Title
Applying ESI-GEMMA towards the Study of Large Protein Complexes

Author
Woodruff, Kristina

Publication Date
2012

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Applying ESI-GEMMA towards the Study of Large Protein Complexes

A thesis submitted in partial fulfillment
of the requirements for the degree Master of Science
in Biochemistry and Molecular Biology

by

Kristina Pan Woodruff

2012

ABSTRACT OF THE THESIS

Applying ESI-GEMMA towards the Study of
Large Protein Complexes

by

Kristina Pan Woodruff

Master of Science in Biochemistry and Molecular Biology
University of California, Los Angeles, 2012
Professor Joseph A. Loo, chair

Electrospray ionization (ESI) coupled to gas-phase electrophoretic mobility molecular analysis (GEMMA) is a soft ionization technique capable of characterizing large protein complexes. Particles are separated by their mobility in air and their size is determined as electrophoretic mobility diameter (EMD). This data can be converted to molecular weight information by modeling the particles as spheres and taking into account their average density. Here we adapt ESI-GEMMA towards the study of 1.5 MDa E2 protein cages and 9 MDa vault nanoparticles. Recombinant vaults with N- and C- terminal tags of varying mass were used as standards to calibrate the GEMMA data specifically for the study of vault complexes. We used this improved calibration to evaluate the composition and integrity of different vault preparations. Additionally, we quantified the capacity of vaults to be loaded with protein drugs and monitored the stability of these formulations over time.

The thesis of Kristina Pan Woodruff is approved.

Emil Reisler

Leonard H. Rome

Joseph A. Loo, Committee Chair

University of California, Los Angeles

2012

Table of Contents

List of Equations and Tables

Recalibration of GEMMA for the Study of Vault Proteins

Structural Changes to the Vault upon Transition to the Gas Phase

Quantitation of Proteins Encapsulated inside the Vault

List of Figures

Effects of Vault Sample Preparation on GEMMA Spectra

GEMMA for the Study of E2 Proteins

List of Abbreviations

CPC : condensation particle counter

DMA : differential mobility analyzer

EMD: electrophoretic mobility diameter

ESI : electrospray ionization

GEMMA : gas-phase electrophoretic mobility molecular analysis

LC : liquid chromatography

MS : mass spectrometry

MVP : major vault protein

SEC: size exclusion chromatography

TEP1: telomerase-associated protein 1

VLP : virus-like particle

vPARP: vault poly ADP ribose polymerase

vRNA: vault RNA

Acknowledgments

I would like to thank Joe Loo and Rachel Loo for their excellent technical advice and project guidance. I thank Shirley Lomeli for her suggestions regarding GEMMA sample preparation. GEMMA instrumentation training was provided by Thi Nguyen and Carly Ferguson. I am grateful towards Nalaka Rannulu for providing LC-MS data. I acknowledge my collaboration with the Rome laboratory at UCLA and thank Professor Leonard Rome as well as Professor Emil Reisler for serving as members of my committee.

I am grateful towards Valerie Kickhoefer and Jan Mrazek for providing the purified vault samples. E2 samples were obtained in collaboration with Mareike Posner at the University of Bath. I acknowledge the work of Yang et al. (Copyright 2010 American Chemical Society), Kickhoefer et al. (Copyright 2005 National Academy of Sciences, USA), Peng et al. (Copyright 2012 American Chemical Society), Bacher et al. (Copyright 2001 John Wiley and Sons), Kaddis et al. (Copyright 2007 Elsevier), and Tanaka et al. (Copyright 2009 The American Association for the Advancement of Science), from which figures were reproduced with permission.

Chapter 1

Introduction

1.1 Gas Phase Electrophoretic Mobility Molecular Analysis (GEMMA)

The ability to probe the structure and composition of protein complexes as well as the dynamics of their assembly is central to understanding biological processes. However, accurate determination of this information for large protein assemblies remains an ongoing challenge (Freeke et al., 2010). Techniques that provide high-resolution structural data, such as NMR and X-ray crystallography, require high sample purity and impose limitations on protein size (Loo 1997). Mass spectrometry differs from these methods in its sensitivity, facilitating the study of less concentrated samples (Kirshenbaum et al., 2010). Proteins need not be purified to homogeneity, allowing one to concurrently examine discrete species in a sample and determine their relative abundances. Lastly, the rapidity of mass spectrometry assays allows for real-time monitoring.

ESI-GEMMA (electrospray ionization coupled to gas-phase electrophoretic mobility molecular analysis) is a method capable of acquiring size information for large protein complexes. This soft ionization approach preserves non-covalent bonds, permitting the study of intact macromolecular assemblies. However, as protein size increases, the number of possible charge states for the ESI-generated gas-phase ions also escalates. The consequential abundance of multiply charged ions makes spectra interpretation difficult (Scalf et al., 1999). The ESI-GEMMA setup addresses this impediment by utilizing a ^{210}Po α ionization source to neutralize multiply charged ions to singly charged species (Ebeling et al., 2000). The instrument separates these neutralized particles according to their mobility in air and detects their size as

electrophoretic mobility diameters (EMD) (Thomas et al., 2004). EMD can be converted to molecular weight by modeling the particles as spheres and taking into account their average density values (Kaufman 1998). The features of ESI-GEMMA discussed above make the technique amenable to the study of large protein complexes such as 9 MDa vaults.

Although standards of known molecular weight are readily available for smaller proteins, considerably less GEMMA data has been compiled for larger complexes (Kaddis et al., 2007). It will be important to establish whether the EMD-mass correlation exhibited by small proteins holds true for assemblies exceeding 1 MDa. We are also interested in determining to what degree the structure of a protein influences its EMD. GEMMA is performed in the gas phase, and proteins may undergo configurational changes upon desolvation. We questioned if raw EMD measurements require correction before they can accurately reflect a native protein in liquid media.

1.2 Vault Protein Complexes

Vaults are hollow 13 MDa ribonucleoproteins with an interior volume large enough to encapsulate hundreds of proteins (Mikyas et al., 2004; Kong et al., 1999). The abundance of vaults in the cytosol and their highly conserved nature renders them unlikely to trigger an immune response (Suprenant et al., 2002). The vault complex self-assembles from 78 copies of the 97 kDa major vault protein (MVP), which contributes to more than 70% of the vault particle mass (Kickhoefer 1999a,b; Tanaka et al., 2009). Endogenous vaults are packaged with vPARP (vault poly ADP ribose polymerase), which ADP ribosylates itself and MVP. TEP1 (telomerase-associated protein 1) occupies the vault interior in lower quantities than vPARP and assists in loading vault RNAs into vaults. The small untranslated vRNAs (vault RNAs) associated with

TEP1 are transcribed by PolIII. One of these vRNAs has been shown to regulate CYP3A4, a drug-metabolizing enzyme, by an miRNA-like mechanism (Persson et al., 2009).

As demonstrated in Figure 1, the MVP subunits may be modified at the N- and C- termini to generate vaults with pharmaceutical applications (Kickhoefer et al., 2005). These modifications generally maintain the vault's shape and stability, despite their effect on increasing molecular mass. Vaults bearing various tags can be expressed and purified as MVP complexes devoid of vPARP, TEP1, and vRNA (Stephen et al., 2001). Accordingly, recombinant vaults can potentially serve as molecular weight standards to calibrate ESI-GEMMA for the analysis of large protein complexes. By investigating the relationship between the theoretical mass of vault variants and their experimental EMD values, we derived a relationship better suited for the mass determination of vaults. Evaluation of vault structure allowed us to ascertain whether or not this correlation could be applied towards large protein complexes in general.

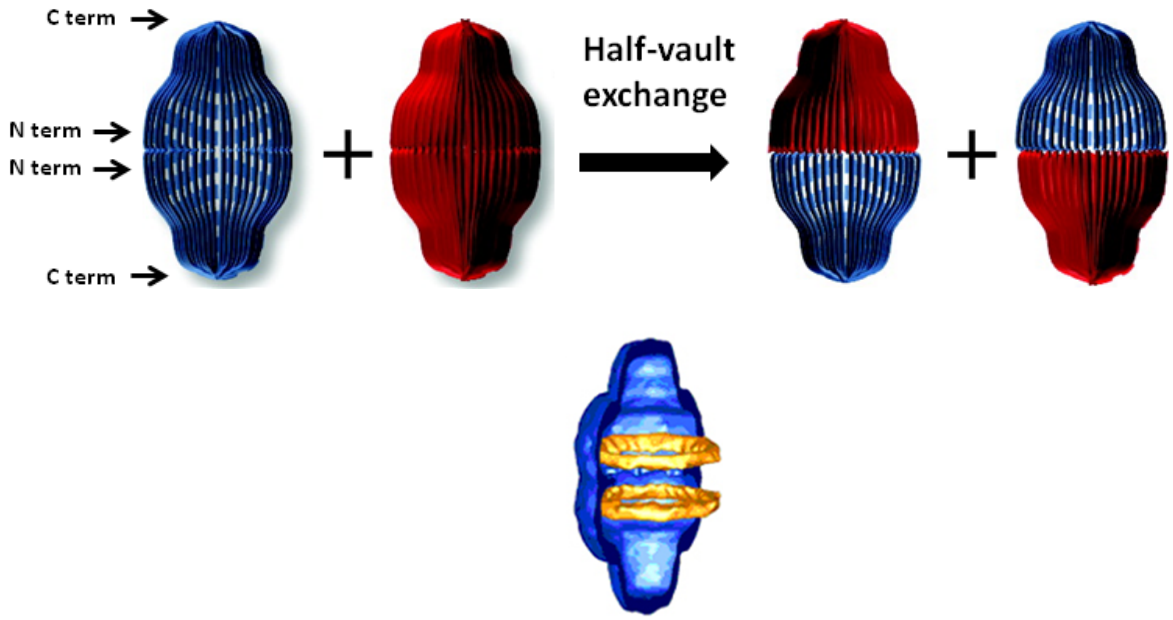

Figure 1. Structure of the recombinant vault (from Yang et al., 2010 and Kickhoefer et al., 2005). Top: Each vault half consists of 39 MVP subunits. The C- termini may be modified to include antibody-binding domains. N-terminal modifications alter the vault's ability to separate at the waist. Bottom: Blue represents the MVP "shell". Yellow represents encapsulated proteins that interact with MVP through fused INT domains.

Vaults have been implicated to participate in many pathways including cell signaling, nuclear-cytoplasmic transport, multidrug resistance, and innate immunity (Berger et al., 2009). Although the precise function of these complexes remains elusive, vaults have emerged as potential drug delivery agents due to their large hollow structure (roughly 40 x 40 x 67 nm) and biocompatibility (Tanaka et al., 2009). In drug delivery efforts, proteins of interest (i.e. chemotherapeutic CCL21 and model antigen ovalbumin) can be fused to the INT domain of vPARP (amino acids 1563-1724) (Yang et al., 2004; Song et al., 1997; Kickhoefer et al., 2005). This domain binds to the MVP subunits at the interior of the vault, and co-incubating the INT fusion protein with MVP during the vault purification process is sufficient for encapsulation (Poderycki et al., 2006) (Figure 1). However, the propensity of the complex to split at the waist during half-vault exchange raises the question of whether or not INT-fused therapeutics will escape from the vault before arrival at the target destination (Yang et al., 2010). ESI-GEMMA provides a convenient method to quantitate initial drug loading of vaults and monitor loss of encapsulated drug over time.

1.3 E2 Protein Complexes

E2 is the dihydrolipoyl acyltransferase component of the pyruvate dehydrogenase of *Bacillus stearothermophilus* (Domingo et al., 2001). This protein is also expressed in mammals, yeast, fungi, and gram-positive and gram-negative bacteria. E2 proteins assemble into cages, with subunit organization varying by species (Patel and Korotchkina, 2006). E2 obtained from *B. stearothermophilus* is often referred to as virus-like particles (VLPs) despite a lack of sequence homology and distinct function from viruses (Trovato et al., 2012). This comparison arises from structural similarities; E2 complexes, HIV, rice dwarf virus, and coronavirus spike protein all

consist of trimer intermediates that assemble into dodecahedral configurations (Hill et al., 1996; Iwasaki et al., 2008; Delmas et al., 1990; Izard et al., 1999).

The 28 kDa catalytic core of E2 contains a C-terminal domain that facilitates organization into trimers (Trovato et al., 2012). Subsequently, twenty trimers aggregate to form a 60-chain core (Perham 2000) (Figure 2). Extensive hydrophobic and electrostatic interactions amongst the subunits impart stability to the scaffold (Peng and Lim, 2011). The complex's molecular weight exceeds 1.5 MDa, comprising an outer diameter of 24 nm, a hollow core with a diameter of 12 nm, and 12 openings of 5 nm each (Peng and Lim, 2011). *In vivo,* up to 60 copies of E1 (2-oxoacid decarboxylase, ~150 kDa) and E3 (dihydrolipoamide dehydrogenase, ~100 kDa) enzymes noncovalently associate with the 60-mer E2 (Domingo et al., 2001; Perham 1991, 2000, 2002).

