Integrative Structure Validation Report

October 09, 2025 - 04:37 PM PDT

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

IHMValidation Version 3.0
Python-IHM Version 2.5

PDB ID	8ZZ1 | pdb_00008zz1
PDB-Dev ID	PDBDEV_00000001
Structure Title	Structure of the Nup84 sub-complex of the Nuclear Pore Complex
Structure Authors	Shi, Y.; Fernandez-Martinez, J.; Tjioe, E.; Pellarin, R.; Kim, S.J.; Williams, R.; Schneidman-Duhovny, D.; Sali, A.; Rout, M.P.; Chait, B.T.
Deposited on	2016-08-31

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 2 model(s). A total of 28 dataset(s) were used to build this entry.

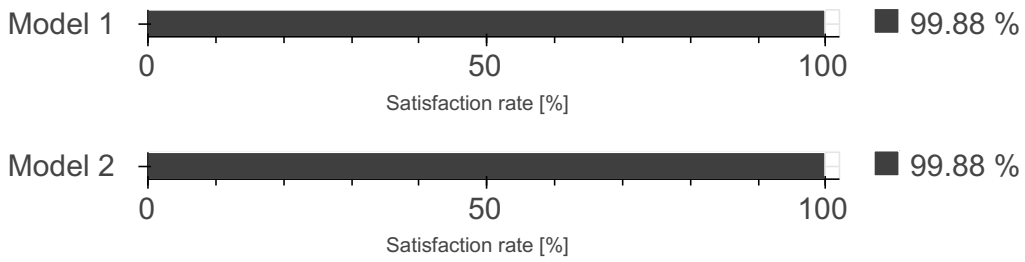
Name	Type	Count
2DEM class average	Experimental data	1
Crosslinking-MS data	Experimental data	2
EM raw micrographs	Experimental data	1

Name	Type	Count
Experimental model	Starting model	17
Comparative model	Starting model	7

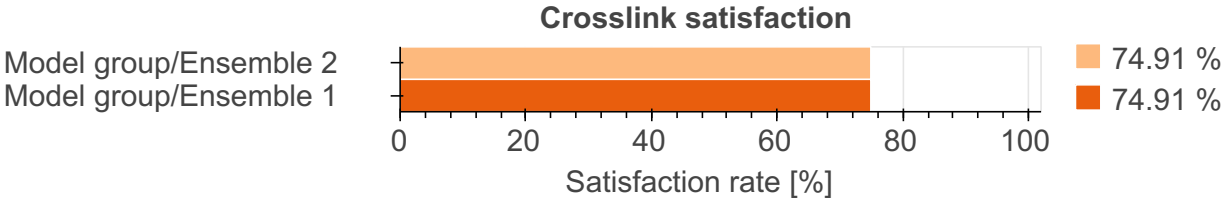
1.2. Overall quality

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



Fit to Data Used for Modeling



2. Model Details

2.1. Ensemble information

This entry consists of 2 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

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
1	1-2	1	Nup84	A	726	7-20, 27-80, 96-126, 136-364, 372-483, 506-562, 575-726	1-6, 21-26, 81-95, 127-135, 365-371, 484-505, 563-574	100.00 / 89.39	Multiscale: Coarse-grained: 1 - 20 residue(s) per bead
		2	Nup85	B	744	67-122, 135-427, 461-529, 533-602, 620-671, 680-743	1-66, 123-134, 428-460, 530-532, 603-619, 672-679, 744	100.00 / 81.18	Multiscale: Coarse-grained: 1 - 20 residue(s) per bead
		3	Nup120	C	1037	1-29, 53-212, 221-305, 311-429, 440-710, 711-712, 727-781, 805-892, 903-910, 921-1010, 1023-1037	30-52, 213-220, 306-310, 430-439, 713-726, 782-804, 893-902, 911-920, 1011-1022	100.00 / 88.91	Multiscale: Coarse-grained: 1 - 20 residue(s) per bead
		4	Nup133	D	1157	56-78, 86-125, 133-144, 162-184, 193-200, 206-249, 258-480, 490-763, 772-1155	1-55, 79-85, 126-132, 145-161, 185-192, 201-205, 250-257, 481-489, 764-771, 1156-1157	100.00 / 89.11	Multiscale: Coarse-grained: 1 - 20 residue(s) per bead
		5	Nup145c	E	712	126-144, 151-175, 182-553	1-125, 145-150, 176-181, 554-712	100.00 / 58.43	Multiscale: Coarse-grained: 1 - 20 residue(s) per bead
		6	Seh1	F	349	1-248, 288-346	249-287, 347-349	100.00 / 87.97	Multiscale: Coarse-grained: 1 - 20 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
		7	Sec13	G	297	2-158, 166-296	1, 159-165, 297	100.00 / 96.97	Multiscale: Coarse-grained: 1 - 7 residue(s) per bead

