A Thesis

entitled

Conformational analysis of E. coli DnaT and the complex with the N-terminal domain of PriA

by Lindsey M. Easthon

Submitted to the Graduate Faculty as partial fulfillment of the requirements for the Masters of Science in Chemistry

Dr. Timothy Mueser, Committee Chair

Dr. Wendell P. Griffith, Co-Advisor

Dr. Donald Ronning, Committee Member

Dr. Patricia Komuniecki
College of Graduate School

The University of Toledo
May 2010
An Abstract of

Conformational analysis of *E. coli* DnaT and the complex with the N-terminal domain of PriA

Lindsey M. Easthon

Submitted as partial fulfillment of the requirements for
the Master of Science in Chemistry

The University of Toledo

May 2010

The fidelity of DNA replication is essential to maintaining genomic integrity. Therefore, pathways to restart DNA replication upon fork arrest are essential to an organism’s survival. In *E. coli* there are seven protein involved in the formation of the replication restart primosome, PriA, PriB, PriC, DnaT, DnaB, DnaC and SSB. To date, with the exception of PriB, there is very little structural information available for these proteins or the complexes formed during primosome assembly.

Electrospray ionization mass spectrometry (ESI-MS) can provide a great amount of structural information, beyond a simple mass measurement. Since ESI-MS is a soft ionization technique it allows for investigation of proteins under native folding conditions as well as non-covalent complexes. Combination of ESI-MS with ion mobility spectroscopy and collision induced dissociation experiments provide information about the shape and size of the protein and stability of non-covalent complexes.
This work focused on conformational analysis of *E. coli* DnaT and the N-terminal domain of PriA using ESI-MS. The oligomerization states of the proteins were determined under native conditions. For DnaT, metal binding studies, CID, IMS, and limited proteolysis studies were used to analyze the conformation of the proteins in solution. An attempt was made to investigate the complex formed between the N-terminal domain of PriA and DnaT but no complex was observed. Finally, a buffer screen was developed to determine volatile buffers and salts for metal binding studies suitable for use in ESI-MS experiments.
Acknowledgements

I would like to take this opportunity to thank my research advisors, Dr. Timothy Mueser and Wendell Griffith. Their expertise was essential to the success of this project. Furthermore, I am beyond grateful for them taking the time to share their knowledge. I am a better chemist for having worked for them. I would also like to thank Dr. Ronning for serving on my committee as well as providing advice when I would wander into his lab with questions. Many thanks to Dr. Funk for providing advice and insight into many of the experiments in this thesis.

My labmates have made my three years at Toledo much more enjoyable. Thank you to all the Mueser and Griffith and lab members past and present. I am extremely grateful for having such a fun, helpful and caring group to work with. I would also like to say a special thank you to Amanda Bryant-Friedrich. She has provided a great amount of encouragement and advice both in science and life in general. Finally, I would like to thank my family for their support and encouragement over the past three years. I am lucky to have such a tremendous support system.
### Table of Contents

ABSTRACT ...................................................................................................................... iii  
ACKNOWLEDGEMENTS ................................................................................................. v  
TABLE OF CONTENTS .................................................................................................... vi  
LIST OF TABLES ............................................................................................................. x  
LIST OF FIGURES ........................................................................................................... xi  
LIST OF ABBREVIATIONS ............................................................................................... xvi  

#### CHAPTER 1:

**Introduction** ................................................................................................................. 1  

1.1: Research goals ............................................................................................................ 1  

1.2: DNA replication ......................................................................................................... 1  

1.3: DNA replication restart ............................................................................................. 5  

1.4: Project aims ............................................................................................................... 9  

#### CHAPTER 2:

**Methods** ....................................................................................................................... 10  

2.1 Protein preparation and purification ........................................................................... 10  

2.1.1: Large-scale protein preparation ........................................................................... 10  

2.1.2: Cell lysis ............................................................................................................. 11  

2.1.3: SDS-PAGE ........................................................................................................ 11  

2.1.4 FPLC/HPLC protein purification ............................................................................ 12
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1.6 RPLC Chromatography</td>
<td>14</td>
</tr>
<tr>
<td>2.2 ESI-MS</td>
<td>14</td>
</tr>
<tr>
<td>2.2.1 Introduction</td>
<td>14</td>
</tr>
<tr>
<td>2.2.2 Sample Preparation for ESI-MS</td>
<td>16</td>
</tr>
<tr>
<td>2.2.3 Limited Proteolysis Using Trypsin</td>
<td>16</td>
</tr>
<tr>
<td>2.2.4 ESI-MS Analysis</td>
<td>17</td>
</tr>
<tr>
<td>2.2.5 Ion Mobility Mass Spectrometry</td>
<td>17</td>
</tr>
<tr>
<td>2.2.6: Collision induced dissociate of intact proteins</td>
<td>18</td>
</tr>
</tbody>
</table>

CHAPTER 3: *E. coli* DnaT

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Introduction</td>
<td>20</td>
</tr>
<tr>
<td>3.2 Protein Preparation</td>
<td>21</td>
</tr>
<tr>
<td>3.2.1 Protein Expression and Lysis</td>
<td>21</td>
</tr>
<tr>
<td>3.2.2 Protein Purification</td>
<td>23</td>
</tr>
<tr>
<td>3.2.3 Preparation of ESI-MS Samples</td>
<td>24</td>
</tr>
<tr>
<td>3.3 ESI-MS Studies</td>
<td>25</td>
</tr>
<tr>
<td>3.3.1: Protein Oligomerization Studies by ESI-MS</td>
<td>26</td>
</tr>
<tr>
<td>3.3.2: Effects of protein preparation and salt concentration of oligomerization state of DnaT</td>
<td>28</td>
</tr>
<tr>
<td>3.4 Conformational analysis of DnaT</td>
<td>32</td>
</tr>
<tr>
<td>3.4.1: Ion Mobility Data for DnaT</td>
<td>32</td>
</tr>
<tr>
<td>3.4.2 Metal Binding Studies of DnaT</td>
<td>33</td>
</tr>
</tbody>
</table>
3.4.3: Limited proteolysis studies....................................................36
3.4.4: Collision induced dissociation studies.................................43

3.5. Analysis of DnaT degradation product...................................47
3.5.1: Identification of degradation product.................................47
3.5.2: Crystallization studies of DnaT.............................................49

3.5: Conclusions...........................................................................53

Chapter 4: PriA N-Terminal Domain and Complex Studies with DnaT....54

4.1 Introduction.............................................................................54

4.2 Protein Preparation...............................................................58
4.2.1 Protein expression..............................................................58
4.2.2 Cell lysis............................................................................59
4.2.3 HPLC Purification..............................................................60

4.3: ESI-MS Studies of PriA N-Domain.......................................64
4.3.1: Native Conformation Studies............................................64
4.3.2 Complex studies with DnaT...............................................65

4.4 Conclusions.............................................................................66

CHAPTER 5: Volatile Buffer Screen for ESI-MS..............................68

5.1 Introduction.............................................................................68

5.2: Volatile Buffer Screen...........................................................69

5.3 Salt Studies..............................................................................74

5.4 Additive Screen.......................................................................81
5.4.1 Methods.............................................................................81
5.4.2: Effects of glucose on data quality.......................................81
5.4.3: Effects of sucrose on data quality

5.4.4: Effects of glycerol on data quality

5.5: Conclusions

REFERENCES
# List of Tables

Table 3.1: Physical Properties of DnaT.........................................................20

Table 3.2: Summary of the peptides identified by limited proteolysis studies of DnaT. ..........................................................43

Table 3.3: Mass, sequence and residue numbers for the peptide cleaved in DnaT degradation product.............................................52

Table 4.1: Physical properties of PriA N-terminal domain........................58

Table 4.2: Ion-exchange purification buffers for PriA N.................................61

Table 5.1: Test proteins.................................................................................70

Table 5.2: Test Buffers..................................................................................70
List of Figures

Figure 1.1: Schematic of a DNA replication fork………………………………………………3
Figure 1.2: DnaA binding at oriC and recruitment of DnaB•DnaC complex………………4
Figure 1.3: Assembly of proteins at replication fork…………………………………………4
Figure 1.4: DNA substrates for PriA and PriC…………………………………………………6
Figure 1.5: DNA replication fork containing terminal 3’ end for PriA recognition……………………………………………………………………………………………7
Figure 1.6: Assembly of the PriA dependent pathway…………………………………………8
Figure 2.1: Schematic of charged droplet formation .........................................................15
Figure 3.1: SDS-PAGE results for expression and extraction of DnaT…………………………22
Figure 3.2: Q-Sepharose chromatogram for DnaT…………………………………………..23
Figure 3.3: Poros HS chromatogram for DnaT…………………………………………………24
Figure 3.4: S200 purification and SDS-PAGE results for DnaT…………………………………25
Figure 3.5: ESI-MS results for DnaT at 1 and 5 uM in 10 mM NH₄H₂CCOO pH 7.0…………………………………………………………………………………………………26
Figure 3.6: ESI-MS data for 5 µM DnaT in 10 mM NH₄OAc buffers at varying pH values…………………………………………………………………………………………………27
Figure 3.7: ESI-MS data for 5 uM DnaT in 10 mM NH₄OAc pH 7.0 showing decreased dimer formation……………………………………………………………………………29
Figure 3.8: ESI-MS data for 5 µM DnaT with NH₄Cl addition……………………………………30
Figure 3.9: ESI-MS data for DnaT at concentration from 5-40 µM in 10 mM NH₄OAc pH 7.0……………………………………………………………………………………31
Figure 3.9: Ion mobility spectra for DnaT at pH 7.0 (a) and 3.0 (b) ......................32

Figure 3.10: Figure 3.11: ESI-MS data for 5 µM DnaT at varying Mg$^{2+}$ concentrations ..................................................................................34

Figure 3.11: ESI-MS data showing the +7 charge state for DnaT in the presence of Mg$^{2+}$ ions. .................................................................35

Figure 3.11: Plot of relative intensity for +7 charge state vs Mg$^{2+}$ concentration..............................................................................36

Figure 3.13: Zero hour ESI-MS spectra for 2µM DnaT in 50%H$_2$O/50% acetonitrile and 0.1% FA prior to limited proteolysis.........................38

Figure 3.14: ESI-MS data for 2 µM limited proteolysis sample for DnaT after 20 minutes ..............................................................................39

Figure 3.15: Fragmentation data for the 958.5 ion from 20 minute DnaT limited proteolysis. .................................................................40

Figure 3.16: ESI-MS data for 2 µM limited proteolysis sample of DnaT at 60 minutes ...............................................................................41

Figure 3.17: ESI-MS fragmentation data for 1036.6 peak for DnaT limited proteolysis at 60 minutes ............................................................42

Figure 3.18: Amino acid sequence for DnaT showing region cleaved by trypsin in limited proteolysis studies ........................................43

Figure 3.19: CID data for the 2761.4 peak for 5 µM DnaT 10 mM NH$_4$OAc at pH 7.0 (a) and pH 3.0 (b). .........................................................44

Figure 3.20: CID data for 5 µM DnaT in 10 mM NH$_4$OAc at varying collision energies for the 1208.6 peak at pH 7.0 (a) and 3.0 (b) ..............46

Figure 3.21: ESI-MS data for 3 µM DnaT sample containing degradation product in denaturing buffer and deconvoluted masses.............48

