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
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, 1395-1414
  • 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
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.)

Atomic Coordinates
Universal for All Methods

Experimental Data
MX: Diffraction data
EM: Maps, Half maps, Mask, FSC
NMR: Chemical Shifts, Restraints
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

The following experimental techniques were used to determine the structure: X-RAY DIFFRACTION
The reported resolution of this entry is 2.59 Å.

Per centile scores (ranging from 0 to 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.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Percentile Rank</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall quality</td>
<td>overall</td>
<td></td>
</tr>
<tr>
<td>Chain</td>
<td>chain</td>
<td></td>
</tr>
<tr>
<td>Ramachandran outliers</td>
<td>Ramachandran outliers</td>
<td></td>
</tr>
<tr>
<td>Sidechain outliers</td>
<td>Sidechain outliers</td>
<td></td>
</tr>
<tr>
<td>RSSE outliers</td>
<td>RSSE outliers</td>
<td></td>
</tr>
</tbody>
</table>

The table below summarizes 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 remain outliers for >3, 2, 1 and 0 types of geometric quality criteria, respectively. A grey segment represents the fraction of residues that are not analysed. 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.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Length</th>
<th>Quality of chain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>1,50</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5%</td>
<td>25%</td>
</tr>
<tr>
<td>2</td>
<td>C</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5%</td>
<td>1%</td>
</tr>
</tbody>
</table>

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

---

<table>
<thead>
<tr>
<th>Metric</th>
<th>Percentile Ranks</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rfree</td>
<td></td>
<td>0.209</td>
</tr>
<tr>
<td>Clashscore</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Ramachandran outliers</td>
<td></td>
<td>0.3%</td>
</tr>
<tr>
<td>Sidechain outliers</td>
<td></td>
<td>1.8%</td>
</tr>
<tr>
<td>RSRZ outliers</td>
<td></td>
<td>4.2%</td>
</tr>
</tbody>
</table>

Worse

Better

☑ Percentile relative to all X-ray structures
☑ Percentile relative to X-ray structures of similar resolution

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
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

<table>
<thead>
<tr>
<th>Metric</th>
<th>Percentile Ranks</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rfree</td>
<td></td>
<td>0.209</td>
</tr>
<tr>
<td>Clashscore</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Ramachandran outliers</td>
<td></td>
<td>0.3%</td>
</tr>
<tr>
<td>Sidechain outliers</td>
<td></td>
<td>1.8%</td>
</tr>
<tr>
<td>RSRZ outliers</td>
<td></td>
<td>4.2%</td>
</tr>
</tbody>
</table>
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

\[ \text{Molprobity + PDB software} \]

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


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

<table>
<thead>
<tr>
<th>Mol</th>
<th>Type</th>
<th>Chain</th>
<th>Res</th>
<th>Link</th>
<th>Bond lengths</th>
<th>Bond angles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Counts</td>
<td>RMSZ</td>
</tr>
<tr>
<td>3</td>
<td>TTT</td>
<td>A</td>
<td>403</td>
<td>-</td>
<td>25,25,25</td>
<td>1.58</td>
</tr>
</tbody>
</table>

List of component outliers

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Res</th>
<th>Type</th>
<th>Atoms</th>
<th>Z</th>
<th>Observed(A)</th>
<th>Ideal(A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>A</td>
<td>403</td>
<td>TTT</td>
<td>C1-N2</td>
<td>5.31</td>
<td>1.45</td>
<td>1.34</td>
</tr>
</tbody>
</table>

Graphical depiction of outliers

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

<table>
<thead>
<tr>
<th>Metric</th>
<th>Percentile Ranks</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rfree</td>
<td></td>
<td>0.209</td>
</tr>
<tr>
<td>Clashscore</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Ramachandran outliers</td>
<td></td>
<td>0.3%</td>
</tr>
<tr>
<td>Sidechain outliers</td>
<td></td>
<td>1.8%</td>
</tr>
<tr>
<td>RSRZ outliers</td>
<td></td>
<td>4.2%</td>
</tr>
</tbody>
</table>

How well does the overall structure agree with MX data?

Residues lacking experimental data support

Overall Experimental Data Assessment

MX experimental diffraction data validation

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

<table>
<thead>
<tr>
<th>Space group</th>
<th>P 21 2 2</th>
<th>Depositor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell constants</td>
<td>107.7Å  54.29Å  68.57Å</td>
<td>Depositor</td>
</tr>
<tr>
<td>a, b, c, α, β, γ</td>
<td>90.0°  90.0°  90.0°</td>
<td>Depositor</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50.00 – 1.42</td>
<td>Depositor EDS</td>
</tr>
<tr>
<td>48.48 – 1.42</td>
<td>Depositor EDS</td>
<td></td>
</tr>
<tr>
<td>% Data completeness</td>
<td>95.0 (50.00-1.42)</td>
<td>Depositor EDS</td>
</tr>
<tr>
<td>(in resolution range)</td>
<td>94.7 (48.48-1.42)</td>
<td>Depositor EDS</td>
</tr>
<tr>
<td>$R_{merge}$</td>
<td>0.06</td>
<td>Depositor</td>
</tr>
<tr>
<td>$R_{sym}$</td>
<td>(Not available)</td>
<td>Depositor</td>
</tr>
<tr>
<td>$&lt; I/\sigma(I) &gt;$</td>
<td>1.91 (at 1.42Å)</td>
<td>Xtriage</td>
</tr>
</tbody>
</table>
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