2.3. Datasets used for modeling

There are 28 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	Experimental model	PDB	pdb_00003jro
2	Experimental model	PDB	pdb_00003f3f
3	Experimental model	PDB	pdb_00003iko
4	Comparative model	Zenodo	10.5281/zenodo.1218053
5	Experimental model	PDB	pdb_00003cqc
6	Comparative model	Zenodo	10.5281/zenodo.1218053
7	Experimental model	PDB	pdb_00004lct
8	Experimental model	PDB	pdb_00002qx5
9	Experimental model	PDB	pdb_00003ewe
10	Comparative model	Zenodo	10.5281/zenodo.1218053
11	Experimental model	PDB	pdb_00003f7f
12	Experimental model	PDB	pdb_00003hxr
13	Experimental model	PDB	pdb_00004fhn
14	Comparative model	Zenodo	10.5281/zenodo.1218053
15	Experimental model	PDB	pdb_00004q9t
16	Comparative model	Zenodo	10.5281/zenodo.1218053
17	Experimental model	PDB	pdb_00003i4r
18	Experimental model	PDB	pdb_00003kfo
19	Comparative model	Zenodo	10.5281/zenodo.1218053
20	Experimental model	PDB	pdb_00003bg1
21	Experimental model	PDB	pdb_00003bg0
22	Comparative model	Zenodo	10.5281/zenodo.1218053
23	Experimental model	PDB	pdb_00003f3f

ID	Dataset type	Database name	Data access code
24	Experimental model	PDB	pdb_00002pm7
25	Crosslinking-MS data	Zenodo	10.5281/zenodo.1218053
26	Crosslinking-MS data	Zenodo	10.5281/zenodo.1218053
27	EM raw micrographs	Zenodo	10.5281/zenodo.58025
28	2DEM class average	Zenodo	10.5281/zenodo.1218053

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	500	False	True
2	1	Sampling	Replica exchange monte carlo	Not available	5000	False	True

There are 6 software packages reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	Integrative Modeling Platform (IMP)	develop-0a5706e202	integrative model building	https://integrativemodeling.org
2	IMP PMI module	67456c0	integrative model building	https://integrativemodeling.org
3	HHpred	2.0.16	protein homology detection	https://toolkit.tuebingen.mpg.de/hhpred
4	PSIPRED	4.00	secondary structure prediction	http://bioinf.cs.ucl.ac.uk/psipred/
5	DISOPRED	3	disorder prediction	http://bioinf.cs.ucl.ac.uk/psipred/?disopred=1
6	MODELLER	9.12	comparative modeling	https://salilab.org/modeller/

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.

3.4. 2DEM class average

Validation for this section is under development.

3.4. EM raw micrographs

Validation for this section is under development.

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	9165621	10899	99.88
2	9165621	10935	99.88

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.

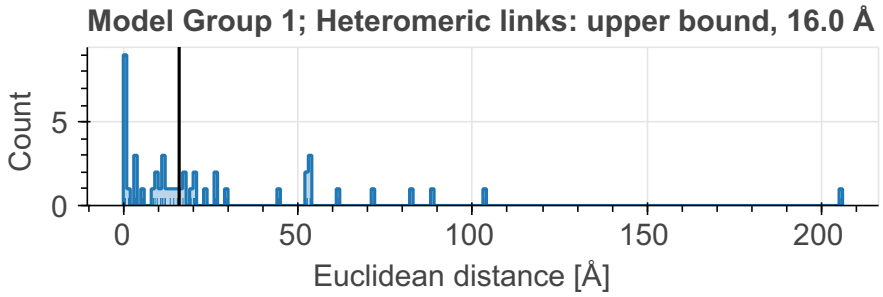
There are 291 crosslinking restraints combined in 291 restraint groups.