Figure 3.22: RPLC chromatogram for DnaT containing degradation product......49

Figure 3.23: Crystals of DnaT used in ESI-MS studies.........................................................51

Figure 3.24: ESI-MS data for the protein obtained from the DnaT crystals...........52
Figure 4.1: DNA substrates recognized by PriA .............................................55
Figure 4.2: Crystal structure of PriA N-terminal domain (PDB id:2D7E) ..........57
Figure 4.3: SDS-PAGE results for expression of PriA N-terminal domain ......59
Figure 4.4: SDS-PAGE results for lysis of PriA N-terminal domain .............60
Figure 4.5: SP-Sepharose chromatogram for PriA N ...............................61
Figure 4.6: HA chromatogram for PriA N .............................................62
Figure 4.7: Poros HS chromatogram for PriA N .................................63
Figure 4.8: S75 chromatogram for PriA N in 100 mM NH₄OAc ..................64
Figure 4.9: ESI-MS data for 5 µM PriA N domain in 10 mM NH₄H₃C COO buffers at varying pH values. .................................................65
Figure 4.10: ESI-MS spectrum for 6 µM PriA N domain complex with 6µM DnaT in 10 mM NH₄OAc pH 7.0 .................................66
Figure 5.1: ESI-MS buffer screen results for 3 µM myoglobin .....................71
Figure 5.2: ESI-MS buffer screen results for 5 µM transferrin ....................72
Figure 5.3: ESI-MS buffer screen results for 10 µM BSA ............................73
Figure 5.4: ESI-MS results for 3µM myoglobin in 10 mM NH₄OAc pH 7.0 with 50 mM Mg²⁺ salts .........................................................74
Figure 5.5: ESI-MS results for 3 µM myoglobin in 10 mM NH₄OAc pH 7.0 with 50 mM Ca²⁺ salts .........................................................75
Figure 5.6: ESI-MS results for 5 µM transferrin in 10 mM NH₄OAc pH 7.0 with 50 mM Mg²⁺ salts .........................................................75
Figure 5.7 ESI-MS results for 5 µM transferrin in 10 mM NH₄OAc pH 7.0 with 50 mM Ca²⁺ salts .........................................................76
Figure 5.8: ESI-MS results for 10 µM BSA in 10 mM NH₄OAc pH 7.0 with 50 mM Mg²⁺ salts .........................................................76
Figure 5.9: ESI-MS results for 10 µM BSA in 10 mM NH₄OAc pH 7.0 with 50 mM Ca²⁺ salts .........................................................77
Figure 5.10: ESI-MS data for the +8 charge state for 3 µM myoglobin with the 50 mM Mg$^{2+}$ salts investigated
Figure 5.10: ESI-MS data for the +8 charge state for 3 µM myoglobin with the 50 mM Ca$^{2+}$ salts investigated
Figure 5.11: ESI-MS data for the +19 charge state of 5 µM transferrin with 50 mM Mg$^{2+}$ salts
Figure 5.12: ESI-MS data for the +19 charge state of 5 µM transferrin with 50 mM Ca$^{2+}$ salts
Figure 5.13: ESI-MS data for the +16 charge state of 10 µM BSA with 50 mM Mg$^{2+}$ salts
Figure 5.14: ESI-MS data for the +16 charge state of 10 µM BSA with 50 mM Ca$^{2+}$ salts
Figure 5.15: ESI-MS data for 3 µM myoglobin in 10 mM NH$_4$OAc pH 7.0 with varying ratios of glucose
Figure 5.16: ESI-MS data for the +8 charge state of myoglobin in 10 mM NH$_4$OAc with varying ratios of glucose
Figure 5.17: ESI-MS data for 5 µM transferrin in 10 mM NH$_4$OAc pH 7.0 with varying ratios of glucose
Figure 5.18: ESI-MS data for the +19 charge state of transferrin in 10 mM NH$_4$OAc with varying ratios of glucose
Figure 5.19: ESI-MS data for 10 µM BSA in 10 mM NH$_4$OAc pH 7.0 with varying ratios of glucose
Figure 5.20: ESI-MS data for the +16 charge state of BSA in 10 mM NH$_4$OAc with varying ratios of glucose
Figure 5.21: ESI-MS data for 3 µM myoglobin in 10 mM NH$_4$OAc pH 7.0 with varying ratios of sucrose
Figure 5.22: ESI-MS data for the +8 charge state of myoglobin in 10 mM NH$_4$OAc with varying ratios of sucrose
Figure 5.23: ESI-MS data for 5 µM transferrin in 10 mM NH$_4$OAc pH 7.0 with varying ratios of sucrose
Figure 5.24: ESI-MS data for the +19 charge state of transferrin in 10 mM NH₄OAc with varying ratios of sucrose

Figure 5.25: ESI-MS data for 10 µM BSA in 10 mM NH₄OAc pH 7.0 with varying ratios of sucrose

Figure 5.26: ESI-MS data for the +16 charge state of BSA in 10 mM NH₄OAc with varying ratios of sucrose

Figure 5.27: ESI-MS data for 3 µM myoglobin in 10 mM NH₄OAc pH 7.0 with varying ratios of glycerol

Figure 5.28: ESI-MS data for the +8 charge state of myoglobin in 10 mM NH₄OAc with varying ratios of glycerol

Figure 5.29: ESI-MS data for 5 µM transferrin in 10 mM NH₄OAc pH 7.0 with varying ratios of glycerol

Figure 5.30: ESI-MS data for the +19 charge state of transferrin in 10 mM NH₄OAc with varying ratios of glycerol

Figure 5.31: ESI-MS data for 5 µM BSA in 10 mM NH₄OAc pH 7.0 with varying ratios of glycerol

Figure 5.32: ESI-MS data for the +16 charge state of BSA in 10 mM NH₄OAc with varying ratios of glycerol
List of Abbreviations

% (v/v)………….% volume/volume
% (w/v)………….% weight/volume
AEBSF………….4-(2-Aminoethyl)-benzylsulfonylfluoride
Amp……………..Ampicillin
ATP……………...Adenosine Triphosphate
A_x……………….Absorption, at x nm
BME……………β-Mercaptoethanol
Cam……………...Chloramphenicol
CID……………….Collision induced dissociation
DTT……………..Dithiothreitol
E. coli……………Escherichia coli
EDTA…………....Ethylenediaminetetraacetic acid
ESI……………….Electrospray Ionization
HEPES………….4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HLPC…………….High-Performance Liquid Chromatography
IMS……………….Ion mobility spectroscopy
IPTG……………Isopropyl-β-D-thiogalactopyranoside
Kan……………..Kanamycin
LB……………….Luria Broth
MALDI…………..Matrix Assisted Laser Desorption Ionization
MW……………..Molecular Weight
MWCO…………..Molecular Weight Cut-Off
OB-fold…………..oligosaccharide/oligonucleotide-binding fold
OD_x…………….Optical Density, at x nm
PAS…………....Primosome Assembly Site
PDB………………Protein Data Bank
PEI………………Poly(ethylenimine)
pI………………….Isoelectric point
PIES…………….Piperazinbis(ethansulfonic acid)
RT………………Room Temperature
SDS-PAGE……..Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis
SSB…………….Single-Stranded DNA-Binding protein
ssDNA…………Single-stranded deoxyribonucleic acid
TAPS…………..N-Tris(hydroxymethyl)methyl-3-aminopropanesulfonic acid
TEDG…………..Tris, EDTA, DTT, Glycerol
TOF…………….Time-of-Flight
Q-TOF…………Quadrupole-Time-of-Flight
Tris…………….Tris(hydroxymethyl)aminomethane
UV-Vis…………..Ultraviolet-Visible
ε_x………………….Molar extinction coefficient, at x nm
Chapter 1:

Introduction

1.1 Research goals

DNA replication, repair and restart are highly coordinated processes necessary for organism survival. Previous work in the Mueser laboratory has focused on using X-ray crystallography and several other biophysical techniques to study these pathways in archaeal and prokaryotic organisms. The focus of this thesis was the DNA replication restart proteins found in *E. coli*. These seven proteins, PriA, PriB, PriC, DnaT, DnaB, DnaC and SSB are responsible for the reinitiating of DNA synthesis at stalled replication forks and recombination intermediates\(^1\)\(^-\)\(^3\). Specifically, electrospray ionization mass spectrometry (ESI-MS) was used to investigate the conformation of DnaT and the N-terminal domain of PriA.

1.2 DNA Replication

All the genetic information for a given organism is contained within its DNA. DNA is responsible for encoding all RNA and subsequent proteins, which in turn, coordinate all cellular processes. Survival of an organism is therefore dependent on its ability to successfully replicate its genetic code. Given the importance of DNA replication, it is a highly coordinated process involving numerous proteins.
The bacterial chromosome is a large circular, double stranded piece of DNA. Replication is a semiconservative process in which both parental strands serve as the template for two complementary daughter strands. Each strand of DNA has a polarity. On one end of the strand is the 3’ hydroxyl group of the ribose moiety of the nucleotide and the other is the 5’ phosphate of the final nucleotide. Complementary strands of duplexed DNA are antiparallel to each other. DNA chain elongation is catalyzed by the enzyme DNA polymerase III that adds the incoming dNTP in the 5’ to 3’ direction of the growing strand. Replication occurs at a DNA structure called a replication fork shown in Figure 1.1. In order for DNA polymerase to catalyze chain elongation several requirements need to be met. First, DNA polymerase must have a single stranded template DNA strand. The mechanism by which duplex DNA is opened for initiation or replication is discussed below. Secondly, DNA polymerase must have a terminal hydroxyl group to add the 5’ phosphate of the incoming dNTP. This is accomplished by the use of RNA primers which are synthesized and annealed to the DNA by the primase, DnaG⁴.

As previously mentioned, chain elongation always occurs in the 5’ to 3’ direction. As the chains run antiparallel, only one chain can be synthesized continuously. This strand is termed the leading strand. The second chain, the lagging strand, must be synthesized discontinuously in short segments called Okazaki fragments. Short stretches of RNA primers are laid down on the lagging strand by DnaG, the primase. DNA polymerase can then act on the terminal hydroxyl group with the incoming dNTP.
Figure 1.1: Schematic of a DNA replication fork

Initiation of replication in *E. coli* occurs at the oriC and proceeds bidirectionally. DnaA recognizes the oriC in a sequence specific manner. DnaA polymerization, along with several other proteins, on the DNA causes positive supercoiling of the duplex DNA resulting in duplex opening of the DNA unwinding element (DUE). Six copies of DnaC, the hexameric DNA helicase loader, can then load a hexamer of DnaB, the primary 5’-3’ DNA helicase, onto the single stranded regions of the DUE. Figure 1.2 shows a schematic of DnaA binding at the oriC.
DnaB is the primary helicase and travels in the 5’ to 3’ direction of the single stranded DNA. Two replication forks are therefore formed by the loading of two DnaB hexamer and can travel in opposite directions. Once DnaB has been loaded onto the DNA, DNA polymerase, the primase and single stranded binding proteins can then associate and replication proceeds\textsuperscript{4,6}. The assembly of the proteins required for replication can be seen in Figure 1.3.

**Figure 1.3**: Assembly of proteins at replication fork.

---

**Figure 1.2**: DnaA binding at oriC and recruitment of DnaB•DnaC complex

![Diagram of DnaA binding at oriC and recruitment of DnaB•DnaC complex](image-url)
1.3: DNA Replication Restart

While DNA replication is a highly precise process, it is unlikely that the entire genome can be replicated without encountering DNA damage that results in replication fork arrest\(^7\). DNA damage can occur by a variety of mechanisms including nucleotide mismatch, degradation, and oxidative damage. To ensure survival, organisms have evolved a variety of complex DNA repair mechanisms. Once DNA damage has been repaired, DNA replication must restart in an origin independent manner.

