Integrative Structure Validation Report

March 13, 2025 - 12:39 PM PDT

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

Integrative Modeling Validation Version 2.0
Python-IHM Version 1.8

PDB ID	9A8M
PDB-Dev ID	PDBDEV_00000386
Structure Title	Modeling of Yeast NPC basket
Structure Authors	Singh, D.; Soni, N.; Hutchings, J.; Echeverria, I.; Shaikh, F.; Duquette, M.; Suslov, S.; Li, Z.; van Eeuwen, T.; Molloy, K.; Shi, Y.; Wang, J.; Guo, Q.; Chait, B.T.; Fernandez-Martinez, J.; Rout, M.P.; Sali, A.; Villa, E.
Deposited on	2024-07-07

This is a PDB-IHM IM Structure Validation Report for a publicly released PDB-IHM entry.

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

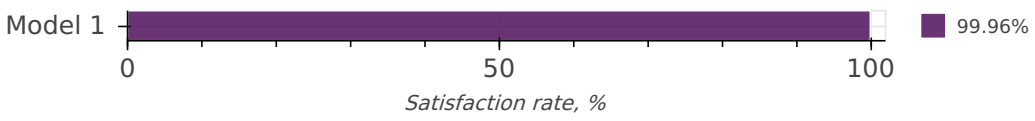
A user guide is available at https://pdb-ihm.org/validation_help.html with specific help available everywhere you see the ? symbol.

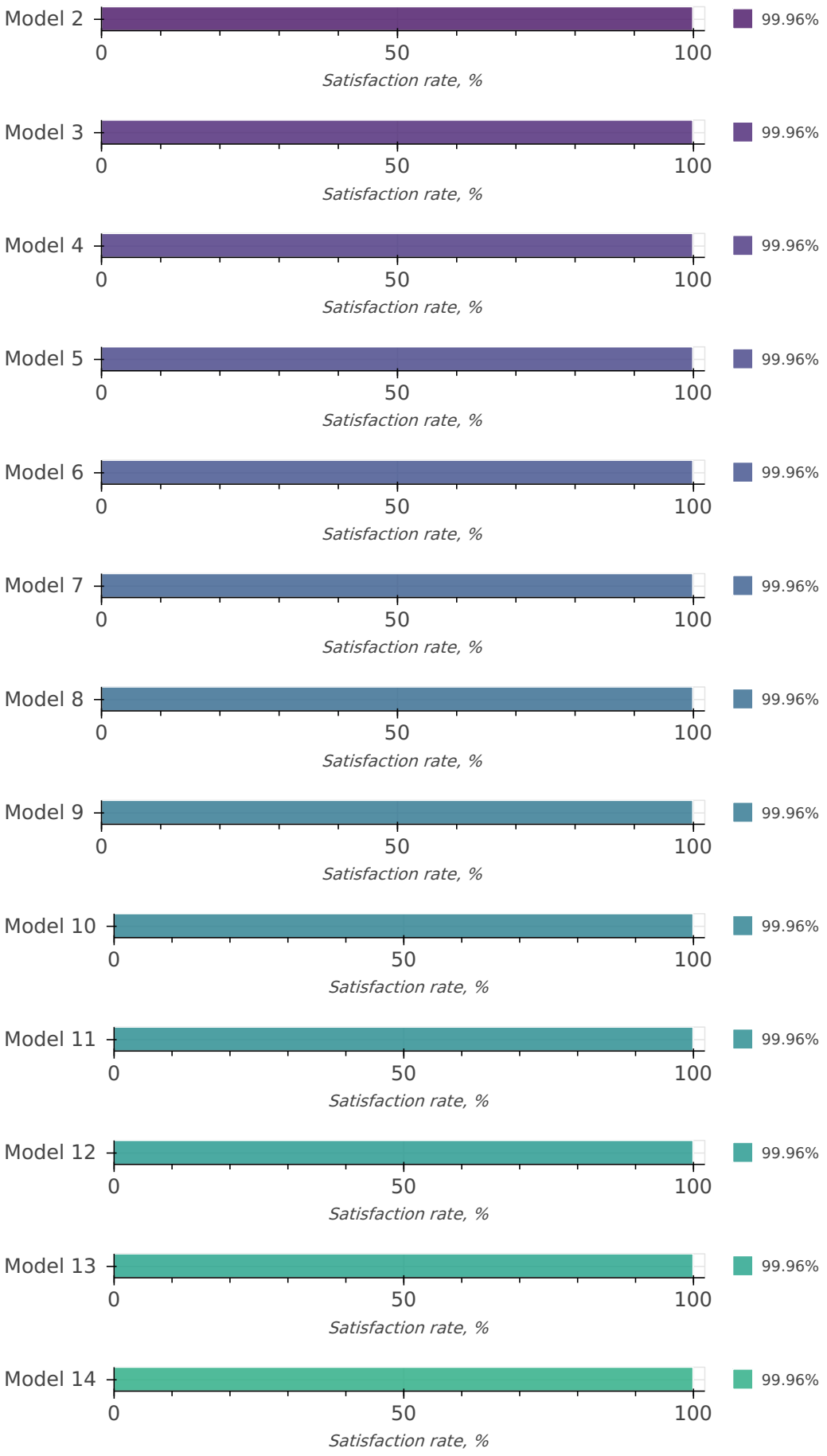
List of references used to build this report is available here.

