Geometric Modeling and Quantitative Visualization of Virus Ultra-structure

Chandrajit Bajaj

Department of Computer Science & Institute of
Computational Engineering and Sciences,
Center for Computational Visualization,
University of Texas at Austin,
201 East 24th Street, ACES 2.324A,
Austin, TX 78712-0027
Phone: +1 512-471-8870, Fax: +1 512-471-0982
Email : bajaj@cs.utexas.edu

Total number of actual words:

Total number of figures:

Total word equivalents:

Acknowledgements: This work was supported in part by NSF-ITR grants ACI-022003, and EIA-0325550, and grants from the NIH 0P20 RR020647, and R01 GM074258.

1. Introduction

Viruses are one of the smallest parasitic nano-organisms that are agents of human disease [89]. They have no systems for translating RNA, ATP generation, or protein, nucleic acid synthesis, and therefore need the subsystems of a host cell to sustain and replicate [89]. It would be natural to classify these parasites according to their eukaryotic or prokaryotic cellular hosts (e.g. plant, animal, bacteria, fungi, etc.), however there do exist viruses which have more than one sustaining host species [89]. Currently, viruses are classified simultaneously via the host species(Algae, Archae, Bacteria, Fungi, Invetebrates, Mycoplasma, Plants, Protozoa, Spiroplasma, Vetebrates), the host tissues that are infected, the method of virial transmission, the genetic organization of the virus (single or double stranded, linear or circular, RNA or DNA), the protein arrangement of the protective closed coats housing the genome (helical, icosahedral symmetric nucleo-capsids), and whether the virus capsids additionally have a further outer envelope covering (the complete virion)[89]. Table 1 summarizes a small yet diverse collection of viruses and virions [98]. The focus of this article is on the computational geometric modeling and visualization of the nucleo-capsid ultrastructure of plant and animal viruses

exhibiting the diversity and geometric elegance of the multiple protein arrangements. Additionally, one computes a regression relationship between surface area v.s. enclosed volume for spherical viruses with icosahedral symmetric protein arrangements. The computer modeling and quantitative techniques for virus capsid shells ultra-structure that we review here are applicable for atomistic, high resolution (less than 4 A) model data, as well as medium (5 A to 15 A) resolution map data reconstructed from cryo-electron microscopy.

2. The Morphology of Virus Structures

Minimally viruses consist of a single nucleocapsid made of proteins for protecting their genome, as well as in facilitating cell attachment and entry. The capsid proteins magically self-assemble, into often a helical or icosahedral symmetric shell (henceforth referred to as capsid shells). There do exist several examples of capsid shells which do not exhibit any global symmetry [98], however we focus on only the symmetric capsid shells in the remainder of this article.

Different virus morphologies that are known, (a small sampling included in Table 1) are distinguished by

optional additional outer capsid shells, the presence or lack of a surrounding envelope for these capsid shells (derived often from the host cell's organelles membranes), as well as additional proteins within these optional capsids and envelopes, that are necessary for the virus lifecycle. The complete package of proteins, nucleic acids and envelopes is often termed a virion.

Fig. 2.1. Organization of Helical Viruses

The asymmetric structural subunit of a symmetric capsid shell may be further decomposable into simpler and smaller protein structure units termed protomers. Protomers could be a single protein in monomeric form (example TMV), or form homogeneous dimeric or trimeric structure units (example RDV). These structure units also often combine to form symmetric clusters, called capsomers, and are predominantly distinguishable in visualizations at even medium and low resolution virus structures. The capsomers and/or protomeric structure units pack to create the capsid shell in the form of either helical or icosahedral symmetric arrangements, with a greater propensity for icosahedral symmetry.

Fig. 2.2. Organization of Icosahedral Viruses

The subsequent sub-sections dwell on the geometry of the individual protomers, and capsomers, as part of a hierarchical arrangement of symmetric capsid shells.

2.1 The Geometry of Helical Capsid Shells

Helical symmetry can be captured by a 4 x 4 matrix transformation $H_{(\vec{a},\phi,L)}$ parameterized by $\vec{a}=(a_x,a_y,a_z)$, a unit vector along the helical axis, by θ, an angle in the plane of rotation, and by the pitch L, the axial rise for a complete circular turn.

$$H_{(\vec{a},\phi,L)}=\begin{bmatrix} a_x^2(1-\cos\theta)+\cos\theta & a_xa_y(1-\cos\theta)-a_z\sin\theta & a_xa_z(1-\cos\theta)+a_y\sin\theta & \frac{a_xL\theta}{2\pi} \\ a_xa_y(1-\cos\theta)+a_z\sin\theta & a_y^2(1-\cos\theta)+\cos\theta & a_ya_z(1-\cos\theta)-a_x\sin\theta & \frac{a_yL\theta}{2\pi} \\ a_xa_z(1-\cos\theta)-a_y\sin\theta & a_ya_z(1-\cos\theta)+a_x\sin\theta & a_z^2(1-\cos\theta)+\cos\theta & \frac{a_zL\theta}{2\pi} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

If P is the center of any atom of the protomer, then P' is the transformed center, and $P'=H*P$. Repeatedly applying this transformation to all atoms in a protomer yields a helical stack of protomeric units. The desired length of the helical nucleo-capsid shell

is typically determined by the length of the enclosed nucleic acids. The capsid shell of the tobacco mosaic virus (TMV) exhibits helical symmetry (Fig. 2.1), with the asymmetric protein structure unit or the protomer consisting of a single protein (pdb id 1EI7)

Fig. 2.3 Helical Symmetry Axis

2.2 The Geometry of Icosahedral Capsid Shells

More often, the virus structure is icosahedrally symmetric. The advantage over the helical symmetry structure is the efficient construction of a capsid of a given size using the smallest protein subunits. An icosahedron has 12 vertices, 20 equilateral triangular faces, and 30 edges, and exhibits 5:3:2 symmetry. A 5-fold symmetry axis passes through each vertex, a 3-fold symmetry axis through the center of each face, and a 2-fold axis through the midpoint of each edge (see Fig. 2.3).

Fig. 2.4 Icosahedral Symmetries and Axes

A rotation transformation around an axis $\vec{a} = (a_x, a_y, a_z)$ by an angle θ is described by the 4x4 matrix

$$R_{(\vec{a},\phi)} = \begin{bmatrix} a_x^2(1-\cos\theta)+\cos\theta & a_x a_y(1-\cos\theta)-a_z\sin\theta & a_x a_z(1-\cos\theta)+a_y\sin\theta & 0 \\ a_x a_y(1-\cos\theta)+a_z\sin\theta & a_y^2(1-\cos\theta)+\cos\theta & a_y a_z(1-\cos\theta)-a_x\sin\theta & 0 \\ a_x a_z(1-\cos\theta)-a_y\sin\theta & a_y a_z(1-\cos\theta)+a_x\sin\theta & a_z^2(1-\cos\theta)+\cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The vertices of a canonical icosahedron are given by {(0,±1,±Φ), (±1,±Φ,0), (±Φ,0,±1)}, where Φ = (1+√5)/2 is the golden ratio. For a 5-fold symmetry transformation around the vertex (0,±1,±Φ) the normalized axis of rotation is $\vec{a} = (0, 0.52573, 0.85064)$ and the angle of rotation is $\theta = \frac{2\pi}{5}$ yielding a five fold symmetry transformation matrix

$$R_{(5-fold)} = \begin{bmatrix} 0.30902 & -0.80902 & 0.5000 & 0 \\ 0.80902 & 0.5000 & 0.30902 & 0 \\ -0.5000 & 0.30902 & 0.80902 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Similarly, one is able to construct five fold symmetry transformation matrices for the other icosahedron vertices. Using the generic rotational transformation matrix $R_{(\vec{a},\phi)}$, one is able to construct the three fold transformation matrices via the rotation axis passes

through the centroid of the triangular faces of the icosahedron and an angle of rotation of $\theta = \frac{2\pi}{3}$. Consider the triangular face with corners at (0,1,Φ),(0,-1,Φ) and (Φ,0,1). The centroid is at (Φ/3,0,(2Φ+1)/3) and the normalized axis of rotation is $\vec{a} = (0.356822,0,0.934172)$ and the transformation matrix

$$R_{(3-fold)} = \begin{bmatrix} 0.30902 & -0.80902 & 0.5000 & 0 \\ 0.80902 & -0.5000 & -0.30902 & 0 \\ 0.5000 & 0.30902 & 0.80902 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

A polyhedron with faces all equilateral triangles is called a deltahedron. Deltahedra with icosahedral symmetry are classified as icosadeltahedra. Any icosadeltahedron has 20T facets, where T is the *triangulation number* given by $T = Pf^2$, where $P = h^2 + hk + k^2$, for all pairs of integers h and k which do not have common factor, and f is any integer [15]. The possible values of P are $1, 3, 7, 13, 19, 21, 31, 37, \cdots$. In Fig. 2.4 we display triangles with different triangulation numbers, for icosahedral virus structures.

Fig. 2.5 Architecture of Icosahedral Viruses, Caspar-Klug Triangulation Numbers, Asymmetric structure units

The greater the T number, the larger the size of the virus capsid. Each triangular portion of the icosahedral virus capsid is easily subdivided into its three asymmetrical units, with each unit containing some combination of protein structure units (protomers). In total an icosahedral virus capsid has 60T asymmetrical units with numerous proteins structures inter-twined to form a spherical mosaic. In Fig. 2.4 we see that when T=1, each vertex is at the center of a pentagon, and the capsid proteins are in an equivalent environment, i.e. five neighbors cluster at a common vertex. However, for icosadeltahedra with larger triangulation numbers, e.g. T=3, there are pentagons and hexagons in the capsid mosaic (Fig. 2.5). Therefore, even though the capsid proteins (protomers) may be chemically identical, some cluster into a local 5-fold neighborhood and the others into a local 6-fold neighborhood. Such locally symmetric clusterings of protomers are alternatively termed capsomers. In these situations, the proteins are no longer global symmetrically equivalent, but only *quasi-equivalent* [15].

