

THE UNIVERSITY of TEXAS

SCHOOL OF HEALTH INFORMATION SCIENCES AT HOUSTON

Virus Capsids and Icosahedral Reconstruction

For students of HI 6001-125 "Computational Structural Biology"

Willy Wriggers, Ph.D. http://biomachina.org/courses/structures/10.html

Overview and Biological Relevance

Virus Structure

- crystallization of a virus was first reported in the 1930s.
- first atomic resolution structure of a virus was 1978, tomato bushy stunt virus (Stephen Harrison):



Example

Hepatitis A virus.



Role of Virus Capsids

- Function of the outer shell (**capsid**) of a virus particle is to protect the fragile nucleic acid genome from:
- Physical damage Shearing by mechanical forces.
- Chemical damage- UV irradiation (from sunlight) leading to chemical modification. Enzymatic damage - Nucleases derived from dead or leaky cells or deliberately secreted by vertebrates as defence against infection.
- Protein subunits in a virus capsid are **multiply redundant**, i.e. present in many copies per particle. Damage to one subunit may render that subunit non-functional, but does not destroy the infectivity of the whole particle.

Membrane Envelopes



Membrane envelopes acquired from a cellular structure during release. Membranes are modified by proteins. Matrix proteins are found inside the envelope. Glycoproteins traverse the envelope.

Infection

- The outer surface of the virus is responsible for **recognition of the host cell**. Initially, this takes the form of binding of a specific **virus-attachment protein** to a **cellular receptor molecule**. The capsid also has a role to play in initiating infection by delivering the genome from its protective shell in a form in which it can interact with the host cell.
- To form an infectious particle, a virus must overcome two fundamental problems:
- 1. assemble the particle utilizing only the information available from the components which make up the particle itself (capsid + genome).
- 2. Form regular geometric shapes, even though the proteins from which they are made are irregularly shaped.

Tobacco mosaic virus (helical).



© http://www-micro.msb.le.ac.uk/109/109Structure2.ppt

HIV (complex globular, enveloped)



Vesicular stomatitis virus: bullet shaped



© http://www-micro.msb.le.ac.uk/109/109Structure2.ppt

Cytoplasmic polyhedrosis virus (icosahedral)



Bacteriophage T4 (icosahedral and helical)



Bacteriophage T4

Head consists of an <u>icosahedral</u> shell attached via a collar to a <u>helical</u> tail. At the end of the tail is a plate which functions in attachment to the bacterial host.In addition thin protein fibres are attached to the plate, again involved in binding to host.





Ebola (irregular)



© http://www-micro.msb.le.ac.uk/109/109Structure2.ppt

Icosahedral Symmetry

The Five Platonic Solids

From **equilateral triangles** you can make: with 3 faces at each vertex, a **tetrahedron**

with 4 faces at each vertex, an octahedron

with 5 faces at each vertex, an icosahedron

From squares you can make: with 3 faces at each vertex, a cube

From pentagons you can make: with 3 faces at each vertex, a dodecahedron



Virus Structure: Icosahedra



• A common way of building a virus capsid is to arrange protein subunits in the form of a hollow quasi-spherical structure, enclosing the genome within.

Crick &Watson (1956), after seeing electron microcraphs, were the first to suggest that virus capsids are composed of numerous identical protein sub-units arranged either in helical or icosahedral symmetry.

In order to construct a capsid from repeated subunits, a virus must 'know the rules' which dictate how these are arranged. For an icosahedron, the rules are based on 2-3-5 rotational symmetry.

The Icosahedron



No of vertices : 12 (5-fold symmetry)

© Timothy S. Baker, UCSD

The Icosahedron



No of faces : 20 (3-fold symmetry)

© Timothy S. Baker, UCSD

The Icosahedron



No of edges : 30 (2-fold symmetry)

© Timothy S. Baker, UCSD





30 dimers

© Eisenberg and Crothers



20 trimers

© Eisenberg and Crothers



12 pentamers

Virus Structure: Icosahedra

- 20 equilateral triangles arranged into a sphere.
- bacteriophage ØX174. 60 identical subunits form a capsid. 3 protein subunits per triangular face (T=1). This is the simplest case; most viruses have more subunits per face (higher T number).





