

STRUCTURAL DETERMINATION AND COMPUTATIONAL PREDICTION
OF ADENOVIRAL SURFACE EPITOPES PROVIDE INSIGHT INTO
THE MECHANISMS OF ADENOVIRAL INACTIVATION

BY HUMAN DEFENSIN 5

By

Robert Jaewook Kim

Thesis

Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Chemical and Physical Biology

May, 2012

Nashville, TN

Approved:

Professor Phoebe L. Stewart

Professor Albert H. Beth

Professor Anne K. Kenworthy

ACKNOWLEDGEMENTS

First, I would like to especially thank my mentor, Dr. Phoebe Stewart, for her guidance and support during my graduate career at Vanderbilt. She has been an excellent guide throughout my education, taking countless hours to explain the techniques and theory of CryoEM, and troubleshooting seemingly intractable problems that I came across in my research. Her unique combination of patience and enthusiasm made research a pleasure even on the most difficult days. Dr. Stewart is a truly inspiring scientist, and I am grateful to have spent my graduate years under her tutelage.

I am similarly indebted to Dr. Dewight Williams and Dr. Susan Saban, for providing me with training and assistance with the more technical aspects of research, particularly in the preparation and imaging of samples. The learning process was a long and hard one, but through their efforts, I believe I've become a competent microscopist.

On the computational side of my research, I'd like to give special mention to Steffen Lindert, a recent graduate of the Stewart and Meiler labs. Although we didn't interact much in person, I had a particularly stimulating relationship with the image processing scripts he wrote. It was primarily through observation of his extremely well-commented code that I learned to script in Perl. On the personal side, he was always enjoyable company and a wonderful resource when I had programming questions.

I'd also like to thank my fellow grad students Seth and Tara, for brightening the days with jokes and donuts (respectively), and I'd like to especially thank Justin Flatt for his continual friendship and support, in and out of the lab. He is one of the most down-to-earth, honest, hard-working people I've ever met, and I'm grateful for his camaraderie.

I'd like to thank the CPB department for guiding me through grad school; Lindsay Meyers for taking care of all my administrative needs; Dr. Hassane Mchaourab for his support as my DGS, and my committee for taking the time to review my progress.

Finally, I'd like to thank my family for always being unconditionally supportive throughout this experience.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS.....	ii
LIST OF TABLES.....	vi
LIST OF FIGURES.....	vii
Chapter	
I. INTRODUCTION	
Human Adenovirus.....	1
Human Defensins.....	1
Structure of Adenovirus.....	3
Adenoviral Cell Entry and Replication.....	3
Human Defensin 5 Activity on Adenovirus.....	4
Aim 1.....	7
Aim 2.....	8
Aim 3.....	9
Selection of Technique for Structure Determination.....	9
II. MATERIALS AND METHODS.....	12
Preparation of Adenovirus Samples.....	12
Preparation of Vitrified Grids.....	12
Screening of Grids.....	13
CryoEM Imaging.....	13
Image Processing and Reconstruction.....	14
Modeling.....	28
III. CRYO-EM STRUCTURE DETERMINATION OF ADENOVIRUS (AD5.F35) WITH HD5.....	34
Aim.....	34
Experimental Design.....	34
CryoEM Structure determination.....	35
CryoEM Structural Analysis.....	37
Discussion.....	39
IV. CRYO-EM STRUCTURE DETERMINATION OF ADENOVIRUS (AD5.PB/GYAR) WITH HD5.....	42
Aim.....	42
Experimental Design.....	42
CryoEM Structure determination.....	43
CryoEM Structural Analysis.....	47
Discussion.....	49
V. PREDICTION OF HD5 BINDING ACTIVITY THROUGH MOLECULAR DYNAMICS SIMULATIONS.....	52

Aim.....	52
Experimental Design.....	52
Simulation and Results.....	53
Discussion.....	58
VI. SUMMARY AND CONCLUSIONS.....	62
Significance of Determining a Subnanometer Structure of Ad5.F35 + 5 μ M HD5.....	63
Significance of Determining a Subnanometer Structure of Ad5.PB/GYAR + 5 μ M HD5.....	64
Significance of Predicting Binding Sites Via Molecular Dynamics Simulation.....	65
REFERENCES.....	68

LIST OF TABLES

Table	Page
1. List of finestep levels used in progressively scaling refinements for higher resolution information.....	17
2. List of displacements used when testing image centers during rescue rounds.....	21

LIST OF FIGURES

Figure	Page
1. Adenovirus N-terminal Fiber Sequence Alignment.....	6
2. Rescue Round Center Displacements.....	18
3. Atomic Model of the Ad5.F35 vertex region.....	32
4. Ad5.F35 + 5 μ M HD5 Capsid Reconstruction.....	36
5. Differential Comparison of Ad5.F35 Vertex Regions.....	38
6. Ad5.PB/GYAR + 5 μ M HD5 Capsid Reconstruction.....	44
7. Ad5.PB/GYAR Control Capsid Reconstruction.....	46
8. Differential Comparison of Ad5.PB/GYAR Vertex Regions.....	48
9. Comparison of Ad5.F35 Vertex Regions after Simulation.....	54
10. Ad5.F35 Vertex Atomic Model Overlayed with CryoEM density.....	55
11. Ad5.F35 Vertex Region Simulation Results.....	57

CHAPTER I

INTRODUCTION

Human Adenovirus

Adenoviridae is a family of non-enveloped icosahedral viruses approximately 100 nm in diameter with a double-stranded DNA genome. Human Adenoviruses (HAdVs) are a major cause of acute respiratory, intestinal and ocular infections. Aside from their purely pathological effects, these viruses are of particular interest as candidate gene therapy vectors because of their ability to infect a wide range of tissues and the relative ease of production in the laboratory. [1], [2] There are 57 known serotypes of HAdV, exhibiting a wide range of cell targeting and immune system evasion attributes. Understanding the molecular mechanisms underlying the host immune responses to Adenovirus is of great importance in the design of novel, well-targeted vectors. [3]

Human Defensins

Defensins are small (<50 amino acids), cationic proteins that bind to and inactivate bacterial and viral pathogens through various membrane and capsid interactions. Defensins are characterized by a triple-stranded beta sheet fold, and six disulfide linked cysteines. Human defensins are split into three groups: alpha-defensins, beta-defensins, and theta-defensins. Only the alpha and beta

defensins are naturally expressed; the theta defensin encoding genes have a premature stop codon, preventing reliable expression. Defensins exhibit strong anti-microbial effects against a broad range of pathogens, including many types of bacteria, fungi, and viruses. The anti-bacterial effects are generally attributed to perforation of cellular membranes by defensin molecules, an interaction mediated by the charge properties of defensins and cellular membrane potentials. The perforated membrane leaks electrolytes and other cellular contents, thereby disrupting the cytoplasmic environment and killing the cell. Due to the lack of a membrane potential, dormant bacteria and fungi tend to be more resistant to defensin inactivation. Defensin activity is also notably curtailed in high-salt environments, due to the tendency of salts to stabilize charged molecules. [4]

The antiviral activities of defensin are more diverse in nature. Experiments with HIV-1 have revealed that Defensins inactivate virions prior to attachment, inhibit cell fusion, and block nuclear import and transcription. [5] In other enveloped viruses, defensins have been shown to exhibit membrane-perforating behavior on the viral envelope. In some cases, defensins will accumulate in an infected cell and induce death through concentration-based cytotoxicity. Until recently, the mode of interaction of defensins against non-enveloped viruses has been rather poorly understood, but recent studies have revealed potential mechanisms of action against BK Virus, Human Papillomavirus, and Adenovirus. [5]

Structure of Adenovirus

The full capsid structure of Adenovirus has been resolved to atomic detail through X-ray crystallography and CryoEM techniques [6] , [7] , creating a solid platform on which to model capsid interactions with external elements. The adenoviral capsid is primarily composed of 7 distinct proteins. The major capsomer is protein II (copy number 720), a trimeric hexon protein that forms the bulk of the exterior of the capsid. Protein IX (copy number 240) weaves in between the exterior of hexons, aiding in capsid stability. The twelve vertices of the capsid are sealed by protein III (copy number 60), a pentameric protein, often referred to as the penton base. Above each penton base is the trimeric protein IV(copy number 36), or fiber protein, and directly below the penton base is the pentameric protein IIIa (copy number 60) which plays a role in capsid assembly and disassembly. Protein VIII (copy number 120) provides interior capsid stability. Protein VI is an interior protein which mediates endosomal penetration. These capsid proteins are arranged in an icosahedral fashion, with twelve vertices, each marked by a penton base and fiber protein.

Adenoviral Cell Entry and Replication

In order to understand how human defensins inactivate Adenovirus, an understanding of the method of cellular uptake is required. When a virion comes

in close proximity to a cell, the knob domain of the adenoviral fiber protein (Protein IV) is recognized by the Coxsackie virus and Adenovirus Receptor (CAR receptor). Following recognition, $\alpha_v\beta$ integrins cluster around the penton base protein (protein III), and initiate clathrin-mediated endocytosis. Once inside the endosome, the low pH environment stimulates a conformational change in the capsid, shedding the CAR receptor, releasing the fiber and penton base proteins, as well as protein IIIa and membrane lytic protein VI. [8] Protein VI perforates the endosomal membrane before the endosome transforms into a lysosome, thereby releasing the partially disassembled capsid into the cytoplasm. [9] The capsid is then translocated to the nucleus via microtubules, and docks with the nuclear pore complex, releasing its genome into the nucleus for subsequent transcription. [10]


Human Defensin 5 (HD5) Activity on Adenovirus

The mechanisms of defensin neutralization of Adenovirus have largely been elucidated in studies conducted by the Nemerow lab at Scripps. [11] Human Defensin 5 attaches to virions before cellular uptake, but does not prevent endocytosis. Instead, HD5 neutralizes Adenovirus by preventing virus uncoating and release of the membrane lytic protein VI after endosomal uptake. Most serotypes of Human Adenovirus are sensitive to HD5, but not all. Sensitivity to HD5 neutralization was found for many types from HAdV species A, B, C, and E, while all of the types that were tested from HAdV species D and F were

resistant. The primary sequences of the capsid proteins allowed for predictions to be made regarding the binding site of HD5. This was accomplished through comparison of sensitive and insensitive chimeras in a sequence alignment of the fiber proteins of HAdVs [11].

Adenovirus Fiber N-terminal Sequence Alignment

A (12)	DFNPVYPF DPFDTS -DVPFVTPPFTS
B1 (3,16)	SFNPVYPY EDESS -QHPFINPGFIS
B1 (7)	SFNPVYPY EDESTS -QHPFINPGFIS
B2 (11,14,35)	SFNPVYPY EDESTS -QHPFINPGFIS
C (2,5)	TFNPVYPY DTETGP PTVPFLTPPFVS
D (19c,25,37)	DFNPVYPY GYANQ -NIPFLTPPFVS
E (4)	GFDPVYPY DADNDR PCPSSTLPSFSS
F (41,S)	DFNPVYPY DTFSTP -SIPYVAPPFVS
F (41,L)	DFNPVYPY ERTNPL -DIPFITPPFAS



critical residues

Fig.1: A sequence alignment of the N-terminal tails of the fiber protein of various serotypes of Human Adenovirus. HD5 sensitive serotypes have sequences colored in red, while HD5 insensitive serotypes have sequences colored in yellow.

The sequence alignment revealed that four residues are critical to whether HD5 inactivates a virion or not. Heavily negatively charged sequences (-2 or -3) sequences are correlated with HD5 sensitivity, while relatively uncharged or positively charged sequences are correlated with insensitivity to HD5. A 12Å resolution cryoEM structure of the Ad5.F35 vector in complex with HD5 (produced by mixing the virus with 20µM HD5) showed numerous binding sites over the HAdV capsid [10] (over 3000 binding sites). The cryoEM structure in combination with the sequence alignment of fiber proteins from sensitive and resistant HAdV types led to a hypothesis that the critical HD5 binding site is located at the region between the fiber and penton base proteins.