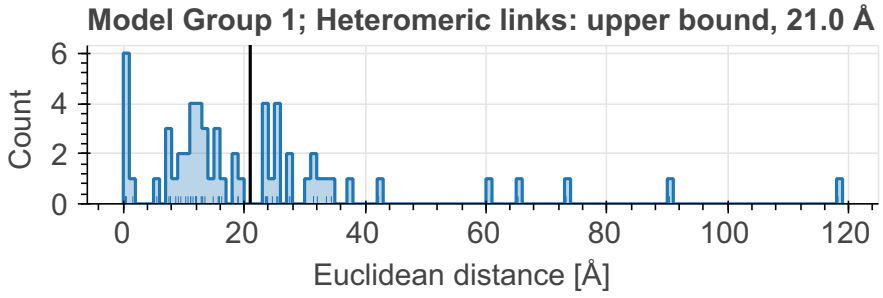
Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSS	LYS	CA	LYS	CA	upper bound	21.00	80
DSS	LYS	coarse-grained	LYS	coarse-grained	upper bound	21.00	46
DSS	LYS	coarse-grained	SER	coarse-grained	upper bound	21.00	4
DSS	LYS	coarse-grained	THR	coarse-grained	upper bound	21.00	15
DSS	LYS	coarse-grained	MET	coarse-grained	upper bound	21.00	3
DSS	LYS	CA	VAL	CA	upper bound	21.00	8
DSS	LYS	coarse-grained	VAL	coarse-grained	upper bound	21.00	3
DSS	LYS	CA	MET	CA	upper bound	21.00	3
DSS	MET	coarse-grained	THR	coarse-grained	upper bound	21.00	1
DSS	GLN	coarse-grained	MET	coarse-grained	upper bound	21.00	1
EDC	GLU	CA	LYS	CA	upper bound	16.00	36
EDC	ASP	coarse-grained	LYS	coarse-grained	upper bound	16.00	23
EDC	GLU	coarse-grained	LYS	coarse-grained	upper bound	16.00	24

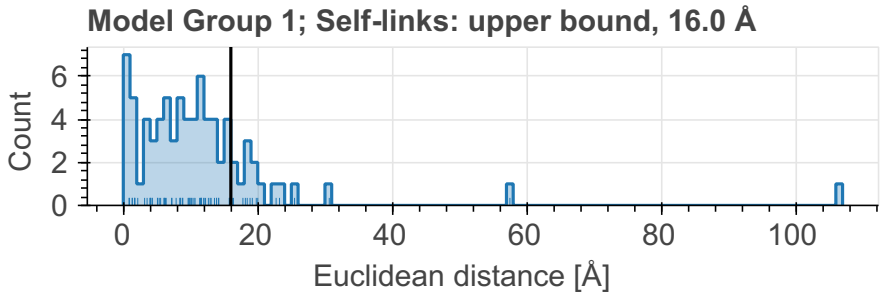
Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
EDC	ASP	CA	LYS	CA	upper bound	16.00	16
EDC	ASN	coarse-grained	LYS	coarse-grained	upper bound	16.00	1
EDC	SER	coarse-grained	THR	coarse-grained	upper bound	16.00	1
EDC	ILE	coarse-grained	THR	coarse-grained	upper bound	16.00	1
EDC	ASP	coarse-grained	THR	coarse-grained	upper bound	16.00	1
EDC	ASP	coarse-grained	MET	coarse-grained	upper bound	16.00	3
EDC	MET	coarse-grained	THR	coarse-grained	upper bound	16.00	1
EDC	ASP	CA	GLU	CA	upper bound	16.00	1
EDC	ASP	CA	VAL	CA	upper bound	16.00	3
EDC	GLU	coarse-grained	VAL	coarse-grained	upper bound	16.00	1
EDC	GLU	CA	VAL	CA	upper bound	16.00	4
EDC	ASP	coarse-grained	VAL	coarse-grained	upper bound	16.00	5
EDC	GLU	coarse-grained	MET	coarse-grained	upper bound	16.00	1
EDC	ASP	coarse-grained	GLN	coarse-grained	upper bound	16.00	2
EDC	GLN	coarse-grained	GLU	coarse-grained	upper bound	16.00	1
EDC	LYS	coarse-grained	THR	coarse-grained	upper bound	16.00	2