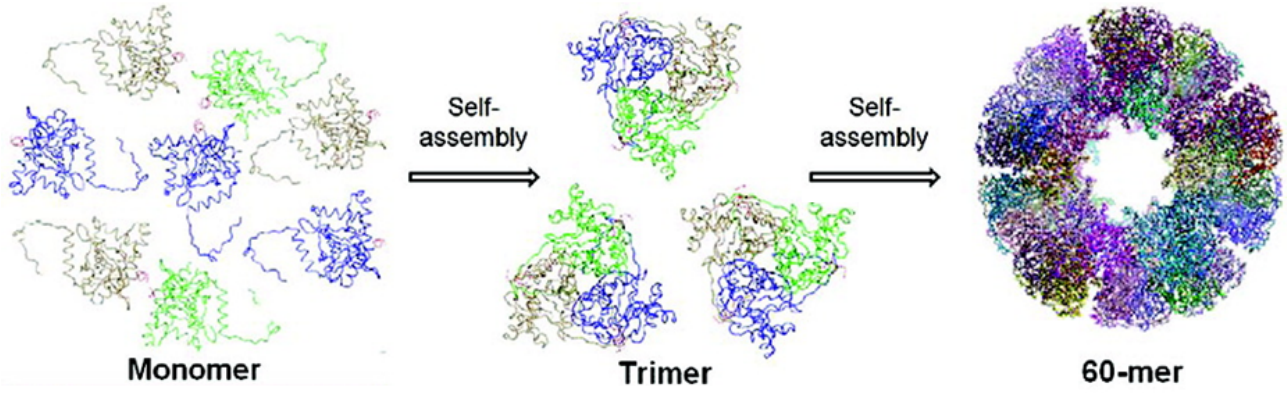

Figure 2. Structure of the E2 complex (from Peng et al., 2012).

Clearly, the organized E2 scaffold possesses a high capacity for accommodating proteins on its surface. As long as the C- termini are maintained to guide the formation of the 60-mer, the N-termini of the E2 subunits can be modified to display exogenous peptides (Domingo et al., 2001).

Introducing residues at the N- and C- termini or at the intratrimer interfaces may also generate E2 complexes with controlled assembly and altered pH sensitivity (Peng and Lim, 2011). For example, Peng and Lim have demonstrated the irreversible pH-triggered disassociation of E2 proteins containing N-terminal truncations (Peng and Lim, 2011).

Although sensitive to pH, E2 complexes are resilient to heat. *B. stearothermophilus* are viable at temperatures ranging from 30 to 75 °C (Peng and Lim, 2011), conferring E2 with stability up to 85 °C (Dalmau et al., 2008). Indeed, the robust intratrimer interactions impede attempts to isolate E2 monomers, despite the use of high concentrations of chaotropic salts and denaturants (Peng et al., 2012). Chaperonins are not required for *in vitro* renaturation of the 60-mer VLP from its constituent trimers (Lessard et al., 1998). The study of these complexes is also facilitated by the fact that E2 can be expressed in *E. coli*, bypassing the need for more expensive mammalian or baculovirus cell culture (Trovato et al., 2012).

E2's natural function as a structural scaffold, its high thermal stability, and its pH-dependent assembly have generated interest in the fields of engineering and vaccines. For example, De Berardinis and colleagues have successfully produced E2 scaffolds that display HIV antigens (Trovato et al., 2012). Mice immunized with these particles mounted strong, sustained antibody responses. Another recent effort involves designing pH-responsive complexes with enhanced endosomal escape properties (Peng and Lim, 2011). The ability to implement E2 fusion proteins as therapeutics will undoubtedly require intensive characterization of their properties. ESI-GEMMA could provide an insightful tool to monitor the composition, assembly, and disassembly of these complexes.

1.4 Aims of Study

This project utilized ESI-GEMMA technology to probe the composition and dynamics of vault protein complexes. The vault samples that we analyzed in this study have been extensively described by other techniques. This insight allowed us to adapt GEMMA specifically towards vault proteins and to improve the accuracy of molecular weight derivations. Additionally, we evaluated potential applications of ESI-GEMMA for the study of E2 protein complexes.

This study consisted of six specific aims:

1. Recalibrate GEMMA for the study of vault proteins
2. Elucidate structural changes to the vault upon transition to the gas phase
3. Quantify proteins encapsulated inside the vault
4. Evaluate retention of vault-encapsulated proteins over time
5. Investigate the effects of vault sample preparation on GEMMA spectra
6. Apply GEMMA towards the study of E2 protein complexes

Chapter 2

Methodology

2.1 GEMMA Instrumentation

ESI-GEMMA instrumentation (TSI Inc., St. Paul, MN) consists of three components as depicted in Figure 3: an electrospray ionization (ESI) unit, a differential mobility analyzer (DMA), and a condensation particle counter (CPC). Details regarding the instrumentation have been described elsewhere (Bacher et al., 2001; Kaufman et al., 1996; Kaufman 1998).

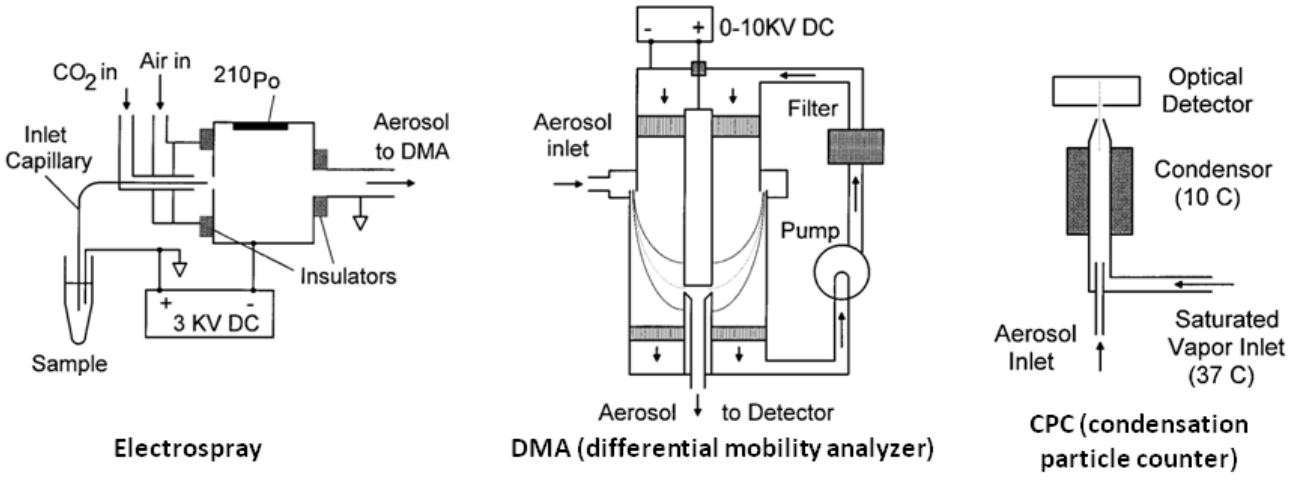

Figure 3. Schematic of GEMMA instrumentation (from Bacher et al., 2001).

ESI unit with neutralizing chamber (operated at atmospheric pressure and room temperature)

The sample solution enters the instrument through a fused silica capillary (24 cm long, 25 µm I.D. and 150 µm O.D.) that is connected to a high-voltage source. On the other end, the capillary

is ground to a conical shape at an angle of 75° to form the ESI spray tip. A thin Pt electrode immersed in the sample solution serves as the return connection to the positive side of the electrospray voltage supply.

At the entrance to the electrospray chamber, the spray tip is surrounded by filtered air (flow rate of 1 to 2 L/min) and a concentric flow of CO_2 (0.1 L/min) that stabilizes against corona discharge. Upon pressurizing the sample vial compartment, the liquid sample is aspirated through the capillary and enters the electrospray chamber. The shape of the emerging droplet can be adjusted by regulating the electrospray voltage. For these studies, measurements were performed in "cone jet" mode, with an operating voltage ranging from 1.5 to 2.5 kV and currents ranging from 200 to 300 nA.

This process produces multiply charged aerosol droplets that continuously decrease in size due to evaporation. The droplets are swept into the neutralizing chamber, where they encounter bipolar ions. These bipolar ions result from the reaction between gases present in the compartment and a ^{210}Po α ionization source (5 mCi, model P-2042 Nucleospot local air ionizer; NRD, Grand Island, NY) (Ebeling et al., 2000). The ESI-generated multiply charged analyte species are largely reduced to neutral charge upon exposure to the bipolar ions. At this stage, further evaporation has occurred as a consequence of the interactions between primary gas ions and multiply charged analyte molecules.

DMA unit (operated at atmospheric pressure and room temperature)

In the DMA, the analyte particles are separated by their electrophoretic mobility in air. The ions flow through a system of two coaxial cylindrical electrodes. The potential difference between the central electrode rod (connected to a negative power supply) and the outer electrode (grounded) gives rise to an electric field. Along with an orthogonal laminar flow of a sheath of

air (15-20 L/min), this applied radial electric field directs the ions into an exit flow. This process selects for singly charged ions. Negatively charged species are repelled by the inner electrode. Neutral particles exit the DMA along with the sheath air, which is recirculated after filtration. Particles of specific electrophoretic mobilities (EM) can be selected by adjusting the voltage difference between the two electrodes. In these studies, we sampled the 2 to 56 nm EMD range.

CPC unit (operated at atmospheric pressure and temperatures of 10 and 37 °C)

The CPC unit detects and counts the selected particles. The separated ions are exposed to saturated *n*-butanol vapor at 37 °C. Upon cooling to 10 °C, the vapor condenses around the ions. The *n*-butanol-coated particles increase in size to the extent that they can be measured by light scattering (in the micrometer range).

Data Generation

Spectra were generated using Aerosol Instrument Manager Software (TSI Inc.), which scans the DMA voltage and records data. These studies employed a 135 s scan (120 s of increasing voltage and 15 s to return to the original voltage). Ten consecutive scans over the entire EM range were complied to produce one GEMMA spectrum. No smoothing algorithm was applied.

2.2 GEMMA Spectra Interpretation

Counts (equivalent to abundance) of particles are displayed on the y-axis (Figure 4). The x-axis corresponds to the EMD (electrophoretic mobility diameter, in nm) of the particles. The EMD listed for each sample is the centroid of the peak of interest.

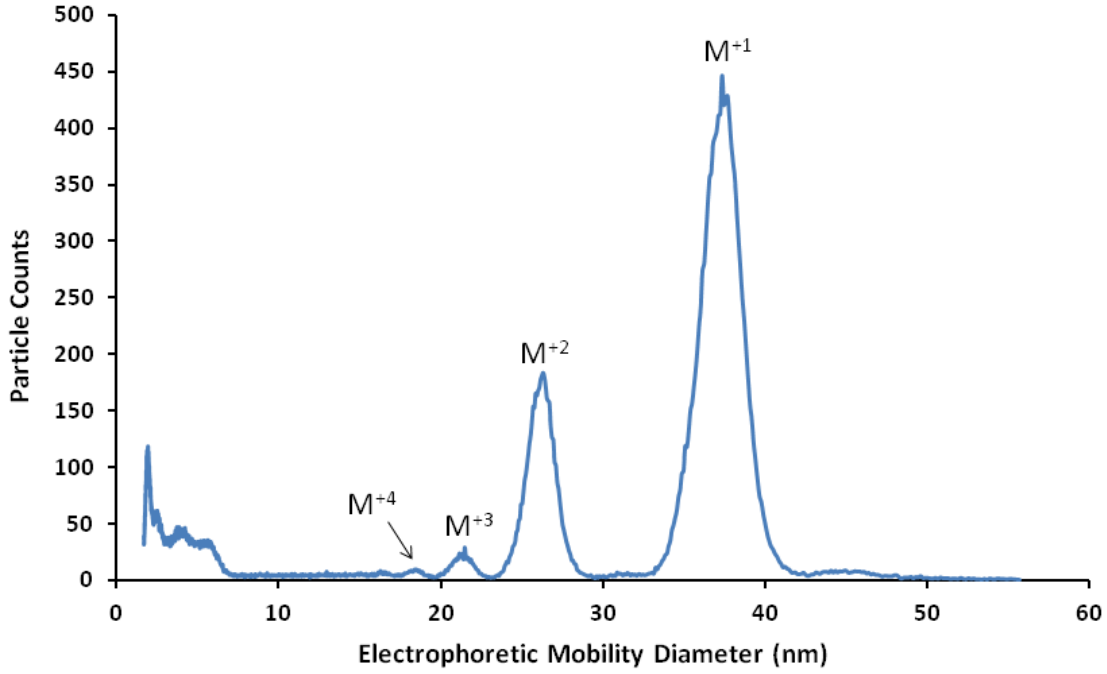

Figure 4. Sample GEMMA spectrum of a CP2 vault. The CP2 vault is a stable, ~7.6 MDa protein complex. The peak at 37.3 nm represents the singly charged ion of the intact vault. Peaks of lower abundance represent multiply charged ions.

2.3 Sample Preparation and Other Methods

Vault Sample Preparation:

Vaults were expressed and purified as described previously (Poderycki et al., 2006). Unless otherwise noted, non-volatile salts were removed from samples using Millipore Microcon centrifugal filter devices (YM-100). Vaults were exchanged into a 20 mM ammonium acetate, pH 7.4 buffer. Samples were typically run at a concentration of 50 μg/mL. Other desalting techniques included Thermo Scientific Slide-A-Lyzer Dialysis Cassettes (7K MWCO) and Millipore C18 ZipTips. The GEMMA instrumentation was operated under the conditions described by Kaddis et al. (Kaddis et al., 2007).