Aim 1

In order to better understand the interaction between Adenovirus and Human Defensin 5, a subnanometer structure of Ad5.F35 incubated with 5µM HD5 was determined. Ad5.F35 is a chimeric virus composed of an Adenovirus serotype 5 capsid with a substituted fiber protein (protein IV) from Adenovirus serotype 35. This particular chimera was chosen because its activity is neutralized by HD5; also, it has the short serotype 35 fiber, which is more conducive to 3D reconstruction due to its limited flexibility. A concentration of 5µM HD5 was used because in a previous reconstruction [12], 20µM HD5 concentration completely saturated the external capsid and all ~3000 binding sites regardless of binding affinity. 5µM HD5 is still a neutralizing concentration,

and is less likely to show high occupancy in low affinity binding sites in the 3D reconstruction. Alpha-helices can generally be resolved at a resolution of 9Å , thus necessitating a resolution approaching this in order to dock in crystal structures of capsid proteins with relatively high accuracy. Comparing the structure of Ad5.F35 + 5µM HD5 with a control CryoEM structure of Ad5.F35 previously determined by Saban et. al. [13] showed distinct differences in electron density that can be presumed to represent HD5 binding to the Ad5.F35 capsid.

Aim 2

Structures of Ad5.PB/GYAR with and without 5µM HD5 were determined in order to visualize the binding of HD5 to an HD5-insensitive mutant of adenovirus. Ad5.PB/GYAR is a mutant of Adenovirus consisting of an Adenovirus serotype 5 capsid with a penton base substitution (protein III), to have penton base from serotype 19 and with a sequence modification to the Ad5 fiber with a the DTET sequence mutated to GYAR. 3D reconstruction of Ad5.PB/GYAR + 5µM HD5 yielded a structure which showed no significant density where the Ad5.F35 structure showed increased density. This suggests that the critical binding sites responsible for adenovirus inactivation are not present on the Ad5.PG/GYAR chimeric vector. This result, however, will still remain somewhat inconclusive until refinement of the Ad5.PB/GYAR control structure is complete.

Aim 3

Although a good deal of insight can be gained from resolving the structure of HD5 binding to Adenovirus via CryoEM techniques, it is not possible to fully determine the nature of the binding interaction using CryoEM alone. In order to understand the interaction with finer detail, atomic models of the Ad5.F35 vertex region were created using existing crystal structures and sequence information. Flexible regions of the penton base protein (protein III) were modeled using Rosetta structural prediction methods. The software package MDFF (Molecular Dynamics Flexible Fitting) [14] was used to energy minimize the atomic model into the previously determined CryoEM structure, as well as simulate Brownian dynamics of the vertex region in complex with HD5. Simulations reveal candidate residues important for HD5 binding, which can be further studied using chimeric Adenovirus mutants in the future.

Selection of Technique for Structure Determination

Multiple structural techniques have been used to study Adenovirus and its component proteins, particularly X-ray crystallography and CryoEM [6] , [7] . X-ray crystallography in particular has been highly successful in producing structures of component capsid proteins with atomic detail [15], [16], [17], [18], [19], [20], [21] which Aim 3 of this work, in particular, heavily depends upon. It is

more difficult, however, to use this technique to obtain representative structures of flexible protein interactions given the crystallization requirement of the technique. CryoEM has an advantage in this regard, in that samples can be imaged in a near-native state, allowing for more physiologically relevant sets of interactions to take place.

CryoEM is a technique that involves flash-freezing small aqueous samples down to liquid nitrogen temperatures (-196°C) in order to prevent cubic or hexagonal crystalline ice from forming. The ice formed from this process is known as amorphous or vitreous ice; in vitreous ice, the water molecules remain in random orientations as they were in a liquid state, effectively creating a snapshot in time of the natural state of the specimen. In this state, sample particles are typically distributed in random orientations, with occasional preferred orientations due to surface tension effects created by the very low thickness of the layer of ice (<300 nm). Imaging Adenovirus particles in random orientations provides image projections covering the majority of angular space, and thus the virus can be reconstructed through 3D processing techniques [22]

Electron microscope technology is now sufficiently advanced to provide resolutions approaching those obtainable by X-ray crystallography. The FEI Polara microscope used in this study, has a point resolution limit of 2.36Å and an information limit of 1.6Å. The Polara employs a Field Emission Gun (FEG) driven at 300 kV, providing a highly coherent beam that allows such high resolution. The

camera used is a 4000 x 4000 (16 megapixel) Gatan CCD, allowing for the acquisition of micrographs with extremely high resolution information.

Processing capacity is currently advanced enough to refine 3D structures at least up to 960x960x960 pixels, using computing clusters employing Intel Nehalem processors racks with 16 GB of RAM per node. Further advancements in GPU accelerated processing using the NVIDIA CUDA platform promises to further increase the speed of refinement a minimum of tenfold [23], although the use of GPU processing in CryoEM is not yet widespread.

CHAPTER II

MATERIALS AND METHODS

Preparation of Adenovirus Samples

The chimeric Ad5.F35 and Ad5.PB/GYAR were prepared as described in [12] by J.G. Smith at Scripps. Samples were incubated with 5 μ M HD5.

Preparation of Vitrified Grids

CryoEM grids were prepared using Quantifoil holey carbon grids (R 2/4). Grids were prepared by first applying a thin carbon layer on the grid via vapor deposition and then treated with alternating applications (3x40 min.) of chloroform/dichloroethane and acetone/ethyl acetate mixtures on a thin layer chromatography plate, to dissolve and remove the underlying plastic substrate on the grids, leaving behind a lacey carbon matrix. Grids were then conditioned under a high intensity electron beam overnight in a Tecnai T12 electron microscope, before deposition of a second layer of carbon directly prior to sample application. Virus samples are then applied to the grid in 2x2 μ L droplets. The first droplets were applied to the front of the grid and blotted from the rear in order to pull the mixture through the grid to ensure saturation. The second droplets were subsequently applied and blotted for varying amounts of time, between 5 and 10 seconds, before being manually plunged into an ethane slush

to ensure rapid vitrification of each grid. Grids were subsequently stored under liquid nitrogen before screening.

Screening of Grids

Vitrified samples were subjected to two rounds of screening. The first screening sessions were performed on a Tecnai T12 microscope, in order to assess general grid integrity and density of Adenovirus particles. Damaged grids with obviously peeling carbon layers were discarded. The second round of screening is carried out in the FEI Polara microscope, under imaging conditions (300kV beam). Assessment of ice thickness and homogeneity is performed, and “good” areas of the grid are marked and saved for later imaging. Good samples are removed from the microscope and stored under liquid nitrogen in order to prevent sample devitrification and degradation under vacuum.

CryoEM Imaging

Data was collected on an FEI Polara electron microscope (300kV Field Emission Gun). Samples were kept at liquid nitrogen temperatures during imaging. Images were acquired using a Gatan UltraScan 4k x 4k CCD. All data was collected at an absolute magnification of 397,878x. This magnification

produces a sub-angstrom pixel size, allowing for the resolution of very fine details.

Image Processing and Reconstruction

Particle images were picked from each dataset using in-house scripts employing the EMAN [24], IMAGIC [25], and CTFFIND3 software packages. Original scripts were written by Steffen Lindert specifically for cutting out and processing single Adenovirus particles from each micrograph at the given magnification (approx. 397,878kX). Scripts were later heavily modified by the author for use with micrographs of arbitrary magnification, binning, and particle size, with an added provision for selection of multiple particles per micrograph. Refinement of each dataset was then carried out using the FREALIGN [26] software package.

Refinement of Ad5.F35 + 5 μ M HD5

The Ad5.F35 + 5 μ M HD5 dataset was composed of 3515 particle images, manually selected and centered using in-house scripts described above. Initial defocus and astigmatism parameters were determined by the CTFFIND3 program for each micrograph. Micrographs were binned 16x, and centers were selected by manually clicking the center of each particle in the micrograph. Once all centers were selected, the scripts use the EMAN and IMAGIC software packages to crop a 1280x1280 pixel region around each center point, remove

noise, normalize, and create mrc image stacks binned at 320x320, 640x640, 960x960, and 1280x1280 resolutions. The dataset was split into seven image stacks, each stack containing approximately 500 images. This split was used so that refinement processes could be distributed among seven processor cores to expedite refinement. The software package FREALIGN v7.07 was used for all subsequent refinement steps. All refinement was either conducted locally on AMD Opteron or Intel Nehalem processors, or on the Vanderbilt ACCRE computing cluster on Intel Nehalem processors.

Initial refinement was conducted at 320x320 resolution in order to maximize speed of initial low-resolution structure determination. A previously determined map of Ad5.F35 [13] was used as a template structure for beginning refinement. The first round of refinement was the initial step to determine the rough angles and shifts of each image relative to the Ad5.F35 map. Once this initial refinement was carried out, a series of new maps were calculated from the preliminary angles and shifts, each with a different threshold cutoff. The map threshold with the best resolution according to the Fourier Shell Correlation (FSC) 0.5 value (75) was used as a threshold for the following refinement step, "rescue rounds".

Once this initial round was completed, a set of rescue rounds were initiated in order to compensate for human error in picking particle centers. These

rescue rounds involved displacing the picked center for each particle according to the following table:

Table 1: List of displacements used when testing image centers during rescue rounds.

Distance (pixels)	x1	y1	x2	y2	x3	y3	x4	y4
5	0	5	0	-5	5	0	-5	0
10.61	7.5	7.5	7.5	-7.5	-7.5	7.5	-7.5	-7.5
10	0	10	0	-10	10	0	-10	0
17.68	12.5	12.5	12.5	-12.5	-12.5	12.5	-12.5	-12.5
15	0	15	0	-15	15	0	-15	0
24.75	17.5	17.5	17.5	-17.5	-17.5	17.5	-17.5	-17.5
20	0	20	0	-20	20	0	-20	0
31.82	22.5	22.5	22.5	-22.5	-22.5	22.5	-22.5	-22.5
25	0	25	0	-25	25	0	-25	0
38.89	27.5	27.5	27.5	-27.5	-27.5	27.5	-27.5	-27.5
16.77	-15	7.5	-15	-7.5				
26.10	-25	7.5	-25	-7.5				

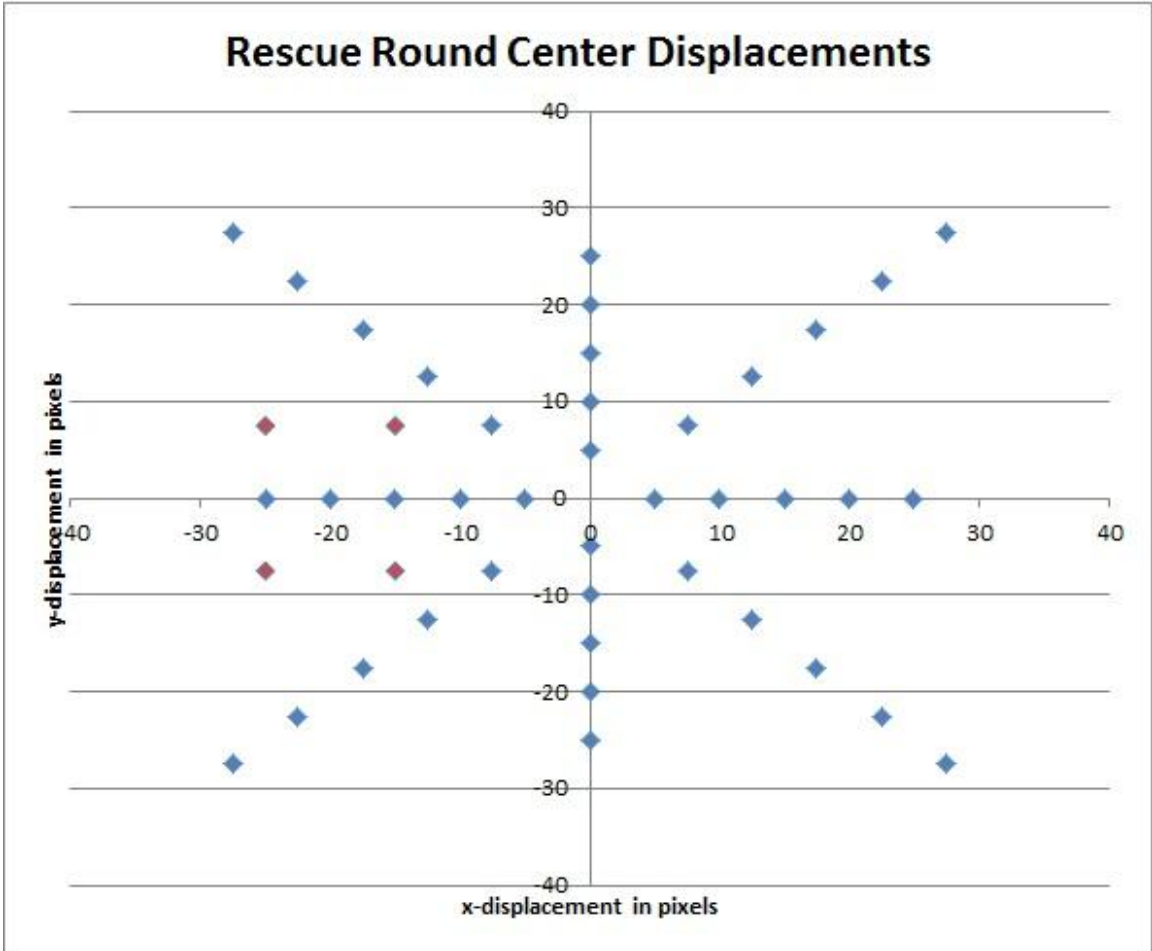


Fig.2: Picked center displacements used in rescue rounds. Each blue point corresponds to a tested displacement from the original center picked in each image. Red points correspond to additional tested displacements, and are on the left of the plot to compensate for human bias during picking.

