Integrative Structure Validation Report

October 09, 2025 - 04:40 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	9A2A | pdb_00009a2a
PDB-Dev ID	PDBDEV_00000152
Structure Title	Integrative structure of the human MHM complex
Structure Authors	Arvindekar S; Jackman MJ; Low JKK; Landsberg MJ; Mackay JP; Viswanath S
Deposited on	2022-07-30

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 17 dataset(s) were used to build this entry.

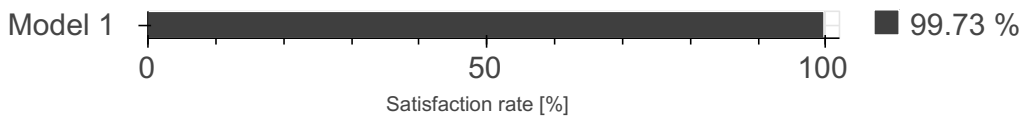
Name	Type	Count
3DEM volume	Experimental data	2

Name	Type	Count
Crosslinking-MS data	Experimental data	2
De Novo model	Starting model	2
Comparative model	Starting model	4
Experimental model	Starting model	7

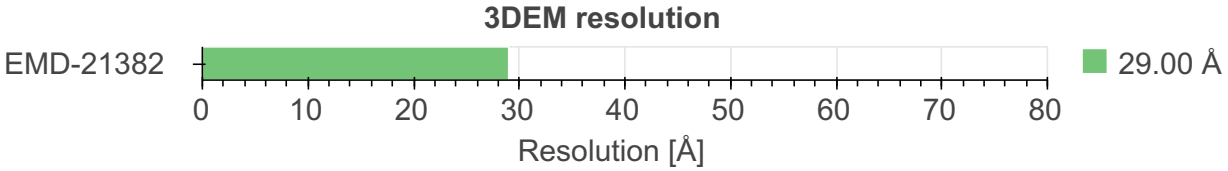
1.2. Overall quality

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.

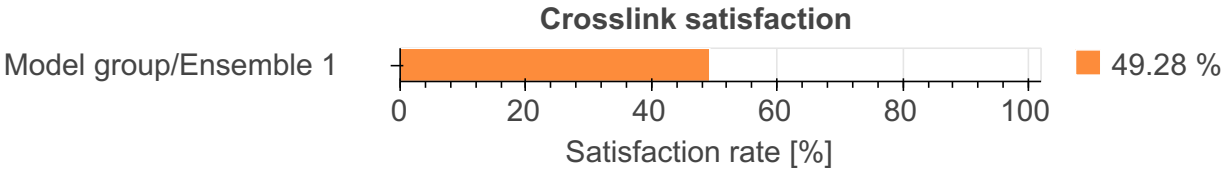
Model Quality: Excluded Volume Analysis



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	MTA1	A	431	1-164, 165-333, 334-353, 389-431	229-236, 354-388	100.00 / 91.88	Multiscale: Coarse-grained: 1 - 30 residue(s) per bead
				B					
		2	HDAC1	C	482	8-376	1-7, 377-482	100.00 / 76.56	Multiscale: Coarse-grained: 1 - 30 residue(s) per bead
				D					
		3	MBD3	E	291	1-71, 221-249	72-220, 250-291	100.00 / 34.36	Multiscale: Coarse-grained: 1 - 30 residue(s) per bead
				F					
		4	P66A	G	633	137-178	136	6.79 / 97.67	Coarse-grained: 1 residue(s) per bead
				H					

2.3. Datasets used for modeling

There are 17 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	Experimental model	PDB	pdb_00002fvu
2	Comparative model	Zenodo	10.5281/zenodo.6674232
3	Experimental model	PDB	pdb_00004bkx
4	De Novo model	Zenodo	10.5281/zenodo.6674232
5	Experimental model	PDB	pdb_00002gat
6	Comparative model	Zenodo	10.5281/zenodo.6674232
7	Comparative model	Zenodo	10.5281/zenodo.6674232
8	De Novo model	Zenodo	10.5281/zenodo.6674232
9	Comparative model	Zenodo	10.5281/zenodo.6674232
10	Experimental model	PDB	pdb_00006cc8
11	Experimental model	PDB	pdb_00002l2l
12	Experimental model	Zenodo	10.5281/zenodo.6674232
13	Experimental model	Zenodo	10.5281/zenodo.6674232
14	Crosslinking-MS data	Zenodo	10.5281/zenodo.6674232
15	Crosslinking-MS data	Zenodo	10.5281/zenodo.6674232
16	3DEM volume	EMDB	EMD-21382
17	3DEM volume	Zenodo	10.5281/zenodo.6674232

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	600000	False	True

There are 2 software packages reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	IMP PMI module	2.16.0	integrative model building	https://integrativemodeling.org
2	Integrative Modeling Platform (IMP)	2.16.0	integrative model building	https://integrativemodeling.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-21382

