

Integrative Structure Validation Report ?

February 18, 2025 - 08:26 AM PST

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

Integrative Modeling Validation Version 2.0

Python-IHM Version 1.8

PDB ID	8ZZ3
PDB-Dev ID	PDBDEV_00000003
Structure Title	Molecular architecture of the yeast Mediator complex
Structure Authors	Robinson PJ; Trnka MJ; Pellarin R; Greenberg CH; Bushnell DA; Davis R; Burlingame AL; Sali A; Kornberg RD
Deposited on	2016-08-31

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

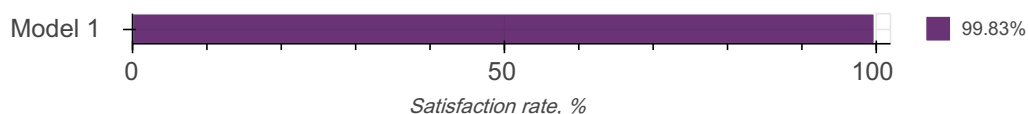
A user guide is available at https://pdb-ihm.org/validation_help.html with specific help available everywhere you see the ? symbol.

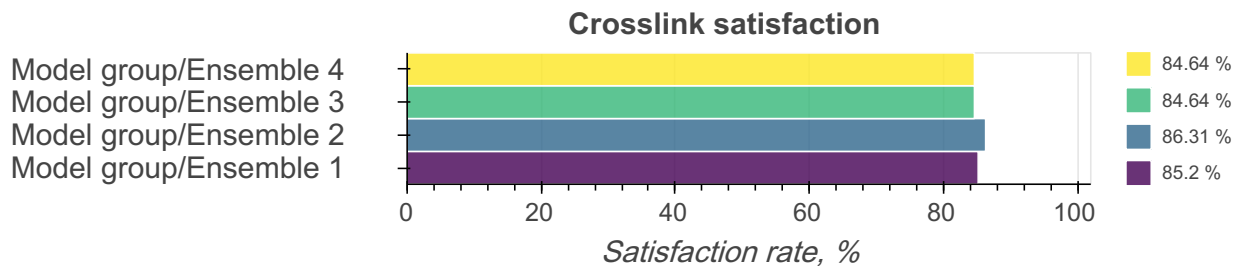
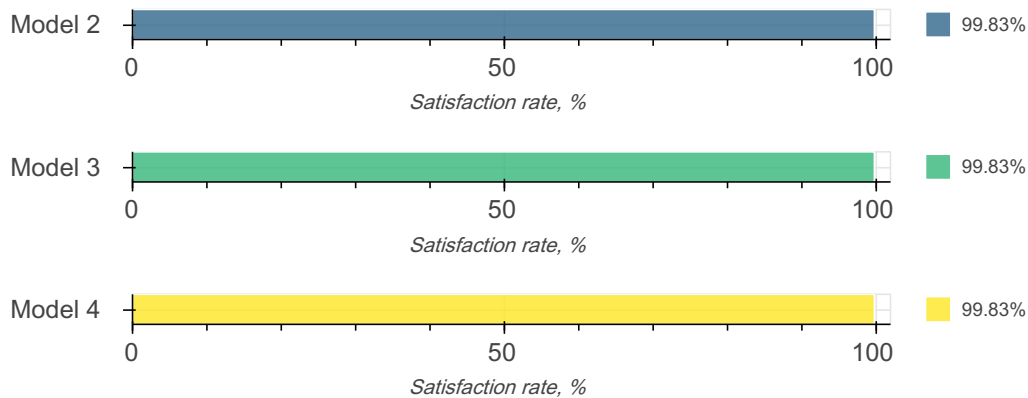
List of references used to build this report is available [here](#).

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





Ensemble information ?

This entry consists of 4 distinct ensemble(s).

Summary ?