Initial angles and shifts were calculated again for each image against the Ad5.F35 map using each of the displaced centers. Once the initial parameters were determined, the phase residuals from each of the centers were compared. The center with the best phase residual was saved as the proper displacement for each image. This operation improved the phase residuals of 3005 out of 3515 images, and ended with 1622 particle images with phase residuals below 75. The best parameters for each particle were then combined into a single parameter file, used for the ensuing refinement steps.

After rescue rounds, nine more refinement rounds were conducted at 320x320 resolution. Each round was comprised of a refinement step in which angles and shifts were calculated against the previous map, followed by a map calculation step. Maps of varying thresholds were calculated in order to determine the best resolution map for use in the following round of refinement. Refinement at 320x320 did not employ the Defocus and Astigmatism refinement options available in frealign; employing these parameters too early tends to cause problems. If proper angles and shifts of an image are not locked in to a reasonable degree of accuracy, adjusting defocus and astigmatism is often counterproductive. Other refinement parameters are adjusted between rounds, most importantly the resolution ranges of refinement. The starting resolution range was 100.0Å - 10.0Å. This ensures that all features between these sizes are considered when fitting an image to the map. Another parameter adjusted between rounds is the step sizes of angles and shifts to consider. As refinement

progresses, smaller and smaller step sizes are considered, utilizing the assumption that angles and shifts for each image are converging. These step sizes are detailed in the following table:

Table 2: List of finestep levels used in progressively scaling refinements for higher resolution information.

Finestep Level	Angles	Shifts
0	0.1	0.01
1	0.05	0.01
2	0.025	0.01
3	0.01	0.01
4	0.005	0.01
5	0.0025	0.01
6	0.001	0.005
7	0.0005	0.0025
8	0.00025	0.001
9	0.0001	0.0005

Following refinement at 320x320, the parameter file shifts were multiplied by two and applied to the 640x640 image stacks. Refinement at 640x640 included the options of Defocus and Astigmatism refinement. The first round at 640x640 (round 12) had a resolution range between 50.0Å - 8.0Å with finestep level 4. Refinement at 640x640 continued until round 50, at which the point the resolution range was set between 22.5Å and 7.5Å with finestep level 8. Further refinement at 640x640 yielded little or no improvement, thus requiring a progression to the next larger image binning, 960x960.

Parameter file shifts were multiplied by 1.5 to accommodate the 960x960 binned stacks. For the first two rounds at 960x960, Defocus and Magnification refinement were initially turned off, to give the new binning parameters time to settle. Round 51 began with the same refinement parameters as round 50: resolution range was set between 22.5Å and 7.5Å with finestep level 8. At round 53, Defocus refinement was enabled. At round 58, Magnification refinement was enabled as well. Refinement continued through round 100, however, the best map was acquired at round 94, with a refinement range between 16.5Å and 7Å at finestep level 9.

Refinement of Ad5.PB/GYAR + 5µM HD5 using Frealign

The Ad5.PB/GYAR + 5µM HD5 dataset was composed of 3620 particle images, manually selected and centered using in-house scripts described above. Initial defocus and astigmatism parameters were determined by the CTFFIND3

program for each micrograph. Micrographs were binned 16x, and centers were selected by manually clicking the center of each particle in the micrograph. Once all centers were selected, the scripts use the EMAN and IMAGIC software packages to crop a 1280x1280 pixel region around each center point, remove noise, normalize, and bin at 320x320, 640x640, 960x960, and 1280x1280 resolutions. The dataset was split into seven image stacks, each stack containing approximately 500 images. This split was used so that refinement processes could be distributed among seven processor cores to expedite refinement. The software package FREALIGN v7.07 was used for all subsequent refinement steps. All refinement was either conducted locally on AMD Opteron or Intel Nehalem processors, or on the Vanderbilt ACCRE computing cluster on Intel Nehalem processors.

Initial refinement was conducted at 320x320 resolution in order to maximize speed of initial low-resolution structure determination. A previously determined map of Ad5.F35 [13] was used as a template structure for beginning refinement. The first round of refinement was the initial step to determine the rough angles and shifts of each image relative to the Ad5.F35 map. Once this initial refinement was carried out, a series of new maps were calculated from the preliminary angles and shifts, each with a different threshold cutoff. The map threshold with the best resolution according to the Fourier Shell Correlation (FSC) 0.5 value (75) was used as a threshold for the following refinement step, “rescue rounds”.

Initial angles and shifts were calculated again for each image against the Ad5.F35 map using each of the displaced centers (see Fig.1). Once the initial parameters were determined, the phase residuals from each of the centers were compared. The center with the best phase residual was saved as the proper displacement for each image. The first round of refinement following rescue rounds retained 2160 images out of 3620 total, producing a 10.7Å map as measured by the FSC 0.5 criteria.

After rescue rounds, seven more refinement rounds were conducted at 320x320 resolution. Each round was comprised of a refinement step in which angles and shifts were calculated against the previous map, followed by a map calculation step. Maps of varying thresholds were calculated in order to determine the best resolution map for use in the following round of refinement. Refinement at 320x320 did not employ the Defocus and Astigmatism refinement options available in frealign. Other refinement parameters are adjusted between rounds, most importantly the resolution ranges of refinement. The starting resolution range was 100.0Å - 10.0Å. This ensures that all features between these sizes are considered when fitting an image to the map. Another parameter adjusted between rounds was the step sizes of angles and shifts to consider (see Table 1).

Following refinement at 320x320, the parameter file shifts were multiplied by two and applied to the 640x640 image stacks. The first round at 640x640 (round 10) set its resolution range between 70.0Å - 7.5Å with finestep level 3. Defocus refinement was activated during this round, but magnification refinement was left off. Magnification refinement was activated on round 27, refining at a resolution range between 37.0Å - 7.5Å at finestep level 6. Refinement at 640x640 continued until round 50, at which the resolution range was set between 29.5Å - 6.5Å with finestep level 7. At this point, further refinement at 640x640 yielded little or no improvement, thus requiring a progression to the next larger image binning, 960x960.

Parameter file shifts were multiplied by 1.5 to accommodate the 960x960 binned stacks. For the first two rounds at 960x960, Defocus and Magnification refinement were initially reset to off, to give the new binning parameters time to settle. Round 51 began with the same refinement parameters as round 50: resolution range was set between 29.5Å and 6.5Å with finestep level 7. At round 53, Defocus refinement was enabled with no change in resolution range.. At round 56, Magnification refinement was enabled. Refinement continued through round 93, with a refinement range between 22.0Å - 6.5Å at finestep level 9. The best resolution map was acquired on round 92, at a resolution of 8.0Å.

Refinement of Ad5.PB/GYAR (control) using Frealign

The Ad5.PB/GYAR + 5 μ M HD5 dataset was composed of 3029 particle images, manually selected and centered using in-house scripts described above. Initial defocus and astigmatism parameters were determined by the CTFFIND3 program for each micrograph. Micrographs were binned 16x, and centers were selected by manually clicking the center of each particle in the micrograph. Once all centers were selected, the scripts use the EMAN and IMAGIC software packages to crop a 1280x1280 pixel region around each center point, remove noise, normalize, and bin at 320x320, 640x640, 960x960, and 1280x1280 resolutions. The dataset was split into six image stacks, each stack containing approximately 500 images. This split was used so that refinement processes could be evenly distributed among six processor cores to expedite refinement. The software package FREALIGN v7.07 was used for all subsequent refinement steps. All refinement was either conducted locally on AMD Opteron or Intel Nehalem processors, or on the Vanderbilt ACCRE computing cluster on Intel Nehalem processors.

Initial refinement was conducted at a 320x320 resolution in order to maximize speed of initial low-resolution structure determination. A previously determined map of Ad5.F35 [13] was used as a template structure for beginning refinement. The first round of refinement was the initial step to determine the rough angles and shifts of each image relative to the Ad5.F35 map. Once this initial refinement was carried out, a series of new maps were calculated from the preliminary angles and shifts, each with a different threshold cutoff.

Initial angles and shifts were calculated again for each image against the Ad5.F35 map using each of the displaced centers (see Fig.1). Once the initial parameters were determined, the phase residuals from each of the centers were compared. The center with the best phase residual was saved as the proper displacement for each image. The first round of refinement following rescue rounds retained 2731 out of 3029 particles, resulting in an 18.2Å resolution map as determined by the FSC 0.5 criteria.

After rescue rounds, eight more refinement rounds were conducted at 320x320 resolution. Each round was comprised of a refinement step in which angles and shifts were calculated against the previous map, followed by a map calculation step. Maps of varying thresholds were calculated in order to determine the best resolution map for use in the following round of refinement. Refinement at 320x320 did not employ the Defocus and Astigmatism refinement options available in frealign. Other refinement parameters are adjusted between rounds, most importantly the resolution ranges of refinement. The starting resolution range was 100.0Å - 10.0Å. This ensures that all features between these sizes are considered when fitting an image to the map. Another parameter adjusted between rounds was the step sizes of angles and shifts to consider.

Following refinement at 320x320, the parameter file shifts were multiplied by two and applied to the 640x640 image stacks. The first round at 640x640

(round 11) set its resolution range between 90.0Å - 12.0Å with finestep level 1. Defocus and Magnification refinement were left off during this round.. Defocus refinement was activated on round 28, refining at a resolution range between 63.0Å - 10.0Å at finestep level 4. Magnification refinement was activated on round 32, refining at a resolution range between 60.0Å - 10.0Å. Refinement at 640x640 continued until round 34, at which the resolution range remained between 60.0Å - 10.0Å with finestep level 5. Further refinement on this dataset appeared to yield little or no improvement; more images will have to be collected on the microscope if a better resolution is to be achieved. The best resolution map was calculated at round 32; 13.1Å by the FSC 0.5 criteria, with a total of 806 particle images out of 3029 total images included.

Modeling

Calibration of map size

Following image processing and reconstruction, the electron density maps were stored in MRC file format. An MRC file is a 3 dimensional voxel grid, with each voxel assigned an electron density value. MRC files are a generally accepted standard format for CryoEM structures, and are compatible with most major modeling packages, e.g. Chimera and Visual Molecular Dynamics (VMD).

During image processing and reconstruction, FREALIGN uses user-supplied information to determine the magnification of the images. In real-world

imaging conditions however, the actual magnification of the microscope may vary depending on day-to-day alignments, sample thickness, and defocus. FREALIGN attempts to compensate for these differences by employing algorithms to determine the actual magnifications of images relative to one another. While the final reconstructed structure is generally self-consistent in terms of magnification relative to its component images, the actual Angstrom/voxel ratio may be off by several percent.

In order to correct variances in voxel size, the reconstructed maps were all compared to a crystal structure of human adenovirus serotype 5, RCSB PDB ID: 1VSZ [6]. This crystal structure is comprised of the asymmetric unit of the capsid, primary composed of 4 hexons, and a single subunit of the pentameric penton base protein. Comparison was accomplished by docking the crystal structure into the electron density maps manually in chimera, attempting to line up electron density rods with the alpha helices of the crystal structure. Once a crude alignment was completed, a built in tool called "Fit in Map" was applied, which determines the most likely positioning of the crystal structure through electron density simulation and recursive energy minimization. After fitting, the tool provides an Average map value. A higher average map value corresponds to a better correlation between the crystal structure and electron density map. This procedure of docking and fitting was repeated for each structure in steps of .01 Å/voxel, until the best average map value was determined for each structure. It should be noted that docking into alpha helices manually was only possible in the

Ad5.F35 + HD5 and Ad5.PB/GYAR structures, because they were both at subnanometer resolution, 9.8Å and 8.0Å respectively. Ad5.PB/GYAR control structure was only refined to 13.1Å, a resolution at which alpha helices are not clearly resolvable. For this structure, and quite possibly for all structures, it may have been more accurate to compare the map fit with the full biological assembly of the crystal structure. This however was not possible with our current computational resources, requiring a local compute node with over 24 GB of RAM to even render the full biological assembly alongside the CryoEM densities, much less perform fitting computations on them.

