Covalent Modification that Enhances Protein Properties and Functions and Folds Random Coil into Alpha-Helix

Deepali Prashar
Syracuse University

Follow this and additional works at: http://surface.syr.edu/che_etd

Part of the Chemistry Commons

Recommended Citation

This Dissertation is brought to you for free and open access by the College of Arts and Sciences at SURFACE. It has been accepted for inclusion in Chemistry - Dissertations by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.
ABSTRACT

Understanding the fundamental science of aggregation and surface adsorption of proteins is critical for developing protein-based drugs, understanding aggregation-induced diseases and for preventing biofouling that impacts a wide range of industries. In spite of the advances in sensitive and high-throughput analytical methods, the commercialization of protein drugs has been impeded by poor solubility, inevitable aggregation, and difficulty in purification. The primary aim of this dissertation is to modify proteins covalently to induce folding in random coil sequences, prevent aggregation, surface adsorption and for increasing the thermal stability against aggregation of proteins.

The first chapter describes the chemistry for modifying the lysine residues in proteins with alditols (organic kosmotropes), via an amide bond using the traditional N-hydroxysuccinimide (NHS) activated carboxylic acids or via an efficient water-driven ligation generating a squaramide linkage. Conjugating such organic kosmotropes to proteins via a squaramide linkage, results in the formation of a hydrolytically stable bond together with higher amount of modification when compared to amide bond formation using NHS-activated carboxylic acids.

The second chapter presents a general approach for preventing both protein aggregation and surface adsorption by modifying proteins with β-cyclodextrins (βCD) via a squaramide linkage. As compared to native unmodified proteins, the β-cyclodextrin-modified proteins (lysozyme and RNase A) exhibited significant reduction in aggregation, surface adsorption and increase in thermal stability against aggregation.

The third chapter describes the covalent modification of a random coil mutant (L44A) of the coiled coil domain of cartilage oligomeric matrix protein (COMPcc) with squarate derivatives. The covalent modification with βCD-amino squarates or methyl squarate alone folded the random coil into an α-helix. The COMPcc is capable of binding small molecule such as
curcumin, retinol and vitamin D. Curcumin exhibited an increase in the fluorescence signal upon binding to L44A-(βCD)<sub>n</sub> as compared to the wild type protein suggesting that the folding stability of L44A is either enhanced or preserved.

The fourth chapter describes the synthesis and development of an adamantane-platinum conjugate (analog of carboplatin) hosted in β-cyclodextrin. This inclusion complex formed by βCD and the adamantane-platinum conjugate exhibited higher cytotoxicity towards neuroblastoma (SK-N-SH) cells than carboplatin. This approach provides versatility for enhancing drug properties via structural modification and targeted delivery using protein-drug conjugates.
Covalent Modification that Enhances Protein Properties and Functions and Folds Random Coil into Alpha-Helix

by

Deepali Prashar

M.Phil. (Chemistry), Syracuse University, 2009
M.Sc. (Chemistry), University of Delhi, New Delhi, India, 2005
B.Sc. (Chemistry), University of Delhi, New Delhi, India, 2003

Dissertation
Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemistry.

Syracuse University
August 2012
ACKNOWLEDGEMENT

This dissertation would not have been possible without the guidance and support of several individuals, who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

I would like to take the opportunity to offer my sincere thanks to my advisor Professor Yan-Yeung Luk for his constant encouragement and support throughout my PhD studies. Professor Luk’s enthusiasm for science and intense drive for developing new ideas has been a constant source of inspiration through my five years in his laboratory. I have never felt the slightest hesitation about approaching Professor Luk for discussing science or even personal matters. My time in his laboratory has been a great learning experience for me both academically and professionally.

I would also like to thank all the members of my defense committee, Professors Raina, Dabrowiak, Braiman, Borer and Maye for finding time out of their busy schedule to be on my defense committee. I am grateful to them for critically reviewing of my PhD dissertation. In particular, I am thankful to Professors Borer, Maye, Dabrowiak and Korendovych for allowing me to use various instruments and equipments in their laboratories.

I would like to thank members of the Luk group Sijie, Nisha, Gauri and Nischal for being great friends and colleagues. Their company in the laboratory and even outside the laboratory enabled the smooth progression of research during my graduate study. A special thanks to my friend and colleague Yi Shi for all the intellectual discussions and collaboration. I have truly enjoyed working and spending time with her.

I also thank the Chemistry department staff members, Jodi, Cathy, Deb, Linda and Joyce who have contributed in many ways to my wonderful time in graduate school.
I would like to offer my heartfelt gratitude to my parents Mr. Raj Kumar Prashar and Mrs. Rita Prashar for their unconditional love and constant encouragement through my PhD studies. Both my parents with a very heavy heart supported my decision to step out to another country for pursuing higher studies. I am completely indebted to my parents for all the sacrifices they have made for me. I would also like to thank my brother Anshul Prashar and my sister-in-law Neelmani Bharadwaj for their support and being pillars of strength during my time in the United States. I would like to thank my grandparents Mr. Manohar Lal Prashar and Mrs. Kamla Wati Prashar for being an inspiration in my life. Sincere thanks to my father-in-law and mother-in-law Late Krishna Kumar Bandyopadhayay and Debapriya Bandyopadhyay for their prayers and good wishes. Special thanks to my sister-in-law Dr Panchali Lal for being more of a friend to me. Her constant encouragement has played a major role during my PhD studies. I am also grateful to my nephews Ojaswi and Pramit and my niece Neelanshi for bringing a smile on my face even when the times were difficult. Most importantly, I would like share the credit of my work with my husband Debjyoti without whose love, support, encouragement and guidance the completion of my PhD would not have been possible. I have loved every minute I have spent with him discussing chemistry in the laboratory or even at home. Words alone cannot express what I owe him for his encouragement and patient love, which has enabled me to reach where I am today.

Lastly, I thank the omnipresent God, for answering my prayers and for giving me the strength to reach this day.

Deepali Prashar
# TABLE OF CONTENTS

## Chapter 1

A Kosmotrope Approach for Prevention of Protein Aggregation  

Summary  

1.1 Background and Significance  
   1.1.1 Proteins as pharmaceutical drugs  
   1.1.2 The kosmotrope theory – a new basis for preventing protein aggregation  
   1.1.3 Drawbacks of protein pegylation  
   1.1.4 Protein adsorption and denaturation  
   1.1.5 Strategies and theories for preventing protein aggregation  

1.2 Results and Discussion  
   1.2.1 Synthesis of alditol-terminated carboxylic acids  
   1.2.2 ESI-MS confirms lysozyme modification with D-gulitol  
   1.2.3 Squaramide is a stable linkage compared to amide bond  
   1.2.4 Purification of alditol modified Lysozyme  
   1.2.5 Cyclodextrins- another class of organic kosmotropes  

1.3 Conclusions and Perspectives  

1.4 Experimental section  

## Chapter 2

Covalent Grafting of Proteins with Cyclodextrins Prevents Aggregation and Surface Adsorption and Increases Thermal Stability  

Summary  

2.1 Background and Significance  
   2.1.1 Cyclodextrins can prevent protein aggregation  
   2.1.2 Cyclodextrins versus poly(ethylene glycol) for increasing protein stability  
   2.1.3 Different methods of protein glycosylation  

2.2 Results and Discussion  
   2.2.1 Synthesis of βCD-derivatized amino squarate reagents  
   2.2.2 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and MALDI-TOF confirms the degrees of modification in lysozyme with βCD  
   2.2.3 Modification of lysozyme using NHS/EDC coupling  
   2.2.4 Prevention of protein aggregation  
   2.2.5 Dialysis in water aggregates only unmodified proteins  
   2.2.6 Dynamic light scattering on lysozyme-(βCD)<sub>n</sub> and aged lysozyme for protein size determination  
   2.2.7 Thermo-induced aggregation of unmodified lysozyme and lysozyme-(βCD)<sub>n</sub> using dynamic light scattering (DLS).  
   2.2.8 Surface adsorption of lysozyme-(βCD)<sub>n</sub> and native lysozyme using surface plasmon resonance.  
   2.2.9 Measuring activity of lysozyme-(βCD)<sub>n</sub> and native lysozyme using
turbidimetric method
2.2.10 Purification of lysozyme-(βCD)$_n$ by protein crystallization 72
2.2.11 Modification of Bovine Serum Albumin (BSA) with βCD 73
2.2.12 Covalent modification of aniline grafted proteins with amino
squarate based reagents
2.3 Conclusions and Perspectives 83
2.4 Experimental Section 83

Chapter 3

Inducing Folding in an Inherently Unstructured Protein by Covalent Modification
with Squarate Derivatives 104

Summary 104
3.1 Background and Significance 104
  3.1.1 Protein folding 104
  3.1.2 Inducing peptide and protein folding 105
  3.1.3 Cartilage oligomeric matrix protein (COMP) - function and uses 106
  3.1.4 Studying folding of an inherently unfolded protein by covalent
       modification with β-cyclodextrins 107
  3.1.5 Protein energy profile 108
3.2 Results and Discussion 109
  3.2.1 MALDI-TOF confirms modification of L44A with βCD 109
  3.2.2 Circular dichroism indicates folding of L44A-(βCD)$_n$ 111
  3.2.3 Dynamic light scattering to measure size of L44 wt, L44A and
       L44-(βCD)$_n$ 114
  3.2.4 Influence of groups like acetyl, polyethylene glycol (PEG) and
       methyl squarate on folding of L44A 115
  3.2.5 Small molecule binding in L44A-(βCD)$_n$ 121
  3.2.6 Breaking the assembly formation in L44A-(βCD)$_n$ 124
  3.2.7 Purification of L44-(βCD)$_n$ using reverse phase HPLC 126
  3.2.8 Effect on thermal stability of lysozyme by modification with compounds
       2 and 3 129
  3.2.9 Generality of our system 131
3.3 Conclusions and Perspectives 131
3.4 Experimental Section 133

Chapter 4

Synthesis of an Adamantane-Platinum Conjugate Hosted in β-cyclodextrin for
Enhanced Activity and Targeted Drug Delivery 141

Summary 141
4.1 Background and Significance 142
  4.1.1 Why Beta-Cyclodextrin (βCD)? 142
  4.1.2 Platinum compounds as anticancer agents 143
4.1.3 β-cyclodextrin as drug carrier

4.2 Results and Discussion

4.2.1 Synthesis of the βCD adamantane-platinum host-guest complex (6)

4.2.2 Studying cytotoxicity of host-guest complex (6) and carboplatin using colorimetric assay

4.2.3 Cytotoxicity of host-guest complex (6) and carboplatin increases with increasing incubation times with SK-N-SH cells

4.2.4 Cyclodextrin enters cells through pinocytosis

4.2.5 SDS-PAGE confirms host-guest complex (6) binds DNA more than carboplatin

4.2.6 βCD alone does not bind DNA

4.2.7 Platinum uptake studies using ICP-MS

4.2.8 Achieving targeted drug delivery

4.3 Conclusions and Perspectives

4.4 Experimental Section

References
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>Chapter 1</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>The ions (cations or anions) towards the right are called chaotropes and act to salt proteins into solution, whereas ions to the left are called kosmotropes and act to salt proteins out of solution.</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>Schematic illustrations for (A) unmodified proteins aggregating in solution, (B) Proteins modified with kosmotropes in solution prevents aggregation by templating a structured interfacial water layer around each protein.</td>
</tr>
<tr>
<td>Figure 1.3</td>
<td>ESI-MS spectra of lysozyme-(D-gulitol)$_n$, (n can range from 1 to 6) indicating that up to two lysine residues in lysozyme were modified with D-gulitol. Peak for the unmodified lysozyme and dimer of unmodified lysozyme was also observed.</td>
</tr>
<tr>
<td>Figure 1.4</td>
<td>Gel image from SDS-PAGE for lysozyme modified with D-gulitol. Unmodified lysozyme (Lane 1), sample containing mixture of unmodified lysozyme and lysozyme-(D-gulitol)$_n$ (Lanes 2 and 3). No separate bands corresponding to different modifications of lysozyme with D-gulitol were observed.</td>
</tr>
<tr>
<td>Figure 1.5</td>
<td>Hydrogen bonding ability and high dielectric constant of water increases the stability of zwitterionic resonance structure and electrophilicity of squarate in water versus organic solvent.</td>
</tr>
<tr>
<td>Figure 1.6</td>
<td>ESI-MS spectra for lysozyme-(N-allyl squaramide)$_n$, indicating that up to three lysine residues are modified with N-allyl squaramide with a relatively small mass peak for unmodified lysozyme.</td>
</tr>
<tr>
<td>Figure 1.7</td>
<td>MALDI-TOF spectra for (A) lysozyme modified with compound 16 indicating that up to two lysine residues in lysozyme are modified with D-mannitol squaramide and a relatively small mass peak for unmodified lysozyme (B) lysozyme modified with compound 6 using NHS/EDC chemistry showing that only one lysine residue is modified with D-gulitol.</td>
</tr>
<tr>
<td>Figure 1.8</td>
<td>Strategy for purifying proteins modified with alditols using boronic acid-laden porous hydrogels.</td>
</tr>
</tbody>
</table>
Figure 2.1  A schematic representation demonstrating the ability of the multiple hydroxyl groups of β-cyclodextrin to generate a structured solvation shell by hydrogen bonding with surrounding water molecules.

Figure 2.2  SDS-PAGE showing separate bands corresponding to one and two modifications of lysozyme (0.05 mg/mL) with βCD using compounds 6, 7 and 8. Lane A is lysozyme; lane B, C and D are for lysozyme modified with 8, 7 and 6 respectively. Inset above shows SDS-PAGE with high concentration of modified lysozyme (0.5 mg/mL) showing separate bands corresponding to one, two and three modifications with βCD.

Figure 2.3  MALDI-TOF spectra for lysozyme-(βCD)$_n$ (A) before and (B) after dialysis in deionized water.

Figure 2.4  (A) MALDI-TOF spectra for RNase A-(βCD)$_n$ (B) MALDI-TOF spectra for RNase A-(βCD)$_n$ after dialysis in water.

Figure 2.5  Number (A and B) and volume (C and D) distribution of DLS showing two replicates each for 1 mg/mL of lysozyme-(βCD)$_n$ and lysozyme in PBS buffer pH 7.4 at 25 °C. Volume distribution showed only monomeric proteins for lysozyme A-(βCD)$_n$ (A) and only aggregates for aged lysozyme (B). Number distribution showed only monomeric proteins for lysozyme A-(βCD)$_n$ (C), and only aggregates for aged lysozyme (D). The aging conditions are: 1 mL solution of lysozyme (1mg/mL) was prepared in 1× PBS buffer pH 8.2. The protein solution was then placed in a 28 ºC incubator shaker for 72 h. The sample was then dialyzed in water overnight and then freeze dried to obtain a white solid.

Figure 2.6  Volume (A and B) and Number (C and D) distribution of DLS showing two replicates each for 1 mg/mL of RNase A-(βCD)$_n$ and aged RNase A in 1× PBS buffer pH 7.4 at 25 °C. Volume distribution showed only monomeric proteins for RNase A-(βCD)$_n$ (A) and only aggregates for aged RNase A (B). Number distribution showed only monomeric proteins for RNase A-(βCD)$_n$ (C), and only aggregates for aged RNase A (D). The aging conditions are: 1 mL solution of RNase A (1 mg/mL) was prepared in 1× PBS buffer pH 8.2. The protein solution was then placed in a 28 ºC incubator shaker for 72 h. The sample was then dialyzed in water overnight and then freeze dried to obtain a white solid.

Figure 2.7  Number distribution of DLS of unmodified lysozyme and lysozyme-(βCD)$_n$ (0.5 mg/mL in 1 × PBS, pH 7.4). Number
distribution for unmodified lysozyme kept at (A) 57 °C, (B) at 67 °C for 25 min and for lysozyme-(βCD)$_n$ (C) at 67 °C and (D) at 87 °C for 25 min.

Figure 2.8 Four replicates of surface plasmon resonance (SPR) of the adsorption of (A) lysozyme and lysozyme-(βCD)$_n$, (B) RNase A and RNase A-(βCD)$_n$, (0.5 mg/mL of protein in 1× PBS, pH 7.4). The SPR intensity (pixels) is plotted against time (seconds) as protein solutions are flowed over a SPR micro array.

Figure 2.9 Decrease in optical density with time of suspensions of Micrococcus Lysodeikticus following addition of lysozyme and lysozyme-(βCD)$_n$. The absorbance of the sample was measured at 450 nm.

Figure 2.10 Lysozyme crystals induced by (A) 5’DSCG (figure adapted from reference 106) and (B) NaCl additives in hanging droplets. Lysozyme-(βCD)$_n$ crystals induced by (C) 5’DSCG and (D) NaCl additives in hanging droplets. Scale bar = 76 µm.

Figure 2.11 MALDI-TOF spectra for BSA-(βCD)$_n$.

Figure 2.12 Number distribution of DLS of BSA-(βCD)$_n$ (1.0 mg/mL) in 1 × PBS, pH 7.4 at 25°C.

Figure 2.13 Four replicates of surface plasmon resonance (SPR) of the adsorption of BSA and mixture of unmodified BSA and BSA-(βCD)$_n$, (0.5 mg/mL of protein in 1× PBS, pH 7.4). The SPR intensity (pixels) is plotted against time (seconds) as protein solutions are flowed over a SPR micro array.

Figure 2.14 MALDI-TOF spectra for (A) lysozyme modified with compound 14 and (B) aniline modified lysozyme further modified with compound 13 (C) RNase A modified with compound 14 and (D) aniline modified RNase A further modified with compound 13.

Figure 2.15 (A) Absorption spectra for the ligation of aniline with compound 6 in water at increment of 1 h from 1 to 7 h. Reaction conditions: Aniline (50 µM in H$_2$O), 6 (1eq), rt (B) MALDI-TOF spectra for the reaction mixture of aniline with compound 6 in water after 24 h.

Figure 2.16 MALDI-TOF spectra for lysozyme (46.67 µM) modified with aniline and further modified with compound 6. Modification of lysozyme-aniline with compound 6 was run at pH 6.04 for 72 h.
Chapter 3

Figure 3.1 (A) Structure and amino acid sequence of the monomer of COMPcc
(B) Chemical structures of small molecules that can bind to the hydrophobic cavity of COMPcc.

Figure 3.2 (A) Amino acid sequence of COMPcc monomer, where the mutation site leucine 44 is indicated in red. (B) Amino acid sequence of COMPcc monomer, where leucine 44 is converted to alanine 44. All lysine residues are represented in bold black font.

Figure 3.3 Proposed energy landscape for (A) L44A and (B) L44A-(βCD)$_n$ modification.

Figure 3.4 MALDI-TOF spectra for L44A-(βCD)$_n$ (A) after dialysis in deionized water without filtering through 0.2 µm pore filter (B) after dialysis in deionized water with filtering through 0.2 µm pore filter.

Figure 3.5 Circular dichroism spectra for L44wt, L44A, L44A-(βCD)$_n$, and L44A-(βCD)$_n$ purified. Circular dichroism was conducted by Jasmin Hume in Professor Jin Montclare’s laboratory at Polytechnic Institute of New York University.

Figure 3.6 Circular dichroism spectra for L44A treated with free βCD. Circular dichroism was conducted by Jasmin Hume in Professor Jin Montclare’s laboratory at Polytechnic Institute of New York University.

Figure 3.7 Circular dichroism spectra for L44wt treated with free βCD. (Circular dichroism was conducted by Jasmin Hume in Professor Jin Montclare’s laboratory at Polytechnic Institute of New York University).

Figure 3.8 Number distribution of DLS showing two replicates each of (A) L44A (B) L44 wt (C) L44-(βCD)$_n$ and (D) lysozyme in PBS buffer pH 7.4 at 25 °C.

Figure 3.9 Lysines of L44A modified with (A) compound 2, (B) compound 3 and (C) compound 4.

Figure 3.10 MALDI-TOF spectra for L44A-(acetyl)$_n$ indicated that up to five acetyl groups were conjugated to L44A.

Figure 3.11 MALDI-TOF spectra for (A) L44A-(methyl squarate)$_n$ indicated that up to three methyl squares were conjugated to the lysine residues of L44A.
Figure 3.12 MALDI-TOF spectra for L44A-(PEG)$_n$ indicated up to three PEG chains were conjugated to the lysine residue of L44A.

Figure 3.13 Circular dichroism spectra for L44A modified with compounds 1, 2, 3 and 4. Concentration of L44A-(methyl sq)$_n$ = 29.65 µM, L44A-(acetyl)$_n$ = 8.27 µM, L44A-(PEG)$_n$ = 23.45 µM. Circular dichroism was conducted by Jasmin Hume in Professor Jin Montclare’s laboratory at Polytechnic Institute of New York University.

Figure 3.14 Fluorescence spectra for curcumin binding in the hydrophobic pocket of L44A-(βCD)$_n$. PB in the fluorescence spectra stands for phosphate buffer (pH 8.0). (Fluorescence study has been conducted by Jasmin Hume in Professor Jin Montclare’s laboratory at Polytechnic Institute of New York University).

Figure 3.15 Strategy for preventing assembly formation of L44A-(βCD)$_n$. (A) Assembly formation in L44A-(βCD)$_n$ due to hydrogen bonding between βCD molecules (B) Repulsions between the TEG chains conjugated to adamantane, which forms inclusion complex with covalently grafted βCD molecules, keeps the proteins monomeric.

Figure 3.16 MALDI-TOF spectra for purified (A) L44A-(βCD)$_1$ and (B) L44A-(βCD)$_2$ using reverse phase HPLC.

Figure 3.17 MALDI-TOF spectra for lysozyme modified with (A) compound 2 and (B) compound 3. In both the modifications, some amount of unmodified lysozyme was also observed.

Chapter 4

Figure 4.1 $[^1$H, $^{15}$N] HSQC NMR of βCD adamantane carboplatin analog host-guest complex (6) in 5% D$_2$O/H$_2$O: δ (ppm) $^1$H = 4.13, δ $^{15}$N = -82.0.

Figure 4.2 Plots of survival (%) of SK-N-SH cell treated with (A) carboplatin, and (B) 6, as a function of concentration with different recovery times at 0 h, 24 h, 48 h and 72 h.

Figure 4.3 Survival (%) of SK-N-SH cells when treated with 500 µM carboplatin and 6 at different incubation times of 2 h, 4 h, 6 h, 8 h, and 12 h. The recovery period was 0 h for all incubation times tested.

Figure 4.4 Confocal fluorescence image of SK-N-SH cells incubated with (A) 150 nM of FITC-tagged βCD and (B) FITC alone. The cells were
fixed with 4% paraformaldehyde before viewing under the microscope. Scale bar = 76 µm.

Figure 4.5  Agarose gel electrophoresis image of plasmid DNA (pBR322) (33.3 µM base pairs) after incubation with 6 or carboplatin in 10 mM HEPES, pH 7.4 for 24 h. Lanes 2 through 11 contained 30, 35, 40, 45, 50, 60, 70, 80, 90, and 100 µM of 6, and lanes 12 through 18 contained 30, 40, 50, 70, 80, 90 and 100 µM of carboplatin, respectively. Lane 1 (control) contained only plasmid DNA (pBR322).

Figure 4.6  Agarose gel electrophoresis image of plasmid DNA (pBR322) (33.3 µM base pairs) after incubation with βCD in 10 mM HEPES, pH 7.4 for 24 h. Lanes 2 through 12 contained 30, 35, 40, 45, 50, 60, 70, 80, 90, and 100 µM of βCD respectively. Lane 1 (control) contained only pBR322 DNA.

Figure 4.7  Platinum (amol) taken up per SK-N-SH cell, when exposed to carboplatin (500 µM) and 6 (500 µM) as a function of time. The rate of uptake (amol of Pt cell⁻¹ h⁻¹) was found to be 16.0 ± 0.56 and 64.9 ± 1.4 for cells treated with 6 and carboplatin, respectively.

Figure 4.8  Interaction between directing ligands (antibody, folic acid and galactose) and cancer cell surface receptors (antigen, folic acid receptor and galactose receptor).
## LIST OF SCHEMES

### Chapter 1

| Scheme 1.1 | Covalent modification of lysozyme with D-gulitol -terminated carboxylic acid (6) using NHS/EDC coupling | 7 |
| Scheme 1.2 | Covalent modification of lysozyme with L-gulitol -terminated carboxylic acid (13) using NHS/EDC coupling | 8 |
| Scheme 1.3 | Covalent modification of lysozyme with N-allyl O-phenoxy amino squarate using compound 15 | 11 |
| Scheme 1.4 | Covalent modification of lysozyme with compound 16 | 13 |
| Scheme 1.5 | Synthesis of boronic acid monomer (18) | 16 |

### Chapter 2

| Scheme 2.1 | Different methods of protein glycosylation | 56 |
| Scheme 2.2 | Synthesis of βCD-derivatized amino squarate reagents | 59 |
| Scheme 2.3 | Modification of lysozyme using NHS/EDC coupling | 62 |
| Scheme 2.4 | Heterocyclic squarate based compounds for protein modification | 77 |
| Scheme 2.5 | Sequential modification of proteins with NHS derivatized aniline (14) and compound 13 | 77 |
| Scheme 2.6 | Reaction of aniline with compound 6 | 80 |

### Chapter 3

| Scheme 3.1 | L44A modification with βCD | 109 |
| Scheme 3.2 | Molecules synthesized to study the influence of groups like acetyl, polyethylene glycol (PEG) and methyl squarate on folding of L44A | 115 |
Chapter 4

Scheme 4.1 Synthesis of the βCD adamantane-platinum host-guest complex (6) 144

Scheme 4.2 Synthesis of cis [Pt(NH$_3$)$_2$(H$_2$O)$_2$]$^{+2}$ (5) and conjugation of compound 5 to adamantane carboxylic acid (4). 145

Scheme 4.3 Reduction of WST-8 by dehydrogenases to produce yellow colored dye formazan. 147

Scheme 4.4 Synthesis of FITC tagged βCD (10) 150

Scheme 4.5 Conjugation of compound 14 with cyclodextrin-galactose conjugate (15) 156

Scheme 4.6 Synthesis of cyclodextrin-galactose conjugate (15) 157
LIST OF TABLES

Chapter 2

Table 2.1  Amino acid sequence of lysozyme and RNase A. Lysine (K) is indicated in red font and the lysine residues in the active site are highlighted in blue font.