E2 Sample Preparation:

E2 samples were purified by M. Posner at the University of Bath, UK. Lyophilized samples were reconstituted in filtered water. E2 samples were desalted according to the same procedures used for vaults, except that filter devices with a lower molecular weight cut off (YM-10) were selected. In addition to Microcon centrigual filter devices, Thermo Scientific Slide-A-Lyzer Dialysis Cassettes and Bio-Rad Micro-BioSpin size exclusion columns (6K MWCO) were used.

SDS-PAGE Gels:

Samples were treated with NuPage LDS sample buffer and NuPage reducing agent and boiled before loading. 12% NuPage Bis-Tris gels were run in MES SDS running buffer for approximately 50 minutes at 200V, 125mA, and 100W.

LC-MS:

In-solution trypsin digestion

5 µg of vault protein was reduced in 2 mM freshly prepared dithiothreitol (DTT) for 30 min, at 60 ºC. The mixture was cooled to room temperature and the cysteine residues were alkylated in the dark, in 20% molar excess of iodoacetamide, for 45 min at 45 ºC. The reaction was quenched by adding DTT to a final concentration of 2 mM. Modified trypsin was added at an enzyme/substrate ratio of 1:50 (w/w) and proteolysis proceeded overnight with stirring at 37 ºC. Digested samples were dried in a SpeedVac and redissolved in 50 µL of 0.1% of formic acid.

Liquid chromatography

Chromatographic separation was achieved using Waters nanoACQUITY UPLC BEH C18 column (1.7 µm, 75 µm x 100 mm, 10K psi). The mobile phase (flow rate of 0.3 µL/min) comprised a gradient mixture of (A) 0.1% formic acid in water and (B) 0.1% formic acid in acetonitrile. A mobile phase with 97% concentration of (A) was applied for 10 min to desalt the

peptides. The concentration of (A) was decreased to 50% over a period of 100 min to elute the peptides. Mobile phase A was further decreased to 2% over a period of 2 minutes, then held at 2% for 10 minutes to wash off any remaining peptides. The concentration of (A) was then increased to 97% over a period of 15 min to equilibrate the column. The column temperature was set at 28 ºC.

Mass spectrometry analysis

Mass spectrometry of tryptic vault peptides was performed utilizing a Waters Xevo quadrupole time of flight (Q-TOF) mass spectrometer coupled directly to a Waters nanoACQUITY UPLC system. All analysis was performed using positive mode electrospray ionization (ESI). Liquid chromatography mass spectrometry (LC-MS) data was acquired by alternating low energy MS with elevated energy MS/MS (MS^E). In low energy MS mode, data was collected at a constant collision energy of 6 eV. In elevated energy MS/MS mode, the collision energy was ramped from 15 to 40 eV laboratory frame energy to collect the product ions of all precursors identified in the MS scan.

LC-MS and LC-MS/MS data analysis

The LC-MS and LC-MS/MS data were processed using ProteinLynx global server version 2.5 (Waters corporation). Proteins were identified using MS/MS peak lists of MS^E data-independent collision-induced fragmentation that were generated from the sequences of the vault and encapsulated proteins. Protein identifications were accepted with more than three fragment ions per peptide, seven fragment ions per protein, and one unique peptide per protein identified. Carbamidomethyl cysteine was set as a fixed modification and oxidized methionine was set as variable modification. Trypsin was specified as the proteolytic enzyme and up to two missed cleavages were allowed, with a false positive rate of 4%.

Chapter 3

Recalibration of GEMMA for the Study of Vault Proteins

Previously, the Allmaier and Loo laboratories reported the GEMMA EMD measurements of more than 40 protein complexes (Bacher et al., 2001; Kaddis et al., 2007). These samples ranged greatly in size, including the 93 kDa enolase dimer and the 4.6 MDa cowpea chlorotic mottle virus. These proteins served as molecular weight standards from which a relationship between EMD and mass could be derived.

The following equation models the particles as spheres in order to calculate volume from measured EMD, as reflected by the $(\pi/6)(EMD^3)$ term. Incorporation of a density term (δ) allows for conversion of volume to mass. Avogadro's number (N_0) is included to convert the mass of a single particle into molar mass.

$$MW = (\delta)(N_0)(\pi/6)(EMD^3) \quad (1)$$

As illustrated in Figure 5, Loo and colleagues plotted the molecular weights of the samples against the experimental EMD data (Kaddis et al., 2007). Fitting this data to equation 1 gave rise to an average particle density of 0.58 g/cm^3.

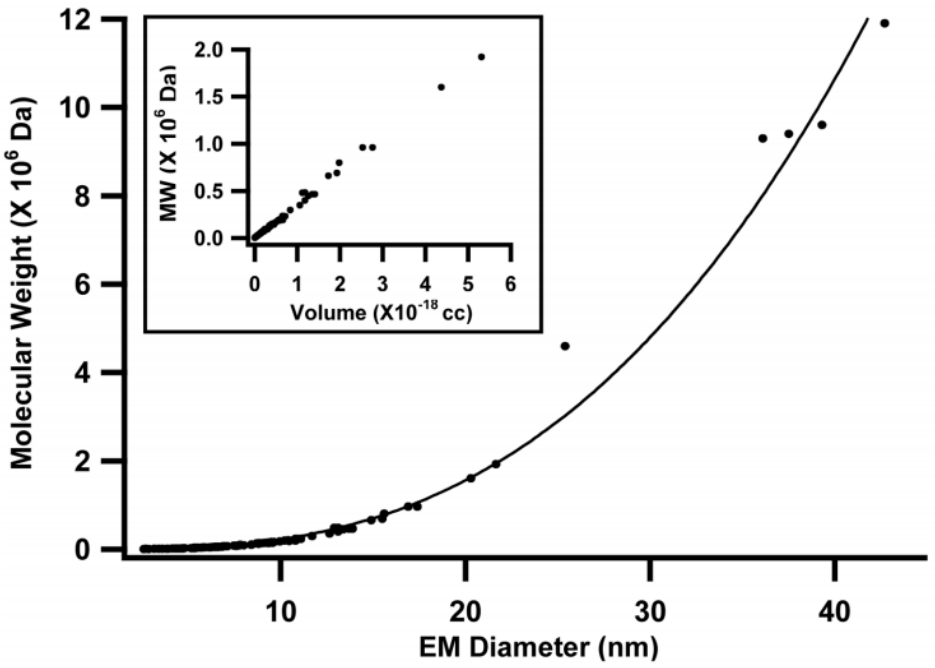

Figure 5. Correlation between electrophoretic mobility measured by ESI-GEMMA and molecular weight (from Kaddis et al., 2007). Data represents a range of proteins and noncovalent protein complexes.

This calibration was performed including four vault variants as high molecular weight standards. At the time of publication, the number of MVP subunits in the vault complex was accepted to be 96, based on a crystal structure solved by Eisenberg and colleagues (Anderson et al., 2007). However, a more recent, higher resolution crystal structure has determined this copy number to be 78 (Tanaka et al., 2009). The actual theoretical mass of vaults is therefore lower than the previously accepted mass that was used to calibrate GEMMA mass analysis. The formerly derived average density of 0.58 g/cm^3 thus stands to be corrected, as it is the crucial liaison between GEMMA EMD measurements and mass determination.

A more accurate calibration may be obtained by simply amending the vault data points in Figure 5 to reflect their true molecular weight. However, in this study, we propose that the accuracy of the curve may actually be improved upon omission of vaults standards. The

electrophoretic mobility of vaults deviates from expected values due to the vault's distinct shape and "hollowness" (discussed in Chapter 4), raising concerns over whether or not the complex accurately represents a typical large protein. Nevertheless, it is still possible to conceive a corrected curve that can be applied uniquely towards the investigation of vault proteins.

To recalibrate the EMD-mass relationship specifically for vault proteins, vault samples of various theoretical mass were measured by GEMMA (Table 1). Although elongated in shape, vault particles were modeled as spheres (equation 1) to simplify data interpretation. A density of 0.45 g/cm^3 was derived from the best-fit curve (Figure 6). This new density value provides an improved EMD-mass correlation that is specific to vaults. The value of 0.45 g/cm^3 accounts for the non-spherical shape and "hollowness" of the vault, allowing standard GEMMA techniques to be applied towards vaults. The ability to accurately convert EMD data to mass will prove useful in characterizing unknown samples and quantifying the amount of protein encapsulated inside vaults.

Table 1. Molecular weight and electrophoretic mobility diameters (EMD) of recombinant vaults analyzed in this study.

Vault[a]	MW of MVP monomer (kDa)[b]	MW of vault (MDa)[c]	EMD (nm), M^{+1}	EMD (nm), M^{+2}
B2	100.62	3.9[d]	29.1	20.7
CP2	96.81	7.6	37.3	26.3
CPZ	100.90	7.9	38.2	26.8
PVT-Z	106.76	8.3	39.3	27.6

[a]Protein tag sequences: MAGCGCPCGC GA (N-term, CP2 and CPZ); MARYRCCRSQ SRSRYYRQRQ RSRRRRRRSC QTRRRAMRCC RPRYRPRCRR H (N-term, PVT-Z); FNMQQQR RFYEALHDPN LNEEQRNAKI KSIRDD (C-term, PVT-Z and CPZ); EEEEEEE (N-term, B2).
[b]As calculated based on sequence.
[c]As determined by multiplying molecular weight of the MVP monomer by 78.
[d]B2 sample consists of half-vaults.

The vaults analyzed in Table 1 migrated at the expected size for intact vaults, except for the B2 sample. The MVP subunits of B2 vaults contain N-terminal glutamic acid-rich tags that create charge repulsion at the waist. Interestingly, all of these particles migrated at the size expected for half-vaults, and no particles representing the full vault were observed.

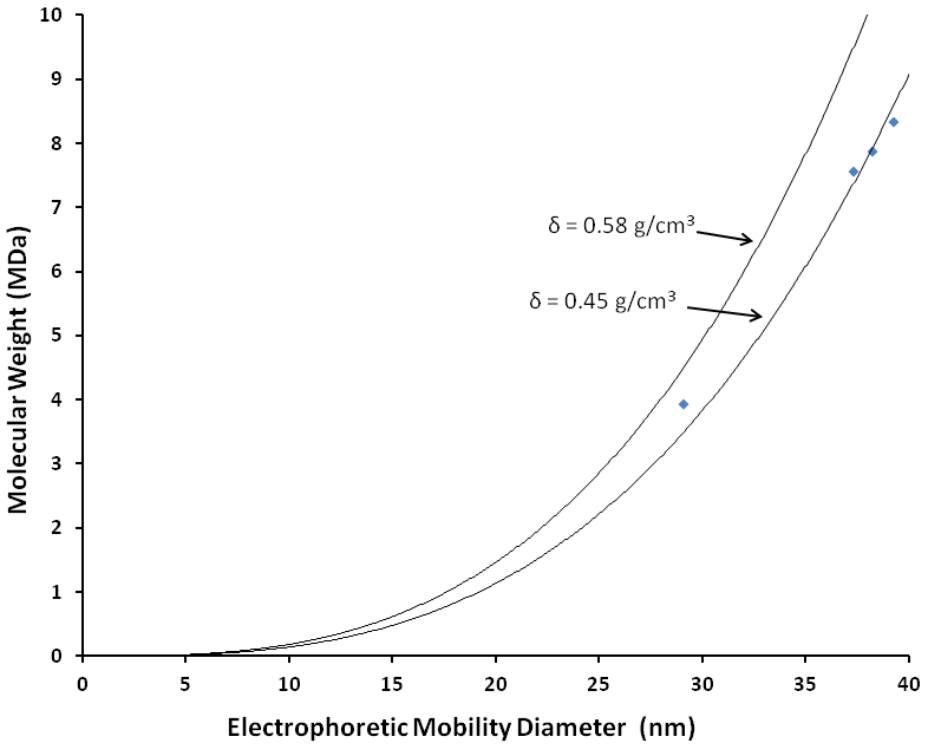

Figure 6. Correlation between electrophoretic mobility measured by ESI-GEMMA and molecular weight for vault proteins analyzed in this study.

As a measure of reproducibility, vault GEMMA measurements obtained in this study were compared to those reported by Kaddis et al. (Table 2 and Figure 7) (Kaddis et al., 2007). The data obtained by Kaddis et al. was amended to reflect the true MVP copy number of 78.

Table 2. Molecular weight and electrophoretic mobility diameters (EMD) of recombinant vaults measured by Kaddis et al. (Kaddis et al., 2007).

Vault[a]	MW of MVP monomer (kDa)[b]	MW of vault (MDa)[c]	EMD (nm)
NT	97	7.6	36.1
CP2	98	7.6	37.5
HT7	100	7.8	39.3
GL	124	9.7	42.7

[a]Protein tag sequences: none (NT); MAGCGCPCGC GA (N-term, CP2); MGSSHHHHHH SSGLVPRGSH MASMTGGQQP W (N-term, HT7); MSKGEELFTG VVPILVELDG DVNGHKFSVS GEGEGDATYG KLTLKFICTT GKLPVPWPTL VTTFTYGVQC FSRYPDHMKQ HDFFKSAMPE GYVQERTIFF KDDGNYKTRA EVKFEGDTLV NRIELKGIDF KEDGNILGHK LEYNYNSHNV YIMADKQKNG IKVNFKIRHN IEDGSVQLAD HYQQNTPIGD GPVLLPDNHY LSTQSALSKD PNEKRDHMVL LEFVTAAGIT HGMDELYKP (N-term, GL).
[b]As determined by MALDI-TOF-MS.
[c]As determined by multiplying molecular weight of the MVP monomer by 78.