3.3.1. Experimental information

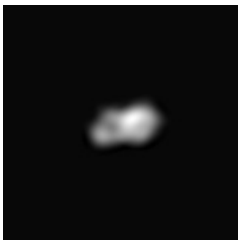
EM reconstruction method:	SINGLE PARTICLE
Resolution:	29.00 Å
Recommended level:	0.120
Estimated volume:	630.40 nm³
Specimen preparation:	Preparation ID 1 Staining: NEGATIVE; Material: Uranyl Acetate
Map-only validation report:	wwPDB validation report

3.3.2. Map visualisation

This section contains visualisations of the EMDB entry EMD-21382. 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

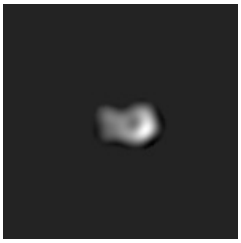


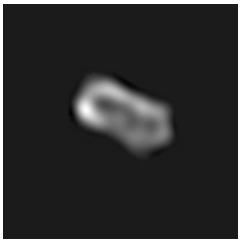
X **Y** **Z**

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

3.3.2.2. Central slices

Primary map



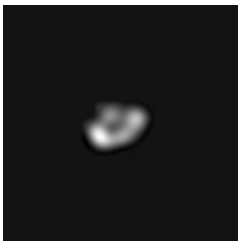


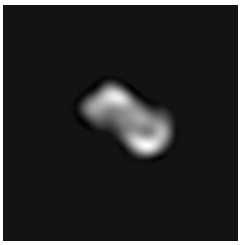
X Index: 74 **Y Index: 74** **Z Index: 74**

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

3.3.2.3. Largest variance slices

Primary map



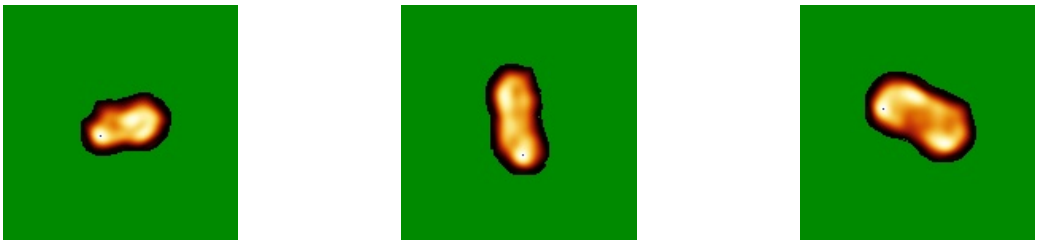


X Index: 87 **Y Index: 84** **Z Index: 67**

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**

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

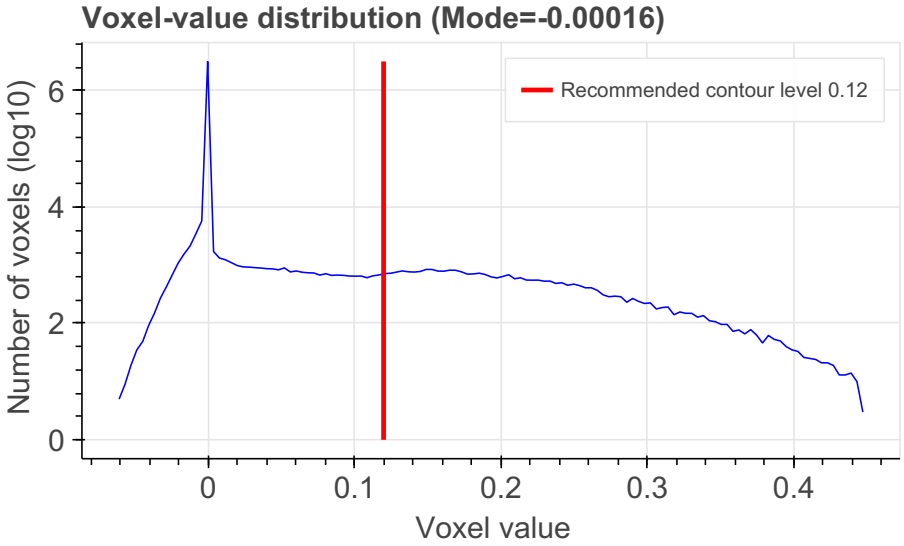
X **Y** **Z**

The images above show the 3D surface view of the map at the recommended contour level 0.120 . These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