In *E. coli*, DNA replication restart is accomplished by seven proteins: PriA, PriB, PriC, DnaB, DnaC, DnaG and SSB. These proteins are referred to as the ΦX174-type primosomal proteins. These proteins were identified through *in vitro* DNA replication studies using the bacteriophage ΦX174 DNA as a template\(^8\)\(^{-10}\). The ΦX174 DNA forms a hairpin structure, which serves as a primosome assembly site (PAS) for the restart proteins\(^11\),\(^12\). The ΦX174-type primosomal proteins provide both the DNA helicase and RNA priming activities necessary for synthesis of the complementary strand by DNA polymerase III\(^13\). These proteins are critical for restart of DNA replication in an origin independent manner. They function not only at stalled replication forks but also at recombination intermediates\(^14\),\(^15\).

Following replication fork arrest, DNA replication continues once DnaB has been reloaded onto the DNA by DnaC. In *E. coli*, there are two proposed pathways for re-initiation of DNA replication. The pathways are dependent of the DNA substrate involved and therefore require different proteins to initiate the loading of DnaB. The primary pathway is initiated by PriA and involves PriB, DnaT and possibly PriC, as
well. The second, less understood pathway involves the initiation of replication restart by PriC\textsuperscript{16}. Figure 1.4 shows the DNA substrates preferred by PriA and PriC.

**PriA substrates**

- **D-loop**

**PriC Substrate**

- **Gapped leading strand**

- **R-Loop**

- **Gapped lagging strand**

**Figure 1.4: DNA substrates for PriA and PriC\textsuperscript{17-19}**

PriA does not have sequence specific binding but is structurally specific\textsuperscript{17, 18, 20, 21}. DNA binding studies have shown that PriA binds to DNA substrates that contain a region of single stranded DNA such as D loops (Figure 1.5) but not DNA templates that contain double stranded regions on the both the leading and lagging strands such as Holliday junctions\textsuperscript{16}. PriA binding to recombination intermediates requires a terminal 3’ hydroxyl group on the leading strand\textsuperscript{18}. Removal or phosphorylation of the 3’ hydroxyl group disrupts binding indicating its
importance in the interaction. This PriA directed assembly of the replication restart primosomal proteins at a replication fork has been shown to require PriB and DnaT. The presence of PriC was not required, however in its presence, the rate of replication was increased by a factor of 2-3 fold for both φX174 complementary strand synthesis and D-loop containing template reactions\textsuperscript{12,22,23}.

Figure 1.5: DNA replication fork containing terminal 3’ end for PriA recognition

Once PriA recognizes and binds to the DNA substrate, PriB and DnaT then bind to the preprimosome. The assembly of these proteins on an SSB coated replication fork can be seen in Figure 1.6. PriB has been found to be a dimer both in Western blotting experiments as well as in crystal structures\textsuperscript{12,24,25}. The oligomerization state of DnaT is unclear. It is possible that the protein functions as a monomer, or as a trimer on every third primosome. PriC can serve to duplicate the function of PriB in the PriA pathway however, restart does not occur in the absence of PriB.
Figure 1.6: Assembly of the PriA dependent pathway

To date, much less is known about the PriC dependent pathway. Genetic analysis indicates an alternate pathway in which PriC can function to load DnaB onto a stalled replication fork\(^{16,26}\). While PriA null mutants have poor viability, are defective in homologous recombination and are sensitive to DNA-damaging agents such as UV radiation\(^{27}\), they are still able to survive indicating a separate mechanism for restart\(^{27-29}\). PriB or PriC knockout cells show only mild defects in replication restart. However, double knockouts for PriB and PriC are more severely affected than PriA mutants\(^{26}\). As shown in Figure 1.4, PriC has been found to function at fork DNA substrates that have large gaps in the leading strand of at least 7 nucleotides but are not highly active at PriA preferred substrates with no gaps in the leading strand\(^{16,19}\).

1.4: Project Aims

Given the importance of replication restart to bacterial viability, a clear understanding of the protein-protein and protein-DNA interactions is critical. Thus far, there is little structural information available for these proteins and complexes, with the exception of PriB\(^{24,30}\). The goal of this project was to utilize ESI-MS to gain structural information on the *E. coli* primosomal proteins in order to facilitate and
compliment X-ray crystallographic experiments. This thesis presents preliminary results for DnaT and PriA N terminal domain.
Chapter 2:

Methods

2.1: Protein preparation and purification

2.1.1 Large-scale protein preparations

Glycerol stocks for the protein were inoculated into 150 mL of LB containing the proper antibiotic. The flasks were shaken at 37°C overnight. The next morning, the cultures were divided among six 1L flasks containing LB and antibiotic. The flasks were again shaken at 37°C for several hours. The OD$_{600}$ of the cultures was monitored frequently. When the OD$_{600}$ reached approximately 0.50-0.70, the cells were induced with 1 mM IPTG for protein expression. Protein expression was monitored by SDS samples taken prior to induction as well as prior to harvest. Cells were harvested by centrifugation at 5525 g. Cell pellets were frozen at -20°C until lysis.

2.1.2 Cell lysis

Protein was extracted from the cells by use of a high-salt lysis. Cell pellets were resuspended in a high-salt lysis buffer (25 mM Tris-HCl pH 7.5, 500 mM NaCl, 2 mM EDTA, 20 mM βME and 20% (w/v) sucrose). Approximately 100 mL of lysis buffer was used for every 10 g cell pellet. After the cells had been resuspended a small amount of lysozyme was added to assist lysis. In order to eliminate protease activity 4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride (AEBSF) was also
added to the slurry. The solution was stirred in ice for 20 minutes followed by sonication for two minutes. Polyethyleneimine (PEI) (0.3%) was added to precipitate DNA. The mixture was then centrifuged at 18,000 g for 30 minutes. SDS samples of the pellet and lysate were taken to verify protein solubility. After centrifugation, the supernatant was collected and the pelleted cell debris discarded. The protein underwent a preliminary purification step using NH₄SO₄ fractionation to precipitate the desired protein. The protein was then centrifuged again at 18,000 g. The protein precipitate was resuspended in Buffer A (25 mM Tris-HCl pH 7.5, 50 mM NH₄Cl) for ion exchange purification with the addition of 10% v/v glycerol. The lysates were frozen at -80°C until further purification.

**2.1.3 SDS-PAGE**

In order to monitor protein expression and solubility samples were prepared for SDS-PAGE. SDS-PAGE is a denaturing electrophoresis method that separates proteins based upon molecular weights.

For the large-scale protein preparations described below, 1 mL samples of the cell culture were pelleted prior to induction. This sample was termed the 0 hr sample. After induction, a 3 hr sample was taken. The samples were centrifuged at 14000 g for 3 minutes to pellet the cells. The supernatant was discarded and the pelleted cells were resuspended in 50 µL Bug Buster™ Protein Extraction Reagent (Novagen, Gibbstown, NJ) and 50 µL of 2X NuPAGE™ SDS sample buffer (Invitrogen, Carlsbad, CA) were added as well. The gel samples were boiled for 10 minutes to further denature the proteins. They were then loaded onto a pre-cast NuPAGE™ 4-12% Bis-Tris gel. The gel was run in a NuPAGE® 1X MES SDS Running Buffer
(Invitrogen, Carlsbad, CA) at 200V for 35 minutes. The gel was then removed from the plastic casing and stained with Bio-Safe™ Commissie dye (BioRad, Hercules, CA). The gels were stained for several hours. In order to visualize the protein bands, the stain was removed using a destain solution containing 30% methanol, 10% glacial acetic acid.

2.1.4 FPLC/HPLC protein purification

Highly purified protein was obtained using FPLC. In FPLC, a pressure resistant column is packed with a stationary phase through which an aqueous buffer is passed to elute the protein. Primarily two FPLC chromatographic techniques were used to purify the proteins in this study: ion-exchange and size-exclusion chromatography.

All FPLC runs were conducted at 4°C using a BioLogic DuoFlow™ HPLC system (BioRad). Protocols and running buffers were optimized for each type of column (Table 2.1). In general, ion exchange protocol utilized a linear gradient from low to high salt concentration to elute the protein and size exclusion purification was performed in the proteins’ optimal buffers. Fractions were collected using the automated fraction collector. Protein elution was monitored by UV/VIS spectroscopy at 280 and 260 nm.

Following the lysis, proteins were purified using a low-resolution ion exchange Sepharose resin. Sepharose (Amersham Biosciences) resins are composed of cross-linked agarose-dextran matrix that can be functionalized to select for anions or cations. Proteins with acidic pI’s were purified using Q-Sepharose. Q-Sepharose resins are derivatized with a quaternary amine group to select for negatively charged
proteins. Proteins with basic pI’s were purified using SP-Sepharose. SP-Sepharose resins contain a sulfolpropyl group (O(CH₂)₃-SO₃⁻) to select for positively charged proteins. For both types of ion exchange, proteins were eluted using a linear gradient from low to high salt concentration.

After the low-resolution purification, a second round of ion exchange chromatography was used to purify the proteins. Poros Perfusion Chromatography utilizes ion exchange chromatography in combination with the capability of sustaining very high pressures to obtain high-resolution purification. Poros resins are polysytrenedivinylbenzene particles that contain pores through which molecules can pass. They are also chemically modified to have ion selectivity. The two types of Poros resins are Poros HS, used for cation exchange and Poros HQ used for anion exchange. The HS resin is derivatized with sulfopropyl groups for cation exchange and HQ resign is a quarternized polyethylenemine derivative for anion exchange.

In order to remove DNA from proteins which bind DNA, CHT™ ceramic hydroxyapatite resin (Bio-Rad 157-0040) was used. HA resin, Ca₁₀(PO₄)₆(OH)₂, contains positively charged calcium ions which compete for the negatively charged phosphate backbone of DNA. Increasing concentrations of phosphate were used to elute the protein from the column.

The final purification step utilized size exclusion chromatography. SEC separates molecules based upon size, with the larger proteins eluting from the column first. Two SEC resins, Superdex75 and Superdex200 were used. Superdex75 has an upper mass limit of 75 kDa for which it can effectively resolve molecules and Superdex200, a 200 kDa upper mass limit. The size exclusion runs were also utilized
for buffer exchange into NH$_4$OAc when the proteins were being prepped for ESI-MS experiments.

2.1.6 RPLC Chromatography

Reversed phase liquid chromatography was used to separate proteins under denaturing conditions. RPLC uses a nonpolar stationary phase, C$_8$, in these experiments. A gradient of increasing solvent hydrophobicity is used to elute analytes based upon differences in polarities. Flow rates were altered in an attempt to optimize separation. The elution of the analytes was monitored using a UV detector at wavelengths of 214 and 280 nm. Fractions were collected manually in 500 µL aliquots.