This entry consists of 4 model(s). A total of 16 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	1-4	1	med6	A	295	-	1-60, 61-82, 83-192, 193-295	100.00 / 57.63	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
		2	med8	B	223	-	1-22, 23-173, 174-181, 182-214, 215-223	100.00 / 82.51	Multiscale: Coarse-grained: 1 - 20 residue(s) per bead
		3	med11	C	115	-	1-3, 4-115	100.00 / 97.39	Multiscale: Coarse-grained: 1 - 3 residue(s) per bead
		4	med17	D	687	-	1-122, 123-181, 182-371, 372-377, 378-661, 662-669, 670-687	100.00 / 71.62	Multiscale: Coarse-grained: 1 - 20 residue(s) per bead
		5	med18	E	307	-	1, 2-110, 111-157, 158-301, 302-307	100.00 / 82.41	Multiscale: Coarse-grained: 1 - 20 residue(s) per bead
		6	med20	F	210	-	1, 2-210	100.00 / 99.52	Coarse-grained: 1 residue(s) per bead
		7	med22	G	121	-	1-121	100.00 / 100.00	Coarse-grained: 1 residue(s) per bead
		8	med4	H	284	37-127	1-36, 128-284	100.00 / 32.04	Multiscale: Coarse-grained: 1 - 20 residue(s) per bead
		9	med7	I	222	12-84, 112-206	1-11, 85-111, 207-222	100.00 / 75.68	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
		10	med9	J	149	65-149	1-64	100.00 / 57.05	Multiscale: Coarse-grained: 1 - 20 residue(s) per bead
		11	med31	K	127	19-110	1-18, 111-127	100.00 / 72.44	Multiscale: Coarse-grained: 1 - 18 residue(s) per bead
		12	med21	L	140	2-128	1, 129-140	100.00 / 90.71	Multiscale: Coarse-grained: 1 - 12 residue(s) per bead
		13	med10	M	157	-	1-157	100.00 / 0.00	Multiscale: Coarse-grained: 17 - 20 residue(s) per bead
		14	med1	N	566	-	1-566	100.00 / 0.00	Multiscale: Coarse-grained: 6 - 20 residue(s) per bead
		15	med14	O	1082	-	1-1082	100.00 / 0.00	Multiscale: Coarse-grained: 11 - 40 residue(s) per bead
		16	med19	P	220	-	1-220	100.00 / 0.00	Coarse-grained: 20 residue(s) per bead
		17	med2	Q	436	-	1-436	100.00 / 0.00	Multiscale: Coarse-grained: 36 - 40 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
		18	med3	R	401	-	1-401	100.00 / 0.00	Multiscale: Coarse-grained: 1 - 40 residue(s) per bead
		19	med5	S	1146	-	1-1146	100.00 / 0.00	Multiscale: Coarse-grained: 26 - 40 residue(s) per bead
		20	med15	T	1094	-	1-1094	100.00 / 0.00	Multiscale: Coarse-grained: 14 - 40 residue(s) per bead
		21	med16	U	986	8-49, 94-150, 165-174, 231-406, 437-476, 503-538	50-93, 151-164, 175-230, 407-436, 477-502, 539-986	99.29 / 36.87	Multiscale: Coarse-grained: 1 - 40 residue(s) per bead

Datasets used for modeling ?

There are 16 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	Experimental model	PDB	4GWP
2	Comparative model	Not available	10.1093/nar/gkt704
3	Comparative model	Zenodo	10.5281/zenodo.802915
4	Experimental model	PDB	3FBI
5	Experimental model	Zenodo	10.5281/zenodo.802915
6	Experimental model	PDB	1YKH
7	Experimental model	Zenodo	10.5281/zenodo.802915
8	Experimental model	PDB	4BZK
9	Comparative model	Zenodo	10.5281/zenodo.802915
10	Mass Spectrometry data	MASSIVE	MSV000079237

ID	Dataset type	Database name	Data access code
11	Crosslinking-MS data	Zenodo	10.5281/zenodo.802915
12	3DEM volume	EMDB	EMD-2634
13	3DEM volume	Zenodo	10.5281/zenodo.802915
14	3DEM volume	Zenodo	10.5281/zenodo.802915
15	3DEM volume	Zenodo	10.5281/zenodo.802915
16	3DEM volume	Zenodo	10.5281/zenodo.802915

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	20000	False	True

There are 5 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	Protein Prospector	5.13.1	mass spectrometry	http://prospector.ucsf.edu/
4	Situs	2.70	density map fitting	http://situs.biomachina.org/
5	Phyre2	2.00	protein homology modeling	http://www.sbg.bio.ic.ac.uk/~phyre2/

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.

Mass Spectrometry

Validation for this section is under development.

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-particle or particle-atom pairs for which excluded volume was analysed.

Model ID	Analysed	Number of violations	Excluded Volume Satisfaction (%)
1	3673405	6332	99.83
2	3673405	6318	99.83
3	3673405	6347	99.83
4	3673405	6337	99.83