3. Surface and Volumetric Modeling and Visualization

3.1 Atomistic Resolution Model Structures

Numerous schemes have been used to model and visualize bio-molecules and their properties [47]. All these different visual representation are often derived from an underlying geometric model constructed from the positions of atoms, bonds, chains, and residues information deposited as part of an atomic resolution structure of the protein or nucleic acid in the Protein Data Bank (PDB). Hence, structural models are designed to represent the primary (sequence), secondary (e.g. α- helices, β-sheets), tertiary (eg. α – β barrels), and quaternary (fully folded) geometric structures of the protein or nucleic acid.

An early approach to molecular modeling is to consider atoms as hard spheres, and their union as an attempt to capture shape properties as well as spatial occupancy of the molecule. This is similar to our perception of surfaces and volume occupancy of macroscopic objects. The top two pictures in Figure 2.4 shows hard-sphere model visualizations of the twin Rice Dwarf capsid shells, with individual proteins colored differently. Solvated versions of these

molecular surfaces have been proposed by Lee – Richards, Connolly, et. al. for use in computational biochemistry and biophysics. Much of the preliminary work, along with later extensions focussed on finding fast methods of triangulating this molecular surface (or as sometimes referred to as the solvent contact surface). Two prominent obstacles in modeling are the correct handling of surface self-intersections (singularities) and the high communication bandwidth needed when sending tessellated surfaces to the graphics hardware.

Figure 3.1 Analytic surface models of capsid shells of icosahedral viruses

A more analytic and smooth description of molecular surfaces (without singularities) is provided by a suitable level set of the electron density representation of the molecule. Isotropic Gaussian kernels have been traditionally used to describe atomic electron density due to their ability to

approximate electron orbitals. The electron density of a molecule with M atoms, centered at $\overrightarrow{x_j}, j \in \{1 \ldots M\}$ can thus be written as $F_{elec\_dens}(\vec{x}) = \sum_{j=1}^{M} \gamma_j K(\vec{x} - \overrightarrow{x_j})$ where γ_j and K are typically chosen from a quadratic exponential description of atomic electron density

$$Atom(\vec{x}) = e^{-\frac{d}{r^2}((\vec{x}-\vec{y})^2 - r^2)} = e^d e^{-\frac{d}{r^2}(\vec{x}-\vec{y})^2}$$

$$= AK^q(\vec{x} - \vec{y})\gamma_{elec\_dens}(\vec{x})$$

$$= e^{-\frac{d}{r^2}((\vec{x}-\overrightarrow{x_j})^2 - r^2)} = \gamma_j K(\vec{x} - \overrightarrow{x_j})$$

The atomic electron density kernels are affected by the radius r of individual atoms and the decay parameter d. Smooth and molecular surface models for individual proteins, structure units, as well as entire capsid shells can be easily constructed as a

fixed level set of $F_{elec\_dens}(\vec{x}) = \sum_{j=1}^{M} \gamma_j K(\vec{x} - \vec{x_j})$. An array of such structural molecular model visualizations are shown as Figures 2.1 – 2.5 as well as figure 3.1. Some of them use transparency on the solvated molecular surface and show the protein back-bone structure (folded chains of α-helices and β-sheets).

3.2 Structure Elucidation from 3D Maps

Electron Microscopy (EM) and in particular single particle reconstruction using cryo-EM, has rapidly advanced over recent years, such that several virus structures (tertiary and secondary) can be resolved routinely at low resolution (10-20 A) and in some cases at sub-nanometer (intermediate) resolution (7-10 A) [7,8].

Figure 3.2 Structure Elucidation from 3D Maps of Icosahedral Viruses

Symmetry of the virus capsid shells are exploited both in the 3D Map reconstruction from raw 2D EM images, as

well in structure elucidation in the 3D Map. In many cases, the 3D maps are of spherical viruses, with protein capsid shells exhibiting icosahedral symmetry. In these cases, the global symmetry detection can be simplified to computing the location of the 5-fold rotational symmetry axes, passing through the twelve vertices of the icosahedron, from which the 3-fold symmetry axis for the twenty icosahedron faces and the 2-fold symmetry axis for the thirty icosahedron edges can be easily derived. However determining the local symmetries of the capsomers (structure units) are more complicated, as they exhibit varied k-fold symmetry, and their detection requires a modified correlation based search algorithm [94]. Volumetric segmentation methods are additionally utilized to partition, color and thereby obtain a clearer view into the macromolecules architectural organization. Furthermore, electronically dissecting the local structure units from a 3D Map allows for further structural interpretation (tertiary and secondary folds). Visualizations from the afore-mentioned local symmetry

detection and automatic segmentation, applied to a 3D volumetric Map of the Turnip Yellow Mosaic virus (pdbid 1AUY), are shown in Figure~3.2.

4. Quantitative Visualization

The geometric modeling of virus capsids and the individual virus structure units, can be further augmented by the computation of several global and local shape metrics. While integral, topological and combinatorial metrics capture global shape properties, differential measures such as mean and Gaussian curvatures have also proved useful to an enhanced understanding and quantitative visualization of macromolecular structures.

4.1 Integral Properties

Integral shape metrics include the area of the molecular capsid surface defining the capsid, the volume enclosed by closed capsid shells, and the gradient integral on the molecular capsid surface. Given our smooth analytic level set definition of the

molecular surface from section 3, $F_{elec\_dens}(\vec{x}) = \sum_{j=1}^{M} \gamma_j K(\vec{x} - \overrightarrow{x_j}) = const$, for all the atoms that make up either an individual structure unit, or the entire virus capsid, an efficient and accurate integration computation for these metrics is given by the contour spectrum [5]. The surface integrations can be performed by adaptively sampling the capsid surface using a technique known as contouring [5]. Contouring is often performed by first decomposing (meshing) the space surrounding the capsid surface into either a rectilinear Cartesian grid mesh, a tetrahedral or a hexahedral mesh. For a tetrahedral mesh, the surface area for the portion of the level set inside a tetrahedron can be represented by a quadratic polynomial B-spline [5]. Summing these B-splines over all of the tetrahedra containing the capsid surface yields the capsid surface area. The volume enclosed by a closed capsid surface is determined by the definite integration of the surface area polynomial B-splines.

Fig. 4.1 Area, Volume Relationship for Icosahedral Viruses

In Figure 4.1 we display the results of surface area and volume calculations, and a regression relationship between the two, for a selection of spherical icosahedral capsids for virus structures summarized in Table 2. The analytic molecular surfaces were first computed, and then surface area and enclosed volume were estimated through B-spline evaluation as stated above.

4.2 Differential Properties

The gradient function of our smooth analytic capsid surface is simply $\nabla F_{elec\_dens}(\vec{x}) = \sum_{j=1}^{M} \gamma_j \nabla K(\vec{x} - \overrightarrow{x_j})$, the summation of the vector of first derivatives of the atomic electron density function. This gradient function is non-zero everywhere on the virus capsid surface (i.e. no singularity). The second derivatives of the molecular surface capture additional

differential shape properties and provide suitable metrics. Popular metrics are the magnitudes of ***Mean Curvature*** **H** and the ***Gaussian curvature*** **G**. These are given directly as **H** = ½(k_{min} + k_{max}) and **G** = $k_{min}k_{max}$ and are respectively the average and the product of the twin principal curvatures, namely, k_{min} and k_{max}, also sometimes known as the minimum and maximum curvatures at a point on the surface. Again for our level set based analytic molecular surface $F_{elec\_dens}(\vec{x}) = cons = f$, the twin curvatures **H** and **K** can be evaluated as **H** = ($\sum$(f_x^2 (f_{yy} + f_{zz})) −2 * $\sum$ ($f_x f_y f_{xy}$)) / (2 * ($\sum$ (f_x^2))$^{1.5}$) and **G** = (2 * $\sum$(f_x f_y (f_{xz} f_{yz} − f_{xy} f_{zz}))) / (($\sum$ (f_x^2))^2) where $\sum$ represents a cyclic summation over x, y and z, and where additionally $f_{x,\ etc.,}$ denotes partial differentiation with respect to those variables.

Displaying the magnitude of the gradient function and its variation, as expressed by the mean and Gaussian curvature functions over a molecular surface helps quantitatively visualize the bumpiness or lack thereof

of an individual protomer, a structure unit or the entire viral capsid. In Figures 2.1 the bottom two pictures display the mean and Gaussian curvature functions of the Tobacco Mosaic virus asymmetric protomer surface, exhibiting and enhancing the bumpiness of the surface.

4.3 Topological and Combinatorial Properties

Affine invariant topological structures of volumetric functions f , such as our smooth analytic electron density function of section 3, include the Morse complex [28, 56] and the contour tree (CT) [46]. Both the Morse complex and contour tree are related to the critical points of the volumetric function f , i.e., those points in the domain M where the function gradient vanishes $\nabla f = 0$. The *functional range* of f is the interval between the minimum and maximum values of the function $f : [f_{min}, f_{max}]$. For a scalar value $w \in [f_{\min}, f_{\max}]$, the level set of the field f at the value w is the subset of points $L(w) \subset M$ such that $f(x) = w \forall x \in L(w)$.

A level set may have several connected components, called contours. The topology of the level set $L(w)$ changes only at the critical points in M, whose corresponding functional values are called critical values. A *contour class* is a maximal set of continuous contours which have the same topology and do not contain critical points. Without loss of generality, the critical points are assumed to be non-degenerate, i.e. only isolated critical points. This assumption can be enforced by small perturbations of the function values. If the critical points are non-degenerate, then the *Hessian* $H(a)$ at a critical point a has non-zero real eigenvalues. The *index* of the critical point a is the number of negative eigenvalues of $H(a)$. For a 3D volumetric function, there are four types of critical points: index 0 (minima), indices 1 and 2 (saddle points), and index 3 (maxima).