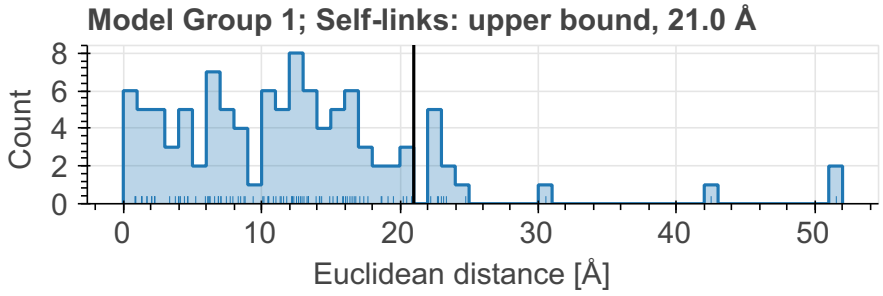
Distograms of individual restraints

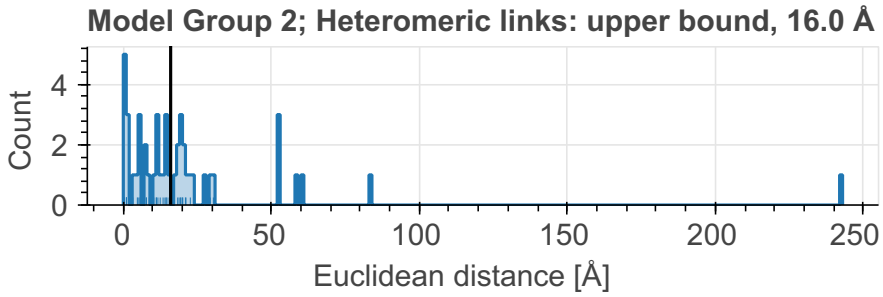
Distograms (i.e., histogram plots of distances) provide an overview of distributions of distances between residues for which chemical crosslinks were identified. The shift of the distogram relative to the threshold value may indicate a poor model. Restraints with identical thresholds are grouped into 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.

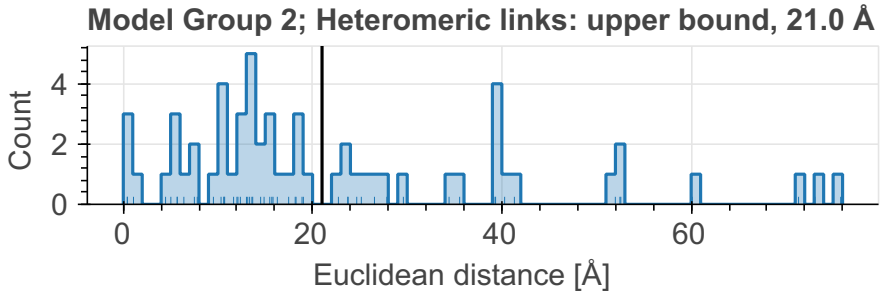


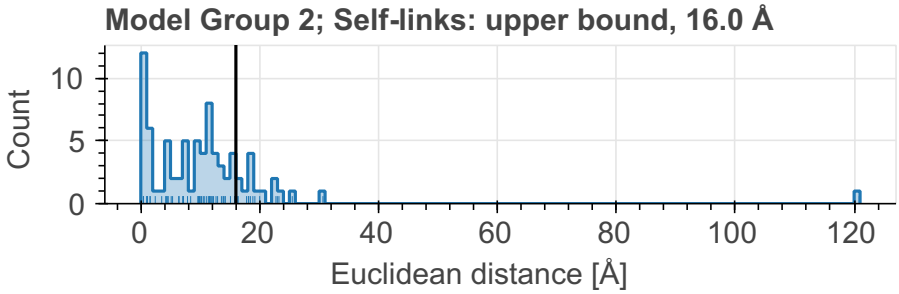


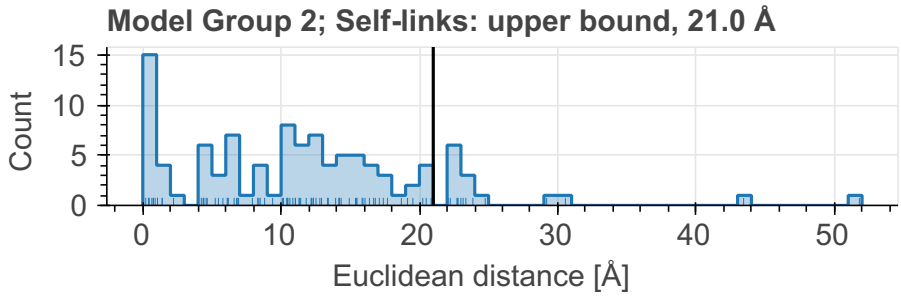

Model Group 1; Self-links: upper bound, 16.0 Å
Count
0
2
4
6
0
20
40
60
80
100
Euclidean distance [Å]