2.2 ESI-MS

2.2.1 Introduction

Since its development in the 1980’s, electrospray ionization has proven to be a highly effective tool for the study of proteins and other biological macromolecules$^{31,32}$. In ESI, large non-volatile molecules are volatilized and ionized for analysis directly from solution. ESI is considered to be a “soft” ionization method as it does not result in fragmentation of the analyte molecules. Ionization occurs due to an electrical potential applied to a heated metal capillary, through which the analyte is injected into the mass spectrometer. The electrical potential results in the formation of a Taylor cone from which the now charged droplets are sprayed into the instrument. Figure 2.1 shows a schematic of charged droplet formation. The ions pass into the desolvation chamber where they are thermally and collisionally desolvated. Once the droplets have been completely desolvated the “naked” ions
pass into the mass analyzer where they are separated based upon their $m/z$ ratio. Once the ions have been separated they then pass to the detector.

**Figure 2.1: Schematic of charged droplet formation**

ESI is especially useful for analysis of protein stability and dynamics as it produces multiply charged species as opposed to the singly charged species that predominate in MALDI. The charge state distributions provide information on protein stability and dynamics. Broad charge state distributions at low $m/z$ values generally correspond to proteins in an open, unfolded conformation whereas narrow charge state distributions are representative of compact, native folding. ESI is a gentle enough method that it can also be utilized to study non-covalent protein-ligand and large multiprotein complexes.
2.2.2 Sample Preparation for ESI-MS

Sample preparation in ESI is critical to data quality. Analyte samples must be thoroughly desalted prior to introduction into the mass spectrometer. Typically, protein samples are prepared in dilute volatile buffers, the most common being NH₄OAc. Buffer exchange can be achieved by a variety of methods including extensive dialysis, the use of centrifuge filtration devices and liquid chromatography. Following the buffer exchange, the protein concentration was determined using UV-VIS spectroscopy absorbance at 280 nm.

2.2.3 Limited Proteolysis Using Trypsin

Limited proteolysis was used to identify regions of structural disorder. Proteins were buffer exchanged into stock solutions of 200 µM protein in 100 mM NH₄HCO₃ buffers at pH values between 8.0 and 8.5. The enzyme:substrate ratio varied between 1:200 and 1:2000 to optimize identification of the first cleavage products. Reactions were carried out between 21° C and 37° C using a dry block. Samples were taken at varying time increments and diluted to a final concentration of 2 µM in a 50% H₂O/ 50% acetonitrile and 0.1% formic acid (FA) buffer on ice to quench the reaction.

Data collection was performed using a Synapt HDMS quadrupole time-of-flight mass spectrometer (Waters, Manchester, UK) in positive ion mode. The analyte was ionized using homemade fused silica capillaries. In the ion trap, the trajectory of the precursor ion is stabilized by alternating the rf frequency of the quadrupole, allowing for isolation of the desired ion. Fragmentation of the precursor ion is achieved by in the collision cell and mass spectrum is acquired by the TOF³³.
2.2.4 ESI-MS Analysis

Data were collected in positive ion mode on a Synapt HDMS Q-TOF ion mobility mass spectrometer. The instrument utilizes a nanospray source with homemade nano-spray emitters. Ionization source voltage and flow rate were optimized to generate the best signal and were dependent on the buffer used as well as the sample. For all data sets, 300 scans were averaged. Working solutions between 1 and 5 µM were prepared from a 200 µM stock solution in 100 mM NH₄OAc. The protein was desalted by filtration using Amicon® Ultra 15 mL concentrator tubes with a molecular weight cutoff of 10,000 Da (Millipore).

2.2.5 Ion Mobility Mass Spectrometry

The Synapt HDMS Q-TOF mass spectrometer is equipped to perform ion mobility mass spectrometry. Ion mobility mass spectrometry has the ability to separate molecules not only on the mass to charge ratio but as well as their interaction with a buffer gas. The separation achieved upon interaction with the gas cloud is based upon the collisional cross section of the analyte molecules. Molecules in an extended conformation, such as unfolded proteins, have little interaction with the gas. However, proteins in a natively folded conformation will have a larger collisional cross section and therefore, their migration through the instrument will be impeded.

For IMS experiments, the same instrument was used for data collection as previously described. Data was collected in a mass range of 50 ≤ m/z ≤ 4000 after external calibration with a cesium iodide standard. All data sets are an average of 600 scans.
2.2.6: Collision induced dissociation of intact proteins

Collision induced dissociation is a method of fragmentation performed within the mass spectrometer. In CID, the analyte molecules are accelerated within the quadrupole by a voltage potential. The increased velocity of the analyte molecules increases the force with which they interact with the argon gas molecules in the collision chamber. The increase force of the collision can result in fragmentation of the protein molecule.

If a protein contains regions of structural disorder, the areas that are not ordered will be more susceptible to fragmentation. The triple quadrupole of the Synapt HDMS mass spectrometer allows for isolation of peaks that correspond to specific folding states of the protein. Ions that represent the folded state of the protein and the unfolded state of the protein were selected and subjected to CID at varying levels of collision energy. Different solution conditions were also tested to investigate the effect of pH on the folding of the protein. Fragments were identified using Fragpro software.
Chapter 3:

**E. coli DnaT**

### 3.1 Introduction

*E. coli* DnaT is an essential component of the DNA replication restart primosome, however little is known about its structure and specific role in replication restart. The physical properties of DnaT are summarized in Table 3.1. Solution studies indicate that the protein is a trimer in solution and CD data is consistent with α+β secondary structure as well as a dependence on Mg$^{2+}$ ions to stabilize secondary structure.

**Table 3.1 Physical Properties of DnaT**

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecular Weight</td>
<td>19324 Da</td>
</tr>
<tr>
<td>Amino Acids</td>
<td>178</td>
</tr>
<tr>
<td>Extinction Coefficient</td>
<td>27946 M$^{-1}$cm$^{-1}$</td>
</tr>
<tr>
<td>pI</td>
<td>5.1</td>
</tr>
</tbody>
</table>

DnaT is believed to function in the PriA mediated primosome assembly. As previously mentioned, DnaT was found to be a trimer in solution. However, it is believed that only a monomer of DnaT functions in the primosome$^{12,22}$. While DnaT does not have any DNA binding capability, it functions in critical protein-protein interactions between the DNA•PriA•PriB complex and the DnaB helicase. It has
previously been shown that at high concentrations, DnaT can bypass the role of PriB and bind to the DNA•PriA complex\textsuperscript{34}.

3.2 Protein Preparation

3.2.1 Protein Expression and Lysis

The \textit{E. coli} DnaT gene (EG 10244, 540 bp), which was a gift from Dr. Nakai, had been cloned into pET21a expression vector and transformed into Rosetta cells by Aude Izac\textsuperscript{35}.

\textit{E. coli} DnaT was expressed in Rosetta cell lines in LB growth medium. Small-scale cultures were started the night before by inoculating 150 mL of LB with DnaT glycerol stock. Rosetta cells containing the DnaT expression plasmid were selected for by the addition of 1 mM ampicillin. The overnight cultures were grown at 37\textdegree C with gentle shaking. The next morning, the cell growth was transferred to 6-1 L flasks of LB containing 1 mM ampicillin. The cells were grown at 37\textdegree C to an \textit{OD}\textsubscript{600} of 0.40-0.60 before induction with 1 mM IPTG. Following induction, the cells were grown for an additional 3 hours prior to harvesting by centrifugation. Protein expression was monitored by SDS-PAGE.

The protein was extracted by a high-salt lysis buffer consisting of 25 mM Tris-HCl pH 7.5, 500 mM NaCl, 10 mM \(\beta\)ME, 2 mM EDTA and 10\% w/v sucrose. The cells were resuspended in the lysis buffer with stirring and lysis was catalyzed by the addition of lysozyme. After stirring for 20 minutes, the solution was sonicated for 2 minutes and 3\% PEI added. The lysate was then centrifuged at 18000 g for 30 minutes to remove cellular debris. DnaT was precipitated using NH\textsubscript{4}SO\textsubscript{4} fractionation. The solution was then centrifuged at 18000 g for 15 minutes. DnaT
was resuspended using a 25 mM Tris HCl pH 7.5 and 50 mM NH₄Cl buffer. The protein was then frozen at -80°C with 20% glycerol. Figure 3.1 shows SDS-PAGE results for the expression and extraction of DnaT.

Figure 3.1 SDS-PAGE results for expression and extraction of DnaT

3.2.2 Protein Purification

Prior to HPLC purification, the DnaT lysate was thawed and filtered. The conductivity of the protein was lowered to approximately 5.0 mS/cm to match that of
the low-salt buffer A for Q-Sepharose purification using a 20 mM Tris-HCl pH 7.5 buffer.

Initial low-resolution purification was performed using a Q-Sepharose column. The protein was eluted with a linear gradient from 100% buffer A (25 mM Tris-HCl pH 7.5, 50 mM NH₄Cl) to 100% buffer B (25 mM Tris-HCl pH 7.5, 50 mM NH₄Cl, 1.5 M NaCl) over 60 minutes. Protein elution was monitored by UV spectroscopy at 280 and 260 nm. Figure 3.2 shows the chromatograms for DnaT Q-Sepharose purification.

Figure 3.2: Q-Sepharose chromatogram for E. coli DnaT

Following Q-Sepharose purification, the protein was dialyzed against 25 mM Tris-HCl pH 7.5 and 50 mM NH₄Cl solution in order to reduce the salt concentration for Poros HS purification. The protein solution was then loaded onto the Poros HS column and again eluted using a linear gradient with the same protocol as previously
used for the Q-Sepharose column. A representative chromatogram for Poros HS purification of DnaT is shown in Figure 3.3.

Figure 3.3: Poros HS chromatogram for *E. coli* DnaT

The fractions containing DnaT were concentrated and injected onto a Superdex 200 column for further purification as well as buffer exchange. The S200 column was run in 100 mM NH₄OAc buffer for mass spectrometry studies. Again the collected fractions were run on a SDS-PAGE gel to identify the protein of interest and purity. Figure 3.4 shows the S200 chromatogram.
3.2.3 Preparation of ESI-MS Samples

_E. coli_ DnaT protein that had been previously prepared by Deepa Singh was buffer exchanged into 100 mM NH₄OAc buffer using Amicon® Ultra 15 mL concentrator tubes with a molecular weight cutoff of 10,000 Da (Millipore) by centrifugation at 3500 g at 10°C. Protein that had been S200 purified in 100 mM NH₄OAc and frozen at -80°C with glycerol was prep for ESI-MS studies by using 500 µL Amicon® Ultra centrifugal filters with a 10 kDA MWC (Millipore) to remove glycerol with a 100 mM NH₄OAc buffer.

3.3 ESI-MS Studies

ESI-MS spectra for DnaT were collected using a Synapt HDMS (Waters, Manchester, UK) quadrupole time of flight (Q-TOF) mass spectrometer. All mass
spectra were collected in positive ion mode. The following instrument parameters were kept constant: cone voltage, 40V; cone gas, 0 L/h; collision energy 6kV. The capillary voltage and flow rate were optimized to obtain the highest quality data.

3.3.1: Protein Oligomerization Studies by ESI-MS

ESI-MS was used to analyze protein conformation under varying solvent conditions. To investigate protein oligomerization state as a function of concentration, DnaT solutions were made at 1 and 5 µM protein in 10 mM NH₄OAc, pH 7.0. The spectra for the two concentrations of DnaT can be seen in Figure 3.5.

