the clot initiation phase. This integral membrane protein is also activated upon extracellular matrix exposure and contact with plasma.\textsuperscript{5} TF plays a crucial role in initiating the cascade of enzymatic processes that are later involved in the propagation, amplification, and stabilization of a clot.\textsuperscript{7} The blood coagulation cascade is commonly used to represent the enzymes, proteins and cofactors involved in this process (Figure 1).
Blood clots are formed following a series of events involving several enzymes organized as a cascade. The blood coagulation cascade can be initiated via two pathways: the intrinsic and extrinsic pathways (Figure 1). The intrinsic or contact pathway is less common and activated via negatively charged molecules including naturally occurring extracellular nucleic acids and polyphosphates, misfolded proteins, and negative artificial surfaces e.g glass, silica. The extrinsic or tissue factor pathway is activated upon injury. Both the intrinsic and extrinsic cascades merge and proceed through the “common pathway”. Although the biochemistry leading up to the formation of the clot has been extensively studied,
there are still many unanswered questions especially pertaining to the modes of regulation and target for each enzyme involved in this cascade.

The blood coagulation cascade is made up mainly of serine proteases Factor II –XII (numbered in roman numerals according to their order of discovery and not their role in the cascade). These enzymes are involved in defined processes leading up to the formation of a stable blood clot. Fibrinogen (Factor I) and Factor XIII are the only two proteins in the cascade that are not serine proteases. Most enzymes in the cascade exist in their inactive zymogen states and require one or more components for activation (denoted with “a” upon activation e.g II to IIa). Calcium (Ca$^{2+}$) serves as a cofactor to a majority of the enzymes and is required for the activation of most zymogens in the cascade (Figure 1). For example, calcium is critical in the activation of the transglutaminase Factor XIII found at the final stage. Activated Factor XIII is responsible for introducing crosslinks within the clot and crosslinking other components needed to stabilize the clot. As expected, a tight regulation in this cascade leading to a stable blood clot formation is needed to maintain homeostasis. A disruption in one or more of the enzymes in the cascade can lead to detrimental effect.

The intrinsic or contact factor pathway is initiated in response to a host pathogen defense mechanism and is triggered by negatively charged surfaces and components (e.g excessive DNA, polyphosphates, misfolded proteins, glass, silica, etc). Contact with artificial surfaces e.g cardiopulmonary bypass, extracorporeal membrane oxygenation (ECMO), and hemodialysis, have also been shown to trigger the intrinsic pathway (Figure 1). There has recently
been a renewed interest in Kallikriken research, particularly pertaining to its role in the formation of Factor XII and possible implications in thrombosis in cancer patients. In cancer for example, where there are numerous indications of excessive DNA, it is proposed that the presence of these negatively charged bodies promote activation of Factor XII leading to an overproduction of the blood coagulation cascade components. In the propagation phase of coagulation, activated Factor XII promotes the activation of Factor XI in the presence of Ca$^{2+}$. This in turn activates Factor IX. It should be noted here that Hemophilia B is a result of deficiency in Factor IX. Meanwhile, Haemophilia A is as a result of a deficiency in Factor VIII in the Tissue factor pathway or extrinsic pathway.

The extrinsic or tissue factor pathway is predominant in tissue damage i.e in the wound healing process. This occurs externally and leads the formation of a thrombus if the damage is within a blood vessel. This pathway is named after the tissue factor responsible for initiating this process. Tissue factor together with Factor VII form the (TF: VIIa) complex. In the propagation phase, this complex activates Factor IX and Factor X which leads to a large burst of thrombin for the common pathway. Both the intrinsic and extrinsic pathways contribute to the activation of Factor Xa, and ultimately lead to increased production of thrombin for the common pathway (Figure 1).

In the common pathway, Factor Xa activates Prothrombin (Factor II) to Thrombin (II). Thrombin is one of the major proteins involved in the formation of blood clots and has been studied extensively for drug design. Thrombin is also responsible for creating the positive feedback effect in blood coagulation by
activating many other clotting factors. Thrombin activates Factor V, Factor VIII, Factor XIII and Fibrinogen, amplifying these components of the coagulation cascade. This amplification results in the production of more Factor Xa and, in turn, more thrombin is generated. This positive feedback by thrombin sustains clot production in the amplification phase. Such an amplification is also seen in platelets whereby, thrombin activates platelets surfaces and the components that promote further platelet aggregation including platelet activated receptors (PARs) and the glycoprotein (GpIbα). Activated platelets aggregate effectively with the help of a fibrin layer from fibrinogen (Factor I). Through limited proteolysis, thrombin cleaves fibrinogen to the form fibrin monomers during the stabilization phase of blood coagulation cascade.

Figure 2: Key players at the final stabilization stage of the Blood Coagulation cascade
The stabilization stage, also referred herein as the final stage of the blood coagulation cascade, is composed of three major players: thrombin, fibrinogen and Factor XIII (Figure 2). Thrombin converts fibrinogen to its active form fibrin. This active fibrin self-polymerizes into a fibrous network. The meshy fibrin network forms the mold of a clot, provides an anchoring surface for platelets, and traps red blood cells and other enzymes needed for clot stability. To further stabilize the clot, thrombin activates Factor XIII, a transglutaminase which introduces covalent bonds within the fibrin clot network. Factor XIII also crosslinks inhibitors that prevent clot degradation. Overall, thrombin plays a crucial role in connecting several clotting components at the later stage of the clotting cascade. Thrombin has therefore been viewed as a critical enzyme for researchers and in the pharmaceutical industry.

Several anticoagulant and procoagulant drugs designed to treat heart disease conditions target steps leading up to the formation of thrombin or thrombin itself. Typically these drugs often have a high bleeding risk as their side effects. As expected, targeting and inhibiting thrombin also affects all other clotting factors that rely on thrombin for their activation. Nonetheless, nature has provided mechanisms to control excessive bleeding and insufficient clotting. Inhibitors are harbored in the vascular endothelium and released, as needed, to regulate haemostasis. For example, antithrombin inhibits thrombin and indirectly inhibits Factor Xa, IXa, and XI in the intrinsic pathway. Tissue factor plasmin inhibitor (TFPI) is stored in platelets and inhibits the extrinsic pathway. The tight network formed by fibrin and Factor XIII is also degraded through a process
known as fibrinolysis. This process is characterized by the cleavage of fibrin into fibrin degradation products by the enzyme plasmin. Tissue plasmin activator (tPA) activates plasmin from its inactive form, plasminogen. Inhibitors of tPA have been designed and are currently being used to prevent fibrinolysis in cases of excessive bleeding. Another mechanism of regulation is through Protein C. Thrombomulin-bound thrombin accelerates Protein C activation. Activated Protein C (APC), inhibits several clotting enzymes including Factor Va and Factor VIIIa. Although nature’s procoagulant and anticoagulant mechanisms are in place, these mechanisms do not suffice in the case of deficiencies affecting key players in the blood coagulation cascade. Disrupting any stage of this process can lead to pathological thrombosis or haemorrhage. There is a significant gap in the research and drug designs that target enzymes beyond thrombin.

Deficiencies especially in fibrinogen and Factor XIII are growing and have posed a unique challenge in drug discovery. Researchers in this field are constantly searching and improving knowledge about biomarkers, tools, and drugs specific to each problem-causing enzyme in the cascade. In this research, fibrinogen (Factor I) and Factor XIII are studied to understand their interplay at the final stage of the blood coagulation cascade. The next sections that follow will focus on structural and functional properties of fibrinogen and Factor XIII as examined for the purposes of this dissertation.
Fibrinogen

Fibrinogen is found in the blood plasma of all vertebrates. In humans, it is the most abundant protein involved in the blood coagulation cascade and circulates at concentrations ranging between 1-3 g/L.\(^{22, 23}\) The activated form of fibrinogen, fibrin, is the fundamental fabric of blood clots, holding all the components of the clot together. Fibrin polymerization results in a meshy network that provides an adhesive surface for platelets and traps red blood cells to form the hemeostatic plug.\(^1\) Fibrinogen also provides an interacting and binding surface for other components required for the formation of a stable clot and subsequent clot degradation/fibrinolysis.\(^{22, 23}\) To understand the functions of fibrinogen in the blood coagulation process, it is important to explore its unique structural features.

Fibrinogen Structure

Fibrinogen is a 340kDa glycoprotein that contains two sets of A\(_\alpha\) (67kDa), B\(_\beta\) (56kDa) and \(\gamma\) (47kDa) chains held together by 29 disulfide bonds.\(^{22, 24, 25}\) Each fibrinogen molecule is a dimer arranged into a unique globular structure with four major regions, which include a central E region, a helical coiled-coil region, the distal D region, and the \(\alpha\)C region (Figure 3).\(^{24, 26}\) The central E region consists of the N-termini of each of the A\(_\alpha\), B\(_\beta\), and \(\gamma\) chains. The N-termini of the A\(_\alpha\) and B\(_\beta\) chains contain fibrinopeptides A (FpA) and fibrinopeptides B (FpB), respectively. The helical coiled-coil region is an extension of all three chains from the central E region to the distal D regions.\(^{24, 27}\)
Figure 3: Crystal Structure of Fibrinogen (PDB: 3GHG) showing the central E region and two distal D domains. Two Aα chains (green and yellow), Bβ (cyan and orange) and γ chains (pink and grey) are also shown to demonstrate the assembly of the dimer.

The D regions, composed mainly of the C-termini of the β and γ chains, flank the ends of the fibrinogen molecule and contain an anchoring site utilized during fibrin polymerization (Figure 3). A majority of the C-terminal portion of the Aα chain exists as a very flexible region known as the αC region (221-610) (Figure 4). The αC region is further divided into a more flexible N-terminal αC connector (221-391) and a more compact C-terminal αC domain (392-610). The intrinsic disordered nature of the αC region has made it difficult to identify from X-ray crystallography studies of the fibrinogen molecule. Nonetheless, this region has been proposed to play several roles during the assembly of the fibrin network. The structural representation of the αC region is shown in Figure 4.
Prior to fibrin polymerization, the \( \alpha C \) region is tethered to the central E region.\textsuperscript{30, 31} Fibrinogen is converted to active fibrin upon thrombin cleavage of fibrinopeptides A and B, on the A\( \alpha \) and B\( \beta \) chains respectively. Cleavage of fibrinopeptides A and B initiates fibrin polymerization by exposing knobs “a” and “b” in the central E region which interact with the complementary holes “A” and “B” in the D region of another fibrin molecule.\textsuperscript{18} The \( \alpha C \) region is also released from the central E region following cleavage of fibrinopeptides A and B, making this region available to participate in fibrin polymerization (Figure 5).\textsuperscript{18} Although the \( \alpha C \) region has been proposed to play important roles in fibrin polymerization, little structural information is available on this region.

Recent approaches, including atomic force microscopy (AFM) and nuclear magnetic resonance (NMR) spectroscopy, have been used to explore the structural features of the \( \alpha C \) region that contributes to its functional roles.\textsuperscript{31, 32} NMR spectroscopy experiments have demonstrated that a majority of the \( \alpha C \)
region is intrinsically disordered. Meanwhile, recent AFM studies provided a visual representation of this flexible αC region in the full-length fibrinogen, highlighting its role in the lateral aggregation of protofibrils.31

Fibrinogen Function

![Fibrin Polymerization](image)

Figure 5: Fibrin Polymerization illustrated (adapted from ref 30)

Fibrinogen provides an interacting surface for enzymes, proteases, and inhibitors involved in the blood coagulation process. The αC region of fibrinogen is a particularly attractive region for these interactions as shown by multiple research studies. The plasmin inhibitor, alpha-2-antiplasmin, binds to the αC region of fibrinogen and has been shown to be specifically crosslinked to K303 in the αC region.33, 34 Alpha-2-antiplasmin’s inhibitory activity prevents the premature degradation of clots.
During the early stages of blood coagulation, thrombin cleaves fibrinopeptides A and B from the central E region. Cleavage of fibrinopeptides A and B converts fibrinogen \((\alpha\beta\gamma)^2\) to active fibrin \((\alpha\beta\gamma)^2\), releases the \(\alpha\gamma\) region, and exposes knobs “a” and “b” in the E region which have complementary binding pockets in the D regions. Hydrophobic knob-hole interaction sets the foundation for the formation of a soft clot made up of staggered protofibrils (Figure 5).\(^{35}\)

The carrier B subunit of Factor XIII relies on fibrinogen as an interacting surface to bring the catalytic A subunit of Factor XIII in close proximity to its substrate for the transamidation (crosslinking) reaction.\(^{19}\) A proper structural assembly of fibrinogen is necessary for its function and other components involved in the blood coagulation cascade that rely on fibrinogen for the formation of a stable clot. Fibrin is also the primary substrate for Factor XIII.\(^{25}\) Although fibrin polymerization can proceed after activation of fibrinogen, a hard clot is only formed in the presence of FXIIIa – a transglutaminase.\(^{28, 36, 37}\) Factor XIII introduces \(\gamma\gamma, \alpha\alpha, \alpha\gamma\), crosslinks between the \(\alpha\) and \(\gamma\) chains of fibrinogen.\(^{38, 39}\) A clot that is more stable and resistant to degradation is formed as a result of this crosslinking reaction.\(^{40, 41}\)

**The Transglutaminase Factor XIII**

Transglutaminases (TGases) are a group of enzymes that catalyze calcium and thiol-dependent post-translational protein modifications. These enzymes are best known for their role in blood coagulation, skin barrier formation and extracellular matrix formation due to their ability to create stable lysis-
resistant networks via inter and intra molecular crosslinks. The catalytic actions of TGase involve transamidation, esterification, and hydrolysis reactions.\textsuperscript{42, 43}

Factor XIII is the only member of the transglutaminase family that exists as a dimer in its inactive zymogen form. In plasma, Factor XIII exists as a 320kDa tetrameric molecule (Factor XIII A\textsubscript{2}B\textsubscript{2}), with two enzymatic A-subunits (~83kDa) and two carriers B-subunits (~80kDa).\textsuperscript{11, 42, 44} In plasma Factor XIII A\textsubscript{2}B\textsubscript{2}, the B-subunits of Factor XIII have been proposed to play a protective role, preventing the degradation and premature activation of the catalytic A-subunit.\textsuperscript{19} Cellular Factor XIII is found in platelets and placenta, and is only composed of the catalytic A-subunits (Factor XIII A\textsubscript{2}).\textsuperscript{42} Deficiencies in Factor XIII have been associated with severe bleeding diathesis. Mutations affecting the catalytic A-subunit of Factor XIII have been shown to be more severe.\textsuperscript{45} In the absence of the carrier B-subunits, the levels of Factor XIII A\textsubscript{2} in circulation were lowered but did not produce a drastic an effect as that observed in the Factor XIII A\textsubscript{2} knockout mice.\textsuperscript{46} In circulation, the carrier B-subunits have also been shown to play an important role in positioning the catalytic A-subunit in an ideal location for crosslinking through its interactions with Fibrinogen.\textsuperscript{43} The crystal structure of the Factor XIII A\textsubscript{2}B\textsubscript{2} has not been resolved. The structural features involved in the interaction between Factor XIII A\textsubscript{2} and B\textsubscript{2} subunits are still unclear. However, structural information on the catalytic A subunit has been successful in understanding of this transglutaminase.\textsuperscript{44, 47}
**Factor XIII Structure**

Structurally, each catalytic A-subunit of Factor XIII (Factor XIIIa) contains an activation peptide, a β-sandwich domain, a catalytic core, and two C-terminal β-barrel domains (β-barrel 1 and β-barrel 2) as shown in Figure 6. During the activation of Factor XIII in plasma, the carrier B-subunits dissociate from the catalytic A-subunits following thrombin cleavage of the activation peptide and upon calcium binding.

![Figure 6: Structure of the catalytic A subunit of Factor XIII (PDB: 4KTY). Adopted from references 44 and 48](image)

The catalytic core of Factor XIII consists of a Cysteine, Histidine, and Aspartate triad, also conserved in all catalytic transglutaminases. In the presence of thrombin, the activation peptide is cleaved off the A-subunit to form the active enzyme FXIIIa. Three calcium-binding sites have been shown to exist on Factor XIII. In the presence of calcium, Factor XIII A adopts a suitable
conformation required for activity.\textsuperscript{50} Factor XIII A can also be activated non-proteolytically in the presence of high millimolar calcium concentrations. Although the concentrations required for this non-proteolytic activation are significantly higher than physiological levels of calcium, cellular Factor XIII has been proposed to be activated non-proteolytically by calcium. A recent study by Klebe et al., crystal structures of Factor XIII A using an inhibitor in the active site and high calcium concentrations (40 mM) suggest that the β-barrels are moved away from the core structure to expose the active site (Figure 7).\textsuperscript{44} However, additional research is needed to understand this mode of activation of Factor XIII including its transition from an inactive dimeric to an active monomeric form.

Figure 7: Non-proteolytic activation of Factor XIII in the presence of high calcium concentrations
Factor XIII Function

In blood coagulation, activated Factor XIII (Factor XIIIa) is known to crosslink specific reactive glutamines within fibrinogen to lysines to form γ-glutamyl-ε-lysyl covalent isopeptide bonds.\textsuperscript{19, 51} Hydrolysis results in the deamidation of glutamine to form glutamate. Esterification is more common in other transglutaminases whereby the side chains of glutamines can be modified by addition of fatty acids.\textsuperscript{42} Fewer reports of esterification have been shown with Factor XIII. Strikingly, each transglutaminase has a unique specificity towards reactive glutamines. Factor XIII’s selectivity for reactive glutamines is more specific than lysines.\textsuperscript{52, 53} However, Factor XIII’s specificity towards one reactive glutamines over another is still unclear.

During catalysis, activated Factor XIIIa selects a specific glutamine containing substrate to serve as the first acyl acceptor (Figure 8). In addition to the catalytic triad Cys 314, His 373 and Asp 396, a conserved tryptophan (Trp279) has been shown to serve as an essential stabilizer employed during the transamidation reaction.\textsuperscript{13, 50, 54} The γ-carboxyamide group on the reactive glutamine reacts with cysteine to form the acyl enzyme intermediate γ-glutamylthioester with the release of ammonia. This is the rate-limiting step.\textsuperscript{54} The second substrate lysine, attacks and cleaves the thioester bond in a deacylation step. The crosslinked γ-glutamyl-ε-lysyl product is then released (Figure 8).
Figure 8: Proposed Reaction Mechanism for Factor XIII transamidation reaction
Reactive glutamines in the \( \gamma \) and \( \alpha \) chains of fibrinogen are the primary physiological substrates of Factor XIIIa.\(^{55-57}\) Two reactive glutamines in the \( \gamma \)-chain (Q398 and Q399) have been identified as substrates of Factor XIIIa. By contrast, multiple reactive glutamines in the \( \alpha \) chains (Q221, Q223, Q328, Q366) have been shown to be crosslinked by Factor XIIIa. Factor XIII’s ability to form, \( \gamma \)-\( \alpha \) dimers and \( \alpha \)-\( \alpha \) polymers promotes lateral aggregation between fibrin monomers resulting in a tighter fibrin network and a hard clot.\(^4, 37, 58\) Both \( \gamma \)-\( \gamma \) and \( \alpha \)-\( \alpha \) crosslinks contribute separately to unique clot properties.\(^{59}\) Factor XIIIa also crosslinks several clot-stabilizing proteins which protect the clot from plasmin-mediated degradation e.g alpha-2-antiplasmin and plasmin activator inhibitor 1 (PAI1).\(^{40, 60}\)

Crosslinking by Factor XIII plays an important role in the overall nature of the clot, altering clot properties to variable extents. In the absence of Factor XIII, fibrin fibers have been shown to have thicker fibers.\(^{37, 61}\) Meanwhile in the presence of Factor XIII, crosslinking results in thinner fibers, with increased fiber compaction, and a less porous clot.\(^{62}\) Factor XIII crosslinking is suggested to produce a tighter fibrin network through chain oligomerization in protofibrils,
which in turn affects the rates of diffusion of fibrinolytic enzymes through the fibrin clot network. Understanding the nature of Factor XIII crosslinking is therefore needed to design specific Factor XIII inhibitors to alter clot properties to different extents.

**Cardiovascular Relevance and Significance of Research**

Fibrinogen levels are currently used as a risk factor for cardiovascular disease.\textsuperscript{63, 64} There is increasing evidence suggesting that the nature of the fibrin clot plays an important role in thrombosis.\textsuperscript{18, 22, 59} Fibrinogen is the most abundant component involved in blood coagulation with plasma concentrations ranging from 1 to 3 mg/mL in plasma. Consequently, blood clots are composed of red blood cells trapped in tight fibrin networks. Fibrin has a higher concentration in clots than circulation and is the major components in all thrombi (venous or arterial). Fibrin is therefore an ideal target for detecting and characterizing thrombosis, particularly because of its high specificity and sensitivity in all thrombi. The fibrin network also provides binding platforms for several clot-stabilizing factors (e.g. alpha-2-antiplasmin and factor XIII), which either enables or prevent clot lysis.\textsuperscript{65} Factor XIIIa, for example, introduces fibrin-fibrin intra and intermolecular covalent bonds within the fibrin network. Factor XIII also crosslinks other substrates to fibrin that, renders blood clots more resistant to fibrinolysis.\textsuperscript{11, 66} Deficiencies in fibrinogen and Factor XIII have been identified across the world. Most cases are presented with abnormal clot properties and especially recurring miscarriages in women with certain fibrinogen and Factor XIII
deficiencies. \(^{22, 45}\) Interestingly, a majority of substrates responsible for resistance in clot degradation are crosslinked to the flexible \(\alpha C\) region of fibrinogen. \(^{58}\) Several mutations associated with disease have also been found in this region, making it a candidate for deciphering the fibrin network at a molecular level.