3.3.3. Map analysis

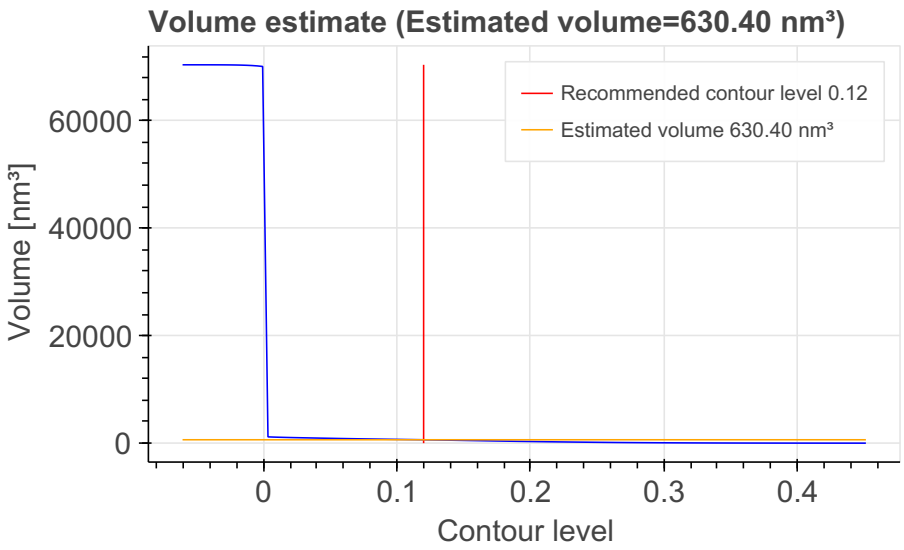
This section contains the results of statistical analysis of the map.

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.

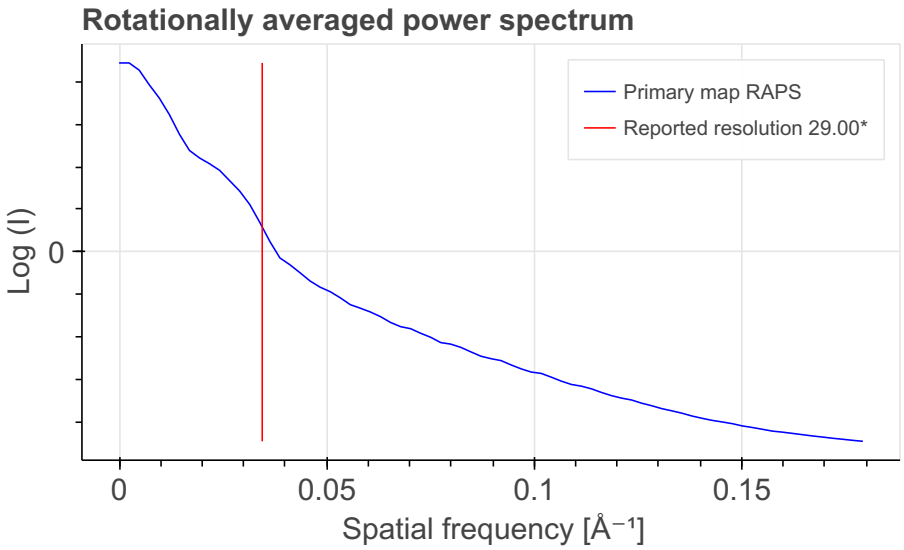
3.3.3.2. Volume estimate



The volume at the recommended contour level is 630.40 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 0.034 Å⁻¹

3.3.4. Fourier-Shell correlation

3.3.4.2. Resolution estimates

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	29.00	-	-
Author-provided FSC curve	28.57	30.30	29.15

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	1677196	4505	99.73

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.

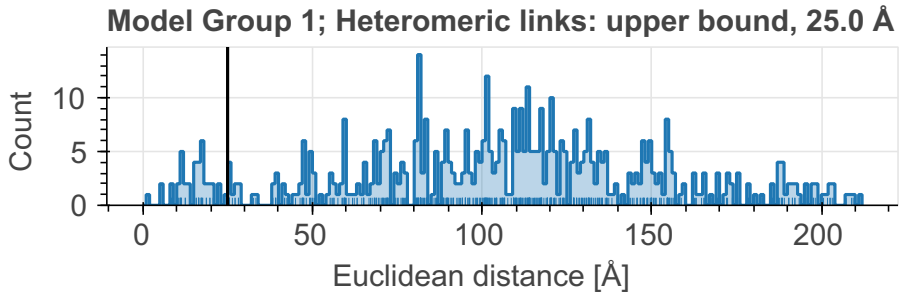
There are 1672 crosslinking restraints combined in 418 restraint groups.