Building an Atomic Model of the Ad5.F35 Fiber region

Once candidate binding sites of HD5 were determined on the Ad5.F35 + HD5 structure, an atomic model had to be built in order to predict the true nature of the binding activity between Adenovirus and HD5. A model of the vertex region of Adenovirus was created using the Adenovirus serotype 2 penton base crystal structure as a basis, due to high sequence homology with the Ad5.F35 penton base protein [27]. The crystal structure of the penton base protein did not, however, contain the 78 amino acid RGD loop region of each monomer, due to high inherent flexibility. These loop regions were modeled using Rosetta structure prediction software. 50 of the best scoring candidates were examined, and 5 of these modeled loop regions were selected for attachment to the penton base crystal structure. Next, the crystal structure of an Adenovirus serotype 5 fiber [28] was used as a template for the serotype 35 fiber, which does not have

a complete crystal structure. In order to account for the added length of the type 35 fiber, an extra repeat of the fiber shaft was added manually in SwissPDB viewer. The long, flexible N-terminal chains of the fiber trimer were also constructed in SwissPDB viewer. After construction, all subunits of the Ad5.F35 vertex region were combined in chimera, along with three HD5 crystal structures, placed in order to simulate potential binding modes with the base of the fiber/shaft and penton base.

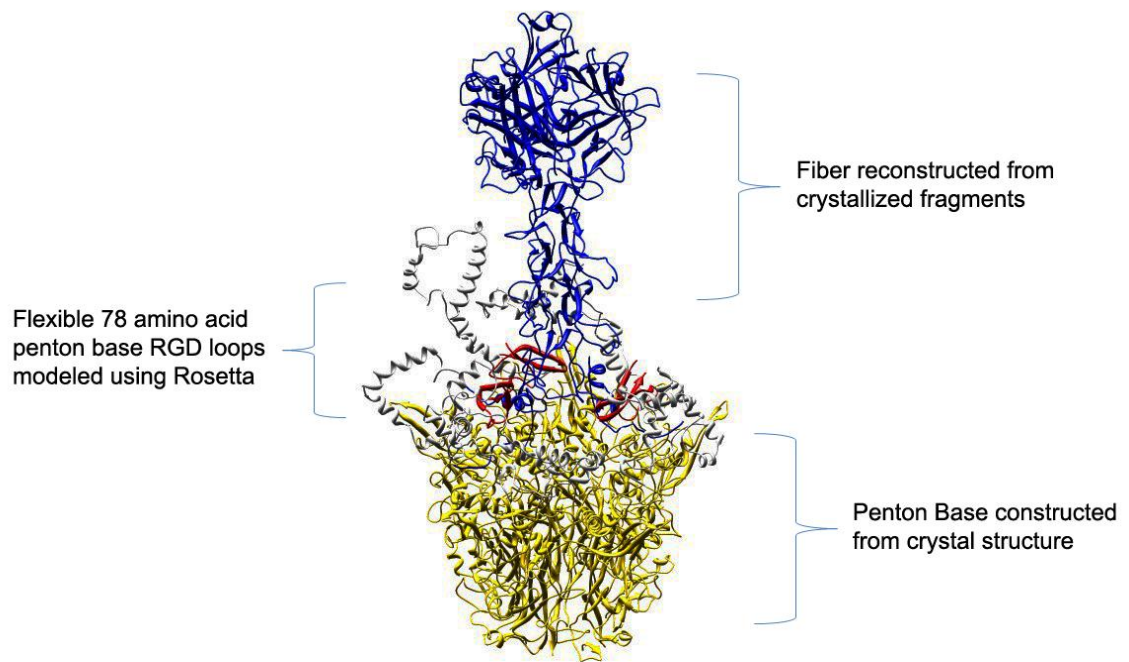


Fig.3: The completed atomic model of the Ad5.F35 vertex region with 3 HD5 molecules prior to simulations. Fiber protein is colored blue, penton base protein is colored gold (penton base RGD loops colored white), and the 3 HD5 molecules are colored red.

Molecular Dynamics Simulations

Simulations of the molecular dynamics of the Ad5.F35 atomic model were conducted using the software packages VMD (Visual Molecular Dynamics) [29], NAMD [30], and MDFF (Molecular Dynamics Flexible Fitting) [14]. The atomic model of the Ad5.F35 fiber region was docked into the Ad5.F35 + HD5 CryoEM structure manually using the Chimera software package. Once docked, the map was thresholded and subjected to a binary mask, to remove noise external to the vertex region. The masked map was exported to SITUS file format for use in VMD and MDFF. Once both the model and map were imported into VMD, an initial energy minimization was carried out followed by a 100 picosecond dynamic relaxation step using the CryoEM density as a restraint. The final recording was exported back to Chimera for viewing and analysis.

CHAPTER III

CRYO-EM STRUCTURE DETERMINATION OF ADENOVIRUS (AD5.F35) COMPLEXED WITH HD5

Aim

The primary capsid structure of Human Adenovirus has been previously determined through both CryoEM and X-ray Crystallography studies. More recent studies have been devoted to determining the structures of Adenovirus-antigen complexes. [31] The goal of Aim I was to gain insight into the binding localization and behavior of Human Defensin 5. General binding localizations can be determined with a subnanometer CryoEM structure, and further analysis of capsid protein interactions can be conducted given such a structure.

Experimental Design

The Ad5.F35 vector was selected for this study because it holds a number of advantages as an experimental model. Ad5.F35 is an Adenovirus serotype 5 virus, pseudotyped with an Adenovirus serotype 35 fiber. Firstly, there already exists a subnanometer CryoEM structure of Ad5.F35 [13], facilitating downstream structure comparison. Second and more importantly, Ad5.F35 has been shown to be inactivated by HD5 [11].

CryoEM Structure Determination

The structure of Ad5.F35 + 5 μ M HD5 was determined by plunging samples of Ad5.F35 + 5 μ M HD5 onto EM grids, and collecting micrographs on an FEI Polara microscope. Micrographs were computationally processed and resulting image stacks underwent 99 rounds of refinement. The best resolution achieved according to the FSC 0.5 criteria was 9.6A from 1014 out of 3515 processed micrographs.

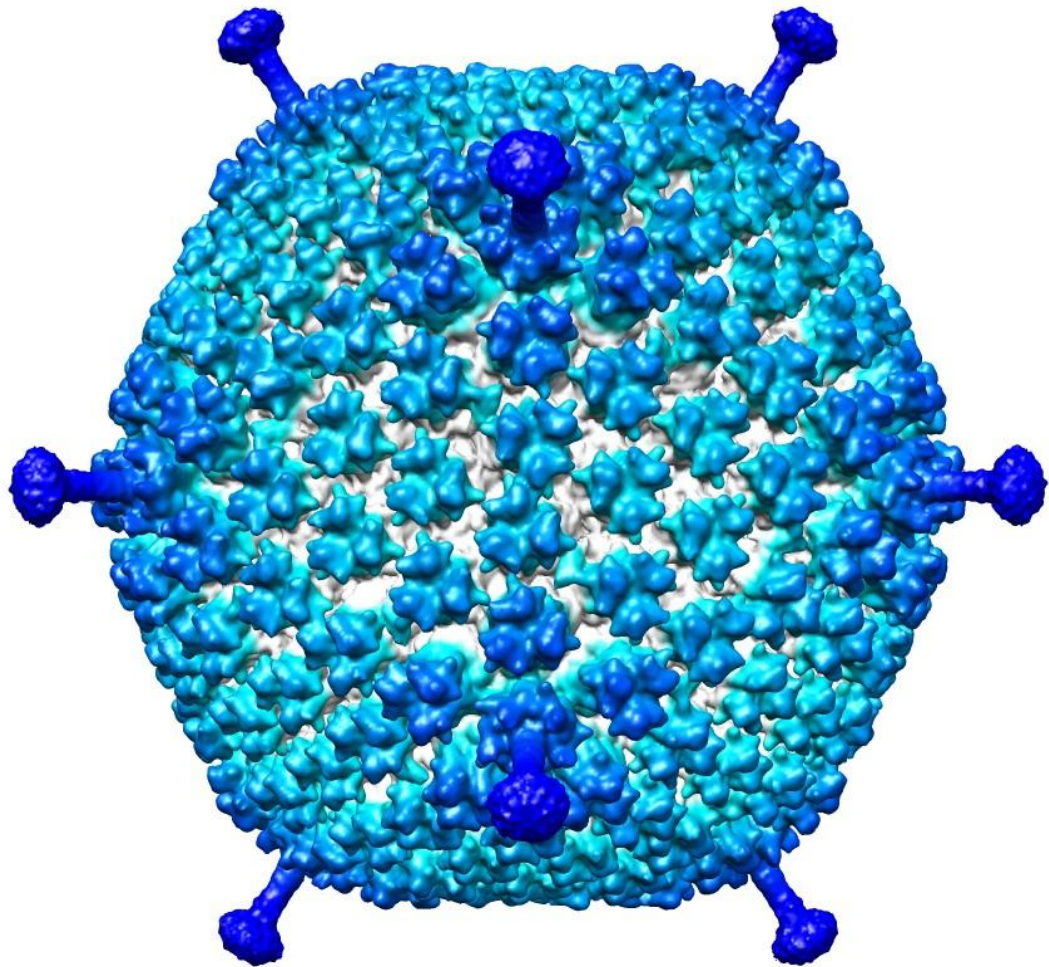


Fig.4: A 9.6Å reconstruction of Ad5.F35 + 5μM HD5

CryoEM Structural Analysis

In order to determine potential binding sites of HD5 on the Ad5.F35 capsid, The calculated structure was compared to an existing structure of Ad5.F35. [13] Structures were compared by computationally assisted alignment in Chimera, and contouring to match features between the two structures. Differences in electron density were noted between the structures, suggesting binding of HD5.

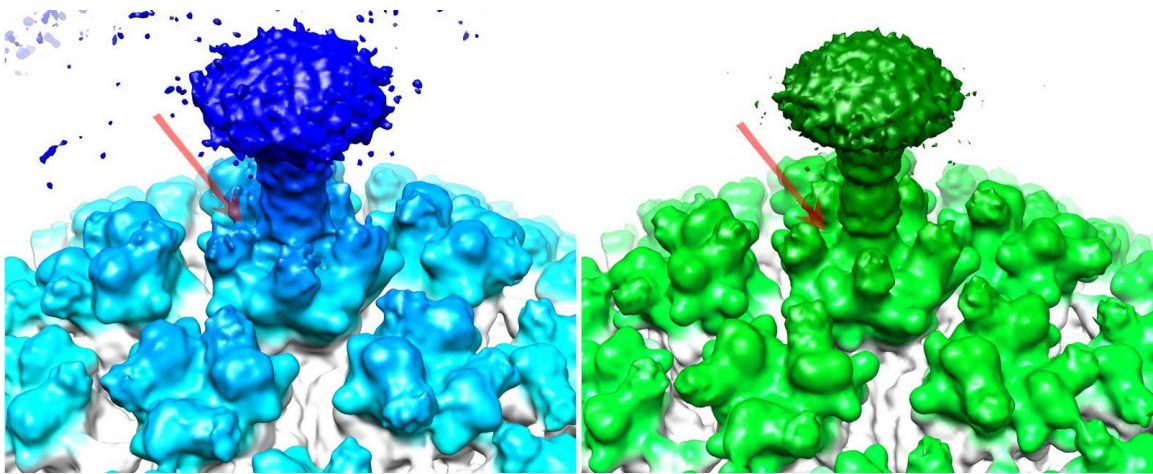


Fig.5: A comparison of the vertex region of Ad5.F35 + 5µM HD5 (Blue) with the vertex region of Ad5.F35 without HD5. The red arrows indicate the primary area of differential density between the two structures. The density in question consists of a density bridge between the protruding lobes of the penton base protein and the fiber shaft.