### Role of the \(\alpha C\) region Fibrinogen

In recent years, studies have emerged exploring the \(\alpha C\) region as a potential biomarker in several pathological conditions. \(^{67}\) Previous studies have shown several truncations and mutations in the \(\alpha C\) region associated with bleeding disorders. \(^{58, 68}\) Examples involving specific reactive glutamines include Fibrinogen Otago, Fibrinogen Keokuk and Fibrinogen Seoul II. \(^{68-70}\) Fibrinogen Otago (\(A\alpha 270\)) was discovered in a patient with mild bleeding and recurrent miscarriages. Here, removal of about 60% of the C-terminus of \(\alpha C\) region (\(A\alpha 271-610\)) resulted in abnormal clotting conditions, decreased circulating fibrinogen, and a reduced rate of polymerization. \(^{71}\) In Fibrinogen Keokuk, a point mutation at Gln328 into a stop codon results in a 46% truncation of the \(\alpha C\) region. \(^{69}\) This mutation and truncation resulted in abnormal bleeding conditions and multiple miscarriages in two patients. In Fibrinogen Seoul II, a single mutation in the reactive glutamine Gln328 to a proline results in impaired fibrin crosslinking in a patient diagnosed with myocardial infarction. \(^{68}\) A recent study by Duval et al, isolated the contributions of fibrin \(\alpha\) and \(\gamma\) crosslinking. These findings suggest that crosslinking from the alpha chain are responsible for fiber stiffness and clot stiffness. \(^{59}\) It is therefore important to fully characterize the
selectivity for each reactive glutamines to understand how Factor XIIIa recognizes this intrinsically disordered αC region as a substrate. Most studies have made use of physiological assays to understand patterns and trends involved in dysfibrinogenemia. However, in depth evaluation of each crosslinking glutamine has not been explored. The goal of this dissertation is to understand crosslinking in the αC region of fibrinogen by ranking each reactive glutamine and to use this information to elucidate Factor XIIIa’s substrate specificity. The knowledge gained from this study can be used to develop new diagnostic tools and therapeutic strategies to alter the nature of clots.

**Structural features of the αC region**

The αC region (221-610) makes up two thirds of the alpha chain of fibrinogen. The intrinsically disordered nature of this region is promoted by a large number of serines (17.7%) glycines (14.8%) and prolines (6.9%).

Figure 10: Amino acid sequence of Fibrinogen αC region (221-610). Five reactive glutamines (red) crosslinked by Factor XIII and two Cysteine residues (bold) located C-terminal highlighted.
The flexible αC connector (221-391) contains a majority of the reactive glutamines crosslinked by Factor XIII. The αC domain (392-610) has more structural features compared to the αC connector. Several truncations have been explored to map these structural features in the αC domain. Medved et al., demonstrated that there are two anti-parallel beta chains within αC (406-483) that form a β-hairpin. This is promoted by the presence of a disulfide bond, Cys423-Cys453 in the αC domain. 28

A segment of the αC region, Fibrinogen αC (233-425) previously shown to bind to Factor XIII is explored in this dissertation. Factor XIII’s ability to crosslink three reactive glutamines in this region was examined. This segment was also used to explore any structural features located close to the Factor XIII’s binding site (αC (389-402).

Hypothesis and scope of research

Intrinsically disordered regions of proteins have been increasingly associated with diseases. In cardiovascular disease, the αC region of fibrinogen would be one substrate to investigate and consider given that is the most flexible region in fibrinogen. Research exploring fibrinogen deficiencies involving the αC region already provide some evidence that this disordered region should be explored extensively. Moreover, Factor XIII selects most reactive glutamines in this region for crosslinking its substrates. 26 Substrates of Factor XIII do not align with any specific consensus sequences that demonstrate a preferred selectivity for one reactive glutamine over another, on a particular sequence. 42 We hypothesize that Factor XIII’s substrate specificity in the αC region is not random.
but rather results from the flexible nature of this region and the presence of a proximal binding site for Factor XIII. To test this hypothesis a combination of mass spectrometry and NMR spectroscopy techniques were used.

A brief summary of these techniques and the detailed application for the purpose of this study is discussed in Chapter 2. The first research goal to rank the reactive glutamines in the αC region is discussed in Chapter 3. The role of a previously proposed residue αCE396 on Factor XIII crosslinking ability is explored in Chapter 4. In Chapter 5, initial structural studies to identify pre-structured motifs in the αC connector region are discussed and future directions for this research are presented in the concluding Chapter 6.
CHAPTER 2
METHODS AND APPLICATIONS

Protein expression and purification of recombinant Fibrinogen αC (233-425)

The DNA encoding Fibrinogen αC (233-425) was obtained from Smith et al. as a pGEX-6P-1 construct containing a GST-tag and a precision protease recognition and cleavage site (Leu Glu Val Leu Phe Gln | Gly Pro). The expression and purification steps were performed as previously described by Smith et al.\textsuperscript{72}

Briefly, the DNA was transformed into BL21-Gold (DE3) \textit{E.coli}. Starter cultures were grown in 180 mL terrific broth (TB: 1.2% (w/v) tryptone, 2.4% (w/v), 0.8% (w/v) glycerol), 20 mL phosphate buffer (171 mM KH$_2$PO$_4$, 942 mM K$_2$HPO$_4$) and 200 µL ampicillin (100 mg/mL) and incubated at 37°C for 16 h in a Brunswick Scientific shaker (Edison, NJ, USA) at 225 rpm. The starter culture was used to inoculate a mixture of 1800 mL Terrific Broth, 200 mL phosphate buffer and 2 mL 100 mg/mL ampicillin. Cell growth was monitored for OD$_{600}$ 0.7-0.9, when protein expression was induced by adding isopropyl-β-D-thiogalactopyranoside (IPTG) to 1 mM. The cell culture was incubated with shaking at 180 rpm for 16 h at 30°C in a Labline Incubator shaker. Cells were
harvested by centrifugation (5500 rpm, for 10 min in a Beckman Coulter JA-10 rotor). The pellets were resuspended in 180 mL wash buffer (20 mM Tris acetate, 50 mM NaCl pH 8), and the resuspended cells were then pelleted by centrifugation at 5000 rpm for 1 h at 4°C. The cell pellets were quick frozen at and stored at -20°C.

Cells were lysed in the same manner as described Smith et al.\textsuperscript{72} Cells were resuspended in 88 mL phosphate buffer saline (PBS: 137 mM NaCl, 2.7 mM KCl, 10 mM NaHPO\textsubscript{4}.2H\textsubscript{2}O, 2 mM K\textsubscript{2}HPO\textsubscript{4} pH 7.4), incubated with 1 mg/ml lysozyme, 1 mM dithiotreitol (DTT) for 30 min at 25°C. 4 mM Benzamidine, 2 μg/mL aprotinin, 1 μM pepstatin-A and 10 μM leupeptin were added to the mixture and incubated for 30 min at 4°C with gentle stirring. After the addition of phenylmethansulphonylfluoride (PMSF) to 0.5 mM and sodium deoxycholate to 0.05% (v/v), this mixture was incubated at 25°C for another 30 min. To this lysate triton X-100 was added to 0.02% (v/v) with vigorous agitation. DNase I (to 5 μg/mL) and MgCl\textsubscript{2} to 5 mM were added and the mixture incubated for 10 min at 4°C and quenched with the addition of ethylenediamine tetraacetic acid (EDTA) to 6 mM and streptomycin (1 g). The cell lysate was then spun at 13,500 rpm for 20 min at 4°C. The soluble protein in the supernatant was filtered through 0.45 μm and then 0.2 μm vacuum filters (Nalgene), and purified with a GST- affinity resin column (GE Healthcare) on an AKTAprime™ FPLC system (Amersham Scientifics) (Figure 11).
Fibrinogen αC (233-425) was cleaved off the GST-tag using a PreScission™ Protease at 2 U per 100 ug protein as recommended by GE Healthcare and eluted in PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 2 mM KH₂PO₄ pH 7.4) at room temperature. An additional GST-trap column was used during the elution to ensure high protein purity (Figure 11). This catch column traps any uncleaved GST-tagged protein during the elution step, thereby increasing the protein purity of the αC (233-425) fractions collected.
Figure 12: A representative SDS-PAGE gel (15%) of fractions with Fibrinogen αC (233-425) eluted from GST-Trap column following on column cleavage (Lanes 1-8) shows protein band at 20kDa. Fraction eluted with glutathione (Lane 9) shows position of GST-tag (25 kDa) relative to protein fractions. MW = molecular weight marker.

Complete cleavage of the GST tag from αC was confirmed by SDS – PAGE (Figure 12). Fibrinogen αC (233-425) and GST have similar molecular weights; SDS PAGE alone could not definitively distinguish between these two bands. Western blot analysis was used to verify the presence of αC only and no free GST or uncleaved αC-GST in the protein fractions collected after cleavage by the PreScission™ Protease. SDS-PAGE gels were prepared for Western blot transfer. Protein fractions and samples from the glutathione elution step were loaded on the gel. Following separation, the gel was then subjected to Western blotting using an anti-GST-HRP Conjugate (Amersham™, GE Healthcare). All Western blot set-up components were equilibrated in Western Blot transfer buffer (25 mM Tris, 192 mM glycine, and 20% methanol) for 10 min at room temperature. The gel was equilibrated in the same buffer then sandwiched
between two nitrocellulose membranes. Transfer was performed at 100 V and run for 1 h at 4°C. The transfer nitrocellulose membrane was developed as specified by the manufacturer. Non-specific binding was reduced by shaking the nitrocellulose membrane overnight at room temperature in a blocking solution containing 3% BSA in PBST buffer (80 mM Na₂HPO₄, 20 mM NaH₂PO₄, 100 mM NaCl, 0.001% Tween-20). The membrane was washed with fresh PBST buffer and transferred to a solution containing 6 µL anti-GST HRP conjugate in 25 µL PBST which was shaken for 1 h and rinsed in PBST, before the bands were developed using TMB blotting buffer. The membrane was quenched with water. Images of all gels and blots were taken on a Biorad Image system (Figure 12 and 13). Although the anti-GST antibody was not highly specific, comparison of samples containing GST-tagged proteins and SDS-PAGE analysis confirmed that the GST-cleaved Fibrinogen αC (233-425) protein fractions were free of the GST-tag.

Figure 13: Western Blot analysis of Fibrinogen αC (233-425) protein fractions (Lane 2 and 3) and GST-tag fractions (Lane 5 and 6). Anti-GST antibody shows no bands in fractions containing purified proteins. Fraction eluted with glutathione (Lane 5 and 6) shows bands containing GST-tag
The purified recombinant Fibrinogen αC (233-425) was concentrated in a Vivaspin 2 concentrator (Sartorius Stedim Biotech, Goettingen, Germany), and the concentration was determined using $A_{280\text{nm}}$ determined with a Cary 100 Bio UV-Vis Spectrophotometer (Varian) and an extinction coefficient of $41480 \text{ M}^{-1}\text{cm}^{-1}$, calculated from ExPASy (http://web.expasy.org/protparam). Protein yields of 6-11 mg were obtained from 4 L cultures. Concentrated protein was aliquoted for storage at 4°C. All proteins with point mutations were expressed as described above. A similar expression protocol with slight modifications was also adopted for $^{15}\text{N}$ and $^{13}\text{C}$-enriched αC (233-425) as discussed in Chapter 5.

**Mass Spectrometry**

Recent advances in mass spectrometry have been motivated by the need to analyze large quantities of complex samples, identify biomarkers for disease, and develop assays with very low limits of detection. Several mass spectrometry approaches specific for proteomics have been developed. Mass spectrometry allows for the separation of complex molecular species by their mass-to-charge ratio upon ionization. MALDI-TOF (Matrix Assisted Laser Desorption/Ionization-Time Of Flight) mass spectrometry and LC-MS (Liquid chromatography - mass spectrometry) approaches were used for experiments in this dissertation.

MALDI-TOF mass spectrometry is a matrix-assisted approach whereby sample ionization is enabled by the transfer of energy from the matrix. The matrix absorbs energy from the laser and transfers this energy to the sample that is ionized. The ionized species are then accelerated via a voltage grid through a
time of flight tube toward a detector with reflectron capabilities. Sample preparation is a key step in mass spectrometric analysis, and in the case of MALDI-TOF MS, the choice of matrix is also important.  

Two general approaches are utilized in proteomics for sample preparation; the “top-down approach” and the “bottom-up approach.” The top-down approach explores proteins in their intact state whereas in the bottom-up approach, proteins are digested and the data obtained can be pooled to characterize the whole protein. For this study, the transamidation reaction catalyzed by Factor XIII was monitored using the bottom-up approach. Here, peptide fragments containing the reactive glutamines of interest were identified and monitored after Chymotrypsin, and GluC digests. Two common matrices for peptide/protein analysis, α-cyno-4-hydroxycinnamic acid (αCHCA) and ferulic acid matrices were optimized and used for the chymotrypsin and GluC digests, respectively (Figure 14).
Figure 14: Chymotrypsin and GluC digests of Fibrinogen αC (233-425) showing peaks containing all three reactive glutamines (Q237, Q328, and Q366)

MALDI-TOF MS on Fibrinogen αC (233-425)

To identify peaks containing reactive glutamines in Fibrinogen αC (233-425), this sample was separately digested with chymotrypsin (1 h) and GluC protease (2 h) at 25°C. Each sample was then zip-tipped using a C18 resin tip, and α-cyano-4-hydroxycinnamic acid (αCHCA) and ferulic acid matrices were used for chymotrypsin and GluC protease digests, respectively. A MALDI-TOF mass spectrometer (Applied Biosystems Voyager DE-PRO) was employed to analyze these samples. The spectra obtained were compared with theoretical digests to identify peaks containing reactive glutamines.
One peak containing Q328 was observed when αC (233-425) was digested with chymotrypsin, and the two other reactive glutamines were observed in peaks from the GluC digest. It should be noted that peaks containing Q328 and Q366 were confirmed using the theoretical digest whereas Q237 was identified using LC-MS/MS (Waters Synapt G2-Si). The peptide fragment with Q237 was shown to have five extra residues (GPLGS) from the precision protease cleavage site, which are not a part of the wild type αC sequence. Two reactive glutamines, Q398 and Q399, on the Fibrinogen gamma module γ(148-411) were also identified following a GluC digest. A summary of all peptide fragments discussed here and in subsequent chapters is listed in Table 1.
Table 1: Summary of peptide fragments containing reactive glutamines, Q→N mutations, and other important residues characterized by MALDI-TOF mass spectrometry

<table>
<thead>
<tr>
<th>Species</th>
<th>Peptide Fragment</th>
<th>Average m/z</th>
<th>Digest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q237</td>
<td>GPLGSTDMPQMRM</td>
<td>1550</td>
<td>GluC</td>
</tr>
<tr>
<td>Q237N</td>
<td>GPLGSTDMPNMRM</td>
<td>1535</td>
<td>GluC</td>
</tr>
<tr>
<td>Q328</td>
<td>NSGSSGTGSTGQNPNPGPRPGSTGTW</td>
<td>2448</td>
<td>chymotrypsin</td>
</tr>
<tr>
<td>Q328N</td>
<td>NSGSSGTGSTGNNPNPGPRPGSTGTW</td>
<td>2347</td>
<td>chymotrypsin</td>
</tr>
<tr>
<td>Q366</td>
<td>SSVSGSTGQWHSE</td>
<td>1349</td>
<td>GluC</td>
</tr>
<tr>
<td>Q366N</td>
<td>SSVSGSTGNWHSE</td>
<td>1334</td>
<td>GluC</td>
</tr>
<tr>
<td>E396</td>
<td>SGSFRPDPSPGSGNARPNNPNDWGTFAE</td>
<td>2776</td>
<td>GluC</td>
</tr>
<tr>
<td>E396A</td>
<td>SGSFRPDPSPGSGNARPNNPNDWGTFAE</td>
<td>2647</td>
<td>GluC</td>
</tr>
<tr>
<td>γ398γ399</td>
<td>GQQHHHLGGAKQAGD</td>
<td>1501</td>
<td>GluC</td>
</tr>
</tbody>
</table>

**MALDI-TOF MS Kinetic Assay**

This kinetic assay was adapted from a previously described work on FXIII substrate specificity for Q-containing peptides substrates. The Q-containing protein substrate was combined with the lysine mimic glycine ethyl ester (GEE) in the presence of activated Factor XIII. At different time points, aliquots of the reaction were quenched and peptide digests performed to monitor for fragments containing reactive glutamines and its crosslinked product (Figure 15).
MALDI-TOF MS Kinetic Assay on Fibrinogen αC (233-425)

The assay was prepared in a 1.5 mL centrifuge tube with a final volume of 250 µL. Recombinant FXIII (500 nM) was activated with bovine thrombin (8.4 U/mL) in the presence of CaCl₂ (4 mM), glycine ethyl ester (GEE, 17 mM) and buffer [100 mM Tris-acetate, 150 mM NaCl and 0.1% PEG, pH 7.4] for 10 min at 37°C. Thrombin was inhibited with 200 nM PPACK (D-phenylalanyl-L-prolyl-L-arginine chloromethyl ketone).

The transglutaminase reaction was initiated by adding Fibrinogen αC (233-425) at a final concentration of 13.6 µM. 25 µL aliquots were removed at various time points and quenched in EDTA (0.5 mM). 6 µL of each aliquot were digested separately with chymotrypsin (1 h) or GluC protease (2 h). Each sample
was zip-tipped using a C18 resin tip. Chymotrypsin and GluC samples were analyzed with α-cyano-4-hydroxycinnamic acid (αCHCA) and ferulic acid matrices respectively, using a MALDI-TOF mass spectrometer (Applied Biosystems Voyager DE-PRO). The Factor XIII-mediated reaction between each reactive glutamine with GEE results in an additional 86 m/z seen as a product peak on each spectrum. A representative MALDI-TOF time course spectra showing the reactant disappearing and product peak appearing is shown in Figure 16.

![Figure 16: Representative data showing reactant and product peak formation at different time points from MALDI-TOF MS kinetic assay.](image)

The peak height ratio method (see equation below) was used to determine the extent to which Factor XIIIa incorporated GEE over time.
\[ \frac{\sum \text{Reactant peak height}}{\sum \text{Reactant peak height} + \sum \text{Product peak height}} \times \text{peptide concentration} \]

**MALDI-TOF MS Kinetic Assay on Fibrinogen Gamma Module γ (148-411)**

The gamma chain of fibrinogen contains two reactive glutamines (Q398 and Q399) known to be crosslinked by Factor XIII. To test this MALDI-TOF MS Kinetic Assay, a fibrinogen gamma module γ (148-411) sample was obtained from the lab of Dr. Alisa Wolberg (UNC Chapel Hill). Using a GluC digest, a peptide fragment containing both reactive glutamines Q398 and Q399 was identified with a 1501 m/z (GQQHHLGGAKQAGD) (Table 1 and Figure 17). The MALDI Kinetic Assay was tested using this gamma module as the Q-substrate. After optimizing the assay for this new substrate, a higher concentration of Factor XIII and longer reaction times were required to observe the crosslinking reaction in the gamma module compared to the αC (233-425) segment. Two peaks corresponding to one (1587 m/z) and two (1673 m/z) glycine ethyl esters crosslinked to the reactive glutamines were observed by 30 min (Figure 17). The peak containing one reactive glutamine showed increasing intensity at 60 and 120 min.
Figure 17: MALDI TOF Kinetic Assay was also used to demonstrate crosslinking in gamma module $\gamma$ (148-411) of fibrinogen

**Liquid Chromatography – Mass Spectrometry**

This assay and sample preparation approach was also compatible with the liquid chromatography-mass spectrometry (LC-MS). Liquid chromatography provides an additional separation step with unique retention times for each peptide fragment. When coupled with mass spectrometry, specific peptide fragments of interest can then be identified by their mass-to-charge ratio. In this project, the LC-MS method provided higher throughput and a more automated approach than the MALDI-TOF MS to characterize the crosslinking reaction between reactive glutamines in the $\alpha$C region.
LC-MS Experimental Approach

Sample Preparation

The crosslinking reaction was carried out in the same way as described for the MALDI TOF MS kinetic assay. Samples obtained at each time point were prepared for LC–MS analysis as summarized below. A volume of each sample was diluted with three volumes of 2% v/v acetonitrile (ACN) / 0.1% v/v trifluoracetic acid (TFA) to a final volume of 30 µL. A p10 ZipTip® was wetted with 3 x 10 µL 80% v/v ACN / 0.1% v/v TFA, equilibrated with 3 x 10 µL 2% v/v ACN / 0.1% v/v TFA, the sample was cycled in the tip 15 x 10 µL, the tip was washed with 4 x 10 µL 2% v/v ACN / 0.1% v/v TFA, and the sample was eluted into 30 µL 80% v/v ACN / 0.1% v/v formic acid. The samples were frozen and then dried in a SpeedVac. The dried samples were dissolved in 30 µL 2% v/v ACN / 0.1% v/v formic acid and 4 µL of each sample were analyzed.

Liquid Chromatography

A 250 µm ID x 15 cm Radel R tube (Idex Health & Science LLC, Oak Harbor, WA, USA) was packed in-house with Aeris Peptide 3.6 µm XB-C18 material (Phenomenex, Torrance, CA, USA). 2 µm stainless steel Frit-in-a-Ferrule™ Super Flangeless™ PEEK fittings (Idex) were used to retain the material at both ends of the column. An Acquity M-Class UPLC® system (Waters Corporation, Milford, MA, USA) was used with buffer A (water with 0.1% v/v formic acid) and buffer B (acetonitrile with 0.1% v/v formic acid) as mobile phases. Following injection of the sample onto the column, separation was accomplished with a 40 min linear gradient from 2% B to 40% B, followed by a 5
min linear gradient from 40% B to 85% B, and a 5 min wash with 85% B. A Low Flow Electron Spray Ionization (ESI) probe was used to introduce sample into a ZSpray LockSpray source (Waters).

**Mass Spectrometry**

A Synapt G2-Si mass spectrometer (Waters) was used to collect data from the LC eluate. An MSe Continuum method was created in MassLynx v4.1 SCN924 (Waters) operating in positive mode. A 1.5 s low energy continuum scan without trap or transfer collision energy was acquired, followed by a 1.5 s high energy continuum scan (15 to 40 V ramp trap collision energy; no transfer collision energy). The scan range was 50 to 2000 Da for GluC-digested samples and 50 to 3000Da for chymotrypsin-digested samples. Leucine enkephalin (556.2771 Da/e) at 300 pg/µL in 1:1 acetonitrile: water was used as the lock mass during data acquisition.

