

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Structural characterization of viral capsid maturation in bacteriophage HK97

Permalink

<https://escholarship.org/uc/item/1gg9s15k>

Author

Gertsman, Ilya

Publication Date

2009

Supplemental Material

<https://escholarship.org/uc/item/1gg9s15k#supplemental>

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Structural Characterization of Viral Capsid Maturation in Bacteriophage HK97

A dissertation submitted in partial satisfaction of the requirements for the degree of
Doctor of Philosophy.

in

Biochemistry

by

Ilya Gertsman

Committee in charge

Professor John E. Johnson, Chair
Professor Elizabeth A. Komives, Co-chair
Professor Simpson Joseph
Professor Amitabha Sinha
Professor Douglas E. Smith
Professor Deborah H. Spector

2009

Copyright

Ilya Gertsman, 2009

All rights reserved.

The Dissertation of Ilya Gertsman is approved, and it is acceptable in quality and form
for publication on microfilm and electronically:

Co-chair

Chair

University of California, San Diego

2009

TABLE OF CONTENTS

Signature Page	iii
Table of Contents	iv
List of Abbreviations	viii
List of Figures	x
List of Tables	xxiii
Acknowledgements	xxiv
Vita and Publications	xxv
Abstract of the Dissertation	xxvi
Chapter 1 Introduction To Viral Capsid Maturation	1
1. Introduction	1
a. Sequential packaging of genome segments accompanied by conformational changes in an RNA bacteriophage	4
b. HK97, a model system for understanding dsDNA maturation in bacteriophage	9
c. Capsid maturation in other dsDNA bacteriophage and Herpes Virus	16
d. Understanding the structural mechanism in HK97 capsid maturation	20
2. References	24
Chapter 2 X-ray Structure Determination of HK97 Prohead II	28
1. Introduction	28
2. Methodology and Results.....	29

	a. W336F/E-loop truncation mutation and Crystallization	29
	b. Indexing and scaling	35
	c. Particle orientation and the self-rotation function	41
	d. Molecular replacement and averaging	42
	e. Model building and refinement	45
	f. Summary of 5.2 Å structure data collection and processing	47
	g. Comparison of 3.65 Å and 5.2 Å structures	49
	3. Acknowledgements	53
	4. References	54
Chapter 3	Application of H/ ² H exchange to studies of Viral Capsid Dynamics and maturation.....	56
	1. Introduction	56
	2. Methodology and Results	63
	a. Optimization of H/ ² H exchange for HK97	63
	b. H/ ² H exchange of capsid maturation	69
	3. Acknowledgements	74
	4. References	75
Chapter 4	An Unexpected Twist in Viral Capsid Maturation	77
	1. Introduction	77
	2. Methods and Materials	78
	a. Mutagenesis	78
	b. Crystal Growth and Structure determination	79
	c. H/ ² H exchange	81

3. Results and Discussion	82
4. Acknowledgements	98
5. References	99
Chapter 5 Resolving the molecular reorganization and refolding of the maturing HK97 viral capsid	101
1. Abstract	101
2. Introduction	102
3. Methods and Materials	106
a. Protein Preparation and Crystallization	106
b. Crystallographic processing	107
c. H/ ² H exchange	108
4. Results and Discussion	110
a. Crystallographic comparison between P-II and H-II	110
b. H/ ² H exchange	121
5. Conclusions	134
6. Acknowledgements	136
7. References	137
Chapter 6 Characterizing the relevance of 3-fold interactions in capsid assembly, stability, and expansion mechanism.....	141
1. Introduction	141
2. Methods and Materials	143
a. Mutagenesis, Protein expression, and analysis	143
b. Hexon and Penton Inter-conversion	144
c. Fluorescence kinetics	144

	d. Differential Scanning Calorimetry	145
3.	Results	145
	a. Expression and morphologies	145
	b. 3-fold mutant expansion	149
	c. Differential Scanning Calorimetry	154
4.	Discussion	158
5.	References	163
Chapter 7	Thesis Conclusions	165
	a. References	171
Appendix A	Peptide identification by Tandem Mass Spectrometry	172

List of Abbreviations

Δ -domain	Delta-domain (residues 2-103)
Å	Angstroms
A-domain	Axial-domain
A-loop	Axial-loop
B.S.A.	Buried Surface Area
Cryo	Cryogenic
dsDNA	double-stranded DNA
EI	Expansion Intermediate
E-loop	Extended-loop
EM	Electron Microscopy
gp3	gene product 3
gp4	gene product 4
gp5	gene product 5
gp5*	gene product 5 after proteolytic cleavage (residues 104-385)
H/ ² H	Hydrogen/Deuterium
H-I	Head I
H-II	Head II
HK97	Hong Kong 97
kD	kilodalton
ms/ms	Tandem Mass Spectrometry

N-arm	N-terminal arm
NCS	Non-crystallographic symmetry
Pdb	Protein data bank
P-domain	Peripheral-domain
PEG	Polyethylene glycol
P-II	Prohead II
SAXS	Small angle X-ray scattering
Tof	Time of flight
WT	Wild Type

List of Figures

- Figure 1.1: **Genome packaging mechanisms.** (A) A cutaway view of Flock House Virus with its RNA shown in the center, structured as a dodecahedral cage. The Flock House Virus capsid assembles around its RNA genome. (B) P22, a dsDNA bacteriophage, packages its genome into a pre-assembled particle through a 5-fold portal complex. Concentric rings of packaged dsDNA are shown in green. (C) Phi6, a dsRNA bacteriophage, is thought to package its genome through symmetrically related openings around the icosahedral capsid. The above figures are adapted from references 1-3 in the main text..... 2
- Figure 1.2: **Viral quasi-equivalence.** Icosahedral viruses are composed of symmetrically related building blocks. The icosahedron can be created from a hexagonal lattice that is shown in figure a. The icosahedron is composed of 20 identical triangular faces which can vary in size based on the number of hexagonal points included in the face. The formula: $T=h^2+hk+k^2$, calculates the T number of a virus particle based on the number of lattice points used in the h and k directions. (B) The construction of a T=3 capsid particle. (C) A T=1 particle, the simplest Icosahedral particle, composed of 60 identical copies of a single subunit. The particle is composed entirely of pentamers of subunits (12 of the in all). (D) A T=7l particle, the same triangulation number as HK97. The particle is composed of 7 asymmetric subunits. When applying the 60 icosahedral matrices to number of asymmetric subunits, one generates the entire capsid particle. Casper and Klug originally described quasi-equivalence for viruses that had a larger triangulation number than 1. In those cases, the T-number is equal to the number of asymmetric subunits in the icosahedron, each with potential variations, therefore referred to as quasi-equivalence. Illustrations above were taken from the virus database: <http://viperdb.scripps.edu>..... 6
- Figure 1.3: **Figure 1.3: Phi6 capsid maturation.** (A) The phi6 Cryo-EM density is shown with its corresponding NC shell and PC shell colored in orange and yellow respectively. The poorly ordered P4 hexamers are colored in green. The unexpanded PC shell (B) and expanded shell (C) depict the major conformational changes

	occurring in the P1 protein interfaces following RNA packaging, Panel A has been adapted from reference 41, while panels B and C are from reference 14.....	8
Figure 1.4:	In vivo assembly of HK97. Capsid subunits assemble around protease proteins and portal complex to form the procapsid particle. DNA is packaged in an ATP dependant pathway resulting in particle expansion from a spherical shell to the icosahedral form. An EM micrograph of a native HK97 bacteriophage is shown at the right of the schematic.....	11
Figure 1.5:	In vitro HK97 expansion. Electron density maps of distinct intermediate characterized by EM (P-II and EI), and crystallography (Balloon, and Head II). Intermediate particles were attained <i>in vitro</i> by low pH treatment of Prohead II particles, assembled in an <i>E.coli</i> expression system. K169Y mutant is crosslink defective causing the expansion process to end at the Head I state. Head I is the morphological equivalent of WT Balloon particles, except no crosslinks are present.....	12
Figure 1.6:	(A) Mature subunit of HK97. Residues N356 and K169, which are involved in covalent crosslinking, are shown as colored spheres. (B) Isopeptide bond formation, catalyzed by a nearby glutamate at residue 363.(C) Isosurface rendering of Head II structure. The region highlighted in blue represents the asymmetric subunit of the T=7 capsid, which includes the hexamer subunits and a penton subunit. The full capsid can be reproduced by applying the 60 icosahedral matrices to these 7 subunits. The hexon subunits that form crosslinks around the pentamers are highlighted in orange and shown in (D).....	14
Figure 1.7:	HK97, a common fold in dsDNA bacteriophage and Herpes Virus. (a) Hk97 subunit. (B) HK97 subunit fit into phi29 Cryo-EM reconstruction. (C) HK97 subunit fit into P22 Cryo-EM density. (D) Gp24 subunit (pentamer subunit) of T4. (E) HK97 fit into Cryo-EM reconstruction of Lambda. (F) Secondary structure elements fit into Cryo-EM density of Herpes virus (HSV-1). The figures above were adapted from references 2,20,34, and 42.....	17
Figure 1.8:	Capsid maturation. The particles on the left represent procapsid states of HK97, P22, HSV-1, and phi29. The particles on the right are the matured capsid forms. (A) HK97 rendered from crystallographic coordinates. (B) Cryo-EMreconstructions of P22. (C) Cryo-EM reconstructions of HSV-1. (D) Cryo-EM reconstructions of phi29 prolate procapsid and mature capsid particles. Panels b-d were adapted from references 34,12, and 43	

	respectively.....	21
Figure 2.1:	E-loop truncation (A) Diagram showing multiple steps associated with splicing by overlap extension. Primers coding for the mutation are in the middle (b,c), while flanking primers (a,d) are used to create two intermediate segments of dna with overlapping sequences (AB,CD). After the overlapping fragments hybridize and are completed using the flanking primers, the complete AD fragment with the mutation can be inserted into a vector. B) Prohead subunit D showing the region of the truncated E-loop in orange. This region, spanning residues 159-171 were replaced with an APGD sequence.....	30
Figure 2.2:	P-II crystallization at 3.65 Å. (A) A representative P-II crystal (ranging from 0.4-1.0mm in size). (B) NDSB-211, the additive used in crystallization. C) The mother liquor solution (100mM CHES pH 9.0, 200mM Manganese chloride) before the precipitant has settled.....	34
Figure 2.3:	Diffraction pattern of P-II. (A) Full diffraction image as recorded from the detector and (B) a zoomed-in region corresponding to high resolution data. The image was taken at APS, beamline 14-BMC.....	36
Figure 2.4:	Diffraction intensity. (A) The change in average intensity of recorded reflections as a function of resolution. (B) Intensity/sigma (I) as a function of resolution.....	38
Figure 2.5:	Packing of the P-II crystals in the I222 cell. 1 full particle is located at the unit cell center, while 1/8 of each particle is located at each corner. The cell dimenstions are a=553.0 Å, b=574.4 Å, c=587.4 Å.....	39
Figure 2.6:	Self-rotation functions (A) 2-fold rotation function. Scaled data between 12 and 8 Å was used for the search. Peaks are a result of crystallographic 2-fold axes. (B) 3-fold rotation function. 12-8 Å data was also used for this search. Peaks represent 3-fold icosahedral axes.(C) 5-fold rotation function. Peaks indicate 5-fold icosahedral axes. Data only spanning 3.95-3.65 Å was used in the search. Using only the high-resolution recorded reflections in the rotation search provides an effective tool for evaluating the high-resolution data. If expected peaks in the rotation function search are not present using the high resolution data, the data likely will not contribute to structure determination. The 5-fold search shown above shows all expected peaks.....	40

- Figure 2.7: **Crystallographic statistics.** Data in the highest resolution shell is shown in brackets. (A) Preliminary scaling statistics for data processed without a sigma (I) cutoff. The R_{merge} value in the highest resolution shell is extremely high, indicating data in this ranges makes little or no valid contribution. (B) Final Statistics for data processed with the sigma (I) cutoff. Statistics for experiments other than scaling are shown as well, including Averaging, refinement, and stereochemistry. The following equations are used to calculate R_{cryst} , CC_{ave} , and R_{ave} respectively:

$$R_{\text{cryst}} = S(|F_{\text{obs}}| - k |F_{\text{calc}}|) / S|F_{\text{obs}}| ; k = S F_{\text{obs}} / S F_{\text{calc}} ;$$

$$CC_{\text{ave}} = S(|F_{\text{obs}}| - \langle |F_{\text{obs}}| \rangle) (|F_{\text{calc}}| - \langle |F_{\text{calc}}| \rangle) / (S(|F_{\text{obs}}| - \langle |F_{\text{obs}}| \rangle)^2 S(|F_{\text{calc}}| - \langle |F_{\text{calc}}| \rangle)^2)^{1/2}$$
 R_{ave} is the same as R_{cryst} except F_c is calculated from a 15-fold NCS averaged electron density..... 43
- Figure 2.8: **Averaging and phase extension statistics.** The plot shows the change in correlation coefficient for various cycles of averaging. The consecutive cycles use increasing amounts of scaled diffraction data. After each cycle, additional higher resolution data is incorporated into the averaging process. As seen, the highest resolution data, which is usually weaker than the lower resolution data causes a significant decrease in the correlation coefficient. Though the correlation coefficient drops to lower than 0.1 at 3.65 Å, data in this resolution range still contributes positively to the structure determination, verified both by the rotation function using only high resolution data, and direct electron density map comparisons at different resolutions..... 44
- Figure 2.9: **P-II stereochemistry.** (A) Ramachandran plot of 3.65 Å P-II structure. Percentages of residues with favorable, allowed, and unfavorable phi/psi values are listed below the plot. Each spot on the plot represents a specific residue. Red regions are most favored, yellow are additionally allowed, the tan colored zone corresponds to generously allowed, while the white area is the disallowed region. (B) Ramachandran plot of 5.2 Å P-II structure..... 48
- Figure 2.10: **Comparison of 3.65 Å and 5.2 Å electron density maps.** (A) Spine helix of subunit F fit into the 3.65 Å electron density. Notice the sharp grooves of the map in the helix region as well as the extensive density for the side chain atoms. (B) Subunit F fit into 5.2 Å electron density. Virtually no density is resolved for side chains residues, allowing only a confident fitting of the protein backbone..... 52
- Figure 3.1: **Rates of H/²H exchange is limited by structure and interactions.** Schematic representations of varying rates of exchange for various

	structural elements such as (A) alpha-helices, and (B) loops. (C) Protection from solvent exchange can result from quaternary interactions, or residues buried in binding pocket.....	58
Figure 3.2:	pH determines amide proton exchange rate. (A) The rate of exchange increases logarithmically below pH ~2, and above pH ~3. The logarithmic changes in rate are due to two different exchange mechanisms, (B) and acid catalyzed mechanism, C) and a base catalyzed reaction.....	59
Figure 3.3:	Gaussian envelopes. Mass envelopes of a 19 residue peptide spanning residues 288-306 of the HK97 capsid protein sequence. The first trace shows the non-deuterated mass envelope, while the next tow traces show shifted mass envelopes after 30 seconds and 10 minutes of deuterium exchange respectively.....	61
Figure 3.4:	Protocol of H/²H exchange as optimized for HK97. After incubation of protein with deuterium solution, the exchange is quenched with non-deuterated acidic solution, digested with pepsin beads, mixed with alpha-C matrix and analyzed on the MALDI-TOF Voyager.....	64
Figure 3.5:	Fragments identified and analyzed in P-II. The total number of residues covered represents 68% of the sequence. Arrows stacked on top of each other represent unique peptides that have overlapping coverage in sequence, providing increased resolution of just the residues corresponding to the overlap.....	67
Figure 3.6:	Salt concentration affects mass spectra: Traces of Prohead II with and without KCL are shown. The trace in red, corresponding to no KCL, shows better coverage of the higher masses while the purple trace (w/KCl) shows better sensitivity in the lower masses. The trace in green shows the spectra of K169Y mutant EI in the presence of isobutanol and KCl.....	68
Figure 3.7:	Using salt to select for specific peptides. The first trace shows two nearby envelopes that appear when P-II is flown with no salt. The envelopes correspond to two separate peptides with different sequences, which when deuterated, results in an overlap of mass envelopes (shown in trace 2) that cannot be confidently deconvoluted. We were only interested in the first envelope, which corresponded to a region of particular interest. By increasing the concentration of salt to ~150mM KCl, the second peptide was no longer visible as seen in the non-deuterated spectra (trace 3), and the deuterated form (trace 4).....	70

- Figure 3.8: **Crosslinking hinders WT mass spectra.** (A) A 10% SDS PAGE gel showing the increase in crosslinking that occurs during WT HK97 expansion. Various oligomeric forms of covalent interactions are seen, ranging from covalent dimers to 6-mers. (B) Mass spectra were run after various times of expansion. Comparing the spectra to the gel, it is clear that crosslink formation greatly inhibits the detection of peptides..... 72
- Figure 4.1: **HK97 assembly and morphology.** (a) The 384-residue gp5 subunit initially assembles into hexameric and pentameric oligomers, termed capsomers, that first assemble to form the Prohead I capsid (P-I). The T=7 *laevo* particle is composed of 12 pentamers and 60 hexamers and encapsidates approximately 60 copies of gp4 protease [9-11]. Expression with a defective protease produces a P-I particle that can be disassembled *in vitro* into free capsomers and re-assembled when exposed to specific chemical treatments [12]. When active gp4 is present, particles spontaneously mature to the 13-MDa Prohead II (P-II) form following digestion of residues 2-103 from all subunits. Crosslinking occurs in the WT particle following formation of the EI state. Crosslinks (isopeptide bond) form between Lys 169 and Asn 356 located on different subunits. A crosslink defective mutant K169Y expands to Head I, a state nearly identical to Balloon minus the crosslinks. WT Balloon undergoes a final expansion step to Head II in which the pentons become more protruded and form one last class of crosslinks, with a molecular topology similar to chainmail [2,13]. (b) Crystal structure of subunit D of Prohead II at 3.65 Å (c) 3.65 Å electron density map (displayed as a solid surface) of the full Prohead II capsid, contoured at $\sim 1\sigma$ in chimera. The Prohead II hexamers and pentamers are shown alongside the capsid with the 7 subunits of the viral asymmetric subunit labeled A-F for the hexamers and G for the pentamers. (d) A calculated electron density map of the Head II capsid shown at 3.65 Å, also rendered at $\sim 1\sigma$. (e) Prohead II and Head II hexamers shown tangential to the capsid surface (rotated 90 degrees from view in 1C and 1D)..... 84
- Figure 4.2: **P-loops located at 3-fold axes act as invariant pivot points.** (a) Ribbon representation of P-II. The orange triangle represents an icosahedral 3-fold, black and magenta triangles represent two quasi-3-fold positions. A zoomed in view of subunits at a quasi-3 fold axis are shown viewed from outside of the capsid. Residues 345-353 of the P-loop are colored lime-green, and represent the peptide fragment analyzed by H/²H exchange. (b) Side-chain interactions at

3-folds that remain invariant during expansion, viewed from the interior of the capsid directly underneath a quasi-3-fold axis, 180 degrees from the view of the trimer in figure a. Electron density is contoured at 1σ . The 3 outer circles highlight salt bridges while the center circle highlights three glutamates (E348) coordinated at a putative metal binding site. (c) Subunit G of P-II (yellow) and H-II (green) have been aligned by the region of the P-domain which remains invariant (blue). (This motion is best captured in Supplementary Movie 1). (d) A P-II trimer (subunits A,F,G as shown in a) is aligned on the H-II trimer (green) by the regions that remain invariant (blue), illustrating the rotational motions in respect to the fixed trimeric interactions. The upper view is looking down a quasi 3-fold while the lower figure shows a perpendicular view, tangential to the capsid surface. (e) Table shows the angles of rotation (θ) of each subunit from P-II to H-II as illustrated in both c and d. (f) H^2/H exchange curve of a peptide fragment spanning residues 345-353 of the P-loop (colored lime-green in a) shown for P-II, H-I, and free capsomer states. Time points are taken from 30 seconds to 10 minutes, with error bars representing standard deviations from the average of 3 independent experiments, with 2-3 measurements per experiment (6-9 total measurements for each timepoint)..... 87

Figure 4.3: **Spine helix bends during maturation.** (a) The spine helix (yellow) is shown for subunit F of Prohead II both in its corresponding electron density on the left (1σ), and aligned with Head II on the right (green). The subunits from the two states were aligned using the subunit core that acts mostly as a rigid body (residues 230-383) with an r.m.s.d of 1.3 Å or better for each alignment. The region colored blue represents the fragment spanning residues 206-216, analyzed by H^2/H exchange. (b) H^2/H exchange rate curves comparing deuterium exchange between P-II, EI-I, H-I, and free capsomer helix fragment..... 93

Figure 4.4: **A working hypothesis for the formation, meta-stability, and subsequent maturation of HK97, represented with a single hexamer.** Individual subunits are first assembled into hexamers and pentamers (far left figure is a hypothetical representation of the initial subunit organization). Based on H^2/H exchange data subunit tertiary structures are distorted in free capsomers and we hypothesize that the hexamers are skewed (second figure). Prohead I is formed by assembly of hexamers and pentamers into a T=7 particle with Δ -domains attached. Following proteolysis of the Δ -domains to form Prohead II, the skewed hexamers and distorted tertiary structures are preserved by quaternary structure interactions

in the particle, raising the free energy of the particle to a meta-stable state maintained in a local minimum. Perturbation of these particles by dsDNA packaging (in vivo) or lowering the pH (in vitro) lowers the energy barrier, leading to an exothermic expansion of the particles producing symmetric hexamers and undistorted subunit tertiary structures..... 94

Figure 4.5: **The P-loop is solvent protected throughout maturation.** Mass envelopes of the peptide fragment spanning residues 345-353 of the P-loop are shown at various expansion states following 1 minute deuterium incubations. Row 1 shows the non-deuterated control, while the next three rows shows the P-II, EI and H-I states respectively. The bottom row shows the mass envelope of free capsomers. The free capsomers exchanged much more deuterium into the P-loop peptide, so much so that the envelope merged into the next peak in the spectrum. The spectral window has been truncated at the next peak for ease of viewing..... 95

Figure 5.1: **HK97 maturation.** (A) The schematic depicts the assembly and expansion of HK97 in an *E.coli* expression system, lacking the portal protein and other machinery required for genome packaging. The expansion pathways of both WT and crosslink defective mutant, K169Y, are shown in response to chemical perturbation *in vitro*. (B) Structural model of P-II, subunit A, color coded by the four major domains labeled in bold. (C) P-II and H-II particle forms are shown rendered as a surface representation of the corresponding crystal structures..... 111

Figure 5.2: **Buried surface area calculations.**(A) Graph shows the changes in buried surface area (B.S.A) of quaternary interactions for subunits A (orange trace), B (purple trace), F (sky-blue trace), and G (yellow trace) during expansion from P-II to H-II states, using calculations derived from their respective crystal structures. Regions of interest are circled and labeled. The color of the label corresponds to the color of the region in panel B. (B) The structure of the P-II super-penton (5 hexons surrounding a penton) is colored in regions that have significant changes in their quaternary interfaces during expansion (excluding the P-loop, which maintains nearly identical 3-fold contacts). The structure is, color-coded according to the domain labels in the graph above. Note, The N-arm undergoes a major re-ordering during expansion but is barely visible in the super-penton, as it is positioned inside the P-II capsid. The view is from the exterior of the capsid, looking down the 5-fold axis..... 113

Figure 5.3:	Quaternary interactions of A-loops and spine helix. (A) Stereo-view of Prohead II hexon with subunits labeled A-F. Quaternary Interactions between subunits are shown for the A-loop (center circle) of the A-domain (panel B shows zoomed in view), and the interface between the spine helix of subunit F with the neighboring subunit A (outer circle, with zoomed-in view in panel C). (D) Stereo-view of Head II hexon. (E) Quaternary interactions between subunits A and F, zoomed-in view of central circle from panel D. (F) Interface between spine helix in subunit F with subunit A, zoomed-in view of outer circle from panel D. Side chain interactions, distances, and corresponding residues are labeled in panels B and E. The salt bridge between R210 of the spine helix and E153 of the neighboring E-loop is labeled in both panels C and F.....	118
Figure 5.4:	Refolding in A-loop and spine helix (A-D) Each subunit from P-II has been individually aligned by least squares fitting with its corresponding subunit in H-II. Panels A-D show A-loop differences for subunit A,B,F, and G respectively. (E-H) Panels showing difference between subunits A,B,F, and G in the spine helix.....	120
Figure 5.5:	Fragments used in the H^2H exchange study for the P-II, EI, and H-I intermediate states. Arrows represent the sequence of a single fragment. Arrows stacked upon each other represent separate fragments that have overlapping sequences. All fragments were generated by pepsin digestion following native capsid incubation with deuterium solution. Identical fragments spanning 138/282(49%) possible residues have been characterized for the three expansion states.....	122
Figure 5.6:	H^2H exchange kinetics. Deuterium exchange of identified peptide fragments is plotted versus time for the P-II(■), EI(○), and H-I(Δ) states, shown in the next two pages. The corresponding residues and domains pertaining to each fragment are shown above the curves. Fragments are grouped into separate columns according to domains or sub-domains of interest. Curves were fit to either a single or double exponential rate, as described in the methods.....	124
Figure 5.7:	H^2H exchange of P-II and H-I. (a) Subunit E of P-II is color coded (spectrum below) according to the percentage of amide protons that exchange with deuterium after 10-min for each peptide fragment. Regions of overlap between peptide fragments are also color coded according to their exchange percentage. (b) Subunit E of H-II is color coded based on the percentage of amide protons	

	exchanging in H-I state after 10-min incubations. The H-II model was used since it has been previously determined that H-II subunit E is nearly structurally identical to that of H-I.....	127
Figure 5.8:	H/²H exchange envelopes of N-arm and A-loop. (A) Mass envelopes of a fragment spanning residues 117-126 in the N-arm, shown at 1-min and 10-min time-points for P-II, and 1-min time-points for EI and H-I. (B) Mass envelopes of a fragment spanning residues 288-306 in the A-loop, shown for various time-points in the P-II state. (C) Mass envelopes shown for the A-loop fragment at different intermediate states. The mass envelopes correspond to 10-minute timepoints.....	128
Figure 5.9:	H/²H exchange envelopes of spine helix fragment spanning residues 204-216. (A) Mass envelopes of helix fragment in P-II state after 0, 1, 5, and 10 minutes of exchange. (B) Mass envelopes of the helix fragment after 1-minute of exchange for P-II, EI, and H-I expansion states. (C) Bimodal envelope from P-II 1min spectra is fitted to 2 gaussian distributions. (D) Table showing calculations from the bimodal gaussian fit in panel C. The centroids of the resolved gaussians are shown as well as the number of amide protons corresponding to each envelope (after 0 min subtraction and back exchange correction). Finally, the number of subunits attributed to each gaussian is calculated by dividing the area of each gaussian by the total area and multiplying the value by 7, which corresponds to the number of subunits in the viral asymmetric unit.....	132
Figure 6.1:	3-fold mutations. The P-II structural models of subunits A,F, and G, looking from the interior of the capsid through the quasi-3fold axis. Only one of the three sets of quasi-equivalent interactions has been highlighted above with the corresponding side chain residues. Five mutations have been characterized in the study: E348A (labeled 3 above), E344A (labeled 2), E344Q (labeled 2), R194N (labeled 1), and a double mutant, E344A/E348A.....	142
Figure 6.2:	R194N characterization. (A) SDS Acrylamide gel comparing pellets from initial lysis following protein purification of E344A and R194N constructs (grown in E.coli.). The R194N prep only contained the uncleaved gp5 subunit unlike E344A which shows primarily gp5*. (B) Native acrylaide gel (7.5%) showing WT capsomers that have been previously treated to convert to predominately hexon and penton forms respectively. The R194N sample from initial purification shows that the oligomeric form is capsomers (mixture of hexons and pentons). (C) Native acrylaide	

	gel (7.5%). R194N capsomers were concentrated (first lane), resulting in a change in ratio between hexon and penton. The sample was then treated with either hexon converting or penton converting conditions as WT capsomers in B.....	147
Figure 6.3:	Mechanical perturbation. (A) 1% Agarose gel shows that ultra-centrifugation (35krpm in 50.2Ti rotor) of E344Q particles results in very slight expansion, while a much greater fractions of E344A undergoes expansion during spinning. (B) Agarose gel shows no mechanically induced expansion in double mutant, as normally expected for P-II particles.....	148
Figure 6.4:	Low pH (Acid) induced expansion assays. (A) 1% agarose gels comparing change in particle forms as a function of time following incubation of samples in pH 4.0 buffer. The upper band represents expanded particles and the lower band represents P-II. The gel compares WT P-II and E348A P-II expansion. (B) WT-P-II and E344Q expansion, similar to A. (C) WT-P-II and E344A/E348A (referred to as double mutant or DM) expansion.....	150
Figure 6.5:	Fluorescence monitored expansion. P-II particles of the various mutants and WT are monitored for their change in spectral center of mass following incubation with acidic buffer (pH 4.0). The time for half of the WT particles to expand ($t_{1/2}$) is shown in the table, as are the rates for mutant expansion as a fraction of the WT expansion rate.....	152
Figure 6.6:	Isobutanol induced expansion. 1% agarose gels comparing WT expansion (after treatment with a 10% final concentration of isobutanol) with (A) E344Q mutant, (B) E348A mutant, and (C) E344A/E348A.....	153
Figure 6.7:	Differential scanning calorimetry of 3-fold mutants. DSC scans and corresponding endotherms of (A) WT-P-II, (B) E344Q P-II, (C) E348A P-II, (D) E344A/E348A (double mutant), and (E) E344A. (F) A modified figure of WT-P-II result from previous study [6]. (G) Overlay of Scans from A-E and corresponding melting temperatures. There is an asterisk (*) next to E344A as the endotherm looks significantly different from the other mutants and WT and should not be quantitatively compared.....	156
Figure 6.8:	Resolving DSC endotherms. (A) The figure shows three overlapped traces corresponding to three separate scans. First, WT-P-II (blue trace) was heated until full denaturation, showing a small shoulder before the peak corresponding to particle melting. A fresh	

	WT-P-II sample was scanned until reaching 72°C (green trace), than cooled to room temperature. The same sample was then re-scanned until full denaturation (red), showing that the shoulder no longer contributed to the endotherm. (B) A 1% agarose gel showing that the 72°C sample was in fact still P-II in morphology (as could be resolved by gel-chromatography), not an expanded particle, which would run in a similar place as the sample in lane 2 (an expanded control).....	157
Figure A.1:	Fragment 1560(m/z). The fragment was identified as residues 275-287 of gp5 with the sequence: IVLNPRDWHNIAL. Letters correspond to amino acids of internal fragments.....	175
Figure A.2:	Fragment 1240. The fragment corresponds to residues 117-126 of gp5, with the sequence: IQPMQIPGIIM. (A) The non-adducted species identified by the more common daughter ions including, y+, b+, and internal fragments. (B) The adducted species, unusually containing many a+ ions, adducted with sodium. The sodium adduct appears to have greatly contributed to the fragmentation mechanism of the parent ion.....	176
Figure A.3:	Fragment 1862 (m/z). The fragment was identified as residues 275-287 of gp5 with the sequence: SASGIVLNPRDWHNIAL.....	177
Figure A.4:	Fragment 1701(m/z). The fragment was identified as residues 248-263 of gp5 with the sequence: NATGDTRADIIAHAIY.....	178
Figure A.5:	Fragment 2114 (m/z). The fragment was identified as residues 288-306 of gp5, with the sequence: LKDNEGRYIFGGPQAFTSN. The peptide was identified with mostly y+ ions and b+ ions, many of which had losses of ammonia.....	179
Figure A.6:	Fragment 1327(m/z). The fragment was identified as residues 206-216 of gp5, with the sequence: YINNRLMYGLA. Most of the daughter ions contained losses of ammonia.....	180
Figure A.7:	Fragment 1137(m/z). The fragment was identified as residues 345-353 of gp5, with the sequence: VSREDRDNF.....	181
Figure A.8:	Fragment 1542(m/z). The fragment was identified as residues 204-216 of gp5, with the sequence: QSYINNRLMYGLA. Similarly to the 1327 fragment, which has most of the same sequence, the fragmentation resulted in many daughter species with losses in ammonia.....	182

Figure A.9: **Fragment 1123(m/z)**. The fragment was identified as residues 331-339 of gp5, with the sequence: MASQVWDRM..... 183

Figure A.10: **Fragment 1811(m/z)**. The fragment was identified as residues 288-303 of gp5, with the sequence: LKDNEGRYIFGGPQAF..... 184

List of Tables

Table 2.1:	Virus crystal screen. Each buffer is mixed with each of the salt conditions, resulting in 192 total conditions. The mixtures are all at 2x concentration. For the purposes of screening P-II, the mixtures were diluted to 1x with a final concentration of 3.0% Peg 4K, which was previously shown to induce crystallization.....	32
Table 2.2:	Crystallographic statistics of 5.2 Å structure. Values in brackets represent the highest resolution shell.....	51
Table 3.1:	Peptide detection optimized by salt conditions. The table shows the conditions that allows for optimal analysis of the corresponding peptide fragment. Optimal concentrations for the salts listed above are at or near 100mM.....	66
Table 4.1:	Crystallographic statistics for Prohead II. Values in brackets indicate highest resolution shell. $R_{\text{cryst}} = \frac{S(F_{\text{obs}} - k F_{\text{calc}})}{S F_{\text{obs}} } ; k = S F_{\text{obs}} / S F_{\text{calc}} ; CC_{\text{ave}} = \frac{S(F_{\text{obs}} - \langle F_{\text{obs}} \rangle) (F_{\text{calc}} - \langle F_{\text{calc}} \rangle)}{(S(F_{\text{obs}} - \langle F_{\text{obs}} \rangle))^2 S(F_{\text{calc}} - \langle F_{\text{calc}} \rangle)^2}^{1/2}$ R_{ave} is the same as R_{cryst} except F_c is calculated from a 15-fold NCS averaged electron density map. R_{free} was not calculated as previous studies have shown this to be an incorrect validation method for data with high non-crystallographic symmetry.....	96
Table 5.1:	Buried surface area comparisons between 3.65 Å Prohead II crystal structure and the 3.44 Å Head II crystal structure. Values are in Å ² . Measurements are based on buried surface area within each hexon and penton (intra-capsomer), or between neighboring capsomers (inter-capsomer).....	115

Acknowledgements

I would like to thank Jack Johnson for his thoughtful mentorship, guidance, and support throughout my graduate career. Jack's exuberance for science is inspiring and is a constant source of motivation. I would also like to thank everyone that I have worked with in the Johnson Lab. I am most indebted to Lu Gan, a previous graduate student who mentored me in crystallography, and whose passion for science is equally as contagious as Jack's. Lu's encouragement and patience has enabled me to learn crystallography, a feat that would have likely take another 10 or so years without his guidance. Along with Lu, Kelly Lee and Rick Huang have also been invaluable collaborators with whom I have greatly enjoyed working with. I would like to thank Jeff Speir and Vijay Reddy who have been extremely helpful in my crystallographic studies. Also, our collaborators in Pittsburgh, namely Roger Hendrix, Bob Duda, Crystal Moyer, and Brian Firek have been crucial to the success of our work with HK97. I have also had the good fortune to collaborate with Elizabeth Komives and various members of her lab including Miklos Guttman and Julia Keoppe, in order to apply H/²H exchange and mass spectrometry to my studies. I am very thankful to Betsy and her lab, as they have been extremely supportive in both training and guidance. Lastly, but most importantly, I would like to thank my family, especially my mother and father for their guidance and endless support.

Chapter 2, in part, is from a manuscript in preparation. The dissertation author is the primary investigator of the study.

Chapter 4, in full, appears in Nature, 2009, Ilya Gertsman, Lu Gan, Miklos Guttman, Kelly Lee, Jeffrey Speir, Robert L.Duda, Roger Hendrix, Elizabeth A. Komives, and John E. Johnson. The dissertation author is the primary investigator of the study.

Chapter 5, in full, will be submitted to Journal of Molecular Biology, co-authored with Elizabeth A. Komives and John E. Johnson.

Vita

2000 University of New Mexico
2000-2003 Research Assistant, University of Pennsylvania
2003-2009 University of California, San Diego

Publications

Gertsman, I., Gan, L., Guttman, M., Lee, K., Speir, J. A., Duda, R. L., Hendrix, R. W., Komives, E. A. & Johnson, J. E.. An unexpected twist in viral capsid maturation. *Nature* **458**, 646-50 (2009).

Gertsman I, Duda RL, Hendrix R, Komives EA, Johnson JE. Resolving the molecular reorganization and refolding of the maturing HK97 viral capsid (sending to *Journal of Molecular Biology*).

Gertsman I, Huang R, Johnson, JE. Development of a crystallographic screen for macromolecular complexes (in process as a communication to *Acta Cryst.*).

Abstract of the Dissertation

Structural Characterization of Viral Capsid Maturation in Bacteriophage HK97

Biochemistry

By

Ilya Gertsman

Doctor of Philosophy

University of California, San Diego, 2009

Professor John E. Johnson, Chair
Professor Elizabeth A. Komives, Co-chair

Most complex viruses undergo a capsid maturation process from a precursor, or procapsid particle form to a more stable, infectious particle. dsDNA bacteriophage in particular, exhibit enormous conformational changes in their capsids, as hundreds of protein subunits must undergo a coordinated rearrangement that enables the particle to package and withstand high pressures of DNA. The reorganization in these phage often includes expansion of the capsid diameter, and significant thinning of the capsid shell, facilitated by subunit rotations. *In vitro* maturation experiments have shown that for many systems, the process can be mimicked without DNA packaging, instead using heat or perturbation agents to stimulate maturation, characterized as a highly exothermic event. Unlike most motor proteins such as ATP synthase, myosin, and many others, the process is uncoupled to hydrolysis of NTP's, instead using a trapped assembly intermediate as the driving force for the process. The structural mechanism for this behavior has been unknown for systems that undergo large-scale rearrangements such as

dsDNA bacteriophage. We have been able to crystallize the pre-expanded particle form of bacteriophage HK97 at near atomic resolution and have been able to compare it to a previously solved structure of the mature HK97 capsid, enabling a first look at the structural mechanism associated with HK97 capsid maturation. The study has enabled us to identify the Prohead II state as an intermediate folded state that undergoes refolding and tertiary twisting motions to accommodate particle maturation, in contrast to the previously posited theories implicating rigid body movements as the method of transition. We have also used H/²H exchange coupled to MALDI mass spectrometry to follow the refolding transitions during expansion, as subunits hinge around newly identified 3-fold interactions that remain fixed throughout expansion, stabilizing inter-capsomer contacts and leveraging the transition. The study provides insight into a mechanism of expansion that consists of local high energy conformations maintained by quaternary interactions in the procapsid state, that when perturbed to expand, undergo an irreversible refolding transition that stabilizes the particle.

Chapter 1

Introduction To Viral Capsid Maturation

Viral capsids are genomic containers that function in the protection and delivery of the viral genome to its respective host. Viruses have evolved various methods of capturing their genome within their capsid shell, either by individual capsid proteins assembling around the genome as in most plant and animal viruses, genome packaging through a unique capsid vertex in an NTP driven process as found in dsDNA bacteriophage and some animal viruses, or through multiple symmetrically related openings accessible in the initially assembled or procapsid state as seen in bacteriophage phix174 and phi6 [1-3] (figure 1.1). Delivery of the viral genome into the host is often related to the method of packaging. For example, dsDNA bacteriophage that package highly pressurized DNA through a single vertex, utilize the same vertex to deliver the genome into the host cell membrane. Viral capsids that assemble around their corresponding genomes often undergo no further rearrangement of capsid proteins while in other cases local conformational changes are required in order to achieve the infectious particle form [4-6]. On the other hand, icosahedral bacteriophage and viruses that package their genome into preassembled particles often undergo large scale rearrangements to accommodate a genome that is often packed to near crystalline density [2,7]. Regardless of the magnitude of capsid reorganization, the process is an essential step in the maturation process of most complex viruses. Figure 1.2 illustrates several virus systems in which the morphologies associated with maturation have been characterized.

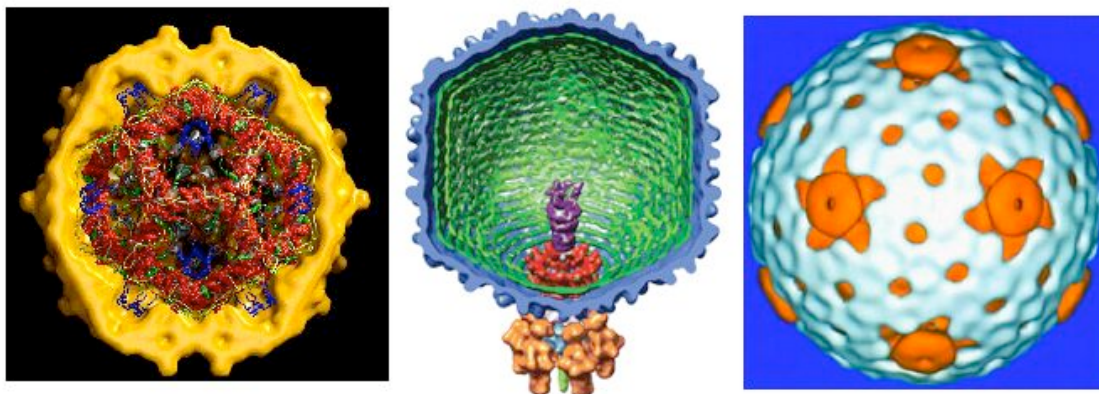


Figure 1.1: **Genome packaging mechanisms.** (A) A cutaway view of Flock House Virus with its RNA shown in the center, structured as a dodecahedral cage. The Flock House Virus capsid assembles around its RNA genome. (B) P22, a dsDNA bacteriophage, packages its genome into a pre-assembled particle through a 5-fold portal complex. Concentric rings of packaged dsDNA are shown in green. (C) Phi6, a dsRNA bacteriophage, is thought to package its genome through symmetrically related openings around the icosahedral capsid. The above figures are adapted from references 1-3 in the main text.

Structural data of capsids that package their genome has shown that the supra-molecular protein complex that forms the capsid is actually a molecular machine of its own, that undergoes enormous conformational rearrangements including refolding, rotational and translation motions, and in some cases results in over 100Å increase in particle diameter and an increase in the particle volume by over 2-fold the capacity of the precursor state [8]. The structural and biochemical study of capsids is therefore not only relevant to the understanding of viruses, but has a general application to the understanding of macromolecular assembly and dynamics associated with complex molecular machines. Systems that have been of particular interest in structural biology for both their complex assembly states and ability to coordinate large scale conformational changes, include but are not limited to the ribosome, ATP synthase, proteasome, GroEL, RNA polymerase, as well motor proteins such as dynein, myosin, and kinesin.

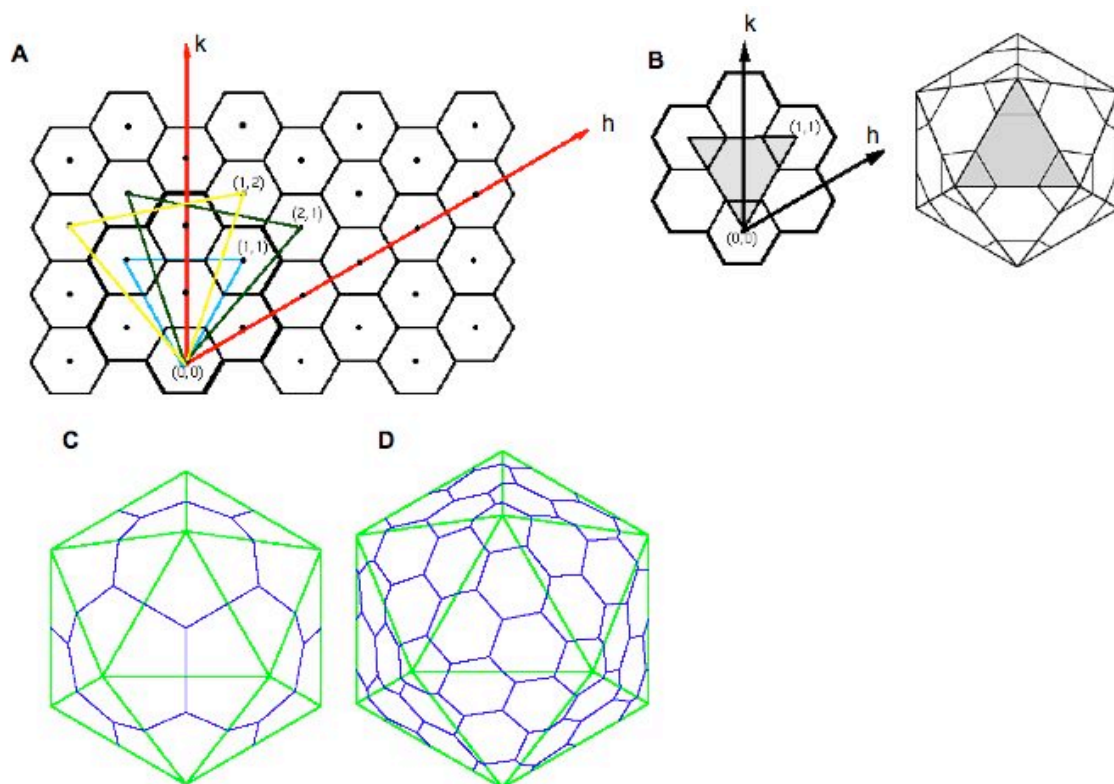
The conformational changes of many phage can be induced without genome packaging, instead using various chemical or thermal perturbations to induce capsid maturation. The motions concomitant with maturation of such viral capsids are therefore often uncoupled from energy sources such as NTP's, relying instead on the assembly and structure of the immature particle form as the energy storage for conformational rearrangement. The self-contained nature of this molecular machine combined with the enormity of their protein complexes, which has been quantified in the thousands of subunits for several viruses [44], makes the viral capsid a unique and important subject for protein studies. *In vitro* capsid maturation techniques has allowed structural and thermodynamic comparisons of the immature state, termed the procapsid, with the mature

capsid state, using a number of biophysical tools including x-ray crystallography, Cryo-EM, calorimetry, raman spectroscopy, Small angle x-ray scattering, and Hydrogen/deuterium exchange to name a few. The high level of symmetry associated with icosahedral viruses enables higher resolution Electron Microscopy studies and improves data quality in crystallographic studies as well [9]. A brief description of quasi-equivalence, triangulation number and icosahedral symmetry, initially described by Casper and Klug, is explained in figure 1.2 [10-11]. In the cases of several viral systems including Herpes, HK97, and phi 6, intermediate states have been identified which have distinct conformations from their procapsid and mature capsid counterparts [12-14]. The intermediate particle forms have been implicated both biochemically and structurally to be on-pathway expansion intermediates that are often temporarily trapped in particle forms too transient for detection *in vivo*. The previous structural studies detailing capsid maturation in several well studied viral systems, including HK97, are summarized below.