Table 2.2  Different conditions tried for sequential conjugation of lysozyme and RNase A with compound 14 followed by modification with compound 6.

Chapter 3

Table 3.1  Secondary structure and hydrodynamic diameter of modified L44A with different chemical agents.

Table 3.2  Elution times for L44A-(βCD)₁ and L44A-(βCD)₂ from a C18 HPLC column (eluents: water-0.1%TFA/acetonitrile-water-0.1% TFA).

Table 3.3  Thermal stability of chemically modified lysozyme using dynamic light scattering (DLS).
Chapter 1
A Kosmotrope Approach for Prevention of Protein Aggregation

Summary

Protein structural instability, misfolding, and aggregation are significant challenges faced by the biopharmaceutical industries. There is no well-proposed theory for prevention of protein aggregation. This section of the dissertation is based on the hypothesis that modifying proteins with kosmotropes can help prevent protein aggregation. This hypothesis is based on the century old-observation that certain ions classified as “kosmotropes” preserve protein structure, whereas other ions classified as “chaotropes” tend to denature protein structure. Amine groups on surfaces are believed to be chaotropic in character and are responsible for non-specific adsorption of proteins on surfaces. I hypothesize that the lysine residues on the surface of proteins are potentially chaotropic in character and may be responsible for protein aggregation. Hence, modifying the lysine residues with “organic kosmotropes” can mask the chaotropic character of the lysine residues and can prevent protein aggregation. To test this hypothesis, the lysine residues of proteins like lysozyme were modified with non-natural sugar alcohols (alditols), via an amide bond and a squaramide linkage. Although modification of lysozyme with alditols via amide bond was successful, the extent of modification was low due to the hydrolytic instability of the amide bond. Lysozyme modified with D-mannitol via squaramide linkage using squarate-based reagents exhibited much higher degree of modification compared to the amide bond formation. Hence, squarate based chemistry provides a great opportunity to modify proteins with sugars using a stable bond.
1.1 Background and Significance

1.1.1 Proteins as pharmaceutical drugs

Proteins and nucleic acids have long been used as biopharmaceutical drugs for treatment of diseases. The development and manufacturing of protein therapeutics is increasing each year with the product required globally and the number of new drugs undergoing clinical trials. Amidst the great therapeutic potential of peptide or protein-based drug candidates, 96% are abandoned during preclinical or clinical development due to solubility or aggregation issues. The stability and efficacy of the protein drugs is often lost during production cycle and storage due to self-association leading to formation of aggregates. Aggregation can possibly occur during any of the multiple stages of protein drug processing including fermentation, purification, formulation, and even storage, which are conditions that would subject the proteins under stress. Apart from potential decrease in efficacy and reduced bioavailability, aggregation of protein therapeutics can even lead to serious detrimental effects such as renal failure and immunogenicity. The surface adsorption of protein drugs also causes loss of efficacy of therapeutic proteins in syringes for intravenous delivery. Preventive strategies can help to overcome these degradative pathways and ensure efficacy and stability of protein therapeutics.

1.1.2 The kosmotrope theory – a new basis for preventing protein aggregation

Certain compounds such as glycerol, sucrose and polyethylene glycol can increase the stability of proteins when present in solution. Different theories have been proposed to rationalize how the native protein structure is stabilized in solution by such additives. In 1888, Franz Hofmeister described the effect of salts on protein solubility and arranged salts in the order of their ability to either salt protein out of solution or salt protein into solution. This series came to be known as the Hofmeister series (Figure 1.1).
**Figure 1.1** The ions (cations or anions) towards the right are called chaotropes and act to salt proteins into solution, whereas ions to the left are called kosmotropes and act to salt proteins out of solution.

Certain ions stabilize protein structure or salt proteins out of solution and are classified as kosmotropes, whereas other ions destabilize protein structure or salt proteins into solution and are classified as chaotropes. The idea of such ions present as solutes in protein solutions was extended to small organic molecules called osmolytes produced by bacteria, plants and fish to combat osmotic stress. Osmolytes are believed to be kosmotropic for their ability to preserve the naturally folded structure and function of proteins. The mechanism of the Hofmeister effect is largely empirical and not entirely understood. Considering the effect of additives on protein folding, recent studies suggest that while kosmotropic anions stabilize protein structures by direct interactions, the effect of nonionic organic osmolytes on proteins is likely indirect.

1.1.3 **Drawbacks of protein PEGylation**

Modifying proteins covalently with poly(ethylene glycol) (PEG) is called protein PEGylation and is a process of growing interest to enhance the therapeutic activity of proteins. PEG modified proteins possess reduced toxicity, immunogenicity and proteolysis; reduced rates of kidney clearance; and enhanced stability and solubility. PEGylation not only increases the stability of proteins but can also prevent protein aggregation. Jiang and co-workers have reported the conjugation of α-chymotrypsin with poly(ethylene glycol) and zwitterionic polymers such as poly(carboxybetaine) (pCB). They report that conjugation with pCB improves the stability in a
manner similar to PEGylation, but the pCB conjugates retain or even improve the binding
affinity as a result of enhanced protein-substrate hydrophobic interaction. Poly(ethylene glycol)
(PEG) being a synthetic polymer is polydispersed, and the multiple ether groups of the PEG
units can form undesired complexes with metal ions and lead to medical complications in vivo.
Considering the diverse structures and properties of proteins, a different and competent
chemistry for preventing protein aggregation and surface adsorption is highly desired.

1.1.4 Protein adsorption and denaturation

Denaturation is the alteration of the three-dimensional structure of a protein when subjected to
some form of stress such as heat, acid or alkali, in such a way that it will no longer be able to
carry out its function. Denatured proteins can exhibit a wide range of characteristics ranging
from loss of solubility to aggregation. Proteins have a high tendency to denature to facilitate
hydrophobic interaction between their non-polar residues and their environment, such as a
hydrophobic surface. Protein adsorption on a surface is a complex process, and involves
multistep processes. Regardless of the nature of the surface, protein adsorption involves among
other processes a) reversible adsorption of the native protein, b) partial denaturation and
orientation change, c) desorption of the partially denatured protein from the surface and d)
complete denaturation and immobilization on the surface. Since protein aggregation and
surface adsorption are intimately associated with protein denaturation, I have explored a general
working hypothesis, that covalent modification of proteins with kosmotropic organic molecules
can help prevent protein aggregation and surface adsorption.
1.1.5 Strategies and theories for preventing protein aggregation

Self-assembled monolayers (SAMs) presenting polyols such as D-mannitol are known to prevent protein adsorption on gold surface.\textsuperscript{25} It is believed that SAMs presenting polyols such as D-mannitol can directly template a structured interfacial water layer, which makes the surface invisible to a solvated protein.\textsuperscript{25-27} Using the same hypothesis, covalent modification of proteins with polyols can help template a structured interfacial water layer around a protein making one protein invisible to the other and thus preventing protein aggregation (Figure 1.2).

\textbf{Figure 1.2} Schematic illustrations for (A) unmodified proteins aggregating in solution, (B) Proteins modified with kosmotropes in solution prevents aggregation by templating a structured interfacial water layer around each protein.

Surfaces presenting amines are known to promote non-specific adsorption of proteins.\textsuperscript{28} I hypothesize that the potentially chaotropic character of the lysine residues in proteins is responsible for protein aggregation.\textsuperscript{29} The contribution of lysine residues to the stability of proteins also depends on the nature of counter anions.\textsuperscript{9, 13-15, 30-32} Thus, utilizing the large
abundance of the lysine residues in proteins and their chemical reactivity, modifying the lysine residues in proteins may help increase protein stability.\textsuperscript{33}

1.2 Results and Discussion

1.2.1 Synthesis of alditol-terminated carboxylic acids

I synthesized D-gulitol-terminated carboxylic acid to modify the lysine residues of lysozyme via an amide bond using N-hydroxysuccinimide (NHS) and 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) chemistry (Scheme 1.1). Luk and co-workers have suggested that SAMs formed by enantiomeric and diastereomeric alditol-terminated alkanethiols on gold template a different interfacial water structure and hence have different ability to resist protein adsorption and mammalian cell adhesion.\textsuperscript{34} Based on these results, the enantiomeric L-gulitol-terminated carboxylic acid was also synthesized (Scheme 1.2) using procedures similar to those used for the synthesis of D-gulitol-terminated carboxylic acid and was further used for covalent modification of lysozyme.
Scheme 1.1 Covalent modification of lysozyme with D-gulitol-terminated carboxylic acid (6) using NHS/EDC coupling

The synthesis of molecule 6 is given Scheme 1.1. Briefly, compound 3 was synthesized by literature reported procedures. 35 Nucleophilic substitution on 4-bromovaleronitrile with compound 3 afforded compound 4. Hydrolysis of compound 4 in ethanolic NaOH afforded the acid 5. Deprotection of the acetonide groups in compound 5 with 1N HCl provided the D-gulitol-terminated carboxylic acid, 6. Treatment of lysozyme with activated carboxylic acid 6 (prepared in situ with EDC and NHS) afforded the product 7.
Scheme 1.2 Covalent modification of lysozyme with L-gulitol-terminated carboxylic acid (13) using NHS/EDC coupling

1.2.2 ESI-MS confirms lysozyme modification with D-gulitol

Electron spray ionization mass spectroscopy (ESI-MS) was used to monitor the degree of modification of lysine residues in the protein lysozyme with D-gulitol. Mass spectra showed that one (lysozyme-(D-gulitol)_1) or two (lysozyme-(D-gulitol)_2) lysine residues were modified on the protein with D-gulitol. The mass corresponding to the dimer of unmodified lysozyme was also observed. Such dimer formation is likely due to the aggregation of the unmodified proteins. Interestingly, mass corresponding to the dimer for lysozyme-(D-gulitol)_n was not observed (Figure 1.3). This result is consistent with the hypothesis that modifying proteins with organic kosmotropes can prevent protein aggregation.
Figure 1.3 ESI-MS spectra of lysozyme-(D-gulitol)$_n$ (n can range from 1 to 6), indicating that up to two lysine residues in lysozyme were modified with D-gulitol. Peak for the unmodified lysozyme and dimer of unmodified lysozyme was also observed.

To further characterize the products of lysozyme modification with the alditols, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was conducted on lysozyme modified with D-gulitol. However, no separate bands for lysozyme or modified lysozyme (lysozyme-(D-gulitol)$_n$, (n can range 1 to 6) were observed on the SDS-PAGE. This may be due to the very small difference in mass between unmodified lysozyme and lysozyme-(gulitol)$_n$, making the resolution of the individual bands impossible (Figure 1.4).
Figure 1.4 Gel image from SDS-PAGE for lysozyme modified with D-gulitol. Unmodified lysozyme (Lane 1), sample containing mixture of unmodified lysozyme and lysozyme-(D-gulitol)$_n$ (Lanes 2 and 3). No separate bands corresponding to different modifications of lysozyme with D-gulitol were observed.

1.2.3. Squaramide is a stable linkage compared to amide bond

Luk and coworkers have reported that chemoselective reaction between amino squarate derivatives and the amino acid cysteine or unprotected peptides with N-terminus cysteine proceeds efficiently in aqueous media at physiological pH. Amino squarate derivatives reacted exclusively with cysteine at neutral pH in water containing a mixture of five amino acids (L-cysteine, L-serine, glycine, L-tyrosine, and L-lysine). As the pH was increased to 8.5, L-lysine also started to react suggesting that the rate of reaction for L-lysine with the amino squarate derivative increases with increasing pH. It is believed that the substitution reaction of squarate reagents with nucleophiles in water is enhanced by the high dielectric constant and hydrogen bonding ability of water (Figure 1.5).
Figure 1.5 Hydrogen bonding ability and high dielectric constant of water increases the stability of zwitterionic resonance structure and electrophilicity of squarate in water versus organic solvent.

While amino squarates are reported to be stable against hydrolysis in water and in the presence of esterases and cell lysates, \textsuperscript{36} amide bond formed using derivatives of (NHS)-activated carboxylic acids are prone to hydrolysis in water. Hence conjugating organic kosmotropes to proteins using a squaramide bond rather than an amide bond may be a better method to achieve a hydrolytically stable linkage with high degree of modifications.

Scheme 1.3 Covalent modification of lysozyme with N-allyl squaramide using compound 15

A model reaction was run where lysozyme was modified covalently with N-allyl O-phenoxy amino squarate, 15 (Scheme 1.3). \textsuperscript{37} The primary amine groups of the lysines in lysozyme can displace the phenoxy group bonded to the reactive N-allyl amino squarate moiety in compound
This conjugation reaction was conducted at slightly basic pH of 7.8 in PBS buffer with the aim of keeping the lysine residues potentially deprotonated (pKa of lysine side chain is 10.53).

Figure 1.6 ESI-MS spectra for lysozyme-(N-allyl squaramide)$_n$ indicating that up to three lysine residues are modified with N-allyl squaramide with a relatively small mass peak for unmodified lysozyme.

The ESI-MS spectra of lysozyme-(N-allyl squaramide)$_n$ showed that up to three lysine residues were successfully modified with N-allyl squaramide at pH 7.8 (Figure 1.6). Interestingly, the peak for unmodified lysozyme was smaller in comparison to the peaks of lysines modified with N-allyl squaramide. Since modification of lysozyme using N-allyl O-phenoxy amino squarate, (15) was successful, I attempted to modify the lysine residues in lysozyme with compound 16\textsuperscript{37}.
in which D-mannitol was conjugated to a phenoxy squarate moiety (Scheme 1.4). This approach would ensure that the potentially kosmotropic polyol D-mannitol, is conjugated to the protein via a squaramide linkage which is hydrolytically more stable than the amide bond formation using the NHS activated carboxylic acids.

**Scheme 1.4** Covalent modification of lysozyme with compound 16

Matrix assisted laser desorption ionization time of flight (MALDI-TOF) mass spectrometry was used to determine the degree of modification of the lysine residues in lysozyme with compound 16. The modification of the lysine residues with compound 16 was compared to the degree of modification of the lysine residues with compound 6. Interestingly, the intensity of the product peaks in the mass spectrum corresponding to different degree of modification of the lysine residues with compound 16 was higher than those observed with compound 6 (Figure 1.7). In addition, the peak for unmodified lysozyme was very small. Hence, this chemistry provides a great opportunity to modify proteins with sugars efficiently via the squaramide bond.
Figure 1.7 MALDI-TOF spectra for (A) lysozyme modified with compound 16 indicating that up to two lysine residues in lysozyme are modified with D-mannitol squaramide and a relatively small mass peak for unmodified lysozyme. (B) lysozyme modified compound 6 using NHS/EDC chemistry showing that only one lysine residue is modified with D-gulitol.

1.2.4 Purification of alditol modified lysozyme

In order to study the stability of proteins against denaturation and aggregation, it is imperative to separate the modified protein from the unmodified protein. The cis-diol groups in ribonucleic acid as well as other sugars are known to form complexes with borate anion. This property has been exploited to effect chromatographic separation of nucleosides, oligonucleotides and tRNA by using dihydroxyborylphenyl groups attached covalently to cellulose, methacrylic acid polymer, or polyacrylamide. Based on the same principle, the cis-diols of the alditols can potentially bind to borate anion and allow separation of proteins grafted with alditols from unmodified proteins. The science of preparation of biocompatible porous hydrogels based on polymer stabilization of lyotropic liquid crystal phase has been developed in Luk laboratory.
Such porous hydrogels with pendant boronic acid groups can be used for the isolation and further purification of proteins modified with alditols (Figure 1.8).

**Figure 1.8** Strategy for purifying proteins modified with alditols using boronic acid-laden porous hydrogels.

Boronic acid monomer, (18) was synthesized using literature reported procedure (Scheme 1.5). Boronic acid-laden porous hydrogels were prepared and lysozyme-(D-gulitol)$_n$ was treated with these hydrogels for 24 h. After 24 h, it was expected that lysozyme-(D-gulitol)$_n$ would bind to the pendant borate ion on the hydrogel. The hydrogel was isolated and further washed with buffer (PBS, pH 5.0) to displace lysozyme-(D-gulitol)$_n$. Analysis of the isolated protein sample by ESI-MS indicated that purified lysozyme-(D-gulitol)$_n$ was not obtained. The failure to isolate purified lysozyme-(D-gulitol)$_n$ using the above strategy could be either due to insufficient modification of lysozyme with D-gulitol or the inadequate binding of the cis-diols of the alditols covalently attached to lysozyme.
**Scheme 1.5** Synthesis of boronic acid monomer (18)

High performance liquid chromatography (HPLC) was also used to attempt the separation of lysozyme-(D-gulitol)$_n$ from unmodified lysozyme. However, the purified protein could not be obtained probably due to the small mass difference between modified and unmodified proteins.

1.2.5 *Cyclodextrins* - another class of organic kosmotropes

Cyclodextrins have been reported to stabilize protein structure and aid in protein folding.$^{46-48}$ Cyclodextrins are another class of organic kosmotropes, which have rigid structures and conformations. When covalently conjugated to proteins the multiple hydroxyl groups of cyclodextrins can act as a template to build a structured solvation shell around the protein and help prevent protein aggregation. The ability of cyclodextrins to prevent protein aggregation and surface adsorption will be discussed in detail in chapter 2 of this dissertation. $\beta$-cyclodextrin ($\beta$CD) has a mass close to 1.1 kDa. Thus, based on the difference in mass between unmodified proteins and proteins modified with $\beta$CD, one can potentially effect the separation of modified proteins from unmodified proteins and isolate individual modifications as well.
1.3 Conclusion and Perspectives

The work presented in this section of the dissertation explores the hypothesis that modifying proteins with organic molecules (organic kosmotropes) can stabilize native protein structure, can also prevent protein aggregation and surface adsorption. Lysine residues are believed to be chaotropic and hence are considered responsible for protein aggregation. The chaotropic lysine residues in lysozyme were modified with non-natural sugar alcohols such as D- and L- gulitol using NHS/EDC chemistry. This was done by forming amide bonds with the primary amine groups of the lysine residues. I discovered that conjugating organic kosmotropes to proteins using squaramide derivatives is a better method than making hydrolytically unstable derivatives of (NHS)-activated carboxylic acids. Interestingly, MALDI-TOF spectra showed that the intensity of the peaks corresponding to different degree of modification of the lysine residues using amino squarate conjugated to the polyol D-mannitol, was higher than what was observed when lysozyme was modified with the polyols D-gulitol or L-gulitol using NHS/EDC chemistry. This new squaramide based conjugation chemistry hence provides an opportunity to modify proteins with sugars efficiently with a hydrolytically stable bond. Modification of proteins with larger kosmotropes such as cyclodextrins will be described in the following chapters of this dissertation.

1.4 Experimental section

Materials and Methods

Chemicals. Chicken egg white lysozyme (catalog number L6876) was obtained from SigmaChemical Company (St. Louis, MO) and were used without further purification. Sinapinic acid was obtained from ProteoChem (Denver, 105 CO). Phosphate buffered saline, PBS (2.7 mM potassium chloride, 137 mM sodium chloride, 8 mM sodium phosphate dibasic, 1.48 mM
potassium phosphate monobasic), was used for all protein reactions unless stated otherwise (pH of the prepared buffers was adjusted using 1 N HCl and 1 N NaOH). Ammonium persulfate (APS 99 + %), 3-(trimethoxysilyl) propyl methacrylate, 98% were purchased from Acros Organics USA (Morris Plains, NJ). N,N,N’,N’-Tetramethylethylenediamine (TEMED) was purchased from Sigma Chemical Company (St. Louis, MO). N,N’-Methylenebisacrylamide (bis-acrylamide) was purchased from Fluka Chemicals (Milwaukee, WI). Acrylamide was purchased from MP Biomedicals (Solon, OH). All other chemicals were obtained from Aldrich Chemicals (Milwaukee, WI). Water used for preparing all buffers and solutions had resistivity of 18 MΩ cm (Millipore, Billerica, MA).

**General Methods for Synthesis**

The processes involving reactants sensitive to moisture or air were executed under an atmosphere of argon using oven-dried glassware. Reagents and solvents were reagent grade and used as supplied unless otherwise mentioned. Solvents were distilled under reduced pressure using Buchi rotary evaporator below 40 °C. EMD Silica Gel 60 F254 precoated plates (0.25-mm thickness) were used for TLC and a solution of phosphomolybdic acid/meric sulfate/sulfuric acid (10g : 1.25g : 8% 250 mL), followed by charring at ~ 150 °C, was used for visualization. Flash column chromatography was performed using SILICYCLE, Silia-P Flash Silica Gel with 40-63μ mesh size. $^1$H, $^{13}$C NMR spectra were recorded on 300 MHz, $^1$H chemical shifts are reported in ppm relative to CDCl$_3$ δ 7.26, MeOH δ 4.87, 3.31 and δ D$_2$O 4.80. $^{13}$C chemical shifts are reported relative to CDCl$_3$ δ 77.23, MeOH δ 49.15 and (High Resolution Mass Spectra) HRMS was recorded by positive ion electrospray on a Bruker 12 Tesla APEX –Qe FTICR-MS with Apollo II ion source. Protein ESI-MS spectra were obtained on Mass spectrometer: ESI-Q-TOF II (Waters, Milford, MA, USA) Electrospray: Nanoflow-LC ESI.
**Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)**

SDS-PAGE was conducted using 15% separating gel (3.4 mL of water, 7.5 mL of 30% gel concentrate, 3.8 mL of 1.5 M Tris buffer, pH 8.8, 0.15 mL of 10% sodium dodecyl sulfate (SDS), 0.15 mL of 10% ammonium persulphate (APS) and 6.0 µL of tetramethylethylenediamine (TEMED) and 3% stacking gel (1.32 mL of water, 30% gel concentrate, 0.5 mL of tris buffer, pH 6.8, 0.24 mL of 10% SDS, 0.04 mL of ammonium persulfate and 2.4 µL of TEMED). The 3% staking gel was slowly poured over the 15% separating gel and the comb was gently inserted before the gel polymerized. After 30 min, the comb was removed slowly to load the protein samples in the gel. The protein sample (15 µL, 0.5 mg/mL) was mixed with 10 µL 0.25% of bromophenol blue (BPB) and loaded on the gel. A protein ladder, (Precision Plus Dual Color from BIO-RAD (BIO-RAD, Hercules, CA) was added in one of the wells for molecular weight reference. Tris-glycine (1×) was used as the running buffer for gel electrophoresis. The gels were run at 30 V for 3 h. The gel was stained with coomassie blue for 24 h and then destained in deionized water overnight. The image of the gel was taken on the *Alpha DigiDoc* (Alpha Innotech, San Leandro, CA).

**Dialysis**

The protein samples were dialyzed in 1× PBS (dialysis media), using Slide-A-Lyzer Dialysis Cassettes 7K molecular weight cut off (MWCO) purchased from Thermo Scientific (Rockford, IL). Before use, the Slide-A-Lyzer Dialysis Cassette was hydrated in the dialysis media for 5 min. The sample was injected in the cassette and placed in 250 mL of dialysis media to eliminate unreacted compound 6. The buffer was changed twice every 2 h for 6 h and then left overnight,
after which time the sample was removed from the dialysis cassette and protein sample was analysed by MALDI-TOF or ESI-MS.

**Polymerization of Boronic Acid Monomers into Hydrogels**

Hydrogels were prepared in 2 mL vials (1.25 cm × 4.5 cm) blown with nitrogen gas. Polymerization of boronic acid monomers into hydrogels was done by addition of 0.4% w/w of APS (from 40% w/w stock solution prepared in water) and bis-acrylamide (from 2% w/w stock solution in water) as a crosslinker, to 6% w/w of boronic acid monomers (50% w/w stock solution of compound 17 and acrylamide in water). The sample was vortexed for 1 min, followed by the initiation of polymerization by addition of 0.2% w/w of TEMED (from 20% w/w stock solution prepared in water). TEMED accelerates the polymerization reaction by catalyzing the formation of free radicals from APS. The sample was left undisturbed at rt for 24 h to allow complete polymerization.

**Synthetic procedures and Spectral data**

The synthetic procedures and spectral data for the compounds successfully synthesized are given below:

*Synthesis of compound 2*

To a solution of D-(-)-gulonic-γ-lactone (3.00 g, 16.8 mmol) in dry acetone (30 mL) was added 2, 2-dimethoxypropane (10.0 g, 96.0 mmol) and added TsOH monohydrate (catalytic amount).
The mixture was heated at 50 °C for 3 h. Reaction was quenched with 0.2 mL of triethylamine. Crude product was obtained as a white solid after evaporation of solvent. The crude product was purified using silica gel column (35% EtOAc in Hexane) to give compound 2 (2.84 g, 65%) as a white solid. Rf = 0.56, (50% EtOAc in Hexane). 1H NMR (300 MHz, CDCl3): δ 4.84-4.71 (m, 2H), 4.45-4.41 (m, 2H), 4.24-4.21 (m, 1H), 3.84-3.76 (m, 1H), 1.43 (s, 3H), 1.41 (s, 3H), 1.39 (s, 3H), 1.37 (s, 3H). 13C NMR (75 MHz, CDCl3): δ 173.0, 114.9, 110.7, 81.1, 76.2, 75.9, 75.4, 65.4, 26.9, 26.8, 26.1, 25.4. HRMS found 281.0990 [M + Na]+ calcd for [C12H18O6+Na]+ = 281.0995.

