Integrative Structure Validation Report

October 09, 2025 - 04:38 PM PDT

The following software was used in the production of this report:

IHMValidation Version 3.0
Python-IHM Version 2.5
EMDB validation analysis Version 0.0.1.dev127
ChimeraX Version 1.9
Chimera Version 1.19
MapQ Version 1.8.1

PDB ID	9A0D | pdb_00009a0d
PDB-Dev ID	PDBDEV_00000049
Structure Title	In-cell architecture of an actively transcribing-translating expressome from M. pneumoniae
Structure Authors	OReilly FJ; Xue L; Graziadei A; Sinn L; Lenz S; Tegunov D; Bloetz C; Singh N; Hagen WJH; Cramer P; Stuelke J; Mahamid J; Rappsilber J
Deposited on	2020-05-15

This is a PDB-IHM Structure Validation Report.

We welcome your comments at helpdesk@pdb-ihm.org

A user guide is available at https://pdb-ihm.org/validation_help.html with specific help available everywhere you see the ? symbol.

List of references used to build this report is available here.

1. Overview

1.1. Summary

This entry consists of 1 model(s). A total of 27 dataset(s) were used to build this entry.

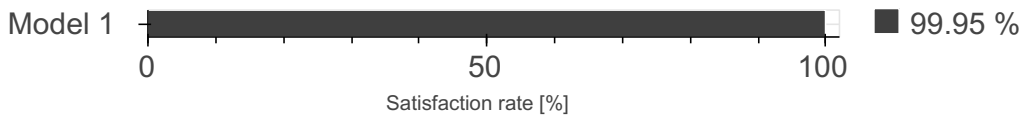
Name	Type	Count
3DEM volume	Experimental data	2

Name	Type	Count
Crosslinking-MS data	Experimental data	2
Comparative model	Starting model	18
Experimental model	Starting model	5

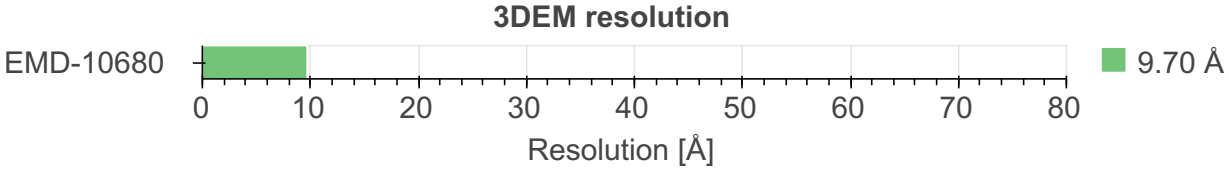
1.2. Overall quality

This validation report contains model quality assessments for all structures, data quality and fit to model assessments for SAS and crosslinking-MS datasets. Data quality and fit to model assessments for other datasets and model uncertainty are under development. Number of plots is limited to 256.

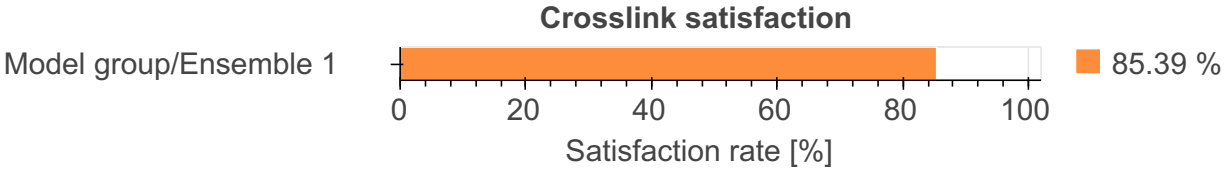
Model Quality: Excluded Volume Analysis



Data Quality



Fit to Data Used for Modeling



2. Model Details

2.1. Ensemble information

This entry consists of 1 distinct ensemble(s).

2.2. Representation

This entry has 1 representation(s).

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
1	1	1	P75591	A	540	12-141, 145-212, 216-364	1-11, 142-144, 213-215, 365-540	100.00 / 64.26	Multiscale: Coarse-grained: 1 - 25 residue(s) per bead
		2	Q50295	B	327	29-233, 234-239, 266-325	1-28, 240-265, 326-327	100.00 / 82.87	Multiscale: Coarse-grained: 1 - 25 residue(s) per bead
		2	Q50295	C	327	29-233, 234-325	1-28, 240-265, 326-327	100.00 / 90.83	Multiscale: Coarse-grained: 1 - 25 residue(s) per bead
		3	P78013	D	1391	14-980, 990-998, 1006-1356	1-13, 225-342, 398-498, 981-989, 999-1005, 1357-1391	100.00 / 95.40	Multiscale: Coarse-grained: 1 - 25 residue(s) per bead
		4	P75271	E	1290	14-1280	1-13, 137-237, 266-272, 658-690, 836-848, 1027-1047, 1056, 1064-1068, 1117-1120, 1281-1290	100.00 / 98.22	Multiscale: Coarse-grained: 1 - 25 residue(s) per bead
		5	P75049	F	150	7-146	1-6, 48-91, 147-150	100.00 / 93.33	Multiscale: Coarse-grained: 1 - 5 residue(s) per bead
		6	Q50301	G	219	70-212	1-69, 213-219	100.00 / 65.30	Multiscale: Coarse-grained: 1 - 25 residue(s) per bead
		7	P75560	H	294	20-241	1-19, 242-294	100.00 / 75.51	Multiscale: Coarse-grained: 1 - 25 residue(s) per bead
		8	P75581	I	108	12-107	1-11, 108	100.00 / 88.89	Multiscale: Coarse-grained: 1 - 11 residue(s) per bead
		9	P41205	J	273	2-206	1, 207-273	100.00 / 75.09	Multiscale: Coarse-grained: 1 - 25 residue(s) per bead