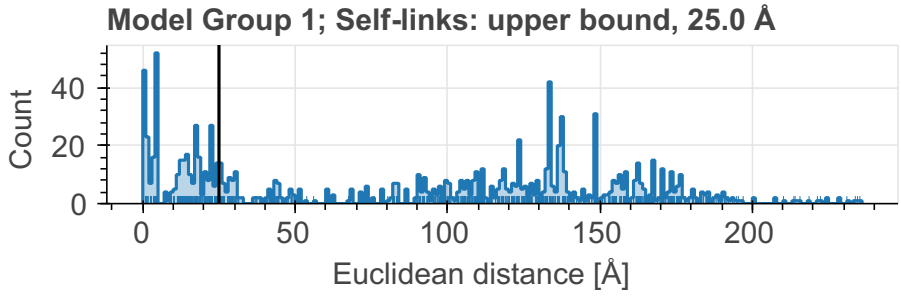
Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
BS3	LYS	CA	SER	CA	upper bound	25.00	220
BS3	LYS	coarse-grained	SER	coarse-grained	upper bound	25.00	364
BS3	SER	coarse-grained	SER	coarse-grained	upper bound	25.00	80
BS3	SER	CA	TYR	CA	upper bound	25.00	20
BS3	LYS	coarse-grained	LYS	coarse-grained	upper bound	25.00	356
BS3	LYS	coarse-grained	THR	coarse-grained	upper bound	25.00	92
BS3	ARG	CA	SER	CA	upper bound	25.00	24
BS3	LYS	CA	LYS	CA	upper bound	25.00	128
BS3	SER	coarse-grained	TYR	coarse-grained	upper bound	25.00	8
BS3	ARG	CA	ARG	CA	upper bound	25.00	8
BS3	SER	CA	SER	CA	upper bound	25.00	60
BS3	LYS	CA	TYR	CA	upper bound	25.00	68
BS3	SER	coarse-grained	THR	coarse-grained	upper bound	25.00	28
BS3	LYS	CA	THR	CA	upper bound	25.00	60
BS3	GLU	CA	LYS	CA	upper bound	25.00	12
BS3	LYS	CA	MET	CA	upper bound	25.00	12
BS3	SER	CA	THR	CA	upper bound	25.00	16
BS3	ARG	CA	THR	CA	upper bound	25.00	12

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
BS3	ARG	CA	GLN	CA	upper bound	25.00	4
BS3	LYS	coarse-grained	TYR	coarse-grained	upper bound	25.00	24
BS3	PRO	coarse-grained	VAL	coarse-grained	upper bound	25.00	4
BS3	THR	CA	TYR	CA	upper bound	25.00	4
BS3	ALA	CA	ARG	CA	upper bound	25.00	4
BS3	TYR	CA	TYR	CA	upper bound	25.00	4
BS3	ARG	CA	MET	CA	upper bound	25.00	4
BS3	GLU	CA	THR	CA	upper bound	25.00	8
BS3	ARG	CA	LYS	CA	upper bound	25.00	20
BS3	MET	CA	TYR	CA	upper bound	25.00	8
BS3	GLN	CA	LYS	CA	upper bound	25.00	4
BS3	ALA	CA	LYS	CA	upper bound	25.00	4
BS3	GLU	CA	TYR	CA	upper bound	25.00	4
BS3	MET	CA	SER	CA	upper bound	25.00	4
BS3	ARG	coarse-grained	LYS	coarse-grained	upper bound	25.00	4

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.





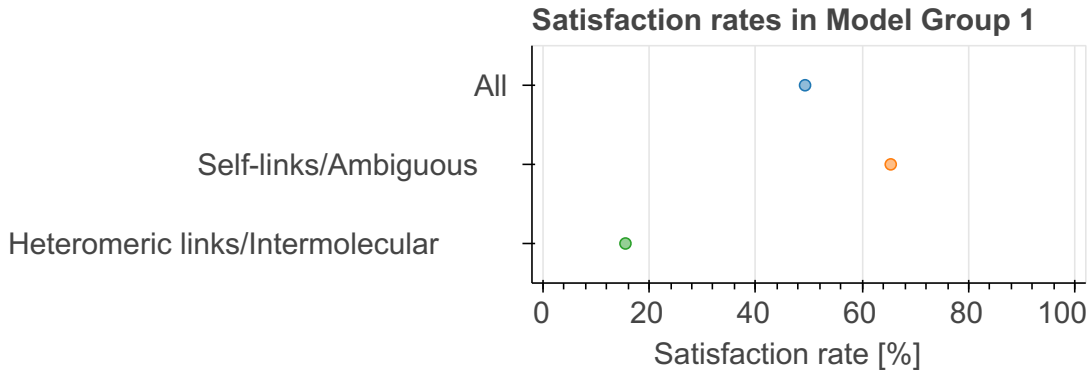
5.2.2. Satisfaction of restraints

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=418)
1	1	1	1/28451	All	49.28	50.72	418
				Self-links/ Ambiguous	65.37	34.63	283
				Heteromeric links/ Intermolecular	15.56	84.44	135

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.