© http://mmtsb.scripps.edu/viper/

T=3 Triangulation



T=4 Triangulation



T=7 Triangulation



Example: Norwalk Virus Structure







3D Reconstruction



2D

3D

Why CryoEM?

- Well suited for large macromolecules
- Resolution limit near 5Å
- Sample is frozen in vitreous ice and imaged at liquid nitrogen temperatures
- Imaging thousands of individual particle randomly orientated on a thin substrate
- Computer reconstructions to generate 3D structure

CryoEM

Sample : ~2-3 µl at 1-5 mg/ml Specimen support: holey carbon film (1-2 µm) Microscope: 200-300 keV with FEG Defocus range: 1-3 µm underfocus Dose: 10-20 e⁻/Å² Film: SO-163 (12 min, full strength) Micrographs: 25-100 Particles: 10³-10⁴ Target resolution: 12 - 6 Å



Basic Assumptions

Specimen consists of stable particles with 'identical' structures (else averaging is invalid)

Programs test for and *assume* presence of icosahedral (532) symmetry

T (triangulation) symmetry is not incorporated into or enforced by the 3D reconstruction algorithms

Hence, T symmetry emerges as a result of a properly performed 3D reconstruction analysis
Overview of Reconstruction Scheme









Digitize Micrograph





Extracted



Masked



Floated



Apodized



Floated



Square mask; unfloated



Circular mask; unfloated



Circular mask; floated



Circular mask; floated & apodized





Pre-Process Images

Remove blemish, Remove Gradient Normalize means/variances, Apodize Determine CTF parameters Create Initial Parameter Files





Gradient removed

Pre-Process images Remove blemish, Remove Gradient Normalize means/variances. Apodize Determine CTF parameters Create Initial Parameter Files



Gradient not removed





Pre-Process Images Remove blemish, Remove Gradient Normalize means/variances, Apodize Determine CTF parameters Create Initial Parameter Files



Pre-Process Images Remove blemish, Remove Gradient Normalize means/variances, Apodize Determine CTF parameters Create Initial Parameter Files



Pre-Process Images Remove blemish, Remove Gradient Normalize means/variances, Apodize Determine CTF parameters Create Initial Parameter Files



Extracted







Masked



Apodized

Pre-Process Images Remove blemish, Remove Gradient Normalize means/variances, Apodize Determine CTF parameters Create Initial Parameter Files



Pre-Process Images Remove blemish, Remove Gradient Normalize means/variances, Apodize Determine CTF parameters Create Initial Parameter Files

	X→× FFT - CTF Estimation				- • • • ×
			🔷 Linear	↓ 1024 × 1024	Intensity
			A Log	♦ 512 × 512	Recalc FFT
			A Man of finance	♦ 258 × 258	
			V ana.rrat	↓128 × 128	Lock Scrolls
			Min 🖂	l Nu	7.61
		1. A.	Max 🖾	1 N ¹	8.52
			TFac 🗐		0
			Range Min:	3.16 Max:	12.69
			Back Transform	Generate Defa	ult Param File
		P	ointer Detail		
			- 1 C		
	where the second second second second			Defeeur Defin	ananti
		Ove	erlay Intensity:	Derocus Ref I	rement
	and the second second second second		I X		
			Zoom		_
		Disp 1D CA	Contour 1D CircAva	ScStp(um) 14	Mag33019
			······	PxSiz(nm) 0.42	4 Wiener 0.2
	_ Pick CTF Pts Flicker Determine Image CTF			AmpCont 10.07	AngMaj 0.0
	Ang Major 🖾 🧵 🧎	CTF Node Num:		Volts(kV) <u>300</u>	tFac jo
	Focus Major		FFT Average	Wave (nm) 10.00	1 Mode <u>i</u> 1
	Focus Minor	<u> I</u> I	i Ing#	Cs(mm)2.0	ctf 1/ctf
			Average FFTs	FocMaj(um) 👖 0.0	PixelSize i
		0 I	Incoherent Hvg	FocMin(um) 👖 0.0	nanometers
© Timothy S. Baker	UCSD				Taxaataa Taxaataa

Pre-Process Images Remove blemish, Remove Gradient Normalize means/variances, Apodize Determine CTF parameters Create Initial Parameter Files