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
		10	P46775	K	205	2-202	1, 203-205	100.00 / 98.05	Multiscale: Coarse-grained: 1 - 3 residue(s) per bead
		11	Q50304	L	142	10-142	1-9	100.00 / 93.66	Multiscale: Coarse-grained: 1 - 9 residue(s) per bead
		12	P75179	M	132	6-132	1-5	100.00 / 96.21	Multiscale: Coarse-grained: 1 - 5 residue(s) per bead
		13	DNA1N	N	39	1-38	16-24, 39	100.00 / 97.44	Multiscale: Coarse-grained: 1 - 5 residue(s) per bead
		14	longRNAR1	O	46	1-9	10-11	23.91 / 81.82	Multiscale: Coarse-grained: 1 - 2 residue(s) per bead
		14	longRNAR1	P	46	37-46	-	21.74 / 100.00	Coarse-grained: 1 residue(s) per bead
		15	DNA1T	Q	39	2-39	1	100.00 / 97.44	Coarse-grained: 1 residue(s) per bead
		16	P75090	R	85	9-74	1-8, 75-85	100.00 / 77.65	Multiscale: Coarse-grained: 1 - 5 residue(s) per bead
		17	30SsubunitE	S	92	1-92	-	100.00 / 100.00	Coarse-grained: 1 residue(s) per bead
		18	30SsubunitF	T	153	1-153	-	100.00 / 100.00	Coarse-grained: 1 residue(s) per bead
		19	30SsubunitJ	U	118	1-118	-	100.00 / 100.00	Coarse-grained: 1 residue(s) per bead

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
		20	30SsubunitK	V	135	1-135	-	100.00 / 100.00	Coarse-grained: 1 residue(s) per bead
		21	30SsubunitL	W	119	1-119	-	100.00 / 100.00	Coarse-grained: 1 residue(s) per bead
		22	30SsubunitM	X	60	1-60	-	100.00 / 100.00	Coarse-grained: 1 residue(s) per bead
		23	30SsubunitN	Y	84	1-84	-	100.00 / 100.00	Coarse-grained: 1 residue(s) per bead
		24	30SsubunitO	Z	80	1-80	-	100.00 / 100.00	Coarse-grained: 1 residue(s) per bead
		25	30SsubunitP	AA	83	1-83	-	100.00 / 100.00	Coarse-grained: 1 residue(s) per bead
		26	30SsubunitQ	AB	71	1-71	-	100.00 / 100.00	Coarse-grained: 1 residue(s) per bead
		27	30SsubunitR	AC	83	1-83	-	100.00 / 100.00	Coarse-grained: 1 residue(s) per bead
		28	30SsubunitT	AD	1544	1-1544	-	100.00 / 100.00	Coarse-grained: 1 residue(s) per bead
		29	30Ssubunitj	AE	77	1-77	-	100.00 / 100.00	Coarse-grained: 1 residue(s) per bead
		30	30SsubunitZ	AF	77	1-77	-	100.00 / 100.00	Coarse-grained: 1 residue(s) per bead

2.3. Datasets used for modeling

There are 27 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	Comparative model	Zenodo	10.5281/zenodo.3837625

ID	Dataset type	Database name	Data access code
2	Experimental model	PDB	pdb_00006flq
3	Comparative model	Zenodo	10.5281/zenodo.3837625
4	Comparative model	Zenodo	10.5281/zenodo.3837625
5	Comparative model	Zenodo	10.5281/zenodo.3837625
6	Comparative model	Zenodo	10.5281/zenodo.3837625
7	Experimental model	PDB	pdb_00003j9w
8	Comparative model	Zenodo	10.5281/zenodo.3837625
9	Comparative model	Zenodo	10.5281/zenodo.3837625
10	Comparative model	Zenodo	10.5281/zenodo.3837625
11	Comparative model	Zenodo	10.5281/zenodo.3837625
12	Comparative model	Zenodo	10.5281/zenodo.3837625
13	Comparative model	Zenodo	10.5281/zenodo.3837625
14	Comparative model	Zenodo	10.5281/zenodo.3837625
15	Comparative model	Zenodo	10.5281/zenodo.3837625
16	Comparative model	Zenodo	10.5281/zenodo.3837625
17	Comparative model	Zenodo	10.5281/zenodo.3837625
18	Comparative model	Zenodo	10.5281/zenodo.3837625
19	Comparative model	Zenodo	10.5281/zenodo.3837625
20	Comparative model	Zenodo	10.5281/zenodo.3837625
21	Crosslinking-MS data	PRIDE	PXD017695
22	Crosslinking-MS data	PRIDE	PXD017711
23	3DEM volume	EMDB	EMD-10680
24	3DEM volume	Zenodo	10.5281/zenodo.3837625
25	Experimental model	PDB	pdb_00006flq
26	Experimental model	PDB	pdb_00006c6u
27	Experimental model	PDB	pdb_00003j9w

2.4. Methodology and software

This entry is a result of 1 distinct protocol(s).