The contour tree (CT) was introduced by Kreveld et al. [46] to find the connected components of level sets for contour generation. The CT captures the topological changes of the level sets for the entire

functional range $[f_{min}, f_{max}]$ of f; each node of the tree corresponds to a critical point and each arc corresponds to a contour class connecting two critical points. As an example, the contour tree for a virus capsid is shown in 4.2. Each leaf node of the CT represents the creation or deletion of a component at a local minimum or maximum and each interior node represents the joining and/or splitting of two or more components or topology changes at the saddle points. A cut on an arc of the tree $(v_1, v_2) \in T$ by an isovalue $v_1 \leq w \leq v_2$ represents a contour of the level set $L(w)$. Therefore, the number of connected components for the level set $L(w)$ is equal to the number of cuts to the CT at the value w. The CT can be enhanced by tagging arcs with topological information such as the Betti numbers of the corresponding contour classes [46]. Betti numbers $\beta_k \quad (k = 0,1,\ldots)$ intuitively measure the number of k-dimensional holes of a virus capsid surface or of any individual structure unit. Only the first three Betti numbers $(\beta_0, \beta_1, \beta_2)$ of a smooth

surface are non-zero: β_0 corresponds to the number of connected components; β_1 corresponds to the number of independent tunnels; β_2 represents the number of voids enclosed by the surface. For example, a sphere has the Betti numbers $(\beta_0, \beta_1, \beta_2) = (1,0,1)$ while a torus has $(\beta_0, \beta_1, \beta_2) = (1,2,1)$. Betti number computations for virus capsid surfaces provide useful topological and combinatorial structural information.

5. Conclusion

Ultra-structure modeling and visualization of virus capsids are clearly just a couple of the steps in a computational modeling pipeline for determining structure to function relationships for such nano-organisms. Efforts are underway by several groups for virus energetics in solvated environs, atomistic and coarse grained virus dynamics, as well as interactions and binding of various ligands and proteins to the nucleo-capsids.

Acknowledgements: Sincere thanks to my students S. Goswami, S. Siddahanvalli, J. Wiggins and W. Zhao for their help with this manuscript. Thanks also to invaluable discussions with Dr. Tim Baker at Univ. of California, San Diego.

6. REFERENCES

[1] Aloy P, Russell RB (2002) The third dimension for protein interactions and complexes. Trends Biochem Sci. 27(12): 633-638.

[2] Anderer FA, Schlumberger HD, Koch MA, Frank H, Eggers HJ (1967) Structure of simian virus 40. II. Symmetry and components of the virus particle. Virology. 32(3): 511-23.

[3] (8)Arnold E, Rossmann MG (1988) The use of molecular-replacement phases for the refinement of the human rhinovirus 14 structure. Acta Crystallogr., Sect.A 44:270-282.

[4] Aroul-Selvam R, Hubbard T, Sasidharan R (2004) Domain insertions in protein structures. J Mol Biol. 338(4): 633-641.

[5] Bajaj C, Pasucci V, Schikore D (1997) The Contour Spectrum *IEEE Visualization '97*.

[6] Bajaj CL, Pasucci V, Schikore D (1998) Visualization of scalar topology for structural enhancement *IEEE Visualization '98*, D. Ebert, H. Hagen, and H. Rushmeir, Editors. 51-58.

[7] Baker TS, Olson NH, Fuller SD (1999) Adding the third dimension to virus life cycles: Three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. Microbiol. Mol. Biol. Rev. 63(4): 862-922.

[8] Belnap DM, Kumar A, Folk JT, Smith TJ, Baker TS (1999) Low-resolution density maps from atomic models:

how stepping "back" can be a step "forward". J Struct Biol. 125(2-3): 166-175.

[9] Belnap DM, Olson NH, Cladel NM, Newcomb WW, Brown JC, Kreider JW, Christensen ND, Baker TS (1996) Conserved features in papillomavirus and polyomavirus capsids. J Mol Biol. 259(2): 249-263.

[10] Berger B, Shor PW, Tucker-Kellogg L, King J (1994) Local rule-based theory of virus shell assembly. Proc Natl Acad Sci U S A 91(16): 7732-7736.

[11] Bhyravbhatla B, Watowich SJ, Caspar DL (1998) Refined atomic model of the four-layer aggregate of the Tobacco Mosaic Virus coat protein at 2.4-A resolution. Biophysics J. 74: 604-615.

[12] Burley SK, Bonanno JB (2002) Structuring the universe of proteins. Annu Rev Genomics Hum Genet. 3: 243-262. Epub 2002 Apr 15.

[13] (14)Canady MA, Larson SB, Day J, McPherson A (1996) Crystal structure of turnip yellow mosaic virus. Nat.Struct.Biol. 3:771-781.

[14] Carfi A, Willis SH, Whitbeck JC, Krummenacher C, Cohen GH, Eisenberg RJ, Wiley DC (2001) Herpes simplex virus glycoprotein D bound to the human receptor HveA. Mol.Cell 8:169-179.

[15] Caspar DL, Klug A (1962) Physical principles in the construction of regular viruses. Cold Spring Harb Symp Quant Biol. 27: 1-24.

[16] Caston JR, Trus BL, Booy FP, Wickner RB, Wall JS, Steven AC (1997) Structure of L-A virus: a specialized compartment for the transcription and replication of double-stranded RNA. J Cell Biol. 138(5): 975-985.

[17] Ceulemans H, Russell RB (2004) Fast fitting of atomic structures to low-resolution electron density

maps by surface overlap maximization. J Mol Biol. 338(4): 783-793.

[18] Chen WN, Oon CJ (1999) Human hepatitis B virus mutants: significance of molecular changes. FEBS Lett. 453(3): 237-242.

[19] Chen XS, Stehle T, Harrison SC (1998) Interaction of polyomavirus internal protein VP2 with the major capsid protein VP1 and implications for participation of VP2 in viral entry. EMBO J. 17:3233-3240.

[20] Cheng RH, Caston JR, Wang GJ, Gu F, Smith TJ, Baker TS, Bozarth RF, Trus BL, Cheng N, Wickner RB, Steven AC (1994) Fungal virus capsids, cytoplasmic compartments for the replication of double-stranded RNA, formed as icosahedral shells of asymmetric Gag dimers. J Mol Biol. 244(3): 255-258.

[21] Crowther RA (1971) Procedures for three-dimensional reconstruction of spherical viruses by

Fourier synthesis from electron micrographs. Philos Trans R Soc Lond B Biol Sci. 261(837): 221-30.

[22] Crowther RA, DeRosier DJ, Klug A (1970) The reconstruction of a three-dimensional structure from projections and its application to electron microscopy. Proc. Roy. Soc. Lond. A 317: 319-340.

[23] Dar AC, Sicheri F (2002) X-ray crystal structure and functional analysis of Vaccinia virus K3L reveals molecular determinants for pkr subversion and substrate recognition. Molecular Cell. 10:295-305.

[24] de Haas F, Paatero AO, Mindich L, Bamford DH, Fuller SD (1999) A symmetry mismatch at the site of RNA packaging in the polymerase complex of dsRNA bacteriophage phi6. J Mol Biol. 294(2): 357-372.

[25] DeRosier DJ, Klug A (1968) Reconstruction of three dimensional structures from electron micrographs. Nature 217: 130-134.

[26] Dong G, Lu G, Chiu W (2001) Electron cryomicroscopy and bioinformatics suggest protein fold models for rice dwarf virus. Nat Struct Biol. 8(10):868-873.

[27] Dormitzer PR, Sun Z-YJ, Wagner G, Harrison SC (2002) The Rhesus Rotavirus VP4 Sialic Acid Binding Domain has a Galectin Fold with a Novel Carbohydrate Binding Site Embo J. 21:885-897.

[28] Edelsbrunner H, Harer J, Zomorodian A (2001) Hierarchical Morse complexes for piecewise linear 2-manifolds, *Symposium on Computational Geometry*. 70-79.

[29] Enemark EJ, Stenlund A, Joshua-Tor L (2002) Crystal structures of two intermediates in the assembly of the papillomavirus replication initiation complex. EMBO J. 21:1487-1496.

[30] Fuller SD, Butcher SJ, Cheng RH, Baker TS (1996) Three-dimensional reconstruction of icosahedral particles--the uncommon line. J Struct Biol. 116(1): 48-55.

[31] Goldsmith-Fischman S, Honig B (2003) Structural genomics: computational methods for structure analysis. Protein Sci. 12(9): 1813-1821.

[32] Grimes J, Basak AK, Roy P, Stuart D (1995) The crystal structure of bluetongue virus VP7. Nature 373(6510): 167-170.

[33] Grimes JM, Burroughs JN, Gouet P, Diprose JM, Malby R, Zientara S, Mertens PP, Stuart DI (1998) The atomic structure of the bluetongue virus core. Nature 395(6701): 470-478.

[34] Grimes JM, Jakana J, Ghosh M, Basak AK, Roy P, Chiu W, Stuart DI, Prasad BV (1997) An atomic model of the outer layer of the bluetongue virus core derived

from X-ray crystallography and electron cryomicroscopy. Structure 5(7): 885-893.

[35] Groft CM, Burley SK (2002) Recognition of eIF4G by rotavirus NSP3 reveals a basis for mRNA circularization. Mol.Cell 9:1273-1283.

[36] Grunewald K, Desai P, Winkler DC, Heymann JB, Belnap DM, Baumeister W, Steven AC (2003) Three-dimensional structure of herpes simplex virus from cryo-electron tomography. Science 302(5649): 1396-1398.

[37] He J, Schmid MF, Zhou ZH, Rixon F, Chiu W (2001) Finding and using local symmetry in identifying lower domain movements in hexon subunits of the herpes simplex virus type 1 B capsid. J Mol Biol. 309(4): 903-914.

[38] Helgstrand C, Munshi S, Johnson JE, Liljas L (2004) The Refined Structure of Nudaurelia Capensis

Omega Virus Reveals Control Elements for a T = 4 Capsid Maturation Virology 318:192-.

[39] Hopper P, Harrison SC, Sauer RT (1984) Structure of tomato bushy stunt virus. V. Coat protein sequence determination and its structural implications. J. Mol. Biol. 177:701-713.

[40] Jiang W, Li Z, Zhang Z, Baker ML, Prevelige PE Jr, Chiu W (2003) Coat protein fold and maturation transition of bacteriophage P22 seen at subnanometer resolutions. Nat Struct Biol. 10(2):131-135.

[41] Johnson JE (1996) Functional implications of protein-protein interactions in icosahedral viruses. Proc Natl Acad Sci U S A 93(1): 27-33.

[42] Johnson JE, Speir JA (1997) Quasi-equivalent viruses: a paradigm for protein assemblies. J Mol Biol. 269(5): 665-675.