X→ FFT - CTF Estimation					• • ×
★-* FFT - CTF Estimation		<pre>↓ Linear ↓ Log ↓ Non-Linear Min ↓ Hax ↓ TFac ↓ Range Min: ↓ Back Transform</pre>	 ✓ 1024 × 102 ◆ 512 × 512 ✓ 256 × 256 ✓ 128 × 128 14.71 Generate 	4 Intensity Recalc FFT Lock Scro 14.71 17.42 0 4ax: 25.16	
		Pointer Detail	Defocus R	efinement	11e
Pick CTF Pts Flicker Determine Image CTF] Disp 1D CA	_ Zoom _ Contour <u>1D CircAvg</u>	ScStp(um) ji PxSiz(nm) 0 AmpCont jo	4 Mag .424 Wiener	<u>33019</u> 0.2
Ang Hajor H	CTF Node Num:	FFT Average	Volts(kV) 3 Wave(nm) [0 Cs(nm)]2	00 tFac .001 Mode	jo j1 /ctf
© Timothy S. Baker. UCSD	DI	Average FFTs	FocMaj(um) [0 FocMin(um) [0	. 0 Pi	kelSize i anometers

Pre-Process Images Remove blemish, Remove Gradient Normalize means/variances, Apodize Determine CTF parameters Create Initial Parameter Files

×	-» FFT - CTF Estimation				• • ×
			<pre>verlay Intensity:</pre>	1024 x 1024 512 x 512 255 x 255 255 x 128 128 x 128 14.71 Hat Generate De Defocus Ref	Intensity Recalc FFT Lock Scrolls 14.71 17.42 0 x: 25.16 fault Param File
		Disp 1D CA] Zoom] Contour <u>1D CircAvg</u>	ScStp(um) 14	Mag <u>j</u> 33019
	Pick CTF Pts Flicker Determine Image CTF	CTE Nodo Nur:	٦	AmpCont <u>io.</u> (Volts(kV) <u>š</u> o()7 AngMaj [359.4) tFac [0
	Focus Major 2 1 2 12.16 Focus Minor 2 1 2 13.31		FFT Average	Wave(nm) [0.(Cs(mm) [2.(001 Mode ji
© Timothy S Baker		o I	Average FFTs	FocMaj(um) <u>1</u> 2.1 FocMin(um) <u>1</u> 3.3	6 PixelSize i nanometers

Pre-Process Images Remove blemish, Remove Gradient Normalize means/variances, Apodize Determine CTF parameters Create Initial Parameter Files

X-₩ FFT - CTF Estimation		×→ Automatic Defocus Refinement
	↓ Linear ↑ Log ↓ Ron-Li Min M	a0: 1 b0: 1 a1: 10 b1: 10 a2: 10 0 b2: 10 a3: 10 b3: 10 a4: 10 b4: 10
	Hax H TFac R Range Back Transfo inter Detail	Highest Resolution (Angs): je sfc Estimated Defocus (um): stil Chi squared: J
		Select different plots: Circular Average S S S S Circular Average S S S S S S S S S S S S S
	oom ontour 1D Ci	Ci Change Overlay Intensity to see CTF curves.
		AmpCont Ta az AngHaj Tara 4
Ang Majon		
Focus Minor	FT Average	e Wave (nm) [0.001 Hode]1 (mg# Cs (mm) [2.0 ctf 1/ctf
	Average FFTs Incoherent	FTS FocMaj (um) I2.16 ent Hvg FocMin (um) I3.31

Pre-Process Images Remove blemish, Remove Gradient Normalize means/variances, Apodize Determine CTF parameters Create Initial Parameter Files









Determining Origin and Orientation

Convention of Coordinate System



BPV Projections: Icosahedral ASU


BPV Projections: ¹/₂ Icosahedral ASU



Icosahedral Particle Reconstruction Scheme



How do we determine the (θ , ϕ , ω , x, y) parameters? Two methods:

1. Common lines

New or unknown structure

2. Model-based (template) matching

General features of structure are known or a crude model can be generated

Icosahedral Particle Reconstruction Scheme

Determine Origin and Orientation (θ , ϕ , ω ,x,y)

Common Lines

The 'gospel' according to Tony Crowther (*Phil. Trans. R. Soc. Lond. B.*(1971) 261:221-230)

"[Common lines] arise as follows:"

"An observed section of the transform intersects an identical symmetry-related section in a line, along which the transform must have the same value in both sections"

"The common line lies in the original section."

