UNIVERSITY OF CALIFORNIA,
IRVINE

E2 Protein Cage as a Multifunctional Nanoplatform

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY
in Chemical and Biochemical Engineering

by

Mercè Dalmau Mallorquí

Dissertation Committee:
Professor Szu-Wen Wang, Chair
Professor Nancy A. Da Silva
Professor Kenneth J. Shea

2010
To my parents and my sister

To Alfred
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE OF CONTENTS</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xi</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>xii</td>
</tr>
<tr>
<td>CURRICULUM VITAE</td>
<td>xiv</td>
</tr>
<tr>
<td>ABSTRACT OF THE DISSERTATION</td>
<td>xvii</td>
</tr>
<tr>
<td>CHAPTER 1:  INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Motivations</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Objectives</td>
<td>8</td>
</tr>
<tr>
<td>CHAPTER 2:  LITERATURE REVIEW</td>
<td>10</td>
</tr>
<tr>
<td>2.1 Nanoparticles as drug delivery vehicles</td>
<td>11</td>
</tr>
<tr>
<td>2.2 Protein nanoscaffolds</td>
<td>15</td>
</tr>
<tr>
<td>2.3 Model scaffold: E2 subunit of pyruvate dehydrogenase</td>
<td>19</td>
</tr>
<tr>
<td>CHAPTER 3:  CHARACTERIZATION AND THERMOSTABILITY OF THE E2 SUBUNIT OF THE PYRUVATE DEHYDROGENASE MULTIENCE COMPLEX</td>
<td>23</td>
</tr>
<tr>
<td>3.1 Abstract</td>
<td>24</td>
</tr>
<tr>
<td>3.2 Introduction</td>
<td>25</td>
</tr>
<tr>
<td>3.3 Materials and Methods</td>
<td>29</td>
</tr>
<tr>
<td>3.3.1 Materials</td>
<td>29</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3.3.2 Gene Design and Construction of Plasmid</td>
<td>29</td>
</tr>
<tr>
<td>3.3.3 Protein Expression and Purification</td>
<td>31</td>
</tr>
<tr>
<td>3.3.4 Structure Confirmation</td>
<td>33</td>
</tr>
<tr>
<td>3.3.5 Thermal Stability Characterization</td>
<td>33</td>
</tr>
<tr>
<td>3.4 Results and Discussion</td>
<td>36</td>
</tr>
<tr>
<td>3.5 Conclusions</td>
<td>43</td>
</tr>
<tr>
<td>3.6 Acknowledgements</td>
<td>44</td>
</tr>
<tr>
<td>CHAPTER 4: DESIGN OF A pH-DEPENDENT ASSEMBLY MECHANISM ON A PROTEIN SCAFFOLD</td>
<td>45</td>
</tr>
<tr>
<td>4.1 Abstract</td>
<td>46</td>
</tr>
<tr>
<td>4.2 Introduction</td>
<td>47</td>
</tr>
<tr>
<td>4.3 Materials and Methods</td>
<td>51</td>
</tr>
<tr>
<td>4.3.1 Design of a pH-Dependent Molecular Switch on the E2 Scaffold</td>
<td>51</td>
</tr>
<tr>
<td>4.3.2 Mutagenesis</td>
<td>54</td>
</tr>
<tr>
<td>4.3.3 Protein Expression and Purification</td>
<td>54</td>
</tr>
<tr>
<td>4.3.4 MALDI-TOF Mass Spectrometry</td>
<td>56</td>
</tr>
<tr>
<td>4.3.5 Structure Confirmation</td>
<td>57</td>
</tr>
<tr>
<td>4.3.6 Circular Dichroism Characterization</td>
<td>57</td>
</tr>
<tr>
<td>4.3.7 Dialysis</td>
<td>58</td>
</tr>
<tr>
<td>4.3.8 Reversibility of pH-Induced Dissociation</td>
<td>58</td>
</tr>
<tr>
<td>4.4 Results and Discussion</td>
<td>59</td>
</tr>
<tr>
<td>4.5 Conclusions</td>
<td>70</td>
</tr>
</tbody>
</table>
CHAPTER 5: STUDY OF A pH-TRIGGERED DISASSEMBLY MECHANISM ON A CAGED PROTEIN COMPLEX

5.1 Abstract ..............................................................................................................72
5.2 Introduction .......................................................................................................73
5.3 Materials and Methods .....................................................................................78
  5.3.1 Materials ......................................................................................................78
  5.3.2 Design of pH-Triggered Disassembly .....................................................78
  5.3.3 Protein Expression and Purification .......................................................82
  5.3.4 Assembly Confirmation ........................................................................83
  5.3.5 Secondary Structure Characterization ...................................................84
  5.3.6 Characterization of pH-Dependent Structural Stability .......................84
5.4 Results and Discussion .....................................................................................86
5.5 Conclusions .....................................................................................................103

CHAPTER 6: MOLECULAR ENCAPSULATION WITHIN AN ENGINEERED PROTEIN CAGE

6.1 Abstract ..........................................................................................................106
6.2 Introduction .....................................................................................................107
6.3 Materials and Methods ..................................................................................111
  6.3.1 Design and Construction of Internal Mutants for Chemical Conjugation
  .............................................................................................................................111
  6.3.2 Design and Construction of Internal Hydrophobic Mutants ..............113
  6.3.3 Protein Expression and Purification .......................................................120
  6.3.4 Evaluation of Correctly Self-assembled Nanoparticles .....................121
A.10 Hydrophobic mutants scores........................................................................199

A.11 Matlab in house program to determine protein melting temperature ($T_m$) and onset of unfolding ($T_o$)...........................................................................................................................................201
## LIST OF FIGURES

| Figure 1.1 | Schematic of native E2 domain | 9 |
| Figure 2.1 | Cryo-reconstruction of the Sulfolobus turreted icosahedral virus (STIV) | 16 |
| Figure 2.2 | Model for active-site coupling in a hypothetical E1E2E3 complex | 20 |
| Figure 2.3 | Electron micrographs of core wild-type E2 negatively stained | 21 |
| Figure 3.1 | Plasmid map of pE2 expression vector | 31 |
| Figure 3.2 | SDS-PAGE showing wild-type E2 expression | 36 |
| Figure 3.3 | Size measurement of wild-type E2 scaffold | 37 |
| Figure 3.4 | Structural analysis of wild-type E2 scaffold | 38 |
| Figure 3.5 | Stability data of wild-type E2 scaffold | 39 |
| Figure 3.6 | Far-UV CD scans of E2-WT at pH 7.4 | 40 |
| Figure 4.1 | Schematic model of pH-mediated disassembly of E2 caged protein | 48 |
| Figure 4.2 | Molecular structure highlighting the domains and residues involved in the engineered pH-dependent molecular switch | 52 |
| Figure 4.3 | SDS-PAGE gel shows the soluble fraction of both wild-type and N-deletion mutant before and after purification | 56 |
| Figure 4.4 | Hydrodynamic diameters determined by dynamic light scattering | 59 |
| Figure 4.5 | Transmission electron micrographs of E2 protein stained with 2% uranyl acetate | 61 |
| Figure 4.6 | Far-UV circular dichroism spectra for protein assemblies at pH 7.4 and pH 5.0, measured at 25°C | 63 |
| Figure 4.7 | Temperature profiles of protein assemblies at pH 7.4 and pH 5.0 | 65 |
Figure 4.8  Transmission electron micrographs of E2-ΔN dialyzed from pH 5.0 back to pH 7.4 and stained with 2% uranyl acetate..................................................................................................................67

Figure 4.9  pH-unfolding profile for the E2-ΔN mutant...........................................................................68

Figure 5.1  Molecular structure highlighting relevant protein domains and mutation sites ...........................................................................................................................................79

Figure 5.2  pH-unfolding profile for E2-ΔN at two different ionic strengths. Using CD measurements to evaluate α-helical unfolding, the pH values at which 50% unfolding occurs are 7.2 and 6.2 in 100 and 500 NaCl (with 50 mM potassium phosphate buffer), respectively. ........................................................................................................................................................................88

Figure 5.3  All N-terminal mutants yield correctly-assembled E2 particles at pH 7.4 and aggregates without intact particles at pH 5.0..........................................................................................................................90

Figure 5.4  Far-UV circular dichroism spectra for N-terminal E2 mutants at pH 7.4 and 25°C ..........................................................................................................................................................91

Figure 5.5  pH-unfolding profiles for N-terminal mutants based on molar ellipticity at 208 nm (50 mM potassium phosphate, 500 mM NaCl) ..................................................................................................................................................92

Figure 5.6  Representative hydrodynamic diameter measurements using dynamic light scattering for E2-ΔN at pH 6.0 .................................................................................................................................94

Figure 5.7  pH-solubility profiles for N-terminal mutants based on the volume percentage of the DLS particle size peak for intact, soluble particles at 25 nm............94

Figure 5.8  Far-UV circular dichroism spectra for E2-WT and E2-M425H at pH 7.4 and 5.0 ........................................................................................................................................................99

Figure 5.9  Transmission electron micrographs of E2-WT and E2-M425H ..........99
Figure 5.10 Temperature stability profiles of E2-M425H protein assemblies at pH 7.4 and 5.0 ............................................................................................................................. 101

Figure 6.1 E2 model highlighting the mutagenesis sites in the internal cavity .......... 112

Figure 6.2 Hexamer from the three-fold axis of the E2 subunit highlighting the mutagenesis sites in the internal cavity ........................................................................... 114

Figure 6.3 Schematic of sucrose gradient preparation ............................................. 126

Figure 6.4 Transmission electron micrographs of mutant scaffolds showing dodecahedral structure ........................................................................................................... 130

Figure 6.5 Conjugation of fluorescein-5-maleimide (F5M) and Alexa Fluor 532 C5-maleimide (AF532M) to the D381C mutant and the wild-type control ......................... 132

Figure 6.6 Transmission electron micrographs of D381C scaffolds which have been reacted with (A) fluorescein-5-maleimide and (B) Alexa Fluor 532 C5-maleimide ...... 133

Figure 6.7 Transmission electron micrographs of internal hydrophobic mutants showing dodecahedral structure .................................................................................. 139

Figure 6.8 Representative temperature profiles for one triple and one quadruple internal mutants ............................................................................................................. 140

Figure 6.9 DCF solubility limits in 20 mM Tris/HCl buffer (pH 8.7) at room temperature. ..................................................................................................................... 142

Figure 6.10 DCF content in sucrose gradient fractions was determined by fluorescence intensity measurements ......................................................................................... 145

Figure 6.11 SDS-PAGE gel shows wild-type control without DCF and mutant loaded with DCF ............................................................................................................. 146
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 5.1</td>
<td>Summary of E2 variants created to investigate pH-dependent assembly.</td>
<td>82</td>
</tr>
<tr>
<td>Table 5.2</td>
<td>Hydrodynamic diameters (in nm) for E2-WT and N-terminal mutants (E2-ΔN, E2-ΔN/H218A, and E2-ΔN/H222A) at pH 7.4 and 5.0 (50 mM potassium phosphate, 500 mM NaCl)</td>
<td>89</td>
</tr>
<tr>
<td>Table 5.3</td>
<td>Summary of particle sizes and midpoint unfolding temperatures for E2-WT and E2-M425H at pH 7.4 and 5.0 (50 mM potassium phosphate, 100 mM NaCl)</td>
<td>98</td>
</tr>
<tr>
<td>Table 6.1</td>
<td>Oligomers used to construct hydrophobic single mutants.</td>
<td>116</td>
</tr>
<tr>
<td>Table 6.2</td>
<td>Oligomers used to construct hydrophobic double mutants.</td>
<td>117</td>
</tr>
<tr>
<td>Table 6.3</td>
<td>Oligomers used to construct hydrophobic triple mutants.</td>
<td>118</td>
</tr>
<tr>
<td>Table 6.4</td>
<td>Oligomers used to construct hydrophobic quadruple mutants.</td>
<td>119</td>
</tr>
<tr>
<td>Table 6.5</td>
<td>Volumes of each sucrose concentration used to build a discontinuous sucrose gradient.</td>
<td>127</td>
</tr>
<tr>
<td>Table 6.6</td>
<td>Summary of size and thermostability properties for wild-type (control) and single internal mutants.</td>
<td>129</td>
</tr>
<tr>
<td>Table 6.7</td>
<td>Summary of expression results at 37 and 30°C, and thermostability test of phenylalanine mutants</td>
<td>134</td>
</tr>
<tr>
<td>Table 6.8</td>
<td>Summary of size and thermostability properties for wild-type and hydrophobic mutant scaffolds</td>
<td>138</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

I want to dedicate some words of acknowledgement to those who were close to me and helped make this work possible. First I would like to thank my advisor, Prof. Szu-Wen Wang, for giving me this opportunity and her guidance through these years. I would also like to thank Prof. Nancy A. Da Silva, not only for being a member of my committee and helping me with revising my dissertation work, but also for her help during this time. Thanks to Prof. Kenneth J. Shea for his help in revising my dissertation work.

I am also thankful to Prof. Pierre Baldi and Dr. Arlo Z. Randall for collaboration in the design of internal hydrophobic mutants.

I also would like to thank Mr. Pete Balsells, for providing me with a way to attend the University of California, Irvine, through the Balsells Fellowship, and for his valuable initiative. Also thanks to Prof. Roger H. Rangel for his continuous and unconditional help.

I am also thankful to Dr. John Greaves for teaching and helping me with mass spectrometry and to Dr. Wytze van der Weer for teaching and helping me with the dynamic light scattering and circular dichroism instruments.

My most sincere gratitude to my former advisors at MIT, Dr. Jean-François Hamel, and at the Ramon Llull University-IQS, Dr. Mercedes Balcells, Prof. Rosa Nomen and Prof. Salvador Borrós. Today I am here thanks to all of them.

Thanks to the past and present members of the Wang and Da Silva groups for all their help. Special thanks to Dr. Sierin Lim for teaching me some of the biology I know,
for training me in several techniques and for all the moments we shared. Thanks to Cesar Ruiz for his work on Differential Scanning Calorimetry (DSC) and to Helen Chen for her work on the chemical conjugation of dyes to the cysteine mutant. Also thanks to Mike Shen and Matt Shindel for their valuable discussions and their support.

Thanks to the National Science Foundation and the University of California, Irvine for funding this project.

Per acabar, vull donar les gràcies als meus amics de Califòrnia, els que hi han passat i als que encara hi són. Amb tots vosaltres hem compartit molts moments que han ajudat a fer que aquesta etapa hagi esdevingut més enriquidora. Als amics de Cassà i Girona, per mantenir-me al dia des de la distància. I també a la gent de la meva etapa a Barcelona i Boston.

Finalment, l’agraïment més especial és pels meus pares Àngel i Agnès, sense els quals això no hauria estat possible. A ells els dec l’educació que m’han donat, la formació com a persona, i especialment l’estimació que han demostrat i tot el seu suport incondicional en tot moment. També vull agrair a la meva germana Anna els ànims i ajuda que sempre m’ha donat i sobretot, l’estima. I per acabar, un agraïment especial per l’Alfred, per haver estat al meu costat durant aquests anys i per tot el seu suport, paciència i estima mostrats.
CURRICULUM VITAE

Mercè Dalmau Mallorquí

Education:

2005-2010 Ph.D. in Chemical and Biochemical Engineering, University of California, Irvine
Thesis advisor: Professor Szu-Wen Wang

2005-2007 M.S. in Chemical and Biochemical Engineering, University of California, Irvine
Thesis advisor: Professor Szu-Wen Wang

2003-2004 M.S. in Chemical Engineering, Universitat Ramon Llull/Massachusetts Institute of Technology
Thesis advisors: Dr. Jean-François Hamel and Dr. Mercedes Balcells Camps

1996-2002 B.S. in Chemical Engineering, Universitat Ramon Llull

1987-1995 Bachelor of Music in Piano Performance, Isaac Albéniz Music Conservatory

Employment History:

Spring 2009 University of California, Irvine
Teaching Assistant in Separation Processes.

Winter 2007 University of California, Irvine
Teaching Assistant in Heat and Mass Transfer.
Fellowships:

2005-2006 Balsells Fellowship - Generalitat de Catalunya

Publications:


3. Dalmau M.; Lim S.; Chen H. C.; Ruiz C.; Wang S.-W.; Thermostability and molecular encapsulation within an engineered caged protein scaffold; Biotechnology

**Presentations: (* marks poster presentations)**


3. Dalmau M.; Lim S.; **Wang S.-W.**; Engineering Non-Native Functionality in a Caged Protein Scaffold; 2009 ICBE, Santa Barbara, CA.


ABSTRACT OF THE DISSERTATION

E2 Protein Cage as a Multifunctional Nanoplatform

By

Mercè Dalmau Mallorquí

Doctor of Philosophy in Chemical and Biochemical Engineering

University of California, Irvine 2010

Professor Szu-Wen Wang, Chair

Caged protein systems such as viral capsids, heat shock proteins, and ferritin are spherical structures that occur naturally in living organisms and are a growing class of biomimetic templates used to create new materials in nanotechnology. Such systems have been proposed as general drug carriers since they form highly symmetric nanoscale architectures that offer the potential to be tailored according to the desired application. Within this framework, this dissertation focuses on the design and development of a new drug delivery nanoplatform based on the E2 subunit of the pyruvate dehydrogenase protein from *Bacillus stearothermophilus*. This scaffold forms a 25-nm nanocapsule structure with a hollow cavity. We produced a variant of this protein consisting only of the structural core, and found the thermostability of this self-assembled scaffold to be unusually high, with an onset unfolding temperature of $81.1 \pm 0.9$ °C and an apparent midpoint unfolding temperature of $91.4 \pm 1.4$ °C. To evaluate the potential of this scaffold for encapsulation of guest molecules in the internal cavity, we made variants which altered the physicochemical properties of the hollow internal surface. These
mutants, yielding up to 240 mutations within this cavity, assembled into correct architectures and exhibited high thermostability that was also comparable to the wild-type scaffold. To show the applicability of this scaffold we coupled two drug-like small molecules to the internal cavity. We also developed a new strategy for encapsulation of small hydrophobic drug molecules. This method is based on hydrophobic differences between the interior cavity and the external buffer to nucleate drug-like agents inside the protein cage. We demonstrate that internal mutations can introduce non-native functionality and enable molecular encapsulation within the cavity while still retaining the dodecahedral structure. Another surface amenable to modifications is the interface between subunits. Such a region was modified to introduce pH-dependent scaffold disassembly ability to assist drug release upon endocytosis inside the cells. Moreover, we demonstrated that modulation of the pH at which disassembly occurs can be achieved by modulation of electrostatic interactions through mutagenesis or changing ionic strength. Together, these results demonstrate the potential of our scaffold as a robust nanoscale platform for biomedical applications.
CHAPTER 1

INTRODUCTION
1.1 Motivations

Nanoscale science and engineering is today one of the most dynamic technologies that integrates physical sciences, biology, materials science and medicine. Nanotechnology has provided the tools to better understand nature’s biological systems, and biology offers the models to design new synthetic materials in a wide range of applications, such as detection and treatment of illnesses, regenerative medicine, targeted delivery of drugs, nanostructure materials, and nanoelectronics (Davis et al. 2002; Padilla et al. 2001; Whitesides and Boncheva 2002). However, understanding of nature’s building blocks at the nanoscale has not yet been completely decoded.

Current methods in drug delivery present specific problems that scientists strive to identify and solve. One of the major problems that pharmaceutical companies are facing today in high-throughput drug discovery and development is the poor water solubility of small molecule drug candidates (Lipinski 2002; Radtke 2001). It has been estimated that at some pharmaceutical companies, up to 90% of these candidates show poor solubility in aqueous solutions (Stegemann et al. 2007). This issue translates to serious problems such as lower bioavailability, chemical and physical instability upon circulation inside the body, and unacceptable stability, shelf life, and storage conditions. Because of these limitations many potential drugs remain untested and are discarded for use. Many different strategies in drug delivery are being investigated to overcome these limitations. Drug delivery systems are designed to improve pharmacological properties of drugs and facilitate higher therapeutic efficiency delivery to desired targets while protecting the
body from harmful side effects through entrapment of drugs inside water-soluble scaffolds.

Traditionally, drug solubility has been enhanced quite effectively through modifications. However, these modifications can cause serious side effects. In particular, the use of surfactants is a common method that helps to increase the solubility of drug compounds, but on the other hand it is also known to cause adverse effects. An example is the case of the surfactant Cremophor® EL used in the formulation of the anticancer drug Taxol, where it is known to lead to hypersensitivity reactions (He et al. 2003). Also, the use of cosolvents or adjustment of pH can cause systemic effects. New advances in the development of artificial carriers such as micro and nanoparticles (amphiphilic micelles, polymer-based microparticles, and nanocrystals) have been shown to improve drug solubility in some cases. Nonetheless they have often shown some disadvantages like acute toxicity, drug instability, inefficient drug encapsulation and failure in controlled release (Soppimath et al. 2001; Yang and Alexandridis 2000).

Another approach used to improve drug solubilization is based on the formation of inclusion complexes between molecular assemblies and drug molecules. This is the case of cyclodextrins, a family of cyclic oligosaccharids with a hydrophilic outer surface and a lipophilic inner core that form relatively large assemblies (on the order of 50-110 µm). These molecules are commonly used in the pharmaceutical industry as molecular cages to entrap drug agents at the center cavity since they are FDA-approved and have shown to increase the aqueous solubility of poorly soluble drugs, their bioavailability and stability (Loftsson and Brewster 1996). But functionalization of these molecules is limited to the hydroxyl groups present on the outer surface.
Besides drug solubilization and bioavailability, another important functionality that a drug delivery system should possess is the ability to attack only diseased cells, diminishing severe side effects. This functionality can be easily introduced through the fusion of cell-targeting peptides to the outer surface of microparticles. Such ability has been shown to concentrate drugs at malignant cells (Harasym et al. 1998; Jaafari and Foldvari 2002; Lestini et al. 2002). However, such functionality has not yet been introduced to molecular complexing systems such as cyclodextrins.

An ideal drug delivery system should be envisioned to tune release kinetics, to regulate biodistribution and to minimize toxic side effects, thereby decreasing drug doses and improving patient compliance. More specifically, the goals of drug delivery are (1) to overcome the limitations associated with biomacromolecular therapeutics, that is insufficient stability, short plasma half-life and potential immunogenicity, and (2) to maximize the therapeutic potential while reducing the toxicity of side effects (Christie and Grainger 2003; Rosler et al. 2001).

Advantages of using nanoparticles in the range of 10-100 nm for drug delivery and non-invasive imaging are (1) nanoparticles, due to their small size, penetrate even small capillaries and are taken up inside cells (Vinogradov et al. 2002), reducing the risk of undesired elimination from the body through the liver or spleen (Svenson and Tomalia 2005) and minimizing their uptake by the reticuloendothelial system (RES) (Qiu and Bae 2006), (2) small particles have a higher surface area to volume ratios, thereby have greater dissolution rates, and (3) particles smaller than 50 nm can be internalized within endocytic vesicles (endosomes), penetrating inside the cell (Jiang et al. 2008; Zhang et al. 2009). In conclusion, a target size range for drug delivery nanoparticles is 25 to 50 nm.
The overall goal of my Ph.D. work was to introduce different functionalities on a protein cage for drug delivery applications. The research focused on a self-assembling protein scaffold, since hollow protein architectures have shown promise in many applications. For example, a heat-shock protein has been shown to demonstrate controlled encapsulation and release of a chemotherapeutic compound (Flenniken et al. 2005). Nanometer scale protein cage architectures have also been used as catalytic sites for materials synthesis (Varpness et al. 2005). Virus-based nanoparticles have been shown to have the capacity to link and display imaging moieties (such as quantum dots or fluorescent dyes) for therapeutic imaging (Manchester and Singh 2006). Therefore, in order to introduce different functionalities on protein cages and given the harsh conditions to which they are submitted in different applications, it is important to have a stable and intact assembly of the protein nanocage. Hence, my Ph.D. work studied the feasibility of introducing non-native, novel functionality, as well as the impact of these modifications on the structural stability of our engineered scaffold.

One of the functionalities that a protein cage architecture should possess for use as a drug carrier is the ability for controlled drug encapsulation. Protein cages offer their interior cavity for potential encapsulation of large amounts of therapeutic agents. This cargo is sheltered on the interior until release is selectively triggered by an external stimulus. Flenniken et al. reported successful attachment of an antitumor agent (doxorubicin) at the interior surface of a genetically modified small heat shock protein cage (Flenniken et al. 2005). Our protein system, the E2 subunit from pyruvate dehydrogenase, self-associates into a dodecahedron cage, and three-dimensional
crystallographic data shows a hollow core which can potentially be used for molecular entrapment of drug molecules (Izard et al. 1999; Milne et al. 2002). Hence, we introduced alternative physicochemical properties into the cavity, through both chemical and genetic manipulation. This allowed us to evaluate the encapsulation ability for each modification, as well as the stability properties of the whole assembly.

One of the desired functionalities of protein nanoscaffolds in drug delivery is the ability to self-assemble and disassemble. This ability makes them very unique as they can naturally assist drug release inside the cells. Since the formation of these nanostructures result from precise self-assembly of a distinct number of identical subunits held together by interactions at the interface between subunits, we proposed the study of these interactions between protein subunits at the interface in order to better understand and potentially impart control of the self-assembly of these architectures (Uchida et al. 2007). Natural caged systems show polymorphism when critical regions of each subunit (mainly at the N and C termini) are modified, and the resulting assemblies yield nanoparticles of different sizes (Calhoun et al. 2007; Hsu et al. 2006; Tang et al. 2006). In many viruses, such as CCMV, brome mosaic virus, and tomato bushy stunt virus, the N-termini of the subunits have been shown to be involved in controlling the assembly mechanism (Hsu et al. 2006; Larson et al. 2005; Tang et al. 2006). However, we are not aware of any natural or synthetic protein assembly that uses a pH-mediated strategy to control assembly. Such ability would enable release of drugs inside the cells. We desired a capsule that remains intact at neutral pH, but disassembles at pH 5.0 in the endosomes where the pH drops upon merging with lysosome during natural protein degradation. In our work, we were able to introduce a pH-dependent ability onto our
caged protein by altering the interface between subunits. This new construct was structurally intact at neutral pH and disassembled at acidic conditions, assembly behavior which has not yet been observed or engineered. This pH-dependent assembly behavior is especially important for applications in intracellular therapeutic delivery.

The dynamics and assembly of some viral cages can be modified by altering their surrounding conditions, such as ionic strength and pH. The capsid of the Cowpea chlorotic mottle virus (CCMV) undergoes a pH- and metal-ion dependent pore gating mechanism (Speir et al. 1995). Other examples include pH-dependent conformational changes on viral scaffolds which cause these capsid proteins to uncoat or the fusion with host membranes (Dryden et al. 1993; Greber 1998; McClelland et al. 2002). Hence, we further studied the pH-dependent disassembly mechanism introduced on our protein cage and its modulation via ionic strength and pH.
1.2 Objectives

The overall objective of my Ph.D. work was to engineer a generalized nanometer-scale protein assembly that possesses multiple functionalities to (1) direct encapsulation of a broad range of small drug molecules, (2) direct cage assembly/disassembly to assist drug release inside the cell, and (3) study the nature of the disassembly mechanism, to increase drug solubility and stability, retain its therapeutic activity and decrease toxicity of side effects (Figure 1.1). This research also establishes the techniques required to characterize the correct-assembly of the E2 scaffold and its thermostability. In addition, the E2 protein cage was modified at the interior cavity both chemically and genetically for encapsulation of small drug-like molecules and the effects of these modifications were evaluated. Finally, key interactions at the intersubunit interface were evaluated to enable pH-dependent disassembly, and the nature of this mechanism was studied.

The specific objectives of the work were:

1. To characterize the E2 assembly scaffold and study its thermostability.

2. To engineer a pH-dependent assembly mechanism of the E2 complex.

3. To evaluate the mechanism of pH-dependent disassembly of the E2 system and determine key interactions which allow control of assembly by variations of pH.

4. To modify the interior cavity of the E2 protein for encapsulation of small drug-like molecules inside and elucidate the effects of these modifications on the overall protein stability.
Figure 1.1 Schematic of native E2 domain. After genetic engineering the lipoyl domain and the peripheral binding subunit domain were removed to yield a core scaffold (E2-WT) and this protein cage offers three regions susceptible to be modified.
CHAPTER 2

LITERATURE REVIEW
2.1 Nanoparticles as drug delivery vehicles

Drug carriers are often envisioned as a strategy to overcome non-ideal properties of therapeutic compounds by improving their solubility, circulation time, drug release and cell-specific targeting (Allen and Cullis 2004). Besides altering pharmacokinetics (PK) and biodistribution (BD), nanoparticles preferentially accumulate in tumor tissue due to the enhanced permeability and retention effect (EPR), also called passive targeting (Allen and Cullis 2004; Duncan 1999). It has been shown that in certain pathological conditions, including tumor growth, the permeability of the vasculature increases and particulate carriers can effuse from the blood vessels and accumulate in the tissue intrastitial space (Allen and Cullis 2004; Duncan 1999). This observation is especially relevant for drug delivery systems that target tumors. As the tumor grows, it requires the formation of new blood vessels (angiogenesis). These newly formed vessels have large spaces (600-800 nm) between endothelial cells, in contrast to the tight packing of endothelial cells in normal healthy mature blood vessels (Allen and Cullis 2004). Passive targeting of drug delivery systems can result in a ten-fold increased accumulation as compared to free drug. This is due in part to the poor drainage of tumor lymphatic systems (Allen and Cullis 2004).

