

Cryo-EM Map Challenge, Map Deposition and Archiving

Cathy Lawson
Rutgers University

SEMC Cryo-EM Course
April 15, 2019

Unified Data Resource for 3DEM



**Stanford
University/SLAC**



Rutgers University



**European Bioinformatics
Institute**

- Established 2007 under NIGMS Support (R01GM079429)
- Develop Data Archives for 3DEM (EMDB + PDB)
- Promote Community Development of Validation and Standards

www.emdataresource.org

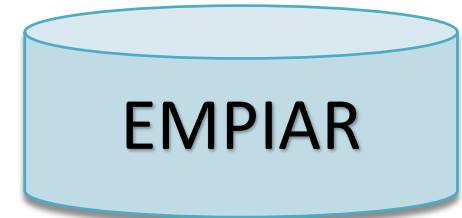
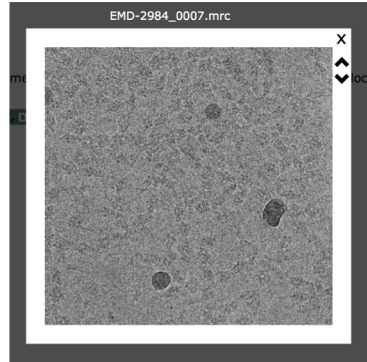


CryoEM Data Archives

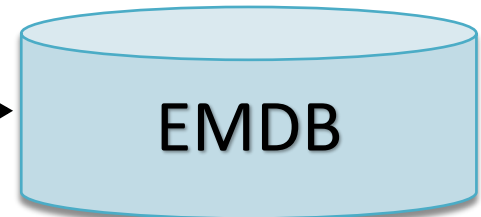
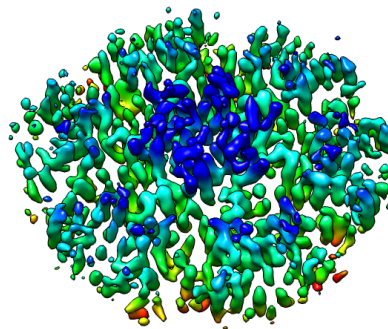
Empiar, EMDB, PDB

Data Archives: What data is found where...

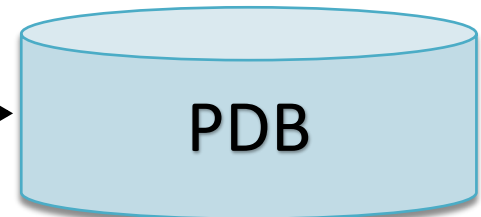
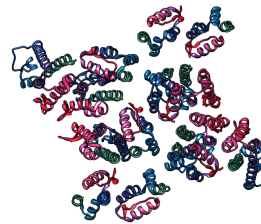
2D Raw
images



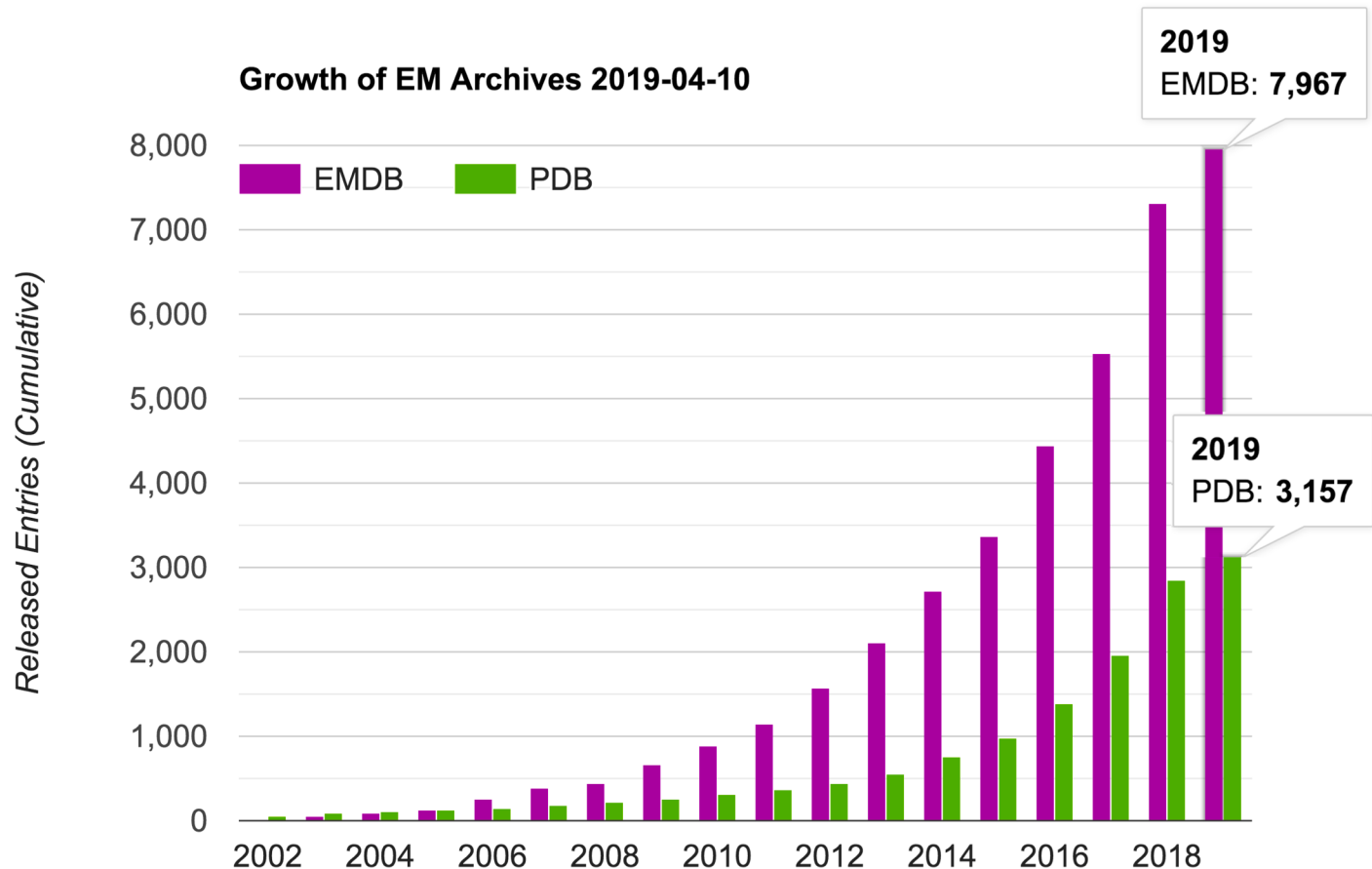
3D Volumes



Fitted
models



Growth of EM Structures



Statistics updated every week:

www.emdataresource.org/statistics.html

CryoEM (R)Evolution

Molecular Shapes

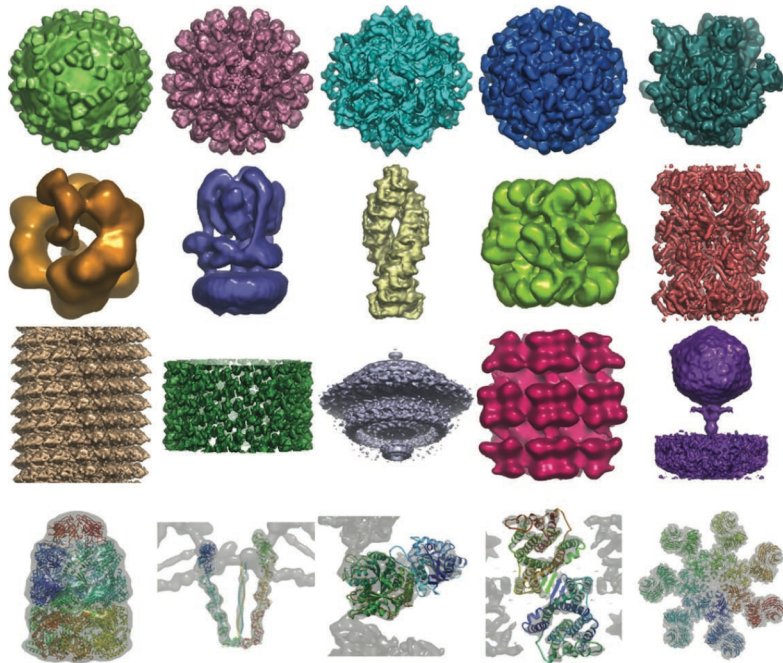
PRINT ISSN: 0305-1048
ONLINE ISSN: 1362-4962

2011

Nucleic Acids Research

VOLUME 39 DATABASE ISSUE JANUARY 1, 2011

www.nar.oxfordjournals.org



OXFORD
UNIVERSITY PRESS

Open Access

No barriers to access – all articles freely available online



Mixed Sources
Product group from well-managed
forests and other controlled sources
www.fsc.org Cert no. SA-COC-000079
© 1996 Forest Stewardship Council

Traceable Densities

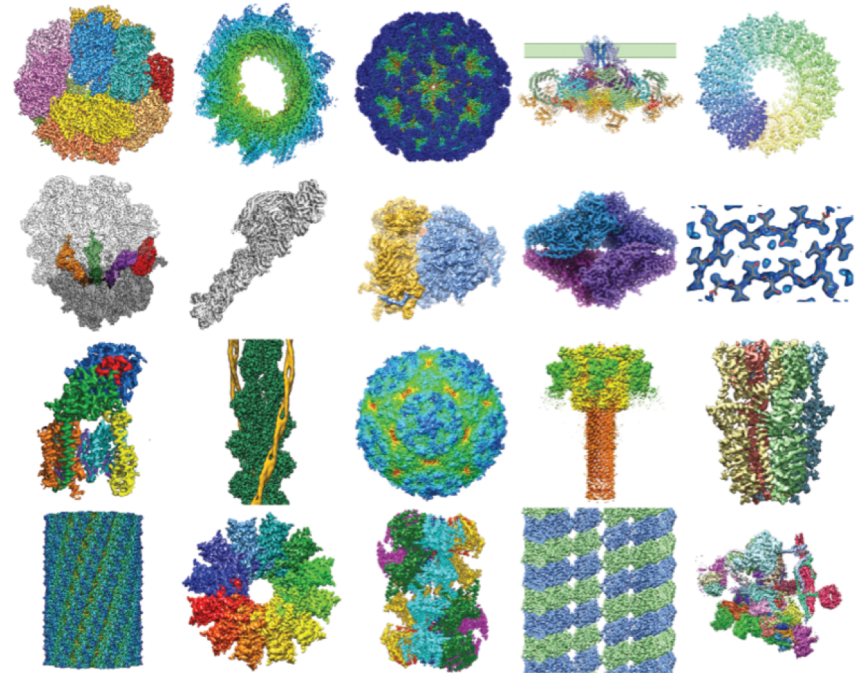
PRINT ISSN: 0305-1048
ONLINE ISSN: 1362-4962