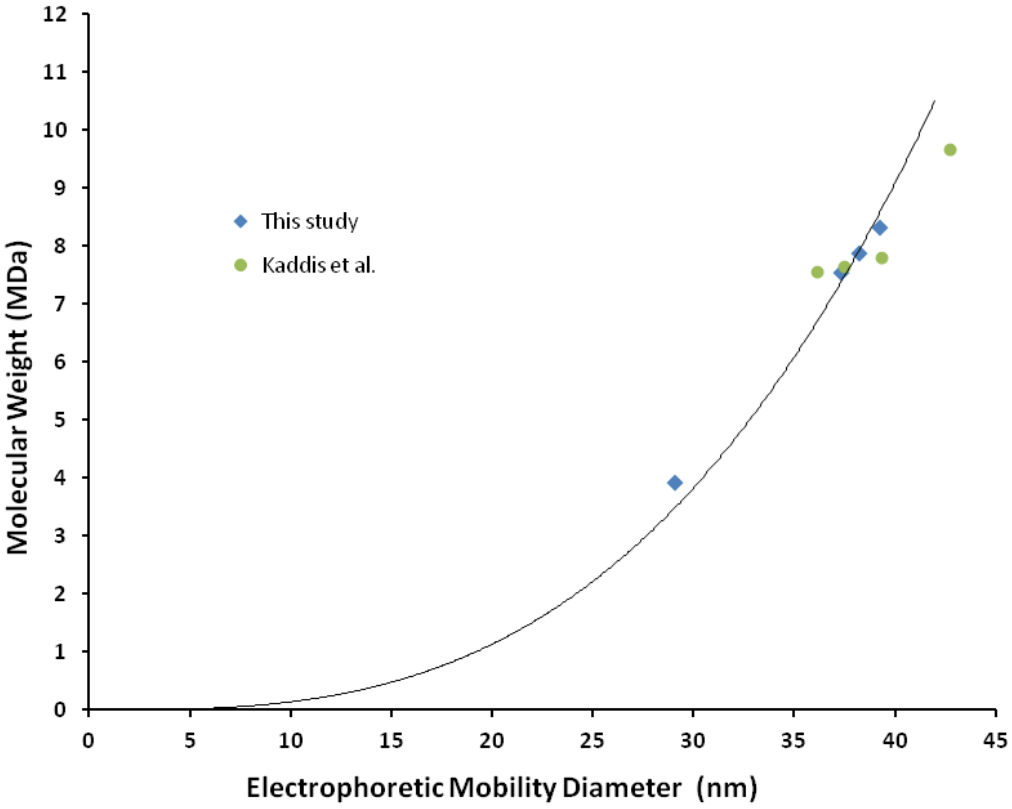

Figure 7. Comparison of data obtained in this study to values reported by Kaddis et al. (Kaddis et al., 2007).
Curve represents new MW/EMD correlation determined in this study, reflecting an average density of 0.45 g/cm^3.

Chapter 4

Structural Changes to the Vault upon Transition to the Gas Phase

Previous studies have indicated that vault complexes may collapse and increase in density upon transition to the gas phase (Kaddis et al., 2007). With regards to GEMMA, this phenomenon could give rise to EMD values that underestimate the actual size of the vault. If this increased density were not accounted for during conversion of EMD to molecular weight (equation 1), the calculated molecular weight would underestimate the actual mass of the sample. This discrepancy impedes the identification of vault complexes of unknown mass. Accuracy is especially required to quantify the amount of drug encapsulated inside vault samples. We probed this issue by investigating the densities of vaults under three different conditions:

1) The theoretical density of an empty vault complex in solution, calculated from crystal structure dimensions
2) The average density of empty vault samples in the gas phase (measurements obtained from GEMMA)
3) The average density of protein-filled vault samples in the gas phase (measurements obtained from GEMMA)

The crystal structure of the vault (Tanaka et al., 2009) indicates dimensions of approximately 400 x 400 x 740 nm. By modeling the vault as an ellipsoid, a particle volume of 5.5989×10^{-17} cm^3 is obtained. Given that the mass of a single vault particle (MVP only) is 1.2408×10^{-17} g, the density of a vault particle in aqueous solution corresponds to 0.22 g/cm^3 (Table 3).

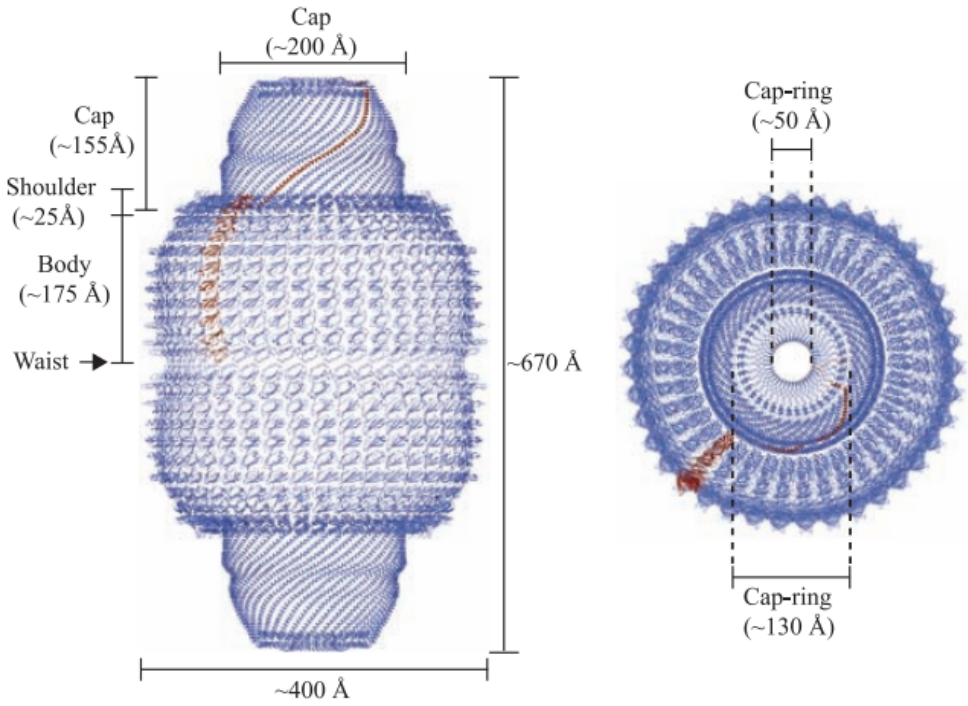

Figure 8. Crystal structure of vault (from Tanaka et al., 2009).

As established by the data presented in Figure 6, vaults subjected to ESI-GEMMA conditions exhibit an average density 0.45 g/cm^3 (Table 3).

Should the MVP shell be rigid, the addition of proteins inside the vault would not be expected to alter the dimensions of the entire complex. However, CP2 vaults that encapsulate CCL21-INT migrate at a higher EMD than empty CP2 vaults, indicating a larger EMD (37.3 vs. 38.9 nm; refer to Figure 10 for data). The packaged CCL21-INT seemingly protects against the compaction that the gas-phase empty vault may undergo. If the mass of the packaged CCL21-INT is excluded in calculations, the filled CP2 vault has a reduced density of 0.41 g/cm^3, owing to its larger size (Table 3). The mass of each encapsulated CCL21-INT is 33.3 kDa (to be discussed in Chapter 5).

Table 3. Density of the vault under various conditions. The average density of particles analyzed by ESI-GEMMA is 0.58 g/cm^3 (Kaddis 2007)

Conditions	Vault density (g/cm^3)
Aqueous	0.22
Gas phase (GEMMA)	0.45
CCL21-INT packaged inside	0.41

We believe that the difference in density between the empty and CCL21-INT-filled vaults is significant and not due to stochastic fluctuations in measurement. Although there is some variance amongst biological and technical replicates, these discrepancies are not large enough to account for the difference in EMD which we observe between the CP2 and CP2-CCL21-INT vaults.

For example, we found the EMD of CP2 vaults to be 37.3 nm, while Kaddis et al. reported a value of 37.5 nm (Kaddis et al., 2007). This 0.2 nm difference appears insignificant when compared to the 1.6 nm difference in the EMD values of the CP2 (37.3 nm) and CP2-CCL21-INT (38.9 nm) vaults. Moreover, as discussed in Chapter 6, variable parameters such as vault protein concentration, salt concentration, and sample age have negligible effect on measured EMD. The CP2-CCL21-INT samples were also analyzed by LC-MS (Chapter 5), with results that support our GEMMA data.

Interestingly, the ability of encapsulated proteins to affect the vault's dimensions appears limited to the gas phase. Vaults packaged with only TEP1 or both vPARP and TEP1 exhibit EMD values of 38.0 and 40.1 nm, respectively, when analyzed by GEMMA (Kaddis et al., 2007). As discussed in Chapter 5, these numbers signify that one copy of the 294 kDa TEP1 is inside the TEP1-encapsulating vault (Table 4). In addition to one copy of TEP1, the TEP1/vPARP vault contains six copies of the 195 kDa vPARP.

Cryo-EM studies report the native dimensions of both TEP1 and TEP1/vPARP vault samples to be the same (72.5 x 41.0 nm) despite their different amount of contents (Mikyas et al., 2004). We therefore dismiss the possibility that encapsulated proteins cause the vault to bulge, leading to increased dimensions in both the solution and gas phases. Rather, we propose that packaged proteins prevent collapse of the vault under non-native (i.e. gas-phase) conditions (Figure 9). This hypothesis reconciles the observations that both empty and packaged vaults are of identical size in solution, yet exhibit different size (EMD values) in the gas phase.

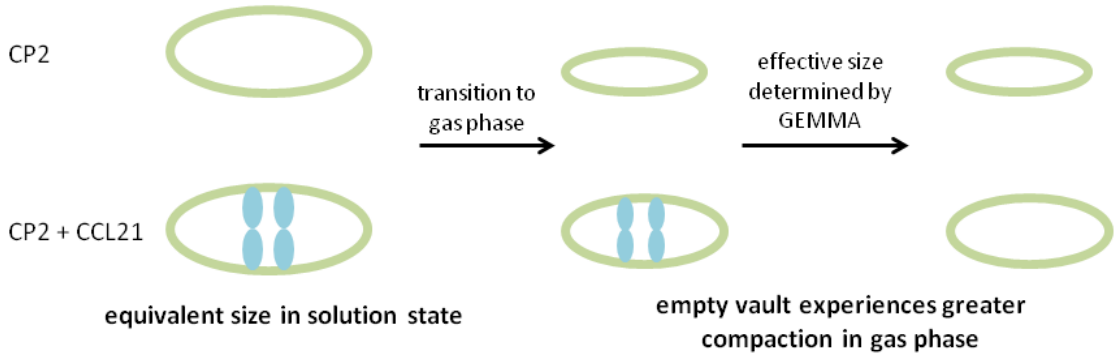

Figure 9. Protein-encapsulating vaults experience less compaction upon transition to the gas phase.

Chapter 5

Quantitation of Proteins Encapsulated inside the Vault

In designing a vault drug delivery platform, therapeutic proteins can be fused to the INT domain of vPARP. This domain interacts with MVP subunits on the inner shell of the vault, facilitating encapsulation of the fusion protein (Kickhoefer et al., 2005). As previously mentioned, loading of vaults with proteins such as CCL21-INT prevents complete contraction of the vault upon transition to the gas phase, leading to an EMD increase that is detectable by GEMMA. Figure 10 demonstrates that the empty CP2 vault has an EMD of 37.3 nm whereas the vault loaded with CCL21-INT exhibits an EMD of 38.9. As discussed in Chapter 4, we believe that this shift in EMD is significant.

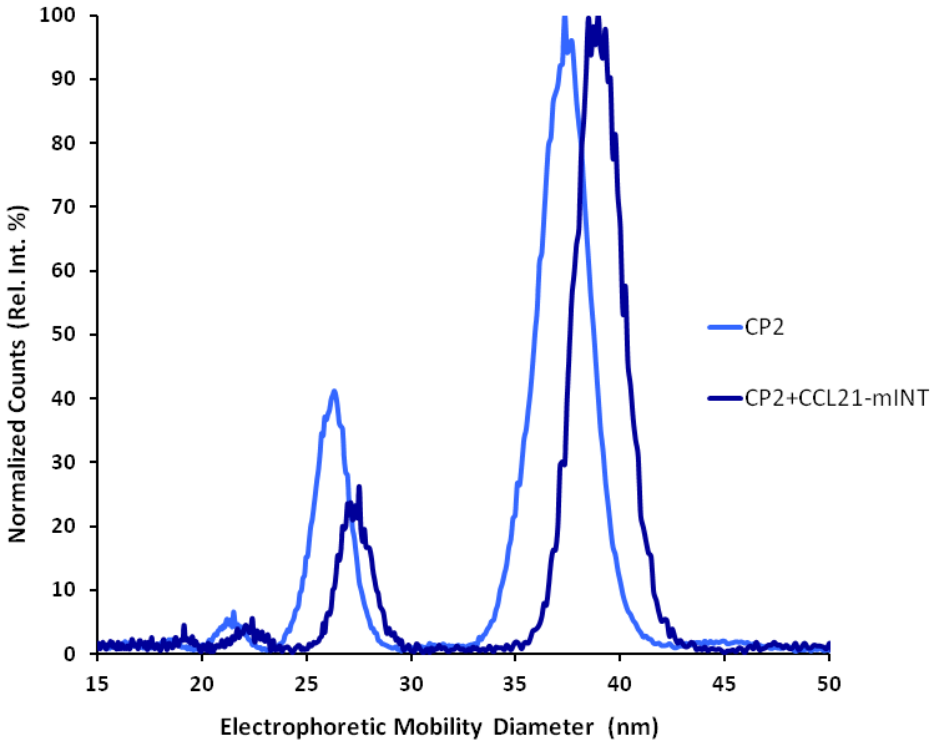

Figure 10. GEMMA spectrum of CP2 vaults packaged with CCL21-INT compared to spectrum of empty vaults. Centroid of CP2 peak: 37.3 nm. Centroid of CP2+CCL21-INT peak: 38.9 nm.