Many nanoparticle drug delivery systems (nanoparticles of ~200 nm or less) are composed of lipids (liposomes, micelles, lipid emulsions, lipid-drug complexes) or polymers (polymer-drug conjugates, polymer microspheres). In addition, many of them are conjugated to either polyethylene glycol (PEG) or antibodies since they have been shown to be advantageous (Allen and Cullis 2004). PEG is a hydrophilic flexible
polymer that usually improves the circulation time of conjugated therapeutics and nanoparticle systems (Allen and Cullis 2004). Other nanoscale drug delivery systems also being investigated include: silica nanoparticles, polysaccharide colloids, polyamidoamine (PAMAM) dendrimer clusters, hydrogel dextran nanoparticles, biodegradable poly(D,L-lactide-co-glycolide) (PLG) microspheres, latex and polystyrene beads, polyalkyl-cyanoacrylate nanocapsules, and nanoparticles made up of fatty acid polymers and titanium dioxide particles (Allen and Cullis 2004; Arap et al. 1998; Braslowski et al. 1990; Braslowski et al. 1991; Brigger et al. 2002; Choi et al. 2005; Ellerby et al. 1999; Emerich and Thanos 2003; Fundarò et al. 2000; Gurdag et al. 2006; Hallahan et al. 2003; Jana et al. 2002; Janes et al. 2001; Janes et al. 2001; Laakkonen et al. 2004; Laakkonen et al. 2002; Lanza et al. 2002; Müller 2001; Na and Bae 2002; Nahde T 2001; Rae et al. 2005; Roy et al. 2003; Wickline SA 2002; Winter et al. 2003; Yoo 2000). Moreover, antibody-mediated drug delivery has been shown to be successful both in laboratory and clinical trials and antibodies themselves may serve as therapeutic agents (Willner 1993). PAMAM dendrimers are repeatedly hyperbranched synthetic macromolecules with amine groups present at their molecular surface, and are currently being studied as target-specific magnetic resonance imaging (MRI) contrast agents (Kobayashi et al. 2002; Portney and Ozkan 2006). Furthermore, nanoparticle delivery of currently used anticancer drugs have been shown to achieve higher therapeutic effects due to the evasion of therapeutic resistance (Gurdag et al. 2006).

When comparing different drug delivery systems some important characteristics to be assessed are the ability to synthesize them in sufficient amounts, their therapeutic cargo capacity, their biodistribution, biocompatibility and the ability to incorporate
simultaneous functionalities at the same time. For example, liposomes have a high therapeutic drug-molecule cargo capacity (in the order of tens of thousands), however drug solubility may be a limiting factor (Allen and Cullis 2004). Hydrophilic molecules can be readily entrapped inside the aqueous interior of liposomes and hydrophobic weak bases (like doxorubicin) can be loaded through a pH or chemical gradient across the liposome bilayer, but neutral hydrophobic drugs tend to be rapidly released in the presence of cell membranes (Rice et al. 2001). One advantage of PEG-coupled therapeutic agents is that the clearance rate via the kidneys is usually decreased (Rice et al. 2001). Polymeric micelles have been proved to be suitable for water insoluble agents delivery due to their hydrophobic core (Aliabadi and Lavasanifar 2006). Antibodies are excellent cell-specific targeting agents whose biocompatibility has been significantly improved by using human-based antibodies and antibody fragments (minibodies, dibodies, etc), but their cargo capacity is only a few drug molecules per antibody (Allen and Cullis 2004; Doronina et al. 2003; Hamblett et al. 2004). One advantage of dendrimers is that they provide accessible functional groups whereas glycopolymers have their functional groups folded inside (Portney and Ozkan 2006). On the other hand, as the size of PAMAM dendrimers is increased, so does their \textit{in vitro} cytotoxicity. However, this cytotoxic effects can be diminished by PEGylation which has been shown to reduce immunogenicity and increase circulation time. Dendrimers have been used to entrap anticancer drugs (doxorubicin and methotrexate), but their release rates are slow with only 15% of the drug released after 24 hours (Cloninger 2004; Kobayashi et al. 2002; Portney and Ozkan 2006).
In summary, there exist many drug delivery systems, each having its own inherent advantages and disadvantages, many of which can be overcome through different approaches. The field of targeted therapeutic delivery is a rapidly growing field. Furthermore, every disease will require specialized drug delivery strategies. Therefore, a lot of research is focused toward the development of new drug delivery systems, but there might appear new systems which might improve some features over other systems for specific applications.
2.2 Protein nanoscaffolds

As current research seeks to design novel materials at nanoscale ranges, nature has already provided with precise nanostructures. Spherical protein cage assemblies are precisely self-assembled containers that occur naturally in different living organisms. These architectures consist of multiple copies of a specific number of protein subunits that self-assemble into 5-100 nm-diameter spherical cages. These structures play crucial biological roles. For example, viral capsids serve to house, protect, and deliver nucleic acid genomes to specific host cells. Heat shock proteins act similarly to chaperones preventing protein denaturation, and ferritins are known to store iron and release it in a controlled fashion (Harrison 1996; Narberhaus 2002). Although all these structures have evolved to perform a unique biological function, they are similar in structure. Essentially, they are protein containers with three well-defined surfaces (interior, exterior, and subunit interface) which offer potential sites to introduce different functionalities (Figure 2.1).
Figure 2.1 Cryo-reconstruction of the Sulfolobus turreted icosahedral virus (STIV). STIV is used to highlight the structural features of protein cage architectures that can be chemically and genetically modified to introduce therapeutic functionalities. Image reproduced from Uchida et al.; Biological Containers: Protein Cages as Multifunctional Nanoplatforms. Advanced Materials, 2007, 19 (8): 1025-1042. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission.

These self-assembling caged systems which form hollow cavities have shown promise in a wide range of applications (Uchida et al. 2007). These biological systems have been investigated to assess their ability, when genetically or chemically manipulated, to incorporate several functionalities without changing the overall assembly (Douglas and Young 2006). Therefore, they can incorporate simultaneously multiple functionalities within a single nanoparticle (Flenniken et al. 2005; Wang et al. 2002). This tunability together with their ideal size (between 25-50 nm) and confined hollow cavity, allow these nanoscale containers to be utilized as platforms for scaffolds and templates.

Many different examples exist to show the potential of such structures. Caged protein systems at the nanoscale, including viral capsids (Beames and Lanford 1995;
Douglas and Young 1998; Helene Braun 1999; Kathryn M. Taylor 2000; Peabody 2003; Schmidt et al. 2001; Wang et al. 2002), heat-shock proteins (Flenniken et al. 2003; Flenniken et al. 2006) and ferritin (Meldrum et al. 1991) have been proposed as general drug carriers for biomedical therapeutics and as imaging agents in medical diagnosis (Brannon-Peppas and Blanchette 2004; Flenniken et al. 2006; Manchester and Singh 2006). These self-assembling systems have been investigated to assess their ability, when genetically manipulated, to incorporate onto the exterior surface of the cage exogenous biologically active molecules (Flenniken et al. 2006; Schmidt et al. 2001). Studies demonstrate the ability of these caged architectures to nucleate and attach metal nanoparticles both inside and outside (Dixit et al. 2006; Manchester and Singh 2006; Uchida et al. 2006; Varpness et al. 2005).

Protein-based systems have great potential as platforms for drug delivery, with the advantage of a high degree of spatial organization and well-characterized self-assembly motifs that enable the control of functional group density and orientation (Gupta et al. 2005). Furthermore, it was shown that the effects of the immune system can be diminished by masking the recognition surface through PEGylation (Raja et al. 2003) or by using humanized proteins.

To date, a few investigations have shown the ability of such scaffolds to covalently attach guest molecules (fluorescent molecules or drugs) inside the protein cage (Flenniken et al. 2005; Flenniken et al. 2003; Milne et al. 2002; Milne et al. 2006). But the ability of these protein systems for simultaneously targeting cells, in vivo imaging, and encapsulating and delivering drugs inside mammalian cells has not yet been reported.
Therefore this research will focus towards the development of such platform that is able to solubilize drug molecules and deliver them in targeted cells.
2.3 Model scaffold: E2 subunit of pyruvate dehydrogenase

The overall goal of this research was to develop a protein cage architecture that serves as a therapeutic delivery system while simultaneously gaining understanding of protein cage assembly mechanism and dynamics. The template of this research was the E2 subunit of pyruvate dehydrogenase multienzyme complex from *Bacillus stearothermophilus* (Figure 2.2). The pyruvate dehydrogenase (PDH) multienzyme complex is a key component in cellular metabolism. PDH is responsible for the conversion of pyruvate to acetyl-CoA and NADH, linking the glycolysis to the citric acid cycle (Lessard et al. 1998; Perham 1991). PDH complexes with icosahedral symmetry are among the largest of cellular machines and are found in the mitochondria of eukaryotes and in Gram-positive bacteria such as *Bacillus* (Milne et al. 2006). The detailed structure of PDH complexes from eukaryotes and bacteria have been elucidated using 3-D crystal diffraction (Izard et al. 1999; Milne et al. 2002; Milne et al. 2006). The PDH complex consists of multiple copies of three different enzymes: pyruvate decarboxylase (E1), dihydrolipoamide acetyltransferase (E2), and dihydrolipoamide dehydrogenase (E3) (Figure 2.2) (Izard et al. 1999; Milne et al. 2002).

This E2 protein has many interesting properties that make it a suitable scaffold for drug delivery applications. First of all, the E2 subunit can be heterologously expressed in *E. coli* (Lessard et al. 1998). Another important feature of this system is that it self-assembles into a dodecahedron cage that consists of 60 identical subunits. Therefore, it is a robust nanostructure which has precisely defined interior and exterior surfaces. Since this protein comes from a thermophilic organism, the association of the subunits in the
multienzyme complex has been shown to be stable at elevated temperatures (Jung et al. 2002). Moreover, the E2 scaffold was tested \textit{in vivo} showing that it is stable to proteolytic degradation in serum and is able to penetrate cells (Domingo et al. 2003; Domingo et al. 2001). Furthermore, foreign peptides and proteins can be genetically fused to the outer surface without affecting the self-assembly of the E2 complex (Domingo et al. 2003; Domingo et al. 2001). Finally, when assemblies displaying different exogenous peptides are mixed, denatured, and reassembled, these different peptides can be displayed all on the same E2 scaffold (Domingo et al. 2003; Domingo et al. 2001). This ability is especially important for incorporating different functionalities at the same time.

\textbf{Figure 2.2} Model for active-site coupling in a hypothetical E1E2E3 complex. Three E1 tetramers (purple) and three E3 dimers (yellow) are shown located in the outer protein shell above the inner icosahedron (gray) formed by 60 E2 catalytic domains. Image reproduced from Milne et al.; Molecular Structure of a 9-kDa Icosahedral Pyruvate Dehydrogenase Subcomplex Containing the E2 and E3 Enzymes Using Cryoelectron Microscopy. Journal of Biological Chemistry, 2006, 281 (7): 4364-4370. Copyright The American Society for Biochemistry and Molecular Biology. Reproduced with permission.
The E2 component of the *B. stearothermophilus* PDH complex is composed of a 28-kDa catalytic (acetyltransferase) domain, a 4-kDa peripheral subunit-binding domain (PSBD), and a 9-kDa lipoyl domain, which are connected by long stretches of polypeptide chain (Milne et al. 2002; Milne et al. 2006). The E2 component self-assembles into an icosahedron (Figure 2.3) that consists of 60 identical subunits forming a spherical bulky ball-like shell with an outer diameter of 25 nm, 12 large openings of 5 nm diameter across the fivefold axes, and an inner core of 12 nm in diameter shown by crystallographic data (Izard et al. 1999; Milne et al. 2002). In our research group and others, we have shown that the PSBD and lipoyl domain can be genetically removed to yield a core scaffold domain (Dalmau et al. 2008; Domingo et al. 2001).

![Figure 2.3](image)

**Figure 2.3** Electron micrographs of core wild-type E2 negatively stained with uranyl acetate show correct assembly of 24-nm capsule. (A) A general view of a field of E2 molecules (scale bar 50nm). (B-D) Projections of E2 single molecules along the two, three, and five-fold axis. (Dalmau et al. 2008)

Since the *B. stearothermophilus* PDH complex comes from a thermophilic organism, it has an intrinsic stability under extreme conditions (Domingo et al. 2001). In
contrast, with other self-assembled spherical structures, such as icosahedral virus capsids, the attachment of foreign peptides to the surface still enables the assembly of the E2 core (Domingo et al. 2001). This property is analogous to its natural configuration where the E2 core is fused to two independently-folding domains at its surface. Therefore, the E2 scaffold shows the potential to be used as a drug carrier for drug delivery purposes, as well as an imaging agent for therapeutic diagnosis.
CHAPTER 3

CHARACTERIZATION AND THERMOSTABILITY OF THE E2 SUBUNIT OF THE PYRUVATE DEHYDROGENASE MULTIENTZYME COMPLEX

This chapter has been slightly modified and published as: Mercè Dalmau, Sierin Lim, Helen C. Chen, Cesar Ruiz, Szu-Wen Wang; Thermostability and Molecular Encapsulation Within an Engineered Caged Protein Scaffold. Biotechnology and Bioengineering, 2008, 101 (4): 654-664.
3.1 Abstract

Self-assembling biological complexes such as viral capsids have been manipulated to function in innovative nanotechnology applications. The E2 component of pyruvate dehydrogenase from *Bacillus stearothermophilus* forms a dodecahedral complex and provides a potential platform for these purposes. In this investigation, we show that this protein assembly exhibits unusual stability. To distill the core E2 protein down to its structural scaffold core, we synthesized a truncated gene optimized for expression in *Escherichia coli*. The correct assembly and dodecahedral structure of the resulting scaffold was confirmed with dynamic light scattering and transmission electron microscopy. Using circular dichroism and differential scanning calorimetry, we found the thermostability of the complex to be unusually high, with an onset temperature of unfolding at $81.1 \pm 0.9 \, ^\circ C$ and an apparent midpoint unfolding temperature of $91.4 \pm 1.4 \, ^\circ C$. The unusually robust nature of this scaffold reveal its potential for nanoscale applications.
3.2 Introduction

Self-assembling protein systems which form hollow architectures have shown promise in a wide range of nanotechnology applications (Singh et al. 2006; Uchida et al. 2007). Proteins are readily manipulated to incorporate various functionalities through chemical synthesis or recombinant DNA technology. This molecular-level tunability, together with their ideal size (~10-100 nm) and constrained hollow cavity, have enabled these nanoscale containers to serve as platforms for scaffolds and templates. Numerous examples exist for the potential of these structures. Among these applications include the encapsulation and targeted delivery of therapeutic molecules (Flenniken et al. 2005; Yang et al. 2007), loading of foreign DNA for potential gene delivery (Braun et al. 1999), and materials synthesis (Douglas and Young 1998; Kramer et al. 2004; Meldrum et al. 1991) which range in function from catalytic sites (Varpness et al. 2005) to medical imaging (Liepold et al. 2007; Uchida et al. 2006). The most commonly manipulated protein scaffolds with icosahedral symmetry have been viruses, a heat-shock protein, and ferritin (Singh et al. 2006; Uchida et al. 2007). In this work, we present a new, engineered scaffold based on the E2 subunit of the pyruvate dehydrogenase multienzyme complex. Our scaffold has been designed to include only the structural, core component of the naturally-occurring E2 subunit, but we demonstrate that it is still very stable and amenable to specific modifications.

For this general class of proteins, novel functionality is dictated by the ability to modify both the internal and external surfaces of the nanocapsule while still remaining correctly assembled. These modifications are performed by introducing amino acid
changes, by chemical modifications, or a combination of these two strategies. Internal surface modifications of the protein-based containers have been shown to expand their cargo-carrying capability to include guest entities such as drug molecules (Flenniken et al. 2005) or inorganic nanoparticles (Kramer et al. 2004). The exterior surface can be decorated with foreign peptides or molecules to enable cell-targeting ability (Flenniken et al. 2006), the attachment of metallic nanoparticles (Wang et al. 2002), or polymer-based modulation of immunogenic properties (Raja et al. 2003). To introduce multiple functionalities, dual modification of both the internal and external surfaces has been shown as well to introduce drug-like molecules inside the capsule while expressing targeting peptides (Flenniken et al. 2006) or polyethylene glycol (Kovacs et al. 2007) on the outside surface. The feasibility of multiple functionalities has also been demonstrated by the display of two different multivalent ligands on the external surface (Gillitzer et al. 2006); this takes advantage of the large number of subunits which comprise these icosahedral protein assemblies.

Not only is the ability to functionalize the scaffold important, but given the loading requirements and often relatively harsh conditions needed for materials synthesis, the stability and intact assembly of the scaffold is also imperative. Therefore, we have investigated both the feasibility of introducing non-native, functional modifications, as well as the impact of these alterations on the structural and thermal stabilities of our engineered scaffold. While the catalytic mechanisms and biology of the pyruvate dehydrogenase complex from *Bacillus stearothermophilus* has been well-studied (Allen and Perham 1997; Domingo et al. 2003; Domingo et al. 2001; Izard et al. 1999; Jung et al. 2002; Lessard et al. 1998; Milne et al. 2002; Milne et al. 2006; Perham 1991), the
potential application of its E2 structural scaffold as a molecular carrier has not been investigated.

The pyruvate dehydrogenase complex comprises three subunits (E1, E2, and E3), and together these proteins catalyze the oxidative decarboxylation of 2-oxo acids (Lessard et al. 1998; Perham 1991). In this complex, the E2 subunit forms the center core upon which E1 and E3 are bound (Milne et al. 2002; Milne et al. 2006). This central E2 protein has many attractive features as a generalizable scaffold for nanotechnology applications. For example, the E2 complex self-associates into a dodecahedron cage that consists of 60 identical subunits, and 3-dimensional crystallographic structure shows this assembly is approximately 24 nm in diameter with 12 openings of 5 nm each, leading to a hollow core which can potentially be utilized for molecular encapsulation (Izard et al. 1999; Milne et al. 2002). Since the source of this particular complex is derived from a thermophilic organism, the association of the subunits in the multi-enzyme complex has been shown to be stable under extreme environmental conditions (Jung et al. 2002). Furthermore, foreign peptides and proteins can be genetically fused to the E2 complex’s surface while still enabling the self-assembly of the complex (Domingo et al. 2003; Domingo et al. 2001). Finally, when assemblies displaying different chimeric peptides are mixed, denatured, and reassembled, these different peptides can be displayed all on the same E2 scaffold core (Domingo et al. 2003; Domingo et al. 2001); this ability is particularly important for incorporating different functionalities to the scaffold.

In this investigation, we introduce a new nanoscale protein cage as a novel platform for nanotechnology applications. Our work utilizes a synthetic E2 gene which has been truncated to its structural protein core and optimized for *E. coli* expression.
This work introduces a new caged protein platform which has the robust properties needed for applications such as nanoscale materials synthesis and molecular targeting.
3.3 Materials and Methods

3.3.1 Materials

*E. coli* strains DH5α (Zymo Research) and BL21(DE3) (Stratagene) were used as host cells. All oligonucleotides were purchased from Integrated DNA Technologies. Restriction enzymes (*HindIII*, *NdeI*, and *BamHI*), T4 DNA ligase, DNase, RNase, and calf intestinal alkaline phosphatase were obtained from New England Biolabs. *PfuUltra* High-Fidelity DNA polymerase was from Stratagene. All buffer reagents were supplied by EMD Chemicals, except for isopropyl β-D-thiogalactopyranoside (IPTG) and potassium phosphate (Fisher Scientific), phenylmethylsulfonyl fluoride (PMSF, Pierce), and sodium azide (Merck KGaA).

3.3.2 Gene Design and Construction of Plasmid

Although the full protein sequence of the E2 subunit of pyruvate dehydrogenase from *Bacillus stearothermophilus* includes three separate domains linked by flexible regions (which interact with other proteins in the enzyme complex), we desire to use this protein as a structural scaffold. Therefore, we truncated the full-length protein down to its structural core. The wild-type protein sequence of this core structure (E2-WT) was designed according to the core domain defined by Perham and coworkers (Domingo et al. 2001), and includes amino acids Pro174-Ala427 (using numbering scheme from PDB code 1b5s) preceded by an N-terminal linker sequence, LSVPG.
The DNA sequence of this wild-type E2 scaffold was optimized by CODA Genomics (Laguna Hills, CA) and the gene was synthesized as follows. In this algorithm, both synthetic gene assembly and expression in the heterologous organism (E. coli) are optimized by determining favorable oligonucleotide hybridization temperatures and codon usage, respectively (Hatfield and Roth 2007; Lathrop and Hatfield 2007). We assembled twenty-five oligonucleotides encoding the E2-WT gene into four intermediate fragments using PCR cycling with *PfuUltra* High-Fidelity DNA polymerase and CODA Blue primer mix (CODA Genomics). Each fragment was then cloned into the vector CODA Blue. These four intermediate segments were then PCR amplified, and equal amounts of each intermediate PCR product were mixed and subjected to another round of PCR extension and amplification. The resulting full-length gene was digested with *HindIII*, ligated into the CODA Blue vector, and sequenced.

To incorporate a stop codon (TAA) and *NdeI* and *BamHI* restriction endonuclease sites for cloning, the synthetic gene was PCR-amplified using *PfuUltra* High-Fidelity DNA polymerase from the assembly vector with primers 5’-gggaattccatATGCTGTCTGGTCCC–3’ (forward) and 5’-cgcggatcTTAAGCTTCCATCAGCAGCAG–3’ (reverse). The E2-WT DNA sequence is denoted in capital letters and the *NdeI* and *BamHI* sites are underlined.

To construct the expression vector pE2 containing the wild-type scaffold gene, we digested both the vector pET-11a (Novagen) and the amplified synthetic gene with *NdeI* and *BamHI*. The digested vector was treated with calf intestinal alkaline phosphatase and both vector and amplified synthetic gene were purified by agarose gel electrophoresis. We constructed the plasmid by ligating the vector backbone and synthetic gene with T4
DNA ligase. Selection of the construct was performed by transforming the ligation mixture into DH5α (E. coli) grown on LB plates containing ampicillin. DNA manipulations were performed according to standard procedures (Sambrook and Russell 2001). The final expression plasmid (pE2) containing the optimized, wild-type scaffold is shown in Figure 3.1.

![Figure 3.1](image)

**Figure 3.1** Plasmid map of pE2 expression vector. “E2-WT” encodes the synthetic E2 gene which has been truncated down to its core scaffold.

### 3.3.3 Protein Expression and Purification

Our protein was produced using a previously-reported protocol for E2 purification as a departure point (Allen and Perham 1997). *E. coli* strain BL21(DE3), transformed with the wild-type or mutant gene, was cultured in Luria-Bertani medium supplemented with ampicillin (100 μg/mL) under constant shaking at 37 °C. Five-mL from an
overnight culture (100 mL) of cells was used to inoculate 1 L of medium. Expression was induced by the addition of 1 mM IPTG when the culture reached an optical density of 0.7 – 0.9 at 600 nm. Cells were harvested three hours after induction by centrifugation at 6500 g for 15 min and stored frozen. Thawed cells were resuspended in buffer (20 mM Tris pH 8.7, 0.02% sodium azide, 2 U/mL DNase, 2 U/mL RNase, 1 mM MgCl₂ and 1 mM PMSF) and lysed in a French pressure cell (Thermo Scientific) under a pressure of 60 MPa. The insoluble fraction was removed by ultracentrifugation at 185,000 g for 1 hour.

The resulting supernatant was heated at 70 °C for 20 minutes and the native, denatured *E. coli* protein aggregates were removed by ultracentrifugation at 197,500 g for 1 hour. The recovered supernatant was loaded onto a HiPrep Q-sepharose anion exchange column (GE Healthcare) previously equilibrated with Tris buffer (20 mM Tris pH 8.7, 0.02% sodium azide, and 5 mM EDTA). Fractions containing the E2 protein were pooled, concentrated using ultrafiltration with 100 kDa molecular weight cut-off membranes (Biomax Ultrafiltration Discs, Millipore), and loaded onto a Superose 6 Prep Grade XK 16/70 size exclusion column (GE Healthcare) equilibrated with 50 mM potassium phosphate (pH 7.4), 100 mM NaCl, 0.02% sodium azide, and 5 mM EDTA. Fractions containing E2 were pooled and concentrated to a final concentration of 0.5-1 mg/mL. For long-term storage, proteins were kept at -80 °C.

Protein concentrations were determined using a Micro BCA Protein Assay Kit (Pierce; described in Appendix A.6) with bovine serum albumin as a standard. SDS-PAGE analyses were run on 12% Tris-HCl gels (described in Appendix A.5). Purified proteins (1 mg/mL) were dialyzed using drop dialysis ("V" Series Membrane, Millipore)
for 20 min with deionized water prior to MALDI-TOF analysis (Applied Biosystems Voyager-DE STR) to confirm molecular mass.

### 3.3.4 Structure Confirmation

The sizes of the resulting E2 assemblies were determined by measuring 1 mg/mL of the protein sample in phosphate buffer (50 mM potassium phosphate, 100 mM NaCl, pH 7.4) using dynamic light scattering (DLS) with a Zetasizer Nano ZS instrument (Malvern). We further confirmed correct assembly and symmetry of the protein complex with transmission electron microscopy (TEM). Protein samples (0.05 mg/mL) were negatively stained for 2 minutes with 2% uranyl acetate on carbon-coated electron microscopy grids (Ted Pella, Inc.) and images were obtained with a Philips CM-20 transmission electron microscope operating at 80 kV.

### 3.3.5 Thermal Stability Characterization

Stability of the recombinant protein was initially screened by heating the soluble fraction of freshly lysed cells to 37 °C, 55 °C, 75 °C, and 95 °C for 20 minutes. Denatured, aggregated protein was removed by centrifugation, and the remaining supernatant was loaded onto a Tris-HCl SDS-PAGE gel.

We also used far-UV circular dichroism (CD) to evaluate and quantify protein folding and thermal stability. CD scans were performed on a Jasco 810 spectropolarimeter equipped with a Jasco peltier temperature controller. Protein samples
at a concentration of 0.06 mg/mL in 50 mM potassium phosphate and 100 mM NaCl (pH 7.4) were scanned between 180-260 nm at 25 °C at a scanning speed of 10 nm/min in 0.1 cm pathlength quartz cells. For thermal unfolding, we monitored the ellipticity at 222 nm from 40 to 95 °C at a heating rate of 1 °C/min. At least three scans of each condition were obtained and results were averaged.

The onset of protein unfolding (T_o) and the apparent midpoint of the unfolding transition temperature (T_m) were determined by fitting the experimental thermal profile to standard two-state unfolding equations (Greenfield 2006). We used an in-house nonlinear least squares fitting routine programmed in MATLAB to determine the best-fit curve and to calculate the resulting first-derivative (see Appendix A.11). We defined T_m as the temperature at which 50% of the protein is unfolded (and where the first-derivative of the curve fit is a maximum) and T_o as the temperature at which 5% of the protein is unfolded. Although this thermodynamic analysis is strictly derived for two-state, reversible folding processes, it is also used to determine relative stability temperatures in irreversible processes (Greenfield 2006), and we apply it in this manner.

We also investigated thermal events of the wild-type E2 scaffold by differential scanning calorimetry (DSC). DSC scans were taken using a VP-DSC instrument (MicroCal). Solutions of protein in 20 mM Tris buffer at pH 8.7 and of buffer alone were degassed for 5 minutes, and each cell was filled with 600 µL of the sample and reference solutions, respectively. The temperature range for each scan was 40 to 100 °C. Rescans were not possible due to the irreversible aggregation of the E2 protein. The dependence of scan rate (0.5, 1, and 2 °C/min) on thermal stability, evaluated by comparing the transition temperatures (T_1, T_2), was examined using 2 mg/mL protein, and the
dependence of these temperatures on protein concentration was also studied using 1, 2, and 3 mg/mL of wild-type E2 protein. Thermal transition temperatures were determined by the DSC software Origin (version 5.0), and \( T_1 \) and \( T_2 \) were defined to be the maximum points on an endothermic peak in the temperature profile. Averages obtained reflect replicates of four scans.
3.4 Results and Discussion

Expression of wild-type E2 scaffold yields soluble complex with correct structure. The plasmid pE2, containing the optimized wild-type gene for the truncated E2 scaffold, was successfully constructed. Expression of the codon-optimized gene showed protein in both the soluble and insoluble fractions of the crude cell lysate (Figure 3.2). Although the degree of overexpression appears to result in a large fraction of proteins presented as misfolded inclusion bodies, a significant amount of the protein is also soluble. We continue purification with the soluble fraction, and obtain a final yield after purification of ~20 mg per liter of growing media. This amount could be significantly increased if the insoluble fraction is refolded and purified as well.

Figure 3.2 SDS-PAGE showing wild-type E2 expression. Lanes: (1) Protein molecular weight standards (in kDa); (2) negative control – Lysate of E. coli BL21(DE3) cells with no plasmid; (3) lysate of cells transformed with pE2 plasmid before induction; (4) lysate of cells transformed with pE2 3 h after induction with 1 mM IPTG. “P” indicates insoluble pellet, “S” indicates soluble fraction, and “E2” shows expected molecular weight at ~ 28 kDa.
The molecular mass of the E2-WT subunit was confirmed by MALDI-TOF analysis and yielded 28100 ± 6 Da, which is within 0.1% of the theoretical mass of 28,117 Da and the sensitivity of the instrument. An assembly of 60 E2 subunits has been determined by crystal structure to yield a complex with an external diameter of 24 nm (Izard et al. 1999). Using DLS, we show that the hydrodynamic diameter of our wild-type scaffold is 26.6 ± 0.6 nm (Figure 3.3), which corresponds to the predicted size. Confirmation of correct assembly is given by the dodecahedral structure observed by TEM (Figure 3.4). In these micrographs, projection views along the two, three, and five-fold axes verify that our resulting protein scaffold possesses the reported icosahedral symmetry.