2016

Nucleic Acids Research

VOLUME 44 DATABASE ISSUE JANUARY 4 2016

www.nar.oxfordjournals.org



OXFORD
UNIVERSITY PRESS

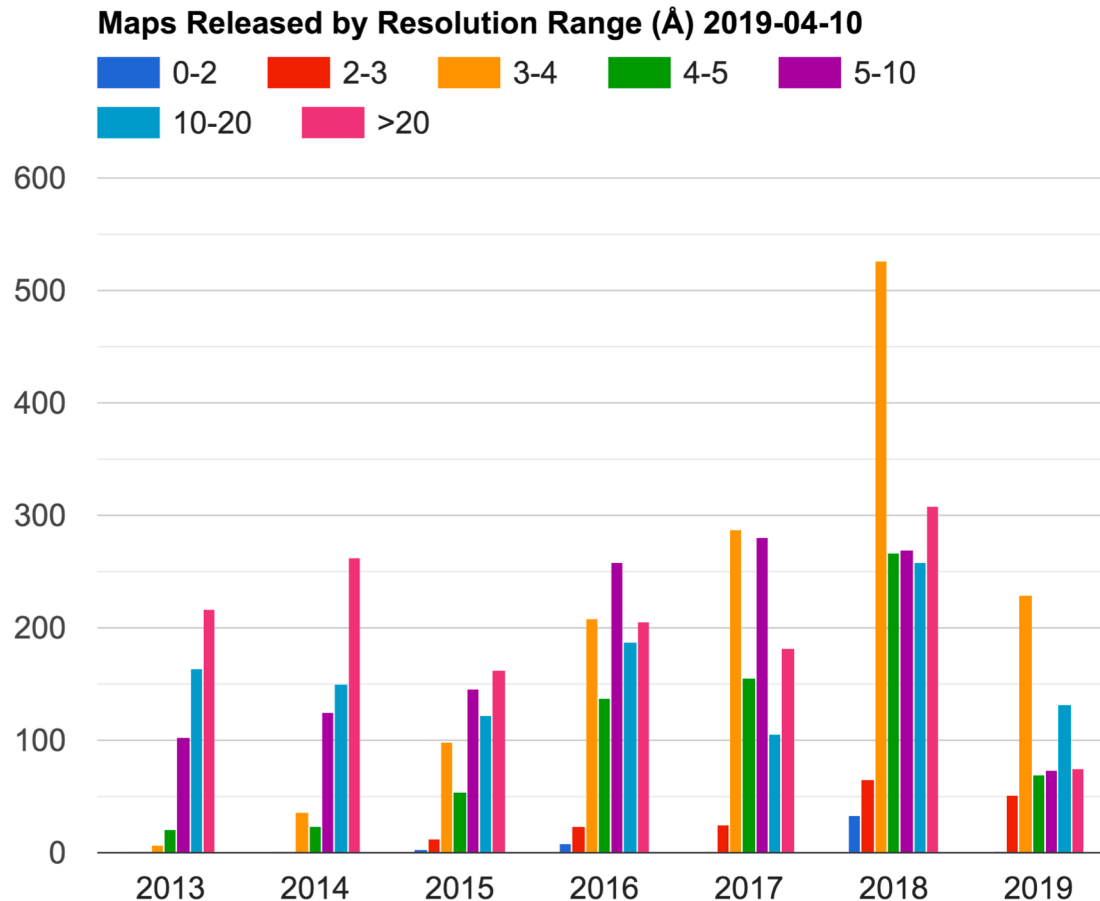
Open Access

No barriers to access – all articles freely available online



MIX
Paper from
responsible sources
FSC® C007207

EMDB entries by year and resolution



Statistics updated every week:

www.emdataresource.org/statistics.html

Finding Cryo-EM Structures

emdataresource.org/search.html

EMDB maps

- [EMDR Search](#)
- [Advanced Search @ EBI](#)

PDB EM models

- [RCSB-PDB](#)
- [PDBe](#)

EMDB + PDB

- [EM Navigator @ PDBj](#)
- [3DBioNotes @ CSIC](#)

Raw image data

- [EMPIAR @ EBI](#)

EMDR Search (New!)

emdataresource.org/search.html

ribosome

Show/Hide Search Options

Found 947 EMDB map entries for query **ribosome** (Released Entries, Any Method, Any Specimen Type, Any Resolution Status)


Copy CSV Excel Column visibility Show 10 rows

Filter Results:

| EMDB id | Status | Entry Title | Entry Authors | Deposit Date | Resolution (Å) | Associated PDB |
|----------|--------|--|---|--------------|----------------|----------------|
| EMD-0047 | REL | Cryo-EM reconstruction of yeast 80S ribosome in complex with mRNA tRNA and eEF2 (GMPPCP/sordarin) | Pellegrino S, Demeshkina N, Mancera-Martinez E, Melnikov S, Simonetti A, Myasnikov A, Yusupov M, Yusupova G, Hashem Y | 2018-06-07 | 4.4 | 6gq1 |
| EMD-0048 | REL | Cryo-EM reconstruction of yeast 80S ribosome in complex with mRNA tRNA and eEF2 (GDP+AlF4/sordarin) | Pellegrino S, Yusupov M, Yusupova G, Hashem Y | 2018-06-07 | 3.9 | 6gqb |
| EMD-0049 | REL | Cryo-EM reconstruction of yeast 80S ribosome in complex with mRNA tRNA and eEF2 (GMPPCP) | Pellegrino S, Yusupov M, Yusupova G, Hashem Y | 2018-06-08 | 4 | 6gqv |
| EMD-0055 | REL | Cryo-EM reconstruction of yeast 80S ribosome in complex with mRNA tRNA and eEF2 (GDP+AlF4/sordarin): SSU-focused map | Pellegrino S, Yusupov M, Yusupova G, Hashem Y | 2018-06-14 | 4.3 | |

EMDR Search: Setup

emdataresource.org/solrsearch.html

**EMDataResource**
Unified Data Resource for 3DEM

?

Q

X

Show/Hide Search Options

Filters

Released Entries ▾ Any Method ▾ Any Specimen Type ▾ Any Resolution Status ▾

Results

☒ Entry Title

☒ Entry Authors

☒ Deposition Date

☐ Header Release Date

☐ Map Release Date

☐ Method

☐ Specimen type

☐ Sample Molecular Weight

☐ Microscope Model

☐ Detector

☐ Energy Filter

☐ Electron Dose

☐ Number of Particles

☐ Imposed Symmetry

☒ Reported Resolution

☒ Fitted Model PDB

☐ Citation Authors

☐ Citation Title

☐ Citation Year

☐ Journal Name

☐ PubMed

☐ DOI

EMDR Search: Keywords and Ranges

emdataresource.org/solrsearch.html

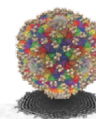
Keyword Help



find

| | |
|---|--|
| Deposition Date | deposited:[2002 TO 2004] |
| Map Release Date | released:2018-10-31 |
| Sample Molecular Weight (MDa) | sampleMW:[10 TO *] |
| Microscope Model | microscope:JEOL |
| Detector | detector:CCD |
| Energy Filter | energyfilter:omega |
| Electron Dose (e/Åsq) | dose:[5 to 10] |
| Number of Particles | numparticles:[1 to 250] |
| Imposed Symmetry | symmetry:D* |
| Reported Resolution(Å) | resol:[1.5 TO 3.5] |
| Citation Year | pubyear:2019 |
| Citation Journal Name | journal:"Mol Cell" |
| Citation PubMed ID | PMID:27654913 |
| Raw Images EMPIAR ID | EMPIAR:10250 |
| Intersection | ribosome and human |
| Multiple Topics | ribosome or tRNA |
| Exclude Subtopic | ribosome and not human |
| Combine Keyword, Topic | released:2018 and spliceosome |
| Combine hint: add quotes around authornames with initials or multiword topics | "Frank J" AND resol:[7 TO 8] AND symmetry:C2 |

Note: use square and/or curly brackets for date and numerical ranges, e.g., [x TO y]: equal to or more than x, less than or equal to y, {x TO y}: more than x, less than or equal to y



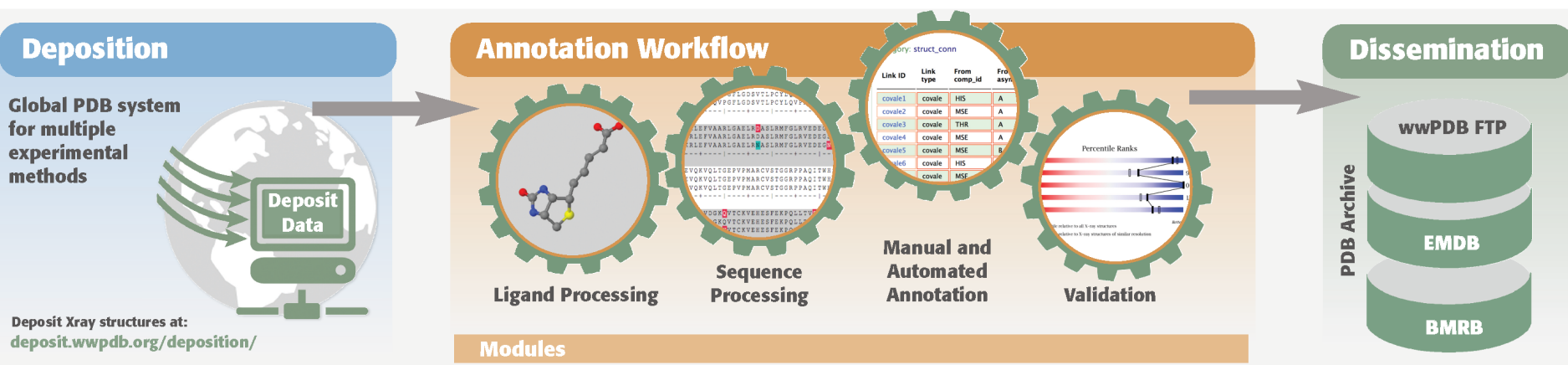
EMDataResource
Unified Data Resource for 3DEM

