RCSB PDB PROTEIN DATA BANK

IQB

VIRTUAL WEBINAR

UNDERSTANDING PDB VALIDATION: WHICH EXPERIMENTAL STRUCTURES SHOULD I RELY ON?

Tuesday May 14th 2024
2-3pm Eastern | 11am-12pm Pacific

Stephen K. Burley, M.D., D.Phil.

Chenghua Shao, Ph.D.

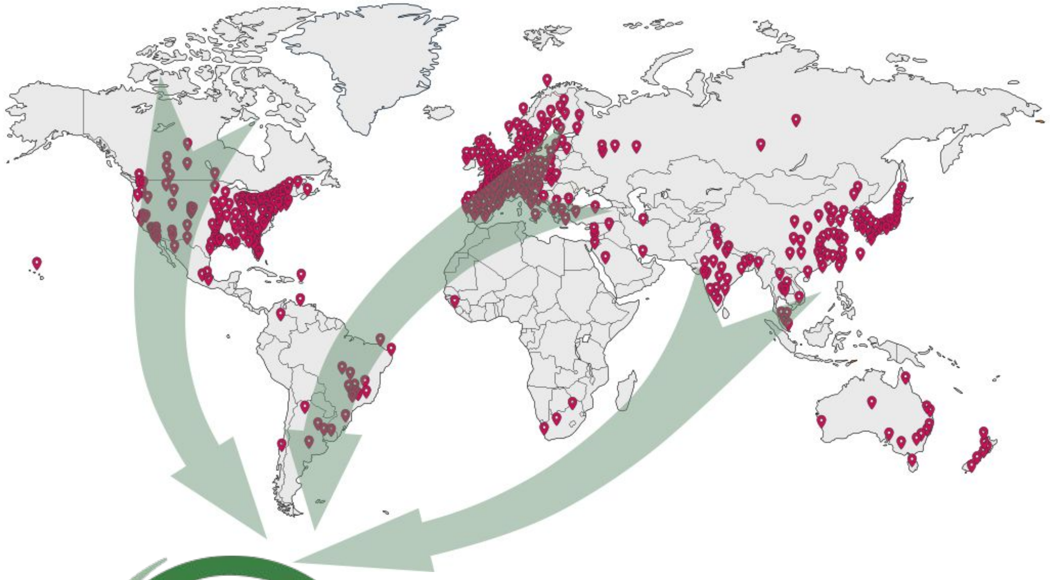
Rutgers, The State University of New Jersey

Outline

- Protein Data Bank (PDB) and quality variation of structures
- Worldwide Protein Data Bank (wwPDB) validation of structure quality
    - Validation Overview
    - Validation of chemical geometry for PDB structures determined by all methods
    - Validation of macromolecular crystallography (MX) structures
    - Validation of 3D electron microscopy (3DEM) structures
- RCSB.org access to validation reports and quality review in 3D
- RCSB.org structure confidence review in 3D
- RCSB.org interactive ligand quality review

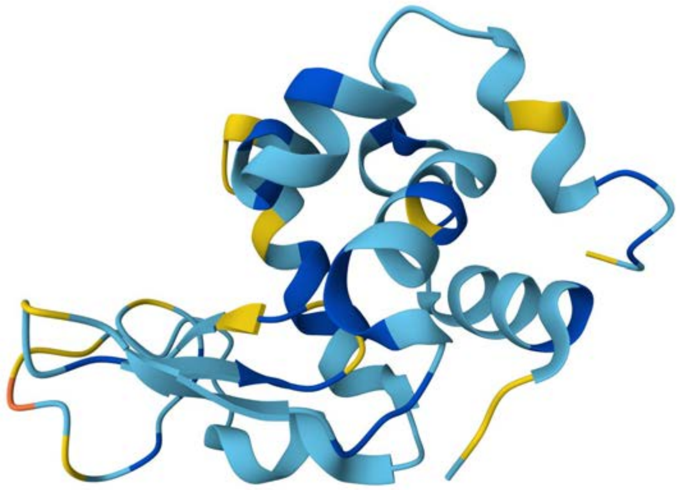
Protein Data Bank (PDB) Archive

- 1st open access digital data resource in all of Biology established in 1971
- Single global archive for protein and nucleic acid experimental structures with ~220,000 structures
- Managed jointly by Worldwide PDB (wwPDB) regional partners
    - RCSB PDB (US)
    - Protein Data Bank in Europe (PDBe)
    - PDB Japan (PDBj)
    - Associate Member: PDB China (PDBc)
    - Plus EMDB and BMRB
- All PDB data are validated, deposited, and biocurated using OneDep



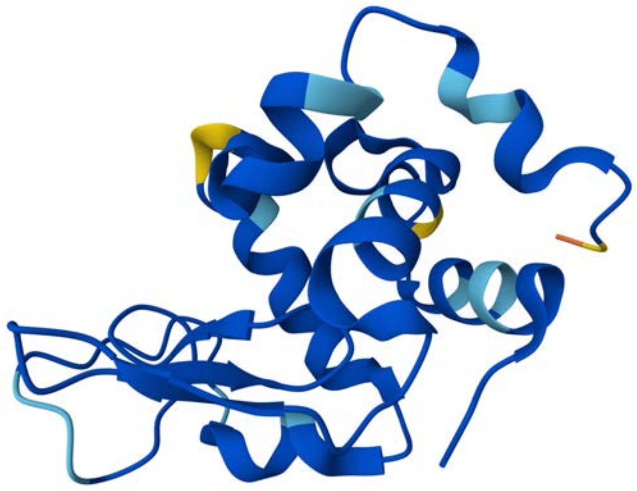
PDB Structure Quality Varies

- Structure quality depends on the experimental data, structure determination, and other factors
- Quality metrics and visual inspection can tell you a lot about structure quality
- wwPDB Validation Report calculates these metrics and provides review
- RCSB.org provides additional metrics and tools to perform quality reviews tailored to your needs



PDB ID 5F81
at 2.1 Å

Colored by Confidence
- Very high
- High
- Low
- Very Low



PDB ID 5HNL
at 2.4 Å

5F81: Roessler *et al.* Structure 24: 631-640
5HNL: Tsukui *et al.* Acta Crystallogr D Struct Biol 72: 823-829

Validation Overview

Chenghua Shao, Ph.D.

Structure Validation is Key to PDB Archive

- wwPDB method-specific Validation Task Forces published recommendations
    - Macromolecular Crystallography (MX): Read *et al*. (2011) *Structure* 19, 1395-141
    - 3D Electron Microscopy (3DEM): Henderson *et al*. (2012) *Structure* 20, 205-214
    - NMR Spectroscopy (NMR): Montelione *et al*. (2013) *Structure* 21, 1563-1570
- OneDep launched by wwPDB in 2014
- wwPDB/CCDC/D3R Ligand Validation Workshop in 2016
    - Adams *et al.* (2016) *Structure* 24, 502-508
    - New ligand validation implemented with code from Global Phasing Limited
    - wwPDB Validation 2.0 launched 2019

CellPress

Structure

Resource

OneDep: Unified wwPDB System for Deposition, Biocuration, and Validation of Macromolecular Structures in the PDB Archive

Jasmine Y. Young,[1,12,*] John D. Westbrook,[1] Zukang Feng,[1] Raul Sala,[1] Ezra Peisach,[1] Thomas J. Oldfield,[2,9] Sanchayita Sen,[2] Aleksandras Gutmanas,[2] David R. Armstrong,[2] John M. Berrisford,[2] Li Chen,[1] Minyu Chen,[3] Luigi Di Costanzo,[1] Dimitris Dimitropoulos,[1,10] Guanghua Gao,[1] Sutapa Ghosh,[1] Swanand Gore,[2] Vladimir Guranovic,[1] Pieter M.S. Hendrickx,[2] Brian P. Hudson,[1] Reiko Igarashi,[3] Yasuyo Ikegawa,[3] Naohiro Kobayashi,[3] Catherine L. Lawson,[1] Yuhe Liang,[1] Steve Mading,[4] Lora Mak,[2] M. Saqib Mir,[2] Abhik Mukhopadhyay,[2] Ardan Patwardhan,[2] Irina Persikova,[1] Luana Rinaldi,[2] Eduardo Sanz-Garcia,[2] Monica R. Sekharan,[1] Chenghua Shao,[1] G. Jawahar Swaminathan,[2,11] Lihua Tan,[1] Eldon L. Ulrich,[4] Glen van Ginkel,[2] Reiko Yamashita,[3] Huanwang Yang,[1] Marina A. Zhuravleva,[1] Martha Quesada,[1] Gerard J. Kleywegt,[2] Helen M. Berman,[1] John L. Markley,[4] Haruki Nakamura,[3] Sameer Velankar,[2] and Stephen K. Burley[1,5,6,7,8]

[1]RCSB Protein Data Bank, Center for Integrative Proteomics Research, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA

[2]Protein Data Bank in Europe (PDBe), European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK

[3]PDBj, Institute for Protein Research, Osaka University, Osaka, 565-0871, Japan

[4]BMRB, BioMagResBank, University of Wisconsin-Madison, Madison, WI 53706, USA

[5]RCSB Protein Data Bank, San Diego Supercomputer Center

[6]Skaggs School of Pharmacy and Pharmaceutical Sciences University of California San Diego, La Jolla, CA 92093, USA

[7]Institute for Quantitative Biomedicine, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA

[8]Rutgers Cancer Institute of New Jersey, Rutgers, The State University of New Jersey, New Brunswick, NJ 08903, USA

[9]Present address: Dotmatics, The Old Monastery, Windhill, Bishops Stortford, Herts CM23 2ND, UK

[10]Present address: Google, Irvine, CA 92612, USA

[11]Present address: Illumina Cambridge Ltd., Chesterford Research Park, Little Chesterford CB10 1XL, UK

[12]Lead Contact

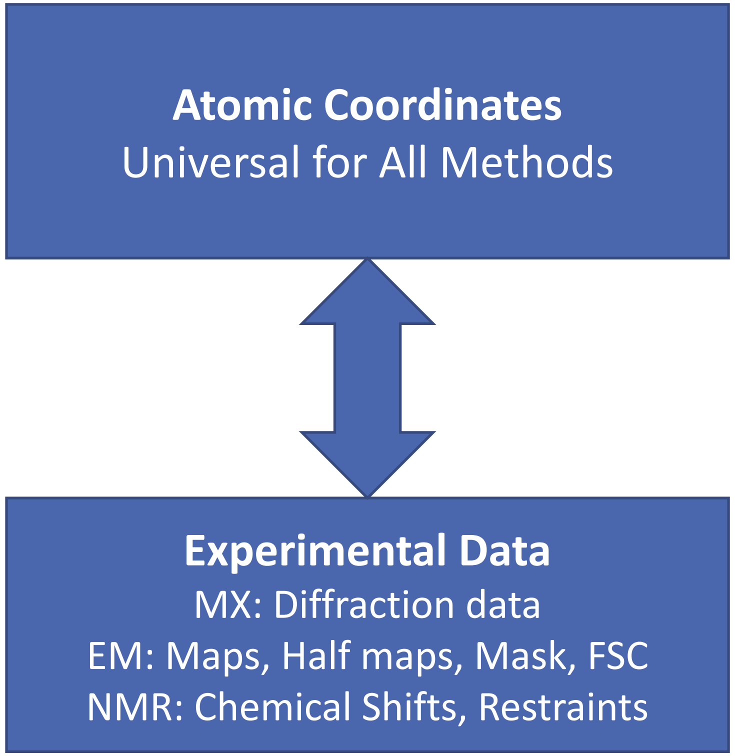
*Correspondence: jasmine.young@rcsb.org

http://dx.doi.org/10.1016/j.str.2017.01.004

Young *et al.* (2017) *Structure 25*, 536-545;
Gore *et al*. (2017) *Structure 25*, 1916-1927

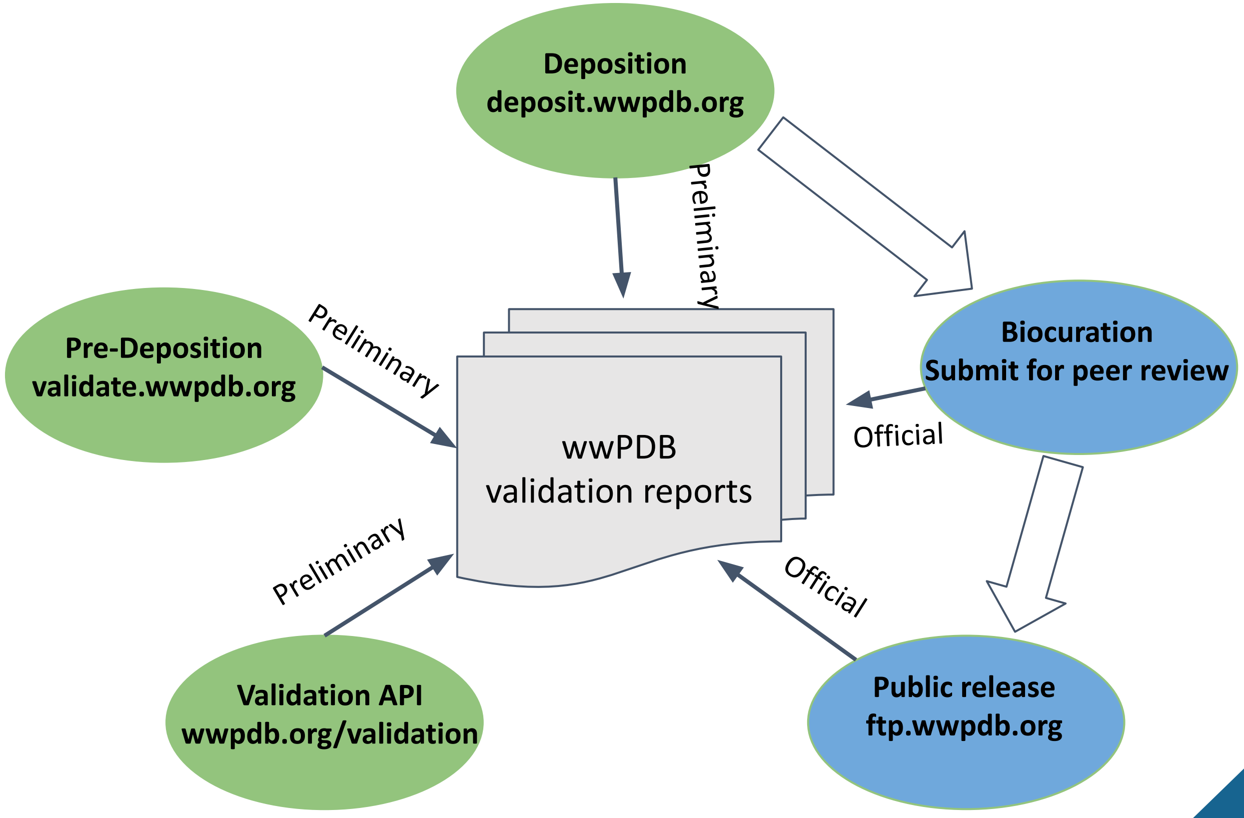
wwPDB Validation Scope

- Molecular geometry agreement with established chemical references (bond lengths, bond angles, *etc*.)
- Experimental data quality
- Goodness-of-fit between atomic coordinates and method-specific experimental data
- Global *vs.* local structure validation
- Validation for distinct molecular components (polymers, ligands, etc.)