**Synthesis of compound 3**

To an ice cooled solution of compound 2 (2.84g, 11.0 mmol) in MeOH (28 mL), NaBH4 (1.45 g, 38.5 mmol) was added portion wise with stirring. The mixture was stirred at rt for 2 h. While cooling the reaction mixture in an ice bath with vigorous stirring, 1.0 M AcOH solution in water was added drop wise to neutralize the mixture. After evaporation of solvents the residue was suspended in brine and extracted with EtOAc until TLC showed no product in the aqueous phase. The combined organic phase was dried over anhydrous Na2SO4 to obtain the crude product as colorless oil after evaporation of solvent. The crude product was purified using silica gel column (75% EtOAc in Hexane) to give compound 3 (2.454 g, 85%) as a colorless oil. Rf = 0.25 (80% EtOAc in Hexane). 1H NMR (300 MHz, CDCl3): δ 4.27-4.21 (m 2H), 4.21-4.02 (m, 2H), 3.85-3.74 (m, 4H), 3.02 (br s, 1H), 1.50 (s, 3H), 1.43 (s, 3H), 1.36 (s, 3H), 1.35 (s, 3H). 13C

Synthesis of compound 4

Compound 3 (2.45 g, 9.36 mmol) was dissolved in THF (24 mL) and the solution was, cooled to -75 °C and charged with NaH (0.430 g, 60% by wt in mineral oil, 17.9 mmol). The temperature was allowed to rise to 0 °C slowly and the mixture was allowed to stir at 0 °C for 90 min. Then 5-bromovaleronitrile (1.82 g, 11.2 mmol) was added drop wise. The mixture was allowed to stir at rt for 2 d. The reaction was quenched with water and extracted with EtOAc. The organic phase was dried over anhydrous Na₂SO₄. The crude product was obtained as colorless oil after evaporation of solvent. The crude product was purified using flash silica gel column (70% EtOAc in Hexane) to obtain compound 4 (1.276 g, 56%) as colorless oil. R_f = 0.52 (80% EtOAc in Hexane). ¹H NMR (300 MHz, CDCl₃): δ 4.27-4.04 (m, 2H), 4.03-3.95 (m, 2H), 3.78-3.45 (m, 5H), 2.75 (d, J_H-H = 6.49, 1H), 1.67-1.66 (m, 2H), 1.43 (m, 3H), 1.37 (m, 3H), 1.30 (m, 3H), 1.29 (m, 3H). ¹³C NMR (75 MHz, CDCl₃): δ 119.6, 109.6, 108.8, 105.9, 77.7, 77.3, 77.0, 76.9, 76.4, 75.9, 74.4, 70.6, 69.8, 69.7, 68.9, 66.0, 54.6, 29.7, 28.5, 27.0, 26.6, 25.4, 25.1, 22.5, 17.0. HRMS found = 366.1884[M + Na]⁺ calcd for [C₁₂H₁₈O₆+Na]⁺ = 366.1887.
Synthesis of compound 5

To a solution of compound 4 (1.27 g, 3.71 mmol) in EtOH (60 mL) added 25 N NaOH (13 mL) was heated under reflux for 18 h. Reaction mixture was neutralized using concentrated HCl. Solvent was evaporated under reduced pressure to yield a white residue. The residue was washed with THF and the washings were filtered through a pad of celite. The crude product was obtained as colorless oil after evaporation of solvent. The crude product was purified using flash silica gel column (50% EtOAc in Hexane and then eluting the pure compound in MeOH) to obtain compound 5 (0.987, 73%) as a colorless oil. R_f = 0.21 (80% EtOAc in Hexane). 1H NMR (300 MHz, CDCl_3): \( \delta \) 4.30-4.22 (m, 2H), 4.07-4.02 (m, 2H), 3.83-3.58 (m, 4H), 3.47 (br s, 2H), 2.25 (br s, 2H) 1.93 (br s, 1H), 1.60 (br s, 4H), 1.49 (m, 3H), 1.43 (m, 3H), 1.36 (m, 3H), 1.34 (m, 3H). 13C NMR (75 MHz, CDCl_3): \( \delta \) 109.8, 108.9, 76.9, 75.9, 71.7, 70.1, 69.2, 66.1, 35.1, 29.1, 27.0, 26.6, 25.4, 25.2, 22.0. HRMS found = 385.1828[M + Na]^+, calcld for [C_{12}H_{18}O_6Na]^+ = 385.1832.
Synthesis of compound 6

Compound 5 (0.987 g, 2.72 mmol) was suspended in 15 mL of 1N HCl and refluxed for 1 h. The reaction mixture was brought to rt and neutralized by adding 1N NaOH drop wise. Solvent was evaporated to yield a white residue. The residue was washed with THF and the washings were filtered through a pad of celite. The filtrate was concentrated to yield compound 6 (0.376 g, 49%) as a hygroscopic solid. $^1$H NMR (300 MHz, CDCl$_3$): $\delta$ 3.97-3.30 (m, 10H), 2.27-2.25 (br m, 2H), 1.64 (br s, 4H). $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 180.2, 73.8, 72.9, 72.1, 70.9, 70.4, 69.1, 65.7, 63.0, 35.0, 29.0, 22.0, 14.2. HRMS found = 305.1206 [M + Na]$^+$, calcd for [C$_{12}$H$_{18}$O$_6$+Na]$^+$=305.1206.

Conjugation of lysozyme with compound 6

Added 0.30 mg EDC and 0.20 mg of NHS, to 1.00 mg of compound 6 dissolved in 2-(N-morpholino)ethanesulfonic acid (MES) buffer, at pH 6.0. Compound 6 was reacted with NHS and EDC for 15 min at rt. To the reaction mixture added 0.70 µl of mercaptoethanol to quench excess of EDC. Added 500 µL of 1 mg/mL solution of lysozyme in MES buffer at pH 6.0, and the protein reaction was run at rt for 2 h. The protein reaction was then dialysed in PBS buffer,
pH 7.4, to eliminate unreacted reagents. Degree of protein modification was analysed by ESI-MS.

_Synthesis of compound 9_

![Chemical structure of compound 9](attachment:image.png)

To a solution of L- (+)-gulonic-γ-lactone (2.00 g, 11.2 mmol) in dry acetone (40 mL) added 2, 2-dimethoxypropane (6.66 g, 64.0 mmol) and TsOH monohydrate (catalytic amount). The mixture was heated at 50 °C for 3 h. Reaction was quenched with 0.1 ml of triethylamine. Crude white solid was obtained after evaporation of solvent. The crude product was purified using silica gel column (35% EtOAc in Hexane) to give compound 9 (1.99 g, 69%) as a white solid. R_f = 0.56, (50% EtOAc in Hexane). $^1$H NMR (300 MHz, CDCl$_3$): $\delta$ 4.84-4.72 (m, 2H), 4.45-4.39 (m, 2H), 4.24-4.19 (m, 1H), 3.84-3.79 (m, 1H), 1.47 (br s, 3H), 1.39 (s, 3H), 1.37 (s, 3H). $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 173.0, 114.9, 110.5, 81.1, 76.2, 75.9, 75.4, 26.9, 26.8, 26.0, 25.4. HRMS found = 281.0991 [M + Na]$^+$, calcd for [C$_{12}$H$_{18}$O$_6$+Na]$^+$ = 281.0995.

_Synthesis of compound 10_

![Chemical structure of compound 10](attachment:image.png)

To an ice cooled solution of 9 (1.99 g, 7.73 mmol) in MeOH (20 mL), NaBH$_4$ (1.02 g, 27.0 mmol) was added portion wise with stirring. The mixture was stirred at rt for 2 h. While cooling
the reaction mixture in an ice bath with vigorous stirring, 1.0 M AcOH solution in water was added drop wise to neutralize the mixture. After evaporation of solvents the residue was suspended in brine and extracted with EtOAc until TLC showed no product in the aqueous phase. The combined organic phase was dried over anhydrous Na$_2$SO$_4$ to obtain the crude product as colorless oil after evaporation of solvent. The crude product was purified using silica gel column (75% EtOAc in Hexane) to give compound 10 (1.619 g, 80%) as a colorless oil. $R_f$ = 0.25 (80% EtOAc in Hexane). \[^1\]H NMR (300 MHz, CDCl$_3$): $\delta$ 4.27-4.19 (m, 2H), 4.08-4.00 (m, 2H), 3.83-3.73 (m, 4H), 3.10 (br s, 1H), 1.48 (s, 3H), 1.41 (s, 3H), 1.33 (s, 3H), 1.33 (s, 3H). \[^{13}\]C NMR (75 MHz, CDCl$_3$): $\delta$ 109.8, 108.7, 77.5, 76.2, 69.7, 66.0, 61.2, 27.1, 26.6, 25.4, 25.2. HRMS found = 285.1302 [M + Na]$^+$ calcld for [C$_{12}$H$_{18}$O$_6$+Na]$^+$ = 285.1308.

**Synthesis of compound 11**

![Compound 11](image)

Compound 10 (0.600 g, 2.28 mmol) was dissolved in THF (15 mL) and the solution was, cooled to -75 °C and charged with NaH (0.105 g, 60% by wt in mineral oil, 4.37 mmol). The temperature was allowed to rise to 0 °C slowly and the mixture was allowed to stir at 0 °C for 90 min. Added 5-bromovaleronitrile (0.320 g, 1.97 mmol) drop wise. The mixture was allowed to stir at rt for 2 d. The reaction was quenched with water and extracted with EtOAc. The organic phase was dried over anhydrous Na$_2$SO$_4$. The crude product was obtained as colorless oil after evaporation of solvent. The crude product was purified using flash silica gel column (70% EtOAc in Hexane) to obtain compound 11 (0.342 g, 62%) as colorless oil. $R_f$ = 0.52 (80% EtOAc
in Hexane). $^1$H NMR (300 MHz, CDCl$_3$): $\delta$4.33-4.20 (m, 2H), 4.18-4.00 (m, 2H), 3.82-3.46 (m, 5H), 2.76 (d, $J_{H-H} = 6.53$, 1H), 2.38-2.33 (m, 2H), 1.73-1.70 (m, 4H), 1.48 (s, 3H), 1.42 (s, 3H), 1.35 (s, 3H), 1.34 (s, 3H). $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 119.6, 109.7, 108.9, 70.6, 69.9, 69.7, 29.3, 28.5, 27.0, 26.7, 25.4, 25.1, 22.5, 17.1. HRMS found = 366.1874 [M + Na]$^+$; calcd for [C$_{12}$H$_{18}$O$_6$+Na]$^+$ = 366.1887.

**Synthesis of compound 12**

![Diagram of compound 12]

To a solution of 11 (0.342 g, 0.995 mmol) in EtOH (16 mL) added 25 N NaOH (3.0 mL) and heated under reflux for 18 h. Reaction mixture was neutralized using concentrated HCl. Solvent was evaporated under reduced pressure to yield a white residue. The residue was washed with THF and the washings were filtered through a pad of celite. The crude product was obtained as colorless oil after evaporation of solvent. The crude product was purified using flash silica gel column (50% EtOAc in Hexane and then eluting the pure compound in MeOH) to obtain compound 12 (0.245, 87%) as a colorless oil. R$_f$ = 0.21 (80% EtOAc in Hexane). $^1$H NMR (300 MHz, CDCl$_3$): $\delta$4.24-4.20 (m, 2H), 4.02-3.97 (m, 2H), 3.79-3.58 (m, 5H), 3.41-3.36 (m, 2H), 2.16 (br s, 2H), 1.53 (br s, 4H), 1.38 (s, 3H), 1.35 (s, 3H), 1.31 (s, 3H), 1.30 (s, 3H). $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$109.7, 109.4, 108.9, 71.8, 70.1, 69.2, 66.2, 66.1, 35.6, 29.2, 27.0, 26.6, 25.6, 25.5, 25.2, 22.1. HRMS found = 385.1827 [M + Na]$^+$; calcd for [C$_{12}$H$_{18}$O$_6$+Na]$^+$ = 385.1832.
Synthesis of compound 13

Compound 12 (0.245 g, 0.678 mmol) was suspended in 10.0 mL of 1N HCl and refluxed for 1h. The reaction mixture was brought to rt and neutralized using 1N NaOH drop wise. Solvent was rotaevaporated to yield a white residue. The residue was washed with THF and the washings were filtered through a pad of celite. The filtrate was concentrated to yield 13 (quantitative) as a hygroscopic solid. \(^1\)H NMR (300 MHz, CDCl\(_3\)): \(\delta 3.94-3.29\) (m, 10H), 2.21-2.19 (br s, 2H), 1.64 (br s, 4H). \(^{13}\)C NMR (75 MHz, CDCl\(_3\)): \(\delta 182.3, 73.7, 72.7, 72.5, 72.1, 71.4, 71.0, 70.3, 69.2, 63.0, 36.7, 29.2, 22.7\). HRMS found = 305.1207 [M + Na]\(^+\), calcd for \([\text{C}_{15}\text{H}_{18}\text{O}_6\text{+Na}]^+\) = 305.1206.

Conjugation of lysozyme with compound 13

Added 100 µL of 30 mg/mL solution of EDC and 6.6 µL of 30 mg/mL solution of NHS to 500 µL of 1 mg/mL solution of compound 13 in MES buffer, at pH 6.0 and reacted for 15 min at rt. To the reaction mixture 0.7µl of mercaptoethanol was added to quench EDC, followed by addition of 500 µL of 1 mg/mL solution of lysozyme in MES buffer at pH 6.0, which was reacted for 2 h at rt. ESI-MS found = 14570.00[M]\(^+\). The protein reaction was then dialysed in
PBS buffer, pH 7.4, to eliminate unreacted reagents. Degree of protein modification was analysed by ESI-MS.

*Synthesis of compound 17*

![Chemical structure of compound 18](image)

Compound 17 (0.100 g, 0.576 mmol) and NaHCO₃ (0.193 g, 2.29 mmol) were dissolved in a 2:1 vol/vol mixture of water and THF. To this solution acryloyl chloride (0.104 g, 1.15 mmol) was added at 0 °C, and stirred mixture for 1.5 h. Product was extracted with EtOAc. After evaporation of solvent the product was recrystallized with water to obtain compound 18 (0.020 g, 18%) as a white crystalline solid. ¹H NMR (300 MHz, CDCl₃): δ 7.87 (s, 1H), 7.71 (d, J₆₋₇ = 7.05, 1H), 7.68-7.30 (br m, 2H), 6.48-6.31 (m, 2H), 5.77-5.73 (dd, J₆₋₇ = 9.52, J₇₋₈ = 9.46, 1H). HRMS found = 214.0649 [M + Na]^+, calcd for [C₁₂H₁₆O₆Na]^+ = 214.0645.
(13)
Chapter 2

Covalent Grafting of Proteins with Cyclodextrins Prevents Aggregation and Surface Adsorption and Increases Thermal Stability*

Summary

This chapter describes a general approach for modifying proteins with β-cyclodextrins (βCD) which prevents their aggregation and adsorption on surfaces, and increases their thermal stability. Using water-driven ligation, βCD is covalently tethered to the lysine residues of proteins via a squaramide linkage. The cyclodextrin-modified proteins exhibited significant reduction in aggregation, surface adsorption and an increase in thermal stability compared to unmodified proteins. These results are consistent with the hypothesis that molecules that can stabilize protein structure can also reduce protein aggregation and surface adsorption when grafted covalently to the surface of proteins. The ability of βCD to aid in protein folding will further be utilized to conjugate antibodies for targeted drug delivery and for increasing the stability of the antibody-conjugated drug.

2.1 Background and Significance

Aggregation and surface adsorption of protein-based biopharmaceuticals not only decreases the efficacy of a biopharmaceutical drug, but also causes undesired and sometimes fatal immunoresponse by the body. Both aggregation and surface adsorption of proteins are daunting problems faced by several industries, but are mostly studied as two separate problems. However,

in spite of their differences, both aggregation and surface adsorption of proteins involve some degree of protein denaturation. Covalent grafting of kosmotropic polyols such as β-cyclodextrins (βCD) on the surface of proteins can help preserve the folding stability of proteins and prevent protein denaturation. In this section of the dissertation, the chemistry for reducing both aggregation and surface adsorption and increasing thermal stability of proteins by modifying proteins with β-cyclodextrins is presented.

2.1.1 Cyclodextrins can prevent protein aggregation

Carbohydrates and their derivatives are another class of molecules known to stabilize protein folding, and in some cases prevent protein adsorption. In particular, cyclodextrin\textsuperscript{52} in solution has the ability to function as an artificial chaperone that refolds protein structures,\textsuperscript{46-47, 53} and inhibits aggregation. Cyclodextrin dimers have also been shown to disrupt protein aggregation.\textsuperscript{54} These dimers cause the disruption of hydrophobic interactions between proteins, or between proteins and surfactants, strongly suggesting that the presence of cyclodextrin facilitates protein folding and refolding. In addition to the hydrophobic patches on proteins that can cause protein association and aggregation,\textsuperscript{55-56} chaotropic additives can also promote protein denaturation and aggregation. Covalent grafting of cyclodextrins onto proteins offers an entirely new approach for preventing protein adsorption and aggregation and enhancing their thermal stability.

2.1.2 Cyclodextrins versus poly(ethylene glycol) for increasing protein stability

Unlike oligo(ethylene glycol) (OEG) or poly(ethylene glycol) (PEG), cyclodextrins have well-defined molecular weights and rigid conformations. Cyclodextrins have both hydrogen bond acceptors and donors, whereas PEG only possesses hydrogen bond acceptors. Thus, the solvation of βCD and PEG in water will be different. When grafted on proteins, βCD can exhibit a fixed conformation with limited orientational variations, whereas PEG can exhibit a large ensemble of
conformations and orientational variations. These differences suggest that both the enthalpy and entropy for aggregation and surface adsorption are different for βCD-grafted and PEG-grafted proteins. The mechanisms involved in reducing both protein aggregation and surface adsorption for βCD-grafted and PEG-grafted proteins will also be different.

Recent studies suggest that while kosmotropic anions stabilize protein structures by direct interactions,9,12-13,16 the effect of nonionic organic osmolytes on proteins is likely indirect.9 Particularly for sugar molecules that stabilize protein structures,57 by exerting a long range effect on solvent water, the hydrogen bonding associated with molecular solvation is likely multivalent in nature. For instance, sucrose gradients commonly used for purification of biomolecules, consists of visible layers of different density that can be easily prepared and do not readily become a homogeneous solution.58 For this reason, I believe that the multiple hydroxyl groups in cyclodextrin can hydrogen bond with water molecules and can act as a template to build a structured solvation shell that contributes to two plausible mechanisms for preventing aggregation and surface adsorption. First, folding stability is increased by grafting βCD on protein surfaces, thus reducing the denaturation-initiated aggregation and surface adsorption. Second, βCD creates less dynamic hydration shells on proteins, which prevent proteins moving into close contact with each other.
Figure 2.1 A schematic representation demonstrating the ability of the multiple hydroxyl groups of β-cyclodextrin to generate a structured solvation shell by hydrogen bonding with surrounding water molecules.

It is also possible that the βCD groups grafted covalently on the surface of proteins form intramolecular inclusion complexes with nearby hydrophobic residues on the same protein. However, such binding may also induce intermolecular binding that would lead to aggregation. The mechanism of this intramolecular inclusion does not exclude the model presented in this chapter. Both can be effective under different conditions.
2.1.3 Different methods of protein glycosylation

Glycosylation of proteins has been used to enhance the solubility, increase stability and activity of proteins.⁶⁰-⁶³ A number of methods exist for conjugating monosaccharides or oligosaccharides to proteins and peptides.⁶⁴ One approach is chemical modification of reactive amino acid residues in proteins and peptides with sugar derivatives. These methods include formation of acetimidine,⁶⁵ glycosylmaleimides,⁶⁶ diazo,⁶⁷ thiourea,⁶⁸-⁶⁹ amide⁶², ⁷⁰-⁷² and disulfide bond formation.⁷³ Lemieux and co-workers have reported that bovine serum albumin can be chemically glycosylated via amide bonds, to generate semisynthetic antigens for raising carbohydrate specific antibodies.⁷⁰ Hindsgaul and coworkers reported squarate-based chemical modification, to couple oligosaccharide amines to proteins using diethyl squarate (Scheme 2.1).⁷⁴
**Scheme 2.1** Different methods of protein glycosylation

(a) $\text{AcO} \rightarrow \text{Lys}$ 
   - $\text{(H}_2\text{N})\text{CS}$
   - $\text{CICH}_2\text{CN}$

(b) $\text{HO} \rightarrow \text{Lys}$
   - $\text{NaBH}_3\text{CN}$

(c) $\text{HO} \rightarrow \text{Lys}$
   - $\text{HNO}_2$

(d) $\text{HO} \rightarrow \text{Lys}$
   - $(\text{Im})\text{CS}$

(e) $\text{HO} \rightarrow \text{Lys}$
   - $\text{L-GLu-pNA}$

(f) $\text{HO} \rightarrow \text{Lys}$
   - $\text{R=OCH}_3$, $\text{R=NH}_{3}\text{H}_2$

(g) $\text{HO} \rightarrow \text{Lys}$
   - $(\text{i})$ Hal oxidation
   - $(\text{ii})$ NaOH (aq)

(h) $\text{HO} \rightarrow \text{Lys}$
   - Asp, Glu

(i) $\text{HO} \rightarrow \text{Lys}$
   - $\text{R=OCH}_3$, $\text{R=NH}((\text{CH})_2\text{NH}_2$
Another approach for glycosylating proteins and peptides involves a combination of site-directed mutagenesis and chemical modification. Ullman and co-workers have introduced cysteine residues into proteins, to facilitate the coupling of the thiol group of cysteine specifically with glycomethanethiosulfonate (glycoMTS) reagents via disulfide linkages. Copper catalyzed click reactions between alkyne modified proteins and sugar azides have been used by Carell and co-workers for glycosylation of proteins. Bertozzi and co-workers demonstrated that enzymatic methods can be used to introduce aldehyde groups at selective positions of peptides. These aldehyde groups were further reacted with aminoglycosides to produce glycosylate peptides via oxime bonds. While significant advances have been made in bioconjugation in recent years, modification of proteins with cyclodextrins, particularly with a stable and well-characterized bond identity remains a challenge.

In this work, I report the direct covalent modification of lysine residues of proteins with β-cyclodextrin (βCD), using water-driven ligation with activated amino squarate reagents. The lysine residues of lysozyme and RNase A were covalently modified with βCD with the aim of preventing protein aggregation and surface adsorption and to promote thermal stability.
Table 2.1 Amino acid sequence of lysozyme and RNase A. Lysine (K) is indicated in red font and the lysine residues in the active site are highlighted in blue font.

Source: NCBI Sequence Viewer


<table>
<thead>
<tr>
<th>Protein</th>
<th>Accession number</th>
<th>Source</th>
<th>Amino acid sequence</th>
<th>Amino acids in the active site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysozyme</td>
<td>1AZF_A</td>
<td>Chicken</td>
<td>1KVFGRCCELAA AMKRHGLDNY RGYSLGNWVC AAKFSENFNT OATNMTDGS TDYGLQINS 61 RWCCNDGRTGP GSRINCNIPC SALLSSDITA SVNNCAKIVS DGNGMNAWWA WRNRCKGTDV 121 GAWIRGCCR</td>
<td>Glu35, Asp52</td>
</tr>
<tr>
<td>RNaseA</td>
<td>AAB35594</td>
<td>Cattle</td>
<td>1PSSLGGKETTAATA KFEROHMDSST TAAAANYNCS QMMKSRNLTT KDRCKPVNTF VHESLADVQAA 61 VCSQKNVACK NGQTNCCQSY STMSTDCRE TGSSKYIPCA YKTTQANKHI IVACEGNPPYV 121 PVMFUDAS</td>
<td>Lys41, Phe120, Asp121, Gln11</td>
</tr>
</tbody>
</table>

Table 2.1 gives the amino acid sequence of lysozyme and RNase A. Lysozyme and RNase A have six and ten lysine residues respectively. Although lysozyme does not have any lysine residues in its active site, one lysine residue is present in the active site of RNase A. If the lysine residue in the active site of the proteins is covalently modified with βCD, the activity of the protein may be lost. Furthermore, the large βCD group, which has an outer diameter of 1.54 nm, might itself block the active site of the proteins and may lead to reduced protein activity.

2.2 Results and Discussion

Amino squarates are stable in water, but have unique reactivity for substitution that is driven by the hydrogen bonding ability of water.36 The unique stability and reactivity of squarate reagents in water provides an opportunity to modify proteins effectively with βCD-based amino squarate reagents.
2.2.1 Synthesis of βCD-derivatized amino squarate reagents

Scheme 2.3 shows the optimized synthesis of βCD amino squarate reagents 6, 7 and 8. Briefly, compounds 1,\(^4\) and mono-6-amino-6-deoxy-β-cyclodextrin, \(^5\) were synthesized using literature-reported procedures. Compounds 2 and 3 were obtained by coupling 1 with 3, 5-difluorophenol and 4-fluorophenol, respectively, in the presence of triethyl amine (TEA). Compounds 6, 7 and 8 were obtained by mixing 5 with compound 2, 3 or 4, respectively, in water with minimal amounts of N,N'-dimethylformamide (DMF) and acetone.

**Scheme 2.2** Synthesis of βCD-derivatized amino squarate reagents

βCD amino squarates 6, 7 and 8 (Scheme 2.2) were synthesized to covalently modify the lysine residues in the proteins lysozyme and RNase A with βCD. The introduction of an electron withdrawing fluorine group on the phenyl moiety in the βCD amino squarate reagent may cause the lysine residues of proteins to displace the fluorinated phenoxy substituent more efficiently than a phenoxy substituent. After synthesizing the compounds, the degree of modification of the lysine residues in lysozyme (lysozyme-(βCD)\(_n\), n can range from 1 to 6) with reagents 6, 7 and 8 at 5 h, 10 h, 24 h, 48 h and 72 h was studied using matrix assisted laser desorption ionization
time of flight (MALDI-TOF) mass spectrometry. It was observed that all three amino squarate reagents modified the same number of lysine residues, but the intensity of the mass peaks in the MALDI spectrum corresponding to each degree of modifications ((lysozyme-(βCD))<sub>1</sub>, (lysozyme-(βCD))<sub>2</sub>, (lysozyme-(βCD))<sub>3</sub> etc.) was different. After 5 h, two lysine residues of lysozyme were modified with βCD using 6, 7 and 8; however the intensity of peaks corresponding to different degree of modifications using reagent 6 was higher than that by using reagents 7 or 8. After 10 h, number of lysine residues modified with 6, 7 and 8 increased to three, but the intensity of the mass peaks corresponding to different degree of modifications were highest for 6, followed by 7 and 8. Prolonged reaction time (72 h) increased the degree of modification of lysine residues to four, with the intensity of peaks corresponding to different degree of modifications being similar for all three compounds 6, 7, and 8.

2.2.2 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and MALDI-TOF confirms the degrees of modification in lysozyme with βCD

To further characterize the products of lysozyme modification with βCD, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was conducted on lysozyme modified with 6, 7 and 8. Separate bands corresponding to one, two and three lysine residues modified with βCD, were observed on the SDS-gel (Figure 2.2). This observation is consistent with the results from MALDI-TOF analysis, where multiple βCD modifications were observed in the protein lysozyme.
Figure 2.2 SDS-PAGE showing separate bands corresponding to one and two modifications of lysozyme (0.05 mg/mL) with βCD using compounds 6, 7 and 8. Lane A is lysozyme; lane B, C and D are for lysozyme modified with 8, 7 and 6 respectively. Inset above shows SDS-PAGE with high concentration of modified lysozyme (0.5 mg/mL) showing separate bands corresponding to one, two and three modifications with βCD.