Step number	Protocol ID	Method name	Method type	Method description	Number of computed models	Multi state modeling	Multi scale modeling
1	1	Sampling	Replica exchange monte carlo	Not available	14400000	False	True

There are 6 software packages reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	IMP PMI module	2.12.0	integrative model building	https://integrativemodeling.org

ID	Software name	Software version	Software classification	Software location
2	Integrative Modeling Platform (IMP)	2.12.0	integrative model building	https://integrativemodeling.org
3	SWISS-MODEL	2019-11-21	comparative modeling	https://swissmodel.expasy.org/
4	MODELLER	9.21	comparative modeling	https://salilab.org/modeller/
5	SWISS-MODEL	2.0.0	protein homology modeling	https://swissmodel.expasy.org/
6	SWISS-MODEL	1.3.0	protein homology modeling	https://swissmodel.expasy.org/

3. Data quality

3.2. Crosslinking-MS

At the moment, data validation is only available for crosslinking-MS data deposited as a fully compliant dataset in the PRIDE Crosslinking database. Correspondence between crosslinking-MS and entry entities is established using pyHMMER. Only residue pairs that passed the reported threshold are used for the analysis. The values in the report have to be interpreted in the context of the experiment (i.e. only a minor fraction of in-situ or in-vivo dataset can be used for modeling).

Crosslinking-MS dataset is not available in the PRIDE Crosslinking database.

3.3. 3DEM

This section describes quality of the 3DEM datasets

EMD-10680

3.3.1. Experimental information

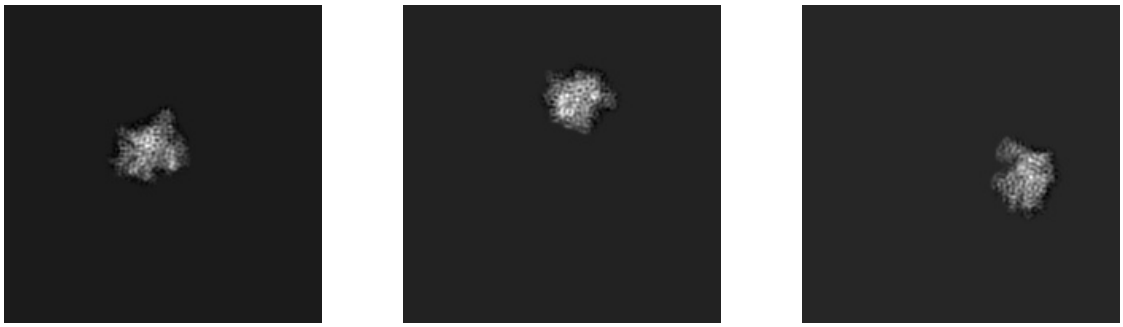
EM reconstruction method:	SUBTOMOGRAM AVERAGING
Resolution:	9.70 Å
Recommended level:	0.750
Estimated volume:	463.06 nm³
Specimen preparation:	Preparation ID 1 Vitrification
Map-only validation report:	wwPDB validation report

3.3.2. Map visualisation

This section contains visualisations of the EMDB entry EMD-10680. These allow visual inspection of the internal detail of the map and identification of artifacts. Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