"However, regarded as lying in the symmetry-related section it must have been generated by the symmetry operation from some other line in the original section."

Common Lines

Electron Images of Virus Particles

Equivalent data in Fourier space





 $\ensuremath{\mathbb{C}}$ Wah Chiu and Hong Zhou

Common Lines

3D Object

Projection Image

Fourier Transform



© Wah Chiu and Hong Zhou

Icosahedral Particle Reconstruction Scheme

↓ Determine Origin and Orientation (θ,φ,ω,x,y)

Common Lines

The 'gospel' continued:

"We therefore have a pair of lines in the original transform plane along which the transform must have identical values"

"A similar pair of lines will be generated by each possible choice of pairs of symmetry operations"

"The angular positions of these lines are dependent on the orientation of the particle."



2D Fourier Transform

Simple example: object with single three-fold axis of symmetry



- ABCD = 2D transform of image from particle **not** viewed along an axis of symmetry
- Let z-direction coincide with **3-fold** axis of symmetry
- 3-fold operation generates **two** additional FT sections (only A'B'C'D' shown)

Both planes have **common values** along the **line** (1,2,3) of their intersection



Adapted from Moody (1990) Fig. 7.68, p.245

Adapted from Moody (1990) Fig. 7.69, p.246



Original Transform Plane





Symmetry-Related Transform Plane



Orientation Determination by Common Lines Ok, that's easy (simple object with single 3-fold axis) What about an object with 532 symmetry?

For a **general view**, icosahedral symmetry generates:

5-folds:
$$\frac{12}{2} \times 2 = 12$$
 pairs
3-folds: $\frac{20}{2} \times 1 = 10$ pairs
2-folds: $\frac{30}{2} \times 1 = \frac{15}{2}$ real lines
37 common lines



[©] Timothy S. Baker, UCSD





What is (θ, ϕ, ω) for this particle?



ω



ω



ω



ω



ω





ω



ω ↓ (80,11,10)



(80,11,15)

ω



(80,11,30)

ω



↓ (80,11,60)

ω



(80,11,90)

ω



(80,11,135)

ω



(80,11,180)

ω



Metric: Identify ω that gives lowest phase residual



Repeat process for all possible (θ,ϕ,ω) combinations



> 250,000 combinations for 1° angular search intervals



Common Lines

The (θ, ϕ, ω) that results in the lowest phase residual is selected as the best estimate for the particle view orientation

The 'common lines' procedure is similarly used to determine the particle phase origin (x, y)

Minimizing Phase Differences Among Common Lines

$$P_{i}(\phi,\theta,\omega,x,y) = \frac{\sum_{j=1}^{N} \sum_{k=1}^{k_{\max}(j)} \sum_{R=R_{\min}}^{R_{\max}} |\psi_{i}(R,x_{i},y_{i},\alpha_{i,j,k}) - \psi_{j}(R,x_{j},y_{j},\alpha_{j,i,k})| \times w(R,\alpha_{i,j,k},\alpha_{j,i,k})}{(R_{\max} - R_{\min}) \sum_{j=1}^{N} k_{\max}(j)}$$

P is the phase residual
Ψ is the phase value in Fourier space
i and j refer to particles or symmetry-related sections
x and y define the phase origin
k refers to the common line
αijk are the angles of the *k*-th common-lines
Rmin and Rmax define the frequency range within which the phase residuals are evaluated
kmax is total number of common lines (here: 37)

N is the number of particles

w defines a weighting function for the Fourier elements at different frequency and orientations

Icosahedral Particle Reconstruction Scheme



Recall: two methods to determine (θ , ϕ , ω , x, y):

1. Common lines

- 2. Model-based (template) matching
 - Many structures now solved this way
 - Same as reference based alignment in singleparticle processing in earlier session...