2.2.3 Modification of lysozyme using NHS/EDC coupling

The modification of lysozyme with βCD using N-hydroxy succinimide/1-Ethyl-3-[3-dimethylaminopropyl] carbodiimide hydrochloride (NHS/EDC) or N,N'-dicyclohexylcarbodiimide (DCC) chemistry was also attempted. Compound 12 was synthesized to conjugate βCD to the lysine residues in lysozyme using NHS coupling chemistry. The synthesis of compound 12 is shown in Scheme 2.3. Briefly, compound 9 was activated with NHS/DCC to give the ester 10. Compound 10 was further treated with the amine 5 to afford the amide 11. Base hydrolysis of compound 11 gave the acid 12. Treatment of lysozyme with activated 12 (prepared in situ with EDC and NHS) should have afforded lysozyme-(βCD)$_n$. 
However, MALDI-TOF analysis on the dialysed sample for lysozyme modification with 12 indicated absence of any modification. It is possible that lysine residues could not be modified with βCD due to the hydrolytic instability of the amide bond linking βCD to the lysine residues of proteins. Cyclodextrins presumably can also catalyze hydrolysis of ester bonds. Hence, modification of proteins with βCD using the squarate chemistry is a better method compared to NHS/EDC chemistry, due to the hydrolytic stability of the squaramide linkage.

**Scheme 2.3** Modification of lysozyme using NHS/EDC coupling

![Scheme 2.3](image)

2.2.4 Prevention of protein aggregation

Lysozyme and RNase A are the most studied model proteins and their aggregation propensity has been widely reported. The aggregation process largely depends on pH, temperature and protein concentration. Earlier reports using equilibrium sedimentation have shown that 1 mM of hen egg white lysozyme forms dimers in the pH range 5–9 and higher oligomers at pH 10–11. Hen egg white lysozyme has also been reported to form fibrils at acidic pH and elevated temperatures. Aggregation of lysozyme at micro molar concentrations has been shown to be initiated at an alkaline pH of 12.2. Dimers of RNase A have been reported to form at 37 °C and
65 °C at pH 6.5.\textsuperscript{93, 97} Ribonuclease A (RNase A) has also been shown to form aggregates when stored in the freeze-dried form at elevated temperatures.\textsuperscript{94}

2.2.5 Dialysis in water aggregates only unmodified proteins

After the reaction of lysozyme with the βCD-amino squarate reagent 6, the unreacted reagents were removed by dialysis in phosphate buffered saline (PBS), pH 7.4. Interestingly, further dialysis of the samples in deionized water instead of in buffer, followed by filtration of the protein solution through a 0.2 µm syringe filter, resulted in a significant reduction in the amount of unmodified protein. MALDI-TOF spectra showed that up to four lysine residues in lysozyme were modified with βCD (Figure 2.3) using compound 6. The MALDI-TOF spectrum of protein samples dialyzed overnight in deionized water showed mass peaks corresponding mainly to the modified protein. RNase A was also modified with βCD using compound 6. Similar degree of modification was obtained with RNase A (Figure 2.4). MALDI-TOF analysis showed only small amount of unmodified RNase A relative to the βCD modified RNase A. Furthermore, MALDI-TOF analysis on samples of modified RNase A dialysed in deionized water indicated the presence of small amount of unmodified RNase A, while the peaks of βCD modified RNase could still be observed. The effect of ionic strength on protein stability depends on the nature of the protein.\textsuperscript{91, 99} These results suggest that unmodified lysozyme and RNase A form aggregates due to low ionic strength of deionized water. I believe that the modification of protein with βCD stabilizes the folded structure of the protein and prevents denaturation and aggregation, thus only the unmodified proteins formed aggregates during the stress of dialysis.
Figure 2.3 MALDI-TOF spectra for lysozyme-(βCD)$_n$ (A) before and (B) after dialysis in deionized water.

Figure 2.4 (A) MALDI-TOF spectra for RNase A-(βCD)$_n$ (B) MALDI-TOF spectra for RNase A-(βCD)$_n$ after dialysis in water.
2.2.6 Dynamic light scattering on lysozyme-(βCD)_n and aged lysozyme for protein size determination

Dynamic light scattering (DLS) is a technique used to measure the size distribution profile of small particles in suspension or polymers in solution. When the light source hits the particles, fluctuation in the scattering intensity is observed which the instrument detects.

Dynamic light scattering was conducted to measure the aggregation behavior of βCD-modified proteins and unmodified proteins, which were aged under same reaction conditions and stress factors including aging, freeze-drying, and dialysis. I discovered that dialysis in pure water was the major cause for complete aggregation of unmodified lysozyme, but βCD-modified proteins remained monomeric under the same conditions. After dialysis in pure water lysozyme-(βCD)_n showed only monomeric proteins with a hydrodynamic diameter of about 3.8 nm in the number and volume distributions (Figure 2.5A). Under the same conditions of dialysis in pure water, the entire sample of unmodified lysozyme formed aggregates, with a hydrodynamic diameter about 698 nm (Figure 2.5B).
Figure 2.5 Number (A and B) and volume (C and D) distribution of DLS showing two replicates each for 1 mg/mL of lysozyme-(βCD)$_n$ and lysozyme in PBS buffer pH 7.4 at 25 °C. Volume distribution showed only monomeric proteins for lysozyme A-(βCD)$_n$ (A) and only aggregates for aged lysozyme (B). Number distribution showed only monomeric proteins for lysozyme A-(βCD)$_n$ (C), and only aggregates for aged lysozyme (D). The aging conditions are: 1 mL solution of lysozyme (1 mg/mL) was prepared in 1× PBS buffer pH 8.2. The protein solution was then placed in a 28 °C incubator shaker for 72 h. The sample was then dialyzed in water overnight and then freeze dried to obtain a white solid.

The same results were observed for RNase A, where RNase A-(βCD)$_n$ showed only monomeric proteins in the number and volume distributions whereas, RNase A under the same conditions showed only aggregates in the number and volume distribution (Figure 2.6).
Figure 2.6 Volume (A and B) and Number (C and D) distribution of DLS showing two replicates for 1 mg/mL of RNase A-(βCD)$_n$ and aged RNase A in 1× PBS buffer pH 7.4 at 25 °C. Volume distribution showed only monomeric proteins for RNase A-(βCD)$_n$ (A) and only aggregates for aged RNase A (B). Number distribution showed only monomeric proteins for RNase A-(βCD)$_n$ (C), and only aggregates for aged RNase A (D). The aging conditions are: 1 mL solution of RNase A (1 mg/mL) was prepared in 1× PBS buffer pH 8.2. The protein solution was then placed in a 28 ºC incubator shaker for 72 h. The sample was then dialyzed in water overnight and then freeze dried to obtain a white solid.

2.2.7 Thermo-induced aggregation of unmodified lysozyme and lysozyme-(βCD)$_n$ using dynamic light scattering (DLS).

Past studies have shown that trypsin modified with α, β and γ CD are stable when heated to 70 °C, whereas native trypsin was totally inactivated at 60 °C. During thermo-induced aggregation studies of unmodified lysozyme and lysozyme-(βCD)$_n$ using DLS, it was observed that heating unmodified lysozyme (0.5 mg/mL in 1 × PBS buffer at pH 7.4) at 67 °C for 25 min resulted in complete aggregation, whereas only monomeric proteins were observed when lysozyme-(βCD)$_n$ was treated under the same conditions. While aggregates begin to appear after heating lysozyme-(βCD)$_n$ at 77 °C for 25 min, the sample contained mostly monomers.
Complete aggregation of lysozyme-(βCD)$_n$ was observed only when the same sample was heated at 87 °C for 25 min (Figure 2.7).

**Figure 2.7** Number distribution of DLS of unmodified lysozyme and lysozyme-(βCD)$_n$ (0.5 mg/mL in 1 × PBS, pH 7.4). Number distribution for unmodified lysozyme kept at (A) 57 ºC, (B) at 67 ºC for 25 min and for lysozyme-(βCD)$_n$ (C) at 67 ºC and (D) at 87 ºC for 25 min.

2.2.8 *Surface adsorption of lysozyme-(βCD)$_n$ and native lysozyme using surface plasmon resonance (SPR).*

I believe that aggregation and surface adsorption of proteins share some common mechanism. Thus, modification that results in preventing aggregation should also reduce or prevent surface adsorption. Surface plasmon resonance (SPR) is a widely used technique to measure the amount of specifically or non-specifically adsorbed protein on the surface of gold.$^{101-102}$ Surface plasmons are surface electromagnetic waves, which run parallel to the metal air/solution interface. These waves are very sensitive to the changes occurring at the interface, such as the adsorption of molecules or biomolecules on the metal surface. When proteins adsorb on the metal surface a decrease in the intensity of the reflected light is observed which the SPR instrument measures.
To measure the effect of βCD modification on surface adsorption of proteins, I used SPR to directly monitor the adsorption of modified and unmodified proteins on gold surfaces of 1 mm diameter spots (on a 16-spot microarray). The experiment was performed by first flowing phosphate buffered saline (1× PBS, pH 7.42) through the flow cell followed by flowing 0.5 mg/mL solutions of lysozyme-(βCD)_n and lysozyme in separate channels in the same buffer and then finally flowing PBS again (Figure 2.8A). The amount of protein which adsorbs irreversibly on gold was estimated by comparing the SPR intensity in pixel intensity units before and after the gold surface was exposed to the protein solutions. From the SPR data, the pixel intensity units (PIUs) was calculated by taking the average of all runs performed for lysozyme and lysozyme-(βCD)_n adsorptions. By assuming the adsorption exhibited by lysozyme to be 100%, the % monolayer (% ML) for lysozyme-(βCD)_n was calculated to be 6.75 ± 2.4% using the formula % ML = \( \frac{\Delta \text{PIU}_{\text{lysozyme-(βCD)_n}}}{\Delta \text{PIU}_{\text{lysozyme}}} \) × 100. Hence, it can be said that the lysozyme adsorption was reduced by ~93% after modification of lysozyme with βCD. This result proves lysozyme-(βCD)_n is stable against adsorption on gold surfaces. For RNase A, %ML for adsorbed RNase A-(βCD)_n on gold surface was found to be 21.26 ± 2.62% over eight replicates, corresponding to a reduction of adsorption by 80% (Figure 2.8 B). I noted that whereas the adsorption of lysozyme-(βCD)_n appeared to have reached equilibrium, adsorption of RNase A did not. This result suggests that, at least for RNase A, modification of proteins with βCD can reduce only the rate of protein adsorption.
Figure 2.8 Four replicates of surface plasmon resonance (SPR) of the adsorption of (A) lysozyme and lysozyme-(βCD)$_n$, (B) RNase A and RNase A-(βCD)$_n$, (0.5 mg/mL of protein in 1× PBS, pH 7.4). The SPR intensity (pixels) is plotted against time (seconds) as protein solutions are flowed over a SPR micro array.

2.2.9 Measuring activity of lysozyme-(βCD)$_n$ and native lysozyme using turbidimetric method

An enzymatic assay was used to determine the activity of lysozyme-(βCD)$_n$ and native lysozyme. The assay involves the use of Micrococcus lysodeikticus bacterial cells as a substrate for the enzyme lysozyme, where the enzyme is responsible for degrading the bacterium’s cell wall. A suspension of the M. lysodeikticus cells is mixed with a small amount of the protein. As the enzyme acts on the suspension of bacterial cells, the turbidity of the suspension and therefore the absorbance of the sample decreases. The rate of decrease in the turbidity is a relative measure of the activity of lysozyme present in the sample. The more the amount of lysozyme present in the sample, the faster is the reaction and the decrease in absorbance. The activity of lysozyme-(βCD)$_n$ was found to be lower than that of native lysozyme (Figure 2.9).

The turbidimetric method indicated that the activity of lysozyme-(βCD)$_n$ had reduced significantly compared to the native lysozyme. However, it is important to note that the activity
in general depends on individual proteins. Since a mixture of proteins with different degree of modifications is being examined, some modification may be of lower activity than others. The reduced activity of lysozyme-(βCD)$_n$ could also be due to the blocking of the lysozyme active site by the large βCD molecules. Masuda and co-workers have reported that if the positive charge on the lysine residues in lysozyme is modified, the lytic activity of lysozyme towards *M. lysodeikticus* tends to decrease. Hence, the blocking of the positive charges on the lysines by the βCD squaramide may also be responsible for the reduced activity of lysozyme-(βCD)$_n$.

![Figure 2.9](image)

**Figure 2.9** Decrease in optical density with time of suspensions of *Micrococcus Lysodeikticus* following addition of lysozyme and lysozyme-(βCD)$_n$. The absorbance of the sample was measured at 450 nm.
2.2.10 Purification of lysozyme-(βCD)$_n$ by protein crystallization

The science for aiding the crystallization of proteins (lysozyme), using a class of nonamphiphillic organic molecules that can form liquid crystal phases in water is reported in the literature and has been developed in Professor Luk’s laboratory.\textsuperscript{106} Based on the above knowledge 5’ disodium cromoglycate (5’ DSCG) and sodium chloride (NaCl) was used for crystallizing lysozyme and lysozyme-(βCD)$_n$. Lysozyme and lysozyme-(βCD)$_n$ were crystallized via vapor diffusion using the hanging drop method.\textsuperscript{107} The crystallization of lysozyme and lysozyme-(βCD)$_n$ was tried using sodium chloride (958 mM) and 5’DSCG (5.47 mM). Tetragonal shaped lysozyme crystals were observed in the drops containing 5’DSCG and NaCl (Figure 2.10). However lysozyme-(βCD)$_n$ crystals were not observed in the drops containing both 5’DSCG and NaCl. Crystallization of proteins starts with the formation of stable nuclei in a supersaturated protein sample. Formation of the stable nuclei is followed by coming together of more protein molecules to generate a macromolecular network. This network consists of channels, which are surrounded by “water bound” protein and possibly ions of the precipitant salt. The “bound water” forms an ordered hydration shell around the protein and situates itself adjacent to the protein molecules in the crystal lattice. This bound water participates in the extended hydrogen bonding that mediates indirect interaction with the protein molecules.\textsuperscript{108-109} Lysozyme-(βCD)$_n$ failed to show any crystal formation may be because a stable nuclei could not form due to the presence of the large βCD groups on the surface of the protein. Even if a stable nuclei can form in case of lysozyme-(βCD)$_n$, the large βCD groups on the surface of the protein can possibly inhibit the protein molecules from coming together to form a network of macromolecules for inducing crystallization. This further supports the hypothesis that covalent
grafting on proteins with βCD prevents proteins molecules from coming together to form assemblies or possibly aggregates.

Figure 2.10 Lysozyme crystals induced by (A) 5’DSCG (figure adapted from reference 106) and (B) NaCl additives in hanging droplets. Lysozyme-(βCD)_n crystals induced by (C) 5’DSCG and (D) NaCl additives in hanging droplets. Scale bar = 76 µm.

2.2.11 Modification of Bovine Serum Albumin (BSA) with βCD

Modification of bovine serum albumin (BSA) with βCD using compound 6 was also conducted. Bovine serum albumin (BSA) is a large protein with 583 amino acids and close to 50 lysine residues. Even though BSA has fifty lysine residues, MALDI-TOF analysis indicated that only one lysine residue in BSA was modified with compound 6 (Figure 2.11).
This could be due to less number of lysine residues being present on the surface of the protein BSA. The removal of unmodified BSA from BSA-(βCD)$_n$ was tried by dialyzing the sample containing the mixture of modified and unmodified BSA in deionized water. However, MALDI-TOF analysis indicated that dialysis in deionized water could not remove the unmodified BSA completely. Dynamic light scattering was conducted on the mixture containing modified and unmodified BSA (Figure 2.12). The number distribution indicated that the mixture of proteins was monomeric and the observed hydrodynamic diameter was consistent with the reported hydrodynamic diameter of ~7.0 nm for native BSA.
Figure 2.12 Number distribution of DLS of BSA-(βCD)$_n$ (1.0 mg/mL) in 1 × PBS, pH 7.4) at 25°C.

Surface plasmon resonance was also conducted to measure the surface adsorption of the mixture of BSA-(βCD)$_n$ and unmodified BSA versus unmodified BSA. Relative to the unmodified BSA, %ML for the mixture of BSA-(βCD)$_n$ and unmodified BSA was found to be 80.05 ± 6.5% (Figure 2.13). These results suggest that modification of BSA with βCD could only reduce surface adsorption by ~20% and that higher reduction in surface adsorption might be possible on achieving higher degree of modification of lysine residues with βCD.
Figure 2.13 Four replicates of surface plasmon resonance (SPR) of the adsorption of BSA and mixture of unmodified BSA and BSA-(βCD)$_n$, (0.5 mg/mL of protein in 1× PBS, pH 7.4). The SPR intensity (pixels) is plotted against (seconds) as protein solution is flowed over a SPR micro array.

2.2.12 Covalent modification of aniline grafted proteins with amino squarate based reagents

Francis and coworkers have developed an efficient protein modification reaction that couples new functionality to anilines under conditions mild enough to preserve the function of most biomolecules.$^{110}$ Luk and coworkers have synthesized heterocyclic amino squarate based compounds such as 13 (Scheme 2.4), to covalently modify aniline-grafted proteins under neutral conditions in aqueous solutions. N-hydroxysuccinimide (NHS) ester derivatives cannot be used for this covalent modification strategy since anilines do not react with NHS esters but only with squarates. Furthermore, while water does not readily hydrolyze squarates, it can easily hydrolyze N-hydroxysuccinimide (NHS) esters.
**Scheme 2.4** Heterocyclic amino squarate based compounds for protein modification

![Chemical structure of compound 13 and 6](image)

Motivated by the science developed by Francis and coworkers, lysine residues of proteins modified covalently with anilines were further conjugated with either cyclic amino squarates having pendant pyridinium group \((13)\) or with βCD amino squarate \((6)\) (Scheme 2.5). The alkyl group \((-R)\) of the pyridinium moiety in the cyclic amino squarates \((13)\) can act as a handle for conjugating ligands or small molecule based drugs, while the βCD conjugated to the squarate can serve as a host to carry and deliver small molecule drugs.

**Scheme 2.5** Sequential modification of proteins with NHS derivatized aniline \((14)\) and compound \(13\)†

![Sequential modification steps](image)

† Compounds \(13\) and \(14\) were synthesized by Dr DaWei Cui in Professor Y-Y Luk’s lab.
Proteins lysozyme and RNase were first grafted with aniline groups by modifying the lysine residues with NHS ester derivatized aniline (14) (Scheme 2.5, Step I). The aniline-grafted proteins were further conjugated with compound 13 (Scheme 2.5 Step II). Proteins (58 µM lysozyme and 60 µM RNase A) in 10 mM of PBS buffer at pH 8.2 were treated with 100 equivalents of compound 14,\textsuperscript{111} for about 5 h. The aqueous solutions of the reaction mixture were dialyzed in 1× PBS buffer, pH 6.0 to remove small molecule reagents. MALDI-TOF revealed that up to five lysine residues on lysozyme and up to eight lysine residues on RNase A were modified with aniline derivatives (Figure 2.9). The aniline modified proteins were further mixed with 50 equivalents of a cyclic amino squarate, 13 (in 200 µL of PBS, pH 6.0) for 16 h. For proteins presenting multiple copies of anilines, a few unmodified anilines were also present. All the reactions of the aniline modified proteins with compound 13 were conducted at pH 6.0 to keep the lysine residues in the proteins protonated and to ensure that the reaction happens only with the anilines bound covalently to the proteins. MALDI-TOF analysis showed mass adducts exhibiting high level of modification of multiple aniline groups with 13 for both RNase A and lysozyme (Figure 2.14).
Figure 2.14 MALDI-TOF spectra for (A) lysozyme modified with compound 14 and (B) aniline modified lysozyme further modified with compound 13 (C) RNase A modified with compound 14 and (D) aniline modified RNase A further modified with compound 13.

Before reacting aniline grafted proteins with compound 6, the reactivity of aniline with 6 in water was first tested and studied by UV-Vis and MALDI-TOF (Scheme 2.6). When aniline reacts with compound 6, compound 15 is formed as product and compound 16 is formed as a byproduct in the reaction. The reaction was monitored by measuring the UV absorption every 5
min for the first 30 min and then every hour for the next 7 h. A small change was observed in the UV-Vis absorption spectrum over 7 h, suggesting that the rate of reaction was very slow. MALDI-TOF confirmed that there was indeed product formation after reaction of aniline with 6 in water (Figure 2.15).

**Scheme 2.6** Reaction of aniline with compound 6

![Scheme 2.6](image)

**Figure 2.15** (A) Absorption spectra for the ligation of aniline with compound 6 in water at increment of 1 h from 1 to 7 h. Reaction conditions: Aniline (50 µM in H₂O), 6 (1 eq), rt (B) MALDI-TOF spectra for the reaction mixture of aniline with compound 6 in water after 24 h.
Table 2.2 lists the different conditions tried for the sequential conjugation of proteins lysozyme or RNaseA with compound 14, followed by compound 6. The different degrees of modifications for the aniline grafted proteins were characterized by MALDI-TOF. When the aniline activated lysozyme was modified with 6 at pH 8.2, up to three lysine residues got modified with βCD. However, at pH 8.2 it is not possible to determine whether the reaction is happening at the free lysines residues or with anilines bound covalently to the protein. Hence, all the reactions for the aniline activated proteins with compound 6 were conducted at pH 6.0 to keep the lysine residues on the protein protonated. When lysozyme was directly modified with compound 6 at pH 6.0, only negligible modification of the lysine residue with βCD was observed even after 72 h. When aniline activated lysozyme and RNase A were modified with compound 6, up to three (Figure 2.16) and five lysine residues were modified with βCD respectively after 72 h.

Table 2.2 Different conditions tried for sequential conjugation of lysozyme and RNase A with compound 14 followed by modification with compound 6.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Protein</th>
<th>Reaction with 14&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Reaction with 6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Time (h)</td>
<td>Modifications</td>
</tr>
<tr>
<td>1.</td>
<td>Lysozyme</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>2.</td>
<td>Lysozyme</td>
<td>N.A</td>
<td>N.A</td>
</tr>
<tr>
<td>3.</td>
<td>Lysozyme</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>4.</td>
<td>RNase A</td>
<td>5</td>
<td>7</td>
</tr>
</tbody>
</table>

<sup>a</sup> The reaction of proteins with compound 14 was conducted at pH 8.2.
Figure 2.16 MALDI-TOF spectra for lysozyme (46.67 µM) modified with aniline and further modified with compound 6. Modification of lysozyme-aniline with 6 was run at pH 6.04 for 72 h.

The covalent modification of aniline grafted proteins by oxidative method,\textsuperscript{110} appears to be slower than using amino squarate based reagents. While reaction with compound 13 requires 5 to 16 h, reaction with compound 6 requires 72 h. This strategy of incorporation of anilines into proteins and peptides provides a powerful solution to the problem of inability to achieve protein modification in a single location due to the absence of a reactive residue. Furthermore, this strategy can be used for site specific incorporation of poly(ethylene glycol) or other functional groups for protein-based drugs.
2.3 Conclusions and Perspectives

A water-driven chemistry that modifies proteins efficiently with cyclodextrin has been developed. Cyclodextrin-modified proteins exhibited significant reduction in aggregation and surface adsorption and increase in thermal stability as compared to unmodified proteins. These results are consistent with a unifying hypothesis that grafting proteins with molecules that stabilize protein structure can also reduce protein aggregation and surface adsorption. This chemistry also helps us to understand and validates future studies on protein stability and protein engineering at selected sites for design properties and activities. Furthermore, cyclodextrin can form noncovalent host-guest inclusion complexes with hydrophobic groups in water. Tethering the hydrophobic groups with small molecules or drugs enables combined drug therapeutics as well as targeted drug delivery. These studies along with artificial glycosylation and the mechanistic studies of aggregation and surface adsorption of modified proteins are the subjects of ongoing research in Professor Luk’s laboratory.

2.4 Experimental Section

Materials and Methods

Chemicals. Chicken egg white lysozyme (catalog number L6876) and RNase A (catalog number R6513) were obtained from Sigma Chemical Company (St. Louis, MO) and were used without further purification. Sinapinic acid was obtained from ProteoChem (Denver, CO). Phosphate buffered saline, 1× PBS (2.7 mM potassium chloride, 137 mM sodium chloride, 8 mM sodium phosphate dibasic, 1.48 mM potassium phosphate monobasic) was used for all protein reactions unless mentioned (pH of the prepared buffers was adjusted using 1N HCl and 1N NaOH). Cameo 3A Syringe Filters, Acetate 0.22 Micron, 3 mm, were obtained from GE
Water & Process Technologies (Trevose, PA). Disodium cromoglycate (5 ′DSCG ) was purchased from MP Biomedicals (Solon, OH). All other chemicals were obtained from Aldrich Chemicals (Milwaukee, WI). Deionized water used for preparing all buffers and solutions had resistivity of 18 MΩ cm (Millipore, Billerica, MA).

**Modification of lysozyme with βCD**

Compound 6 (4.68 mg, 50 eq.) was added to a solution of lysozyme/RNase A (1 mg/mL) prepared in 1× PBS pH 8.2 (reaction buffer). The reaction mixture was stirred in a 28 ºC incubator shaker for 72 h. The reaction mixture was dialyzed in 250 mL 1× PBS, pH 7.4. MALDI-TOF was conducted on the sample to see the number of modification on the protein. The modified protein sample was freeze dried to obtain a light yellow solid.