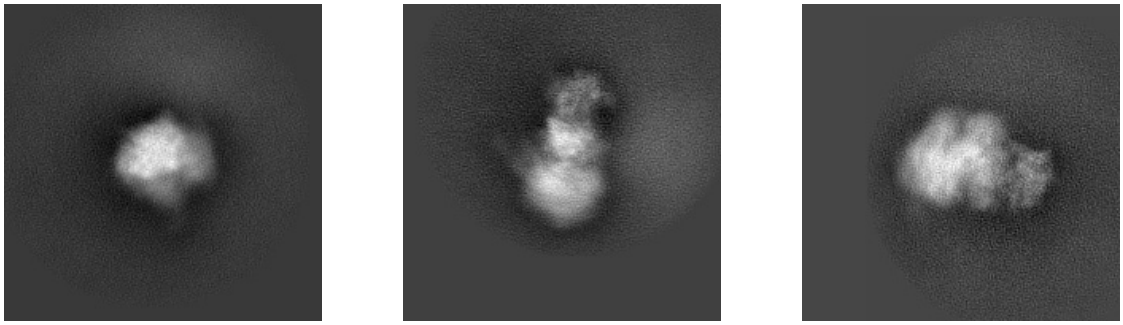
3.3.2.1. Orthogonal projections

Primary map



X **Y** **Z**

Raw map

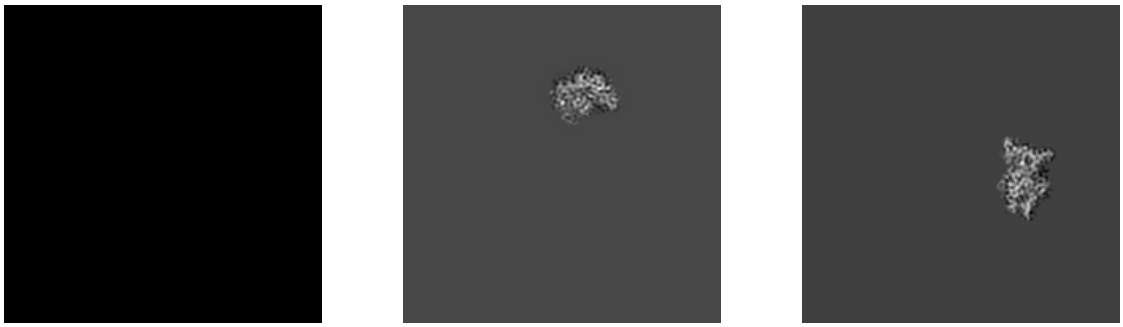


X **Y** **Z**

The images above show the map projected in three orthogonal directions.

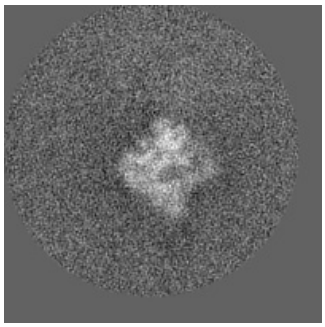
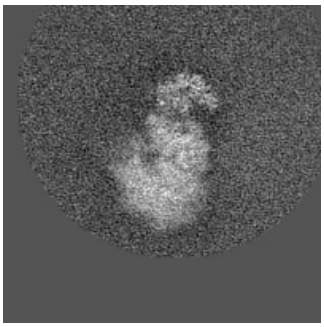
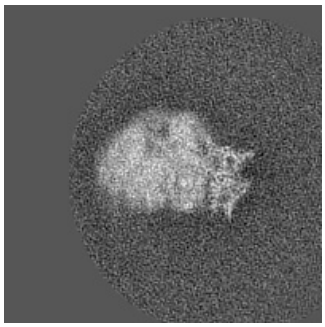
3.3.2.2. Central slices

Primary map



X Index: 100 **Y Index: 100** **Z Index: 100**

Raw map

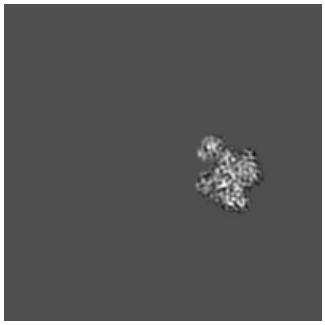
X Index: 100 **Y Index: 100** **Z Index: 100**

The images above show central slices of the map in three orthogonal directions.

3.3.2.3. Largest variance slices ?

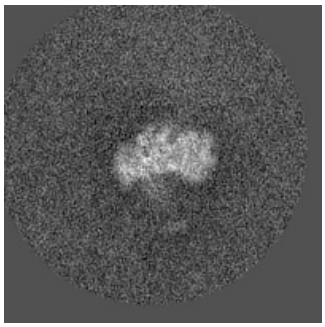
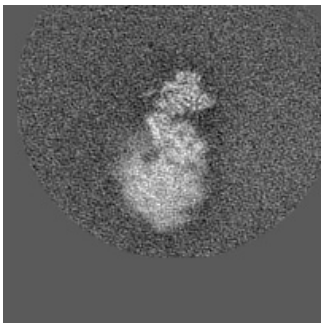
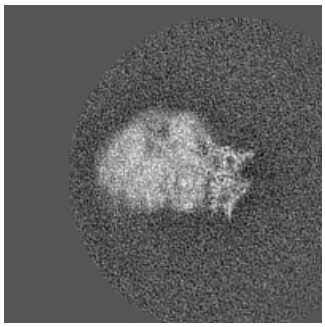
Primary map





X Index: 143 **Y Index: 90** **Z Index: 109**

Raw map

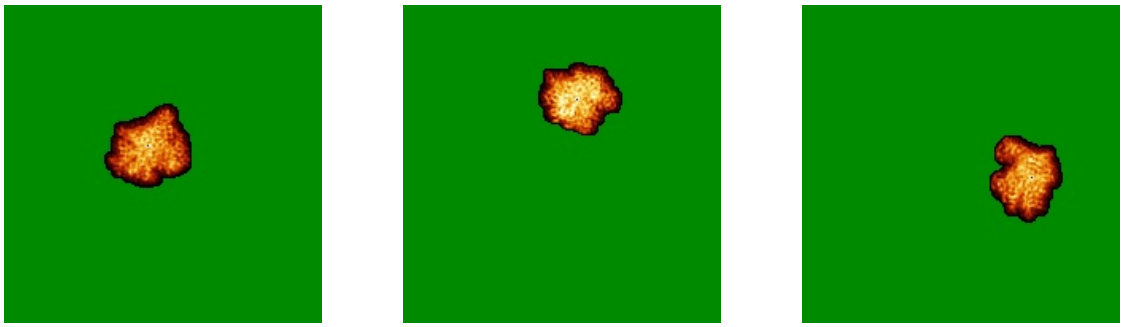




X Index: 116 **Y Index: 96** **Z Index: 100**

The images above show the largest variance slices of the map in three orthogonal directions.