wwPDB Validation Reports Tailored to Different Audiences

- Data Authors/Depositors: Can generate and access watermarked reports pre-/post-deposition
  - Deposition site
  - Standalone validation server
  - Application Programming Interface (API)
- Journals: Supporting peer review
  - Authors provide reports to journals
  - Journals provide reports to referees
  - Required by many journals
- Data Consumers
  - Access reports on all wwPDB partner sites
  - CIF/XML/PDF formatted reports available for download and analysis



wwPDB Validation Supports Peer Review

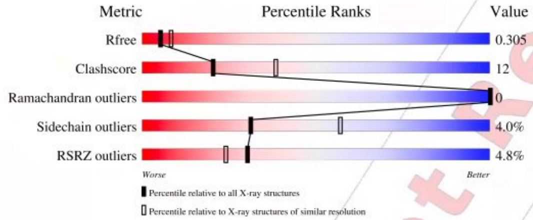
- PDB policy requires mandatory experimental data deposition for method-specific validation
- wwPDB Validation Report with special watermark provided to Journal together with manuscript by Authors
- Many scientific journals require wwPDB validation report for manuscript submission, including
    - Cell
    - IUCr journals
    - J Biol Chem
    - Nature
    - PLoS One
    - Protein Sci
    - Science
    - Structure

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.59 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	159	2%; 60% 16% • 23%
2	C	14	21%; 7% 21% 71%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Validation Report Slider for Overall Quality at a Glance (X-ray Crystallographic Example)

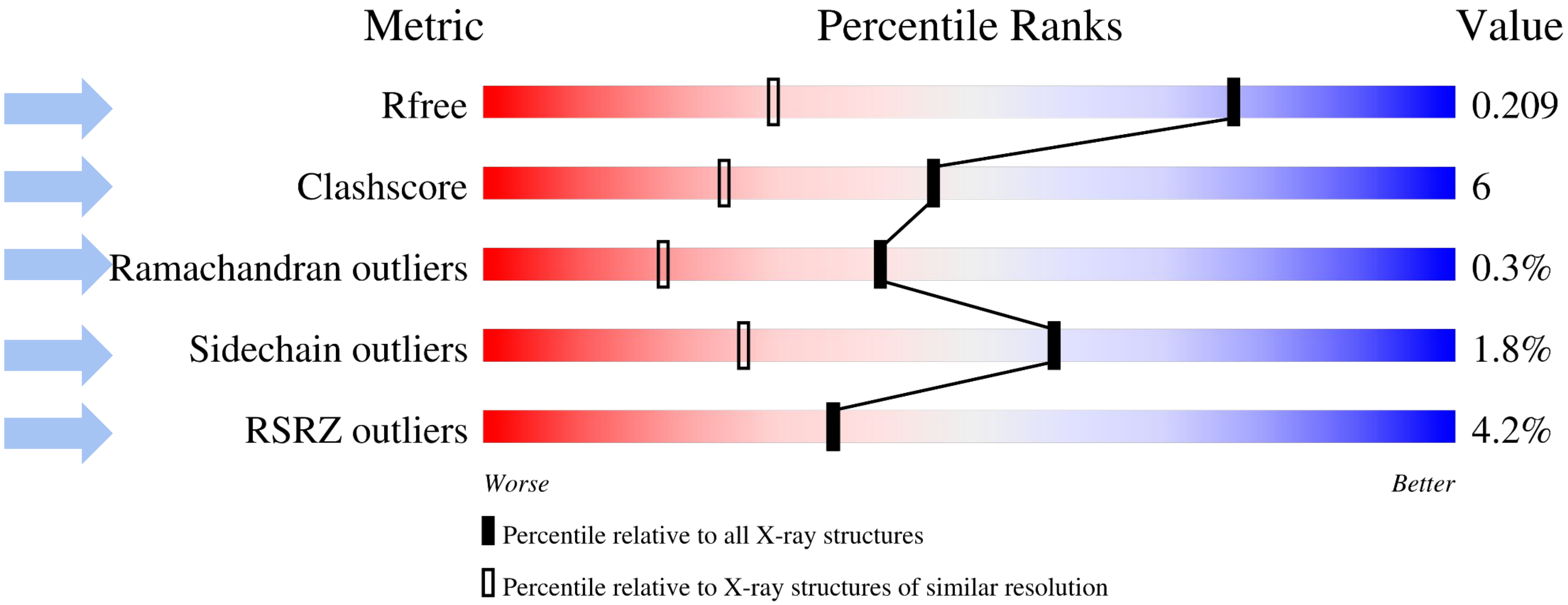
How well does the overall structure agree with MX data?

Atom clashes/1000 residues

Residues with unusual **main chain** torsion angles

Residues with unusual **sidechain** torsion angles

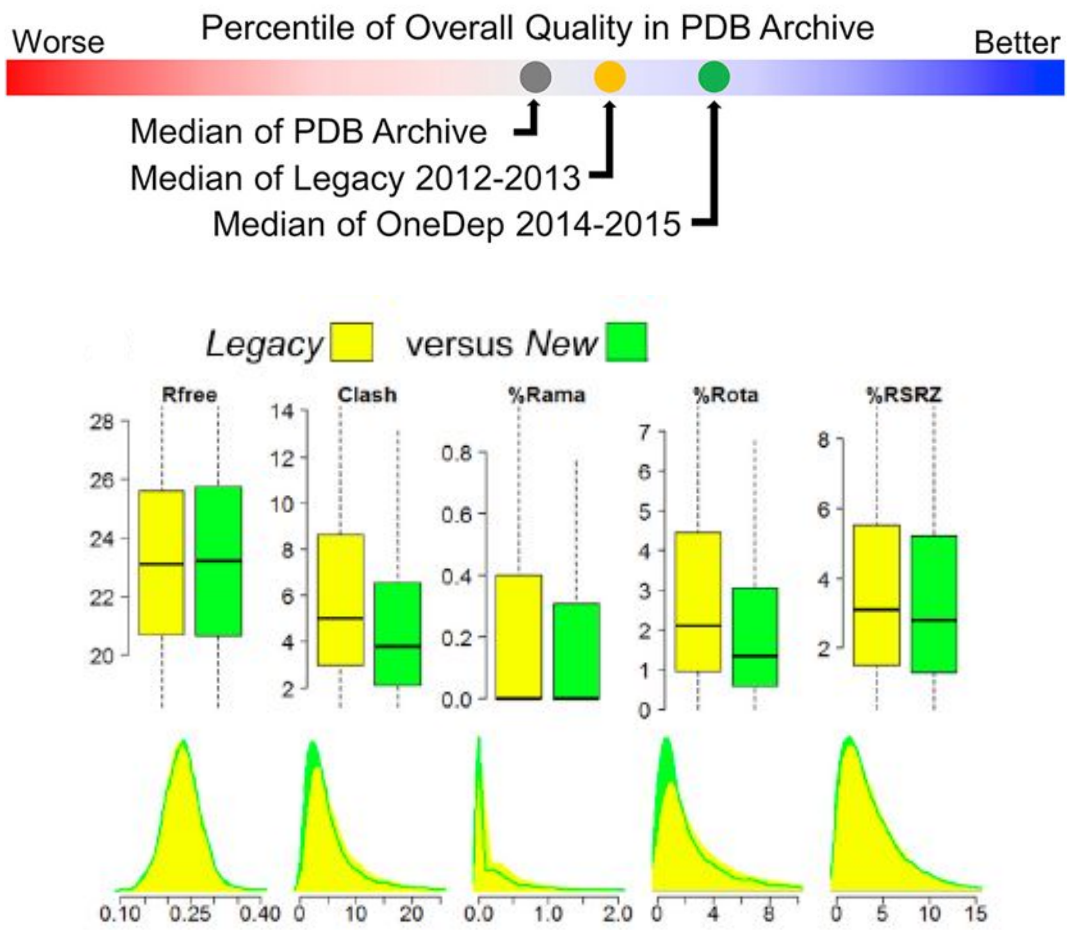
Residues lacking experimental data support



PDB ID 2HYV: Shao *et al.* J Biol Chem 281: 31689-31695

wwPDB OneDep Validation Processes in OneDep Improved PDB Structure Quality

- Structures processed with Legacy (2012-2013) *vs*. OneDep (2014-2015) deposition, annotation and validation system
- Overall Structure Quality improved after OneDep deployment
- Clashscores, % Rotamer Outliers, and % Real Space R-factor Z score (RSRZ) Outliers improved modestly



Shao *et al*. Structure 25, 458-468

RCSB PDB
PROTEIN DATA BANK

rcsb.org

Validation of Chemical Geometry for PDB Structures Determined by All Methods

Sections in the PDF report

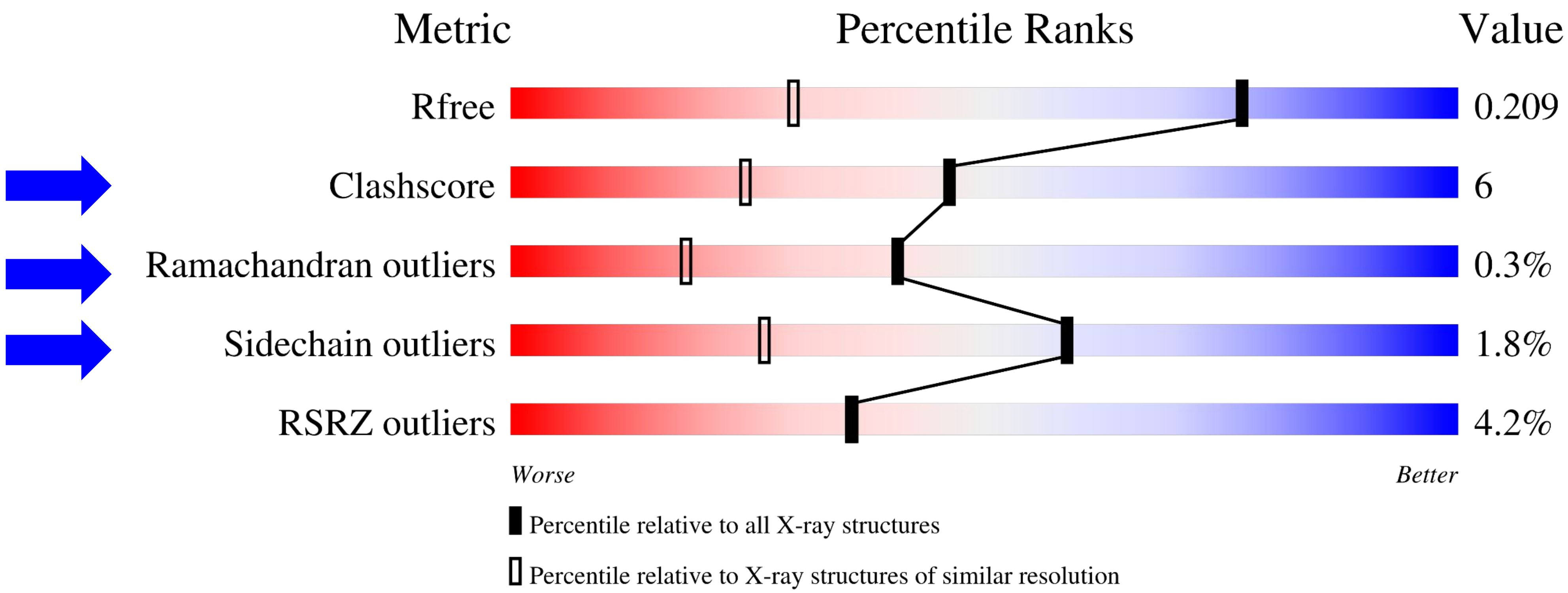
- ➢ Residue-property plots
- ➢ Model quality

Validation Report Slider for Overall Quality at a Glance (X-ray Crystallographic Example)

Atom clashes/1000 residues

Residues with unusual **main chain** torsion angles

Residues with unusual **sidechain** torsion angles



PDB ID 2HYV: Shao *et al.* J Biol Chem 281: 31689-31695

Chemical Geometry Analysis

Polymers are analysed for the following geometry issues

- Bond Lengths
- Bond Angles
- Atom Clashes
- Ramachandran Outliers
- Sidechain Conformers
- Chirality Issues
- RNA backbone quality