![Volume vs. Size](image)

**Figure 3.3** Size measurement of wild-type E2 scaffold. Representative DLS scan shows a hydrodynamic diameter ~ 26 nm for the wild-type scaffold.
Wild-type E2 scaffold shows remarkable thermostability. Since the scaffold is based on a protein from the thermophilic source *Bacillus stearothermophilus*, we expected that it should be stable and correctly folded up until the source organism’s optimal growing temperature range, which has been reported to be 45 – 65 °C (Wu and Welker 1991). We are, however, using a significantly truncated version of the native protein, so this stability could not be assumed *a priori*. The solubility of the protein was used to generate the initial thermostability profile. Multi-component protein assemblies often exhibit irreversible unfolding upon heating (Backmann et al. 1998; Barroso et al. 2005; Blaauwen and Driessen 1998), and we observe this as well when our system is heated. As presented in Figure 3.5A, heating the protein for 20 minutes at elevated temperatures does not affect protein solubility until the regime between 75-95 °C. To quantify this transition temperature, we used CD and DSC.
Figure 3.5 Stability data of wild-type E2 scaffold demonstrates that the protein remains intact and soluble up to at least 75 °C. A: SDS-PAGE gel shows the soluble fraction of cell lysis after heating the protein to elevated temperatures for 20 min. Lanes: (1 and 7) Protein molecular weight standards (in kDa); (2) 22°C; (3) 37°C; (4) 55°C; (5) 75°C; (6) 95°C. B: Representative DSC scan of purified wild-type scaffold shows multiple thermal events starting at ~ 78 °C and indicates relatively close coupling of protein dissociation, unfolding, and aggregation.

Figure 3.5B shows a representative DSC scan of a wild-type E2 scaffold. From this thermal profile, we see that there appears to be two primary endothermic events with peaks at $T_1 = 78.8 \pm 0.8$ °C and $T_2 = 88.7 \pm 0.7$ °C, and an exothermic event at the intermediate temperature range which likely corresponds to protein aggregation. This profile shows us that the unfolding process does not fit a simple 2-state model, but that several intermediate folding states exist. This is not unexpected, as the protein complex can be identified at several levels of organization, from the overall 60-mer complex down to individual folding domains within a single subunit. Furthermore, after heating to 95 °C and cooling, we can visually observe that the resulting protein aggregation is not reversible; this is additionally supported by a DSC cooling profile which lacks noticeable thermal events after a single heating scan. Since these events are irreversible and so closely coupled, thermodynamic values cannot be accurately determined. Additional support of the kinetically-controlled nature of denaturation is seen when protein
concentrations and scan rates are varied, resulting in shifts in peak $T_1$ and $T_2$ temperature values. This DSC data clearly shows that the truncated E2-WT scaffold is stable at least until an onset transition temperature at approximately 78 °C.

The crystal structure of the E2 subunit reveals approximately one-third of the protein to be $\alpha$-helical, which is also its predominant secondary structure (PDB file 1b5s). This is confirmed by a far-UV CD scan of our E2-WT scaffold, which shows the characteristic spectrum and minima at 208 and 222 nm for $\alpha$-helices (Figure 3.6A). Using the presence of $\alpha$-helices as a gauge for overall protein folding, we can estimate the apparent midpoint of protein unfolding ($T_m$) and onset temperature of unfolding ($T_o$) by monitoring the reduction of the ellipticity at 222 nm as temperature is increased (Figure 3.6B). With the unfolding model and definitions described in the Materials and Methods section, we determined $T_o$ to be 81.1 ± 0.9°C and $T_m$ to be 91.4 ± 1.4 °C.

![Figure 3.6](image)

**Figure 3.6** Far-UV CD scans of E2-WT at pH 7.4. **A:** Molar ellipticity versus wavelength. This spectrum is an average of spectra obtained from three E2-WT samples at room temperature, and shows the characteristic minima at 208 nm and 22 nm for a protein rich in $\alpha$–helices. **B:** Temperature profile of one sample of E2-WT. Ellipticity at 222 nm is measured as sample is heated and indicates a high $T_m$ at approximately 90 °C.
Our DSC data shows distinct dissociation, aggregation, and unfolding events, and it also reveals that these events occur relatively close together in temperature. Our first-derivative analysis of the CD folding profile, however, yields only one maximum (at $T_m$), which suggests that there is only one apparent folding transition (Greenfield 2006). Together, this data demonstrates that the individual thermal events observed by DSC occur rapidly in succession such that the process of unfolding is essentially continuous. Therefore, we are still able to fit the unfolding profile quite well to the 2-state model to obtain estimates of midpoint unfolding temperatures.

Both our CD and DSC results indicate remarkable thermostability of the protein scaffold, with the onset of dissociation and unfolding starting between ~78-81°C. The exact temperature of the thermal events varies by a few degrees, depending on the determination method used, and these differences can be explained by several factors. First, using DSC we found that the protein concentration contributes to differences in the temperature profiles, with protein concentrations at 1 mg/mL exhibiting an increased shift in $T_1$ and $T_2$ by approximately 4 ºC (relative to 2 mg/mL), indicating higher thermostability at lower protein concentrations. However, the sensitivity of DSC and CD requires vastly different protein concentrations of 1 mg/mL and 0.06 mg/mL, respectively. We would expect, therefore, that CD data will reflect a higher stability than DSC. Furthermore, we observed that the protein was more stable at near-physiological ionic strengths than in solutions with low salt, as would be expected by increasing electrostatic screening. This can also account for the higher thermostability seen with CD, since CD data was obtained at 100 mM NaCl in phosphate buffer while the protein solution for DSC contained no salt. Despite these minor differences in the stability
temperatures, the general overall conclusion of unusually high thermal stability in our scaffold is confirmed by both methods, and the general temperature range of stability is also consistent with each other.
3.5 Conclusions

Our investigations demonstrate that the E2 structural scaffold from the multienzyme complex pyruvate dehydrogenase is robust and has the characteristics of an ideal caged-protein platform for molecular transport and encapsulation. We confirm that the truncated core scaffold exhibiting wild-type sequence self-assembles into a structure with icosahedral symmetry and is highly thermostable. The high thermostability of this system, together with previously-demonstrated ability to attach and display multiple foreign peptides on the outside surface (Domingo et al. 2001), suggest that this platform can potentially be a powerful and versatile tool in bionanotechnology, with potential applications in broad areas such as materials synthesis and therapeutics.
3.6 Acknowledgements

I gratefully acknowledge Dr. Sierin Lim for the construction of the E2 gene and development of expression and purification protocols, Cesar Ruiz for performing the thermostability studies with Differential Scanning Calorimetry (DSC).
CHAPTER 4

DESIGN OF A pH-DEPENDENT ASSEMBLY MECHANISM ON A PROTEIN SCAFFOLD

This chapter has been slightly modified and published as: Mercè Dalmau, Sierin Lim, Szu-Wen Wang; Design of a pH-Dependent Molecular Switch in a Caged Protein Platform. Nano Letters, 2009, 9 (1): 160-166.
4.1 Abstract

Self-assembling protein cages provide a wide range of possible applications in nanotechnology. We report the first example of an engineered pH-dependent molecular switch in a virus-like particle. By genetically manipulating the subunit-subunit interface of the E2 subunit of pyruvate dehydrogenase, we introduce pH-responsive assembly into a scaffold that is natively stable at both pH 5.0 and 7.4. The redesigned protein module yields an intact, stable particle at pH 7.4 which dissociates at pH 5.0. This triggered behavior is especially relevant for applications in therapeutic delivery.
4.2 Introduction

Viruses and viral-like particles are a growing class of biomimetic templates used to create and functionalize new materials in nanotechnology. Such systems are potentially flexible in application, as these macromolecular systems are defined in size, can range from approximately ten nanometers to tens-of-nanometers (depending on the specific virus), self-assemble from smaller identical subunits, and can be specifically altered through genetic or chemical approaches (Singh et al. 2006; Uchida et al. 2007). Several studies have demonstrated their utility as constrained containers for synthesizing inorganic nanoparticles (Douglas and Young 1998; Kramer et al. 2004; Meldrum et al. 1991; Slocik et al. 2005). These investigations show that the resulting particles can be used in applications spanning widely from catalysis for hydrogen production (Varpness et al. 2005) to contrast agents for medical imaging (Liepold et al. 2007). Therapeutic molecules have also been confined to the interior compartment of the scaffolds and these resulting particles have been explored in targeted drug delivery (Flenniken et al. 2006).

In all these studies of caged protein systems, the functionality has been imparted by altering either the interior or exterior surfaces. Conjugating polymers to the exterior surface, for example, can help modulate a particle’s immunogenic response (Raja et al. 2003), and attaching specific metal-binding peptides to the interior cavity creates nucleation sites for inorganic nanoparticle synthesis (Klem et al. 2005; Slocik et al. 2005). As recently pointed out by Douglas and coworkers, since the formation of these icosahedral nanostructures result from precise self-assembly of modular subunits, redesign of the subunit-subunit interface is another potential avenue for controlling their
architecture and assembly behavior (Uchida et al. 2007). However, the only type of behavior which we are aware of that uses this strategy to redirect natural assembly is polymorphism, resulting in a virus particle with a smaller volume (Calhoun et al. 2007; Hsu et al. 2006; Tang et al. 2006).

In this investigation, we alter the interface between subunits of a virus-like particle and demonstrate that we can introduce a pH-dependent molecular switch into a caged protein that is normally stable across physiologically relevant pH values. This redesigned scaffold, which is structurally intact and unusually stable at pH 7.4, will disassemble at pH 5.0 (Figure 4.1). This pH profile is particularly important for applications in drug delivery, as macromolecules entering a cell via the endocytotic pathway will encounter an environment initially at pH 7.4 that is eventually acidified to pH 5.0 in the lysosome (Christie and Grainger 2003). Disassembly dependent on the pH environment could modulate important parameters such as drug release and endosomal escape.

![Figure 4.1](image)

**Figure 4.1** Schematic model of pH-mediated disassembly of E2 caged protein. At pH 7.4, nanoparticles with icosahedral symmetry remain self-assembled, and at pH 5.0 they irreversibly disassemble.
The dynamics and assembly of some natural virus capsids can be altered by modulating environmental conditions such as ionic strength and pH. For example, the capsids of the cowpea chlorotic mottle virus (CCMV) and Norwalk virus are known to disassemble into monomers at high pH and reassemble at low pH (Ausar et al. 2006; Bancroft et al. 1967; Johnson and Speir 1997). CCMV also exhibits a pH-dependent pore-gating mechanism (Johnson and Speir 1997). Other natural pH-mediated processes in viruses include conformational changes on external scaffolding shells, which allow these capsid-supporting proteins to uncoat or the viral envelope to fuse with host membranes (Dryden et al. 1993; Greber 1998; McClelland et al. 2002). However, we are not aware of any viral capsids or caged proteins (natural or engineered) that remain assembled at pH 7.4 but dissociate at pH 5.0.

Our model protein is the enzyme dihydrolipoyl acetyltransferase (E2) of the pyruvate dehydrogenase multienzyme complex from *Bacillus stearothermophilus* (Allen and Perham 1997; Perham 1991). The E2 subunit forms the central core and binds two other peripheral enzymes to form the multienzyme complex that catalyzes the oxidative decarboxylation of 2-oxo acids. Sixty of these identical E2 subunits will associate to form a hollow dodecahedron with icosahedral 532 symmetry (Izard et al. 1999). In recent work, we described an E2 scaffold that includes only the truncated, core structure (with the peripheral enzyme-binding regions deleted), and demonstrated that this scaffold and its variants are unusually stable up to approximately 85°C (Dalmau et al. 2008). Its potential for transporting guest molecules was illustrated by encapsulating drug-like molecules in the hollow cavity without detriment to its structural integrity. In the current
work, we use this truncated core structure as the template upon which we introduce pH-dependent assembly.
4.3 Materials and Methods

4.3.1 Design of a pH-Dependent Molecular Switch on the E2 Scaffold

To design an E2 scaffold that would enable triggered disassembly, we examined the subunit-subunit interfaces of naturally-occurring virus capsids which have been reported to exist in alternate architectural configurations under different conditions. One paradigm was recurrent among numerous viruses; these were examples in which the amino terminal (N-terminal) end of virus capsid subunits is involved in assembly mechanisms (Calhoun et al. 2007; Dong et al. 1998; Hsu et al. 2006; Kumar et al. 1997; Larson et al. 2005; Sangita et al. 2004; Savithri and Erickson 1983; Schmitz and Rao 1998; Tang et al. 2006). The N-terminal domains in these viruses modulate the icosahedral architecture; all these viral capsids consist of 180 subunits (to form a T=3 structure), but deletion of this N-terminal domain result in either no assembly, or more remarkably, a smaller (T=1) structure comprising 60 subunits. These terminal peptides are usually disordered in crystallographic structure, revealing a dynamic nature. This flexibility is supported by the observation that the N-terminal regions can be removed by endogenous proteolytic degradation. When these regions are resolved in the crystallographic data, they are shown to embrace an adjacent subunit (Calhoun et al. 2007; Hsu et al. 2006; Tang et al. 2006). The N-terminal domain in this class of viruses directs the dynamic assembly of the capsid structure.

Although our E2 protein complex is unrelated to viral capsids with respect to both sequence homology and natural function, it surprisingly has features which are similar to
viruses exhibiting this N-terminal control of assembly. For example, the first 11 N-terminal residues of the truncated, core E2 scaffold are not resolved in the crystallographic structure, presumably due to its flexible nature. The structural data of the portion of this domain which is available enfolds an adjacent subunit within a trimer cluster (see Figure 4.2) (Izard et al. 1999). Furthermore, we have observed that the N-terminus of our E2 scaffold is also prone to endogenous proteolysis in insufficiently purified cell lysates. Despite proteolysis the dodecahedron is still able to assemble correctly. These features suggest that there are structural parallels between the E2 scaffold and viral particles and that selective deletion of this N-terminal arm could also modulate assembly and architecture.

Figure 4.2  Molecular structure highlighting the domains and residues involved in the engineered pH-dependent molecular switch. (A) Monomer subunit of E2 (green) shows His218 (red), His222 (blue), and the N-terminal residues (184-217; purple) that embrace an adjacent subunit (grey) and that are deleted in the mutant. Close-up view of the histidine trimers involved in the pH-controlled disassembly (B) at the 3-fold axis of symmetry and (C) at a 90° side view relative to the 3-fold axis of symmetry. The structures are displayed using PyMOL (DeLano, 2002) and PDB crystallographic data (1b5s).
Since we desired to create a pH-dependent scaffold, we needed to incorporate interactions that retain stability at neutral pH but destabilize the complex at pH 5.0. We hypothesized that this could be accomplished by introducing adjacent histidines at the subunit-subunit interfaces, as the pKa of this amino acid is typically reported to be near 6.5 within proteins (Stryer 1988). Therefore, histidine is likely to be neutral at pH 7.4, but its positive charge at pH 5.0 would yield repulsive interactions at sufficiently close range. Such a strategy of introducing imidazole (the chemical moiety of histidine) has been explored previously in polymeric condensation and endosomal escape in gene delivery (Park et al. 2006; Putnam et al. 2001).

Upon examination of the crystallographic structure, we noted two native histidines (His218 and His222) that both converge around the 3-fold axis. The structure also shows that the imidazole groups of both His218 and His222 are geometrically within an estimated 0.6 nm proximity of the other two respective histidines in their trimeric cluster (see Figures 4.2B, C). We desired to retain the potential interactions of these two sets of histidine triplets, and we therefore truncated the N-terminus at amino acid 217 using genetic engineering. This N-terminal deletion protein (E2-ΔN, spanning amino acids His218-Ala427) and the full-length wild-type protein scaffold (E2-WT control, spanning amino acids Pro174-Ala427 and preceded by a linker peptide as previously described) (Dalmau et al. 2008; Domingo et al. 2001) were expressed in *E. coli* and purified. To assess the purity of the recombinant protein and the correct assembly of 60 subunits into a caged structure with icosahedral symmetry at pH 7.4, the resulting protein capsules were characterized by size exclusion chromatography, SDS polyacrylamide gel
electrophoresis, MALDI-TOF mass spectrometry, dynamic light scattering (DLS), and
transmission electron microscopy (TEM).

4.3.2 Mutagenesis

To construct the N-terminal deletion mutant (E2-ΔN), the gene encoding the wild-
type scaffold (E2-WT) was PCR-amplified using \textit{PfuUltra} High-Fidelity DNA
polymerase with the following oligonucleotides: 5’–
gggaattccatATGCACACCGCGACACGTTACC-3’ (forward) and
5’-cgccggatccTTAAGCTTCCATCAGCAGCAG–3’ (reverse). The resulting clone was
screened on LB + ampicillin plates and cloned into the expression vector pET-11a at the
\textit{NdeI} and \textit{BamHI} sites (underlined above). To construct the expression vector pE2NΔ50,
we digested the vector pET-11a and the amplified gene with \textit{NdeI} and \textit{BamHI} restriction
enzymes. Both vector and synthetic gene were purified with agarose gel electrophoresis
and ligated with T4 DNA ligase. The sequence of the mutant E2-ΔN gene was confirmed
by DNA sequencing. Details of standard molecular cloning techniques follow those
described by Sambrook & Russell (Sambrook and Russell 2001).

4.3.3 Protein Expression and Purification

One-liter culture of \textit{E. coli} strain BL21(DE3) (Stratagene), transformed with
plasmid containing either the E2-WT or E2-ΔN gene, was cultured in LB medium
supplemented with 100 µg/mL ampicillin at 37 °C and 30 °C for the wild-type and
mutant, respectively. Five-mL from an overnight culture (100 mL) of cells was used to inoculate 1 L of medium. Expression was induced by the addition of 1 mM IPTG when the culture reached an optical density of 0.7 – 0.9 at 600 nm. Cells were harvested three hours after induction by centrifugation at 6500 g for 15 min and stored at -80°C. Thawed cells were resuspended in buffer (20 mM Tris pH 8.7, 0.02% sodium azide, 2 U/mL DNase, 2 U/mL RNase, 1 mM MgCl₂ and 1 mM PMSF) and lysed in a French pressure cell at 60 MPa (Thermo Scientific). The insoluble fraction was removed by ultracentrifugation at 185,000 for 1 hour.

The resulting supernatant was loaded onto a HiPrep Q-sepharose anion exchange column (GE Healthcare) previously equilibrated with Tris buffer (20 mM Tris pH 8.7, 0.02% sodium azide, and 5 mM EDTA). Fractions containing the E2 protein were pooled, concentrated using ultrafiltration with 100 kDa molecular weight cut-off membranes (Biomax Ultrafiltration Discs, Millipore), and loaded onto a Superose 6 Prep Grade XK 16/70 size exclusion column (GE Healthcare) equilibrated with 50 mM potassium phosphate (pH 7.4), 100 mM NaCl, 0.02% sodium azide, and 5 mM EDTA. Fractions containing E2 were pooled and concentrated to a final protein concentration of about 0.5 mg/mL. For long-term storage, proteins were kept in 50 mM potassium phosphate (pH 7.4), 250 mM NaCl, 0.02% sodium azide, and 5 mM EDTA.

Protein concentrations were determined using a Micro BCA Protein Assay Kit (Pierce) with bovine serum albumin as a standard. SDS-PAGE analyses were run on 12% Tris-HCl gels, and typical purification results are shown below in Figure 4.3.
**Figure 4.3** SDS-PAGE gel shows the soluble fraction of both wild-type and N-deletion mutant before and after purification. Lanes: (1 & 6) Protein molecular weight standards; (2) purified E2-ΔN; (3) soluble fraction of E2-ΔN before purification; (4) purified E2-WT; (5) soluble fraction of E2-WT before purification.

4.3.4 MALDI-TOF Mass Spectrometry

To determine purity and molecular weights of the proteins, we used MALDI-TOF mass spectrometry. Purified protein (1 mg/mL) was dialyzed into deionized water prior to mass spectrometry analysis. The experimental molecular weight of a E2-WT subunit was 28101 ± 10 Daltons, compared to the theoretical value of 28117 Daltons. The average measured mass of the E2-ΔN mutant was 23045 ± 7 Daltons while theoretically expected mass is 23073 Daltons. Measured and theoretical molecular weights are within 0.1%, which is the accuracy of the instrument, and confirm correct protein mass.
4.3.5 Structure Confirmation

The hydrodynamic diameters of the E2 scaffolds were determined by dynamic light scattering (DLS) with a Zetasizer Nano ZS instrument (Malvern) at a protein concentration of 1 mg/mL in phosphate buffer (50 mM potassium phosphate, 250 mM NaCl, pH 7.4). We further confirmed correct assembly of the protein complexes with transmission electron microscopy (TEM). Protein samples (0.05 mg/mL) were negatively stained for 2 minutes with 2% uranyl acetate on carbon-coated electron microscopy grids (Ted Pella, Inc.) and images were obtained with a Philips CM-20 transmission electron microscope operating at 80 kV.

4.3.6 Circular Dichroism Characterization

We used far-UV circular dichroism (CD) to evaluate and quantify protein folding/unfolding and thermal stability. CD scans were performed on a Jasco 810 spectropolarimeter equipped with a Jasco peltier temperature controller. Protein samples at a concentration of 0.06 mg/mL in 50 mM potassium phosphate and 250 mM NaCl (pH 7.4) were scanned between 180-260 nm at 25 °C at a scanning speed of 10 nm/min in 0.1 cm path length quartz cells. For thermal unfolding, we monitored the ellipticity at 222 nm from 40 to 95 °C at a heating rate of 1 °C/min. For pH-mediated unfolding, we used ellipticity at 208 nm over pH 5-8. At least three scans of each condition were obtained and results were averaged.

The apparent midpoint of unfolding temperature (T_m) was determined by fitting the experimental thermal profile to standard two-state unfolding equations (Greenfield
We used an in-house nonlinear least squares fitting routine programmed in MATLAB to determine the best-fit curve and to calculate the resulting first-derivative (see Appendix A.11). We defined $T_m$ as the temperature at which 50% of the protein is unfolded (and where the first-derivative of the curve fit is a maximum).

### 4.3.7 Dialysis

Buffer exchange to alter pH was performed with drop dialysis ("V" Series Membrane, Millipore) of the protein solutions (1 mg/mL protein in 50 mM potassium phosphate and 250 mM sodium chloride with pH adjusted) in an enclosed, hydrated dish for 20 min.

### 4.3.8 Reversibility of pH-Induced Dissociation

To test the reversibility of the unfolding process, protein solutions were dialyzed to pH 5.0, analyzed, and then dialyzed back to the initial pH conditions (pH 7.4). Figure 4.4 shows an electron micrograph of one sample of E2-$\Delta$N that was dialyzed back from pH 5.0 to pH 7.4 and confirms the irreversibility of the pH-induced unfolding process.
4.4 Results and Discussion

Caged nanocapsules consisting of the truncated N-terminal mutant are correctly assembled at pH 7.4 but dissociate at pH 5.0. After purification, E2-WT and E2-ΔN at 1 mg/mL were stored in phosphate buffer at pH 7.4 (50 mM potassium phosphate, 250 mM NaCl). To investigate the E2-ΔN design for its predicted pH-sensitive behavior, we dialyzed the ΔN-E2 mutant and the E2-WT control into pH 5.0 buffer (50 mM potassium phosphate, 250 mM NaCl) according to protocol described in section 4.3.7. The hydrodynamic diameters (measured by DLS) for the E2-WT control were 27.4 ± 0.7 nm and 28.0 ± 0.8 nm at pH 7.4 and pH 5.0, respectively (Figure 4.4), which are consistent with previous values for the baseline wild-type scaffold (Dalmau et al. 2008). This invariance in particle size after the pH-shift suggests that E2-WT remains correctly assembled as a dodecahedron and intact at the different pH values.

![Figure 4.4](image)

**Figure 4.4** Hydrodynamic diameters determined by dynamic light scattering. The unaltered scaffold (E2-WT) yields similar particle sizes at pH 7.4 (27.4 ± 0.7 nm) relative to pH 5.0 (28.0 ± 0.8 nm). The mutant protein (E2-ΔN) shows a particle size of 32.4 ± 3.0 nm at pH 7.4 but aggregates at pH 5.0 (1843 ± 818 nm).
In contrast, characterization of the hydrodynamic diameter of E2-ΔN after dialysis to pH 5 revealed a significant increase in size for the N-terminal deletion protein, from 32.4 ± 3.0 nm (at pH 7.4) to greater than 1 µm (at pH 5.0) (Figure 4.4). It should be noted that the numerical value at the acidic pH (1843 ± 818 nm) is inexact due to the high polydispersity observed, but it is clear that aggregates exist since they are a couple of orders of magnitude larger than soluble, correctly-assembled particles with icosahedral symmetry. Visual inspection of the sample at pH 5.0 shows aggregation as well, with the solution showing turbidity. We also note that although the hydrodynamic radius of the deletion mutant at pH 7.4 is slightly larger than that for E2-WT, it appears that deviations are related to batch-to-batch variability in purification. Therefore, the large size disparity between E2-ΔN at pH 7.4 and pH 5.0 support the aggregation of E2-ΔN at acidic pH. To confirm that the aggregation of proteins is due to disassembly of the caged protein particles and not simply precipitation or crystallization of intact particles, we used circular dichroism (CD) and TEM to probe the proteins’ secondary and quaternary structures, respectively.

The wild-type and mutant proteins at pH 7.4 and pH 5.0 were imaged with TEM (Figure 4.6). Intact assembly and correct architecture of the E2-WT control were seen at both pH 7.4 and 5.0 (Figures 4.6A, B), supporting the DLS data. The E2-ΔN mutant also showed distinct, correctly assembled dodecahedral particles at pH 7.4, as expected from the hydrodynamic diameter measurements (Figure 4.5C). However, at pH 5.0, extensive aggregation of proteinaceous material was observed without clearly intact particles (Figure 4.5D). These images are similar to those of E2-WT samples at pH 7.4 which have been heated to 95 °C (Figure 4.5E), conditions which have been determined by
differential scanning calorimetry and circular dichroism to result in subunit-subunit
dissociation, protein unfolding, and aggregation (Dalmau et al. 2008).

**Figure 4.5** Transmission electron micrographs of E2 protein stained with 2% uranyl acetate. (A) E2-WT at pH 7.4, (B) E2-WT at pH 5.0, and (C) E2-ΔN at pH 7.4 show intact capsules, while (D) E2-ΔN at pH 5.0 shows aggregation. (E) Thermally-denatured E2-WT exhibit unfolding and aggregation which appears similar to the N-terminal mutant at pH 5.0. (F) E2-ΔN at an intermediate pH of 6.5. (G) E2-ΔN at an intermediate pH of 7.0 shows aggregates that are (a) primarily unfolded protein agglomerates and (b) some regions of proteins which are neither fully intact nor fully denatured. Scale bar is 50 nm.
In contrast to images obtained at pH 5.0 and pH 7.4, the intermediate pH of 7.0 (Figure 4.5G) shows partially assembled, but not completely intact, particles in coexistence with large aggregates, similar to those at pH 5.0, are observed (Figure 4.5F). The distinction between images of E2-ΔN at pH 5.0, 7.4, and 7.0 demonstrates that the insoluble proteins observed at pH 5.0 are not simply aggregates of fully formed particles or partially intact particles but are aggregates of disassembled protein complexes.

We used far-UV circular dichroism spectroscopy (CD) to confirm further and quantify the pH-triggered dissociation of the N-terminal deletion mutant. Techniques such as gel electrophoresis or size exclusion chromatography for determination of subunit dissociation require that the monomeric units are soluble. Since disassembly is closely coupled to irreversible aggregation in our protein (Dalmau et al. 2008) (as is typical of multimeric protein complexes) (Backmann et al. 1998; Barroso et al. 2005), we required an alternate technique for probing disassembly (Izard et al. 1999). CD is commonly used to detect unfolding of proteins with α-helices, since α-helical secondary structure demands the presence of precise interactions and conformational geometries (Berova et al. 2000; Creighton 1993; Martin et al. 2008). Furthermore, this technique can be used to evaluate denaturation and aggregation of insoluble proteins (Benjwal et al. 2006; Calloni et al. 2008). Studies which correlate CD with other techniques that probe subunit dissociation (such as size exclusion chromatography, gel electrophoresis, fluorescence spectroscopy, or nuclear magnetic resonance) have established that CD can also be used as a method of evaluating multimer dissociation (Chiaraluce et al. 1997; Oliveira et al. 2000). Therefore, since the crystallographic structure of the E2 protein cage reveals that it is approximately one-third α-helical (PDB file 1b5s), we probed protein unfolding and
disassembly of the caged nanoparticles by monitoring $\alpha$-helical content via circular dichroism.

Figure 4.6 shows ellipticity as a function of wavelength for both the E2-WT control and the E2-$\Delta$N mutant at pH 7.4 and 5.0. As expected for an intact, fully-folded assembly, the spectra for E2-WT at both pH conditions show the characteristic profile of an $\alpha$-helical-rich protein with minima at 208 and 222 nm (Figure 4.6A). Together with the TEM images, this confirms that E2-WT is correctly assembled at both pH values, and that no secondary structural changes are observed upon dialysis into lower pH. This set of data demonstrates that E2-WT is stable at both pH values.

Figure 4.6 Far-UV circular dichroism spectra for protein assemblies at pH 7.4 and pH 5.0, measured at 25°C. (A) Molar ellipticity vs. wavelength for E2-WT. Correctly-assembled E2-WT shows the two characteristic minima at 208 nm and 222 nm for proteins high in $\alpha$-helical content. (B) Molar ellipticity vs. wavelength for pH-sensitive mutant. E2-$\Delta$N reveals the loss of the 208-nm minimum and a shift in the 222-nm minimum at pH 5.0 relative to pH 7.4, which is indicative of unfolding. All spectra are an average of data from three samples.
For E2-ΔN at pH 7.4, we also observe a similar CD profile with minima at 208 nm and 222 nm (Figure 4.6B), indicating that the protein is folded at normal physiological pH. However, data at pH 5.0 reveal a noticeable level of unfolding of the protein subunits. The level of α-helicity decreases, which is evidenced by the significantly altered spectra; the characteristic 208-nm minimum is lost and the 222-nm minimum is shifted (Figure 4.6B). Our set of data therefore demonstrates that our engineered, pH-triggered E2-ΔN mutant remains fully folded and assembled at pH 7.4, but shows unfolding at pH 5.0. Using differential scanning calorimetry, it was previously found that the E2 assembly’s unfolding is closely coupled to dissociation and aggregation, and that these events occur rapidly in succession (Dalmau et al. 2008). Hence, we expect that unfolding in E2-ΔN is linked to dissociation and aggregation as well.