**Data Analysis and Results**

The precursors with the reactant and product peaks from +1, +2 and +3 charged states from both reactant and product peaks at each time point were extracted using Skyline Lab Software. The Q-GEE modification was incorporated into the search library for quantification. The peak height ratio method was also used to characterize each reactive glutamine from the peaks detected in the LC-MS assay. A comparison of the data obtained from the MALDI kinetic assay and the LC-MS kinetic assay obtained in triplicate is shown in Figure 18. This data
suggests that both MALDI-TOF MS and LC-MS approaches produce a similar trend for each reactive glutamine.

Figure 18: Ranking reactive glutamines in Fibrinogen using (A) MALDI-TOF mass spectrometry versus (B) LC-MS approach
This LC-MS approach provided a faster throughput, an improved detection limit, and an efficient data analysis package to quickly monitor the crosslinking reaction. The MALDI-TOF MS approach was employed to monitor crosslinking trends for each reactive glutamine (Chapters 3 and 4). The LC-MS approach was used to characterize each remaining glutamine following Q to N mutations of one or more reactive glutamines (Chapter 4).

**Nuclear Magnetic Resonance (NMR) Spectroscopy**

In the absence of X-ray crystal structure information on proteins, two-dimensional NMR can reveal through-space interactions of amino acid residues within a protein. Several 2D experiments have been developed to obtain enough information to elucidate structural characteristics of proteins.\(^{75}\) In fibrinogen, the \(\alpha\)C region is very flexible, and no crystal structure information is available for this region. Previous studies have relied on NMR spectroscopy to obtain clues about these structural features of the \(\alpha\)C region.\(^{28, 32}\) Similar strategies were adopted for the current studies with Fibrinogen \(\alpha\)C (233-425).

**Heteronuclear Single Quantum Coherence Spectroscopy (HSQC)**

In protein chemistry, two categories of two-dimensional (2D) NMR experiments are generally employed—homonuclear and heteronuclear correlated spectroscopy. Homonuclear experiments provide information about interactions between the same nuclei (e.g., proton-proton or carbon-carbon) and heteronuclear experiments monitor spins from different nuclei (e.g., proton-nitrogen and proton-
carbon). To produce magnetically sensitive nitrogen and carbon species, proteins are typically enriched (100%) using [U-$^{13}$C]glucose and [$^{15}$N]ammonium chloride to overcome low natural abundance levels for these nuclei species. Nonetheless, sensitivities of both $^{13}$C and $^{15}$N are still significantly lower than the sensitivity of protons ($^1$H).\textsuperscript{75, 76}

![Diagram of pulse sequence for a typical HSQC experiment](image)

Figure 19: Pulse sequence for a typical HSQC experiment. Each 90 and 180 degrees pulse applied is represented in the proton and nitrogen. The delay time applied, $\tau = 1/4J$ and the time of evolution $t$ shown. Polarization steps are shown with the INEPT.

In 2D HSQC experiments, the intense spin polarization of the proton is transferred to the heteronuclear spin ($^{15}$N for example), through a series of pulse sequences and a polarization step known as the INEPT (Insensitive Nuclei Enhanced by Polarization Transfer). To increase the sensitivity further, the polarization is transferred from the heteronuclear spin back to the proton for detection by reverse INEPT.\textsuperscript{75, 77} A representation of the pulse sequence for the HSQC experiment is shown in Figure 19. A typical 2D $^1$H-$^{15}$N HSQC spectrum consists of cross peaks at the chemical shifts of the $^1$H directly attached to an $^{15}$N. Unique chemical environments provide distinct peak positions in both the proton and nitrogen dimension (Figures 20 and 21). Compared to homonuclear
experiments, heteronuclear experiments produce less spectral overlap for each amino acid (i.e greater dispersion). In this dissertation, mainly heteronuclear experiments were employed predominantly $^1$H-$^{15}$N HSQC. A brief summary of this approach is described below.

**2D HSQC NMR experiment to monitor Factor XIII’s Crosslinking Reaction**

The HSQC experiments complement the mass spectrometry assay data. In the mass spectrometry assay, Factor XIII’s crosslinking ability is monitored by the addition of $86 \text{ m/z}$ to each peptide fragment containing a reactive glutamine. An isotopically labeled substrate was used in the complementary 2D $^1$H-$^{15}$N HSQC NMR approach to produce a signal that would monitor the transamidation reaction only. Factor XIII incorporates this labeled substrate into αC (233-425) when crosslinked to each reactive glutamine 1 (Figure 20). The $^{15}$N- labeled glycine ethyl ester 2, $[^{15}\text{N}]\text{GEE}$, is used for the 2D $^1$H-$^{15}$N HSQC experiment. $[^{15}\text{N}]\text{GEE}$ is covalent crosslinked to each reactive glutamine and is observed as a single peak. In the 2D $^1$H-$^{15}$N HSQC experiment, each peak corresponds to the $^1$H attached to the $^{15}$N upon crosslinking. The proton attached to the $^{15}$N-labeled nitrogen can be observed following the transamidation reaction. Due to the unique chemical environment adopted upon crosslinking, only species involved in the crosslinking reaction are observed. The unique chemical environment of each reactive glutamine is also reflected in the spectra obtained from the 2D $^1$H-$^{15}$N HSQC experiment.
Figure 20: Transamidation reaction catalyzed by Factor XIII monitored using a 2D $^1$H-$^{15}$N HSQC NMR approach. Factor XIIIa crosslinks $^{15}$N-labeled GEE to each reactive glutamine in Fibrinogen $\alpha$C (233-425). Three distinct peaks corresponding to the proton attached to the $^{15}$N appear upon crosslinking (Top Panel). Mutation studies confirm the identity of each reactive glutamine (Q237, Q328, and Q366) crosslinked and its unique chemical environment (Bottom Panels).

Figure 20 shows the reaction monitored in the complementary 2D $^1$H-$^{15}$N HSQC NMR experiment and results obtained from crosslinking the WT $\alpha$C (233-425) containing all three glutamines. The top right panel shows three peaks corresponding to all three reactive glutamines crosslinked by Factor XIIIa. The identity of each individual reactive glutamine was confirmed following point mutations of two glutamines (Q) to asparagines (N) (Figure 20, Bottom panel). Each peak appears in a unique position as a result of the chemical environment surrounding these reactive glutamines.
The complementary NMR assay was also tested on the fibrinogen gamma module γ(148-411). Two distinct peaks corresponding to two reactive glutamines, Q398 and Q399, were observed (Figure 21). One peak was more intense than the other suggesting that one of the reactive glutamines is crosslinked to a greater extent than the other. Although these reactive glutamines are neighboring identical residues (Q398, Q399), their HSQC crosspeaks do not overlap. The ability to distinguish the surrounding chemical environments of these two residues is an additional strength for the NMR approach. Future studies with point mutations to one of the reactive glutamines will help to identify which peak is crosslinked to a faster extent.

![Complementary 2D HSQC NMR in fibrinogen gamma module (γ148-411). Two peaks corresponding to two reactive glutamines Q398 and Q399 are observed in unique positions corresponding to their chemical environments.](image)

Figure 21: Complementary 2D HSQC NMR in fibrinogen gamma module (γ148-411). Two peaks corresponding to two reactive glutamines Q398 and Q399 are observed in unique positions corresponding to their chemical environments.
Overall, the 2D $^1$H-$^{15}$N HSQC experiments provide a direct method to quickly monitor Factor XIII’s ability to crosslink reactive glutamines in fibrinogen. The lower sensitivity of NMR experiments compared to mass spectrometry poses a technical challenge. Higher protein concentrations are needed to observe these peaks in the NMR assay compared to mass spectrometry.

2D HSQC NMR for Structural Characterization

In the second NMR approach, the entire backbone and side-chain nitrogens in the protein were labeled to obtain global information about the protein’s structure. Fibrinogen αC (233-425) was expressed in minimal media containing $[^{15}$N]ammonium chloride. 2D $^1$H-$^{15}$N HSQC NMR experiments with this labeled sample confirm that Fibrinogen αC (233-425) is mostly intrinsically disordered (Figure 22). Intrinsically disordered proteins (IDPs) have a characteristic 2D HSQC NMR profile with peaks clustered at 8-9 ppm in the amide region. Although obtaining structural information from NMR studies is typical, the intrinsically disordered nature of the αC region of Fibrinogen makes this very challenging to achieve.

Fibrinogen αC (233-425), like most intrinsically disordered proteins, contains a high percentage of prolines. Prolines do not have an amide proton and are therefore not detected in the 2D HSQC experiments $^1$H-$^{15}$N. To obtain additional information for peak assignment, proteins are expressed with both $^{13}$C and $^{15}$N labels. Fibrinogen αC (233-425) containing both $^{13}$C and $^{15}$N label was expressed and purified. A combination of 2D and 3D experimental approaches to
characterize the unique structural features for Fibrinogen αC (233-425) are described in Chapter 5.

Figure 22: Representative backbone structure showing all labeled residues and the proline amide nitrogen highlighted (Top Panel). 2D HSQC NMR spectra from $^{15}$N-labeled Fibrinogen αC (233-425) showing characteristic profile for intrinsically disordered proteins.
CHAPTER 3

RANKING REACTIVE GLUTAMINES IN THE FIBRINOGEN αC REGION THAT ARE TARGETED BY BLOOD COAGULANT FACTOR XIII

Introduction

Abnormal fibrinogen levels and fibrin clot structure have been associated with several pathological conditions including cardiovascular disease, arteriosclerosis, and bleeding disorders. The N-termini of the fibrinogen (AαBβγ)\textsubscript{2} chains form the central E region, which extends via a coil-coiled region to two terminal D regions. A flexible portion of the Aα chain—the αC region—extends from the ends of the D regions and is tethered to the central E region.

To initiate fibrin polymerization, thrombin cleaves fibrinopeptides A and B from the Aα and Bβ chains of fibrinogen to form fibrin monomers (αβγ)\textsubscript{2}. These cleavages lead to protofibril formation among fibrin monomers and also release the αC region.

At the final stage of blood coagulation, Factor XIIIa (FXIIIa) introduces γ-glutamyl-ε-lysyl isopeptide bonds between selective glutamines (Q) and lysines (K) within fibrin to form a clot with enhanced elastic properties. Moreover, the overall fibrin network has decreased fiber diameter, increased fiber density, reports of thinner fibers, and increased clot stiffness.
crosslinks fibrin γQ398 or γQ399 to fibrin γK406 thereby forming γ-γ dimers. Crosslinks involving the fibrin α chain appear later and involve the generation of α-α dimers, γ-α hybrids, and higher α-α polymers.55-57, 79 Five known reactive glutamines in the αC region include Q221 (and/or 223), Q237, Q328, and Q366. 59, 80-82 FXIIIa also crosslinks α2-antiplasmin (α2AP), plasmin activator inhibitor 2 (PAI-2), and fibronectin, into the fibrin clot network.21, 83-85 FXIII(a) and fibrinogen play additional supporting roles in the presence of red blood cells (RBC). FXIII binding to fibrinogen γ390-396 and subsequent crosslinking of the alpha chain helps mediate RBC retention in venous thrombi.46, 86

The αC region of fibrinogen contributes its own critical roles in the assembly and properties of clot structure.26, 31, 58, 87, 88 This region has also been directly correlated with adhesion events.89-91 Fibrinogen αC region (221-610) is composed of a flexible αC connector (221-391) and a more structured αC domain (392-610).28, 29 During fibrin polymerization, crosslinking in the αC region has been shown to promote lateral aggregation and protofibril staggering.26, 30, 31 Studies with native fibrinogen versus the mutant (γQ398N, γQ399N, αK406R) have revealed independent contributions coming from α-α crosslinks.59, 92 Such crosslinks play major roles in promoting clot stiffness, fiber straightening, and hindering fibrinolysis.

Truncations to the αC region have been shown to alter the nature of the clot formed.58 With Fibrinogen Otago, the Aα chain ends at P270 resulting in severe hypofibrinogenemia. Further analysis revealed that loss of (271-610) leads to impaired fibrin polymerization and the production of much thicker fibers. 58, 70, 88
Fibrinogen Seoul II (Q328P mutation) exhibits impaired fibrin polymerization and a reduced extent of α-α crosslinking due in part to loss of a reactive glutamine within the αC region.\textsuperscript{68}

Several studies have identified specific crosslinking sites in the αC region responsible for lateral aggregation, but little is known about the individual reactive glutamines and the roles played by the surrounding residues.\textsuperscript{55-57, 81, 82, 93, 94} Antibody studies by Procyk \textit{et al.}, identified a FXIIIa binding site within αC (242-424) and then narrowed the site more specifically to αC (389-402).\textsuperscript{95} Using recombinant αC (233-425), Smith \textit{et al.} demonstrated that the key contact region for FXIII A\textsubscript{2} (thrombin-activated) and A\textsubscript{2}B\textsubscript{2} involves αC (371-425) with E396 playing an important role.\textsuperscript{41, 72} In close proximity to this FXIII binding region, αC (233-425) contains three reactive glutamines (Q237, Q328, and Q366) that each become crosslinked to a selective set of lysines within the fibrin clot network.\textsuperscript{57, 82}

The aim of the current study was to characterize the glutamines within the fibrinogen αC region (233-425) that are known to be crosslinked by FXIIIa under physiological conditions. Matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometry and 2D Heteronuclear Single Quantum Coherence (HSQC) NMR assays were used to directly monitor FXIIIa-catalyzed reactions between the substrate glutamines in αC (233-425) and a set of lysine mimics.\textsuperscript{96} The results demonstrate that Q237 is the most reactive glutamine of fibrinogen αC (233-425) followed by Q366 and Q328. Moreover, no glutamine is dependent on another to react first in the series. New knowledge gained about
these three glutamines is considered relative to the conformational features of fibrinogen and to the mutants Fibrinogen Otago and Fibrinogen Seoul II.

**Materials and methods**

*Proteins and chemicals*

Human cellular FXIII was a generous gift from the late Dr. Paul Bishop. Stock aliquots of FXIII were prepared in 18 MΩ deionized water and the concentration measured on a Cary 100 Bio UV-Visible spectrophotometer using an extinction coefficient of 1.49 (ml mg⁻¹ cm⁻¹) and stored at -70 °C. Thrombin and glycine ethyl ester (GEE) were obtained from Sigma Aldrich (St. Louis, MO). ¹⁵NH₄Cl and [¹⁵N]GEE were obtained from Cambridge Isotope Laboratories (Andover, MA).

*Site-directed mutagenesis of αC (233-425) to generate Q to N substitutions*

The PGEX-6P-1 plasmid vector contained cDNA for GST-tagged αC fragment (233-425). This αC sequence was derived from the full-length fibrinogen Aα chain cDNA. Individual glutamine substitutions, Q237N, Q328N, and Q366N, were introduced into αC (233-425) using the QuickChangeII Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, CA). The resulting substitutions were confirmed with DNA sequencing. Plasmids encoding the different αC (233-425) variants were transformed into BL21 Gold (DE3) *E. coli* cells.
**Fibrinogen αC (233-425) expression**

GST-αC (233-425) and its mutants were expressed as described previously by Smith *et al.*\(^7\) with few modifications. Briefly, GST-αC expressing cells were incubated in Terrific Broth (1.2% typtone, 2.4% yeast extract, 0.4% glycerol, 742 mM K\(_2\)HPO\(_4\), 171 mM KH\(_2\)PO\(_4\), and 0.1 mg/ml ampicillin) at 37°C on a rotary shaker until the optical density at 600 nm reached 0.9. αC expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) followed by incubation at 30°C for 16 h. Cells were harvested by centrifugation at 5000 g at 4°C. Biomass was resuspended in Wash Buffer (20 mM Tris Base pH 8.0, 50 mM NaCl) and centrifuged at 5000 g at 4°C. Resulting pellets were stored at -20°C.

Frozen pellets were resuspended in phosphate buffer saline (PBS: 137 mM NaCl, 2.7 mM KCl, 10 mM Na\(_2\)HPO\(_4\), 2 mM KH\(_2\)PO\(_4\), pH 7.4) and incubated with 1 mM dithiothreitol and 1 mg/ml lysozyme, followed by the addition of 2 µg/ml aprotinin, 1 µM peptatin A, 10 µM leupeptin, 4 mM benzamidine hydrochloride, 0.5 mM phenylmethansulphonylfluoride (PMSF), 0.05% sodium deoxycholate, 0.02% triton X-100, 5 µg/ml Dnase I, and 5 mM MgCl\(_2\). The Dnase I reaction was quenched with 6 mM EDTA, and the lysate was centrifuged at 22,000 g at 4°C. Supernatant was supplemented with 1% (w/v) streptomycin sulfate and centrifuged at 22,000 g at 4°C. The resulting soluble fraction was passed through a 0.2 µm membrane filter and loaded on a GST-affinity column using an AKTAprime™ system (Amersham Biosciences, Piscataway, NJ).
An on-column procedure employing PreScission™ protease was used to cleave αC (233-425) from the GST-tag. The released αC (233-425) was eluted with PBS buffer. Complete cleavage of the GST tag was confirmed by SDS–PAGE and Western blot analysis (Amersham™ anti-GST-HRP conjugate antibody). The resultant αC (233-425) was concentrated in a 5,000 MWCO Vivaspin 2 concentrator (Sartorius, Goettingen). Protein concentrations were determined using an extinction coefficient of 41480 M\(^{-1}\)cm\(^{-1}\) calculated from ExPASy (www.expasy.org).

**MALDI-TOF mass spectrometry kinetic assay: Monitoring crosslinking reaction of glutamines in αC (233-425)**

The ability of FXIIIa to crosslink each reactive glutamine in αC (233-425) to the lysine mimic GEE was monitored using our previously optimized MALDI-TOF mass spectrometry kinetic assay.\(^{48}\) Final concentrations of 500 nM FXIII, 17 mM GEE, 4 mM CaCl\(_2\), and MALDI assay buffer [100 mM Tris-acetate, 150 mM NaCl and 0.1% PEG\(_{8000}\) pH 7.4] were incubated at 37°C in a 1.5ml reaction tube. The FXIII was activated with bovine thrombin (8.4 U/ml final) for 10 min and then 200 nM PPACK (D- phenylalanyl-L-prolyl-L-arginine chloromethyl ketone) was added to inhibit the thrombin. The crosslinking reaction between αC (233-425) and GEE was initiated by adding a final concentration of 13.6 µM αC (233-425) to a total assay volume of 250 µL. At different time points, a 25 µL aliquot of this reaction mixture was removed and quenched in 1.6 µL of 160 mM EDTA (10 mM final). Control experiments were performed in the absence of Factor XIII.
All samples were subjected to proteolytic digestion with chymotrypsin and GluC (Roche, Indianapolis, IN). For the digests, 6 µL of either the control or time point sample were combined with 6 µL chymotrypsin buffer (100 mM Tris-HCl, 10 mM CaCl$_2$ pH 7.4) and 1.5 µL chymotrypsin (1 µg/µL) and incubated for 1 h at 25°C. This digest was quenched with 2 µL 5% TFA. GluC digest was performed in a similar manner with 6 µL of the sample, 6 µL GluC buffer (25 mM NH$_4$HCO$_3$ pH 7.8) and 1.5 µL GluC (0.05 µg/µL) incubated for 2 h at 25°C and quenched with 5% TFA. All digested samples were zip-tipped and analyzed using a MALDI-TOF mass spectrometer (Applied Biosystems Voyager DE-PRO).$^{48}$ α-Cyano-4-hydroxycinnamic acid (αCHCA) matrix was employed for chymotrypsin digests and ferulic acid matrix for GluC. This assay was done in triplicate using αC (233-425) from three independent expression trials. The peak-height ratio method was utilized to determine the extent of crosslinking of GEE to each reactive glutamine with time. The amount of reactant left at each time was calculated as follows:

$$\frac{\sum \text{Reactant Peak Height}}{\sum \text{Reactant Peak Height} + \sum \text{Product Peak Height}} \times \text{Peptide concentration}$$

$^{15}$N-HSQC NMR spectroscopy: Validating αC (233-425) crosslinking with $^{15}$N-labeled substrates

FXIIIa-catalyzed crosslinking of $^{15}$N-labeled NH$_4$Cl or $^{15}$N[GEE] to reactive glutamines in αC (233-425) was monitored using $^{15}$N- heteronuclear single quantum coherence (HSQC) NMR spectroscopy.$^{96}$ All assays were carried out in a total volume of 400 µL with 10% D$_2$O added for the NMR deuterium lock. Final
concentrations are presented. For crosslinking reactions with $^{15}$N[GEE], 800 nM FXIII was proteolytically activated with 21 U/ml bovine thrombin in the presence of 5 mM CaCl$_2$ and 20 mM Borate buffer. Thrombin activity was quenched with 460 nM PPACK. 10 mM $^{15}$N[GEE] and 35 µM αC (233-425) were added and incubated for 1 h at 37°C. For the $^{15}$NH$_4$Cl exchange reactions, 400 nM FXIII was non-proteolytically activated in the presence of excess calcium (50 mM) and 20 mM borate buffer (pH 8) for 10 min at 37°C. Then, 100 mM $^{15}$NH$_4$Cl and 40 µM αC (233-425) were added and the mixture was incubated for 1 h at 37°C. 10% D$_2$O was added and the mixtures transferred into Shigemi NMR tubes. All reactions were analyzed at 25°C using a single z-axis gradient triple resonance (HCN) cryoprobe run on a 700 MHz Varian Inova NMR spectrometer. The parameters for the 1D HSQC were number of transients (nt) = 512, number of increments (ni) = 1, number of points (np) = 2048, and sweep width (sw) = 7022.5. For the 2D HSQC, the parameters used were nt = 64, ni = 64, np = 2048 and sw = 7022.5. The spectra were processed using NMRPipe and nmrDraw.
Results

Preparation and characterization of recombinant αC (233-425)

Recombinant αC (233-425) containing an N-terminal GST-tag and a PreScission™ protease cleavage site was overexpressed in E.coli and purified. SDS-PAGE results and Western Blot analysis using an anti-GST antibody confirmed that the absence of GST-αC and free GST protein in the αC samples. The amino acid sequence of αC (233-425) can be found in Figure 23A.