**Dialysis**

The protein samples were dialyzed in 1× PBS and/or in water (dialysis media), using Slide-A-Lyzer Dialysis Cassettes 7K MWCO purchased from Thermo Scientific (Rockford, IL). Before use, the Slide-A-Lyzer Dialysis Cassette was hydrated in the dialysis media for 5 min. Sample was injected in the cassette and placed in 250 mL of dialysis media to get rid of unreacted compound 6. The 1× PBS was changed twice every 2 h for the first 6 h and then left overnight. The same reaction mixture was further dialyzed in a new dialysis bag in 250 mL of deionized water to promote aggregation of unreacted protein. The deionized water was changed twice every 2 h and then left overnight. Reaction mixture was then passed through a syringe filter (0.22 µm) to eliminate aggregated protein.
Aging of Unmodified lysozyme/RNase A

A solution of lysozyme/RNase A (1 mg/mL) was prepared in 1× PBS buffer pH 8.2. The protein solution was then placed in a 28 °C incubator shaker for 72 h. The sample was then dialyzed in water overnight and then freeze-dried to obtain a white solid. This aging condition was the same as the reaction conditions for modifying proteins with βCD, except that compound 6, 7 or 8 were not added. The sample was not passed through the syringe filter.

Matrix Assisted Laser Desorption Ionization (MALDI)

A saturated solution of sinapinic acid (matrix) was prepared in a mixture acetonitrile and aqueous 0.1% TFA (1:2). Protein solution was mixed with aqueous 0.1% TFA solution and sinapinic acid solution in a 1:1:1 ratio and spotted on a MTP 384 target ground steel plate (Bruker Daltonics, Billerica, MA). Mass spectra were obtained on a MALDI TOF/TOF Mass spectrometer, Bruker AutoFlex III instrument (Bruker Daltonics, Billerica, MA). Protein mass was measured using a positive ion linear mode.

Dynamic Light Scattering (DLS)

Dynamic light scattering (DLS) experiment was performed on MALVERN zetasizer nanoseries (MALVERN Instruments Ltd. Worcestershire, UK) at a scattering angle of 173°. Protein solution (1 mg/mL) of lysozyme-(βCD)_n/RNase A-(βCD)_n and aged lysozyme/RNase A in 1× PBS buffer pH 7.4 at 25 °C was used for all DLS experiments.

Surface Plasmon Resonance (SPR)

All surface plasmon resonance (SPR) experiments were performed on a SPRimager® II Array System (GWC Technologies, Madison, WI). Protein solution (0.5 mg/mL) of lysozyme-(βCD)_n/RNase A-(βCD)_n and lysozyme/RNase A in 1× PBS buffer, pH 7.4 at 25 °C was used
for all SPR experiments. SpotReady™ 16 chip (GWC Technologies, Madison, WI) were mounted in a dual channel flow cell and experiment was performed by first flowing 1× PBS, pH 7.4 through the cell followed by a solution of lysozyme-(βCD)$_n$/RNase A-(βCD)$_n$ and unmodified lysozyme/RNase A in the same buffer and then finally flowing PBS. The amount of protein which adsorbs irreversibly on the gold was estimated by comparing the SPR intensity before and after the gold spots were exposed to the protein solution.

**Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)**

SDS-PAGE was conducted using BIO-RAD 7.5% Mini-PROTEAN® TGX™ Precast Gel (BIO-RAD, Hercules, CA). The protein sample (15 µL, 0.5 mg/mL or 0.05 mg/mL) was mixed with 5 µL of deionized water and 5 µL of 5 × SDS-loading dye (Tris-Cl 0.25 M, pH 6.8, 10% SDS, 50% glycerol, and 0.25% bromophenol blue). The mixture was heated to 90 ºC for 5 min and 20 µL of each sample was loaded on to the gel. A protein ladder, Precision Plus Dual Color from BIO-RAD (BIO-RAD, Hercules, CA) was added in one of the wells for molecular weight reference. Tris-glycine (1×) was used as the running buffer for gel electrophoresis. The gel was run at 100 V for 80 min, and was washed with deionized water three times and stained with GelCode blue stain reagent (Thermo Scientific, Rockford, IL) for 1 h and then destained in deionized water overnight. The gel image was taken on the Alpha DigiDoc (Alpha Innotech, San Leandro, CA).

**Enzymatic Assay of lysozyme-(βCD)$_n$ and lysozyme**

Micrococcus lysodeikticus cells suspension (0.015% (w/v)) was prepared in 66 mM potassium phosphate buffer, pH 6.24. A solution containing 300 units/mL of lysozyme in cold potassium phosphate buffer was also prepared. A 2.5 mL of *Micrococcus lysodeikticus* cells
suspension was mixed with 0.10 mL of lysozyme by inversion and the decrease in absorbance at 450 nm was recorded for approximately 5 min.

**Crystallization of Lysozyme and Lysozyme-(βCD)\textsubscript{n}**

Lysozyme and lysozyme-(βCD)\textsubscript{n} crystallized via vapor diffusion methods as hanging drops in 48 well VDX plates (Hamilton Research, Aliso Viejo, CA) with sealent. All solutions were prepared using 50 mM sodium acetate buffer, (pH 4.59) at 25 °C. A stock solution of lysozyme (75.0 mg/mL) and lysozyme-(βCD)\textsubscript{n} (1.0 mg/13.5 μL ~74.1 mg/mL) was prepared. Four reservoirs containing solutions of sodium chloride (958 mM) and 5’DSCG (5.47 mM) were separately prepared and sterile filtered through 0.22 μm syringe filters. 2.5 μL of lysozyme solution and lysozyme-(βCD)\textsubscript{n} were mixed with equal volume of reservoir solutions separately. The solution consisting of lysozyme and lysozyme-(βCD)\textsubscript{n} and reservoir solution was suspended on a siliconized coverslip as an inverted drop over a well containing 350 μL of the corresponding reservoir solution. Crystals were grown for a period of 2-4 d at rt and were viewed under the light microscope.

**General Information for Synthesis of βCD Amino Squarate Reagents**

Reagents and solvents were reagent grade and used as supplied unless otherwise mentioned. Solvents were distilled under reduced pressure using Buchi rotary evaporator below 40 °C. EMD Silica Gel 60 F254 precoated plates (0.25-mm thickness) were used for TLC. Flash column chromatography was performed using SILICYCLE, Silica-P Flash Silica Gel with 40-63 μ mesh size. Proton and carbon-13 NMR spectra were recorded on 300 MHz, \textsuperscript{1}H chemical shifts were reported in ppm relative to CDCl\textsubscript{3} δ 7.26 or D\textsubscript{2}O δ 4.80; \textsuperscript{13}C chemical shifts were reported relative to CDCl\textsubscript{3} δ 77.23, and High Resolution Mass Spectra (HRMS) was recorded by positive
ion electrospray on a Bruker 12 Tesla APEX –Qe FTICR-MS with Apollo II ion source, unless otherwise reported.

**Synthesis of compound 2**

![Diagram of compound 2]

Compound 1\(^{85}\) (1.40 g, 9.38 mmol) was dissolved in benzene (10 mL) and the solution was cooled to 0 °C. 3,5-Difluorophenol (2.44 g, 18.7 mmol) was added to the above solution in one portion, followed by slow addition of excess of TEA (2.09 g, 20.6 mmol) and the reaction mixture was stirred at room temperature overnight. The reaction mixture was filtered and the residue obtained was washed with benzene to yield a white solid. The solid obtained was purified using flash silica gel column (20% EtOAc in hexanes) to give compound 2 (0.80 g, 27%) as a white solid. \(R_f = 0.59\) (20% EtOAc in Hexanes). \(^1\)H NMR (300 MHz, CDCl\(_3\)): \(δ 6.90-6.88\) (m, 4H), 6.84-6.78 (m, 2H). \(^{13}\)C NMR (75 MHz, CDCl\(_3\)): \(δ 186.2, 182.4, 165.3, 165.1, 161.9, 161.7, 154.0, 153.6, 103.7, 103.5, 103.4, 103.3, 103.2, 102.9\). HRMS found = 339.0280 [M + H]\(^+\), calcd for \([C_{16}H_6F_4O_4 + H]\)^+ = 339.0275.

**Synthesis of compound 6**

![Diagram of compound 6]

Compound 5\(^{89}\) (10.0 mg, 0.008 mmol) was dissolved in water (0.5 mL). A solution of compound 2 (6.00 mg, 0.017 mmol) in DMF (0.1 mL) was added to the above solution drop wise. Addition
of the solution of compound 2 to the reaction mixture was accompanied by precipitation of 2 as solid. Acetone was added drop wise to the reaction mixture to dissolve the precipitate and the reaction mixture was stirred at room temperature for 2 h. Excess acetone was added to the reaction mixture to precipitate out the crude product. The precipitated solid was washed thoroughly with acetone and dried under high vacuum to yield the pure compound 6 (7.60 mg, 64%) as a pale yellow solid. $^1$H NMR (500 MHz, D$_2$O): $\delta$ 7.19-7.15 (t, 1H, $J_{H-H}$ =10.0 Hz), 6.79-6.64 (m, 2H), 5.13-5.01 (m 7H), 4.04-3.91 (m, 24H), 3.87-3.68 (m, 18H). Analyzed by MALDI-TOF and matched [M + Na]$^+$ at 1364.64. Expected Mass of sodium adduct: 1364.38.

**Synthesis of compound 3**

![Compound 3](image)

Compound 3 was synthesized and purified by following procedures similar to those used for the synthesis of compound 2. R$_f$ = 0.51 (20% EtOAc in hexane). $^1$H NMR (300 MHz, CDCl$_3$): $\delta$ 7.26-7.23 (m, 4H), 7.17-7.09 (m, 4H), $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 187.3, 182.9, 162.7, 159.4, 149.2, 120.6, 120.5, 117.2, 116.9. Low resolution GC-MS found = 302 [M]$^+$, calcd for [C$_{16}$H$_8$F$_2$O$_4$]$^+$ = 302.04.

**Synthesis of compound 7**

![Compound 7](image)

Compound 7 was synthesized and purified by following procedures similar to those used for the synthesis of compound 6. $^1$H NMR (500 MHz, D$_2$O): $\delta$ 7.38-7.35 (m, 2H), 7.22-7.20 (m, 2H),

*Synthesis of compound 8*

![Chemical structure of compound 8](image)

Compound 8 was synthesized and purified by following procedures similar to those used for the synthesis of 6. ^1^H NMR (500 MHz, D_2_O): δ 7.66-7.64 (m, 2H), 7.56-7.53 (m, 1H), 7.19-7.18 (m, 2H), 5.07-5.05 (m, 7H), 3.98-3.83 (m, 24H), 3.69-3.62 (m, 18H). Analyzed by MALDI and matched on [M + Na]^+ at 1329.24. Expected Mass of sodium adduct: 1329.14.

*Synthesis of compound 10*

![Chemical structure of compound 10](image)

To a solution of compound 9 in anhydrous DCM (3 mL), added NHS (0.068 g, 0.593 mmol) and DCC (0.122 g, 0.591 mmol) at 0 °C. Stirred reaction mixture at 0 °C for 30 min and allowed the reaction mixture to come to rt. The reaction mixture was stirred at room temperature overnight. The reaction mixture was filtered to remove dicyclohexyl urea (DCU). The filtrate was concentrated under vacuum to remove solvent and the crude product was obtained as colorless oil. The crude product was purified using flash silica gel column (gradient 10% → 40% EtOAc in Hexane) to obtain compound 10 as a white solid in quantitative yield. R_f = 0.44 (50% EtOAc in Hexane). ^1^H NMR (300 MHz, CDCl_3): δ 3.62 (s, 1H), 2.80 (s, 4H), 2.56 (t, J_H-H = 6,
2H), 2.27(t, J_{H-H} = 6, 2H), 1.73-1.58 (m, 2H), 1.37-1.29 (m, 4H). $^{13}$C NMR (75 MHz, CDCl$_3$): δ 174.3, 169.3, 168.7, 51.6, 34.0, 31.0, 28.9, 28.8, 28.6, 25.7, 24.9, 24.6. HRMS found = 322.1261 [M + Na]$^+$ calcd for [C$_{12}$H$_{18}$O$_6$+Na]$^+$ = 322.1258.

**Synthesis of compound 11**

![Chemical structure of compound 11](image)

To a solution of compound 10 (0.053 g, 0.177 mmol) in DMF (2 mL) added compound 5 (0.240 g, 0.211 mmol). The reaction mixture was stirred at rt for 48 h. The crude product was precipitated by acetone. The product was washed with acetone 2-3 times to obtain a compound 11 as white solid (0.090 g, 40%). $R_f$ = 0.46 (t-BuOH: EtOH: H$_2$O, 3.3:2.6:1). $^1$H NMR (300 MHz, DMSO): δ 7.56 (s, 1H), 5.77-5.64 (m, 14H), 4.80 (s, 1H), 4.45-4.42 (m, 6H), 3.64-3.60 (br m, 19H), 3.56 (s, 3H), 3.34-3.29 (br m, 12H), 2.25 (t, J$_{H-H}$ = 9, 2H), 2.23-2.063 (m, 2H), 1.47-1.43 (m, 3H), 1.21 (s, 4H). HRMS found = 1340.4831 [M + Na]$^+$ calcd for [C$_{12}$H$_{18}$O$_6$+Na]$^+$ = 1340.4849.

**Synthesis of compound 12**

![Chemical structure of compound 12](image)

To a solution of compound 11 in DMSO (60 µL), added 1N NaOH (440 µL). Stirred the reaction mixture at rt for 3 h. Neutralized the reaction mixture with 1N HCl. Product was precipitated out using acetone. The product was washed with acetone 2-3 times to obtain
compound **12** as white solid. \( R_f = 0.40 \) (t-BuOH: EtOH: H\(_2\)O, 3.3:2.6:1). \(^1\)H NMR (300 MHz, DMSO): \( \delta \) 7.72 (s, 1H), 5.78 (br m, 14H), 4.87-4.79 (m, 6H), 3.6 (br m, 37H), 2.50-2.47 (m, 4H), 2.05-1.90 (m, 3H), 1.40-1.18 (m, 4H). HRMS found = 1326.4738 [M + Na] \(^+\) calcd for [C\(_{12}\)H\(_{18}\)O\(_6\)+Na] \(^+\) = 1326.4692.
Chapter 3

Inducing Folding in an Inherently Unstructured Protein by Covalent Modification with Squarate Derivatives

Summary

Motivated by the fact that covalent modification with βCD can stabilize the folded structure of proteins, the ability of βCD to induce folding in entirely unfolded proteins was explored. Collaborating with Professor Jin Montclare at New York University-Polytechnic, I used cartilage oligomeric matrix protein (COMP) in which a single amino acid mutation of leucine 44 to alanine in the coiled coil domain of COMP (COMPcc) caused a complete loss of the α-helical structure. Chemical modification of the lysine residues of COMPcc mutant L44A with β-cyclodextrins (βCD) via a squaramide linkage folded the protein into a α-helix. Interestingly, modification of L44A with methyl squarate alone also folded the protein into a α-helix. Small molecules such as curcumin exhibited an increase in the fluorescence signal upon binding to L44A-(βCD)_n, which likely suggests that the folding stability of the protein is either enhanced or preserved. These results suggest that the chemical modification of L44A with βCD or methyl squarate changes the energy landscape of the protein and creates a new global minimum.

3.1 Background and Significance

3.1.1 Protein folding

Protein folding is a process in which an unfolded protein having a distorted structure, folds into a well-defined tertiary structure with a specific function. Understanding the science of how proteins fold from a random coil to a three dimensional structure is highly desired due to the large implications of this process in the field of biology, biotechnology and human diseases. Often folding involves the establishment of secondary structures, particularly alpha helices and beta sheets followed by establishment of tertiary structures. The de novo design of peptides...
and proteins has recently emerged as an approach for investigating protein structure and function.\textsuperscript{115}

\textit{3.1.2 Inducing peptide and protein folding}

A variety of designed helical peptides and proteins provide great model systems to study the different interactions that stabilize the tertiary structure of proteins. A protein can fold spontaneously during or after biosynthesis or can be forced to fold using artificial chaperones or under varying salt concentrations and temperatures. Chaperones prevent misfolding or aggregation which may occur due to changes in the cellular environment.\textsuperscript{116}

“Peptide stapling” is a strategy in which an all-hydrocarbon cross-link is generated within natural peptides. It is a new approach for dramatically increasing the helicity and potency of peptides. This strategy of stapling peptides has largely been applied to B-cell lymphoma 2 (Bcl-2) family proteins which are dominant activators and inhibitors of apoptosis. It has been demonstrated that stapled helices are viable drug candidates for a wide variety of human cancers.\textsuperscript{117-118} Hodges and coworkers have reported the synthesis for three peptide analogs of the helix-loop-helix Ca\textsuperscript{2+} binding unit, which have a sequence similar to rabbit skeletal troponin C site (TnC). They observed that addition of Ca\textsuperscript{2+} induced conformational changes and induced secondary structure in TnC.\textsuperscript{119} Uversky, Fink and coworkers have studied the anion-induced folding of \textit{Staphylococcal Nuclease}.\textsuperscript{120}

The amyloid β-protein (Aβ) is a peptide, which is required for the activation of kinase enzymes, protection against oxidative stress, regulation of cholesterol transport, transcription and anti-microbial activity. Besides these important functions in the body, the Aβ fibrils are also the principal protein component in Alzheimer’s disease (AD). It is believed that a small number of stacked, extended, ribbon-like β-sheets form the Aβ fibril.\textsuperscript{121} Teplow and co-workers have used
circular dichroism to show that a transient increase in helicity is observed immediately before the appearance of β-structure. Understanding the oligomerization of Aβ is important for developing strategies to block this process Aβ fibril formation and is of significant therapeutic value for treating AD.\textsuperscript{122-123}

Carbohydrates and their derivatives are known to stabilize protein folding. For example, artificial chaperones such as cyclodextrins are known to assist in protein folding and refolding. Gellman and coworkers have used cyclodextrins as artificial chaperones that can refold proteins, likely by stripping the detergents that are associated with the denatured proteins.\textsuperscript{46-47, 53} Here, the disruption of hydrophobic interaction between proteins or between protein and surfactant prevents aggregation and supports protein refolding.

3.1.3 Cartilage oligomeric matrix protein (COMP) - function and uses

Cartilage oligomeric matrix protein (COMP) is glycogenic protein present in the cartilage, tendons, ligaments and osteoblasts and has a pentameric bouquet like structure with a molecular mass of 524 kDa. The COMP is comprised of various domains and is assembled into a homopentamer via a N-terminal coiled-coil domain (COMPcc). The coiled-coil domain is a common structural motif found in proteins and consists of two to five α-helices wrapped around each other.\textsuperscript{124} The pentameric bundle of COMP is built of 230 amino acid residues. Five chains, each with 46 residues, form a parallel coiled-coil structure.\textsuperscript{125} The coiled-coil domains in proteins are capable of selectively storing and delivering small molecules. COMPcc has the ability to store and bind molecules such as vitamin D, all-\textit{trans} retinol (ATR), retinoic acid, curcumin, benzene and cyclohexane (Figure 3.1).\textsuperscript{126-127} One major role of COMP is storage and delivery of regulatory molecules in bone metabolism.
Figure 3.1 (A) Structure and amino acid sequence of the monomer of COMPcc. (B) Chemical structures of small molecules that can bind to the hydrophobic cavity of COMPcc.

3.1.4 Studying folding of an inherently unfolded protein by covalent modification with β-cyclodextrins

Montclare and coworkers have reported that the N-terminal aliphatic residues L37, L44, V47, and L51 of the coiled-coil region of COMP are necessary for maintaining its helical content, pentameric structure, stability and small molecule binding ability.\textsuperscript{127} The original COMPcc sequence has two cysteine residues near the C-terminus, which form disulfide bridges to stabilize the pentameric structure. In order to avoid complications due to oxidation, the N-terminal cysteines of COMPcc were mutated to serines to produce the mutant COMPcc\textsuperscript{s}. The L44A is a mutant of COMPcc\textsuperscript{s}, where leucine 44 is mutated to alanine (Figure 3.2). The replacement of leucine 44 with alanine results in a significant loss of the α-helicity of L44A.\textsuperscript{127} Montclare and coworkers suggest that the leucine residue probably interacts with residues at the i+3 positions to stabilize the α-helical structure. Hence, mutation of leucine 44 to alanine prevents α-helix formation. Recently, Luk and coworkers have reported that covalent grafting of proteins with β-cyclodextrins (βCD) can prevent aggregation, surface adsorption and increase thermal stability of proteins.\textsuperscript{128} In this chapter of the dissertation, covalent modification of the COMPcc\textsuperscript{s} mutant
L44A with βCD using water driven ligations is presented. The ability of βCD to promote folding and induce α-helical structure into the completely unstructured protein L44A is demonstrated.

Figure 3.2 (A) Amino acid sequence of COMPcc\textsuperscript{8} monomer, where the mutation site leucine 44 is indicated in red. (B) Amino acid sequence of COMPcc\textsuperscript{8} monomer, where leucine 44 is converted to alanine 44. All lysine residues are represented in bold black font.

3.1.5 Protein energy profile

The native protein conformation is a stable three-dimensional structure that determines its biological function. Denatured proteins exist in a set of partially folded states. The folding process is described using a free-energy funnel where the unfolded states have high energy and entropy. As protein folding begins, the free energy drops and the number of conformational states decreases. The metastable protein can be trapped in the local minima’s along the way, which can slow down its progress towards the free energy minimum. As the protein folds, the narrowing of the energy funnel indicates the decrease in the number of conformations exhibited by the protein and the bottom of the well is referred to as the global minimum.\textsuperscript{129-130} I believe L44A does not have an existing global minimum, but covalent modification of L44A with βCD will create a new global minimum, lowering the entropy and energy of the protein (Figure 3.3). This mechanism of induced folding in unstructured proteins has potential to introduce new functions in the proteins and make not-so-useful proteins useful.
3.2 Results and Discussion

3.2.1 MALDI-TOF confirms modification of L44A with βCD

The lysine residues in L44A were covalently modified with βCD using the βCD amino squarate compound 1 (Scheme 3.1). The hydrolytic stability and reactivity of amino squarates makes these molecules ideal for modifying proteins with βCD.

Scheme 3.1 L44A modification with βCD

I used matrix assisted laser desorption ionization time of flight (MALDI-TOF) spectroscopy to determine the degree of modification of L44A with βCD. I observed that up to three lysine residues in L44A were successfully conjugated to βCD using compound 1 (Figure 3.3A).
Nevertheless, some unreacted L44A was still observed by MALDI-TOF. After the reaction of L44A with compound 1, the unreacted reagents and other small molecules were removed by dialysis in phosphate buffered saline (PBS), pH 7.4. Dialyzing the sample subsequently in deionized water resulted in significant reduction in the mass peak of unmodified L44A (Figure 3.4B). This suggests that unmodified L44A formed aggregates on dialysis in water, which were removed by filtration, whereas the βCD modification stabilized the protein structure and prevented the modified proteins from forming aggregates. It is interesting to note that prior to dialysis in water, when the L44A-(βCD)$_n$ is in PBS buffer, the intensity of the mass peak in the MALDI-TOF spectrum corresponding to the first modification is higher than that of the second modification. However, after dialysis in water and subsequent filtration of the protein solution through a 0.2 µm syringe filter, the intensity of the mass peak in the MALDI-TOF spectrum corresponding to the second modification is considerably higher than that of the first modification. The above observation suggests that modification of L44A with at least two βCD groups is essential to prevent aggregation. Only the unmodified L44A and some of the L44-(βCD)$_1$ formed aggregates during the stress of dialysis in deionized water.
Figure 3.4 MALDI-TOF spectra for L44A-(βCD)$_n$ (A) after dialysis in deionized water without filtering through 0.2 µm pore filter (B) after dialysis in deionized water with filtering through 0.2 µm pore filter.

3.2.2 Circular dichroism indicates folding of L44A-(βCD)$_n$

Circular dichroism was conducted to determine if covalent modification of L44A with βCD can induce secondary structure in L44A. Alpha-helix, beta-sheet, and random coil structures each give rise to a characteristic shape and magnitude in the circular dichroism spectrum.\textsuperscript{131} While L44wt shows the characteristic shape of α-helix, L44A has the characteristic shape of random coil. Circular dichroism was conducted on samples containing a mixture of unmodified L44A+L44A-(βCD)$_n$ and purified L44A-(βCD)$_n$ alone (where n can range from 1 to 4). The results from circular dichroism on both the samples indicated presence of α-helical structure. The lower signal intensity in purified L44A-(βCD)$_n$ may be due to difference in the concentration of the two proteins (Figure 3.5). I believe that folding is induced in L44A-(βCD)$_n$ because the multiple hydroxyl groups on βCD act as a template to generate a structured solvation shell as a result of low dynamic hydrogen bonding between the hydroxyl groups on βCD and the water.
molecules. Another possibility for induced folding in L44A-βCD$_n$ could be non-covalent stapling resulting from intra-molecular interaction between hydrophobic amino acid residues and βCD.

Figure 3.5 Circular dichroism spectra for L44wt, L44A, L44A-βCD$_n$, and L44A-βCD$_n$ purified. Circular dichroism was conducted by Jasmin Hume in Professor Jin Montclare’s laboratory at Polytechnic Institute of New York University.

To determine the significance of covalent modification of L44A with βCD to induce folding or any structural change, control experiments were conducted where free βCD was added to L44A (29 μM) and L44wt protein (29 μM) (Figure 3.6 and 3.7). L44A and L44wt were treated with increasing concentrations of free βCD and circular dichroism was conducted on the protein samples. In general, addition of βCD to L44wt and L44A resulted in a decrease and increase in the circular dichroism signal with increasing concentrations of added βCD respectively, but did
not bring any change to the protein structure. However covalent conjugation of βCD to L44A induced α-helical structure in the protein.

**Figure 3.6** Circular dichroism spectra for L44A treated with free βCD. Circular dichroism was conducted by Jasmin Hume in Professor Jin Montclare’s laboratory at Polytechnic Institute of New York University.