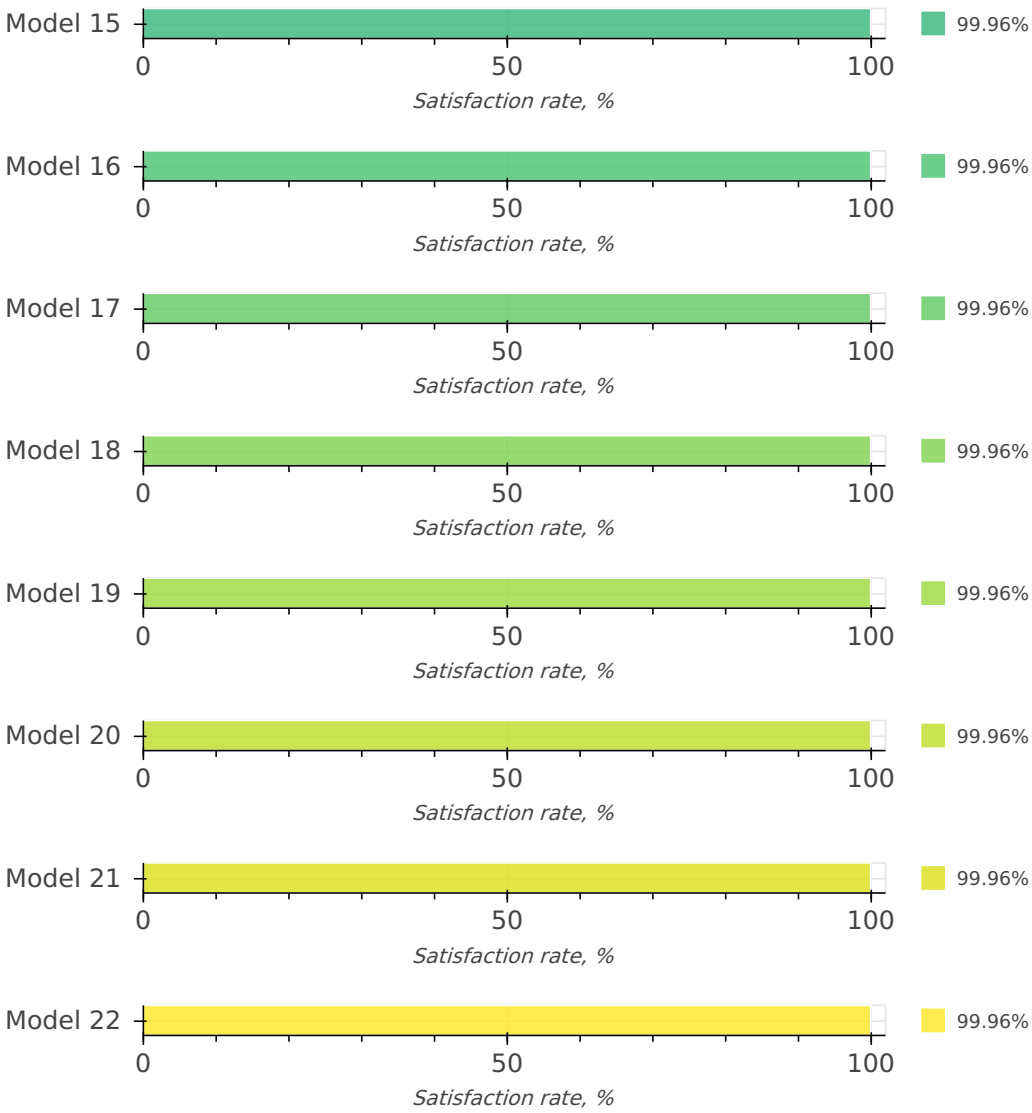
Overall quality

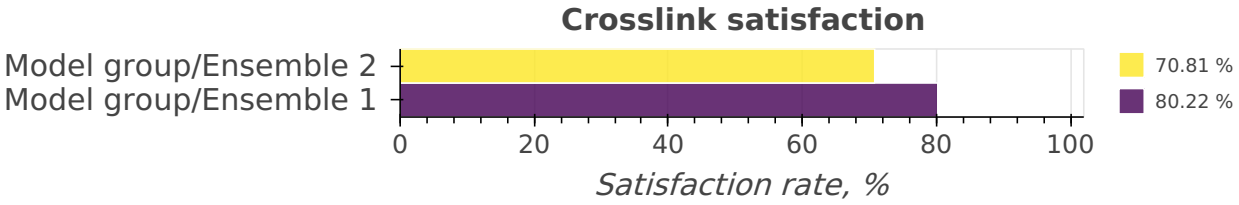
This validation report contains model quality assessments for all structures, data quality and fit to model assessments for SAS and crosslinking-MS datasets. Data quality and fit to model assessments for other datasets and model uncertainty are under development. Number of plots is limited to 256.

Model Quality: Excluded Volume Analysis




Model 2
99.96%
0
50
100
Satisfaction rate, %
Model 3
99.96%
0
50
100
Satisfaction rate, %
Model 4
99.96%
0
50
100
Satisfaction rate, %
Model 5
99.96%
0
50
100
Satisfaction rate, %
Model 6
99.96%
0
50
100
Satisfaction rate, %
Model 7
99.96%
0
50
100
Satisfaction rate, %
Model 8
99.96%
0
50
100
Satisfaction rate, %
Model 9
99.96%
0
50
100
Satisfaction rate, %
Model 10
99.96%
0
50
100
Satisfaction rate, %
Model 11
99.96%
0
50
100
Satisfaction rate, %
Model 12
99.96%
0
50
100
Satisfaction rate, %
Model 13
99.96%
0
50
100
Satisfaction rate, %
Model 14
99.96%
0
50
100
Satisfaction rate, %





Ensemble information

This entry consists of 1 distinct ensemble(s).

Summary

This entry consists of 22 model(s). A total of 11 datasets were used to build this entry.

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	1-22	1	Unknown MLP Protein	A B	1875	71-183, 197-221, 239-281, 285-324, 340-371, 435-463, 535-570, 575-617, 622-668, 689-717, 745-773, 791-840, 844-893, 930-986, 990-1088, 1092-1127, 1143-1199, 1212-1251, 1254-1286, 1290-1339, 1343-1385, 1408-1457	1-70, 184-196, 222-238, 282-284, 325-339, 372-434, 464-534, 571-574, 618-621, 669-688, 718-744, 774-790, 841-843, 894-929, 987-989, 1089-1091, 1128-1142, 1200-1211, 1252-1253, 1287-1289, 1340-1342, 1386-1407, 1458-1875	100.00 / 54.99	Multiscale: Coarse-grained: 1 - 50 residue(s) per bead