MALDI-TOF mass spectral analysis of separate chymotrypsin and GluC digests of αC (233-425) showed peptide fragments containing all three reactive glutamines. Theoretical protease digests (Protein Prospector, UCSF) and MS/MS analyses were used to identify the αC fragments. The chymotrypsin digest exhibited a peak containing Q328 (2448 m/z) (Figure 23B) whereas the GluC digest showed peaks for Q366 (1349 m/z) and for Q237 (1550 m/z) (Figure 23C).
Figure 23: Proteolytic digests of fibrinogen αC (233-425) show three reactive glutamines in a MALDI-TOF MS. (A) αC (233-425) sequence showing position of reactive glutamine in bold (B) MALDI-TOF MS spectrum of αC (233-425) digested with chymotrypsin contained a fragment with one reactive glutamine Q328 at 2448.7 m/z (C) GluC digest of αC (233-425) showed two peaks (1550.5 m/z and 1349.3 m/z) containing Q366 and Q237, respectively in the MALDI-TOF MS

Monitoring crosslinking of Q237, Q328, and Q366 with GEE

Our MALDI-TOF mass spectrometry kinetic assay was used to monitor FXIIIa-catalyzed crosslinking of each reactive αC (233-425) glutamine to the lysine mimic GEE. The crosslinking reaction was quenched at distinct time points and then digested separately with appropriate protease. A representative set of mass spectral data for a chymotrypsin (Figure 24A) and a protease GluC digest (Figure 24B) show the reactant peak containing the Q of interest and the
subsequent product peak (+86 m/z for the new Q-GEE). No reactions products were observed in the absence of FXIIIa.

Figure 24: Crosslinking of reactive glutamines in fibrinogen αC (233-425) with lysine mimic glycine ethyl ester (GEE). Representative MALDI-TOF spectra showing crosslinking reaction between reactive glutamine in αC (233-425) and the lysine mimic GEE catalyzed by FXIIIa (A) Chymotrypsin digest at time points 0, 5, 10 and 15 min shows that the fragment peak containing reactive glutamine Q328 (2448.5 m/z) was crosslinked to GEE by FXIIIa to form a product peak Q328-GEE (2534.5 m/z). Each spectrum was obtained from the same starting assay and represents a quenched time point followed by a chymotrypsin digest analyzed using MALDI-TOF MS (B) A similar plot with GluC digest at time points 0, 0.4, 5, and 10 mins shows reactant and product peaks containing Q237 and Q366 and their respective GEE-crosslinked products (1636.5 m/z and 1434.3 m/z). Q237 is mostly crosslinked to product by 20 seconds as compared to Q366, which was only crosslinked to product by 10 mins.
The amount of reactant left after each time point was calculated using the peak height ratio method. As shown in Figures 24B and 25, the peptide containing Q237 was rapidly crosslinked by 20 seconds and completely depleted by 5 minutes of reaction with GEE. FXIIIa was able to very effectively crosslink the lysine mimic GEE to the Q237 side chain amine. Q328 and Q366 could also undergo FXIII-catalyzed crosslinking with GEE but at a slower rate than Q237. A plot displaying all three αC glutamines indicates that the order of reactivity toward GEE crosslinking was Q237 >> Q366 ≈ Q328. Reactions with Q366 and Q328 could be brought closer to completion by increasing the FXIIIa concentration by 3-fold and the incubation time to 120 minutes (Figure 25B). When the original FXIIIa concentration was reduced 15-fold, the Q237 reaction rate became much slower with a 50% loss of reactant occurring by 5 minutes (data not shown). There have been previous reports on FXIIIa-catalyzed crosslinking pairs that involve α-chain glutamine residues and their respective lysine partners. Using the current assay strategy, the individual reactivities of each glutamine could, for the first time, be directly monitored.
Figure 25: Plot showing combined reactivities of Q237, Q328 and Q366 in fibrinogen αC (233-425)
(A) A combined graph showing the rate of consumption of all three reactive glutamines Q237 (diamonds ♦), Q328 (squares ■), and Q366 (triangles ▲). The peak height ratio method was used to calculate the amount of reactant left in triplicate experiments and plotted as mean ± STD. These reactive glutamines were ranked as Q237 ≈ Q328 > Q366. (B) An increase in FXIIIa concentration by 3-fold shows that Q366 reacts comparably with Q328 even at higher time points.

2D $^{15}$N HSQC NMR confirms that $^{15}$N labeled substrates crosslink to reactive glutamines

$^{15}$N-labeled substrates have been used to follow reactions and conformational changes in proteins. Here, 2D HSQC NMR was used to further examine the crosslinking of $^{15}$N-labeled substrates ($^{15}$N[GEE] and $^{15}$NH$_4$Cl) to the reactive glutamines on αC (233-425). With the $^{15}$N[GEE] reaction, the magnetically silent
$^{14}$NH$_2$ on the $\alpha$C glutamine side chains was replaced with NMR active $^{15}$NH amide product that is part of the FXIIIa-catalyzed GEE dependent reaction (Figure 26A). The 2D HSQC spectrum of the FXIIIa-catalyzed reaction shows three crosspeaks with each peak corresponding to an amide $^1$H (8.4-8.1 ppm) directly attached to a $^{15}$N (114 ppm range) (Figure 26B). The observed HSQC peaks revealed that FXIIIa had successfully crosslinked $^{15}$N[GEE] into the three reactive glutamines within $\alpha$C (233-425). One crosspeak clearly had an intensity greater than the other two (Figure 26B). Additional HSQC studies demonstrated that $^{15}$NH$_4$Cl could also be utilized as a lysine replacement (Appendix 3).

HSQC experiments with $^{15}$N[GEE] were then carried out using $\alpha$C (233-425) mutants in which a single Q was replaced each time with a non-reactive N. As shown in Figure 26C, each mutant displayed two as opposed to the three crosspeaks seen in wild-type $\alpha$C (233-425). A comparison of chemical shift positions and the mutant employed revealed that the strong intensity peak at (114 ppm, 8.35ppm) corresponds to the Q237. Q328 occurs at (113.8 ppm, 8.20 ppm) and Q366 at (113.9 ppm, 8.16 ppm) (Figure 26B and 26C). The high intensity peak for Q237 matches well with the highest reactivity observed with the MALDI-TOF mass spectrometry kinetics assay and followed by Q366 and Q328.
Figure 26: Reactive glutamines in αC (233-425) are crosslinked by FXIIIa in wild-type αC and following single Q to N substitutions (A) Reaction scheme for FXIIIa-catalyzed crosslinking reaction of $^{15}$N-GEE with reactive glutamines in αC (233-425) as monitored by NMR spectroscopy. B) 2D $^1$H-$^{15}$N HSQC NMR spectrum for the reactions between αC (233-425) and $^{15}$N-GEE. The three NMR peaks observed correspond to successful crosslinking reactions with the three glutamines Q237, Q328, and Q366. C) FXIIIa-catalyzed reactions between $^{15}$N-GEE and the individual αC mutants Q237N, Q328N, and Q366N as monitored by 2D $^1$H-$^{15}$N HSQC NMR. In αC Q237N, peaks corresponding to Q238 and Q366 were seen (top). Q237 reacted fastest following either Q328N or Q366N substitutions (middle and bottom panels).
Discussion

Within a blood clot, FXIIIa introduces γ-glutamyl-ε-lysyl crosslinks between specific reactive glutamines and lysines located in the fibrin γ and α chains.\textsuperscript{25, 40, 57, 93} Little is known about the individual contributions of each reactive glutamine and how FXIII selects one glutamine over another. We characterized and ranked three reactive glutamines in fibrinogen αC (233-425) for their abilities to be crosslinked by FXIIIa to the lysine mimic GEE. MALDI-TOF mass spectrometry and NMR strategies were employed that would allow us to directly measure individual reactivities.

Several studies have demonstrated that FXIIIa first introduces γ-γ crosslinks into the blood clot followed by α-α crosslinks.\textsuperscript{61, 79} Putative binding sites for FXIIIa have been identified within α and γ fibrinogen chains, and such sites are located either upstream or downstream from the crosslinking glutamines.\textsuperscript{46, 72} γ-γ crosslinking plays key roles in fiber appearance time and the generation of fiber density whereas α-α crosslinking is a major determinant in fibrin clot stiffness and hindering fibrinolysis ability.\textsuperscript{38, 59} Interestingly, there are far more covalent Q-K pairs involving the α chains than the γ.\textsuperscript{82} FXIIIa crosslinks fibrin γQ398 or γQ399 to γK406.\textsuperscript{79} By contrast, the αC region contains multiple reactive glutamines [Q221 (and/or Q223), Q237, Q328, and Q366] and multiple reactive lysines [K418, K448, K508, K539, K556, K563, K580, and K601].\textsuperscript{81, 82, 93}

Curiously, there is no obvious consensus sequence for the Q-containing substrates of FXIIIa.\textsuperscript{52, 98, 99} The reactive glutamines are often found in flexible regions of the substrates. It is important to point out that not all freely available
glutamines are good FXIII substrates. Fibronectin contains three glutamines toward its N-terminus but only Q3 is highly reactive, minor crosslinking involves Q7 and Q9. α2-Antiplasmin (N1) contains two reactive glutamines near the N-terminus but only Q2 is utilized. Reactive glutamines of α2-antiplasmin and PAI-2 become crosslinked to reactive lysines (K230, K303, and K413) within αC (233-425). α2-Antiplasmin is one of the few FXIIIa substrates that can be studied employing short peptides. Sequences based on α2-antiplasmin (N1, 1-15) have been used to characterize individual glutamines, evaluate roles of surrounding residues, and assess effects of peptide length. By contrast, similar sized peptides based on the fibrinogen Aα and γ chains make poor Q-substrates for FXIIIa. These observations suggest that fibrinogen chains require a larger, more protein-like environment to promote substrate specificity.

αC (233-425), containing close to 200 amino acids, has been shown to be a highly promising system for probing the FXIIIa substrate specificity toward the Aα chain. αC (233-425) has three reactive glutamines and a putative FXIIIa binding site. GEE serves successfully as a lysine mimic and targets physiological, reactive glutamines. Unlike spectrophotometric assays that indirectly monitor crosslinking activity or SDS PAGE gels that provide more general conclusions about crosslinking, our mass spectrometry kinetic assay directly follows each reactive glutamine. By focusing on αC (233-425), the challenges of working with multiple competing glutamines and lysines both on fibrinogen Aα and γ chains were also avoided.
Using the MALDI-TOF mass spectrometry assay, we were able, for the first time, to rank the individual glutamine players in fibrinogen αC (233-425). Results revealed that Q237 was the most reactive followed by Q366 and Q328. Our complementary NMR techniques validated these findings. Interestingly, the αC region is mostly disordered and no X-ray crystallographic information is available. Despite this disordered nature, all three glutamines in αC (233-425) could still serve as FXIIIa substrates. Other FXIIIa substrates may take advantage of even more defined structural features. Fibronectin, α2-antiplasmin, and PAI-2 each contain a flexible region for the reactive glutamine(s) that is further supported by a folded protein architecture. Unlike Aα based peptides, we propose that the larger αC (233-425) exhibits some localized conformational environment that allows or promotes interactions with FXIIIa.

The αC reactivity ranking Q237 >> Q366 ≈ Q328 provides new information to help interpret crosslinking pairs reported recently for full length fibrinogen.82 For that study, the FXIIIa catalyzed crosslinking reactions were followed by trypsin digests and mass spectral analysis. Wang demonstrated that Q237 could be crosslinked to more lysine pairs (αK418, αK508, αK539 αK556 and αK601) than all other reactive glutamines that were probed.82 The high reactivity documented for Q237 suggests that this amino acid is well positioned to crosslink a variety of reactive lysines. Unlike Q237, Wang reported that αC Q366 only crosslinks to K539. It should be noted that Q328 was not identified in Wang’s study. Previous investigations, however, have shown that Q328 is an important reactive glutamine.68, 81 In this assay, we were able to characterize Q328 following a
chymotrypsin digest. The ability to monitor all three reactive glutamines simultaneously in αC (233-425) is enhanced by digesting each sample time point separately with chymotrypsin and Glu C (Figure 23).

Using an NMR assay, we were able to confirm that FXIIIa crosslinks both $[^{15}\text{N}]$GEE and $^{15}\text{NH}_4\text{Cl}$ to reactive glutamines (Figure 26 and Appendix 3) in a similar order as seen in the MALDI-TOF kinetic assay. Furthermore, FXIIIa still crosslinks Q237 fastest following Q328N or Q366N mutations (Figure 26). Also in the absence of the most reactive glutamine Q237N, both Q328 and Q366 are shown to be crosslinked to GEE by FXIIIa (Figure 26). Thus, FXIIIa does not need to crosslink the fast acting αC Q237 before moving on to the other reactive glutamines. Further mutant combinations reveal that if any of the three Qs is replaced with a nonreactive N, the other two Qs can still participate in the crosslinking reactions (Figure 26). The data collected thus far suggest there is no obvious sequentiality in FXIIIa crosslinking.

Our current results can also be considered relative to disease-associated fibrinogen truncations and mutants. In Fibrinogen Otago, a major portion of the αC region (271-610) is absent and in Fibrinogen Keokuk, a single point mutation changes Q328 to a stop codon generating a larger truncation.\(^69,70\) In both cases, Q221, Q223, and Q237 are still available for crosslinking by FXIIIa but Q328, Q366, and multiple reactive lysines are missing. Our MS assay suggests that Q237 is highly reactive and the Wang studies revealed that this glutamine interacts with several lysines. FXIIIa may catalyze a heterodimeric reaction between fibrin αC Q237 and fibrin γK406. In normal fibrinogen, the α-γ
heterodimer has been reported to make up 2% of the fibrin crosslink contributions, but the reactive Q and K partners have not been identified.\textsuperscript{104} αC Q237 (and/or Q221, Q223) might therefore help support the dominant fibrin γ-γ crosslinks observed in SDS-PAGE gels of Fibrinogen Otago treated with FXIIIa and calcium.\textsuperscript{68}

In Fibrinogen Seoul II, a single point mutation results in αC Q328 being replaced with proline (Q328P). With this mutation, fibrin γ-γ crosslinking occurs but α-α is impaired.\textsuperscript{68} Considering our data, Q237 and Q366 would still be available in the Seoul II patients along with potential K residues in the more C-terminal portion of αC. Multiple crosslinks involving highly reactive Q237 may, however, not be sufficient to overcome the introduction of Q328P.

Recombinant variants that model Fibrinogen Seoul II revealed that Q328P/Q366P exhibited greater impairment of fibrin polymerization than the single mutants Q328P or Q366P.\textsuperscript{71} The current work with αC (233-425) suggests Q328 and Q366 exhibit similar reactivities. With each single P substitution, the pyrrole group may induce a conformational change that hinders fibrin architecture and subsequent crosslinking. A greater disruption occurs with the double mutant. Fibrin α-α crosslinking, however, was still observed and likely involved Q237, Q221, and/or Q223. Unlike working with Q328P and Q366P, Q to N substitutions within αC (233-425) provides a more conservative strategy for understanding the contributions of the different reactive Q residues. Moreover, any glutamines still present within αC mutants can be individually monitored and kinetically ranked. Q237 always remained the most reactive glutamine but no obvious sequentiality
in FXIIIa crosslinking was observed. Park and coworkers propose that conformational problems resulting from P328 help explain the consequences of Seoul II. Another possibility to consider is that crosslinks containing Q328 (and/or Q366) play structural roles within the fibrin clot that differ from those of Q237, Q221, and Q233. These issues can be further probed in the future using full-length fibrinogens.

In summary, we have characterized and ranked three known reactive glutamines Q237, Q328, and Q366 in fibrinogen αC (233-425) using a MALDI-TOF mass spectrometry kinetics assay and a complementary 2D HSQC NMR approach. αC Q237 was shown to be the most reactive glutamine followed by Q366 and Q328. Moreover, FXIIIa can independently crosslink these reactive glutamines. Additional studies will help better understand the individual crosslinking contributions from each reactive Q and their unique effects on clot properties. Continuing investigations on the αC region are anticipated to provide new clues about sources of FXIII substrate specificity and their applications in future drug or assay substrate designs.¹

¹ This chapter is reproduced in part from the published manuscript: Mouapi, K. N., Bell, J. D., Smith, K. A., Ariens, R. A., Philippou, H., and Maurer, M. C. (2016) Ranking reactive glutamines in the fibrinogen alphaC region that are targeted by blood coagulant factor XIII, Blood 127, 2241-2248.
CHAPTER 4

INVESTIGATING THE ROLE OF KEY MUTATIONS IN FIBRINOGEN αC (233-425) IN CONTROLLING FACTOR XIII’S ABILITY TO CROSSLINK REACTIVE GLUTAMINES (Q237, Q328, Q366)

Introduction

In the blood coagulation cascade, three proteins—fibrinogen, Factor XIII and thrombin—play important roles at the end of the cascade to form a stable clot that is resistant to fibrinolysis. The most abundant of these three clotting factors is fibrinogen. Fibrinogen has been shown to play multiple roles in the clotting process. First, the fibrin forms a meshy network at the site of injury and is important in the formation of the haemostatic plug. Fibrinogen also provides a suitable surface for binding several components required for the stability of a clot including platelets and antifibrinolytic enzymes. Recent studies have emphasized the role of the αC region of fibrinogen in promoting several functions of the fibrinogen molecule.

Duval et al. showed that Factor XIII-catalyzed crosslinking reactions involving the αC region provided unique contributions to the overall clot stability. Fibrin α-chain crosslinking resulted in increased fiber thickness and tautness and
decreased fibrin clot lysis rate.\textsuperscript{59} Crosslinking in the αC region was found to increase clot stiffness to a greater extent than crosslinking in the γ chain. Byrnes et al. also demonstrated that in the absence of the αC region, red blood cell retention in the fibrin clot is decreased.\textsuperscript{86} Together, these studies confirm the unique functional contributions of the αC region of fibrinogen. Few reports have elucidated the relationship between Factor XIII’s ability to bind to the αC region and the role of the binding site in promoting Factor XIII’s crosslinking activity.

Previous studies by Credo et al. showed that activation of plasma Factor XIII A\textsubscript{2}B\textsubscript{2} was enhanced by a segment on the αC region, αC (242 - 424).\textsuperscript{106, 107} Procyk et al. later localized the placental Factor XIII A\textsubscript{2} binding site on fibrinogen to αC (389-402) using antibodies specific for segments of the Aα, Bβ, and γ chains of fibrinogen. Of all the antibodies monitored, the anti-Aα 389-402 was shown to significantly decrease Factor XIII’s ability to bind to fibrinogen suggesting that the binding site for Factor XIII was located within this region.\textsuperscript{95} Using Surface Plasmon Resonance (SPR), Smith et al. and showed that plasma Factor XIII A\textsubscript{2}B\textsubscript{2} binds tighter than Factor XIII A\textsubscript{2} to Fibrinogen αC (233-425). Smith et al. also demonstrated that αC (389-402), with the sequence (PDWGTFEEVSGNVS), is highly conserved across different species.\textsuperscript{72} Residue E396 (highlighted in bold above) is located within this binding site and was proposed to promote Factor XIII’s binding interaction with fibrinogen. Smith et al. later showed that a synthetic peptide (Pep1-PDWGTFEEVSGNVS) containing E396 and the surrounding amino acids derived from αC (389-402) substantially inhibited Factor XIII’s binding interaction with αC and delayed crosslinking.
activity. Other synthetic peptide segments that were derived from Pep1 containing the αC E396A mutation and scrambled amino acids did not result in the same inhibitory effect. Together, these previous findings suggest that E396 and other surrounding residues promote Factor XIII’s ability to bind to the fibrinogen molecule. Although Factor XIII binding was significantly reduced in the Fibrinogen αC E396A, it remains unclear whether this reduction also alters Factor XIII’s ability to crosslink each reactive glutamine.

In the previous chapter, crosslinking of individual reactive glutamines Q237, Q328, and Q366 in Fibrinogen αC (233-425) were characterized and ranked using a combination of mass spectrometry and NMR spectroscopy assays. Our results suggested that the reactive glutamines are crosslinked to different extents, now ranked as Q237 >> Q366 ≈ Q328. This ranking aligns with the number of glutamine–lysine crosslinking pairs identified by Wang et al. in full-length fibrinogen. Q237 was shown to be involved in more lysine crosslinking pairs than either Q366 or Q328. In this chapter, the MALDI kinetic assay used to further characterize the role of a critical αC residue, E396, located within the binding site of Factor XIII. A combination of mass spectrometry and NMR was used to monitor crosslinking of each reactive glutamine in the presence of the E396A mutation within Fibrinogen αC (233-425). Our findings suggest that in the presence of Fibrinogen αC (233-425) E396A, Factor XIII can still crosslinks all three reactive glutamines to similar extents as compared to Fibrinogen αC (233-425) WT. This effect was the same whether recombinant
Factor XIII A₂ or plasma Factor XIII A₂B₂ were used in the presence or absence αC E396A.

To further probe the role of each individual reactive glutamine, mutations Q237N, Q237N/Q328N, and Q237N/Q366N were introduced into WT αC (233-425) and crosslinking patterns involving the remaining Q328 and Q366 residues were monitored. These two reactive glutamines are located upstream from the Factor XIII binding site αC (389-402) and were previously modeled to experience the most effect in the presence of Pep1. For this study, the mass spectrometry kinetic assay was further optimized for use on a Waters Synapt LC – MS system for higher throughput. Reactive glutamines, Q328 and Q366, were monitored individually and in the presence of the Q237N mutation. Studies with Q237N showed that Q366 is crosslinked to different extents in the absence of one or more reactive glutamines.

**Materials and Methods**

**Proteins and Materials**

Factor XIII A₂ was a generous donation from the late Dr. Paul Bishop. Factor XIII A₂B₂ was purchased from Enzymes Research Laboratories. Thrombin, glycine ethyl ester and other assay components were purchased from Sigma Aldrich. ¹⁵N-labeled glycine ethyl ester was purchased from Cambridge Isotopes.