Cryo-EM Structure Deposition

EMDB, PDB

wwPDB neDep System

- X-ray, NMR, and EM Methods (since 2016)
- EM Methods: Deposit map to EMDB with associated model to PDB
- Validation report produced



3DEM Deposition: Method

Experimental method

[deposit.wwpdb.org](https://deposit ww.pdb.org)

- ☐ X-Ray Diffraction
- ☒ Electron Microscopy
 - ☐ Helical
 - ☒ Single particle
 - ☐ Subtomogram averaging
 - ☐ Tomography
- ☐ Solution NMR
- ☐ Neutron Diffraction
- ☐ Electron Crystallography
- ☐ Solid-state NMR
- ☐ Fiber Diffraction

3DEM Deposition: ID Assignment

deposit.wwpdb.org

Are you depositing coordinates with this submission?

☐ No, experimental data only

☒ Yes

Has the associated map been deposited previously?

☒ No

☐ Yes

~~Coming Soon:~~ Here!: 5 digit EMDB ids
e.g. EMD-12345

Requested accession codes



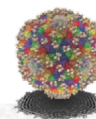
PDB



EMDB



BMRB



EMDataResource

Unified Data Resource for 3DEM

File uploads: 3DEM map/model submission in OneDep

✓ Select file type...

0) Coordinates

Coordinates (mmCIF format)

Coordinates (PDB format)

1) Main map (mandatory)

EM map (MRC/CCP4 format)

2) Image for EMDB (mandatory)

Entry image for public display

3) Additional maps

Additional EM map (MRC/CCP4 format)

4) Masks

EM mask (MRC/CCP4 format)

5) Half (even-odd) maps

EM half map (MRC/CCP4 format)

6) Structure Factors

mmCIF (structure factors)

MTZ

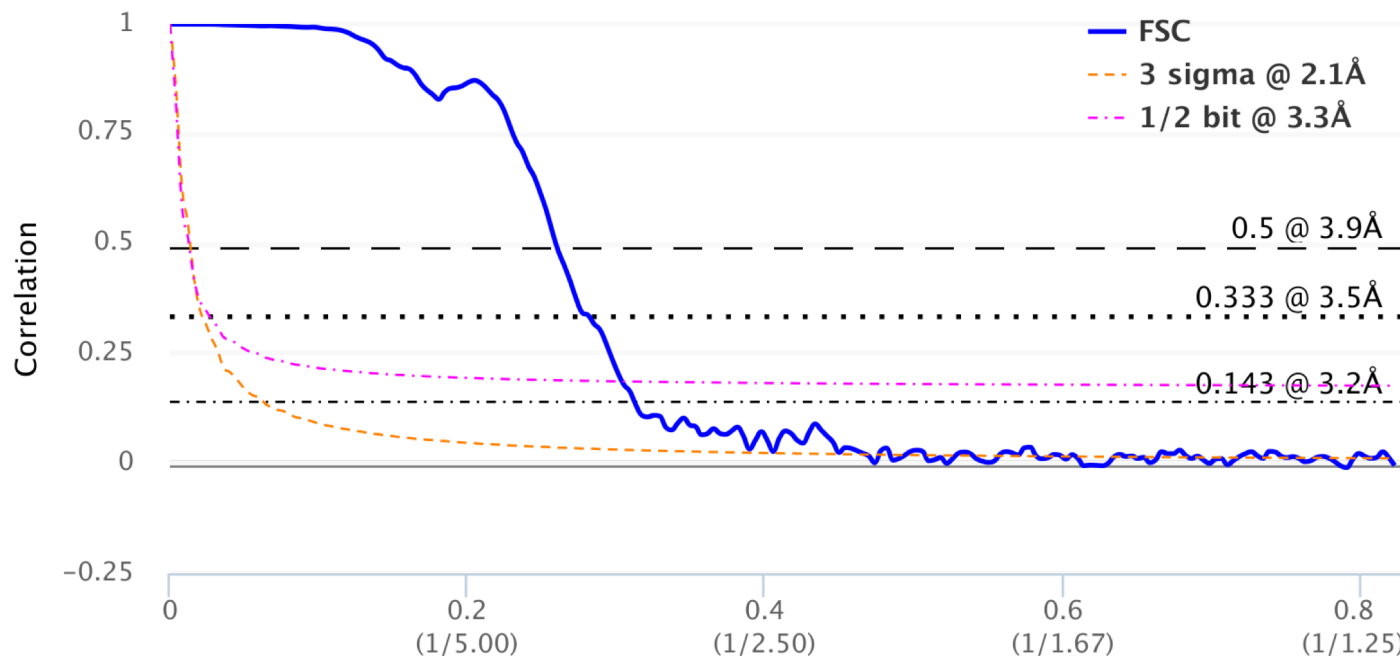
Other Files

FSC file (XML format)