Sequential packaging of genome segments accompanied by conformational changes in an RNA bacteriophage

Phi 6 is an enveloped virus that packages a segmented RNA genome. Three segments of dsRNA are packaged sequentially into the polymerase complex which consists of a T=1 protein shell composed of 60 asymmetric P1 dimers (A and B) as well P2, P4, and P7 proteins. P2 is the RNA-dependant polymerase, P4 is the NTPase which serves as the packaging motor and sits at the PC vertices, and P7 acts as a packaging factor. An outer protein shell (NC) with an icosahedral T=13 lattice surrounds the inner core (PC) and is composed of mostly P8 protein [3,14]. Surrounding the P8 layer, or

Figure 1.2: **Viral quasi-equivalence.** Icosahedral viruses are composed of symmetrically related building blocks. The icosahedron can be created from a hexagonal lattice that is shown in figure a. The icosahedron is composed of 20 identical triangular faces which can vary in size based on the number of hexagonal points included in the face. The formula: $T=h^2+hk+k^2$, calculates the T number of a virus particle based on the number of lattice points used in the h and k directions. (B) The construction of a T=3 capsid particle. (C) A T=1 particle, the simplest Icosahedral particle, composed of 60 identical copies of a single subunit. The particle is composed entirely of pentamers of subunits (12 of them in all). (D) A T=7 laevo particle, the same triangulation number as HK97. The particle is composed of 7 asymmetric subunits. When applying the 60 icosahedral matrices to number of asymmetric subunits, one generates the entire capsid particle. Casper and Klug originally described quasi-equivalence for viruses that had a larger triangulation number than 1. In those cases, the T-number is equal to the number of asymmetric subunits in the icosahedron, each with potential variations, therefore referred to as quasi-equivalence. Illustrations above were taken from the virus database: <http://viperdb.scripps.edu>.



nucleocapsid, is the viral envelope consisting of a lipid membrane as well as integral proteins. Though the P8 shell and envelope are necessary for the formation of the infectious particle, the polymerase complex (PC) alone is sufficient for genome packaging and can be studied independently of the other components [15]. A long standing question had been how the complex is able to package a specific form of dsRNA in a sequential fashion, with the S type packaged first followed by the M and then L type. Cryo-EM reconstructions of various conformational states of Phi 6 have revealed a large scale expansion process that drastically alters the conformations of P1 dimers of the PC shell, therefore also altering the potential binding sites for RNA recognition and packaging [3,14] (figure 1.3). Though high-resolution structures of the intermediate complexes have not been determined, a recent 7.5 Å Cryo-EM structure of the expanded particle has allowed secondary structure determination of the P1 protein domains and visualization of the alteration of contacts between P1 subunits during expansion. The expansion from a 460Å particle to a 500Å particle appears mainly facilitated by rigid body rotations (between 35 and 46 degrees) that results in a nearly 2.4 fold increase in the particle's volume. The conformational flexibility seen from Cryo-EM studies has helped to explain the particles capability of packaging specific dsRNA segments in a pre-determined sequence. It has been shown that the three types of RNA's (S, M, and L) packaged have different *pac* sites approximately 50 bases from the 5' end of the RNA that recognize the P1 subunits of the core [16]. An empty particle can only package an S strand of RNA, followed by a conformational switch that exposes a binding site specifically for the M strand. Only upon packaging of the M strand does another

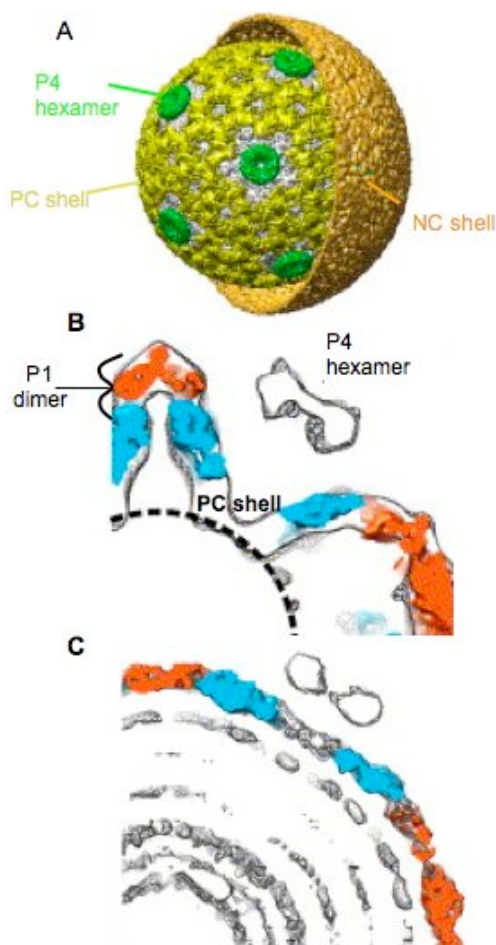


Figure 1.3: **Phi6 capsid maturation.** (A) The phi6 Cryo-EM density is shown with its corresponding NC shell and PC shell colored in orange and yellow respectively. The poorly ordered P4 hexamers are colored in green. The unexpanded PC shell (B) and expanded shell (C) depict the major conformational changes occurring in the P1 protein interfaces following RNA packaging. Panel A has been adapted from reference 41, while panels B and C are from reference 14.

conformational switch lead to the recognition and packaging of the L type RNA. Large scale rotations of both A and B P1 monomers explains the potentially different binding surfaces presented by the particle during expansion. Interestingly, the size of the packaged segment is of great importance in activating the switch that changes binding surfaces. When shortened S type RNA fragments were packaged, multiple copies of the fragments were needed to induce the transition to the M type recognition conformation [17]. The data indicates that the volume of the packaged segment is related to the mechanism of the conformational change, not simply the binding or packaging event. The expansion intermediates appear to fall into local energy minima that require the packaging of additional RNA to overcome the energetic barriers necessary for further conformational transitions.

HK97, a model system for understanding dsDNA maturation in bacteriophage

Studying detailed conformational changes of viral capsids has been a challenging obstacle requiring isolation of distinct intermediate states followed by moderate to high-resolution structural analysis. An atomic model of the capsid subunit is especially important in the determination of refolding and hinging events as well as changes in interactions that facilitate the global stability of the capsid during morphological transitions. The structure of the HK97 mature capsid was previously solved to 3.44 Å resolution [18-19] (figure 1.6), providing a first glimpse into a new subunit fold, that has recently been implicated in many other phage and even Herpes viruses following Cryo-EM studies of these systems [20]. The ability to express HK97 VLP's (virus like particles), mature them *in vitro*, and isolate expansion intermediates as described below,

has made HK97 an ideal model system in which to study the structural changes associated with capsid maturation.

HK97 is a tailed lambda like bacteriophage that infects *E.coli*. HK97, like other dsDNA phage, assembles into a precursor or procapsid form before packaging its genome through one unique vertex of an icosahedral shell. The capsid shell is composed of 415 identical copies of the gp5 subunit that assemble into capsomers of hexamers and pentamers before organizing into the full particle [21]. Unlike many other phage though, HK97 does not employ scaffolding proteins or accessory proteins to assemble or stabilize the capsid. The first 102 residues of each gp5 subunit, called the Δ -domain is thought to perform the scaffolding function. The 415 copies of gp5 subunits assemble along with 12 copies of portal protein (gp3), and ~60 copies of protease into a precursor particle [22]. The portal complex sits at one 5-fold vertex of the capsid and functions in genome packaging. The internally packaged protease gp4 cleaves the 102-residue Δ -domain fragment from each gp5 subunit. The protease also cleaves itself and exits the procapsid along with fragments of the Δ -domain, leaving behind the Prohead II particle. *In vivo*, particle maturation is induced from the Prohead II state after terminase protein binds to and packages dsDNA through the portal complex in an ATP driven process (figure 1.4). Alternatively, the maturation process can be mimicked *in vitro* by assembling the procapsid particle without portal protein, resulting in the substitution of the packaging vertex with another pentamer of gp5 subunits. Following Δ -domain cleavage and attainment of the Prohead II state, the particle can be matured *in vitro* using a variety of chemical and low pH induced perturbations [23] (figure 1.5). Most of the characterization of intermediate HK97 particles has followed expansion using acidic pH

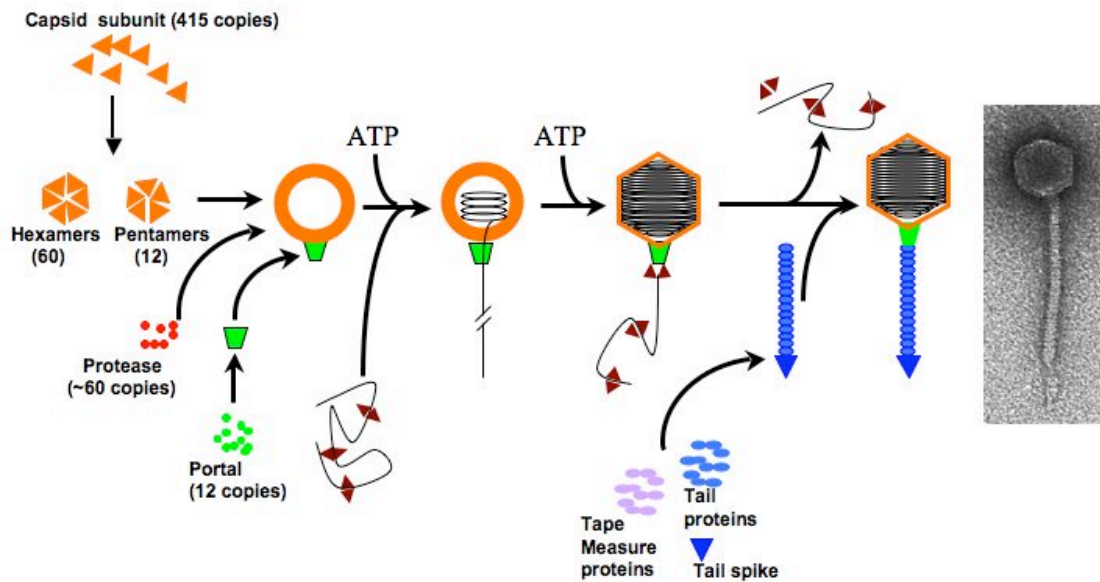


Figure 1.4: **In vivo assembly of HK97.** Capsid subunits assemble around protease proteins and portal complex to form the procapsid particle. DNA is packaged in an ATP dependant pathway resulting in particle expansion from a spherical shell to the icosahedral form. An EM micrograph of a native HK97 bacteriophage is shown at the right of the schematic.

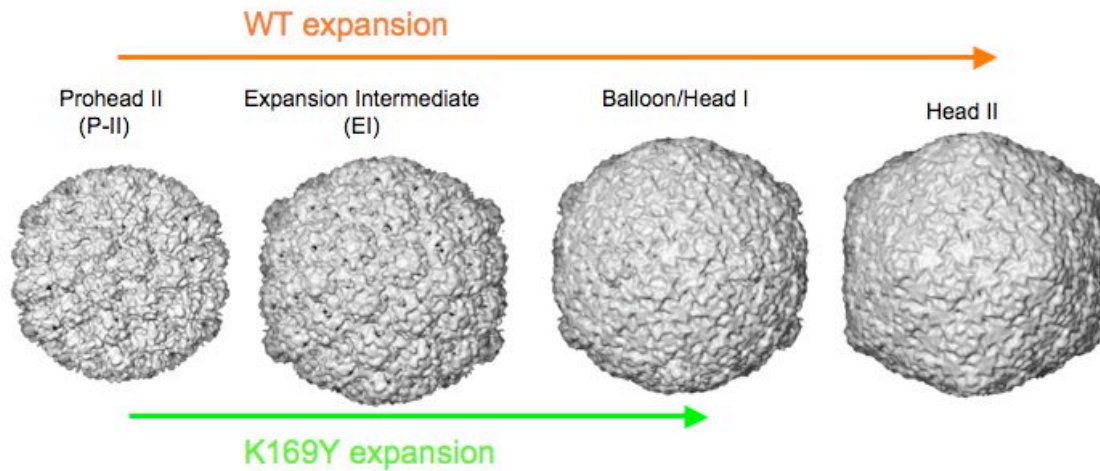


Figure 1.5: ***In vitro* HK97 expansion.** Electron density maps of distinct intermediate characterized by EM (P-II and EI), and crystallography (Balloon, and Head II).

Intermediate particles were attained *in vitro* by low pH treatment of Prohead II particles, assembled in an *E.coli.* expression system. K169Y mutant is crosslink defective causing the expansion process to end at the Head I state. Head I is the morphological equivalent of WT Balloon particles, except no crosslinks are present.

conditions, most commonly 50mM Na-Acetate, pH 4.0, with 200mM potassium chloride [13,24-27].

In vitro maturation of the Prohead II induces subunit reorganization and expansion of the particle nearly 60 Å to the first expansion intermediate state, EI. Previous Cryo-EM studies show that the surface of the particle transforms from one that is corrugated and has a skewed arrangement of its hexamers in P-II, to a more symmetrized form in EI [28] (figure 1.8A). This first expansion event is a highly cooperative two-state transition [26], shown by single molecule expansion measurements to occur in less than half of a second (data unpublished), the time resolution of the measurement. Following the first major transition, the E-loops of subunits are able to sample a conformation that allows an auto-catalyzed isopeptide bond to form between Lysine 169 of one subunit with an Asparagine 356 of another subunit located on an adjacent capsomers [18,23]. The linkage has been shown to be catalyzed by a neighboring Glu 363 residue [29] (figure 1.6). Lys 169 is located near the tip of an extended β -hairpin or E-loop, while the other two residues are at subunit interfaces near local three-fold axes of symmetry on particles. Increasing numbers of cross-links accumulate, reaching a threshold that ultimately ratchets the expansion intermediate into further expansion, resulting in the EI-IV state [28]. One final morphological transition occurs as EI-IV expands to the mature Head II state [27]. Crystal structures of EI-IV and Head II were previously solved to 3.44 Å and 3.8 Å respectively [19,27], a study that showed the final expansion step is attributed to conformational sampling of penton subunits that resulted in a final class of crosslink formation and further protrusion of pentons. Hexon conformations were shown to exist in a nearly mature form by the EI-IV

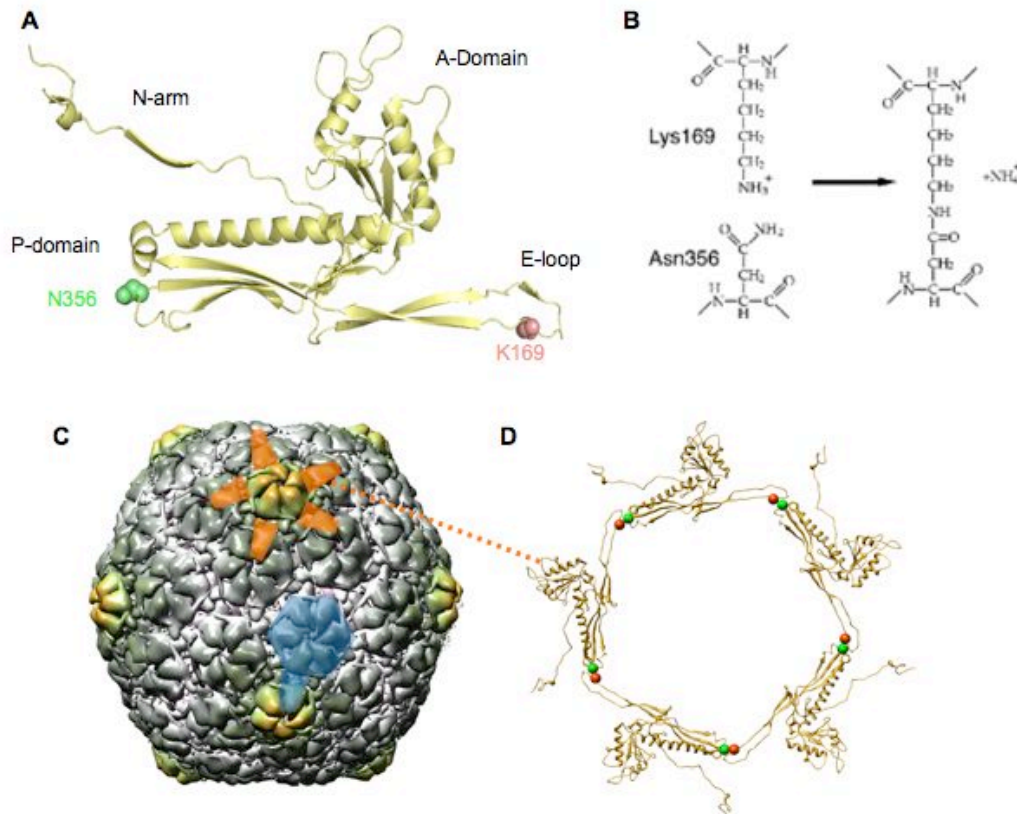


Figure 1.6: (A) Mature subunit of HK97. Residues N356 and K169, which are involved in covalent crosslinking, are shown as colored spheres. (B) Isopeptide bond formation, catalyzed by a nearby glutamate at residue 363. (C) Isosurface rendering of Head II structure. The region highlighted in blue represents the asymmetric subunit of the T=7 capsid, which includes the hexamer subunits and a penton subunit. The full capsid can be reproduced by applying the 60 icosahedral matrices to these 7 subunits. The hexon subunits that form crosslinks around the pentamers are highlighted in orange and shown in (D).

stage. The crosslinking through isopeptide bonds ultimately results in a molecular topology that is similar to that of chainmail.

Previous studies have used a moderate resolution Cryo-EM structure of P-II, at 12 Å [8], to compare with the near atomic resolution late intermediate states as a way to understand the structural changes associated with expansion. The studies had implicated rigid body movements as playing the major role in the reorganization of subunits during the process and revealed an enormous increase in inter-protein interactions in addition to crosslink formation as the mode of stabilizing the mature capsid form. In the Prohead II state, the individual subunits appear in a more radial orientation, resulting in a thicker capsid shell (50 Å thick), while the subunits in the Head II state have a flattened orientation (18 Å thick), but are much more heavily inter-digitated with neighboring subunits. In Head II, one subunit makes contact with up to nine other subunits. It is obvious from the heavily intertwined topology of Head II that such interactions cannot result from a single assembly event, which demonstrates the necessity for an intermediate or procapsid particle. It is from this initial particle form, that further reorganization can lead to increased capsid stability. Calorimetric studies have shown that crosslink formation leads to an enhanced thermal stability for the fully mature Head II structure as compared to the pre-expanded Prohead II particle [30]. No such chainmail morphology has been reported for other families of bacteriophage or viruses, which often use accessory proteins, expressed as separate gene products, that decorated and stabilize the outer capsid shell.

Though crosslink formation is clearly the method of stabilization in HK97, a crosslink defective mutant, K169Y, is able to undergo *in vitro* expansion as well. K169Y

can expand to a form similar to EI-IV, termed Head I, but cannot attain the mature Head II morphology, clearly showing that cross-linking is crucial for overcoming the energetic barrier associated with the final expansion event. Complementation assays used to evaluate particle assembly *in vivo*, have shown that the K169Y mutant does not form particles that can package DNA [30], likely succumbing to the extreme pressures (measured at over 50 atm. of pressure in a similar phage, phi29 [31]) of packaging a 40kb genome without crosslinks to stabilize the capsid .

Capsid maturation in other dsDNA bacteriophage and Herpes Virus

Though a structure of sufficient resolution has not been determined for other dsDNA bacteriophage or viruses, the recent sub-nanometer Cryo-EM studies of various capsid forms has revealed a possible evolutionary link to HK97. Indeed, well studied phage such as P22, T4, T7, phi29, lambda, and even Herpes virus, all share a common subunit fold as HK97 (figure 1.7). The 3.44 Å structure of the mature HK97 form revealed a novel fold at the time of its discovery, that is now considered to likely be the most common fold found in nature, considering that phage are the most abundant organism on the planet (estimated at over 10^{31}). In addition to the Cryo-EM studies, a crystal structure for the T4 penton subunit alone, gp24, has revealed a very similar fold to HK97 [32] (figure 1.7D). The mature HK97 subunit has been modeled into the various Cryo-EM density maps from the aforementioned studies, with relatively good fit, though differences are clearly seen in various domains of the fold. The differences are expected since there is considerable variation in the sequence lengths of the capsid subunits in the

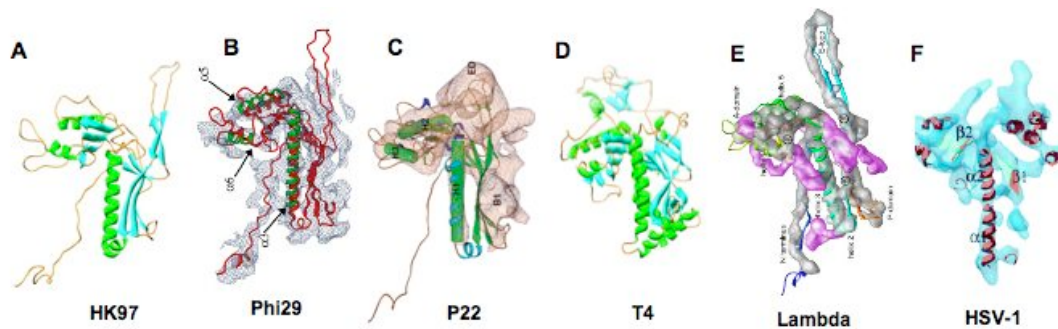


Figure 1.7: **HK97, a common fold in dsDNA bacteriophage and Herpes Virus.** (a) Hk97 subunit. (B) HK97 subunit fit into phi29 Cryo-EM reconstruction. (C) HK97 subunit fit into P22 Cryo-EM density. (D) Gp24 subunit (pentamer subunit) of T4. (E) HK97 fit into Cryo-EM reconstruction of Lambda. (F) Secondary structure elements fit into Cryo-EM density of Herpes virus (HSV-1). The figures above were adapted from references 2,20,34, and 42.

various systems. Capsid subunits range from 31kD in mature HK97 gp5, to 47kD and 50kD in P22 and Phi29 respectively. Though the fold seems strikingly similar, there is less than 15% sequence homology between these bacteriophage [21].

A combination of Raman spectroscopy, Cryo-EM studies, and Hydrogen/Deuterium (H^2/H) exchange (coupled to mass spectrometry) has characterized the maturation of P22 particles [33-36]. The studies showed local refolding events occurring in various domains of the P22 gp5 subunits. The H^2/H exchange study indicated that the bulk of the changes in solvent protection were attributed to the N-terminal half of the subunit, while C-terminal region remained essentially unchanged [35]. The study supported the Cryo-EM data of P22, which showed that the H0 helix appeared to partially unfold during maturation [34]. After fitting the HK97 subunit to the P22 Cryo-EM density, the H0 helix was implicated as being the equivalent of the N-arm in HK97, which also refolds during maturation (figure 1.7C). The H3 region of the subunit was implicated in undergoing a tilting motion, attributed to the symmetrization and increased interactions seen in the center of hexamers as they matured from the skewed symmetry seen in the procapsid. The conclusion from these studies indicated that P22 capsid subunits have considerable flexibility in various domains in addition to the rigid body movements implicated in HK97 as the basis for conformational changes during maturation. Less extensive structural studies have also shown that T=7 bacteriophage Lambda and bacteriophage T7 also display skewed hexamer subunits in their procapsid particles, which undergo symmetrization during conformational rearrangements [37-38]. The breakdown in hexameric symmetry in the immature particle forms appears to be a common theme in many of tailed bacteriophage, and may be

indicative of the tenuous arrangements present in the procapsid structures. Since all of these bacteriophage require a separate scaffolding protein to assist in particle assembly, or in this case of HK97, a scaffolding domain fused to the coat protein subunit, the 2-fold instead of 6-fold symmetry could be a direct result of the interactions mediated between scaffolding protein and coat protein. An exception to the alteration of hexameric symmetry during dsDNA bacteriophage maturation is phi29 [39]. Cryo-EM comparisons of prohead and mature heads of phi29 (figure 1.8D) reveal no difference in symmetry resulting from maturation. The reason for such an outlier may be due to the elongated morphology of phi29, which forms a T=3, prolate shell. The prolate shell, unlike the true icosahedral form is elongated longitudinally due to additional hexamers assembling in the equatorial region of the capsid. The addition of “extra” hexamers to the icosahedral lattice results in a non-uniform conformation between the different hexamers, with those furthest from the tail exhibiting little to no recognizable skew, while the others maintain some distortion [39]. Differences in assembly between the prolate and icosahedral systems may result in alterations in the morphological changes accompanying maturation, thereby diminishing the changes in symmetry seen in studies of the latter. The smaller capsid lattice of the phi29 particles (T=3 as opposed to T=7) may also result in the reduced distortion, which may be correlated to a diminished conformational rearrangement and expansion. Further studies of distinct maturational intermediates of other bacteriophage and viral systems will be required to fully answer whether a breakdown in symmetry is related both to the shape and metastability of procapsids.

In addition to dsDNA bacteriophage, the animal virus Herpes (HSV-1) also shares an HK97 fold (figure 1.7F). HSV-1 capsid is considerably larger than HK97, with a T=16

icosahedral lattice, and composed of two other proteins, VP19c and VP23, besides the major coat protein, VP5, that forms hexamers and pentamers (figure 1.8). These last two gene products constitute the triplex proteins that surround the hexamers and pentamers and are considered to provide a 3-fold symmetric stabilization in the capsid lattice.

Though there are differences between the two, systems, Cryo-EM studies of HSV-1 have shown that floor domain of the 150kD VP5 subunit has the HK97 fold [20].

Reconstructions of various maturation intermediates of HSV-1 show that the floor domains are the main source of intra-capsomer interactions in the Procapsid state, and that they undergo rotations of ~ 40 degrees during maturation, resulting in new associations of the floor-domains with their neighboring capsomers (i.e. inter-capsomer contacts) [12]. Interestingly, the extent of rotation of this domain is similar to what was documented for HK97 capsid subunit rotation [8]. Though an animal virus, HSV-1 shares another common feature with many bacteriophage, it packages a dsDNA genome in a pre-assembled particle, resulting in significant capsid reorganization as described above. The similarities that have been documented point to an evolutionary link between HSV-1 and dsDNA bacteriophage. Though HSV-1 infects a completely different domain of life, a common ancestor may exist between these viral systems before further evolution events lead to divergences in host selectivity.

Understanding the structural mechanism in HK97 capsid maturation

Though Cryo-EM and other biophysical tools have made major break-throughs in characterizing the structural changes associated with capsid maturation, there are many

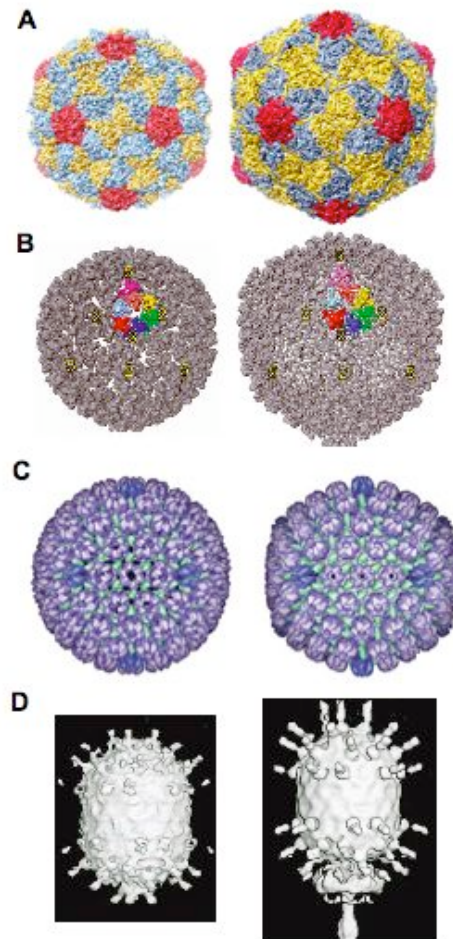


Figure 1.8: **Capsid maturation.** The particles on the left represent procapsid states of HK97, P22, HSV-1, and phi29. The particles on the right are the matured capsid forms. (A) HK97 rendered from crystallographic coordinates. (B) Cryo-EM reconstructions of P22. (C) Cryo-EM reconstructions of HSV-1. (D) Cryo-EM reconstructions of phi29 prolate procapsid and mature capsid particles. Panels b-d were adapted from references 34,12, and 43 respectively.

unanswered questions remaining, one of which is the structural basis for the metastability and irreversibility of the maturation process. For systems in which major changes in capsid conformation occur, such as HK97, only the mature form of the capsid has been solved to high resolution. There is no near atomic resolution structure of a procapsid form to compare to the mature capsid and to distinguish the detailed structural refolding and other transitions that occur during this process. In fact, there is only one example of a near atomic resolution structure for any viral procapsid (including all eukaryotic, prokaryotic and archaea infecting viruses), this being the ssDNA bacteriophage phiX174 [40]. PhiX174 undergoes a much less dramatic rearrangement than most dsDNA bacteriophage and the procapsid structure solved was considered to be an off-pathway maturation intermediate [40]. A major goal of this thesis was to solve the structure of the Pre-expanded particle form of HK97, Prohead II, enabling the structural characterization of an intermediate assembled viral capsid, and its relation to the mature intermediate forms of HK97 previously solved to better than 4Å resolution. In addition to characterizing the static pre-expanded and expanded states, we studied the dynamics of the intermediate transitions of the maturation process using H/²H exchange coupled to mass spectrometry.

The following thesis presents the crystallization, and structure determination of Prohead II at 3.65 Å resolution. In addition, Hydrogen deuterium exchange data, which measures the solvent accessibility of amide protons in response to secondary, tertiary, and quaternary structure, was used for the native characterization of intermediate forms of HK97 maturation. In combination with the crystallographic study, we have discovered a structural mechanism for HK97 expansion that dispels the previous hypothesis of largely

rigid body rotations and translations as the basis for the conformational changes, previously identified with lower resolution Cryo-EM data. The new data demonstrates a tertiary twisting of each subunit with respect to the newly identified 3-fold interaction sites, that remain fixed during expansion. Each subunit is involved in this binding interface at all quasi and icosahedral 3-fold axes. Concomitant with the twisting is a refolding of the major helix of each subunit as well as local refolding in all major domains of the subunit. Most importantly, the structure reveals an intermediately folded structure that appears to be the energy source for capsid reorganization. We further explored the interaction interfaces that change or remain constant during maturation using a combination of mutagenesis and biophysical techniques, providing a structural basis for capsid maturation and a model for how the structural reorganization is propagated through the capsid lattice. Though this mechanism is the first to be identified in viral capsids, it may in fact be shared in a variety of bacteriophage and possibly even animal viruses that share the HK97 fold. Future characterizations of other viral systems at similar or higher resolutions will ultimately be required to determine the extent in which such systems are related.

References

1. Tihova, M., Dryden, K. A., Le, T. V., Harvey, S. C., Johnson, J. E., Yeager, M. & Schneemann, A. (2004). Nodavirus coat protein imposes dodecahedral RNA structure independent of nucleotide sequence and length. *J Virol* 78, 2897-905.
2. Lander, G. C., Tang, L., Casjens, S. R., Gilcrease, E. B., Prevelige, P., Poliakov, A., Potter, C. S., Carragher, B. & Johnson, J. E. (2006). The structure of an infectious P22 virion shows the signal for headful DNA packaging. *Science* 312, 1791-5.
3. Butcher, S. J., Dokland, T., Ojala, P. M., Bamford, D. H. & Fuller, S. D. (1997). Intermediates in the assembly pathway of the double-stranded RNA virus phi6. *Embo J* 16, 4477-87.
4. Canady, M. A., Tihova, M., Hanzlik, T. N., Johnson, J. E. & Yeager, M. (2000). Large conformational changes in the maturation of a simple RNA virus, nudaurelia capensis omega virus (NomegaV). *J Mol Biol* 299, 573-84.
5. Allison, S. L., Schalich, J., Stiasny, K., Mandl, C. W., Kunz, C. & Heinz, F. X. (1995). Oligomeric rearrangement of tick-borne encephalitis virus envelope proteins induced by an acidic pH. *J Virol* 69, 695-700.
6. Salanueva, I. J., Carrascosa, J. L. & Risco, C. (1999). Structural maturation of the transmissible gastroenteritis coronavirus. *J Virol* 73, 7952-64.
7. Jardine, P. J., McCormick, M. C., Lutze-Wallace, C. & Coombs, D. H. (1998). The bacteriophage T4 DNA packaging apparatus targets the unexpanded prohead. *J Mol Biol* 284, 647-59.
8. Conway, J. F., Wikoff, W. R., Cheng, N., Duda, R. L., Hendrix, R. W., Johnson, J. E. & Steven, A. C. (2001). Virus maturation involving large subunit rotations and local refolding. *Science* 292, 744-8.
9. Lee, K. K. & Johnson, J. E. (2003). Complementary approaches to structure determination of icosahedral viruses. *Curr Opin Struct Biol* 13, 558-69.
10. Caspar, D. L. & Klug, A. (1962). Physical principles in the construction of regular viruses. *Cold Spring Harb Symp Quant Biol* 27, 1-24.
11. Caspar, D. L. (1980). Movement and self-control in protein assemblies. Quasi-equivalence revisited. *Biophys J* 32, 103-38.

12. Heymann, J. B., Cheng, N., Newcomb, W. W., Trus, B. L., Brown, J. C. & Steven, A. C. (2003). Dynamics of herpes simplex virus capsid maturation visualized by time-lapse cryo-electron microscopy. *Nat Struct Biol* **10**, 334-41.
13. Lata, R., Conway, J. F., Cheng, N., Duda, R. L., Hendrix, R. W., Wikoff, W. R., Johnson, J. E., Tsuruta, H. & Steven, A. C. (2000). Maturation dynamics of a viral capsid: visualization of transitional intermediate states. *Cell* **100**, 253-63.
14. Huiskonen, J. T., de Haas, F., Bubeck, D., Bamford, D. H., Fuller, S. D. & Butcher, S. J. (2006). Structure of the bacteriophage phi6 nucleocapsid suggests a mechanism for sequential RNA packaging. *Structure* **14**, 1039-48.
15. Gottlieb, P., Strassman, J., Qiao, X. Y., Frucht, A. & Mindich, L. (1990). In vitro replication, packaging, and transcription of the segmented double-stranded RNA genome of bacteriophage phi 6: studies with procapsids assembled from plasmid-encoded proteins. *J Bacteriol* **172**, 5774-82.
16. Qiao, X., Qiao, J. & Mindich, L. (1997). Stoichiometric packaging of the three genomic segments of double-stranded RNA bacteriophage phi6. *Proc Natl Acad Sci U S A* **94**, 4074-9.
17. Mindich, L., Qiao, X. & Qiao, J. (1995). Packaging of multiple copies of reduced-size genomic segments by bacteriophage phi 6. *Virology* **212**, 213-7.
18. Wikoff, W. R., Liljas, L., Duda, R. L., Tsuruta, H., Hendrix, R. W. & Johnson, J. E. (2000). Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* **289**, 2129-33.
19. Helgstrand, C., Wikoff, W. R., Duda, R. L., Hendrix, R. W., Johnson, J. E. & Liljas, L. (2003). The refined structure of a protein catenane: the HK97 bacteriophage capsid at 3.44 Å resolution. *J Mol Biol* **334**, 885-99.
20. Baker, M. L., Jiang, W., Rixon, F. J. & Chiu, W. (2005). Common ancestry of herpesviruses and tailed DNA bacteriophages. *J Virol* **79**, 14967-70.
21. Xie, Z. & Hendrix, R. W. (1995). Assembly in vitro of bacteriophage HK97 proheads. *J Mol Biol* **253**, 74-85.
22. R. L. Duda, K. Martincic, and R. W. Hendrix. Genetic basis of bacteriophage HK97 prohead assembly. *J Mol Biol* **247**, 636-647 (1995).
23. Duda, R. L., Hempel, J., Michel, H., Shabanowitz, J., Hunt, D. & Hendrix, R. W. (1995). Structural transitions during bacteriophage HK97 head assembly. *J Mol Biol* **247**, 618-35.

24. Lee, K. K., Gan, L., Tsuruta, H., Hendrix, R. W., Duda, R. L. & Johnson, J. E. (2004). Evidence that a local refolding event triggers maturation of HK97 bacteriophage capsid. *J Mol Biol* **340**, 419-33.
25. Gan, L., Conway, J. F., Firek, B. A., Cheng, N., Hendrix, R. W., Steven, A. C., Johnson, J. E. & Duda, R. L. (2004). Control of crosslinking by quaternary structure changes during bacteriophage HK97 maturation. *Mol Cell* **14**, 559-69.
26. Lee, K. K., Tsuruta, H., Hendrix, R. W., Duda, R. L. & Johnson, J. E. (2005). Cooperative reorganization of a 420 subunit virus capsid. *J Mol Biol* **352**, 723-35.
27. Gan, L., Speir, J. A., Conway, J. F., Lander, G., Cheng, N., Firek, B. A., Hendrix, R. W., Duda, R. L., Liljas, L. & Johnson, J. E. (2006). Capsid conformational sampling in HK97 maturation visualized by X-ray crystallography and cryo-EM. *Structure* **14**, 1655-65.
28. Lee, K. K., Gan, L., Tsuruta, H., Moyer, C., Conway, J. F., Duda, R. L., Hendrix, R. W., Steven, A. C. & Johnson, J. E. (2008). Virus capsid expansion driven by the capture of mobile surface loops. *Structure* **16**, 1491-502.
29. Dierkes, L. E., Peebles, C. L., Firek, B. A., Hendrix, R. W. & Duda, R. L. (2009). Mutational Analysis of a Conserved Glutamic Acid Required for Self-Catalyzed Cross-Linking of Bacteriophage HK97 Capsids. *J Virol* **83**, 2088-2098.
30. Ross, P. D., Cheng, N., Conway, J. F., Firek, B. A., Hendrix, R. W., Duda, R. L. & Steven, A. C. (2005). Crosslinking renders bacteriophage HK97 capsid maturation irreversible and effects an essential stabilization. *Embo J* **24**, 1352-63.
31. Smith, D. E., Tans, S. J., Smith, S. B., Grimes, S., Anderson, D. L. & Bustamante, C. (2001). The bacteriophage straight phi29 portal motor can package DNA against a large internal force. *Nature* **413**, 748-52.
32. Fokine, A., Leiman, P. G., Shneider, M. M., Ahvazi, B., Boeshans, K. M., Steven, A. C., Black, L. W., Mesyanzhinov, V. V. & Rossmann, M. G. (2005). Structural and functional similarities between the capsid proteins of bacteriophages T4 and HK97 point to a common ancestry. *Proc Natl Acad Sci U S A* **102**, 7163-8.
33. Tuma, R., Prevelige, P. E., Jr. & Thomas, G. J., Jr. (1998). Mechanism of capsid maturation in a double-stranded DNA virus. *Proc Natl Acad Sci U S A* **95**, 9885-90.
34. Jiang, W., Li, Z., Zhang, Z., Baker, M. L., Prevelige, P. E., Jr. & Chiu, W. (2003). Coat protein fold and maturation transition of bacteriophage P22 seen at subnanometer resolutions. *Nat Struct Biol* **10**, 131-5.

35. Kang, S. & Prevelige, P. E., Jr. (2005). Domain study of bacteriophage p22 coat protein and characterization of the capsid lattice transformation by hydrogen/deuterium exchange. *J Mol Biol* **347**, 935-48.
36. Tuma, R., Coward, L. U., Kirk, M. C., Barnes, S. & Prevelige, P. E., Jr. (2001). Hydrogen-deuterium exchange as a probe of folding and assembly in viral capsids. *J Mol Biol* **306**, 389-96.
37. Dokland, T. & Murialdo, H. (1993). Structural transitions during maturation of bacteriophage lambda capsids. *J Mol Biol* **233**, 682-94.
38. Agirrezabala, X., Velazquez-Muriel, J. A., Gomez-Puertas, P., Scheres, S. H., Carazo, J. M. & Carrascosa, J. L. (2007). Quasi-atomic model of bacteriophage t7 procapsid shell: insights into the structure and evolution of a basic fold. *Structure* **15**, 461-72.
39. Xiang, Y., Morais, M. C., Battisti, A. J., Grimes, S., Jardine, P. J., Anderson, D. L. & Rossmann, M. G. (2006). Structural changes of bacteriophage phi29 upon DNA packaging and release. *Embo J* **25**, 5229-39.
40. Dokland, T., McKenna, R., Ilag, L. L., Bowman, B. R., Incardona, N. L., Fane, B. A. & Rossmann, M. G. (1997). Structure of a viral procapsid with molecular scaffolding. *Nature* **389**, 308-13.
41. Jaalinoja, H. T., Huiskonen, J. T. & Butcher, S. J. (2007). Electron cryomicroscopy comparison of the architectures of the enveloped bacteriophages phi6 and phi8. *Structure* **15**, 157-67.
42. Morais, M. C., Choi, K. H., Koti, J. S., Chipman, P. R., Anderson, D. L. & Rossmann, M. G. (2005). Conservation of the capsid structure in tailed dsDNA bacteriophages: the pseudoatomic structure of phi29. *Mol Cell* **18**, 149-59.
43. Tao, Y., Olson, N. H., Xu, W., Anderson, D. L., Rossmann, M. G. & Baker, T. S. (1998). Assembly of a tailed bacterial virus and its genome release studied in three dimensions. *Cell* **95**, 431-7.
44. Xiao, C., Chipman, P. R., Battisti, A. J., Bowman, V. D., Renesto, P., Raoult, D. & Rossmann, M. G. (2005). Cryo-electron microscopy of the giant Mimivirus. *J Mol Biol* **353**, 493-6.

Chapter 2

X-ray Structure determination of HK97 P-II

Introduction

A crucial component in understanding the structural changes occurring during HK97 maturation is the structure determination of the pre-expanded particle form, P-II. Diffraction data of W336F Prohead II were previously collected and indexed to nearly 5.5 Å by William Wikoff as described previously [1]. A W336F mutation was used for the crystallization, as the mutant was found to slow particle expansion, resulting in more homogenous preparations of Prohead II. Though an electron density map was created, the data were not processed to their maximum potential nor was a model refined into the electron density. We reprocessed the data from the initial indexing stage and refined a 5.2Å structure of Prohead II as described in this chapter. Upon re-dissolving the crystals used for diffraction, it was determined by mass spectrometry that an unintended cleavage had occurred at the E-loop domain of the P-II subunits. A previous 12 Å Cryo-EM study of Prohead II demonstrated that the E loop was disordered and pointing away from the protein shell, indicating that it was not involved in quaternary interactions with neighboring subunits. The cleavage was identified in the crystallographic map, as no density was visible for ~12 residues at the tip of the loop. All the diffracting crystals came from a single preparation in which the unusual proteolytic cleavage occurred. Uncleaved W336F P-II particles resulted in non-diffracting crystals. Efforts were made to reproduce the fortuitous cleavage by performing limited proteolysis of purified P-II

preps. A number of proteases were tried, including trypsin, chymotrypsin, and elastase. The proteolysis did in fact lead to cleavage of the E-loop, though diffracting crystals were not obtained from the experiments. A mutant construct was therefore generated (in collaboration with Robert Duda and Roger Hendrix at Univ. of Pitt.) that truncated the disordered region of the loop and replaced this sequence with a 4 residue sequence known to promote a reverse turn. Extensive screening of crystallization conditions of the W336F/E-loop truncation mutant resulted in significantly improved crystals leading to a 3.65 Å structure of the P-II state. The increase in resolution allowed visualization of side chains in the P-II proteins, resulting in the discovery of key quaternary interactions stabilizing capsomers during expansion, as well comparisons with the mature H-II form that has identified differences in the subunit fold between the two states. The details of the 3.65 Å structure and their biological implications are presented in chapters 4 and 5. The following is a summary of the crystallization, processing, and refinement of the 3.65 Å P-II structure. In addition a brief summary of the 5.2 Å data processing is described, as well as a brief comparison between the P-II structures at the two different resolutions.