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.*

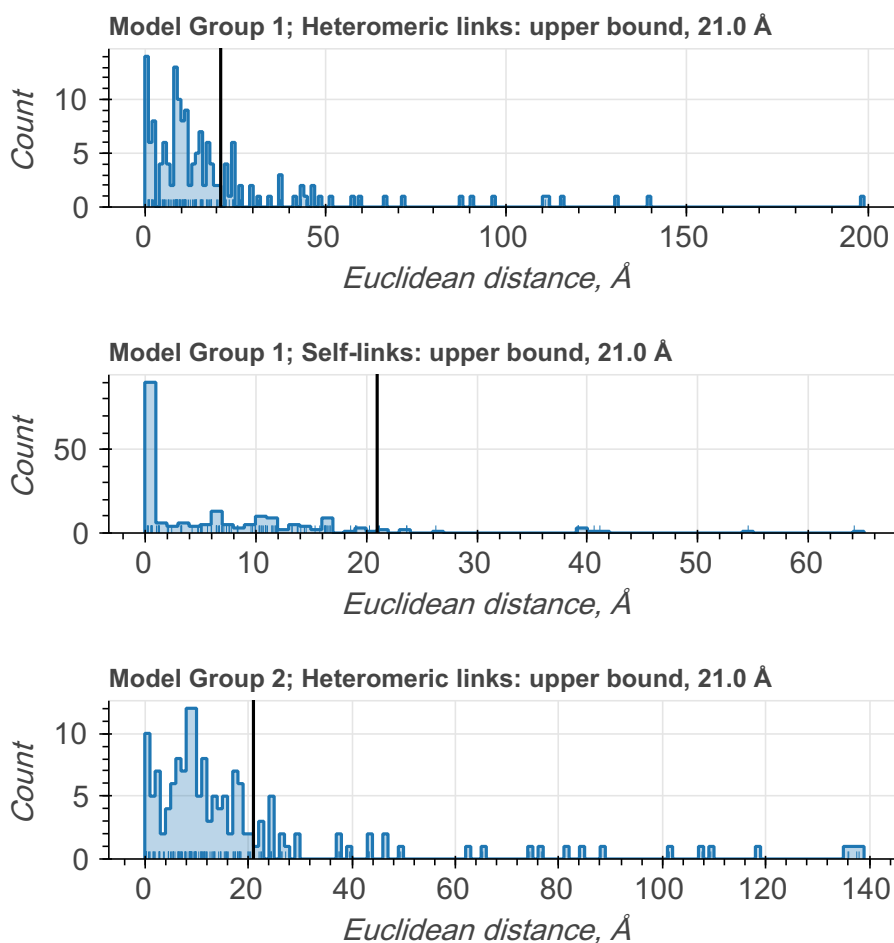
There are 359 crosslinking restraints combined in 359 restraint groups.