Ligand Image

FSC Curve Upload

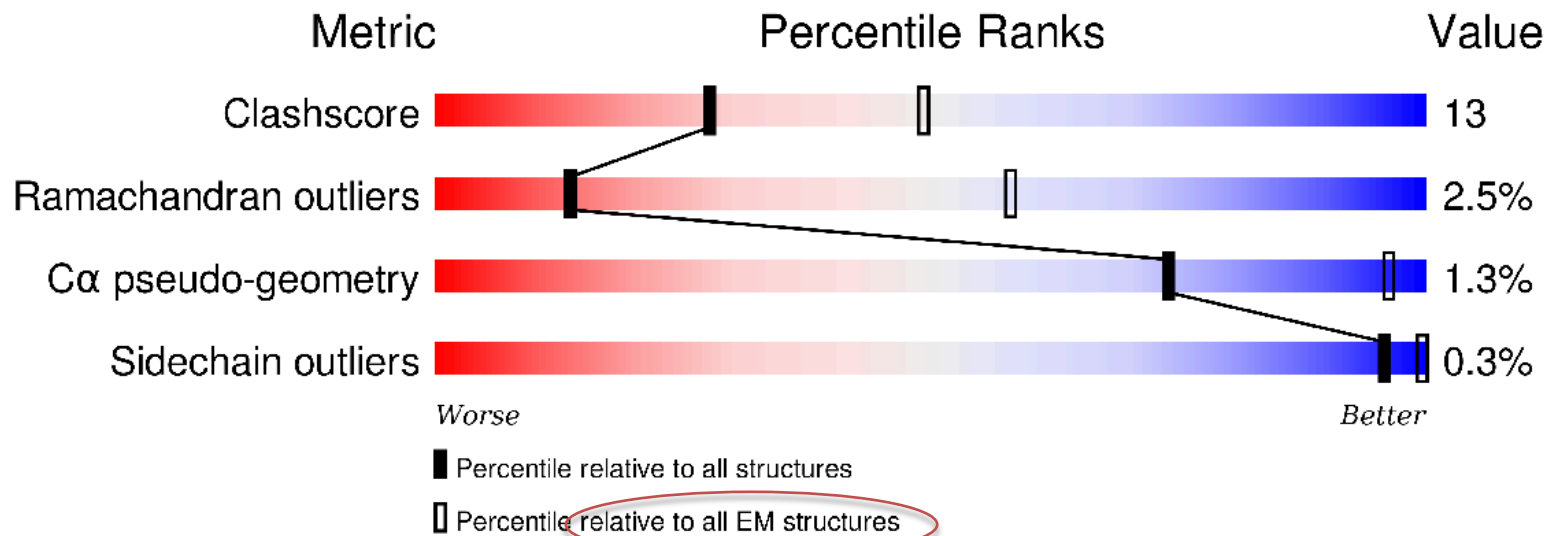
- Create xml format file using a software package (e.g., Relion, EMAN), or...
- Use PDBe's Server: PDBe.org/FSC
- Community recommendation: upload half-maps



EM Validation Report

- “Table 1” + EM model metrics
- Comparative statistics updated annually
- Planned improvements: images/statistics

| Property | Value | Source |
|---------------------------------|-----------------|-----------|
| Reconstruction method | SINGLE PARTICLE | Depositor |
| Imposed symmetry | I | Depositor |
| Number of images | 30000 | Depositor |
| Resolution determination method | FSC 0.143 | Depositor |
| CTF correction method | Not provided | Depositor |
| Microscope | JEOL 3200FSC | Depositor |
| Voltage (kV) | 300 | Depositor |
| Electron dose (| | |
| Minimum defoc | | |
| Maximum defoc | | |
| Magnification | | |
| Image detector | | |

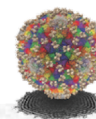


Archive Files and Data Dictionaries

- EMDB produces EMDB/xml format files
- PDB produces PDBx/mmCIF files
- Underlying dictionaries are equivalent!

Example: Vitrification Instruments

```
<xs:simpleType name="vitrInstrType">
  <xs:restriction base="xs:string">
    <xs:enumeration value="BAL-TEC HPM 010"/>
    <xs:enumeration value="EMS-002 RAPID IMMERSION FREEZER"/>
    <xs:enumeration value="FEI VITROBOT"/>
    <xs:enumeration value="FEI VITROBOT MARK I"/>
    <xs:enumeration value="FEI VITROBOT MARK II"/>
    <xs:enumeration value="FEI VITROBOT MARK III"/>
    <xs:enumeration value="FEI VITROBOT MARK IV"/>
    <xs:enumeration value="GATAN CRYOPLUNGE 3"/>
    <xs:enumeration value="HOMEMADE PLUNGER"/>
    <xs:enumeration value="LEICA PLUNGER"/>
    <xs:enumeration value="LEICA EM GP"/>
    <xs:enumeration value="LEICA EM CPC"/>
    <xs:enumeration value="LEICA EM HPM100"/>
    <xs:enumeration value="LEICA EM PACT"/>
    <xs:enumeration value="LEICA EM PACT2"/>
    <xs:enumeration value="LEICA KF80"/>
    <xs:enumeration value="NONE"/>
    <xs:enumeration value="REICHERT-JUNG PLUNGER"/>
    <xs:enumeration value="ZEISS PLUNGE FREEZER CRYOBOX"/>
    <xs:enumeration value="OTHER"/>
    <xs:enumeration value="SPOTITON"/>
  </xs:restriction>
```



CryoEM Validation Challenges

EM VTF 2010 Recommendations

- Full FSC curve from independent half-maps
- Model Stereochemistry (same as X-ray / NMR)
- Other Metrics: **More Research Needed**

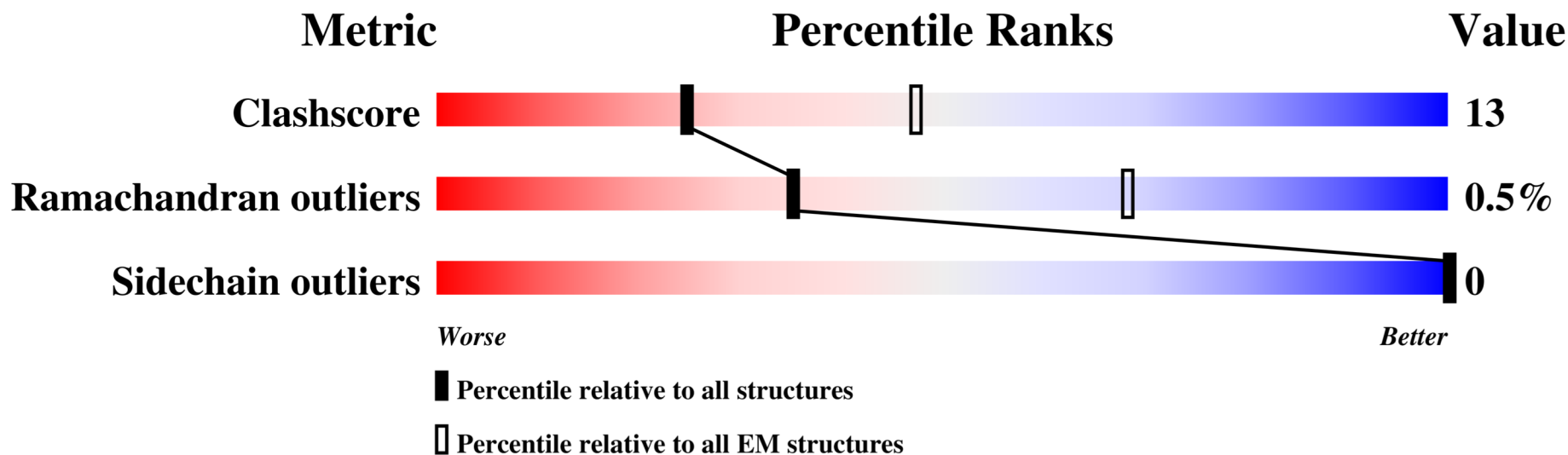


Henderson *et al.* (2012) **Structure** 20, 205-214
<http://www.ncbi.nlm.nih.gov/pubmed/22325770>

vtf.emdataresource.org

Current Validation Report for EM Structures

- Resolution value reported by depositor
- Model statistics

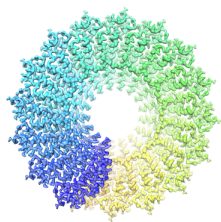


2015/2016 Map, Model Challenges

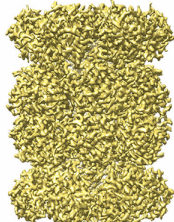
■ Goals:

- Develop benchmarks
- encourage development of best practices
- evolve criteria for validation
- compare and contrast different approaches

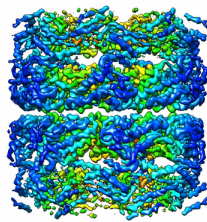
■ Using recent data archived in EMPIAR, EMDB, PDB:



Tobacco
Mosaic Virus



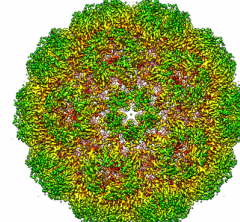
T20S
Proteasome



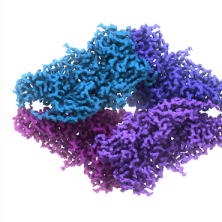
GroEL



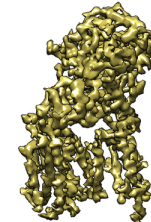
TrpV1
channel



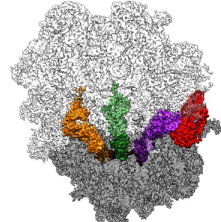
Brome
Mosaic Virus



β -galacto-
sidase

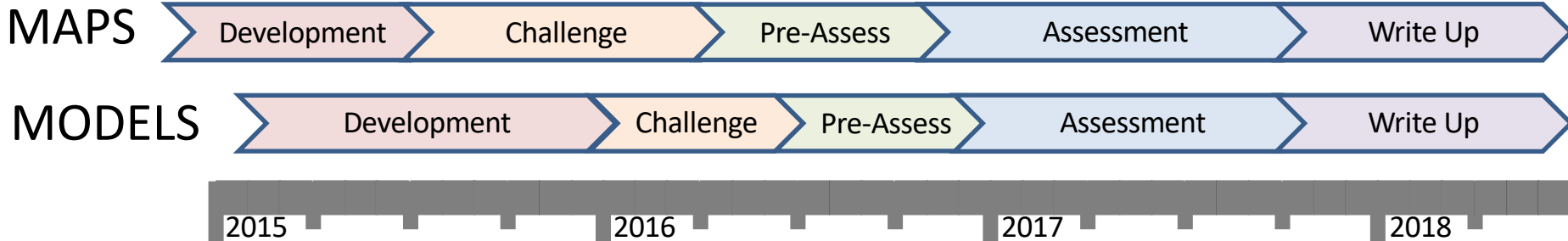
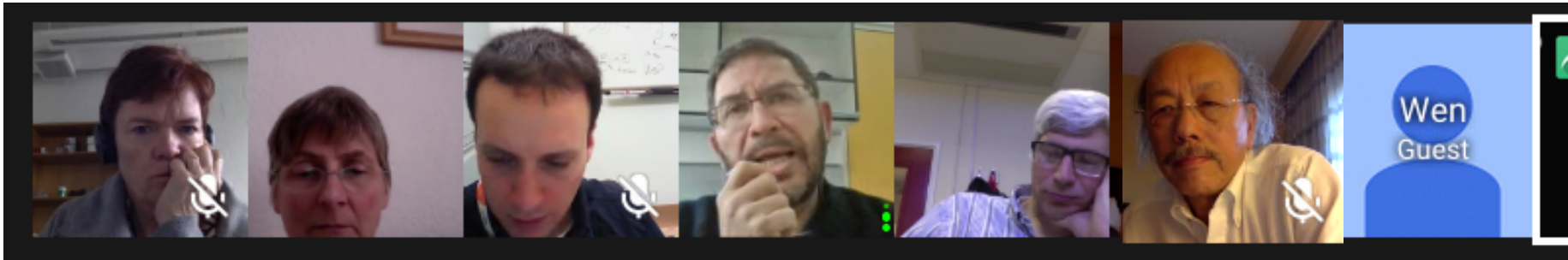