Molprobity + PDB software

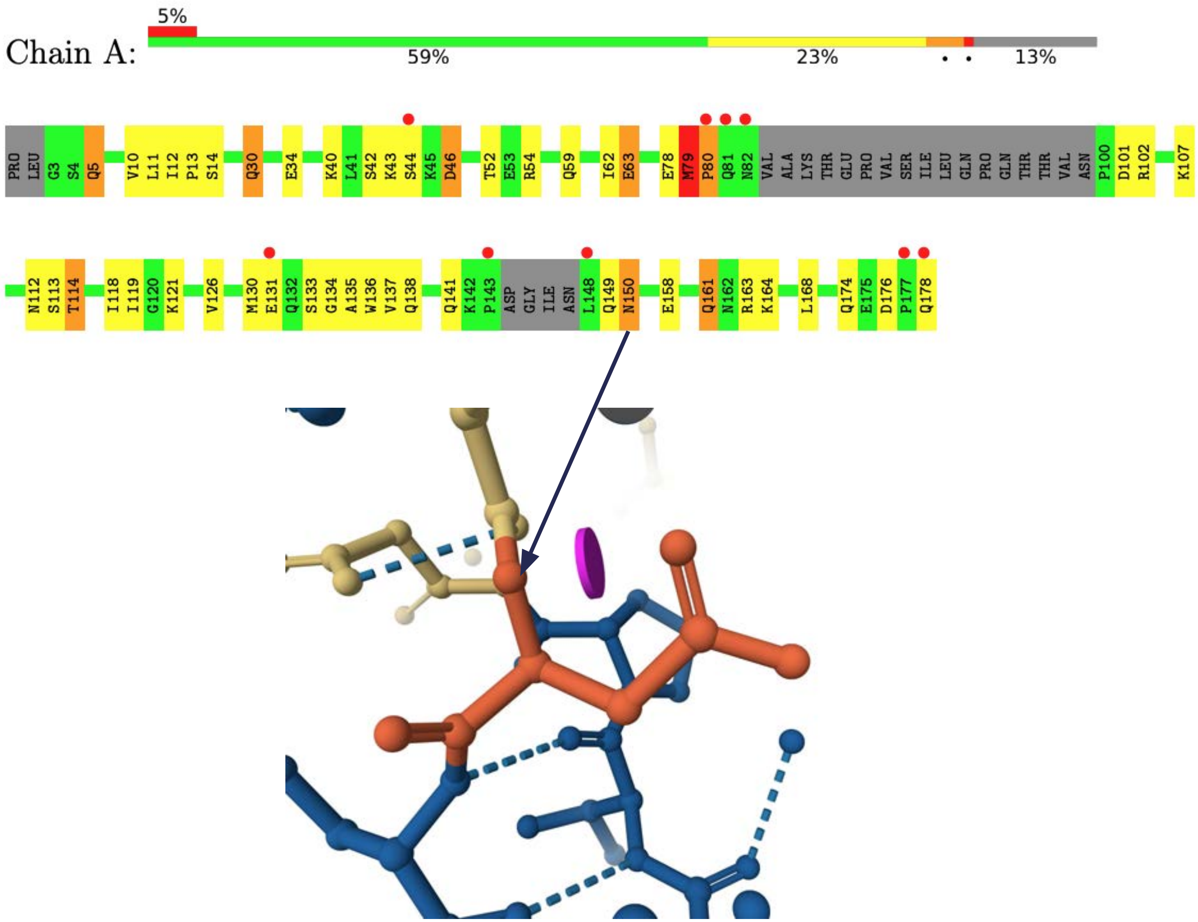
- Model quality
  - Standard Geometry
  - Too-close Contacts
  - Torsion Angles
  - Polymer Linkage Issues
- Residue-property plots

Reference for Protein: Engh, R. A. and Huber, R., Accurate bond and angle parameters for X-ray protein structure refinement, Acta Cryst. A47:392-400, 1991; Engh, R. A. and Huber, R., Structure quality and target parameters, International Tables for Crystallography (2006). Vol. F, ch. 18.3, pp. 382-392

Reference for nucleic acid: Parkinson, G.N., *et. al.*, New parameters for the refinement of nucleic acid containing structures. Acta Cryst., D52:57–64, 1996

Polymer Chemical Geometry: Overall Structure and Individual Residues

- Green, yellow, orange and red color coding indicates the fraction of residues with 0, 1, 2, ≥3 chemical geometry outliers, respectively
- Grey segment indicates residues present in the sample but not modelled
- Red dot indicates poor fit to electron density (MX, to be discussed in later slides)



PDB ID 2ANR: Teplova *et al.* Structure 19: 930-944

Ligand Chemical Geometry

- PDB validation focuses on Ligand Of Interest (LOI) designated by authors or potential LOI with MW > 250 Da
- Agreement with known chemistry in Cambridge Structural Database (CSD) of small molecule crystal structures
  - Bond Lengths: RMSZ, # |Z|>2 Bond Angles: RMSZ, # |Z|>2
  - Analyses of Chirality, Torsions, Rings
- 2D graphical depiction for geometrical metrics
  - Green: within normal range
  - Magenta: statistical outlier
  - Gray: not applicable, or insufficient chemical reference data to assess

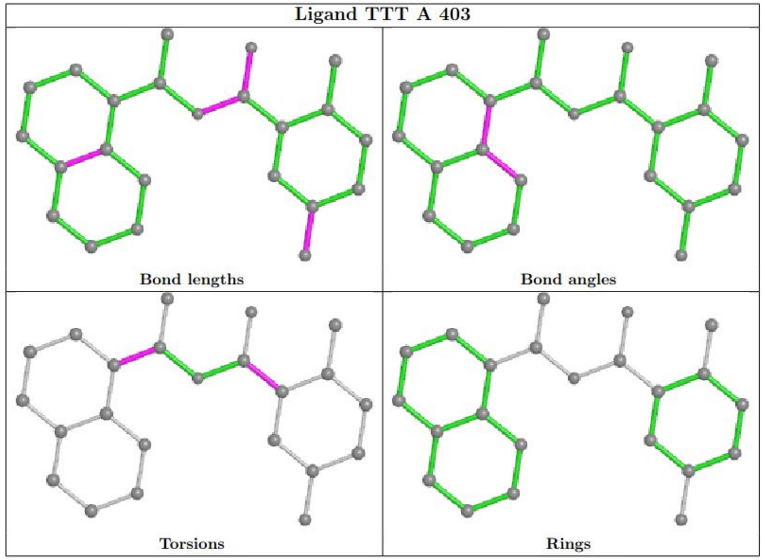
Summary

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	TTT	A	403	-	25,25,25	1.58	4 (16%)	35,35,35	1.04	1 (2%)

List of component outliers

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	403	TTT	C1-N2	5.34	1.45	1.34

Graphical depiction of outliers



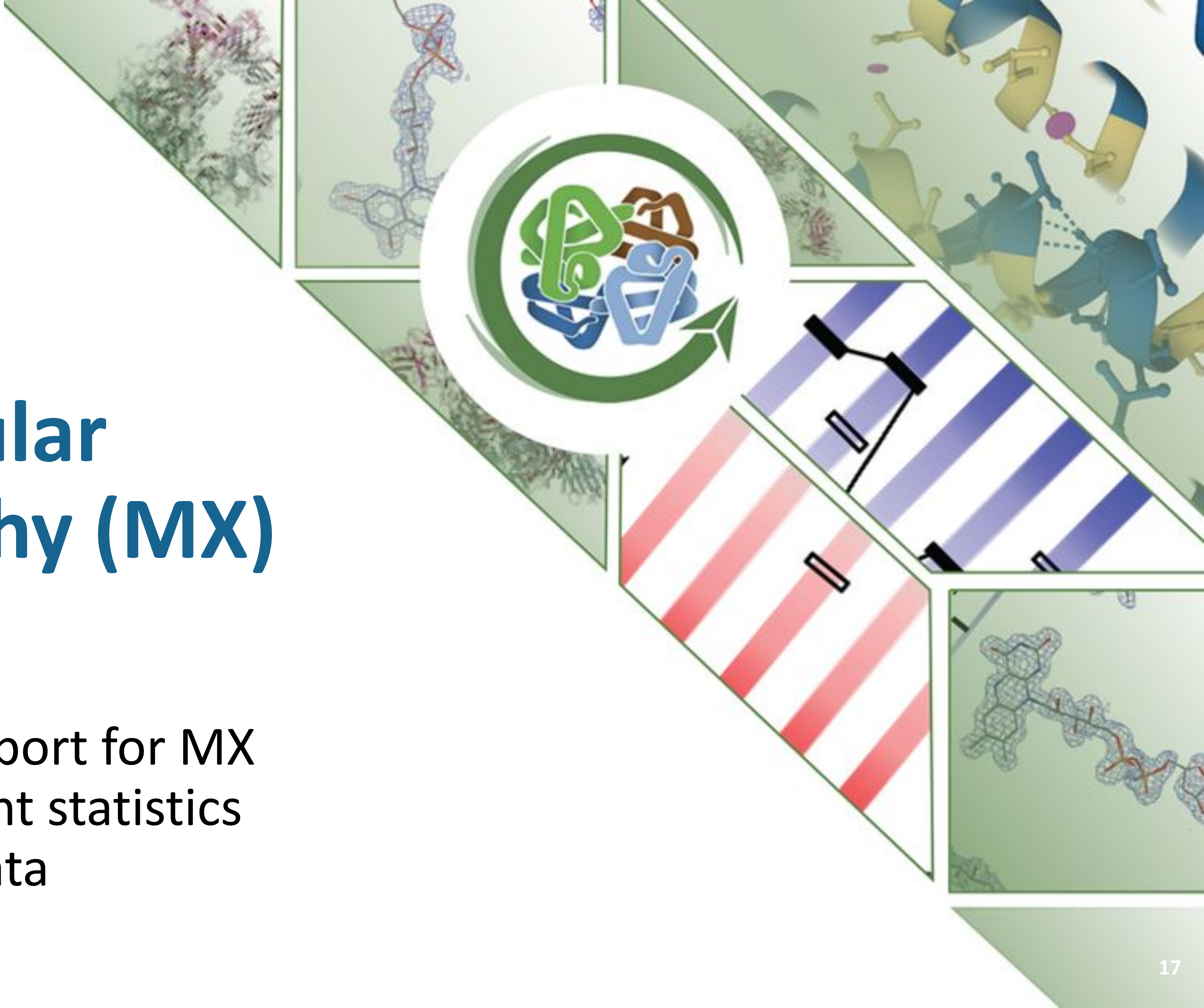
PDB ID 7SKQ: Freitas *et al.* ACS Infect Dis 8: 596-611

rcsb.org

Validation of Macromolecular Crystallography (MX) Structures

Sections in the PDF report for MX

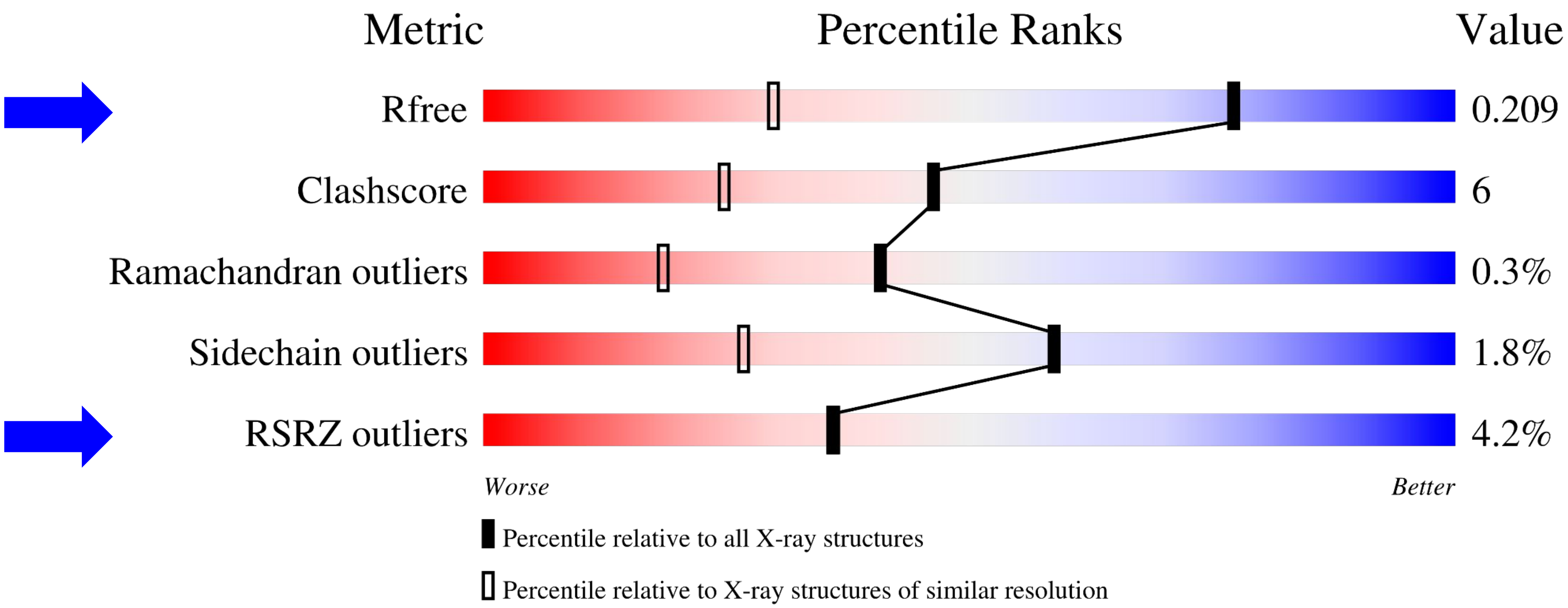
- ➢ Data and refinement statistics
- ➢ Fit of model and data



Validation Report Slider for Overall Quality at a Glance (X-ray Crystallographic Example)

How well does the overall structure agree with MX data?

Residues lacking experimental data support



PDB ID 2HYV: Shao *et al.* J Biol Chem 281: 31689-31695

Overall Experimental Data Assessment

MX experimental diffraction data validation

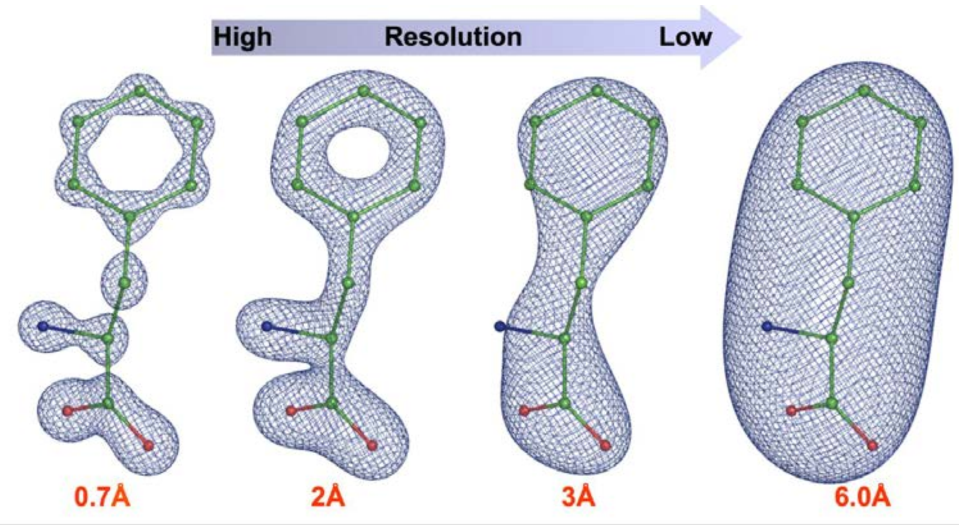
- Resolution limit (Å)
- Diffraction data completeness (%)
- Diffraction data Consistency (R_{merge})
- Signal-to-noise ($I/\sigma(I)$)