		2	Nucleoporin NUP1	C	1076	1-32, 85-104, 106-123	33-84, 105, 124-335	31.13 / 20.90	Multiscale: Coarse-grained: 1 - 30 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
		3	Nucleoporin NUP2	D E	720	83-136, 602-720	51-82	28.47 / 84.39	Multiscale: Coarse-grained: 1 - 30 residue(s) per bead
		4	Nucleoporin NUP60	F G	539	27-47, 91-104, 106-119, 121-140, 142-162	1-26, 48-90, 105, 120, 141, 163-398, 505-539	80.33 / 20.79	Multiscale: Coarse-grained: 1 - 30 residue(s) per bead
		5	Nucleoporin NUP120	H I	1037	1-29, 53-305, 311-711, 714-1036	30-52, 306-310, 712-713, 1037	100.00 / 97.01	Multiscale: Coarse-grained: 1 - 23 residue(s) per bead
		6	Nucleoporin NUP85	J	744	47-126, 132-230, 235-436, 451-744	1-46, 127-131, 231-234, 437-450	100.00 / 90.73	Multiscale: Coarse-grained: 1 - 46 residue(s) per bead
		6	Nucleoporin NUP85	K L	744	47-126, 132-230, 235-436, 451-739	1-46, 127-131, 231-234, 437-450, 740-744	100.00 / 90.05	Multiscale: Coarse-grained: 1 - 46 residue(s) per bead
		7	Nucleoporin NUP145C	M N	712	119-712	1-118	100.00 / 83.43	Multiscale: Coarse-grained: 1 - 50 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
		8	Protein transport protein SEC13	O P	297	8-157, 170-293	1-7, 158-169, 294-297	100.00 / 92.26	Multiscale: Coarse-grained: 1 - 12 residue(s) per bead
		9	Nucleoporin SEH1	Q R	349	1-248, 288-346	249-287, 347-349	100.00 / 87.97	Multiscale: Coarse-grained: 1 - 39 residue(s) per bead
		10	Nucleoporin NUP84	S T	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 - 22 residue(s) per bead
		11	Nucleoporin NUP133	U	1157	63-183, 198-480, 490-763, 772-1155	1-62, 184-197, 481-489, 764-771, 1156-1157	100.00 / 91.79	Multiscale: Coarse-grained: 1 - 50 residue(s) per bead
		11	Nucleoporin NUP133	V	1157	56-77, 86-125, 133-144, 162-184, 193-200, 206-249, 258-480, 490-763, 772-1155	1-55, 78-85, 126-132, 145-161, 185-192, 201-205, 250-257, 481-489, 764-771, 1156-1157	100.00 / 89.02	Multiscale: Coarse-grained: 1 - 50 residue(s) per bead