Model Group 1; Self-links: upper bound, 21.0 Å
Count
0
2
4
6
8
0
10
20
30
40
50
Euclidean distance [Å]


Model Group 2; Heteromeric links: upper bound, 16.0 Å
Count
0
2
4
0
50
100
150
200
250
Euclidean distance [Å]


Model Group 2; Heteromeric links: upper bound, 21.0 Å
Count
0
2
4
0
20
40
60
Euclidean distance [Å]


Model Group 2; Self-links: upper bound, 16.0 Å
Count
0
5
10
0
20
40
60
80
100
120
Euclidean distance [Å]



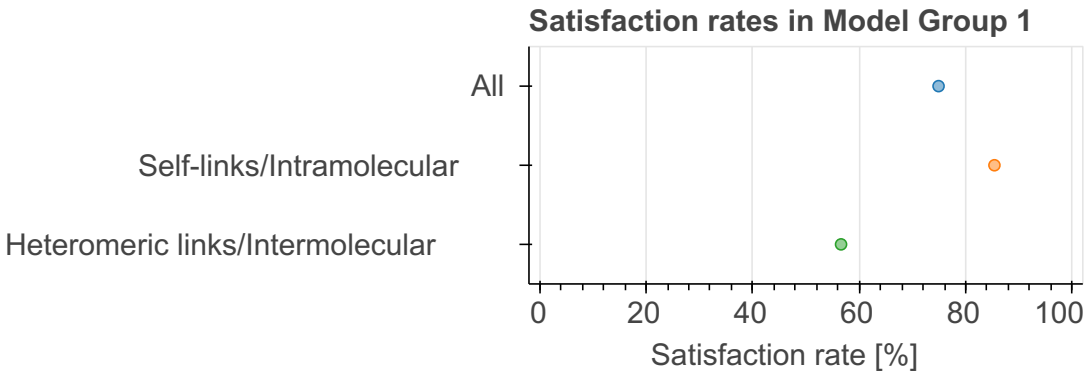
5.2.2. Satisfaction of restraints

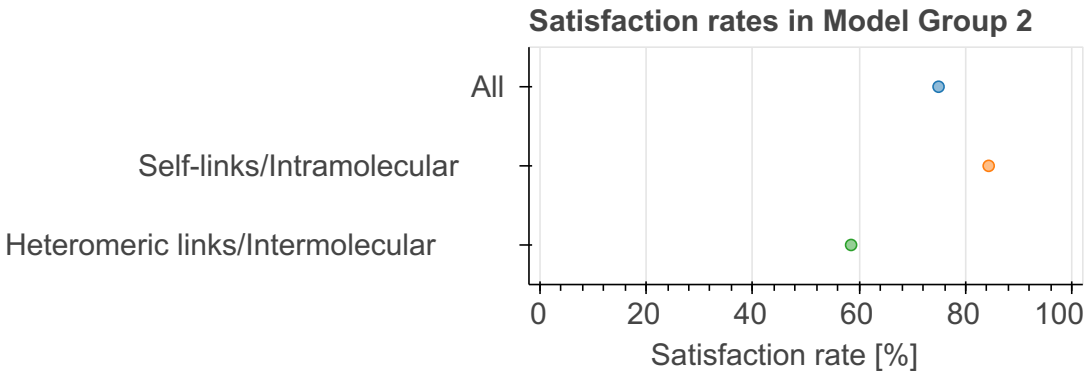
Satisfaction of restraints is calculated on a restraint group (a set of crosslinking restraints applied collectively in the modeling) level. Satisfaction of a restraint group depends on satisfaction of individual restraints in the group and the conditionality (all/any). A restraint group is considered 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=291)
1	1	1	1/1257	All	74.91	25.09	291
				Self-links/ Intramolecular	85.41	14.59	185
				Heteromeric links/ Intermolecular	56.60	43.40	106
1	1	2	1/1010	All	74.91	25.09	291
				Self-links/ Intramolecular	84.32	15.68	185
				Heteromeric links/ Intermolecular	58.49	41.51	106

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.





5.4. 2DEM class average

Validation for this section is under development.

5.4. EM raw micrographs

Validation for this section 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.