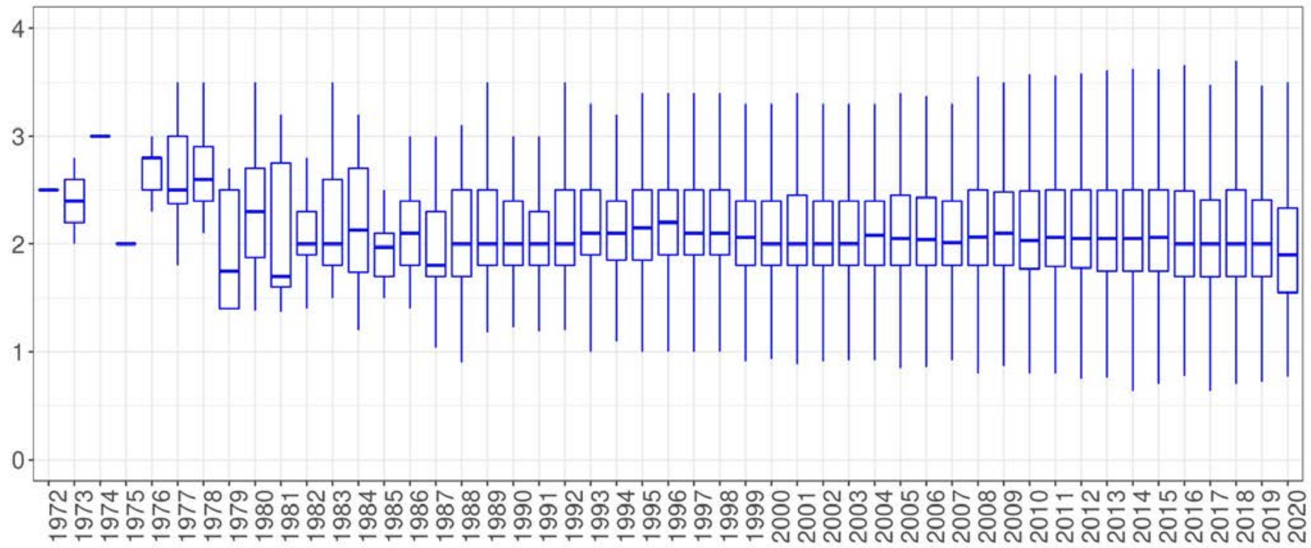
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	107.71Å 54.29Å 68.57Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 1.42 48.48 – 1.42	Depositor EDS
% Data completeness (in resolution range)	95.0 (50.00-1.42) 94.7 (48.48-1.42)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.91 (at 1.42Å)	Xtriage

PDB ID 2HYV: Shao *et al.* J Biol Chem 281: 31689-31695

Resolution: Primary MX Data Quality Metric

- The spacial limit of observed diffraction data (smaller value indicates higher resolution)
- Measures the level of details in the electron density map
- Median PDB resolution ~2Å
- No significant change in the past four decades as it depends on the crystal





Overall Structure Goodness-of-Fit Assessment

Goodness-of-fit validated on overall structure through re-calculated

- R/R_{free}
- F_o *vs.* F_c correlation

Refinement program	CNS	Depositor
R, R_{free}	0.212 , 0.223 0.200 , 0.209	Depositor DCC
R_{free} test set	6113 reflections (8.12%)	wwPDB-VP
F_o,F_c correlation	0.96	EDS
Total number of atoms	3086	wwPDB-VP
Average B, all atoms (Å^2)	22.0	wwPDB-VP

PDB ID 2HYV: Shao *et al.* J Biol Chem 281: 31689-31695

Local Polymer Goodness-of-Fit to Experimental Data Assessment

- Local goodness-of-fit to experimental data per residue assessed using Real Space R-factor Z score (RSRZ)
- RSRZ compares experimental electron density to computed electron density (calibrated against other structures at similar resolution)
- Surface, terminal, and loop residues may be of higher RSRZ due to their flexibility

RSRZ summary per chain

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	B	24/25 (96%)	-0.30	0	100	100	16, 22, 29, 45	0
2	A	153/178 (85%)	0.46	9 (5%)	22	28	14, 24, 39, 46	2 (1%)
All	All	177/203 (87%)	0.36	9 (5%)	28	35	14, 24, 39, 46	2 (1%)

List of RSRZ outliers

Mol	Chain	Res	Type	RSRZ
2	A	143	PRO	6.8
2	A	148	LEU	4.0
2	A	80	PRO	3.9
2	A	178	GLN	3.0

PDB ID 2ANR: Teplova et al. Structure 19: 930-944

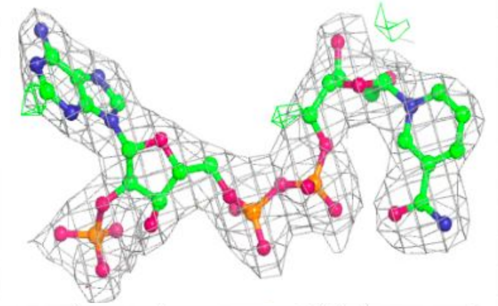
Ligand Goodness-of-Fit to Experimental Data Assessment

- Atomic coordinates agreement with experimental MX data (Electron Density map)
    - Real Space R-factor (RSR) measures difference between (A) modeled ligand and (B) experimental electron density.
    - Real Space Correlation Coefficient (RSCC) measures consistency between A and B.
- Map-Model overlay on Ligand of Interest (LOI)
- Tabular report for validation metrics

PDB ID 5ZIX with good NADP Map-Model Fit

Electron density around NAP B 401:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)

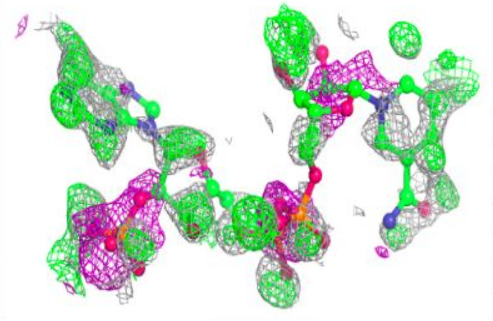


Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
3	NAP	B	401	48/48	0.96	0.14	27,42,60,76	0

PDB ID 1ZK4 with poor NADP Map-Model Fit

Electron density around NAP A 1270:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)



Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
3	NAP	A	1270	48/48	-0.06	0.67	87,96,100,100	0

5ZIX: Khanppnavar *et al.* Biochim Biophys Acta Gen Subj 1863: 1547-1559
1ZK4: Schlieben *et al.* J Mol Biol 349: 801-813

RCSB PDB
PROTEIN DATA BANK

rcsb.org

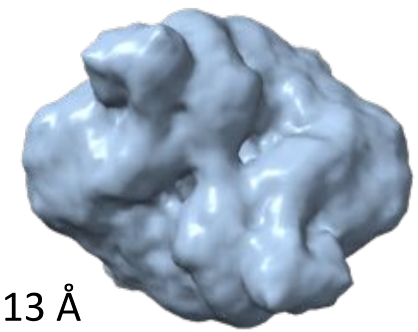
Validation of 3D Electron Microscopy (3DEM) Structures

Sections in the PDF report for 3DEM

- ➢ Experimental information
- ➢ Map visualization
- ➢ Map analysis
- ➢ Fourier-shell correlation
- ➢ Map-model fit

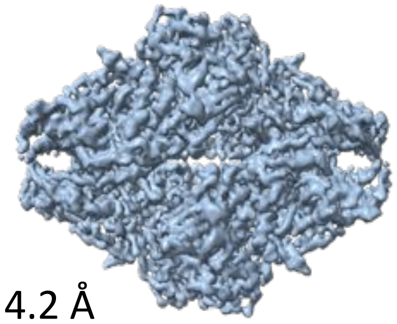
3DEM Resolution Revolution

Beta-galactosidase
EMD-2548



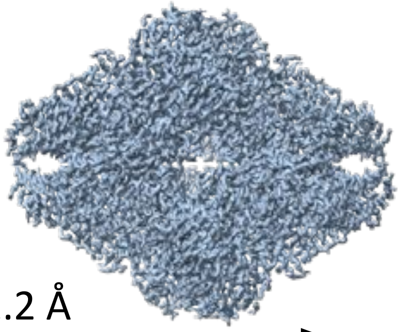
13 Å

EMD-2824



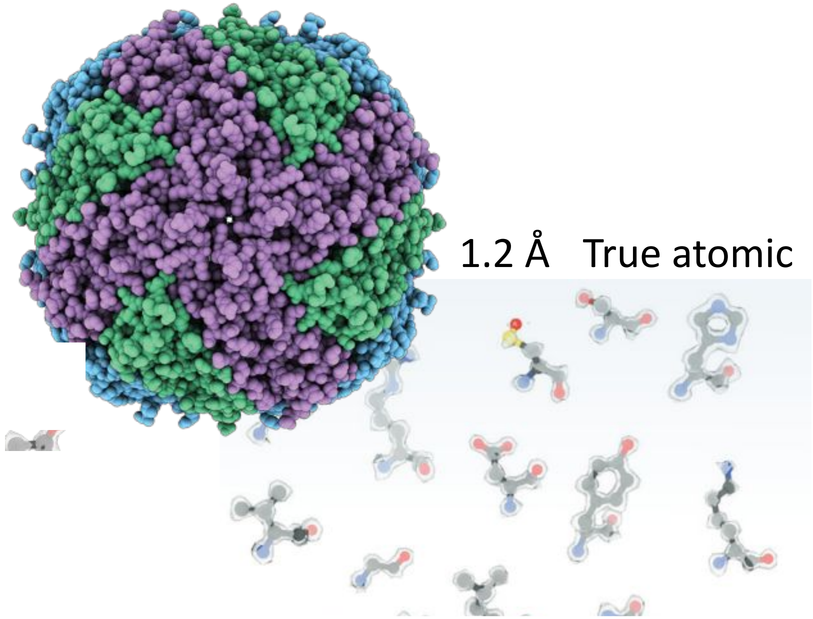
4.2 Å

EMD-2984, PDB ID 5A1A



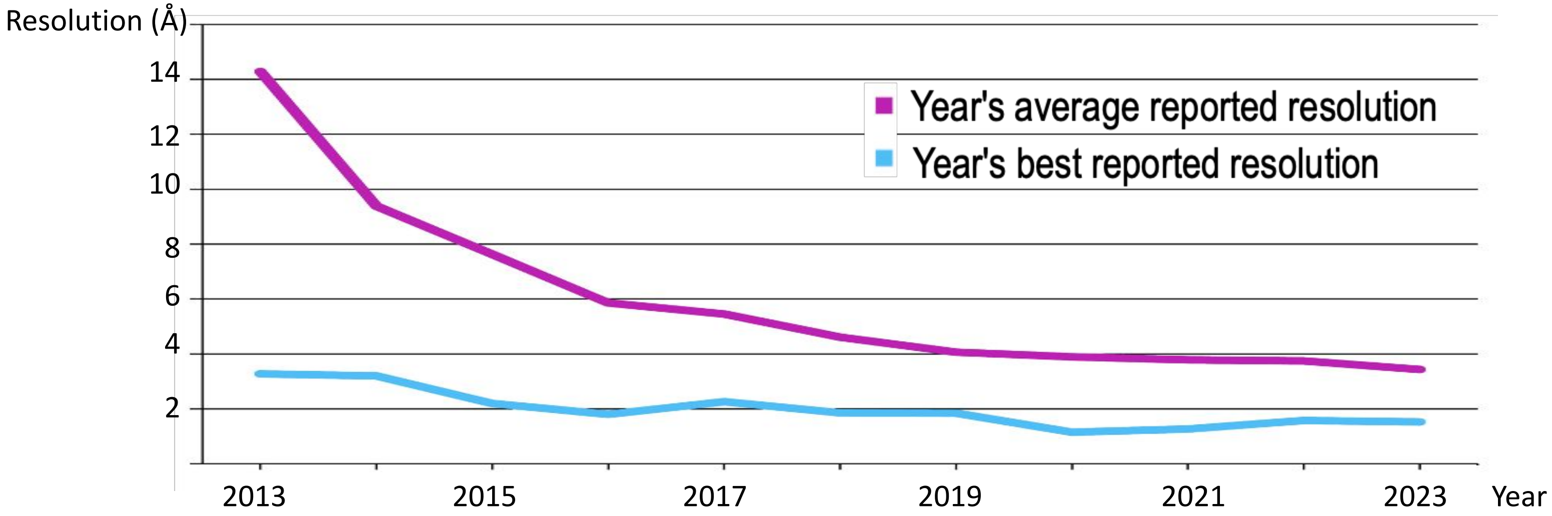
2.2 Å

Apoferritin
EMD-11103, PDB ID 6Z6U



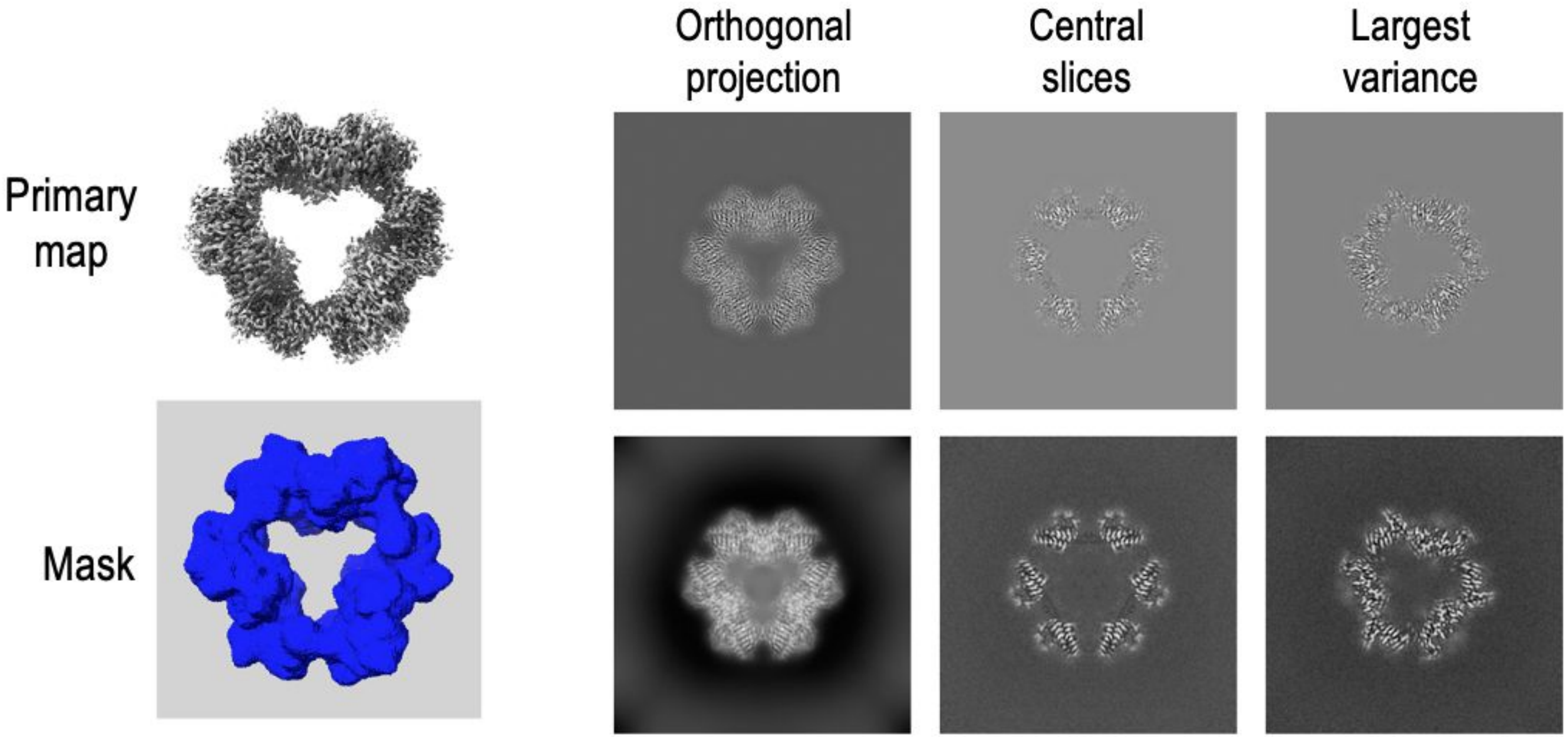
1.2 Å True atomic

Shape | Fold/Polypeptide | Side-chains Waters/Carbonyls



EMD-2548: Vinothkumar et al. Structure 22:621-627
EMD-2824: Scheres J Struct Biol 189:114-122
PDB 5A1A / EMD-2984: Bartesaghi et al. Science 348:1147-1151
PDB 6Z6U / EMD-11103: Yip et al. Nature 587:157-161