[43] Kang BS, Devedjiev Y, Derewenda U, Derewenda ZS (2004) The PDZ2 domain of syntenin at ultra-high resolution: bridging the gap between macromolecular and small molecule crystallography. J Mol Biol. 338(3): 483-493.

[44] Kidd-Ljunggren K, Zuker M, Hofacker IL, Kidd AH (2000) The hepatitis B virus pregenome: prediction of RNA structure and implications for the emergence of deletions. Intervirology 43(3): 154-164.

[45] Klug A, Finch JT (1965) Structure of viruses of the papilloma-polyoma type. I. Human wart virus. J Mol Biol. 11: 403-423.

[46] van Kreveld M, van Oostrum R, Bajaj C, Pascucci V, Schikore D (1997) Contour Trees and small seed sets for isosurface traversal *Proc. 13th Annual Sympos. Comput. Geom.* 212-220.

[47] Leach AR (1996) Molecular Modeling. Principles and Applications, Longman.

[48] Lee MS, Feig M, Salsbury FR Jr, Brooks CL 3rd (2003) New analytic approximation to the standard molecular volume definition and its application to generalized Born calculations. J Comput Chem. 24(11): 1348-1356.

[49] Li N, Erman M, Pangborn W, Duax WL, Park CM, Bruenn J, Ghosh D (1999) Structure of Ustilago maydis killer toxin KP6 alpha-subunit. A multimeric assembly with a central pore. J.Biol.Chem. 274:20425-20431.

[50] Liddington RC, Yan Y, Moulai J, Sahli R, Benjamin TL, Harrison SC (1991) Structure of simian virus 40 at 3.8-A resolution. Nature 354(6351): 278-284.

[51] Liemann S, Chandran K, Baker TS, Nibert ML, Harrison SC (2002) Structure of the reovirus membrane-

penetration protein, Mu1, in a complex with is protector protein, Sigma3. Cell 108:283-295.

[52] Liljas L, Strandberg B (1984) The Structure of Satellite Tobacco Necrosis Virus. Bio. Macromol. Assem. 1:97-.

[53] Lu G, Zhou ZH, Baker ML, Jakana J, Cai D, Wei X, Chen S, Gu X, Chiu W (1998) Structure of double-shelled rice dwarf virus. J Virol. 72(11): 8541-8549.

[54] Lwoff A, Horne R, Tournier P (1962) A system of viruses. Cold Spring Harb Symp Quant Biol. 27: 51-55.

[55] Mancini EJ, Clarke M, Gowen BE, Rutten T, Fuller SD (2000) Cryo-electron microscopy reveals the functional organization of an enveloped virus, Semliki Forest virus. Molecular Cell 5(2): 255-266.

[56] Mavrakis M, Mccarthy AA, Roche S, Blondel D, Ruigrok RWH (2004) Structure and Function of the C-

Terminal Domain of the Polymerase Cofactor of Rabies Virus J.Mol.Biol 343:819.

[57] (7) Mckenna R, Xia D, Willingmann P, Ilag LL, Krishnaswamy S, Rossmann MG, Olson NH, Baker TS, Incardona NL (1992) Atomic structure of single-stranded DNA bacteriophage phi X174 and its functional implications. Nature 355:137-143.

[58] Milnor J (1963) Morse Theory, Princeton University Press, Princeton, New Jersey.

[59] Naitow H, Tang J, Canady M, Wickner RB, Johnson JE (2002) L-A virus at 3.4 A resolution reveals particle architecture and mRNA decapping mechanism. Nat.Struct.Biol. 9:725-728.

[60] Nakagawa A, Miyazaki N, Taka J, Naitow H, Ogawa A, Fujimoto Z, Mizuno H, Higashi T, Watanabe Y, Omura T, Cheng RH, Tsukihara T (2003) The Atomic Structure of

Rice dwarf Virus Reveals the Self-Assembly Mechanism of Component Proteins STRUCTURE 11:1227-1238.

[61] Namba K, Pattanayek R, Stubbs G (1989) Visualization of Protein-Nucleic Acid Interactions in a virus: Refined structure of intact Tobacco Mosaic Virus at 2.9 A resolution by X-ray fiber diffraction Mol. Bio. 208:307-325.

[62] Natarajan P, Johnson JE (1998) Molecular packing in virus crystals: geometry, chemistry, and biology. J Struct Biol. 121(3): 295-305.

[63] Navaza J, Lepault J, Rey FA, Alvarez-Rua C, Borge J (2002) On the fitting of model electron densities into EM reconstructions: a reciprocal-space formulation. Acta Crystallogr D Biol Crystallogr. 58(Pt 10 Pt 2): 1820-1825. Epub 2002 Sep 28.

[64] Ng KK, Cherney MM, Vazquez AL, Machin A, Alonso JM, Parra F, James MN (2002) Crystal structures of

active and inactive conformations of a caliciviral RNA-dependent RNA polymerase. J.Biol.Chem. 277:1381-1387.

[65] Padilla JE, Colovos C, Yeates TO (2001) Nanohedra: using symmetry to design self assembling protein cages, layers, crystals, and filaments. Proc Natl Acad Sci U S A 98(5): 2217-2221. Epub 2001 Feb 20.

[66] Poranen MM, Tuma R (2004) Self-assembly of double-stranded RNA bacteriophages. Virus Res. 101(1): 93-100.

[67] Prasad BV, Hardy ME, Dokland T, Bella J, Rossmann MG, Estes MK (1999) X-ray crystallographic structure of the Norwalk virus capsid. Science 286:287-290.

[68] Qu F, Morris TJ (1997) Encapsidation of turnip crinkle virus is defined by a specific packaging signal and RNA size. J Virol. 71(2): 1428-1435.

[69] Qu F, Ren T, Morris TJ (2003) The coat protein of turnip crinkle virus suppresses posttranscriptional gene silencing at an early initiation step. J Virol. 77(1): 511-522.

[70] Rapaport DC, Johnson JE, Skolnick J (1999) Supramolecular self-assembly: Molecular dynamics modeling of polyhedral shell formation. Computer Physics Communications 121: 231-235.

[71] Reddy VS, Natarajan P, Okerberg B, Li K, Damodaran KV, Morton RT, Brooks CL 3rd, Johnson JE (2001) Virus Particle Explorer (VIPER), a website for virus capsid structures and their computational analyses. J Virol. 75(24):11943-11947.

[72] Reinisch KM, Nibert ML, Harrison SC (2000) Structure of the reovirus core at 3.6 A resolution. Nature 404:960-967.

[73] Roberts MM, White JL, Grutter MG, Burnett RM (1986) Three-dimensional structure of the adenovirus major coat protein hexon. Science 232(4754): 1148-1151.

[74] Rossmann MG (2000) Fitting atomic models into electron-microscopy maps. Acta Crystallogr D Biol Crystallogr. 56(Pt 10): 1341-1349.

[75] Rossmann MG, Blow DM (1962) The detection of sub-units within the crystallographic asymmetric unit. Acta Cryst. 15: 24-31.

[76] Schwartz R, Prevelige PE Jr, Berger B (1998) Local rules modeling of nucleation-limited virus capsid assembly. MIT-LCS-TM 584.

[77] Simpson AA, Chipman PR, Baker TS, Tijssen P, Rossmann MG (1998) The structure of an insect parvovirus (Galleria mellonella densovirus) at 3.7 A resolution. Structure 6:1355-1367.

[78] Smith TJ, Chase E, Schmidt T, Perry KL (2000) The structure of cucumber mosaic virus and comparison to cowpea chlorotic mottle virus. J.Virol. 74:7578-7586.

[79] Speir JA, Munshi S, Wang G, Baker TS, Johnson JE (1995) Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. Structure 3:63-77.

[80] Stehle T, Gamblin SJ, Yan Y, Harrison SC (1996) The structure of simian virus 40 refined at 3.1 A resolution. Structure 4(2): 165-182.

[81] Stehle T, Harrison SC (1997) High-resolution structure of a polyomavirus VP1-oligosaccharide complex: implications for assembly and receptor binding. EMBO J. 16(16): 5139-5148.

[82] Tao Y, Farsetta DL, Nibert ML, Harrison SC (2002) RNA Synthesis in a Cage-Structural Studies of Reovirus Polymerase lambda3 Cell (Cambridge,Mass.) 111:733-745.

[83] Tendulkar AV, Joshi AA, Sohoni MA, Wangikar PP (2004) Clustering of protein structural fragments reveals modular building block approach of nature. J Mol Biol. 338(3): 611-629.

[84] Trus BL, Heymann JB, Nealon K, Cheng N, Newcomb WW, Brown JC, Kedes DH, Steven AC (2001) Capsid structure of Kaposi's sarcoma-associated herpesvirus, a gammaherpesvirus, compared to those of an alphaherpesvirus, herpes simplex virus type 1, and a betaherpesvirus, cytomegalovirus. J Virol. 75(6): 2879-2890.

[85] van Heel M, Gowen B, Matadeen R, Orlova EV, Finn R, Pape T, Cohen D, Stark H, Schmidt R, Schatz M, Patwardhan A (2000) Single-particle electron cryo-

microscopy: towards atomic resolution. Quarterly Reviews Biophysics 33(4): 307-369.

[86] Wang J, Simon AE (2000) 3'-End stem-loops of the subviral RNAs associated with turnip crinkle virus are involved in symptom modulation and coat protein binding. J Virol. 74(14): 6528-6537.

[87] Wynne SA, Crowther RA, Leslie AG (1999) The crystal structure of the human hepatitis B virus capsid. Mol.Cell 3:771-780.

[88] Weissenhorn W, Carfi A, Lee KH, Skehel JJ, Wiley DC (1998) Crystal structure of the EBOLA membrane-fusion subunit, GP2, from the envelope glycoprotein ectodomain. Molecular Cell. 2: 605-616.

[89] White DO, Fenner (1994) FJ *Medical Virology – 4th Edition*

[90] Wikoff WR, Johnson JE (1999) Virus assembly: Imaging a molecular machine. Curr Biol. 9(8): R296-300.

[91] Wodak SJ, Mendez R (2004) Prediction of protein-protein interactions: the CAPRI experiment, its evaluation and implications. Curr Opin Struct Biol. 14(2): 242-249.