Results and Discussion

W336F/E-loop truncation mutation and Crystallization

The truncated mutant was engineered by Robert Duda at the University of Pittsburgh, and created using the splicing by overlap extension technique. The technique uses two internal primers containing the desired mutation, as well as two outside flanking primers to create two new short segments of mutated DNA that can be hybridized,

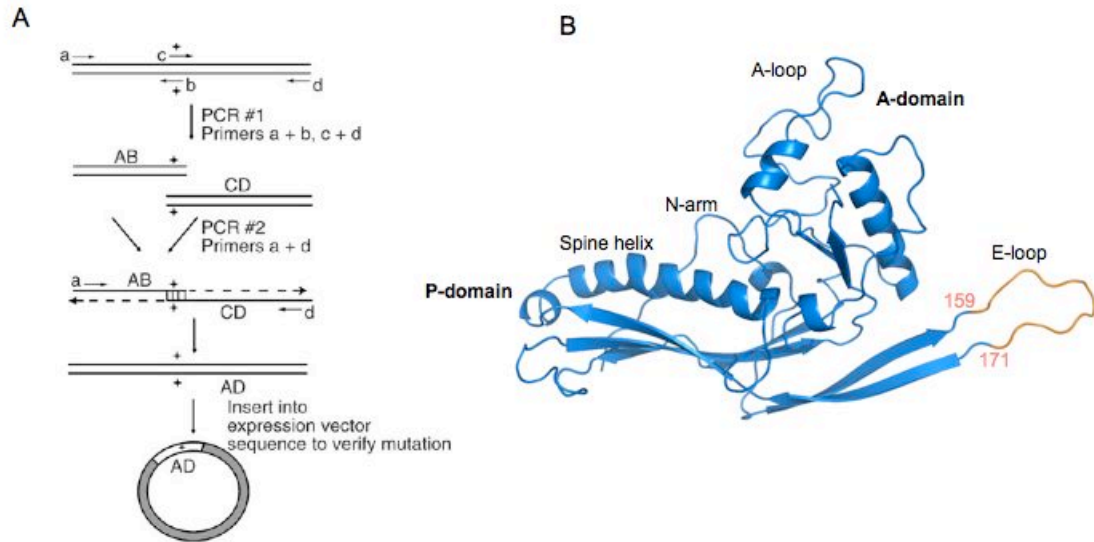


Figure 2.1: **E-loop truncation** (A) Diagram showing multiple steps associated with splicing by overlap extension. Primers coding for the mutation are in the middle (b,c), while flanking primers (a,d) are used to create two intermediate segments of dna with overlapping sequences (AB,CD). After the overlapping fragments hybridize and are completed using the flanking primers, the complete AD fragment with the mutation can be inserted into a vector. B) Prohead subunit D showing the region of the truncated E-loop in orange. This region, spanning residues 159-171 were replaced with an APGD sequence.

extended, and later inserted into a vector of choice (shown in figure 2.1) [2]. The primers were constructed to substitute a four residue sequence, APGD, which is known to promote the formation of a reverse turn [3], for residues 159-171 of the HK97 gp5 sequence. The truncated region of the protein is color coded on the protein subunit in figure 2.1.

The crystallization condition used in the 5.2 Å dataset (100mM Hepes, 200mM CaCl₂, 2-3% Peg 4K) did not produce diffracting crystals with the truncated construct. Various commercial screens were tried, but were unsuccessful, in large part due to their salt and precipitant concentrations, which are intended for crystallization of smaller proteins. Viruses, such as HK97, are among the largest complexes successfully crystallized and therefore require modified mother liquor compositions. The major difference is the concentration of precipitant, which is generally too high in commercial screens. Simple dilution of commercial solutions alters the salt and buffer concentrations, which are often optimal as is. The alternative to commercial screens is the manual creation of mother liquor solutions.

In collaboration with Rick Huang, we developed a comprehensive crystallization screen intended for large protein complexes such as viruses. The screen uses a matrix of combinations of buffers and salt conditions intended to be used in conjunction with further precipitant screens. Table 2.1 lists the individual buffers and salt conditions used in the screen, which were mixed in a combinatorial fashion, resulting in 192 unique conditions. The buffers, salts, and concentrations were based on previous conditions found in the literature. Many of these conditions were compiled from the ViperdB website [4], an online server that contains crystallization conditions for all virus

Table 2.1: **Virus crystal screen.** Each buffer is mixed with each of the salt conditions, resulting in 192 total conditions. The mixtures are all at 2x concentration. For the purposes of screening P-II, the mixtures were diluted to 1x with a final concentration of 3.0% Peg 4K, which was previously shown to induce crystallization.

<i>Buffer</i>	<i>Salts 400mM except for NiCL</i>
1) 200mM Na-Citrate pH 4.0	A) Calcium chloride
2) 200mM Na-Acetate pH 4.6	B) Calcium acetate
3) 200mM Hepes pH 7.5	C) Magnesium chloride
4) 200mM Mops pH 7.5	D) Magnesium acetate
5) 200mM Tris pH 8.0	E) Sodium chloride
6) 200mM MES pH 6.5	F) Potassium chloride
7) 200mM Na-Cacodylate pH 7.0	G) Lithium chloride
8) 100mM Histidine pH 6.0	H) Lithium sulfate
9) 200mM Na-phosphate pH 7.0	I) Ammonium sulfate
10) 200mM Imidazole pH 7.0	J) Magnesium sulfate
11) 200mM Bicine pH 9.0	K) 200mM Nickel chloride
12) 200mM CHES pH 9.0	L) Manganese chloride
	M) Potassium tartrate
	N) Sodium malonate
	O) Sodium formate
	P) Sodium nitrate

structures deposited to the website. The pH ranges of the buffers correspond to the pH range from which nearly all viruses were previously crystallized. Most of the buffers had optimal pH ranges between 6.0 and 8.0, which is the most common protein crystallization pH range [5]. It was determined by separate screening of P-II crystals that 3.0% Polyethylene glycol was an optimal concentration of precipitant to induce crystallization without immediate aggregation. All 192 conditions were used to screen crystallization of the mutant P-II with a final concentration of 3.0% Peg 4K in each condition.

One condition produced large, sharp edged crystals between 0.4 and 1.0 mm in size (figure 2.2). The condition was 100mM CHES pH 9.0, 200mM Manganese chloride, 3.0% Peg 4K. Mixing the components of the mother liquor caused immediate precipitation of much of the manganese. The solution was left to settle overnight, and the colloidal supernatant was used as the mother liquor. Crystals grown in this condition diffracted to 7.5 Å resolution and were in the I222 space group. The condition was optimized using an additive screen in combination with the aforementioned mother liquor solution. Each additive, at 1/10th volume, was added to the mother liquor and protein mixture. A single additive, NDSB-211, improved diffraction. NDSB-211, a zwitterionic compound, is shown in figure 2.2. All crystals were grown by the hanging drop and vapor diffusion method.

Protein crystallography commonly relies on frozen crystals for data collection. In most cases, crystals are either grown in conditions that result in cryogenic protection from freezing damage, or are briefly soaked in such a solution directly prior to freezing. Common cryo-protectants include glycerol, PEG 400, sucrose, hexanediol, and MPD to name a few. Unfortunately nearly all of the commonly used cryo-protectants resulted in

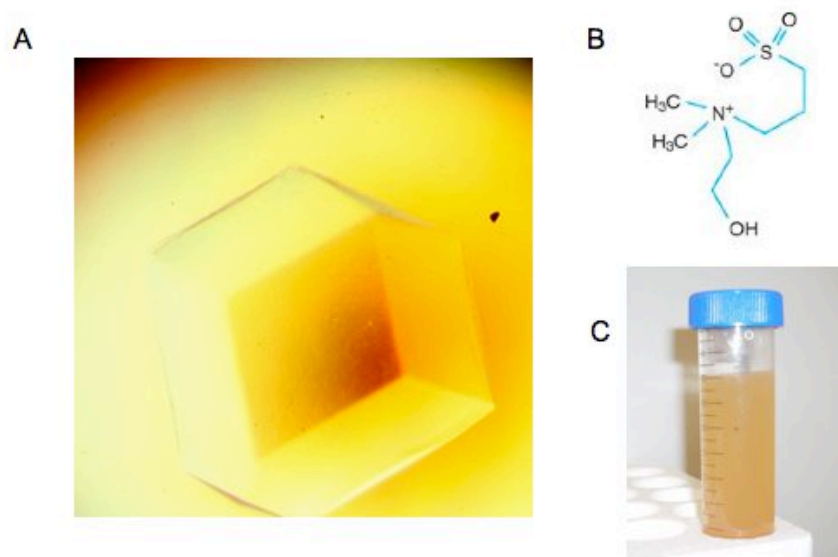


Figure 2.2: **P-II crystallization at 3.65 Å.** (A) A representative P-II crystal (ranging from 0.4-1.0mm in size). (B) NDSB-211, the additive used in crystallization. (C) The mother liquor solution (100mM CHES pH 9.0, 200mM Manganese chloride) before the precipitant has settled.

the maturation expansion of the metastable P-II capsids. Accidental expansion was monitored by agarose gel following incubation with the various cryo-protectants. Glycerol was the only reagent tested that did not induce particle maturation. Crystals either grown in the presence of glycerol, or soaked in a glycerol containing buffer prior to freezing resulted in poorer diffraction than the room temperature crystals. Data therefore had to be collected at $\sim 18^{\circ}\text{C}$. Diffraction data from more than 60 crystals were collected at the Advanced Photon Source on beam lines 14BMC and 23IDD. Crystals diffracted to a maximum of $\sim 3.5\text{ \AA}$. A diffraction pattern is shown in figure 2.3.

Indexing and Scaling

The room temperature crystal data were indexed with the program HKL2000 [6]. Initial indexing and scaling were done with C2 (monoclinic) symmetry, however I222 (orthorhombic) symmetry produced similar processing and scaling statistics indicating the higher symmetry space group. The lattice dimensions were $a=553.0\text{ \AA}$, $b=574.4\text{ \AA}$, $c=587.4\text{ \AA}$, consistent with the expected particle size. Unlike a data set collected from a single frozen crystal, data collected at room temperature require many different crystals, due radiation damage. Any crystals that showed defects in quality or that did not scale well with other crystals were discarded. Between 2 and 15 frames were indexed for each crystal, though automated, continuous, indexing could not be performed in HKL2000 due to an increase in mosaicity between consecutive frames because of prolonged radiation exposure. Individual frames were usually indexed separately and subsequently scaled with other frames. 29 crystals were used in the final scaling, which was performed in SCALEPACK [6].

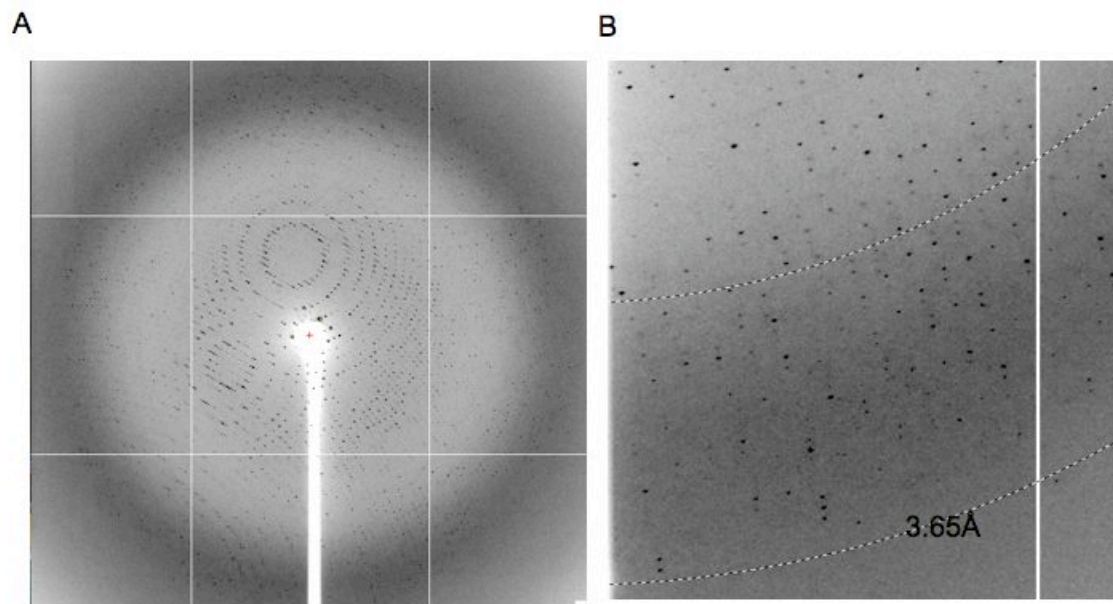


Figure 2.3: **Diffraction pattern of P-II.** (A) Full diffraction image as recorded from the detector and (B) a zoomed-in region corresponding to high resolution data. The image was taken at APS, beamline 14-BMC.

Reflections in the 4-3.5 Å range were weak as is typical for protein crystals, however, the problem is accentuated in crystals with very large unit cells. The exceptionally large unit cells common to crystals of large viruses results in a significantly lower number of unit cells per crystal, which combined with the greater solvent content due to the empty particles, results in less scattering from protein and weak diffraction. This combined with the increase in background in the detector in the 4-3.5 Å range due to the scattering from water results in a decrease in Intensity (I) /sigma (I). Due to these obstacles, using all of the recorded data higher than 4 Å resulted in poor high-resolution data quality. To overcome this, all recorded reflections that had an I/sigma (I) of 0 or lower were discarded during scaling. Though negative I/sigma (I) reflections do in fact contribute to statistically valid data, the large percentage of such reflections dominated the processing statistics. The I/sigma (I) >0 cutoff greatly improved the statistics as shown in figure 2.7. Table 2.7 shows the comparison of statistics of diffraction data processed with and without this cutoff. The R_{merge} from scaling is clearly significantly improved in the higher resolution shell of the data using the cutoff as compared to data without it. The cost of discarding the very weak recorded data is a decrease in data completeness, but due to the high symmetry of the space group and a high non-crystallographic symmetry of 15, the gain in overall I/sigma (I) in the higher resolution range was justified. Plots of intensity vs. resolution and I/sigma (I) vs. resolution are shown in figure 2.4. Using the I/sigma(I)>0 cutoff enabled additional, previously un-scaleable data, to be incorporated into the processing. In addition to scaling the fully measured reflections, partial reflections were summed from successive images and scaled up to whole reflections with SCALA [7].

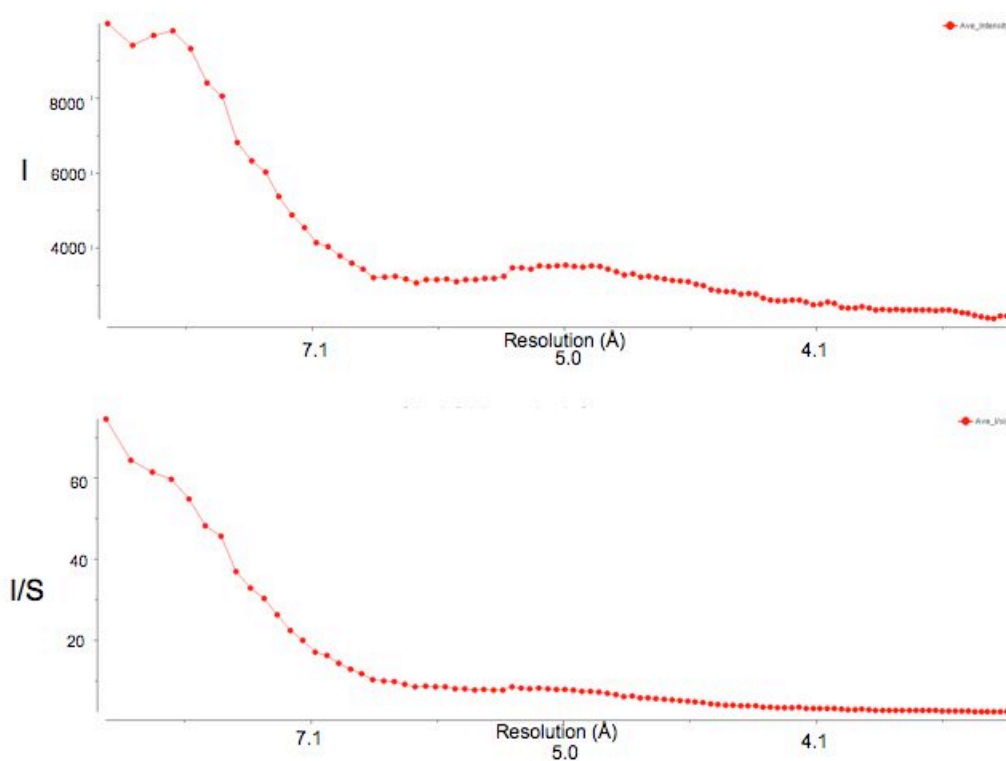


Figure 2.4: **Diffraction intensity.** (A) The change in average intensity of recorded reflections as a function of resolution. (B) Intensity/sigma (I) as a function of resolution.

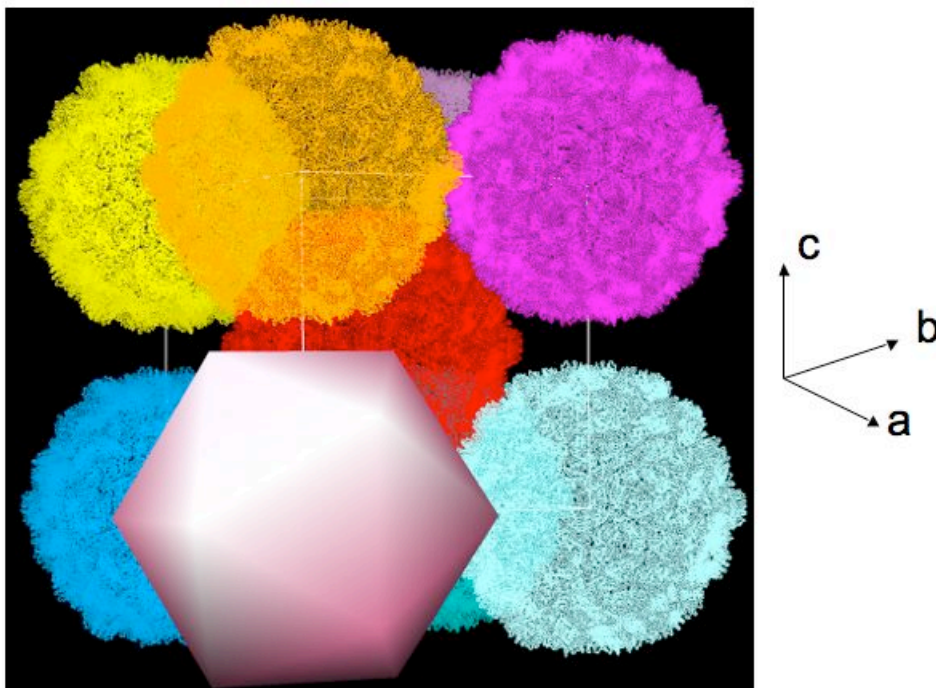


Figure 2.5: **Packing of the P-II crystals in the I222 cell.** 1 full particle is located at the unit cell center, while 1/8 of each particle is located at each corner. The cell dimensions are $a=553.0 \text{ \AA}$, $b=574.4 \text{ \AA}$, $c=587.4 \text{ \AA}$.

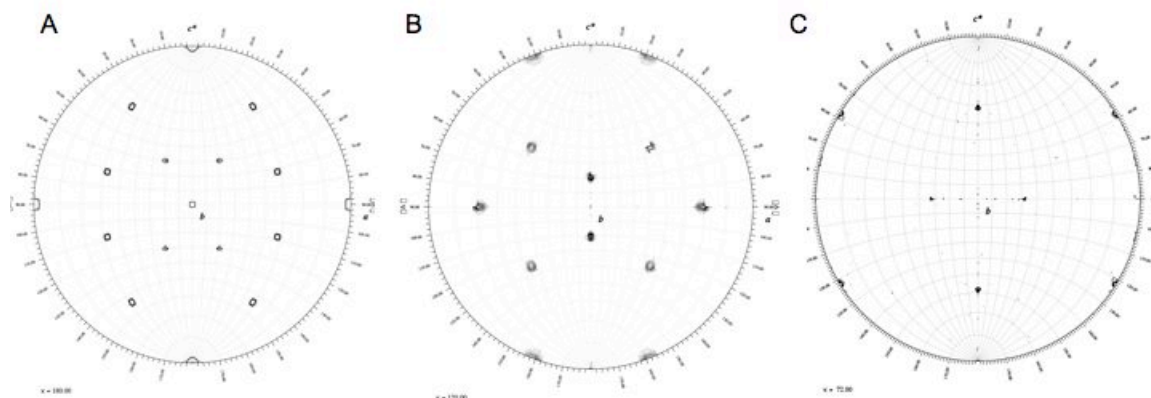


Figure 2.6: **Self-rotation functions** (A) 2-fold rotation function. Scaled data between 12 and 8 Å was used for the search. Peaks are a result of crystallographic 2-fold axes. (B) 3-fold rotation function. 12-8 Å data was also used for this search. Peaks represent 3-fold icosahedral axes. (C) 5-fold rotation function. Peaks indicate 5-fold icosahedral axes. Data only spanning 3.95-3.65 Å was used in the search. Using only the high-resolution recorded reflections in the rotation search provides an effective tool for evaluating the high-resolution data. If expected peaks in the rotation function search are not present using the high resolution data, the data likely will not contribute to structure determination. The 5-fold search shown above shows all expected peaks.

Particle orientation and the self-rotation function

The volume of the unit cell, $186,584,000 \text{ \AA}^3$, indicated that 2 particles were present (figure 2.5). The estimated volume of a particle based on a sphere with a radius of $250 \text{ \AA}$ is $65,450,000 \text{ \AA}^3$. Based on the symmetry requirements of the I222 space group, and the size of the cell, there could only be 2 or 4 virus particles per cell. Four particles would significantly exceed the cell volume, therefore indicating that 2 particles occupied the cell. The symmetry properties of I222 requires 8 asymmetric subunits in the unit cell. A single crystallographic asymmetric subunit is therefore $\frac{1}{4}$ particle. Since the asymmetric unit of the virus is $\frac{1}{60^{\text{th}}}$ of a particle, the non-crystallographic symmetry is 15-fold. The space group symmetry requires that the 3 orthogonal icosahedral 2-fold axes be aligned with the crystallographic 2-fold axes indicating that the 222 sub point group symmetry of the 532 icosahedral symmetry is incorporated into the crystal lattice. Both particles are in identical orientations, but there are two possibilities depending on the assignment of the axial directions. A self-rotation function computed with data between 12 and $8 \text{ \AA}$ resolution and a $300 \text{ \AA}$ radius of integration was performed using the GLRF program [8] to determine the particle orientation. Figure 2.6 shows the contour plots of the 2-fold, 3-fold, and 5-fold rotation functions. Positions in the diffraction pattern corresponding to the non-crystallographic symmetry (NCS) are systematically searched and correlation coefficients computed in the rotation function, giving significant peaks on the contour plots where NCS is detected. A 3-fold and 5-fold rotation function will therefore search for the 3-fold and 5-fold icosahedral axes of the virus particle respectively, since there are no crystallographic axes corresponding to that symmetry. The 2-fold search will result in peaks dominated by the 2-fold symmetry in

the space group I222. The peaks in the 3-fold and 5-fold search indicated that the particle was oriented in the standard convention (z35x), and due to space group constraints, must be centered at 0,0,0 and 1/2,1/2,1/2.

Molecular replacement and averaging

In order to determine the structure of a protein in a crystal using the integrated diffraction intensities, phases are required. In virus crystallography, initial phases are either estimated using a lower resolution Cryo-EM model, another virus structure with structural similarities, or even in some cases a spherical shell model [9,10,11]. The 12 Å Cryo-EM model of P-II provided the initial phases for the 5.2 Å structure [12]. The phases of the refined 5.2 Å structure were then used for the initial phases of the currently described 3.65 Å structure. Averaging and phase extension were used to refine the phases and calculate an electron density map. Following the calculation of the initial map from observed amplitudes of low-resolution data and the phases from the starting model, the phases are subsequently refined during NCS averaging in real space. After repeated cycles of the process (6 in this case), the phases converge to a more correct approximation. The electron density is then back transformed into fourier space, where additional high resolution data is added and the process repeated again. The process is repeated in increments using additionally added higher resolution data until all of the data has been incorporated and the phases at maximum resolution have been refined. Statistics that reflect the convergence of the phases as well as the quality of data include the averaging R-factor (R_{ave}), as well the Correlation coefficient (CC_{ave}) (figure 2.8). The values for both of these are found in figure 2.7 with the equations used for their

A		B	
Space Group	I222	Space Group	I222
Unit cell (a), (b), (c)	553.0, 574.4, 587.4	Unit cell (a), (b), (c)	553.0, 574.4, 587.4
Non-crystallographic symmetry (NCS)	15	Non-crystallographic symmetry (NCS)	15
Resolution Range (Å)	35.0-3.7 [3.8-3.7]	Resolution Range (Å)	35.0-3.65 [3.74-3.65]
Measured reflections	1,298,008	Measured reflections	1,243,749
Unique reflections	666,458	Unique reflections	656,590
Completeness (%)	68.5 [13.4]	Completeness (%)	64.8 [14.8]
Overall R _{merge} (%)	24.5 [100.0]	Overall R _{merge} (%)	18.4 [43.1]
<I/σ(I)>	1.9 [0.6]	<I/σ(I)>	2.5 [1.4]
Multiplicity	1.9 [1.1]	Multiplicity	1.9 [1.1]
		R _{cryst}	0.365
		CC _(ave) (%)	80.7
		R _{ave}	0.284
		R.m.s. deviations	
		Bond lengths (Å)	0.005
		Bond angles (°)	1.16

Figure 2.7: **Crystallographic statistics.** Data in the highest resolution shell is shown in brackets. (A) Preliminary scaling statistics for data processed without a sigma (I) cutoff. The R_{merge} value in the highest resolution shell is extremely high, indicating data in this ranges makes little or no valid contribution. (B) Final Statistics for data processed with the sigma (I) cutoff. Statistics for experiments other than scaling are shown as well, including Averaging, refinement, and stereochemistry. The following equations are used to calculate R_{cryst}, CC_{ave}, and R_{ave} respectively:

$$R_{\text{cryst}} = S(|F_{\text{obs}}| - k |F_{\text{calc}}|) / S|F_{\text{obs}}| ; k = S F_{\text{obs}} / S F_{\text{calc}};$$

$$CC_{\text{ave}} = S(|F_{\text{obs}}| - \langle |F_{\text{obs}}| \rangle) (|F_{\text{calc}}| - \langle |F_{\text{calc}}| \rangle) / (S(|F_{\text{obs}}| - \langle |F_{\text{obs}}| \rangle)^2 S(|F_{\text{calc}}| - \langle |F_{\text{calc}}| \rangle)^2)^{1/2}$$

R_{ave} is the same as R_{cryst} except F_c is calculated from a 15-fold NCS averaged electron density

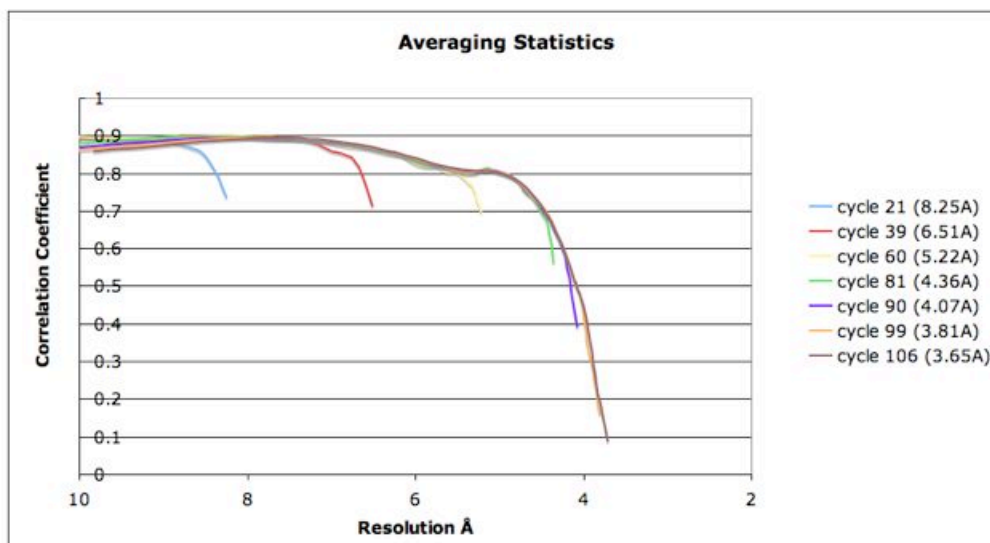


Figure 2.8: **Averaging and phase extension statistics.** The plot shows the change in correlation coefficient for various cycles of averaging. The consecutive cycles use increasing amounts of scaled diffraction data. After each cycle, additional higher resolution data is incorporated into the averaging process. As seen, the highest resolution data, which is usually weaker than the lower resolution data causes a significant decrease in the correlation coefficient. Though the correlation coefficient drops to lower than 0.1 at 3.65 Å, data in this resolution range still contributes positively to the structure determination, verified both by the rotation function using only high resolution data, and direct electron density map comparisons at different resolutions.

calculations described in the table legend. A sharpened map was also created, which increases the contribution of higher resolution reflections that are normally down-weighted in proportion to the Wilson B-factor. Applying a negative Wilson B-factor, (performed in the CCP4 program) increases the contribution of the high-resolution data.

The result was an electron density map with better resolved features that was particularly helpful in the model building.

Model building and refinement

The 3.65 Å structure was initially interpreted by rigid body fitting of the 5.2 Å model into the 3.65 Å electron density. The 5.2 Å structure was determined with the Head II model serving as a starting point that was fitted as a rigid body into the 5.2 Å density. Described below are the major steps in refining the 5.2 Å structure from the initial starting model of Head II [13], followed by the steps involved in the further refinement of the model using the 3.65 Å electron density.

The 5.2 Å Prohead II electron density allowed an unambiguous fit of the main chain of the Head II subunit model, however no side chain density was visible at this resolution. Following rigid body refinement of the Head II model with CNS it was clear that the core of each subunit, including most of the Axial domain (A-domain), remained the same, while various other regions of the subunit underwent hinging and or significant refolding during maturation. The details of the structural differences between P-II and H-II and their relevance are discussed in chapters 4-6. Manual rebuilding of the refolded regions was performed in the program “O” followed by the following refinement procedures.

Simulated annealing and energy minimization refinement were performed in CNS [14] using a least squares residual target function similar to the following:

$$E^{LSQ} = E_{\text{restraints}} + w_a \sum_{hkl} (|F_o| - k|F_c|)^2,$$

where F_o are the experimental structure factors, F_{calc} are the calculated structure factors, w_a are the weights, and $E_{\text{restraints}}$ are the geometric restraints which include bond length, bond angle, and atomic repulsion constraints. The reasoning for using this target function during refinement instead of the more commonly used maximum-likelihood with cross validation function is due to the high non-crystallographic symmetry in the data, common for most viral crystal refinements [14]. The cross validation involves the separation of the data into two components, the test set and the working set, which are refined separately. This allows the measurement of an R_{free} , which is commonly used as an overall indicator of refinement, and more specifically over-refinement. It has been shown, however, that high NCS (over 5) prevents a valid determination of R_{free} due to symmetry related data in the working and test sets [15]. Therefore, the use of cross validation target functions was not valid and would have resulted in less data for the refinement. The aforementioned reciprocal space refinement was performed in CNS. Simulated annealing was also performed in real space with the program Rsref2000 [16].

Identical software packages (CNS and Rsref2000) were used to refine the 3.65 Å structure. The program Coot [17] was used for manual building and visualization of the structure, instead of “O.” Coot has the advantage of performing real space refinement of rebuilt residues immediately after building, as well as more comprehensive statistical tools that allow for rapid evaluation of the geometry of the structure. Since side chain

residues were not modeled in the 5.2 Å P-II structure, there was considerable effort required to position side chains during the 3.65 Å model development. At this resolution, automated side chain refinement in CNS and RSref was not always optimal. Regions that appeared to be more disordered such as the N-terminal part of the spine helix, A-loops, E-loop, and N-arm, required considerable manual fitting in Coot followed by further cycles of energy minimization refinement. Also, regions involved in quaternary interactions required extensive manual refinement despite the use of NCS operators in CNS. During manual fitting and refinement, geometry was heavily weighted in order to maintain good stereochemistry in the model. The Ramachandran plot, the key indicator of stereochemistry, of both the 3.65 Å structure and 5.2 Å structure are shown in figure 2.9.

Summary of 5.2 Å structure data collection and processing

Crystals of W336F P-II were grown in 0.1M Hepes with either 0.2M CaCl₂ or 0.2M MgCl₂ at pH's between 6.0 and 8.5, along with 2-3% PEG 4000. The 5.2 Å structure was determined from diffraction data from 5 room temperature crystals. Data was collected at the synchrotron at Argonne National Laboratories, 14-BMC. The room temperature crystals were cooled at 4°C during data collection, enabling longer collection on each crystal. A similar procedure did not help in the collection of the 3.65 Å data. The diffraction data was indexed in Denzo [6]. The data processing was improved from the previous processing initially done by William Wikoff using a different indexing process. Instead of integrating individual frames of data independently, as was previously done, images were integrated consecutively for each crystal. This alleviates the need for the program to pick up a new crystal orientation for every frame. In

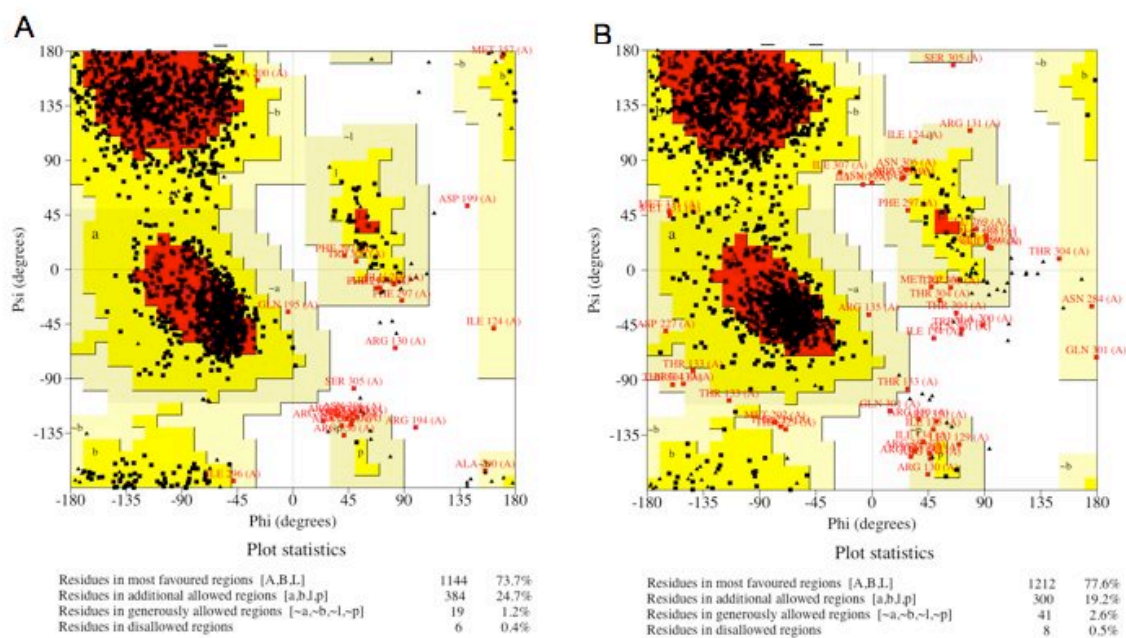


Figure 2.9: **P-II stereochemistry.** (A) Ramachandran plot of 3.65 Å P-II structure.

Percentages of residues with favorable, allowed, and unfavorable phi/psi values are listed below the plot. Each spot on the plot represents a specific residue. Red regions are most favored, yellow are additionally allowed, the tan colored zone corresponds to generously allowed, while the white area is the disallowed region. (B) Ramachandran plot of 5.2 Å P-II structure.

addition, the mosaicity was carefully evaluated and set for batches of images that showed similar mosaicities. The mosaicity will commonly increase during room temperature data collection due to the radiation damage in the crystal perturbing the homogeneity of crystal packing. Most importantly, these efforts increased the completeness of the data set by allowing careful measurement and summation of partial reflections. Reflections with a partiality of 0.5 or greater were also scaled by multiplying $1/(\text{partiality})$ by their observed intensity value. The data, which were originally processed to 5.5 Å showed a significant amount of data going to 5.2 Å. This higher resolution data was included in the processing. To improve the overall data quality, measurements with an $I/\sigma(I)$ value lower than -1.5 were discarded, as well as were diffraction maxima with sigma values lower than zero. This was done with a perl script written by Gabe Lander.

The crystals had space group symmetry $P2_13$, with a lattice constant of 706.9 Å. Four particles were packed into the unit cell with the particle centered at $(1/4, 1/4, 1/4)$ and symmetry related positions. Scaling was done using Scalepack. Both self-rotation functions and locked self-rotation functions were performed to determine the orientation of the particle in the cell. Molecular replacement, and averaging was performed as previously described, using the 12 Å Cryo-EM model for initial phases. Refinement was performed as described earlier in the chapter. The overall statistics are shown in table 2.2.

Comparison of 3.65 Å and 5.2 Å structures

Improving the resolution from 5.2 Å to 3.65 Å may seem modest, but in actuality it is a monumental improvement. Though the 5.2 Å structure identified regions that refolded and hinged during maturation, the extent of these movements, and the

interactions that discriminate the pre-mature and mature particle forms were not defined in detail. This was due to the virtual lack of side chain density at 5.2 Å resolution, compared to nearly complete side chain density at 3.65 Å. The ~1.6 Å increase in resolution allowed us to visualize the side chain interactions that stabilize the quaternary and tertiary structure of P-II and evaluate how they change in response to maturation, further described in chapters 4 and 5. Figure 2.10 shows a comparison of the electron density of the spine helix at both resolutions.

The unconventional techniques used to crystallize and solve the structure of P-II may be important to solve the structures of large macromolecular assemblies, which often require novel crystallization conditions, room-temperature data collection, and extensive fine tuning of the processing and scaling procedures that have become virtually automated for most of the smaller proteins studied by crystallography. Though frozen crystals are greatly preferable due to their durability in withstanding high doses of radiation that allows entire data sets to be collected from a single crystal, metastable proteins and complexes such as viral procapsids are not be amenable to standard freezing and processing techniques.

Table 2.2: **Crystallographic statistics of 5.2 Å structure.** Values in brackets represent the highest resolution shell.

Crystallographic Statistics	
Space Group	P2 ₁ 3
Unit cell (a), (b), (c)	706.9
α , β , γ	90°
Resolution Range (Å)	30-5.20 [5.34-5.20]
Measured reflections	2,012,144
Unique reflections	373,697
Completeness (%)	83.9 [30.1]
Overall R _{merge} (%)	17.8 [36.6]
$\langle I/\sigma(I) \rangle$	3.7 [1.7]
Multiplicity	5.4 [1.2]
R _{cryst}	0.41
Correlation coeff.	0.91
R.m.s. deviations	
Bond lengths (Å)	1.34583
Bond angles (°)	0.008744

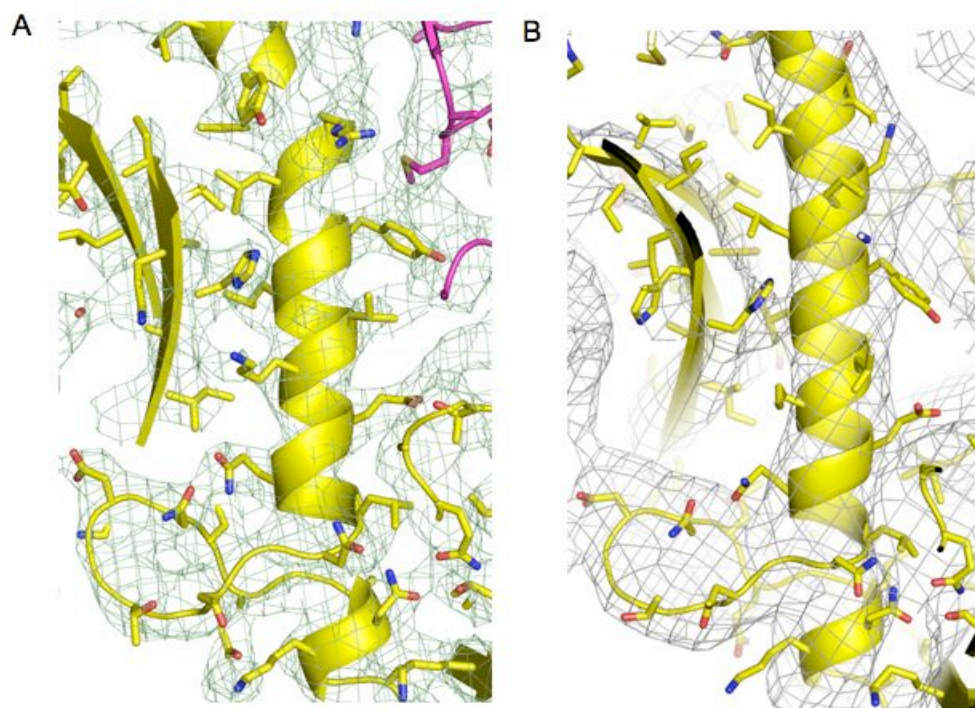


Figure 2.10: **Comparison of 3.65 Å and 5.2 Å electron density maps.** (A) Spine helix of subunit F fit into the 3.65 Å electron density. Notice the sharp grooves of the map in the helix region as well as the extensive density for the side chain atoms. (B) Subunit F fit into 5.2 Å electron density. Virtually no density is resolved for side chains residues, allowing only a confident fitting of the protein backbone.

Acknowledgements

I would like to thank Jeff Speir and Vijay Reddy for assistance in the crystallographic studies. I would also like to thank the staff at beamlines 14-BMC and 23-IDD at the Argonne National Laboratories.

References

1. Wikoff, W. R., Che, Z., Duda, R.L., Hendrix, R.W., and Johnson, J.E. (2003). Crystallization and preliminary analysis of dsDNA bacteriophage capsid intermediate; Prohead II pf HK97. *Acta crystallographica* **D59**, 1-5.
2. Heckman, K. L. & Pease, L. R. (2007). Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat Protoc* **2**, 924-32.
3. Dyson, H. J., Bolinger, L., Feher, V. A., Osterhout, J. J., Jr., Yao, J. & Wright, P. E. (1998). Sequence requirements for stabilization of a peptide reverse turn in water solution--proline is not essential for stability. *Eur J Biochem* **255**, 462-71.
4. Shepherd, C. M., Borelli, I. A., Lander, G., Natarajan, P., Siddavanahalli, V., Bajaj, C., Johnson, J. E., Brooks, C. L., 3rd & Reddy, V. S. (2006). VIPERdb: a relational database for structural virology. *Nucleic Acids Res* **34**, D386-9.
5. Kantardjieff, K. A. & Rupp, B. (2004). Protein isoelectric point as a predictor for increased crystallization screening efficiency. *Bioinformatics* **20**, 2162-8.
6. Otwinowski, Z., and Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology* **276**, 307-326.
7. Diederichs, K. & Karplus, P. A. (1997). Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat Struct Biol* **4**, 269-75.
8. Tong, L. R., M.G. . (1990). The locked rotation function. *Acta crystallographica* **A46**, 783-792.
9. Lee, K. K. & Johnson, J. E. (2003). Complementary approaches to structure determination of icosahedral viruses. *Curr Opin Struct Biol* **13**, 558-69.
10. Baker T, Johnson J: Principles of virus structure determination. In Structural Biology of Viruses. Edited by Chiu W. Oxford: Oxford University Press; 1997:38-79.
11. Tsao, J., Chapman, M. S., Wu, H., Agbandje, M., Keller, W. & Rossmann, M. G. (1992). Structure determination of monoclinic canine parvovirus. *Acta Crystallogr B* **48 (Pt 1)**, 75-88.
12. Conway, J. F., Wikoff, W. R., Cheng, N., Duda, R. L., Hendrix, R. W., Johnson, J. E. & Steven, A. C. (2001). Virus maturation involving large subunit rotations and local refolding. *Science* **292**, 744-8.

13. Helgstrand, C., Wikoff, W. R., Duda, R. L., Hendrix, R. W., Johnson, J. E. & Liljas, L. (2003). The refined structure of a protein catenane: the HK97 bacteriophage capsid at 3.44 Å resolution. *J Mol Biol* **334**, 885-99.
14. Adams, P. D., Pannu, N. S., Read, R. J. & Brunger, A. T. (1997). Cross-validated maximum likelihood enhances crystallographic simulated annealing refinement. *Proc Natl Acad Sci U S A* **94**, 5018-23.
15. Fabiola, F., Korostelev, A. & Chapman, M. S. (2006). Bias in cross-validated free R factors: mitigation of the effects of non-crystallographic symmetry. *Acta Crystallogr D Biol Crystallogr* **62**, 227-38.
16. Korostelev, A., Bertram, R. & Chapman, M. S. (2002). Simulated-annealing real-space refinement as a tool in model building. *Acta Crystallogr D Biol Crystallogr* **58**, 761-7.
17. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-32.

Chapter 3

Application of H/²H exchange to studies of Viral Capsid

Dynamics and maturation

(The chapter describes the methodology and optimization of H/²H exchange coupled with Maldi mass spectrometry to HK97, in guidance from the Komives lab at UCSD)

Introduction

Hydrogen/Deuterium (H/²H) exchange studies have been used to characterize the behavior of proteins for over 50 years [1,2]. The technique has been especially popular in the field of NMR, using deuterium exchanged amide protons as a method for studying structural details as well as protein folding mechanisms [3]. In such experiments, the protein is incubated in a deuterated water from which amide protons as well as side chain protons can exchange with deuterium in the solvent, though only amide proton exchange is considered. The rate of exchange will be determined by the structural properties of the protein, including secondary structure, as well as tertiary and quaternary structure, which may limit accessibility of solvent deuterium to particular regions of the protein [4]. NMR assignments of protein residues are very efficient for smaller proteins <30kD, but larger proteins as well as protein complexes impart difficult technical obstacles [5].