3DEM Experimental Density Map Review



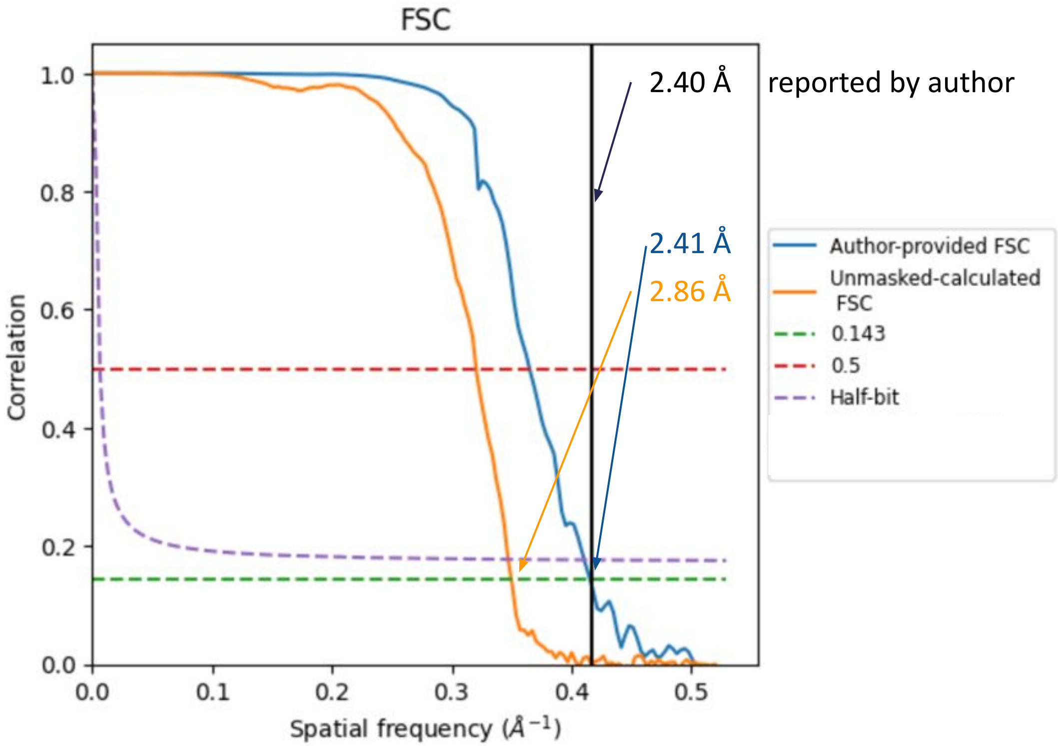
PDB ID 7ZNJ / EMD-14803: Pacheco-Fiallos *et al.* Nature 616: 828-835

3DEM Resolution Estimation: FSC Curve

- 3DEM resolution estimated by analyzing Fourier Shell Coefficient (FSC); FSC generated from 3DEM Experimental Density Maps, i.e.

  Map -> FSC -> Resolution

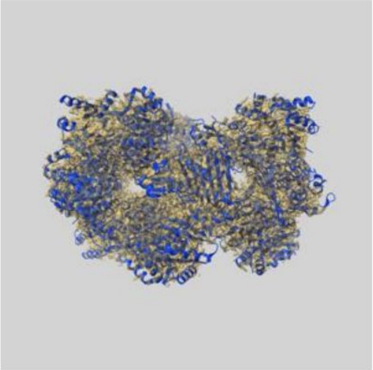
- FSC curve drops from low to high resolution; The cut off to decide resolution limit varies, but usually set at 0.143
- FSC calculation also depends on masking of the map (Caution: Not Objective!)
- wwPDB validation reports both Author-provided and OneDep-estimated resolution based on deposited FSC and maps



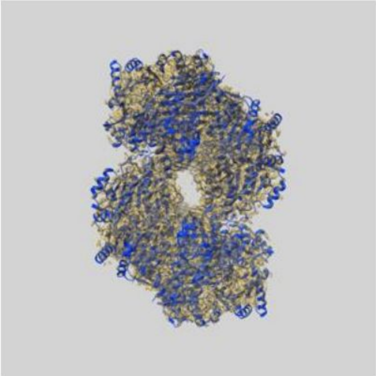
PDB ID 7ZNJ / EMD-14803: Pacheco-Fiallos et al. Nature 616: 828-835

Experimental 3DEM Map *vs*. Atomic Model: Visualization

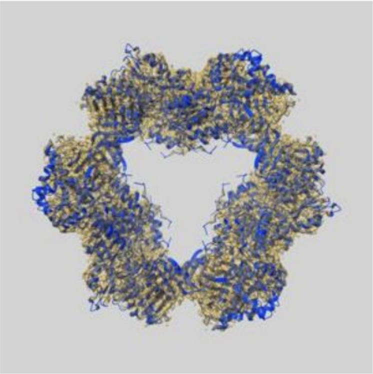
- Projection views of the Experimental 3DEM Map (yellow, at author-selected contour)
- Ribbon representation of the Atomic Coordinates (blue)
- Regions with poor fitting to the map indicate insufficient experimental support



X-axis



Y-axis

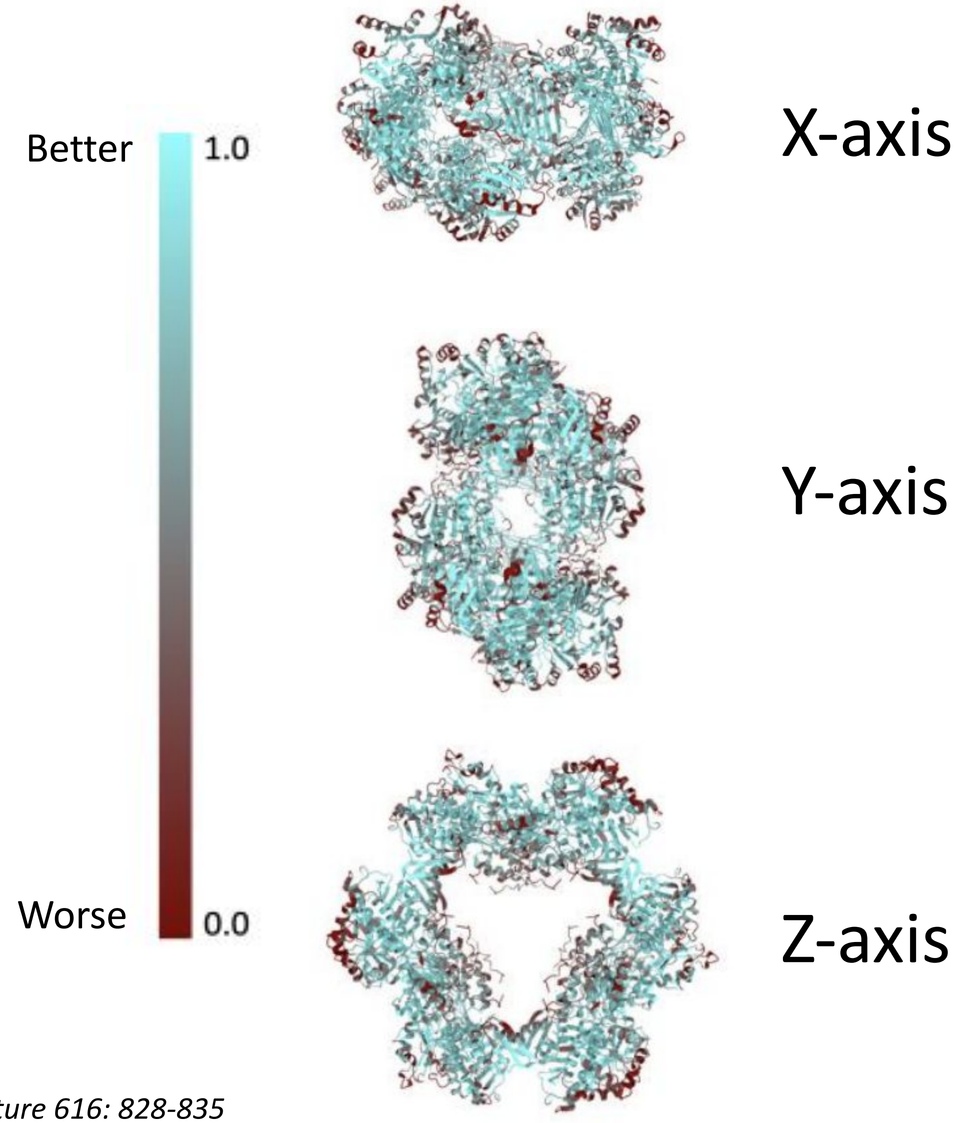


Z-axis

PDB ID 7ZNJ / EMD-14803: Pacheco-Fiallos et al. Nature 616: 828-835

Experimental 3DEM Map *vs*. Atomic Model: Atom Inclusion

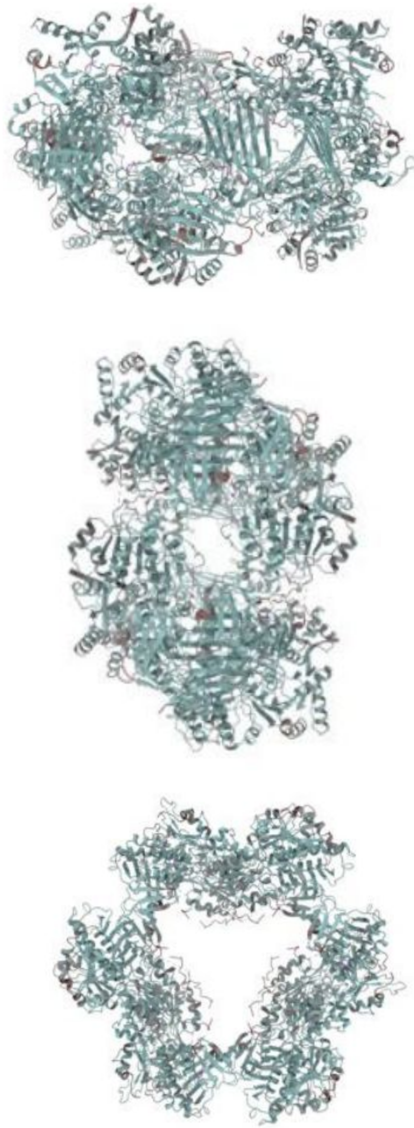
- Atom inclusion calculated for each residue in the map at the author-selected contour.
- Residues with high atom inclusion (better) are shown in cyan while low (worse) in brown. Regions with low atom inclusion lack experimental data support
- wwPDB validation report also includes average atom inclusion for each polymer chain and the overall structure



PDB ID 7ZNJ / EMD-14803: Pacheco-Fiallos et al. Nature 616: 828-835

Experimental 3DEM Map *vs*. Atomic Model: Q-Score

- Q-score calculated for each residue on atom resolvability based on 3DEM Map
  - Not subject to author-selected contour
  - Depends on resolution
- Residues with high Q-score (better) are shown in cyan while low (worse) in brown/purple. Regions with low Q-score lack experimental data support
- wwPDB validation report also includes average Q-score for each polymer chain and the overall structure

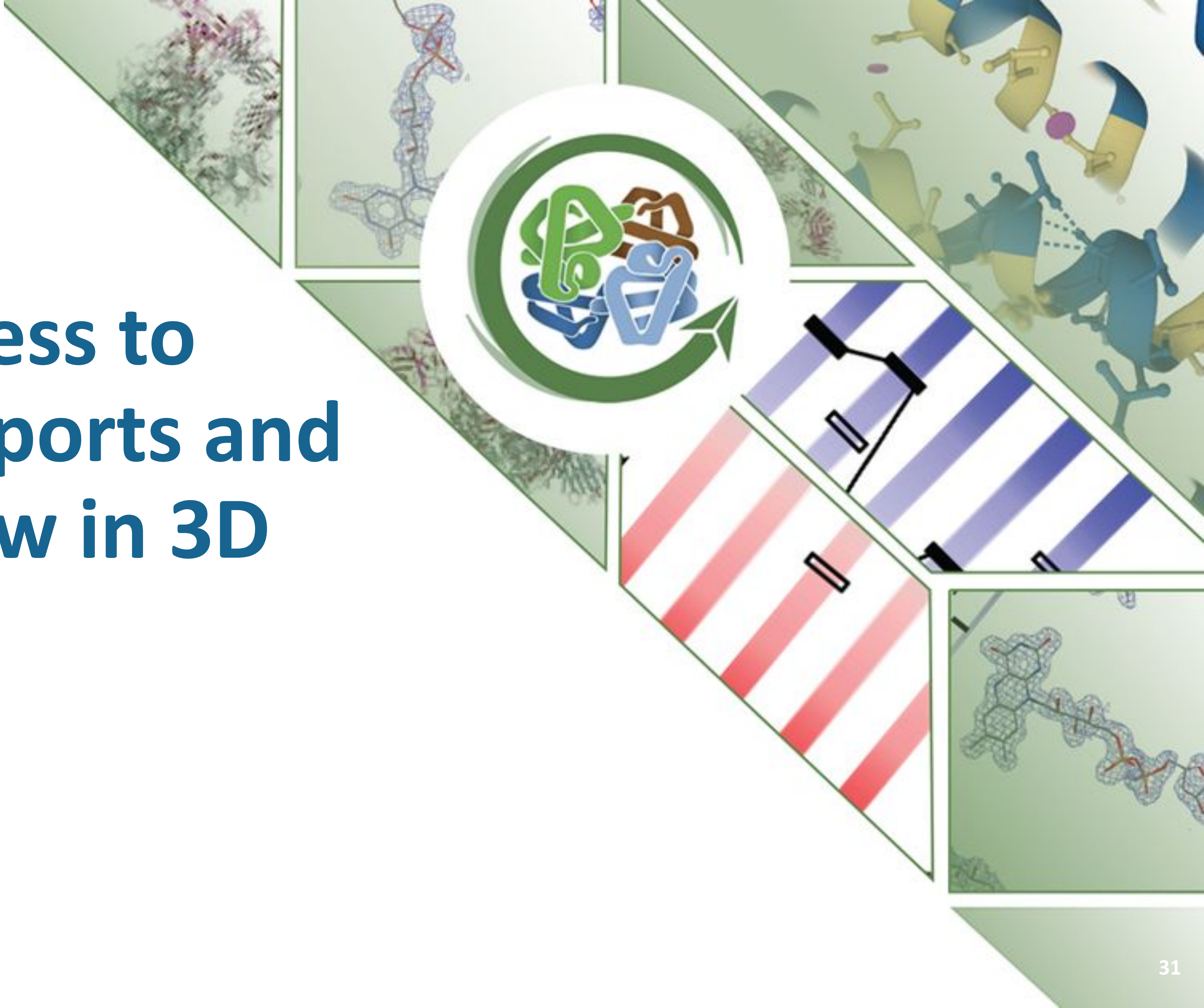


.