Fibrinogen αC (233-425) WT (Fbg αC WT) was expressed and purified as described in Chapter 2. To insert an E396A mutation into this fibrinogen segment, a single nucleotide base change was required. The codon GAG for E396 was mutated to GCG to code for alanine (A396). The following primers
were designed for this purpose, Forward 5’-CTGGGGCACATTGAA**GCG**GTGAGGAAATG-3’ and Reverse 5’-CATTTCCTGACACCC**G**TTCAAATGTGCCCCAGTCTGG-3’. The QuikChange II site-directed mutagenesis kit was used to insert this mutation using the WT cDNA as the template. The E396A point mutation was confirmed by DNA sequencing. This DNA was then transformed into BL21-Gold(DE3) *E.coli* cells, expressed, and purified under the same conditions as the WT Fbg αC (233-425). For simplicity, Fibrinogen αC (233-425) expressed with E396A is referred in this chapter as Fbg αC E396A. Other variants of Fibrinogen αC (233-425) used in this Chapter include: Q237N, Q237N/Q366N (or “Q328 only”), and Q237N/Q328N (or “Q366 only”). The primers employed to introduce these Q to N substitutions are listed in Appendix 1. All variants of Fbg αC (233-425) were expressed and purified using a GST affinity trap column and on-column cleavage using human rhinovirus 3C protease with a GST tag. SDS PAGE and Western Blots were used to confirm the purity of protein fractions collected. Protein yields in the 3-6 mg/L range were obtained. When necessary, protein fractions were concentrated in a 5000 MWCO Vivaspin 2 concentrators. Protein concentrations were determined using the absorbance at 280 nm and an extinction coefficient of 41480 M⁻¹cm⁻¹. The presence of each mutation in the protein was confirmed by analyzing proteolytic digests (GluC or chymotrypsin) of protein fractions using MALDI-TOF mass spectrometry.
2D $^1$H-$^{15}$N HSQC NMR Assay

To test the ability of Factor XIIIa to crosslink all three reactive glutamines in Fbg αC E396A, the 2D HSQC NMR experiment was performed as previously described in Chapter 2. Briefly, Factor XIII A$_2$ was activated by incubating 800 nM Factor XIII with thrombin (21 U/mL) and CaCl$_2$ (5 mM) for 10 min at 37˚C. Thrombin was quenched using 200 nM PPACK. The crosslinking reaction was initiated by adding the lysine mimic $^{15}$N-labeled glycine ethyl ester (10 mM $[^{15}$N]GEE) and Fbg αC E396A (40 µM) to activated Factor XIII. The total assay volume was 400 µL in 20 mM borate buffer (pH 8). This reaction mixture was incubated for 30 min at 37˚C. D$_2$O (10%) was added to the sample for the NMR deuterium lock and then transferred to a Shigemi NMR tube.

This reaction was analyzed using a Varian Inova 700 MHz NMR spectrometer at 25˚C with a single z-axis gradient triple resonance (HCN) cryoprobe. The following parameters were used to obtain the 2D HSQC results; nt = 64, ni = 64, np = 2048 and sw = 7022.5. The spectrum was processed using NMRpipe and nmrDraw.

MALDI-TOF mass spectrometry kinetic assay

The previously described MALDI–TOF mass spectrometry kinetic assay was used to monitor the Factor XIIIa-catalyzed crosslinking reaction between reactive glutamines in Fbg αC (233-425) and the lysine mimic GEE (Chapter 2). This assay allowed us to compare crosslinking of each reactive glutamine in Fbg αC WT versus E396A. The assay was performed as described previously with a
few modifications. Factor XIII A$_2$ or Factor XIII A$_2$B$_2$ (500 nM) was incubated in the presence of GEE (17 mM) and Fbg αC WT or E396A (13.6 µM final concentration) for 5 min at 37˚C in MALDI Kinetic buffer (100 mM Tris acetate, 150 mM NaCl and 0.1% PEG 8000 pH 7.4). To activate Factor XIII A$_2$ or Factor XIII A$_2$B$_2$ and thereby initiate the reaction, thrombin (8.4 U/ml) and CaCl$_2$ (4 mM) were added to the reaction mixture. At different time points, 25 µL aliquots from this reaction were removed and quenched in 1.6 µL of 160 mM EDTA (10 mM). Samples from this reaction were digested separately with Chymotrypsin and GluC and analyzed on the MALDI TOF Voyager DE/Pro mass spectrometer as previously described in Chapter 2. The peak height ratio method was used to analyze and compare crosslinking in Fbg αC WT versus E396A.

Liquid chromatography- mass spectrometry (LC-MS) kinetic assay

Our previous NMR HSQC studies (Chapter 3) suggested that reactive glutamines can be crosslinked in the absence of one or more of these glutamines. In this initial study, one or more reactive glutamines (Q) where mutated into an inactive asparagine (N) residue. An LC-MS analysis approach was used to monitor crosslinking in Fbg αC (233-425) with the following Q mutations: Q237N, Q237N/Q366N, and Q237N/Q328N. Factor XIII’s ability to crosslink Q328 and Q366 in these variants was compared with Fbg αC WT. The kinetic assay was performed in the same manner as described for the MALDI TOF MS assay above. Final concentrations of Factor XIII A$_2$ and Fbg αC (233-425) WT and its variants were maintained at 500 nM and 13.6 µM, respectively.
Aliquots quenched from different time points were digested separately with Chymotrypsin and GluC and analyzed using an LC-MS approach described below.

Digested samples were prepared as follows for analysis on the LC-MS system. A P10 ZipTip® with 0.6 µL of C18 resin (Merck Millipore Corp., Billerica, MA, USA) was used to prepare the samples. One volume of each sample was diluted with three volumes of 2% v/v acetonitrile (ACN) / 0.1% v/v trifluoracetic acid (TFA) to a final volume of 30 µL. A ZipTip® was wetted with 3 x 10 µL 80% v/v ACN / 0.1% v/v TFA, equilibrated with 3 x 10 µL 2% v/v ACN / 0.1% v/v TFA, the sample was cycled in the tip 15 x 10 µL, the tip was washed with 4 x 10 µL 2% v/v ACN / 0.1% v/v TFA, and the sample was eluted into 30 µL 80% v/v ACN / 0.1% v/v formic acid. The samples were frozen and then dried in a SpeedVac. The dried samples were dissolved in 30 µL 2% v/v ACN / 0.1% v/v formic acid and 4 µL of each sample were analyzed.

A 250 µm ID x 15 cm Radel R tube (Idex Health & Science LLC, Oak Harbor, WA, USA) was packed in-house with Aeris Peptide 3.6µm XB-C18 material (Phenomenex, Torrance, CA, USA). An Acquity M-Class UPLC® system (Waters Corporation, Milford, MA, USA) was used with buffer A = water with 0.1% v/v formic acid and buffer B = acetonitrile with 0.1% v/v formic acid as mobile phases. Following injection of the sample onto the column, separation was accomplished with a 40min linear gradient from 2% B to 40% B, followed by a 5min linear gradient from 40% B to 85% B, and then a 5-minute wash with 85%
B. A Low Flow Electron Spray Ionization (ESI) probe (Waters) was used to introduce sample to a ZSpray LockSpray source (Waters).

A Synapt G2-Si mass spectrometer (Waters) was used to collect data from the LC eluate. An MSe Continuum method was created in MassLynx v4.1 SCN924 (Waters). The spectrometer was operated in positive resolution mode. A 1.5 sec low energy continuum scan without trap or transfer collision energy was acquired, followed by a 1.5 sec high energy continuum scan (15 to 40V ramp trap collision energy; no transfer collision energy). The scan range was 50 to 2000Da for GluC-digested samples and 50 to 3000 Da for chymotrypsin-digested samples. Leucine enkephalin (556.2771 Da/e) at 300 pg/µL in 1:1 acetonitrile : water was used as the lock mass during data acquisition. Skyline data analysis software was used to obtain precursor ions for each reactant peak and product peak species. The peak height ratio method was used to evaluate the extent of crosslinking of each reactive glutamine over time.

**Results**

*Factor XIIIa crosslinks all three reactive glutamines in Fibrinogen αC (233-425) E396A.*

A previously employed 2D $^1\text{H} - ^{15}\text{N}$ HSQC NMR assay was used to monitor Factor XIII’s ability to crosslink all three reactive glutamines in Fbg αC E396A. The three reactive glutamines (Q237, Q328, and Q366) in αC (233-425) are all located N-terminal to the E396A mutation. The 2D $^1\text{H} - ^{15}\text{N}$ HSQC NMR approach serves as a direct method to quickly access whether the loss of the acidic E396
residue in αC (233-425) hinders the reactivity of Factor XIII towards these three glutamines. With this heteronuclear NMR method, a single peak for each reactive glutamine that becomes crosslinked to [\(^{15}\text{N}\)]GEE is expected. This peak corresponds to the proton that is attached to the NMR active nitrogen (\(^{15}\text{N}\)). In the absence of crosslinking, no peaks should be observed.

Our results on Fbg αC E396A revealed three distinct peaks, with each 2D crosspeak appearing at a similar \(^1\text{H}\) and \(^{15}\text{N}\) chemical shift as observed with Fbg αC WT. The individual reaction products involving Q237, Q328 and Q366 from Fbg αC E396A are shown in Figure 27. The results suggest that the E396A mutation does not affect Factor XIII’s ability to crosslink all three reactive glutamines in Fibrinogen αC (233-425). Although the E396A mutation was previously shown to reduce binding for αC (389 - 403) to Factor XIII, this mutation did not inhibit Factor XIII’s ability to crosslink the reactive glutamines under the conditions employed in our assay. To monitor the crosslinking of each reactive glutamine to GEE over time, the MALDI-TOF MS kinetic assay was employed. With such an approach, more subtle differences in reactivities of the three glutamines could be probed.
MALDI-TOF mass spectrometry kinetic assay to monitor Q237

Amongst the three reactive glutamines in Fibrinogen αC (233-425), Q237 was the most reactive glutamine using our MALDI-TOF MS kinetic assay. However, using this assay design, Q237 was completely crosslinked to lysine mimic GEE within 10 minutes of the reaction. To address this issue, the assay was modified for shorter time points. Factor XIII’s ability to crosslink GEE to the reactive glutamine Q237 in Fbg αC WT versus Fbg αC E396A could now be better compared. Under these solution conditions, the observed reactivities for Q328 and Q366 would be very weak.

In the previous assay design (Chapter 3), Factor XIII A2 was pre-activated and the reaction was initiated by adding the glutamine substrate source Fbg αC (233-425). For this assay, Factor XIII (500 nM final) was incubated with Fbg...
αC WT or E396A, GEE, and MALDI kinetic buffer for 5 min at 25°C to promote Factor XIII binding and interaction with the Fibrinogen αC substrate. The assay reaction was then initiated by adding thrombin and calcium to activate Factor XIII. This experimental strategy is referred to here as the in-assay activation approach. The time points obtained were digested separately with Chymotrypsin and GluC as described in Chapter 3. The peak height ratio method was used to quantify the amount of reactant left with time from the reactant and product peaks obtained using the MALDI -TOF mass spectrometer.

Results demonstrate that although Q237 was crosslinked following the E396A mutation, the crosslinking activity appeared to be slightly slower (Figure 28). This assay was monitored every 2 minutes for WT and every minute for the E396A variant for a total of 10 min. Although the reaction with E396A appeared to be slower, this reaction was not substantially different from the WT. The results obtained indicate that full crosslinking for Q237 still occurs within 10 min for the Fbg αC E396A mutation (Figure 28).
Reactive glutamines Q328 and Q366 are crosslinked to similar extents in Fbg αC (233-425) WT and E396A.

To examine the effect of Fbg αC E396A on the other two reactive glutamines Q328 and Q366, a higher concentration of Factor XIII (2 µM) was used and reactivities monitored over an hour. Factor XIII was incubated with all reaction components using the in-assay activation approach described above. Thrombin and calcium were added to initiate the reaction, and aliquots were quenched at different time points during the course of the assay for comparisons between Fbg αC WT and Fbg αC E396A.
The results confirm that reactive glutamines Q328 and Q366, are crosslinked in Fbg αC E396A with a similar ranking as in the WT. During the first 20 minutes of the reaction, the amount of reactant left from reactive glutamines Q328 and Q366 in Fbg αC E396A appears to be slightly greater compared to the WT (Figure 29). This effect indicates that under the same reaction conditions, Factor XIII tends to crosslink reactive glutamines in Fbg αC E396A slightly slower than the WT. It is possible that in the presence of the αC E396A mutation, Factor XIII is not as well anchored to the binding site as observed with WT αC. Nonetheless, the ability for Factor XIIIa to crosslink all reactive glutamines, with a similar ranking to the WT, is still maintained following the αC E396A mutation. To further investigate the role of this mutation in controlling Factor XIII’s activation and subsequent crosslinking ability, this assay was performed with the plasma Factor XIII A2B2.
Figure 29: Comparing crosslinking for each reactive glutamine in WT versus E396A using 2µM Factor XIII A₂

**MALDI –TOF Kinetic assay comparing crosslinking with Factor XIII A₂B₂ in Fbg αC (233-425) WT and E396A.**

The Factor XIII binding site within Fbg αC (389-402) exhibits a higher affinity for Factor XIII A₂B₂ than Factor XIII A₂. Moreover, this binding site has been previously proposed to play an important role in the activation of Factor XIII A₂B₂. The MALDI-TOF MS Kinetic assay was used to rank reactive glutamines in Fbg αC WT versus E396A as zymogen Factor XIII A₂B₂ was converted to activated Factor XIIIa. During this activation process, the regulatory Factor XIII B₂ subunits are released, and the catalytic Factor XIII A₂ is activated. The assay was performed as described for Factor XIII A₂. Factor XIII A₂B₂ (500 nM final)
was incubated with Fbg αC (233-425) WT or E396A in the presence of MALDI kinetic buffer and glycine ethyl ester (GEE) using the in-assay activation approach. The reaction was initiated by adding calcium (4 mM) and thrombin (8.4 U/ml). Aliquots obtained at different time points were quenched and analyzed using the MALDI-TOF mass spectrometry approach. To compare the extent of crosslinking for the most reactive glutamine Q237, time points were collected with one-minute intervals (Figure 30).

Results suggest that Q237 can be crosslinked to GEE in the presence of Factor XIII A2B2 for both Fbg αC E396A and WT. Although the extent of crosslinking was slightly reduced in Q237 with Fbg αC E396A, this difference
was not significant when compared to the extent of crosslinking in the WT (Figure 30).

Reactive glutamines Q328 and Q366 were also monitored at higher time points and the extent of crosslinking compared in Fbg αC WT and E396A. Results from these experiments confirm that all three reactive glutamines can still be crosslinked with a similar ranking when plasma Factor XIII A2B2 is used (Figure 31). The presence of E396A does not affect the ranking of these reactive glutamines. Again, the extent of crosslinking for each reactive glutamine is slightly reduced in the E396A. However, this effect is not significantly different from the extent of crosslinking observed in the Fbg αC WT.

Figure 31: Comparing all three reactive glutamines crosslinked by Factor XIII A2B2 (500 nM) in Fibrinogen αC WT and E396A
Crosslinking trends in Q to N mutations differ slightly amongst reactive glutamines

The complementary NMR assay confirmed that in the absence of one or more reactive glutamines, following a Q to N mutation, Factor XIIIa is still able to crosslink any remaining reactive glutamines (Chapter 2 and 3). To monitor for differences in the extent of crosslinking for Fbg αC (233-425) containing one, two, or the original three reactive glutamines, an LC-MS based assay was employed. This LC-MS assay was adapted from the MALDI-TOF MS kinetic assay. The earlier NMR experiments provided a quick method to confirm that crosslinking occurred after a 30-minute incubation time for each Q to N species (Chapter 3). The LC-MS approach would now permit a more in depth analysis of the interplay of each individual reactive glutamine. Moreover, the LC-MS assay allows us to better monitor low intensity peaks and provides a faster throughput.

For these experiments, the pre-activated Factor XIII activation strategy described in Chapter 3 was used. Briefly, Factor XIII A2 (500 nM) was pre-activated in the presence of thrombin (8.4 U/ml) and calcium (4 mM). Crosslinking for each Q to N mutation was then monitored following separate Chymotrypsin and/or GluC digests. The LC-MS parameters used are detailed in Materials and Method Chapter 2. The retention times and m/z values observed for each peptide fragment containing a reactive glutamine are summarized in Table 2. The peak intensity for each precursor ion detected by the mass spectrometer was determined using Skyline data analysis software, and the peak
height ratio method was employed to compare the extent of crosslinking across different Q to N mutations.

Table 2: Average retention times and \(m/z\) values for peptide fragments containing reactive glutamines analyzed by LC-MS

<table>
<thead>
<tr>
<th>Reactive Glutamine monitored</th>
<th>Average retention time (min)</th>
<th>(m/z +1)</th>
<th>(m/z +2)</th>
<th>(m/z +3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q237</td>
<td>19.4</td>
<td>1549.6807</td>
<td>775.3440</td>
<td>517.2318</td>
</tr>
<tr>
<td>Q237-GEE</td>
<td>22.1</td>
<td>1635.7175</td>
<td>818.3624</td>
<td>545.9107</td>
</tr>
<tr>
<td>Q328</td>
<td>12.7</td>
<td>2448.0770</td>
<td>1224.5422</td>
<td>816.6972</td>
</tr>
<tr>
<td>Q328-GEE</td>
<td>18.3</td>
<td>2534.1138</td>
<td>1267.5606</td>
<td>845.3761</td>
</tr>
<tr>
<td>Q366</td>
<td>9.7</td>
<td>1348.5764</td>
<td>674.7918</td>
<td>450.1970</td>
</tr>
<tr>
<td>Q366-GEE</td>
<td>13.5</td>
<td>1434.6132</td>
<td>717.8102</td>
<td>478.8759</td>
</tr>
</tbody>
</table>

In the absence of the most reactive glutamine Q237 (i.e Q237N), the remaining reactive glutamines Q328 and Q366 were crosslinked to a similar but slower extent than the WT (Figure 32). Crosslinking of the most reactive glutamine Q237 may have a modest influence on the ability of Factor XIII to crosslink the other two reactive glutamines Q328 and Q366. Overall, reactive glutamines Q328 and Q366 experience the slowest extent of crosslinking in Q237N of all the αC series (Figure 32 and Figure 33).
To further probe for differences in the extent of crosslinking to Q328 and Q366, these reactive glutamines were monitored individually using the double mutations Q237N/Q366N (“Q328 only”) and Q237N/Q328N (“Q366 only”). Reactivities were compared for Q328 in Fbg αC WT (Q237, Q328, Q366), Q237N (Q328 and Q366) and finally with Q237N/Q366N (“Q328 only”). Of all three αC variants, Q328 was crosslinked the fastest with Fbg αC WT. By contrast, Q328 was crosslinked the slowest with Fbg αC Q237N where Q328 competes with Q366 for the lysine mimic GEE. When Q328 only (i.e Fbg αC Q237N/Q366N) was monitored, the extent of crosslinking was slow at the start of the reaction. At higher time points, the extent of crosslinking became comparable to the Fbg αC WT (Figure 32).
Figure 33: Comparison of crosslinking trends for Q328 in WT and following Q to N mutations

The reactivities of Q366 in Fbg αC WT, Q237N, and Q237N/Q328N (“Q366 only”) was also evaluated. Similar to Q328, the extent of crosslinking involving Q366 was significantly reduced in the absence of the most reactive glutamine Q237. Surprisingly, in the absence of the other two reactive glutamines (“Q366 only”), the extent of crosslinking involving Q366 occurred more rapidly over time (Figure 34). Intriguingly, at higher time points the reactivity of “Q366 only” became greater than Q366 in WT.

These studies suggest that the ability of Factor XIIIa to crosslink Q366 or Q328 is affected by the αC variant employed. With Fbg αC WT, Q366 begins to show greater reactivity than Q328 at higher time points. When Q237N is removed, Q328 and Q366 both exhibit similar slower reactivities than with WT. With the double Q to N mutation, “Q366 only” clearly exhibited greater reactivity.
than “Q328 only” (Figures 33 and 34). It remains unclear what features enhance the crosslinking extent in Q366.

Figure 34: Comparison of crosslinking trends for Q366 in WT and following Q to N mutations using LC-MS

Discussion

The αC region of fibrinogen has recently been shown to play several roles in formation of a stable clot.\textsuperscript{83, 106, 109} Factor XIIIa-catalyzed crosslinking in the αC region contributes to increased fiber thickness, increased tautness, and decreased clot lysis rate.\textsuperscript{59} Factor XIII has been shown to crosslink several reactive glutamines in this αC region compared to only two reactive glutamines in the Fibrinogen gamma chain. In our previous study (Chapter 3), a mass spectrometry and a complementary NMR spectroscopy assay were developed to monitor and rank the extent of crosslinking for three known reactive glutamines (Q237, Q366, and Q328) within a segment of the αC region (αC 233-425).\textsuperscript{108} Fbg
αC (233-425) was previously used to localize and characterize a Factor XIII binding site within the fibrinogen Aα chain. A residue within this αC region, Aα E396, was identified to be critical for the binding of Factor XIII. To further probe the role of E396, we introduced an E396A mutation in αC (233-425) and monitored Factor XIII’s ability to crosslink the reactive glutamines located upstream of its binding site αC (389-402).

Using this MALDI-TOF mass spectrometry kinetic assay and complementary NMR spectroscopy assay the presence of the E396A mutation in αC (233-425) was confirmed and Factor XIIIa was shown to still crosslink all three reactive glutamines. To ensure that Factor XIIIa was exposed to αC (233-425) prior to activation, an in-assay activation approach was adopted for all experiments comparing crosslinking in WT versus E396A. Compared to the pre-activation approach where activated Factor XIIIa was added to αC (233-425), in the in-assay activation approach, all components for Factor XIII activation were added in the presence of αC (233-425). No significant differences were detected in the ranking of reactive glutamines for the in-assay and pre-activated Factor XIII. At 500 nM, Factor XIIIa, Q237 was crosslinked to a slightly slower extent in E396A compared to WT at earlier time points. It is still unclear whether this delay is a result of the E396A mutation hindering Factor XIII’s ability to bind properly on this site.