With recent advances in mass spectrometry, proteins have found a new method for analysis. Ionization methods, specifically Electrospray (ESI), and Matrix assisted laser desorption (MALDI), have enabled efficient ionization of proteins into the gaseous phase. Other advances, which include the time of flight tube, as well as fourier

deconvolution techniques have enabled high resolution measurements of protein species. In the last 20 years, many researches have used the hydrogen deuterium exchange technique and have coupled it to mass spectrometry, allowing for the exchange rates of large proteins and complexes to be quantified [6,7,8]. The technique has since been applied to the study of protein interactions, protein folding, conformational dynamics, and allosteric regulation [8,9,10]. Intact virus capsids, which are usually composed of hundreds of protein subunits are among the macromolecular assemblies successfully studied by this technique. We have applied Hydrogen/Deuterium exchange H^2/H coupled to MALDI mass spectrometry in order to study the structural changes and dynamics associated with HK97 capsid expansion. Various expansion intermediates including P-II, EI, and Head I were analyzed by H^2/H , and then compared to each other, as well as known structural information obtained from the crystal structures of P-II and Head I. The H^2/H data for the various particle forms and the implications are described in chapters 4 and 5. In addition to a brief explanation of H^2/H exchange theory in relation to protein study, the rest of this chapter describes the methodologies used to enable the study of HK97 capsid expansion with this technique.

As mentioned previously, the structure and environment of the protein are crucial aspects that determine the rate of backbone amide proton exchange. If a residue is part of a helix or a Beta-sheet, its amide hydrogen is involved in a hydrogen bond, and will therefore exchange solvent protons slower than residues involved in a loop region or dynamic peptide for example. Also, if the residue is buried in a binding pocket of a complex, or a region in the core of a protein protected from solvent, its exchange rate will

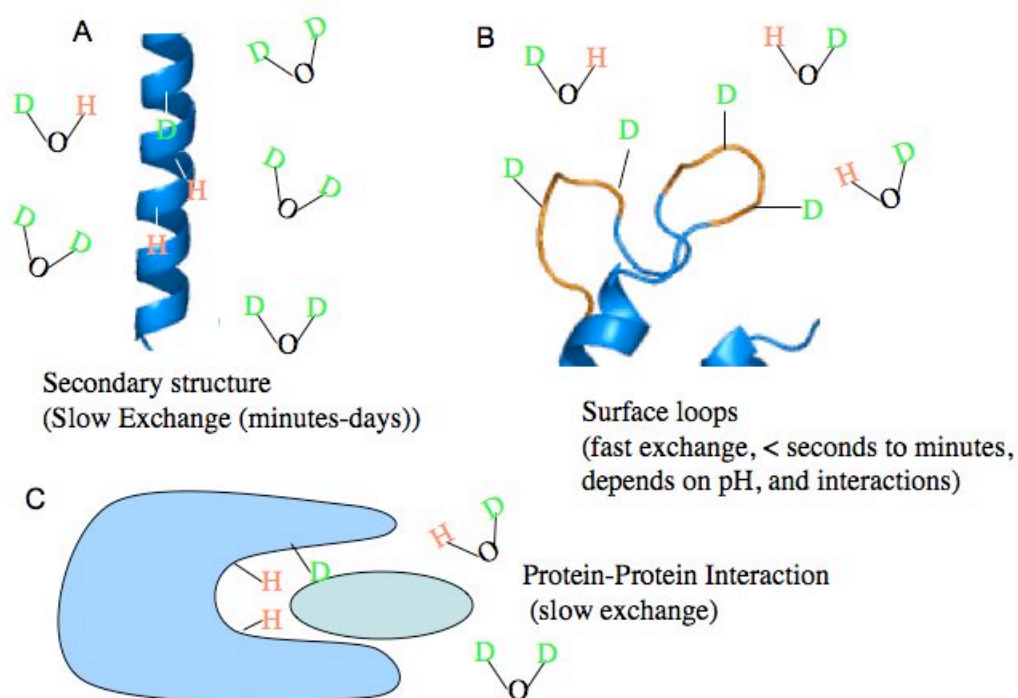


Figure 3.1: **Rates of H^2/H exchange is limited by structure and interactions.**

Schematic representations of varying rates of exchange for various structural elements such as (A) alpha-helices, and (B) loops. (C) Protection from solvent exchange can result from quaternary interactions, or residues buried in binding pocket.

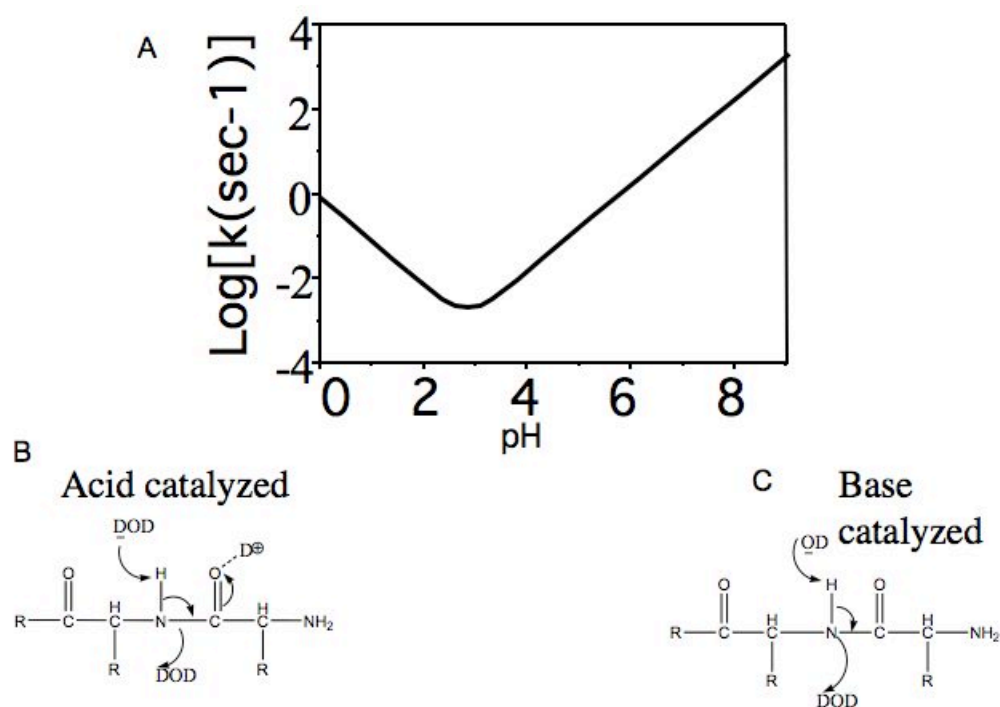


Figure 3.2: **pH determines amide proton exchange rate.** (A) The rate of exchange increases logarithmically below pH ~2, and above pH ~3. The logarithmic changes in rate are due to two different exchange mechanisms, (B) and acid catalyzed mechanism, (C) and a base catalyzed reaction.

also be slower (figure 3.1). Two of the main factors that determine the overall rates are pH and temperature [4]. As shown in figure 3.2, the rate of backbone exchange has a chevron type profile, with the slowest rate of exchange occurring between pH 2 and 3, and increasing logarithmically on either side of the minimum. Increasing the temperature of the reaction will increase the dynamics of the protein and the susceptibility of the amide protons to exchange with solvent. Proteins are generally exposed to deuterium at close to native conditions, which is normally between pH 7.0 and 8.0. Most experiments are also performed at or near room temperature (22° C).

Due to the dynamic properties of proteins, there are local unfolding events that allow solvent to penetrate all regions of the protein, though as mentioned before, these rates will vary. In most protein systems, the intrinsic rate of chemical exchange is considerably slower than the closing rates of the local unfolding events. This type of exchange mechanism is known as EX2. The distribution of exchange with deuterium will therefore appear Gaussian, since all of the protein molecules will not all exchange the same amount of deuterium at any given time-point. During longer incubation times, more of the protein will be exchanged and the Gaussian envelope will shift to higher masses. This can be seen for a peptide in figure 3.3. The first trace shows the spectra of a 19-residue peptide with a standard isotopic distribution from the natural presence of isotopes such as ^{13}C , and ^{15}N . The two traces below show the same peptide after a 30 second and 10 minute exposure to D_2O . The other type of exchange mechanism, in which the intrinsic rate of exchange is much quicker than the closing rate of the transiently unfolded regions is known as EX1 [11], but is rarely seen in native protein studies.

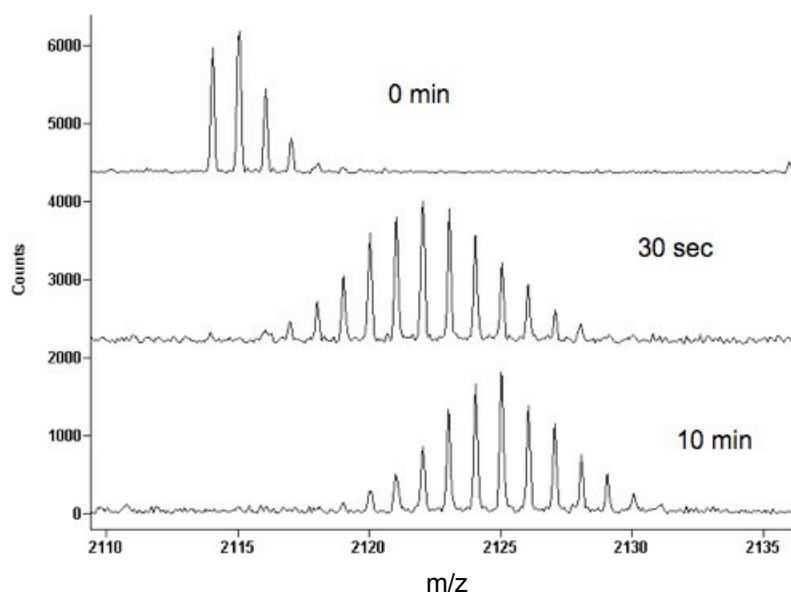


Figure 3.3: **Gaussian envelopes.** Mass envelopes of a 19 residue peptide spanning residues 288-306 of the HK97 capsid protein sequence. The first trace shows the non-deuterated mass envelope, while the next two traces show shifted mass envelopes after 30 seconds and 10 minutes of deuterium exchange respectively.

Though studies have been done in which the entire protein is flown on the mass spectrometer after deuterium exchange, the experiment is not able to disseminate which regions of the protein correspond to either fast or slow exchange rates. In order to increase the resolution of the experiment, deuterium exchange of a given incubation will be quenched using a pH ~2.5 solution, usually trifluoroacetic acid, followed by digestion with a protease that can leave at very low pH, most notably pepsin. Pepsin is very effective, as it is able to cleave after a number of different residues and is most active at acidic pH. Usually the pepsin is bound to immobilized beads and spun down from the protein solution after several minutes of digestion. The result is a number of peptides that can be flown either by ESI or MALDI, and analyzed for any changes in mass in comparison to the non-deuterated standards (figure 3.4). The identities of each peptide are previously characterized without the presence of deuterium, usually by tandem mass spectrometry (MS/MS) in which individual peptides are isolated, fragmented, analyzed, and then carefully matched to a specific sequence of the known protein. The centroid of the deuterated mass envelope can then be quantified after integration of the Gaussian envelope, followed by subtraction of the non-deuterated envelope, enabling calculation of the number of deuterons exchanged on the particular peptide. There is some back exchange that occurs with H₂O both after quenching with low pH during pepsin digestion and while in the mass spectrometer, but this is accounted for during the calculations. Back exchange is normally corrected by fully deuterating a protein either by very long incubations (>24hrs) or at high temperature or in the presence of a denaturant (i.e. Guanidine-HCL), to increase the rate of exchange. Analyzing a fully deuterated fragment on the mass spec will allow one to identify the average number of deuterons that have

back exchanged.

By quantifying the number of exchanged deuterons at different time-points, a kinetic exchange curve can be fit to the data. The curve may either fit single exponential rate, in which all of the amide protons of that fragment exchange at the same rate, or multiple exponential rates where different amide protons of a fragment exchange at different rates. The exchange behavior for most proteins fits either a single, double, or triple exponential rate curve, shown by the following equations.

$$1) D = N_{\text{fast}}(1 - e^{-k_{\text{fast}}t})$$

$$2) D = N_{\text{fast}}(1 - e^{-k_{\text{fast}}t}) + N_{\text{inter}}(1 - e^{-k_{\text{inter}}t})$$

$$3) D = N_{\text{fast}}(1 - e^{-k_{\text{fast}}t}) + N_{\text{inter}}(1 - e^{-k_{\text{inter}}t}) + N_{\text{slow}}(1 - e^{-k_{\text{slow}}t})$$

D is the total of deuterons exchanged, N_x is the number of amides exchanging at a particular rate (fast, intermediate, and slow, though even more are theoretically possible), k_x is the actual rate of exchange, and t is the time of deuterium incubation. Generally, the fast rate of exchange is $> 1 \text{ min}^{-1}$, the intermediate exchange rate is between 0.01 to 1 min^{-1} , and the slow rate is considered $< 0.01 \text{ min}^{-1}$ [12].

Results and Discussion

Optimization of $\text{H}/^2\text{H}$ exchange for HK97

We have applied $\text{H}/^2\text{H}$ exchange with MALDI mass spectrometry to the study of HK97. Unlike ESI in which peptides can be separated using liquid chromatography and eluted separately, all of the peptides in the MALDI scan are analyzed at once. Also, the ionization method in MALDI results in primarily +1 ions, unlike ESI in which various charge states are generated. The negative aspect of flying everything at once is

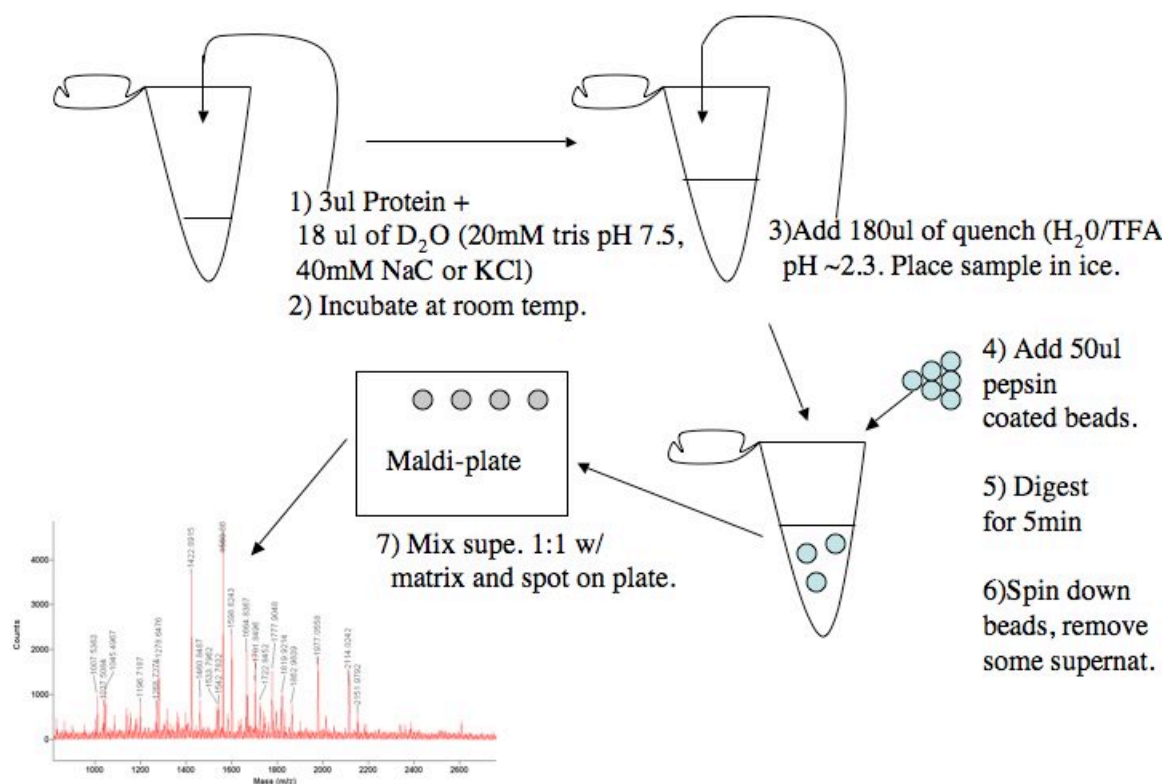


Figure 3.4: **Protocol of H^2/H exchange as optimized for HK97.** After incubation of protein with deuterium solution, the exchange is quenched with non-deuterated acidic solution, digested with pepsin beads, mixed with alpha-C matrix and analyzed on the MALDI-TOF Voyager.

suppression of certain ions, and due to a potential crowding of the peptide envelopes, certain fragments may not be resolved as well. The advantage to MALDI is that the same back exchange correction can be applied to all of the peptides uniformly since there is no staggered elution. The protein solution is mixed directly with a solution of [alpha]-cyano-4-hydroxycinnamic acid in 0.1% TFA, 33% ethanol, 33% acetonitrile (1:1 ratio) at low pH, spotted directly on a cold Maldi plate (low temperature reduces back exchange rate), and flown immediately after a quick drying process under vacuum (Figure 3.4) [19].

One of the main challenges in adapting H/²H exchange for a particular protein system is identifying conditions that allow for the most extensive coverage of the protein of interest. After the protein is digested with pepsin, the goal is to be able to visualize as many of the peptides during the scan as possible, but most importantly, visualizing the peptides spanning regions of the protein that are of most interest. In the case of HK97, there is considerable aggregation that occurs when particles are treated with low pH (<4), unless additional salt is added, and or the concentration of protein is reduced [13]. We flew non-deuterated standards of HK97 at various protein concentrations and salt conditions (sodium chloride and or potassium chloride ranging between 20mM and 300mM in concentration). The scans showed differences in abundances and specificity of fragments. After identifying the sequence of the protein to which each specific peptide was attributed to using ms/ms (appendix), the optimal conditions for each peptide were noted and listed in table 3.1. 70% of the gp5* sequence was resolved in the Prohead II state and shown in figure 3.5. The percentage of identical peptides that could be analyzed on all of the expansion states, was 50%. The optimal final concentration of

Table 3.1: **Peptide detection optimized by salt conditions.** The table shows the conditions that allows for optimal analysis of the corresponding peptide fragment.

Optimal concentrations for the salts listed above are at or near 100mM.

Fragment mass	Residue #'s	Sequence	Best condition
936.48	337-344	DRMDATVE	
941.4733	150-156	YVREEVF	NaCl
999.5376	367-375	ALAHYRPTA	
1007.5274	204-211	QSYINNRL	NaCl
1123.5029	331-339	MASQVWDRM	NaCl
1137.5289	345-353	VSREDRDNF	KCl
1143.5621	206-214	YINNRLMYG	H ₂ O/NaCl
1196.7228	127-136	PGLRRLTIRD	H ₂ O
1240.6798	116-125	IQPMQIPGIIM	H ₂ O
1327.6833	206-216	YINNRLMYGLA	NaCl/H ₂ O
1358.6527	204-214	QSYINNRLMYG	H ₂ O
1422.8909	127-138	PGLRRLTIRDLL	NaCl/H ₂ O
1456.6708	150-161	YVREEVFTNNAD	KCl
1542.7872	204-216	QSYINNRLMYGLA	NaCl
1560.8651	275-287	IVLNPRDWHNIAL	all
1664.8396	288-302	LKDNEGRYIFGGPQA	KCl
1701.8560	248-263	NATGDTRADIIAHAIY	H ₂ O/KCl
1811.9081	288-303	LKDNEGRYIFGGPQAF	H ₂ O/NaCl
1862.9877	271-287	SASGIVLNPRDWHNIAL	H ₂ O/NaCl
1977.0558	367-385	ALAHYRPTAIKGTFFSSGS	all
2010.0561	270-287	FSASGIVLNPRDWHNIAL	H ₂ O/NaCl
2114.0307	288-306	LKDNEGRYIFGGPQAFTSN	all
2226.1307	268-287	SEFSASGIVLNPRDWHNIAL	H ₂ O
2323.0375	324-344	FTVGGFDMASQVFDRMDATVE	NaCl
2358.1552	288-308	LKDNEGRYIFGGPQAFTSNIM	H ₂ O
2375.1479	248-269	NATGDTRADIIAHAIYQVTESE	H ₂ O
2522.2163	248-270	NATGDTRADIIAHAIYQVTESEF	H ₂ O
2537.4276	127-149	PGLRRLTIRDLLAQGRSSNALE	H ₂ O
2607.3353	362-385	CEERLALAHYRPTAIKGTFFSSGS	H ₂ O
3025.5352	177-203	SKQTANVKTIAHWVQASRQVMDDAPML	H ₂ O

Figure 3.5: **Fragments identified and analyzed in P-II.** The total number of residues covered represents 68% of the sequence. Arrows stacked on top of each other represent unique peptides that have overlapping coverage in sequence, providing increased resolution of just the residues corresponding to the overlap.

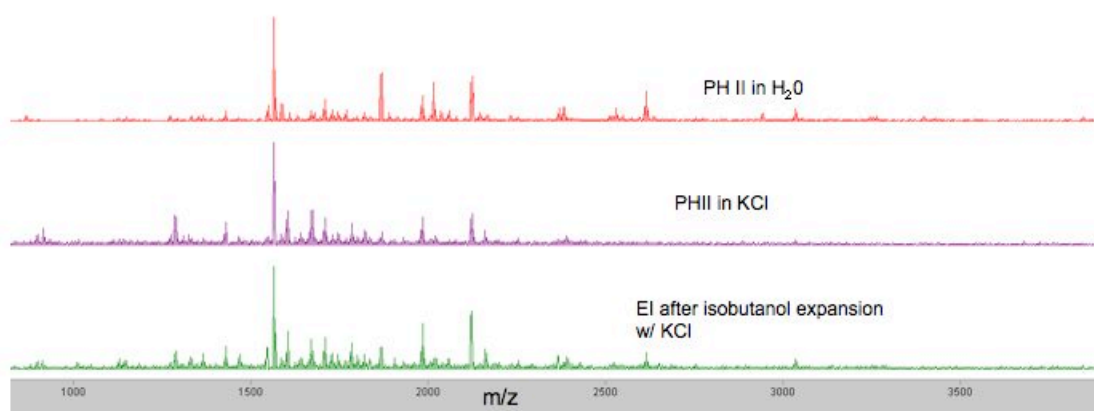


Figure 3.6: **Salt concentration affects mass spectra:** Traces of Prohead II with and without KCL are shown. The trace in red, corresponding to no KCL, shows better coverage of the higher masses while the purple trace (w/KCl) shows better sensitivity in the lower masses. The trace in green shows the spectra of K169Y mutant EI in the presence of isobutanol and KCl.

HK97 after deuterium incubation, quenching, and digestion, was ~13uM or 0.4mg/ml. The stock concentration of HK97 used was between 20 and 40mg/ml which was diluted 10-fold in a deuterium solution buffered with 20mM tris pH 7.5 and containing either sodium chloride or potassium chloride at varying concentrations. It was determined that low salt concentrations were best suited for the peptides of higher mass (2kdaltons to 3.5kdaltons) while higher salt concentrations resulted in better detection of smaller peptides (1kdaltons-1.5kdaltons), seen both in table 3.1 and figure 3.6. Detection of specific peptides of interest required significant variations in salt conditions. For example, the fragment with the m/z 1240.68, spanning residues 117-126 of the N-arm, is barely detectable. However, when flown under high salt conditions, either with sodium chloride, or potassium chloride, the peptide is very abundant in the mass spectra, detected at mass of either 1262.7 or 1278.7, corresponding to the mass of the adducted species containing the respective cation. The N-arm fragment was seen to be fully deuterated within 30 seconds of exchange and was used to calculate the back exchange in all additional experiments. In addition to improving the signal of peptides, salt was also varied to eliminate overlapping mass envelopes that inhibited centroid calculation. For example, the deuterated mass envelope of the fragment spanning residues 288-302 (m/z of 1664) overlaps with a nearby peptide under low salt conditions, but when flown in high KCL concentrations, the contaminating peptide is undetectable, allowing to cleanly resolve the 1664 fragment (figure 3.7).

H/²H exchange of capsid maturation

The major goal of the H/²H study was to probe the changes in folding and quaternary

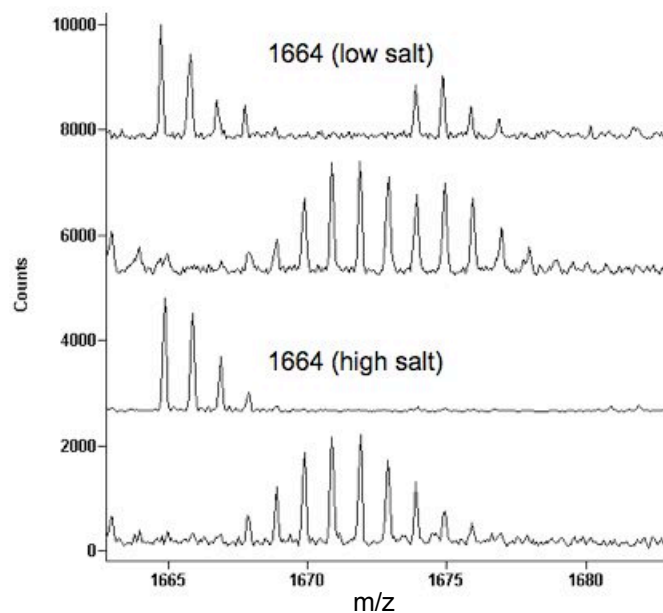


Figure 3.7: **Using salt to select for specific peptides.** The first trace shows two nearby envelopes that appear when P-II is flown with no salt. The envelopes correspond to two separate peptides with different sequences, which when deuterated, results in an overlap of mass envelopes (shown in trace 2) that cannot be confidently deconvoluted. We were only interested in the first envelope, which corresponded to a region of particular interest. By increasing the concentration of salt to $\sim 150\text{mM}$ KCl, the second peptide was no longer visible as seen in the non-deuterated spectra (trace 3), and the deuterated form (trace 4).

reorganizations associated with capsid maturation. The *in vitro* expansion protocol previously developed for HK97 allows distinct intermediate states to be trapped and studied by various biophysical methods including Cryo-EM, SAXS, and crystallography [14,15,16]. P-II must be treated with one of many known chemical reagents in order to induce particle expansion. Low pH (~pH 4.0) is the most commonly used perturbation, but due to the incredibly slow intrinsic exchange rate seen at pH 4.0 (figure 3.1), low pH was a poor candidate for the experiment. Other perturbations such as high guanidine concentrations also could not be used due to its poor compatibility with MALDI. Isobutanol, a very effective expansion inducing reagent, proved to be very compatible with the MALDI method and was therefore chosen for all future expansion experiments.

Previous data has shown that nearly 60% of the particle expansion to the first intermediate state [17], EI, occurs in less than 1 second for a single particle. Ensemble measurements have shown that nearly 100% of the particles have reached the EI state within ~5 minutes of treatment [13,15]. SAXS data has shown that this particle morphology remains stable until further perturbation conditions cause further maturation to later intermediated forms. In the case of acid induced expansion, neutralization to pH 7.0 or greater causes further expansion to the Head II state. In the case of isobutanol, dilution of isobutanol to under 1% results in further expansion. Though the morphology of the EI state remains similar for long periods of time in WT particles, a crosslinking reaction between the lysine on the E-loop and Asparagine on the P-domain occurs after reaching EI and crosslinks progressively accumulate. Conformational flexibility of the E-loop allows the lysine to sample the proper conformation necessary for crosslinking,

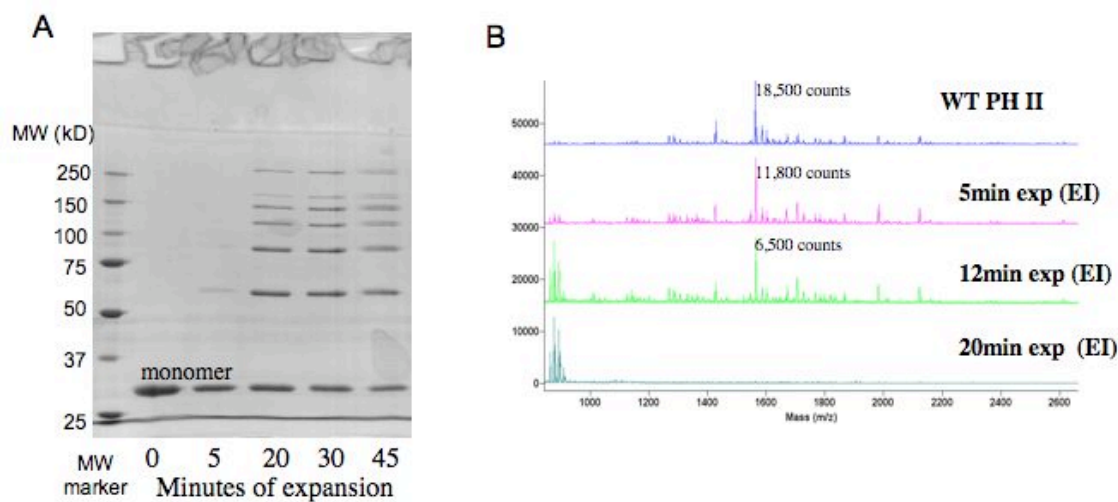


Figure 3.8: Crosslinking hinders WT mass spectra. (A) A 10% SDS PAGE gel showing the increase in crosslinking that occurs during WT HK97 expansion. Various oligomeric forms of covalent interactions are seen, ranging from covalent dimers to 6-mers. (B) Mass spectra were run after various times of expansion. Comparing the spectra to the gel, it is clear that crosslink formation greatly inhibits the detection of peptides.

but only after reaching the EI-form. The crosslinking unfortunately greatly hindered the digestion and analysis of the peptides by mass spectrometry as shown in figure 3.8. After ~10 minutes of expansion, when dimers and trimers could start to be seen by gel electrophoresis (due to crosslinks), the detection of peptides post-pepsin digestion was negligible. It became clear that WT was a poor candidate to examine the kinetic rates of exchange for its corresponding peptides. We instead turned to the K169Y mutant, which has been previously characterized to undergo a similar maturation transition under similar perturbations, though unable to form crosslinks [18]. The EI state could therefore be trapped and analyzed without inhibition of thorough digestion and mass spec analysis. In addition to the EI state, the K169Y Head I state was also analyzed in the study, a particle form that has further expanded from EI to a particle form crystallographically shown to be identical to WT EI-IV [16]. Though the Head I particle form is not in the fully mature head II conformation, it is very close, with the only major difference being the penton vertices which are more extruded in Head II. Nearly all of the hexon subunits have identical conformations between the two states. Head I also showed identical sensitivity to mass spec detection since it was also free of covalent crosslinks. The K169Y mutation therefore allowed for comparisons of the pre-expanded P-II particles, the intermediate EI state, and nearly fully mature Head I state with H^2/H exchange. Analysis of all data was performed using MALDI Voyager software for mass calibration, and a centroiding program called “CAPP,” which was written by Jeffrey Mandell in the Komives lab [8]. The results of the study are presented in chapters 4 and 5.

Acknowledgements

We would like to thank Miklos Guttman, Julia Koeppel, Gina Nubile, and Elizabeth Komives for assistance in training and optimization of H/²H exchange and mass spectrometry for the study of HK97

References

1. Hvidt, A. & Nielsen, S. O. (1966). Hydrogen exchange in proteins. *Adv Protein Chem* **21**, 287-386.
2. Elliott, A. & Hanby, W. E. (1958). Deuterium exchange in polypeptides. *Nature* **182**, 654-5.
3. Englander, S. W., Mayne, L., Bai, Y. & Sosnick, T. R. (1997). Hydrogen exchange: the modern legacy of Linderstrom-Lang. *Protein Sci* **6**, 1101-9.
4. Englander, S. W., Sosnick, T. R., Englander, J. J. & Mayne, L. (1996). Mechanisms and uses of hydrogen exchange. *Curr Opin Struct Biol* **6**, 18-23.
5. Wider, G. (2005). NMR techniques used with very large biological macromolecules in solution. *Methods Enzymol* **394**, 382-98.
6. Zhang, Z. & Smith, D. L. (1993). Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. *Protein Sci* **2**, 522-31.
7. Katta, V. & Chait, B. T. (1991). Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry. *Rapid Commun Mass Spectrom* **5**, 214-7.
8. Mandell, J. G., Baerga-Ortiz, A., Akashi, S., Takio, K. & Komives, E. A. (2001). Solvent accessibility of the thrombin-thrombomodulin interface. *J Mol Biol* **306**, 575-89.
9. Gu, Z., Rao, M. K., Forsyth, W. R., Finke, J. M. & Matthews, C. R. (2007). Structural analysis of kinetic folding intermediates for a TIM barrel protein, indole-3-glycerol phosphate synthase, by hydrogen exchange mass spectrometry and Go model simulation. *J Mol Biol* **374**, 528-46.
10. Anand, G. S., Hotchko, M., Brown, S. H., Ten Eyck, L. F., Komives, E. A. & Taylor, S. S. (2007). R-subunit isoform specificity in protein kinase A: distinct features of protein interfaces in PKA types I and II by amide H/2H exchange mass spectrometry. *J Mol Biol* **374**, 487-99.
11. Fu, C. Y. & Prevelige, P. E., Jr. (2006). Dynamic motions of free and bound O29 scaffolding protein identified by hydrogen deuterium exchange mass spectrometry. *Protein Sci* **15**, 731-43.

12. Kang, S. & Prevelige, P. E., Jr. (2005). Domain study of bacteriophage p22 coat protein and characterization of the capsid lattice transformation by hydrogen/deuterium exchange. *J Mol Biol* **347**, 935-48.
13. Gan, L., Conway, J. F., Firek, B. A., Cheng, N., Hendrix, R. W., Steven, A. C., Johnson, J. E. & Duda, R. L. (2004). Control of crosslinking by quaternary structure changes during bacteriophage HK97 maturation. *Mol Cell* **14**, 559-69.
14. Lata, R., Conway, J. F., Cheng, N., Duda, R. L., Hendrix, R. W., Wikoff, W. R., Johnson, J. E., Tsuruta, H. & Steven, A. C. (2000). Maturation dynamics of a viral capsid: visualization of transitional intermediate states. *Cell* **100**, 253-63.
15. Lee, K. K., Tsuruta, H., Hendrix, R. W., Duda, R. L. & Johnson, J. E. (2005). Cooperative reorganization of a 420 subunit virus capsid. *J Mol Biol* **352**, 723-35.
16. Gan, L., Speir, J. A., Conway, J. F., Lander, G., Cheng, N., Firek, B. A., Hendrix, R. W., Duda, R. L., Liljas, L. & Johnson, J. E. (2006). Capsid conformational sampling in HK97 maturation visualized by X-ray crystallography and cryo-EM. *Structure* **14**, 1655-65.
17. Lee, K. K., Gan, L., Tsuruta, H., Moyer, C., Conway, J. F., Duda, R. L., Hendrix, R. W., Steven, A. C. & Johnson, J. E. (2008). Virus capsid expansion driven by the capture of mobile surface loops. *Structure* **16**, 1491-502.
18. Ross, P. D., Cheng, N., Conway, J. F., Firek, B. A., Hendrix, R. W., Duda, R. L. & Steven, A. C. (2005). Crosslinking renders bacteriophage HK97 capsid maturation irreversible and effects an essential stabilization. *Embo J* **24**, 1352-63.
19. Mandell, J. G., Falick, A. M. & Komives, E. A. (1998). Measurement of amide hydrogen exchange by MALDI-TOF mass spectrometry. *Anal Chem* **70**, 3987-95.

Chapter 4

An Unexpected Twist in Viral Capsid Maturation

The study is presented as published in Nature (currently as an advanced online publication). Co-authored with Lu Gan, Miklos Guttman, Kelly Lee, Jeffrey Speir,

Robert L.Duda, Roger Hendrix, Elizabeth Komives, and John E. Johnson

Introduction

Lambda-like dsDNA bacteriophage undergo massive conformational changes in their capsid shell during the packaging of their viral genomes. Capsid shells are complex organizations of hundreds of protein subunits that assemble into intricate quaternary complexes that ultimately are able to withstand over 50 atm. of pressure during genome packaging [1]. The extensive integration between subunits in capsids requires the formation of an intermediate complex, termed a procapsid, from which individual subunits can undergo the necessary refolding and structural rearrangements needed to transition to the more stable capsid. Though various mature capsids have been characterized at atomic resolution, no such procapsid structure is available for a dsDNA virus or bacteriophage. We present a procapsid x-ray structure at 3.65Å resolution, termed Prohead II, of the lambda like bacteriophage HK97, whose mature capsid structure was previously solved to 3.44 Å [2]. A comparison of the two largely different capsid forms has unveiled an unprecedented expansion mechanism that describes the transition. Crystallographic and Hydrogen/Deuterium exchange data presented here

demonstrates that the subunit tertiary structures are significantly different between the two states, with twisting and bending motions occurring in both helical and β -sheet regions. We also identified subunit interactions at each 3-fold axis of the capsid that are maintained throughout maturation. They sustain capsid integrity during subunit refolding and leverage the rotational and translational motions during maturation. Previously published calorimetric data of a closely related bacteriophage, P22, showed that capsid maturation was an exothermic process that resulted in a release of 90KJ/mol of energy [3]. We propose the major tertiary changes presented in this study reveal a structural basis for an exothermic maturation process likely present in many dsDNA Bacteriophage and possibly viruses such as Herpes which share the HK97 subunit fold [4].

Methods and Materials

Mutagenesis

The W336F point mutation was generated from WT plasmid by site-directed mutagenesis. The W336F mutation suppresses the spontaneous expansion observed in wild-type Proheads and therefore increased the homogeneity of Prohead II preparations. The E-loop was truncated between residues 159-171 to improve crystallization, as previous studies showed that the tip of the full-length E-loop was partially disordered and protruded from the capsid surface [6]. Expression of gp5 capsid protein and gp4 protease was completed in an E.coli T7 expression system and purification of HK97 Prohead II was performed as previously described [27]. W336F mutant assembles into an intact capsid with similar efficiency as wild type. W336F Prohead II is able to crosslink and

expand under acidic conditions, but at a slower rate than wild type Prohead II (unpublished).

Crystal Growth and Structure determination

Crystals were grown using the hanging drop vapor diffusion method with a mother liquor consisting of 0.1M CHES buffer, pH 9.0, 200mM Manganese chloride and 2.3-3.0% Peg 4000. A 200mM final concentration of NDSB-211 (Hampton Research) was added to the drop. An atomic model for the Prohead II structure was initially derived by rigid-body fitting of the refined 3.44 Å structure of the mature Head II coordinates (PDB ID: 1OHG) into the Prohead II electron density. Crystallographic data was collected at the Advanced Photon Source synchrotron at Argonne National Laboratories, beamlines 23-IDD and 14-BMC. The Room-temperature diffraction data from twenty nine crystals was indexed, integrated, and scaled using HKL2000 suite [28]. The crystals belong to the space group, I222. Reflections with an I/sigma (I) of less than 0 were discarded during scaling. Partial reflections were then scaled up to whole reflections using CCP4's SCALA program, followed by scaling of all reflections also with SCALA. The orientation of the particle was confirmed with a 5-fold self-rotation function using the GLRF program [29] to determine which of 2 possible particle orientations is maintained in the unit cell. An atomic model for the Prohead II structure was initially derived by rigid-body fitting of the refined 3.44 Å structure of the mature Head II coordinates (PDB ID: 1OHG) into the Prohead II electron density. Averaging and phase extension was done with CCP4 and RAVE using a previously determined 12 Å Cryo-EM model [6] to build the mask used for molecular replacement. The Head II structure was

manually fit into the Prohead II density followed by rigid body refinement in CNS. The majority of each subunit fit well into the Prohead II map, but certain regions in the A-domain loops, P-domain, and E-loop required significant additional adjustments to improve the fit to the experimental density. These regions were manually fit in “Coot” [30] followed by simulated annealing in real space using Rsref2000 [31]. Energy minimization and geometry refinement of the rebuilt residues was then done in CNS. The regions that required conformation refitting included residues 289-294 of subunits A and D, 298-305 of A-G, and residues 193-215 of the spine helix. Regions that were hinged include the E-loop, N-arm, and P-domain β -sheets. Residues 104-118 of the N-arm were disordered with no visible electron density. The first residue of the N-arm for which electron density could be seen varied for each subunit. The most N-arm density was seen for subunit A, which was visible starting from residue 119, while density for the other subunits started between residues 120 and 127. Electron density for the tip of the E-loops, residues 158-162, also appears weakly, likely due to conformational flexibility in this region.

Bending angles between subunits from Prohead II and Head II states (Figure 4.2) were calculated by deconvoluting matrices needed to align coordinates representing the refined Prohead II structure with that of Head II structure fit into the Prohead II map. The two particle states were initially least squares aligned by the P-loops (346-357), which remained fixed during expansion. R.M.S. deviations ranged from 0.67 Å to 1 Å. Matrices were then calculated for the alignment of the subunit cores (residues 230-383) of Prohead II and Head II from the initial P-loop aligned state. R.M.S. deviations for these least squares alignments ranged from 1.1 Å to 1.3 Å. Residues in this core region

remain mostly rigid during expansion and are all located C-terminal to the E-loop.

Residues N-terminal to the E-loop were not used for alignment, as these residues are all involved in major structural movements between the two states.

H/²H exchange

Samples were incubated in an 85% D₂O solution at pH 7.5 for various time periods, then quenched by the addition of a pH 2.5 non-deuterated quench solution containing Trifluoroacetic acid (final D₂O concentration of 9.0%). Following quench, all samples were kept on ice in a 4°C fridge. Protein was digested with 50ul Pepsin coated beads (Pierce) for 5min. Beads were removed by centrifugation, while supernatant was flash frozen in liquid N₂. Samples were thawed individually, mixed 1:1 with alpha C matrix and vacuum crystallized on a Maldi plate. H/²H exchanged samples were analyzed on a DE-STR MALDI-TOF mass spectrometer. The total number of deuterons exchanged was calculated by subtracting the centroid of the mass envelope from the non-deuterated control from the centroid of the deuterated mass envelopes. The error (standard deviation) was estimated from the average of 3 independent experiments with 2-3 measurements recorded for each experiment (Total of 6-9 measurements for each timepoint). Back exchange was calculated as 42% using a peptide in the N-arm region (residues 117-126), which exchanged amide protons for deuterium completely within 20 seconds. A separate control for back exchange was performed using an 11 residue unstructured synthetic polypeptide, which showed similar back exchange values as the N-arm fragment (117-126). N-terminal amide protons of peptide fragments were not

considered exchangeable residues as they are not expected to retain deuterium after quenching, nor were proline residues.

Deuterium incubations were performed for up to 10 minutes, enabling measurement of amide protons exchanging at both a fast ($> 1 \text{ min}^{-1}$) and intermediate rate (0.01 to 1 min^{-1}) [32]. Longer incubations measuring slow exchange rates ($< 0.01 \text{ min}^{-1}$) were not done in this study. Amide protons exchanging at intermediate to slow rates are generally a result of solvent protection due to either secondary structure or protein-protein interactions. The measured data for all fragments was best fit to either a single or two-exponential model accounting for deuterons exchanging at only a fast rate, or both a fast and intermediate rate respectively. The following equation represents the two-exponential fit:

$$D = N_{\text{fast}}(1 - e^{-k_{\text{fast}}t}) + N_{\text{inter}}(1 - e^{-k_{\text{inter}}t})$$

where D is the total number of deuterons exchanged at time t , N_{fast} is the number of deuterons exchanging at a fast rate, k_{fast} , and N_{inter} is the number of deuterons exchanging at an intermediate rate, k_{inter} . The fast exchanging amide protons had nearly all exchanged by the first timepoint, so k_{fast} was estimated as described previously [17]. All fragments were identified with tandem mass spectrometry (MS/MS) using a Q-star mass spectrometer.

Results and Discussion

HK97 is a favorable system for studying capsid maturation as capsid particles can be assembled in *E.coli* from the expression of just two viral gene products, gp4 (protease)

and gp5 (capsid subunit), and maturation can be triggered and analyzed *in vitro* (figure 4.1) using chemical or low pH treatments [2, 5-8] as opposed to the packaging of dsDNA, which induces maturation *in vivo*. During the maturation, subunit reorganization facilitates a particle expansion from 540 Å (Prohead II) to 660 Å (Head II) in diameter (Figure 4.1). The kinetics of maturation were previously studied using time-resolved solution X-ray scattering [14] and the structures of the intermediates were determined with Cryo-EM [6,7,15] and X-ray crystallography [2,5]. Near atomic resolution structures have characterized the late maturation states (Balloon, Head II), but only lower resolution Cryo-EM models were previously available for the Procapsid and EI forms, which used the 3.44 Å Head II structure as a basis for pseudo atomic models. The previous 12 Å resolution Cryo-EM study of Prohead II suggested that the majority of the capsid structural changes in expansion were the result of rigid-body rotations and translations of the central domains of the subunit, while the E-loop and N-arm regions moved independently⁶.