The truncated N-terminal mutant is highly thermostable at pH 7.4. Virus-like particles have been employed in applications such as materials synthesis and therapeutic delivery (Douglas and Young 1998; Flenniken et al. 2006; Kramer et al. 2004; Liepold et al. 2007; Meldrum et al. 1991; Slocik et al. 2005), tasks in which the extent of robustness could limit utility at harsh conditions. Therefore, it is important for a scaffold which exhibits triggered behavior to be stable under conditions prior to the desired activated response. In designing our pH-sensitive complex, we have removed a domain on the subunit-subunit interface which is anticipated to modulate assembly; hence, we expected that this deletion would likely destabilize the complex. To investigate the thermostability of E2-ΔN relative to E2-WT, we monitored the ellipticity
at 222 nm as the protein is heated and calculated apparent midpoint unfolding temperatures ($T_m$) (Dalmau et al. 2008).

Previous work demonstrated that the E2-WT scaffold is unusually stable (Dalmau et al. 2008b), and our investigation here shows this as well. For the wild-type control, we observe unfolding transitions ($T_m$) of $91.2 \pm 0.9 \, ^\circ C$ and $85.1 \pm 1.1 \, ^\circ C$ at pH 7.4 and 5.0, respectively (Figure 4.7A). Above these melting temperatures, the proteins unfold and subsequently aggregate, as demonstrated by TEM (Figure 4.5E) and calorimetry (Dalmau et al. 2008). While the stability of this scaffold at pH 5.0 is lower than at pH 7.4, the curve is still sigmoidal with a definite $T_m$, and the decrease in $T_m$ is only approximately 6 °C. Overall, this is a small shift relative to its deviation from the physiological temperature of 37 °C. This data then confirms that the E2-WT scaffold is very thermostable at both pH 5 and pH 7.4.

**Figure 4.7** Temperature profiles of protein assemblies at pH 7.4 and pH 5.0. Molar ellipticity is measured to evaluate the level of unfolding. (A) E2-WT shows an apparent unfolding transition temperature of $91.2 \pm 0.9^\circ C$ and $85.1 \pm 1.1^\circ C$ at pH 7.4 and pH 5.0, respectively. (B) E2-$\Delta$N also exhibits a high midpoint unfolding temperature at pH 7.4 ($88.0 \pm 1.0^\circ C$), but shows no apparent unfolding events at pH 5.0. Profiles are an average of data from three samples.
For the pH-responsive E2-ΔN mutant, the thermal profile at pH 7.4 is also a sigmoidal curve, and it yields a high T\textsubscript{m} of 88.0 ± 1.0 °C, which is close to that for the E2-WT control (Figure 4.7B). The small difference between the T\textsubscript{m} of E2-WT and that of E2-ΔN is surprising because it demonstrates that although a relatively large (44 amino acid) domain at the subunit-subunit interface has been removed, the thermostability of the assembled particle is still high at typical physiological pH. However, this N-terminal region is clearly involved in modulating assembly, as the E2-WT scaffold (which contains this domain) does not show pH-dependent structure.

Since the TEM micrographs and CD spectra demonstrate that the E2-ΔN protein at pH 5.0 is unfolded and aggregated (Figures 4.5D & 4.6B), we expected that no apparent cooperative unfolding would be detected in its thermal profile. Our data does, in fact, support this prediction. At acidic pH, the profile lacks clearly folded and unfolded regions or a corresponding transition inflection between these 2-folding states to yield a midpoint transition temperature. This further supports an unfolded state at pH 5.0.

**Disassembly at acidic pH is irreversible.** Our data shows the aggregates at pH 5 to be unfolded protein rather than protein crystals or agglomerates of assembled nanoscaffolds. Since the unfolding of large protein complexes is usually irreversible (Backmann et al. 1998; Blaauwen and Driessen 1998), and E2-WT also remains aggregated after heating to high temperatures (Dalmau et al. 2008), we predicted that dissociated complexes at pH 5.0 could not be reassembled simply by buffer-exchange back to pH 7.4. Indeed, when E2-ΔN was dialyzed to pH 5.0 and re-dialyzed to pH 7.4, we obtained a final average hydrodynamic diameter in the size range of aggregates at pH
5.0 (2341 ± 575 nm, with high polydispersity). Transmission electron micrographs also corroborate this data, showing that only disassembled aggregates remain even after return of the protein to pH 7.4 (Figure 4.8). This supports our expectations of irreversibility in pH-triggered disassembly.

![Figure 4.8](image)

**Figure 4.8** Transmission electron micrographs of E2-ΔN dialyzed from pH 5.0 back to pH 7.4 and stained with 2% uranyl acetate. This image shows unfolded protein aggregates similar to those observed at pH 5.0, and confirm the irreversibility of the disassembly process. Scale bar is 50 nm.

**Intermediate-pH Results Reveal that Dissociation of Nanoparticles Is Likely Due to Electrostatic Repulsion Generated by Protonation of Histidines.** As previously described, the design of the pH-responsive scaffold (E2-ΔN) was conceived to tune self-assembly by protonating the two sets of histidines (His218 and His222) located at the subunit-subunit interface (Figure 4.2B). Through crystalloographic structure, the distance between these histidines side chains is estimated to be less than 0.6 nm. Since the Debye length at our ionic strength is also near 0.5-0.6 nm between pH 5.0 to 7.4 (Israelachvili 1992), we hypothesized that the repulsive interactions arising from the protonation of the imidazole groups on these neighboring histidines could be within the
electrostatic screening distance and provide repulsive interactions to disassemble the scaffold.

To test this model, we examined the unfolding of E2-ΔN as a function of pH. Since disassembly of the scaffold causes the loss of the characteristic α-helical minimum at 208 nm on the CD spectra (Figure 4.6B), we monitored ellipticity at this wavelength over the pH range of 5.0 to 8. Figure 4.9, together with TEM images in Figure 4.6, show that the nanoparticles are fully folded and intact at pH 7.4 and above. At pH 6.5 and below, the protein is disassembled and TEM images reveal unfolded protein aggregates (Figures 4.5D and 4.5F).

![Figure 4.9 pH-unfolding profile for the E2-ΔN mutant. Majority of pH-related unfolding and dissociation occur near pH 7.0.](image)

The major transitional loss of protein structure and assembly primarily occurs near pH 7.0. In fact, TEM micrographs demonstrate an intermediate level of dissociation. While aggregates at this pH primarily contain unfolded protein agglomerates, we also observe occasional regions of indistinct, partially-dissociated
protein assemblies which are neither fully intact nor fully denatured (Figure 4.5G). This pH is near the expected pKa of histidine’s imidazole group and supports our model of histidine-triggered dissociation. As pH is decreased, a greater fraction of protonated imidazole side chains would result in greater repulsive forces. Assuming a typical pKa of 6.5 for histidine (Stryer 1988), our first-order estimation gives 24% protonation at pH 7.0. This suggests that complete protonation is not necessary to create repulsive interactions which are considerable enough to destabilize the protein complex.
4.5 Conclusions

In summary, our results demonstrate that we have engineered a non-native, pH-dependent molecular switch into a caged protein assembly. While the wild-type scaffold of the E2 subunit of pyruvate dehydrogenase exhibits remarkable stability at both pH 5.0 and 7.4, our redesigned pH-responsive protein complex is stable at a typical physiological pH of 7.4 but irreversibly dissociates at pH 5.0. The design of our scaffold incorporates deletion of the N-terminal domain (which appears to be important for modulating assembly) and the introduction of repulsive electrostatic forces between adjacent histidines at a subunit-subunit interface. As the pH nears imidazole’s pKa, the disassembly of the 60-subunit complex is triggered. Presumably the pH profile could be more finely tuned by adjusting parameters such as the length of the N-terminal deletion or the number of interacting electrostatic interactions at the intersubunit interface.

Using the N-terminus to modulate assembly is a paradigm that many viruses appear to employ. As our investigation with a nonviral system suggests, this domain could be a broader structural and functional motif for caged protein particles. This suggests that our general design strategy could potentially be applicable to regulate assembly in other caged protein systems, which have been widely utilized in nanotechnology, such as cowpea chlorotic mottle virus (Douglas and Young 1998; Slocik et al. 2005), brome mosaic virus (Huang et al. 2007), or heat-shock protein (Flenniken et al. 2006; Klem et al. 2005). These protein assemblies, by examination of the crystallographic structures (Kim et al. 1998; Lucas et al. 2002; Speir et al. 1995), appear to also exhibit similar N-terminal domains that reside at the subunit-subunit interface.
CHAPTER 5

STUDY OF A pH-TRIGGERED DISASSEMBLY MECHANISM ON A CAGED PROTEIN COMPLEX

This chapter has been slightly modified and published as: Mercè Dalmau, Sierin Lim, Szu-Wen Wang; pH-Triggered Disassembly in a Caged Protein Complex. Biomacromolecules, 2009, 10: 3199-3206.
5.1 Abstract

Self-assembling protein cage structures have many potential applications in nanotechnology, one of which is therapeutic delivery. For intracellular targeting, pH-controlled disassembly of virus-like particles and release of their molecular cargo is particularly strategic. We investigated the potential of using histidines for introducing pH-dependent disassembly in the E2 subunit of pyruvate dehydrogenase. Two subunit interfaces likely to disrupt stability, an intra-trimer interface (the N-terminus) and an inter-trimer interface (methionine-425), were redesigned. Our results show that changing the identity of the putative anchor site 425 to histidine does not decrease stability. In contrast, engineering non-native pH-dependent behavior and modulating the transition pH at which disassembly occurs can be accomplished by mutagenesis of the N-terminus and by ionic strength changes. This observed pH-triggered disassembly is due to electrostatic repulsions generated by histidine protonation. These results suggest that altering the degree of electrostatic repulsion at subunit interfaces could be a generally-applicable strategy for designing pH-triggered assembly in protein macromolecular structures.


5.2 Introduction

Caged proteins are precisely assembled biomacromolecular complexes that occur in living organisms. Their architectures consist of multiple protein subunits which interact to form spherical cages with icosahedral symmetry. These virus-like particle systems have been utilized as a new class of biomimetic templates in nanotechnology, with their potential applications including catalysis (Varpness et al. 2005), medical imaging (Liepold et al. 2007; Uchida et al. 2009), nanoparticle synthesis (Kramer et al. 2004), and drug delivery vehicles (Flenniken et al. 2006). In these investigations, novel functionality is introduced to the complexes by the modulation of the protein cages at the exterior and interior surfaces.

We have recently reported design of a non-viral protein cage, the E2 subunit of the pyruvate dehydrogenase multienzyme complex from *Bacillus stearothermophilus*. The pyruvate dehydrogenase complex consists of three different enzymes: pyruvate decarboxylase (E1), dihydrolipoamide acetyltransferase (E2), and dihydrolipoamide dehydrogenase (E3) (Lessard et al. 1998; Perham 1991). The hollow, caged component of the complex, E2, is a dodecahedron that consists of 60 identical subunits. It forms a spherical shell with an outer diameter of 24 nm, 12 openings of 5 nm at the fivefold axes, and an inner core diameter of 12 nm (Izard et al. 1999; Milne et al. 2002). Our previous work has generated an E2 scaffold that consists of only the core, caged structure. We demonstrated that this scaffold can tolerate multiple mutations in the interior surface of the inner cavity while remaining unusually stable up to approximately 85°C, and this cavity can be engineered to accommodate drug-like molecules (Dalmau et al. 2008).
Perham and coworkers have shown that the outer surface of the E2 scaffold can be functionalized with antigenic peptides for vaccine development (Domingo et al. 2003; Domingo et al. 2001). We have also demonstrated that by altering the subunit-subunit interface, this structure can serve as the template upon which a pH-triggered molecular switch can be designed (Dalmau et al. 2009). This pH-dependent behavior yields stability at pH 7.4 but dissociation at pH 5.0, which is a specific profile not yet observed in any natural or engineered caged protein complex.

In several viral complexes, the assembly behavior and dynamics can be altered by changing solvent conditions or modifying protein domains. For example, the pore-gating mechanism of the cowpea chlorotic mottle virus (CCMV) can be controlled by pH, and polymorphism of this complex is dependent on ionic strength, pH, and the presence of divalent metal ions (Bancroft et al. 1967; Johnson and Speir 1997). In numerous viruses, including CCMV (Tang et al. 2006), brome mosaic virus (Larson et al. 2005), and tomato bushy stunt virus (Hsu et al. 2006), the N-terminus of the viral capsid is involved in assembly mechanisms, and its removal yields capsids of smaller sizes. Although our E2 scaffold is not a virus, it exhibits behavior similar to such viral particles (Dalmau et al. 2008) and represents the first example of a caged protein system that can be engineered for correct assembly at physiological pH but disassembly at pH 5.0.

This pH-assembly profile is relevant for applications in intracellular therapeutic delivery. For such delivery systems, drug carriers should remain intact until they have reached the site of therapy. One strategy for drug release and particle disassembly has been to utilize the increase in acidity as these macromolecules are taken up by the endocytotic pathway (Christie and Grainger 2003). Prior investigations have introduced
imidazole groups into polymeric carriers to impart pH-dependent assembly, as their protonation near pH 6-6.5 introduces repulsive electrostatic interactions capable of dissociating the complex (Park et al. 2006; Putnam et al. 2001). Since imidazole is the side chain of histidine (His), we propose that introduction of histidines at targeted locations within subunit-subunit interfaces could potentially yield the same disassembly effect in multimeric protein complexes. The role of histidine in the pH-dependent inhibition of natural multimeric protein receptors has been speculated, and these interactions also appear to affect the electrostatic repulsion (and consequently stabilization) at dimeric interfaces (Huggins and Grant 2005). We show in our work that this general approach of using histidine to modulate stability and assembly can also be used as a protein engineering strategy to control assembly of large protein complexes.

Others have speculated that redesigning subunit-subunit interactions in a caged protein complex can be a potential strategy for controlling its architecture and assembly (Uchida et al. 2007). In this study, we examine two possible subunit-subunit interfaces of the E2 protein complex structure for introducing pH-triggered disassembly: (1) the intratrimer interface and (2) the intertrimer interface. Our previous work (Chapter 4) had demonstrated that, using the first strategy, pH-modulated disassembly could be engineered by removing the N-terminal arm (yielding the E2-$\Delta N$ mutant) (Dalmau et al. 2009). This N-terminus region appears to be structurally analogous to similar motifs that control assembly in viral particles (Calhoun et al. 2007; Hsu et al. 2006; Tang et al. 2006). In contrast to the wild-type protein scaffold (E2-WT), which is correctly-assembled with high thermostability at both pH 7.4 and pH 5.0, the pH-responsive E2-$\Delta N$ is stable at pH 7.4, but dissociates at pH 5.0 (Dalmau et al. 2008; Dalmau et al.
2009). The degree of unfolding was evaluated over the pH range from 5.0 to 7.4, and we found that the transition for unfolding occurs near histidine’s pKa value. This unfolding and dissociation profile had suggested that the electrostatic repulsions between adjacent histidines are critical for disassembly. Both histidine-218 (His218) and histidine-222 (His222) appeared to be likely candidates for inducing disassembly at low pH, as they each converge to form a trimeric cluster within 0.6 nm of each other around the three-fold axis of symmetry (at the intratrimer interface).

In this investigation, we probe the potential of modulating these putative electrostatic interactions using both ionic strength and mutagenesis to obtain finer control over the pH at which disassembly occurs. The effect of increasing the electrostatic screening on the interactions by changing salt concentration is investigated. We also remove the two sets of histidines at the intratrimeric interface (His218 and His222) to study the effect of decreasing the number of potential histidine interactions on multimeric protein dissociation.

This study also probes the feasibility of introducing pH-dependent assembly by altering the intertrimeric interface. Izard et al. had found that the E2 structures with icosahedral symmetry are governed by the principles of quasi-equivalence. When comparing residues and architectures of E2 from different species, they found one residue in each which was central to almost all the quasi-equivalent interactions and was conserved in its identity and hydrophobic nature. In E2 from Bacillus stearothermophilus, this putative “anchor” residue is methionine-425 (Met425) and is located in a hydrophobic pocket of an adjacent trimer (Izard et al. 1999). Based on these
prior observations, we also explore the potential of using site 425 as a location for introducing a pH-dependent molecular switch.
5.3 Materials and Methods

5.3.1 Materials

We used the host *E. coli* strains DH5α (Zymo Research) and BL21(DE3) (Stratagene). Our oligonucleotides were synthesized by Integrated DNA Technologies. Restriction enzymes (*NdeI*, and *BamHI*), T4 DNA ligase, DNase, RNase, and calf intestinal alkaline phosphatase were purchased from New England Biolabs, and *PfuUltra* High-Fidelity DNA polymerase was from Stratagene. All buffer reagents were supplied by EMD Chemicals, except for isopropyl β-D-thiogalactopyranoside (IPTG) and potassium phosphate (Fisher Scientific), phenylmethylsulfonyl fluoride (PMSF, Pierce), and sodium azide (Merck KGaA).

5.3.2 Design of pH-Triggered Disassembly

To introduce pH-triggered disassembly, we investigated the use of histidine as a potential switch for inducing disassembly at acidic conditions. The pKa of histidine has typically been reported to be between pH 6-7 (Creighton 1993; Edgcomb and Murphy 2002; Stryer 1988). Therefore, the protonation of the imidazole side chain in histidine could be a mechanism for decreasing protein stability. We focused on subunit-subunit interactions to destabilize the complex, and investigated two different interfaces on the E2 complex: the intratrimer and the intertrimer interfaces.
Design and mutagenesis of N-terminal mutants at intra-trimer interface. To determine possible locations within the intratrimer interface of the E2 scaffold, we noted that one common feature existed among several icosahedral virus capsids: an N-terminal arm which embraces an adjacent subunit and modulates protein assembly (Calhoun et al. 2007; Hsu et al. 2006; Tang et al. 2006). We found that this structural motif is also present in our E2 scaffold (Figure 5.1A).

Figure 5.1 Molecular structure highlighting relevant protein domains and mutation sites. (A) The altered sites at the intratrimer interface includes deletion of an N-terminal domain (dark blue) and mutagenesis of interacting histidines (His218, yellow; His222, magenta) of one monomer subunit (green). (B) Close-up of His218 and His222 shows the interaction of each set of trimeric histidines 90° to the 3-fold axis of symmetry. (C) Residue 425 (red) resides at the intertrimer interface. Regions colored in orange and light blue are trimers comprising three E2 subunits. The structures are displayed using PyMOL (DeLano, 2002).

In our recent work, we reported the design of non-native, pH-dependent behavior in our caged protein that is based on the modification of this interface. We proposed that the deletion of the first 50 amino acids of the N-terminal arm and the presence of adjacent histidines at the subunit-subunit interface would destabilize the scaffold at pH 5.0 (Dalmau et al. 2009). Dissociation of the protein nanoparticles at low pH was indeed observed and was likely due to the electrostatic repulsion arising from the protonation of
the neighboring imidazole groups at His218 and His222. The protein structure shows that these two sets of histidines form trimeric clusters that converge around the 3-fold axis, and their side chains are within a proximity of 0.6 nm (Figure 5.1B) (Dalmau et al. 2009).

We desired to investigate the role of these individual histidines (His218 and His222) in modulating disassembly. To replace these histidines, we chose alanine, a nonpolar residue with a small methyl side chain which is unlikely to cause steric hindrance. Using cassette mutagenesis, we constructed four N-terminal mutants for expression in *E. coli*: E2-ΔN, E2-ΔN/H218A, E2-ΔN/H222A and E2-ΔN/H218A/H222A. All these variants have the first 50 amino acids deleted from the wild-type (E2-WT) scaffold, and they have His218 and His222 replaced by alanine as indicated by the name.

To construct these mutants, the gene encoding E2-WT was PCR-amplified using PfuUltra High-Fidelity DNA polymerase with the following oligonucleotides: 5′–gggaattccatATGCACACCCGCGCCACACGTTACC-3′ (forward) and 5′-cgcggatccTTAAGCTTCCATCAGCAGCAG–3′ (reverse) for E2-ΔN; 5′–gggaattccatATGCGGACCGCGCCACACGTTACC-3′ (forward) and 5′–cgcggatccTTAAGCTTCCATCAGCAGCAG–3′ (reverse) for E2-ΔN/H218A; 5′–gggaattccatATGCAACCGCGCCACACGTTACC-3′ (forward) and 5′–cgcggatccTTAAGCTTCCATCAGCAGCAG–3′ (reverse) for E2-ΔN/H222A; and 5′–gggaattccatATGCGGACCGCGCCACACGTTACC-3′ (forward) and 5′–cgcggatccTTAAGCTTCCATCAGCAGCAG–3′ (reverse) for E2-ΔN/H218A/H222A. The histidine-to-alanine mutations are highlighted in bold.
To construct the expression vectors pE2ΔN50, pE2ΔN50/H218A, and pE2ΔN50/H218A pE2ΔN50/H222A, we digested the vector pET-11a and the mutated genes with NdeI and BamHI restriction enzymes (underlined above). Both vector and synthetic genes were purified with agarose gel electrophoresis and ligated with T4 DNA ligase. The sequences of the mutated genes were confirmed by DNA sequencing. Details of standard molecular cloning techniques follow those described by Sambrook & Russell (Sambrook and Russell 2001).

**Design and mutagenesis of methionine-425 mutant at inter-trimer interface**

*site.* Another interaction that we investigated is within the intertrimer interface. Izard et al. has reported one residue around which all the quasi-equivalent interactions of all known E2 species are centered, and this residue is conserved (Izard et al. 1999). In the E2 from *Bacillus stearothermophilus*, this residue is Met425, and it is located in a hydrophobic pocket of an adjacent trimer (Figure 5.1C).

Since we desired to modulate assembly with pH, we mutated Met425 to a histidine, an amino acid which is never observed at any corresponding “anchor” sites in known E2 complexes. Since histidine has a pKa around 6.5, we hypothesized that the protonation of His425 would destabilize the assembly by disrupting the hydrophobic pocket.

To perform site-directed mutagenesis at Met425, we followed a modified version of the Stratagene QuickChange protocol using E2-WT cloned into pGEM-3Z and an oligonucleotide with the sequence 5’– TCC GAC CCG GAA CTG CTG CTG CAT GAA GCT TAA gga tcc tct a ga gtc-3’. The methionine-to-histidine mutation at site 425 is highlighted in bold. The resulting clones were screened and the mutated E2 genes
were cloned into the expression vector pET-11a at the *Nde*I and *Bam*HI sites as described above for the N-terminal mutant genes. The sequences of the mutants were confirmed by DNA sequencing.

### 5.3.3 Protein Expression and Purification

Our proteins containing the mutations are summarized in Table 5.1 and were expressed and purified using a previously-reported protocol for E2 production (Allen and Perham 1997; Dalmau et al. 2008). Briefly, *E. coli* strain BL21(DE3) was transformed with the mutant genes, cultured in Luria-Bertani medium supplemented with 100 µg/mL ampicillin, and induced by the addition of 1 mM IPTG at an optical density (OD$_{600}$) of 0.7 – 0.9. Cells were harvested three hours after induction by centrifugation at 6500 g for 15 min and stored at -20°C. Thawed cells were resuspended in lysis buffer (20 mM Tris at pH 8.7, 0.02% sodium azide, 2 U/mL DNase, 2 U/mL RNase, 1 mM MgCl$_2$ and 1 mM PMSF), disrupted in a French pressure cell at 60 MPa (Thermo Scientific), and ultracentrifuged at 185,000 g for 1 hour to remove the insoluble fraction.

<table>
<thead>
<tr>
<th>E2 variant</th>
<th>Mutated interface</th>
<th>Expressed in <em>E. coli</em></th>
<th>Soluble after purification</th>
<th>pH-dependent disassembly</th>
</tr>
</thead>
<tbody>
<tr>
<td>E2-WT</td>
<td>None, scaffold with wild-type sequence</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>E2-NΔ50</td>
<td>Intra-trimer</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>E2-NΔ50/H218A</td>
<td>Intra-trimer</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>E2-NΔ50/H222A</td>
<td>Intra-trimer</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>E2-NΔ50/H218A/H222A</td>
<td>Intra-trimer</td>
<td>Yes</td>
<td>No</td>
<td>n/a</td>
</tr>
<tr>
<td>E2-M425H</td>
<td>Inter-trimer</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

**Table 5.1** Summary of E2 variants created to investigate pH-dependent assembly.
The resulting supernatant was loaded onto a HiPrep Q-sepharose anion exchange column (GE Healthcare) that had been equilibrated with Tris buffer (20 mM Tris at pH 8.7, 0.02% sodium azide, and 5 mM EDTA). Fractions containing the E2 protein were concentrated and loaded onto a Superose 6 Prep Grade XK 16/70 size exclusion column (GE Healthcare) equilibrated with 50 mM potassium phosphate (pH 7.4), 100 mM NaCl, 0.02% sodium azide, and 5 mM EDTA. Fractions containing E2 were pooled and concentrated to a final protein concentration of 0.5-1 mg/mL, and the purified proteins were stored at -80 °C.

Protein concentrations were determined using a Micro BCA Protein Assay Kit (Pierce). SDS-PAGE analyses were run on 12% Tris-HCl gels. The molecular mass of purified proteins was determined using MALDI-TOF mass spectrometry (Applied Biosystems Voyager-DE STR).

5.3.4 Assembly Confirmation

The diameters of the E2 assemblies were determined by dynamic light scattering (DLS) (Zetasizer Nano ZS, Malvern) with 1 mg/mL protein in phosphate buffer (50 mM potassium phosphate, 500 mM NaCl, pH 7.4). We confirmed structure of the protein complex with transmission electron microscopy (TEM). Protein samples were stained for 2 minutes with 2% uranyl acetate on carbon-coated electron microscopy grids (Ted Pella, Inc.) and images were obtained with a Philips CM-20 transmission electron microscope.
5.3.5 Secondary Structure Characterization

We used far-UV circular dichroism (CD) to evaluate and quantify protein folding and thermal stability. CD scans were performed on a Jasco 810 spectropolarimeter equipped with a Jasco peltier temperature controller. Protein samples (0.06 mg/mL) in 50 mM potassium phosphate and 500 mM NaCl (pH 7.4) were scanned at 25 °C at 10 nm/min in 0.1 cm pathlength quartz cells. For thermal unfolding, we monitored the ellipticity at 222 nm from 40 to 95 °C at a heating rate of 1 °C/min. Results are an average of at least three scans.

5.3.6 Characterization of pH-Dependent Structural Stability

Two methods were used for evaluating physical pH-stability of the protein scaffolds: determination of the pH-unfolding profile using circular dichroism (CD) and evaluation of pH-dependent solubility by dynamic light scattering (DLS). In our recent work, we showed that the unfolding of the E2 scaffold and its dissociation and aggregation are closely coupled (Dalmau et al. 2008; Dalmau et al. 2009). Both E2-WT and mutant E2-ΔN at the normal physiological pH of 7.4 gave the characteristic CD profile of an alpha-helical rich protein with minima at 208 and 222 nm, indicating correct folding and assembly. The pH-dependent disassembly profile can then be evaluated by measuring the loss of alpha-helical content at 208 nm (Dalmau et al. 2009), and was determined for both E2-WT and mutants. For E2-WT, protein at 0.06 mg/mL was scanned on the spectropolarimeter at 25°C over pH 5.0 – 8.0 (50 mM potassium...
phosphate in 100 or 500 mM NaCl) at a scanning speed of 10 nm/min in 0.1 cm quartz cells. The ellipticity at 208 nm was plotted with pH. The data was fit to a sigmoidal function using a non-linear least squares routine in MATLAB (see Appendix A.11). We define the “pH transition of unfolding” as the pH of the inflection point between the folded and unfolded states. For pH-unfolding profiles of E2 mutants, scans were performed over various pH values. Data points are an average of three samples.

Since the differential scanning calorimetry and CD profiles of E2-WT showed that the complex’s unfolding, dissociation, and aggregation are tightly coupled, (Dalmau et al. 2008) our second method of evaluating physical stability of the protein was to measure the fraction of 25-nm particles remaining in solution. As pH was changed, E2 nanoparticles remained either in solution (at 25 nm) or became aggregated (>1 micron). This bimodal distribution of particle sizes can be observed by DLS (Figure 5.6). We determined the percent volume of 25-nm particles remaining in solution as pH was varied, and we fit the data to a standard sigmoidal curve using a non-linear least squares fitting routine. The “pH transition of solubility” corresponds to the inflection point of this curve, which describes the pH between the fully-soluble and fully-aggregated states. Data points are an average of readings from three protein samples.
**5.4 Results and Discussion**

**Protein Expression and Purification.** A summary of the E2 variants is presented in Table 5.1. All proteins were expressed in the soluble fraction, with the intratrimer (N-terminal) mutants having a slightly lower expression relative to E2-WT and the intertrimer variant (E2-M425H) yielding a similar expression level as E2-WT. Although E2-$\Delta$N/H218A/H222A was equately expressed in the soluble fraction of *E. coli*, this mutant aggregated to particle sizes greater than 2 µm during the purification process, and conditions could not be found to keep the protein soluble. Subsequent investigations, therefore, could not be performed with this variant.

The remaining mutants were purified and their molecular weights were confirmed by mass spectrometry. E2-$\Delta$N, E2-$\Delta$N/H222A, and E2-M425H all yielded molecular weights within 0.1% of the calculated theoretical value, confirming the expected subunit sizes and sequences. However, a 0.6% decrease in molecular weight relative to the calculated expected value was observed for E2-$\Delta$N/H218A. This difference in molecular weight corresponds to a single methionine and is consistent with protein which has had its translational start-methionine cleaved. The removal of the N-terminal methionine by *E. coli* methionine aminopeptidase occurs approximately 95% when the penultimate residue is glycine or alanine (Hirel et al. 1989; Miller et al. 1987). In E2-$\Delta$N/H218A, this second amino acid is the alanine that was originally a histidine. Therefore, the protein sequence for E2-$\Delta$N/H218A is as-coded by its DNA sequence, except for the truncation of the N-terminal methionine.
Transition pH of Unfolding in pH-Triggered Disassembly can be Modulated by Ionic Strength. The mechanism of pH-triggered disassembly in E2-ΔN was proposed to be due to electrostatic repulsions between adjacent histidines at the intratrimer interface (Figure 5.1 A,B) (Dalmau et al. 2008). To test the electrostatic nature of this disassembly, we altered the ionic strength of the buffer, which can screen electrostatic interactions within a protein. The calculated Debye length was near 0.6 nm in 100 mM NaCl and approximately 0.4 nm in 500 mM NaCl (both in 50 mM potassium phosphate, at pH 5.0 and 7.4) (Israelachvili 1992). In comparison, the distance between the adjacent histidines in the two sets of trimeric side chains (His218 and His222) was estimated to be near 0.6 nm. We expected that higher salt concentrations will be more effective at screening adjacent histidine charges, therefore stabilizing the protein complex and yielding a lower unfolding transition pH.