A higher Factor XIII A2 concentration (2 µM) was used to monitor crosslinking in Q328 and Q366 over a longer time. These two reactive glutamines are located upstream of the Factor XIII binding region within Fbg αC
(389-402) and could experience larger differences in crosslinking upon introducing E396A. A molecular modeling study by Smith et al. predicted that the synthetic peptide (Pep1- PDWGTFFEEVSGNVS), derived from the αC (389-402) binding residues, would bind on Factor XIII. The results from this modeling study suggested that αC (389-402) binding directs Q366 to a position closer to the Factor XIII active site compared to Q328. The present results showed that although Q366 seemed to be crosslinked to a larger extent than Q328, the presence of the E396A mutation does not significantly affect the extent of crosslinking. Together, our findings suggest that the binding interaction promoted by the key amino acid E396 does not affect the ability of Factor XIII to perform its catalytic role.

Previous studies using a synthetic peptide (Pep1) derived from the αC binding site for Factor XIII showed that Pep1 delayed Factor XIII-catalyzed crosslinking. Smith et al. proposed that this delay in crosslinking was a result of Pep1’s inhibitory effect on a site associated with the Factor XIII activation peptide (AP) cleft. Molecular modeling studies using Pep1 suggested that E396 forms a salt bridge with Arg 158 on Factor XIII, which further emphasized the role of E396 in anchoring Factor XIII to this binding site. Surface Plasmon Resonance (SPR) studies demonstrated that non-activated Factor XIII A2B2 bound to αC (233-425) with a $K_d = 7.3$ nM, meanwhile, thrombin cleaved FXIII A2B2 bound with a $K_d = 21$ nM in the presence of calcium. The αC region has also been proposed to promote activation of Factor XIII A2B2. To investigate the effect of
E396 on Factor XIII’s activation and subsequent crosslinking, plasma Factor XIII A₂B₂ containing the carrier B-subunits was used.

The results from thrombin-cleaved Factor XIII A₂B₂ in the presence of calcium showed that the catalytic subunit of Factor XIII A can still crosslink all three reactive glutamines to a similar extent and ranking in both WT and E396A αC (233-425). This feature suggests that the activation of Factor XIII A₂B₂ is not affected by E396. Compared to recombinant Factor XIII A₂, crosslinking in plasma Factor XIII A₂B₂ was modestly slower for all three reactive glutamines. An SDS-PAGE gel analysis monitoring over time the activation of Factor XIII A in recombinant Factor XIII A₂ and Factor XIII A₂B₂ showed that Factor XIII A₂ was completely activated within 5 minutes, but Factor XIII A₂B₂ was activated fully only by 15 min (Appendix 4). This delay could be as a result of the presence of the carrier B-subunits. Nonetheless, activation did not affect the extent to which each reactive glutamine was crosslinked in both WT and E396A in αC (233-425).

We propose that although Factor XIII uses E396 as an anchoring site for binding, its catalytic activity does not depend on this residue alone. Other surrounding residues within the αC (389-402) must contribute to the overall binding interaction. A truncation of this region will provide more understanding on the role of this region in Factor XIII’s catalytic activity.

In addition to crosslinking all three reactive glutamines in Fibrinogen αC (233-425) WT, Factor XIIIa has been previously shown to crosslink each reactive glutamine independently following Q to N substitutions of one or more reactive glutamines (Chapter 2 and 3). In this study, changes in Factor XIIIa’s ability to
crosslink Q328 and Q366 were monitored using an LC-MS strategy. These results demonstrate that Q366 is most affected by Q to N substitutions. In the absence of Q237, both Q328 and Q366 are crosslinked the slowest compared to WT. Surprisingly, when both reactive glutamines Q237 and Q328 are substituted with inactive asparagines, Factor XIIIa crosslinks Q366 (Q366 only) to a greater extent compared to WT. In “Q328 only”, the differences across these mutations are not as apparent as in Q366. Previous molecular modeling studies proposed that Q366 is positioned closer to the active site compared to Q328. Factor XIIIa could therefore more readily recognize and crosslink Q366, in the absence of the competing Q237, since Q366 is located closest to Factor XIII’s binding site on Fbg αC (233-425). By contrast, the extent of crosslinking in WT, Q237N, and “Q328 only” were slightly different for Q328. These results suggest that Q328 is least affected following substitutions of other reactive glutamines. The central location of Q328 within the Fbg αC (233-425) may contribute to this effect. Q237 is located N-terminus of αC (233-425), but Q366 is located closest to Factor XIII’s binding site. Factor XIII may therefore use a combination of proximity to the binding site and substrate accessibility to select its preferred substrate for crosslinking.

Together our results suggest that Factor XIII uses a surrounding region within the binding site to promote its crosslinking activity. Although E396A was previously proposed to play an important role in Factor XIII’s binding activity, this substitution does not affect the catalytic role of either recombinant Factor XIII A₂ and plasma Factor XIII A₂B₂. More studies are needed to probe the role of αC
(389-402) in controlling Factor XIII binding and/or catalytic activity. Fbg αC (233-425) could be truncated to αC (233-388) thus removing the proposed Factor XIII binding site. Alternatively, additional mutations could be introduced into the αC (389-402) segment of αC (233-425).
CHAPTER 5

STRUCTURAL CHARACTERIZATION OF FIBRINOGEN αC (233-425) USING NMR SPECTROSCOPY

Introduction

Intrinsically disordered regions and proteins have been increasingly associated with diseases including cancer, Alzheimer’s, and cardiovascular disease. The αC region of fibrinogen is very flexible, and the N-terminal connector behaves very similar to intrinsically disordered proteins (IDPs). Truncations and mutations in this region have also been correlated with disease conditions associated with IDPs. Furthermore, many of the reactive glutamines crosslinked by Factor XIIIa are located in loop and flexible regions. In Chapters 3 and 4, three reactive glutamines in Fibrinogen αC (233-425) were ranked to understand Factor XIII’s ability to crosslink these reactive glutamines. The role of a mutation, E396A, within Factor XIII’s binding site was also explored.

Recent studies have proposed and demonstrated that some IDPs have Pre-Structured Motifs (PreSMos) known to have some degree of structure. Molecular Recognition Features (MoRFs)—short protein regions that undergo disorder-to-order transitions upon interaction with binding partners—have also been shown in IDPs. These types of structural features have been
hypothesized to enhance substrate binding and activity in IDPs. Whether the αC region undergoes structural changes upon binding Factor XIII is not known. Thus, the intrinsically disordered nature of Fibrinogen αC (233-425) is explored in this chapter and the presence of pre-structured motifs are investigated using NMR spectroscopy.

Table 3: Amino acid composition of Fibrinogen αC (233-425)

<table>
<thead>
<tr>
<th>Residue</th>
<th>Number of residues in expressed Fibrinogen αC (233-425)</th>
<th>Percent Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala (A)</td>
<td>4</td>
<td>2%</td>
</tr>
<tr>
<td>Arg (R)</td>
<td>12</td>
<td>6.1%</td>
</tr>
<tr>
<td>Asn (N)</td>
<td>13</td>
<td>6.6%</td>
</tr>
<tr>
<td>Asp (D)</td>
<td>4</td>
<td>2%</td>
</tr>
<tr>
<td>Cys (C)</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Gln (Q)</td>
<td>3</td>
<td>1.5%</td>
</tr>
<tr>
<td>Glu (E)</td>
<td>13</td>
<td>6.6%</td>
</tr>
<tr>
<td>Gly (G)</td>
<td>40</td>
<td>20.2%</td>
</tr>
<tr>
<td>His (H)</td>
<td>3</td>
<td>1.5%</td>
</tr>
<tr>
<td>Ile (I)</td>
<td>1</td>
<td>0.5%</td>
</tr>
<tr>
<td>Leu (L)</td>
<td>4</td>
<td>2%</td>
</tr>
<tr>
<td>Lys (K)</td>
<td>4</td>
<td>2%</td>
</tr>
<tr>
<td>Met (M)</td>
<td>3</td>
<td>1.5%</td>
</tr>
<tr>
<td>Phe (F)</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>Pro (P)</td>
<td>18</td>
<td>9.1%</td>
</tr>
<tr>
<td>Ser (S)</td>
<td>40</td>
<td>20.2%</td>
</tr>
<tr>
<td>Thr (T)</td>
<td>21</td>
<td>10.6%</td>
</tr>
<tr>
<td>Trp (W)</td>
<td>7</td>
<td>3.5%</td>
</tr>
<tr>
<td>Tyr (Y)</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>Val (V)</td>
<td>4</td>
<td>2%</td>
</tr>
</tbody>
</table>
Intrinsically disordered proteins are composed of a high percentage of glycines, prolines, serines, and threonines. As shown in Table 3, these residues make up 60% of the amino acid composition of Fibrinogen αC (233-425). A previous NMR study using fragments of the αC connector proposed that this region has residues with poly-proline type II (PxxP) propensity.\textsuperscript{114} An initial screening of Fibrinogen αC (233-425) using two disorder predictors PrDOS and PONDR\textsuperscript{115} predict that a majority of this sequence is intrinsically disordered (Figure 35). Interestingly, residues surrounding reactive glutamine Q366 and additional residues (DWGTE\textsuperscript{*}E) that contain E396 (*E) within Factor XIII’s binding region were predicted to have a more ordered propensity (Figure 35). Those residues exhibiting increased ordered propensities could contain PreSMos. Furthermore, the residues surrounding the Factor XIII binding site within αC (233-425) could contain MoRFs for Factor XIII. For this chapter, the intrinsically disordered nature of Fibrinogen αC (233-425) is explored and the presence of PreSMos investigated using NMR spectroscopy. Initial 3D NMR experiments for peak assignments were also examined.
Figure 35: Fibrinogen αC (233-425) sequence highlighting the number of glycines (blue) and prolines (mostly bold) and reactive glutamines (red). PONDR graphic representation of Fibrinogen αC (233-425) shows residues with higher score and probability to be disordered and a few residues with some degree of structural propensity (underlined in red and shown in graph). Residues from PrDOS are underlined in black.

Methods

IPAP NMR Experiments

2D $^1$H-$^{15}$N HSQC NMR experiments on the αC fragment were measured in the absence and presence of Pf1 phage alignment. This was followed by IPAP experiments, which were measured in the absence or presence of the alignment media Pf1 phage. Pf1 phage was purchased from ASLA biotech (distributed by CEDARLANE) at 1 g/mL stored in a solution. 500 μL of the phage
was obtained and centrifuged at 90,000rpm for 1 h at 4°C. The phage pellet obtained was resuspended and washed twice in the same NMR buffer used for the protein sample (20 mM K$_2$HPO$_4$, 150 mM NaCl, pH 6.5). After the washes, $^{15}$N–labeled αC (233-425) at a concentration of 200 µM was added to the phage pellet and then resuspended by vortexing. The final concentration of Pf1 phage in the sample was diluted to concentrations between 25-31 mg/mL (v/v), following manufacturers guidelines. 10% D$_2$O was added to this sample and then carefully transferred to a Shigemi NMR tube using a teflon tube attached to a pipette tip. The following parameters were employed to obtain the representative HSQC spectrum. Number of transients (nt) = 64, number of t$_1$ increments (ni) = 64, sweep width = 7022.5. IPAP experiments were measured using the 800 MHz Varian Inova with a triple resonance probe and pulsed field triple axis gradient. The data obtained from all experiments were processed using nmrPipe.$^{97}$ These experiments were performed with the help of T. Michael Sabo.

**Improved Protein Expression for $^{15}$N and $^{15}$N / $^{13}$C labeled αC (233-425)**

For structural characterization and backbone assignment of fibrinogen αC (233-425), the protein was expressed in minimal media containing $^{15}$NH$_4$Cl and $^{13}$C-glucose. The resultant αC (233-425) would thus be double labeled for solution NMR experiments. Using a procedure adopted from a previous study, a strategy to express high-yield labeled protein was optimized.$^{118}$ For these experiments, starter cultures were first grown in 40 mL LB Broth and then transferred into a larger LB culture (2- 4L) until cell growth attained an OD of 0.7.
Cells were then harvested by centrifuging at 5000 rpm. The cell pellets obtained were washed and resuspended in minimal media (M9 salts, MgSO₄, thiamine HCl, and glucose).

For the next steps, a 4:1 ratio was used for protein expression; i.e, pellets from each 4L culture in LB broth were used to inoculate 1L of minimal media containing 0.25 g/L of $^{15}$NH₄Cl and/or 4 g/L $^{13}$C-glucose. To boost cell growth, Bioexpress cell growth media with U-15N or U-15N/U-13C (Cambridge Isotopes Laboratories Inc, Andover, MA) respectively, was spiked into the media (500 µL volume for each 500 mL minimal media). This media significantly improves the yield of labeled protein. Cells grown in minimal media containing $^{15}$NH₄Cl or both $^{15}$NH₄Cl and $^{13}$C-glucose were harvested by centrifugation at 5000 rpm. The cells were then resuspended in PBS buffer and lysed in the same manner as unlabeled αC (233-425) (Chapter 2).

An alternative to this lysis method was sonication. The cell lysate was centrifuged at 13500 rpm for 20 min. The soluble supernatant was collected and then filtered through a 0.4 µm and then a 0.2 µm filter. This soluble supernatant was loaded unto three 5-mL GST Trap columns connected in series. The bound GST-αC was subjected to in-solution digestion with human rhinovirus 3C protease to cleave the GST-tag. The released αC (233-425) was collected as the flow-through from the GST-column, and SDS-PAGE was used to evaluate the purity of each fraction (2 mL). All other GST-tagged components were eluted using 20 mM glutathione and fractions collected.
Fractions containing $^{15}$N-labeled αC or doubly labeled $^{15}$N/$^{13}$C Fibrinogen αC (233-425) were individually combined and concentrated to approximately 150 µM for $^{15}$N/$^{13}$C Fibrinogen αC (233-425) and >200 µM for the $^{15}$N-labeled protein using a vivaspin (Sartorius, MWCO 5000) concentrator. Initially, these concentrator units were washed twice with deionized water, followed by two washes with PBS (same buffer used for protein elution). The Fibrinogen αC (233-425) containing samples were spun at 5500 rpm for 15-minute intervals, and the concentration of the protein measured regularly. When a concentration of 200 µM was attained, the protein samples were transferred into a dialysis cassette bag (3500 MWCO) for buffer exchange. The protein was dialyzed into 20 mM $\text{K}_2\text{HPO}_4$, 150 mM NaCl, pH 6.5. The concentration of the sample was measured and further concentrated as needed (150-360µM) using the vivaspin concentrator. For all NMR experiments, 10% D$_2$O was added and the mixture transferred to a Shigemi NMR tube.

**Results and Discussion**

Fibrinogen αC (233-425) was successfully expressed in minimal media containing $^{15}$NH$_4$Cl, as described in Chapter 2. Initial 2D [$^1$H,$^{15}$N]-HSQC NMR experiments confirmed that this region shows characteristic spectra for intrinsically disordered proteins (Figure 36). Such a unique spectral profile is characterized by the clustering of peaks within a narrow $^1$H chemical shift range in the amide proton region. In more structured proteins, peaks in the amide protein region are dispersed with $^1$H shifts between 6-9 ppm. As shown in Figure
36, the 2D \([{}^1\text{H},{}^{15}\text{N}]\)-HSQC NMR spectra for Fibrinogen \(\alpha\text{C} (233-425)\) exhibits peaks clustered between 8.0-8.5 ppm. Peaks corresponding to the amide nitrogen from the side chains for glutamines and asparagines are highlighted at 7-8 ppm (blue circle). This cluster contains the 3 reactive glutamines and 13 asparagines. The indole NH side chains from aromatic tryptophan residues are typically seen as peaks in the 10 ppm range (Figure 36, gray circle).

Figure 36: 2D \([{}^1\text{H},{}^{15}\text{N}]\)-HSQC NMR spectra for \(^{15}\text{N}\)-labeled Fibrinogen \(\alpha\text{C} (233-425)\). Areas on spectrum containing side chains of tryptophan highlighted in gray circle (10 ppm) and regions with asparagine and glutamine side chain residues also shown (blue circle 7-8 ppm)
Recent studies have shown improved NMR approaches to characterize protein structure and their binding interactions with ligands. One approach is to utilize residual dipolar couplings (RDC), which has been employed as a complement to distance constraints obtained from the nuclear Overhauser effect (nOe). Such RDC based projects have been successfully used to characterize IDPs and other proteins. The magnitude of the RDC depends on the distance between the coupled nuclei, the orientations of the vector formed by the two nuclei relative to the alignment tensor, and the amplitude of dynamics. Weak alignment is achieved in solution NMR using liquid crystals, gels, phages and bicelles. By contrast, dipolar couplings average to zero in isotropic conditions due to all possible orientations of the vector being accessible during the course of the measurements. An anisotropic condition e.g phages, bicelles, gels is a requirement for RDC measurements in solution NMR. In this weakly orienting environment, NMR dipolar coupling are usually less than 100 Hz in anisotropic conditions.

RDCs are complementary to nOe’s in that they provide information about the orientation of a molecule with both long range and short-range interactions. Unlike nOe’s, which provide distance restraints between atoms, RDCs have information about the angles formed by a vector connecting the two atoms within a tensor axis system. Through space interactions that arise between any two magnetically active nuclei are measured under isotropic and anisotropic conditions using an 2D [¹H,¹⁵N]-in-phase/anti-phase (IPAP) HSQC NMR.
approach which will be used to obtain RDCs (Figure 37).\textsuperscript{76} 2D [$^1$H,$^{15}$N]-IPAP HSQC experiments are modified HSQC experiments designed to record NMR signals that allow observation of the residual dipolar coupling and separation from J coupling.\textsuperscript{127,128} The IPAP experiments are performed in the absence and presence of the alignment media.\textsuperscript{119} The presence of alignment media provides anisotropic conditions that result in a measurable dipolar coupling constant.\textsuperscript{121}

![Figure 37: In-phase and anti-phase (IPAP) HSQC experiment. Scheme illustrating two peaks observed in an HSQC and its corresponding IPAP HSQC in-phase (dark peak) and anti-phase (dotted) peaks. Peak separation with and without alignment media Pf1 phage, illustrated to demonstrate the source of residual dipolar coupling constant $D_{\text{NH}}$ for example (right) (adapted from reference 75)](image)

For Fibrinogen $\alpha$C(233-425), RDC measurements were carried out with the goal of probing for PreSMos. To test the feasibility with IPAP experiments, Pf1 phage was used as an alignment media in the presence of $^{15}$N–labeled Fibrinogen $\alpha$C (233-425).\textsuperscript{129} Spectral splittings obtained in the presence and absence of the Pf1 phage alignment was measured to determine residual dipolar couplings, $D_{\text{NH}}$ (Figure 38). Negative $D_{\text{NH}}$ RDCs are found for segments in which the NH vector is largely oriented perpendicular to the polypeptide chain. For
example in extended conformations and loops $D_{NH}$ values are mainly negative. Meanwhile, positive $D_{NH}$ values are found in alpha helical segments.\textsuperscript{78, 125, 129}

The 2D HQSC spectrum obtained for $^{15}$N $\alpha$C (233-425) in the presence of alignment media Pf1 phage showed a spectral profile similar to that without Pf1 phage (Figure 38). Representative spectra collected In-Phase and Anti-phase in the IPAP experiments are displayed in Figure 39. The sum of peaks from In-Phase and Anti-phase experiments are shown as well as the spectra obtained from the difference of these peaks (Figure 40). To obtain RDCs, resolved peaks from the 2D HSQC experiments were assigned arbitrary numbers 1-164 using NMRpipe. For each peak, the frequency in the sum and differences IPAP experiments was obtained using NMRPipe and recorded. Spectral splittings were recorded from the difference between a corresponding IPAP sum and IPAP difference frequency (i.e J coupling constant (Hz) = IPAP difference – IPAP sum).
Figure 38: 2D HSQC NMR spectra for $^{15}$N-labeled Fibrinogen αC (233-425) in the presence of Pf1 Phage
Figure 39: Representative IPAP HSQC spectra showing peaks In-Phase (top) and Anti-phase (bottom) from the 2D HSQC experiments using Fibrinogen αC (233-425)
Figure 40: Representative Sum and Difference spectra from IPAP experiments with Fibrinogen αC (233-425)
To obtain RDCs, the spectral splitting for each peak was calculated from the sum and difference for both IPAP experiments measured +/- phage. For the peaks in the absence of the phage, typical $J_{NH}$ values were between 93-94 Hz. The spectral splitting for most peaks ranged between 93-98 Hz in the presence of phage. The difference in splittings in the presence and absence of phage are the residual dipolar coupling ($D_{NH}$) (Figure 41).

A majority of the peaks within Fibrinogen $\alpha$C (233-425) have small RDCs between 0.1-2.5 Hz. A few peaks showed negative RDC values. Interestingly, only a couple of peaks exhibited RDCs values >4Hz (Figure 41) which potentially indicate that the residues confer propensity for order and thus are targets for further study. Preliminary dilution experiments with the sample containing the Pf1 phage suggest that the optimum phage concentration is approximately 18 mg/mL substantially lower than the 31 mg/mL used in these experiment. Future studies with this optimized phage concentration may provide further indications of more residues that are in ordered regions. Additional studies in the presence of Factor XIII will also highlight $\alpha$C residues within Fibrinogen $\alpha$C (233-425) that interact with Factor XIII.
With this initial success with RDC experiments, it was then necessary to assign the HSQC peaks for each residue within Fibrinogen αC (233-425). Peak assignments would enable the RDCs to be assigned to specific αC residues. By knowing the identity of the peaks, residues with the potential to possess residual structure as indicated by the RDCs could be checked for involvement with the Factor XIII binding region.
Typical peak assignment experiments with NMR spectroscopy require both $^{15}$N and $^{13}$C labeling of the protein. Our initial experiments to express the double-labeled $^{15}$N and $^{13}$C Fibrinogen αC (233-425) using the same protocol as the $^{15}$N-labeled protein, resulted in lower yields. For NMR studies with the $^{15}$N-labeled Fibrinogen αC (233-425) yields of 2 mg/L were obtained from cultures containing $^{15}$NH$_4$Cl. When the same approach was used to obtain double-labeled proteins ($^{15}$NH$_4$Cl and $^{13}$C-glucose), the yields decreased to 0.1 mg/L. To optimize this yield, a new approach using Bioexpress growth media was tested for both $^{15}$N and $^{15}$N/$^{13}$C-labeled Fibrinogen αC (233-425) described in methods above.

