Characterization of the Archaeal Thermophile Sulfolobus Turreted Icosahedral Virus Validates an Evolutionary Link among Double-Stranded DNA Viruses from All Domains of Life

Presenter: Justin Bradshaw
Co-Presenter: Alice Mei Lee
Viruses

- Latin, poison

- Non-living (?) obligate parasites

Possible Origins

- Acytota: Possible domain

Classified using ICTV, Baltimore

Orthopoxvirus Cowpox virus
Viral Phylogeny

- Absent from fossil record
- No rRNA
- Topological similarity used to determine common ancestry

DI Stuart et. al. uses a PHYLIP analysis of probabilities of equivalence between pairs of residues for viral proteins responsible for the capsid shell (pretty picture)
- dsDNA bacteriophage
- Icosahedral
- Lipid bilayer
- Pentameric “spikes” at icosahedral facets
Sulfolobus solfataricus P2

- Archaeaum isolated from many volcanic areas (~80°C, pH ~2-3)

- Lithoautotrophic or chemoheterotrophic

- Tetraether lipid monolayer
Sulfolobus Turreted Icosahedral Virus

- 17,663 bp dsDNA genome
- 36 predicted ORFs
- Infects S. Solfataricus P2
STIV as a PRD1-like virus

- Almost identical protein fold structure
- STIV “turrets” similar to PRD1 “spikes”
- Inner electron-dense layers similar to PRD1 lipid layer
- What say proteomics?
SDS-PAGE
Schematic of a standard proteome analysis by 2DE-MS

Gygi and Aebersold Current opinion in chemical biology 2000
Methods

• Virus purification
• SDS-PAGE/IEF + SDS-PAGE/In-gel digestion

• MALDI-TOF MS

• Nanospray LC-MS/MS
• Homology searches for STIV structural proteins
• Protein modeling
• Glycoprotein analysis
• Lipid Analysis
What is a Mass spectrometer?
And what does it do?

• Analytical device determines MW
• Mass-to-charge ratio (m/z)
• Can ID/quantify unknown compounds, structure, chemical properties of molecules.
• Small quantities, high sensitivity.
• Steroids, monitor patients by anesthesiologists, structures of biomolecules, detect dioxins, sequence of biopolymers, forensic analysis, environmental pollutants...
Parts of a Mass Spectrometer

What Is a Mass Spectrometer and What Does It Do?

- ionization source
  - electron ionization
  - fast atom bombardment
  - laser desorption
  - electrospray

- mass filter/analyzer
  - quadrupoles
  - magnetic sector
  - time-of-flight
  - ion trap

- ion detector
  - electron multiplier
  - scintillation counter

- sample introduction
- GC Column
- HPLC column
- solid probe
MALDI-TOF MS
Matrix-Assisted Laser Desorption Ionization Time Of Flight Mass Spectrometry

KE = $1/2 MV^2$
BiflexIII MALDI-TOF MS

http://www.biotech.wisc.edu/ServicesResearch/MassSpec/biflex.html
MALDI-TOF MS/MS
ESI MS
ElectroSpray Ionization Mass spectrometry

- Produces gaseous ionized molecules from a liquid solution
- Done by creating a fine spray of highly charged droplets in the presence of a strong electric field.

http://www.newobjective.com/electrospray/
Nano ESI (LC-MS/MS)

- Droplets <10 μm in diameter
- Flowrate is 20-40 nl/min

Siuzdak Mass spectrometry for biotechnology 1996
Nano ESI (LC-MS/MS)

Figure 1.8 Ion formation from electrospray ionization source.

Siuzdak Mass spectrometry for biotechnology 1996
Nano ESI Needle point

www.sitemaker.umich.edu/.../files/esi_spray.jpg
Agilent XCT-plus ion trap mass spectrometer
Tandem MS (MS/MS)

Figure 3.12. Structural characterization of a mass-selected ion by tandem mass spectrometry. Precursor ion mass analysis determines the m/z of the peptide ion of interest. That ion is mass-selected by the first stage of mass analysis and is activated by collision with a neutral gas molecule to induce a fragmentation reaction. The ionic products of the fragmentation reaction are mass-analyzed in the second stage of mass analysis to produce a product ion spectrum.
Protein

molecular weight determination

Electrospray
MALDI
FAB

Proteolytic or chemical digestion
followed by LC separation or LC-MS.

MS of LC fractions

tandem MS on individual ions

figure 4.22 Protein characterization by mass spectrometry.
Structural prediction used to identify proteins