Figure 5.2 shows the pH-unfolding profile of E2-ΔN at 100 and 500 mM NaCl. We observe that there is a significant shift in the transition pH of unfolding when salt concentration is increased, with the transition at pH 7.2 and 6.2 for 100 and 500 mM NaCl, respectively. Both these pH values are within proximity of the typical pKa for histidine (Creighton 1993; Stryer 1988). Increasing ionic strength, therefore, does indeed stabilize the pH-dissociating protein complex, as we expect for a disassembly mechanism that is electrostatic in nature. These results confirm the ability to modulate protein assembly behavior using ionic strength and further support the electrostatic mechanism of pH-induced disassembly in the E2-ΔN scaffold.
**Figure 5.2** pH-unfolding profile for E2-ΔN at two different ionic strengths. Using CD measurements to evaluate α-helical unfolding, the pH values at which 50% unfolding occurs are 7.2 and 6.2 in 100 and 500 NaCl (with 50 mM potassium phosphate buffer), respectively.

**ΔN-Based Mutants All Demonstrate pH-Triggered Disassembly.** To determine the relative contributions of the specific histidines (His218 and His222) at the intratrimeric interface in pH-triggered disassembly, and to investigate the feasibility of modulating the transition pH for disassembly by manipulating these histidines, we replaced these residues by sitedirected mutagenesis. We obtained variants with two sets of histidines (E2-ΔN) and one set of histidines (E2-ΔN/H218A and E2-ΔN/H222A) at the intratrimeric interface. As previously mentioned, mutant E2-ΔN/H218A/H222A (zero histidines at the interface) did not remain soluble after purification and therefore was not further used in this study.
Table 5.2 Hydrodynamic diameters (in nm) for E2-WT and N-terminal mutants (E2-ΔN, E2-ΔN/H218A, and E2-ΔN/H222A) at pH 7.4 and 5.0 (50 mM potassium phosphate, 500 mM NaCl). The large particle sizes for the set of ΔN mutants at pH 5.0 are inexact due to high polydispersity index values, but qualitatively show aggregation. In contrast, the E2-WT control remains soluble and assembled.

<table>
<thead>
<tr>
<th></th>
<th>E2-WT</th>
<th>E2-ΔN50</th>
<th>E2-ΔN/H218A</th>
<th>E2-ΔN/H222A</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH 7.4</td>
<td>26.6 ± 0.4</td>
<td>25.0 ± 0.9</td>
<td>33.3 ± 2.2</td>
<td>24.0 ± 0.6</td>
</tr>
<tr>
<td>pH 5.0</td>
<td>26.5 ± 0.7</td>
<td>2428 ± 577</td>
<td>2677 ± 563</td>
<td>2887 ± 1169</td>
</tr>
</tbody>
</table>

Correct structure of these mutants was first evaluated by particle size using DLS. A summary of hydrodynamic diameters for N-terminal mutants at pH 7.4 and 5.0 is presented in Table 5.2. E2-ΔN, E2-ΔN/H218A, and E2-ΔN/H222A exhibit diameters close to the value of the wild-type scaffold at pH 7.4 and support correctly assembled particles. Although the average value of E2-ΔN/H218A is slightly larger than the others, we believe this is due to the purity of the sample rather than a size difference in the mutant, as the maximum peaks in the size histograms primarily range from 24-28 nm. At pH 5.0, all ΔN variants aggregate to particles >2 µm, while E2-WT remains intact and soluble. These observations are supported by TEM micrographs, which show intact particles with icosahedral symmetry at pH 7.4, but protein aggregates at pH 5.0 for the ΔN-based mutants (Figure 5.3). In contrast, TEM micrographs of E2-WT confirm correct structure at both pH 7.4 and 5.0 (Dalmau et al. 2009). We also note that the large aggregates of the E2-ΔN variants at acidic pH show no distinct identifiable particles, and prior protein folding analyses of E2-ΔN yielding similar TEM images showed that these particles are indeed partially unfolded (Dalmau et al. 2009).
Figure 5.3 All N-terminal mutants yield correctly-assembled E2 particles at pH 7.4 and aggregates without intact particles at pH 5.0, as shown in these transmission electron micrographs. (A) E2-ΔN at pH 7.4, (B) E2-ΔN/H218A at pH 7.4, (C) E2-ΔN/H222A at pH 7.4, (D) E2-ΔN at pH 5.0, (E) E2-ΔN/H218A at pH 5.0, and (F) E2-ΔN/H222A at pH 5.0. Single nanoparticles are highlighted in red. Samples are stained with 2% uranyl acetate and scale bars are 50 nm.

**pH-Triggered Disassembly is Affected by the Number of Adjacent His Interactions.** To identify the pH range over which disassembly occurs, we evaluated two characteristics between pH 5.0 and 7.4: (1) the loss of the minimum ellipticity peak at 208 nm and (2) the percentage volume of nonaggregated, intact particles remaining in solution. Although both strategies measure structural changes, the first approach requires unfolding of the R-helices in the protein, while the second strategy is potentially more sensitive to subtle structural variations. Prior results had shown that our particle size measurements with DLS detected aggregation readily and that solubility of the E2-WT
protein corresponded with correct dodecahedral structure (Dalmau et al. 2008). Furthermore, data from differential scanning calorimetry revealed unfolding of the E2 scaffold to be tightly coupled with dissociation and aggregation (Dalmau et al. 2008), indicating that the dual methods measuring unfolding and solubility can determine critical transition pH values.

Ellipticity Versus pH. A correctly folded and assembled E2-WT complex is relatively rich in \( \alpha \)-helices (Izard et al. 1999), and we have shown that this protein yields CD spectra with the characteristic 208 and 222 minima for \( \alpha \)-helical folding (Dalmau et al. 2008). We can therefore measure ellipticity at 208 nm as pH is varied to obtain a pH-folding profile (Dalmau et al. 2009). For all our pH-responsive (N-terminal) mutants at pH 7.4, we observe the characteristic profile of proteins rich in \( \alpha \)-helices, further confirming that these proteins are correctly folded at normal physiological conditions (Figure 5.4).

![Figure 5.4](image)

**Figure 5.4** Far-UV circular dichroism spectra for N-terminal E2 mutants at pH 7.4 and 25°C. Molar ellipticity vs. wavelength for mutants E2-\( \Delta N \), E2-\( \Delta N/H218A \), and E2-\( \Delta N/H222A \) show the two characteristic minima at 208 nm and 222 nm for proteins high in \( \alpha \)-helical content and demonstrate correctly-folded nanoparticles at pH 7.4. All spectra are an average of data from three samples.
The loss of the minimum at 208 nm is measured over the pH range between 5.0 and 8.0 to determine the pH transition point. Figure 5.5 shows that these mutants undergo a transition from a folded state (at pH 7.0 and above) to an unfolded state in the typical pKa range for histidine. By fitting this data to a sigmooidal curve, we determined the transition pH of unfolding to be 6.18, 6.16, and 6.17 for E2-ΔN, E2-ΔN/H218A, and E2-ΔN/H222A, respectively. Based on CD data, the three mutants appear to unfold at a common transition pH between 6.1 and 6.2. Our pH-unfolding profiles suggest that either no difference in the transition pH exists or that any differences between transition pH values (based on folding) are small and ellipticity is not sensitive enough to detect subtle conformational changes. To distinguish between these two possibilities, we therefore used an alternative method (solubility based on DLS) to probe structural differences.

**Figure 5.5** pH-unfolding profiles for N-terminal mutants based on molar ellipticity at 208 nm (50 mM potassium phosphate, 500 mM NaCl). Using CD measurements to evaluate α-helical unfolding, the pH values at which 50% unfolding occurs are 6.18, 6.16, and 6.17 for E2-ΔN, E2-ΔN/H218A, E2-ΔN/H222A, respectively. Individual data points are an average of three samples.
We note that the mutant E2-ΔN/H222A shows a higher amount of R-helical content at high pH values (Figures 5.4 and 5.5) relative to both E2-WT and E2-ΔN/H218A. Increased helicity in an alanine mutant is consistent with prior investigations that found alanine to have the highest helix propensity relative to all other amino acids (Pace and Scholtz 1998). The observation that E2-ΔN/H218A, which also has an alanine substituted for histidine, does not display greater helicity can be explained by its location; amino acids very close to the N- and C-termini are typically more dynamic in nature, and this flexibility may alter their ability to participate in local folding (Derewenda 2004).

**Volume of Intact, Soluble Particles Versus pH.** To determine the amount of soluble and correctly assembled protein, we measured the percentage volume of the 25 nm peak between pH 5.0 and 8.0 for the N-terminal mutants using DLS. As pH is decreased and aggregation occurs, there is a loss of the 25 nm peak (for correctly assembled structure) and a corresponding growth of large (>1 μm) peaks, indicating aggregated protein. (Representative DLS data is shown in Figure 5.6). The pH-solubility profile was determined and the transition pH of solubility, based on intact, nonaggregated particles, was 6.4, 5.9, and 5.7 for mutants E2-ΔN, E2-ΔN/H218A, and E2-ΔN/H222A, respectively (Figure 5.7).
Figure 5.6 Representative hydrodynamic diameter measurements using dynamic light scattering for E2-ΔN at pH 6.0. This distribution shows a soluble, 25-nm particle population and another population of protein aggregates > 1 μm, and yields a percent volume of 57.5% and 42.5%, respectively. A bimodal profile is also typical of those observed for the other N-terminal mutants at intermediate pH values between pH 5.5 and 6.5, and the average percent volume of the soluble fractions (25-nm peak) are the values plotted in Figure 5.7.

Figure 5.7 pH-solubility profiles for N-terminal mutants based on the volume percentage of the DLS particle size peak for intact, soluble particles at 25 nm. When these size measurements are used, the pH values at which 50% of the proteins are soluble are 6.4, 5.9, and 5.7 for E2-ΔN, E2-ΔN/H218A, E2-ΔN/H222A, respectively. Individual data points are an average of three samples.

The difference in these transition pH values obtained by DLS is distinguishable between the variants with two sets of histidines compared to one set of histidines. E2-ΔN,
which contains histidines at both sites 218 and 222, starts to aggregate at a higher transition pH (pH 6.4) than either E2-ΔN/H218A or E2-ΔN/H222A, indicating that it is more sensitive to protonation than the single-histidine mutants. Because an E2 complex is comprised of 60 subunits, the presence of each histidine interaction is amplified. Our DLS data, therefore, supports the premise that the pH-dependent behavior is related to the strength of electrostatic repulsions between subunit-subunit interfaces and that His218 and His222 are involved in these interactions.

Furthermore, the transition pH for E2-ΔN/H218A is slightly higher (pH 5.9) than E2-ΔN/H222A (pH 5.7), despite the same number of histidines at the intratrimeric interface for these mutants. This small difference might be explained by the removal of the start-methionine in E2-ΔN/H218A. Although the noncharged methyl side chain of alanine replaces the ionizable imidazole of histidine at location 218, after truncation of the start-methionine, this alanine becomes the first amino acid in this mutant and therefore has a positively charged amine in its backbone. In addition to the potential electrostatic interactions between adjacent His222, then, these protonated amines in E2-ΔN/H218A could also introduce a small amount of additional repulsion between subunits (though to a weaker degree due to the increased distance).

**pH-Triggered Dissociation.** By the dual strategies of ionic strength changes and interface modifications via mutagenesis, our investigation elucidates the role of repulsive histidine interactions at acidic pH. Screening these interactions by increasing salt concentrations can stabilize the complex and significantly decrease the pH at which disassembly occurs, demonstrating the importance of electrostatics in the disassembly mechanism. These electrostatic interactions likely involve histidines, as histidine is the
only amino acid with a pKa within the pH transition range of our set of E2-ΔN experiments. This approach does not, however, identify the specific interactions that are involved in disassembly.

To probe whether our hypothesized histidine interactions (His218 and His222) are involved, we used protein engineering. Our data demonstrates that the mutant with two sets of histidines in the N-terminal interface is less stable than the ones with one histidine, with disassembly occurring at a pH value closer to neutral physiological conditions when more histidine interactions are present at that interface. However, this transition-pH difference between the mutants appears to be small relative to the differences when ionic strength was modulated and, in fact, was not large enough to be measured by circular dichroism. Therefore, although the strength of the repulsive histidine interactions at sites 218 and 222 is an important parameter for modulating assembly, additional interactions, either between subunits or within subunits, may also be important. We predict that these interactions will still involve histidine because disassembly occurs near its typical pKa, and will be within the Debye length. Although we did not identify additional native His-His interactions that fit this distance criteria, several potential histidine-lysine interactions could be present within the complex. Since the amine side chain of lysine (Lys) remains protonated between pH 5 and 8, an adjacent histidine could also introduce repulsion at acidic pH. His308-Lys312 and Lys211-His215 are potential interactions within the same subunit, while Lys187-His362 and Lys217-His222 are interactions between different subunits.

**Non-Native Residue at “Anchor” Site Does Not Disrupt Dodecahedral Structure at pH 5.0 and 7.4.** The amino acid Met425 had been previously shown to be
an “anchor” residue around which the quasi-equivalent interactions are centered (Figure 5.1C) (Izard et al. 1999). Through structure-based sequence alignment, this anchor residue is conserved over all known E2 species and is located at a trimer-trimer interface. It has been found to be exclusively hydrophobic and aliphatic, with the identity of methionine or leucine conserved over almost all known species. Furthermore, this amino acid resides in a hydrophobic pocket of an adjacent trimer (Izard et al. 1999).

We mutated the methionine at site 425 to a histidine (mutant E2-M425H), reasoning that if this site is indeed a conserved anchor point with requisite aliphatic, hydrophobic character, then the stability of the protein is likely to be affected upon changing the physicochemical character at this location. This anchor site is never observed to be histidine (or any ionizable residue) in any known species (Izard et al. 1999), suggesting that the introduction of a charge at this location could alter the assembly of E2. Furthermore, the ringed nature of histidine contrasts the structure of Met or Leu, also potentially creating a steric disruption to histidine’s fit within the conserved hydrophobic pocket. Steric or electrostatic disruption of the hydrophobic interaction in this conserved region could potentially alter nanoparticle assembly and stability and serve as a molecular switch for pH-triggered assembly.
Table 5.3 Summary of particle sizes and midpoint unfolding temperatures for E2-WT and E2-M425H at pH 7.4 and 5.0 (50 mM potassium phosphate, 100 mM NaCl). *Data from E2-WT at pH 7.4 is taken from Dalmau et al (Dalmau et al. 2008).

<table>
<thead>
<tr>
<th></th>
<th>E2-WT</th>
<th>E2-M425H</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH 7.4</td>
<td>Diameter (nm)</td>
<td>26.6 ± 0.6</td>
</tr>
<tr>
<td></td>
<td>T_m (°C)</td>
<td>91.4 ± 1.4</td>
</tr>
<tr>
<td>pH 5.0</td>
<td>Diameter (nm)</td>
<td>29.3 ± 0.4</td>
</tr>
<tr>
<td></td>
<td>T_m (°C)</td>
<td>83.0 ± 2.3</td>
</tr>
</tbody>
</table>

We recombinantly produced the E2 scaffold containing the His mutation at position 425. Its structure, stability, and assembly at both pH 7.4 and 5.0 were characterized with DLS, CD, and TEM, and these results are summarized in Table 5.3 and Figures 5.8 and 5.9. E2-M425H yields a hydrodynamic diameter of 28.4 ± 1.4 and 28.3 ± 1.2 nm at pH 7.4 and 5.0, respectively, demonstrating that no pH-dependent disassembly is observed for this mutant. These properties are identical to those of the wild-type E2 scaffold, which also shows no change in particle size between pH 7.4 and 5.0 (26.6 ± 0.6 and 29.3 ± 0.4 nm, respectively) (Dalmau et al. 2008).
Figure 5.8 Far-UV circular dichroism spectra for E2-WT and E2-M425H at pH 7.4 and 5.0. (A) Molar ellipticity vs. wavelength for E2-M425H mutant. (B) Molar ellipticity vs. wavelength for E2-WT. Correctly-assembled E2-M425H and the E2-WT control show the two characteristic minima at 208 and 222 nm for proteins high in \( \alpha \)-helical content. All spectra are an average of data from three samples.

Figure 5.9 Transmission electron micrographs of E2-WT and E2-M425H. Intact capsules are observed for (A) E2-WT at pH 7.4, (B) E2-WT at pH 5.0, (C) E2-M425H at pH 7.4, and (D) E2-M425H at pH 5.0, indicating that E2-M425H does not demonstrate pH-dependent assembly. Samples are stained with 2% uranyl acetate and scale bars are 50 nm.
This invariance in particle size after the pH-shift suggests the dodecahedral structure of the assembly remains intact at pH 5.0, and this is further supported by secondary structure data. Circular dichroism was used to measure the ellipticity as a function of wavelength at both pH 7.4 and 5.0 (Figure 5.8). As expected for correctly folded E2 scaffolds, the E2-M425H spectra at both pH conditions are nearly identical to each other and show the characteristic profile of $\alpha$-helical structure, with minima at 208 and 222 nm (Figure 5.8A). Furthermore, the spectra are also similar to those of E2-WT (Figure 5.8B) (Dalmau et al. 2009). Our data demonstrate that no secondary structural changes in either E2-M425H or E2-WT are detected at the acidic conditions, which are expected for histidine protonation. Intact protein structure is confirmed by the TEM micrographs (Figure 5.9). This entire set of structural data demonstrates that both E2-WT and E2-M425H are stable at both pH conditions and do not display pH-responsive disassembly and structure. Although the identity of the site-425 “anchor” residue is conserved in nature and likely involves hydrophobic interactions between subunits, we show that it can be modified without disrupting the overall structure of the protein assembly.

Non-Native Residue at “Anchor” Site Does Not Disrupt Thermostability at pH 5.0 or 7.4. Although the structure of E2-M425H is not affected by pH changes, it is possible that stability is decreased by the mutation at the putative anchor site. To evaluate the effect of changing this site on the protein’s thermostability, we monitored the decrease in ellipticity at 222 nm as the protein is heated at both pH 7.4 and 5.0 (Figure 5.10). The loss of $\alpha$-helical structure then enables the determination of an apparent
midpoint unfolding temperature, $T_m$ (Dalmau et al. 2008; Greenfield 2006). Both temperature profiles for E2-M425H at pH 7.4 and 5.0 yielded sigmoidal curves, with calculated $T_m$ values of $91.3 \pm 0.8$ and $88.3 \pm 1.2 \, ^\circ C$ for pH 7.4 and 5.0, respectively (Table 5.3). These values are comparable to the $T_m$ for the wild-type complex at the same solution conditions, which are $91.4 \pm 1.4 \, ^\circ C$ at pH 7.4 and $83.0 \pm 2.3 \, ^\circ C$ at pH 5.0 (Dalmau et al. 2008). Therefore, the histidine mutation at site 425 does not appear to disrupt the native high thermostability.

![Temperature stability profiles of E2-M425H protein assemblies at pH 7.4 and 5.0. Molar ellipticity at 222 nm is measured to evaluate the level of unfolding as temperature is increased. E2-M425H shows an apparent unfolding transition temperature of $91.3 \pm 0.8$ and $88.3 \pm 1.1 \, ^\circ C$ at pH 7.4 and 5.0, respectively. These values are similar to E2-WT and demonstrate the assemblies’ high thermostability. Scans are an average of data from three samples.](image)

**Figure 5.10** Temperature stability profiles of E2-M425H protein assemblies at pH 7.4 and 5.0. Molar ellipticity at 222 nm is measured to evaluate the level of unfolding as temperature is increased. E2-M425H shows an apparent unfolding transition temperature of $91.3 \pm 0.8$ and $88.3 \pm 1.1 \, ^\circ C$ at pH 7.4 and 5.0, respectively. These values are similar to E2-WT and demonstrate the assemblies’ high thermostability. Scans are an average of data from three samples.

Since the scaffold is based on a protein from a thermophilic organism (*Bacillus stearothermophilus*), it is not surprising that the wild-type E2 scaffold at neutral physiological pH exhibits high thermostability. It is, however, unexpected that E2-M425H shows the same level of structural integrity and high thermostability, given the
reported conserved nature of the amino acids (methionine and leucine) that naturally exist at this location (Izard et al. 1999). Because histidine is never observed at this “anchor” site for any known E2 forms, one might expect that ideal hydrophobic packing at this location is optimal with methionine or leucine. Our data at pH 7.4, however, shows that there is more flexibility at this site than the conserved sequence indicates.

Furthermore, it is surprising that E2-M425H remains nearly as stable at pH 5.0 as at pH 7.4 (and, in fact, is even slightly more thermostable than E2-WT at pH 5.0). This implies that the introduction of a positive charge at site 425 does not disrupt this hydrophobic intertrimer interaction or the overall complex assembly and that this “anchor” site is not as critical for assembly as suggested by structural and sequence analyses (Izard et al. 1999). One possibility for the unexpected high stability of E2-M425His that histidine-425 has a very acidic local pKₐ or is not titratable, because site 425 is buried in a hydrophobic environment. In fact, E2-M425H remains stable down to at least pH 3.0, where the protein remains soluble and the hydrodynamic diameter has been measured to be 29.3 ± 2.7 nm. If the imidazole at site 425 is indeed not titratable at pH 3.0 and above, the hydrophobic nature of the site would be retained and remain uncharged. Studies have reported that histidines in an apolar, hydrophobic environment have a widely varying range of pKₐ values (Edgcomb and Murphy 2002) and can remain unprotonated even at pH 2.3 (Geierstanger et al. 1998).
5.5 Conclusions

In this study, we targeted two different subunit-subunit interfaces to test the feasibility of designing a pH-triggered molecular switch for disassembly of a caged protein complex. In one strategy, we disrupted the intratrimer interface and specific electrostatic interactions within this interface by changing ionic strength and by mutagenesis of the protein’s N-terminus. This set of mutants did indeed demonstrate pH-dependent behavior, and we investigated the degree to which this pH-triggered assembly is tunable. Our results reveal that the engineered pH-dependent molecular switch in the E2-ΔN mutant set does indeed involve electrostatic interactions. The transition point of pH-dependent assembly was strongly modulated by ionic strength and, to a lesser degree, by the presence of repulsive histidine interactions at sites 218 and 222. These combined results suggest that pH-induced repulsive interactions between residues in other locations of the protein, in addition to these two histidine trimer sets, are likely involved in the disassembly mechanism.

We also mutated a putative anchor site, Met425, to a nonnative amino acid that had not been observed at this anchor location. Since we were aiming for assembly at pH 7.4, but disassembly at acidic pH values, we introduced a histidine at this location, potentially to disrupt the existing hydrophobic pocket. We found that changing this residue from methionine to histidine did not affect overall structure or the high thermostability, even at low pH values expected to introduce charge and disruption of the pocket. Therefore, although Met425 was previously identified as a structural anchor, our work suggests that its identity does not appear to be critical in assembly.
Our results also demonstrate that modulating the degree of electrostatic repulsion through changing the number of histidine interactions at subunit interfaces could be a generally applicable strategy for designing pH-triggered assembly in protein macromolecular structures. To accomplish this, the identification of key interactions through mutagenesis is critical, as not all putative interaction sites resulted in the same degree of pH-dependent behavior.
CHAPTER 6

MOLECULAR ENCAPSULATION WITHIN AN ENGINEERED PROTEIN CAGE
6.1 Abstract

Self-assembling protein cages have been proposed as general drug carriers since they form highly symmetric nanoscale architectures that offer the potential to be tailored according to the desired application. We have genetically modified the E2 subunit of pyruvate dehydrogenase to yield a highly thermostable dodecahedral protein complex. This scaffold has a hollow interior cavity which can be modified, both chemically and genetically. To show the applicability of our scaffold, we conjugated two different fluorescent dye molecules to an internal-cysteine variant. Moreover, we designed multiple variants that offer accessible hydrophobic regions in the interior hollow cavity and developed a general strategy intended to nucleate hydrophobic small drug-like agents inside our protein complex. We demonstrated that these mutations can introduce non-native functionality in the E2 scaffold, while retaining its self-assembling ability and stability. The molecular-loading ability of these internal hydrophobic variants is comparable to the wild-type control.
6.2 Introduction

Nature provides many nanometer scale protein cages, including viral capsids (Douglas et al. 1998; Douglas et al. 2006; Hsu et al. 2006; Larson et al. 2005; Manchester et al. 2006; Wang et al. 2002), heat shock proteins (Hsp) (Flenniken et al. 2003, Flenniken et al. 2006), and ferritins (Klem et al. 2005; Uchida et al. 2009). The interior cavities of these self-assembled nanocaged structures are hollow and thus are potentially suitable to serve as platforms for encapsulation and release of guest molecules. Current approaches include chemical and genetic manipulation. One of the most used methods is the covalent attachment to site-specifically engineered residues on the interior surface. Thus, a cysteine residue genetically introduced at the inner surface presents a reactive thiol group in the assembled protein cage. Flenniken et al. covalently modified Hsp cage interior and selectively attached fluorescein molecules to the reactive cysteines with a pH-sensitive linker (Flenniken et al. 2003; Flenniken et al. 2006). The utility of this approach has been demonstrated with the attachment and release of the anticancer drug doxorubicin using a virus-like cage (Flenniken et al. 2005) and imaging agents (Lewis et al. 2006). Schmidt et al. inserted the WW domain (a small and compact globular structure that is characterized by two highly conserved tryptophan residues and interacts with proline-rich ligands) of mouse formin binding protein 11 (FBP 11) to the N-terminus of the outer shell of the murine polyomavirus VP1 (which faces the interior of the capsid), and attached proline-rich ligands to the inner surface of the virus-like particles (Schmidt et al. 2001). These virus-like encapsulated scaffolds were able to penetrate mammalian cells and deliver peptides and proteins in vitro.
Another approach for molecular entrapment using virus-like scaffolds is a pH-dependent gating mechanism used by Douglas et al. (Douglas and Young 1998). Many viruses can experience reversible structural changes that open or close the pores on the scaffold. Under swollen conditions, free molecular exchange between the virus cavity and the medium allows the entrapment of materials. They showed the ability to couple this swelling mechanism of the cage with a pH-dependent inorganic polymerization for the entrapment and crystallization of two mineral particles. Also, they demonstrated the electrostatic interactions between the scaffold and the guest molecules by a pH-dependent encapsulation of an anionic organic polymer. Thus, this approach was shown to be versatile for guest molecules encapsulation.

Other protein-based systems used in drug encapsulation include human serum albumin (HSA) and protein-based polymers. Microcapsules of HSA can be produced by the stabilization of an emulsified solution of albumin in the presence of drug, by heating or addition of a cross-linking agent (Eatock et al. 1999; Kreuter et al. 2007; Weber et al. 2000). Thus, drug solubility and stability can increase and alter pharmacokinetic profiles (Eatock et al. 1999; Stehle G et al. 1999). However, the use of HSA as a pharmaceutical excipient is still very empirical. One of the most exploited protein-based polymers are the elastin-like polymers (ELP) (Urry 1988). Nanoparticles of ELP can be loaded with significant amounts of a model drug dexamethasone phosphate (DMP) when self-assembly is carried out in an aqueous co-solution of polymer-DMP by raising the temperature from 4°C to 37°C. This temperature change induces the polymer to fold into a non-random structure called a β-spiral. In this folded state the polymer chain loses all the water structures from hydrophobic hydration and this water is mostly placed inside
the hydrophilic channel within the β-spiral, where the DMP is entrapped (Herrero-Vanrell et al. 2005; Urry 1988).

On the other hand, the development of high-throughput screening technology in drug discovery allows the massive emergence of new drug candidates. Nevertheless, it has been reported that poor aqueous solubility is the predominant problem associated with high-throughput tested drug candidates (Lipinski 2002). Therefore, there is a need to develop a generalized encapsulation assay for these new emerging drug agents.