Datasets used for modeling

There are 11 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
11	3DEM volume	Zenodo	10.5281/zenodo.13131753
7	Crosslinking-MS data	Zenodo	10.5281/zenodo.13131753
8	Crosslinking-MS data	Zenodo	10.5281/zenodo.13131753
1	De Novo model	Zenodo	10.5281/zenodo.13131753
2	De Novo model	Zenodo	10.5281/zenodo.13131753
3	De Novo model	Zenodo	10.5281/zenodo.13131753
4	De Novo model	Zenodo	10.5281/zenodo.13131753
6	Experimental model	Zenodo	10.5281/zenodo.13131753
10	3DEM volume	Zenodo	10.5281/zenodo.13131753
5	Experimental model	PDB	7N84
9	3DEM volume	EMDB	EMD-44377

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	None	1000	False	True
2	1	Sampling	Replica exchange monte carlo	None	15467933	False	True

There are 3 software packages reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	IMP PMI module	2.19.0	integrative model building	https://integrativemodeling.org
3	COCONUT	1.0.0	Coiled-coil model building	https://github.com/neeleshsoni21/COCONUT
2	Integrative Modeling Platform (IMP)	2.19.0	integrative model building	https://integrativemodeling.org

Data quality

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.

3DEM volume

Validation for this section is under development.

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.

Excluded volume satisfaction

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	81071011	30075	99.96
2	81071011	29943	99.96
3	81071011	29978	99.96
4	81071011	30152	99.96
5	81071011	30057	99.96
6	81071011	29968	99.96
7	81071011	29990	99.96
8	81071011	29886	99.96
9	81071011	29843	99.96
10	81071011	29962	99.96
11	81071011	30032	99.96
12	81071011	30039	99.96
13	81071011	30066	99.96
14	81071011	29936	99.96
15	81071011	30002	99.96
16	81071011	30198	99.96

Model ID	Analysed	Number of violations	Excluded Volume Satisfaction (%)
17	81071011	29946	99.96
18	81071011	29795	99.96
19	81071011	29705	99.96
20	81071011	30271	99.96
21	81071011	30271	99.96
22	81071011	30075	99.96

Fit of model to data used for modeling

Fit of model(s) to crosslinking-MS data

Restraint types

Restraint types are summarized in the table below. 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.