Figure 3.5: ESI-MS results for DnaT at 1 and 5 uM in 10 mM NH₄OAc pH 7.0. Peaks for the monomer (M) and dimer (D) are labeled with the corresponding charge state.
The experimentally determined mass for the DnaT monomer was 19343.9 Da which is consistent with the sequence with the N-terminal methionine residue having been cleaved after synthesis. DnaT shows three charge state distributions in NH₄OAc pH 7.0. The low m/z values indicate a largely disordered protein conformation. DnaT exhibits two oligomeric states, one for a monomer and a second for a dimer. The peaks corresponding to the monomer and dimer are labeled in Figure 3.5.

DnaT structure as a function of pH was also analyzed by utilizing 10 mM NH₄OAc buffers at pH values between 3 and 10. Working solutions of 5 µM protein were used in all experiments. An overlay of these spectra can be seen in Figure 3.6.

![Figure 3.6: ESI-MS data for 5 µM DnaT in 10 mM NH₄OAc buffers at varying pH values. Peaks for the monomer (M) and dimer (D) are labeled with corresponding charge state.](image)
As would be expected at either of the pH extremes, there is an increase in the unfolded state of the protein. However, DnaT still maintains a significant amount of native folding even under very acidic or basic conditions. This indicates that the intramolecular interactions that are responsible for maintaining protein structure are more likely to be van der Waals forces as opposed to ionic interactions, which would be disrupted upon titration with acid or base.

3.3.2: Effects of protein preparation and salt concentration of oligomerization state of DnaT

It was noticed that the oligomeric state of DnaT in 10 mM NH₄OAc buffers was different between protein preparations. The data collected in § 3.3.1 had been stored in a higher ionic strength buffer containing 25 mM Tris HCl pH 7.5, 50 mM NH₄Cl, 2 mM EDTA and 10% glycerol. Freshly prepped protein for the following studies was S200 exchanged in 100 mM NH₄OAc with 10% glycerol. Figure 3.7 shows a mass spectrum for 5 µM DnaT in 10 mM NH₄OAc pH 7.0 collected with the same instrument parameters as previously described (§3.3.1).
Figure 3.7: ESI-MS data for 5 uM DnaT in 10 mM NH₄OAc pH 7.0 showing decreased dimer formation. Peaks for monomer (M) and dimer (D) are labeled with corresponding charge state.

To investigate the effects of salt concentration on dimer formation for DnaT, NH₄Cl was added to the working solution to increase ionic strength while maintaining a volatile buffer. ESI-MS data was collected as previously described. Working solutions were prepared that contained 0, 5, 7.5 and 10 mM NH₄Cl in addition to the 10 mM NH₄OAc. The data for these experiments are shown in Figure 3.8.
Figure 3.8: ESI-MS data for 5 μM DnaT with NH₄Cl addition. Peaks for monomer (M) and dimer are labeled with corresponding charge state.

The data show that increasing the ionic strength to 15 mM with the addition of NH₄Cl does not result in the formation of the dimer as previously seen. However, there appears to be a slight increase in protein folding at 5 mM NH₄Cl. Increasing the salt concentration of the working solution above 15 mM using NH₄Cl results in a significant decrease in data quality and therefore should be avoided.

To further study the effects of salt concentration and protein preparation on DnaT, additional protein was prepared. Rather than buffer exchanging the protein into NH₄OAc prior to freezing, the protein was stored in 25 mM Tris HCl pH 7.5, 100 mM NaCl and 10% glycerol. Prior to data collection, the protein was buffer exchanged by filtration into 100 mM NH₄OAc. ESI-MS data was collected in the
same manner as previously described (§3.3.1). Various concentrations of DnaT from 5-40 µM were analyzed to investigate dimer formation. The data are summarized in Figure 3.9.

Figure 3.9: ESI-MS data for DnaT at concentration from 5-40 µM in 10 mM NH₄OAc pH 7.0. Peaks for monomer (M) and dimer are labeled with corresponding charge state.

The concentration profile for DnaT shows no dimer formation between 5 and 40 µM. Likewise, there is no change in the charge state distribution, indicating that there is no increase in the concentration of the folded monomer with increasing concentration. Based upon the data in Figures 3.8 and 3.9, it can be concluded that the dimer previously observed was not dependent on the concentration of salt in solution nor the preparation of the protein and is not reproducible. Therefore, it is
most likely that DnaT is a monomer within the concentration range used for these ESI-MS studies.

3.4 Conformational analysis of DnaT

3.4.1: Ion Mobility Data for DnaT

Ion mobility ESI-MS was used to determine the number of different conformations of DnaT at in 10 mM NH₄H₂CCOO buffers at pH 7.0 and 3.0. Ion mobility IMS gas was set to 20.0. The data for DnaT at pH 7.0 is shown in Figure 3.10a and 3.10b for pH 3.0.

Figure 3.10: Ion Mobility Spectra for DnaT in 10 mM NH₄OAc at pH 7.0 (a) and 3.0 (b)
Based upon the ion mobility data, it appears that three different conformations of DnaT are present at pH 7.0. The peaks corresponding in short drift times represent the unfolded conformation of DnaT. The intermediate peaks represent a partially folded species. The peaks at the longer drift times correspond to a compactly folded monomer. Interestingly, the distribution is mostly skewed to that of the fully folded and disordered conformations with only a minor contribution from the intermediate state. When comparing the data at pH 7.0 with that of 3.0, the data are very similar. However, there are slight differences in the drift times and intensities of the partially fully folded states. As would be expected, DnaT at pH 7.0 has a larger contribution from natively folded protein as can be seen by the peak at the long drift time for the peaks between approximately 2761.4 and 3228.1 m/z. The IMS data for DnaT are interesting in that the unfolded protein passes through the drift tube at shorter times than the compactly folded conformer. This is likely due to the high charge of the unfolded protein relative to the compact conformer.

3.4.2 Metal Binding Studies of DnaT

CD data indicate that the presence of Mg$^{2+}$ is critical to maintaining secondary structure for DnaT$^{36}$. As the initial ESI-MS studies show a large contribution due to disordered protein, the effects of the addition of Mg$^{2+}$ needed to be investigated. ESI-MS data were collected for 5 µM DnaT in the presence of varying concentrations of Mg$^{2+}$. The data in Figure 3.11 show the spectra obtained for DnaT with the addition of Mg ions. Figure 3.12 shows the +7 charge state for DnaT at varying concentrations of Mg$^{2+}$. 
Figure 3.11: ESI-MS data for 5 µM DnaT at varying Mg$^{2+}$ concentrations. Mg$^{2+}$ concentrations are normalized to 1 µM DnaT. Peaks for the monomer (M) and dimer (D) are labeled with corresponding charge state.
Figure 3.12: ESI-MS data showing the +7 charge state for DnaT in the presence of Mg\textsuperscript{2+} ions. Mg\textsuperscript{2+} concentrations are normalized to 1 µM DnaT. Peaks are labeled showing the addition of Mg\textsuperscript{2+}.

The data show that DnaT does bind Mg\textsuperscript{2+}. However, there is not a noticeable increase in the folding of the protein. The stoichiometry of binding is not clear from the data due to the possibility of nonspecific adduct formation. To better analyze the binding, the relative intensity of the +7 charge state peak vs Mg\textsuperscript{2+} was analyzed. The data are presented in Figure 3.14. The data show that up to 5 Mg\textsuperscript{2+} ions may be bound by the protein over the concentration range used in these experiments. However, the titration needed to be extended to higher concentrations as the data do not show saturation of binding.
3.4.3: Limited proteolysis studies

Since separation of the full-length and degradation product of DnaT was not possible, limited proteolysis was used to identify regions of protein disorder. As opposed to a complete protein digest, limited proteolysis studies involved the digestion of the protein on a much slower time scale so as to identify the regions of the protein that are most vulnerable to cleavage. By identifying highly flexible regions of the protein, the core domain of the protein can be identified. This technique is especially applicable to protein crystallography since truncated forms of the protein can be cloned which contain only the ordered, core region of the protein of interest.
DnaT was buffer exchanged into 100 mM NH₄HCO₃, pH 8.0. The stock protein solution was 200 µM with a final volume of 100 µL. Trypsin was added to achieve an enzyme to substrate ratio of 1:2000. This ratio was optimized based on previous experiments at 1:200, 1:500: 1:1000 and 1:5000 enzyme:substrate ratios. The reaction was carried out at room temperature. Samples were taken every minute for 60 minutes. In order to ensure that the reaction was quenched, the samples were added to a solution of 50% H₂O/50% acetonitrile, 0.1 % FA and diluted to 2 µM. The samples were kept on ice until data collection. All data were collected within a few hours to the digest to avoid any continued proteolysis of the protein.

The digest was monitored using ESI-MS using the same instrument as previously described. The instrument was calibrated externally using a cesium iodide standard in the mass range of 50 ≤ m/z ≤ 3000 m/z. Digest samples were directly injected into the instrument. The following parameters were maintained at a constant value for each data set: capillary voltage 3.5 kV, sample cone 40 kV; cone gas, 0 L/h. For data sets collected without fragmentation the trap collision energy was maintained at 6 V. To induce fragmentation of peptides, the collision energy was adjusted for each peptide of interest to values between 20 and 35 kV. For each data set 300 scans were averaged. An in silico digest using Protein Prospector was used to predict peptides.

Prior to proteolysis, a “zero hour” sample was taken to judge whether the protein was sufficiently desalted for data collection as well as to ensure that there was no presence of the degradation product previously observed. The “zero hour” spectrum is shown in Figure 3.13.
The first peptide was identified at 20 minutes at 958.5 m/z. The peptide was isolated and sequenced. The spectrum for the 20 minute sample is shown in Figure 3.14 and the sequencing data for the peptide in Figure 3.15. As can be seen by the data, the majority of the peaks correspond to the full-length protein and only a minor peak is observed for the peptide.
Figure 3.14: ESI-MS data for 2 μM limited proteolysis sample for DnaT after 20 minutes. The peptide identified at the 958.5 peak is labeled with an asterisk (*).
Figure 3.15: Fragmentation data for the 958.5 ion from 20 minute DnaT limited proteolysis. Precursor ion is labeled with an asterisk (*). Fragment ions are labeled accordingly.

After 60 minutes, the largest peak in the data corresponding to DnaT is that of the 958.5 peak first identified. The spectrum for the 60 minute sample can be seen in Figure 3.16. The 958.5 and 1036.6 peaks are indicated by asterisk (*).
Figure 3.16: ESI-MS data for 2 µM limited proteolysis sample of DnaT at 60 minutes. Peaks for peptides at 958.5 and 1036.6 are labeled with an asterisk (*)

Additionally after 60 minutes, two more peaks corresponding to peptides of DnaT were identified and subjected to sequencing. The fragmentation data for the peak at 1036.6 are showing in Figure 3.17.
Figure 3.17: ESI-MS fragmentation data for 1036.6 peak for DnaT limited proteolysis at 60 minutes. Precursor ion at 1036.6 is labeled with an asterisk (*). Fragment ions are labeled accordingly.