The most prominent area of differential density can be seen in the area between the penton base and fiber proteins, manifested as five separate chunks of density. Binding orientation cannot be directly inferred from this structure, due to flexibility of the binding pocket and averaging effects from the imposition of icosahedral symmetry. It should be noted that even without the imposition of symmetry, it is unlikely that these binding pockets could be better resolved through CryoEM, due to stochastic binding effects.

Experimental data has determined that N-terminal residues of the trimeric fiber protein are important to whether Adenovirus is inactivated by HD5. The penton base protein has also been shown to play a role in this interaction, but it is not yet known which residues in particular are important to this interaction. Cursory analysis of the CryoEM structure narrows down the range of possible binding sites on the surface of the penton base protein but does not provide any insight into the behavior of the flexible RGD loops extending from the outer radius of the penton base.

Discussion

A subnanometer resolution structure of Ad5.F35 + 5 μ M HD5 was determined in this study. The Ad5.F35 vector has been shown to be neutralized by a 5 μ M concentration of HD5 in previous experiments, suggesting that 5 μ M is an appropriate concentration to observe binding. A previous reconstruction [12]

of Ad5.F35 + 20 μ M was conducted to understand the binding behavior of HD5 to Adenovirus. This experiment was successful in determining a 12.3Å structure of an HD5 saturated capsid. The experiment, however, could not determine the location of specific critical binding sites responsible for HD5 inactivation due to the aforementioned saturation; there are over 3000 binding sites for HD5 on an Adenovirus capsid but not all of them yield any particular activity [11].

Of particular note is that in the previous 20 μ M HD5 study, the vertex region reconstruction showed a lack of fiber knob density. It has been hypothesized that attachment of HD5 to a fiber shaft results in the canting of the fiber. Imposing icosahedral symmetry on Adenovirus particles in the reconstruction process would result in the averaging away of canted fiber knobs. The presence of a distinct fiber knob in the current 5 μ M HD5 study suggests that although binding was present as evidenced by the extra density between the penton base and fiber, saturation was by no means achieved. It may very well have been the case that few of the vertex regions experienced much binding of HD5 at all, allowing the preservation of the fiber knob upon icosahedral reconstruction. Another possibility is that the fiber does not cant when HD5 binds to the active site. Perhaps there are other binding activities associated with the fiber knob that cause this canting in high concentrations of HD5. This question will probably require advanced solvent-based molecular dynamic simulations to solve, and is outside the scope of this study.

This subnanometer structure provides primary visual evidence that the critical binding site of HD5 to Ad5.F35 is at the base of the fiber protein. Previous experiments had suggested that this was the case, but structural observation has confirmed this beyond a reasonable doubt. Using this information, further studies can be directed toward understanding the molecular mechanism of this binding and inactivation, as well as the role of protein III in this interaction.

CHAPTER IV

CRYO-EM STRUCTURE DETERMINATION OF ADENOVIRUS (AD5.PB/GYAR) COMPLEXED WITH HD5

Aim

The previous aim dealt primarily with determining the potential binding sites of HD5 on sensitive serotypes of Human Adenovirus. In order to paint a clearer picture of the binding behavior of HD5 to Adenovirus in general, it was necessary to observe the binding of HD5 to an HD5 insensitive strain of Adenovirus. The goal of Aim 2 was to gain insight into the binding behavior of Human Defensin 5 to HD5 insensitive serotypes of Adenovirus. General binding localizations can be determined with a subnanometer CryoEM structure, and further analysis of capsid protein interactions can be conducted given such a structure.

Experimental Design

The Ad5.PB/GYAR vector was selected for this study primarily due to its high homology with Adenovirus serotype 5 and complete insensitivity to HD5 exposure. Ad5.PB/GYAR is an Adenovirus serotype 5 virus, pseudotyped with an Adenovirus serotype 19 Penton Base. Four residues of the variable region in the fiber protein are also modified from DTET to GYAR. Ad5.F35 and Ad5.PB/GYAR

have identical hexon protein sequences and structure; therefore it could be assumed that any differences in binding would occur at the vertex regions of the virus. Two structures were determined in this aim, that of Ad5.PB/GYAR + 5 μ M HD5, and a control structure of Ad5.PB/GYAR without HD5.

CryoEM Structure Determination

The structure of Ad5.PB/GYAR + 5 μ M HD5 was determined by plunging samples of Ad5.PB/GYAR + 5 μ M HD5 onto EM grids, and collecting micrographs on an FEI Polara microscope. Micrographs were computationally processed and resulting image stacks underwent 93 rounds of refinement. The best resolution achieved was 8.0Å from 1730 out of 3620 processed micrographs. The final reconstruction is shown here:

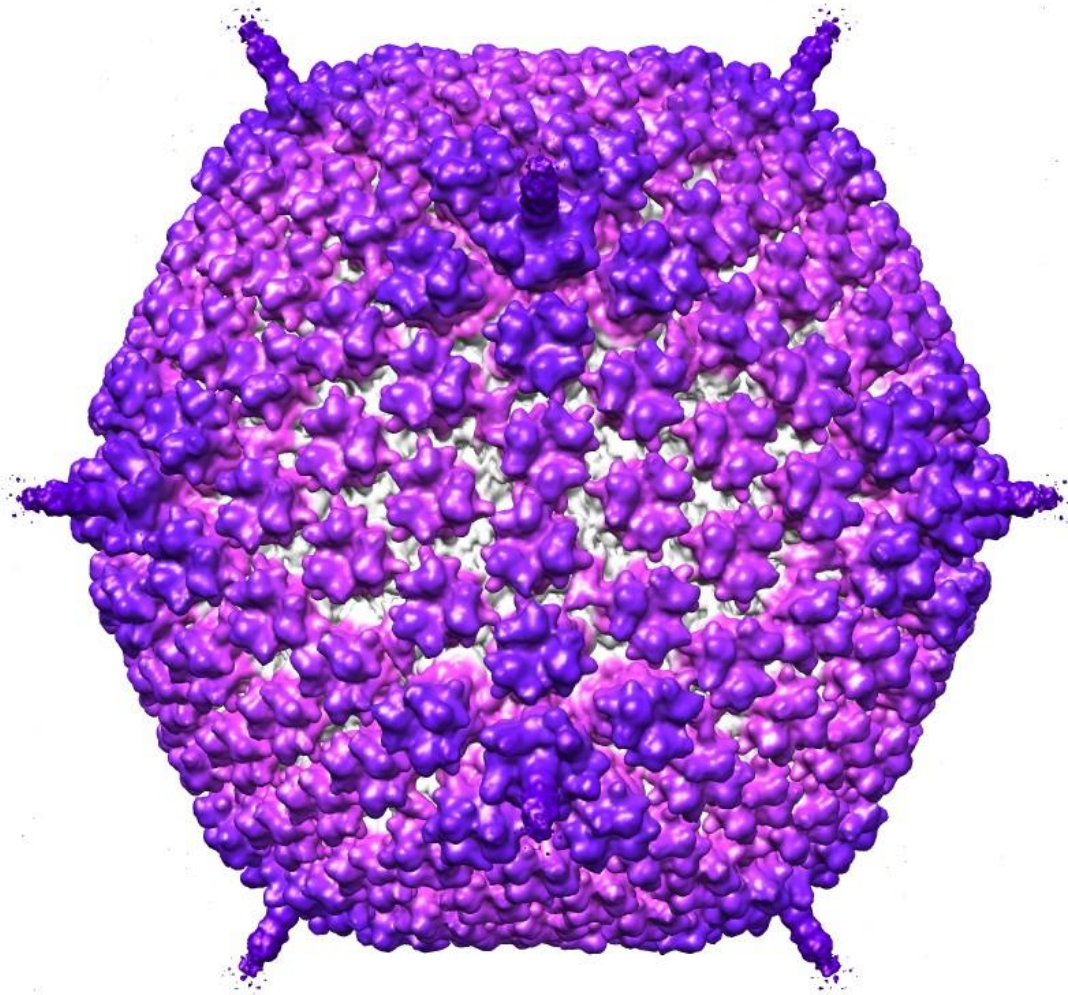


Fig.6: A 8.0Å reconstruction of Ad5.PB/GYAR + 5μM HD5

The structure of Ad5.PB/GYAR control was determined by plunging samples of Ad5.PB/GYAR onto EM grids, and collecting micrographs on an FEI Polara microscope. Micrographs were computationally processed and resulting image stacks underwent 32 rounds of refinement. The best resolution achieved was 13.1Å from 806 out of 1029 processed micrographs. The final reconstruction is shown here:

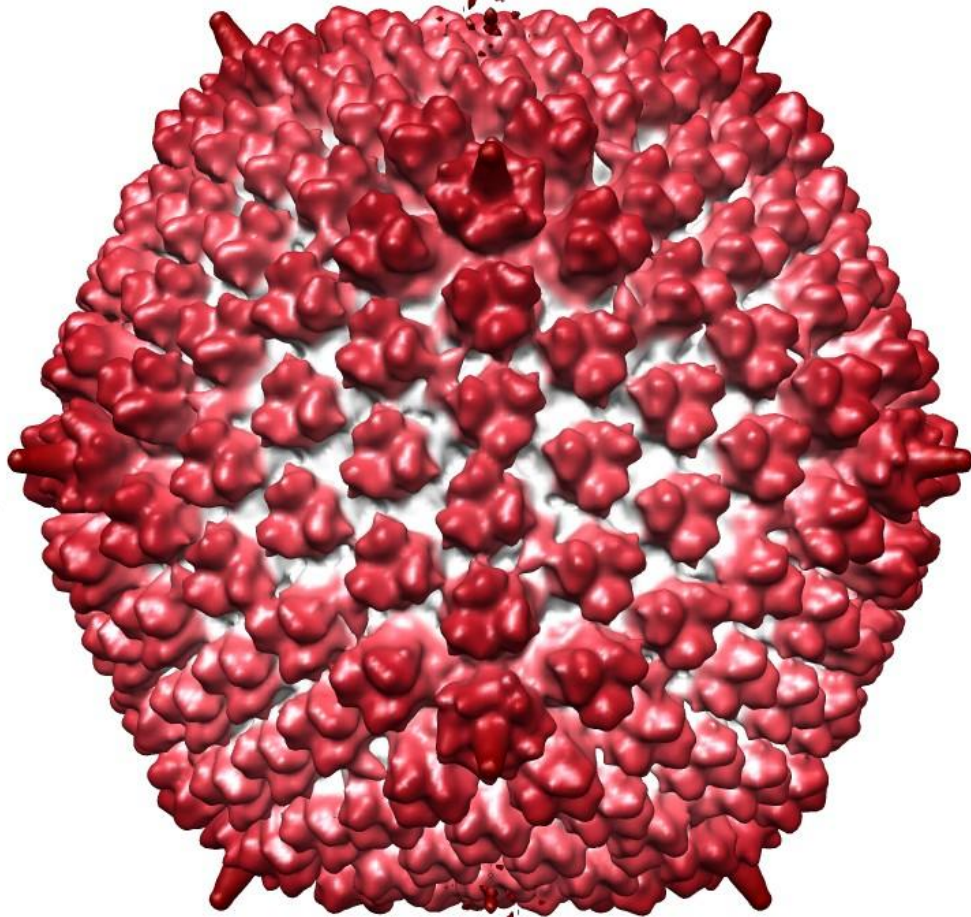


Fig.7: A 13.1Å reconstruction of Ad5.PB/GYAR with no HD5.

CryoEM Structural Analysis

When final reconstructions of Ad5.PB/GYAR+ 5 μ M HD5 and Ad5.PB/GYAR control were complete, the structure was compared with the Ad5.F35 and Ad5.F35 + 5 μ M HD5 density maps. Structures were aligned in Chimera using built-in alignment and sizing tools. Differences in electron density were noted among the maps.