Thus, our objective was to design a generalized protein scaffold to encapsulate and solubilize a broad range of drug molecules while decreasing toxicity of side-effects, maintaining drug stability, and retaining therapeutic activity. Therefore, we first had to analyze generalized drug-protein interactions. These types of interactions have been studied in multi-drug efflux transporters. Efflux pumps are known to be present in most living cells. In human cells, pumps such as P-glycoprotein (P-gp) were found to protect cells from toxic molecules (by ejecting foreign molecules outside the cell) and to decrease intracellular accumulation of drugs. Since P-gp has broad substrate specificity, with over 100 different known substrates (Seelig 1998), the key point was to elucidate the nature of this substrate recognition. Further studies reported that the nature of this chemical specificity is mostly hydrophobic with some electrostatic interactions (aromatic stacking and Van der Waals forces) (Gottesman 2002; Yu et al. 2003; Zgurskaya and Nikaido 2000). Therefore we hypothesize that the protein nanocapsule design for hydrophobic encapsulation must incorporate a significant amount of accessible hydrophobic residues in the internal cavity wall.
In this investigation, we altered the surface of the interior cavity of a virus-like particle both chemically and genetically, and we demonstrate that we can attach drug-like molecules through chemical coupling. Moreover, we designed multiple variants that incorporate accessible hydrophobic residues at the hollow cavity, and we developed a general assay intended to nucleate a broad range of hydrophobic drug-like compounds inside these protein cages. To date, most of the procedures employed to position the drug inside the cavity of protein scaffolds have been chemical-reaction-based. Therefore, the physicochemical properties and therapeutic activity of drugs can be easily affected. Hence, our strategy consists in nucleating drug particles utilizing hydrophobic differences between the inner cavity of the protein assembly and the medium outside.
6.3 Materials and Methods

6.3.1 Design and Construction of Internal Mutants for Chemical Conjugation

To determine locations within the hollow internal cavity for mutagenesis, the E2 core domain complex was visualized in Protein Explorer (Martz 2002) and PyMOL (DeLano 2002). Crystallographic data of the E2 subunit was obtained from the Protein Data Bank (PDB code 1b5s). For preserving the assembly of the complex, we looked for residues within the surface of the internal cavity which did not appear to interact with adjacent subunits and could support the covalent attachment of guest molecules. Both the native interior and exterior surfaces of the nanocapsule are abundant in charged, hydrophilic residues, and we selected lysine 239 (K239) and aspartate 381 (D381) (Figure 6.1). For mutagenesis of these residues, we chose amino acids that would introduce alternative physicochemical properties into the cavity - a small, slightly hydrophobic side chain (alanine, Ala, A), a large hydrophobic side chain (phenylalanine, Phe, F), and a side chain which introduces a thiol to enable chemical modification (cysteine, Cys, C). It is important to note that although single mutations on the E2 subunit are made, the protein complex results from an assembly of 60 subunits; therefore, a single mutation results in 60 modifications in the interior of capsule and two mutations result in 120 modifications (Figure 6.1B).
To perform site-directed mutagenesis, we cloned the E2-WT gene into pGEM-3Z (Promega) and followed a modified version of the Stratagene QuickChange protocol (described in the Appendix A.2). We used oligonucleotides with the sequence 5’–ACCAAACTGGTTGCGCACCCTNNNAATTCAAGGCATTCGTCGGGCG-3’ and 5’-GCCGAAAAGCCGATCGTTCGTTGGT3’ for mutants at positions 239 and 381, respectively. NNN indicates the 239 or 381 mutation site and was replaced with GCG, TTT, or TGC for Ala, Phe, or Cys, respectively. For constructing the double-Phe mutant, site-directed mutagenesis was performed consecutively. The resulting clones were screened and the mutated E2 genes were cloned into the expression vector pET-11a at the NdeI and BamHI sites as described above for the wild-type gene. The sequences of the mutant and wild-type genes were confirmed by
DNA sequencing. The resulting mutants were K239A, K239F, and K239C (for position 239), D381A, D381F, and D381C (for position 381), and the double-Phe mutant (K239F/D381F).

6.3.2 Design and Construction of Internal Hydrophobic Mutants

The design of internal hydrophobic mutants was carried out in collaboration with Prof. Baldi and Dr. Randall. To identify optimal locations at the interior cavity in order to increase the accessible hydrophobic surface area, we ranked all residues taking into account the solvent accessible surface area that they offer, the charge of the residue to mutate and the distance of the amino acid to the center of the nanocapsule in order to maximize the increase in hydrophobicity (data in Appendix A.10). Altogether we identified nine potential sites: aspartate 230 (D230), alanine 236 (A236), lysine 239 (K239), glutamate 375 (E375), arginine 380 (R380), aspartate 381 (D381), glycine 382 (G382), glutamate 383 (E383), alanine 386 (A386) (Figure 6.2). In order to increase the hydrophobicity of the interior cavity, we chose to introduce a large hydrophobic residue (phenylalanine, Phe, F) since in nature P-gp uses Phe residues for binding diverse hydrophobic drug molecules (Loo and Clarke 1993; Loo et al. 2006).
Single mutants were constructed by site-directed mutagenesis as described in Section 6.3.1. The primers used for the mutagenesis are listed in Table 6.1. The resulting mutants were D230F, A236F, K239F, E375F, R380F, D381F, G382F, E383F, and A386F. After mutagenesis, these mutants were tested for expression at 30 and 37 °C and the new variants were screened by a quick thermostability test described in Section 6.3.5. The stability results together with the steric hindrance (both within the monomer...
and oligomer) were included in the design and construction of double mutants. This procedure was repeated until reaching a total of four mutations per subunit. That is a total of 240 modifications per scaffold. A list of all constructs with their corresponding primers can be found in Tables 6.2, 6.3, and 6.4. Based on expression and thermostability results together with solvent accessibility, distance to center of the particle and charge of the mutated site (described in section 6.4), we chose eight mutants (four triple and four quadruple) to perform hydrophobic drug loading experiments.
<table>
<thead>
<tr>
<th>Mutant</th>
<th>Primers sequence (from 5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D230F</td>
<td>Fwd GTT ACC CTG ATG GAT GAA GCA <strong>TTT</strong> GTT ACC AAA CTG GTT GCG CAC</td>
</tr>
<tr>
<td></td>
<td>Rev GTG CGC AAC CAG <strong>TTT</strong> GGT AAC GAA TGC TTC ATC CAT CAG GTC TGC</td>
</tr>
<tr>
<td>A236F</td>
<td>Fwd GCA GAC GTT ACC AAA CTG <strong>TTT</strong> CAC CGT AAA AAA TTC AAG GCG</td>
</tr>
<tr>
<td></td>
<td>Rev CGC CTT GAA <strong>TTT</strong> TTT ACG GTG <strong>GAA</strong> AAC CAG TTT GGT AAC GTC TGC</td>
</tr>
<tr>
<td>K239F</td>
<td>Fwd ACC AAA CTG GTT GCG CAC CGT <strong>TTT</strong> AAA TTC AAG GCG ATT GCG GCG</td>
</tr>
<tr>
<td></td>
<td>Rev CGC CGC AAT CGC CTT GAA <strong>TTT</strong> AAA ACG GTG CGC AAC CAG TTT GGT</td>
</tr>
<tr>
<td>E375F</td>
<td>Fwd GCG ATC CGT GGT ATT GGT CGT <strong>TTT</strong> AAG CCG GCT GCT CGT CGT</td>
</tr>
<tr>
<td></td>
<td>Rev ACG AAC GAT CGG CTT <strong>AAA</strong> GGC TAT ACG ACC AAT ACC CAG GAT CGC</td>
</tr>
<tr>
<td>R380F</td>
<td>Fwd GCC GAA AAG CCG ATC GTT <strong>TTT</strong> GAC GTT GAA ATC GTT GCT GCT CGG</td>
</tr>
<tr>
<td></td>
<td>Rev CGG AGC AGC AAC GAT TTC ACC GTC <strong>AAA</strong> AAC GAT CGG CTT TTT GGC</td>
</tr>
<tr>
<td>D381F</td>
<td>Fwd GCC GAA AAG CCG ATC GTT <strong>TTT</strong> GGT GAA ATC GTT GCT GCT CGG</td>
</tr>
<tr>
<td></td>
<td>Rev CGG AGC AGC AAC GAT TTC ACC <strong>AAA</strong> ACG AAC GAT CGG CTT TTT GGC</td>
</tr>
<tr>
<td>G382F</td>
<td>Fwd GCC GAA AAG CCG ATC GTT GAC <strong>TTT</strong> GAA ATC GTT GCT GCT CGG</td>
</tr>
<tr>
<td></td>
<td>Rev CGG AGC AGC AAC GAT TTC <strong>GAA</strong> GTC ACG AAC GAT CGG CTT TTT GGC</td>
</tr>
<tr>
<td>E383F</td>
<td>Fwd GAA AAG CCG ATC GTT <strong>TTT</strong> CGT GAC GGT GGT GCT GCT CGT ATG</td>
</tr>
<tr>
<td></td>
<td>Rev CAT CGG AGC AGC AAC GAT <strong>AAA</strong> ACC GTC ACG AAC GAT CGG CTT TTT</td>
</tr>
<tr>
<td>A386F</td>
<td>Fwd GTT CGT GAC <strong>TTT</strong> GGT GAA ATC GTT GCT GCT CGG ATG CTG GCC CGT TC</td>
</tr>
<tr>
<td></td>
<td>Rev GAC AGG GCC AGC ATC GGA GCA <strong>AAA</strong> ACG ATT TCA CGG TCA CGA AC</td>
</tr>
</tbody>
</table>

Table 6.1 Oligomers used to construct hydrophobic single mutants.
<table>
<thead>
<tr>
<th>Mutant</th>
<th>Primers sequence (from 5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D230F-R380F</td>
<td>Used primers for single mutant R380F and used pE2-D230F as DNA template</td>
</tr>
<tr>
<td>A236F-K239F</td>
<td>Fwd: GTT ACC AAA CTG GTT <strong>TTT</strong> CAC CGT <strong>TTT</strong> AAA TTC AAG GCG ATT GCG</td>
</tr>
<tr>
<td></td>
<td>Rev: CGC AAT CGC CTT GAA <strong>TTT</strong> <strong>AAA</strong> ACG GTG <strong>AAA</strong> AAC CAG TTT GGT AAC</td>
</tr>
<tr>
<td>E375F-D381F</td>
<td>Used primers for single mutant E375F and used pE2-D381F as DNA template</td>
</tr>
<tr>
<td>E375F-A386F</td>
<td>Used primers for single mutant E375F and used pE2-A386F as DNA template</td>
</tr>
<tr>
<td>R380F-D381F</td>
<td>Fwd: GCC GAA AAG CCG ATC GTT <strong>TTT</strong> <strong>TTT</strong> GGT GAA ATC GTT GCT GCT CCG</td>
</tr>
<tr>
<td></td>
<td>Rev: CGG AGC AGC AAC GAT TTC ACC <strong>AAA</strong> <strong>GAA</strong> AAC GAT CCG CTT TTC GGC</td>
</tr>
<tr>
<td>D381F-G382F</td>
<td>Fwd: GCC GAA AAG CCG ATC GTT CGT <strong>TTT</strong> <strong>TTT</strong> GAA ATC GTT GCT GCT CCG</td>
</tr>
<tr>
<td></td>
<td>Rev: CGG AGC AGC AAC GAT TTC <strong>GAA</strong> <strong>AAA</strong> ACG AAC GAT CCG CTT TTC GGC</td>
</tr>
<tr>
<td>D381F-E383F</td>
<td>Fwd: GAA AAG CCG ATC GTT CGT <strong>TTT</strong> <strong>GTT</strong> <strong>TTT</strong> ATC GTT GCT GCT CCG ATG</td>
</tr>
<tr>
<td></td>
<td>Rev: CAT CGG AGC AGC AAC GAT <strong>AAA</strong> <strong>ACC</strong> <strong>AAA</strong> ACG AAC GAT CCG CTT TTC</td>
</tr>
<tr>
<td>G382F-A386F</td>
<td>Fwd: CCG ATC GTT CGT GAC <strong>TTT</strong> GAA ATC GTT <strong>TTT</strong> GCT CCG ATG CTG GCC</td>
</tr>
<tr>
<td></td>
<td>Rev: GGC CAG CAT CGG AGC <strong>AAA</strong> <strong>AAC</strong> GAT TTC <strong>AAA</strong> GTC ACG AAC GAT CCG</td>
</tr>
</tbody>
</table>

**Table 6.2** Oligomers used to construct hydrophobic double mutants.
<table>
<thead>
<tr>
<th>Mutant</th>
<th>Primers sequence (from 5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D230F-R380F-D381F</td>
<td>Fwd GAA AAG CCG ATG CTT TTT TTC GGT GAA ATC GTT TTC CAC</td>
</tr>
<tr>
<td></td>
<td>Rev GTG GAA AAC GAT TTC ACC GAA AAA AAC GAT CGG CTT TTC CAC</td>
</tr>
<tr>
<td>D230F-A236F-K239F</td>
<td>Fwd ACC CTG ATG GAT GAA GCA TTC GTT ACC AAA CTG GTT TTC CAC</td>
</tr>
<tr>
<td></td>
<td>Rev GTG GAA AAC CAG TTT GGT AAC GAA TGC TTC ATC CAT CAG GGT</td>
</tr>
<tr>
<td>E375F-R380F-D381F</td>
<td>Fwd GGT ATT GGT CGT ATA GCC TTT AAG CCG ATC GTT TTC TTC</td>
</tr>
<tr>
<td></td>
<td>Rev AAA GAA AAC GAT CGG CTT AAA GGC TAT ACG ACC AAT ACC</td>
</tr>
<tr>
<td>E375F-D381F-A386F</td>
<td>Fwd TTT AAG CCG ATC GTT CTT GGT GAA ATC GTT TTC CAC</td>
</tr>
<tr>
<td></td>
<td>Rev GTG GAA AAC GAT TTC ACC AAA AC G AAC GAT CGG CT TAAA</td>
</tr>
<tr>
<td>R380F-D381F-G382</td>
<td>Fwd AAG CCG ATC GTT TTC TTC TTC GAA ATC GTT GCT GCT CCG</td>
</tr>
<tr>
<td></td>
<td>Rev CGG AGC AGC AAC GAT TTC GAA AAA GAA AAC GAT CGG</td>
</tr>
<tr>
<td>D381F-G382F-E383F</td>
<td>Fwd CCG ATC GTT CGT TTC TTC TTC ATC GTT GCT GCT CCG ATG</td>
</tr>
<tr>
<td></td>
<td>Rev CAT CGG AGC AGC AAC GAT GAA GAA AAA AAC GAT CGG</td>
</tr>
<tr>
<td>R380F-D381F-E383F</td>
<td>Fwd GCC GAA AAG CCG ATC GTT TTC TTC GGT TTC ATC GTT GCT</td>
</tr>
<tr>
<td></td>
<td>Rev AGC AAC GAT AAA ACC AAA GAA AAC GAT CGG CTT TTC GGC</td>
</tr>
<tr>
<td>G382F-E383F-A386F</td>
<td>Fwd CCG ATC GTT CGT GAC TTT TTC ATC GTT TTC GCT CCG ATG</td>
</tr>
<tr>
<td></td>
<td>Rev CAT CGG AGC AAA AAC GAT GAA AAA GTC ACG AAC GAT CGG</td>
</tr>
<tr>
<td>A236F-K239F-E383F</td>
<td>Fwd CCG ATC GTT CGT GAC GGT TTC ATC GTT GCT GCT CCG ATG</td>
</tr>
<tr>
<td></td>
<td>Rev CAT CGG AGC AGC AAC GAT AAA ACC GTC ACG AAC GAT CGG</td>
</tr>
<tr>
<td>E375F-E383F-A386F</td>
<td>Fwd CCG ATC GTT CGT GAC GGT TTC ATC GTT TTC GCT CCG ATG</td>
</tr>
<tr>
<td></td>
<td>Rev CAT CGG AGC AAA AAC GAT AAA ACC GTC ACG AAC GAT CGG</td>
</tr>
</tbody>
</table>

Table 6.3 Oligomers used to construct hydrophobic triple mutants.
<table>
<thead>
<tr>
<th>Mutant</th>
<th>Primers sequence (from 5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D230F-A236F-</td>
<td>Fwd: GAA AAG CCG ATC GTT CGT <strong>TTT</strong> GGT GAA ATC GTT GCT</td>
</tr>
<tr>
<td>K239F-D381F</td>
<td>Rev: AGC AAC GAT TTC ACC <strong>AAA</strong> ACG AAC GAT CGG CTT TTC</td>
</tr>
<tr>
<td>A236F-K239F-</td>
<td>Fwd: CCG ATC GTT CGT <strong>GAC</strong> <strong>TTT</strong> <strong>TTT</strong> ATC GTT GCT GCT CCG ATG</td>
</tr>
<tr>
<td>G382F-E383F</td>
<td>Rev: CAT CGG AGC AGC AAC GAT <strong>AAA</strong> GAA GTC ACG AAC GAT CCG</td>
</tr>
<tr>
<td>K239F-E375F-</td>
<td>Fwd: <strong>TTT</strong> <strong>TTT</strong> GGA ATC GTT <strong>TTT</strong> GCT CCG</td>
</tr>
<tr>
<td>R380F-D381F</td>
<td>Rev: <strong>AAA</strong> <strong>AAA</strong> ACG AAC GAT CGG</td>
</tr>
<tr>
<td>Used primers for</td>
<td>triple mutant E375F-R380F-D381F and used pE2-K239F as DNA template</td>
</tr>
<tr>
<td>K239F-R380F-</td>
<td>Used primers for triple mutant R380F-D381F-G382F and used pE2-K239F as DNA template</td>
</tr>
<tr>
<td>D381F-E383F-A386F</td>
<td>Fwd: <strong>TTT</strong> <strong>TTT</strong> <strong>TTT</strong> ATC GTT GCT CCG</td>
</tr>
<tr>
<td>E375F-D381F-G382F-A386F</td>
<td>Rev: GGC CAG CAT CGG AGC <strong>AAA</strong> AAC GAT <strong>GAA</strong> <strong>GAA</strong> <strong>AAA</strong> ACG</td>
</tr>
<tr>
<td>Used primers for</td>
<td>triple mutant R380F-D381F-G382F and used pE2-K239F as DNA template</td>
</tr>
<tr>
<td>E375F-R380F-D381F-G382F</td>
<td>Fwd: GGT ATT GGT CGT <strong>ATA</strong> GCC <strong>TTT</strong> AAG CCG ATC GTT <strong>TTT</strong></td>
</tr>
<tr>
<td>Rev: <strong>GAA</strong> AAC GAT CGG CGT <strong>GAA</strong> GGC TAT ACG ACC AAT ACC</td>
<td></td>
</tr>
<tr>
<td>R380F-G382F-E383F-A386F</td>
<td>Fwd: GCC GAA AAG CCG ATC GTT <strong>TTT</strong> <strong>TTT</strong> ATC GTT</td>
</tr>
<tr>
<td>Rev: <strong>AAC</strong> GAT <strong>GAA</strong> <strong>AAA</strong> GTC <strong>GAA</strong> AAC GAT CGG CGT <strong>TTT</strong> <strong>TTT</strong> GGC</td>
<td></td>
</tr>
<tr>
<td>R380F-D381F-G382F-E383F</td>
<td>Fwd: GCC GAA AAG CCG ATC GTT <strong>TTT</strong> <strong>TTT</strong> ATC GTT</td>
</tr>
<tr>
<td>Rev: <strong>AAC</strong> GAT <strong>GAA</strong> <strong>GAA</strong> <strong>AAA</strong> GAA AAC GAT CGG CGT <strong>TTT</strong> <strong>TTT</strong> GGC</td>
<td></td>
</tr>
</tbody>
</table>

*Table 6.4* Oligomers used to construct hydrophobic quadruple mutants.
6.3.3 Protein Expression and Purification

Our proteins were produced using a previously-reported protocol for E2 purification as a departure point (Allen and Perham 1997, Dalmau et al. 2008). *E. coli* strain BL21(DE3), transformed with the wild-type or mutant gene, was cultured in Luria-Bertani medium supplemented with ampicillin (100 µg/mL) under constant shaking at 37 °C for the single mutants, or 30 °C for the triple and quadruple mutants (due to their lower expression level at 37 °C). Five-mL from an overnight culture (100 mL) of cells was used to inoculate 1 L of medium. Expression was induced by the addition of 1 mM IPTG when the culture reached an optical density of 0.7 – 0.9 at 600 nm. Cells were harvested three hours after induction by centrifugation at 6500 g for 15 min and stored frozen for the case of single mutants, and five hours, for the triple and quadruple variants. Thawed cells were resuspended in buffer (20 mM Tris pH 8.7, 0.02% sodium azide, 2 U/mL DNase, 2 U/mL RNase, 1 mM MgCl₂ and 1 mM PMSF) and lysed in a French pressure cell (Thermo Scientific) under a pressure of 60 MPa. The insoluble fraction was removed by ultracentrifugation at 185,000 g for 1 hour.

The resulting supernatant was loaded onto a HiPrep Q-sepharose anion exchange column (GE Healthcare) previously equilibrated with Tris buffer (20 mM Tris pH 8.7, 0.02% sodium azide, and 5 mM EDTA). Fractions containing the E2 protein were pooled, concentrated using ultrafiltration with 100 kDa molecular weight cut-off membranes (Biomax Ultrafiltration Discs, Millipore), and loaded onto a Superose 6 Prep Grade XK 16/70 size exclusion column (GE Healthcare) equilibrated with 50 mM potassium phosphate (pH 7.4), 100 mM NaCl, 0.02% sodium azide, and 5 mM EDTA.
Fractions containing E2 variants were pooled and concentrated to a final concentration within the range of 0.5-1 mg/mL. For long-term storage, proteins were kept at -80 °C.

Protein concentrations were determined using a Micro BCA Protein Assay Kit (Pierce; described in Appendix A.6) with bovine serum albumin as a standard. SDS-PAGE analyses were run on 12% Tris-HCl gels (detailed in Appendix A.5). Purified proteins (1 mg/mL) were dialyzed using drop dialysis (“V” Series Membrane, Millipore) for 20 min with deionized water prior to MALDI-TOF analysis (Applied Biosystems Voyager-DE STR) to confirm molecular mass. For the cysteine mutant, we first reacted the protein with 300 µM tris(2-carboxyethyl)phosphine (TCEP) for 1 hour at room temperature to reduce any disulfide linkages on the cysteine side chain prior to the dialysis step and MALDI-TOF analysis.

### 6.3.4 Evaluation of Correctly Self-assembled Nanoparticles

The hydrodynamic diameters of the resulting E2 variants were evaluated by dynamic light scattering (DLS; Zetasizer Nano ZS, Malvern) with 0.5 mg/mL protein in phosphate buffer (50 mM potassium phosphate, 100 mM NaCl, pH 7.4). We confirmed structure of the protein self-assemblies with transmission electron microscopy (TEM). Protein samples were stained for 2 min with 2% uranyl acetate on carbon-coated electron microscopy grids (TedPella, Inc.) and images were acquired with a Philips CM-20 transmission electron microscope.
6.3.5 Thermal Stability Characterization

Stability of the recombinant protein was initially screened by heating the soluble fraction of freshly lysed cells to 40, 50, 60, 75 and 95 ºC for 20 min. Denatured, aggregated protein was removed by centrifugation at 14000 g for 10 min, and the remaining supernatant was loaded onto a Tris-HCl SDS-PAGE gel.

We also used far-UV circular dichroism (CD) to evaluate and quantify protein folding and thermal stability. CD scans were performed on a Jasco 810 spectropolarimeter equipped with a Jasco peltier temperature controller (Jasco, Easton, MD). Protein samples at a concentration of 0.06 mg/mL in 50 mM potassium phosphate and 100 mM NaCl (pH 7.4) were scanned between 180-260 nm at 25ºC at a scanning speed of 10 nm/min in 0.1 cm pathlength quart cuvette. For thermal unfolding, we monitored the ellipticity at 222 nm from 40 to 95 ºC at a heating rate of 1 ºC/min. At least three scans of each variant were obtained and results were averaged.

The onset temperature of protein unfolding (T_o) and the apparent midpoint of the unfolding transition temperature (T_m) were determined by fitting the experimental thermal profile to standard two-state unfolding equations (Greenfield 2006). We used an in-house nonlinear least squares fitting routine programmed in MATLAB to determine the best-fit curve and to calculate the resulting first-derivative (see Appendix A.11). We defined T_m as the temperature at which 50% of the protein is unfolded (and where the first-derivative of the curve fit is a maximum) and T_o as the temperature at which 5% of the protein is unfolded. Although this thermodynamic analysis is strictly derived for two-
state, reversible folding processes, it is also used to determine relative stability
temperatures in irreversible processes (Greenfield 2006), and we apply it in this manner.

6.3.6 Conjugation of Model Drug Molecule to the Cys Mutant

To examine whether the thiols introduced in the D381C mutant are available for
conjugation of therapeutic molecules, we used two fluorescent molecules, fluorescein-5-
maleimide (F5M) and Alexa Fluor 532 C5-maleimide (AF532M), to represent typical
small-molecule model drugs. All reactions involving D381C and E2-WT (control) were
carried out in buffer containing 50 mM potassium phosphate and 100 mM NaCl at pH 7.4. To ensure that the cysteines were reduced and available for reaction, the samples
were mixed with 8.5 molar equivalents of TCEP at room temperature for 1 hour. Two
concentrations of each dye, 1 and 3 molar equivalents per protein subunit, were reacted
with 0.8 mg/mL E2-WT and D381C at room temperature for 2 hours, followed by
incubation at 4 °C for 20 hours. Unbound F5M and AF532M were separated from the
protein scaffold by dialyzing the mixture for two hours in buffer using drop dialysis (“V”
Series Membrane, Millipore) in an enclosed, hydrated dish. The fluorophore
concentration remaining in the protein solution and the ratio of dye to protein were
calculated by measuring absorbance at 494 and 526 nm for F5M and AF532M,
respectively, and readings were then calibrated to a standard curve under these buffer
conditions. The final protein concentration was rechecked using the previously described
BCA method, and the molecular weight of the resulting protein was determined by mass
spectrometry as described above.
6.3.7 Hydrophobic Drug Loading

We developed a general method for the encapsulation of hydrophobic small drug molecules by non-specific binding of drug molecules. Briefly, the loading strategy is based on hydrophobic differences between the inner cavity of the protein assembly and the medium outside. The encapsulation process utilized the evaporation of an organic solvent (ethanol) in which the drug molecules are initially dissolved, to supersaturate and nucleate drug particles in the interior site of the E2 nanocapsule. A hydrophobic fluorophore, 2’,7’-dichlorofluorescein (DCF), was used as a model drug to carry out these experiments. This model drug has a chemical structure similar to commercially available antitumor drugs, as well as its solubility is within the range of solubilities from small hydrophobic drugs, and it can be detected by means of fluorescence.

Since our main goal was to nucleate DCF inside the E2 protein cage, the experimental conditions for this phenomenon to happen were required. Previous studies in our lab had confirmed the intact structure of our scaffold in 8% (v/v) ethanol solution through transmission electron microscopy (TEM). The solubility limits of DCF at different concentrations of ethanol and at room temperature were obtained. Solid DCF was added in excess into each 4 mL solution of 0, 2, 4, 6 and 8 % ethanol in buffer (20 mM Tris/HCl pH 8.7) and stirred for 2 hours to reach saturation and followed by 2 hours of equilibration (without stirring). After equilibration, samples were syringe-filtered with a 0.2 µm porous size filter (Millipore) and appropriate dilutions with working buffer were done to take absorbance readings at the linear range. Repetitions of three were performed to obtain repeatability.
In order for the nucleation of DCF to happen, the solution needed to be at supersaturated concentration. Such condition was obtained by monitoring precipitation through turbidity measurements (absorbance at 600 nm) of three DCF solutions: supersaturated, saturated and below saturation concentration (was used as a negative control) in 8% (v/v) ethanol in buffer (20 mM Tris/HCl pH 8.7). Six replicates were prepared for each concentration and loaded into a 96-well plate which was stirred for 30 s. The plate was placed in a container saturated with water vapor to ensure just the evaporation of ethanol (not the water). Every 24 hours, the plate was assayed for precipitation using the turbidity assay mentioned above. Those samples with precipitate were spun down, and the soluble fraction was analyzed to confirm expected saturation concentration.

Supersaturated solutions of 3.3 and 3.5 mg/mL of 2’,7’-dichlorofluorescein (DCF) with 8% ethanol in buffer (20 mM Tris/HCl pH 8.7) were used to nucleate this model drug inside the protein scaffold. These solutions were mixed with protein solutions at a concentration of 0.4 mg/mL in a 96-well plate placed inside a container saturated with water vapor to avoid evaporation of the water phase and incubated at room temperature for 3 days.

### 6.3.8 Separation of Loaded Capsules by Sucrose Gradient

A discontinuous sucrose density gradient was prepared by layering successive decreasing sucrose density solutions upon one another; therefore the first solution applied is the 90% sucrose solution (see Table 6.5). A methodical and steady application of the
solutions yields the most reproducible gradient; therefore, the setup described below in Figure 6.3 was followed. First, a thick-wall polycarbonate ultracentrifuge tube (Beckman) was held upright in a tube stand. Next, a gel loading tip (200 µL) was used to load each sucrose solution. Once the 90% solution was loaded, then the following solutions were added by placing the bottom of the tip at the meniscus and gently loading the sucrose solution with a constant flow on top.

![Figure 6.3 Schematic of sucrose gradient preparation.](image)

Once the sucrose gradient was formed, discrete layers of sucrose were visible. The sample was applied to the top of the gradient, and the tube was loaded into the TLS-55 rotor (Beckman) carefully. Ultracentrifugation was started immediately at 111,000 g for 2 h at RT with an acceleration profile of 9 and a deceleration profile of 0.
Table 6.5  Volumes of each sucrose concentration used to build a discontinuous sucrose gradient.

<table>
<thead>
<tr>
<th></th>
<th>Solution</th>
<th>Volume (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (top)</td>
<td>Sample</td>
<td>200</td>
</tr>
<tr>
<td>2</td>
<td>20% sucrose</td>
<td>200</td>
</tr>
<tr>
<td>3</td>
<td>30% sucrose</td>
<td>200</td>
</tr>
<tr>
<td>4</td>
<td>50% sucrose</td>
<td>200</td>
</tr>
<tr>
<td>5</td>
<td>70% sucrose</td>
<td>200</td>
</tr>
<tr>
<td>6 (bottom)</td>
<td>90% sucrose</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1200 µL total</td>
</tr>
</tbody>
</table>

Immediately after the run, the tubes were removed from the rotor, taking care not to disturb the sucrose layers. Fractions (100 µL) were collected from the top using a loading gel tip, by placing the end of the tip right underneath the meniscus.

Collected fractions were then resolved by SDS-PAGE electrophoresis for protein, and the free DCF was determined through fluorescence readings (excitation: 490 nm and emission wavelength: 528 nm).
6.4 Results and Discussion

Expression of variants with single or double mutations in the internal cavity of the E2 scaffold yields complexes with correct structures. The E2 scaffold containing Ala, Phe, and Cys mutations at positions 239 and 381 expressed in the soluble fraction at levels comparable to the wild-type scaffold. The only exception to this was K239C, which did not appear to produce protein in *E. coli* strains BL21(DE3) or BL21(DE3)pLysS at various expression conditions. Mass spectrometry confirmed the successful production of the mutant subunits K239A, K239F, D381A, D381F, D381C, and K238F/D381F, with all molecular weights within 0.1% of the theoretical expected values.