Q-score: Pintile et al. Nature Methods 17, 328-334.
PDB ID 7ZNJ / EMD-14803: Pacheco-Fiallos et al. Nature 616: 828-835

rcsb.org

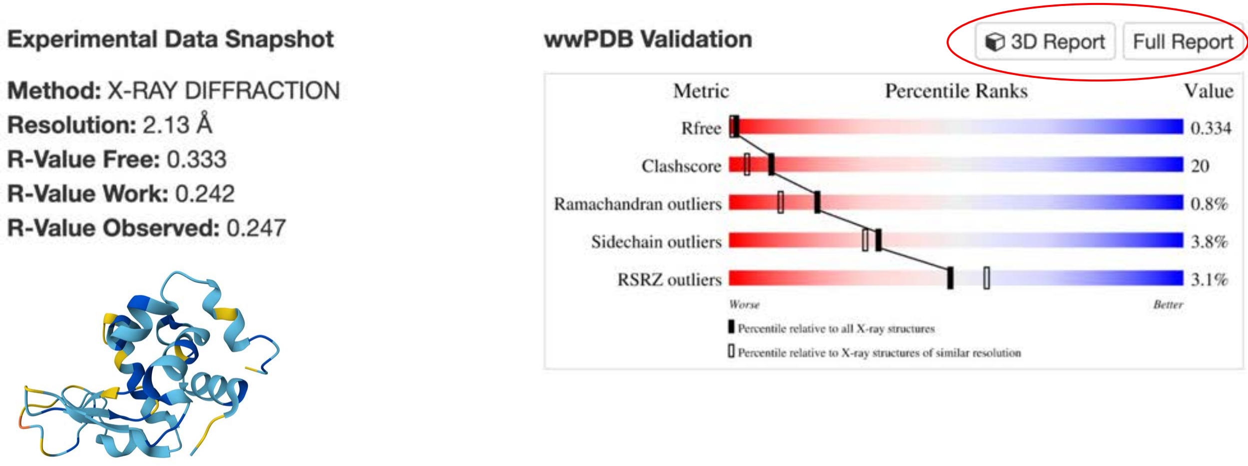
RCSB.org Access to Validation Reports and Quality Review in 3D



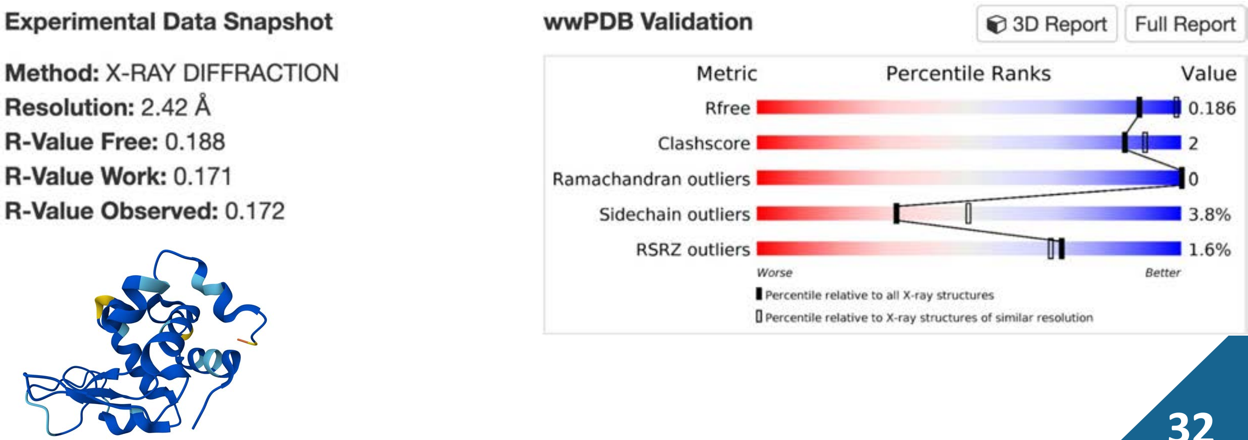
wwPDB Validation Report Access at RCSB.org

- Structure Summary Page shows wwPDB Validation Report Sliders, together with a brief summary
- Buttons above the Sliders provide
    - Full wwPDB Validation Report access/download
    - 3D Report view of the atomic structure integrated with quality assessment

PDB ID 5F81



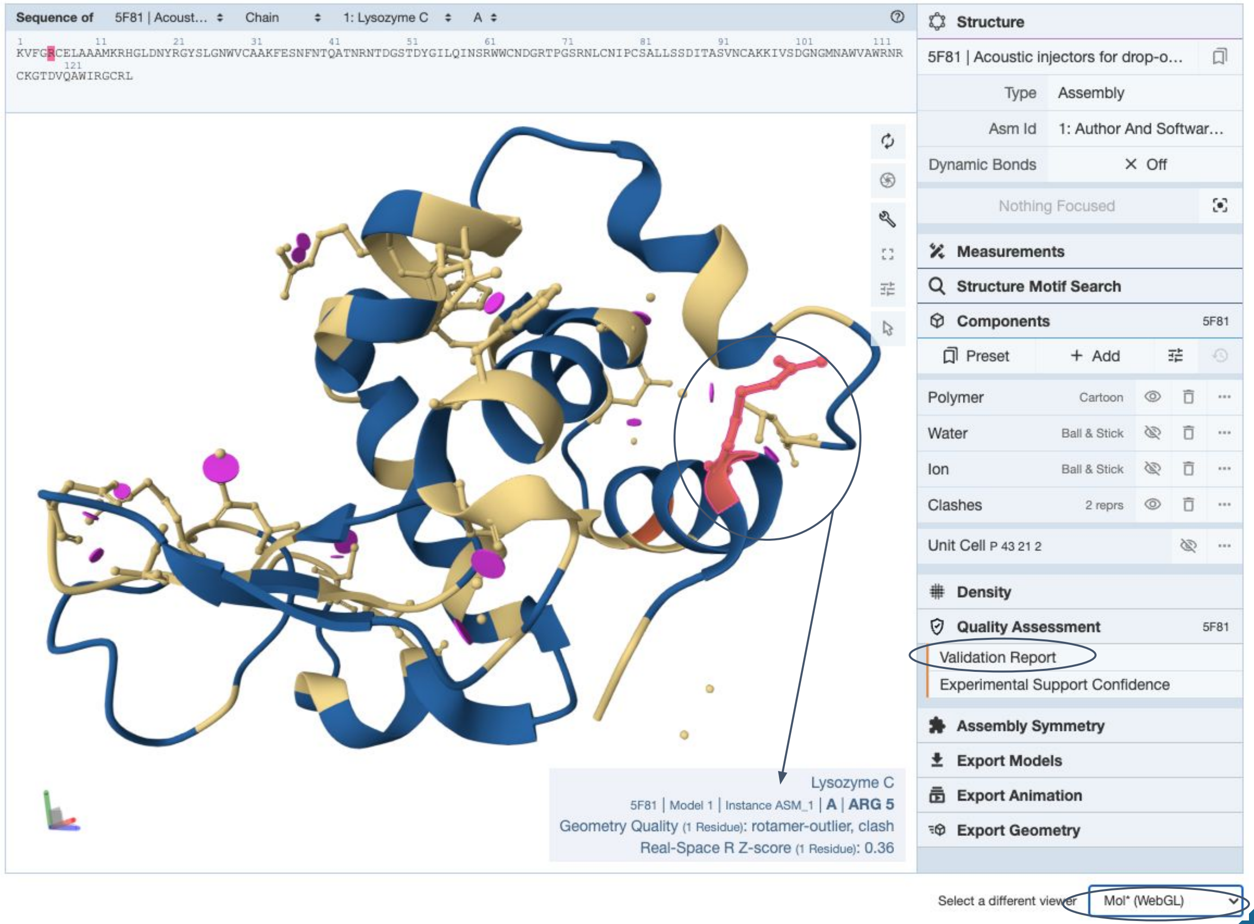
PDB ID 5HNL



5F81: Roessler *et al.* Structure 24: 631-640
5HNL: Tsukui *et al.* Acta Crystallogr D Struct Biol 72: 823-829

Mol* 3D Structure Quality View

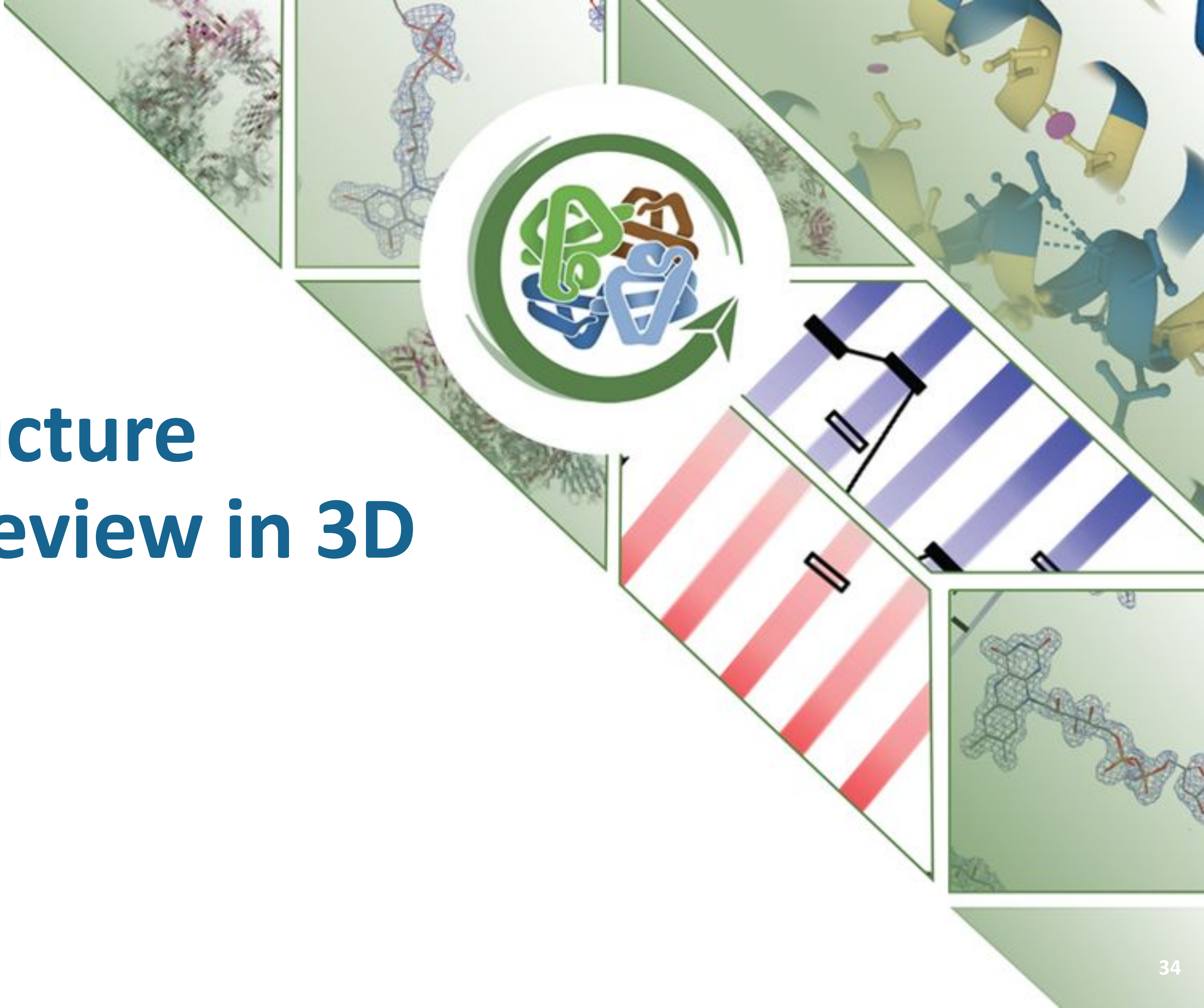
- Mol* is wwPDB Open Source 3D molecular visualization system
- Mol* at RCSB.org provides high-quality 3D views of structures with structure quality information
- wwPDB Validation Report metrics integrated into 3D views, with residues colored by quality
- Hovering cursor over individual residues or components displays quality metrics (lower right corner)