To test for successful expression of the $^{15}$N/$^{13}$C Fibrinogen αC (233-425), 1D $^{13}$C-filtered HSQC experiments were performed. Figure 42 shows peaks from the $^{15}$N/$^{13}$C-labeled protein sample at 150 µM compared to $^{15}$N-labeled protein sample at 360 µM. As expected, the $^{15}$N/$^{13}$C Fibrinogen αC (233-425) spectrum showed higher peak intensities compared to spectra measured with the $^{15}$N-labeled sample, containing a natural abundance of $^{13}$C. The presence of the additional $^{13}$C label was thus confirmed in this sample.
Figure 42: Comparison of 1D $^{13}\text{C}$ HSQC experiment showing improved $^{13}\text{C}$ peaks with 150 µM $^{15}\text{N}/^{13}\text{C}$-labeled sample (cyan) compared to natural abundance of $^{13}\text{C}$ from a 360 µM sample (red) of Fibrinogen αC (233-425)

Similar patterns of clustered peaks were observed with $^{15}\text{N}/^{13}\text{C}$ Fibrinogen αC (233-425) in the 2D $^1\text{H}-^{15}\text{N}$ HSQC experiments. With the incorporation of the $^{13}\text{C}$-label within this protein, protons attached to the $^{13}\text{C}$ could now be monitored to assign each amino acid. In a 2D $^1\text{H}-^{13}\text{C}$ HSQC experiment, each peak represents a proton attached to a $^{13}\text{C}$. Peak clustering was observed in the aliphatic region for experiments with $^{15}\text{N}/^{13}\text{C}$ Fibrinogen αC (233-425). To complete peaks assignments, 3D NMR experiments were performed to help increase peak dispersion to levels needed for peak assignments.
With the $^{15}\text{N}/^{13}\text{C}$-labeled Fibrinogen αC (233-425), it was possible to perform 3D NMR experiments to detect protons ($^1\text{H}$) attached to both $^{15}\text{N}$ and $^{13}\text{C}$. These experiments make use of multiple magnetization transfers though one-bond $J$ coupling. This magnetization is then passed back to $^1\text{H}$ for detection. In the 3D HNCO experiment, the magnetization is transferred from the amide proton $^1\text{H}$ to the $^{15}\text{N}$ and then to the carbonyl $^{13}\text{C}$ of the preceding amino acid (i-1) (Figure 43). This magnetization is sequentially transferred back to the proton ($^1\text{H}$) and detected as a single peak of frequencies corresponding to $\text{HN}_i$, $\text{N}_i$, $\text{CO}_{i-1}$ for each residue. Significant peak clustering was observed with the 3D HNCO experiments. A representative strip from a specific $^{15}\text{N}$ position (123.86 ppm) showed that the peaks observed in this region are well resolved (Figure 44). Peaks in the 123.86 $^{15}\text{N}$ ppm region have higher positive RDC numbers between 2-5 Hz, making this region a good one to monitor for changes in chemical shifts upon Factor XIII binding.
Figure 43: Illustrated magnetization transfer for HNCO experiments (adopted from Protein-nmr.org.uk). Representative 3D HNCO experiment from Fibrinogen αC (233-425) showing position of amide proton peaks $^1$H-N detected in the $^{15}$N and $^{13}$C planes.
Fibrinogen αC (233-425) contains multiple glycines, serines, prolines and threonine residues. Additional 3D experiments such as the HNCACB, HNCOCA, and HNCA will provide information to assign these identical residues due to the connectivities that these experiments observe. These experiments provide addition information about the Cα and Cβ to resolve residues.
Conclusion

Together these results suggest that Fibrinogen αC (233-425) is largely disordered. This segment has now been doubly labeled by expression in quantities that make these NMR experiments practical. Initial RDC measurements confirmed that this approach could be used to identify peaks that have regions with PreSMos. Preliminary 2D and 3D NMR experiments show a tight clustering of peaks, making peak assignments challenging but possible with additional 3D experiments. Future studies might focus on identifying and assigning peaks that experience significant chemical shift changes in the presence of Factor XIII.
CHAPTER 6
RESEARCH SUMMARY AND FUTURE DIRECTIONS

Fibrinogen is the most abundant protein involved in blood coagulation and levels of fibrinogen have been implicated with cardiovascular disease. Fibrinogen is a key player involved in the formation of the fibrin network that makes up blood clots. Blood coagulant Factor XIII functions as a transglutaminase to introduce covalent crosslinks within the fibrin network to create a more stable clot that is more resistant to degradation/fibrinolysis. For many years, the major technique used to evaluate fibrin crosslinking has been SDS PAGE analysis. Although this technique is efficient to quickly monitor global crosslinking trends especially in the gamma chain of fibrinogen, very little information was obtained about crosslinking in the alpha chain. More recently, research findings have suggested that the alpha chain of fibrinogen contributes unique roles to the overall clot structure and stability. The main focus of this dissertation has been to understand fibrin crosslinking at a molecular level.

Two approaches were developed to monitor Factor XIII’s ability to crosslink reactive glutamines in fibrinogen using a lysine mimic. A mass spectrometry-based assay was designed whereby each reactive glutamine can be monitored individually. In this assay, the ability of Factor XIIIa to incorporate a lysine mimic glycine ethyl ester into each reactive glutamine was
monitored over time. Unlike SDS PAGE analysis, information about the role of each individual reactive glutamine could be assessed. Moreover, the assay provides a direct measurement of Factor XIII’s crosslinking ability as opposed to previously designed coupled assays that rely on a secondary reaction for detection. The second complementary instrumental approach used for this research project was NMR spectroscopy. This time $^{15}$N - labeled glycine ethyl ester was used to probe Factor XIII’s ability to target reactive glutamines. Upon crosslinking GEE to the reactive glutamine, a single peak corresponding to the proton ($^1$H) attached to the $^{15}$N is observed in a 2D Heteronuclear Single Quantum Coherence (HSQC) NMR experiment. This approach serves as a quick test for crosslinking. However, due to the low sensitivity of the NMR instrument, higher concentrations of protein are needed for the NMR assay compared to the mass spectrometry approach.

In the sections that follow, findings from chapters 2, 3, 4 and 5, are summarized along with additional plans for this research area. Future studies for structural characterization of Fibrinogen $\alpha$C (233-425) are also presented. Furthermore, other Fibrinogen $\alpha$C and $\gamma$ segments that could provide additional information on Factor XIII’s substrate specificity are discussed.

**Ranking Reactive glutamines crosslinked in Fibrinogen**

Previous researchers identified multiple reactive glutamines within the Fibrinogen A$\alpha$ chain that are selected by Factor XIII for crosslinking reactions. These reactive glutamines are located mainly in the flexible $\alpha$C region of
Fibrinogen and include Q221, Q223, Q237, Q328, and Q366. By contrast, only two reactive glutamines, Q398 and Q399, have been identified in the γ chain thus far. These reactive glutamines are located in the flexible C-terminus of the gamma chain. Information about the role of each reactive glutamine in Fibrinogen has not been available. A combination of NMR spectroscopy and mass spectrometry approaches described in Chapter 2 and 3 was used to characterize and rank three reactive glutamines in Fibrinogen αC (233-425). Our observations correlate with previous Factor XIII crosslinking trends reported for full length Fibrinogen.

Fibrinogen αC (233-425) contains three reactive glutamines (Q237, Q328, and Q366) previously identified as substrates for Factor XIII’s crosslinking activity. This αC segment was expressed in E.coli and purified free of the GST- tag. Variants of Fibrinogen αC (233-425), in which reactive glutamines were substituted for inactive asparagines, were also expressed and purified. A MALDI-TOF mass spectrometry assay was then employed that allowed us to monitor crosslinking over time. An NMR spectroscopy approach was also used to assess crosslinking patterns in WT and variants of Fibrinogen αC (233-425).

From our results, the three reactive glutamines were ranked as Q237 >> Q366 ≈ Q328. Further investigation demonstrated that at higher concentrations of Factor XIII and longer time points, Q366 was crosslinked faster than Q328. These rankings correlate with the crosslinking pairs previously identified in full-length fibrinogen under the conditions studied. Wang et al., demonstrated that Q237 can be crosslinked to up to four lysines in Fibrinogen. This amount was
the most lysine pairs associated with any other reactive glutamine identified in this study. Such an observation suggests that Q237 is playing a major role in Aα chain crosslinking and more likely driving the unique properties associated with crosslinking in this region. When compared to the other reactive glutamines in this segment, Q237 is located at the N-terminus of this αC (233-425). A combination of residue accessibility and the presence of surrounding amino acids residues may contribute to this high selectivity for Q237 by Factor XIII. Only one glutamine-lysine crosslinking pair was identified for Q366 and none for Q328 in full-length fibrinogen. Compared to Q237 and Q366, Q328 is more centrally located and it is the least reactive glutamine. Q366 is located more C-terminally but is closer to a previously proposed binding region for Factor XIII found within αC (389-402). This unique location could contribute to Q366 being a better substrate than Q328.

Factor XIII’s ability to crosslink any remaining reactive glutamines following the mutation of one or more of these reactive glutamines into an inactive asparagine is introduced in Chapter 2, and discussed further in Chapters 3 and 4. In Fibrinogen Seoul, a single point mutation in one reactive glutamine Q328 (AαQ328P) resulted in dysfibrinogenemia. Also, previous studies in which segments of the αC containing reactive glutamines were absent showed overall changes in clot properties and fibrin network. Together, these previous findings suggest that the absence of one or more reactive glutamines results in altered crosslinking properties. For the first time, it was shown that in the absence of one or more reactive glutamine, activated Factor XIII can still crosslink any remaining
glutamine independently. The absence of Q237 for example did not prevent Factor XIII’s ability to crosslink Q366 and Q328. The same is true for Q366 located closer to the Factor XIII binding site αC (389-402). Interestingly, a Q328N mutation also did not affect crosslinking of the other remaining reactive glutamines. A future study with the Q328P mutation, similar to Fibrinogen Seoul II, could provide additional information about the role of the proline in this position.

These conclusions were made because of the ability to monitor each reactive glutamine using MALDI-TOF mass spectrometry and NMR assays. These assays were also applied to a segment of the gamma chain of Fibrinogen γ (148 – 411) that contains two known remaining reactive glutamines Q398 and Q399. The MALDI-TOF mass spectrometry and NMR assays demonstrated that Factor XIIIa crosslinked both reactive glutamines within this γ segment. Similar to reactive glutamines located on the alpha chain of Fibrinogen, the extent to which each reactive glutamine was crosslinked differed (Chapter 2). One reactive glutamine appeared to be crosslinked to a slightly faster extent than the other. However, the most reactive glutamine in this γ segment could not be identified. Future studies with mutations in this region could provide more information about the most reactive glutamine.

It should be noted that higher concentrations of Factor XII and longer incubation times were needed to observe the crosslinking reaction involving Q398 and Q399 in Fibrinogen γ(148–411) compared to the three reactive glutamines Q237, Q328, and Q366 in Fibrinogen αC (233-425). This result is
rather surprising since in full length Fibrinogen, crosslinking among gamma chains is used as a quick indicator for changes in Fibrinogen crosslinking via SDS PAGE analysis. In gel assays, γ-γ crosslinking is observed first followed by α-α and α-γ hybrids in full-length Fibrinogen. Fibrinogen γ(148–411) segment may have a different structural conformation that does not allow for an effective Factor XIIIa crosslinking as in the full length γ chain.

Previous researchers have shown that the primary binding site for Factor XIII A2B2 is located on Fibrinogen residues γ390-396.\textsuperscript{46} This region is positioned only a few residues away from the reactive glutamines, Q398 and Q399, identified on the gamma chain. Mice studies in which residues in Fibrinogen γ390-396 were mutated to alanine (Fibγ\textsuperscript{390-396A}) showed delayed fibrin crosslinking and Factor XIIIa activation.\textsuperscript{133} In the presence of alanines, the binding of Factor XIII to this region could be significantly altered. This change can therefore affect Factor XIII’s ability to crosslink two reactive glutamines on Fibrinogen γ chain. There is increasing evidence that Fibrinogen provides multiple binding sites for Factor XIII on both the alpha and gamma chains.

**Role of E396 and Factor XIII ability to crosslink reactive glutamines**

A second binding site for Factor XIII on Fibrinogen was identified on the Fibrinogen αC region within αC (389-402).\textsuperscript{72} The amino acid E396 was found to be a key contributor to Factor XIII binding interactions within this region. This conclusion was based on the observation that Factor XIII binding was significantly decreased with the αC mutant E396A. To test for Factor XIIIa’s
ability to crosslink reactive glutamines in the absence of this binding residue, we introduced the E396A mutation was introduced into Fbg αC (233-425). These results are discussed in Chapter 4.

In the presence of the E396A mutation, all three reactive glutamines were still crosslinked by Factor XIIIa. Interestingly, these reactive glutamines were crosslinked to a similar extent as in the wild type Fibrinogen αC (233-425). Furthermore, a comparison of pre-activated Factor XIII A_2 and Factor XIII A_2 activated in the presence of Fbg αC E396A showed that all reactive glutamines were crosslinked to the same extent as in WT. The presence of the E396A mutation thus did not affect Factor XIII A_2 activation nor its crosslinking ability.

To further explore the role of E396A on Factor XIII activation, the plasma form of Factor XIII (Factor XIII A_2B_2) that contains the carrier B subunit and the catalytic A subunits was used. In the presence of the B subunits, the extent of crosslinking was similar for all three reactive glutamines in both WT and E396A. Curiously, a slight delay in crosslinking was observed with Factor XIII A_2B_2 compared to Factor XIII A_2. This effect could be a result of a delay in the activation of plasma Factor XIII A_2B_2 compared to the Factor XIII A_2. During the activation of Factor XIII A_2B_2, thrombin cleaves the activation peptide (AP) from the A subunit, and in the presence of calcium, the B subunits dissociate from the A subunits. This additional dissociation step required for the exposure of the catalytic A subunit in Factor XIII A_2B_2 could account for the delay in activation and subsequent crosslinking.
Surprisingly, a recent study by Souri et al., suggests that the B subunits of Factor XIII accelerates fibrin crosslinking. Normal plasma was depleted of Factor XIII and recombinant Factor XIII A$_2$ was added to this plasma. Crosslinking trends in the gamma chain were monitored by SDS PAGE, with and without Factor XIII B$_2$ subunit. Densitometric analysis for γ-γ crosslinking suggests that upon addition of Factor XIII B subunits, crosslinking occurs faster compared to Factor XIII A only. It should be noted that the gels show little evidence of crosslinking involving the alpha chain. Moreover, these results do not correlate with our observations of alpha chain crosslinking. In our work, the presence of the B-subunits, in plasma Factor XIII A$_2$B$_2$, did not accelerate crosslinking for each reactive glutamine in Fibrinogen αC (233-425). Instead, crosslinking in Factor XIII A$_2$B$_2$ was slightly delayed compared to Factor XIII A$_2$ under the same conditions. The crosslinking trends for all three reactive glutamines were similar for both recombinant Factor XIII A$_2$ and plasma Factor XIII A$_2$B$_2$. These differences in crosslinking extents highlight the unique contributions of Fibrinogen crosslinking in the alpha and gamma chains. It is possible that the γ390-396 segment is a more critical player in controlling Factor XIII binding and activation than αC (389-402). Alternatively, the two Fibrinogen chains contribute different regulatory roles. Future studies to distinguish the contributions of the binding sites on the gamma and alpha chain will help in elucidating their individual roles in Factor XIII’s activation and function.

In chapter 4, the results from unique crosslinking trends observed when one or more reactive glutamines in Fibrinogen αC (233-425) were replaced with
inactive asparagines are discussed. One emphasis of this study was to monitor Q328 and Q366 in the absence of the fastest reactive glutamine Q237. Studies later proceeded to αC (233-425) Q→N mutants with “Q328 only” and “Q366 only”. The residues Q328 and Q366 are located closest to the Factor XIII binding site for identified within αC (389-402). For these experiments, an LC–MS method was used to monitor these reactive glutamines because this method provided higher throughput and sensitivity, than the MALDI-TOF MS.

Our results indicated that Q328 and Q366 were crosslinked the slowest in αC Q237N variants. The absence of the most reactive glutamine (Q237) seems to affect the crosslinking ability for the remaining two reactive glutamines. This effect is still unclear since in “Q366 only”—the αC (233-425) mutant where Q237 and Q328 are replaced with inactive asparagines (Q237N/Q328N)—the remaining reactive glutamine Q366 is crosslinked more readily than in WT. Overall, in the absence of Q237, Q366 alone is a better substrate than Q328 alone. For short peptide substrates, Q366 may be a better candidate to examine Factor XIII’s substrate specificity. The proximity of Q366 to the binding site and its peculiar crosslinking trend makes it an attractive substrate for future studies. Additional kinetic studies monitoring each reactive glutamine would provide more information about their unique contributions in the presence of different mutation.
Preliminary Results for the Study of αC (221-391)

Fibrinogen αC (221-391) contains two other reactive glutamines, Q221 and Q223, previously shown to be crosslinked by Factor XIIIa. In addition, this protein sequence lacks the critical residue E396 and surrounding residues of the Factor XIII binding region. Compared to Fibrinogen αC (233-425), this segment can be used to better assess both the role of the binding region αC (389-402) for Factor XIII and also the crosslinking ability for two additional reactive glutamines.

An attempt to express this protein using the previously optimized approach by the Medved lab resulted in very low yields with impurities (Appendix 5). Further exploration for this αC (221-391) segment would provide useful clues to the ability of Factor XIII to crosslink the reactive glutamines Q221 and Q223. With an optimized mass spectrometry and NMR spectroscopy method to monitor crosslinking, these findings would provide additional information about Factor XIII’s substrate specificity. In the absence of the binding site αC (389-402) and with additional reactive glutamines to select, it would be interesting to detect any changes in the crosslinking trends for Q237, Q328, and Q366. In Fibrinogen αC (233-425), Q237 is the most N-terminally located and the first reactive glutamine to be crosslinked. In αC (221-391), Q221 and Q223 are additional reactive glutamines available for crosslinking at the N-terminus. Results from the full length Fibrinogen glutamine –lysine crosslinking pairs would suggest that Q237 would still be the fastest reactive glutamine in the presence of Q221 and Q223. Questions surrounding Factor XIII’s use of this binding site for crosslinking would be answered using this αC (221-391) fragment. Changes in the reactivity of
Q327 in the presence of Q221 and Q223 will be interesting to explore. Information from this series of studies would elucidate the role of the additional reactive glutamines in the N-terminus as well as Factor XIII binding site. Overall, these findings would provide more information about Factor XIII’s selectivity and substrate specificity.

**Preliminary NMR Structural Characterization of αC (233-425)**

Intrinsically disordered proteins and regions have been increasingly associated with disease. In the cardiovascular disease research field, fibrinogen has been identified as a crucial protein in the formation of stable clots. Interestingly, a majority of the binding and crosslinking activity within fibrinogen have been identified to occur within the flexible αC region of the Aα chain and the flexible C-terminal portion of the γ-chain. Binding sites for Factor XIII have been identified in flexible regions within γ390-396 and in αC (389-402). All reactive glutamines crosslinked by Factor XIII are also located in these flexible regions especially in the more disordered αC region. Only two reactive glutamines located in the flexible γ-chain have been shown to be crosslinked. Although the Bβ chain of fibrinogen contains multiple glutamines, none of these glutamines are crosslinked by Factor XIII. A combination of accessibility and substrate specificity is therefore important for Factor XIII’s activity.

One of the challenges in characterizing intrinsically disordered proteins and regions is the absence of suitable techniques. NMR spectroscopy has been adopted more recently to study these intrinsically disordered proteins (IDP). A
unique spectral profile for IDPs is observed in 2D HSQC NMR experiments.\textsuperscript{78} This HSQC spectrum for IDPs is characterized by the narrow clustering of peaks within the backbone amide region of the spectra (8-9 ppm) compared with a spread of peaks for more structured proteins (6-10 ppm). This IDP type of clustering makes it challenging to identify each amino acid residue associated with a particular peak compared to more structurally ordered proteins where peak clustering is less of a concern for peak identification. To aid in the assignment process, the NMR approach requires that the protein be isotopically labeled, which can be quite costly.

We first successfully expressed $^{15}$N-labeled $\alpha$C (233-425) and confirmed using the 2D HSQC NMR approach that this segment is intrinsically disordered (Chapter 5). A long range goal was then to characterize residues within the Factor XIII binding region $\alpha$C (389-402) and monitor for global changes occurring in the presence of Factor XIII. To achieve this goal, chemical shift assignments would be needed for this region within $\alpha$C (233-425). To aid in the assignment process, additional uniform $^{13}$C–labeling of $\alpha$C (233-425) was needed. The conditions optimized for expressing the double-labeled $^{15}$N/$^{13}$C $\alpha$C (233-425) are described in Chapter 5. The double labeled protein sample obtained was concentrated to 150 µM and initial NMR test experiments were performed. 1D and 2D HSQC experiments were used to confirm the presence of both $^{15}$N and $^{13}$C labels on $\alpha$C (233-425).

Initial results from this study are discussed in Chapter 5. Results from a 3D HNCO experiment showed few unresolved peaks. Future NMR experiments
including HNCACO and HNCACB will provide additional information for peak assignments. Once the assignments are complete, NMR titration experiments can then be used to monitor changes occurring upon introducing Factor XIII. In these experiments, the chemical shift and peak characteristics of the αC residues will be examined. Results obtained with Factor XIII A₂ and Factor XIII A₂B₂ can be compared. Additional clues about Factor XIII’s binding interaction within Fibrinogen αC (233-425) can be probed further from these studies.

**Future Research to Explore Factor XIII’s substrate Specificity**

In the previous chapters, Factor’s XIII ability to crosslink reactive glutamines in the αC segment and gamma segment has been explored. The role of a critical residue E396 as well as structural information about αC (233-425) is now available. These findings were made possible using the design of a mass spectrometry and NMR spectroscopy assay approaches that allow each reactive glutamine to be monitored independently. These strategies have facilitated our understanding of the contribution of each reactive glutamine. A similar approach can later be applied to truncated variants of the αC region as well as the full-length chain to gain a better understanding of crosslinking trends in this region.