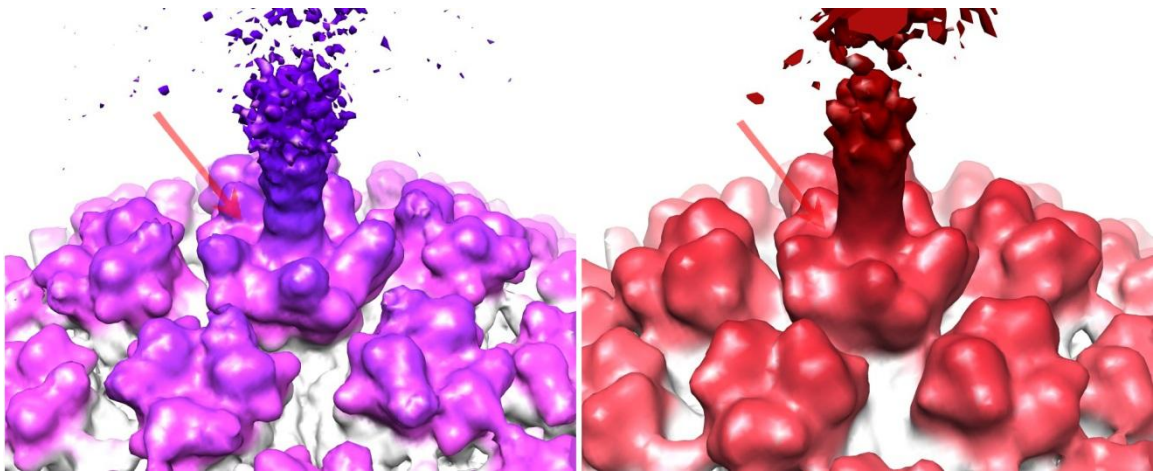


Fig.8: A comparison of the vertex region of Ad5.PB/GYAR + 5 μ M HD5 (Purple) with the vertex region of Ad5.PB/GYAR without HD5(red). The arrows indicate the primary area that HD5 appears to bind in the Ad5.F35 + 5 μ M HD5 structure (see Fig. 5). Density is notably absent in this area on the Ad5.PB/GYAR + 5 μ M HD5 structure. The difference is more difficult to see on the Ad5.PB/GYAR control structure, most likely due to the lower resolution of the reconstruction.

Analysis of the structure reveals little difference between the vertex regions of Ad5.PB/GYAR+ 5 μ M HD5, Ad5.PB/GYAR control, and the Ad5.F35 control structure. Comparing the Ad5.PB/GYAR+ 5 μ M HD5 with the Ad5.F35 + 5 μ M HD5 structure, there is a clear difference in the electron density surrounding the base of the fiber protein. This suggests that HD5 binding in that area only occurs in sensitive serotypes of Adenovirus.

Discussion

A subnanometer resolution structure of Ad5.PB/GYAR + 5 μ M HD5 was determined in this study, along with a medium resolution structure of Ad5.PB/GYAR without bound HD5. The Ad5.PB/GYAR vector has been shown to be immune to HD5 inactivation in previous experiments. The structural reconstruction of Ad5.PB/GYAR + 5 μ M HD5 compared with the reconstructions of Ad5.PB/GYAR control and Ad5.F35 + 5 μ M HD5 further confirms the previous conclusion that the critical binding site of HD5 is at the base of the fiber shaft of Ad5.F35. HD5 density is conspicuously absent from the Ad5.PB/GYAR vertex region.

It should be noted however that the Ad5.PB/GYAR fiber protein differs significantly from the Ad5.F35 fiber protein. The Ad5.PB/GYAR (or Ad19) fiber

has more β -spiral repeats, but each repeat has high homology with corresponding repeats in Ad5.F35. Also, naturally, the critical binding N-terminal residues in the Ad5.PB/GYAR fiber also differ from Ad5.F35. Though one might consider the difference in fiber lengths to be significant and possibly a confounding variable in the study, it would be unlikely that the fiber length has much to do with viral susceptibility to HD5 inactivation. This is a reasonable assumption because there are many long-fibered serotypes of HAdV that are susceptible to HD5 inactivation, and the β -spiral repeats are conserved sequences between most serotypes. The Ad5.F35 chimeric virus was constructed and used in the initial studies particularly because it was more conducive to reconstruction via CryoEM; short fibers appear less flexible because they are well within the persistence length of the β -spiral repeat structure.

In spite of the added flexibility of the fiber shaft of the Ad5.PB/GYAR structure, a higher resolution was achieved with the Ad5.PB/GYAR + 5uM HD5 dataset than with the Ad5.F35 + 5uM HD5 dataset although the two datasets were roughly equivalent in size. This can be attributed to a number of factors. First, imaging quality may have differed between datasets as a result of differing ice thickness, microscope alignment, or temperature induced drift during imaging. Second, and possibly more importantly, the lack of HD5 binding on Ad5.PB/GYAR reduced the structural variability of the primary Adenoviral capsid. Although the fiber in Ad5.PB/GYAR is more variable in orientation, most

refinement steps masked off the majority of the fiber, focusing on the hexons and penton bases for alignment. This negates the effects a flexible fiber may have on reconstruction, making surface features more important than fiber flexibility. The reason for the rather low resolution reconstruction of the Ad5.PB/GYAR control structure, however, can be attributed to poor sample quality, making good image acquisition difficult.

The subnanometer structure of Ad5.PB/GYAR + 5 μ M HD5 provides additional visual evidence that the critical binding site of HD5 to Ad5.F35 is at the base of the fiber protein, as noted by absence of density when compared with both the Ad5.F35 + 5 μ M HD5, and a control structure of Ad5.F35. Using this information, further studies can be directed toward understanding the molecular mechanism of HD5 binding and Adenovirus inactivation.

CHAPTER V

PREDICTION OF HD5 BINDING ACTIVITY THROUGH MOLECULAR DYNAMICS SIMULATIONS

Aim

The previous two aims dealt with determining the location of the critical binding sites responsible for HD5 inactivation of HD5. The information provides a rough estimate of the location, but fails to provide atomic detail into the nature of this interaction. The goal of Aim 3 was to predict the binding activity of HD5 with atomic detail using molecular dynamics simulations.

Experimental Design

In order to conduct a reasonable molecular dynamics simulation of the interaction between the Adenovirus vertex region and HD5, four things are required. First, an atomic model of the entire external vertex region is required, containing the entirety of protein III and IV (penton base and fiber proteins). Second, an atomic model of Human Defensin V is required. Third, a CryoEM map of the interaction is required as a restraint for a plausible simulation. Fourth, realistic simulation software is required. The Ad5.F35 vertex region was constructed as described in the materials and methods section, using various tools including Rosetta structure prediction, SwissPDB viewer, and Chimera. The

atomic structure of HD5 was acquired from the PDB. [32] A subnanometer CryoEM map of Ad5.F35 was created under Aim 1. Finally, the simulation software used was Molecular Dynamics Flexible Fitting, or MDFF.

Simulation and Results

After the Ad5.F35 vertex region atomic model was created the model was manually docked into the CryoEM density. Three HD5 peptides were placed in potential binding locations, close to the negatively-charged EDES residues within the N-terminal region of the fiber[12]. Five separate models for the RGD loops were used in the simulation, to provide a range of possible structures since the true orientations cannot be determined through standard structural techniques. The model was energy minimized against the map, and then allowed to “relax” into the structure while undergoing a simulation of thermal motion, created by imparting random starting velocities to atoms in the structure. The simulation lasted for a real time equivalent of 100 picoseconds.

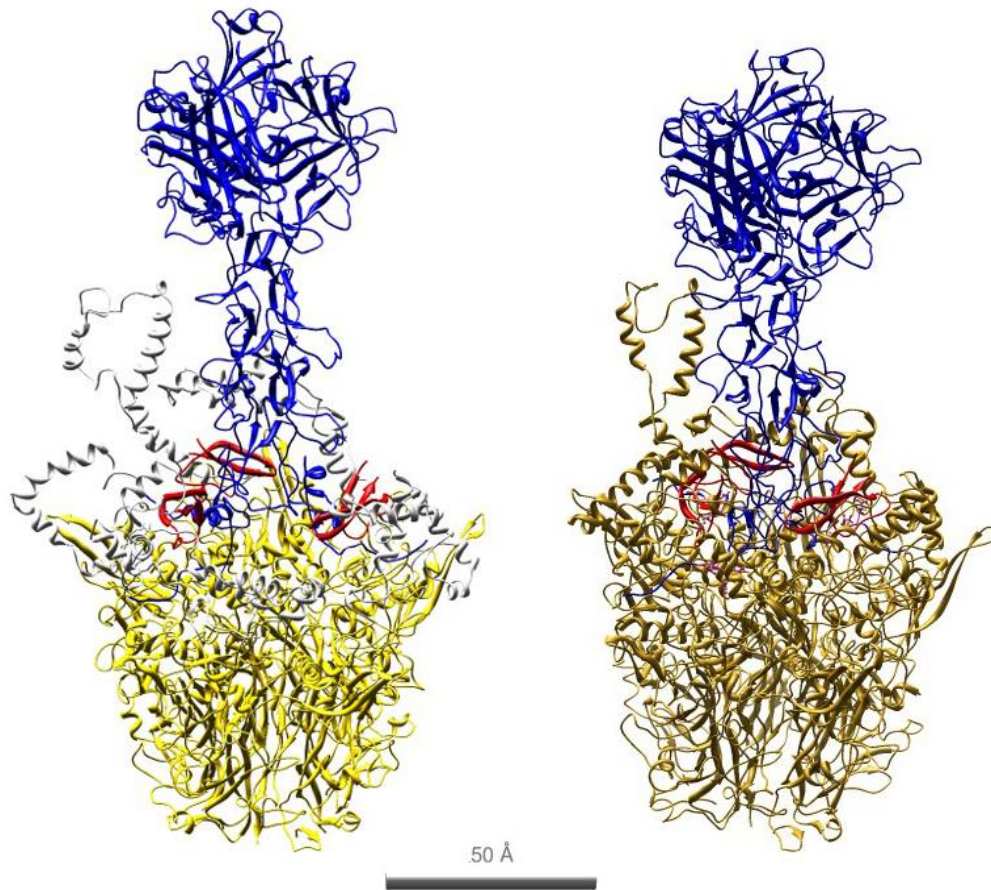


Fig.9: The ribbon structures of the Ad5.F35 vertex model with 3 HD5 molecules before simulation (left) and after 100 picosecond simulation (right). Fiber protein is colored blue, penton base protein is colored gold (penton base RGD loops on starting structure colored white), and the 3 HD5 molecules are colored red.

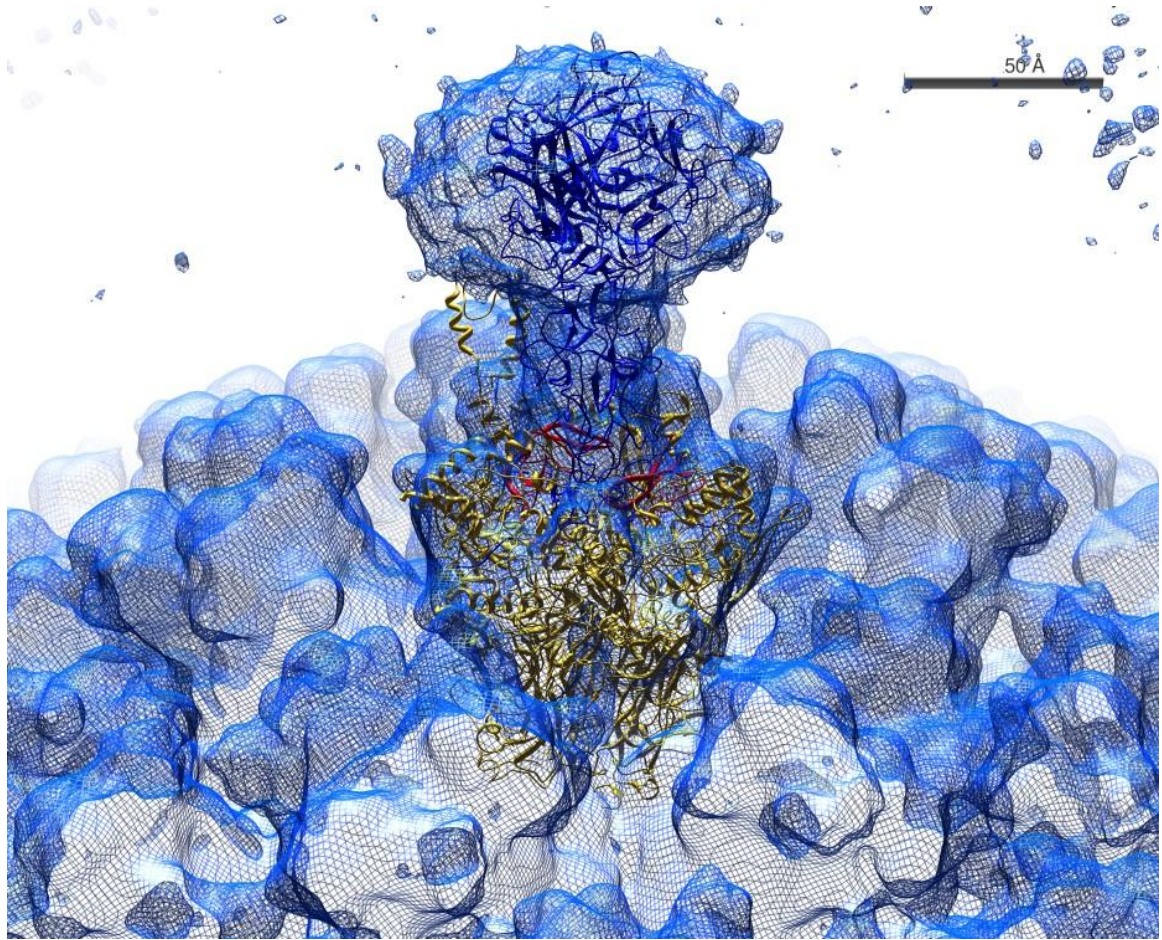


Fig.10: A composite of the Ad5.F35 + 5µM HD5 CryoEM structure used as a restraint in the simulation overlaid with the resulting atomic model of the vertex region.