3D Reconstruction




Goal: weed out 'bad' particle images before computing 3D reconstruction



Goal: combine "good" particle images to compute a 3D density map

Compute 3DR



From Lake (1972), p.174

Compute 3DR

Two dimensional Fourier transform (d) (c) Two dimensional Fourier transform (f) (e) Inverse three dimensional Fourier transform R.J.M. (q) (h)

Overall scheme: $\rho \leftarrow \mathbf{g} \leftarrow \mathbf{G} \leftarrow \mathbf{F}$

Remember? Fourier-Bessel Formalism

$$\rho(r,\varphi,z) = \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} g_n(r,Z) e^{in\varphi} e^{2\pi i z Z} dZ$$

$$g_n(r,Z) = \int_0^\infty G_n(R,Z) J_n(2\pi Rr) 2\pi R dR$$

$$F(R,\Phi,Z) = \sum_{n=-\infty}^{\infty} G_n(R,Z) i^n e^{in\Phi}$$

Steps:

- 1. Compute 2D FFT of each particle image
- 2. Combine all 2D FFTs to build up 3D Fourier-Bessel transform



© Timothy S. Baker, UCSD Crowther, DeRosier and Klug, 1970, p.329

Adapted from Crowther (1971) Fig. 4, p.223

 Φ^{ς}

Icosahedral Particle Reconstruction Scheme f $\rho \leftarrow g \leftarrow G \leftarrow F$

Steps:

- 1. Compute 2D FFT of each particle image
- 2. Combine all 2D FFTs to build up 3D Fourier-Bessel transform
- 3. Compute G_n 's on each annulus $G = (B^{\dagger}B)^{-1}B^{\dagger}F$

solve linear system of equations

- 4. Compute g_n's from G_n's (Fourier-Bessel transform)
- 5. Compute polar density map ($\rho(r, \phi, z)$) from g_n 's
- 6. Convert from polar to Cartesian map ($\rho(r, \phi, z) \rightarrow \rho(x, y, z)$)



Option: correct for CTF effects in particle FFTs before FFTs are merged to form the 3D FFT

Resolution Estimation and Quality Control



Goal: assess resolution of 3D density map to determine what to do next

















Monitor Data Quality



Monitor Data Quality



Monitor Data Quality

Note: quality of 3D density map is not the identical throughout the map

Monitor Data Quality



© Timothy S. Baker, UCSD

Monitor Data Quality



Monitor Data Quality



© Timothy S. Baker, UCSD
















Icosahedral Particle Reconstruction Scheme



References

- Crowther, R.A., Amos, L.A., Finch, J.T., DeRosier, D.J. & Klug, A. Three dimensional reconstructions of spherical viruses by Fourier synthesis from electron micrographs. *Nature* **226**, 421-5 (1970).
- Crowther, R.A. Procedures for three-dimensional reconstruction of spherical viruses by Fourier synthesis from electron micrographs. *Philos Trans R Soc Lond B Biol Sci* **261**, 221-30 (1971).
- Böttcher, B., Wynne, S.A. & Crowther, R.A. Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy. *Nature* **386**, 88-91 (1997).
- Conway, J.F., Cheng, N., Zlotnick, A., Wingfield, P.T., Stahl, S.J. & Steven, A.C. Visualization of a 4-helix bundle in the hepatitis B virus capsid by cryo-electron microscopy. *Nature* **386**, 91-4 (1997).
- Zhou, Z.H., Dougherty, M., Jakana, J., He, J., Rixon, F.J. & Chiu, W. Seeing the herpesvirus capsid at 8.5 Å. *Science* **288**, 877-80 (2000).
- Mancini, E.J., Clarke, M., Gowen, B.E., Rutten, T. & Fuller, S.D. Cryo-electron microscopy reveals the functional organization of an enveloped virus, Semliki Forest virus. *Mol Cell* **5**, 255-66 (2000).
- Baker, T.S. & Cheng, R.H. A model-based approach for determining orientations of biological macromolecules imaged by cryoelectron microscopy. *J Struct Biol* **116**, 120-30 (1996).
- Baker, T.S., Olson, N.H. & Fuller, S.D. Adding the third dimension to virus life cycles: three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. *Microbiol & Mol Biol Rev* **63**, 862-922 (1999).
- Thuman-Commike, P.A. & Chiu, W. Reconstruction principles of icosahedral virus structure determination using electron cryomicroscopy. *Micron* **31**, 687-711 (2000).

Zhou, Z.H. and Chiu, W. (2003). Structural determination of icosahedral viruses by electron cryomicroscopy at subnanometer resolution. *Adv Protein Chem* **64**: 93-130.