<table>
<thead>
<tr>
<th>Identified protein</th>
<th>Mass (Da)</th>
<th>pI</th>
<th>Band no.</th>
<th>2D spot no.</th>
<th>In solution</th>
<th>Predicted structure or function</th>
</tr>
</thead>
<tbody>
<tr>
<td>STIV C557</td>
<td>58,574</td>
<td>5.5</td>
<td>6</td>
<td>4, 7</td>
<td>+</td>
<td>Protein interaction</td>
</tr>
<tr>
<td>C381</td>
<td>41,655</td>
<td>5.56</td>
<td>5</td>
<td>8, 9, 10, 11, 12, 13, 14, 15, 16</td>
<td>+</td>
<td>PRD1 P5 vertex protein</td>
</tr>
<tr>
<td>B345 (coat protein)</td>
<td>37,810</td>
<td>6.17</td>
<td>7</td>
<td>24</td>
<td>+</td>
<td>Major capsid protein</td>
</tr>
<tr>
<td>A223</td>
<td>24,410</td>
<td>4.72</td>
<td>3</td>
<td>17, 18, 19, 20</td>
<td>+</td>
<td>PRD1 P5 vertex protein</td>
</tr>
<tr>
<td>B164</td>
<td>19,025</td>
<td>9.29</td>
<td>1</td>
<td>27, 28</td>
<td>+</td>
<td>Poxy virus ATPase</td>
</tr>
<tr>
<td>B130</td>
<td>13,768</td>
<td>4.97</td>
<td>2, 5</td>
<td>17, 18, 25</td>
<td>+</td>
<td>NSM</td>
</tr>
<tr>
<td>B109</td>
<td>11,969</td>
<td>5.04</td>
<td>2</td>
<td>17, 18, 25</td>
<td>+</td>
<td>NSM</td>
</tr>
<tr>
<td>*A78</td>
<td>9,610</td>
<td>10.79</td>
<td>1</td>
<td></td>
<td>+</td>
<td>NSM</td>
</tr>
<tr>
<td>A55</td>
<td>6,338</td>
<td>4.49</td>
<td>1</td>
<td></td>
<td>+</td>
<td>NSM</td>
</tr>
<tr>
<td>Host 7 DNA-binding protein-kDa (SSO7D)</td>
<td>7,735</td>
<td>9.52</td>
<td>4</td>
<td></td>
<td>+</td>
<td>DNA binding</td>
</tr>
<tr>
<td>Conserved hypothetical protein</td>
<td>25,020</td>
<td>5.86</td>
<td></td>
<td></td>
<td>+</td>
<td>VPS24 vacuolar sorting protein</td>
</tr>
<tr>
<td>SSO0881</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PIC/E a</td>
<td>31,540</td>
<td>5.07</td>
<td>17, 18</td>
<td></td>
<td>–</td>
<td>ASP</td>
</tr>
<tr>
<td>PIC/E</td>
<td>~17,000</td>
<td>~8.5–9.5</td>
<td>29, 30</td>
<td></td>
<td>–</td>
<td>ASP</td>
</tr>
<tr>
<td>PIC/E</td>
<td>~17,000</td>
<td>~4.7</td>
<td>26</td>
<td></td>
<td>–</td>
<td>ASP</td>
</tr>
<tr>
<td>PIC/E</td>
<td>~26,000</td>
<td>~4–5</td>
<td>21, 22, 23</td>
<td></td>
<td>–</td>
<td>ASP</td>
</tr>
</tbody>
</table>

*a* PIC/E, protease inhibitor cocktail and endonucleases added during the sample preparation for analysis by 2D electrophoresis.  
*b* NSM, no significant match.  
*c* ASP, added during sample preparation.
Host Proteins

- SS07D – Stabilizes and packages viral DNA?
  - Suggested histone-like role in *S. solfataricus*
  - Similar protein in SSV1
  - Needed: Similar proteins in other virus particles

- SSO0881 – Cell sorting and trafficking?
  - Similar protein in yeast and mammals

- Consistent cosedimentation, but can't rule out contamination
Genome Mapping

- A78 N-terminus peptide is a valine located 9 nucleotides downstream from A55 amber stop codon (UAG)

- Read-through product?

- Alternative start codon?
PSI-BLAST

- Position Specific Iterative BLAST (Basic Local Alignment Search Tool)
- Position specific scoring
- Scoring based on conservation in related proteins
- “related proteins” determined using traditional BLAST results
- Repeats iteratively
FIG. 4. B164 alignment with several proteins from the family Poxviridae. SwissProt accession numbers are separated by an underscore from species abbreviations (FOWPV, Fowlpox virus; 9POXV, Vultur Gryphus poxvirus; POXVV, Vaccinia virus; MCV1, Molluscum Contagiosum virus subtype 1; SWPV, Swinepox virus; YMTV, Yaba monkey tumor virus; RPOXV, Rabbit fibroma virus). Residues are shaded according to 90% consensus, with white letters on black background signifying perfect conservation. Abbreviations on the consensus line are as follows: h, hydrophobic; s, small; l, aliphatic, b, big; a, aromatic; c, charged.
Surface Overlap Maximization

Consistently placed a single copy of B164 at base of turret complex
B164 as a nucleic acid-packaging ATPase

Three lines of indirect evidence:

- Profile-Profile Comparision
- Position
- 3-nm channel
C381, A223, C557

- Structure and folding of C381, A223 similar to PRD1's P5 vertex protein
- High mass in 1D SDS-PAGE, expected mass in 2D – indication of complex (C381 and A223)
- C557 also located in turrets
- Turret mass ~637 kDa
- C557 + C381 + A223 ~125 kDa (pentameric?)
Glycosylation

- Stained with Pro-Q

- B345 Glycosylated

- Deglycosidases ineffective

- Too stable?
FIG. 6. Total lipid extracts from host and viral membranes. Lipids isolated from host cells (upper) and STIV (lower) were analyzed by negative-mode electrospray mass spectrometry. Viral lipids are a subset of those found in the host.
Analysis of lipids

A. 1858.3745

B. 1858.3703, C_{104}H_{194}O_{24}P_{1}

C. 1858.3341, C_{104}H_{190}O_{24}Cl_{1}

Chemical structures and peaks at different m/z values.
Lipids

- Outer leaflet interacts with B345 (acidic lipids, basic C-terminal of B345)

- Inner leaflet function unknown
  - Thermal stability?
  - Particle assembly/disassembly?
  - Capsid assembly?
Protein, lipid and carbohydrate analysis of STIV indicates that it is similar to tail-less dsDNA viruses from the other domains of life.
Extended Paper Conclusions

- Proteomics-based approaches to protein identification have much utility.

- Relationships between viral proteins can only reliably be detected structurally.

- Strengthens evidence for existence of a common viral ancestor that predates the *Eukarya/Bacteria/Archaea divergence*. 
Questions?