**Figure 3.7** Circular dichroism spectra for L44wt treated with free βCD. (Circular dichroism was conducted by Jasmin Hume in Professor Jin Montclare’s laboratory at Polytechnic Institute of New York University).
3.2.3 Dynamic light scattering to measure the size of L44wt, L44A and L44-\((\beta CD)_n\)

Dynamic light scattering (DLS) was used to measure the hydrodynamic diameter of L44 wt, L44A and L44A-\((\beta CD)_n\). The concentration of all protein solutions used for the study was \(\sim29\) µM. The hydrodynamic diameter of unmodified L44A and L44wt was found to be 18.4 nm 20.4 nm respectively (Figure 3.8). Native lysozyme (14 kDa) which has a molecular weight that is twice that of the molecular weight of L44A (6.9 kDa), has a hydrodynamic diameter of only 4.4 nm. This suggests that L44A has a hydrodynamic diameter which is \(~4\) times larger than that of lysozyme. The hydrodynamic diameter of lysozyme-\((\beta CD)_n\) (\(n\) can range from 1 to 6) and RNase A-\((\beta CD)_n\) (\(n\) can range from 1 to 10) has been reported to be similar to unmodified lysozyme and RNase A respectively.\(^\text{128}\) I believe the relatively large hydrodynamic diameter exhibited by L44A may be due to the lack of a defined structure. It is surprising that modification of L44A with \(\beta CD\) causes such a significant (almost double) increase in the hydrodynamic diameter for L44A-\((\beta CD)_n\). I believe the large hydrodynamic diameter exhibited by L44A-\((\beta CD)_n\) is likely due to formation of a pentameric assembly formed by the folded protein. One likely mechanism for assembly formation by L44A-\((\beta CD)_n\) is non-covalent inter-molecular interaction of the pendant \(\beta CD\) groups with amino acid residues on the adjacent protein.
Figure 3.8 Number distribution of DLS showing two replicates each of (A) L44A (B) L44wt (C) L44A-(βCD)$_n$ and (D) lysozyme in PBS buffer pH 7.4 at 25 °C.

3.2.4 Influence of groups like acetyl, polyethylene glycol (PEG) and methyl squarate on folding of L44A

Modification of the lysine residues in L44A with βCD results in removal of charge due to modification of the lysine residues, addition of a squaramide and βCD moiety. In order to determine which of the above factors are responsible for inducing folding in L44A, I modified L44A with compounds 2, 3 and 4 (Scheme 3.2).

Scheme 3.2 Molecules synthesized to study the influence of groups like acetyl, polyethylene glycol (PEG) and methyl squarate on folding of L44A
Compound 2 was synthesized and used to covalently modify the lysine residues with an acetyl group, thereby removing the positive charge on the lysine residues (Figure 3.8A). Compound 3 was synthesized to covalently modify the lysine residues in L44A, to study the effect of methyl squarate moiety alone on the folding of L44A (Figure 3.8B). Since its development in the 1970s, modification of proteins with poly(ethylene) glycol (PEG) has remained a major approach to reduce protein aggregation.\textsuperscript{132-133} Surfaces functionalized with PEG are known to resist protein adsorption.\textsuperscript{134-135} The resistance of PEG towards protein adsorption on surfaces is believed to be due to the steric repulsion effect, where the PEG chains prevent the protein from reaching the surface for adsorption.\textsuperscript{136} Modification of proteins with PEG is known to increase the stability of proteins and also prevent protein aggregation.\textsuperscript{18} Hence, the lysine residues in L44A were covalently modified with PEG using compound 4 (Figure 3.9C), to study the influence of the pendant PEG groups on folding and stability against aggregation of L44A.
Matrix assisted laser desorption ionization time of flight (MALDI-TOF) indicated that up to five acetyl groups were grafted covalently on to L44A. Since L44A has only three lysine residues, it seems that amino acids other than lysine also reacted with compound 2 and hence up to five modifications were observed (Figure 3.10). Mass spectra also confirmed that up to three lysine residues in L44A were modified with compounds 3 (Figure 3.11) and compound 4 (Figure 3.12). However, unmodified L44A was observed along with the modified proteins in all the three modification experiments.
Figure 3.10 MALDI-TOF spectra for L44A-(acetyl)$_n$ indicated that up to five acetyl groups were conjugated to L44A.

Figure 3.11 MALDI-TOF spectra for (A) L44A-(methyl squarate)$_n$ indicated that up to three methyl squarates were conjugated to the lysine residues of L44A.
Figure 3.12 MALDI-TOF spectra for L44A-(PEG)$_n$ indicated up to three PEG chains were conjugated to the lysine residue of L44A.

To measure the ability of compounds 2, 3 and 4 to induce folding in L44A, circular dichroism was conducted on the samples containing a mixture of unmodified L44A + L44A-(acetyl)$_n$, unmodified L44A + L44A-(methyl squarate)$_n$, and unmodified L44A + L44A-(PEG)$_n$ (Figure 3.13). Interestingly, L44A modified with methyl squarate exhibited α-helix formation. However, L44A modified with acetyl and PEG groups did not exhibit any folding. Hence, this suggests that modifying the unfolded protein sequence L44A with βCD or methyl squarate alone causes the protein to fold into an α-helix, while modifying the same unfolded sequence with PEG does not induce any folding.

The folding exhibited by L44A-(methyl squarate)$_n$ eliminates the possibility that the induced folding in L44A-(βCD)$_n$ might be because of the interaction of βCD with the hydrophobic amino acids in L44A. I believe, due to the rigid structure and presence of the two carbonyl groups in the
methyl squarate, it can strongly hydrogen bond with surrounding water molecules and hence induce folding as hypothesized previously for L44A-(βCD)ₙ.

Figure 3.13 Circular dichroism spectra for L44A modified with compounds 1, 2, 3 and 4. Concentration of L44A-(methyl sq)ₙ = 29.65 µM, L44A-(acetyl)ₙ = 8.27 µM L44A-(PEG)ₙ = 23.45 µM. Circular dichroism was conducted by Jasmin Hume in Professor Jin Montclare’s laboratory at Polytechnic Institute of New York University.

Dynamic light scattering was used to determine the size of L44A-(acetyl)ₙ, L44A-(methyl squarate)ₙ, and L44A-(PEG)ₙ. Keeping L44wt as standard for explaining extent of assembly formation, I believe that L44wt is folded and exists as a pentameric assembly and exhibits a hydrodynamic diameter of 18.4 nm. L44A exhibits a hydrodynamic diameter of 21.7 nm suggesting mild denaturation and aggregation. L44A-(βCD)ₙ which showed folding into an α-helix exhibited a hydrodynamic diameter of 42.9 nm, which was almost twice the diameter of L44wt, indicating folding and formation of a high order assembly. Surprisingly, L44A-(methyl squarate)ₙ which was able to induce folding in L44A, exhibited a hydrodynamic diameter of 152
nm, which was much higher than the hydrodynamic diameter of L44A-(βCD)$_n$. This observation suggests that covalent modification of L44A with methyl squarate results in the formation of a multi order assembly. L44A-(PEG)$_n$ showed a hydrodynamic diameter of 3.7 nm, which is closer to the average diameter of COMPcc monomer. The small hydrodynamic diameter exhibited by L44A-(PEG)$_n$ could be due to the steric repulsions between the PEG chains, which are able to keep L44A-(PEG)$_n$ in monomeric form. L44A-(acetyl)$_n$ showed very high hydrodynamic diameter suggesting complete denaturation and aggregation (Table 3.1).

**Table 3.1** Secondary structure and hydrodynamic diameter of modified L44A with different chemical agents.

<table>
<thead>
<tr>
<th>Protein</th>
<th>*Diameter (nm)</th>
<th>Folding$^b$ vs Denaturation</th>
<th>Assembly vs Denaturation</th>
</tr>
</thead>
<tbody>
<tr>
<td>L44wt</td>
<td>18.4</td>
<td>α-helix</td>
<td>assembly</td>
</tr>
<tr>
<td>L44A</td>
<td>21.7</td>
<td>mild denaturation</td>
<td>mild aggregation</td>
</tr>
<tr>
<td>L44A-(βCD)$_n$</td>
<td>42.9</td>
<td>α-helix</td>
<td>high order assembly</td>
</tr>
<tr>
<td>L44A-(methyl squarate)$_n$</td>
<td>151.9</td>
<td>α-helix</td>
<td>multi order assembly</td>
</tr>
<tr>
<td>L44A-(PEG)$_n$</td>
<td>3.7</td>
<td>denaturation</td>
<td>no aggregation</td>
</tr>
<tr>
<td>L44A-(acetyl)$_n$</td>
<td>2169</td>
<td>denaturation</td>
<td>aggregation</td>
</tr>
</tbody>
</table>

$^a$ determined by dynamic light scattering (DLS)

$^b$ determined by circular dichroism

3.2.5 Small molecule binding by L44A-(βCD)$_n$

The coiled coil domain of COMPcc is capable of binding small molecules such as vitamin D, *trans*-retinol (ATR), retinoic acid and curcumin. Since the covalent modification of L44A with βCD can induce folding in L44A, the ability of L44A-(βCD)$_n$ to bind such small molecules was also studied. The binding of curcumin by L44A-(βCD)$_n$ was monitored using fluorescence
spectroscopy. We discovered that binding of curcumin with L44A-(βCD)ₙ was associated with enhancement of fluorescence suggesting strong binding of curcumin with L44A-(βCD)ₙ.

Past Studies show that curcumin and βCD exhibit weak host guest binding

Tang and coworkers have studied the supramolecular interaction of curcumin and βCD by spectrophotometry. They discovered that βCD binds with curcumin to form a 2:1 host-guest complex with a formation constant of $5.53 \times 10^5$ mol$^{-2}$L$^2$. The formation constant was determined at pH 2.4 in water/acetonitrile medium. Maximum absorbance was observed at 431 nm when the concentration of curcumin and βCD was $1 \times 10^{-5}$ mol/L and $5.0 \times 10^{-3}$ mol/L respectively.

Wagner and coworkers have studied fluorescence enhancement for curcumin upon inclusion into cyclodextrins or modified cyclodextrins. The fluorescence enhancement was studied for all CDs (10 mM) with $1 \times 10^{-5}$ M of curcumin in 1% methanol/water system. They found that for both parent and modified CDs, βCD gave a larger enhancement in the fluorescence signal than αCD. However, γCD showed a higher fluorescence enhancement compared to βCD. The authors discovered that curcumin forms 2:1 host-guest inclusion complexes with these CDs and the strongest binding was exhibited by αCD ($K_1= 3300$ M$^{-1}$, $K_2=5.4$ M$^{-1}$ and $K_1K_2= 18000$ M$^{-2}$) and hydroxypropylated (HP) βCD ($K_1= 3400$M$^{-1}$, $K_2=120$ M$^{-1}$ and $K_1K_2= 41000$ M$^{-2}$). However, βCD alone exhibited weaker binding constants ($K_1= 280$ M$^{-1}$, $K_2=6.6$ M$^{-1}$ and $K_1K_2= 1800$ M$^{-2}$).

Swaroop and coworkers have characterized the 1:1 inclusion complex of curcumin and βCD using absorption and fluorescence spectroscopy. They discovered that the fluorescence intensity increased with increase in concentration of βCD. At 10 mM βCD and 10 µM of curcumin, a well-defined blue shifted band was observed at $\lambda_{max} = 520$ nm. The binding constant for the 1:1 complex was found to be $5.2 \times 10^3$ M$^{-1}$. 
Binding study using L44A-(βCD)$_n$ suggests enhanced binding of curcumin in the hydrophobic pocket of the pentameric assembly and not in the βCD

We studied the binding of curcumin (145 µM) with L44A-(βCD)$_n$ (29 µM) and discovered that curcumin exhibited an enhancement in fluorescence upon binding with L44A-(βCD)$_n$ (Figure 3.14). This suggests that curcumin binds within the hydrophobic pore of the pentameric assembly formed by L44A-(βCD)$_n$. We observed no enhancement in the fluorescence signal when 58 µM of βCD was incubated with 145 µM of curcumin. This suggests that there is no potential binding between βCD and curcumin at the tested concentrations. We note that buffer (phosphate buffer) and L44A alone do not influence the fluorescence signal of curcumin. One major question about modifying proteins with βCD is whether their folding stability or activities is preserved, enhanced or reduced. The increase in fluorescence signal observed during the binding study of L44A-(βCD)$_n$ with curcumin likely suggests that the folding stability of the protein is either enhanced or preserved. In chapter 2 of this dissertation, I have demonstrated that modifying lysozyme with βCD prevents aggregation, surface adsorption and increases thermal stability but reduces the activity of lysozyme-(βCD)$_n$. Here, we discover that modifying L44A with βCD not only folds the random coil sequence but also promotes pentameric assembly formation accompanied by enhanced binding to small molecules such as curcumin by L44A-(βCD)$_n$. The enhancement in fluorescence could also be due to the higher effective concentration of L44A-(βCD)$_n$, which likely results in binding of more curcumin molecules. Given that the concentration of the L44A-(βCD)$_n$ and L44A is same, the enhancement in absorbance is probably due to enhanced folding stability of L44A-(βCD)$_n$. 
Figure 3.14 Fluorescence spectra for curcumin binding in the hydrophobic pocket of L44A-(βCD)$_n$. PB in the fluorescence spectra stands for phosphate buffer (pH 8.0). (Fluorescence study has been conducted by Jasmin Hume in Professor Jin Montclare’s laboratory at Polytechnic Institute of New York University).

3.2.6 Breaking the assembly formation in L44A-(βCD)$_n$

The mixture of unmodified L44A+L44A-(PEG)$_n$ exhibited a hydrodynamic diameter of 3.7 nm, which is believed to be the expected hydrodynamic diameter of L44A and L44wt. In order to break the assembly formation of L44A-(βCD)$_n$, an adamantane tetra(ethylene glycol) (TEG) conjugate (adamantane-TEG) was synthesized using literature reported procedures. The adamantane-TEG can potentially form a non-covalent inclusion complex with the βCD attached covalently to the lysine residues of L44A and render the protein monomeric due to steric repulsions between the TEG chains (Figure 3.15).
Figure 3.15 Strategy for preventing assembly formation of L44A-(βCD)$_n$. (A) Assembly formation in L44A-(βCD)$_n$ due to hydrogen bonding between βCD molecules (B) Repulsions between TEG chains conjugated to adamantane, which forms inclusion complex with covalently grafted βCD molecules, keeps the proteins monomeric.

Several methods were tried to form an inclusion complex between βCD covalently conjugated with L44A and adamantane-TEG. One of the methods involved the treatment of the mixture of L44A+ L44A-(βCD)$_n$ (29 µM) with excess amount of adamantane-TEG (100 eq) for 24 h. Dynamic light scattering indicated that the size of the mixture of L44A+ L44A-(βCD)$_n$ was ~ 88 nm. Addition of adamantane-TEG reduced the size of L44A+ L44A-(βCD)$_n$ to ~ 38 nm. This result suggests that the large assembly size of the L44A+ L44A-(βCD)$_n$ was broken due to formation of inclusion complex between the βCD and the adamantane-TEG, resulting in the decrease in the hydrodynamic diameter by almost half.
3.2.7 Purification of L44A-(βCD)$_n$ using reverse phase HPLC

In order to study how many βCD moieties are required for inducing folding in L44A, it is important to separate out proteins with different degree of modifications and study their folding individually using circular dichroism. The MALDI-TOF results suggested that at least two βCD moieties are essential for preventing aggregation of L44A-(βCD)$_n$. I further believe that at least two βCD moieties may also be essential for stabilizing the folded structure of L44A-(βCD)$_n$.

The separation of different degree of modification in L44A-(βCD)$_n$ was attempted using reverse phase high performance liquid chromatography (HPLC) using a C18 column. The L44A-(βCD)$_1$ and L44A-(βCD)$_2$ were successfully separated out and the different elution times for L44A-(βCD)$_1$ and L44A-(βCD)$_2$ from the C18 HPLC column are shown in table 3.2.

Table 3.2 Elution times for L44A-(βCD)$_1$ and L44A-(βCD)$_2$ from a C18 HPLC column (eluents: water-0.1%TFA/acetonitrile-water-0.1% TFA).

<table>
<thead>
<tr>
<th>Entry</th>
<th>Protein</th>
<th>Elution Time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>L44A-(βCD)$_2$</td>
<td>27.7</td>
</tr>
<tr>
<td>2.</td>
<td>L44A-(βCD)$_1$</td>
<td>28.5</td>
</tr>
</tbody>
</table>

After elution from HPLC, MALDI-TOF was conducted on the eluted fractions to confirm the extent of purification of L44A-(βCD)$_n$. MALDI-TOF indicated that the L44A-(βCD)$_1$ was almost 100% pure whereas the L44A-(βCD)$_2$ had small amounts of L44A-(βCD)$_1$. 
3.2.8 Effect on thermal stability of lysozyme by modification with compounds 2 and 3

Luk and coworkers have reported that covalent modification of lysozyme with βCD increases the thermal stability of the protein against aggregation. Since covalent modification of L44A with methyl squarate alone induced α-helix structure in L44A, I further explored how the thermal stability of lysozyme against aggregation would be affected, if the lysine residues in lysozyme were covalently modified with acetyl and methyl squarate groups.

Gorgani and coworkers have reported that acetylation of the lysine residues in lysozyme promotes amyloid formation resulting in formation of more pronounced fibrils. Hence the proposed study can help us determine whether the positive charge on the lysine residues in lysozyme is essential for maintaining the thermal stability against aggregation of lysozyme.
The positive charges on lysozyme were blocked by acetylating the lysine residues using compound 2 and squarylating the lysine residues using compound 3 at pH 8.2 for 72 h. MALDI-TOF showed that up to four acetyl (Figure 3.17A) and four methyl squarate groups (Figure 3.17B) were conjugated to the lysine residue of lysozyme.

**Figure 3.17** MALDI-TOF spectra for lysozyme modified with (A) compound 2 and (B) compound 3. In both the modifications, some amount of unmodified lysozyme was also observed.

The thermal stability against aggregation for lysozyme-(acetyl)_n and lysozyme-(methyl squarate)_n was compared to lysozyme-(βCD)_n. All the lysozyme modifications reactions with the three chemical agents contained some amounts of unmodified lysozyme. The thermo-induced aggregation study was conducted using dynamic light scattering. During the thermo-induced aggregation of lysozyme-(acetyl)_n and lysozyme-(methyl squarate)_n, I observed that lysozyme-(acetyl)_n formed aggregates after 5 min of heating at 67 °C, whereas lysozyme-(methyl squarate)_n formed aggregates after 5 min of heating at 57 °C. However lysozyme-(βCD)_n formed aggregates at 77 °C after 5 min. Native lysozyme forms aggregates at 67 °C after heating for 25
min. Interestingly, the aggregate size formed by lysozyme-(βCD)$_n$ was smallest compared to lysozyme-(acetyl)$_n$ and lysozyme-(methyl squarate)$_n$ (Table 3.3). These results indicate that the blocking the positive charges on the lysines residues in lysozyme results in a decrease in the thermal stability against aggregation of lysozyme. In case of lysozyme-(βCD)$_n$ the positive charges on the lysines are also getting blocked, but the grafted βCD on lysozyme is responsible for increase in the thermal stability against aggregation. Studying the folding of lysozyme-(acetyl)$_n$ and lysozyme-(methyl squarate)$_n$ using circular dichroism can further help to determine if the structure of lysozyme was still preserved after blocking the positive charges on lysozyme.

Table 3.3 Thermal stability of chemically modified lysozyme using dynamic light scattering (DLS)

<table>
<thead>
<tr>
<th>Entry</th>
<th>Protein</th>
<th>Diameter (nm)</th>
<th>Aggregation Temperature (°C)</th>
<th>Aggregate Size (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lysozyme</td>
<td>4.40</td>
<td>67</td>
<td>455.5</td>
</tr>
<tr>
<td>2</td>
<td>Lysozyme-(βCD)$_4$</td>
<td>2.36</td>
<td>77</td>
<td>579.7</td>
</tr>
<tr>
<td>3</td>
<td>Lysozyme-(methyl squarate)$_4$</td>
<td>3.81</td>
<td>57</td>
<td>910.8</td>
</tr>
<tr>
<td>4</td>
<td>Lysozyme-(acetyl)$_4$</td>
<td>3.57</td>
<td>67</td>
<td>698.5</td>
</tr>
</tbody>
</table>

*a determined by dynamic light scattering (DLS) at 25 °C

3.2.9 Generality of our system

The coiled coil domain is the most frequently occurring motif found in proteins. A characteristic feature of the coiled coil domains is the occurrence of an amino acid heptad repeat (abcdefg)$_n$, with predominantly hydrophobic residues at positions “a” and “d” and polar residues elsewhere. Generally amino acids such as Leu, Ile, Val, and Ala are found at the “a” and “d” positions. Positions “e” and “g” frequently have charged residues that might contribute to
stability and the specificity of helix association.\textsuperscript{142} Despite an apparent simplicity at the sequence level, a large number of structural variations are observed among coiled coils. Thrombospondins 3 and 4 (TSP3 and TSP4) have the same bouquet like structure formed by the COMP. Furthermore, the sequence alignment of the coiled coil domains of COMP, TSP3 and TSP4 are very similar.\textsuperscript{125} This suggests that a single mutation in TSP3 and TSP4 may result in the loss of α-helicity of the protein and covalent modification with βCD and methyl squarate alone can help induce α-helicity in the mutated protein.

Liu and coworkers have reported the engineering of a 52 residue protein (Ala14) that contains only alanine in the “a” and “d” positions of the natural heptad repeat sequence. Ala14 is unfolded under normal conditions and yet forms parallel three standed α-helical coiled coil crystals.\textsuperscript{143} Because hydphobicity is a key feature for determining the stability of proteins, replacing a bulky hydrophobic group such as leucine with alanine is believed to result in a decrease in protein stabilization.\textsuperscript{143-144} Hodges and coworkers have designed de novo a two stranded α-helical coiled coil and found that replacement of one large hydrophobic leucine with alanine decreases the stability of the coiled coils.\textsuperscript{145} Keating and coworkers report that the coiled-coil energy landscape is degenerate, where small changes in sequence can lead to large changes in structure.\textsuperscript{146} Clearly coiled coils are important structural motifs involved in a variety of important interactions that have the potential to be biochemically and therapeutically exploited.

The protein p53 also known as TP53 or tumor protein is a gene that codes for a protein that regulates the cell cycle and hence functions as a tumor suppressor. The N-terminal region consists of an intrinsically disordered transactivation domain (TAD) and a proline-rich region.\textsuperscript{147-148} One of the principle p53 regulator is Mdm2 or Hdm2, which binds directly to the TAD domain of p53 transactivation domain and targets p53 for proteosomal degradation. Upon
binding of the TAD domain of p53 to partner proteins, regions of nascent helical structure that are present in the native state rigidify and become fully folded.\textsuperscript{149} Reactivation of the p53 cell apoptosis pathway through inhibition of the p53-Hdm2 interaction is a viable approach to suppress tumor growth in many human cancers. Researchers have prepared synthetic p53 analogs via a hydrocarbon cross-link (staple) to increase the stabilization of the helical structure.\textsuperscript{117, 150} Since the formation of a stabilized $\alpha$-helix of TAD is important for Mdm2 or Hdm2 binding, I believe covalent modification of the TAD domain with $\beta$CD can stabilize and induce folding in the TAD domain and help us in exploring therapeutically relevant peptides.

### 3.3 Conclusions and Perspectives

The cartilage oligomeric matrix protein (COMP) is a glycogenic protein present in the cartilage, tendons, and ligaments of osteoblasts and has a pentameric bouquet like structure. A single amino acid mutation of leucine 44 to alanine in the coiled coil domain of COMP (COMPcc) caused a complete loss of the $\alpha$-helical structure for the mutant L44A. Covalent modification of L44A with $\beta$CD induced $\alpha$-helical structure in the inherently unstructured L44A. Interestingly, covalent modification of L44A with methyl squarate alone also induced an $\alpha$-helical structure in unstructured L44A. One plausible mechanism for induced folding is that the multiple hydroxyl groups on $\beta$CD probably act as a template to generate a structured solvation shell, due to of low dynamic hydrogen bonding between the hydroxyl groups on $\beta$CD and water molecules. Dynamic light scattering showed that the hydrodynamic diameter of L44A-$(\beta$CD)$_n$ was around 42 nm, which was higher than the hydrodynamic diameter of proteins with higher molecular weights. Modifying L44A with $\beta$CD not only folded the random coil sequence but also promoted pentameric assembly formation and enhanced binding to small molecules such as curcumin within the hydrophobic core of this pentameric assembly.
Another branch of the same work can be inducing folding in unstructured peptides. Anderson and coworkers have designed a 20-residue self-folding peptide using only the standard set of L-amino acids.\textsuperscript{151} Using such small folded proteins, one can design peptides, by decreasing the amino acid residues one at a time until the protein does not fold anymore. Once a peptide with a certain number of amino acids is acquired, refolding of the designed peptide can be studied by covalent modification with βCD. This approach can help in designing proteins with even less than 20 amino acid residues.

The protein p53 is a tumor suppressor protein encoded by the gene TP53. Mdm2 is an important negative regulator of the p53 tumor suppressor, which binds to the transactivation domain (TAD) of p53. The TAD domain of p53 is intrinsically disordered and adopts a α-helical conformation upon binding to Mdm2. I have designed synthetic p53 mimics, which I will covalently modify with βCD to study if folding is induced in the βCD conjugated p53 mimics. If folding is induced in the designed p53 mimics, binding studies with the Mdm2 receptor can be done to explore the therapeutic potential of these peptides.

### 3.4 Experimental Section

**Materials and Methods**

L44A and L44wt were obtained from Professor Jin Montclare at from Polytechnic Institute of New York University. All other chemicals were obtained from Aldrich Chemicals (Milwaukee, WI). Phosphate buffered saline, 1× PBS (2.7 mM potassium chloride, 137 mM sodium chloride, 8 mM sodium phosphate dibasic, 1.48 mM potassium phosphate monobasic) was used for all protein reactions unless mentioned (pH of the prepared buffers was adjusted using 1N HCl and 1N NaOH).
Circular Dichroism

Circular dichroism (CD) spectroscopy was conducted on a Jasco J-815 CD spectrometer. All proteins were at a concentration of 29 µM, unless otherwise mentioned. The wavelength spectrum was measured over a range from 190 to 250 nm with a step size of 1 nm.

General Procedure for Modification of L44A with Different Chemical Agents

Different chemical agents (~100 eq) were added to a solution of L44A (29 µM) in 1× PBS pH 8.2 (reaction buffer). The reaction mixture was stirred in a 28 ºC incubator shaker for 72 h. The reaction mixture was dialyzed in 250 mL 1× PBS, pH 7.4. MALDI-TOF was conducted on the protein samples to see the number of modification on the protein.