[92] Wu H, Rossmann MG (1993) The canine parvovirus empty capsid structure. J.Mol.Biol. 233:231-244.

[93] Yan X, Olson NH, van Etten JL, Begoin M, Rossman MG, Baker TS (2000) Structure and Assembly of large lipid-containing dsDNA viruses. Nature Struct. Bio. 7:101-103.

[94] Yu Z, Bajaj C (2005) Automatic Ultra-structure Segmentation of Reconstructed Cryo-EM Maps of Icosahedral Viruses. IEEE Transactions on Image Processing: Special Issue on Molecular and Cellular Bioimaging, September 2005 (in press).

[95] Zhou ZH, Baker ML, Jiang W, Dougherty M, Jakana J, Spencer JV, Trus BL, Booy FP, Steven AC, Newcomb WW, Brown JC (1997) Structure of the herpes simplex virus capsid: peptide A862-H880 of the major capsid protein is displayed on the rim of the capsomer protrusions. Virology 228(2): 229-235.

[96] Zhang X, Walker SB, Chipman PR, Nibert ML, Baker TS (2003) Reovirus polymerase λ3 localized by cryo-electron microscopy of virions at a resolution of 7.6 A. Nature Struct. Bio. 10:1011-1018.

[97] Zhou ZH, Chen DH, Jakana J, Rixon FJ, Chiu W (1999) Visualization of tegument-capsid interactions and DNA in intact herpes simplex virus type 1 virions. J Virol. 73(4): 3210-3218.

[98] ICTV Database

http://www.ncbi.nlm.nih.gov/ICTVdb/Ictv/index.htm

Appendix I: LIST OF FIGURES

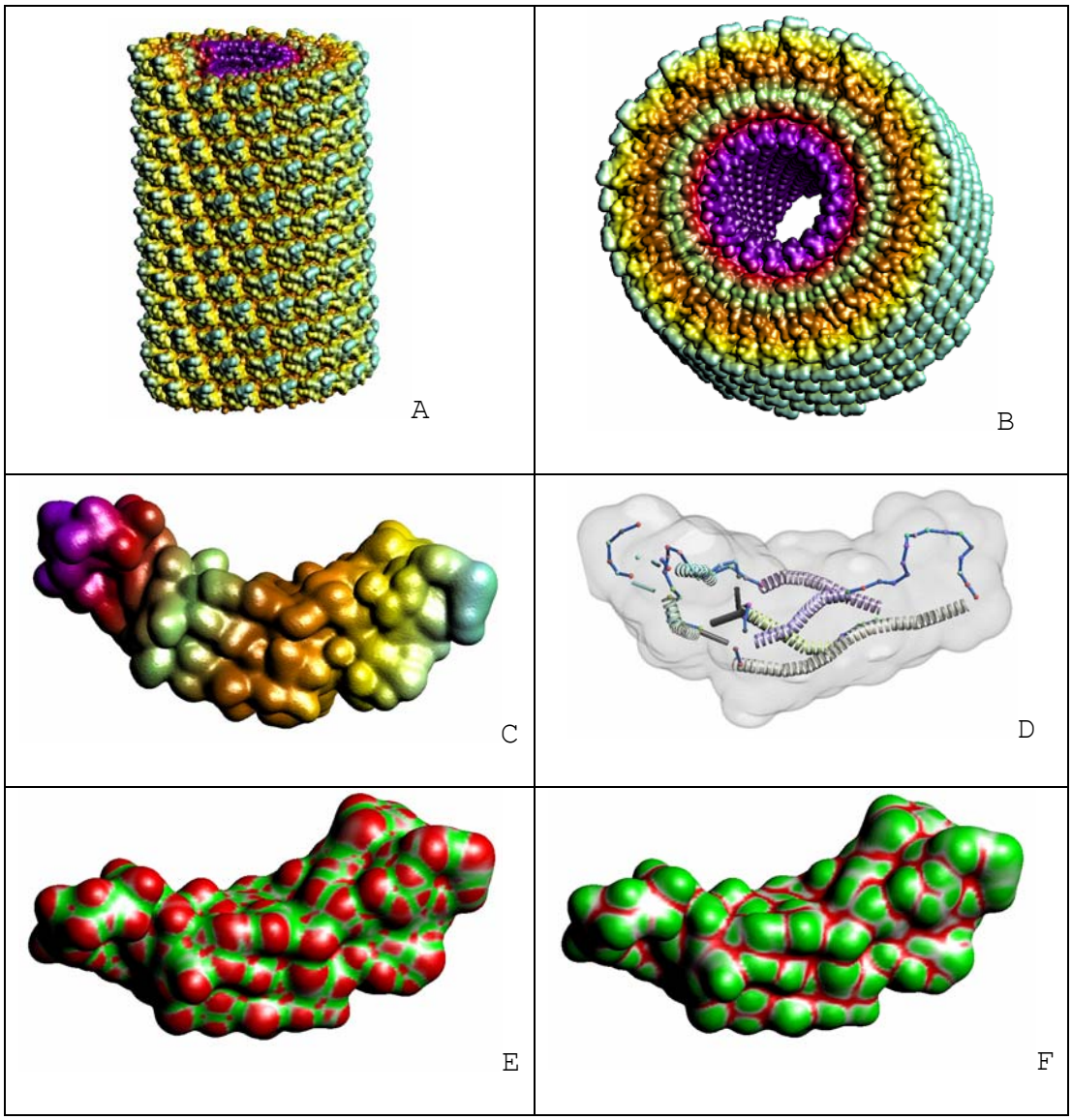
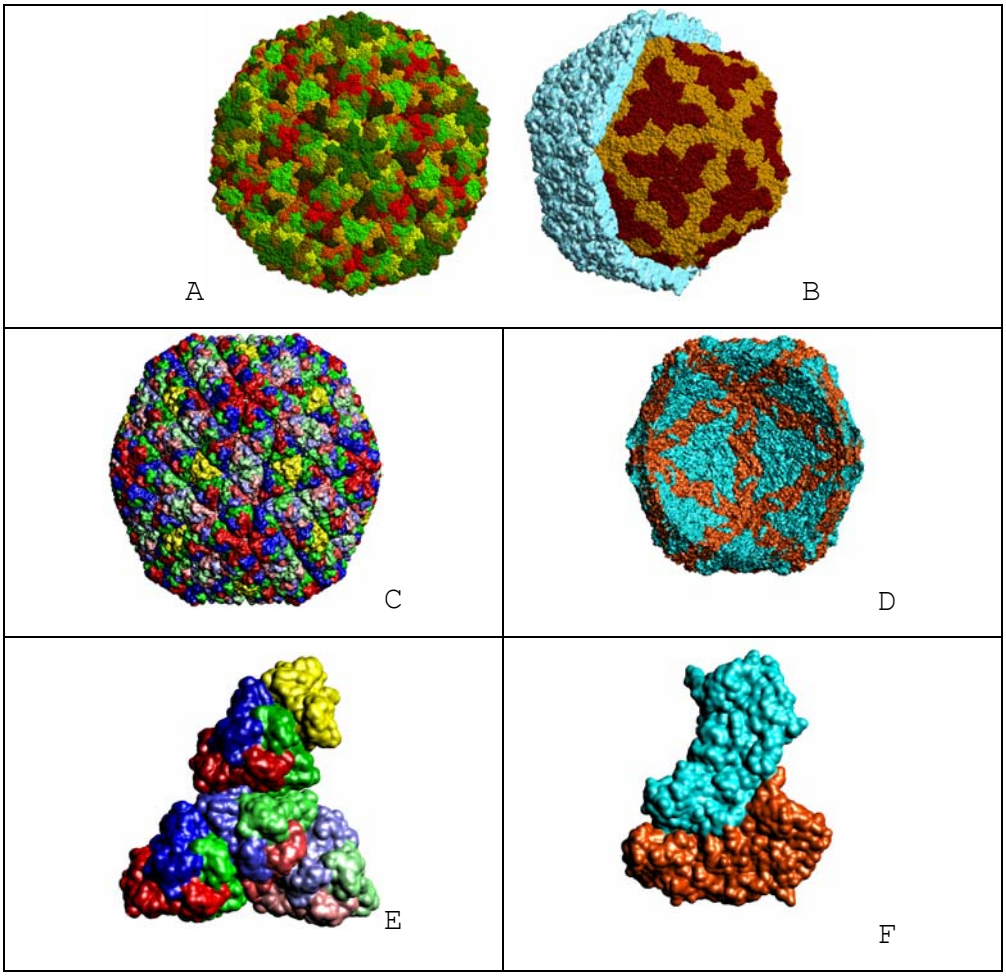


Fig. 2.1 Organization of the Tobacco Mosaic Virus (1EI7) with its helical nucleo-capsid shown in (A) and (B). The asymmetric protomeric structure unit is visualized in (C) as an implicit solvation molecular

surface colored by distance from the helix symmetry axis (D) with a transparent molecular surface and the protein backbone showing helix secondary structures. (E) molecular surface of protomer with the mean curvature function (F) Gaussian curvature function.



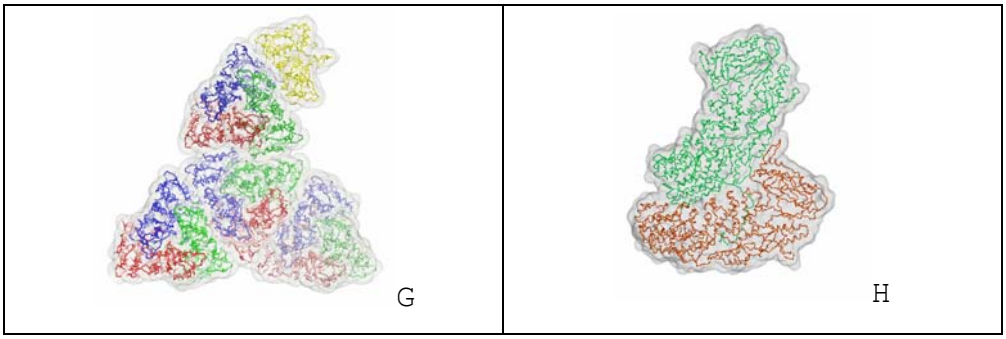


Fig. 2.2. Organization of Rice Dwarf Virus (1UF2) with Icosahedral capsid shells. (A) 2D texture based visualization of the outer capsid shell showing a single sphere per atom, and colored by proteins (B) the outer capsid shell shown as a smooth analytic molecular surface while the inner capsid surface is displayed using 2D texture maps of a union of spheres and colored (C) shows the outer capsid (D) displays the inner capsid (E) shows the icosahedral asymmetric structure unit of the outer unit (F) displays the icosahedral asymmetric structure unit of the inner unit (G) shows the protein backbone of the structure unit shown in (E) and (H) shows the protein backbone of the structure unit show in (F).