PDB ID 5F81: Roessler *et al.* Structure 24: 631-640

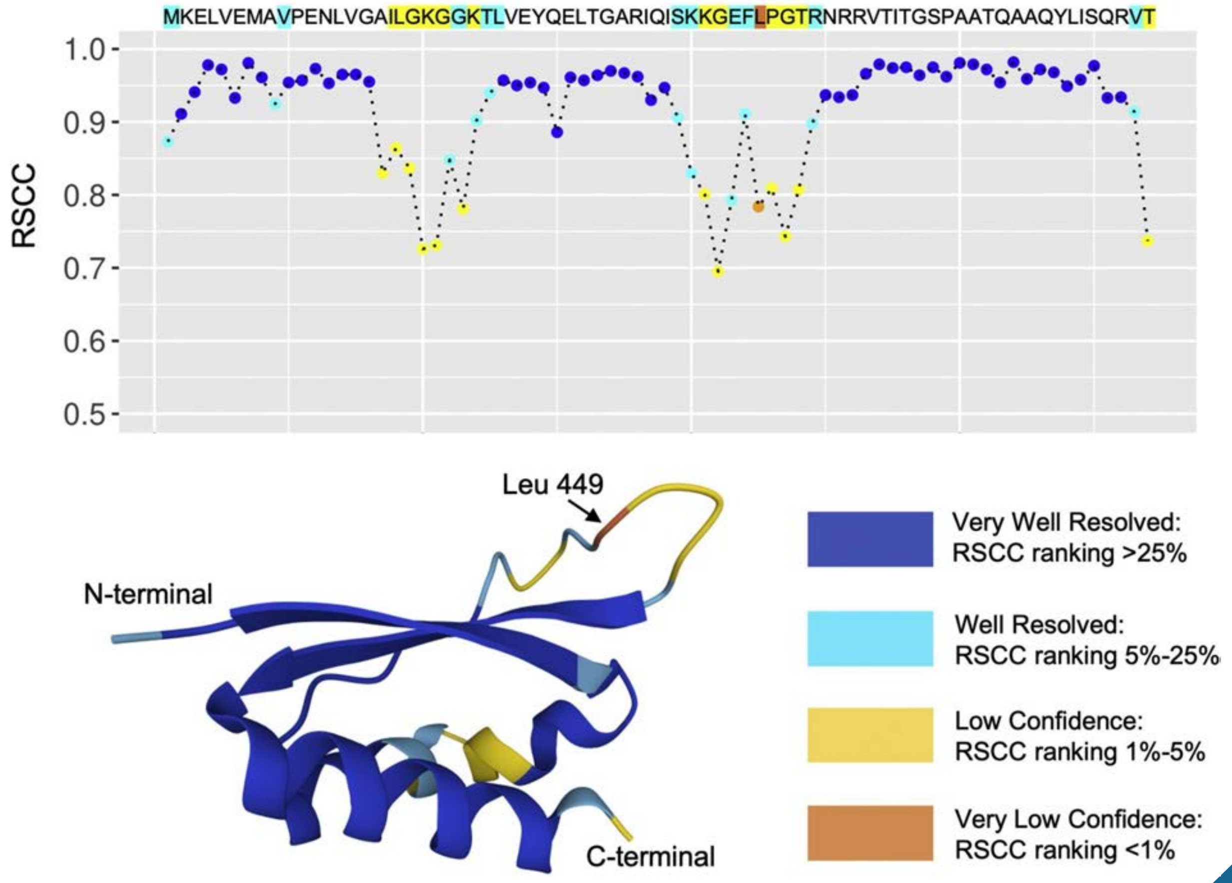
rcsb.org

RCSB.org Structure Confidence Review in 3D



RSCC-Based Structure Confidence

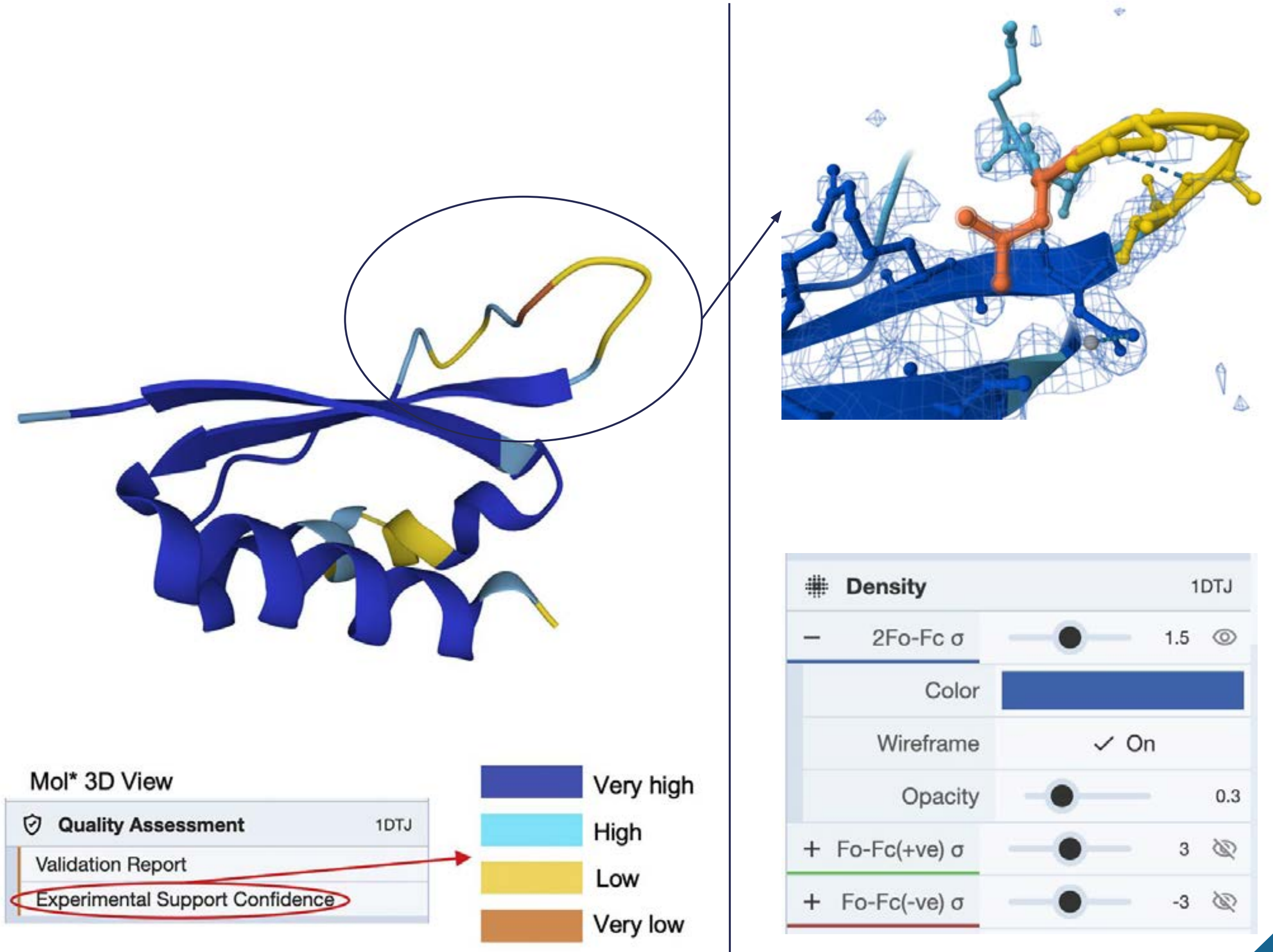
- Real Space Correlation Coefficient (RSCC) measures the agreement between residues atomic coordinates and local MX experimental data map
- Higher RSCC→well resolved→high confidence
- Lower RSCC→poorly resolved→low confidence
- RCSB.org displays color scheme for RSCC-based confidence resembling the pLDDT local confidence score of AlphaFold2 Computed Structure Models



Shao *et al.* Structure 30, 1385-1394.

RCSB.org Mol* View of Structure Confidence and Electron Density Overlay

- Mol* coloring integrated with RSCC-based structure confidence metrics
- Supporting comprehensive quality reviews by RCSB.org Users, together with electron density overlay



PDB ID 1DTJ: Lewis et al. Structure 7: 191-203

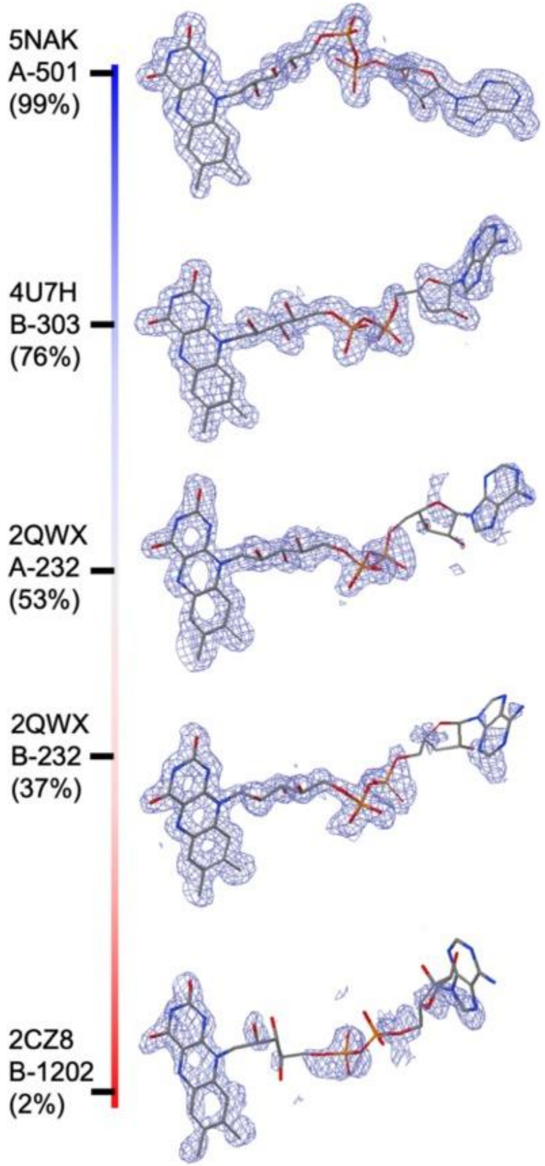
RCSB PDB PROTEIN DATA BANK

rcsb.org

RCSB.org Interactive Ligand Quality Review

RCSB.org Value-Added Ligand Quality Metrics: Principal Component Analysis

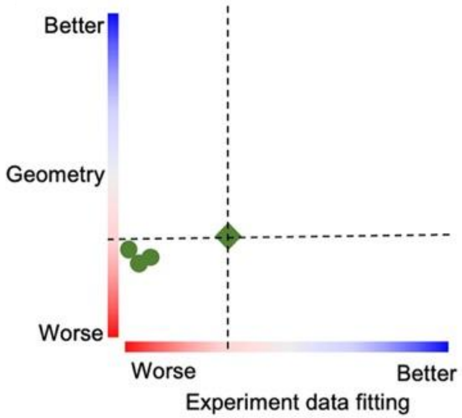
- Principal Component Analysis (PCA) of ligand quality in PDB
  - PC1-fit (1st principal component) percentile ranking of agreement of atomic coordinates with MX experimental density map
  - PC1-geo percentile ranking of agreement of atomic coordinates with known chemical geometry
- PC1-fit and PC1-geo 2D display



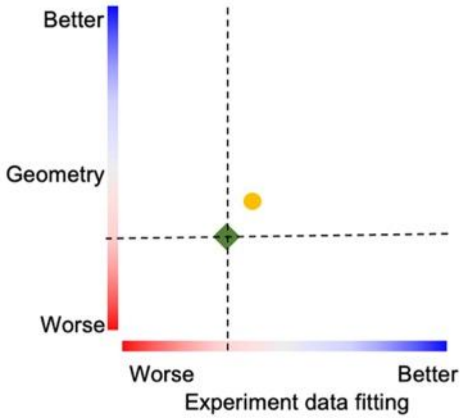
Ranking FAD Ligands
using PC1-fit
Map-Coordinates Fitting

Shao *et al.* Structure 30, 252-262.

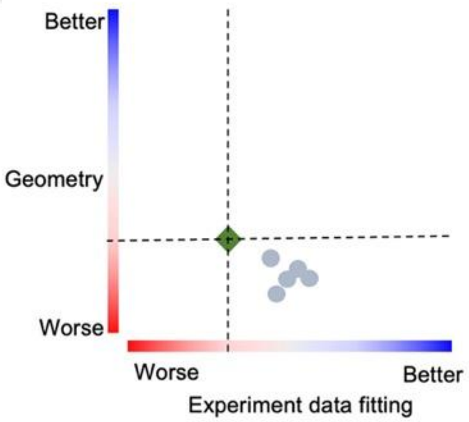
RCSB.org Value-Added Ligand Quality Metrics: 2D Ligand Ranking Goodness-of-fit/Geometry



All Instances in 6WJC



Best fitted Instance in 6WJC vs. best fitted PDB Instances with the same molecular target