The size and thermal stability properties of the wild-type and the mutant assemblies are summarized in Table 6.6. All mutants exhibit diameters close to that of the wild-type scaffold (Table 6.6), and all readings are slightly larger than the 24 nm determined by crystallographic data, which is expected for hydrodynamic values. These size results, together with the solubility of the protein in solution, support the correct assembly of all the single or double internal mutant scaffolds. As shown in Figure 6.1, in a fully-assembled scaffold, a single or double mutation on one subunit results in 60 or 120 mutations within the cavity, respectively.
Table 6.6 Summary of size and thermostability properties for wild-type (control) and single internal mutants. Hydrodynamic diameter is measured by dynamic light scattering. The onset of unfolding (T₀) and midpoint unfolding temperature (Tₘ) are measured by circular dichroism.

<table>
<thead>
<tr>
<th>E2-protein scaffold</th>
<th>Diameter (nm)</th>
<th>Tₘ (°C)</th>
<th>T₀ (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E2-Wild-type (control)</td>
<td>26.6 ± 0.6</td>
<td>91.4 ± 1.4</td>
<td>81.1 ± 0.9</td>
</tr>
<tr>
<td>E2-K239A</td>
<td>25.8 ± 0.4</td>
<td>93.3 ± 1.2</td>
<td>80.2 ± 1.2</td>
</tr>
<tr>
<td>E2-K239F</td>
<td>30.4 ± 0.9</td>
<td>89.1 ± 0.8</td>
<td>77.1 ± 0.8</td>
</tr>
<tr>
<td>E2-D381A</td>
<td>25.8 ± 0.5</td>
<td>90.3 ± 1.0</td>
<td>79.6 ± 0.5</td>
</tr>
<tr>
<td>E2-D381F</td>
<td>25.9 ± 1.0</td>
<td>90.6 ± 1.0</td>
<td>79.8 ± 0.2</td>
</tr>
<tr>
<td>E2-D381C</td>
<td>26.8 ± 0.6</td>
<td>89.5 ± 0.7</td>
<td>79.0 ± 0.6</td>
</tr>
<tr>
<td>E2-K239F/D381F</td>
<td>26.2 ± 0.4</td>
<td>89.0 ± 0.9</td>
<td>77.7 ± 0.9</td>
</tr>
</tbody>
</table>

We then confirmed the correct assembly and symmetry of the mutants by TEM images (Figure 6.4). All micrographs reveal correctly self-assembled dodecahedral structures of expected size. The data demonstrates that internal mutations at our chosen positions which confer alternative, non-native properties do not affect the self-assembled architecture. In fact, the most acute mutations in which 120 positions within the cavity are simultaneously changed from hydrophilic charged amino acids to large hydrophobic uncharged residues (K239F/D381F) still yield a correctly-assembled structure (data not shown).
**Figure 6.4** Transmission electron micrographs of mutant scaffolds showing dodecahedral structure. (A): K239A, (B) K239F, (C) D381A, (D) D381F, (E) D381C, and (F) K239F/D381F. Scale bar is 50 nm.

**E2 scaffold containing single and double mutations in the internal cavity shows comparable thermostability to wild-type scaffold.** Using CD, we determined the apparent midpoint unfolding temperature ($T_m$) and the onset temperature of unfolding ($T_o$) for each of the mutants. As listed in Table 6.6, all mutants have $T_m$ values near 90 ºC, which are comparable to those of the wild-type complex. $T_o$ for all variants is also near 80 ºC and close to the wild-type values. Although the onset of unfolding in K239F and K239F/D381F appears to begin at a slightly lower temperature than the wild-type scaffold (by ~3 ºC), this shift is small relative to the overall thermostability. Our data
therefore demonstrate that the physicochemical changes in these amino acid positions do not significantly decrease the thermostability and integrity of the complex.

**Mutations in the hollow cavity can change the chemical reactivity of the internal surface and demonstrate the applicability of the scaffold to encapsulate guest molecules through chemical coupling.** To assess whether these mutations can impart non-native functionality, we tested the feasibility of covalently attaching drug-like molecules within the internal cavity. For these investigations, we used the D381C mutant, which has a surface-accessible thiol side chain that reacts with our model small-molecules, F5M and AF532M, to form thioether bonds. In our wild-type E2 scaffold, there is only one native cysteine per subunit; however, the crystal structure shows that this native cysteine, residue 343, is buried in the interior of each subunit and is not surface-accessible. Therefore, we expect that any coupling to cysteine side residues will occur at the mutated position 381.

After reacting F5M or AF532M with our protein scaffold and removing unbound dye molecules, we calculated the number of covalently bound molecules to a 60-mer protein scaffold. Results are presented in Figure 6.5 and demonstrate that the cysteine mutant selectively reacts with the two maleimide molecules. Furthermore, the extent of coupling can be controlled by varying the initial dye concentration in the reaction mixture. For a 1:1 F5M-to-subunit molar concentration, the D381C scaffold yields a coupling ratio of $20.8 \pm 1.3$ molecules per 60-mer nanocapsule scaffold. While this is lower than the theoretical maximum ratio of 60, it is more than three times more than the background wild-type control of $5.7 \pm 1.7$. We presume this background is nonzero due
to either incomplete removal of nonspecifically-bound dye or the slower reaction of maleimide with free amines (Smyth et al. 1960); in the latter possibility, F5M could exhibit background reaction to the lysine side chains in the native E2 scaffold.

**Figure 6.5** Conjugation of fluorescein-5-maleimide (F5M) and Alexa Fluor 532 C₅-maleimide (AF532M) to the D381C mutant and the wild-type control. **A:** F5M molecules were reacted in 1:1 and 3:1 molar ratios of dye:protein subunit. **B:** AF532M molecules were reacted in a 3:1 molar ratio of dye:protein subunit. (Dalmau et al. 2008)

When excess dye is added in the ratio of 3:1 dye:subunit molar concentration, the mutant D381C scaffold reacts with 43.3 ± 8.4 molecules of F5M per complex and 50.9 ± 4.6 molecules of AF532M per 60-mer scaffold complex. In contrast, the wild-type control binds approximately four times less background, with 10.7 ± 5.4 molecules of F5M and 15.3 ± 2.35 molecules of AF532M per 60-mer complex. The specific reaction of the maleimide molecules to the Cys mutant is further confirmed by mass spectrometry results. Reaction of the two different maleimide dyes to the D381C introduces an additional peak at 28,532 and 28,873 Da for F5M and AF532M, respectively; these
values are within 0.1% of the expected molecular weights of 28,532 and 28,895 Da for single molecules of F5M and AF532M covalently attached to D381C. TEM micrographs (Figure 6.6) confirm that the size and assembly of the scaffold is not altered by the presence of guest molecules within the complex. These results demonstrate the ability to encapsulate foreign molecules within the internal cavity of this novel scaffold, which is an important component of using this platform in bionanotechnology applications.

![Figure 6.6](image)

**Figure 6.6** Transmission electron micrographs of D381C scaffolds which have been reacted with (A) fluorescein-5-maleimide and (B) Alexa Fluor 532 C₅-maleimide. Images show that scaffolds remain intact even after covalent encapsulation of the molecules in the internal cavity. Scale bar is 50 nm.

**Expression of variants with three and four phenylalanine mutations in the internal cavity of the E2 scaffold yields protein cages with correct structures.** All hydrophobic constructs were screened for expression and thermostability; however not all variants were expressed in *E. coli*, and some of them were not stable. The expression and thermostability test results are summarized in Table 6.7.
<table>
<thead>
<tr>
<th>Mutant</th>
<th>Expression</th>
<th>Thermostability</th>
</tr>
</thead>
<tbody>
<tr>
<td>R380F-D381F</td>
<td>++</td>
<td>T &gt; 75 ºC</td>
</tr>
<tr>
<td>D381F-G382F</td>
<td>++</td>
<td>T &lt; 75 ºC</td>
</tr>
<tr>
<td>D381F-E383F</td>
<td>++</td>
<td>T &gt; 75 ºC</td>
</tr>
<tr>
<td>E375F-D381F</td>
<td>++</td>
<td>T &gt; 75 ºC</td>
</tr>
<tr>
<td>E375F-A386F</td>
<td>+</td>
<td>T &lt; 75 ºC</td>
</tr>
<tr>
<td>G382F-A386F</td>
<td>++</td>
<td>T &gt; 75 ºC</td>
</tr>
<tr>
<td>A236F-K239F</td>
<td>++</td>
<td>T &gt; 75 ºC</td>
</tr>
<tr>
<td>D230F-R380F</td>
<td>+</td>
<td>T &lt; 75 ºC</td>
</tr>
<tr>
<td>R380F-D381F-G382F</td>
<td>+++</td>
<td>T ~ 45 ºC</td>
</tr>
<tr>
<td>D381F-G382F-E383F</td>
<td>+++</td>
<td>T ~ 45 ºC</td>
</tr>
<tr>
<td>R380F-D381F-E383F</td>
<td>+++</td>
<td>T &lt; 45 ºC</td>
</tr>
<tr>
<td>E375F-R380F-D381F</td>
<td>+++</td>
<td>T ~ 75 ºC</td>
</tr>
<tr>
<td>E375F-D381F-A386F</td>
<td>+</td>
<td>T ~ 55 ºC</td>
</tr>
<tr>
<td>G382F-E383F-A386F</td>
<td>++</td>
<td>T ~ 55 ºC</td>
</tr>
<tr>
<td>A236F-K239F-E383F</td>
<td>+++</td>
<td>T ~ 75 ºC</td>
</tr>
<tr>
<td>D230F-R380F-D381F</td>
<td>+</td>
<td>T ~ 45 ºC</td>
</tr>
<tr>
<td>E375F-E383F-A386F</td>
<td>+</td>
<td>T &lt; 45 ºC</td>
</tr>
<tr>
<td>D230F-A236F-K239F</td>
<td>++</td>
<td>T ~ 75 ºC</td>
</tr>
<tr>
<td>R380F-D381F-G382F-E383F</td>
<td>++</td>
<td>T ~ 40 ºC</td>
</tr>
<tr>
<td>R380F-G382F-E383F-A386F</td>
<td>0</td>
<td>T ~ 40 ºC</td>
</tr>
<tr>
<td>E375F-R380F-D381F-G382F</td>
<td>++</td>
<td>T ~ 40 ºC</td>
</tr>
<tr>
<td>D381F-G382F-E383F-A386F</td>
<td>+</td>
<td>T ~ 40 ºC</td>
</tr>
<tr>
<td>K239F-R380F-D381F-G382F</td>
<td>0</td>
<td>T &lt; 40 ºC</td>
</tr>
<tr>
<td>E375F-D381F-G382F-A386F</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td>K239F-E375F-R380F-D381F</td>
<td>+++</td>
<td>T ~ 60 ºC</td>
</tr>
<tr>
<td>A236F-K239F-G382F-E383F</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td>D230F-A236F-K239F-D381F</td>
<td>0</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**Table 6.7** Summary of expression results at 37 and 30ºC, and thermostability test of phenylalanine mutants. “+++” indicates good expression level and “0”, no expression.
Construction of phenylalanine mutants started first by doing double mutations at the interior cavity of the E2 protein cage. As explained in Section 6.3.2, nine residues were identified as the best candidates taking into consideration their solvent accessible surface area, their charge and the distance of the amino acid to the center of the nanocapsule. Each identified residue had an individual score that took into consideration the above mentioned properties (see Appendix A.10). From these nine residues, combinations of two were performed and those eight double-mutated candidates with higher score were chosen for screening. Expression results showed in most of the cases a slightly lower level than the E2-WT (control) at 37°C (Table 6.7). Thermostability results did not show great differences among them, since all of them were stable around 75°C. These expression levels and thermostability results were used to rank the triple-mutation candidates.

Based on the double mutants’ results of expression and thermostability, and their potential for hydrophobic drug loading, ten triple mutants were then selected for screening. Most of these variants had similar expression as wild-type, except for mutants E375F-D381F-A386F, D230F-R380F-D381F, E375F-E383F-A386F which had very low expression at 37°C, but comparable to wild-type control at 30°C. As for thermostability results, it was noticeable the effect of 180 mutations per scaffold in certain variants. New constructs with the three mutations tightly packed (R380F-D381F-G382F, D381F-G382F-E383F, R380F-D381F-E383F) and also D230F-R380F-D381F and E375F-E383F-A386F were shown to be stable up to 45°C, which is about half of the melting temperature of the wild-type control. Again, the same procedure was applied and the
expression and thermostability results from the triple mutants were used to rank the quadruple mutant candidates.

Based on the triple mutants results of expression, thermostability and their potential for hydrophobic drug loading, nine quadruple mutants were chosen for screening. A total of 240 mutations per scaffold was quite detrimental for protein expression (E375F-D381F-G382F-A386F, A236F-K239F-G382F-E383F and D230F-A236F-K239F-D381F were not expressed at all, and mutant R380F-G382F-E383F-A386F was expressed only at 30ºC). Also, all expressed mutants showed a melting temperature around 40ºC with the exception of K239F-E375F-R380F-D381F which had a melting temperature around 60ºC.

Therefore, based on expression, thermostability tests and their potential for hydrophobic drug loading, only eight mutants (four quadruple and four triple mutants) were selected for further analysis (R380F-D381F-G382F-E383F, E375F-R80F-D381FG382F, K239F-E375F-R380F-D381F, D381F-G382F-E383F-A386F, D381F-G382F-E383F-A386F, R380F-D230F-D381F, E375F-D381F-A386F, A236F-K239F-E383F, and D230F-A236F-K239F). These eight mutants were expressed at larger scale (1L batch), and purified as explained in section 6.3.3. However, upon scale-up, mutants R380F-D381F-G382F-E383F and E375F-R80F-D381F-G382F were not expressed. We observed that when new chimeras were successfully expressed, they were stable and suitable for further stability and structural analysis. Mass spectrometry confirmed the successful expression of the mutants’ individual subunits with all molecular weights within 0.005% of the theoretical expected values.
The size and thermal stability properties of the variants are summarized in Table 6.8. All mutants show diameters close to that of the wild-type control (26.6 ± 0.6 nm). These hydrodynamic values are slightly larger than the 24 nm measured by x-ray crystallography (Perham, 1991) as expected for this technique.
<table>
<thead>
<tr>
<th>E2-protein scaffold</th>
<th>Diameter (nm)</th>
<th>$T_m$ ($^\circ$C)</th>
<th>$T_o$ ($^\circ$C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild-type (control)</td>
<td>26.6 ± 0.6</td>
<td>91.4 ± 1.4</td>
<td>81.1 ± 0.9</td>
</tr>
<tr>
<td>E2-K239F-E375F-R380F-D381F</td>
<td>31.2 ± 2.6</td>
<td>94 ± 0.8</td>
<td>75.0 ± 1.3</td>
</tr>
<tr>
<td>E2-D381F-G382F-E383F-A386F</td>
<td>27.4 ± 4.5</td>
<td>85.7 ± 1.2</td>
<td>63.5 ± 1.3</td>
</tr>
<tr>
<td>E2-D230F-R380F-D381F</td>
<td>27.5 ± 0.4</td>
<td>92.7 ± 0.8</td>
<td>73.7 ± 0.8</td>
</tr>
<tr>
<td>E2-E375F-D381F-A386F</td>
<td>27.4 ± 0.6</td>
<td>89.0 ± 0.5</td>
<td>76.0 ± 0.8</td>
</tr>
<tr>
<td>E2-A236F-K239F-E383F</td>
<td>26.0 ± 0.7</td>
<td>91.7 ± 1.1</td>
<td>76.0 ± 1.3</td>
</tr>
<tr>
<td>E2-D230F-A236F-K239F</td>
<td>27.2 ± 3.7</td>
<td>84.8 ± 0.6</td>
<td>74.3 ± 0.8</td>
</tr>
</tbody>
</table>

**Table 6.8** Summary of size and thermostability properties for wild-type and hydrophobic mutant scaffolds. Hydrodynamic diameter is measured by dynamic light scattering. The onset of unfolding ($T_o$) and midpoint unfolding temperature ($T_m$) are measured by circular dichroism.

We further confirmed correct assembly of internal mutants by means of TEM (Figure 6.7). All TEM images show correctly self-assembled dodecahedral structures of the expected size. These data demonstrates that our protein cage can sustain up to 240 mutations at the interior cavity. But analysis of the scaffold’s structure and adequate design are key to yield correctly assembled constructs, as some combinations of mutations can be detrimental for self-assembly. Therefore, the size, TEM, and solubility results for these constructs confirm their correct assembly.
Figure 6.7 Transmission electron micrographs of internal hydrophobic mutants showing dodecahedral structure. A: K239F-E375F-R380F-D381F, B: D381F-G382F-E383F-A386F, C: R380F-D230F-D381F, D: E375F-D381F-A386F, E: A236F-K239F-E383F, and F: D230F-A236F-K239F. Scale bar is 50 nm.

E2 scaffold containing multiple hydrophobic mutations at the internal cavity show high thermostability. We determined the apparent midpoint unfolding temperature and the onset of unfolding for each of the selected hydrophobic mutants. Thermal profiles of one triple mutant, E2-A236F-K239F-E383F, and one quadruple mutant, E2-K239F-E375F-R380F-D381F, are presented in Figure 6.8 and are representative of typical scans of all the internal phenylalanine variants. As listed on Table 6.8, all mutants have $T_m$ values within the range of 85-90 ºC, which are quite close to the wild-type control (91.4 ± 1.4). However, the $T_o$ varies with variants. This variability might give us insight of relative thermostability of these constructs. Hence,
quadruple mutant E2-D381F-G382F-E383F-A386F and triple mutants E2-D230F-R380F-D381F and E2-D230F-A236F-K239F are the least stable ones with an onset temperature around 63, 73 and 74 °C, respectively (see Table 6.8). One clear observation is that the quadruple mutant (E2-D381F-G382F-E383F-A386F) with mutated sites at close locations is the least stable. Hence, this instability might be due to steric hindrance. However, these onset temperatures are still high and therefore demonstrate that the physicochemical changes introduced in those internal sites do not substantially affect the integrity of the protein scaffold.

Figure 6.8 Representative temperature profiles for one triple and one quadruple internal mutants. Molar ellipticity at 222 nm was monitored by far-UV CD as the sample was heated and indicates a high Tₘ around 92 °C for the triple mutant E2-A236F-K239F-E375F-R380F-D381F (red dashed line) and 94 °C for the quadruple mutant E2-K239F-E375F-R380F-D381F (black solid line), respectively. These profiles are typical of those observed for the remaining internal mutants. A complete summary of stability data for all variants is presented in Table 6.8.
**Solubility limits and supersaturation condition determination.** Since we wanted to develop a generalized assay for encapsulation of hydrophobic small drug molecules, we used as a drug-like model: 2’, 7’-dichlorofluorescein (DCF). Its absorbance spectra in pure ethanol and in 20 mM Tris/HCl buffer (pH 8.7) were obtained. DCF showed a maximum absorbance at a wavelength of 503 nm, close to the published 509 nm (Index 2006). Prior to determining the solubility limits, a calibration curve at working conditions (8% ethanol and 20 mM Tris/HCl buffer, pH 8.7) was obtained. Only those measurements below saturation at the linear range, i.e. below 1.0 AU, were considered. Therefore, samples were properly diluted so the absorbance readings were in the linear range. Furthermore, since DCF absorbance intensity is pH dependent, pH was carefully controlled during all assays. Figure 6.9 shows the solubility limits of DCF with respect to concentration of ethanol in buffer (20 mM Tris/HCl, pH 8.7) at room temperature. The solubility of DCF increases with the percentage of ethanol in solution, which agrees with the higher solubility of DCF in ethanol found in literature (DCF solubility in pure ethanol = 30 mg/mL) (The Merck Index, 2006), with a value of 3.30 mg/mL at a concentration of 8% (v/v) ethanol. However, the DCF solubility in buffer alone (0% (v/v) of ethanol) is 1.6 mg/mL, higher than the published solubility in water (< 0.1 mg/mL) (The Merck Index, 2006). We further confirmed this value by using water as a control, and the solubility limit in water at room temperature was consistent with the literature value, 0.086 ± 6·10⁻⁴ mg/mL. Therefore, this higher solubility of DCF in 20 mM Tris/HCl buffer (pH 8.7) is due to the presence of Tris base. However, since Tris buffer helps to stabilize the protein and keep the pH constant, and the DCF absorbance intensity is pH dependent, we decided to pursue these experiments in this buffer.
Figure 6.9  DCF solubility limits in 20 mM Tris/HCl buffer (pH 8.7) at room temperature. DCF solubility limits increase as the percentage ethanol increases.

Once the solubility limits of DCF were obtained, confirmation of the stochastic nature of supersaturation was carried out. Samples of saturated solutions of DCF in 8% (v/v) of ethanol in 20 mM Tris/HCl buffer (pH 8.7) were prepared as explained in section 6.3.7. Ethanol was allowed to evaporate and samples became supersaturated and DCF precipitated between 4-6 days and the stochastic nature of nucleation was observed. Positive controls (samples initially supersaturated at a concentration of DCF of 3.79 mg/mL) precipitated after 24 hours. For those samples below saturation concentration (negative controls), no precipitate was detected in 13 days, and the absorbance at 600 nm remained constant throughout the length of the experiment. Furthermore, the supernatant of supersaturated samples was measured at absorbance 503 nm for DCF concentration. DCF level was determined to be around 2.8 mg/mL, considerably higher than expected.
(1.60 mg/mL of DCF). Thus, it appears that we were not able to evaporate all the ethanol present in each sample.

**E2 scaffold containing multiple hydrophobic mutations at the internal cavity suggest comparable drug encapsulation ability as wild-type control.** To assess the encapsulation ability of the internal hydrophobic mutants compared to wild-type control, the purified protein cages were incubated with DCF solutions as explained in section 6.3.7. Sucrose gradient was used to separate free DCF and non-loaded protein capsules from DCF-loaded-capsules. The results suggest comparable encapsulation ability as E2-WT (control). Figure 6.10A shows a representative DCF fluorescence intensity profile from each sucrose gradient fraction for wild-type and all internal hydrophobic mutants. Most of the DCF eluted in the top fractions, which confirms the separation between free DCF (in top fractions) from entrapped-DCF (in bottom fractions). We observed (Figure 6.10B) that for the bottom fractions (number 12), all hydrophobic variants showed fluorescence intensity for DCF similar to wild-type (control), except for mutants D381F-G382F-E383F-A386F and R380F-D230F-D381F which showed about half the intensity of the control. Since these phenylalanine mutants were designed to improve drug loading ability compared to wild-type control, we were expecting the bottom fractions to have higher DCF fluorescence intensity than E2-WT. Determination of protein elution in the sucrose gradient fractions was performed by means of SDS-PAGE gel (Figure 6.11 A and B). Figure 6.11A presents wild-type control incubated with buffer only. As it can be seen, protein eluted between fractions 4 and 8, whereas when the protein was incubated with DCF, the protein eluted in the bottom six fractions: 7 to 12 (Figure 6.11B). When
correlating the presence of DCF dye with the presence of protein in same fractions (Figure 6.10B), both protein and dye eluted in the last six bottom fractions. Therefore, we demonstrate that through sucrose gradients we were able to separate supposedly loaded from non-loaded protein cages and also free DCF from entrapped-DCF. Furthermore, these results show that potentially loaded capsules are heavier than non-loaded ones and travel further down in the sucrose gradient, as expected. On the other hand, one might think that during the experiment protein might aggregate and therefore elutes in the denser sucrose gradient fractions. These two scenarios: aggregated protein versus loaded-protein could be distinguished by means of analytical ultracentrifugation. However, due to the lack of availability of this equipment, it could not be deciphered. In summary, although most of the internal hydrophobic mutants showed similar nucleation ability as wild-type protein, the incomplete evaporation of ethanol remains an issue and hence, the expected improved encapsulation ability of the phenylalanine mutants was not observed.
Figure 6.10  DCF content in sucrose gradient fractions was determined by fluorescence intensity measurements.  **A**: fluorescence intensity of all sucrose gradient fractions from samples with [DCF] = 3.3 mg/mL and 8% (v/v) ethanol in 20 mM Tris/HCl (pH 8.7) buffer; **B**: Zoom-in of DCF fluorescence intensity of bottom fractions. All internal hydrophobic mutants showed comparable encapsulation ability to wild-type (control), except for mutants D381F-G382F-E383F-A386F and R380F-D230F-D381F which showed about half the intensity of the control. DCF fluorescence was measured at an excitation = 490 nm and emission wavelength = 528 nm.
Figure 6.11  **A**: SDS-PAGE gel shows wild-type control without DCF eluted between fraction 4 and 8. Lanes: (1) Protein molecular weight standards (in kDa); (2) fraction 1 (top), (3) fraction 2, (4) fraction 3, (5) fraction 4, (6) fraction 5, (7) fraction 6, (8) fraction 7, (9) fraction 8, (10) fraction 9, (11) fraction 10, (12) fraction 11, (13) fraction 12 (bottom).  **B**: SDS-PAGE gel shows quadruple mutant K239F-E375F-R380F-D381F loaded with [DCF] = 3.3 mg/mL and 8% (v/v) ethanol in 20 mM Tris/HCl buffer (pH 8.7) eluted between fractions 6 and 12. Lanes: (1) Protein molecular weight standards (in kDa); (2) fraction 1 (top), (3) fraction 2, (4) fraction 3, (5) fraction 4, (6) fraction 5, (7) fraction 6, (8) fraction 7, (9) fraction 8, (10) fraction 9, (11) fraction 10, (12) fraction 11, (13) fraction 12 (bottom).
6.5 Conclusions

Our investigations demonstrate that the E2 scaffold from the pyruvate dehydrogenase multienzyme complex can be modified both chemically and genetically without changing the overall structure. Furthermore, this scaffold can tolerate up to 240 mutations in the internal surface of the hollow cavity and still remain self-assembled in a dodecahedron shape. We found the thermostability of the expressed hydrophobic mutants to be slightly lower than the wild-type control, but it still was considerably high. We show that these internal mutations can be selectively constructed to potentially encapsulate drug-like molecules at the interior cavity to introduce non-native functionality. Encapsulation can be carried out through either chemical reaction or non-specific nucleation of drug molecules. Our results demonstrate that when introducing this substantial amount of mutations per cage, it is crucial to choose the residues to substitute as they can affect protein expression and assembly stability. This high tolerance to modifications of our system makes it a potential platform in many nanotechnology applications, such as biotherapeutics.
6.6 Future experiments

Future experiments need to be conducted towards confirming that the drug-like molecules are encapsulated inside the E2 protein cage. Although there might not be direct techniques, one could use several indirect techniques that converge to the same conclusion. For example, a size-exclusion column with double detector could be used to confirm co-elution of protein and DCF. Other techniques could be based on the extraction of the fluorophore from inside the nanocapsules with an organic solvent for further analysis and quantification by means of high-pressure liquid chromatography (HPLC).

On the other hand, the higher concentration of DCF measured in the supernatant of supersaturated samples could be explained by the presence of non-evaporated ethanol. Hence, the remaining ethanol concentration should be evaluated through some analytical technique such as gas chromatography to confirm this hypothesis. Also, this incomplete evaporation of ethanol could be another factor that might decrease the encapsulation efficiency. Therefore, one possible strategy to improve ethanol evaporation could be the use of a gentle stream of air circulating through the container where the 96-well plate is placed during the encapsulation experiment.

Another issue as discussed in the results section is to distinguish between loaded-capsules from aggregated protein in the bottom fractions. This issue could be easily solved through analytical centrifugation since as the sample is being spun down, the molecular weight of the sedimenting particles is determined. Also, encapsulation of two
drug-like molecules could be used to monitor fluorescence resonance energy transfer (FRET) to distinguish between encapsulated and non-encapsulated cages.

Moreover, other model drugs should be tested to confirm reproducibility of this encapsulation assay and to confirm the general applicability of this encapsulation method.

Finally, the fluorescence measurements did not show higher intensities for the case of hydrophobic mutants relative to the wild-type control as it was expected. This observation could be due to the fact that the phenylalanine mutants constructed did not have the optimum amount of hydrophobic surface area required for non-specific encapsulation ability. Therefore, some improvement towards these hydrophobic constructs could be carried out through combinatorial libraries.
6.7 Acknowledgements

I gratefully acknowledge Dr. Sierin Lim for the construction of internal mutants for chemical conjugation, Helen Chen for carrying out the conjugation experiments, Prof. Baldi and Dr. Randall for collaboration in the design of the internal hydrophobic mutants.
CHAPTER 7

CONCLUSIONS
7.1 Conclusions

The long-term goal of this research is to develop a multifunctional nanoplatform based on a protein nanoscale scaffold to impart targeted delivery of therapeutic and imaging agents. In this work, we have investigated a virus-like particle based on the E2 subunit of pyruvate dehydrogenase multienzyme complex towards the achievement of that goal. Along with engineering this nanoparticle, further insight on protein assembly was acquired.

Our results demonstrate that the E2 structural scaffold is robust and has the features of an ideal protein cage platform for drug delivery applications. We confirmed that the truncated E2 protein cage with its structural core self-assembles into a dodecahedron structure with icosahedral symmetry and has an unusually high thermostability (described in Chapter 3).