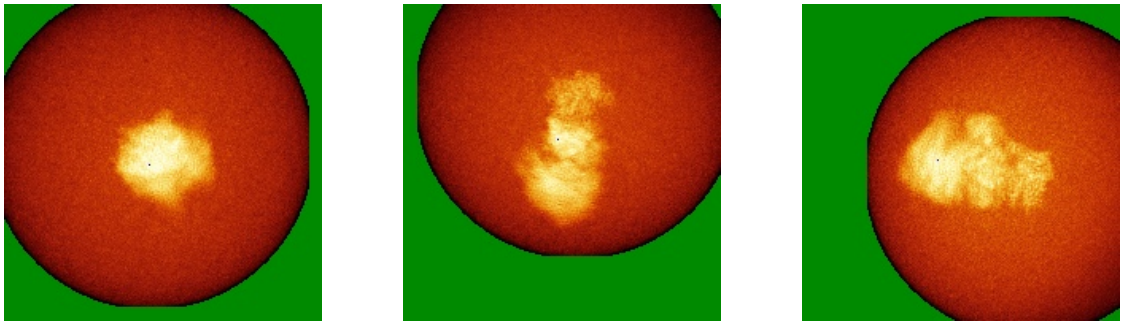
3.3.2.4 Orthogonal standard-deviation projections (false-color) ?

Primary map



X **Y** **Z**

Raw map

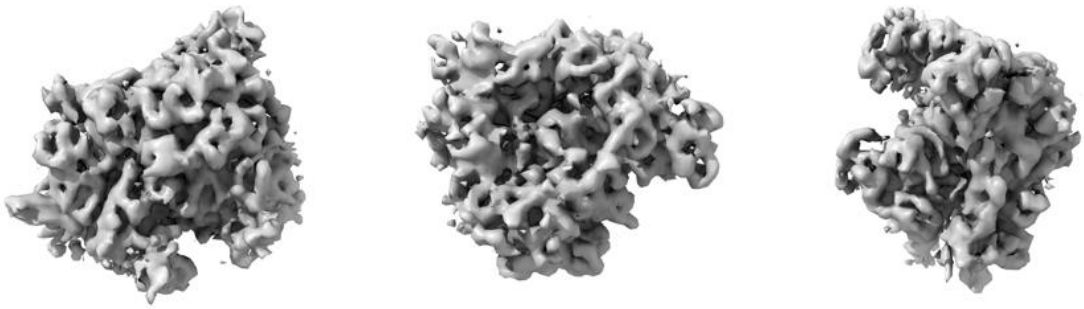


X **Y** **Z**

The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

3.3.2.5. Orthogonal surface views ?

Primary map



X **Y** **Z**

The images above show the 3D surface view of the map at the recommended contour level 0.750 . These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

Raw map

X **Y** **Z**

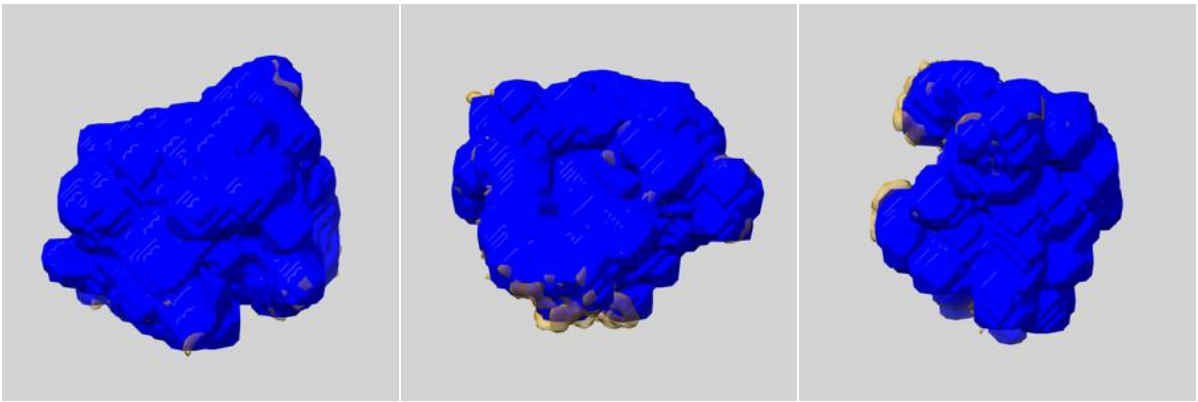
These images show the 3D surface of the raw map. The raw map's contour level 0.340 was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

3.3.2.6. Mask visualisation

This section shows the 3D surface view of the primary map at 50% transparency overlaid with the specified mask at 0% transparency. A mask typically either:

- Encompasses the whole structure;
- Separates out a domain, a functional unit, a monomer or an area of interest from a larger structure.