Upon conclusion of the simulation, the three HD5 peptides were nested into the predicted binding sites, close up against the critical residues (EDES) of the N-terminal regions of the protein IV. Also of note was the behavior of the RGD loops. The RGD loops appeared to fold inwards toward the fiber and envelop the HD5 peptides against the surface of the penton base. This suggests that the RGD loops may play an important role in stabilizing the vertex region upon HD5 binding, preventing disassociation of the fiber and penton base when exposed to acidification in the endosome.

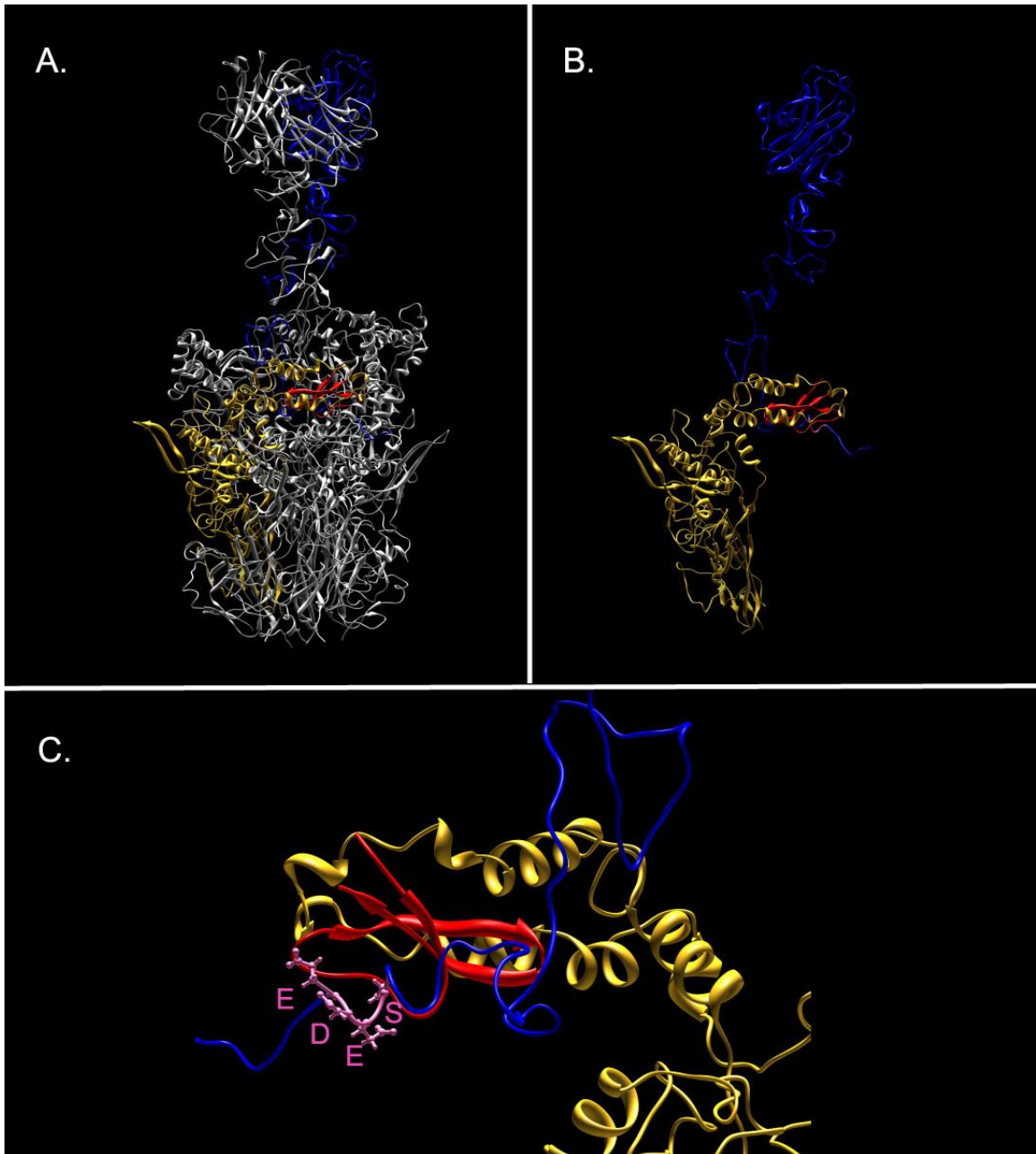


Fig.11: Result of the simulation. Fiber protein (blue), Penton base protein (gold), HD5 (red) **A.** shows the final complete ribbon structure after the simulation with all but the area of interest greyed out. **B.** Removes greyed out area **C.** Close-up of the binding site of HD5, showing the negatively charged EDES sequence in close proximity to the positively charged HD5 molecule.

Discussion

This study conducted a preliminary simulation of the molecular dynamics of Human Defensin 5 binding to the Adenovirus vertex region. The results of the simulation suggest that HD5 binds to the regions previously determined to be important in Adenoviral inactivation. The study also offered a potential role for the flexible RGD loop regions in stabilizing bound HD5 to the viral capsid.

Although the results seem promising in confirming previous experimental results, the predictive value of this model has not yet been determined. This simulation is a preliminary foray into the use of computational simulation for modeling this interaction. There are many issues that need to be addressed before relying on the results provided.

The first issue with this simulation is the quality of the constructed atomic model. Both the fiber (protein IV) and the penton (protein III) contain elements that were manually created and added in. The fiber protein was constructed by adding in an extra β -spiral repeat to a serotype 5 fiber, to lengthen it to match a serotype 35 fiber. Also, the N-terminal tails of the fiber do not have a defined crystal structure, and had to be built by hand in SwissPDB viewer. Building these flexible sections by hand using only sequence data is an exercise in guesswork at best. These flexible tails are known to sit in certain pockets in the penton base. Unfortunately, the models built had a few clashes with the penton base when

combined with the rest of the fiber. This necessitated placing the N-terminal tails in their respective binding pockets while leaving the rest of the fiber lifted several Angstroms above the penton base. The fiber protein was effectively cut at the connection point of the β -spiral to the flexible tails. The cut only extended to the coordinate placement of the atoms, and did not sever the connection in a computational sense; the PDB still considered the fiber protein as one unit, and treated it as such in the simulation. The implications of this are that there are extremely high forces between the tails and spiral pulling them together in the initial moments of the simulation. However crude, this method makes sure that the simulation does not have to deal with initial clashes between atoms. As for the penton base protein, the structures of the 78 amino acid RGD loops were not known, and had to be simulated. Candidate models were generated using Rosetta structure prediction software, but this software is generally not accurate for such a large number of amino acids in a chain without other imposed restraints. Finally, the placement and orientation of the HD5 molecules was determined manually, and only one orientation was tested in each binding site. This issue can be rectified by running simulations with various starting placements and orientations.

A second issue with this simulation is the nature of the map used as a restraint. The density map used was subjected to manual erasing and a binary mask to isolate the regions of interest. This is a step taken to ensure that the molecules in the simulation do not settle into incorrect regions of density, for

example into neighboring hexon density. The application of a binary mask at a certain contour level, however, zeroes out any noise outside the region of interest. This is advantageous for well-defined structures, but for flexible structures like the RGD loops for which no visible electron density exists, a certain level of noise may be necessary. Without the noise, the RGD loops are far more likely to be attracted towards the defined electron density below them, artificially inducing the loops to fold inward toward the penton base.

A third issue with the simulation pertains to the use of randomized energy factors applied to the atoms. Molecular dynamics simulations may include solvent molecules, particularly water, in order to provide a more realistic environmental context for protein relaxation. Solvent was not used in this simulation, particularly because the primary external restraint employed was the electron density map. Randomized energy factors were still employed, creating random momentum vectors for each atom in the initial moment of relaxation. It may be necessary in future endeavors to attempt a solvent-based simulation; the computational cost of this is rather high, however, due to the size of the complexes being studied. Also, the initial randomized energy could be tailored to a temperature level consistent with wet lab experiments; this can be accomplished by imparting a Gaussian distribution of energy levels with the average defined by the relation ($E = 1/2 kbT$). Simply applying this distributed energy is not the only factor to be considered, however; the initial model itself has certain potential energies that must be minimized before a temperature

function can be applied. Once again, this is dependent on the construction of an accurate atomic model of the Ad5.F35 vertex region.

Although molecular dynamics simulations are fraught with difficulties that cannot be completely controlled, it should be stated that the goal of these simulations is not necessarily to generate a physiologically correct model of the Adenovirus interaction with HD5. Rather, the primary goal is to provide insight into *possible* modes of interaction of HD5 with the Adenovirus vertex region. In this regard, the 100 picosecond simulation conducted in this study is a good first step in understanding the interactions between HD5 and the vertex region of Adenovirus.

CHAPTER VI

SUMMARY AND CONCLUSIONS

Understanding the relationship between Human Defensin 5 and Adenovirus inactivation has implications that go far beyond demystifying the mechanism of a single protein-protein interaction. Adenovirus has been an extremely popular model vector for gene therapy studies and trials, given its ease of production and high infectivity. There are two major obstacles to the success of Adenoviral gene therapy however. The first obstacle is deciding which genes to express and how to express it reliably in the Adenovirus genome. The second and arguably more difficult obstacle is targeting the vector to the appropriate organs for treatment, while simultaneously avoiding inactivation by the immune system. Several studies have been conducted to the effect of retargeting Adenovirus [33], [34], [35] [36], [2], [37], as well as reducing immune system activity. [38], [39] Few studies however have examined the structural basis of modifications to Adenovirus. Understanding the structural implications of capsid modification will be essential to the efficient creation of retargeted vectors.

Another benefit of this study is its implications in the fields of general virology and pathology. Human Defensins are known to inactivate a wide variety of viruses through many different modes of attack. Human Defensin 5 in particular has strong antiviral properties against Adenoviruses and

Papillomaviruses[5], [40], [41], [42] Interestingly enough, HD5 inactivates both of these viruses by preventing release from the endosome. It may be possible that HD5 has a more generalized method of attack than previously determined. This could be confirmed by conducting a similar structural study on papillomavirus.

Significance of Determining a subnanometer structure of Ad5.F35 + 5 μ M HD5

Subnanometer CryoEM structures of viruses are becoming more and more common as equipment and techniques become more powerful and sophisticated. This is evidenced by recent accomplishments. Hongrong Liu et. al. determined a structure of Adenovirus at a resolution of 3.6 Å via CryoEM [7]. This is an impressive accomplishment, showcasing the resolving power of the technique. The true power of CryoEM does not lie in high resolution alone, however. The most important advantage of CryoEM is the ability to study the structures of small biological molecular systems in a native aqueous environment. Binding interactions between proteins are by definition flexible; the structural rigidity found in crystal structures of binding interactions is largely an artifact of the crystallization procedure. Crystallization also denies the possibility of observing several different modes of binding, or even asymmetric binding to a protein complex. The CryoEM structure of Ad5.F35 + 5 μ M HD5 is unable to resolve a rigid structure of the HD5 binding interaction, but what it does reveal is the full range of possibilities of binding, presented as what one might interpret as

a probability cloud of electron density. Simply knowing *where* binding occurs is an important first step in determining how binding occurs. Knowledge of these general binding locations opens up possibilities for future experiments, including but not limited to localized mutagenesis studies, and molecular modeling of interactions.