- $R/R_{\text{free}}$
- $F_o$ vs. $F_c$ correlation

<table>
<thead>
<tr>
<th>Refinement program</th>
<th>CNS</th>
<th>Depositor</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R$, $R_{\text{free}}$</td>
<td>0.212, 0.223</td>
<td>Depositor</td>
</tr>
<tr>
<td></td>
<td>0.200, 0.209</td>
<td>DCC</td>
</tr>
<tr>
<td>$R_{\text{free}}$ test set</td>
<td>6113 reflections (8.12%)</td>
<td>wwPDB-VP</td>
</tr>
<tr>
<td>$F_o$,$F_c$ correlation</td>
<td>0.96</td>
<td>EDS</td>
</tr>
<tr>
<td>Total number of atoms</td>
<td>3086</td>
<td>wwPDB-VP</td>
</tr>
<tr>
<td>Average B, all atoms ($\text{Å}^2$)</td>
<td>22.0</td>
<td>wwPDB-VP</td>
</tr>
</tbody>
</table>
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

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Analysed</th>
<th>&lt;RSRZ&gt;</th>
<th>#RSRZ&gt;2</th>
<th>OWAB(Å²)</th>
<th>Q&lt;0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>B</td>
<td>24/25 (96%)</td>
<td>-0.30</td>
<td>0 100</td>
<td>100</td>
<td>16, 22, 29, 45</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>153/178 (85%)</td>
<td>0.46</td>
<td>9 (5%)</td>
<td>22 28</td>
<td>14, 24, 39, 46</td>
</tr>
<tr>
<td>All</td>
<td>All</td>
<td>177/203 (87%)</td>
<td>0.36</td>
<td>9 (5%)</td>
<td>28 35</td>
<td>14, 24, 39, 46</td>
</tr>
</tbody>
</table>

List of RSRZ outliers

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Res</th>
<th>Type</th>
<th>RSRZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>A</td>
<td>143</td>
<td>PRO</td>
<td>6.8</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>148</td>
<td>LEU</td>
<td>4.0</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>80</td>
<td>PRO</td>
<td>3.9</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>178</td>
<td>GLN</td>
<td>3.0</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Mol</th>
<th>Type</th>
<th>Chain</th>
<th>Res</th>
<th>Atoms</th>
<th>RSCC</th>
<th>RSR</th>
<th>B-factors (Å²)</th>
<th>Q &lt; 0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>NAP</td>
<td>B</td>
<td>401</td>
<td>48/48</td>
<td>0.95</td>
<td>0.16</td>
<td>27,12,00,76</td>
<td>0</td>
</tr>
</tbody>
</table>

PDB ID 1ZK4 with poor NADP Map-Model Fit

<table>
<thead>
<tr>
<th>Mol</th>
<th>Type</th>
<th>Chain</th>
<th>Res</th>
<th>Atoms</th>
<th>RSCC</th>
<th>RSR</th>
<th>B-factors (Å²)</th>
<th>Q &lt; 0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>NAP</td>
<td>A</td>
<td>1270</td>
<td>48/48</td>
<td>-0.06</td>
<td>0.67</td>
<td>87,96,100,100</td>
<td>0</td>
</tr>
</tbody>
</table>

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

**Apo ferritin**
- EMD-11103, PDB ID 6Z6U: 1.2 Å True atomic

<table>
<thead>
<tr>
<th>Resolution (Å)</th>
<th>Year's average reported resolution</th>
<th>Year's best reported resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- EMD-2824: Scheres J Struct Biol 189:114-122
- PDB 5A1A / EMD-2984: Bartesaghi et al. Science 348:1147-1151
3DEM Experimental Density Map Review

Primary map

Orthogonal projection  Central slices  Largest variance

Mask

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

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 7ZNI / 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

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

5F81: Roessler et al. Structure 24: 631-640
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)
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.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

Interactive between graph and table

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

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

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Experimental Method</th>
<th>Value for Better Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MX</td>
<td>3DEM</td>
</tr>
<tr>
<td>Resolution</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>R/Rfree</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>clashscore</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td># Ramachandran outliers</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td># sidechain outliers</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td># RSRZ outliers</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>RSR on residue/ligand</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>RSCC on residue/ligand</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Q-score</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Atom inclusion</td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>
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

Register for June 3, 2024
3pm ET | 12pm PT
Virtual Office Hour: Mol*

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

Register for June 13, 2024
1pm ET | 10am PT
Virtual Office Hour: Generating DSN6 and MTZ Files

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