Table 3.2 shows a summary of the peptides from the DnaT proteolysis identified and their location within the sequence. The data shows the \( m/z \) for the peptides, the mass, sequence and the location within the sequence of DnaT. As can be seen for the Table, both peptides that were identified come from the C-terminal end of the protein and differ by only one residue. This is an indication that this portion of the protein is not compactly folded and therefore susceptible to proteolysis. Figure 3.19 shows the sequence for DnaT and the location of the protein found to be most susceptible to limited proteolysis. This portion of the protein is highlighted in bold.
Table 3.2: Summary of the peptides identified by limited proteolysis studies of DnaT

<table>
<thead>
<tr>
<th>m/z</th>
<th>Mass (Da)</th>
<th>Sequence</th>
<th>Residue numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>958.5</td>
<td>1916.0</td>
<td>DVNTVSEPDSQIPPGFRG</td>
<td>161-178</td>
</tr>
<tr>
<td>1036.6</td>
<td>2072.2</td>
<td>RDVNTVSEPDSQIPPGFRG</td>
<td>160-178</td>
</tr>
</tbody>
</table>

SSRVLTPDVGVIDALVHDHQTVLAKAEGGVVAVFANNAPAFYAVTPARLAE LLALEEKRLPGSDLALDDQLYQEPQAAVPVAVPMGKFAMYPDWQPDADFIR LAALWGVALREPVTTTEELASFIAYWQAEGKVFHHVQWQQKLARSQIQGRAS NGGLPKRDVNTVSEPDSQIPPGFRG

Figure 3.18: Amino acid sequence for DnaT showing region cleaved by trypsin in limited proteolysis studies. The region cleaved is highlighted in bold font.

3.4.4: In-source collision studies

Collision induced dissociation (CID) was used to determine more conclusively, the regions of structural disorder for DnaT. DnaT solutions of 5 µM protein were prepared in 10 mM NH₄OAc buffers at pH 7.0 and 3.0. The collision energy was gradually increased and peptide fragments were identified at the different collision energies. CID data for the peak were collected with collision energies of 6, 10, 15, 25, 35, and 40 kV. Figure 3.19 shows the CID experiments for the +7 charge state peak for DnaT at approximately 2761.4. This peak is believed to represent the most intense charge state of the most compact conformer of the protein.
Figure 3.19: Overlay spectra for CID experiments for the 2761.4 peak for 5 \( \mu \)M DnaT 10 mM NH\(_4\)OAc at pH 7.0 (a) and pH 3.0 (b). Precursor ion is marked with an asterisk (*)
Based upon the data presented in Figure 3.19, the native conformation of DnaT appears to be well ordered. No fragmentation can be seen for the +7 charge state peak regardless of pH. These data are consistent with the assumption that the peak corresponds to natively folded protein.

The same procedure was followed for the peak at ~1208.6 (+16 charge state). The peak is believed to represent the disordered population of the protein. The CID experimental data is shown in Figure 3.20 for pH 7 (a) and pH 3 (b). The software Fragpro was used to identify fragments that were cleaved. Based upon the data, once again, the first peptides removed result from the C-terminal end of the protein. The identified fragments are labeled in the spectra.
Figure 3.20: CID data for 5 µM DnaT in 10 mM NH₄OAc at varying collision energies for the 1208.6 peak at pH 7.0 (a) and 3.0 (b). Precursor ion is marked with an asterisk (*)
As would be expected, much more fragmentation is observed for the peak corresponding to disordered protein structure. The data for pH 3.0 and 7.0 show very similar fragmentation patterns. Using FragPro software and the sequence of DnaT, several ions were identified based on the m/z values and a ladder sequence of y-ions from fragmentation at the C-terminus of the protein. Based up the data in Figure 3.20, the fragments were determined to represent the C-terminus of the protein. This is consistent with the limited proteolysis data presented previously. Interestingly, no fragments were conclusively identified from the N-terminal end or core of the protein. This indicates that the protein maintains some ordered folding at the low m/z values.

3.5. Analysis of DnaT degradation product

3.5.1: Identification of degradation product

It was noticed that over time, DnaT solutions that were left at 4°C for several days accumulated a large population of degradation product. Deconvolution of the MS data revealed the mass difference between the full-length protein and the degradation product to be 940.1 Da. Figure 3.21 shows the mass spectrum for a DnaT sample containing degradation product and the deconvoluted mass for the full-length protein and degradation product.
In hopes of separating the full-length protein from the degradation product for sequencing, RPLC was used. The DnaT solutions containing degradation product were run on a C8 column. A gradient of 40%-50%acetonitrile with 0.1% FA of over 10 minutes followed by 50%-60% organic solvent over 50 minutes. A flow rate of 500 µL/ min was used. However, baseline resolution of the two protein forms could not be achieved. The RPLC chromatogram for DnaT can be seen in Figure 3.22. Protein was collected in 500 µL fractions for ESI-MS analysis.
As can be seen in Figure 3.22 baseline separation of the two protein forms could not be achieved. ESI-MS data for the fractions containing protein were collected. The fraction corresponding to the early peak contains primarily the full-length protein and the later fractions contain the degradation product. Fractions between the two peaks contain both protein forms as would be expected. Due to the inability to separate the two protein forms, direct sequencing of the degradation product was not possible.

3.5.2: Crystallization studies of DnaT

Crystals of DnaT had been previously obtained. However, the crystals did produce diffraction quality data and large 3D crystals were difficult to reproduce. Poor crystal diffraction can be attributed to several factors, most commonly poor
crystal packing due to protein structural disorder. Since it was observed that DnaT degrades when stored at 4°C, it is possible that the degradation product could cause the poor crystal diffraction.

In order to determine if the degradation product of DnaT was incorporated into the crystal lattice, crystals trays of DnaT using freshly prepared protein, which did not contain degradation product, were prepared. The same buffer conditions that had previously produced crystals were used. The protein solution consisted of 18 mg/mL protein in 25 mM HEPES pH 7.5, 50 mM NH₄Cl. Hanging drop vapor diffusion method was used. A 2 x12 expansion tray with buffer A containing 100 mM HEPES pH 7.5, 100 mM Li₂SO₄, 100 mM Na/K tartrate and buffer B containing 100 mM HEPES pH 8.3, 100 mM Li₂SO₄, 100 mM Na/K tartrate were set up with 2+2 drops. Two different crystal morphologies were observed within the same drops. The majority of the crystals obtained were plate-like and a small fraction was 3D hexagons. Pictures of the two crystal forms can be seen in Figure 3.23. After a week, the protein crystals were picked out of the drops, washed in the mother liquor to remove protein contaminants from the solutions and dissolved in water. To remove the salt components of the mother liquor, the protein solution was run through the RPLC using the same protocol as previously described.
Figure 3.23: Crystals of DnaT used in ESI-MS studies. Crystallization conditions were 2x12 gradient with buffer A containing 100 mM HEPES pH 7.5, 100 mM LiSO$_4$ and 100 mM Na/K Tartrate and buffer B containing 100 mM HEPES pH 8.3, 100 mM LiSO$_4$ and 100 mM Na/K Tartrate.

The RPLC fractions were run on the ESI-MS in order to determine if the degradation product was incorporated into the crystals. The data for the fractions is shown in Figure 3.24. As can be seen from the data, the crystals seem to contain a small fraction of the degradation product.
Since the protein sequence contains no cysteine residues that may participate in cross-linking, it can be assumed that the protein is being degraded either on the N- or C-termini. Using the sequence and known mass of the degradation product, it was determined that the protein is being degraded on the C-terminal end. This is also consistent with the limited proteolysis data that show that the protein is most disordered on the C-terminus. The identification and location of the cleaved peptide is summarized in Table 3.3.

Table 3.3: DnaT sequence and identification of the peptide cleaved in the degradation product.

<table>
<thead>
<tr>
<th>Mass (Da)</th>
<th>Sequence</th>
<th>Residue numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>958.0</td>
<td>SQIPPGFRG</td>
<td>170-178</td>
</tr>
</tbody>
</table>
3.5: Conclusions

Initial ESI-MS data for *E. coli* DnaT indicated that the protein existed as both a monomer and dimer. Subsequent analysis indicated that the dimer observed is neither reproducible nor dependent on salt concentration. DnaT is a monomer in solution. Ion mobility data show three folding states for the monomer: one largely disordered state, an intermediate, and a compactly folded state. Mg$^{2+}$ binding does not result in an increase in folding of the protein nor formation of the dimer. The protein was found to bind up to 5 Mg$^{2+}$ ions in the concentration range used in this study. CID data show that the C-terminal portion of the protein is susceptible to fragmentation and is therefore likely the disordered portion of the protein. Limited proteolysis data are consistent with the CID data. Furthermore, analysis of crystals of DnaT shows the incorporation of a degradation product that is a 9 amino acid truncation.

Future work for DnaT could include ion mobility data for the protein in solution with Mg$^{2+}$ in order to more conclusively determine if there is any change in the folding state of the protein. If there is a significant change in the folding state of DnaT upon Mg$^{2+}$ binding, crystallization of DnaT in the presence of Mg$^{2+}$ could result in diffraction quality crystals. Likewise, cloning of the C-terminal truncation and crystallization may improve crystal quality.
Chapter 4

PriA N-Terminai Domain and Complex Studies with DnaT

4.1 Introduction

As described in the introduction, PriA is the protein responsible for initiation of origin independent replication restart. PriA contains a N-terminal DNA binding domain and a C-terminal 3'→5' ATP-dependent helicase domain. The helicase domain functions to open a region of duplex DNA for DnaB binding. The two regions are separated by a flexible loop. The loop was found to be hypersensitive to tryptic cleavage upon DNA binding, indicating a large conformational change in the protein upon binding.

PriA has structure specific recognition of DNA rather than sequence specific, as is the case for DnaA. PriA binds to SSB coated ssDNA and initiates the formation of the replisome and subsequent loading of DnaB onto the DNA. The ability of PriA to bind to many different DNA substrates has been extensively investigated. Substrates that contain a three-strand junction such as a D-loop appear to be the preferred substrate for PriA. A region of ssDNA is critical to PriA binding as evidenced by its inability to binding DNA substrates containing entirely duplexed DNA, such a Mu fork and Holiday junctions. Figure 4.1 shows a summary of the DNA substrates recognized by PriA.
Figure 4.1: DNA substrates recognized by PriA

Two modes of PriA binding to DNA have been proposed. In one mode, PriA binds to substrates which contain a 3’-single-stranded extension\textsuperscript{18}. This theory was further supported by studies that showed phosphorylation of the 3’ OH group of ssDNA resulted in a substantial decrease in binding affinity\textsuperscript{37}. The second mode of PriA binding is believed to involve the recognition of the branched DNA structure found at junction such as a D-loop\textsuperscript{18}. While PriA could initiate the formation of the primosome on D-loop structures, it could not function on similar bubble DNA
structures, which contain no free 3’ OH group\textsuperscript{38}. Furthermore, PriA binding a 3’ end at the branch point of an arrested fork was required to stabilize the fork structure. This was believed to be due to the orientation of the helicase domain in a “unwinding deficient” orientation, therefore stabilizing the structure\textsuperscript{39}.

The structural basis for PriA recognition of the 3’ end of DNA was determined by X-ray crystallography. The first 181 residues of PriA were studied by NMR and were found to form a stable tertiary structure. Limited proteolysis by trypsin cleaved the fragment. Two smaller fragments containing residues 1-105 and 109-181 were isolated. NMR analysis of these two constructs also indicated a stable tertiary structure, leading to the conclusion that the N-terminal domain was composed of two independent domains.