Significance of Determining as Subnanometer Structure of Ad5.PB/GYAR + 5 μ M HD5

In the field of structural biology, determining new structures is rarely a wasted effort, even if the structural determination is an end in itself. Structures of a number of serotypes of Adenovirus, both natural and chimeric have previously been determined. Ad5.F35 in particular, has been used due to its short fiber, facilitating grid preparation, imaging, and reconstruction. Ad5.PB/GYAR has no such advantages for CryoEM, but its structure was necessary in the context of determining the mechanisms of HD5 inactivation of Adenovirus. Ad5.PB/GYAR was known to be unaffected by HD5 in terms of infectivity [12], but it was not known whether this was due to a lack of binding, or a lack of activity following binding. Determining the structures of Ad5.PB/GYAR + 5 μ M HD5 and Ad5.PB/GYAR without HD5 definitively showed that there is, in fact, no binding of HD5 in the areas seen in the Ad5.F35 + 5 μ M HD5 structure. This lends credence to the idea that the vertex region is the critical site which determines whether a virus is inactivated by HD5 or not.

Significance of Predicting Binding Sites Via Molecular Dynamics Simulation

Molecular dynamics is a powerful but relatively untapped tool in the field of structural biology. Nowadays there is a large abundance of biomolecular structures that have been determined with atomic detail. This wealth of information is lacking in one respect, however. All of these structures are represented as rigid entities whose ranges of motion and activity can only be speculated through indirect measurements or educated guesswork. Molecular dynamics simulations change this, by allowing these molecular structures to move and change conformation in a simulated environment. Just ten years ago it would have been nearly impossible for a single research group to conduct a molecular dynamics simulation of the scale undertaken in this project, simply because the computational power required did not exist outside of specialized supercomputing facilities. Now that computers are much more powerful and affordable, molecular dynamics will certainly become a widely used tool in all disciplines of molecular and structural biology.

The binding predictions undertaken in this study are not intended to be a replacement for traditional structure determination. Molecular dynamics simulations are far from perfect, and multiple restraints must be placed on simulations to provide realistic structural evolution. In this particular study, the binding of HD5 to the Adenovirus vertex region was simulated using an atomic

model of the vertex region and the CryoEM density of Ad5.F35 + 5 μ M HD5 determined in Aim 1 as a restraint. The goal of this simulation was to observe protein behaviors that could not be visualized through CryoEM or other structural techniques, particularly the behavior of flexible regions of the Ad5.F35 penton base RGD loops. At the completion of the 100 picosecond simulation, the flexible RGD loops could be seen enveloping HD5 peptides against the base of the fiber. This suggests that the RGD loops may in fact serve a role in stabilizing HD5 binding to the base of the fiber protein.

Although the result of the simulation is not definitive in and of itself, it provides lines of inquiry that can be further examined in future experiments. For example, in laboratory studies [12] chimeras of HAdV-5 and HAdV-19c were created, each with varying degrees of inactivation when exposed to HD5. The chimera of HAdV-5 with a substituted HAdV-19c fiber was completely resistant to HD5. The chimera of HAdV-5 with substituted fiber and penton base (Ad5.PB/GYAR) was also completely resistant to HD5. The chimera of HAdV-5 with a substituted HAdV-19c penton base was only partially resistant to HD5 with a threshold level of 10 μ M HD5 required to inactivate the virus. This is of particular interest because this demonstrates that an HD5-sensitive fiber is not in itself enough to ensure inactivation via HD5 exposure. Perhaps this intermediate level of inactivation is due to the difference in RGD loops between the HAdV-19c penton base and the HAdV-5 penton base. The flexible region of the RGD loop in HAdV-5 is over twice as long as the RGD loop in HAdV-19c, 78 amino acids in

HAdV-5 and 32 amino acids in HAdV-19c. It may very well be that the shorter loops are incapable of enveloping and stabilizing HD5 against the base of the fiber, especially since the simulation suggests that the RGD loops responsible for this stabilization are attached to the adjacent penton monomer to where the binding has actually occurred. A CryoEM study of the HAdV-19c penton base substitution followed by molecular dynamics modeling may provide insight into the differences in HD5 binding to the penton base, given differing RGD loop regions.

Without the molecular dynamics simulation undertaken in this aim, the role of the RGD loop regions could only be speculated, and may not have received much consideration at all due to the complete lack of structural information of these flexible regions. Molecular dynamics simulations provide a way to extract insight from processes that cannot be directly observed. Even without definitively accurate simulations, conclusions can still be drawn from the data generated. These simulations will undoubtedly become more and more refined as research in simulation optimization and hardware progresses.

REFERENCES

1. Goncalves, M.A. and A.A. de Vries, *Adenovirus: from foe to friend*. Rev Med Virol, 2006. **16**(3): p. 167-86.
2. Vetrini, F. and P. Ng, *Liver-directed gene therapy with helper-dependent adenoviral vectors: current state of the art and future challenges*. Curr Pharm Des, 2011. **17**(24): p. 2488-99.
3. Nayak, S. and R.W. Herzog, *Progress and prospects: immune responses to viral vectors*. Gene Ther, 2010. **17**(3): p. 295-304.
4. Ganz, T., *Defensins: antimicrobial peptides of innate immunity*. Nat Rev Immunol, 2003. **3**(9): p. 710-20.
5. Ding, J., Y.Y. Chou, and T.L. Chang, *Defensins in viral infections*. J Innate Immun, 2009. **1**(5): p. 413-20.
6. Reddy, V.S., et al., *Crystal structure of human adenovirus at 3.5 Å resolution*. Science, 2010. **329**(5995): p. 1071-5.
7. Liu, H., et al., *Atomic structure of human adenovirus by cryo-EM reveals interactions among protein networks*. Science, 2010. **329**(5995): p. 1038-43.
8. Carson, S.D., *Coxsackievirus and adenovirus receptor (CAR) is modified and shed in membrane vesicles*. Biochemistry, 2004. **43**(25): p. 8136-42.
9. Moyer, C.L., et al., *Functional genetic and biophysical analyses of membrane disruption by human adenovirus*. J Virol, 2011. **85**(6): p. 2631-41.
10. Nemerow, G.R., et al., *Insights into adenovirus host cell interactions from structural studies*. Virology, 2009. **384**(2): p. 380-8.
11. Smith, J.G. and G.R. Nemerow, *Mechanism of adenovirus neutralization by Human alpha-defensins*. Cell Host Microbe, 2008. **3**(1): p. 11-9.
12. Smith, J.G., et al., *Insight into the mechanisms of adenovirus capsid disassembly from studies of defensin neutralization*. PLoS Pathog, 2010. **6**(6): p. e1000959.
13. Saban, S.D., et al., *Visualization of alpha-helices in a 6-angstrom resolution cryoelectron microscopy structure of adenovirus allows refinement of capsid protein assignments*. J Virol, 2006. **80**(24): p. 12049-59.
14. Trabuco, L.G., et al., *Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics*. Structure, 2008. **16**(5): p. 673-83.
15. Rux, J.J. and R.M. Burnett, *Type-specific epitope locations revealed by X-ray crystallographic study of adenovirus type 5 hexon*. Mol Ther, 2000. **1**(1): p. 18-30.
16. Rux, J.J., P.R. Kuser, and R.M. Burnett, *Structural and phylogenetic analysis of adenovirus hexons by use of high-resolution x-ray crystallographic, molecular modeling, and sequence-based methods*. J Virol, 2003. **77**(17): p. 9553-66.
17. Keefe, L.J., et al., *Crystallization and preliminary X-ray diffraction studies of the human adenovirus serotype 2 proteinase with peptide cofactor*. Protein Sci, 1995. **4**(8): p. 1658-60.
18. Devaux, C., P.A. Timmins, and C. Berthet-Colominas, *Structural studies of adenovirus type 2 by neutron and X-ray scattering*. J Mol Biol, 1983. **167**(1): p. 119-32.
19. Devaux, C., et al., *Structure of adenovirus fibre. I. Analysis of crystals of fibre from adenovirus serotypes 2 and 5 by electron microscopy and X-ray crystallography*. J Mol Biol, 1990. **215**(4): p. 567-88.
20. Reddy, V.S., et al., *Crystallization and preliminary X-ray diffraction analysis of human adenovirus*. Virology, 2010. **402**(1): p. 209-14.
21. Baniecki, M.L., et al., *Adenovirus proteinase: crystallization and preliminary X-ray diffraction studies to atomic resolution*. Acta Crystallogr D Biol Crystallogr, 2002. **58**(Pt 9): p. 1462-4.
22. Adrian, M., et al., *Cryo-electron microscopy of viruses*. Nature, 1984. **308**(5954): p. 32-6.
23. Li, X., N. Grigorieff, and Y. Cheng, *GPU-enabled FREALIGN: accelerating single particle 3D reconstruction and refinement in Fourier space on graphics processors*. J Struct Biol, 2010. **172**(3): p. 407-12.

24. Ludtke, S.J., P.R. Baldwin, and W. Chiu, *EMAN: semiautomated software for high-resolution single-particle reconstructions*. J Struct Biol, 1999. **128**(1): p. 82-97.
25. van Heel, M., et al., *A new generation of the IMAGIC image processing system*. J Struct Biol, 1996. **116**(1): p. 17-24.
26. Grigorieff, N., *FREALIGN: high-resolution refinement of single particle structures*. J Struct Biol, 2007. **157**(1): p. 117-25.
27. Zubieta, C., et al., *The structure of the human adenovirus 2 penton*. Mol Cell, 2005. **17**(1): p. 121-35.
28. Xia, D., et al., *Crystal structure of the receptor-binding domain of adenovirus type 5 fiber protein at 1.7 Å resolution*. Structure, 1994. **2**(12): p. 1259-70.
29. Humphrey, W., A. Dalke, and K. Schulten, *VMD: visual molecular dynamics*. J Mol Graph, 1996. **14**(1): p. 33-8, 27-8.
30. Phillips, J.C., et al., *Scalable molecular dynamics with NAMD*. J Comput Chem, 2005. **26**(16): p. 1781-802.
31. DeCaprio, J.A., *How the Rb tumor suppressor structure and function was revealed by the study of Adenovirus and SV40*. Virology, 2009. **384**(2): p. 274-84.
32. Szyk, A., et al., *Crystal structures of human alpha-defensins HNP4, HD5, and HD6*. Protein Sci, 2006. **15**(12): p. 2749-60.
33. Goncalves, M.A., et al., *Targeted chromosomal insertion of large DNA into the human genome by a fiber-modified high-capacity adenovirus-based vector system*. PLoS One, 2008. **3**(8): p. e3084.
34. Knaan-Shanzer, S., et al., *Endowing human adenovirus serotype 5 vectors with fiber domains of species B greatly enhances gene transfer into human mesenchymal stem cells*. Stem Cells, 2005. **23**(10): p. 1598-607.
35. Korokhov, N., et al., *Targeting of adenovirus via genetic modification of the viral capsid combined with a protein bridge*. J Virol, 2003. **77**(24): p. 12931-40.
36. Poulin, K.L., et al., *Retargeting of adenovirus vectors through genetic fusion of a single-chain or single-domain antibody to capsid protein IX*. J Virol, 2010. **84**(19): p. 10074-86.
37. Vetrini, F. and P. Ng, *Gene therapy with helper-dependent adenoviral vectors: current advances and future perspectives*. Viruses, 2010. **2**(9): p. 1886-917.
38. Seregin, S.S. and A. Amalfitano, *Improving adenovirus based gene transfer: strategies to accomplish immune evasion*. Viruses, 2010. **2**(9): p. 2013-36.
39. Thaci, B., et al., *The challenge for gene therapy: innate immune response to adenoviruses*. Oncotarget, 2011. **2**(3): p. 113-21.
40. Nguyen, E.K., G.R. Nemerow, and J.G. Smith, *Direct evidence from single-cell analysis that human {alpha}-defensins block adenovirus uncoating to neutralize infection*. J Virol, 2010. **84**(8): p. 4041-9.
41. Buck, C.B., et al., *Human alpha-defensins block papillomavirus infection*. Proc Natl Acad Sci U S A, 2006. **103**(5): p. 1516-21.
42. Chong, K.T., et al., *High level expression of human epithelial beta-defensins (hBD-1, 2 and 3) in papillomavirus induced lesions*. Virol J, 2006. **3**: p. 75.