γ -Secretase



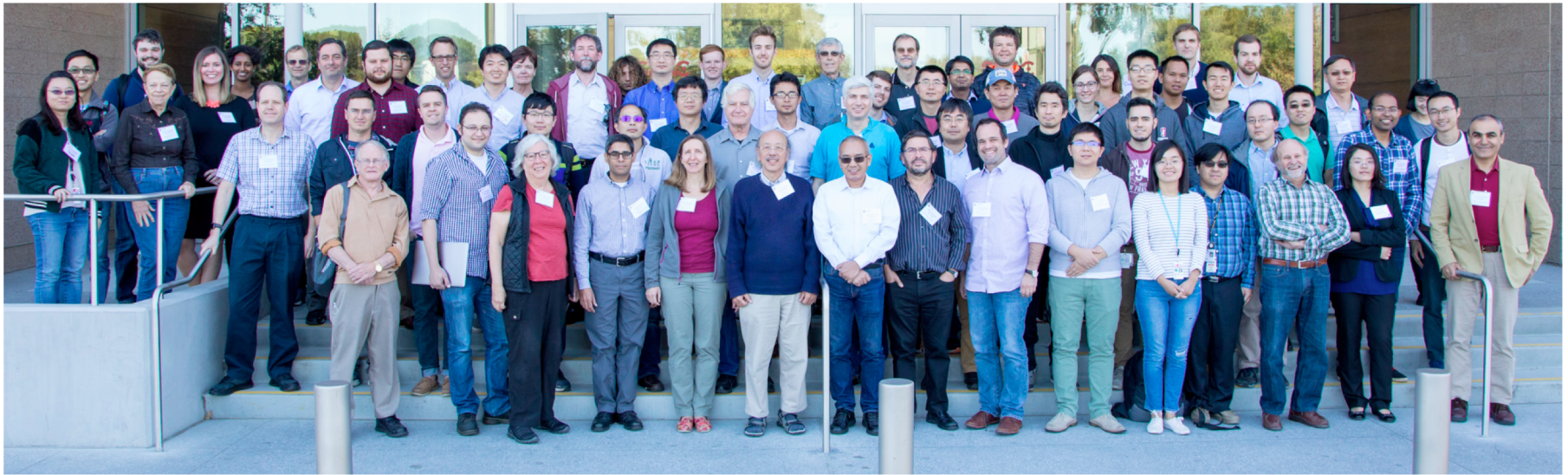
70S
Ribosome

The Process



Wrap Up
Oct 6-8, 2017

Joint Challenges Workshop Oct 6-8, 2017, Stanford/SLAC



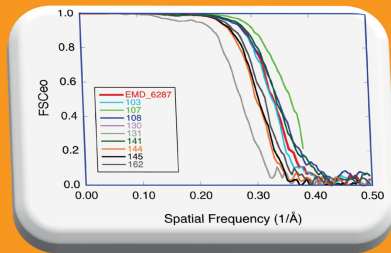
Virtual Special Issue: December 2018



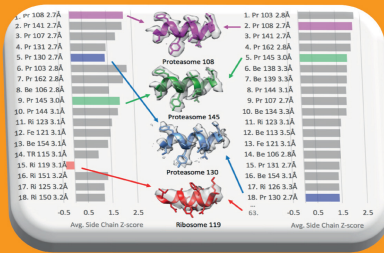
Volume 204, Number 3, December 2018

ISSN 1047-8477

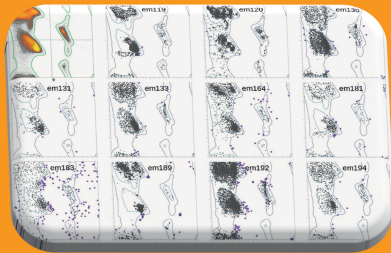
Journal of Structural Biology



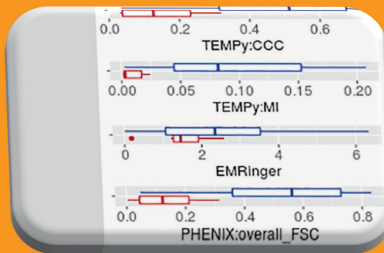
Map Resolution



Map Features



Model Geometry

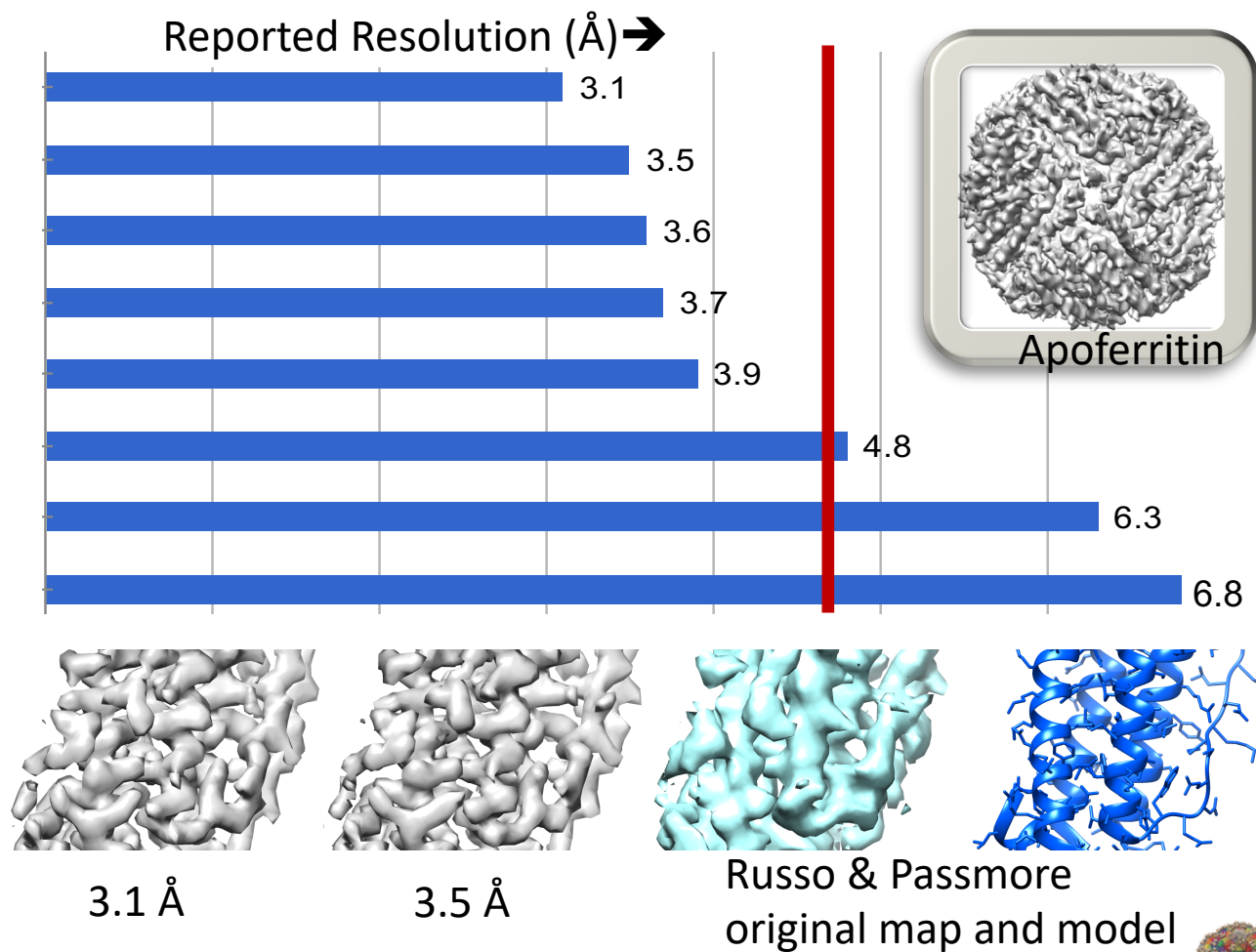


Model Fit-to-Map

The CryoEM Structure Map and Model Challenges

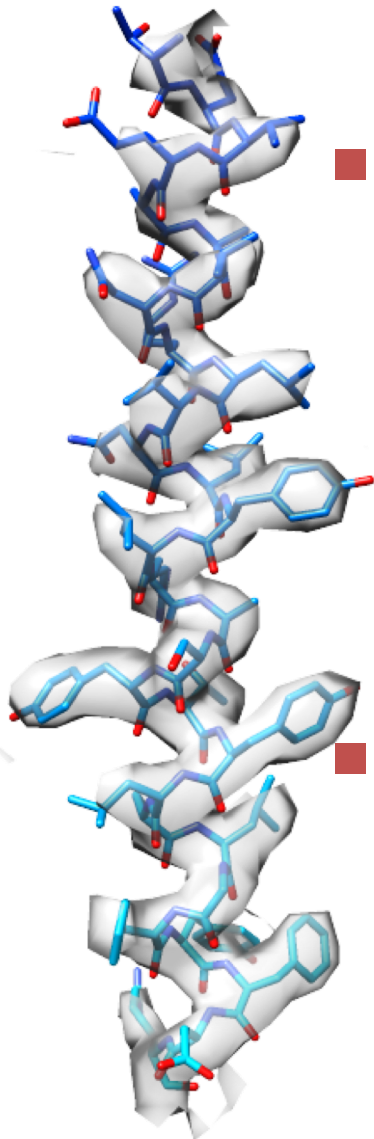
- 18 peer-reviewed articles +editorial summarizing the outcomes of the Oct workshop
- Recommended best practices as well as novel methods for Cryo-EM structure determination and assessment