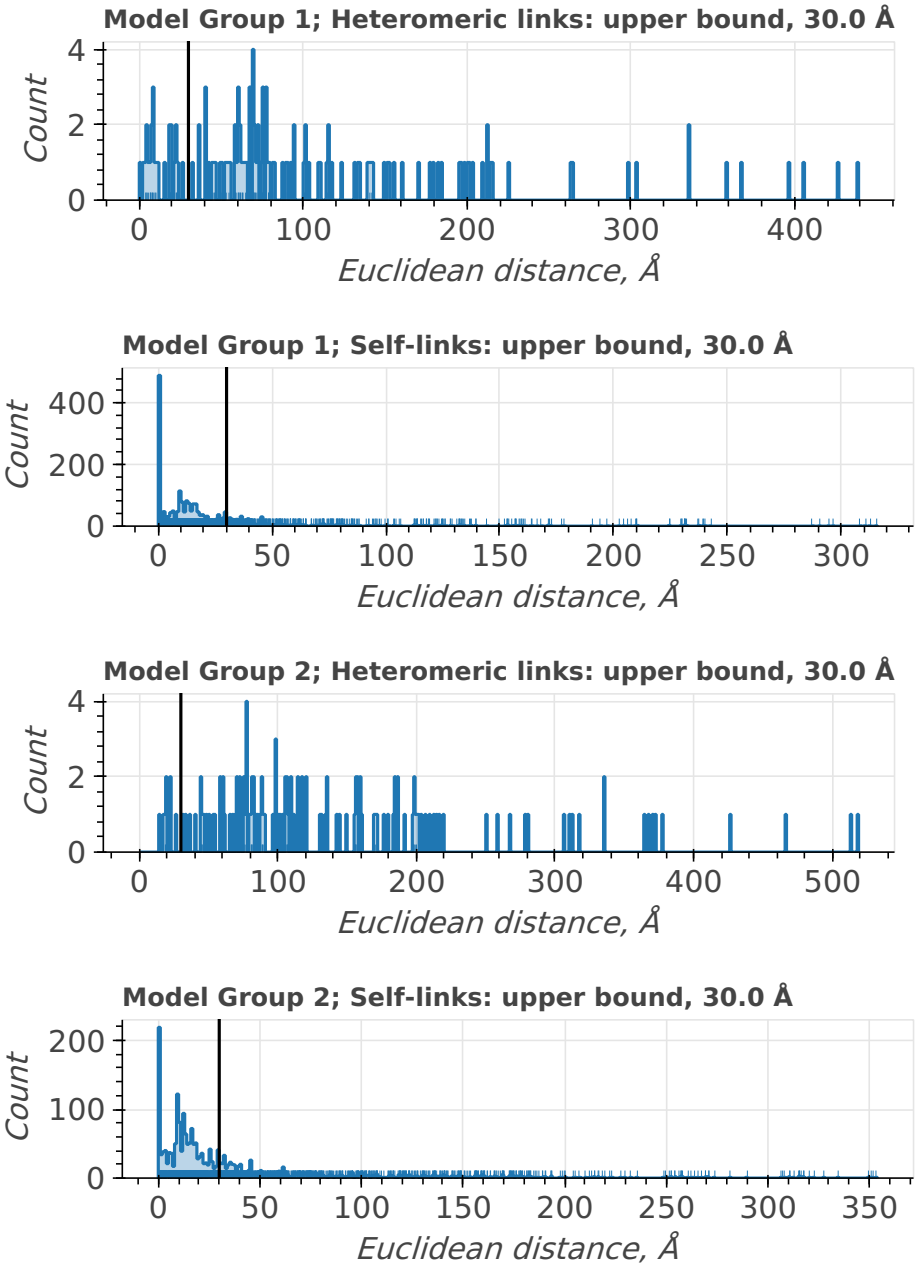
There are 2459 crosslinking restraints combined in 627 restraint groups.

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSS	UNK	CA	UNK	CA	upper bound	30.0	954
DSS	UNK	coarse-grained	UNK	coarse-grained	upper bound	30.0	1240
DSS	LYS	coarse-grained	UNK	coarse-grained	upper bound	30.0	38
DSS	LYS	CA	UNK	CA	upper bound	30.0	16
DSS	LYS	coarse-grained	LYS	coarse-grained	upper bound	30.0	145
DSS	LYS	CA	LYS	CA	upper bound	30.0	40
DSS	HIS	coarse-grained	LYS	coarse-grained	upper bound	30.0	4
DSS	THR	coarse-grained	UNK	coarse-grained	upper bound	30.0	12
DSS	LYS	coarse-grained	THR	coarse-grained	upper bound	30.0	6
DSS	LYS	coarse-grained	MET	coarse-grained	upper bound	30.0	4

Distograms of individual restraints

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.