Our work has also shown that the E2 protein scaffold is amenable to modifications at the interior surface and the interface between subunits. We engineered a non-native, pH-dependent molecular switch into the E2 protein cage assembly (described in Chapter 4). This pH-responsive variant exhibits correct assembly at physiological pH but dissociates irreversibly at pH 5.0. The design of this protein complex includes the deletion of the N-terminal domain and the introduction of repulsive electrostatic forces between adjacent histidines at the interface between subunits. This investigation also gave us insight on how important is the N-terminal domain in modulating assembly. Furthermore, we investigated the degree to which this pH-triggered assembly is tunable (described in Chapter 5). Our data demonstrates that the engineered pH-dependent
variant does involve electrostatic interactions. Indeed, we observed that the pH at which half of the protein is assembled could be tuned through ionic strength and the number of repulsive interactions between adjacent histidines at the intratrimer interface between subunits. Both genetic and chemical modifications at the interior cavity were performed (described in Chapter 6). Our scaffold showed high tolerance to physicochemical modifications of the hollow internal cavity. These changes did not disrupt the overall assembly structure. We demonstrated that these internal mutations can be designed to encapsulate foreign guest molecules within the cavity through chemical conjugation. We also developed a general assay for encapsulation of small hydrophobic drug compounds that showed comparable drug-loading ability to the wild-type control. These characteristics together with previously shown ability to attach targeting peptides on the outer surface (Domingo et al. 2001), suggest that this platform can potentially be a powerful and versatile tool in biomedical applications.
7.2 Future studies

Protein cages show great promise for biomedical applications. The results obtained in the present study support this claim. However, some improvements and further development needs to be done. The non-specific hydrophobic drug loading assay (described in Chapter 6), although it shows promising results, it still requires confirmation of encapsulation that could be accomplished through analytical ultracentrifugation, together with double detection (protein and dye) in liquid chromatography analysis. Moreover, another model drug should be tested to confirm same behavior as the dye used in this project (DCF) and to fully assess the promise of this assay. On the other hand, besides the mutants tested herein, a combinatorial strategy could be used to obtain the optimal cage with greater encapsulation ability.

In addition, receptor-mediated delivery to tumor cells is another area that can be further explored utilizing our virus-like cage. Future work may include other useful functionalities such as therapeutic imaging. Finally, the ideal system should be envisioned as one capable of entrapping the maximum cargo allowed, directing this drug agent to the site of the body where required and lastly, releasing the drug inside the cells. Therefore, one cage would possess three (or more) modifications at the interior, exterior and interface between subunits to become fully functional.
REFERENCES


Greenfield NJ. 2006. Using circular dichroism collected as a function of temperature to
determine the thermodynamics of protein unfolding and binding interactions.

Bioorthogonal Conjugation: A Practical Method for the Ligation of Diverse
16(6):1572-1579.

Gurdag S, Khandare J, Stapels S, Matherly LH, Kannan RM. 2006. Activity of
Dendrimer-Methotrexate Conjugates on Methotrexate-Sensitive and -Resistant

Integrin-mediated targeting of drug delivery to irradiated tumor blood vessels.

Hamblett KJ, Senter PD, Chace DF, Sun MMC, Lenox J, Cerveny CG, Kissler KM,
Bernhardt SX, Kopcha AK, Zabinski RF and others. 2004. Effects of Drug
Loading on the Antitumor Activity of a Monoclonal Antibody Drug Conjugate.

Harasym TO, Bally MB, Tardi P. 1998. Clearance properties of liposomes involving

Harrison PMAP. 1996. The ferritins: molecular properties, iron storage function and

Hatfield GW, Roth DA. 2007. Optimizing Scaleup Yield for Protein Production:
Computationally Optimized DNA Assembly (CODA) and Translation


Lathrop RH, Hatfield GW; 2007. Method for Producing a Synthetic Gene or Other DNA Sequence. USA patent 7,262,031.


Basis of Multiple Drug-Binding Capacity of the AcrB Multidrug Efflux Pump.
Science 300(5621):976-980.


APPENDIX
A.1 Recipes

10x DNA loading dye (20 mL stock)

- 15 mL ddH₂O
- 5 mL glycerol
- 0.08 g bromophenol blue
- 0.08 g xylene cyanol

2-log DNA ladder (500 µL)

- 100 µL new ladder tube (New England Biolabs)
- 50 µL 10x DNA loading dye
- 350 µL ddH₂O

LB Media

25 g LB Broth Miller (EMD Chemicals # 1.10285.0500)

Dissolve in 1 L ddH₂O.

Adjust pH to 7.0.

Autoclave to sterilize.

Media containing Agar

Prepare LB media according to the recipe given.

Just before autoclaving, add 15 g/liter Bacto agar (BD # 214010)

Sterilize by autoclaving on liquid cycle.

Swirl gently to distribute the melted agar evenly.
Allow the media to cool to 50-60ºC before adding thermolabile antibiotic (i.e. ampicillin).
Add 0.1 g/L of ampicillin sodium salt (EMD Chemicals # 2200)
Pour 30-35 mL of media per plate.
To remove bubbles from medium in the plate, flame the surface of the medium with a Bunsen burner before the agar hardens.
Mark the edges of the plates with a color code (one red stripe for LB-ampicillin plates; one blue stripe for LB plates).
When the medium has hardened completely, invert the plates and store them at 4ºC until needed.

A.2 Site-directed Mutagenesis protocol

Site-directed Mutagenesis protocol was derived from the Stratagene’s QuickChange® method (Figure A.1) by Dr. Lim.

![Figure A.1 Schematic of the mutagenesis procedure.](image)

1. Identify site of mutagenesis
2. Design forward and reverse primers of 45 bp each.

   When designing a primer you need to consider:
   - primer should start and end with a G or C
   - primer should have at least 12 bp on each side of mutation
   - GC content should be around 50%
   - Tm, 5-10 °C lower than extension temperature of DNA polymerase (68-72 °C)
   - Check for hairpin formation, online tool [www.idtdna.com](http://www.idtdna.com), Oligo Analyzer

3. PCR to amplify the gene using *PfuUltra* HF DNA Polymerase (Stratagene)

4. PCR clean-up using QIAquick PCR Purification Kit

5. Digest PCR product with *DpnI* enzyme (NEB)

6. Transform into sequencing strain (DH5a)

7. Miniprep 3 colonies for each construct

8. Check digest

9. Sequence new mutant
Primer preparation:

Reagent: 1x TE buffer (100 mL)

- 1M Tris-Cl pH 7.5  1 mL
- 0.5 M EDTA  0.2 mL

Add ddH₂O to a total volume of 100 mL

Filter sterilize with 0.22 µm pore filter (Millipore)

Quick spin oligos at 14000 rpm, for 1 min

Reconstitue each oligo using 1x TE buffer to ~100mM (i.e. 1 nmoles in 10 mL TE buffer)

Let sit for 10 min

Vortex

Quick spin for 1 min

Make primer mix: (20 mM)

10 µL forward primer

10 µL reverse primer

30 µL ddH₂O

50 µL total
PCR reaction

PCR mix:

1 µL DNA template (pGEM plasmid)
1 µL primer mix
2 µL 2.5 mM dNTPs (NEB)
1 µL PfuUltra HF DNA Polymerase
5 µL 10x Pfu buffer
40 µL ddH$_2$O
50 µL total

PCR thermocycling:

95°C 5 min
95°C 20 sec
Lowest Tm – (~5 ºC) 1 min (18x)
72°C 8 min
72°C 10 min
4ºC ∞

PCR cleanup using QIAquick PCR Purification Kit, elute with 30 mL ddH$_2$O

Digestion with DpnI (NEB) (0.5 mL), add 3 mL buffer 4, 37ºC for 2 hr.

Transformation into DH5α (Ca competent cells) with 5 mL of digested PCR product.

Pick 3 colonies for each construct and make 5 mL LB-Amp O/N at 37ºC.

Miniprep each colony using Qiagen kit and in last step elute with 30 mL of ddH$_2$O.
Check DNA concentration using Da Silva’s NanoDrop.

Check Digest at 37 °C for 2 hr:

4 µL Miniprep’d pGEM plasmid
1 µL BamHI buffer
0.5 µL BamHI (NEB)
0.5 µL NdeI (NEB)
0.5 µL BSA (100x)
3.5 µL ddH₂O
10 µL total

(BamHI should be added at the last 45 min)

Run on 1% agarose gel at 70 V for 50 min.

Stain the gel with ethidium bromide for 5-10 minutes.

Visualize the gel on a UV transilluminator.

If check is OK, then send samples for sequencing (www.genewiz.com)
A.3 Preparation of Calcium-competent cells

Protocol adapted from Prof. Wang.

Reagents:

Calcium chloride solution

- 60 mM CaCl₂
- 15% (w/v) glycerol
- 10 mM PIPES, pH = 7.0

Mix and autoclave.

Keep refrigerated in fridge.

Procedure:

1. Innoculate a single colony of *E. coli* into 5 mL LB media.
2. Grow it O/N at 37°C at 250 rpm.
3. Innoculate 0.5 mL of the overnight culture into two 50 mL LB sterile tubes.
4. Grow them for approximately 2.5 hours at 37°C and 250 rpm.
5. Transfer cultures to two 50 mL prechilled, sterile polypropylene tubes and let sit on ice for 5-10 minutes.
6. Centrifuge tubes for 10 minutes at 4°C and 3300 rpm.
7. Pour off supernatant.
8. Resuspend each pellet in 10 mL ice-cold calcium chloride solution.
9. Centrifuge for 10 minutes at 4°C and 3300 rpm.
10. Discard supernatant.
11. Resuspend each pellet in 10 mL ice-cold calcium chloride solution.

12. Keep on ice 30 minutes to 2 hours.

13. Centrifuge for 10 minutes at 4°C and 3300 rpm.

14. Resuspend each pellet in 2 mL ice-cold calcium chloride solution.

15. Dispense into prechilled sterile centrifuge tubes 100-200 µl aliquots.

16. Place immediately on dry ice (-70°C)

17. Store at -80°C freezer.
A.4 Transformation of Calcium-competent cells

Reagents:

**SOB Medium**

- 100 mL ddH₂O
- 2 g peptone
- 0.5 g yeast extract
- 0.05 g NaCl

Mix until the solutes have dissolved.

Adjust pH to 7.0.

Sterilize by autoclaving.

Let cool to room temperature and add 10 mM MgSO₄ and MgCl₂.

Procedure:

1. Thaw frozen competent cells on ice.
2. Aliquot DNA (1 µL) into calcium-competent cells solution and mix gently.
3. Let sit the mixture of DNA and cells on ice for 10 minutes.
5. Chill mixture rapidly on ice for 1-2 minutes.
6. Add 500 µL of SOB media and incubate 1 hour at 37°C and 250 rpm.
7. Plate out culture on plates.
A.5 Characterization of purified protein with SDS polyacrylamide gel electrophoresis (PAGE)

A 20 µL sample is mixed with 4 µL of 6x SDS loading buffer solution. The mixture is boiled for 5 min. Only 10 µL of this mixture are loaded on each well. The gel is run at 120 V for 1 hour and 20 min. SDS-PAGE gel is stained with Imperial Protein stain (Pierce) following standard procedure.

Reagents:

10x SDS Running Buffer
30 g Tris base
144 g Glycine
10 g SDS (sodium dodecyl sulfate)
Add ddH₂O to 1 l and mix till all solutes are dissolved.
DO NOT ADJUST pH

6x SDS-PAGE Loading Buffer
7 mL 0.5 M Tris/HCl pH 6.8 (SDS-PAGE stacking buffer)
3 mL glycerol
1 g SDS (sodium dodecyl sulfate)
0.93 g DTT (dithiothreitol) (0.6 M final)
1.2 mg bromophenol blue
0.5 M Tris/HCl pH 6.8 buffer (200 mL)

12.11 g Tris base
150 mL ddH₂O

Titrate with 6 N HCl to pH 6.8
Add water to final volume of 200 mL

12% resolving gel

3.4 mL ddH₂O
2.5 mL 1.5 M Tris pH 8.8
0.1 mL 10% SDS (sodium dodecyl sulfate)
4 mL Acrylamide (30 : 0.8%)
0.1 mL 10% (w/v) APS (ammonium persulfate)
5 µL TEMED (tetramethylethylenediamine)
10 mL Total

4% stacking gel

844 µL ddH₂O
270 µL 1.5 M Tris pH 8.8
13 µL 10% SDS (sodium dodecyl sulfate)
173 µL Acrylamide (30 : 0.8%)
13 µL 10% (w/v) APS (ammonium persulfate)
1 µL TEMED (tetramethylethylenediamine)
1.3 mL Total
A.6 Determination of protein concentration using Micro BCA Protein Assay

Protein concentrations were determined using a Micro BCA Protein Assay Kit (Pierce) with bovine serum albumin (BSA) as a standard.

The Micro BCA Protein Assay combines the well-known reduction of Cu$^{2+}$ to Cu$^{1+}$ by protein in an alkaline medium with the sensitive colorimetric detection of the cuprous cation (Cu$^{1+}$) by bicinchoninic acid.

First, a fresh working reagent (WR) solution is prepared. The total volume of WR required is determined as follows:

\[
(\text{#standards} + \text{# samples}) \times \text{ (# replicates)} \times 100 \mu\text{L per sample} = \text{total volume of WR required}
\]

The WR solution is prepared by mixing 50 parts of Micro BCA™ Reagent MA and 24 parts of Reagent MB with 1 part of Reagent MC. Solution is vortexed until completely homogeneous.

In a 96-well plate, 100 µL of protein sample is mixed with 100 µL of WR solution and plate is shaken thoroughly. Then plate is covered and incubated at 37°C for 2 hr. Samples are left 5 min at room temperature for equilibration before taking absorbance readings at 562 nm.

BSA standard curve is used to determine protein concentration of each unknown sample.
A.7 Packing SUPEROSE 6 XK 16/70 gel filtration column

Specifications:

- ID = 1.6cm
- Length = 70cm
- Volum = \( \pi r^2 L = 140.73 \) mL
- Resin: Superose 6 Prep Grade (50% slurry)
- Estimated bed height: 64.5cm
- Minimum volume of resin: \( V = \pi (0.8^2) \cdot 64.5 + 15\% \) compression = 149.13 mL

Procedure:

1. Level the column. It is really IMPORTANT to have the column leveled so the plates are leveled.
2. Assemble the bottom of the column. Make sure that the filters are not damaged and that all parts are clean. Wet the bottom filter and mesh with 20% EtOH and place them on top of the bottom adaptor. Make sure that NO air bubbles are present. Screw the bottom adaptor to the bottom of the column, and then close the inner o-ring. Cap the bottom.
3. Pour \( \sim 2 \) mL of 20% EtOH into the column.
4. Attach reservoir to the top of the column. Check that the column is still leveled.
5. Mix the resin with a non-metallic spatula (metallic ones can damage the resin beads) to make the solution as much homogeneous as possible.
6. Pour the resin solution into the column in one swift motion without introducing air bubbles (pour it along the side of the reservoir).

7. Fill the reservoir with 20% EtOH to the brim.

8. Have 2 mL/min 20% EtOH running through the AKTA system. Connect the tubing (line 1) to the top of the reservoir (First connect the tubing to the FPLC valve and then screw it to the reservoir’s cap). Open the bottom of the column.

9. Let it run at 2 mL/min. When the pressure starts getting close to 0.35 MPa, just set it to SystemPumpControlMode. (Under Manual, Alarms&Mon, select Pump, and then SystemPumpControlMode. Select Mode: PressFlowControl, PressLevel: 0.37MPa).

10. When the resin is reaching the top of the column, place the mesh and filter at the bottom of the plunger. Wet mesh and filter with 20% EtOH. Make sure that NO air bubbles are present.

11. Once all the resin is inside the column, cap the bottom of the column and pause the flow rate. Disconnect the tubing from the reservoir, unscrew the reservoir and siphon off the EtOH inside the reservoir. Be careful DO NOT DISTURB the top resin in the column. (Siphon from the side of the reservoir). Connect the top adapter to the column. First tighten the o-ring and then screw the adaptor to the column. Then connect the column inlet to the pump. MAKE SURE THAT NO AIR BUBBLES ARE INSIDE THE TUBING AND/OR COLUMN.

12. Resum the flow rate till the top liquid in the column is clear. Mark the top of the resin level.
13. Then pack the column for 1 hour with setting the maximum pressure to 0.42 MPa (in the SystemPumpControlMode).

14. Once it’s packed, lower the plunger 2 mm inside the resin. (Cap the bottom of the column, then disconnect the tubing from the valve and slightly open the o-ring and lower the plunger. Then close the o-ring and reconnect the tubing to the valve, and open the bottom of the column).

15. Check that the top is packed by running first at low flow rates (i.e. 0.5 mL) and then increase the flow rate to the maximum allowed.

16. The maximum flow rate you can run the column is 70% of the packing flow rate. (For example, if you pack the column at 1.43 mL/min, the maximum flow rate you can run the column is 1 mL/min).

17. If the top resin level lowers a bit, just wait till it’s stable and then repeat the lower the plunger 2-mm inside the resin step.

18. Run ddH₂O through the column and then run the efficiency test (HETP) by injecting 1 mL of 0.5% acetone in water.
A.8 Differential Scanning Calorimetry

Protocol developed by Cesar Ruiz.

Procedure:

Dummy Scan

The first scan of the day is not considered reliable according to the DSC manual. The first scan tends to be shifted more to the right than what it should be. Therefore, with a dummy scan that is set with the parameters of the actual scan tested will put the DSC into the same cycle of thermal history for the sample scan.

- Degas the buffer for 10 minutes.
- Add filling funnel to the cell. It is important to use filling funnel because it will allow the syringe needle to be placed into the cell and be positioned 1-2 mm from the bottom. **Damage to the bottom of the cell will be caused if filling funnel is not used (when removing/adding solution to cell always use supplied filling funnel and syringe needle!) You may switch the syringe depending on the volume you expect to extract as long as you continue to use the supplied syringe needle.
- Cells are loaded with solution from the bottom of the cell to the top.
- Rinse each cell (sample and reference) with DI water using the filling funnel and filling syringe needle. Use a 3-mL syringe (repeat 5 times).
- Rinse each cell once with the degassed buffer.
- Extract 0.7 mL of solution into the syringe. Use a 1-mL syringe.
- Slowly fill the reference cell with degassed buffer (0.6 mL per cell).
- Administer 2-5 shorts burst of solution (0.1 mL) to remove bubbles in the cell.
- Remove the syringe and then the funnel from the cell.
- Pipette excess solution in the cells from the sides of the reservoir (Do not insert tip inside the cell).
- Repeat steps with sample cell.
- Add the pressurizing cap to the cells.
- Secure the lower portion of the cap by tightening all the way (The top part of the cap will help adjust pressure).
- Adjust pressure to 30 psi (This may be difficult because it depends on how well the threads line up) May require multiple tries.
- Enter parameters: Number of scans (usually 1), Initial Temperature, Final Temperature, Scan Rate, Prescan thermostat (time to stay at initial temp), postscan thermostat (time to stay at final temp), and filtering period. Feedback mode is set to none and scan edit mode is set to unique scans. Enter file name (will automatically save).
- Start DSC scan.

Sample Scan

- Once dummy scan is done, remove the buffer from each cell.
- Degas protein sample and buffer for 10 minutes.
- Rinse each cell (sample and reference) with DI water using the filling funnel and filling syringe needle. Use a 3-mL syringe (repeat 5 times).
- Rinse each cell once with the degassed buffer.
- Extract 0.7 mL of solution into the syringe. Use a 1-mL syringe.
- Slowly fill the reference cell with degassed buffer (0.6 mL per cell).
- Slowly fill the sample cell with degassed protein sample (0.6 mL per cell).
- Administer 2-5 shorts burst of solution (0.1 mL) to remove bubbles in the cell.
- Remove the syringe and then the funnel from the cell.
- Pipette excess solution in the cells from the sides of the reservoir (Do not insert tip inside the cell).
- Add the pressurizing cap to the cells.
- Secure the lower portion of the cap by tightening all the way (The top part of the cap will help adjust pressure).
- Adjust pressure to ~30 psi (This may be difficult because it depends on how well the threads line up).
- Enter parameters: Number of scans (usually 1), Initial Temperature (°C), Final Temperature (°C), Scan Rate (°C/hr), Prescan thermostat (min) (time to stay at initial temp), postscan thermostat (min) (time to stay at final temp), and filtering period (sec). Feedback mode is set to none and scan edit mode is set to unique scans. Enter file name (will automatically save).
- Start DSC scan.

Cleaning Scan

- Cleaning the cells after sample scan is very important especially if sample protein precipitates.
- Using a thermovac is preferred (not available) because it provides high pressure/temperature rinsing and it is all automated. Detergent soak is the next best cleaning procedure.
- Remove the solution from the reference and the sample cell.
- Add 0.6 mL of 10-20% aqueous Contrad-70 solution to each cell.
- It is not necessary to remove excess Contrad-70 solution from each cell.
- Fold a lightly damped paper towel (with water) to a size which covers both cells.
- Cover the cells with this damped towel.
- Set initial temp to 61°C.
- Set final temp to 65°C.
- Set prescan thermostat to 45 min.
- Enter file name.
- Start DSC scan.
- After heat treatment, remove the Contrad-70 solution from each cell.
- If there is no access to a thermovac then you must rinse each cell with at least 100 mL of DI water for each cell. Use a 3-mL syringe.
- Remove all water from each cell.
- Place pressurizing cap back on and tighten.
- Shut down computer.

**Contrad-70 is recommended to be used once a day. Too much use of Contrad-70 may damage the cell which is composed of tantalum alloy. Therefore, if your sample precipitates, only one scan a day should be done.**
A.9 Circular Dichroism Spectroscopy

Circular dichroism (CD) is an excellent method for rapidly evaluating the secondary structure, folding and binding properties of proteins.

Reagents:

Samples for CD spectroscopy must be at least 95% pure by the criteria of high-performance liquid chromatography (HPLC), mass spectroscopy or gel electrophoresis. In our case we use purified protein at 0.06 mg/mL in 50 mM potassium phosphate buffer and 100 mM sodium chloride at pH 7.4.

Equipment:

- 0.1 cm path length quartz cuvette (Hellma USA)
- Jasco 810 spectropolarimeter equipped with a Jasco peltier temperature controller.

Procedure:

1. Open valve on liquid nitrogen tank and regulate its flow rate to be around 30 mL/min on the flow meter.

2. Turn on the circulating cooling bath.

3. Turn on Jasco peltier temperature controller.

4. Open CD software and set temperature controller.

5. For a wavelength scan, set following parameters:
   - Start: 260 nm
   - End: 190 nm
   - Sensitivity: 100 mdeg
   - Data pitch: 0.1 nm
6. Press start and let it run.

7. For a temperature scan, set parameters to:
   - Wavelength: 222.0 nm (this wavelength needs to be characteristic of your protein sample).
   - Start: 40.0 ºC
   - End: 95 ºC (range of temperature you desire to analyze)
   - Sensitivity: 100 mdeg
   - Data Pitch: 0.5 ºC
   - Scanning mode: continuous
   - Temperature slope: 1 ºC/min
   - Response: 1 sec
   - Bandwidth: 2 nm
   - Accumulation: 1

8. Press start and let it run.
A.10 Hydrophobic mutants scores

These results were obtained in collaboration with Prof. Baldi and Dr. Randall.

The overall score is obtained from the solvent accessibility score, the distance to the center of the core and the charge of the residues, equally weighted.
<table>
<thead>
<tr>
<th>Mutant</th>
<th>Overall score</th>
<th>Solvent accessibility</th>
<th>Distance to core</th>
<th>Charge status</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Double mutants:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R380F-D381F</td>
<td>12.487</td>
<td>0.59</td>
<td>0.37</td>
<td>1.00</td>
</tr>
<tr>
<td>D381F-G382F</td>
<td>10.362</td>
<td>0.68</td>
<td>0.47</td>
<td>0.50</td>
</tr>
<tr>
<td>D381F-E383F</td>
<td>9.404</td>
<td>0.61</td>
<td>0.51</td>
<td>1.00</td>
</tr>
<tr>
<td>E375F-D381F</td>
<td>6.638</td>
<td>0.43</td>
<td>0.35</td>
<td>1.00</td>
</tr>
<tr>
<td>E375F-A386F</td>
<td>3.764</td>
<td>0.25</td>
<td>0.43</td>
<td>0.50</td>
</tr>
<tr>
<td>A236F-K239F</td>
<td>2.556</td>
<td>0.39</td>
<td>0.57</td>
<td>0.25</td>
</tr>
<tr>
<td>D230F-R380F</td>
<td>1.053</td>
<td>0.43</td>
<td>0.09</td>
<td>0.50</td>
</tr>
<tr>
<td><strong>Triple mutants:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R380F-D381F-G382F</td>
<td>2.892</td>
<td>0.43</td>
<td>0.31</td>
<td>0.50</td>
</tr>
<tr>
<td>D381F-G382F-E383F</td>
<td>1.46</td>
<td>0.45</td>
<td>0.42</td>
<td>0.50</td>
</tr>
<tr>
<td>R380F-D381F-E383F</td>
<td>1.336</td>
<td>0.39</td>
<td>0.33</td>
<td>1.00</td>
</tr>
<tr>
<td>E375F-R380F-D381F</td>
<td>0.727</td>
<td>0.27</td>
<td>0.23</td>
<td>1.00</td>
</tr>
<tr>
<td>E375F-D381F-A386F</td>
<td>0.435</td>
<td>0.23</td>
<td>0.25</td>
<td>0.50</td>
</tr>
<tr>
<td>G382F-E383F-A386F</td>
<td>0.642</td>
<td>0.26</td>
<td>0.50</td>
<td>0.25</td>
</tr>
<tr>
<td>A236F-K239F-E383F</td>
<td>0.046</td>
<td>0.29</td>
<td>0.08</td>
<td>0.50</td>
</tr>
<tr>
<td>D230F-R380F-D381F</td>
<td>0.356</td>
<td>0.21</td>
<td>0.13</td>
<td>1.00</td>
</tr>
<tr>
<td>E375F-E383F-A386F</td>
<td>0.331</td>
<td>0.16</td>
<td>0.37</td>
<td>0.50</td>
</tr>
<tr>
<td>D230F-A236F-K239F</td>
<td>0.01</td>
<td>0.15</td>
<td>0.03</td>
<td>0.50</td>
</tr>
<tr>
<td><strong>Quadruple mutants:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R380F-D381F-G382F-E383F</td>
<td>0.187612</td>
<td>0.28</td>
<td>0.27</td>
<td>0.50</td>
</tr>
<tr>
<td>R380F-G382F-E383F-A386F</td>
<td>0.046378</td>
<td>0.16</td>
<td>0.32</td>
<td>0.25</td>
</tr>
<tr>
<td>E375F-R380F-D381F-G382F</td>
<td>0.030835</td>
<td>0.20</td>
<td>0.19</td>
<td>0.50</td>
</tr>
<tr>
<td>D381F-G382F-E383F-A386F</td>
<td>0.023774</td>
<td>0.24</td>
<td>0.29</td>
<td>0.25</td>
</tr>
<tr>
<td>K239F-R380F-D381F-G382F</td>
<td>0.012903</td>
<td>0.35</td>
<td>0.11</td>
<td>0.50</td>
</tr>
<tr>
<td>E375F-D381F-G382F-A386F</td>
<td>0.009576</td>
<td>0.17</td>
<td>0.20</td>
<td>0.25</td>
</tr>
<tr>
<td>K239F-E375F-R380F-D381F</td>
<td>0.003245</td>
<td>0.22</td>
<td>0.08</td>
<td>1.00</td>
</tr>
<tr>
<td>A236F-K239F-G382F-E383F</td>
<td>0.000655</td>
<td>0.21</td>
<td>0.06</td>
<td>0.25</td>
</tr>
<tr>
<td>D230F-A236F-K239F-D381F</td>
<td>0.000126</td>
<td>0.14</td>
<td>0.02</td>
<td>0.50</td>
</tr>
</tbody>
</table>
A.11 Matlab in house program to determine protein melting temperature ($T_m$) and onset of unfolding ($T_o$)

```matlab
clear all;
load('file.txt');

% variables
T_orig = file((1:108),1);
f_exp(:,1) = file((1:108),2)*1690320/(10*0.06*0.1*15600);

plot(T_orig,f_exp,'k','LineWidth',2);
xlabel('Temperature [ºC]' )
ylabel('\Theta [deg.cm^2/dmol]')

hold on; %grid on;

% initial parameters.
h = -25000; % starting enthalpy in cal/mol
Tm = 80; % starting $T_m$ in deg. C.
u = -4*10^4; % mean residue ellipticity 100% folded
l = -1*10^4; % mean residue ellipticity 100% unfolded
u1=100; %linear correction folded function of Temp
l1=-45; % linear correction unfolded function of Temp

T = T_orig + 273.15; % convert Centigrade to Kelvin
Tm = Tm + 273.15; % convert Centigrade to Kelvin

x0(1) = h;
x0(2) = Tm;
x0(3) = u;
x0(4) = l;
x0(5) = u1;
x0(6) = l1;
x = lsqcurvefit(@f_theo_function,x0,T,f_exp(:,1));

h = x(1)
Tm = x(2)
u = x(3)
l = x(4)
u1 = x(5)
l1 = x(6)
K = K_function(h,T,Tm);
alpha = alpha_function(K)
f_theo = f_function(u,u1,l,l1,alpha,T);
plot(T_orig,f_theo,'b', 'LineWidth', 2)
legend('experimental','fitted curve');
```
for n=1:(size(f_theo)-1)
    \( f_{\text{deriv}}(n,1) = \frac{(f_{\text{theo}}(n+1,1) - f_{\text{theo}}(n,1))}{0.5}; \)
end

figure;
plot (T_orig(2:(size(T_orig))), f_deriv, 'b', 'LineWidth', 2)
title(['CD_1st derivative'])
xlabel('Temperature [°C]')
ylabel('d\Theta/dT')

[C,I]=max(f_deriv);
Tmelt=T_orig(I)

found = 0;
for n=1:(size(f_deriv)-1)
    if f_deriv(n)>=0.5 && found == 0
        f_deriv(n);
        d=n;
        found = 1;
    end
end
Tonset=T_orig(d)

function K = K_function(h,T,Tm)
    for m = 1:size(T,1)
        \[ K(m,1) = e^{(h/(1.987*T(m,1)))*((T(m,1)/Tm)-1)); \] % calculate folding constant at any given temperature.
    end
end

function alpha = alpha_function(K)
    for n=1:size(K,1)
        \[ \text{alpha}(n,1) = K(n,1)/(1+K(n,1)); \] % calculate fraction folded at any given temperature.
    end
end
function f_theo = f_function(u,u1,l,l1,alpha,T)

    for n=1:size(alpha,1)
        f_theo(n,1)=alpha(n,1)*((u+(u1*T(n,1))-
           (l+(l1*T(n,1))))+(l+l1*T(n,1))); % calculate ellipticity at any given
        % temperature.
        end

end