The X-ray crystal structure of amino acids 1-105 (PDB id:2D7E) and the complexes with 2mer and 3mer (PDB id:2D7G and 2D7H) was solved in 2007 and is shown in Figure 4.2. While PriA is believed to function as a monomer at the primosome, the crystal structure indicates that PriA[1-105] forms a dimer in solution. The N-terminal 7 amino acids undergo domain swapping forming an intertwined dimer. This was further supported by gel filtration and analytical ultracentrifugation. SPR and NMR show that dimer formation does not interfere with binding of oligonucleotides\textsuperscript{37}. 
The crystal structure shown in Figure 4.2 reveals a unique amino acid-nucleotide interaction. PriA interacts with the base, the 3’ deoxyribose moiety and the 5’ phosphodiester group of the 3’ terminal residue. The 3’ hydroxyl group is bound through a series of polar interactions. The 3’ terminal binding pocket consists of Phe16, Asp17, Tyr18, Gly37, Leu55 and Lys61.

The DNA binding abilities of full-length PriA and PriA[1-181] were found to be much higher than that of PriA[1-105]. However, attempts to crystallize either construct were unsuccessful. This is believed to be due to the high degree of flexibility within the domains of the protein.

Previous work in the Mueser laboratory, in collaboration with Hiroshi Nakai at Georgetown University, also investigated the domain organization of PriA.
Limited proteolysis experiments identified amino acids 1-200 as the DNA binding domain. The properties of the PriA[1-200] construct are summarized in Table 4.1.

Table 4.1: Physical properties of PriA N-terminal domain

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecular Weight</td>
<td>22508.9 Da</td>
</tr>
<tr>
<td>Amino Acids</td>
<td>200</td>
</tr>
<tr>
<td>Extinction Coefficient</td>
<td>40575 M$^{-1}$cm$^{-1}$</td>
</tr>
<tr>
<td>pI</td>
<td>7.9</td>
</tr>
</tbody>
</table>

The PriA was found to form a complex with DnaT in the presence of DNA. Work by Deepa Singe found that at sufficiently high DnaT concentrations, the complex could be formed in the absence of DNA. This chapter focuses on analysis of the PriA N-terminal domain and complex formation with DnaT utilizing ESI-MS.

4.2 Protein Preparation

4.2.1 Protein expression

The PriA N-domain construct was received as a gift from Dr. Hiroshi Nakai. The gene was cloned into pET 21a and transformed into BL21(DE3) Gold cells for expression by previous group members. Protocol for large-scale preparation was followed as previously mentioned. The cultures were grown at 37° C in LB with 1 mM ampicillin. Protein expression was induced with 1 mM IPTG when the OD$_{600}$ was approximately 0.6. After 3 hours, the cells were harvested by centrifugation and frozen at -20° C until lysis. Protein expression as monitored by SDS-PAGE is shown in Figure 4.3.
4.2.2 Cell lysis

Cells were thawed and resuspended in lysis buffer containing 25 mM tris-HCl pH 7.5, 500 mM NaCl, 2 mM EDTA, 10 mM βME and 10% (w/v) sucrose. Lysozyme was added to assist lysis and the solution was stirred for 20 minutes. The lysate was then sonicated for 2 minutes and PEI added to a final concentration of 3%. The lysate was centrifuged at 18,000 g for 30 minutes. To the supernatant, (NH₄)₂SO₄ was added to precipitate the protein. The solution was again centrifuged at 18,000 g for an additional 15 minutes. The precipitated protein was resuspended in 25 mM tris-HCl pH 7.5, 50 mM NH₄Cl, 2 mM EDTA and 10 mM βME. Glycerol was added to a final concentration of 10% and the protein frozen at -80°C until purification. Figure 4.4 shows the monitoring of PriA N lysis by SDS-PAGE.
4.2.3 HPLC Purification

The first round of purification for PriA N was a low-resolution cation exchange chromatography step using SP Sepharose. The lysate was diluted to a conductivity of approximately 5.0 mS by the addition of 20 mM bis-tris pH 6.5. As was the case for DnaT, a linear salt gradient was used to elute the protein. The buffers used for ion exchange chromatography are summarized in Table 4.2. The SP Sepharose chromatogram for PriA N can be seen in Figure 4.5.
Table 4.2: Ion-exchange purification buffers for PriA N

<table>
<thead>
<tr>
<th>Column</th>
<th>Buffer A</th>
<th>Buffer B</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP Sepharose</td>
<td>25 mM Bis-tris pH 6.5, 50 mM NH₄Cl, 5 mM βME</td>
<td>25 mM Bis-tris pH 6.5, 50 mM NH₄Cl, 1.5 M NaCl, 5 µM βME</td>
</tr>
<tr>
<td>HA</td>
<td>25 mM Tris-HCl pH 7.5, 50 mM NaCl</td>
<td>25 mM Tris-HCl pH 7.5, 50 mM NaCl, 1.5 M (NH₄)SO₄</td>
</tr>
<tr>
<td>Poros HS</td>
<td>25 mM Bis-tris pH 6.5, 50 mM NH₄Cl</td>
<td>25 mM Bis-tris pH 6.5, 50 mM NH₄Cl, 1.5 M NaCl</td>
</tr>
</tbody>
</table>

Figure 4.5: SP Sepharose chromatogram for PriA N. Protein was eluted with a linear gradient with the buffers in Table 4.2. The protein eluted in a broad range between fractions 8-30.

As can be seen by the chromatogram for the SP Sepharose run, the $A_{280}$ for PriA N was very high. This was judged to be due to the DNA binding ability of the
protein. In order to remove DNA, the protein was purified using a hydroxyapatite. Again, a linear gradient was used to elute the protein. The chromatogram for the HA purification is shown in Figure 4.6. Following HA purification, the protein was dialyzed against buffer A for ion exchange chromatography to reduce the salt concentration for Poros HS purification.

![Figure 4.6: HA chromatogram for PriA N.](image)

Poros HS was then used for further purification of the protein. The buffers listed in Table 4.2 were used in a linear gradient. The protein eluted in a very narrow peak. The chromatogram is shown in Figure 4.7.
Figure 4.7: Poros HS chromatogram for PriA N. Protein was eluted with a linear gradient with the buffers in Table 4.2. The protein was found in fractions 11-23.

The final purification step for PriA N was purification by Superdex 75. This column also served as the buffer exchange step in preparation for ESI-MS studies. The protein injected in 500 µL aliquots and eluted at a constant gradient and 100 mM NH$_4$OAc. The S75 chromatogram and for PriA N are shown in Figure 4.8.
Figure 4.8: S75 chromatogram for PriA N domain in 100 mM NH₄OAc. PriA N was found in fractions 9-18.

4.3: ESI-MS Studies of PriA N-Domain

4.3.1: Native Conformation Studies

PriA N that had been buffer exchanged into 100 mM NH₄OAc was concentrated to 250 μM as previously described for DnaT. Two different concentrations of PriA N domain were used to analyze the oligomerization state of the protein. Working solutions of 5 μM in 10 mM NH₄OAc buffers of varying pH. ESI-MS data was collected on a Synapt HDMS Q-TOF mass spectrometer (Waters) in positive ion mode. The flow rate and capillary voltage were adjusted to obtain the best data possible. The following parameters were kept constant for all data sets:
cone voltage, 40 kV; cone gas 0 L/hr and collision energy 6 kV. The data shown are an average of 300 scans.

Figure 4.9: ESI-MS data for 5 µM PriA N domain in 10 mM NH₄OAc buffers at varying pH values. Peaks for the monomer (M) and dimer (D) are labeled with the corresponding charge state.

The data indicate that PriA is monomer. The pH profile shows that PriA N-terminal domain is fairly stable even at low pH values. However, obtaining good data at high pH levels was difficult. This could be due to poor solubility of the protein under these conditions.

4.3.2: Complex studies with DnaT

DnaT and PriA can form a complex under conditions that required the bypass of PriB action at the primosome. Previous work in the Mueser laboratory by Deepa Singh using native gels indicates that DnaT interacts with the PriA N-terminal
domain in the complex. The formation of the complex was studied by ESI-MS. The same instrument parameters as used in §4.3.1 were used for these experiments. Stock protein solutions of 100 µM protein in 120 mM NH₄OAc for both DnaT and PriA N were diluted into working solutions of 6 µM of each protein in 10 mM NH₄OAc pH 7.0. The solution was allowed to equilibrate for 10 minutes prior to data collection.

Figure 4. 10: ESI-MS spectrum for 6 µM PriA N domain complex with 6µM DnaT in 10 mM NH₄OAc pH 7.0. Peaks for DnaT monomer(*), and dimer (*) and PriA N-terminal domain (') are indicated on the spectra.

The data for the PriA N terminal domain in solution with DnaT show no evidence of complex formation. Peaks for both DnaT and PriA can be seen in the spectra but no peaks are observed for a complex between the two.
4.4 Conclusions

The ESI-MS data show that PriA N terminal domain exists as a monomer in solution. This is consistent with the stoichiometry believed to function at the primosome even though the crystal structure of residues 1-105 show dimer formation. Studies on the complex of PriA N terminal domain with DnaT were inconclusive. However, the data do show an increase in dimer formation for DnaT. Future work could investigate using varying concentrations of the proteins and whether DNA binding is required for dimer formation in the larger PriA N-domain.
Chapter 5:

Volatile Buffer Screen for ESI-MS

5.1 Introduction

As mentioned in Chapter 2, ESI-MS requires a volatile buffer, which can be removed upon desolvation of the analyte. Removal of buffer ions is critical to data quality. Buffer ions or contaminants which are not removed during desolvation can form adducts with the protein molecule, decreasing resolution and mass accuracy. There are several known buffers which are amenable to use in ESI-MS, the most commonly used buffers are ammonium acetate (NH₄OAc), ammonium bicarbonate (NH₄HCO₃) and triethylammonium bicarbonate (TEAB). Selection of the buffer used in an ESI-MS experiment is often made based upon the buffer system in which the protein molecules of interest are soluble.

In the case of metal binding experiments, it has been shown that the choice of the anion component of the salt is critical to data quality as well. Typically, chloride or acetate salts are used for metal binding studies. Recently, it was demonstrated that the use of a chelating anion such as tartrate can reduce non-specific metal adducts. This is based upon the theory that chelating salt ions can sequester excess metal ions.
during solvent evaporation, thus preventing adduct formation with carboxylate groups on the protein surface. However, the use of impure salts regardless of anion composition will undoubtedly reduce data quality.

Proteins that are prepared in the laboratory are also subjected to the addition of additives such as glycerol, sucrose and glucose in order to increase protein solubility as well as to serve as a protectant upon freezing at -80° C. These molecules may form adducts with the analyte proteins, changing the mass accuracy or reduce ionization efficiency.

The goal of this chapter was to develop a buffer screen to identify buffer systems that produce high quality data in the instrument available for in-house use. Proteins with well-studied MS behavior were purchased in order to eliminate the time required to prepare proteins. Salts of metals commonly used in metal binding studies were screened to identify salts that not only reduce non-specific adduct formation but also do not hinder data quality. The additives, glycerol, sucrose and glucose were also investigated to study the effect of their presence on data quality and ionization efficiency.