Here we report a 3.65 Å resolution X-ray crystal structure of W336F, E-loop truncated Prohead II (PDB ID: 3E8K) that changes the previous conceptions of capsid maturation (crystallographic statistics listed in Table 4.1). These mutations did not affect the assembly of the capsid or its ability to undergo maturation. The structure reveals that 3-fold contacts between subunits, mediated by “P-loops” as well as their surrounding β -strands on each subunit, are preserved during maturation from Prohead II to Head II. As a result, the previously proposed rigid subunit motions at lower resolution could now be resolved as domain motions corresponding to a twist of the subunit about three β -strands (β D, β J, and β I) (Figure 4.1b), and a simultaneous bending and unwinding of the long

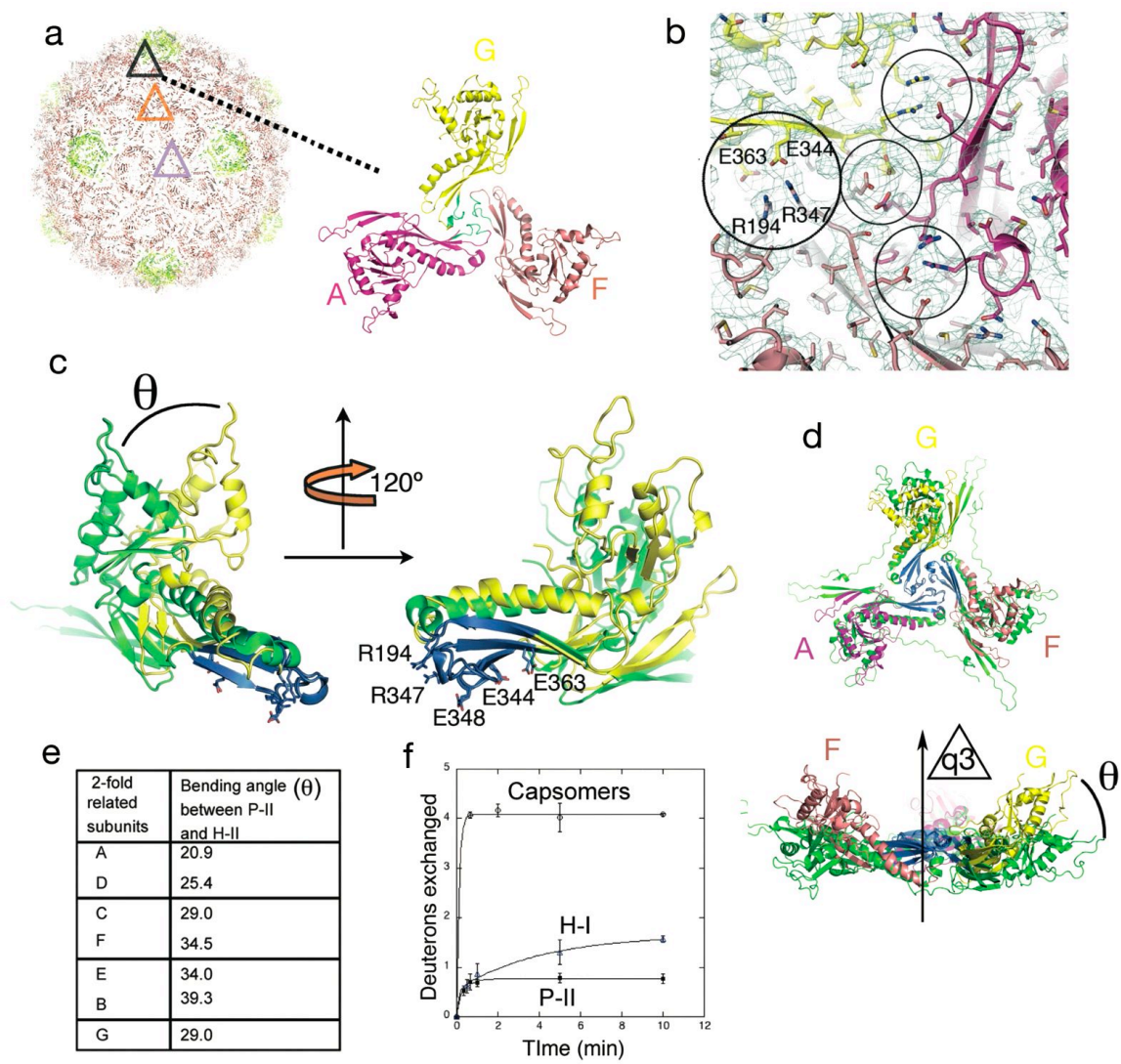
Figure 4.1: HK97 assembly and morphology. (a) The 384-residue gp5 subunit initially assembles into hexameric and pentameric oligomers, termed capsomers, that first assemble to form the Prohead I capsid (P-I). The T=7 *laevo* particle is composed of 12 pentamers and 60 hexamers and encapsidates approximately 60 copies of gp4 protease [9-11]. Expression with a defective protease produces a P-I particle that can be disassembled *in vitro* into free capsomers and re-assembled when exposed to specific chemical treatments [12]. When active gp4 is present, particles spontaneously mature to the 13-MDa Prohead II (P-II) form following digestion of residues 2-103 from all subunits. Crosslinking occurs in the WT particle following formation of the EI state. Crosslinks (isopeptide bond) form between Lys 169 and Asn 356 located on different subunits. A crosslink defective mutant K169Y expands to Head I, a state nearly identical to Balloon minus the crosslinks. WT Balloon undergoes a final expansion step to Head II in which the pentons become more protruded and form one last class of crosslinks, with a molecular topology similar to chainmail [2,13]. (b) Crystal structure of subunit D of Prohead II at 3.65 Å (c) 3.65 Å electron density map (displayed as a solid surface) of the full Prohead II capsid, contoured at $\sim 1\sigma$ in chimera. The Prohead II hexamers and pentamers are shown alongside the capsid with the 7 subunits of the viral asymmetric subunit labeled A-F for the hexamers and G for the pentamers. (d) A calculated electron density map of the Head II capsid shown at 3.65 Å, also rendered at $\sim 1\sigma$. (e) Prohead II and Head II hexamers shown tangential to the capsid surface (rotated 90 degrees from view in 1C and 1D).

(spine) helix with respect to the fixed 3-fold interaction sites. The extent of the subunit twist and the helix bend vary among subunits and depend on their quasi-equivalent position.

The overall morphologies between the Prohead II and Head II states are very distinct. The subunits in Prohead II are oriented radially relative to the capsid surface, but are roughly tangential in Head II (Figure 4.1c-1e). A striking feature of Prohead II, which was seen in the previous Cryo-EM study, is that the skewed hexamers comprised of trimers of subunits with a trapezoidal arrangement give the hexamers a pseudo 2-fold appearance.

The refined P-loop contacts in Prohead II bear a striking similarity to the same contacts in Head II. The P-loop of each subunit is tightly associated with the P-loop of two other subunits from separate capsomers at all 3-fold and quasi-3-fold axes (Figure 4.2). In the previous Cryo-EM-based model, the P-loop of Prohead II was kept fixed relative to the subunit core, changing the trimer associations when compared to those in Head II. It is now clear that the position and quaternary interactions of the P-loops and surrounding β -strands (region colored blue in figure 4.2c) are unchanged during expansion, demonstrating that it functions as a fixed point of subunit interaction in an otherwise highly plastic quaternary structure. Figure 4.2b and c shows salt bridge interactions between Glu 344 and 363 from 2 of the β -strands surrounding the P-loop on one subunit with Arg 194 (located on the turn following the spine helix) and Arg 347 (located on the P-loop) of a neighboring 3-fold related subunit. The salt bridges as well as a putative metal binding site coordinating three glutamate (E348) residues directly underneath each 3-fold axis (figure 4.2b) remain unchanged during capsid maturation.

Figure 4.2: P-loops located at 3-fold axes act as invariant pivot points. (a) Ribbon representation of P-II. The orange triangle represents an icosahedral 3-fold, black and magenta triangles represent two quasi-3-fold positions. A zoomed in view of subunits at a quasi-3 fold axis are shown viewed from outside of the capsid. Residues 345-353 of the P-loop are colored lime-green, and represent the peptide fragment analyzed by H^2H exchange. (b) Side-chain interactions at 3-folds that remain invariant during expansion, viewed from the interior of the capsid directly underneath a quasi-3-fold axis, 180 degrees from the view of the trimer in figure a. Electron density is contoured at 1σ . The 3 outer circles highlight salt bridges while the center circle highlights three glutamates (E348) coordinated at a putative metal binding site. (c) Subunit G of P-II (yellow) and H-II (green) have been aligned by the region of the P-domain which remains invariant (blue). (This motion is best captured in Supplementary Movie 1). (d) A P-II trimer (subunits A,F,G as shown in a) is aligned on the H-II trimer (green) by the regions that remain invariant (blue), illustrating the rotational motions in respect to the fixed trimeric interactions. The upper view is looking down a quasi 3-fold while the lower figure shows a perpendicular view, tangential to the capsid surface. (e) Table shows the angles of rotation (θ) of each subunit from P-II to H-II as illustrated in both c and d. (f) H^2H exchange curve of a peptide fragment spanning residues 345-353 of the P-loop (colored lime-green in a) shown for P-II, H-I, and free capsomer states. Time points are taken from 30 seconds to 10 minutes, with error bars representing standard deviations from the average of 3 independent experiments, with 2-3 measurements per experiment (6-9 total measurements for each timepoint).



Three of these residues (R194, E344, and E363) are proximal to the borders of the region that remains fixed during maturation, defining the boundaries of the pivot points of tertiary rearrangement (Figure 4.2c). In accord with this newly recognized structural constraint, the tertiary structure of the subunit is now seen to have a significant twist about the P-domain β -sheet (Supplementary Movie 4.1).

To corroborate the conclusions from the crystallographic data, we characterized the dynamics of the 3-fold P-loop interactions with H^2H exchange coupled to MALDI mass spectrometry on Prohead II, Head I, and free capsomers. The technique measures the solvent accessibility of amide protons (in native proteins in solution) whose rate of exchange with deuterium is influenced by secondary, tertiary and quaternary structure interactions [17,18]. Following incubation in deuterium, the capsid protein is digested with pepsin protease and the masses of previously determined peptide fragments are quantified. Regions with greater solvent accessibilities will have larger shifts in their mass envelopes, which are quantified as described in the supplementary methods. A K169Y mutant was used instead of WT-Head II for the study since covalent crosslinks inhibited efficient pepsin digestion and subsequent analysis by mass spectrometry. The mutant is able to expand through similar intermediate forms as wild-type Prohead II (Figure 4.1a), though maturation stops at the penultimate, Head I state, which was shown by crystallography to have very similar subunit tertiary structures as wild-type Head II [5]. The crosslink-defective mutant therefore permitted comparisons of H^2H exchange profiles between subunits in Prohead II and subunits in a virtually mature particle form. H^2H exchange was also performed on capsomers that were disassociated from the Prohead I state and were no longer able to form 3-fold P-loop associations. One of the

peptide fragments spanned residues 345-353 of the P-loop (colored lime-green in Figure 4.2a), which lies at the junction of the trimer interface. As seen in Figure 4.2e, this P-loop fragment is highly solvent protected in both the Prohead II and Head I states, while in free capsomers it is nearly five times more solvent accessible. Quaternary interactions are therefore limiting the rate of amide proton exchange in these intact particle forms, while P-loops in the unassociated capsomers are more free to exchange. Data generated for EI (Figure 4.5) yielded nearly identical exchange profiles as seen for Prohead II and Head I, verifying the presence of P-loop interactions during intermediate stages of expansion as well. Consistent with the Prohead II crystal structure, the H^2/H data confirmed that strong interactions remained fixed at the P-loop 3-fold sites, despite the large subunit rotational motions.

The magnitudes of rotation that bring the Prohead II subunit into the Head II conformation (Figure 4.2e) were measured by superimposing the residues behind the fixed region colored blue in Figure 4.2c. The measurements therefore directly relate to the degree of tertiary twisting, which at lower resolution was quantified as whole subunit rotations in previous studies [6]. Subunits closest to the pseudo 2-fold axis (A and D) undergo the least rotation, while those furthest (B and E) undergo the most.

One of the fixed anchor points, Arginine 194, resides several residues N-terminal to the spine helix. The majority of the subunit beyond the fixed P-domain region (colored blue in Figure 4.2c) twists as a rigid unit, causing significant bending of the helix, which is fixed at its N-terminal end. The degree of helix bending is therefore proportional to the extent of β -strand twisting (Supplementary Movie 1). The helix deformation in Prohead II can be seen in figure 4.3a and Supplementary Movie 4.3.

Subunits B,C,E,F and G show dramatic helix bending while subunits A and D exhibit straighter helices as well as smaller twisting motions in the P-domain β -sheet.

To examine the dynamics of the spine helix in solution, H^2H exchange was measured for a peptide spanning residues 206-216, which covers the bent region. The average amount of deuterium exchanged in this region of Prohead II is nearly 5 times greater than in Head I, showing a more canonical helical structure with stronger hydrogen bonding in the mature Head I form (Figure 4.3b). H^2H measurements of the first expansion intermediate, EI, were also performed. The helix peptide shows nearly identical solvent exchange for EI as Head I, indicating that the increased hydrogen bonding in the helix occurs in the initial stage of expansion. The helix in the free capsomer state shows a similar level of solvent accessibility as Prohead II, indicating that the helix distortion is not just a result of the quaternary arrangement enforced in the intact capsid, but is likely occurring at the level of capsomer assembly and facilitated by interactions of the Δ -domain (residues 2-103 that function as a scaffold and are cleaved off of Prohead I to form Prohead II).

Quaternary associations are likely inducing different degrees of strain in the local tertiary structures of the 7 quasi-equivalent subunits. The subunits are not only in a skewed arrangement in the Prohead II hexamer, but they also exhibit different orientations depending on their positions in the hexamer. While the long axes of subunits, B,C,E, and F lie more radial to the capsid surface, subunits A and D lie more parallel to the capsid surface and therefore do not need to rotate as much to assume their orientations in the mature hexamer (Supplementary Movie 4.2). Because P-loop contacts are preserved during maturation, there is a strong correlation between the orientation of

the subunit relative to the capsid surface and the change in tertiary structure between Prohead II and Head II, with the more tangential subunits showing less tertiary structure change and the more radial displaying the larger tertiary structure change.

The combined crystallographic and H^2/H exchange data demonstrate that the large subunit rotations concomitant with expansion from Prohead II to Head II are facilitated by a tertiary structural transition – the twist of the subunit core about a fixed hinge. H^2/H exchange data of the helix, which appears to be bent in concert with the overall hinging motions, indicates that most of the change in tertiary structure occurs during the initial and irreversible expansion from PII to EI [8,19] (Figure 4.3c). This is reasonable considering nearly 60% of the expansion in size occurs in the first transition, as well as the symmetrization of the hexamers [8] (Supplemental Movie 4.2). We propose that the bent helix and twisted β -strand in Prohead II place the subunits in a strained conformation of elevated free energy and that this accounts for both the meta-stability of Prohead II and the driving force for the initial expansion to EI. The 3-fold interactions at the P-loops stabilize inter-capsomer interactions during the expansion. Capsid integrity is augmented after transition to EI, which is competent for covalent crosslinking in the 3-fold region. The energy sources for the distorted tertiary structure in Prohead II likely stems from the initial assembly, in which the Δ -domains (residues 2-103) of each subunit putatively act as molecular scaffolds that promote capsomer assembly. The favorable association of Δ -domains in this early assembly product may induce the strained conformation (Figure 4.4). The high level of deuterium exchange observed in the spine helix of free capsomers supports our hypothesis that the bent subunit conformation exists at the stage of capsomers, not just fully assembled capsid. Δ -domains interact in a trimer

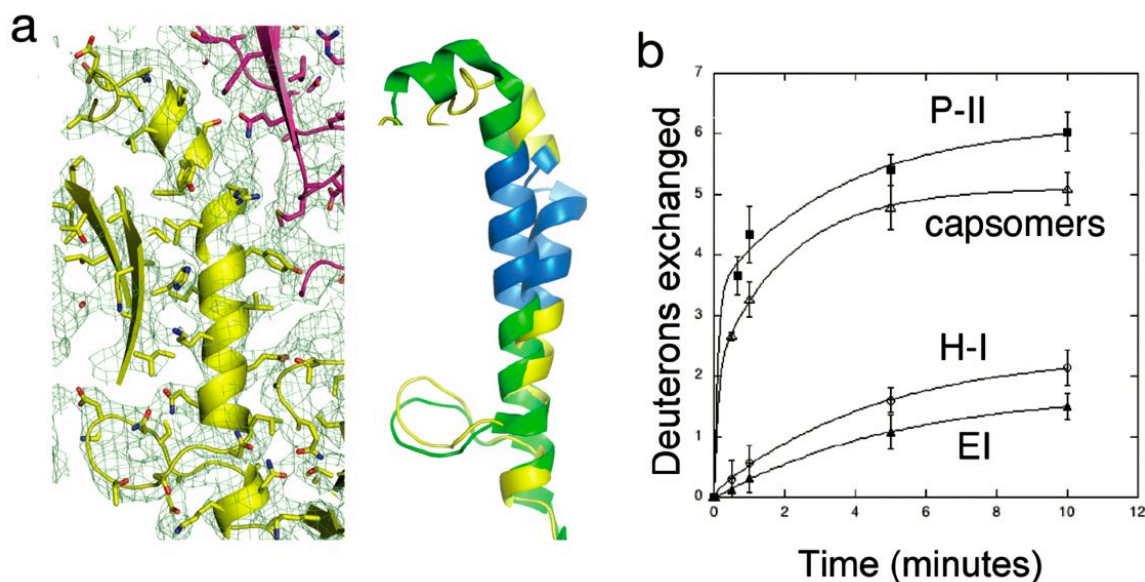


Figure 4.3: **Spine helix bends during maturation.** (a) The spine helix (yellow) is shown for subunit F of Prohead II both in its corresponding electron density on the left (1σ), and aligned with Head II on the right (green). The subunits from the two states were aligned using the subunit core that acts mostly as a rigid body (residues 230-383) with an r.m.s.d of $1.3 \text{ \AA}$ or better for each alignment. The region colored blue represents the fragment spanning residues 206-216, analyzed by H^2/H exchange. (b) H^2/H exchange rate curves comparing deuterium exchange between P-II, EI-I, H-I, and free capsomer helix fragment.

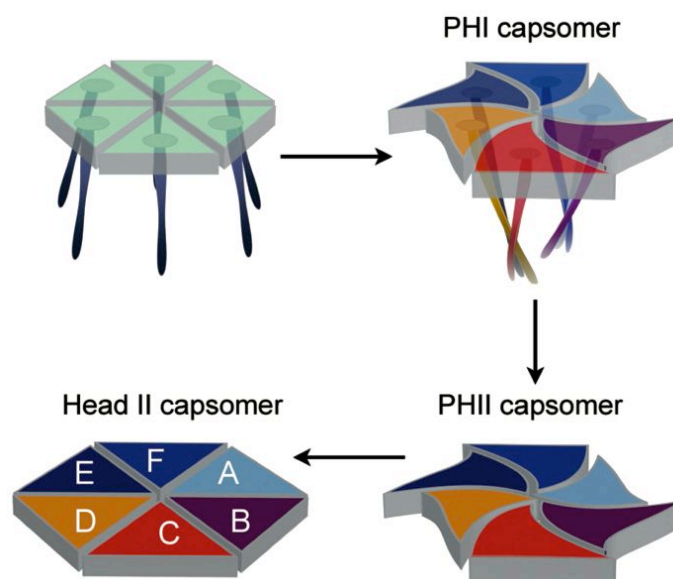


Figure 4.4: **A working hypothesis for the formation, meta-stability, and subsequent maturation of HK97, represented with a single hexamer.** Individual subunits are first assembled into hexamers and pentamers (far left figure is a hypothetical representation of the initial subunit organization). Based on H^2/H exchange data subunit tertiary structures are distorted in free capsomers and we hypothesize that the hexamers are skewed (second figure). Prohead I is formed by assembly of hexamers and pentamers into a $T=7$ particle with Δ -domains attached. Following proteolysis of the Δ -domains to form Prohead II, the skewed hexamers and distorted tertiary structures are preserved by quaternary structure interactions in the particle, raising the free energy of the particle to a meta-stable state maintained in a local minimum. Perturbation of these particles by dsDNA packaging (in vivo) or lowering the pH (in vitro) lowers the energy barrier, leading to an exothermic expansion of the particles producing symmetric hexamers and undistorted subunit tertiary structures.

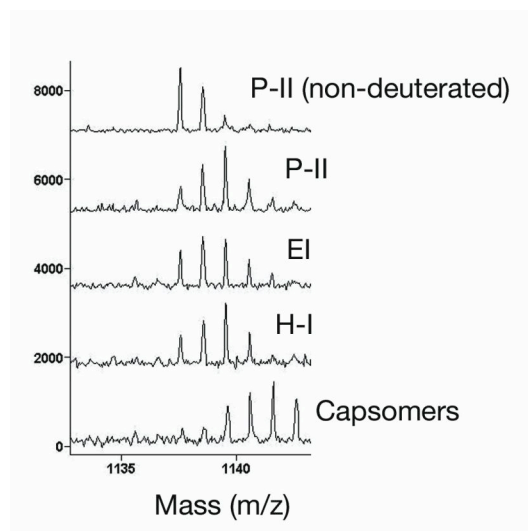


Figure 4.5: **The P-loop is solvent protected throughout maturation.** Mass envelopes of the peptide fragment spanning residues 345-353 of the P-loop are shown at various expansion states following 1 minute deuterium incubations. Row 1 shows the non-deuterated control, while the next three rows shows the P-II, EI and H-I states respectively. The bottom row shows the mass envelope of free capsomers. The free capsomers exchanged much more deuterium into the P-loop peptide, so much so that the envelope merged into the next peak in the spectrum. The spectral window has been truncated at the next peak for ease of viewing.

Table 4.1: **Crystallographic statistics for Prohead II.** Values in brackets indicate highest resolution shell.

$$R_{\text{cryst}} = S(|F_{\text{obs}}| - k |F_{\text{calc}}|) / S|F_{\text{obs}}| ; k = S F_{\text{obs}} / S F_{\text{calc}}; CC_{\text{ave}} = S(|F_{\text{obs}}| - \langle |F_{\text{obs}}| \rangle) (|F_{\text{calc}}| - \langle |F_{\text{calc}}| \rangle) / (S(|F_{\text{obs}}| - \langle F_{\text{obs}} \rangle)^2 S(|F_{\text{calc}}| - \langle |F_{\text{calc}}| \rangle)^2)^{1/2}$$

R_{ave} is the same as R_{cryst} except F_c is calculated from a 15-fold NCS averaged electron density map.

R_{free} was not calculated as previous studies have shown this to be an incorrect validation method for data with high non-crystallographic symmetry.

Space Group	I222
Unit cell (a), (b), (c)	553.0, 574.4, 587.4
Non-crystallographic symmetry (NCS)	15
Data Processing	
Resolution Range (Å)	35-3.65 [3.74-3.65]
Measured reflections	1,243,749
Unique reflections	656,590
Completeness (%)	64.8 [14.8]
Overall R_{merge} (%)	18.4 [43.1]
$\langle I/\sigma(I) \rangle$	2.5 [1.4]
Multiplicity	1.9 [1.1]
Wilson B factor (Å ²)	33.4
$CC_{\text{(ave)}}$ (%)	80.7
R_{ave}	0.284
Refinement	
R_{cryst}	0.365
R.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	1.49
Mean residue B-factor (Å ²)	35.6
Ramachandran Plot	
Allowed	99.6%
Most favoured	73.7%
Additionally allowed	24.7%
Generously allowed	1.2%
Outliers	0.4%

arrangement in the hexamers of Prohead I [20], which assume a skewed symmetry similar to Prohead II. Though Prohead I, prepared without the viral protease, is resistant to expansion when exposed to conditions that expand Prohead II, Cryo-EM of Prohead I particles heated to 55 degrees C showed a reversible transition to an EI like state [20]. That study showed that heating the Prohead I particles causes a disruption of Δ -domain interactions, enabling the particle to expand beyond the Prohead II state to a state in which the hexamers were symmetric. Based on these data we argue that upon disruption of Δ -domain interactions and the formation of symmetric hexamers, the tertiary strain in the subunits is relieved. When the Δ -domains are present, as they are in Prohead I, cooling the particles causes the Δ -domains to re-associate, which induces the skewed hexamers and strained tertiary structure. When the Δ -domain is absent as in Prohead II, the particle is trapped in an elevated local energy minimum until it is perturbed to expand either by DNA packaging *in vivo*, or by chemical perturbation *in vitro*.

Tuma et al. [21] previously proposed a mechanism for P22 expansion where the procapsid subunit existed as a late-folding intermediate that underwent further tertiary changes *en route* to the lower energy, mature conformation. Such a mechanism is now evident in HK97, and may in fact be the driving force for expansion. Systems in which tertiary structure folding events, comparable to those presented here for HK97, have been characterized include the CA domain of the HIV capsid protein, which has been shown to require a kinking of a helix in order to induce dimer activation [22]. Although such tertiary structural changes have not been characterized at high resolution in other dsDNA phage and viruses, they may in fact be present in P22, T4, T7, phi 29 and possibly animal viruses such as Herpes viruses, which all share an HK97-like fold [23-26,4].

Acknowledgements

We thank Vijay Reddy for assistance with crystallographic studies and helpful discussions. We thank Rick Huang for providing HK97 capsomer samples and as well as for helpful discussions, and Tsutomu Matsui for help with x-ray data collection. We thank Brian Firek and Crystal Moyer for mutagenesis of the HK97 constructs used in the study. We also thank Blair Szymczyma for material used in the study. We thank the staffs at beamlines 14-BMC and 23-ID-D of the Advanced Photon Source for assistance in data collection. This work was supported by NIH Grants RO1 AI40101 (to J.E.J), RO1 GM47795 (to R.W.H) and NIH Training Grant GM08326.

References

1. D. E. Smith et al. The bacteriophage straight phi29 portal motor can package DNA against a large internal force. *Nature* **413**, 748-752 (2001).
2. W. R. Wikoff et al. Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* **289**, 2129-2133 (2000).
3. A. C. Steven, J. B. Heymann, N. Cheng, B. L. Trus, and J. F. Conway. Virus maturation: dynamics and mechanism of a stabilizing structural transition that leads to infectivity. *Curr. Opin. Struct. Biol.* **15**, 227-236 (2005).
4. M. L. Baker, W. Jiang, F. J. Rixon, and W. Chiu. Common ancestry of herpesviruses and tailed DNA bacteriophages. *J. Virol.* **79**, 14967-14970 (2005).
5. L. Gan et al. Capsid conformational sampling in HK97 maturation visualized by X-ray crystallography and cryo-EM. *Structure* **14**, 1655-1665 (2006).
6. J. F. Conway et al. Virus maturation involving large subunit rotations and local refolding. *Science* **292**, 744-748 (2001).
7. R. Lata et al. Maturation dynamics of a viral capsid: visualization of transitional intermediate states. *Cell* **100**, 253-263 (2000).
8. K. K. Lee et al. Virus capsid expansion driven by the capture of mobile surface loops. *Structure* **16**, 1491-1502 (2008).
9. R. L. Duda et al. Structural transitions during bacteriophage HK97 head assembly. *J. Mol. Biol.* **247**, 618-635 (1995).
10. R. L. Duda, K. Martincic, and R. W. Hendrix. Genetic basis of bacteriophage HK97 prohead assembly. *J. Mol. Biol.* **247**, 636-647 (1995).
11. J. F. Conway, R. L. Duda, N. Cheng, R. W. Hendrix, and A. C. Steven. Proteolytic and conformational control of virus capsid maturation: the bacteriophage HK97 system. *J. Mol. Biol.* **253**, 86-99 (1995).
12. Z. Xie and R. W. Hendrix. Assembly in vitro of bacteriophage HK97 proheads. *J. Mol. Biol.* **253**, 74-85 (1995).
13. R. L. Duda. Protein chainmail: catenated protein in viral capsids. *Cell* **94**, 55-60 (1998).
14. K. K. Lee, H. Tsuruta, R. W. Hendrix, R. L. Duda, and J. E. Johnson. Cooperative reorganization of a 420 subunit virus capsid. *J. Mol. Biol.* **352**, 723-735 (2005).

15. W. R. Wikoff et al. Time-resolved molecular dynamics of bacteriophage HK97 capsid maturation interpreted by electron cryo-microscopy and X-ray crystallography. *J. Struct. Biol.* **153**, 300-306 (2006).
16. C. Helgstrand et al. The refined structure of a protein catenane: the HK97 bacteriophage capsid at 3.44 Å resolution. *J. Mol. Biol.* **334**, 885-899 (2003).
17. J. G. Mandell, A. Baerga-Ortiz, S. Akashi, K. Takio, and E. A. Komives. Solvent accessibility of the thrombin-thrombomodulin interface. *J. Mol. Biol.* **306**, 575-589 (2001).
18. C. H. Croy, S. Bergqvist, T. Huxford, G. Ghosh, and E. A. Komives. Biophysical characterization of the free IkappaBalpha ankyrin repeat domain in solution. *Protein Sci.* **13**, 1767-1777 (2004).
19. K. K. Lee et al. Evidence that a local refolding event triggers maturation of HK97 bacteriophage capsid. *J. Mol. Biol.* **340**, 419-433 (2004).
20. J. F. Conway et al. A thermally induced phase transition in a viral capsid transforms the hexamers, leaving the pentamers unchanged. *J. Struct. Biol.* **158**, 224-232 (2007).
21. R. Tuma, P. E. Prevelige, Jr., and G. J. Thomas, Jr. Mechanism of capsid maturation in a double-stranded DNA virus. *Proc. Natl. Acad. Sci. U S A* **95**, 9885-9890 (1998).
22. D. Ivanov et al. Domain-swapped dimerization of the HIV-1 capsid C-terminal domain. *Proc. Natl. Acad. Sci. U S A* **104**, 4353-4358 (2007).
23. W. Jiang et al. Coat protein fold and maturation transition of bacteriophage P22 seen at subnanometer resolutions. *Nat. Struct. Biol.* **10**, 131-135 (2003).
24. A. Fokine et al. Structural and functional similarities between the capsid proteins of bacteriophages T4 and HK97 point to a common ancestry. *Proc. Natl. Acad. Sci. U S A* **102**, 7163-7168 (2005).
25. X. Agirrezabala et al. Quasi-atomic model of bacteriophage t7 procapsid shell: insights into the structure and evolution of a basic fold. *Structure* **15**, 461-472 (2007).
26. M. C. Morais et al. Conservation of the capsid structure in tailed dsDNA bacteriophages: the pseudoatomic structure of phi29. *Mol. Cell* **18**, 149-159 (2005).

Chapter 5

Resolving the molecular reorganization and refolding of the maturing HK97 viral capsid

(The study is presented as will be submitted to the Journal of Molecular Biology.

Co-authored with Elizabeth A. Komives and John E. Johnson.)

Abstract

HK97 is a uniquely amenable system for the structural characterization of the major conformational changes associated with capsid maturation in dsDNA bacteriophage. HK97 undergoes a capsid expansion of over 100 Å in diameter associated with major subunit rearrangements during genome packaging. A previous 3.44 Å crystal structure of the mature capsid, Head II, and Cryo-EM studies of other intermediate expansion form of HK97 indicated that primarily rigid body movements facilitated the maturation process. We have recently presented a 3.65 Å structure of the pre-expanded particle form, Prohead II, which has indicated an intermediate fold in the capsid subunit that in fact undergoes significant refolding and twisting to accommodate expansion, in contrast to the previous theory implying rigid body movements. The Prohead II study focused on the major twisting motions in the P-domain, and refolding of the spine helix during the transition. In the current study, we extend the detail of the crystallographic comparison between Prohead II and Head II, characterizing the refolding events occurring in each of the four major domains of the capsid subunit, and the changes in quaternary interactions that stabilize the respective particle states. In addition, an H/²H exchange study coupled

to mass spectrometry has been performed on three distinct capsid intermediates, Prohead II, EI, and the nearly mature Head I. The study has been able to characterize differences in solvent accessibilities of the 7 quasi-equivalent capsid subunits P-II, attributed both to differences in secondary and quaternary structure. The study shows that nearly all differences in solvent exchange between subunits disappears after the first transition to EI, concomitant with symmetrization of hexon subunits, which is shown to be coupled to most of the refolding events, primarily in the spine helix and A-loop regions.

Introduction

Most complex viruses undergo a maturational transition between the initially assembled state, referred to as a procapsid, and the mature, infectious form, termed the capsid. Such systems include the bacteriophage λ , P22, and HK97, and animal viruses such as Herpes and Adenovirus, all of which package their genomes following procapsid assembly [1-5]. The capsid acts as a molecular container that houses the viral genome, and is comprised of a complex arrangement of hundreds of protein subunits. The enormity of the complex requires a fragile precursor assembly state from which further conformational rearrangements and reordering enhance virus stability [6]. Studies such as those done on bacteriophage HK97 have shown that extensive increases in buried surface area occur during the maturation process [7]. The heavily intercalated quaternary interactions that comprise the mature capsid lattice are crucial both to the virus's stability in the extra-cellular environment and in the case of dsDNA bacteriophage, to package a highly pressurized viral genome.

HK97 provides a very accessible system in which to study virus maturation. Isometric particles can be expressed in an *E.coli* expression system, devoid of the

components necessary to package the genome, yet are still able to undergo expansion and molecular rearrangement *in vitro*. The expression system contains gene products for just the capsid protein (gp5) and an internally packaged protease (gp4). The gp5 subunits first assemble into capsomers consisting of either hexons (arrangement of 6-gp5 subunits) or pentons (5-gp5 subunits)[8]. The capsomers then assemble into the initial procapsid state, Prohead I, which contains ~60 copies of gp4 [9], and lacks the portal assembly at one of the 5-fold vertices that along with an ATP driven terminase protein, normally packages a ~40kb dsDNA genome. Proteolytic cleavage of 102 N-terminal residues of each gp5 subunit following assembly results in the pre-expanded particle called Prohead II (P-II), composed entirely of 31kDa gp5* subunits. The expressed P-II capsid is unable to package DNA, but is able to expand upon treatment with low pH buffers, isobutanol and many other conditions [10], resulting in a particle that has a similar morphology and molecular topology as native phage [11,12]. During the expansion, the diameter of the particle increases from 540Å in Prohead II to 660Å in Head II (H-II). Expansion proceeds through distinct intermediate states, the first one termed Expansion Intermediate (EI) (Figure 1). The transition to EI from Prohead II has been previously characterized as a stochastic, highly cooperative two-step process [13] with an ensemble state half-life of ~3 minutes at pH 4, but that occurs on a time scale of less than 1 second per particle (data unpublished). Nearly 60% of the particle expansion occurs during this transition. Cryo-EM and crystallographic studies of the two states revealed that the hexons of P-II are skewed, with their subunits oriented in a more upright position, resulting in thicker capsid shell [7,14]. Cryo-EM data of EI has shown that symmetrization and reorganization of the hexons has occurred during this initial transition [15]. Subsequent to the

morphological change to EI, the E-loop domain of the gp5 subunit is repositioned near the P-domain of a different subunit from a neighboring capsomer. The repositioning allows a self-catalyzed isopeptide bond to form between K169 of the E-loop and N356 of the P-domain [11,12]. These cross-links occur about the 3-fold symmetry axes of the capsid. Increasing numbers of cross-links accumulate, leading to an additional expansion event and a resulting molecular topology that is similar to chainmail. Calorimetric studies have shown that crosslink formation leads to an enhanced thermal stability for the fully mature Head II structure as compared to the pre-expanded Prohead II particle [16].

Previous Raman spectroscopy data of T4 and P22 indicated significant refolding events accompanying capsid maturation [17,18], but structural studies of HK97 with Cryo-EM have implicated rigid body rotations as the predominant event that enables capsid subunit rearrangement [7]. We have solved the crystal structure of the metastable, pre-expanded Procapsid form of HK97, P-II, to 3.65Å resolution [14]. This structure can now be directly compared to the previously solved 3.44Å structure of the mature form, H-II [19,12], allowing a first near atomic resolution look into the structural changes associated with viral capsid maturation. The comparison of the two states indicated that the once hypothesized rigid body rotations are actually a tertiary twist of each subunit accompanied by the bending and refolding of a major helix. The motions of the subunits occur with respect to a set of fixed interactions at all quasi and icosahedral 3-fold sites that acted as inter-capsomer “staples.” A model was proposed implicating quaternary associations in early assembly intermediates acting to trap a high-energy intermediate state primed to undergo conformational changes upon triggering maturation. In the current manuscript we extend the scope of the study of structural transitions to

characterize the roles of each of the subunit domains in the P-II state, and during maturation. Using crystallographic comparisons of the pre-expanded and expanded states, we characterize the tertiary and quaternary contacts that preserve the intermediate folded form in P-II, the contacts that change during the maturation, and the interactions that guide the altering of the quaternary structure. Hydrogen/Deuterium exchange (H^2H) coupled to MALDI-mass spectrometry is used to follow the changes in solvent accessibility during capsid maturation and quantitatively measures the changes in solvent protection of the various domains due to refolding and or changes in binding interfaces. The combined crystallographic and H^2H data show that conformational changes and refolding occurs in all four domains of the capsid subunit (N-arm, A-domain, P-domain, and E-loop), much of which could not be seen in previous lower resolution EM studies.

Hydrogen deuterium exchange (H^2H) coupled to MALDI-mass spectrometry has been a widely used technique probing the structural environment of proteins based on the rate of backbone amide proton exchange with deuterium [20,21]. H^2H exchange has been applied to the study of a variety of protein complexes, including viral capsids, to identify regions of strict protein interactions as well as those of flexibility and disorder [22,23,24]. The method uses pepsin digestion to cleave previously deuterated protein and quantifies the solvent accessibilities of those fragments that appear in subsequent mass spectrometry analysis. H^2H data was taken for three distinct intermediate states, Prohead II, Expansion Intermediate (EI), and H-I. Crosslinking during WT expansion prevents efficient digestion and analysis by mass spectrometry, so a K169Y crosslink defective mutant was used for all H^2H studies. The K169Y mutant has been shown to undergo expansion through similar morphological intermediate states, except that it stops at the Head I form

(Figure 5.1). Crystallographic studies have shown this form to be identical to WT Balloon, and differs from H-II in that the pentons have not fully protruded [25]. The hexon conformations are very similar to H-II, however, and so the analysis of H-I is highly indicative of the behavior of the mature capsid conformation.

The changes in solvent accessibility and dynamics during expansion were measured for multiple expansion states to characterize the sequence and extent of the structural rearrangements. H^2/H data affirmed that refolding was occurring within loops of the A-domain, spine helix, and N-arm, and pinpointed that most of the changes in solvent accessibility were occurring during the first particle transition to EI. H^2/H data showed that exchange in the P-II state occurs in a non-uniform manner in certain domains, due to differences in environments and structure for the 7 quasi-equivalent subunits. The multiple modes of exchange completely disappear after the transition to EI, a state that shows overall more solvent protection than P-II and very similar protection values as the nearly mature H-I state. It therefore appears that the quasi-equivalent differences as seen in P-II have become more equivalent upon reaching EI and that the subunit has attained a nearly mature subunit conformation during this transition.

Methods and Materials

Protein Preparation and Crystallization

Crystals were grown from a W336F/ Δ -E-loop truncation mutant of P-II. Site directed mutagenesis was used to introduce the W336F mutation, a construct that is slower to expand than WT and therefore provides a more homogeneous population of

P-II particles, necessary for crystallization. A 5.2 Å structure was previously solved using the W336F construct, which identified the E-loop as a dynamic region that may prevent efficient crystal packing. Unlike its role in the H-II state, the E-loop makes no quaternary contacts with neighboring subunits, and therefore made this region a good candidate for truncation. Splicing by overlap extension was performed on the W336F construct, replacing residues 159-171 with residues APGD, a sequence known to promote formation of a reverse turn [36]. The result was a truncated loop that was fully visible in the electron density of the crystal structure. Expression of gp5 capsid protein and gp4 protease was completed in an E.coli T7 expression system and purification of HK97 Prohead II was performed as previously described [10]. The mutant assembles into an intact capsid with similar efficiency as wild type and is able to expand, though at a slower rate than WT.

The mother liquor for crystallization contained 0.1 M CHES buffer, pH 9.0, 200 mM manganese chloride and 2.3–3.0% Peg 4000. Upon mixing the manganese chloride with CHES buffer and Peg, much of the manganese chloride precipitated. The precipitant was allowed to settle overnight, and only the supernatant was used for crystallization. To 4ul of mother liquor, 1 ul of 2M NDSB-211 (additive from Hampton Research) was added. 1ul of P-II protein (25mg/ml) was added to 1ul of this mixture and allowed to crystallize by hanging drop and vapor diffusion.

Crystallographic processing

An atomic model for the P-II structure was initially derived by rigid-body fitting of the refined 3.44 Å structure of the mature H-II coordinates (Protein Data Bank 1OHG)

into the prohead II electron density. The initial phases for molecular replacement were derived from a previously solved 12 Å cryo-electron microscopy structure of P-II⁷. Crystallographic data was collected at the Advanced Photon Source synchrotron at Argonne National Laboratories, beamlines 23-IDD and 14-BMC. The Room-temperature diffraction data from twenty nine crystals was indexed, integrated, and scaled using HKL2000 suite [37]. The crystals belong to the space group, I222. Details of Indexing, scaling, structure determination, and refinement were performed as previously described [14].

Buried surface area calculations were performed on the Viper database, using crystallographic coordinates of P-II and H-II [38,39]. Bending angles calculated between the hinging N-arm and E-loop domains were calculated by deconvoluting matrices needed to align coordinates representing the refined P-II structure with that of H-II structure fit into the P-II map. The two particle states were initially least squares aligned by the P-loops (230-383), the subunit core which mainly remains identical except for local refolding in the A-loop. R.M.S. deviations of the alignment ranged from 1.1 Å to 1.3 Å.

H/²H exchange

K169Y P-II and H-I particles were expressed and purified as described previously. EI particles were formed after treating P-II particles with 10% isobutanol for 15 minutes. Isobutanol is one of many conditions that triggers in vitro particle expansion, but was used for this study due to its compatibility with MALDI mass spectrometry. Concentrated K169Y HK97 protein (20-40mg/ml) was diluted 7-fold in an 85% final D₂O concentration, buffered with 20mM tris pH 7.5, and containing either

200mM sodium chloride or 200mM potassium chloride. Samples were incubated in D₂O at pH 7.5 for various time periods followed by addition of pH 2.5 non-deuterated quench solution (final D₂O concentration of 9.0%). Following quench, all samples were handled on ice in a 4°C. Protein was digested with 50ul pepsin coated beads for 5min. The beads were removed by centrifugation, and the supernatant was flash frozen in liquid N₂. Samples were thawed individually, mixed 1:1 with alpha C matrix solution and vacuum crystallized on a Maldi plate, followed by analysis on Maldi-TOF Mass spectrometer [44]. Back exchange was calculated as 42% using the N-terminal peptide (residues 117-126), which exchanged amide protons for deuterium completely within 20 seconds. The sodium adduct of this peptide was present in much greater abundance in the spectra. All future H/²H measurements of this fragment were therefore performed on the adducted species. The total number of deuterons exchanged was calculated by subtracting the centroid of the mass envelope from the non-deuterated control from the centroid of each deuterated mass envelope. The error (standard deviation) was estimated from the average of 3 independent experiments with 2-3 measurements recorded for each experiment (Total of 6-9 measurements for each timepoint). Data analysis and kinetic plots were obtained as previously described [21].

H/²H exchange reactions were performed for up to 15 minutes, enabling measurement of amide protons exchanging at both a fast ($> 1 \text{ min}^{-1}$) and intermediate rate ($0.01 \text{ to } 1 \text{ min}^{-1}$) [24]. Longer incubations measuring slow exchange rates ($< 0.01 \text{ min}^{-1}$) were not done in this study. Amide protons exchanging at intermediate to slow rates are generally a result of solvent protection due to either secondary structure or protein-protein interactions. The measured data for all fragments was best fit to either a single or

two-exponential model accounting for deuterons exchanging at only a fast rate, or both a fast and intermediate rate respectively. The following equation represents the two-exponential fit:

$$D = N_{\text{fast}}(1 - e^{-k_{\text{fast}}t}) + N_{\text{inter}}(1 - e^{-k_{\text{inter}}t})$$

where D is the total number of deuterons exchanged at time t, N_{fast} is the number of deuterons exchanging at a fast rate, k_{fast} , and N_{inter} is the number of deuterons exchanging at an intermediate rate, k_{inter} . The fast exchanging amide protons had nearly all exchanged by the first timepoint, so k_{fast} was estimated as described previously [21]. The mass envelope of helix fragment spanning residues 204-216 could be resolved as a combination of two gaussians at early timepoints. The fitting to multiple gaussians as seen in figure 7 was performed in the software package, Origin. All fragments were identified with tandem mass spectrometry (MS/MS) using a Q-STAR mass spectrometer or a MALDI TOF-TOF (ABI 4800).