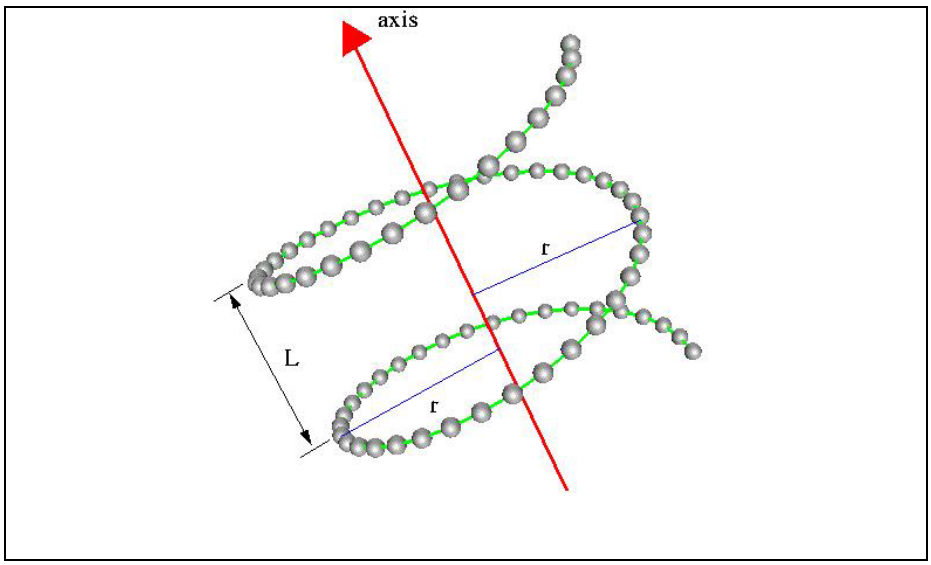


Fig. 2.3. Helical Symmetry Axis

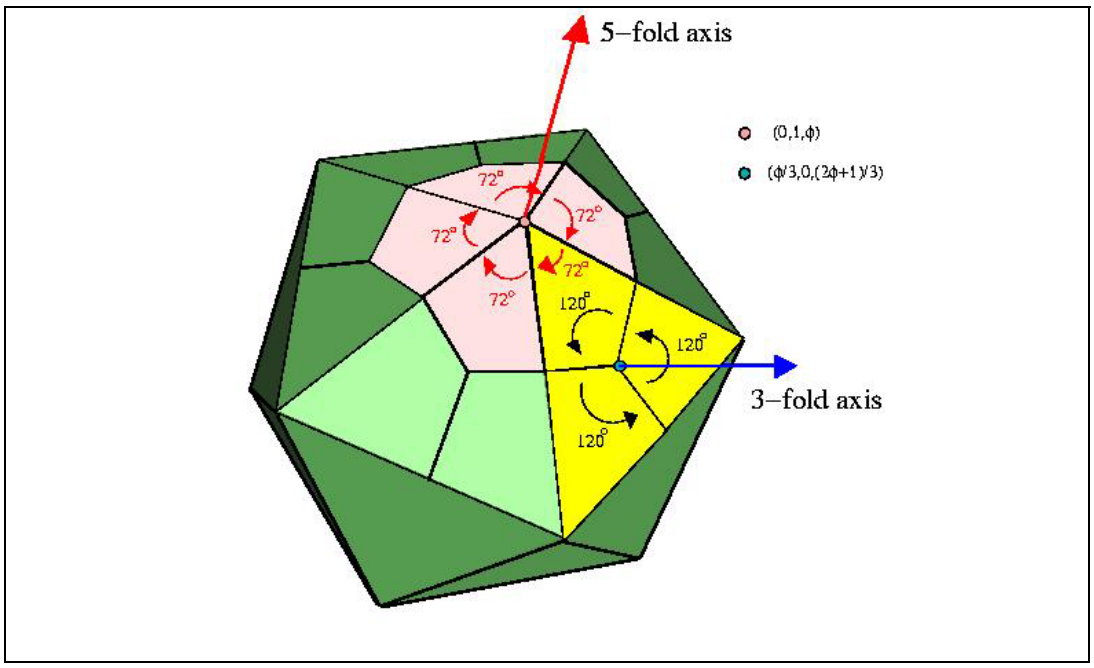


Fig. 2.4 Icosahedral Transformations showing 5-fold and 3-fold Symmetry Axis

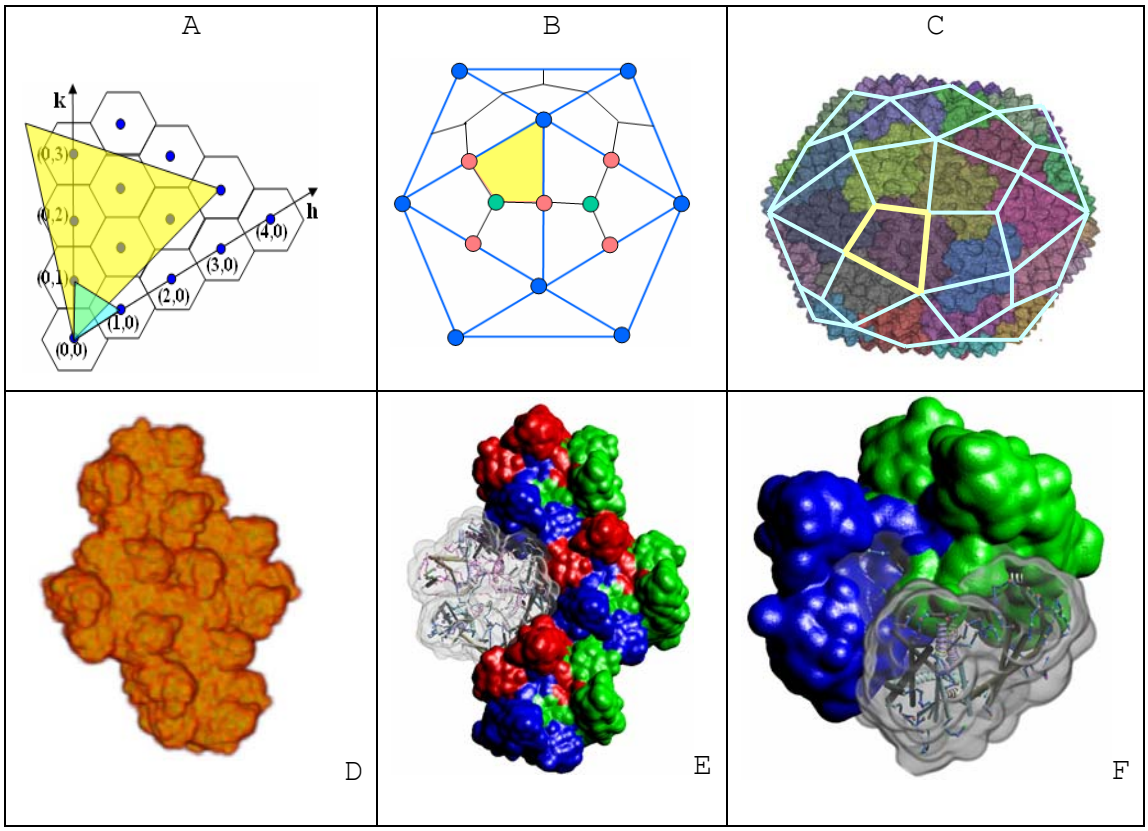


Fig. 2.5 Architecture of Icosahedral Viruses: (A) Caspar-Klug Triangulation Number (T) via a hexagonal lattice. Green triangle has T =1 while yellow represents T = 21 (B) shows the asymmetric unit of an icosahedron, (C) asymmetric structure units of the capsid shall (D) a single asymmetric structure unit (E) asymmetric unit colored by protein as well as showing protein backbone. (F) a capsomere consisting of three proteins. (VIRUS PDB: 1GW8)