emd_10680_msk_1.map

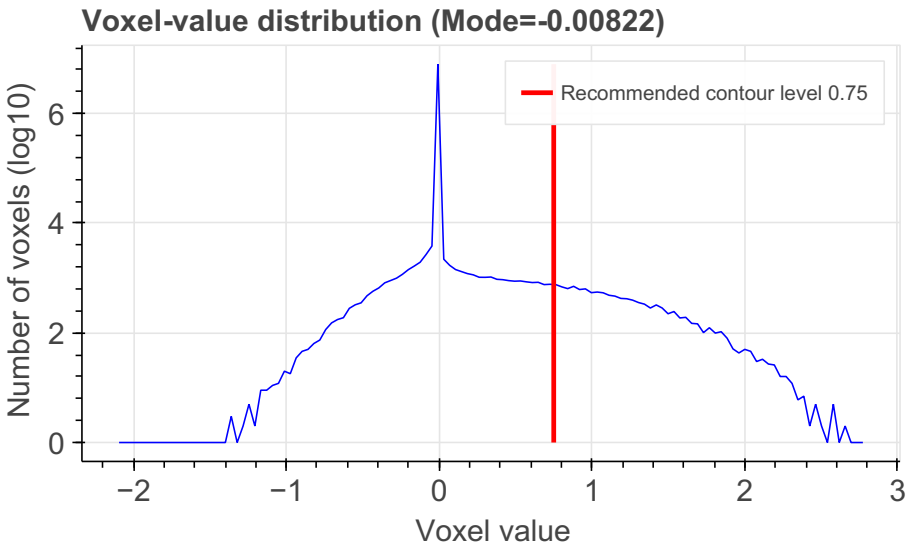


X **Y** **Z**

3.3.3. Map analysis

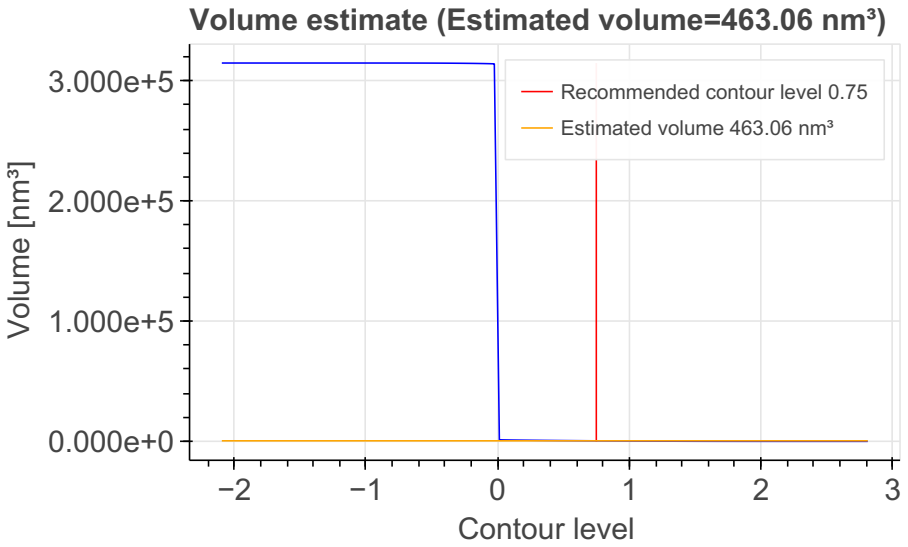
This section contains the results of statistical analysis of the map.

3.3.3.1. Map-value distribution



The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

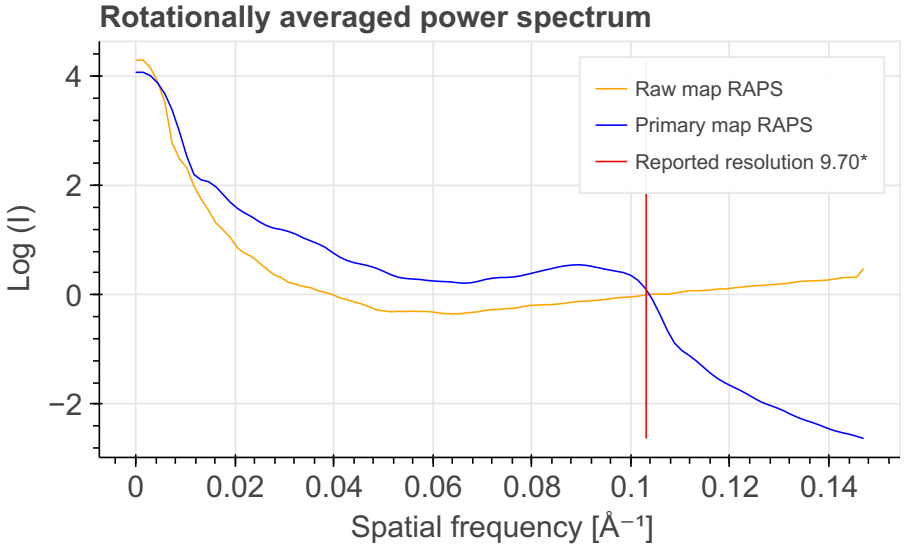
3.3.3.2. Volume estimate ?