Results and Discussion

Crystallographic comparison between P-II and H-II

The 3.65 Å structure of P-II provides the first glimpse into the binding interfaces maintaining an intermediate capsid shell assembly of a dsDNA bacteriophage. 3-fold interactions were previously described as the staples maintaining inter-capsomer linkages, from which individual subunits undergo bending and refolding motions during capsid expansion. A combination of salt bridges and a putative metal binding site were implicated as the basis for the stabilization and the boundary for this fixed region is colored in blue in figure 5.2b. Figure 5.2 shows the changes in buried surface (B.S.A)

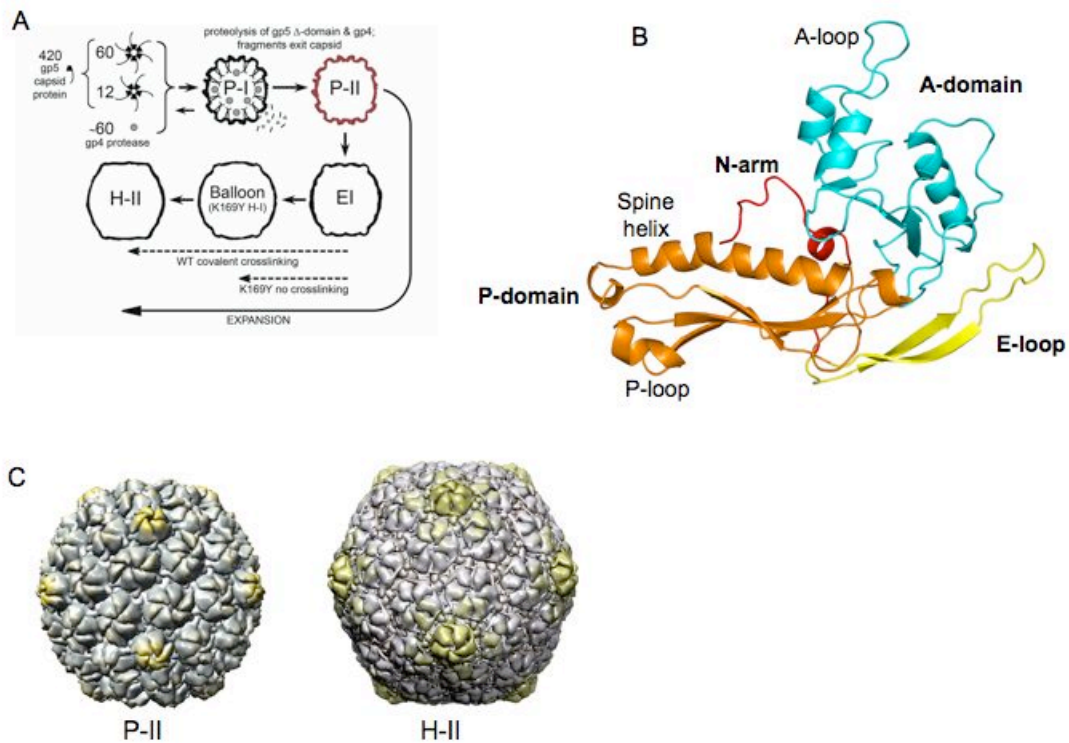
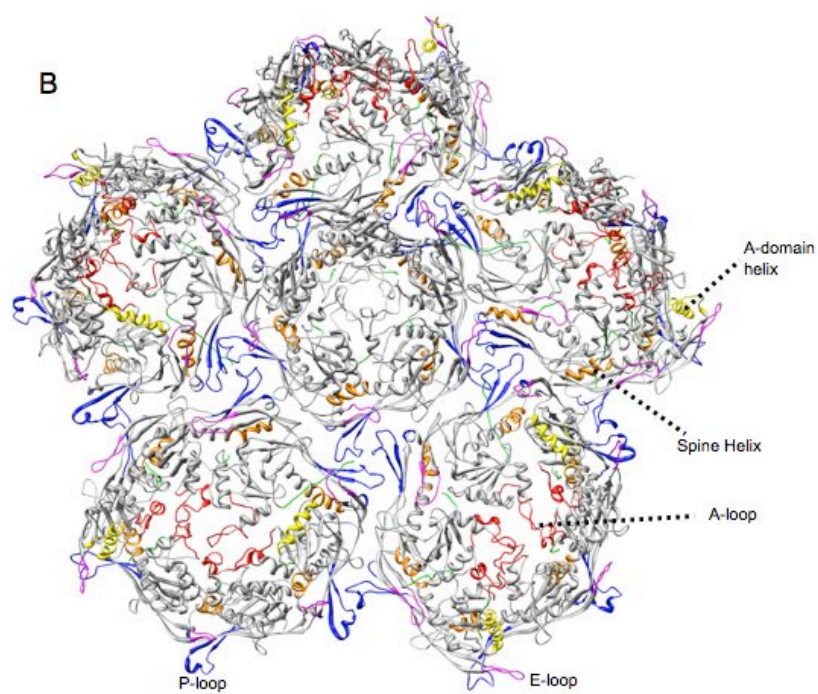
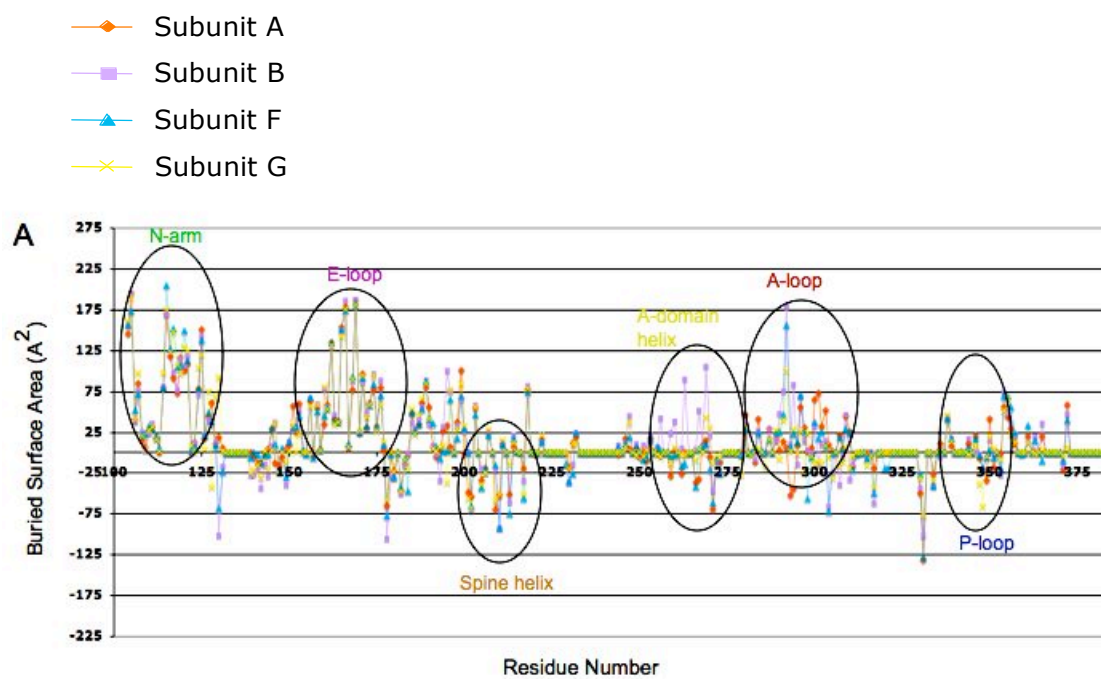


Figure 5.1: HK97 maturation. (A) The schematic depicts the assembly and expansion of HK97 in an *E.coli* expression system, lacking the portal protein and other machinery required for genome packaging. The expansion pathways of both WT and crosslink defective mutant, K169Y, are shown in response to chemical perturbation *in vitro*. (B) Structural model of P-II, subunit A, color coded by the four major domains labeled in bold. (C) P-II and H-II particle forms are shown rendered as a surface representation of the corresponding crystal structures.

Figure 5.2: **Buried surface area calculations.**(A) Graph shows the changes in buried surface area (B.S.A) of quaternary interactions for subunits A (orange trace), B (purple trace), F (sky-blue trace), and G (yellow trace) during expansion from P-II to H-II states, using calculations derived from their respective crystal structures. The color of each trace is also shown in the corresponding legend. Regions of interest are circled and labeled. The color of the label corresponds to the color of the region in panel B. (B) The structure of the P-II super-penton (5 hexons surrounding a penton) is colored in regions that have significant changes in their quaternary interfaces during expansion (excluding the P-loop, which maintains nearly identical 3-fold contacts). The structure is color-coded according to the domain labels in the graph above. Note, The N-arm undergoes a major re-ordering during expansion but is barely visible in the super-penton, as it is positioned inside the P-II capsid. The view is from the exterior of the capsid, looking down the 5-fold axis.



by residue, calculated for quaternary contacts during the transition from P-II to H-II. These were determined computationally from crystallographic coordinates of the structures. The cumulative B.S.A. comparisons between P-II and H-II for inter-capsomer and intra-capsomer contacts are presented in table 5.1. The data shows B.S.A values for four of the seven asymmetric subunits: A, B, F, and G. Subunits in the P-II hexon are not arranged with 6-fold symmetry as seen in H-II, but are instead skewed with a pseudo 2-fold arrangement. The buried surface areas of the two sets of trimers in the skewed arrangement (Figure 5.1C) are nearly identical to each other, so only three of the hexon subunits are shown for clarity. The other subunits are related by pseudo-2-fold symmetry. The changes in the penton subunit (G) are plotted as well. Regions highlighted in figure 5.2, other than the P-loop, represent areas of significant change in both the buried surface as well the intra-capsomer interface as a result of expansion. These regions include the N-arm, E-loop, A-loop (hexon subunits only), spine helix, and A-domain helix of subunits B and E.

Figure 5.3 illustrates the contacts that contribute to the A-loop conformation, with side chains interactions and distances highlighted for the subunits with different binding interfaces. Unlike the symmetric arrangement of hexon subunits in H-II where each A-loop makes extensive interactions with its neighboring subunit's A-loop, there are 3 different environments for the P-II hexon subunits. Subunits A and D sit closer to the center of the hexon, resulting in interactions mostly between its pseudo 2-fold neighbor instead of its neighboring subunit. Subunits F and C A-loops have less extensive contacts while subunits B and E are furthest from the center of the and have the least amount of buried surface area. All hexon subunits have more tenuous A-loop interactions than in

Table 5.1: **Buried surface area comparisons between 3.65 Å Prohead II crystal structure and the 3.44 Å Head II crystal structure.** Values are in Å². Measurements are based on buried surface area within each hexon and penton (intra-capsomer), or between neighboring capsomers (inter-capsomer).

Subunit contacts	Prohead II	Prohead II	Head II	Head II
	Hexamer	Pentamer	Hexamer	Pentamer
Intra-capsomer	19,529	19,212	23, 227	21, 720
Inter-capsomer	17,258	15,523	24, 950	21, 030

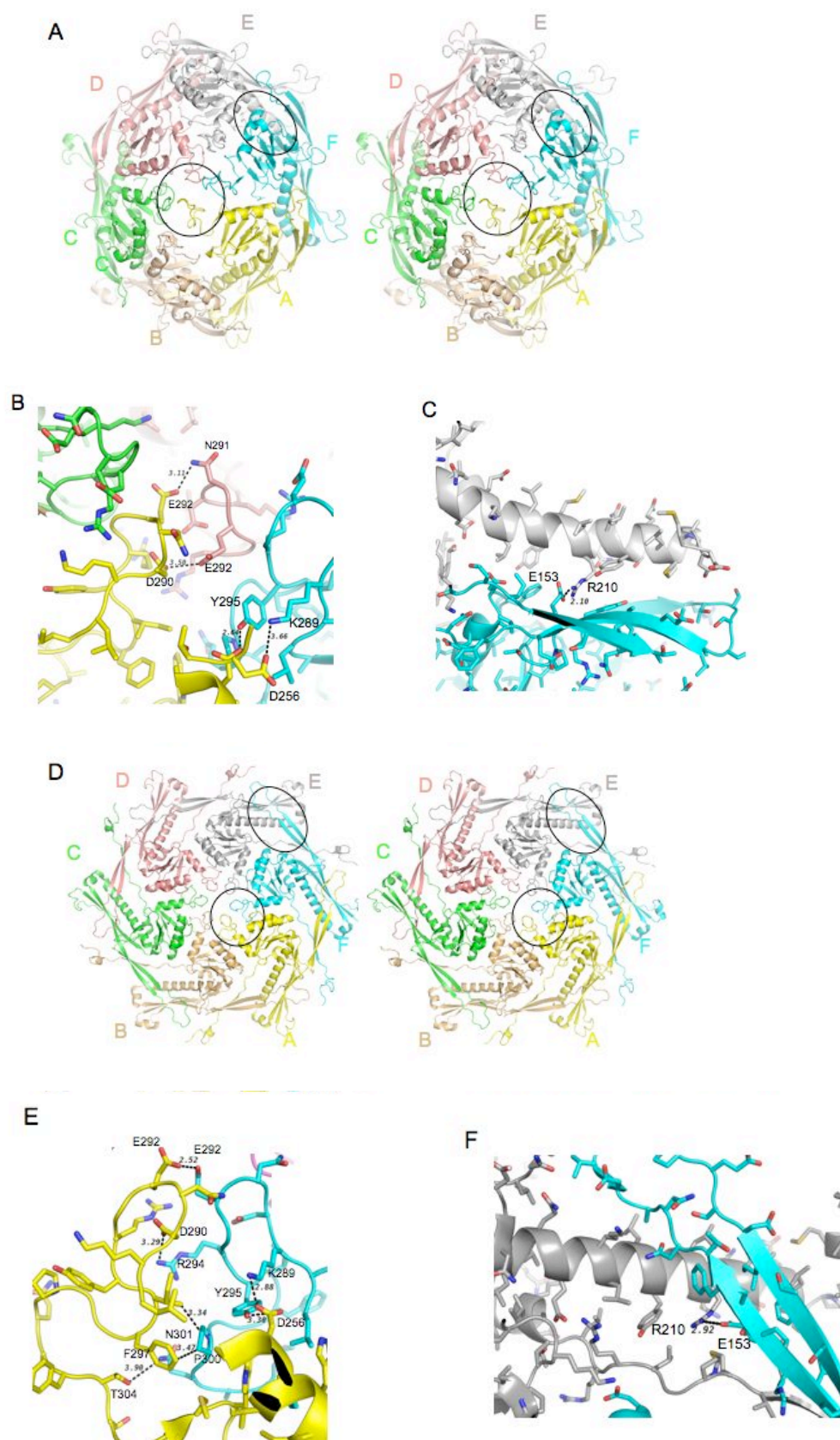
the mature form due to more numerous contacts in the H-II state as shown in Figure 5.3E. In the case of the penton subunit (G), which exhibits 5-fold symmetry in the P-II state, the B.S.A of the A-loops stay essentially constant during expansion (figure 5.2A), as do their conformations (figure 5.4). The results indicate that the differences in A-loop conformations as well as their degree of refolding is dictated by their quaternary associations in the P-II state. A-loops in the penton, which maintain a similar symmetry and similar interactions during expansion, change the least, while A-loops in the hexon, which have the greatest change in binding interface, refold the most (figure 5.4). The A-loop is therefore shown to be a highly flexible region that adapts a conformation suited for assembly of the precursor particle, but able to refold during expansion, resulting in a more buried interface. In addition to the increased interactions seen in the mature form, a sulfate was shown to be coordinated at the center of each hexon in the H-II crystal structure [19]. This would act as an additional source of intra-capsomer stabilization if present in the authentic virus. The dynamics, refolding, and distinct quasi-equivalence of the A-loop was further characterized as discussed in the H/²H exchange section.

The spine helix was previously characterized as bent and partially unfolded in the P-II state (figure 5.4). Here we show differences in buried surface area between the two states. Unlike most of the changes in the B.S.A calculations shown in figure 5.2, where most domains become more buried during expansion, residues 201-213 of the spine helix become less buried by 190-238 Å² for the different subunits. The helix of each hexon subunit in P-II lies directly below the E-loop and part of the P-domain of its neighboring subunit (figure 5.3D). There are multiple interactions including stacking of aromatic rings, a salt bridge, and hydrogen bonds that stabilize the bent conformation (figure 5.3c

and supplementary movie 1). During expansion, the majority of quaternary interactions between the helix and the neighboring subunit disappear, except for a salt bridge interaction (figure 3G) between R210 of the helix and E153 of the E-loop that likely guides the base of the E-loop into an orientation suitable for the crosslinking reaction occurring during expansion in WT particles.

Other regions implicated as having major changes in buried surface area during expansion include the N-arm, and E-loop. The N-arm is mostly disordered in the P-II crystal structure with continuous density beginning at residue 119 in subunit A, 120 for subunit E, and ranging from residues 121 to 128 for the other 5 subunits. Though disordered, significant density can be observed past a hinging region of the N-arm, which was observed to be a 60 degree hinging motion (average of the 7-subunits) between the P-II and H-II conformations. The N-arm refolds and adds a strand to a β -sheet formed with a neighboring subunit in the H-II state. The E-loop also hinges, undergoing an average of 30 degrees rotation during expansion, and is likely guided by the salt bridge between the neighboring helix as described earlier. Both of these domains are peripheral extensions that appear to have no role in maintaining the procapsid structure but play a pivotal role during maturation, in determining quaternary stability, and in the case of the E-loop, crosslink formation [15,25]. The last major binding interface that has significant buried surface area changes is the A-domain helix (colored yellow in figure 5.2), which for 5 of the 7 subunits is unaltered during expansion, but for subunits B and E, greatly increases its binding interface. The reason is due to the recessed position of its neighboring subunits, A and D, respectively (best seen in supplementary movie 2). The other 5 subunits maintain a similar interface in this region, while they rotate towards the

Figure 5.3: **Quaternary interactions of A-loops and spine helix.** (A) Stereo-view of Prohead II hexon with subunits labeled A-F. Quaternary Interactions between subunits are shown for the A-loop (center circle) of the A-domain (panel B shows zoomed in view), and the interface between the spine helix of subunit F with the neighboring subunit A (outer circle, with zoomed-in view in panel C). (D) Stereo-view of Head II hexon. (E) Quaternary interactions between subunits A and F, zoomed-in view of central circle from panel D. (F) Interface between spine helix in subunit F with subunit A, zoomed-in view of outer circle from panel D. Side chain interactions, distances, and corresponding residues are labeled in panels B and E. The salt bridge between R210 of the spine helix and E153 of the neighboring E-loop is labeled in both panels C and F.



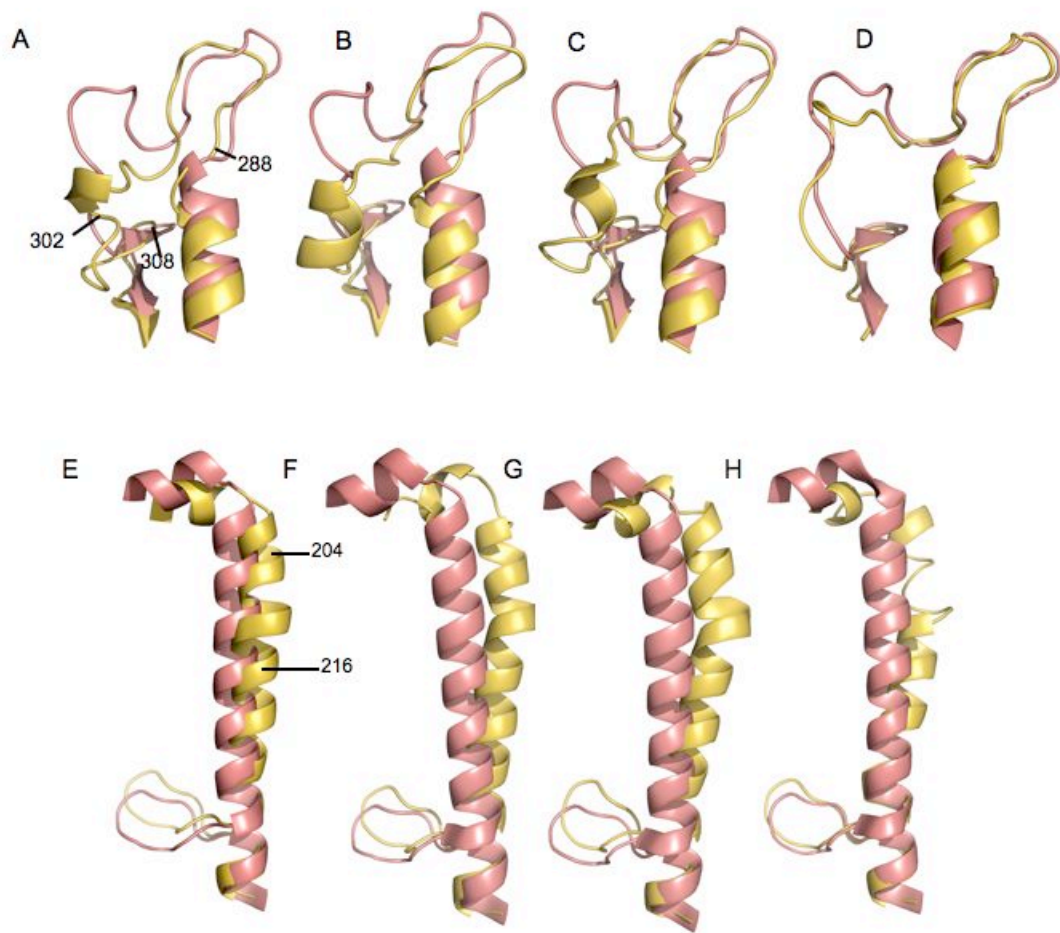


Figure 5.4: **Refolding in A-loop and spine helix (A-D)** Each subunit from P-II has been individually aligned by least squares fitting with its corresponding subunit in H-II. Panels A-D show A-loop differences for subunit A,B,F, and G respectively. (E-H) Panels showing difference between subunits A,B,F, and G in the spine helix.

orientation of subunit A,D. Supplementary movie 2 illustrates the quasi-equivalence in P-II as the subunits are oriented at different angles with respect to each other. Subunits B,C,E,and F undergo large scale rotations during transition to the EI state, while subunits A and D rotate the least, and appear to guide the final position of the other hexon subunits, resulting in a symmetric hexon conformation.

H/²H exchange

The structural refolding and dynamics occurring during expansion were monitored for 3 distinct morphological states, P-II, EI, and H-I, using H/²H exchange coupled with MALDI mass spectrometry. Though the percentage of residues analyzed for any single particle form exceeded 60%, fragments that were present and could be analyzed for all three expansion states represented ~50% of the sequence coverage. Only data for these fragments are presented in the study. Though the amount of coverage is relatively low compared to previous studies [24], the peptides covered represent nearly all of the regions that undergo refolding and conformational changes, except for the E-loop, for which peptides could not be identified. The study is novel in characterizing H/²H exchange kinetics for a trapped expansion intermediate state of a viral capsid, EI, in addition to the pre-expanded, and mature particle forms. The fragments analyzed in this study and their corresponding sequence are shown in figure 5.5. The arrows signify the start and end point of each fragment, while arrows stacked upon each other represent unique peptide fragments that have overlapping sequences. Comparing such fragments provides resolution of just the residues that extend past the overlap. The kinetics corresponding to the rate of deuterium exchange for the various peptide fragments are

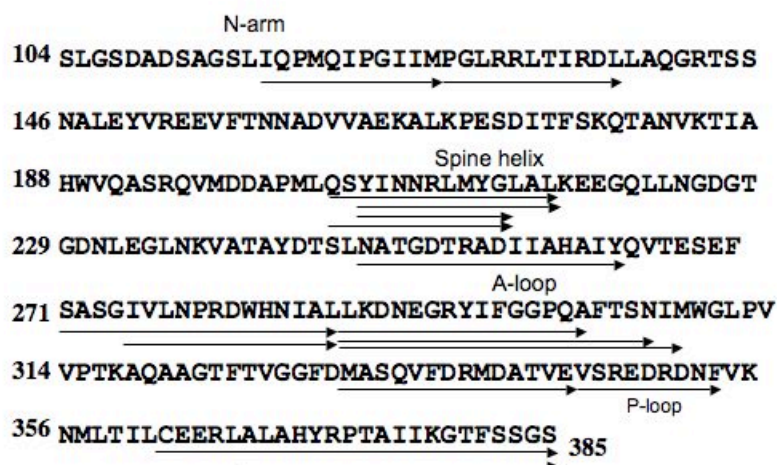


Figure 5.5: Fragments used in the H²/H exchange study for the P-II, EI, and H-I

intermediate states. Arrows represent the sequence of a single fragment. Arrows stacked upon each other represent separate fragments that have overlapping sequences.

All fragments were generated by pepsin digestion following native capsid

incubation with deuterium solution. Identical fragments spanning 138/282(49%) possible residues have been characterized for the three expansion states

plotted in figure 5.6 and are organized by domain. The regions in the subunit structure to which the fragments correspond are color coded in figure 5.7.

A striking feature when comparing the overall change in solvent accessibility between expansion states is that EI is nearly identical to the nearly fully mature H-I state for all covered regions except the N-arm. The N-arm was previously characterized by crystallography of H-I and shown to be quasi-equivalent in terms of its degree of order for the different subunits [25]. The current data supports this conclusion by showing that an N-arm region, which is known to form β -structure with a neighboring subunit's E-loop in the H-II state, is completely solvent exposed in both P-II and the EI states, only becoming more protected in the nearly mature H-I particle form. The kinetic exchange curves nearly overlap between EI and H-I, both for regions where there is no structural changes associated with expansion, as well as in regions such as the spine helix and A-loop, where refolding and conformational rearrangement occurs.

The data show that there is significant protection in the A-loop occurring during the transition to EI (figure 5.8), and the protection remains constant through to the H-I state. The A-loop is much more dynamic in the P-II state, as a fragment spanning residues 288-306 revealed that nearly 15/17 exchangeable amide protons exchange deuterium within 10 minutes. In contrast, only 10/17 amides exchange in the EI and H-I states (figure 5.6). An overlapping fragment spanning residues 288-308 show nearly identical exchange kinetics as the latter fragment, demonstrating that residues 307 and 308 are well protected from exchange. This region directly follows the residues that refold. On the other hand, comparison of another overlapping fragment, spanning residues 288-302, reveals that residues 303-306 are fully exchanged at the

Figure 5.6: **H/²H exchange kinetics.** Deuterium exchange in identified peptide fragments is plotted versus time for the P-II(■), EI(O), and H-I(Δ) states. The data is shown on the following two pages. The data is shown in columns, with the domain name or region listed at the top of the column and the corresponding fragments of the region shown below. Residue numbers pertaining to each fragment are shown on each plot. Curves were fit to either a single or double exponential rate, as described in the methods.

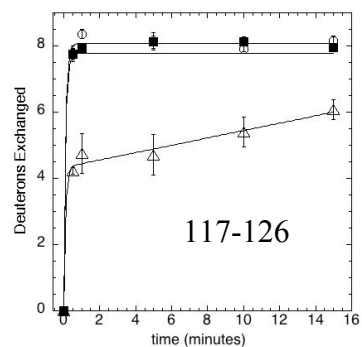
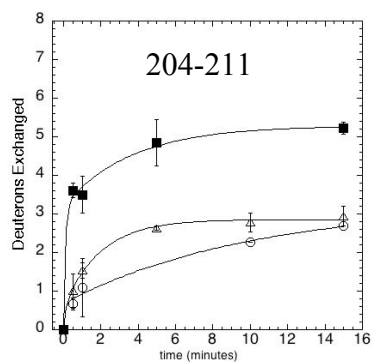
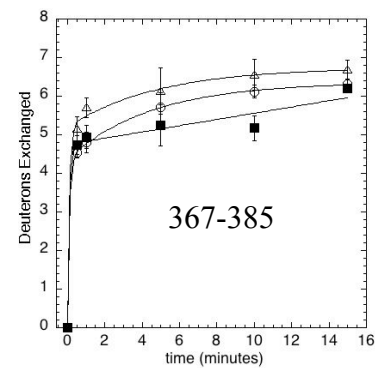
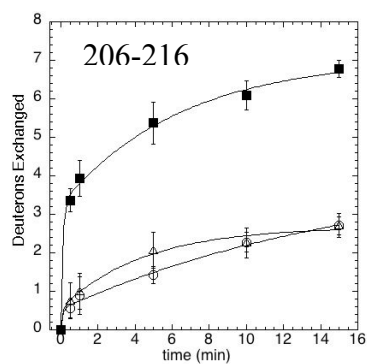
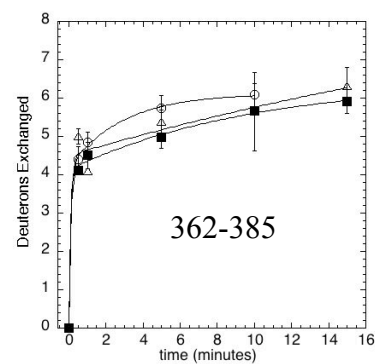
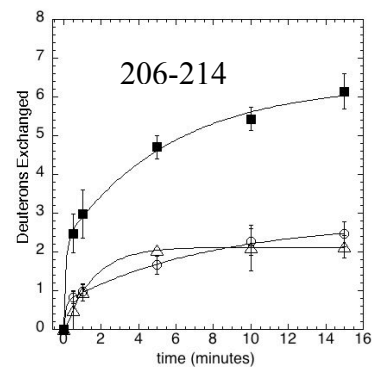
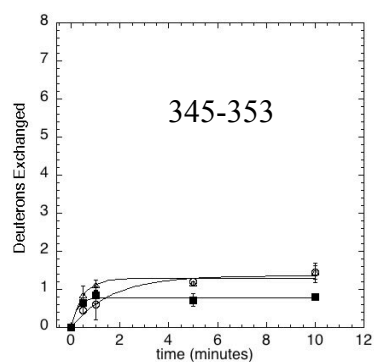
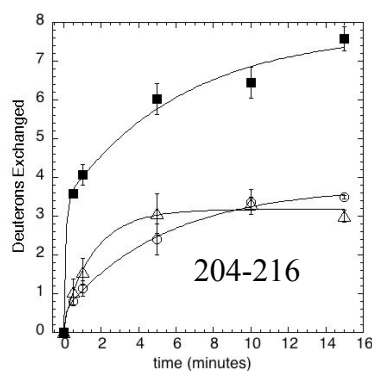
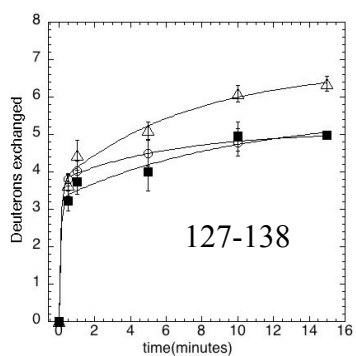
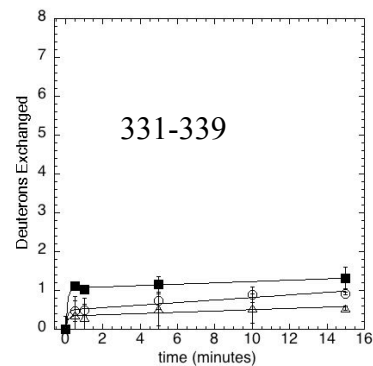
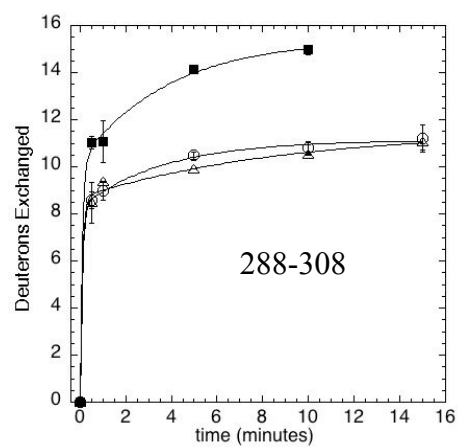
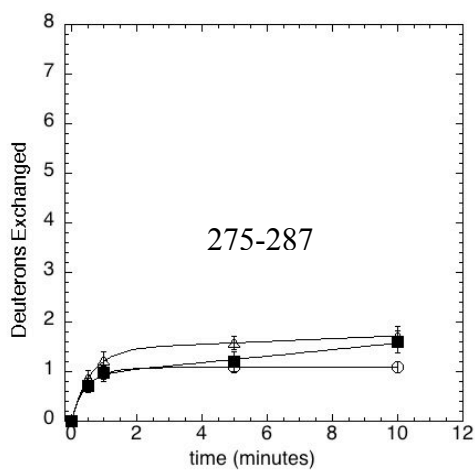
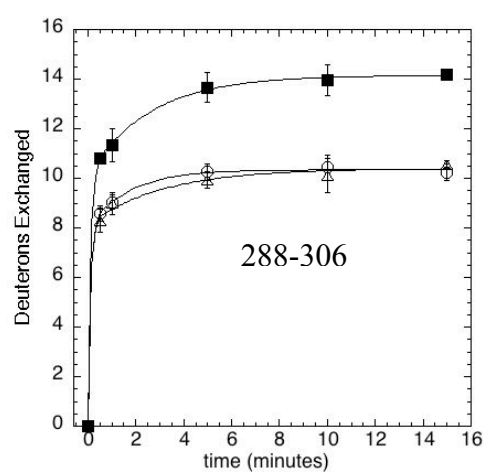
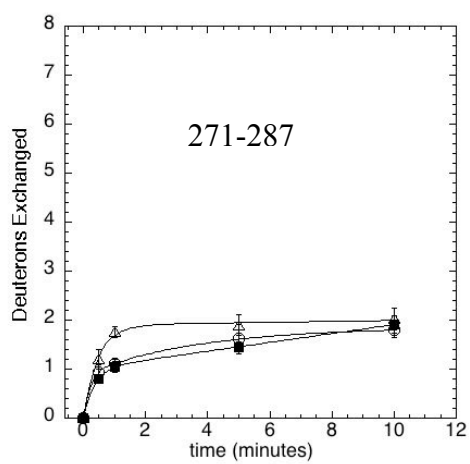
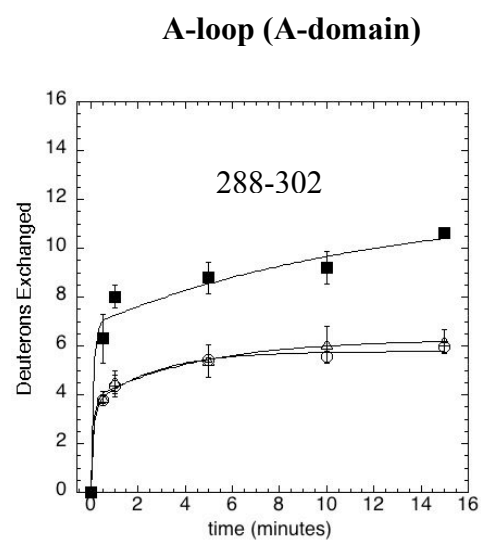
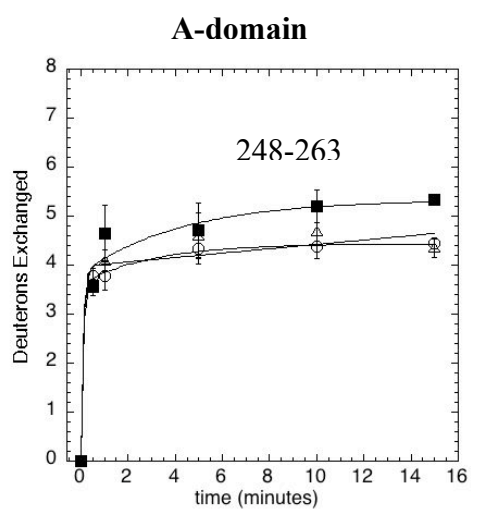
N-arm**Spine Helix****P-domain**

Figure 5.6 continued



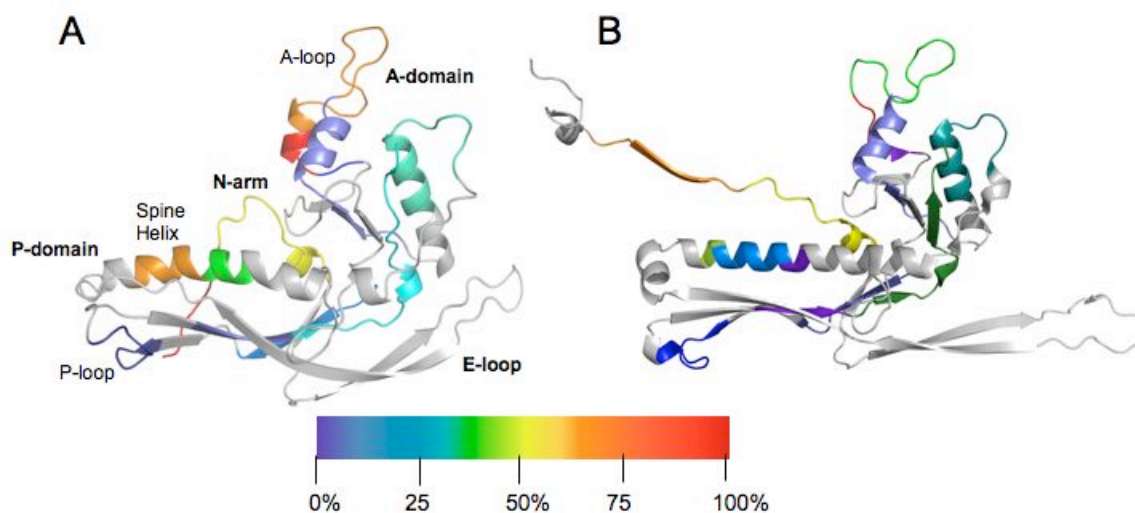


Figure 5.7: $H/{}^2H$ exchange of P-II and H-I. (a) Subunit E of P-II is color coded (spectrum below) according to the percentage of amide protons that exchange with deuterium after 10-min for each peptide fragment. Regions of overlap between peptide fragments are also color coded according to their exchange percentage. (b) Subunit E of H-II is color coded based on the percentage of amide protons exchanging in H-I state after 10-min incubations. The H-II model was used since it has been previously determined that H-II subunit E is nearly structurally identical to that of H-I.

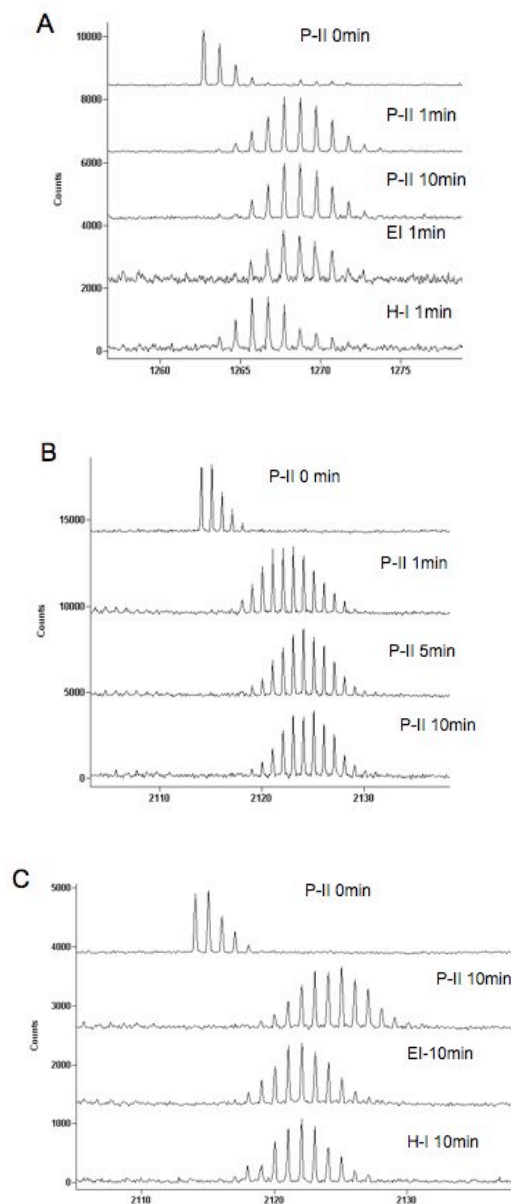


Figure 5.8: H^2/H exchange envelopes of N-arm and A-loop. (A) Mass envelopes of a fragment spanning residues 117-126 in the N-arm, shown at 1-min and 10-min time-points for P-II, and 1-min time-points for EI and H-I. (B) Mass envelopes of a fragment spanning residues 288-306 in the A-loop, shown for various time-points in the P-II state. (C) Mass envelopes shown for the A-loop fragment at different intermediate states. The mass envelopes correspond to 10-minute timepoints.

1-minute time-point (a measure of fast exchange amides) for all expansion states. These residues lie at the C-terminal region of the refolding segment of the A-loop. The data show that even though the P-II structure has a small segment of helix in the refolded region (figure 5.4) that is absent in the H-II structure, there is significant dynamics throughout the A-loop in P-II. The A-loop becomes significantly more solvent protected however (~30 %), despite the apparent unfolding of a short helix segment. The lack of difference in exchange values between EI and H-I for this loop reveals that the refolding and quaternary reorganization in this regions is occurring during the first expansion step, even though the particle is only about 60% through its expansion process. A previous EM map of EI shows the hexons display 6-fold symmetry unlike their P-II counterparts [15]. The symmetrization at the A-loop interface during the first expansion stage therefore likely results in the formation of quaternary associations that remain preserved up to H-II, though the subunits undergo additional rotational motions during the late stages of expansion. The increased dynamics in the P-II state as compared to the H-I state agree with the crystallographic data that shows a much more extensive interface between neighboring A-loops in the 6-fold symmetric mature hexon when compared to the skewed P-II hexon.

A region of particular interest in terms of understanding the structural and thermodynamic mechanism associated with expansion is the spine helix. The spine helix was previously shown to be bent and significantly more solvent exposed in the P-II state compared to the mature subunit conformation [14], indicating that the helix maintained an intermediate folded conformation in P-II. The data presented here illustrates the dynamics of the helix in greater detail, with overlapping peptide fragments resolving

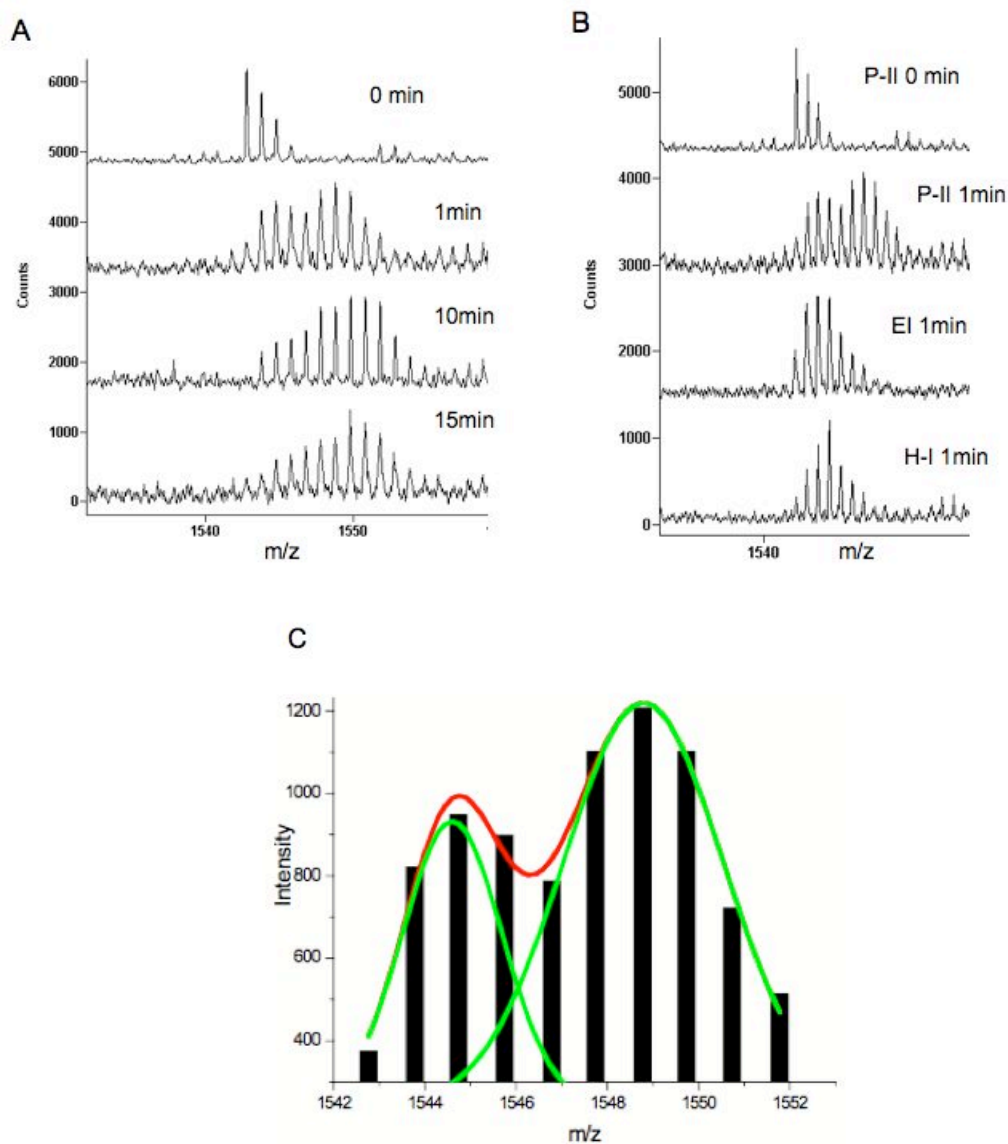
differences in exchange between different parts of the helix. A newly discovered peptide spanning an additional 2 residues that are very solvent exposed (fragment 204-216), has also allowed us to resolve the multiple Gaussians from each other at early time-points, attributed to different subunits in P-II having different rates of exchange in this region. Figure 5.9 shows the two-gaussian fit of the mass envelope of the helix fragment following a 1-minute deuterium incubation.

Bimodal distributions of solvent exchange can be attributed to two possible mechanisms. The first, which is more commonly seen in protein unfolding studies and occasionally protein dynamics studies, involves the EX1 mechanism of exchange [26,27,28]. This exchange mechanism is characterized by the following kinetic equation, $k_{\text{close}} \ll k_{\text{int}}$. The mechanism is predicated on the closing rate of a protein, (k_{close}), to be much slower than the intrinsic chemical exchange rate of amide protons (k_{int}) [29]. This may be seen in an unfolded state of a protein, or may be caused by a local conformational opening, in which all of the amide protons of the peptide fragment being analyzed have completely exchanged its protons before the closing event occurs. The result is a fully deuterated peptide fragment with no intermediate exchange profile during incremental exposures to deuterium. In many systems, the conformational opening event occurs for only a subpopulation of the protein by the time early measurements are taken. The result is a bimodal distribution of a fully deuterated subpopulation that has undergone the opening event, and a solvent protected population that has yet to open. During longer time-points, the fully deuterated envelope grows larger in respect to the protected envelope, with no gaussian distribution indicative of an intermediately

exchanged subpopulation. Macromolecular complexes that have been shown to exhibit such a transition include the scaffold for phage phi29 [30].