**5.2: Volatile Buffer Screen**

Table 5.1 summarizes the proteins used in this study. Transferrin, myoglobin and bovine serum albumin were purchased from Sigma. Protein stock solutions were prepared by dissolving the proteins in HPLC grade water to a final concentration of 200 µM. Buffers were prepared at 10 mM concentrations and the pH values adjusted to ~7.0. Table 5.2 shows the buffers used as well as the acid/base used to adjust the pH.
Table 5.1: Test Proteins

<table>
<thead>
<tr>
<th>Protein</th>
<th>Organism</th>
<th>Form</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myoglobin</td>
<td>Horse Heart</td>
<td>Powder</td>
<td>USB</td>
</tr>
<tr>
<td>Holo-transferrin</td>
<td>Human</td>
<td>Powder</td>
<td>USB</td>
</tr>
<tr>
<td>Serum Albumin</td>
<td>Bovine</td>
<td>Powder</td>
<td>USB</td>
</tr>
</tbody>
</table>

Table 5.2: Test Buffers

<table>
<thead>
<tr>
<th>Buffer</th>
<th>Acid/Base</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH₄H₃CCOO</td>
<td>Acetic acid</td>
</tr>
<tr>
<td>NH₄HCO₃</td>
<td>Acetic acid</td>
</tr>
<tr>
<td>TEAB</td>
<td>Acetic acid</td>
</tr>
<tr>
<td>NH₄Citrate</td>
<td>NH₄OH</td>
</tr>
<tr>
<td>NH₄Oxalate</td>
<td>NH₄OH</td>
</tr>
<tr>
<td>NH₄Tartrate</td>
<td>NH₄OH</td>
</tr>
</tbody>
</table>

As the mass range is optimized to the protein of interest, different parameter files were used for data collection. For myoglobin (~16 kDa), a mass range of 50-5000 m/z was used. Transferrin (~75 kDa) and BSA (~65 kDa) were analyzed in a mass range of 500-8000 m/z. Working solutions of 3 µM myoglobin, 5 µM transferrin and 10 µM BSA were used for all experiments. The data collected for myoglobin are shown in Figure 5.1, transferrin in 5.2 and BSA in 5.3. For all ESI-MS experiments described below, data sets were collected on a Synapt HDMS Q-
TOF mass spectrometer in positive ion mode. Ions were generated using a nanospray source with homemade fused silica capillaries. The capillary voltage and flow rates were optimized for each experiment. The cone voltage and collision energy were kept constant at 40 kV and 6 kV respectively. The cone gas was maintained at 0 L/hr. Data sets are an average of 300 scans.

Figure 5.1: ESI-MS buffer screen results for 3 µM myoglobin
Figure 5.2: ESI-MS buffer screen results for 5 µM transferrin

Figure 5.3: ESI-MS buffer screen results for 10 µM BSA
The data for all three proteins is very consistent. As was expected, \( \text{NH}_4\text{OAc} \) and \( \text{NH}_4\text{HCO}_3 \) buffer systems provided the highest quality data. For TEAB, data was obtained but the ionization efficiency of the proteins was very low and higher concentrations of protein, double those of the other buffers, in order to see protein peaks. TEAB also produced higher \( m/z \) values indicating fewer charges on the proteins. Data could not be obtained for the citrate, oxalate and tartrate buffers regardless if a higher protein concentration was used. Based upon the results of these studies, the best buffer systems to use are \( \text{NH}_4\text{OAc} \) and \( \text{NH}_4\text{HCO}_3 \). TEAB may also be used but higher protein concentrations are necessary to produce data.

### 5.3 Salt Studies

Calcium tartrate was purchased from Sigma. The MgTartrate solution was prepared as described by Konermann *et. al.* by mixing equal molar amounts of \( \text{NH}_4\text{Tartrate} \) and \( \text{Mg(H}_3\text{CCOO)}_2 \). The same protein conditions were run as previously mentioned. Data were collected on a Synapt HDMS q-TOF mass spectrometer in positive ion mode. The capillary voltage and flow rate were optimized to produce the best data. The following instrument parameters were kept constant: collision energy, 6 kV; The data obtained for myoglobin with the \( \text{Mg}^{2+} \) and \( \text{Ca}^{2+} \) salts for each protein are shown in Figures 5.4-5.9.
Figure 5.4: ESI-MS results for 3 μM myoglobin in 10 mM NH₄OAc pH 7.0 with 50 mM Mg²⁺ salts

Figure 5.5: ESI-MS results for 3 μM myoglobin in 10 mM NH₄OAc pH 7.0 with 50 mM Ca²⁺ salts
Figure 5.6: ESI-MS results for 5 µM transferrin in 10 mM NH₄OAc pH 7.0 with 50 mM Mg²⁺ salts
Figure 5.7 ESI-MS results for 5 µM transferrin in 10 mM NH$_4$OAc pH 7.0 with 50 mM Ca$^{2+}$ salts

Figure 5.8: ESI-MS results for 10 µM BSA in 10 mM NH$_4$OAc pH 7.0 with 50 mM Mg$^{2+}$ salts
The data for the Mg salts show that the best results are obtained with chloride and acetate salts. The Ca studies also showed that chloride and acetate salts produced the best data. However, the data for calcium tartrate was much better than that for magnesium tartrate.

Recent studies have shown that using a chelating anion such as tartrate can reduce nonspecific adduct formation. Myoglobin and transferrin have no magnesium or calcium ion binding ability. However, BSA is known to bind calcium ions. Figures 5.10-5.15 show zoom views for the largest peak in each spectrum corresponding to protein for the Mg$^{2+}$ and Ca$^{2+}$ salts.
Figure 5.10: ESI-MS data for the +7 charge state for myoglobin with the 50 mM Mg$^{2+}$ salts investigated.

Figure 5.10: ESI-MS data for the +7 charge state for 3 µM myoglobin with the 50 mM Ca$^{2+}$ salts investigated.
Figure 5.12: ESI-MS data for the +19 charge state of 5 µM transferrin with 50 mM Mg$^{2+}$ salts

Figure 5.12: ESI-MS data for the +19 charge state of 5 µM transferrin with 50 mM Ca$^{2+}$ salts
Figure 5.13: ESI-MS data for the +16 charge state of 10 μM BSA with 50 mM Mg$^{2+}$ salts

Figure 5.14: ESI-MS data for the +16 charge state of 10 μM BSA with 50 mM Ca$^{2+}$ salts
The data for the salt screen are very inconsistent in terms of nonspecific adduct formation. In some cases, tartrate does seem to result in few nonspecific adduct formation and in others produces the most nonspecific adducts. Based upon these results, acetate salts provide the best combination of data quality and nonspecific adduct formation.

5.4 Additive Screen

5.4.1 Methods

The affects of the addition of glucose, sucrose and glycerol on data quality were investigated. For each additive 100x, 10x and 1x molar ratios of additive: protein were studied. Again, the same protein concentrations used in the buffer screen were used for each protein and the working solutions prepared in 10 mM NH₄H₃CCOO pH 7.0.

5.4.2: Effects of glucose on data quality

Working solutions were prepared as described above. ESI-MS data was collected in the same manner as §5.2. Figures 5.16-5.21 show the data for glucose as the additive for each protein. Close-up views of the largest charge state population for each protein are also shown to show adduct formation.
Figure 5.15: ESI-MS data for 3 μM myoglobin in 10 mM NH₄OAc pH 7.0 with varying ratios of glucose

Figure 5.16: ESI-MS data for the +7 charge state of myoglobin in 10 mM NH₄OAc with varying ratios of glucose
Figure 5.17: ESI-MS data for 5 µM transferrin in 10 mM NH₄OAc pH 7.0 with varying ratios of glucose

Figure 5.18: ESI-MS data for the +19 charge state of transferrin in 10 mM NH₄OAc with varying ratios of glucose
Figure 5.19: ESI-MS data for 10 µM BSA in 10 mM NH₄OAc pH 7.0 with varying ratios of glucose

Figure 5.20: ESI-MS data for the +16 charge state of BSA in 10 mM NH₄OAc with varying ratios of glucose
The data for the various proteins in solutions that contain glucose as an additive show that high concentrations of glucose cause adduct formation. While the data for transferrin are inconsistent in terms of data quality with increasing glucose concentration, this is more likely due to the use of a different needle or the inability to find the correct instrument parameters during data acquisition. The effect of glucose on data quality is best demonstrated in the case of BSA as the +16 charge state peak shows severe broadening on the right side. Given that BSA serves as a carrier protein for small molecules, it is not surprising that it is more affected by the additive than myoglobin and transferrin.

5.4.3: Effects of sucrose on data quality

The effects of sucrose as an additive in protein solutions was investigated. Working solutions for each protein with sucrose were prepared as previously described in §5.4.1. Figures 5.22-5.27 summarize the data obtained from these experiments.
Figure 5.21: ESI-MS data for 3 μM myoglobin in 10 mM NH₄OAc pH 7.0 with varying ratios of sucrose

Figure 5.22: ESI-MS data for the +7 charge state of myoglobin in 10 mM NH₄OAc with varying ratios of sucrose
Figure 5.23: ESI-MS data for 5 µM transferrin in 10 mM NH₄OAc pH 7.0 with varying ratios of sucrose

Figure 5.24: ESI-MS data for the +19 charge state of transferrin in 10 mM NH₄OAc with varying ratios of sucrose
Figure 5.25: ESI-MS data for 10 μM BSA in 10 mM NH₄OAc pH 7.0 with varying ratios of sucrose

Figure 5.26: ESI-MS data for the +16 charge state of BSA in 10 mM NH₄OAc with varying ratios of sucrose
Sucrose as an additive in the solution results in not only adduct formation as can be seen in the data for transferrin (Figure 5.24) and BSA Figure (5.26) but also significantly reduces ionization efficiency with increasing concentration. This trend is found for all three proteins.

5.4.5: Effects of glycerol on data quality

The effects of glycerol on data quality was investigated in the same manner as the preceding sections. Figures 5.27-5.32 summarize the data obtained for each protein as well as views of the largest charge state for each protein to examine adduct formation as a function of additive concentration.

Figure 5.27: ESI-MS data for 3 µM myoglobin in 10 mM NH₄OAc pH 7.0 with varying ratios of glycerol
Figure 5.28: ESI-MS data for the +7 charge state of myoglobin in 10 mM NH₄OAc with varying ratios of glycerol

Figure 5.29: ESI-MS data for 5 nM transferrin in 10 mM NH₄OAc pH 7.0 with varying ratios of glycerol
Figure 5.30: ESI-MS data for the +19 charge state of transferrin in 10 mM NH$_4$OAc with varying ratios of glycerol

Figure 5.31: ESI-MS data for 5 µM BSA in 10 mM NH$_4$OAc pH 7.0 with varying ratios of glycerol
Figure 5.32: ESI-MS data for the +16 charge state of BSA in 10 mM NH₄OAc with varying ratios of glycerol

The mass spectra for glycerol show that proteins are most affected by its presence in terms of adduct formation. However, glycerol does not have a significant effect on ionization efficiency, as did sucrose.

5.5: Conclusions

Based upon the results from the buffer screen, NH₄H₃CCOO and NH₄HCO₃ produce the highest quality data. TEAB shows a very low ionization efficiency and produces ions at much higher m/z values than the other buffers. Citrate, oxalate and tartrate buffers should not be used. Data from the salt binding studies show that acetate produces the highest quality data while also prevent nonspecific adduct formation. The data for the additive screen show that higher concentrations of the
additives results in decrease ionization efficiency of the proteins, as would be expected. The addition of sucrose results in the most significant decrease in ionization efficiency. Increased additive concentration also results in severe adduct formation. The presence of glycerol has the most dramatic effect on adduct formation but both glucose and sucrose also resulted in adduct formation. The formation of adducts significantly reduces the mass accuracy and resolution of the data. Therefore, extensive buffer exchange to remove additives used for cryoprotection is required to obtain the highest quality data possible.
References


35. Izaac, A. Crystallographic studies of the *E. coli* DNA replication restart primosome. University of Toledo, 2005.