The volume at the recommended contour level is 463.06 nm³.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

3.3.3.3. Rotationally averaged power spectrum ?

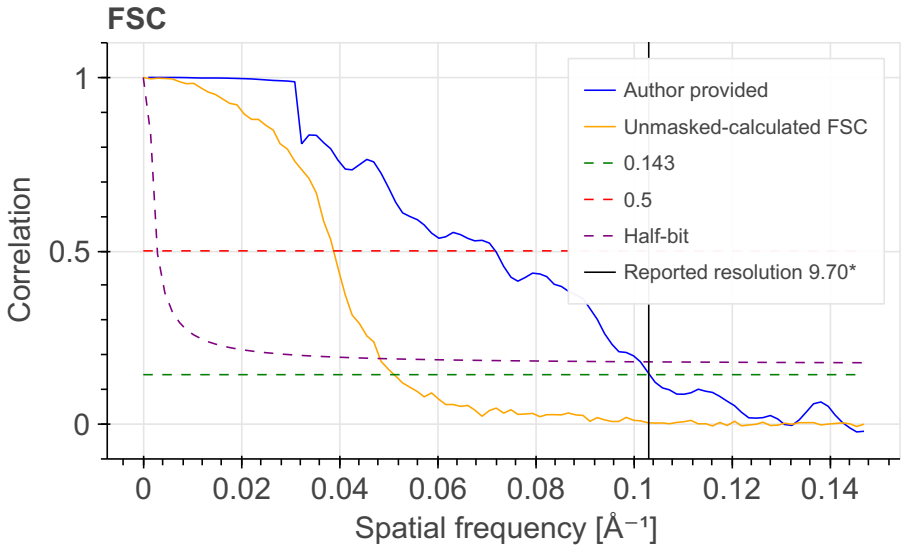


*Reported resolution corresponds to spatial frequency of 0.103 Å^{-1}

3.3.4. Fourier-Shell correlation

3.3.4.1. FSC

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.



*Reported resolution corresponds to spatial frequency of 0.103 Å^{-1}

3.3.4.2. Resolution estimates

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	**0.5**	**Half-bit**
Reported by author	9.70	-	-
Author-provided FSC curve	9.68	13.93	9.86
Unmasked-calculated*	19.57	25.77	20.70

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 19.57 differs from the reported value 9.70 by more than 10%.

4. Model quality

For models with atomic structures, MolProbity analysis is performed. For models with coarse-grained or multi-scale structures, excluded volume analysis is performed.

4.1a. Excluded Volume Analysis

Excluded volume satisfaction for the models in the entry are listed below. The Analysed column shows the number of particle-partice or particle-atom pairs for which excluded volume was analysed.

Model ID	Analysed	Number of violations	Excluded Volume Satisfaction (%)
1	26897445	12929	99.95

5. Fit to Data Used for Modeling Assessment

5.2. Crosslinking-MS

5.2.1. Restraint types

This table summarizes information about crosslinker(s) used for data generation, and how crosslinking information was translated into actual modeling restraints. Restraints assigned "by-residue" are interpreted as between CA atoms. Restraints between coarse-grained beads are indicated as "coarse-grained". Restraint group represents a set of crosslinking restraints applied collectively in the modeling.

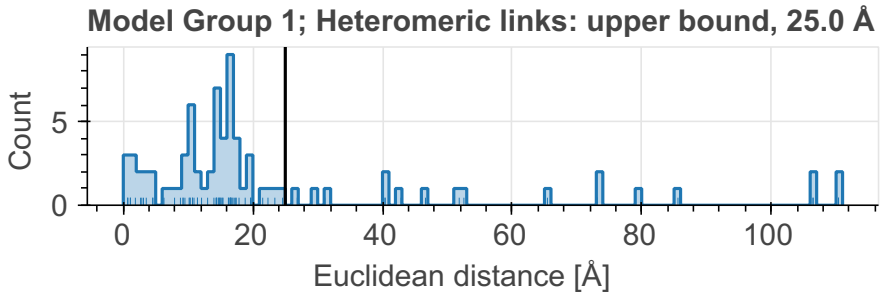
There are 630 crosslinking restraints combined in 479 restraint groups.

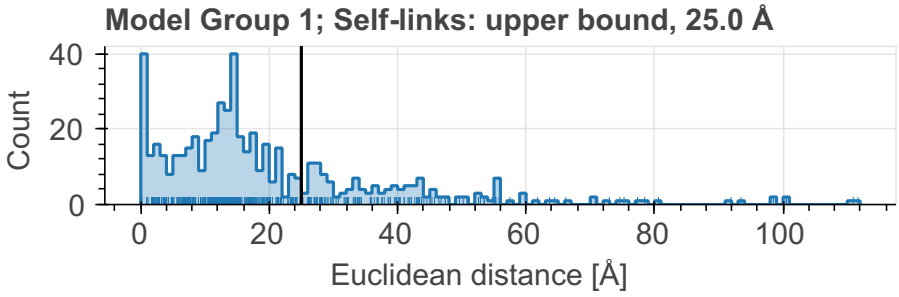
Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSS	LYS	CA	LYS	CA	upper bound	25.00	273
DSS	LYS	coarse-grained	LYS	coarse-grained	upper bound	25.00	173
DSS	LYS	coarse-grained	THR	coarse-grained	upper bound	25.00	5
DSS	LYS	coarse-grained	SER	coarse-grained	upper bound	25.00	5
DSS	LYS	CA	SER	CA	upper bound	25.00	7
DSS	LYS	coarse-grained	MET	coarse-grained	upper bound	25.00	8
DSS	SER	CA	THR	CA	upper bound	25.00	1
DSS	SER	CA	SER	CA	upper bound	25.00	1
DSS	THR	CA	THR	CA	upper bound	25.00	4
DSS	LYS	CA	THR	CA	upper bound	25.00	3
DSS	SER	CA	TYR	CA	upper bound	25.00	2
DSS	LYS	coarse-grained	TYR	coarse-grained	upper bound	25.00	3
DSSO	LYS	CA	LYS	CA	upper bound	25.00	77
DSSO	LYS	coarse-grained	LYS	coarse-grained	upper bound	25.00	65
DSSO	LYS	coarse-grained	MET	coarse-grained	upper bound	25.00	2