In the EX2 mechanism, characterized by the kinetic equation, $k_{\text{close}} \gg k_{\text{int}}$, the peptide region of interest is not fully deuterated after a single opening event, resulting in a mass envelope that progressively shifts towards higher masses during longer exposures to deuterium. By this mechanism, if all of the protein in the sample is behaving uniformly, the envelope will generally have a unimodal appearance. In the case of the P-II state of HK97, it was noticed both by the crystal structure (figure 5.4) and by $\text{H}/^2\text{H}$ exchange that the 7 quasi-equivalent subunits have different structure and environments, unlike the Head II state in which the subunits are virtually identical. The helices of the subunits that undergo the least amount of tertiary rotation, A and D, have a straighter conformation, one that is similar to their H-II form. For the other 5 subunits however, there is a large bend as well as twisting motion in the helix that occurs during expansion. The bimodal distribution is therefore a result of variation between subunits. The definitive proof of an EX2 behavior is seen in the change of the envelopes during incrementally longer exposure times. Both the more solvent exposed and solvent protected envelopes shift to the right, a scenario not possible if the behavior was EX1. The previous study characterized helix dynamics with a peptide spanning residues 206-216, which also showed an unusually wide mass envelope indicative of multiple modes of exchange, but too close together to confidently distinguish. The recently characterized peptide spanning 204-216 has a more definitive distribution, enabling a more confident fitting. Both the centroid and percentage of overall envelope area was calculated for each of the gaussians as shown in the table in figure 5.9. This was only done for the 1 minute

Figure 5.9: **H/²H exchange envelopes of spine helix fragment spanning residues 204-216.** (A) Mass envelopes of helix fragment in P-II state after 0, 1, 5, and 10 minutes of exchange. (B) Mass envelopes of the helix fragment after 1-minute of exchange for P-II, EI, and H-I expansion states. C) Bimodal envelope from P-II 1-minute spectra is fitted to 2 gaussian distributions. (D) Table showing calculations from the bimodal gaussian fit in panel C. The centroids of the resolved gaussians are shown as well as the number of amide protons corresponding to each envelope (after 0 min subtraction and back exchange correction). Finally, the number of subunits attributed to each gaussian is calculated by dividing the area of each gaussian by the total area and multiplying the value by 7, which corresponds to the number of subunits in the viral asymmetric unit.



D

	Centroid Envelope 1	Centroid Envelope 2	# amide exchanged Env. 1	# amide exchanged Env. 2	Subunits in Env. 1 (from % of area)	Subunits in Env. 2 (from % of area)
P-II helix Residues 204-216	1544.60 +/- 0.24	1548.79 +/- 0.17	0 +/- 0.41	7.05 +/- 0.29	2.13	4.83

time-point, a time frame that generally accounts for fast exchanging amides. After back exchange and side chain corrections, the more protected envelope showed that 0 amides exchanged for these subunits, while the larger, more exposed population exchanged nearly 7 deuterons. The percentage of area of the first Gaussian corrected for the number of quasi-equivalent subunits, is equal to 2.13 subunits, while the second gaussian corresponds to 4.83. The result is nearly identical to the crystallographic conclusions, which indicate that 5 of the 7 subunits have much more extensively bent helices than subunits A and D. The helix is considerably more protected upon expansion to the first intermediate, EI, and remains similarly protected in the H-I state. The bimodal envelope disappears during the first transition, indicating that the spine helices in different subunits are nearly equivalent.

Conclusions

The data presented here demonstrate complicated differences in dynamics that characterize an intermediately folded procapsid state of a dsDNA bacteriophage. The large differences in quasi-equivalent subunits is due to the initial assembly event directed by the scaffolding region of the capsid subunit, termed the Δ -domain. Once the Δ -domain is cleaved off by protease in the initially assembled P-I state, a metastable particle form, P-II, is primed to undergo a large rearrangement of protein interactions as well as tertiary and even secondary structure. Data presented here reveals a partially unfolded helix maintained by 5 of the 7 subunits. The B.S.A measurements from the crystal structures also reveal that the helix in the H-II structure has minimal interactions

with neighboring subunits, yet maintains the straight, solvent protected conformation of a canonical helix. The loss of energy in disrupting the hydrogen bonds of the spine helix is counteracted by the increased quaternary interface present in P-II. We therefore believe that the interactions described in this study counterbalance the local high energy conformations, stabilizing the pre-expanded state. Such a model explains the large energy release associated with capsid maturation measured in phage such as P22 (90kJ/mol) [31], and the sensitivity of P-II to undergo expansion, triggered by a variety of chemicals and pH changes *in vitro*. The large differences in quasi-equivalent subunits in the immature state as well as significant differences in orientation and binding interfaces of the subunits revealed a less discriminate arrangement of proteins, compared to the near perfect symmetry seen in mature capsid forms of HK97 as well as all other mature icosahedral bacteriophage. The capsid of HK97 requires extensive flexibility in regions such as the A-loops, to maintain the quaternary arrangement initially present in the procapsid, providing an inherent molecular switch capable of guiding the molecular rearrangements. The mechanism is not coupled to any chemical energy sources such as NTP's, requiring only the packaging of a fraction of the DNA to trigger the event [40,41,42], hypothesized to be a result of electrostatic repulsion with the negatively charged DNA [7]. Such an event therefore likely provides enough energy to force the particle out of a local energy minima, transitioning downhill to the more favorable capsid forms. Interestingly, the study has shown that the major refolding and structural rearrangements occurs during the first transition to the EI state, implicating the symmetrization of the hexons as the basis for the increased and more uniform solvent protection values. Once the EI state is attained, it is the crosslinking mechanism and or

E-loop conformational changes that further alter the energy landscape again, leading to additional expansion [15]. Bacteriophage with related folds such as P22, T4, Phi29, T7, and Lambda [2,32-35], may also share extensive differences in structure among quasi-equivalent subunits in the procapsid state, as documented for HK97. H^2/H exchange studies of P22 have also revealed increased solvent dynamics in the procapsid particle compared to the mature capsid form, but bimodal data illustrating major differences between subunits was not reported [24]. Since a high-resolution crystal structure of P22 and other dsDNA phage has yet to be determined, attributing such data to specific domains is difficult. All of the dsDNA phage studied have a fold homologous to HK97, and therefore also have a spine helix in their structure. We believe that it is likely that they also exhibit a partially unfolded helix in their procapsid forms, a possible mechanism for energy storage and expansion in these systems.

Acknowledgments

We would like to thank Lu Gan and Jeff Speir for help with the P-II crystal structure. We would also like to thank Kelly Lee and Rick Huang for helpful discussions.

References

1. Murialdo, T. & Dokland, H. . Structure of a viral procapsid with molecular scaffolding. *Nature* **389**, 308-313 (1997).
2. W. Jiang, Li, Z., Zhang, Z., Baker, M. L., Prevelige, P. E. & Chiu, W. Protein Fold and Maturation of Bacteriophage P22 seen at Subnanometer Resolutions. *Nature Structural Biology* **10**, 131-135 (2003).
3. A.C. Steven, Spear, P.G. . The herpes simplex virus procapsid; structure, conformational changes upon maturation, and roles of the triplex proteins VP19c and VP23 in assembly. *Journal of molecular biology* **263**, 447-462 (1996).
4. J. B. Heymann, Cheng, N., Newcomb, W.W., Trus, B.L., Brown, J.C. & Steven, A.C. Dynamics of Herpes Simplex Virus Maturation Visualized by Time-lapse Cryo Electron Microscopy. *Nature Structural Biology* **10**, 334-341 (2003).
5. J.C.M. D'Halluin. Adenovirus type 2 assembly analysed by reversible cross-linking labile intermediates. *Journal of Virology* **26**, 357-363 (1978).
6. A. C. Steven, J. B. Heymann, N. Cheng, B. L. Trus, and J. F. Conway. Virus maturation: dynamics and mechanism of a stabilizing structural transition that leads to infectivity. *Curr Opin Struct Biol* **15**, 227-236 (2005).
7. J. F. Conway et al. Virus maturation involving large subunit rotations and local refolding. *Science* **292**, 744-748 (2001).
8. Z. Xie and R. W. Hendrix. Assembly in vitro of bacteriophage HK97 proheads. *J8Mol Biol* **253**, 74-85 (1995).
9. R. L. Duda, K. Martincic, and R. W. Hendrix. Genetic basis of bacteriophage HK97 prohead assembly. *J Mol Biol* **247**, 636-647 (1995).
10. R. L. Duda et al. Structural transitions during bacteriophage HK97 head assembly. *J Mol Biol* **247**, 618-635 (1995).
11. R. L. Duda. Protein chainmail: catenated protein in viral capsids. *Cell* **94**, 55-60 (1998).
12. W. R. Wikoff et al. Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* **289**, 2129-2133 (2000).
13. K. K. Lee, H. Tsuruta, R. W. Hendrix, R. L. Duda, and J. E. Johnson. Cooperative reorganization of a 420 subunit virus capsid. *J Mol Biol* **352**, 723-735 (2005). W. R. Wikoff et al. Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* **289**, 2129-2133 (2000).

14. Gertsman et al. An Unexpected twist in viral capsid maturation. *Nature* (in press).
15. K. K. Lee et al. Virus capsid expansion driven by the capture of mobile surface loops. *Structure* **16**, 1491-1502 (2008).
16. P. D. Ross et al. A free energy cascade with locks drives assembly and maturation of bacteriophage HK97 capsid. *J Mol Biol* **364**, 512-525 (2006).
17. A. C. Steven, H. Greenstone, A. C. Bauer, and R. W. Williams. The maturation-dependent conformational change of the major capsid protein of bacteriophage T4 involves a substantial change in secondary structure. *Biochemistry* **29**, 5556-5561 (1990).
18. R. Tuma, P. E. Prevelige, Jr., and G. J. Thomas, Jr. Mechanism of capsid maturation in a double-stranded DNA virus. *Proc Natl Acad Sci U S A* **95**, 9885-9890 (1998).
19. C. Helgstrand et al. The refined structure of a protein catenane: the HK97 bacteriophage capsid at 3.44 Å resolution. *J Mol Biol* **334**, 885-899 (2003).
20. C. H. Croy, S. Bergqvist, T. Huxford, G. Ghosh, and E. A. Komives. Biophysical characterization of the free IkappaBalpha ankyrin repeat domain in solution. *Protein Sci* **13**, 1767-1777 (2004).
21. J. G. Mandell, A. Baerga-Ortiz, S. Akashi, K. Takio, and E. A. Komives. Solvent accessibility of the thrombin-thrombomodulin interface. *J Mol Biol* **306**, 575-589 (2001).
22. L. Wang and D. L. Smith. Capsid structure and dynamics of a human rhinovirus probed by hydrogen exchange mass spectrometry. *Protein Sci* **14**, 1661-1672 (2005).
23. J. Lanman et al. Key interactions in HIV-1 maturation identified by hydrogen-deuterium exchange. *Nat Struct Mol Biol* **11**, 676-677 (2004).
24. Sebyung Kang, and Previlage P.E. Jr. Domain Study of Bacteriophage P22 Coat Protein and characterization of the capsid lattice transformation by Hydrogen/Deuterium Exchange. *Journal of molecular biology* **347**, 935-948 (2005).
25. L. Gan et al. Capsid conformational sampling in HK97 maturation visualized by X-ray crystallography and cryo-EM. *Structure* **14**, 1655-1665 (2006).

26. Y. Deng and D. L. Smith. Rate and equilibrium constants for protein unfolding and refolding determined by hydrogen exchange-mass spectrometry. *Anal Biochem* **276**, 150-160 (1999).
27. L. Konermann and D. A. Simmons. Protein-folding kinetics and mechanisms studied by pulse-labeling and mass spectrometry. *Mass Spectrom Rev* **22**, 1-26 (2003).
28. D. D. Weis, T. E. Wales, J. R. Engen, M. Hotchko, and L. F. Ten Eyck. Identification and characterization of EX1 kinetics in H/D exchange mass spectrometry by peak width analysis. *J Am Soc Mass Spectrom* **17**, 1498-1509 (2006).
29. S. W. Englander, T. R. Sosnick, J. J. Englander, and L. Mayne. Mechanisms and uses of hydrogen exchange. *Curr Opin Struct Biol* **6**, 18-23 (1996).
30. C. Y. Fu and P. E. Prevelige, Jr. Dynamic motions of free and bound O29 scaffolding protein identified by hydrogen deuterium exchange mass spectrometry. *Protein Sci* **15**, 731-743 (2006).
31. M. L. Galisteo and J. King. Conformational transformations in the protein lattice of phage P22 procapsids. *Biophys J* **65**, 227-235 (1993).
32. A. Fokine et al. Structural and functional similarities between the capsid proteins of bacteriophages T4 and HK97 point to a common ancestry. *Proc Natl Acad Sci U S A* **102**, 7163-7168 (2005).
33. M. C. Morais et al. Conservation of the capsid structure in tailed dsDNA bacteriophages: the pseudoatomic structure of phi29. *Mol Cell* **18**, 149-159 (2005).
34. J. Velázquez-Muriel X. Agirrezabala, P. Gómez-Puertas, S. Scheres, J. Carazo, J. Carrascosa. Quasi-Atomic Model of Bacteriophage T7 Procapsid Shell: Insights into the Structure and Evolution of a Basic Fold. *Structure* **15**, 461-472 (2007).
35. G. C. Lander et al. Bacteriophage lambda stabilization by auxiliary protein gpD: timing, location, and mechanism of attachment determined by cryo-EM. *Structure* **16**, 1399-1406 (2008).
36. H. J. Dyson et al. Sequence requirements for stabilization of a peptide reverse turn in water solution--proline is not essential for stability. *Eur J Biochem* **255**, 462-471 (1998).
37. Z. Otwinowski, and Minor, W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology* **276**, 307-326 (1997).

38. Borelli I. A. Shepherd C. M., Lander G., Natarajan P., Siddavanahalli V., and Johnson J. E. Bajaj C., Brooks, III C. L., and Reddy V. S. VIPERdb: a relational database for structural virology. *Nucleic Acids Research* **34**, D386-D389 (2006).
39. A. Godzik, A. Kolinski, and J. Skolnick. Topology fingerprint approach to the inverse protein folding problem. *J Mol Biol* **227**, 227-238 (1992).
41. M. A. Bjornsti, B. E. Reilly, and D. L. Anderson. Morphogenesis of bacteriophage phi 29 of *Bacillus subtilis*: oriented and quantized in vitro packaging of DNA protein gp3. *J Virol* **45**, 383-396 (1983).
42. P. J. Jardine, M. C. McCormick, C. Lutze-Wallace, and D. H. Coombs. The bacteriophage T4 DNA packaging apparatus targets the unexpanded prohead. *J Mol Biol* **284**, 647-659 (1998).
43. B. Hohn. DNA sequences necessary for packaging of bacteriophage lambda DNA. *Proc Natl Acad Sci U S A* **80**, 7456-7460 (1983).
44. Mandell, J. G., Falick, A. M. & Komives, E. A. Measurement of amide hydrogen exchange by MALDI-TOF mass spectrometry. *Anal Chem* **70**, 3987-95. (1998).

Chapter 6

Exploring the relevance of 3-fold interactions in capsid assembly, stability, and expansion mechanism.

(A study in progress in collaboration with Rick Huang)

Introduction

The crystallographic comparisons between the pre-expanded P-II particles and the expanded H-II have identified quaternary interactions that appear to maintain inter-capsomer integrity during expansion at all quasi and icosahedral 3-fold axes. The major interactions that appear to facilitate the 3-fold staples include 2 sets of salt bridges and a putative metal binding site (figure 6.1). The salt bridge interactions are between residues E344 and R194, and E363 and R347. The putative binding site involves the coordination of 3 glutamates at position 348 interacting directly under the 3-fold axes. The evidence of a metal binding site in the P-II, is the presence of a very strong area of electron density around residues 348 that remains visible at very high contour levels, when the rest of the protein density has all but disappeared. Though such strong density is not present under the 3-folds in crystal structures of the late intermediates, the position of the glutamates remains nearly identical, indicating the presence of a stabilizing interaction. Despite the near identical conformations of the residues at the 3-fold interface, the rest of the subunit undergoes a very large scale twisting motion, causing hinging in all 3 P-domain

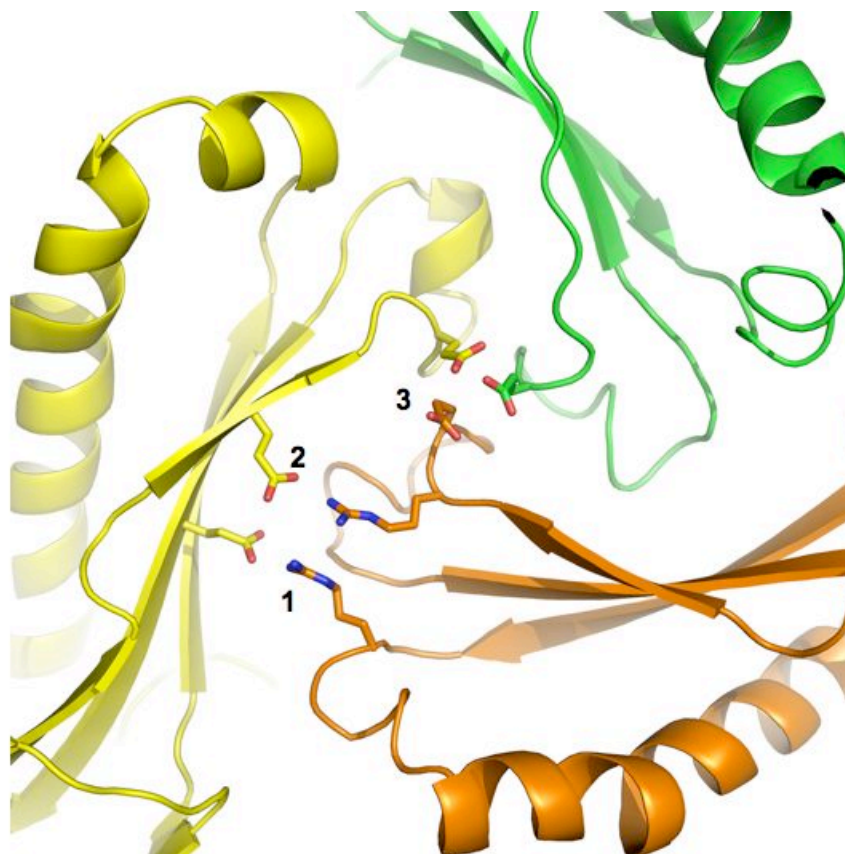


Figure 6.1: **3-fold mutations.** The P-II structural models of subunits A,F, and G, looking from the interior of the capsid through the quasi-3fold axis. Only one of the three sets of quasi-equivalent interactions has been highlighted above with the corresponding side chain residues. Five mutations have been characterized in the study: E348A (labeled 3 above), E344A (labeled 2), E344Q (labeled 2), R194N (labeled 1), and a double mutant, E344A/E348A.

β -strands. The data has suggested that these interactions may be crucial in both stabilizing inter-capsomer contacts, and the relay of tertiary motions that likely facilitates the cooperative expansion event throughout the capsid lattice. In order to explore the true role of the 3-fold interactions, we have done mutagenesis of pertinent residues and characterized the resulting morphologies, thermal stabilities, and maturation kinetics. The mutations described include E348A, E344Q, E344A, R194N, and E348A/E344A.

Methods and Materials

Mutagenesis, Protein expression, and analysis

All 3-fold point mutations were introduced using site directed mutagenesis. Protein expression and purification of all of the mutants was performed as previously described. Purification of the R194N was modified, as no particles were seen after expression and lysis of the BL21(DE3 plysS) cells. Following the lysis and subsequent removal of membrane and other cellular components by centrifugation, viral capsids are normally precipitated and pelleted after exposure to 0.5M NaCl and 6% Polyethylene glycol. R194N were in the capsomer form after cell lysis and therefore did not precipitate during this step. The supernatant from the precipitation step was further purified by filtration in a Centricon (30kD molecular weight cutoff) filter followed by size exclusion chromatography using a sephacryl S-300 column. SDS gels and western blots using HK97 capsid subunit antibodies were used for verification of protein following expression. 1% Agarose gels were used both for identification of particle forms as well as for time dependant expansion assays. A 7.5% native acrylamide gel was used for distinguishing the presence of hexons and pentons in the R194N sample.

Hexon and Penton Inter-conversion

To produce pure hexons, concentrations between 1-3mg/ml were treated with a final concentration of 20mM Tris pH 7.5, 0.5M NaSCN. Samples were incubated at room temperature overnight. The penton conversion solution consisted of 100mM Na-Citrate pH 5.6, 200mM Guanidine-HCl, 1%DMF. The final protein concentration was also between 1-3mg's/ml, and particles were also allowed to incubate overnight at room temperature. WT PH-I was disassociated with the hexon buffer above to initially disassemble the particles into capsomers [1].

Fluorescence kinetics

3ul of a 25mg/ml aliquot of particles were diluted with 50mM Sodium acetate, pH 4.0, 200mM potassium chloride (expansion buffer) [2]. The mixture was quickly aliquot into quartz cuvette followed by measurement using a fluorescence spectrometer. An excitation at 295nm was followed by an emission scan ranging from 310-400nm. The temperature in the cuvette as kept at 20° C during the entire experiment. There was approximately 30 seconds of dead time between mixing and the start of the emission scan. The data was written out in ASCII format and processed in kalaidograph. The spectral center of mass was calculated by the following equation, as previously described:

$$SCM \text{ (spectral center of mass)} = \sum (F_i \lambda_i) / \sum F_i$$

The spectral center of the emission scan is blue shifted during expansion from P-II to EI as shown in the graph in figure 4 [3]. The shift is due to the burial of tryptophan residues and occurs at a similar rate as the overall morphological change in particle size as previously determined [3]. SCM was plotted versus time for each mutant and fit to an exponential decay curve using Kalaidegraph.

Differential Scanning Calorimetry

Protein samples were diluted to 0.5 mg/ml with 100mM Potassium phosphate pH 7.5, 20mM Potassium chloride (degassed before use). Samples were heated in their reservoir cell at a rate of 1.5° K/min in a Microcal Differential scanning calorimeter. The baseline was subtracted from the respective endotherms using Origin software. The melting temperatures (T_m) of each endotherm was quantified with Origin as well.

Results

Expression and morphologies

Expression of the above mentioned mutants all resulted in capsid formation following expression in *E.coli*, except for the R194N mutant. The R194N mutant did not pellet during a salt precipitation and centrifugation step that normally causes viral capsid precipitation. Analysis of the supernatant by SDS gel showed the presence of a ~42kD band, indicative of the uncleaved gp5 subunit. All P-II preparations of the other mutants as well as WT resulted in a nearly pure population of gp5*(31kD), the product of a proteolytic cleavage of the Δ -domain inside the assembled P-I particle form. Previously, only expression of P-I particles deficient in active protease would lead to uncleaved gp5 subunits. The R194N oligomeric state was further investigated after it became clear that the protein was not in a particle form. The most likely oligomeric state was presumed to be capsomers, since soluble monomeric gp5 subunit has never been observed in previous studies. To confirm this hypothesis we analyzed the mutant both by agarose gel (not shown), and native acrylamide gel (figure 6.2B), which is commonly used to distinguish

particles from capsomers, and also distinguishing between hexon and penton states of capsomers. The native acrylamide gel was able to resolve that both hexons and pentons were present in the R194N sample. Treatment of the capsomers by previously determined conditions shifting the equilibrium to either hexon or penton populations, showed a similar result as WT capsomers (figure 6.2C). A tris pH 7.5 buffer containing 0.5M Sodium Thiocyanate shifted the equilibrium to all hexons, while a citrate buffered solution (pH 5.6) containing 0.2M Guanidine-HCL and 1%DMF produced a predominant penton population. The phenotype of R194N indicated that the Arginine 194, involved in the salt bridge near the 3-fold axis is crucial to the assembly of capsomers into full particles. Other than the fact that the mutant capsomers cannot assemble into particles, they exhibit similar behavior as WT capsomers.

The mutants E348A, E344Q, and E344A/E348A appeared to form normal P-II particles during the preparation, which were tested for maturation ability and thermal stability as described later. The E344A mutant produced P-II particles, but exhibited an unusual phenotype when pelleted by high speed centrifugation, one of the final steps used for purification and concentration. It was noticed that nearly two-thirds of the P-II particles had matured during ultra-centrifugation (figure 6.3). This would suggest

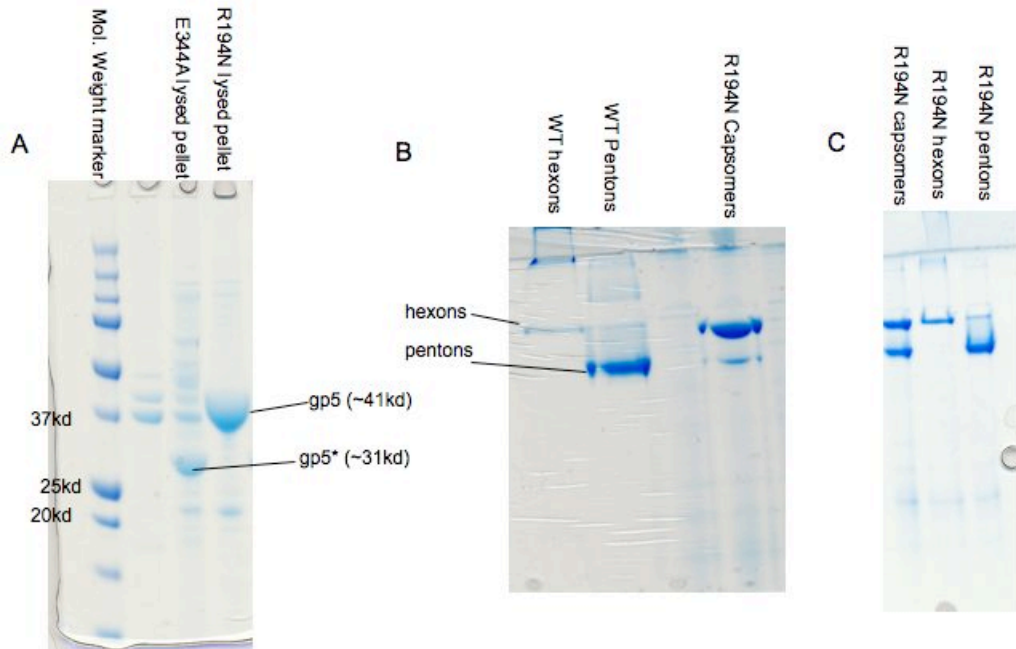


Figure 6.2: **R194N characterization.** (A) SDS Acrylamide gel comparing pellets from initial lysis following protein purification of E344A and R194N constructs (grown in E.coli.). The R194N prep only contained the uncleaved gp5 subunit unlike E344A which shows primarily gp5*. (B) Native acrylamide gel (7.5%) showing WT capsomers that have been previously treated to convert to predominately hexon and penton forms respectively. The R194N sample from initial purification shows that the oligomeric form is capsomers (mixture of hexons and pentons). (C) Native acrylamide gel (7.5%). R194N capsomers were concentrated (first lane), resulting in a change in ratio between hexon and penton. The sample was then treated with either hexon converting or penton converting conditions as WT capsomers in B.

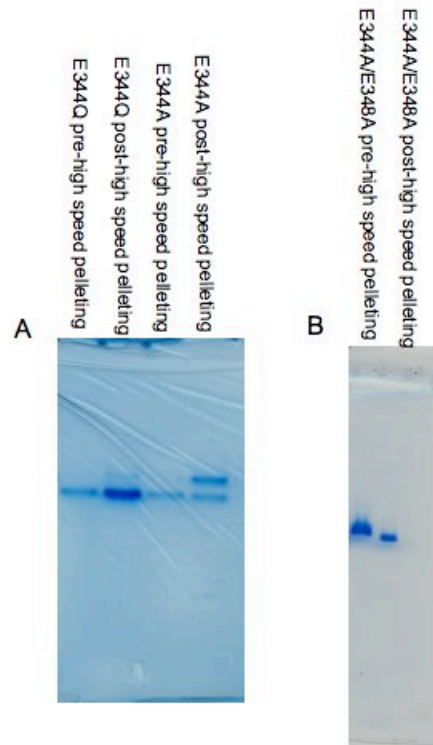


Figure 6.3: **Mechanical perturbation.** (A) 1% Agarose gel shows that ultra-centrifugation (35krpm in 50.2Ti rotor) of E344Q particles results in very slight expansion, while a much greater fractions of E344A undergoes expansion during spinning. (B) Agarose gel shows no mechanically induced expansion in double mutant, as normally expected for P-II particles.

that the E344A mutant is more metastable than WT and excessively prone to maturation, but contradicts chemical expansion assays presented later in the chapter that show that the mutant is much slower to expand during chemical perturbation. The result nonetheless indicates that the E344 mutation causes expansion during mechanical perturbation, the result of high-speed (35krpm) ultra-centrifugation normally done during the purification protocol. Though a very small percentage of the E344Q mutant also exhibited this property, it was very slight compared to E344A. The glutamine may therefore enable a hydrogen bond formation as apposed to the native salt bridge, but complete disruption of the interaction by introduction of an alanine results in a mechanically sensitive particle form.

3-fold mutant expansion

The P-II particles of all of the 3-fold mutants, except for R194N, were isolated and tested for their ability to undergo particle maturation. Low pH has been the most common treatment used to induce *in vitro* particle expansion in previous kinetic and structural studies and was the primary perturbation method in this study as well [4,11,12].

Gel analysis of low pH treated mutant and WT particles show major differences in their expansion rates (figure 6.4). All particle forms were treated with Sodium Acetate pH 4.0 for various amounts of time, and then neutralized to pH ~8.0 with Tris buffer. Upon reaching the EI state, HK97, can no longer revert to the P-II state. The assay therefore measures the amount of particles that have converted to at least the EI form, which mature further to the Head II state upon neutralization [4,5]. EI and Head II particles cannot be distinguished by agarose gel, but P-II and all other expansion states

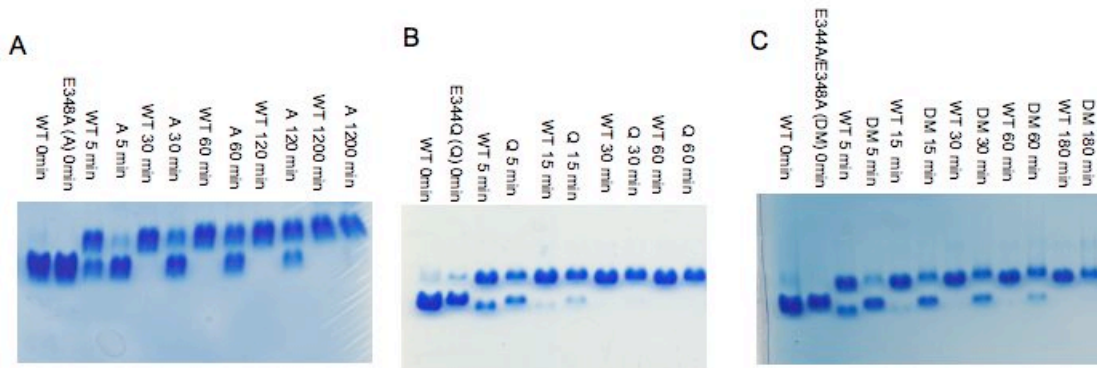


Figure 6.4: **Low pH (Acid) induced expansion assays.** (A) 1% agarose gels comparing change in particle forms as a function of time following incubation of samples in pH 4.0 buffer. The upper band represents expanded particles and the lower band represents P-II. The gel compares WT P-II and E348A P-II expansion. (B) WT-P-II and E344Q expansion, similar to A. (C) WT-P-II and E344A/E348A (referred to as double mutant or DM) expansion.

are clearly distinguishable. The gel shows that within about 10 minutes nearly all of the WT has expanded while all the other mutants require between 30 minutes and several hours for full conversion under acidic conditions. E348A matures slower than E344Q, while E344A/E348A matures the slowest.

To more accurately quantify the expansion rate, fluorescence spectroscopy was used to monitor expansion under acidic conditions (figure 6.5). Previous studies have been able to quantify population kinetics of WT expansion by measuring the change in the spectral emission curve, which is blue shifted for EI as compared to P-II particles, a result of intrinsic tryptophan burial during the event [3]. The spectral central in the emission spectra shifts from 351 nm in P-II to 348 nm in EI. The results show that the E348A, E44Q, E344A, and E344A/E348A all mature slower than WT. E344Q is the next fastest followed by E348A, then E344A, and finally the slowest E344A/E348A. The comparison in expansion rates shows that the more perturbing the mutation is to the 3-fold interactions, the slower the particle expands. For example, E344Q expands slower than WT but considerably faster than E344A. The Glu to Gln mutation likely allows a hydrogen bond to form with Arg 194 instead of the normal salt bridge, while the mutation to the alanine likely completely disrupts this interaction. The data also shows that cumulative mutations that the 3-fold further slow the expansion process. E344A/E348A is slower than both the single mutations E344A or E348A.

The isobutanol expansion data however shows some varying results (figure 6.6). During isobutanol induced expansion, the E344Q mutant matures slightly faster than WT. The E348A mutant shows virtually no expansion even at long (overnight) incubations.

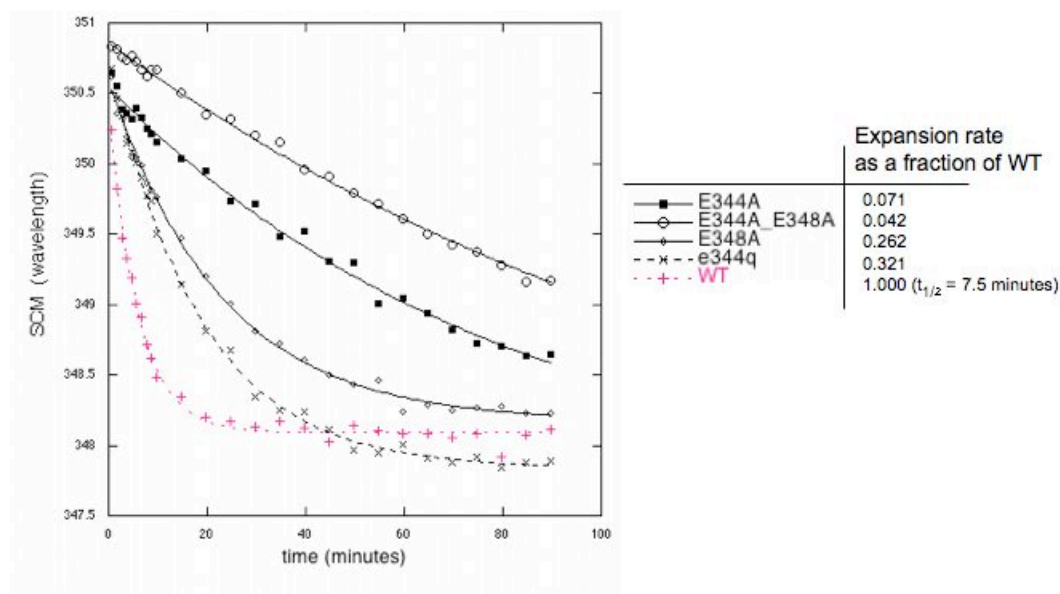


Figure 6.5: **Fluorescence monitored expansion.** P-II particles of the various mutants and WT are monitored for their change in spectral center of mass following incubation with acidic buffer (pH 4.0). The time for half of the WT particles to expand ($t_{1/2}$) is shown in the table, as are the rates for mutant expansion as a fraction of the WT expansion rate.

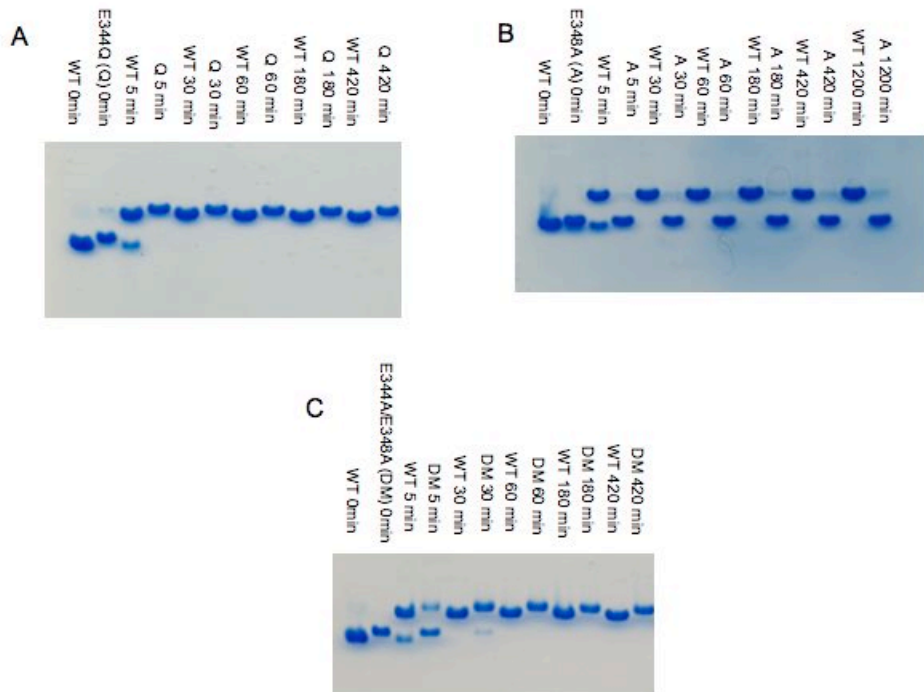


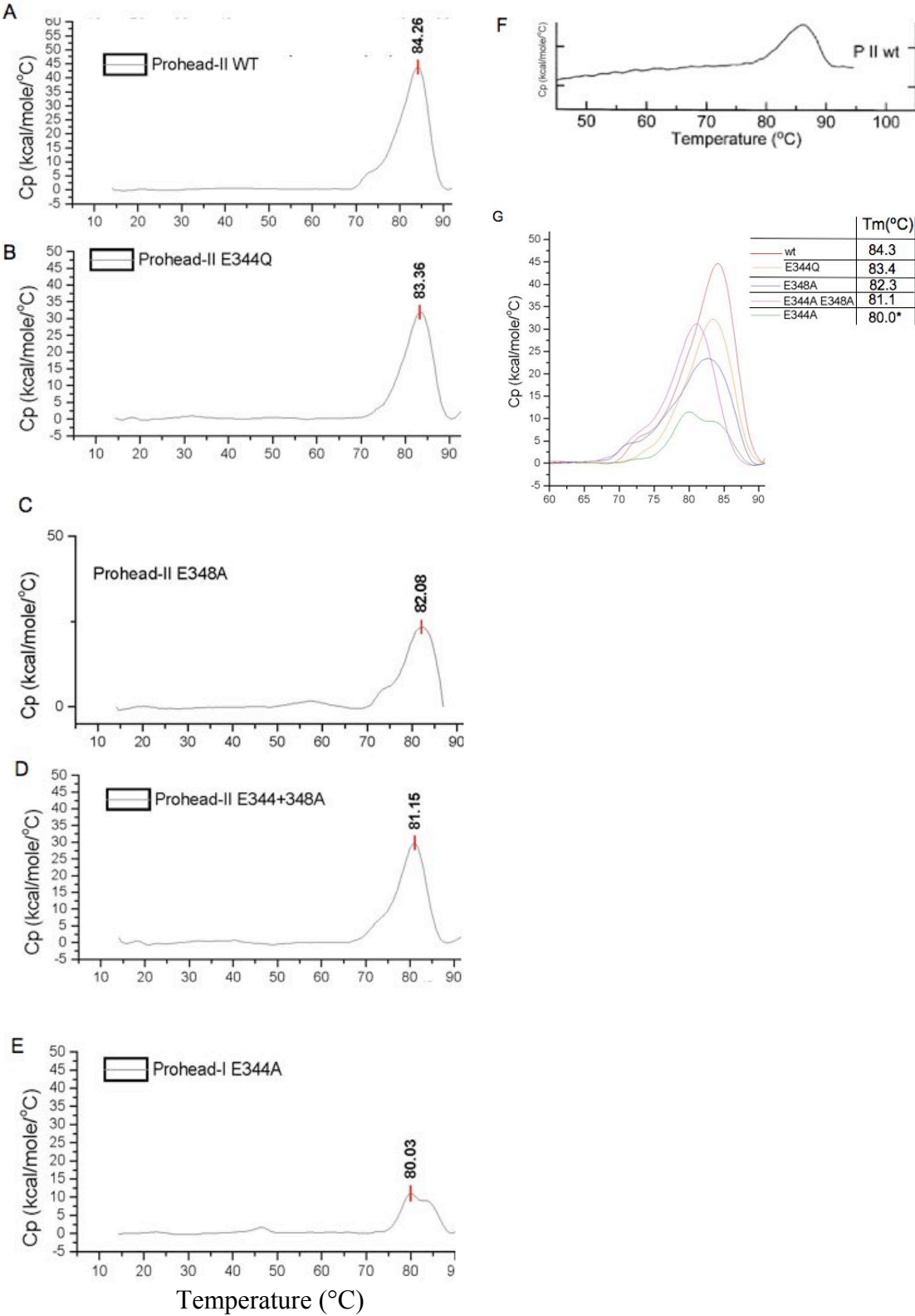
Figure 6.6: **Isobutanol induced expansion.** 1% agarose gels comparing WT expansion (after treatment with a 10% final concentration of isobutanol) with (A) E344Q mutant, (B) E348A mutant, and (C) E344A/E348A.

The E344A/E348A mutant expands slower than WT but faster than it does in low pH assays.

Differential Scanning Calorimetry

Previous calorimetry experiments compared the thermal stabilities of various expansion states of HK97 [6,7]. The studies showed that Head II has a significantly higher melting temperature than both P-II and P-I particles. The study also revealed that the crosslinks in WT H-II greatly stabilize the particle, with a significantly higher denaturation temperature compared to the mature form (H-I) of the crosslink defective K169Y mutant. We hypothesized that the 3-fold interactions were critical to particle stability and that any perturbation to these interactions would result in weaker inter-capsomer association, resulting in a less stable particle. We performed differential calorimetry scans on all of the 3-fold mutations (figure 6.7). The results show that in fact all of the mutants in fact have endotherms with lower temperatures of denaturation. The endotherms of the mutants all have a similar shape as WT except for the E344A mutant. These traces as well as WT P-II have a small shoulder in the beginning of the endotherm indicative of a conformational change or other event preceding the major melting and unfolding transitions. Previous DSC studies of WT showed no such shoulder, but did have a very broad peak (figure 6.7F). The faster heating rate of the previous study (1.98K/min) as compared to the current study (1.5K/min) may have blurred the shoulder that we can now resolve. We further investigated the shoulder by stopping a DSC scan of WT P-II at 72°C, the temperature corresponding to the peak of the shoulder (figure 6.8). Following the scan, the sample was cooled to room temperature, and then scanned to

Figure 6.7: **Differential scanning calorimetry of 3-fold mutants.** DSC scans and corresponding endotherms of (A) WT-P_{II}, (B) E344Q P-II, (C) E348A P-II, (D) E344A/E348A (double mutant), and (E) E344A. (F) A modified figure of WT-P_{II} result from previous study [6]. (G) Overlay of Scans from A-E and corresponding melting temperatures. There is an asterisk (*) next to E344A as the endotherm looks significantly different from the other mutants and WT and should not be quantitatively compared.



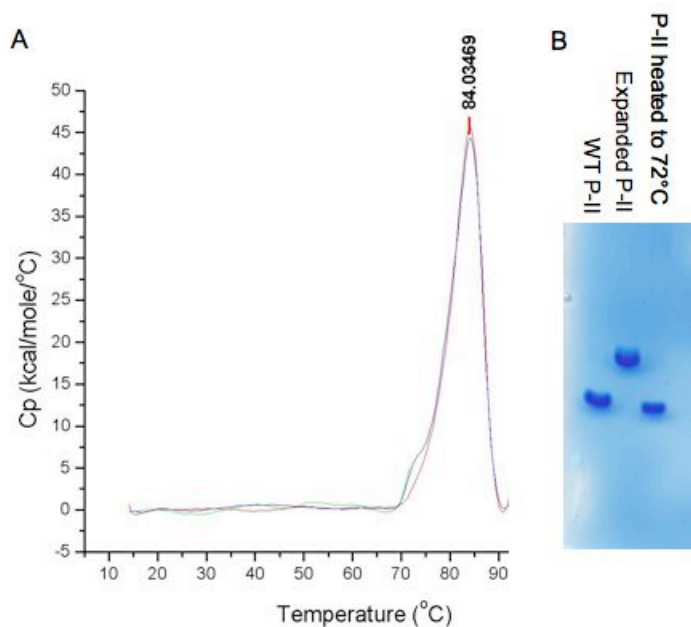


Figure 6.8: **Resolving DSC endotherms.** (A) The figure shows three overlapped traces corresponding to three separate scans. First, WT-P-II (blue trace) was heated until full denaturation, showing a small shoulder before the peak corresponding to particle melting. A fresh WT-P-II sample was scanned until reaching 72°C (green trace), then cooled to room temperature. The same sample was then re-scanned until full denaturation (red), showing that the shoulder no longer contributed to the endotherm. (B) A 1% agarose gel showing that the 72°C sample was in fact still P-II in morphology (as could be resolved by gel-chromatography), not an expanded particle, which would run in a similar place as the sample in lane 2 (an expanded control).

higher temperatures, encompassing the denaturation event. The subsequent scan no longer displayed the shoulder, exhibiting a narrower melting curve (figure x). The 72°C sample was analyzed by agarose gel to determine if the event corresponded to particle expansion. The sample ran similarly to P-II however, affirming that the shoulder did not correspond to an expansion prior to particle denaturation. This endotherm may therefore correspond to a more subtle conformational change or other unknown event which does not induce any noticeable morphological change by conventional chromatographic methods.