Modification of lysozyme with Acetyl and Methyl Squarate Groups

Compounds 2 (1.1 mg, 200 eq), 3 (1.70 mg, 200 eq) and 4 (2.16 mg, 100 eq) were added to a solution of lysozyme (1 mg/mL) prepared in 1× PBS pH 8.2 (reaction buffer) separately. The reaction mixtures were stirred in a 28 ºC incubator shaker for 72 h. The reaction mixtures were dialyzed in 250 mL 1× PBS, pH 7.4 in separate dialysis cassettes. MALDI-TOF was conducted on the samples to see the number of modification on the protein.

Dialysis

The protein samples were dialyzed in 1× PBS and/or in water (dialysis media), using Slide-A-Lyzer Dialysis Cassettes 3K molecular weight cut off (MWCO) purchased from Thermo Scientific (Rockford, IL). Before use, the Slide-A-Lyzer Dialysis Cassette was hydrated in the dialysis media for 5 min. Sample was injected in the cassette and placed in 250 mL of dialysis media to get rid of unreacted compound 6. The 1 × PBS was changed twice every 2 h for the first 6 h and then left overnight. The same reaction mixture was further dialyzed in a new dialysis bag.
in 250 mL of deionized water to promote aggregation of unreacted protein. The deionized water was changed twice every 2 h for first 6 h and then left overnight. Reaction mixture was then passed through a syringe filter (0.22 µm) to eliminate aggregated protein.

**Fluorescence Measurement for Curcumin Binding with L44A-(βCD)$_n$**

Fluorescence measurements of L44A-(βCD)$_n$ and curcumin binding was conducted with 29 µM of L44A-(βCD)$_n$ and 145 µM of curcumin in 10 mM phosphate buffer, pH 8.0. Samples were incubated at rt for 1 h and fluorescence was measured using a SpectraMax Plus M2 instrument with an excitation wavelength of 420 nm and an emission wavelength from 450-600 nm. An emission maxima was observed for L44A-(βCD)$_n$ and L44wt near 525 nm for curcumin binding. The fluorescence spectra was obtained after 2 h of incubation of curcumin with L44A-(βCD)$_n$, L44wt, L44A and βCD.

**Purification of L44A-(βCD)$_n$ Using High Performance Liquid Chromatography (HPLC)**

The separation of different degree of modification in L44A-(βCD)$_n$ were attempted using reverse phase high performance liquid chromatography (HPLC) from Shimadzu (Boston, Marlborough, MA) using a Alltima C18 5µm column (150 mm × 4.6 mm). The column was eluted at 1.75 ml/min with a gradient of water /0.1%TFA as eluent A and acetonitrile/water/0.1% TFA (90:9.9:0.1) as eluent B. The method used for purifying different degrees of modification is given in the table below. Eluted proteins were evaluated at 220 and 280 nm.
Method used for purifying different degrees of modification in L44A-(βCD)$_n$

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>Concentration of B (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>30</td>
</tr>
<tr>
<td>40</td>
<td>50</td>
</tr>
<tr>
<td>41</td>
<td>100</td>
</tr>
<tr>
<td>43</td>
<td>100</td>
</tr>
<tr>
<td>44</td>
<td>0</td>
</tr>
<tr>
<td>45</td>
<td>0</td>
</tr>
</tbody>
</table>

Synthesis and Spectra Data

**Synthesis of compound 2**

To a solution of acetic anhydride (0.088 g, 0.869 mmol) in THF (3 mL) N-hydroxy succinimide (0.100 g, 0.868 mmol) was added. Stirred the reaction mixture at rt for 1 h. Evaporated the THF to obtain pure compound 2 (0.127 g, 93%) as white solid. R$_f$ = 0.59 (70% EtOAc in Hexane). $^1$H NMR (300 MHz, CDCl$_3$): δ 2.78 (s, 4H), 2.30-2.28 (m, 3H). HRMS found = 157.0375 [M]$^+$, calcd for [C$_6$H$_7$O$_4$N]$^+$ =157.0374.
1.6 M methyl lithium (0.70 mL, 1.125 mmol) in Et₂O was added dropwise to a solution of diphenoxy-3-cyclobutene-1,2-dione (0.200 g, 0.751 mmol) in THF (10 mL) at -78 °C. The reaction was stirred -78 °C and was monitored by TLC until no more starting material was observed. Water was added dropwise at -78 °C to quench the reaction and reaction was slowly warmed up to rt. The aqueous phase was extracted with EtOAc. The combined organic layers were dried over anhydrous Na₂SO₄ and evaporation of organic solvents afforded the crude product as yellow oil. The crude product was purified using silica gel column (20% EtOAc in Hexanes) to afford pure compound 3a as colourless oil (0.110 g, 52%). R_f = 0.50 (40% EtOAc in Hexane). ¹H NMR (300 MHz, CDCl₃): δ 7.41-7.18 (m, 7H), 7.11-7.06 (m, 1H), 6.96-6.92 (m, 2H), 4.11 (s, 1H), 1.48 (s, 3H).

**Synthesis of compound 3**

To a solution of compound 3a (0.066 g, 0.233 mmol) in DCM (2 ml) was added 3-5 drops of 12 N HCl at rt. Stirred the reaction mixture at rt for 2 h. The reaction mixture was diluted with DCM and organic phase was washed with water (3 × 10 mL). The organic layer was dried over
anhydrous \( \text{Na}_2\text{SO}_4 \) and evaporation of organic solvents afforded the crude product as yellow oil. The crude product was purified using silica gel column (20% EtOAc in Hexanes) to obtain light yellow solid (0.044g, quantitative yield). The solid was recrystallised with ether, hexane and a few drops of acetone to obtain, white crystals of compound 3. \( R_f = 0.64 \) (40% EtOAc in Hexane).

\( ^1\text{H NMR} \) (300 MHz, CDCl\(_3\)): \( \delta \) 7.46-7.43 (m, 2H), 7.41-7.29 (m, 1H), 7.22-7.18 (m, 2H), 2.08 (s, 3H).

\( ^{13}\text{C NMR} \) (75 MHz, CDCl\(_3\)): \( \delta \) 197.3, 196.3, 191.5, 181.8, 153.5, 130.5, 127.5, 119.5, 10.4.

HRMS found = 188.0468 [M + H]\(^+\), calcd for [\( \text{C}_{11}\text{H}_7\text{O}_3 \)]\(^+\) = 187.0396.
Chapter 4

Synthesis of an Adamantane-Platinum Conjugate Hosted in β-Cyclodextrin for Enhanced Activity and Targeted Drug Delivery†

Summary

In this section of the dissertation, the synthesis of an adamantane-platinum conjugate (analog of carboplatin) hosted in β-cyclodextrin is reported. The cytotoxic activity of the carboplatin analog hosted in β-cyclodextrin was tested towards human neuroblastoma cells (SK-N-SH). This inclusion complex of βCD adamantane carboplatin analog exhibited higher cytotoxicity towards SK-N-SH cells than carboplatin. Agarose gel electrophoresis confirmed that this inclusion complex of βCD adamantane carboplatin analog exhibited higher binding to plasmid pBR322 deoxyribonucleic acid (DNA) than carboplatin. Confocal fluorescence images showed that SK-N-SH cells treated with FITC-tagged βCD exhibited green fluorescence in the vicinity of the nuclei of the cells, suggesting that FITC-tagged βCD were efficiently internalized into the cancer cells. The uptake rate at which carboplatin and the βCD adamantane carboplatin analog were taken up by the cancer cells was measured by inductively coupled plasma mass spectrometry (ICP-MS), which indicated that the uptake rate of carboplatin was about 4 times higher than βCD- adamantane carboplatin analog. The higher cytotoxicity of the inclusion complex than carboplatin towards SK-N-SH cells may be due to its higher DNA binding ability and efficient delivery to the cancer cells as compared to carboplatin.

4.1 Background and Significance

4.1.1 Why Beta-Cyclodextrin (βCD)?

Cyclodextrins (CD) have been known to prevent aggregation during protein refolding.\textsuperscript{46-47, 53} The hydrophobic core of CD is employed for forming molecular inclusion complexes with a wide range of hydrophobic small molecules or drugs.\textsuperscript{52, 152} The ability of CD to allocate hydrophobic drug molecules in its hydrophobic cavity helps in increasing the stability, solubility and bioavailability of such drugs.\textsuperscript{153-154} Furthermore, the primary and secondary hydroxyl groups of CD are potential sites for chemical modification for conjugating targeting ligands or antibodies to achieve targeted drug delivery. Anticancer drugs such as cis-platin, carboplatin and oxaliplatin lack target specificity. The dose-limiting side effects and intrinsic or acquired resistance are the main draw backs associated with these drugs. Targeted drug delivery is important for conventional chemotherapy and for preventing the normal tissue from the toxic effects of the drug. Target specificity can be achieved by either employing antigen-antibody interactions or targeting ligands. To exploit cyclodextrins as targetable drug delivery systems, they can be functionalized with targeting moieties such as peptides, hormones, antibodies and carbohydrates. There are a number of antibodies/ligands, which are capable of recognizing specific antigens/receptors on the tumor cells. Furthermore, a large number of cancer cells over express surface receptors for folic acid\textsuperscript{155-158} and galactose.\textsuperscript{159-160}

4.1.2 Platinum compounds as anticancer agents

Platinum-based compounds are an important class of active chemotherapeutic agents for treating various cancers,\textsuperscript{161-164} yet the mechanism of drug action are still unraveling.\textsuperscript{165-167} Dabrowiak and coworkers have reported that for cisplatin and carboplatin, a platinum carbonato complex rather than the reactive mono-aquo complex may be responsible for cytotoxic effects in
cells.\textsuperscript{168-169} With this growing understanding, the development and optimization of the pharmaceutical properties of platinum-based therapeutic agents has been largely empirical.\textsuperscript{170-171} This work presents a strategy for noncovalent modification via host-guest chemistry to enhance the cytotoxicity of a dicarboxylate chelated platinum (II) complex, which is an analog of the clinically used drug carboplatin. This noncovalent modification strategy employed the cyclic sugar \( \beta \)-cyclodextrin (\( \beta \)CD) to encapsulate an adamantane-derivatized diacarboxylate complexed with a platinum (II) ion.

\textit{4.1.3 \( \beta \)-cyclodextrin as drug carrier}

Noncovalent modification of active agents provides an alternative approach to achieve desired properties for drug development, such as increased water-solubility, cellular uptake and targeted delivery, without severely tampering the activity of the compounds. Kim and coworkers have demonstrated the encapsulation of the anti-cancer drug oxaliplatin in the hydrophobic cavity of cucurbit. Preliminary results indicated that the activity of oxaliplatin decreased after encapsulation in cucurbit.\textsuperscript{172} Brabec and coworkers have studied the cytotoxicity, mutagenicity, cellular uptake and DNA interactions of trans-platinum complexes tethered to 1-adamantylamine.\textsuperscript{171} Cyclodextrins are used in pharmaceutical applications for improving drug properties such as enhancing the solubility, stability and bioavailability of drugs.\textsuperscript{52, 153, 173} Interestingly, Maxfield and coworkers discovered that \( \beta \)-cyclodextrin (\( \beta \)CD) has a high propensity to be taken into mammalian cells presumably through pinocytosis.\textsuperscript{174} This discovery promoted us to explore the possibility of enhancing the properties of a platinum (II)-based compound by noncovalent encapsulation in \( \beta \)CD. I synthesized an adamantane-diacarboxylate chelated with a platinum (II) ion, and used \( \beta \)CD to form a host-guest complex with the adamantane moiety of the chelate (Scheme 4.1). The binding is stoichiometric in water with one
adamantane group being hosted in the hydrophobic annular cavity of a single βCD. The cytotoxicity of this host-guest complex was tested against neuroblastoma cells and compared to the cytotoxicity of carboplatin. Proteins such as antibodies can be conjugated to cyclodextrin to achieve reduced aggregation of proteins and/or targeted drug delivery. This approach provides both chemical versatility for structural modification and potential for protein-drug conjugate therapeutics.

4.2 Results and Discussion

4.2.1 Synthesis of the βCD adamantane-platinum host-guest complex (6)

Scheme 4.1 Synthesis of the βCD adamantane-platinum host-guest complex (6)

Scheme 4.1 shows the synthesis of the βCD adamantane-platinum host-guest complex (6). Briefly, tosylation of adamantane alcohol (1), afforded the tosylate (2). Displacement of tosyl group in (2) with diethylmalonate gave the diethyl ester (3), which on hydrolysis provided the diacid (4). The detailed synthesis of compound 4 has been reported elsewhere. I noted that direct coupling of (4) and cis [Pt(NH₃)₂(H₂O)]⁺² (5), gave a grey solid that was insoluble in all solvents studied (Scheme 4.2).
**Scheme 4.2** Synthesis of cis [Pt(NH$_3$)$_2$(H$_2$O)$_2$]$^{2+}$ (5) and conjugation of compound 5 to adamantane carboxylic acid (4)

Because cyclodextrins are known to form water soluble inclusion complexes with hydrophobic groups, I tried using βCD to possibly extract the product from the insoluble grey solid into water. However, addition of βCD to an aqueous suspension of the grey solid did not result in detectable product. Interestingly, treating the diacid (4) first with 1 equivalent of βCD, followed by addition of [Pt(NH$_3$)$_2$(H$_2$O)$_2$]$^{2+}$ (5), resulted in the desired water-soluble conjugate (6). The details of the synthesis suggest that the exchange rate of the βCD-adamantane host-guest complex is likely slow, and that using βCD likely provides a general approach to render hydrophobic small molecules soluble in water. [$^1$H, $^{15}$N] Heteronuclear single quantum coherence (HSQC) NMR was taken for the βCD adamantane-platinum host-guest complex (Figure 4.1). The chemical shift (δ) in HSQC for $^1$H= 4.13 and $^{15}$N = -82.0, was similar to that observed for carboplatin.
Figure 4.1 $[^1\text{H}, ^{15}\text{N}]$ HSQC NMR of βCD adamantane carboplatin analog host-guest complex (6) in 5% D$_2$O/H$_2$O: δ (ppm) $^1\text{H} = 4.13$, δ $^{15}\text{N} = -82.0$.

4.2.2 Studying cytotoxicity of host-guest complex (6) and carboplatin using colorimetric assay

The βCD adamantane-platinum host-guest complex (6) is structurally close to carboplatin because of the common dicarboxylate motif. To explore the therapeutic potential of the host-guest complex (6), I compared cytotoxicity of 6 and carboplatin toward neuroblastoma cells (SK-N-SH). The neuroblastoma (SK-N-SH) cells were grown for 24 h in 96-well plates, followed by treating the cells with different concentrations of 6 or carboplatin. After 1 h incubation, medium containing 6 or carboplatin was replaced with fresh culture medium. The cells were allowed to grow over a period of 24 h, 48 h, and 72 h (recovery period) in different 96-well plates. At the end of each recovery period, the number of live cells was determined by a colorimetric cell counting assay (CCK-8). The assay utilizes a water soluble tetrazolium salt, which is reduced by dehydrogenases in the live cells to give a yellow colored product (formazan) (Scheme 4.3). The amount of the formazan dye produced in the culture medium is directly proportional to the number of living cells.
Scheme 4.3 Reduction of WST-8 by dehydrogenases to produce yellow colored dye formazan

![Chemical Structures]

The survival (%) was calculated by using the equation \((\text{OD}_{450} \text{ sample})/\text{(OD}_{450} \text{ control}) \times 100\), where the sample OD was obtained from the cells treated with \(6\) or carboplatin, and the control OD was obtained from cells alone.\(^{181}\) Figure 4.2 shows the plots of survival (%) of neuroblastoma (SK-N-SH) cells as the concentrations of \(6\) and carboplatin increases at different recovery periods of 0 h, 24 h, 48 h, and 72 h (starting immediately after host-guest complex (6) or carboplatin were replaced by fresh culture medium). Immediately after 1 h of culturing the cells with \(6\) or carboplatin (zero recovery time), the number of viable cells remained high. However, for longer recovery time, the cells exhibited an unusual dose dependent behavior to carboplatin. The cytotoxicity of carboplatin increases in the drug concentration range 0 µM to \(~500\) µM, reaches a maximum at \(~500\) µM and decreases in the range, 500 µM to 1000 µM (Figure 4.2 A).
Recently, it was shown that carboplatin exists in a monomer-dimer equilibrium in aqueous solution, and there is evidence that the same equilibrium may also exist in culture media.\textsuperscript{182} Since the concentration of a species in monomer-dimer equilibrium is dependent on the total concentration of the agent in the medium, it is possible that the reduced cytotoxicity of carboplatin at high concentration is due to the presence of dimers, which may be less toxic to the cells than the monomeric form of the drug. While being non-toxic at all concentrations for a recovery time of 0 h, \textbf{6} exhibited the typical dose-response expected at longer recovery times with the longest recovery time (72 h) producing the greatest amount of cell death. Comparison of the survival (%) for \textbf{6} with carboplatin indicates that overall \textbf{6} is more toxic to the neuroblastoma cells than is carboplatin.

\textit{4.2.3 Cytotoxicity of host-guest complex (6) and carboplatin against SK-N-SH cells increases with increasing incubation times}

It is intriguing that increasing the concentration of carboplatin did not strictly lead to increase in cytotoxicity. The effect of different incubation times of \textbf{6} and carboplatin on neuroblastoma cells (SK-N-SH) was also examined. Confluent layers of SK-N-SH cells were incubated with...
500 µM of each agent for 2 h, 4 h, 6 h, 8 h and 12 h in 96-well plates. After each incubation time, medium containing 6 or carboplatin was replaced with fresh culture medium, and the number of live cells was immediately determined by CCK-8 assay (recovery time = zero).

**Figure 4.3** Survival (%) of SK-N-SH cells when treated with 500 µM carboplatin and 6 at different incubation times of 2 h, 4 h, 6 h, 8 h, and 12 h. The recovery period was 0 h for all incubation times tested.

Figure 4.3 shows that exposure of the neuroblastoma cells to both carboplatin and 6 for exposure times up to 12 h enhanced the cytotoxic effects of both agents but compound 6 was clearly more toxic than carboplatin. After about 6 h of incubation of the SK-N-SH cells with 500 µM of 6 or carboplatin, I observed 87.04 ± 4.69% survival of SK-N-SH cells treated with 6, and almost no decrease in the survival of SK-N-SH cells treated with carboplatin. Carboplatin appeared to begin showing significant cytotoxic effects (recovery period = 0), only after cells were incubated for about 6 h in the presence of the drug. After 12 h of incubation, 66.29 ± 3.20% and 82.53 ± 3.61% survival was observed for cells treated with 6 and carboplatin respectively. These results indicate that increasing the incubation time of the cells with carboplatin results in
an increase in the cytotoxic effects of the agents. However, higher cytotoxic effect of 6 over carboplatin is evident at longer incubation times (6 h-12 h) of the cells with the agents.

4.2.4 Cyclodextrin enters cells through pinocytosis

Maxfield and coworkers recently discovered that endocytosis of βCD is responsible for cholesterol reduction in Niemann-Pick type C mutant cells.\textsuperscript{174} This effect is believed to be due to the internalization of cyclodextrin in the cell through fluid phase pinocytosis.\textsuperscript{174, 183-184} Pinocytosis is a form of endocytosis in which surrounding fluids and small particles are brought into the cell suspended within small vesicles, which subsequently fuse with lysosomes. To validate the internalization of βCD, I synthesized fluorescein isothiocynate (FITC) tagged cyclodextrin by literature reported procedures (Scheme 4.4).\textsuperscript{185}

**Scheme 4.4** Synthesis of FITC tagged βCD (10)

![Scheme 4.4](image-url)

SK-N-SH cells were treated with FITC or FITC-tagged βCD (10). Confocal fluorescence images showed that SK-N-SH cells treated with compound 10 or FITC alone exhibited green fluorescence (Figure 4.4), suggesting that both are efficiently internalized into the SK-N-SH cells post fixation. Interestingly, compound 10 appeared to be localized around the nucleus of the cells, whereas FITC alone stained the entire cell.
Figure 4.4 Confocal fluorescence image of SK-N-SH cells incubated with (A) 150 nM of FITC-tagged βCD and (B) FITC alone. The cells were fixed with 4% paraformaldehyde before viewing under the microscope. Scale bar = 76 µm.

4.2.5 SDS-PAGE confirms host-guest complex (6) binds DNA more than carboplatin

Because binding to DNA is a key element for platinum-based therapeutic agents, I examined and compared the binding of a plasmid DNA (pBR322) by 6 and by carboplatin using agarose gel electrophoresis. Plasmid DNA (pBR322) exists in different conformations (Form I and Form II), each having a different mobility in an agarose gel. Form I DNA is a circular supercoiled DNA, and Form II DNA is a nicked circular DNA. Form I DNA is wound into a compact structure and hence is the fastest moving conformation of the plasmid. Once the plasmid DNA is treated with the platinum agents, binding and unwinding of Form I DNA takes place, which is manifested by a decrease in the mobility of Form I DNA. The plasmid DNA was incubated for 24 h with different concentrations (30 µM to 100 µM) of 6 or carboplatin. The agarose gel image (Figure 4.5) shows that the mobility of DNA decreased with increase in concentrations of 6 and carboplatin. This decrease in the DNA mobility in the agarose gel is consistent with the unwinding of Form I DNA due to binding of 6 or carboplatin. This result suggests that 6 binds to and unwinds plasmid DNA (pBR322) more than carboplatin. For
example, the DNA in lane 11 (6 at 100 µM) migrated less than in lane 18 (carboplatin at 100 µM). As shown in Figure 4.5, carboplatin and 6 have little effect on the mobility of the nicked circular, Form II DNA (no supercoiling) in the gel.

![Figure 4.5 Agarose gel electrophoresis image of plasmid DNA (pBR322) (33.3 µM base pairs) after incubation with 6 or carboplatin in 10 mM HEPES, pH 7.4 for 24 h. Lanes 2 through 11 contained 30, 35, 40, 45, 50, 60, 70, 80, 90, and 100 µM of 6, and lanes 12 through 18 contained 30, 40, 50, 70, 80, 90 and 100 µM of carboplatin, respectively. Lane 1 (control) contained only plasmid DNA (pBR322). This study was done in conjunction with Yi Shi from Professor Dabrowiak’s lab.

4.2.6 βCD alone does not bind DNA

The potential interaction between βCD and DNA using agarose gel electrophoresis was also studied. The concentrations of βCD used for the DNA interaction studies were same as those used for studying the binding of 6 to pBR322 DNA. No unwinding of supercoiled Form I closed circular DNA was observed for all concentration of βCD used for the study (Figure 4.6). This suggests that either βCD binds DNA but does not affect the mobility of Form I or Form II DNA or βCD does not bind DNA at all. From these results, it is concluded that the decrease in of mobility of DNA on treatment with 6 is only due to the binding of the platinum to DNA.
Figure 4.6 Agarose gel electrophoresis image of plasmid DNA (pBR322) (33.3 µM base pairs) after incubation with βCD in 10 mM HEPES, pH 7.4 for 24 h. Lanes 2 through 12 contained 30, 35, 40, 45, 50, 60, 70, 80, 90, and 100 µM of βCD respectively. Lane 1 (control) contained only pBR322 DNA. This study was done in conjunction with Yi Shi from Professor Dabrowiak’s lab.

4.2.7 Platinum uptake studies using ICP-MS

Because 6 appears to be more cytotoxic than carboplatin for SK-N-SH cells, I examined if the uptake rate of the chemical agents by the SK-N-SH cells possibly contributes to the observed difference in the cytotoxicity. I used inductively coupled plasma mass spectrometry (ICP-MS) to directly measure the platinum content by SK-N-SH cells over time.\textsuperscript{169,190} Figure 4.7 shows the plots for platinum uptake per cell, in attomoles ($10^{-18}$) as a function of incubation time. Six well plates with each well containing $3.5 \times 10^5$ SK-N-SH cells suspended in culture medium, were allowed to adhere on the well surface for 24 h. The number of viable cells in the wells was measured by light microscopy using a hemocytometer and trypan blue. The cells in the wells were then treated with 500 µM of 6 or carboplatin for different incubation times (0, 1, 2, 3, and 4 h). The cells were trypsinized and pelleted by centrifugation and the pellets were washed with phosphate buffered saline (PBS, 10 mM) twice. The cells were then lysed in 70% nitric acid at 50 °C for 48 h and the analyzed for the amount of platinum by ICP-MS. The uptake rate of platinum per cell, in attomoles ($10^{-18}$) per hour per cell, was calculated from the slope of the plot.
The uptake rates of carboplatin and 6 at 500 µM by SK-N-SH cells were found to be 16.0 ± 0.56 and 64.9 ± 1.5 amol of Pt cell⁻¹ h⁻¹ respectively. Comparing the uptake of platinum (Figure 4.7) and the survival (%) of SK-N-SH cells (Figure 4.2) after 4 h incubation, while the uptake of carboplatin was ~ 4.4 times higher than that of 6, there was a decrease of ~ 6% in the survival for SK-N-SH treated with 6 but no observable decrease of survival for cells treated with carboplatin.

**Figure 4.7** Platinum (amol) taken up per SK-N-SH cell, when exposed to carboplatin (500 µM) and 6 (500 µM) as a function of time. The rate of uptake (amol of Pt cell⁻¹ h⁻¹) was found to be 16.0 ± 0.56 and 64.9 ± 1.4 for cells treated with 6 and carboplatin, respectively.

These data indicate that the uptake rate of carboplatin is ~ 4 times higher than 6, while the cytotoxicity experiments suggest that 6 is more cytotoxic than carboplatin. This result is surprising and suggests that the cytotoxic effects of chemical agents are complex. Hence, the amount of platinum entering the cell alone cannot be the basis for the different cytotoxic effects of the compounds. However, since the main target for the platinum agents is believed to be DNA, carboplatin could be more susceptible to reaction with and inactivation by components in the cytosol than 6.
4.2.8 Achieving targeted drug delivery

Many systems have been developed for selective delivery of drugs to their site of action by grafting biorecognizable carbohydrate antennae on to the carrier. Pack and coworkers have synthesized cyclodextrin-ployethylenimine conjugates for targeted \textit{in-vitro} gene delivery. Shinoda and coworkers have investigated the specific interaction between galactose branched-cyclodextrins and hepatocytes \textit{in-vitro}. Davis and coworkers have synthesized transferrin-PEG-adamantane conjugate for nanoparticle modification. The transferrin-modified nanoparticles were further used for tumor targeted gene delivery.