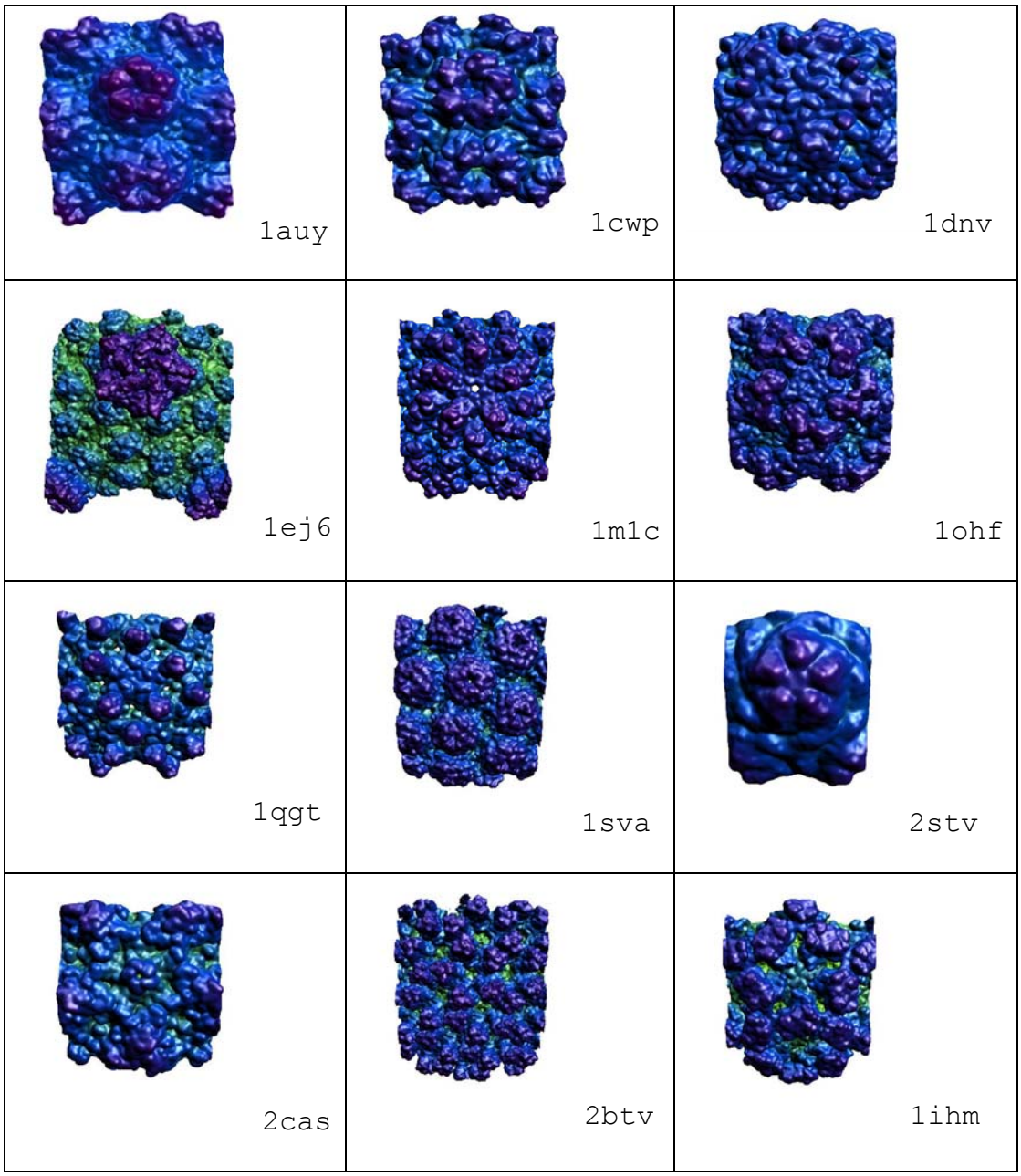


Figure 3.1: Portions of Capsid Shells of Icosahedral Viruses visualized using molecular surface visualization.

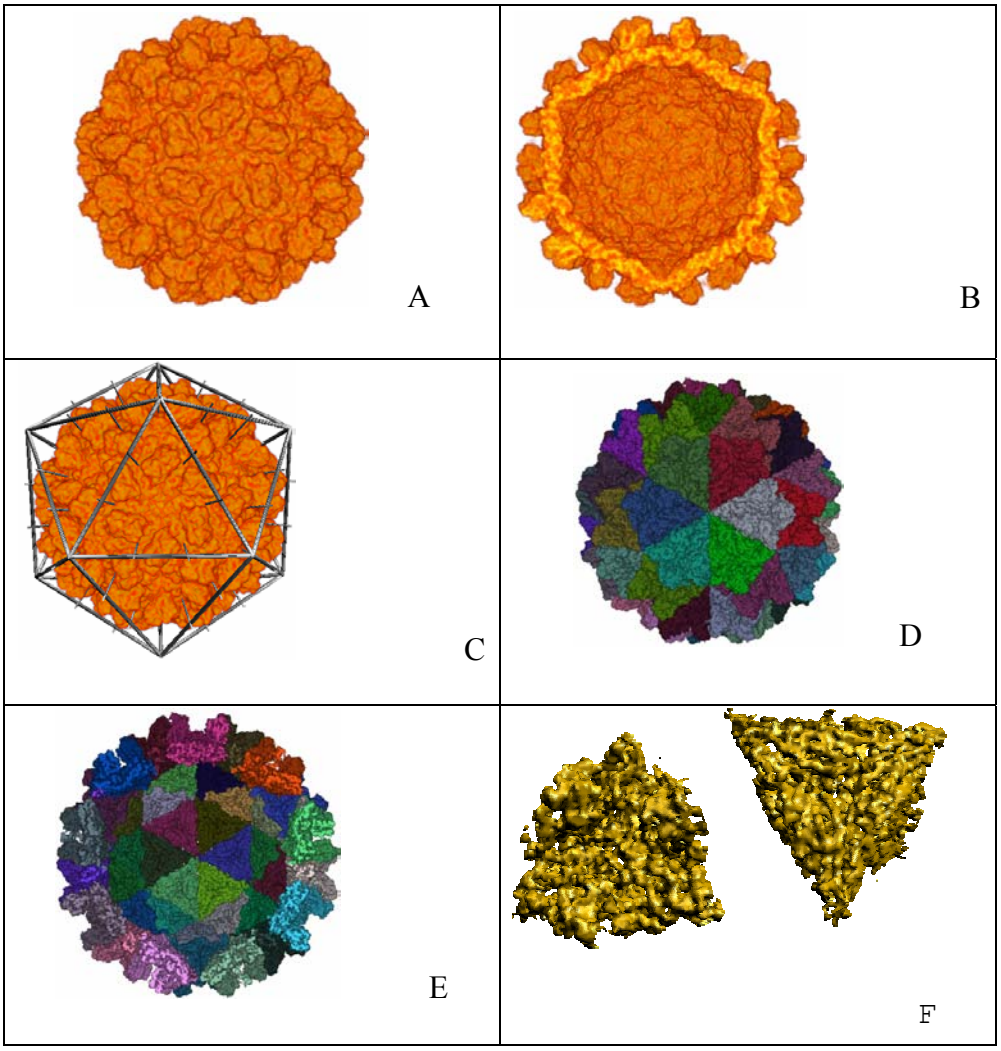


Fig. 3.2 (PDB-ID = 1AUY. Size: 256^3. Resolution: ~4Å). (A) Blurred map (outside view). (B) Blurred map (inside view). (C) Symmetry detected, including global and local 3-fold symmetry axes. (D) Segmented trimers (outside view), with randomly assigned colors. (E) Segmented trimers (inside view). (F) One of the segmented trimers (left-bottom: outside view; right-top: inside view).

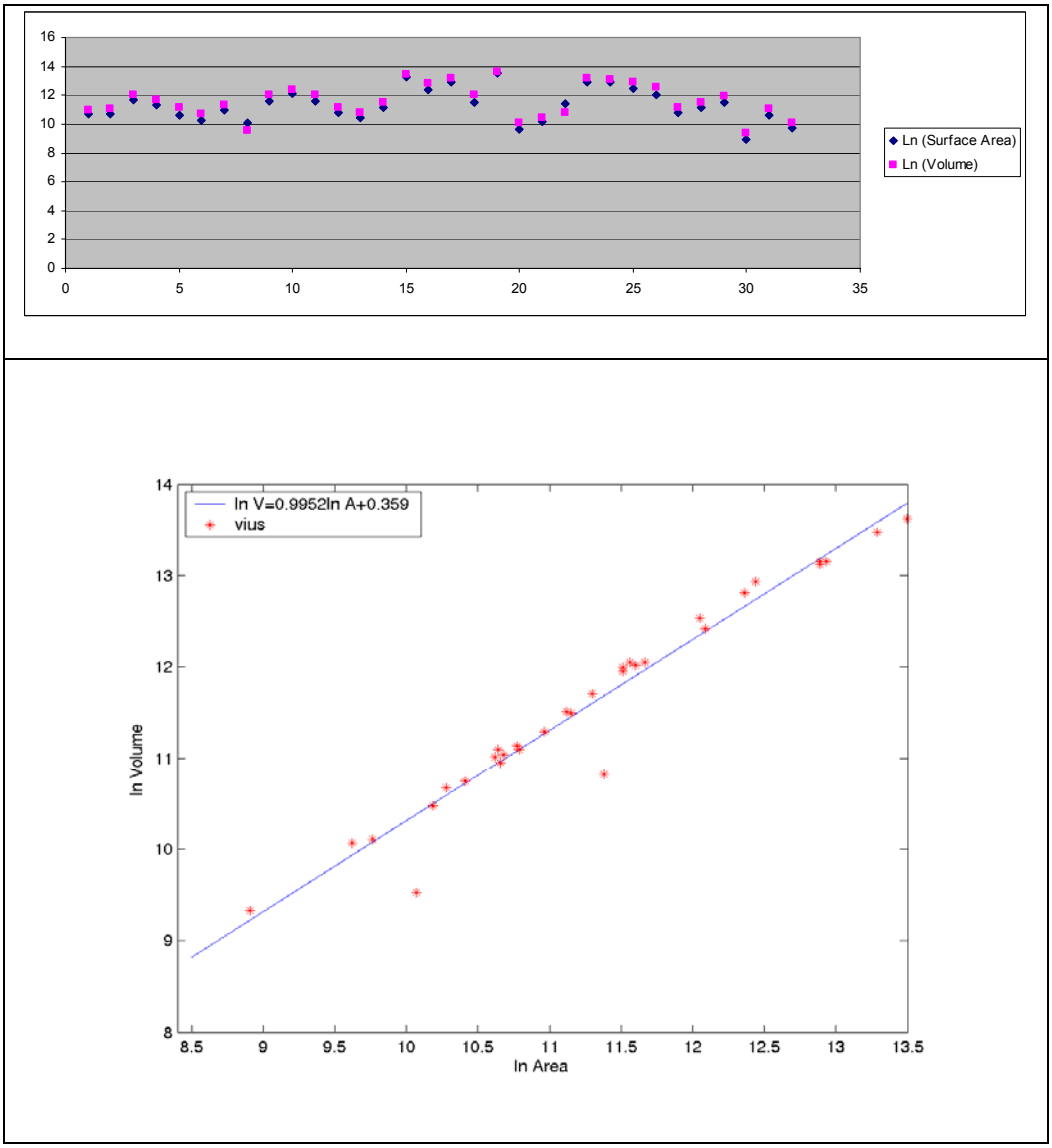


Fig. 4.1 Area, Volume Relationship for Icosahedral Viruses given in Table 1.