The mass contribution of CCL21-INT to the filled vaults can be obtained by calculating the shift in EMD upon CCL21-INT encapsulation. Using equation 1 with the vault-specific GEMMA density of 0.45 g/cm^3, this difference in size can be converted to difference in mass. Taking into account the molecular weight of the CCL21-INT fusion protein (33.3 kDa), the number of proteins encapsulated inside each vault can be assessed.

This GEMMA-based method for quantifying encapsulated proteins was compared to data acquired by LC-MS. In addition, we assayed CP2 vault preparations that had been incubated with varying amounts of CCL21-INT. Testing these samples allowed us to ascertain that our technique was sensitive enough to detect small changes in the quantity of encapsulated CCL21. The results are summarized in Figure 11.

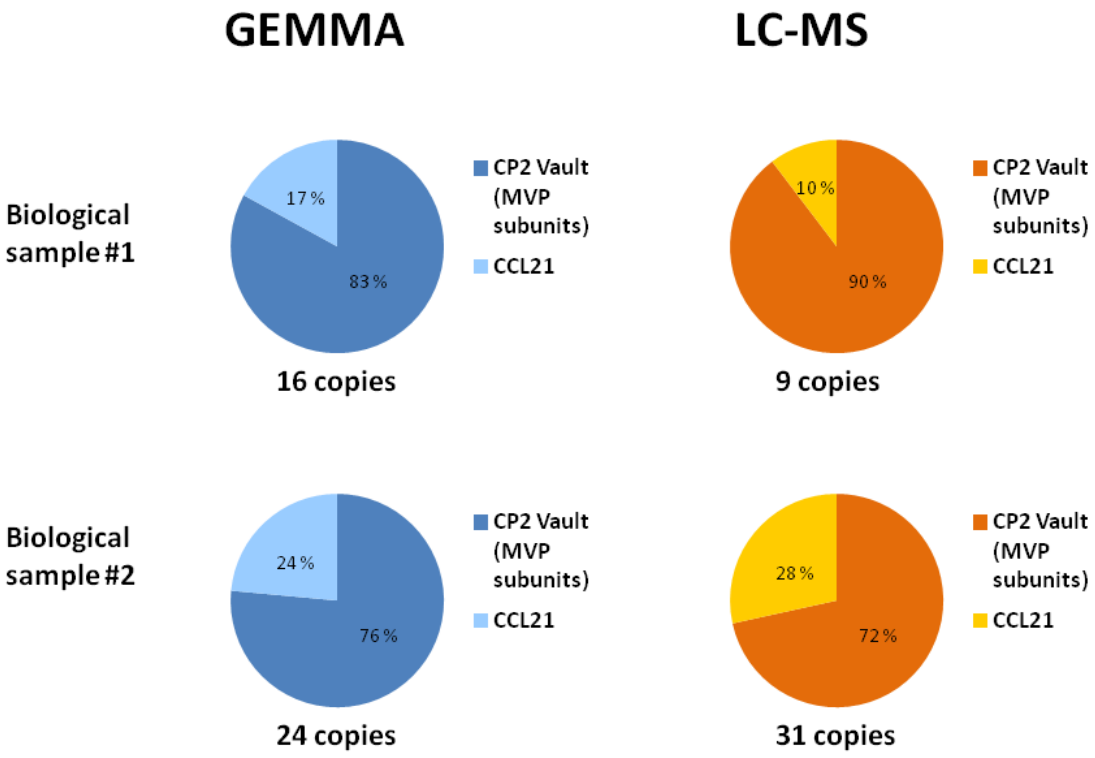

Figure 11. Vault encapsulation quantified by GEMMA and LC-MS. Biological sample #1 was incubated with a 3:1 molar ratio of CCL21-INT:CP2, whereas sample #2 was prepared with a 8:1 molar ratio of CCL21-INT:CP2. Percentages reflect molar ratios.

After performing LC-MS to verify that GEMMA can accurately quantify encapsulated CCL21, we then examined the GEMMA data obtained by Kaddis et al. (Kaddis et al., 2007). These samples were packaged with TEP1 and vPARP, the endogenous interacting proteins of the vault complex. The copy numbers of TEP1 and vPARP inside the vaults were calculated by the method used for CCL21.

Table 4. Molecular weight and electrophoretic mobility diameters (EMD) of recombinant vaults prepared with encapsulated proteins measured by Kaddis et al. (Kaddis et al., 2007).

Vault[a]	MW of MVP monomer (kDa)[b]	MW of vault (MDa)[c]	EMD (nm)	Copies TEP1	Copies vPARP
vMVP/TEP1	99	7.79	38.0	<1	0
vMVP/vPARP/TEP1	99	9.15	40.1	1	6

[a]Peptide tag sequences: MGYTDIEMNRLGKP (v, vsvg-MVP); MDYKDDDDKV NASR (vPARP FLAG tag); MIANVNIAQE QKLISEEDLA QEQKLISEED LAQQSGGGLD (TEP1 double myc sequence).
[b]As determined by MALDI-TOF-MS.
[c]As determined by using the EMD data, eq 1, and the vault-specific GEMMA density of 0.45 g/cm^3.

The empty 78-subunit vMVP vault is expected to have a molecular weight of 7.72 MDa. Due to encapsulation of TEP1 alone or a combination of TEP1 and vPARP, the samples represented in Table 4 exhibit increased masses of 7.79 and 9.15 MDa, respectively. We calculated that less than one copy of the 294 kDa TEP1 is retained inside each vMVP/TEP1 vault. Assuming that one copy of TEP1 is present inside the vMVP/vPARP/TEP1 vaults, we found that the remaining mass corresponds to 6 copies of the 195 kDa vPARP.

Chapter 6

Retention of Vault-Encapsulated Proteins over Time

As discussed in Chapter 5, our GEMMA results verified that both INT-fused proteins (CCL21) and proteins that associate with the endogenous vault (TEP1 and vPARP) can be successfully incorporated inside the vault. However, vaults readily split at the waist during half-vault exchange (Yang et al., 2010). The feasibility of proposed vault-based drug delivery must consider whether the INT-MVP interaction is strong enough to prevent INT-fused proteins from diffusing out of the vault during half-vault exchange.

To probe the strength of the interaction between INT fusion proteins and the vault, we monitored the loss of encapsulated fusion proteins over time. Since packaging of CP2 vaults with CCL21-INT increases the EMD of the vault particle (Figure 10), loss of these encapsulated proteins should confer an EMD closer to that of the empty vault. Remarkably, after storage at 4 °C for 123 days after the initial measurement, the EMD values of both the empty vault and the CCL21-packaged vault were unchanged (37.8 and 38.4 nm, respectively; see Figure 12). Despite the marked decrease in particle abundance, these results indicate no loss of MVP subunits or of encapsulated CCL21-INT fusion proteins over time.

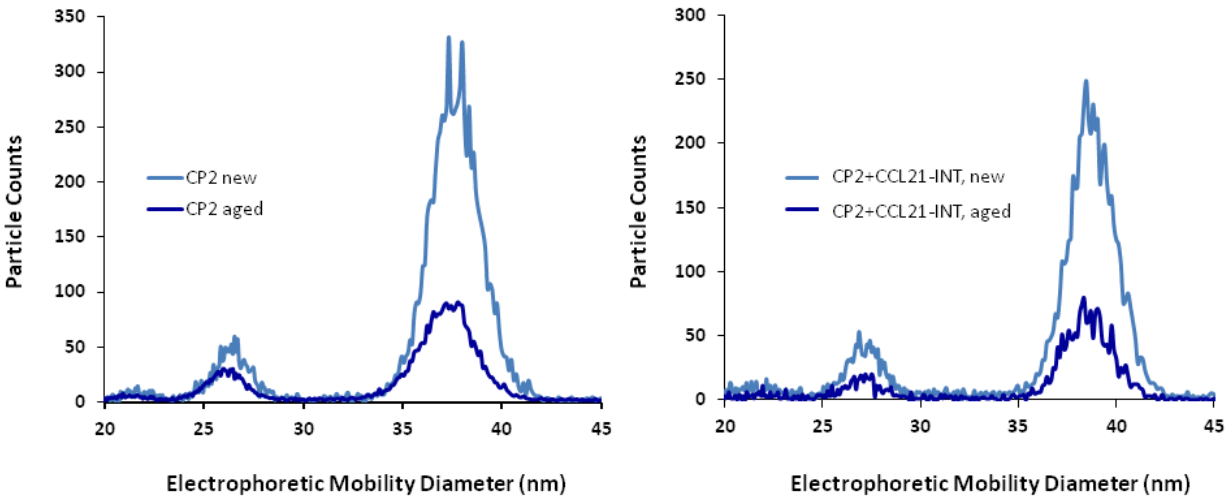

Figure 12. GEMMA spectra of freshly purified and aged vault samples. Left: empty CP2 vaults. Centroid of both peaks is 37.8 nm. Right: vaults packaged with CCL21-INT. Centroid of both peaks is 38.4 nm. Aged vaults were stored at 4 °C for 123 days.

Chapter 7

Effects of Vault Sample Preparation on GEMMA Spectra

The CPZ vault contains a 4.1 kDa, 33 amino acid Z domain at the C-terminus to facilitate antibody binding for receptor-mediated targeting (Kickhoefer et al., 2008). Unfortunately, these vaults proved difficult to assay by GEMMA. Samples prepared by the typical vault desalting protocol (Microcon YM100 desalting into 20 mM ammonium acetate, pH 7.4) displayed no signal above background when tested by GEMMA, except for one instance recorded in Table 1 and Figure 6.

Investigation via SDS-PAGE revealed that Microcon desalting led to significant loss of sample. In Figure 13, the amount of desalted (DS) CPZ sample loaded was 2.3 μg, if assuming zero loss of sample during desalting. However, the weak intensity of the band as compared to the non-desalted sample (neat) indicates that the majority of the CPZ protein was lost during this processing. CPZ vaults packaged with CCL21-INT and Ova-INT exhibited the same trend.

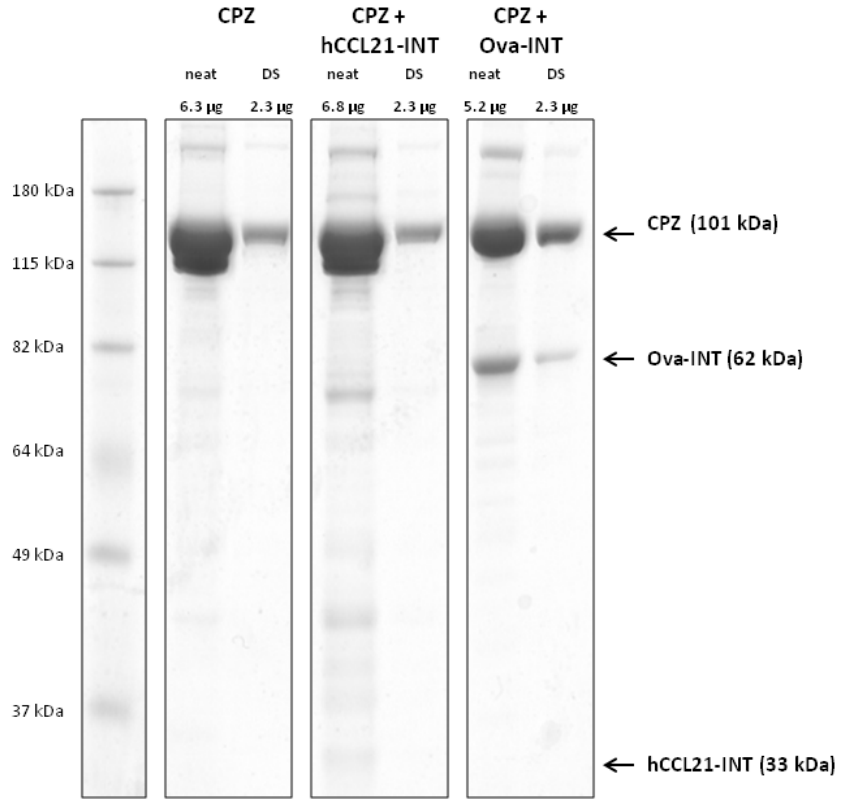

Figure 13. SDS-PAGE analysis of untreated (neat) and desalted (DS) CPZ vault samples. µg amounts of protein reflect the theoretical amount loaded (assuming no protein loss during desalting).