All of these mutants have melting temperatures of ~1-3 degrees lower than WT, indicating lower thermal stabilities. The endotherm of E344A corresponds to P-II melting but appears to be a combination of two highly overlapped peaks, with the center of the first measured at 80 degrees. Though this measurement appears to indicate a lower thermal stability for E344A as compared to the other constructs, a strict comparison cannot be made due to the differences in the shape of the endotherm, which may indicate a modified mechanism of thermal denaturation. Under the peak of each endotherm, multiple irreversible transitions are occurring, mainly the disassembly of the virus particles, and the denaturation of both capsomers and individual subunits. These transitions are occurring at very similar temperatures and are too close to be resolved.

Discussion

The crystallographic comparisons between the pre-expanded and expanded states identified several major inter-capsomer interactions that appeared to be crucial both for the assembly and the maturation properties of HK97. We have made varying mutations to the residues involved to try to perturb the interactions to varying degrees. We have

found that the mutation of R194 to an asparagine completely abolishes the assembly of the capsomers into a full capsid. Mutation to just one of the glutamates from the salt bridge (either E363 or E343) does not prevent capsid assembly. A reason that R194 may be crucial so crucial is that in one of the quasi 3-fold axes, it appears to sit between the two opposing glutamates, enabling interaction with both residues. Further mutagenesis, including a double mutant of Glu363/ Glu344, as well as an R347 mutant will be required in future studies to further characterize the importance of each of these residues in facilitating capsid assembly. The R194N data clearly demonstrates that the 3-fold interactions sites play a crucial role in capsid assembly from the initially assembled capsomers. Further investigation will also be required to determine whether the capsomers from the never assembled R194N form have an identical morphology to WT capsomers that are disassembled from the capsid state. A combination of H/²H exchange and SAXS may distinguish differences between the two forms caused by irreversible structural changes induced during capsid assembly.

The expansion rates of the 3-fold mutants were significantly decreased in comparison to WT during low pH induced expansion. Low pH was the common method for expansion in previous HK97 studies and may better mimic expansion *in vivo* than organic or other solvents. This is because HK97 virions normally expand upon the packaging of negatively charged DNA, which has been hypothesized to induce conformational changes when interacting with the negatively charged interior of the capsid [8]. Increasing the proton ion concentration may act as a perturbant by either neutralizing a charged interaction of an acidic residue, protonating and charging a histidine, or perturbing an entire interface by changing the global electrostatics. The

crystallographic data identified the 3-fold interactions as the main source of contacts between capsomers. Particle expansion was previously characterized as a highly cooperative event that likely proceeds through a wave-like motion through the capsid lattice [9]. Secondary and tertiary structural differences between the pre-expanded and mature particle forms have implicated an energy storage mechanism that results in the rotation of subunits in respect to the fixed 3-fold interactions, resulting in subunit twisting in the Beta-strands and refolding in the spine helix as well as other domains. The 3-fold interactions may therefore be crucial in a highly cooperative mechanism that relays the stress of individual subunit refolding through this inter-capsomer linkage to neighboring capsomers. Perturbing an interaction in this region would result in a decreased ability to propagate the conformational changes through the capsid lattice. The acid expansion assays of the mutants support this hypothesis of cooperativity. The results indicated that increasingly perturbing the 3-fold interactions progressively slows capsid expansion. E344Q expands slower than WT, but faster than E344A, while the double mutant E344A/E348A expands the slowest.

Another possible explanation to the mutants' slow expansion rates during acid incubation is that protonation of glutamates in the WT form is crucial for expansion. This would imply that a partial disruption or perturbation of a salt bridge for example would actually stimulate the subunit refolding and concomitant expansion. This would also imply that the 3-fold interactions that remain nearly identical between P-II and H-II are in fact more dynamic than originally hypothesized. An argument against this scenario is that the E344A mutant expands significantly slower than E344Q, and since protonation should have little significance between these two uncharged residues, this hypothesis

would not be the sole explanation for the mechanism. The differences in rates for the isobutanol based expansion as compared to the acid induced expansion does in fact imply a more complicated mechanism than that assuming the kinetics are solely based on the 3-fold mutants affecting the cooperativity of expansion. Though the method of perturbation may be different when P-II particles are exposed to isobutanol as apposed to low pH, the global mechanism of expansion should be similar, as the particles ultimately reach similar morphological states using various chemical treatments [10, Lee et al, unpublished].

We are currently pursuing methodologies that will further investigate the impact mutations to residues involved in 3-fold interactions have on both capsid stability and expansion. Further investigation are required to determine whether the 3-fold mutations are just altering the energetic barriers that allow the subunits to transition to the more mature conformation, or are they in fact crucial to the cooperativity of capsid expansion. A collaboration with Kelly Lee will be done to measure the single molecule expansion rate of the mutant particles in comparison to WT. Unpublished data has shown that a single particle of HK97 expands to the EI state within the 0.5 seconds of resolution of the fluorescence microscope used in the study (Lee et al). The mutants, especially the double mutant E344A/E348A, will be examined to see if the expansion rate of a single particle is actually affected and if intermediate forms can be trapped due to a breakdown in cooperativity.

As expected, the mutants showed decreased thermal stabilities as compared to WT in DSC experiments, alluding to the aforementioned inter-capsomer contacts playing a significant role in capsid integrity. To fully understand the role of these mutants in

capsid stability, it will be important to investigate them in the context of *in vivo* assembly and DNA packaging though. Our collaborators, Robert L. Duda, and Roger W. Hendrix (University of Pittsburgh) plan to complement HK97 capsid assembly with the mutant capsid protein to test whether such particles can withstand the packaging of the extremely dense concentration of the HK97 genome.

References

1. Xie, Z. & Hendrix, R. W. (1995). Assembly in vitro of bacteriophage HK97 proheads. *J Mol Biol* **253**, 74-85.
2. Duda, R. L., Hempel, J., Michel, H., Shabanowitz, J., Hunt, D. & Hendrix, R. W. (1995). Structural transitions during bacteriophage HK97 head assembly. *J Mol Biol* **247**, 618-35.
3. Lee, K. K., Gan, L., Tsuruta, H., Hendrix, R. W., Duda, R. L. & Johnson, J. E. (2004). Evidence that a local refolding event triggers maturation of HK97 bacteriophage capsid. *J Mol Biol* **340**, 419-33.
4. Gan, L., Conway, J. F., Firek, B. A., Cheng, N., Hendrix, R. W., Steven, A. C., Johnson, J. E. & Duda, R. L. (2004). Control of crosslinking by quaternary structure changes during bacteriophage HK97 maturation. *Mol Cell* **14**, 559-69.
5. Lee, K. K., Gan, L., Tsuruta, H., Moyer, C., Conway, J. F., Duda, R. L., Hendrix, R. W., Steven, A. C. & Johnson, J. E. (2008). Virus capsid expansion driven by the capture of mobile surface loops. *Structure* **16**, 1491-502.
6. Ross, P. D., Cheng, N., Conway, J. F., Firek, B. A., Hendrix, R. W., Duda, R. L. & Steven, A. C. (2005). Crosslinking renders bacteriophage HK97 capsid maturation irreversible and effects an essential stabilization. *Embo J* **24**, 1352-63.
7. Ross, P. D., Conway, J. F., Cheng, N., Dierkes, L., Firek, B. A., Hendrix, R. W., Steven, A. C. & Duda, R. L. (2006). A free energy cascade with locks drives assembly and maturation of bacteriophage HK97 capsid. *J Mol Biol* **364**, 512-25.
8. Lee, K. K., Tsuruta, H., Hendrix, R. W., Duda, R. L. & Johnson, J. E. (2005). Cooperative reorganization of a 420 subunit virus capsid. *J Mol Biol* **352**, 723-35.
9. Conway, J. F., Wikoff, W. R., Cheng, N., Duda, R. L., Hendrix, R. W., Johnson, J. E. & Steven, A. C. (2001). Virus maturation involving large subunit rotations and local refolding. *Science* **292**, 744-8.
10. Duda, R. L., Hempel, J., Michel, H., Shabanowitz, J., Hunt, D. & Hendrix, R. W. (1995). Structural transitions during bacteriophage HK97 head assembly. *J Mol Biol* **247**, 618-35.
11. Lata, R., Conway, J. F., Cheng, N., Duda, R. L., Hendrix, R. W., Wikoff, W. R., Johnson, J. E., Tsuruta, H. & Steven, A. C. (2000). Maturation dynamics of a viral capsid: visualization of transitional intermediate states. *Cell* **100**, 253-63.

12. Duda, R. L. (1998). Protein chainmail: catenated protein in viral capsids. *Cell* **94**, 55-60.

Chapter 7

Thesis Conclusions

HK97 has provided a unique model system in which to study the structural details associated with dsDNA Bacteriophage capsid expansion. The complex reorganization of hundreds of subunits along with a capsid expansion of more than 100 Å has been able to be characterized using multiple biophysical techniques, most notably X-ray crystallography. A 3.44 Å crystal structure of the mature capsid, H-II, provided a near atomic resolution image of the fold, quaternary arrangement, and chemistry that provides the extraordinary stability required to package over 500mg/ml of packaged genome.

Previous Cryo-EM studies of the pre-expanded, P-II particle, implicated mostly rigid body movements of subunits accommodating the large-scale subunit reorganization associated with capsid maturation. The data described in these studies reveal a more complex structural transformation that involves not only changes in quaternary interactions, but changes in secondary and tertiary structure, whose refolding and hinging guides the global conformational transitions. The P-II crystal structure indicated a set of 3-fold interactions seen for all subunits that maintain a fixed contact point between capsomers. The 3-fold interactions, most notably salt bridges as described in chapter 6, leverage a twisting motion in the Beta-strands of each P-domain and a concomitant refolding of the spine helix, structural changes that facilitate between ~20 and ~40 degrees of rotations for the 7 quasi-equivalent subunits. In addition, loops in the A-domain refold as well in response to newly acquired quaternary interaction in the expanded conformation. Both the crystallographic and H/²H exchange data presented

depict the P-II conformation as an intermediate folded state with a locally strained helix conformation preserved by tenuous quaternary contacts, which when perturbed using a variety of chemical condition, spring procapsid expansion to the EI particle form. H^2H exchange has shown that much of the structural changes including refolding of the helix have occurred during this first, irreversible conformational change.

A closely related bacteriophage, P22, shares the same fold as HK97 as seen by Cryo-EM. The thermodynamics of expansion of P22 has been studied with DSC, allowing for the quantitation of the exothermic expansion process, measured at 90kJ/mol [1]. A similar quantitation of energy release in HK97 could not be measured, since HK97 does not undergo heat induced expansion as in the case of P22. ITC measurements of expansion have also proven difficult to collect, as pH changes and the addition of expansion inducing reagents causes perturbations in the ITC baseline measurements. Though direct thermodynamic measurements cannot be quantified for HK97, there are likely many similarities between the expansion processes in the two systems. Raman spectroscopy data of P22 previously indicated an intermediate folded state for the procapsid particle form, which was hypothesized to undergo refolding during expansion.

The structural data presented in the thesis provides near atomic resolution evidence that the procapsids of HK97, and possibly many other related bacteriophage such as P22, lambda, T4, as well as Herpes virus, which all share an HK97 fold, exhibit an intermediate trapped subunit fold that likely provides the energy required for capsid reorganization. In the case of HK97, it is the Δ -domain and the quaternary interactions within both the capsomers and capsid assembly states that preserves local strain in the subunit fold. The subunits in the procapsid are quasi-equivalent though and show

different amounts of helix kinking and bending motions. As shown in the bimodal hydrogen deuterium exchange profile of the spine helix in chapter 5, 5 of the 7 subunits are highly solvent accessible, while the other 2 show near complete protection, data which correlates well with the crystal structures that implicate 5 of the 7 subunits as having major kinks in their spine helix. Though there is no high-resolution data of the Δ -domain from either the P-I or capsomers state, Cryo-Em data implicated a dimer of Δ -domain trimers interacting underneath each hexamer in the P-I state. The interactions are likely inducing the skew in the hexamers, seen in both the P-I and P-II particle forms, resulting in different conformations and different degrees of strain between the subunits. Particles are not prone to chemically induced expansion from the P-I state. When particles are expressed with functional protease, the Δ -domains of each subunit are cleaved, resulting in the metastable particle form P-II that can be triggered to mature using a variety of conditions. The increased metastability of P-II as compared to P-I is likely in large part due to the absence of the Δ -domain quaternary interactions. The transition energy for expansion is therefore lowered, and the twisting and refolding of subunits guides the movement of the subunits during the process. Most related bacteriophage as well as herpes virus have separate scaffolding proteins required for capsid assembly, unlike HK97 which uses the N-terminal region of the coat protein, the Δ -domain, to serve this function. Scaffolding proteins or scaffolding domains in viral assemblies may therefore serve an additional role to that of just guiding capsid assembly. They may in fact enforce and stabilize high-energy intermediate conformations that serves as an energy storage mechanism for the highly exothermic maturation process. In all of the aforementioned systems that package their genomes, their scaffolding proteins

exit prior to capsid maturation. The packaging of the genome may therefore displace the scaffolding proteins and perturb the high-energy capsid conformation.

Interestingly, there are other protein systems other than dsDNA bacteriophage that use conformational strain, with notably kinked helices, as a source for conformational rearrangements in macromolecular complexes. Examples include a variety of ion channel protein complexes, and the CA domain of the HIV capsid protein. Structural studies of potassium channels have shown drastically different conformations corresponding to open, or conductive states, and the closed or non-conductive state [2,3]. The gating mechanism is guided by bending of the helices, resulting in the open form of the channel. Studies have characterized the closed state, with nearly straight inner helices and close helix packing interactions, as the preferred or lower energy state [2]. They believe that the domain that acts as a voltage sensor overcomes this favorable conformation using energy stored in the electric field of the membrane. Subsequent re-polarization results in conformational switching back to the closed state. Helix kinking has been implicated in the gating mechanism of other ion channels as well including the sodium channels [4], the nicotinic acetylcholine receptor [5], and the mechanosensitive channel (MscS) in prokaryotes to name a few [6]. Studies of HIV Gag protein have revealed a kinked helix in the C-terminal domain of the capsid monomer which is straightened in the crystal structure of a domain swapped dimer, a hypothesized intermediate in Gag oligomerization [7]. The study hypothesized that a release in strain during helix straightening provided the energy needed for the formation of the domain swapped dimer. Though the functions of these aforementioned macromolecular complexes may be different, the idea of strained protein conformations providing a

driving force in so called “molecular machines” may in fact be more common phenomenon than expected. Molecular motors such as myosin and kinesin rely on NTP hydrolysis for functional energy, but other protein systems such as capsid shells in maturation dependant viruses may instead harness their energy within intermediate folded and assembly states.

Finally, one of the major differences seen between conformational changes in ion channels for example, and protein capsids in dsDNA bacteriophage, is reversibility. The gating mechanism in ion channels depends on a highly reversible opening and closing of protein conformations, a necessity for its function in maintaining proper membrane polarization. On the other hand, it appears that for HK97 capsid expansion, the process is only one-way. After P-II expands to the EI state, the process cannot be reversed by any known treatment or conditions. Though the free energy between the bent subunit conformation in P-II and the mature H-II subunit is likely within a few kilocalories/mol, the high degree of cooperativity between hundreds of capsids subunit makes reversibility highly improbable. Though one or possibly several subunits may be able to refold between P-II and H-II, once the conformational changes are propagated through the capsid lattice, the energy barrier required to convert hundreds of subunits back to their P-II form is likely too great to induce by conventional methods. The one-way conformational change is well-suited to the function of HK97 and other dsDNA bacteriophage, as their viral capsids are not required to disassemble in order to release their genome, but instead inject their host through a single five fold vertex. The highly stable and energetically favorable conformation of the mature particle is therefore more

of a necessity to dsDNA bacteriophage than the conformational flexibility required for reversibility.

References

1. Galisteo, M. L. & King, J. (1993). Conformational transformations in the protein lattice of phage P22 procapsids. *Biophys J* **65**, 227-35.
2. Yifrach, O. & MacKinnon, R. (2002). Energetics of pore opening in a voltage-gated K(+) channel. *Cell* **111**, 231-9.
3. Jiang, Y., Lee, A., Chen, J., Cadene, M., Chait, B. T. & MacKinnon, R. (2002). The open pore conformation of potassium channels. *Nature* **417**, 523-6.
4. Zhao, Y., Yarov-Yarovoy, V., Scheuer, T. & Catterall, W. A. (2004). A gating hinge in Na⁺ channels; a molecular switch for electrical signaling. *Neuron* **41**, 859-65.
5. Law, R. J., Forrest, L. R., Ranatunga, K. M., La Rocca, P., Tieleman, D. P. & Sansom, M. S. (2000). Structure and dynamics of the pore-lining helix of the nicotinic receptor: MD simulations in water, lipid bilayers, and transbilayer bundles. *Proteins* **39**, 47-55.
6. Akitake, B., Anishkin, A., Liu, N. & Sukharev, S. (2007). Straightening and sequential buckling of the pore-lining helices define the gating cycle of MscS. *Nat Struct Mol Biol* **14**, 1141-9.
7. Ivanov, D., Tsodikov, O. V., Kasanov, J., Ellenberger, T., Wagner, G. & Collins, T. (2007). Domain-swapped dimerization of the HIV-1 capsid C-terminal domain. *Proc Natl Acad Sci U S A* **104**, 4353-8.

Appendix A:

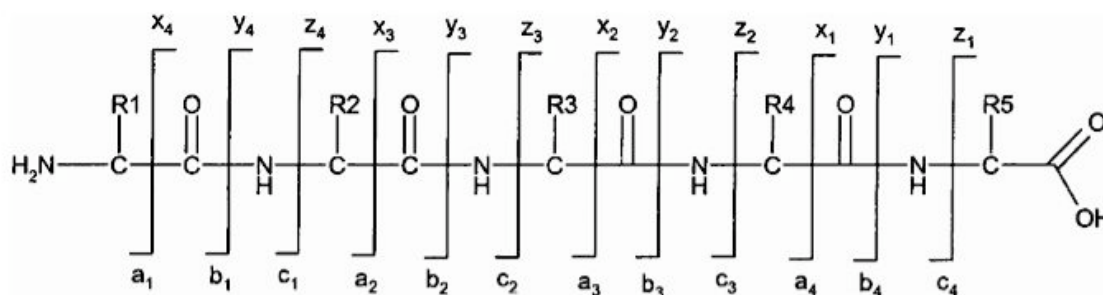
Peptide identification by Tandem Mass Spectrometry

To characterize the pepsin digested fragments in the H/²H study, the identity of the peptide fragments had to first be identified. After calibration of the maldi voyager, the mass accuracy of the peptide fragments less than 4000 daltons in weight were usually within 50 parts per million (ppm). A random theoretical digestion was performed for the gp5* subunit (capsid subunit of HK97) using the online server called “Prospector” (<http://prospector.ucsf.edu>). The mass of each peak seen in the spectra of pepsin digested HK97 was matched against theoretical peptides from simulated cleavages within a certain error range. The following shows an example of a search for all possible fragments that match the mass of a flown peptide that has a m/z of 1560.82, using a 50ppm cutoff:

m/z	MH+	Intensity	Delta	Start	End	Missed	Sequence
Submitted	Matched		ppm			Cleavages	
1560.8200	1560.7500	100.0	44.8	226	241	0	(N)GDGTGDNLEGLNKVAT(A)
1560.8200	1560.7509	100.0	44.3	190	203	0	(W)VQASRQVMDDAPML(Q)
1560.8200	1560.7879	100.0	20.6	300	313	0	(G)PQAFTSNIMWGLPV(V)
1560.8200	1560.7938	100.0	16.8	38	50	0	(K)QLQSDLMKVQEEL(T)
1560.8200	1560.8414	100.0	-13.7	33	46	0	(T)GQVSKQLQSDLMKV(Q)
1560.8200	1560.8645	100.0	-28.5	275	287	0	(G)IVLNPRDWHNIAL(L)
1560.8200	1560.8645	100.0	-28.5	276	288	0	(I)VLNPRDWHNIAL(K)

Though some of the theoretical fragments have closer matches than others, it is crucial to positively identify the correct fragment. Tandem mass spectrometry (ms/ms) is required to accomplish this task. All of the ms/ms data for the study was performed either on a Qstar or MALDI ToF/ToF mass spectrometer. In the Qstar instrument, the process involves isolating the fragment of interest in the first quadrapole followed by collision induced dissociation (CID) in the second quadrapole. CID involves bombardment of the isolated peptide with a collision gas, in this case argon. Various

covalent bonds of the peptide will be broken and chemical modifications of the precursor ions may result as well. The fragmentation pattern is predictable though and can be matched against theoretical fragmentations of peptide sequences. For the example above, the fragmentation pattern of the seven sequences that represent potential matches are compared the experimental fragmentation of the isolated peptide fragment. The most common fragments stem from breakage of the amide bond, resulting in y^+ and b^+ ions, as well as internal ions. These and other common fragments are shown below for an arbitrary peptide:



(figure is adapted from: Paizs, B. & Suhai, S. (2005). Fragmentation pathways of protonated peptides. *Mass Spectrom Rev* **24**, 508-48.)

The ms/ms spectra of the primary peptides used in the study are shown in figures 1-9. The daughter ions of the fragmentation are identified by comparisons to theoretical fragmentations of closely matching sequences as described above. Most of the fragment peptides are labeled as b^+ and y^+ ions though a^+ , z^+ and modified fragments (i.e. loss of water, ammonia, or other modification) are also labeled on the spectra as identified. The most challenging peptide to identify was a region in the N-arm (residues 117-126) which flew very poorly in its normally protonated state at a mass of ~ 1240 , but flew exceptionally well as a sodium adducted species at 1262. Fragmentations of the sodium

adducted ion produced an unusual fragmentation pattern consisting of mostly a^+ ions, many of which carried a sodium ion. Figure 2 shows the identified daughter ions of the fragmented adduct.

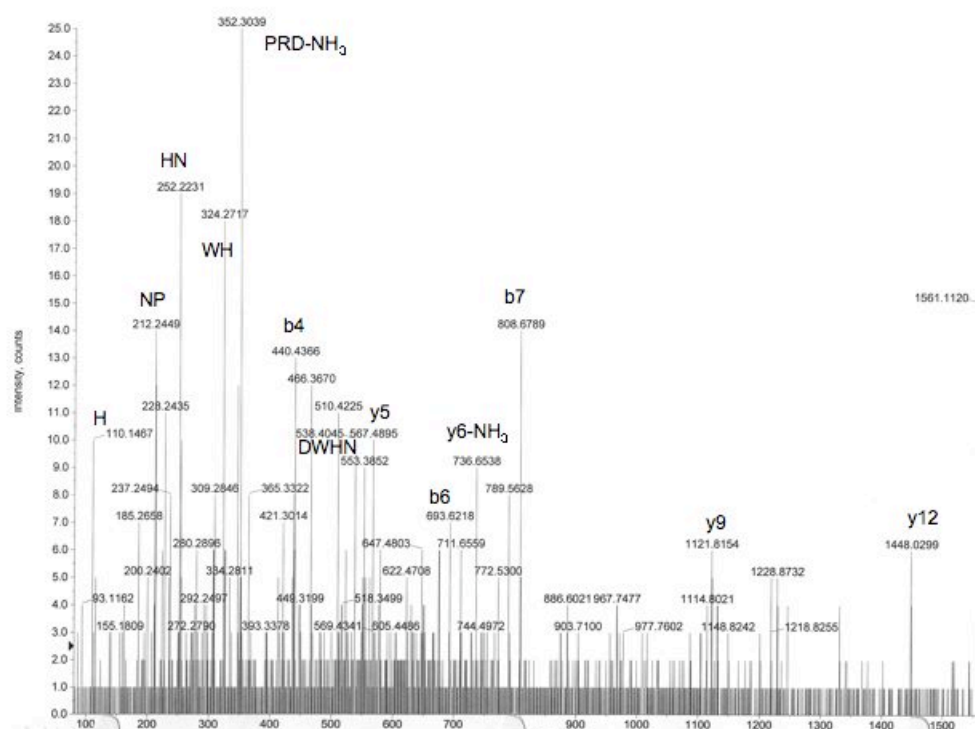
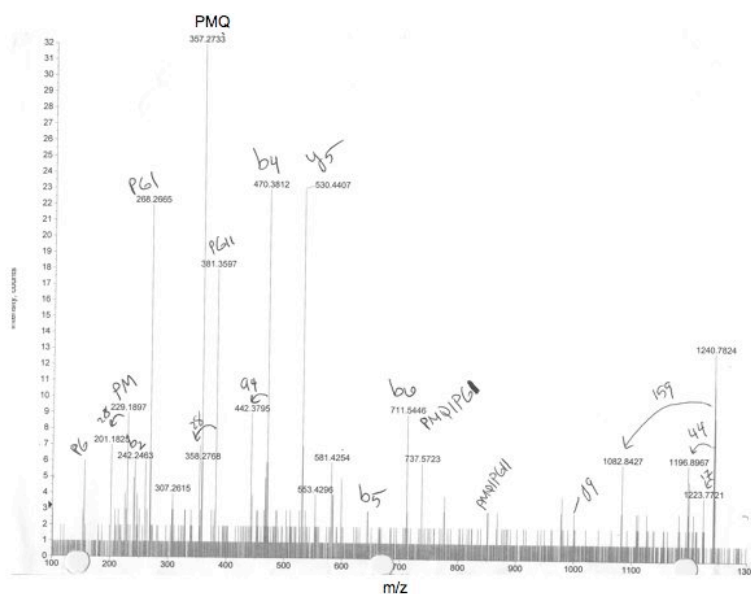


Figure A.1: **Fragment 1560(m/z)**. The fragment was identified as residues 275-287 of gp5 with the sequence: IVLNPRDWHNIAL. Letters correspond to amino acids of internal fragments.

A



B

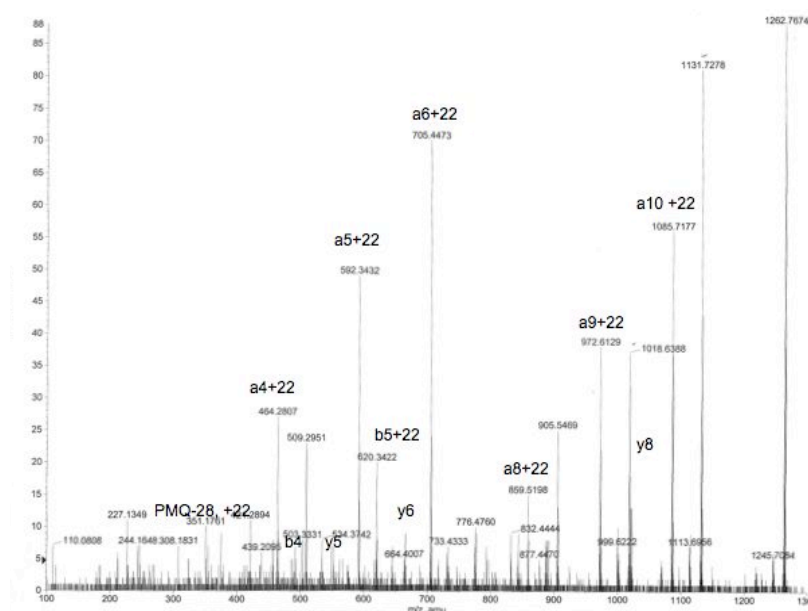


Figure A.2: **Fragment 1240**. The fragment corresponds to residues 117-126 of gp5, with the sequence: IQPMQIPGIIM. (A) The non-adducted species identified by the more common daughter ions including, y^+ , b^+ , and internal fragments. (B) The adducted species, unusually containing many a^+ ions, adducted with sodium. The sodium adduct appears to have greatly contributed to the fragmentation mechanism of the parent ion.

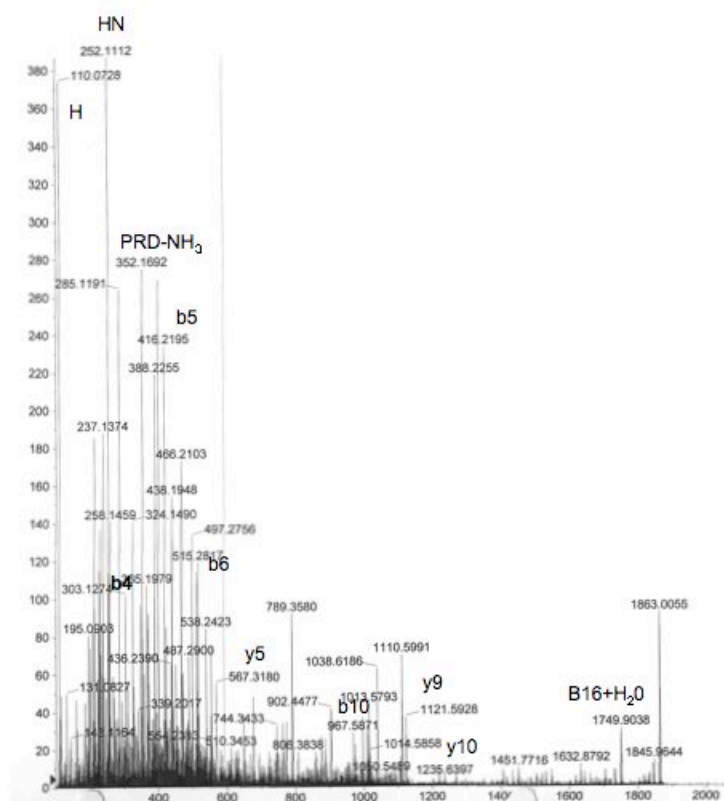


Figure A.3: **Fragment 1862 (m/z)**. The fragment was identified as residues 275-287 of gp5 with the sequence: SASGIVLNPRDWHNIAL.

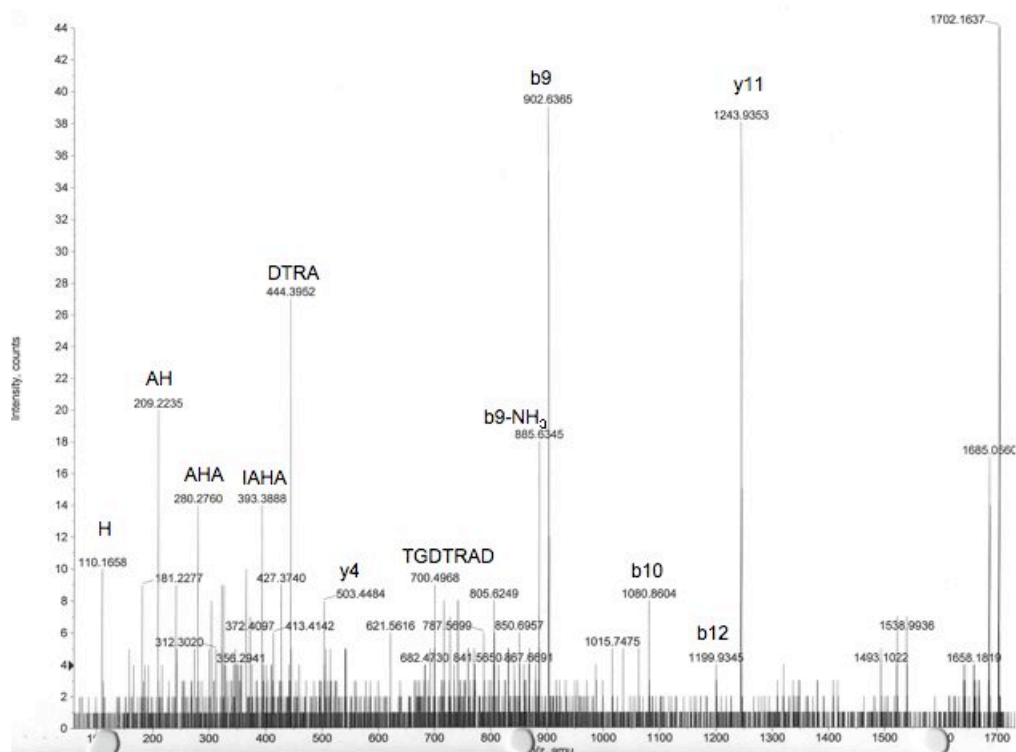


Figure A.4: **Fragment 1701(m/z)**. The fragment was identified as residues 248-263 of gp5 with the sequence: NATGDTRADIIAHAIY.

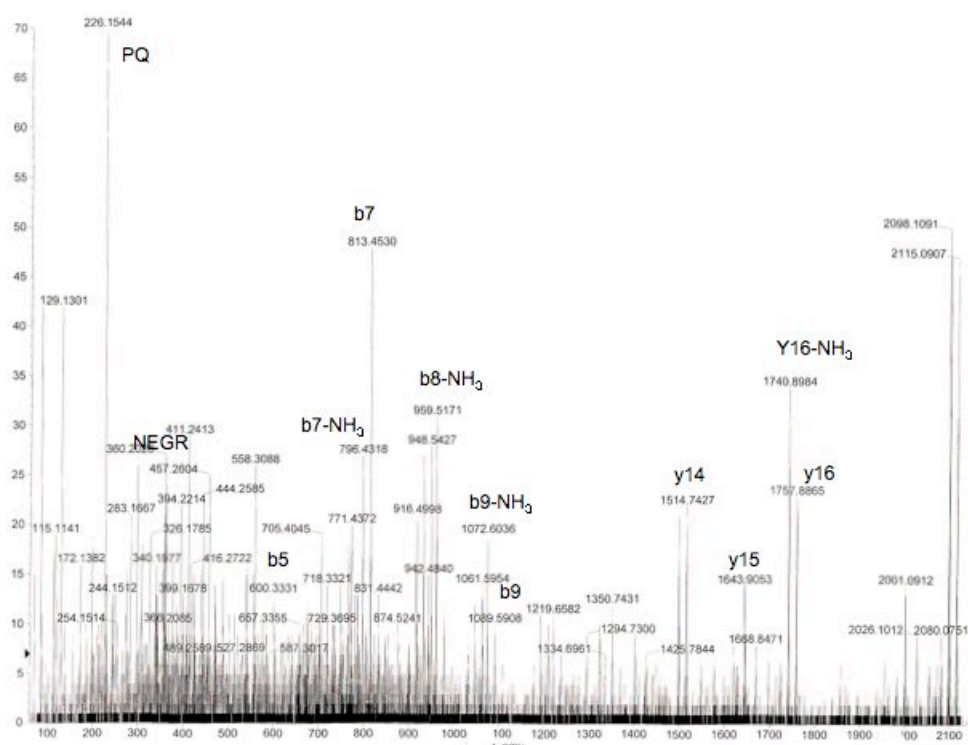


Figure A.5: **Fragment 2114 (m/z)**. The fragment was identified as residues 288-306 of gp5, with the sequence: LKDNEGRYIFGGPQAFTSN. The peptide was identified with mostly y⁺ ions and b⁺ ions, many of which had losses of ammonia.

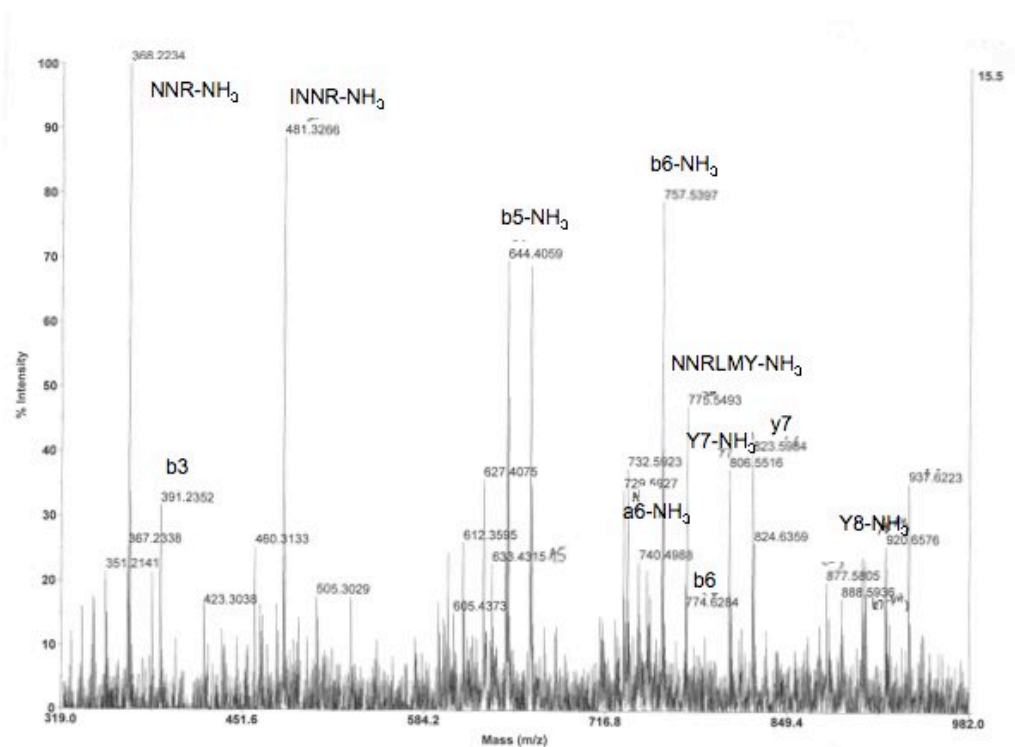


Figure A.6: **Fragment 1327(m/z)**. The fragment was identified as residues 206-216 of gp5, with the sequence: YINNRLMYGLA. Most of the daughter ions contained losses of ammonia.

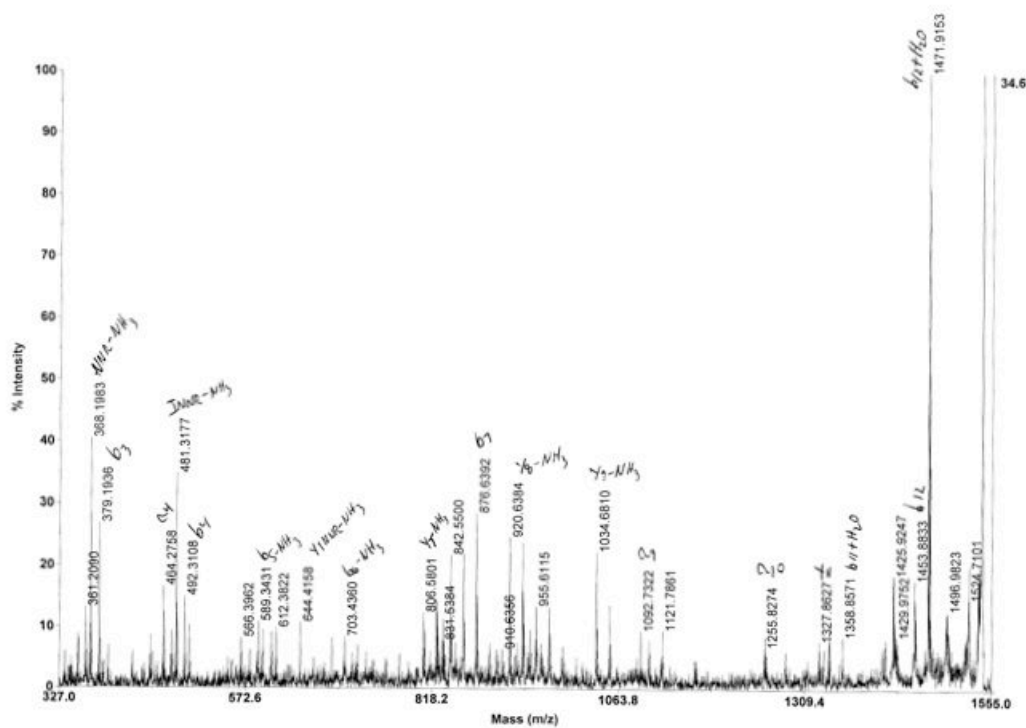


Figure A.8: **Fragment 1542(m/z)**. The fragment was identified as residues 204-216 of gp5, with the sequence: QSYINNRLMYGLA. Similarly to the 1327 fragment, which has most of the same sequence, the fragmentation resulted in many daughter species with losses in ammonia.

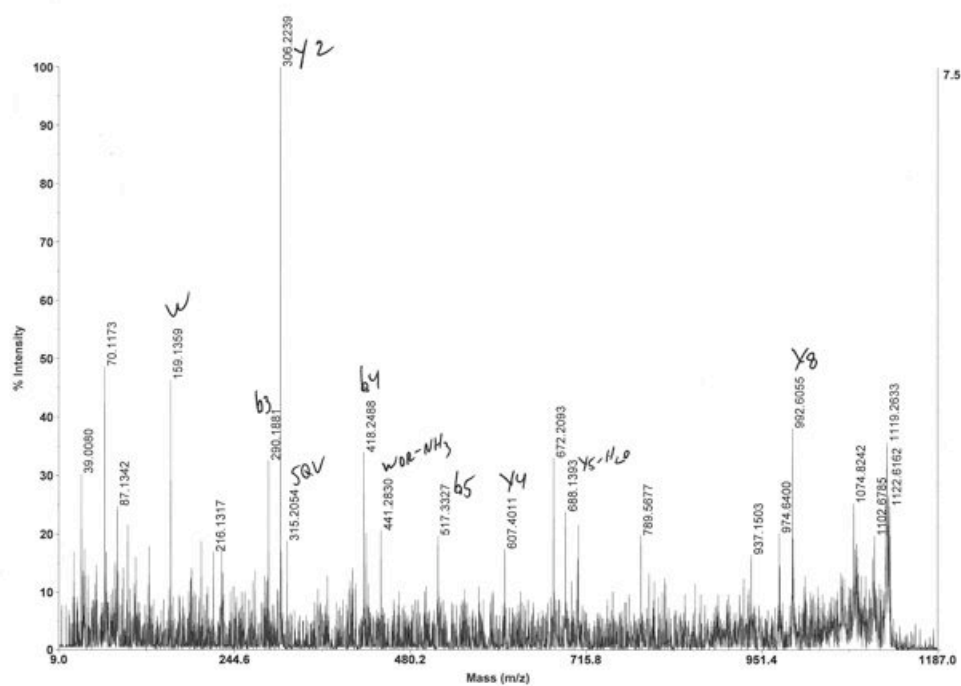


Figure A.9: **Fragment 1123(m/z)**. The fragment was identified as residues 331-339 of gp5, with the sequence: MASQVWDRM.

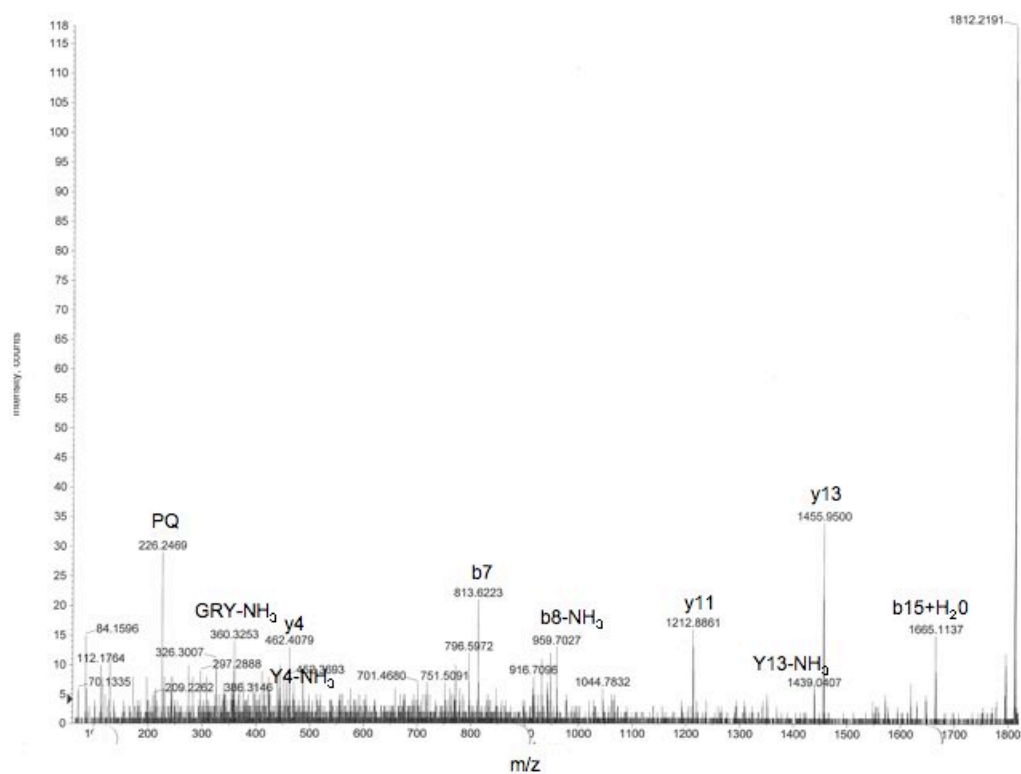


Figure A.10: **Fragment 1811(m/z)**. The fragment was identified as residues 288-303 of gp5, with the sequence: LKDNEGRYIFGGPQAF.