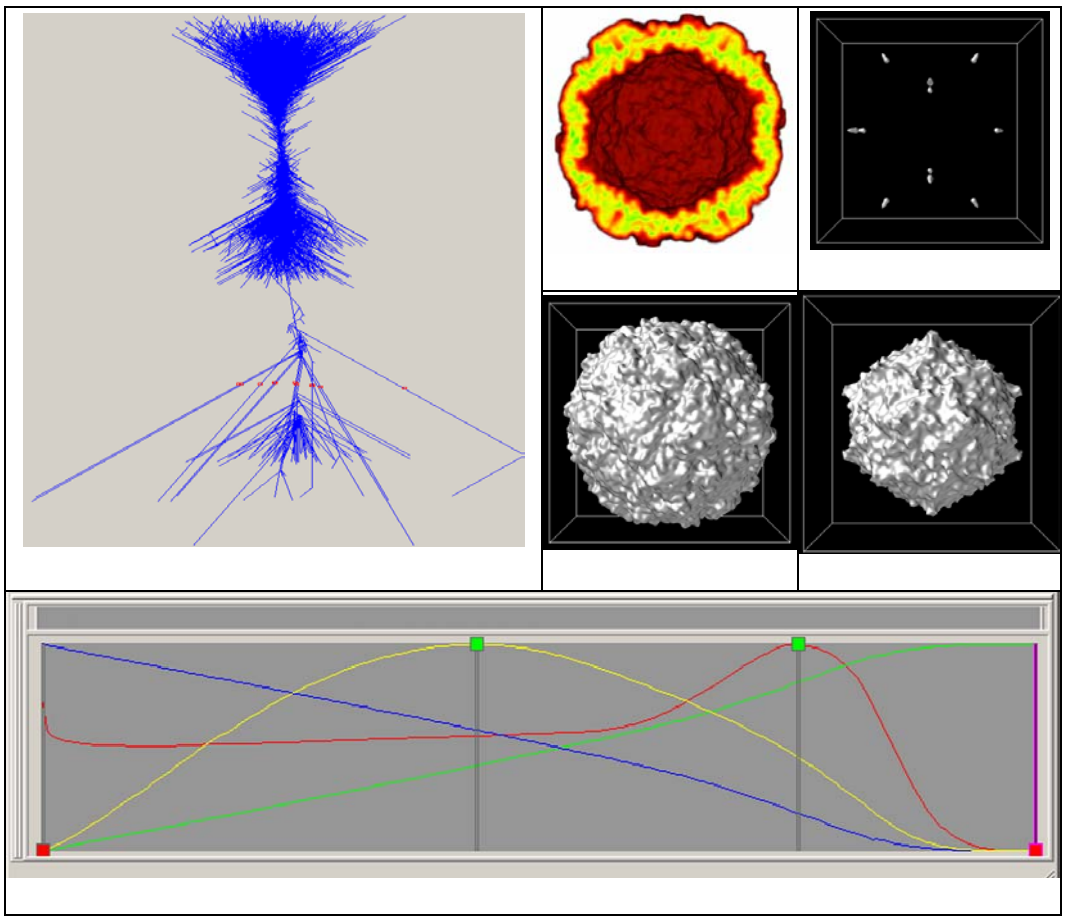


Figure 4.2: The contour tree (upper left) and the contour spectrum (bottom) for the Human Rhinovirus serotype 2 (pdbid: 1 FPN).The red color in the spectrum curve is the graph of molecular surface area, while the blue and green curves are the excluded and enclosed volume.

Appendix: LIST OF TABLES

Table 1 : Helical and Icosahedral Viruses: (1) Name and structure reference of virus given in square brackets (2) Family nomencleature from the ICTV database (3) Host types are P for Plant, V for Vertebrate, I for Invertebrate, F for Fungi (4) Virus Nucleic Acid (NA) type is single stranded RNA (sR) or DNA (sD), double stranded RNA (dR) or DNA (dD) and linear (L) or circular (C) (5) Capsid symmetry is Helix (He) or Icosahedral (Ic) with the triangulation number of each capsid shell in parenthesis (6) The number of capsid shells and whether enveloped (E) or not (n) (7) The acquisition modality X-ray, feature resolution and PDB id in parenthesis.

Name	Family	Host	NA	Capisd Sym. (# T)	#Shell (E?)	Modality (res in A) (pdbid)
Tobacco mosaic [11, 61]	Tobamoviridae	P	sR (L)	He	1(n)	X(2.45) (1ei7)
Ebola [88]	Filoviridae	V	sR (L)	He	1(E)	X(3) (1ebo)
Vaccinia [23]	Poxviridae	V	dD (L)	He	1(E)	X(1.8) (1luz)
Rabies [56]	Rhabdoviridae	V	sR(L)	He	1(E)	X(1.5) (1vyi)
Satellite tobacco necrosis [52]	Tombusviridae	P	sR (L)	Ic (1)	1(n)	X(2.5) (2stv)
L-A (Saccharomyces cerevisiae) [59]	Totiviridae	F	dR (L)	Ic (1)	1(n)	X(3.6) (1m1c)
Canine parvovirus-Fab complex [92]	Parvoviridae	V	sD (L)	Ic (1)	1(n)	X(3.3) (2cas)
T1L reovirus core [72]	Reoviridae	V	dR (L)	Ic (1,1)	2(n)	X(3.6) (1ej6)
T3D reovirus core [82]	Reoviridae	V	dR (L)	Ic (1,1)	2(n)	X(2.5) (1muk)

P4 (Ustilago maydis) [49]	Totiviridae	P	dR (L)	Ic (1)	1(n)	X(1.8) (1kp6)
Tomato bushy stunt [39]	Tombusviridae	P	sR (L)	Ic (3)	1(n)	X(2.9) (2tbv)
Cowpea Chlorotic Mosaic [79]	Bromoviridae	P	sR (L)	Ic (3)	1(n)	X(3.2) (1cwp)
Cucumber mosaic [78]	Bromoviridae	P	sR (L)	Ic (3)	1(n)	X(3.2) (1f15)
Norwalk [67]	Caliciviridae	V	sR (L)	Ic (3)	1(n)	X(3.4) (1ihm)
Rabbit hemorrhagic disease VLP-MAb-E3 complex [64]	Caliciviridae	V	sR (L)	Ic (3)	1(n)	X(2.5) (1khv)
Galleria mellonella denso[77]	Parvoviridae	I	sD (L)	Ic (1)	1(n)	X(3.7) (1dnv)
Semiliki Forest [55]	Togaviridae	I,V	sR	Ic (4,1)	2(E)	C(9) (1dyl)
Polyoma [19]	Papovaviridae	V	dD (C)	Ic (7D)	1(n)	X(2.2) (1cn3)
Simian [80]	Papovaviridae	V	dD (C)	Ic (7D)	1(n)	X(3.1) (1sva)
Papillomavirus Initiation Complex [29]	Papovaviridae	V	dD (C)	Ic (7D)	1(n)	X(3.2) (1ksx)
Blue Tongue [33]	Reoviridae	V	dR (L)	Ic (1,13L)	2(n)	X(3.5) (2btv)
Rice dwarf [60]	Reoviridae	P	dR (L)	Ic (1,13L)	2(n)	X(3.5) (1uf2)
T1L reovirus virion [51]	Reoviridae	V	dR (L)	Ic (1,13L)	2(n)	X(2.8) (1jmu)
Simian rotavirus (SA11-4F) TLP [35]	Reoviridae	V	dR (L)	Ic (1,13L)	2(n)	X(2.38) (1lj2)
Rhesus rotavirus [27]	Reoviridae	V	dR (L)	Ic (1,13L)	2(n)	X(1.4) (1kqr)
Reovirus [96]	Reoviridae	V	dR (L)	Ic(1,13L)	2(n)	C(7.6)
Nudaurelia capensis w [38]	Tetraviridae	I	sR (L)	Ic (4)	1(n)	X(2.8) (1ohf)
Herpes Simplex [14]	Herpesviridae	V	dD (L)	Ic (7L)	1(E)	X(2.65) (1jma)
Chilo Iridescent [93]	Iridoviridae	I	dD (C)	Ic(147)	1(E)	C(13)
Paramecium Bursaria Chlorella [93]	Phycodnaviridae	P	dD (L)	Ic(169D)	1(E)	C(8)
HepBc (human liver) (nHBc) [87]	Hepadnaviridae	V	dD (C)	Ic(4)	1(E)	X(3.3) (1qgt)

Table 2 : Icosahedral Viruses: Capsid Area, Volume and regression relationship

Virus	Surf Area	Vol.	Ln (Surf Area)	Ln (Vol.)
Satellite tobacco necrosis	17401.22	24419.51	9.7643	10.1031
L-A (Saccharomyces cerevisiae)	99643.60	155223.15	11.5094	11.9526
Canine parvovirus-Fab complex	48482.09	66028.46	10.7889	11.0978
T1L reovirus core	412654.67	517093.80	12.9304	13.1560
T3D reovirus core	99627.14	161424.33	11.5092	11.9918
P4 (Ustilago maydis)	7362.92	11269.59	8.9042	9.3299
Tomato Bushy Stunt	69600.33	98169.33	11.1505	11.4944
Cowpea Chlorotic Mosaic	42523.74	56607.48	10.6578	10.9439
Cucumber Mosaic	43317.17	61885.43	10.6763	11.0330
Norwalk	116674.31	170940.17	11.6671	12.0491
Rabbit hemorrhagic disease VLP-MAb-E3 complex	80585.54	121611.94	11.2971	11.7086
Galleria mellonella densovirus	33251.61	46216.49	10.4119	10.7411
Human Rhino	67337.70	99964.30	11.1175	11.5126
HepBc (human liver) (nHBc)	41669.23	65963.21	10.6375	11.0969
Nudaurelia capensis w	170957.88	278225.27	12.0492	12.5362
Semiliki Forest	47586.60	68392.18	10.7703	11.1330
Polyoma	104897.53	171532.31	11.5607	12.0525
Simian	177557.44	246603.02	12.0870	12.4155
Herpes Simplex Virus Glyco-Protein	29035.25	43098.51	10.2763	10.6712
Blue Tongue	590265.25	711692.59	13.2883	13.4754
Rice Dwarf	727228.58	820906.09	13.4970	13.6182
T1L reovirus virion	412654.67	517093.80	12.3640	12.8133
Simian rotavirus (SA11-4F) TLP	26451.69	35311.17	10.1831	10.4720
Rhesus rotavirus	15093.03	23469.18	9.6220	10.0634