ZipTip treatment was also attempted and the prepared samples were characterized by SDS-PAGE (Figure 14). During this procedure, the protein present in a sample aliquot is extracted from solution by aspirating with a pipette tip packed with chromatography media. The ZipTip is washed and the desalted protein is eluted with an 80% acetonitrile, 0.1% trifluoroacetic acid solution. Had all protein been successfully aspirated from solution and retained on the ZipTip, the remaining ZipTip treated solution ("after ZT") should contain a minimal amount of protein. However, the abundance of protein in the "after ZT" sample and the absence of protein in the ZT elution (ZT) indicate that this approach was not successful in initially capturing the sample on the chromatography media. Dialysis desalting was attempted, yet these samples also exhibited no signal on GEMMA.

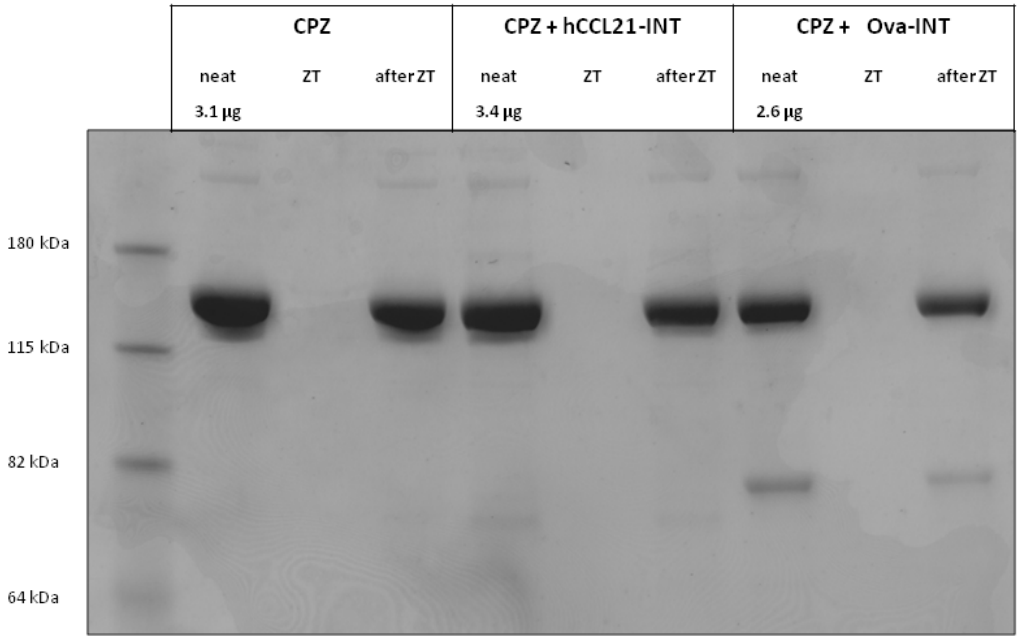

Figure 14. SDS-PAGE analysis of untreated (neat) and ZipTip treated (ZT) CPZ vault samples. µg amounts of ZT protein reflect the theoretical amount loaded (assuming complete efficiency and zero loss during ZipTip processing).

Since we were unable to perform desalting with high sample recovery, we examined the possibility of skipping the desalting step and diluting directly into the GEMMA running buffer (20 mM ammonium acetate). Vault samples are obtained in a variety of buffers, such as 20 mM MES, sterile PBS, or sterile saline. The protein concentration typically exceeds 5000 µg/mL; dilution to 50 µg/mL for GEMMA equates to a 100 fold or higher dilution. For the 20 mM MES buffer, the final salt concentration would be reduced to 200 µM or less.

To establish how the presence of non-volatile salts of micromolar concentration would impact GEMMA, CP2 vaults were assayed. As displayed in Figure 15, bypassing the desalting process resulted in a substantial increase in signal. The difference in counts between non-desalted and desalted CP2 vaults (both of which were initially at equivalent concentrations) signifies that 80% of CP2 protein was lost during desalting. However, the non-desalted sample

exhibits increased noise in the low EMD range, most likely due to low molecular weight contaminants that would have otherwise been excluded during desalting.

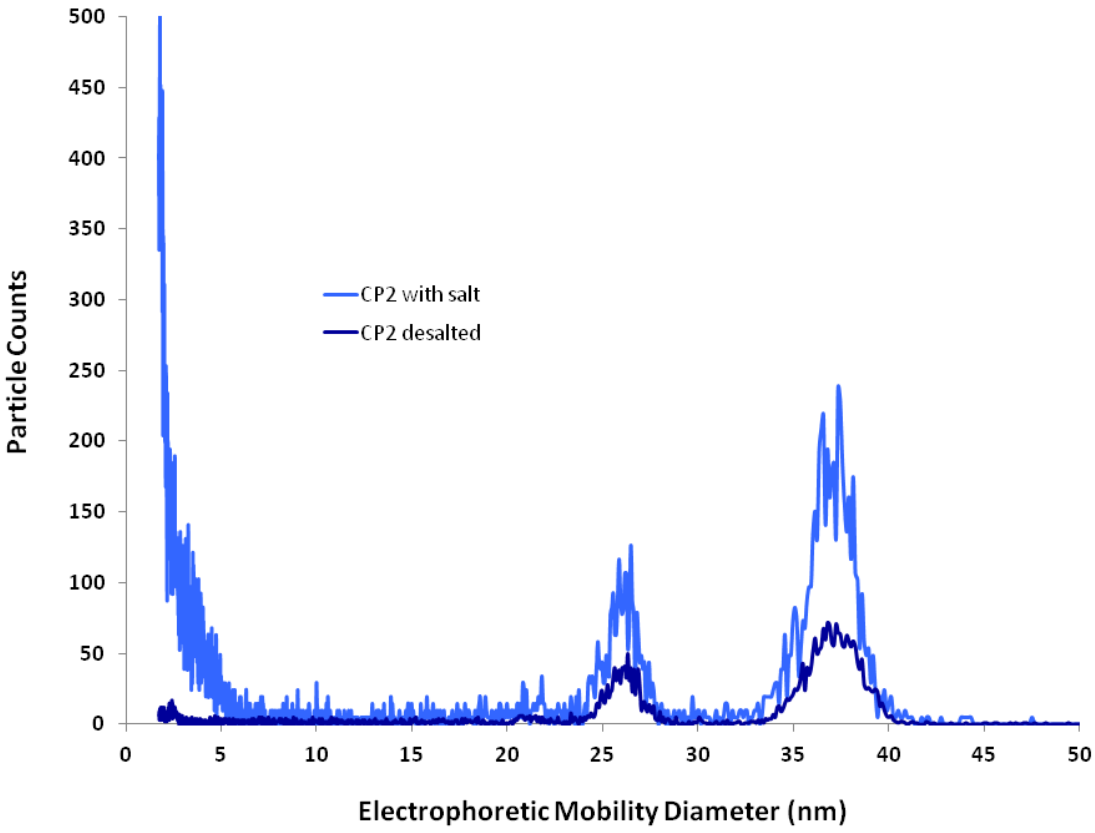

Figure 15. GEMMA spectra of untreated and column desalted CP2 vaults. Samples were prepared from equivalent starting concentrations.

Given our success with the CP2 vaults, we assayed CPZ vaults using the same approach. Despite skipping the desalting step and diluting the vaults into various buffers (ammonium acetate 20 – 100 mM, pH 7, 7.4, and 8), we were unable to obtain a spectrum for CPZ samples.

Chapter 8

GEMMA for the Study of E2 Proteins

Similar to the CPZ vault sample, we encountered difficulties in our attempts to evaluate E2 proteins by GEMMA. Samples prepared by the typical desalting protocol (Microcon YM-10 column desalting into 20 mM ammonium acetate, pH 7.4) displayed no signal above background when analyzed by GEMMA. The use of various buffers (ammonium acetate 20 – 100 mM, pH 7, 7.4, and 8) proved unsuccessful in resolving this predicament. Unlike the CPZ vaults, which were weakly recovered from the desalting columns, the E2 protein appears to have been completely lost (Figure 16). Purifications 1 and 2 of E2 both yielded the same result.

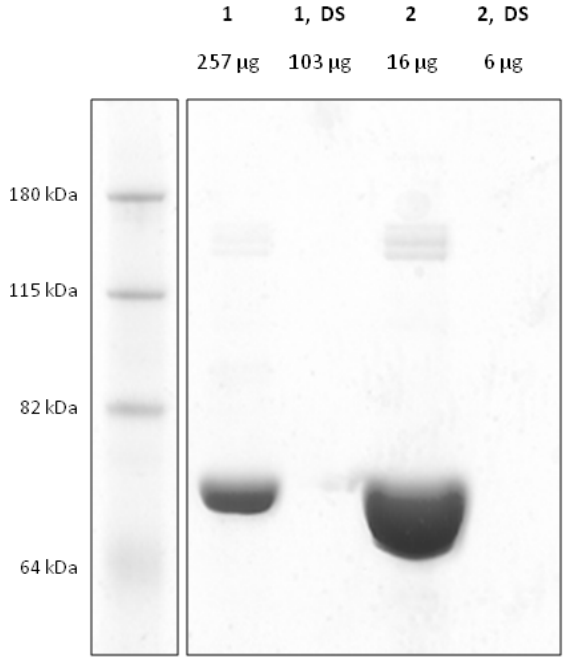

Figure 16. SDS-PAGE analysis of untreated and column desalted (DS) E2 samples. µg amounts of protein reflect the theoretical amount of lyophilized sample loaded (assuming no protein loss during desalting). Mass of lyophilized sample may not directly correlate to mass of E2; depending on purification method, varying amounts of salt also contribute to lyophilized sample mass. Two different preparations, 1 and 2, of E2 were tested.

We attempted dialysis and size exclusion chromatography (SEC), yet these samples also exhibited no signal on GEMMA. SDS-PAGE revealed that SEC, like Microcon column treatment, was unsuccessful in recovering desalted protein (Figure 17). Interestingly, the denaturing conditions of gel electrophoresis were not sufficient to disrupt E2 intratrimer interactions. Titration with increased amounts of reducing agent had no effect on trimer stability. This stable trimer represents the most abundant species in both Figures 16 and 17. A band migrating at the expected range for the hexamer is also observed.

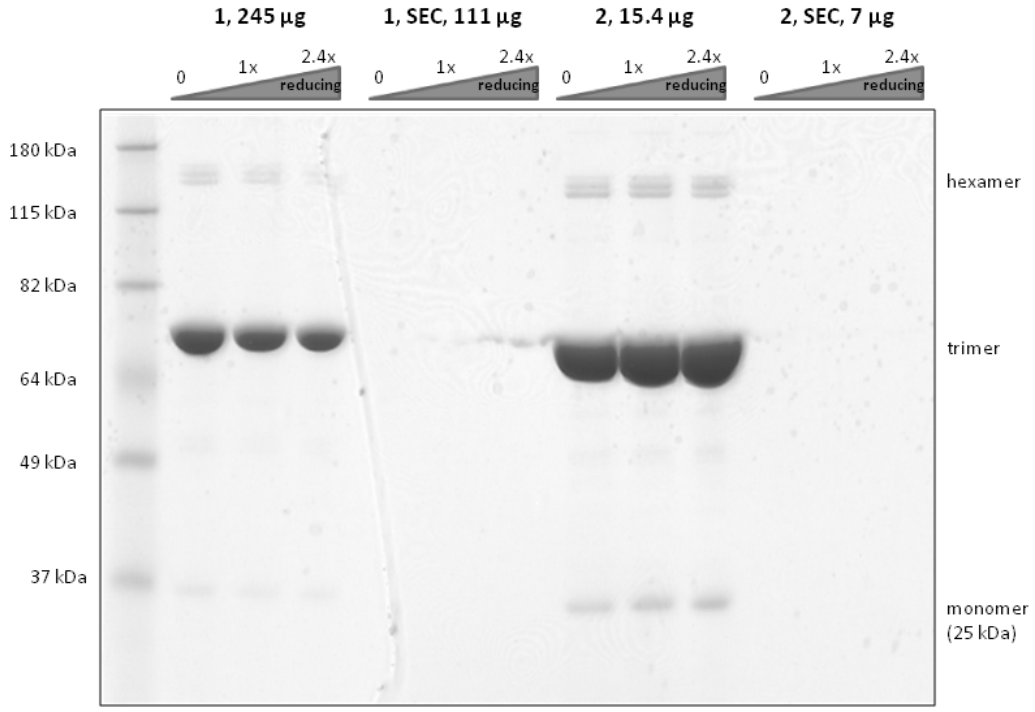

Figure 17. SDS-PAGE analysis of untreated and size-exclusion chromatography (SEC)-prepared E2 samples. µg amounts of protein reflect the theoretical amount of lyophilized sample loaded (assuming no protein loss during desalting). Mass of lyophilized sample and may not directly correlate to mass of E2 (depending on purification method, varying amounts of salt also contribute to lyophilized sample mass). Two different preparations, 1 and 2, of E2 were tested.

The stability of the E2 trimer led us to question whether we could isolate and observe this species via GEMMA. We therefore prepared samples in a variety of denaturing conditions (reducing agent and 0.1% acetic acid) in conjunction with the desalting methods described above. Although desalting previously caused loss of the 60-mer E2 complex (Figures 16 and 17), we reasoned that desalting with denaturing solvents would disrupt interactions amongst subunits, preventing aggregation and increasing recovery. Nevertheless, we were unable to obtain GEMMA spectra for these denatured samples. Unlike the vault preparations, the protein/salt ratio of the E2 samples was quite low. Skipping the desalting step was not feasible and led to high background noise that obscured measurement of the E2 complex.