Map Challenge: Apoferritin Target



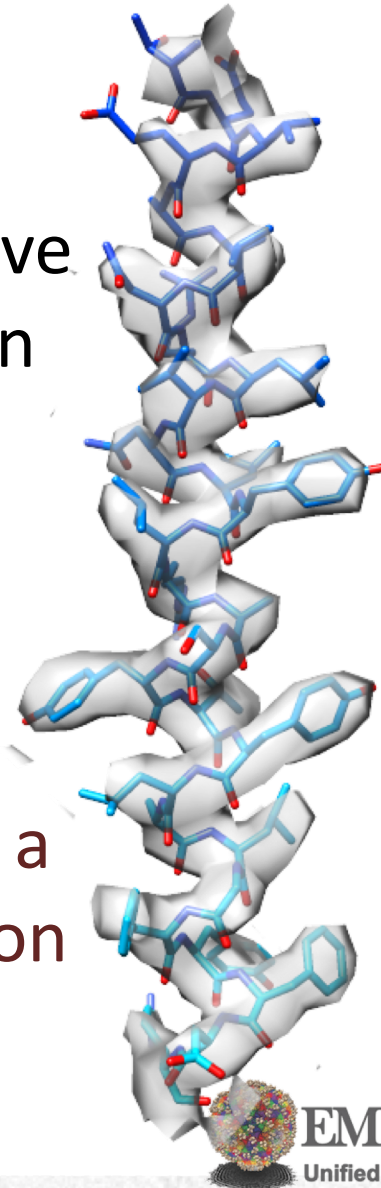
Apo ferritin 3.1 vs 3.5 Å

← Which is which? →



- Different expert practitioners can arrive at different resolution estimates for same level of map detail

- Question: Could the community agree on a standard for resolution estimation?



Challenges Wrap-Up: Maps

■ Results

- All major packages can produce maps of equivalent quality
- But quality can vary considerably among practitioners
- Reported resolution is not fully reliable indicator of resolvability

■ Conclusions

- Current (FSC) practices are inconsistent
- Bullet-proof workflows and best-practice standards needed for reconstruction and resolution estimation

Frontiers in Cryo-EM Validation Workshop

■ January 14-15 at EMBL-EBI in Hinxton, UK



Frontiers Workshop: Some of the Outcomes

- Lots of terrific discussion with near consensus on how to handle most current issues
- Recommendations will be summarized in a white paper
- Key recommendation: EMDB-calculated resolution based on deposited (unmasked, minimally filtered) half-maps

EM Structure Validation Servers

| Map: | Service/Name | Link |
|---------------------|--------------|--|
| Overall Shape, Hand | Tilt-Pair | pdbe.org/tiltpair |
| Resolution | FSC | pdbe.org/FSC |
| Local Resolution | 3DFSC | 3dfsc.salk.edu |
| Local Resolution | Scipion | scipion.cnb.csic.es/m/myresmap# |

| Model: | Service/Name | Link |
|--|--------------|--|
| Stereochemistry, compare with all PDB structures | wwPDB | validate.wwpdb.org |
| Stereochemistry | Molprobit | molprobit.biochem.duke.edu |

| Map/Model Fit: | Service/Name | Link |
|----------------------|--------------|--|
| “backbone bumpiness” | EMRinger | emringer.com (@UCSF) |

Unified Data Resource for 3DEM



**Stanford
University/SLAC**



Rutgers University



**European Bioinformatics
Institute**

Wah Chiu
Greg Pintilie
Andriy Kryshfovych (UC
Davis)

Catherine Lawson
Helen Berman
Brinda Vallat
Brian Hudson
John Westbrook

Ardan Patwardhan
Gerard Kleywegt
Sanja Abbott
Ryan Pye
Osman Salih
Zhe Wang

www.emdataresource.org



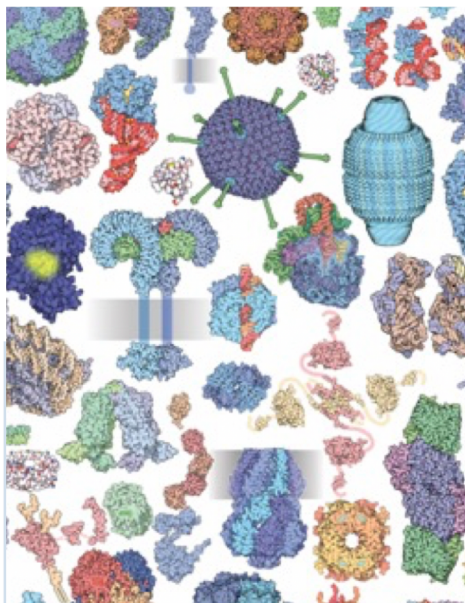
References

- Lawson CL, Chiu W (2018) Comparing cryo-EM structures (Editorial). **J Struct Biol.** **204**, 523-526. [10.1016/j.jsb.2018.10.004](https://doi.org/10.1016/j.jsb.2018.10.004)
- Patwardhan A & Lawson CL (2016). *Databases and Archiving for CryoEM*. **Methods Enzymol** **579**, 393-412. [10.1016/bs.mie.2016.04.015](https://doi.org/10.1016/bs.mie.2016.04.015)
- Henderson R, et al (2012) *Outcome of the first electron microscopy validation task force meeting*. **Structure** **20**, 205-214. [10.1016/j.str.2011.12.014](https://doi.org/10.1016/j.str.2011.12.014)

Join the RCSB PDB Team

<http://www.rcsb.org/pages/jobs>

SOFTWARE DEVELOPERS AND BIOCHEMISTS



Join the RCSB Protein Data Bank Team at Rutgers, The State University of New Jersey

Open positions:

**Biochemical Information &
Annotation Specialist
(Biocurator)**

Curate, validate, and
standardize macromolecular
structures from the PDB
community.

Front End Web Developer

Develop and maintain web
applications, from design to
deployment.