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSSO	LYS	CA	SER	CA	upper bound	25.00	1

Distograms of individual restraints

Distograms (i.e., histogram plots of distances) provide an overview of distributions of distances between residues for which chemical crosslinks were identified. The shift of the distogram relative to the threshold value may indicate a poor model. Restraints with identical thresholds are grouped into one plot. Only the best distance per restraint per model group/ensemble is plotted. Inter- and intramolecular (including self-links) restraints are also grouped into one plot. Distance for a restraint between coarse-grained beads is calculated as a minimal distance between shells; if beads intersect, the distance will be reported as 0.0. A bead with the highest available resolution for a given residue is used for the assessment.





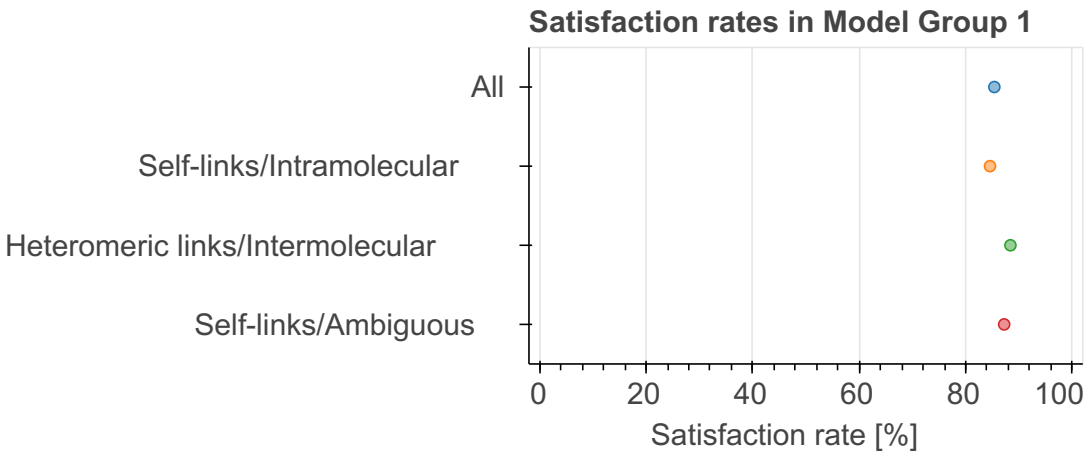
5.2.2. Satisfaction of restraints

Satisfaction of restraints is calculated on a restraint group (a set of crosslinking restraints applied collectively in the modeling) level. Satisfaction of a restraint group depends on satisfaction of individual restraints in the group and the conditionality (all/any). A restraint group is considered satisfied, if the condition was met in at least one model of the model group/ensemble. The number of measured restraints can be smaller than the total number of restraint groups if crosslinks involve non-modeled residues. Only deposited models are used for validation right now.

State group	State	Model group	# of Deposited models/Total	Restraint group type	Satisfied (%)	Violated (%)	Count (Total=479)
1	1	1	1/20131	All	85.39	14.61	479
				Self-links/ Intramolecular	84.57	15.43	363
				Heteromeric links/ Intermolecular	88.41	11.59	69
				Self-links/ Ambiguous	87.23	12.77	47

Per-model satisfaction rates in ensembles

Every point represents one model in a model group/ensemble. Where possible, boxplots with quartile marks are also plotted.



5.3. 3DEM

This section describes fit of models to the 3DEM data. Only results for the representative model, selected as a first model with the largest number of asymmetric units.

3DEM validation for coarse-grained structures is under development.

6. Fit to Data Used for Validation Assessment

Validation for this section is under development.

Acknowledgments

The development of integrative model validation metrics, implementation of a model validation pipeline, and creation of a validation report for integrative structures are funded by NSF awards to the PDB-IHM team (DBI-1756248, DBI-2112966, DBI-2112967, DBI-2112968, and DBI-1756250) and awards from NSF, NIH, and DOE to the RCSB PDB (DBI-2321666, R01GM157729, and DE-SC0019749). The PDB-IHM team and members of the Sali lab contributed model validation metrics and software packages.

Dr. Jill Trewhella, Dr. Dina Schneidman, and members of the SASBDB repository are acknowledged for their advice and support in implementing SAS validation methods. Team members from the labs of Dr. Juri Rappsilber, Dr. Alexander Leitner, Dr. Andrea Graziadei, and members of PRIDE database are acknowledged for their advice and support in implementing crosslinking-MS validation methods. We are grateful to Dr. Shruthi Viswanath for discussions about uncertainty assessment of integrative structural models.

Members of the wwPDB Integrative/Hybrid Methods Task Force provided recommendations and community support for the project.