Chapter 9

Discussion

Development of an Improved Method to Derive Mass from the GEMMA Spectra of Vault Proteins

In an endeavor to improve the accuracy of mass detection techniques for large protein complexes, this project adapted ESI-GEMMA towards the study of vaults. Developing a GEMMA-based method for vault complexes is attractive for numerous reasons. First, vaults bear promise as drug delivery agents and characterization of these nanocapsules will be central to advancing their pharmaceutical development (Kickhoefer et al., 2005; Kickhoefer et al., 2008; Han et al., 2011). Second, recombinant vaults are approximately 9 MDa, and complexes of this size remain largely unexplored by GEMMA or other mass spectroscopic techniques. Third, vault MVP subunits can be stably modified at the N- and C- termini, generating a variety of standards of different theoretical mass. Lastly, vault structure and subunit composition have been well-studied by other techniques such as crystallography and electron microscopy (Tanaka et al., 2009; Kickhoefer et al, 2005). Data from these alternative techniques can be incorporated into the establishment of a GEMMA-based method to accurately measure the mass of vault complexes.

A range of vault samples was tested and EMD values were plotted with their corresponding theoretical masses (Figure 6). Using the method described by Kaddis et al., we modeled the particles as spheres and calculated an average particle density of 0.45 g/cm^3 (Kaddis et al., 2007). This number differs from the previously reported value of 0.58 g/cm^3 due to improved crystal structure data that allowed us to use 78 as the copy number of the vault MVP subunits instead of

96 (Tanaka et al., 2009). Since we were interested in developing a highly accurate correlation for vault complexes only, we excluded other proteins from our calculations. Furthermore, as evinced by the notably low density of vaults in aqueous media (approximately 0.22 g/cm^3) compared to the higher density of the average GEMMA-analyzed protein (0.58 g/cm^3), using vaults as universal weight standards should be approached with caution, as they do not represent the typical protein.

Vault Structure Impacts Ability to Predict Mass from EMD

To evaluate reproducibility of our GEMMA measurements, we re-plotted the vault EMD values and mass reported by Kaddis et al. alongside our own (Table 2 and Figure 7) (Kaddis et al., 2007). Although the molecular weight-EMD correlation derived in this study models the general trend of the samples tested by Kaddis et al., some discrepancies do occur. In one instance (NT vault) our model underestimates the mass of the complex. In other cases, (i.e., GL vault), our model overestimates the mass.

Based on the stability of the vault over time (refer to Chapter 6), we can rule out differences in sample age as a contributor to the discrepancies in GEMMA data. Disparities in sample preparation are equally unlikely to contribute. As discussed in Chapter 7, the vault concentration, salt concentration, and pH of the sample have negligible effect on GEMMA spectra.

Rather, we suspect that the unique identity of the modifications added at the vault N- and C- termini account for the variability in EMD. For example, more ordered N-terminal amino acid tags would be expected to stabilize the vault at the waist. This reinforcement may oppose the collapse and decrease in size typically experienced by the vault upon transition to the gas phase. Conversely, more disordered tags may enhance the propensity of the vault to contract.

The EMD values of NT (no tag) and CP2 (cysteine-rich tag) vaults support this theory. The CP2 vaults migrate at an EMD that is 1.4 nm greater than the EMD of NT vaults, despite just a 1 kDa difference in the mass of the MVP subunits (Table 2). Indeed, the CP2 tag has been shown to stabilize the vault structure, most likely through inter-subunit disulfide bonds at the waist (Mikyas et al., 2004).

The lower than expected EMD of the B2 samples may also be explained by this hypothesis (Table 1 and Figure 7). These complexes exist in the half-vault conformation due to an N-terminal glutamic acid-rich tag that creates charge repulsion at the waist. A half-vault may be more prone to shrinking than an intact vault due to lack of stabilizing interactions at the waist region. Similar to the TEP1 and TEP1/vPARP-encapsulating vaults discussed in Chapter 4, solution-state B2 vaults display no significant aberrations in size in comparison to empty, intact vaults. Furthermore, when analyzed by microscopy B2 complexes bear little indication of compaction. However, during GEMMA analysis, transition to the gas phase may magnify the capacity of different tags to impact the vault's dimensions. This property could provide a means to probe the robustness of different vault structures.

Vault Complexes Compact upon Transition to the Gas Phase

Comparable to reinforcing MVP tags, encapsulated proteins appear to impart stability to the vault structure. Protein-filled vaults migrate at a higher EMD than empty complexes (Figure 10). Since the encapsulated proteins interact directly with the MVP shell, it is feasible that they prevent complete collapse of the vault upon drying from the aqueous state. The densities of these samples (excluding the mass of encapsulated proteins) were 0.45 and 0.41 g/cm^3 for the empty and filled vaults, respectively. Interestingly, calculation of the vault's density in solution revealed a value of 0.22 g/cm^3.

We should point out that the value of 0.22 g/cm^3 is a rough estimation. One must bear in mind that this value derives from modeling the vault as an ellipsoid. The actual density of the vault is expected to be slightly higher, since the contouring of the vault surface leads to an actual volume smaller than that predicted by the ellipsoid model.

Nevertheless, the approximately 0.22 to 0.45 g/cm^3 shift in density equates to a 49% reduction in particle volume upon desolvation. Assuming that all 3 dimensions of the ellipsoidal vault shrink by proportional amounts, the gas-phase dimensions are 79% of those of the aqueous vault. This phenomenon is not unique to vault proteins. Recent studies have revealed similar compaction of the 800 kDa GroEL tetradecamer upon transition to the gas phase (Hogan et al., 2011). On the other hand, some complexes may not experience any structural changes upon desolvation. Our laboratory has previously determined that certain virus capsids, whether empty or containing RNA, migrate at the same EMD, presumably due to structural rigidity (data not shown).

In addition to structural changes induced by desolvation, abnormal protein density may cause deviation from the expected EMD. For example, the 4.6 MDa cowpea chlorotic mottle virus migrates at an exceptionally low EMD for its mass. Applying the density of the average particle to GEMMA data would underestimate the virus' molecular weight; the particle is highly dense due to single stranded RNA packaged inside the viral capsid (Speir et al., 1994). These cases highlight the importance of investigating solution-state protein structure and density before applying the conventional density of 0.58 g/cm^3 to calculate mass from GEMMA data.

Stoichiometry of Protein-Encapsulated Vault Complexes

The susceptibility of the vault structure to packaged proteins can be exploited to quantify the capacity of vaults as drug delivery agents. The difference in EMD between filled and empty

vaults can be converted to mass difference as discussed in Chapter 5. Our GEMMA-derived quantification exhibited agreement with LC-MS results, correlating to 24 and 31 copies of encapsulated CCL21-INT, respectively (Figure 10). Moreover, this data corroborates with the expected stoichiometry of the CCL21-packaged vault. The INT domain of the CCL21 fusion protein binds to the MVP monomer, of which there are 78. Cryo-electron microscopy data has revealed that this binding occurs near the vault waist, manifesting as bands of intense staining on each vault half; refer to Figure 1 for reconstruction of EM data (Kickhoefer et al., 2005). Although this configuration provides for 78 possible CCL21-INT binding sites, steric interactions amongst the adjacent encapsulated CCL21-INT proteins could forestall saturation of binding.

We also measured encapsulation in a vault sample that was prepared with lower concentrations of CCL21. Our measurements yielded 16 and 9 copies of CCL21-INT for GEMMA and LC-MS, respectively (Figure 11). Thus both GEMMA and LC-MS techniques were capable of distinguishing between different preparations of CCL21-INT vault samples.

Using the same technique, we determined the stoichiometry of TEP1 and vPARP/TEP1 packaged vaults (data obtained by Kaddis et al.), concluding that 1 copy of TEP1 and 6 copies of vPARP were packaged inside these samples (Table 4) (Kaddis et al., 2007). Our results agree with the nature of these encapsulated proteins. TEP1's RNA binding function allows it to assist in packaging RNA inside the vault (Kickhoefer et al., 1999a,b). vPARP contains an INT domain that binds directly to MVP (Kickhoefer et al., 1999b). As discussed in Chapter 6, this interaction is highly stable and it is not surprising that more copies of vPARP are incorporated into the vault than TEP1. In support of this presumption, electron microscopy data has revealed that two copies of TEP1 and eight molecules of vPARP per vault are most probable (Kong et al., 2000).

Stability of Vault Complexes

The ability to distinguish between vaults prepared with different amounts and types of encapsulated proteins attests to the sensitivity of the GEMMA technique. Such resolution would be ideal for monitoring subtle changes in vault composition under different conditions, such as long-term storage of samples. In aged samples, GEMMA spectra exhibited decreased overall abundance of protein (Figure 12). Remarkably, we witnessed no change in the EMD of the vaults, whether empty or packaged with CCL21, after storage at 4 °C for a period of 123 days. These results confirm the stability of the MVP shell and the robustness of the association between MVP and INT fusion proteins.

Optimization of Sample Preparation

We were surprised to find that one of our vault samples, CPZ, gave no signal when prepared by conventional methods and assayed by GEMMA. This result is puzzling in the light of the SDS-PAGE data, which revealed that protein, although of low concentration, was present in the desalted sample (Figure 13). As we have established with our studies of the CP2 sample, GEMMA is a fairly sensitive technique and even minor concentrations of vaults should be detectable. Numerous variations on sample preparation (dialysis and ZipTip desalting, altering buffer concentration and pH) were unsuccessful in recovering a signal (Figure 14). We suggest that the intact CPZ vault is refractory towards ionization for reasons unclear but related to the presence of the Z tag at the C-terminus.

Applying GEMMA towards the Study of E2 Protein Complexes

Similar to the CPZ vault, we experienced complications during the analysis of the E2 complexes. SDS-PAGE revealed that the desalting step, despite the use of multiple methods and buffers,

caused complete loss of protein (Figure 16). Unfortunately, skipping the desalting step resulted in high background signal on GEMMA, precluding sample measurement.

We propose that the samples may have aggregated and been excluded during the desalting process. Indeed, E2 fusion proteins are typically purified from the insoluble inclusion bodies of *E. coli* and must be denatured and slowly refolded by dialysis in order to obtain soluble VLPs (Trovato et al., 2012). Although we received pure E2 samples, lyophilization may have influenced protein solubility. Simply reconstituting the lyophilized samples in water may not have been sufficient to restore solubility.

Nevertheless, we did obtain results that attest to the stability of the E2 functional unit. In accord with the results documented by Peng et al., strong denaturing conditions allowed us to observe the E2 trimer by SDS-PAGE (Figures 16 and 17) (Peng et al., 2012). These trimers resisted dissociation into monomers.

Chapter 10

Conclusion

Using ESI-GEMMA (electrospray ionization gas-phase electrophoretic mobility molecular analyzer), we analyzed vault complexes composed of modified MVP subunits. We explored the effects of desolvation on protein compaction and investigated different methods of sample preparation to optimize GEMMA for the study of large protein complexes such as vaults. An EMD-molecular weight correlation specific to vault proteins was established after finding the average particle density to be 0.45 g/cm^3. Using this correlation, we determined the capacity of vaults to be packaged with CCL21 protein therapeutics to be on the order of 20 or 30. Subsequent assays revealed that the concentration of intact vaults decreases over time. However, our results attest to the remarkable stability of the MVP shell, indicating no loss of subunits from the remaining soluble vaults. We also observed complete retention of encapsulated proteins. From the standpoint of employing the vault as a drug delivery agent, this feature is highly desirable because premature release of drug can be avoided. We were unable to obtain GEMMA measurements for both CPZ vaults and E2 complexes. We suggest that alternative sample preparation methods may make such measurements possible. As demonstrated by these studies, GEMMA can be successfully adapted towards the study of high molecular weight protein complexes to yield essential information regarding structure and stoichiometry.

Chapter 11

Future Directions

Recent endeavors have generated modified vaults that bind to monoclonal IgG antibodies targeting the epidermal growth factor receptor (EGFR) (Kickhoefer et al., 2008). Antibody-bound, drug-loaded vaults can then be directed to EGFR-overexpressing cancer cells. Packaging an endosomal lytic protein inside the vault allows for endosomal escape of co-encapsulated therapeutics (Han et al., 2011). However, this mechanism is dependent upon the ability of both the lytic protein and drug to be released from the vault once inside the endosome.

In future studies, it will be interesting to quantify the amount of IgG antibody bound to these EGFR-targeting vaults. It will also be of significance to evaluate vault integrity and release of drug at endosomal pH. The success of attempts to confer vaults with controlled opening and closure properties may also be monitored by GEMMA.

With regards to our unsuccessful characterization of E2 complexes, we suggest that purifying the proteins at a lower salt concentration may facilitate subsequent GEMMA measurements. Should the concentration of E2 be sufficiently high compared to salt, it may be possible to simply dilute the sample into GEMMA running buffer. CP2 vault data demonstrates that skipping the desalting step does not impact the integrity of the spectra and significantly increases signal (Figure 15). As established by this experiment, the desalting step appeared to cause complete loss of E2 protein and should be avoided if possible.