Best fitted Instance in 6WJC vs. best fitted PDB Instances

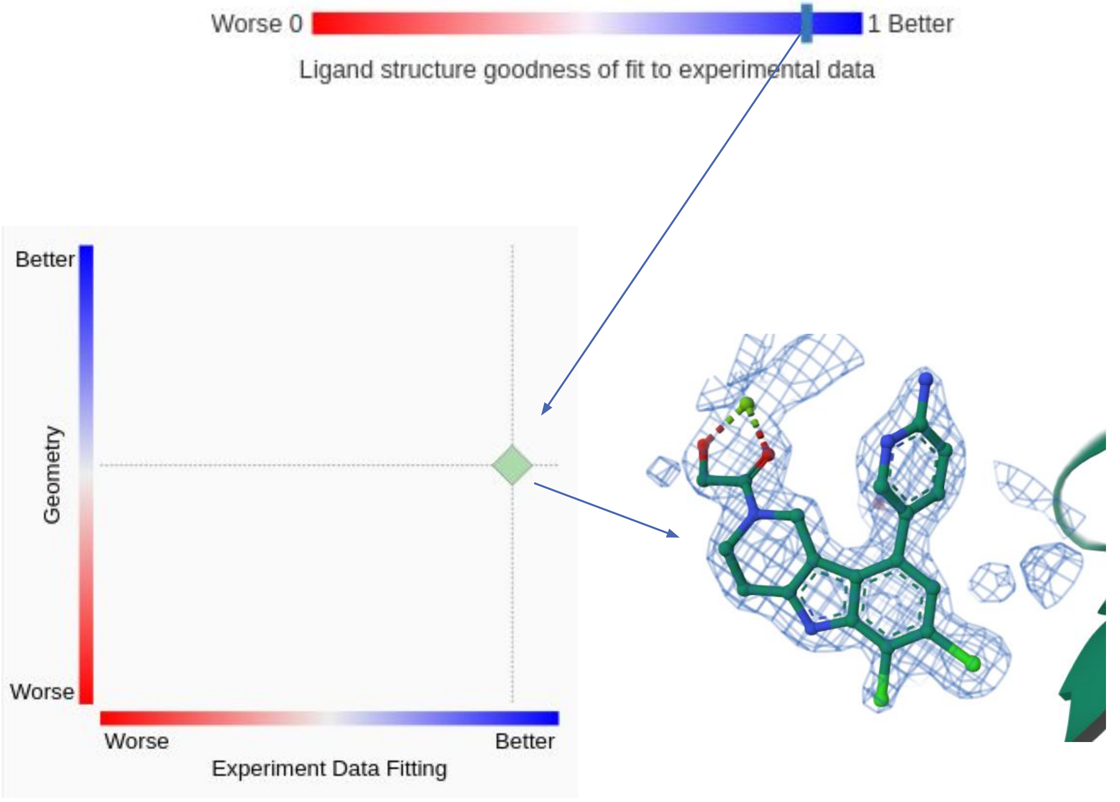
Interactive between graph and table

Identifier	Composite ranking of goodness-of-fit	Composite ranking of geometry	Real space R factor	Real space correlation coefficient	RMSZ-bond-length	RMSZ-bond-angle	Outliers of bond length	Outliers of bond angle	Atomic clashes	Stereo-chemical errors	Model completeness	Average occupancy
6WJC_Y01_A_502	**31.4%**	**35.7%**	**0.227**	**0.892**	**1.31**	**1.34**	**5**	**8**	**0**	**0**	**100%**	**1**
6WJC_Y01_A_503	3.5%	31.5%	0.518	0.871	1.28	1.58	4	12	0	0	100%	1
6WJC_Y01_A_504	2.3%	31.0%	0.422	0.714	1.27	1.62	4	9	3	0	100%	1
6WJC_Y01_A_505	0.4%	33.5%	0.727	0.761	1.32	1.44	4	10	3	0	100%	1
5CXV_Y01_A_502	37.1%	43.2%	0.207	0.899	1.02	1.24	1	5	0	0	100%	1
2Y00_Y01_B_401	51.2%	25.7%	0.195	0.946	1.24	1.94	3	13	1	0	100%	1
2Y01_Y01_A_401	50.3%	27.7%	0.208	0.956	1.23	1.84	3	15	1	0	100%	1
2Y03_Y01_B_401	49.3%	25.3%	0.213	0.957	1.21	1.99	3	13	2	0	100%	1
3ZPR_Y01_A_401	48.5%	17.8%	0.215	0.956	1.24	2.46	3	13	1	0	100%	1
4XNV_Y01_A_1103	47.4%	28.7%	0.175	0.911	1.9	1.15	9	3	0	0	100%	1

https://www.rcsb.org/ligand-validation/6WJC/Y01

RCSB.org Value-Added Ligand Quality Metrics: Interactive 3D Experimental Density Map (Mol*)

- Structure Summary Page ligand quality slider shown below overall structure quality sliders
- Vertical bar representing ligand quality ranking hyperlinked to the 2D ligand quality measures
- Interactive 2D display connects to 3D display of experimental density map/atomic coordinates



Ligand JUJ in PDB ID 7FUR

Groebke-Zbinden et al. DOI:10.2210/pdb7FUR/pdb

Available Resources

- [wwPDB validation documentation]
  - [wwPDB validation report FAQ]
  - [wwPDB Validation Task Forces]
- [RCSB PDB user guide documentation]
  - [How to assess PDB structure overall quality]
  - [How to assess ligand structure quality]
- [RCSB PDB Training Courses at PDB-101]
  - [Mol* Webinar Recording]

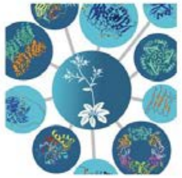
Visualize Biomolecular Structures with Mol*: From Atoms to Movies

Teaching Enzymology with the Protein Data Bank: From Pandemic to Paxlovid

Leveraging RCSB PDB APIs for Bioinformatics Analyses and Machine Learning

Use PDB data to their full extent: Understanding PDBx/mmCIF

Exploring Computed Structure Models on RCSB.org



Using KBase to access PDB Structures and Computed Structure Models

Python Scripting for Biochemistry & Molecular Biology | Part 2

Python Scripting for Biochemistry & Molecular Biology | Part 1

PDB-101 Training Courses: Videos and related materials

RCSB.org Tools for Quality Assessment

- Validation reports for detailed review
  - PDF report for reading
  - CIF/XML report for programmatic parsing
- Structure Summary Page: Experimental data snapshot; Experiment tab
- Sliders for quick review
  - Overall quality slider
  - Ligand of Interest (LOI) quality slider
- Mol* 3D visualization
  - By *validation report* feature: simplified review of chemical geometry
  - By *experimental support confidence*: simplified review of goodness-of-fit (MX)
  - Model-map overlay: expert review of goodness-of-fit
- Dedicated RCSB.org ligand quality page (MX)
  - 2D ligand quality graph: simplified review on ligand quality
  - Interactive 3D model-map overlay on ligands: expert review
  - Comparison among structures with the same ligand: select better ligand structure

Summary: Indication of Better Structure Quality

Indicator	Experimental Method		Value for Better Quality
	MX	3DEM	
Resolution	x	x	smaller
R/Rfree	x		smaller
clashscore	x	x	smaller
# Ramachandran outliers	x	x	smaller
# sidechain outliers	x	x	smaller
# RSRZ outliers	x		smaller
RSR on residue/ligand	x		smaller
RSCC on residue/ligand	x		larger
Q-score		x	larger
Atom inclusion		x	larger

Thank you for joining us!

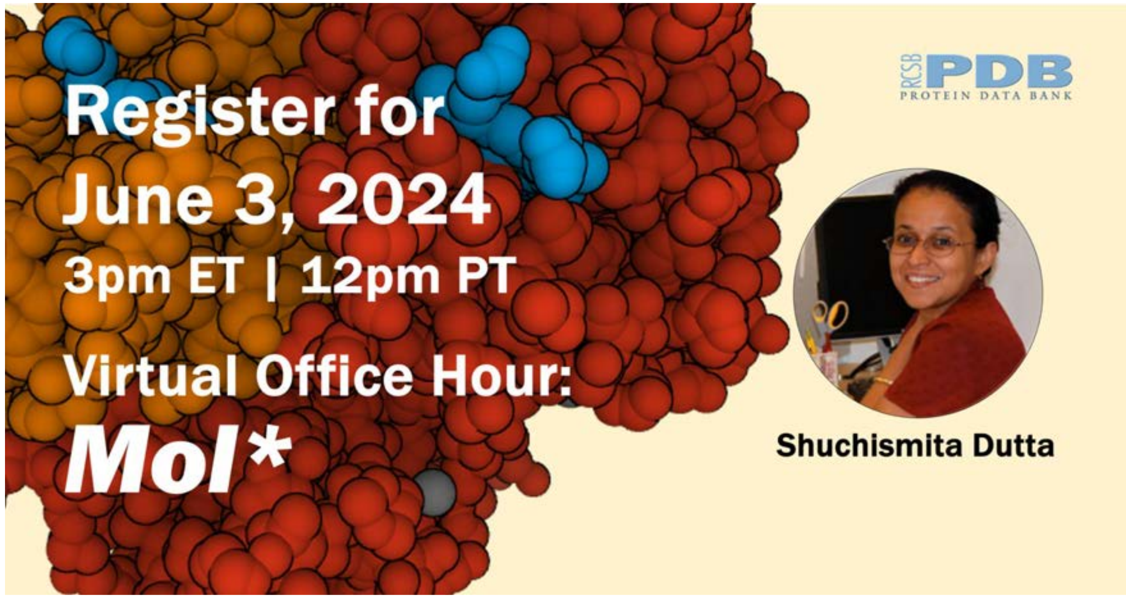
Exit Survey

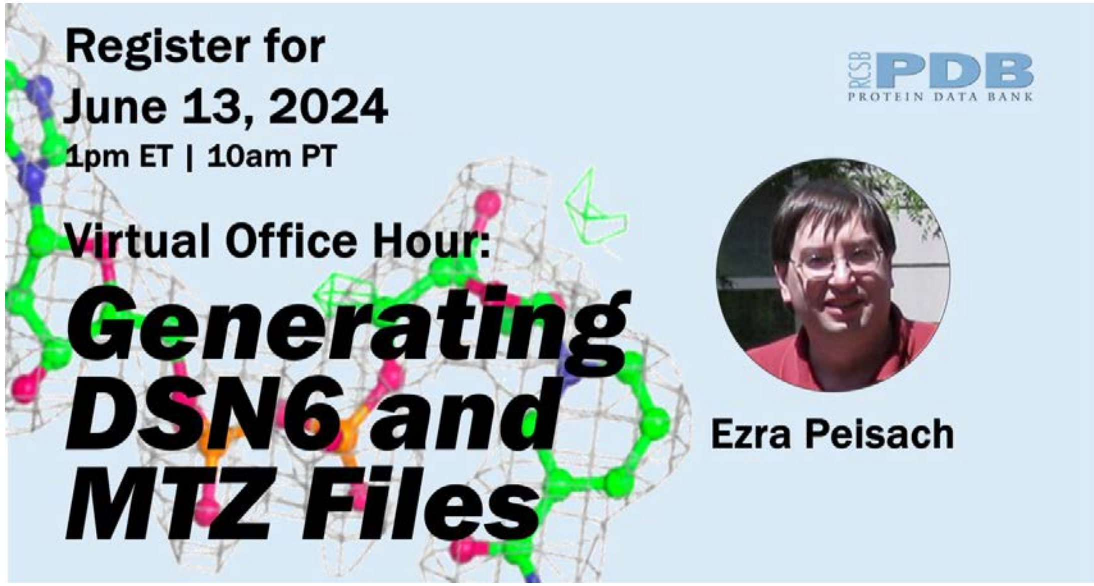
Please take this Exit Survey to help us plan future events and webinars by **Tuesday May 21**

Participation Certificate

You **MUST** complete the Exit Survey in order to receive a participation certificate.

Register at RCSB.org for Upcoming Events





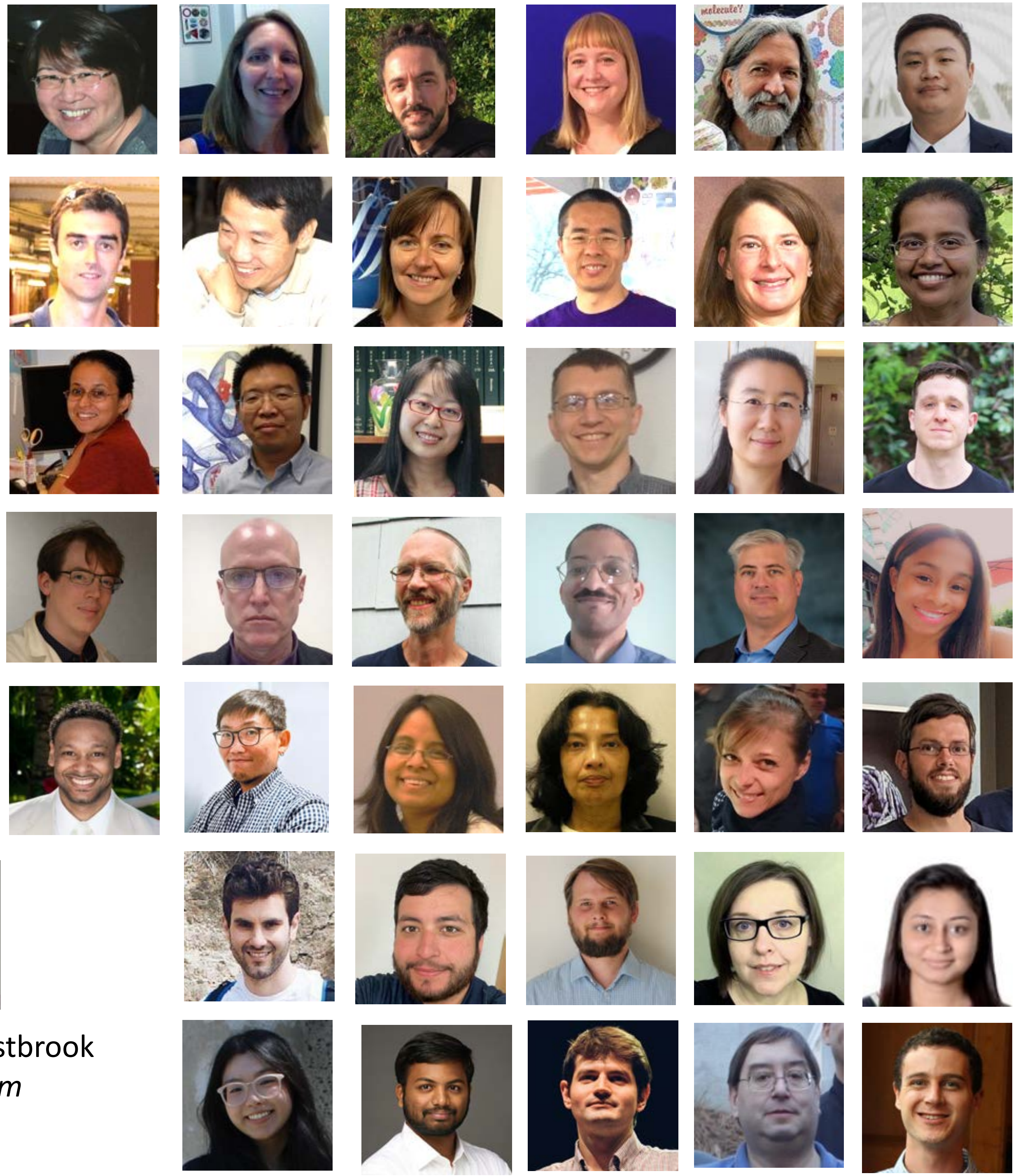
June 3: Quick tips on how to use Mol* in the pairwise structure alignment tool.

June 13: Learn about the impact of the EDMAPS.rcsb.org shut down on DSN6-formatted map files.

RCSB PDB Team

RCSB.ORG

info@rcsb.org



Core Operations Funding

US National Science Foundation (DBI-2321666), National Institute of General Medical Sciences, National Institute of Allergy and Infectious Disease, and National Cancer Institute (NIH R01GM133198), and the US Department of Energy (DE-SC0019749)

Management

RUTGERS THE STATE UNIVERSITY OF NEW JERSEY

UC San Diego

SDSC SAN DIEGO SUPERCOMPUTER CENTER

UCSF University of California San Francisco

Member of the Worldwide Protein Data Bank (wwPDB; **wwpdb.org**)

John D. Westbrook
In memoriam
1957-2021

Follow us

Questions?