Structural characterization for Fibrinogen αC (233-425) especially in the presence of Factor XIII is another area that could provide additional clues for Factor XIII’s substrate specificity. Previous research has sought to identify the role of the binding region of Factor XIII. The NMR approach will provide clues to specific residues on fibrinogen that interact with Factor XIII in both its inactivated
and activated states. The PONDR structure predictor suggests that a few residues within the αC (233-425) region have a propensity for order. Factor XIII binding could promote a more ordered structure within αC (233-425).

Overall, important insights have been gained about Factor XIII’s ability to crosslink each reactive glutamine under different conditions. Fibrinogen αC (233-425) served as a good model to examine trends that have been previously shown in full-length fibrinogen. Additional information from truncated segments of Fibrinogen can be explored in the future to develop drugs that would be specific for Factor XIII. Knowledge gained from this study can be used to design drugs that are specific for Factor XIII. Since fibrinogen is the primary substrate for Factor XIII, additional information about Factor XIII’s ability to select these reactive glutamines can also shed light on how to design fibrinogen-based substrates that produce a desired anticoagulant or procoagulant effect.
REFERENCES


Mouse Models and Inherited Disorders, *Physiological Reviews* 89, 991-1023.


APPENDIX 1

Table 1: Primers designed and used to introduce point mutations for each reactive glutamine in Fibrinogen αC (233-425)

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Forward Primer 5´-3´</th>
<th>Reverse Primer 5´-3´</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q237N</td>
<td>ggatccacagacatgccgaaatgagttgcggatgtggatcc</td>
<td>atgcctccactcatgttcggcatgtctggtc</td>
</tr>
<tr>
<td>Q328N</td>
<td>ggaactggaactactggaacaaccctggaagtccagggttgtttccagtactccagtactcc</td>
<td></td>
</tr>
<tr>
<td>Q366N</td>
<td>tctggtagtactggaactggactgtgaatctggccagattcagagtgcca</td>
<td>ccagattcagtgccagttcaggactactcc</td>
</tr>
</tbody>
</table>

Table 2: Sequence and mass of proteolytically derived peptides containing αC Q237, Q328, and Q366

<table>
<thead>
<tr>
<th>Reactive Glutamine</th>
<th>Amino Acid Sequence</th>
<th>Theoretical m/z</th>
<th>Experimental m/z</th>
<th>Digest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q237*</td>
<td>GPLGSTMPQMRME</td>
<td>1549.7</td>
<td>1550</td>
<td>GluC</td>
</tr>
<tr>
<td>Q328</td>
<td>NSGSSGTGSTGNQNPMSGPRPGSTGTWNPGSSERSAGWH</td>
<td>2448.1</td>
<td>2448</td>
<td>chymotrypsin</td>
</tr>
<tr>
<td>Q366</td>
<td>SSVSGSTGQWHSE</td>
<td>1348.6</td>
<td>1349</td>
<td>GluC</td>
</tr>
</tbody>
</table>

*MS/MS analysis on the peptide containing Q237 revealed that amino acids GPLGS are located just N-terminal to the αC 233-425. These amino acids remain following PreScission protease cleavage of the GST-αC (233-425). The sequences containing Q328 and Q366 were readily identified using a theoretical protease digest (Protein Prospector UCSF).
APPENDIX 2

Analysis of αC (233-425) following Protein Purification

SDS-PAGE

A 10% resolving gel was prepared by combining 4 mL deionized water (dH₂O), 3.4 mL 30% Bis-acrylamide, 2.5 mL 1.5 M Tris (pH 8.8), and 0.1 mL 10% SDS. Gel polymerization was initiated by adding 50 µL Ammonium persulfate (10% APS) and 5.0 µL tetramethylethylenediamine (TEMED). This was allowed to polymerize for 45 min. The stacking was prepared at 4% by combining 6.1 mL dH₂O, 1.3 mL 30% Bis-acrylamide, 2.5 mL 0.5 M Tris (pH 6.8), and 0.1 mL 10% SDS. Gel polymerization was initiated in the same way described above. This was allowed for 30 min and wells formed were rinsed with dH₂O and running buffer. Fractions containing αC and GST eluents were combined with non-reducing buffer, loaded on the gels, and ran at 200 V for 1 h. For c0omassie blue staining, gels were placed in a mixture containing 0.1% coomassie blue, 40% Methanol, 10% Acetic acid, and 50% dl H₂O overnight and destained with a 40% Methanol, 10% Acetic acid and 50% dl H₂O destaining solution.

Western Blot

Western blot analysis on gels was used to confirm the presence of αC only and no GST in the protein fractions collected after cleavage by the PreScission™ Protease. Since αC (233-425) and GST have similar molecular weights, SDS PAGE alone could not be assertive in distinguishing these bands.
It was therefore important to confirm that indeed the fractions collected contained αC (233-425) only. A similar SDS PAGE gel with fractions from αC and GST eluents was run as previously described. This gel then subjected to western blotting using an anti-GST-HRP Conjugate (Amersham™, GE Healthcare). All Western Blot set-up components were equilibrated in Western Blot Transfer buffer (25 mM Tris, 192 mM Glycine, and 20% MeOH) for 10 min at room temperature. The gel was equilibrated as well. This was run at 100V for 1hr at 4°C and the nitrocellulose membrane obtained. Non-specific binding was reduced by shaking the nitrocellulose membrane in a blocking solution in with 3% BSA in PBST buffer (80 mM Na₂HPO₄, 20 mM NaH₂PO₄, 100 mM NaCl, 0.001% Tween-20) overnight at room temperature. The gel was washed with fresh PBST buffer and transferred to a solution containing 6μL anti-GST HRP Conjugate in 25 μL PBST. This was allowed to shake for 1 h, rinsed in PBST and developed in TMB blotting buffer. The staining process was quenched with dIH₂O.

Zip-tipping for Sample Preparation

The process of zip-tipping was done by applying all samples to a C18 zip-tip for desalting. Briefly, each zip-tip was rinsed using 4 μL (3x) in 100% acetonitrile, followed by equilibration in 0.1% TFA (4 μL, 3x). The sample was applied on the zip-tip by pipeting up and down at least 8 times. The unbound material where rinsed using 0.1% TFA (5 μL, 3x). For sample digested with chymotrypsin, the zip-tipped samples were eluted in 5 μL α-cyano-4-hydroxycinnamic acid (αCHCA) matrix (13 mg/ml dissolved in 0.1% TFA, acetonitrile, and ethanol) by pipeting up and down. GluC digested samples were
eluted in 5 µL ferulic acid matrix. 1 µL of each sample was spotted on the MALDI plate. The MALDI-TOF was operated under positive reflectron mode at an optimized manual laser intensity of 1997. Peptide fragments observed were compared to theoretical digests from chymotrypsin and GluC and the peptide peaks containing reactive glutamines were identified. Under crosslinking conditions, peaks containing reactive glutamines and crosslinked to GEE were observed at an additional 86 \( m/z \). Peaks heights for reactant and product peaks were extracted using the Data Explorer analyzer.

**Mutations within αC (233-425) confirmed using MALDI TOF Mass Spectrometry**

Mutations introduced in αC (233-425) were confirmed using DNA sequencing. Following protein purification, the presence of the desired mutation was verified using the MALDI-TOF mass spectrometry. Chymotrypsin digest was used to confirm the presence of Q328N mutation. Meanwhile, peaks identified from the GluC digests of αC (233-425) allowed us to confirm the presence of Q to N mutations in Q237 and Q366 and Q328 as well as the E396A mutation.

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Digested Peptide Fragment</th>
<th>( m/z )</th>
<th>( m/z ) (wt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q237N</td>
<td>GPLGSTDMP_NMRM</td>
<td>1535</td>
<td>1550</td>
</tr>
<tr>
<td>Q328N</td>
<td>NSGSSGTG_STGNN_NPGPRPGSTGTW</td>
<td>2347</td>
<td>2448</td>
</tr>
<tr>
<td>Q366N</td>
<td>SSVSGSTG_NWHSE</td>
<td>1334</td>
<td>1349</td>
</tr>
<tr>
<td>E396A</td>
<td>SGSFRPDSPGSGNARPNNPDWGTFAE</td>
<td>2647</td>
<td>2776</td>
</tr>
</tbody>
</table>
Validating αC (233-425) crosslinking with $^{15}$N-labeled substrates using 2D $^1$H - $^{15}$N HSQC NMR spectroscopy

The exchange of $^{15}$N-labeled NH$_4$Cl and crosslinking of $^{15}$N-labeled GEE with αC (233-425) was monitored using a $^{15}$N- heteronuclear single quantum coherence (HSQC) NMR spectroscopy. All assays were carried out in had a total volume of 400 µL with a 10% v/v D$_2$O added for a deuterium lock on the NMR instrument. For the $^{15}$N-labeled NH$_4$Cl exchange reactions, 400 nM FXIII was non-proteolytically activated in the presence of excess calcium (50 mM) and 20mM borate buffer (pH 8) for 10 min at 37°C. Then, 100 mM $^{15}$NH$_4$Cl and 40 µM αC (233-425) were added and the mixture incubated for 1hr at 37°C. For crosslinking reactions with $^{15}$N-labeled GEE, 800 nM FXIII was proteolytically activated with 21 U/ml bovine thrombin in presence of 5 mM CaCl$_2$ and 20 mM Borate buffer. Thrombin was quenched with 460 nM PPACK. 10 mM $^{15}$N-labeled GEE and 35 µM αC (233-425) was added and incubated for 1 h at 37°C. D$_2$O was added and the mixture transferred to a Shigemi NMR tube. These were analysed on a Varian 700Hz NMR Instrument. The parameters for the 1D HSQC were nt =512, ni = 1, np =2048, and the 2D HSQC nt =64, ni = 64, np = 2048.
2D $^{15}$N HSQC NMR spectrum of $\alpha$C (233-425) incorporating $^{15}$N-labeled NH$_4$Cl in the presence of Ca$^{2+}$ activated FXIII shows three peaks with corresponding crosspeaks. A similar 2D $^{15}$N HSQC NMR spectrum obtained from the crosslinking reaction of reactive glutamines on $\alpha$C (233-425) and $^{15}$N-labeled GEE catalyzed by thrombin-activated FXIII showed three distinct peaks. All reaction schemes for the crosslinking of $\alpha$C (233-425) with labeled substrates are also shown.
APPENDIX 4

Monitoring Factor XIII A₂ and Factor XIII A₂B₂ activation

Thrombin activation of Factor XIII A₂ and Factor XIII A₂B₂ in the presence of used in Chapter 4 were monitored for over time (top and bottom gels respectively). Each gel contains Molecular weight marker and aliquots obtained from time points 0, 1, 2, 5, 10, 15, 30 and 60 minutes respectively. The last lane represents inactivated Factor XIII A₂ and Factor XIII A₂B₂ for each gel. Factor XIII: Thrombin ratios used for these gels were 60.9: 1 for Factor XIII A₂ and 15.1:
1 Factor XIII $A_2B_2$. In both cases, over 50% Factor XIII was activated within 5 minutes of the reaction.
APPENDIX 5

Protein Expression of αC (221-391)

pET20b vector containing cDNA with αC 221-391 was a generous gift obtained from Dr. Leonid Medved (University of Maryland). For specific purification and to isolate only our protein of interest, this DNA was inserted into a pET14b vector with a His tag and sumo protease encoding region. The following primers were designed containing Ndel and HindIII restriction sites. Forward 5’-GGGAATTCCATATGCAGCTGCAGAAA- 3’ and Reverse primer 5’-CGCGGATCCTTACCAATCCGGTTGTCGG -3’. These primers were used to amplify the cDNA encoding for αC 221-391. For the PCR reactions, 200 µM dNTP, 500 nM forward and reverse primers, 112 ng DNA, and 5X Phusion HF buffer (10 µL) were combined. Phusion DNA polymerase (0.5 µL) was added to the mixture and the final volume adjusted to 50 µL with dI water. The following parameters were optimized to amplification; initial denaturation (98°C for 30 sec); Repeated 30 cycles (98°C for 10 sec, 70°C for 25 sec, and 72°C for 25 sec); and final extension (72°C for 10 min) and final hold for 4°C. The PCR product was separated on an agarose gel (1%) and the desired band (~600bp) was cut and isolated following ethidium bromide staining.

To obtain high yield and purity for this protein, the DNA was inserted into a pET14b vector containing a HisTag and a SUMO protease encoding region. The HisTag provides a tag sequence for more specific affinity column protein purification. The SUMO protease is strategically located between the HisTag and
the cDNA for the αC (221-391) protein. SUMO protease cleavage will result in the release of only the αC (221-391) segment and first residue in this αC segment will be the reactive glutamine Q221. For better comparisons with the natural αC sequence, it is important to characterize Q221 free from the HisTag and the SUMO protease segment.

To amplify and extract the cDNA region for αC (221-391) from the original pET20b vector, primers were designed containing BamHI and NdeIII were used. This cDNA was ligated into the pET14b vector and transformed into BL21 Gold DE3 cells. An initial expression test was performed using for optimal conditions needed for IPTG induction. A small scale protein purification was performed. Our preliminary results from purifying αC (221-391) with the both the HisTag and SUMO protease still present, showed high levels of impurities. Gel analysis showed multiple bands from other affinity binding protein. These impurities have a high affinity for the HisTrap column as well. The next strategy will be to add the SUMO protease cleaving enzyme on-column as in our GST-tagged system. Future investigations will optimize the conditions for high yields with low impurities in this segment.
CURRICULUM VITAE

Kelly Njine Mouapi

Department of Chemistry | University of Louisville
2320 South Brook St | Louisville, KY, 40292
k0moua01@louisville.edu

EDUCATION

Ph.D. Chemistry
Department of Chemistry, University of Louisville, Louisville, KY
Advisor: Dr. Muriel Maurer
August 2017

M.S. Chemistry
Department of Chemistry, University of Louisville, Louisville, KY
May 2015

B.A. Chemistry, Spanish minor (cum laude)
Berea College, Berea, KY
May 2012

PUBLICATIONS


RESEARCH EXPERIENCE

Graduate Research Assistant (Department of Chemistry, University of Louisville, Louisville, KY)
Dissertation: Characterizing reactive glutamines in Fibrinogen and elucidating Factor XIII’s substrate specificity. Advisor: Dr. Muriel Maurer (Professor)

Research Intern (Department of Immunology, Mayo Clinic, Rochester, MN)
Epitope mapping of Vimentin in HLA-DR4 arthritis-susceptible *0401 and arthritis-resistant *0402 Transgenic Mice. Advisors: Dr. Chella David (Professor) and Dr. Veena Taneja (Associate Professor)


**RESEARCH EXPERIENCE cont’d**

**Research Intern** (Vanderbilt-Ingram Cancer Center, Vanderbilt University, Nashville, TN)  
Monoclonal Antibody 1A5 abrogates TGF-beta induced Tumor Cell Migration. Advisor: Dr. Andries Zijlstra (Associate Professor)

**Research Intern** (Department of Biology, Berea College, Berea, KY)  
In vitro retraction of the distome body of Proterometra macrostoma. Advisor: Dr. Ronald Rosen (Professor)

- **Technical skills**: Protein expression, Protein purification, Mass spectrometry, MALDI-TOF, LC-MS, Assay development, PCR, NMR spectroscopy, 2D HSQC NMR, HPLC, Bacterial cell culture, Mammalian cell culture, Tissue culture, Mouse models, Western blots, FPLC, ELISA, Grant writing, Biochemistry Research, Teaching & Training laboratory personnel
- **Interests**: Languages, Hiking, Biking, Dance, Reading
- **Languages**: English native, fluent in French, conversational in Spanish

**PRESENTATIONS**

**Invited Speaker**

2016 – Research Talk: “The Fantastic Four in Blood Coagulation” (Berea College, KY)  
2014 – Moderator and Co-chair for “Fibrinogen and Fibrin I” oral session at the XXV Congress of the International Society on Thrombosis and Haemostasis (Toronto, Canada)

**Oral Presentations**

2016 – National Organization for the Professional Advancement of Black Chemists and Chemical Engineers (NOBCChE, Raleigh, NC) – Assisting Science Conference Grant
2015 – International Society of Thrombosis and Haemostasis (ISTH Toronto, Canada) – Session co-chair  
2014 – International Fibrinogen Research Workshop (IFRS Marseille, France) – Young Investigator Award
2011 – Kentucky Academy of Science (Murray State University, KY) – 1st place Physiology and Biochemistry
2011 – Undergraduate Symposium Department of Immunology Mayo Clinic (Rochester, MN)

**Poster Presentations**

2016 – American Chemical Society (San Francisco, CA) – ACS Division of Biological Chemistry Travel Award
2016 – Research!Louisville (Louisville, KY)
2016 – Biophysical Society meeting (Los Angeles, CA) - Education Committee Travel Award  
2015 – Biophysical Society meeting (Baltimore, MD)  
2010 – Kentucky Academy of Science (Western Kentucky University, KY)
2010 – Undergraduate Symposium Vanderbilt-Ingram Cancer Center, Vanderbilt University, (Nashville, TN)
2009 – Kentucky Academy of Science (Northern Kentucky University, KY) – 3rd place Zoology Posters Award

157
FELLOWSHIPS AND GRANTS
2016 –Arno Spatola Endowment Graduate Fellowship; a prestigious and competitive fellowship for Ph.D. candidates by the University of Louisville’s Institute for Molecular Diversity and Drug Design (IMD3), $12,500
2015 –Certification of completion of Grant Writing Academy Training (University of Louisville)
2012 –Olive Ruth Russell Fellowship for outstanding woman in chemistry pursing graduate studies, $1000

TEACHING AND MENTORING EXPERIENCE
Graduate Teaching Assistant (Department of Chemistry, University of Louisville, Louisville, KY) 2012-2014
• Biochemistry Laboratory (CHEM 546)
• Organic Chemistry laboratory (CHEM 343)
• General Chemistry (CHEM 207/CHEM 208)
Mentoring: Trained and mentored 1 graduate and 3 undergraduate students on techniques for research projects during laboratory rotation and summer internship.
• Lucille Wagner (ENSCM Montpellier France Exchange Program – Masters Student)
• Chad Stephens (Undergraduate Honors Thesis –University of Louisville)
• Kelsey Lamb (NSF Research Experience fo Undergraduates -REU summer research)
• Deidre Dillon (NSF Research Experience fo Undergraduates -REU summer research)

Laboratory Director (Department of Chemistry, Berea College, Berea, KY) 2011 - 2012
• Supervised 14 undergraduate teaching assistants and tutors for the department of chemistry work-study program
• Created and implemented operation “C.P.R (Communication, Planning/Preparedness, Rescue)” for effective communication between teaching assistant and faculty, and for organized laboratory classes and tutoring
• Tutored chemistry students, graded laboratory reports/notebooks, and ordered chemicals for laboratory classes

Undergraduate Teaching Assistant (Department of Chemistry, Berea College, Berea, KY) 2009 - 2011
• General Chemistry (2 semesters)
• Organic Chemistry (2 semesters)
• Analytical Chemistry (2 semesters)

AWARDS & HONORS
• 1st Place Research Elevator Pitch “2-minute Drill” NOBCChE (2016)
• Advancing Science Conference Grant, NOBCChE (2016)
• Biophysical Society Education Committee Travel Award, Los Angeles, CA (2016)
• Outstanding Alumnus Award for Berea College African Student Association (2015)
• The M. Celeste Nichols Award from the University of Louisville’s Women Center (2015)
• Young Investigator Award –XXIIIrd International Fibrinogen Workshop, Marseille, France (2014)
• Institute for Molecular Diversity and Drug Design (IMD3) Travel Award (2014)
• Graduate Student Council Travel Awards (University of Louisville, 2014 - 2016)
AWARDS & HONORS cont’d

• Henry W. and Edna Austin Scholar for outstanding academic and personal excellence (2012)
• Class of 1958 Chemistry Academic Achievement Senior Award for highest GPA in Chemistry (2012)
• 1st place in Physiology and Biochemistry student oral presentation at the Kentucky Academy of Science (2011)
• Thomas J. Watson Fellowship Finalist (2011)
• Best Peer Mentor of the year for outstanding mentorship for the Black Cultural Center S.U.C.C.E.S.S (2011)
• Class of 1958 Chemistry Achievement Award for outstanding contributions to undergraduate research (2010)
• Thomas Beebe Organic Chemistry Award for outstanding performance in organic chemistry (2009)
• Academic excellence and high personal standard of living in honor societies Fleur de Lis, Vincit Qui Patitur, and Mortar Board (2009-2011)

LEADERSHIP AND COMMUNITY SERVICE

• Team captain for the American Heart Association 2015 & 2016 HeartWalk in Louisville: Created, organized, and led team “Chemistry Hearts” to participate in the National Heart Walk. Raised a total of over $1000 for research through the American Heart Association (2015-2016)
• Professional Development Committee co-chair for the Minority Association of Graduate Students (M.A.G.S), University of Louisville: Designed, organized, and led professional development programs (CV workshop, Networking Events) to foster professionalism amongst minority graduate students (2016)
• Student Alumni Ambassador, Student Organization for Alumni Relations (SOAR), UofL (2013)
• Volunteer at Park Du Valle Community Health Center (Pharmacy Department) Louisville, KY (2012)
• Created “Operation C.P.R (Communication. Planning/Preparation. Rescue)” adopted by 10-15 students TA for effective laboratory class teaching and tutoring as Laboratory Director (Berea College 2011-2012)
• Volunteer of Berea Buddies Mentor Program and One-on-One tutoring (Berea College 2009-2012)
• Volunteer of Hispanic Outreach Program, Berea College (2010)
• Volunteer of Adopt-a-grandparent, Berea College (2009)
• Bridging-The-Gap chair for the African Student Association: fostering interracial dialogue (2010)

PROFESSIONAL MEMBERSHIP

1. American Heart Association (AHA) 2016 – Present
2. American Association for the Advancement of Science (AAAS) 2016 – Present
3. International Society of Thrombosis and Haemostasis (ISTH) 2014 – Present
5. Biophysical Society (BPS) 2013 – Present
6. American Chemical Society (ACS) 2013 – Present
7. (NOBCChE) National Organization for the Advancement of Black Chemists and Chemical Engineers 2012 – Present