While this study successfully determined an EMD-mass correlation specific to vaults, this method may not be useful for other large and mid-size protein complexes due to the unique

density and shape of the vault. The EMD-mass correlation currently used to interpret the GEMMA spectra of average particles has been derived primarily from the data of small proteins (Figure 5). Data for proteins beyond 1 MDa is uncommon; inclusions of samples of this size would render GEMMA more accurate for application to mid and large size protein complexes.

References

Anderson, D. H.; Kickhoefer, V. A.; Sievers, S.A.; Rome, L.H.; Eisenberg, D. Draft Crystal Structure of the Vault Shell at 9-Å Resolution. *PLoS Biol.* **2007,** *5,* 2661-2670.

Bacher, G.; Szymanski, W. W.; Kaufman, S. L.; Zollner, P.; Blaas, D.; Allmaier, G. Charge-Reduced Nano-Electrospray Ionization Combined with Differential Mobility Analysis of Peptides, Proteins, Glycoproteins, Noncovalent Protein Complexes, and Viruses. *J. Mass Spectrom.* **2001,** *36,* 1038-1052.

Berger, W.; Steiner, E.; Grusch, M.; Elbling, L.; Micksche, M. Vaults and the Major Vault Protein: Novel Roles in Signal Pathway Regulation and Immunity. *Cell Mol. Life Sci.* **2009,** *66,* 43-61.

Dalmau, M.; Lim, S.; Chen, H. C.; Ruiz, C.; Wang, S. W. Thermostability and Molecular Encapsulation within an Engineered Caged Protein Scaffold. *Biotechnol. Bioeng.* **2008,** *101, 654-664.*

Delmas, B.; Laude, H. Assembly of Coronavirus Spike Protein into Trimers and its Role in Epitope Expression. *J. Virol.* **1990,** *64,* 5367-5375.

Domingo, G. J.; Orru', S.; Perham, R. N. Multiple Display of Peptides and Proteins on a Macromolecular Scaffold Derived from a Multienzyme Complex. *J. Mol. Biol.* **2001,** *305,* 259-267.

Ebeling, D. D.; Westphall, M. S.; Scalf, M.; Smith, L. M. Corona Discharge in Charge Reduction Electrospray Mass Spectrometry. *Anal. Chem.* **2000,** *72,* 5158-5161.

Freeke, J.; Robinson, C. V.; Ruotolo, B. T. Residual Counter Ions can Stabilise a Large Protein Complex in the Gas Phase. *Int. J. Mass Spectrom.* **2010,** *298,* 91-98.

Han, M.; Kickhoefer, V. A.; Nemerow, G. R.; Rome, L. H. Targeted Vault Nanoparticles Engineered with an Endosomolytic Peptide Deliver Biomolecules to the Cytoplasm. *ACS Nano* **2011,** *5,* 6128-6137.

Hill, C. P.; Worthylake, D.; Bancroft, D. P.; Christensen, A. M.; Sundquist, W. I. Comparison of the NMR and X-ray Structures of the HIV-1 Matrix Protein: Evidence for Conformational Changes during Viral Assembly. *Prot. Sci.* **1996,** *93,* 3099-3104.

Hogan, C. J., Jr.; Ruotolo, B. T. ; Robinson, C. V. ; Fernandez de la Mora, J. Tandem Differential Mobility Analysis-Mass Spectrometry Reveals Partial Gas-Phase Collapse of the GroEL Complex. *J. Phys. Chem. B* **2011,** *115,* 3614-3621.

Iwasaki, K.; Miyazaki, N.; Hammar, L.; Zhu, Y.; Omura, T.; Wu, B.; Sjoborg, F.; Yonekura, K.; Murata, K.; Namba, K.; Caspar, D. L.; Fujiyoshi, Y.; Cheng, R. H. Pleomorphic Configuration of the Trimeric Capsid Proteins of Rice Dwarf Virus that Allows Formation of Both the Outer Capsid and Tubular Crystals. *J. Mol. Biol.* **2008,** *383,* 252-265.

Izard, T.; Aevarsson, A.; Allen, M. D.; Westphal, A. H.; Perham, R. N.; de Kok, A.; Hol, W. G. Principles of Quasi-Equivalence and Euclidean Geometry Govern the Assembly of Cubic and Dodecahedral Cores of Pyruvate Dehydrogenase Complexes. J. *Proc. Natl. Acad. Sci. U.S.A.* **1999,** *96,* 1240-1245.

Kaddis, C. S.; Lomeli, S. H.; Yin, S.; Berhane, B.; Apostol, M. I.; Kickhoefer, V. A.; Rome, L. H.; Loo, J. A. Sizing Large Proteins and Protein Complexes by Electrospray Ionization Mass Spectrometry and Ion Mobility. *J. Am. Soc. Mass Spectrom.* **2007,** *18,* 1206-1216.

Kaufman, S. L. Analysis of Biomolecules using Electospray and Nanoparticle Methods: The Gas-Phase Electrophoretic Mobility Molecular Analyzer (GEMMA). *J. Aerosol. Sci.* **1998,** *29,* 537-552.

Kaufman, S. L.; Skogen, J. W.; Dorman, F. D.; Zarrin, F.; Lewis, K. C. Macromolecule Analysis Based on Electrophoretic Mobility in Air: Globular Proteins. *Anal. Chem.* **1996,** *68,* 1895-1904.

Kaufman, S. L.; Kuchumov, A. R.; Kazakevich, M.; Vinogradov, S. N. Analysis of a 3.6-MDa Hexagonal Bilayer Hemoglobin from Lumbricus terrestris using a Gas-phase Electrophoretic Mobility Molecular Analyzer. *Anal. Biochem.* **1998,** *259,* 195-202.

Kickhoefer, V. A.; Han, M.; Raval-Fernandes, S.; Poderycki, M. J.; Moniz, R. J.; Vaccari, D.; Silvestry, M.; Stewart, P. L.; Kelly, K. A.; Rome, L. H. Targeting Vault Nanoparticles to Specific Cell Surface Receptors. *ACS Nano* **2008,** *3,* 27-36.

Kickhoefer, V. A.; Stephen, A. G.; Harrington, L.; Robinson, M. O.; Rome, L. H. Vaults and Telomerase Share a Common Subunit, TEP1. *J. Biol. Chem.* **1999a,** *274,* 32712-32717.

Kickhoefer, V. A.; Siva, A. C.; Kedersha, N. L.; Inman, E. M.; Ruland, C.; Steuli, M.; Rome, L. H. The 193-kDa Vault Protein, VPARP, is a Novel Poly(ADP-ribose) Polymerase. *J. Cell Biol.* **1999b,** *146,* 917-928.

Kickhoefer. V. A.; Garcia, Y.; Mikyas, Y.; Johansson, E.; Zhou, J. C. ; Raval-Fernandes, S. ; Minoofar, P. ; Zink, J. I. ; Dunn, B. ; Stewart, P. L. et al. Engineering of Vault Nanocapsules with Enzymatic and Fluorescent Properties. *Proc. Natl. Acad. Sci. U.S.A.* **2005,** *102,* 4348-4352.

Kirshenbaum, N.; Michaelevski, I.; Sharon, M. Analyzing Large Protein Complexes by Structural Mass Spectrometry. *J. Vis. Exp.* **2010,** *40,* e1954.

Kong, L. B.; Siva, A. C.; Kickhoefer, V. A.; Rome, L. H.; Stewart, P. L. RNA Location and Modeling of a WD40 Repeat Domain within the Vault. *RNA* **2000,** *6,* 890-900.

Kong, L. B.; Siva, A. C.; Rome, L. H; Stewart, P. L. Structure of the Vault, a Ubiquitous Cellular Component. *Structure* **1999,** *7,* 371-379.

Lessard, I. A.; Domingo, G. J.; Borges, A.; Perham, R. N. Expression of Genes Encoding the E2 and E3 Components of the Bacillus Stearothermophilus Pyruvate Dehydrogenase Complex and the Stoichiometry of Subunit Interaction in Assembly in Vitro. *Eur. J. Biochem.* **1998,** *258,* 491-501.

Loo, J. A. Studying Noncovalent Protein Complexes by Electrospray Ionization Mass Spectrometry. *Mass Spectrom. Rev.* **1997,** *16,* 1-23.

Mikyas, Y.; Makabi, M.; Raval-Fernandes, S.; Harrington, L.; Kickhoefer, V. A.; Rome, L. H.; Stewart P. L. Cryoelectron Microscopy Imaging of Recombinant and Tissue Derived Vaults: Localization of the MVP N Termini and VPARP. *J. Mol. Biol.* **2004,** *344,* 91-105.

Patel, M. S.; Korotchkina, L. G. The Biochemistry of the Pyruvate Dehydrogenase Complex. *Biochem. Mol. Biol. Educ.* **2006,** *31,* 5-15.

Peng, T.; Lee, H.; Lim, S. Isolating a Trimer Intermediate in the Self-Assembly of E2 Protein Cage. *Biomacromol.* **2012,** *13,* 699-705.

Peng, T.; Lim, S. Trimer-Based Design of pH-Responsive Protein Cage Results in Soluble Disassembled Structures. *Biomacromol.* **2011,** *12,* 3131-3138.

Perham, R. N. Domains, Motifs, and Linkers in 2-Oxo Acid Dehydrogenase Multienzyme Complexes – A Paradigm in the Design of a Multifunctional Protein. *Biochem.* **1991,** *30,* 8501-8512.

Perham, R. N. Swinging Arms and Swinging Domains in Multifunctional Enzymes: Catalytic Machines for Multistep Reactions. *Ann. Rev. Biochem.* **2000,** *69,* 961-1004.

Perham, R. N.; Jones, D. D.; Chauhan, H. J.; Howard, M. J. Substrate Channeling in 2-Oxo Acid Dehydrogenase Multienzyme Complexes. *Biochem. Soc. Trans.* **2002,** *30,* 47-51.

Persson, H.; Kvist, A.; Vallon-Christersson, J.; Medstrand, P.; Borg, A.; Rovira, C. The Non-Coding RNA of the Multidrug Resistance-linked Vault Particle Encodes Multiple Regulatory Small RNAs. *Nat. Cell Biol.* **2009,** *11,* 1268-1271.

Poderycki, M. J.; Kickhoefer, V. A.; Kaddis, C. S.; Raval-Fernandes, S.; Johansson, E.; Zink, J. I.; Loo, J. A.; Rome, L. H. The Vault Exterior Shell is a Dynamic Structure that Allows Incorporation of Vault-Associated Proteins into its Interior. *Biochem.* **2006,** *45,* 12184-12193.

Scalf, M.; Westphall, M. S.; Krause, J.; Kaufman, S. L.; Smith, L. M. Controlling Charge States of Large Ions. *Science* **1999,** *283,* 194-197.

Song, W.; Kong, H.-L.; Carpenter, H., Torii, H.; Granstein, R.; Rafii, S.; Moore, M. A. S.; Crystal, R. G. Dendritic Cells Genetically Modified with an Adenovirus Vector Encoding the cDNA for a Model Antigen Induce Protective and Therapeutic Antitumor Immunity. *J. Exp. Med.* **1997,** *186,* 1247-1256.

Speir, J. A.; Munshi, S.; Wang, G.; Baker, T. S.; Johnson, J. E. Structures of the Native and Swollen Forms of Cowpea Chlorotic Mottle Virus Determined by X-ray Crystallography and Cryo-Electron Microscopy. *Struct. Cell* **1994,** *3,* 63-78.

Stephen, A. G.; Raval-Fernandes, S.; Huynh, T.; Torres, M.; Kickhoefer, V. A.; Rome, L. H. Assembly of Vault-like Particles in Insect Cells Expressing Only the Major Vault Protein. *J. Biol. Chem.* **2001,** *276,* 23217–23220,

Suprenant, K. A. Vault Ribonucleoprotein Particles: Sarcophagi, Gondolas, or Safety Deposit Boxes? *Biochem.* **2002,** *41,* 14447-12254.

Tanaka, H.; Kato, K.; Yamashita, E.; Sumizawa, T.; Zhou, Y.; Yao, M.; Iwasaki, K.; Yoshimura, M.; Tsukihara, T. The Structure of Rat Liver Vault at 3.5 Angstrom Resolution. *Science* **2009,** *323,* 384-388.

Thomas, J., J.; Bothner, B.; Traina, J.; Benner, W. H.; Siuzdak, G. Electrospray Ion Mobility Spectrometry of Intact Viruses. *Spectroscopy.* **2004,** *18,* 31-36.

Trovato, M.; Krebs, S. J.; Haigwood, N. L.; De Berardinis, P. Delivery Strategies for Novel Vaccine Formulations. *World J. Virol.* **2012,** *1,* 4-10.

Yang, J.; Kickhoefer, V. A.; Ng, B. C.; Gopal, A.; Bentolila, L. A.; John, S.; Tolbert, S. H.; Rome, L. H. Vaults are Dynamically Unconstrained Cytoplasmic Nanoparticles Capable of Half Vault Exchange. *A.C.S. Nano* **2010,** *4,* 7229-7240.

Yang, S.-C.; Hillinger, S.; Riedl, K.; Zhang, L.; Zhu, L.; Huang, M.; Atianzar, K.; Kuo, B. Y.; Gardner, B.; Batra R. K. et al. Intratumoral Administration of Dendritic Cells Overexpressing CCL21 Generates Systemic Antitumor Responses and Confers Tumor Immunity. *Clin. Cancer Res.* **2004,** *10,* 2891-2901.