The capacity and specificity of the drug carrying cyclodextrin can be increased by attaching bio-recognizable galactose ligands. Attachment of multiple cyclodextrins carrying the drug will also increase the net amount of platinum reaching the tumor environment and will hence decrease the amount of platinum being used (Figure 4.8). The galactosyl-β-cyclodextrin derivatives can be recognized by galactose binding proteins called galectins. The recognition is dependent upon the length of the spacer chain between the cyclodextrin and the sugar head group and also on the number of sugars presented to the lectin.
Figure 4.8 Interaction between directing ligands (antibody, folic acid and galactose) and cancer cell surface receptors (antigen, folic acid receptor and galactose receptor).

Scheme 4.5 Conjugation of compound 14 with cyclodextrin-galactose conjugate (15)

Since 6 was soluble in water, I attempted the synthesis of a cyclodextrin-galactose conjugate (Scheme 4.5) with the aim of achieving target specificity. The synthesis of the galactose head group connected with a βCD with a spacer chain length of five carbons was tried (Scheme 4.3).
Briefly, compound 14 was synthesized using literature reported procedures. Compound 17 was synthesized by treating compound 16 with sodium azide in DMF. Reduction of compound 17 with Pd/C in ethanol gave compound 18. Compound 18 was condensed with mono-methyl nonanoic acid in the presence of triethyl amine to yield compound 19. Complete ester hydrolysis of 19 was achieved with methanolic sodium hydroxide to afford compound 20. Compound 14 was condensed with compound 20 using DCC/HOBt to obtain compound 15, however pure product was not obtained (Scheme 4.6). Although the $^1$H NMR for the compound 15 was not pure, HRMS confirmed product formation.

Scheme 4.6 Synthesis of cyclodextrin-galactose conjugate (15)
4.3 Conclusion and Perspectives

This work reports the synthesis and development of a βCD encapsulated adamantane-platinum conjugate (6). Conjugate 6 exhibited higher cytotoxicity than carboplatin towards neuroblastoma cells. The uptake of βCD adamantane-platinum inclusion complex is believed to be due to the internalization of βCD through pinocytosis. Agarose gel electrophoresis confirmed the binding of 6 to pBR322 DNA, where 6 exhibited a higher level of binding to DNA than carboplatin. As gel electrophoresis confirmed the binding of 6 to pBR322 DNA, this work suggests that noncovalent modification of a therapeutic agent can be a powerful strategy to control the binding activities and other pharmaco-properties of the therapeutic agent. Interestingly, the ICP-MS measurement indicated that carboplatin had a higher uptake rate than 6 into the cells. These results suggest that other factors such as target receptor binding and heterogeneous distributions of agents inside the cells likely play more important roles than mere cellular uptake for cytotoxic effects. The formulation and structural design for controlling drug activity can be studied and integrated by noncovalent modification of the drug candidates, rather than studied as separate topics.

Two approaches can be used to achieve targeted drug delivery. The first approach involves conjugation of multiple copies of ligands such galactose and folic acid with βCD. The βCD can in turn be used to carry multiple copies of candidates like 6 or other drugs, to achieve combination chemotherapy. The second approach involves conjugation of βCD with antibodies, peptide or protein drugs which are recognized by membrane proteins of cancer cell to achieve targeted drug delivery. Covalent modification of proteins like lysozyme and RNase A with βCD have been reported to prevent aggregation, surface adsorption and to increase the thermal stability of the modified proteins significantly (Chapter 2). Hence conjugation of βCD with
antibodies can help prevent aggregation of the antibodies and ensure efficient delivery of drugs to the affected site.

4.4 Experimental Section

Materials and Methods

Chemicals. Carboplatin was obtained from Sigma-Aldrich (Milwaukee, WI). A stock solution of carboplatin (10 mg/mL, 26 mM) was prepared in water before use. Culture medium, minimum essential medium with Earle’s salts and L-glutamine (EMEM) (10-010) was obtained from Mediatech (Herndon, VA). Cell counting kit 8 (CCK-8) was purchased from Dojindo Molecular Technologies, Inc. (Gaithersburg, MD). For the DNA cleavage studies, HEPES, agarose, Trizma base, boric acid, EDTA, and ascorbic acid were purchased from Sigma-Aldrich (St. Louis, MO), and the plasmid DNA (pBR322), 25 μg/μL, was obtained from Invitrogen (Carlsbad, CA). All other chemicals were obtained from Aldrich Chemicals (Milwaukee, WI). Water used for preparing all buffers and solutions had resistivity of 18 MΩ cm (Millipore, Billerica, MA).

CCK-8 Cytotoxicity Assay

All Cell studies were carried out under standard conditions in a humidified, 37 °C, 5% CO₂ atmosphere. Human neuroblastoma (SK-N-SH) cells in culture medium (EMEM + 10% FBS, 100 μg/mL streptomycin, 100 IU/mL penicillin, and 2.0 mM L-glutamine) were plated in 96-well plates (costar, tissue culture treated, Corning Incorporated, Corning NY), with each well containing 100 μL of cell solution with a concentration of 5×10⁴ cells/mL. The viabilities of SK-N-SH, determined with a hemacytometer under standard trypan blue conditions. After being plated, the SK-N-SH, were given 24 h to adhere to the bottom of the wells. Each plate consisted of six wells containing medium without cells, six wells containing medium with cells, and the remaining wells containing cells with medium and different concentrations of 6 or carboplatin.
After 1 h, the drug-containing medium was removed and the cells were washed twice with fresh culture medium. The cells were then allowed to recover for 24, 48 and 72 h (recovery period), after which the number of live cells was determined by the colorimetric cell counting assay (CCK 8 assay). The assay utilizes a water-soluble tetrazolium salt, which is reduced by dehydrogenases in the live cells to give a yellow colored product (formazan). After 2 h of incubation of the cells with the water soluble tetrazolium salt, the amount of the formazan dye produced in the culture medium is directly proportional to the number of living cells. The number of live cells in the wells was determined by measuring the absorbance of the medium in the wells at 450 nm on ELx 800 absorbance microplate reader from BioTek (Highland Park, Winooski, VA). The survival (%) was calculated using the equation \((\text{OD}_{450 \text{ sample}})/\text{(OD}_{450 \text{ control}}) \times 100\), where the sample OD was obtained from the cells treated with host-guest complex 6 or carboplatin, and the control OD was obtained from cells without any chemical treatment.

**Cleaning of Microscope Cover Glass**

Fisherbrand microscope cover glasses (Fisher Scientific, Pittsburgh, PA), were cleaned with piranha solution prior to cell experiments. The cover glasses were soaked in piranha solution (3 parts of 35% \(\text{H}_2\text{O}_2\) in water and 7 parts of concentrated \(\text{H}_2\text{SO}_4\)) at 70 °C for 45 min. Warning! Piranha solution is extremely corrosive and has the potential for detonation if contaminated with a significant amount of oxidizable material. After cooling, the piranha solution was poured off and the slides were rinsed 20 times with water having a resistivity of 18 MΩ cm.
Confocal Laser Scanning Microscopy (CLSM)

Piranha cleaned cover glasses in a six-well plate were incubated with 1% of SK-N-SH cells in culture medium (EMEM + 10% FBS, 100 μg/mL streptomycin, 100 IU/mL penicillin, and 2.0 mM L-glutamine). The cells were allowed to adhere on the cover glass for 24 h. The adhered cells on the cover glass were then treated with 150 nM solution of βCD-FITC in DMSO for 24 h. The cover glass was then washed thoroughly with 10 mM phosphate buffered saline (PBS, 2.7 mM potassium chloride, 137 mM sodium chloride, 8 mM sodium phosphate dibasic, 1.48 mM potassium phosphate monobasic), pH 7.4. The cells were fixed with 4% paraformaldehyde prepared in 10 mM PBS, for 15 min and the attached cells were again rinsed with 10 mM PBS, before viewing under the microscope. Confocal fluorescence images of neuroblastoma cells (SK-N-SH) treated with FITC-tagged βCD were taken on Carl Zeiss MicroImaging GmbH (Gottingen, Germany).

Platinum Uptake Study by SK-N-SH Cells Treated with βCD adamantane Carboplatin Analog Inclusion Complex (6) and Carboplatin Using ICP-MS

Two six-well plates (plate 1 and plate 2) with each well containing 3.5 ×10^5 of SK-N-SH cells suspended in culture medium (EMEM + 10% FBS, 100 μg/mL streptomycin, 100 IU/mL penicillin, and 2.0 mM L-glutamine) were allowed to adhere on the well surface for 24 h. The number of viable cells in each well for plate 1 and plate 2 after 24 h was found to be 2.04 ×10^5 and 3.85×10^5 respectively. The cells in the wells were then treated with 500 μM of 6 or carboplatin for different incubation times (0, 1, 2, 3, and 4 h). The cells were trypsinised using 1 mL 0.25% trypsin/0.5 mM EDTA solution after being treated with 6 or carboplatin. The suspended cells were then pelleted by centrifugation for 5 min at 3000 rpm in a 15 mL centrifuge tube. After centrifugation, culture medium was removed and the cells were resuspended in 2 mL
of 10 mM PBS and transferred into 1 mL eppendorf tubes. The cells were pelleted a second time by centrifugation, PBS was removed and the cells were lysed for 48 h at 50 °C in 0.5 mL of 70% nitric acid. The nitric acid solutions of the cell lysate were analyzed for amount of platinum by inductively coupled plasma mass spectrometry (ICP-MS) (Department of Engineering, Syracuse University). The uptake rate of platinum per cell, in attomoles \(10^{-18}\) per hour, was calculated from the total number of viable cells (number of viable cells in each well in the six well plates, before treatment with 6 and carboplatin were \(2.04 \times 10^5\) and \(3.85 \times 10^5\) respectively as noted above) and the number of moles of platinum that each cell took up from the culture medium per hour.\(^{169}\)

**Agarose Gel Electrophoresis**

The agarose gel was prepared by dissolving 1 g of agarose in 100 mL of 0.5 × Tris borate EDTA buffer (0.045 M Tris borate buffer and 0.001 M EDTA). Different concentration of 6 (30 µM to 100 µM) and carboplatin (30 µM to 100 µM) were mixed 10 mM HEPES buffer (pH 7.4), plasmid DNA (pBR322, 33.0 µM base pairs) and the volume was made up to 20 µL with distilled water. The samples were heated at 37 °C for 24 h. The solutions of 6 or carboplatin or βCD (20 µL) were further mixed with 2.5 µL of bromphenol blue and 8 µL of each sample was then loaded on to the gel. Electrophoresis was performed at 100 V for 4 h. The gel was then stained with ethidium bromide for 30 min. The gel was destained in water for 15 min and was viewed under UV light using Kodak Gel Logic 100 equipped with Fisher Biotech IT-88A transilluminator.
Synthetic Procedures and Spectral Data

Synthesis of compound 9

\[
\begin{align*}
\text{H}_3\text{^15N} & \quad \text{Pt} \\
\text{H}_3\text{^15N} & \quad \text{I} \\
\end{align*}
\]

To an aqueous solution of potassium tetrachloroplatinate (II) (0.005 g, 1.20 mmol) was added potassium iodide (0.799 g, 4.81 mmol) and stirred at rt for 30 min. When the solution became dark brown in color, \(^{15}\text{NH}_4\text{Cl}\) (0.131 g, 2.40 mmol) was added and the pH was adjusted to 11.0 using 10M KOH solution. The reaction mixture was stirred at room temperature for 1 h. Reaction mixture was filtered to obtain compound 9 as a yellow solid (0.489 g, 84%). \(^{1}\text{H}, {^{15}\text{N}}\) HSQC NMR (500 MHz, 5% D\textsubscript{2}O in water): \(\delta (^{1}\text{H}) = 3.93, \delta (^{15}\text{N}) = -52.0\).

Synthesis of compound 5

\[
\begin{align*}
\text{H}_3\text{^15N} & \quad \text{Pt} \\
\text{H}_3\text{^15N} & \quad \text{H}_2\text{O} \\
\end{align*}
\]

To a suspension of compound 9 (0.042 g, 0.087 mmol) in water added silver nitrate (0.029 g, 0.175 mmol) and stirred the reaction in dark for 20 h. Filtered the reaction mixture to remove silver iodide to yield compound 5 (quantitative) as a colorless liquid. \(^{1}\text{H}, {^{15}\text{N}}\) HSQC NMR (500 MHz, 5% D\textsubscript{2}O in water): \(\delta (^{1}\text{H}) = 4.50, \delta (^{15}\text{N}) = -86.5\).
Synthesis of compound 6

Compound 4 (0.029 g, 0.008 mmol) was dissolved in a mixture of methanol and water (1:3). βCD (0.099 g, 0.008 mmol) was added to the dissolved solution and the reaction mixture was stirred at rt for 1.5 h. A solution of compound 5 (0.387 mmol in 2 mL of water) was added dropwise to the above solution and the reaction mixture was stirred at rt for 3.5 h. The reaction mixture was freeze dried to obtain the crude product as a white solid. The white solid was then suspended in hot water and allowed to precipitate out completely at rt. The supernatant was discarded by decantation to obtain compound 6 (0.122 g, 95%) as white solid. $^1$H NMR (300 MHz, D$_2$O): δ 5.10-5.08 (m, 7H), 3.90-3.80 (m, 26H), 3.64-3.61 (m, 16H), 3.58-3.56 (br s, 2H), 2.98 (br s, 2H), 2.50 (br s, 1H), 2.1 (br s, 3H), 1.92-1.88 (m, 3H), 1.78-1.74 (m, 4H), 1.60 (s, 9H), 1.43 (br s, 3H), 1.32 (br s, 1H). [$^1$H, $^{15}$N] HSQC NMR (500 MHz, 5% D$_2$O in water): δ ($^1$H) = 4.13, δ ($^{15}$N) = -82.0. HRMS found = 873.7782 [M +2+2Na]$^+$/2 calcd for [C$_{12}$H$_{18}$O$_6$+2+2Na]$^+$/2 = 873.7769.

Synthesis of compound 12

Dissolved compound 11 (2.54 g, 0.002 mmol) in 57 mL of water by heating the reaction mixture to 60 °C with vigorous stirring. Reaction mixture was cooled to rt. To the resulting suspension
added finely powered 1-(p-tolunenesulfonyl) imidazole (2.0 g, 0.009 mmol). Reaction mixture was stirred at rt for 2 h. A solution of sodium hydroxide (1.15 g, 0.028 mmol) in 3 mL water was added slowly over a period of 5 min. Reaction mixture was stirred at rt for 10 min and filtered to remove unreacted 1-(p-tolunenesulfonyl) imidazole. Reaction mixture was quenched by addition of ammonium chloride (3.08 g, 0.057 mmol). The reaction mixture was concentrated to half its volume by air blowing overnight. Compound 12 was obtained as white solid (1.138 g, 39%) which was filtered and washed with ice-cold water and acetone. \(^1\)H NMR (300 MHz, DMSO): δ 7.73 (d, \(J_{HH} = 8.1\), 2H), 7.42 (d, \(J_{HH} = 8.1\), 2H), 5.62-5.83 (m, 14H), 4.83 (br s, 4H), 4.51 (br s, 2H), 4.44-4.57 (m, 2H), 4.30-4.38 (m, 2H), 4.15-4.20 (m, 1H), 3.20-3.65 (m, 40H), 2.49 (s, 3H).

\(^{13}\)C NMR (75 MHz, DMSO): δ 145.3, 133.5, 130.5, 128.2, 102.5, 82.3, 82.1, 82.0, 73.6, 73.0, 72.6, 60.5, 21.8. HRMS found 1311.3673 = [M + Na] \(^+\), calcd for \([C_{12}H_{18}O_6Na]\) \(^+\) = 1311.3678.

*Synthesis of compound 13*

Dissolved compound 12 (0.200 g, 0.155 mmol) and sodium azide (0.151 g, 0.002 mmol) in water (1 mL) and stirred reaction mixture at 80 °C for 3 d. Added acetone (10 mL) to precipitate a white solid, which was redissolved in water (10 ml) and reprecipitated with acetone. Repeated this process to obtain compound 13 as a white solid (0.151 g, 84%). \(^1\)H NMR (300 MHz, DMSO): δ 5.71-5.65 (m, 14H), 4.81 (br s, 5H), 4.44-4.40 (m, 6H), 3.62-3.29 (m, 40H). \(^{13}\)C NMR (75 MHz, DMSO): δ 102.9, 102.5, 102.2, 82.4, 82.1, 82.0, 73.6, 73.4, 73.0, 72.8, 72.6, 60.5. HRMS found = 1134.3937 [M + Na] \(^+\) calcd for \([C_{12}H_{18}O_6Na]\) \(^+\) = 1134.3930.
Synthesis of compound 14

Dissolved compound 13 (0.044 g, 0.037 mmol) in 0.5 mL DMF. Added triphenyl phosphine (0.109 g, 0.415 mmol) to the reaction mixture and stirred reaction mixture at rt for 2 h. Added 0.1 ml deionized water and refluxed the solution for 30 min. Cooled reaction mixture to rt and precipitated in acetone to obtain compound 14 (0.032 g, 76%) as white solid. $^1$H NMR (300 MHz, DMSO): 3.34 (m, 30H), 3.56 - 3.65 (m, 14H), 4.43 - 4.46 (m, 6H), 4.83 (d, $J_{H-H} = 2.0$, 6H), 4.89 (d, $J_{H-H} = 2.0$, 1H), 5.61 - 5.77 (m, 14H). $^{13}$C NMR (75 MHz, DMSO): δ 101.9, 82.8, 81.6, 81.5, 73.0, 72.4, 72.0, 59.9, 41.7. HRMS found = 1182.3640 [M + Na]$^{+}$, calcd for [C$_{12}$H$_{18}$O$_6$+Na]$^{+}$ = 1182.3654.

Synthesis of compound 17

Sodium azide (0.352 g, 5.41 mmol) was added to compound 16 (0.300 g, 0.729 mmol) in anhydrous DMF (7 mL) and the reaction mixture was heated at 80 ºC for 6 h. The reaction mixture was suspended in brine and extracted with DCM. Solvent was evaporated and the residue was dried under vacuum to yield compound 17 (0.189 g, 70%) as a white solid. R$_f$ = 0.41 (50% EtOAc in Hexane). $^1$H NMR (300 MHz, CDCl$_3$): δ 5.41-5.40 (m, 1H), 5.18-5.11 (m, 1H), 5.029 (dd, $J_{H-H} = 10.3, J_{H-H} = 8.6$, 1H), 4.58 (d, 1H), 4.16-4.13 (m, 2H), 4.02-3.97 (m, 1H), 2.15
(s, 3H), 2.07 (s, 3H), 2.04 (s, 3H), 1.97 (m, 3H). 170.3, 170.1, 169.9, 169.3, 88.1, 72.7, 70.6, 68.0, 67.2, 66.9, 61.2, 20.6, 20.5, 20.5, 20.4. HRMS found = 396.1007 [M + Na]+. calcd for [C_{12}H_{18}O_{6}+Na]^+ = 396.1013.

Synthesis of compound 18

![18]

Compound 17 (0.189 g, 0.507 mmol) was dissolved in ethanol (20 mL) and added catalytic amount of Pd/C (15.0 mg, 50% wet with water, unreduced 10% Pd) to it. The compound was hydrogenated at rt for 3.5 h. Pd/C was removed by filtration through a pad of silica and the solvent was evaporated. Residue was dried under vacuum to yield compound 18 in quantitative yield as a white solid. R_f= 0.55 (90% DCM in MeOH). \(^1\)H NMR (300 MHz, CDCl$_3$): \(\delta\) 5.38-5.37 (br m, 1H), 5.02-5.00 (m, 2H), 4.14 (d, \(J_{H-H} = 8.25\), 1H), 4.15-4.06 (m, 2H), 3.90-3.85 (m, 1H), 2.13 (s, 3H), 2.06 (s, 3H), 2.03 (s, 3H), 1.96 (s, 3H). \(^{13}\)C NMR (75 MHz, CDCl$_3$): \(\delta\) 170.4, 170.2, 170.0, 85.2, 71.3, 71.2, 69.7, 67.6, 61.7, 20.9, 20.7, 20.6, 20.5. HRMS found = 370.1104 [M + Na]+, calcd for [C_{12}H_{18}O_{6}+Na]^+ = 370.1108.

Synthesis of compound 19

![19]

Monomethyl azelate (0.174 g, 0.863 mmol) was dissolved in DCM (5 mL) and added EDC.HCl (0.198 g, 1.035 mmol) and HOBT (0.116 g, 0.836 mmol) at 0° C to it. Compound 18 (0.300 g,
0.863 mmol) and triethylamine (0.104 g, 1.036 mmol) were added to the reaction mixture at 0 °C. Stirred reaction mixture at 0 °C for 1 h. Continued stirring reaction mixture at rt overnight. Added ethyl acetate to the reaction mixture and washed organic layer with brine, 1N HCl, 5% NaHCO$_3$ and lastly with brine, consecutively. The crude product was obtained after evaporation of solvent. The crude product was purified using flash silica gel column (10%→50% EtOAc in Hexane) to obtain compound 19 (0.217 g, 47%) as a colorless oil. $R_f = 0.423$ (70% EtOAc in hexane). $^1$H NMR (300 MHz, CDCl$_3$): $\delta$ 6.23 (d, $J_{H-H} = 9.12$, 1H), 5.45-5.43 (m, 1H), 5.27-5.21 (m, 1H), 5.13-5.10 (m, 2H), 4.11-4.09 (m, 3H), 3.67 (s, 1H), 2.30 (t, $J_{H-H} = 7.35$, 1H), 2.15-2.00 (m, 6H), 1.61-1.56 (br m, 4H), 1.31-1.30 (m, 6H). $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 174.5, 174.6, 171.6, 170.7, 170.4, 170.1, 78.7, 77.9, 77.4, 77.0, 72.6, 71.2, 68.7, 67.5, 61.4, 51.8, 36.9, 34.4, 25.4, 25.1, 21.1, 21.0, 20.9, 21.0. HRMS found = 554.2197 [M + Na]$^+$, calcd for [C$_{12}$H$_{18}$O$_6$+Na]$^+$ = 554.2207.

*Synthesis of compound 20*

$\text{\includegraphics[width=4cm]{compound20.png}}$

Compound 19 (0.217 g, 0.408 mmol) was dissolved in 0.1 M methanolic NaOH (1 mL) and stirred at rt for 2.5 h. The reaction mixture was neutralized with 1N HCl. The solvent was rota evaporated and dried under vacuum to yield compound 20 (0.108 g, 76%) as a white solid. $^1$H NMR (300 MHz, CDCl$_3$): $\delta$ 3.98 (s, 1H), 3.75-3.60 (br m, 3H), 2.4-2.29 (m, 3H), 1.60 (br s, 4H), 1.13 (br s, 4H). $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 174.2, 173.4, 80.7, 77.4, 75.0, 69.0, 61.2, 52.0, 36.2, 34.1, 29.4, 29.3, 29.2, 25.7, 25.2. HRMS found = 372.1627 [M + Na]$^+$, calcd for [C$_{12}$H$_{18}$O$_6$+Na]$^+$ = 372.1628.
Synthesis of compound 15

To a solution of compound 14 (0.048 g, 0.035 mmol), in dry DMF, added HOBr (0.005 g, 0.042 mmol) and EDC (0.008 g, 0.043 mmol) at 0 °C. Added compound 20 (0.012 g, 0.035 mmol) to the reaction mixture at 0 °C. Reaction mixture was stirred at rt overnight. Crude product 15 was obtained as white solid after precipitation with acetone. HRMS found = 756.2635 [M +2+2Na]^+ /2 calcd [M +2+2Na]^+ /2 = 755.2636.
$$\text{OAc} \quad \text{OAc} \quad \text{OAc} \quad \text{OAc} \quad \text{AcO} \quad (17)$$
References


38. Weith, H. L.; Wiebers, J. L.; Gilham, P. T., Synthesis of cellulose derivatives containing
the dihydroxyboryl group and a study of their capacity to form specific complexes with sugars

39. Rosenberg, M.; Gilham, P. T., Isolation of 3’-terminal polynucleotides from RNA

40. Rosenberg, M.; Wiebers, J. L.; Gilham, P. T., Interactions of nucleotides,
polynucleotides, and nucleic acids with dihydroxyboryl-substituted celluloses. *Biochemistry*
1972, 11, 3623-3628.

41. Schott, H.; Rudloff, E.; Schmidt, P.; Roychoudhury, R.; Koessel, H., Dihydroxyboryl-
substituted methacrylic polymer for the column chromatographic separation of mononucleotides,

42. Duncan, R. E.; Gilham, P. T., Isolation of transfer RNA isoacceptors by chromatography
on dihydroxyboryl-substituted cellulose, polyacrylamide, and glass. *Anal. Biochem.* 1975, 66,
532-539.

43. Ulijn, R. V.; Bibi, N.; Jayawarna, V.; Thornton, P. D.; Todd, S. J.; Mart, R. J.; Smith, A.

44. Simon, K. A.; Sejwal, P.; Gerecht, R. B.; Luk, Y.-Y., Water-in-Water Emulsions
Stabilized by Non-Amphiphilic Interactions: Polymer-Dispersed Lyotropic Liquid Crystals.

45. Kanekiyo, Y.; Sano, M.; Iguchi, R.; Shinkai, S., Novel nucleotide-responsive hydrogels
designed from copolymers of boronic acid and cationic units and their applications as a QCM
1310.


VITA

NAME OF AUTHOR: Deepali Prashar

PLACE OF BIRTH: New Delhi, India

DATE OF BIRTH: July 3rd, 1982

GRADUATE AND UNDERGARUDUATE SCHOOL ATTENDED:
Syracuse University, Syracuse, New York
University of Delhi, New Delhi, India

DEGREES AWARDED:
Master of Philosophy in Chemistry, 2009, Syracuse University, Syracuse, New York
Master of Science in Chemistry, 2005, University of Delhi, New Delhi, India
Bachelor of Science in Chemistry, 2003, University of Delhi, New Delhi, India

AWARDS AND HONORS:
Outstanding Teaching Assistant: William D. Johnson Award, Syracuse University, 2011

PROFESSIONAL EXPERIENCE:
Teaching Assistant, Department of Chemistry, Syracuse University, 2007-2012
Research Assistant, Department of Chemistry, Syracuse University, 2007-2012