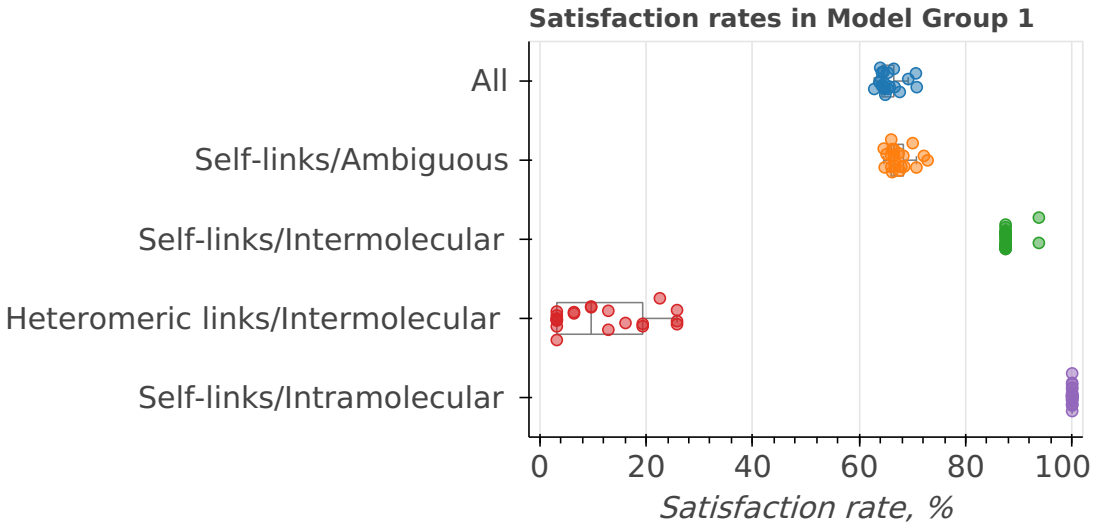
Satisfaction of restraints

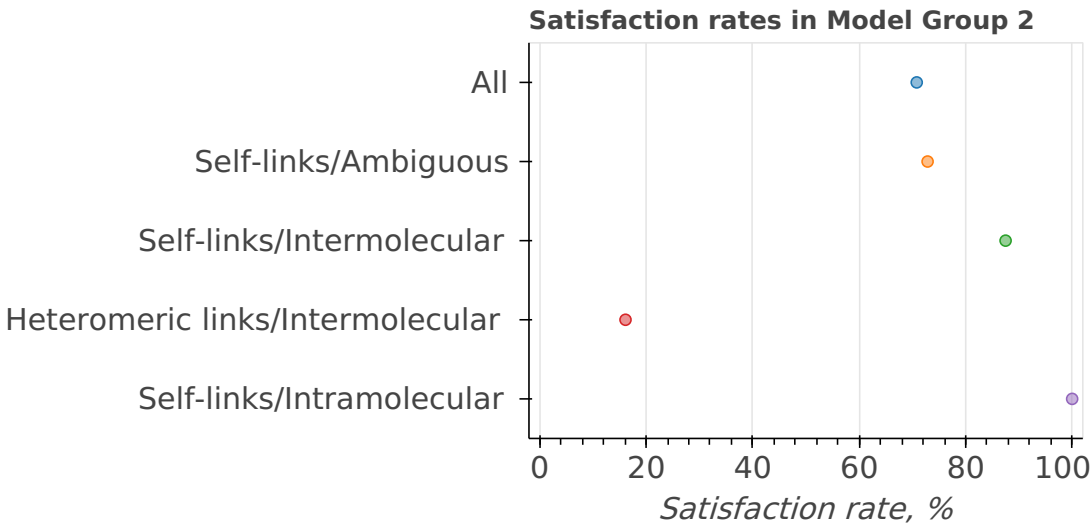
Satisfaction of restraints is calculated on a restraint group (a set of crosslinking restraints applied collectively in the modeling) level. Satisfaction of a restraint group depends on satisfaction of individual restraints in the group and the conditionality (all/any). A restraint group is considered 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=627)
1	1	1	21/21	All	80.22	19.78	627
				Self-links/ Ambiguous	81.79	18.21	571
				Self-links/ Intermolecular	93.75	6.25	16
				Heteromeric links/ Intermolecular	38.71	61.29	31
				Self-links/ Intramolecular	100.00	0.00	9
1	1	2	1/1	All	70.81	29.19	627
				Self-links/ Ambiguous	72.85	27.15	571
				Self-links/ Intermolecular	87.50	12.50	16
				Heteromeric links/ Intermolecular	16.13	83.87	31
				Self-links/ Intramolecular	100.00	0.00	9

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.





3DEM volume

Validation for this section is under development.

Fit of model to data used for validation

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.