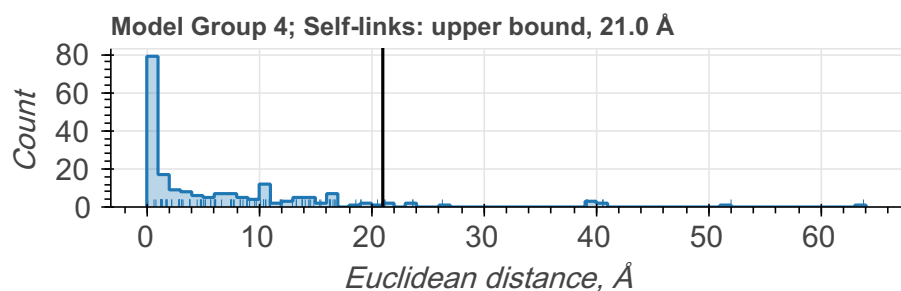
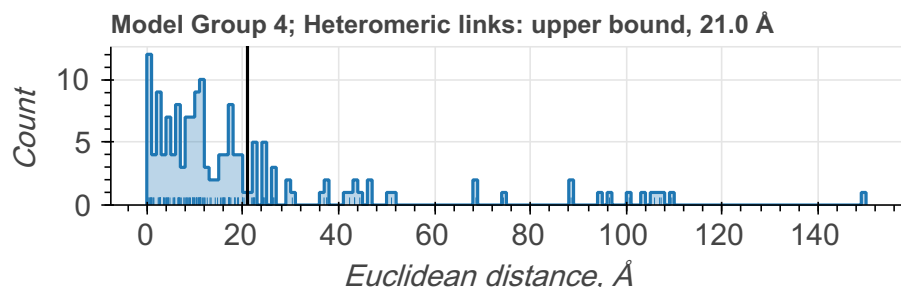
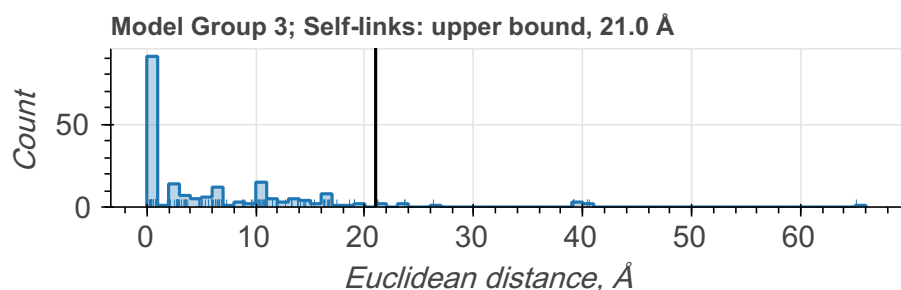
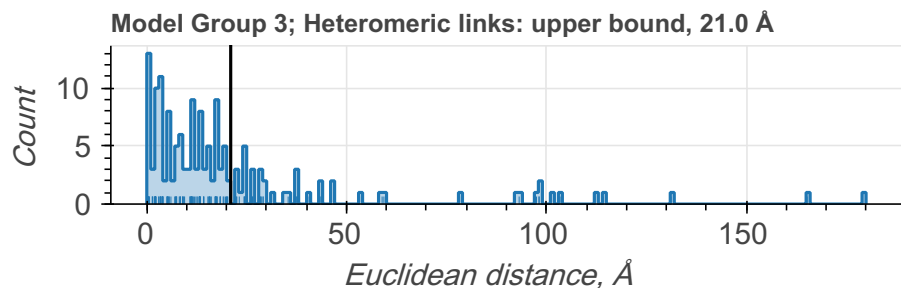
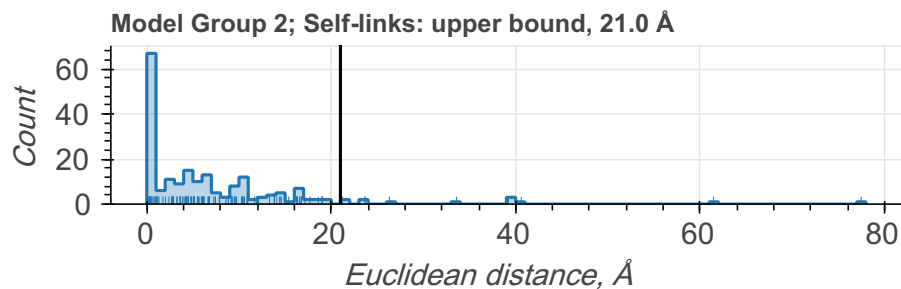
Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSS	LYS	coarse-grained	LYS	coarse-grained	upper bound	21.0	225
DSS	LYS	coarse-grained	MET	coarse-grained	upper bound	21.0	22
DSS	LYS	CA	LYS	CA	upper bound	21.0	86
DSS	GLU	CA	LYS	CA	upper bound	21.0	3
DSS	LYS	CA	VAL	CA	upper bound	21.0	6
DSS	ILE	CA	LYS	CA	upper bound	21.0	4
DSS	LEU	CA	LYS	CA	upper bound	21.0	5

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSS	ILE	CA	VAL	CA	upper bound	21.0	1
DSS	GLU	CA	VAL	CA	upper bound	21.0	2
DSS	ALA	CA	ILE	CA	upper bound	21.0	1
DSS	LYS	CA	MET	CA	upper bound	21.0	1
DSS	LYS	coarse-grained	THR	coarse-grained	upper bound	21.0	2
DSS	LYS	coarse-grained	SER	coarse-grained	upper bound	21.0	1

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.





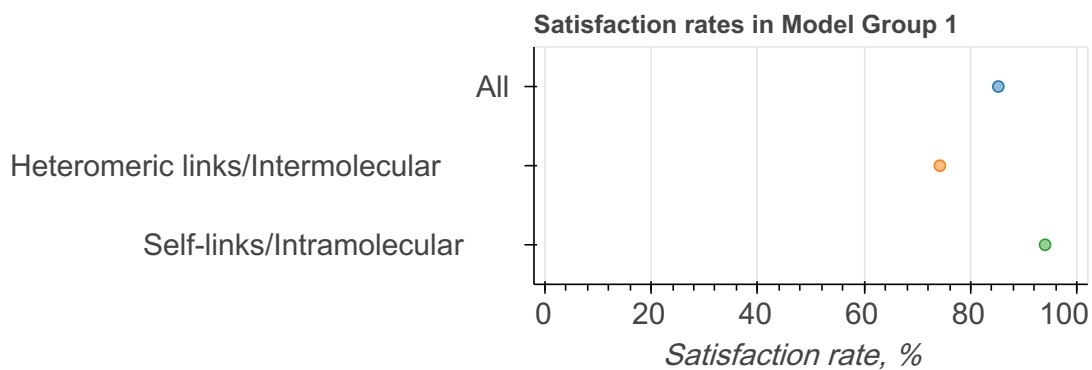
Satisfaction of restraints

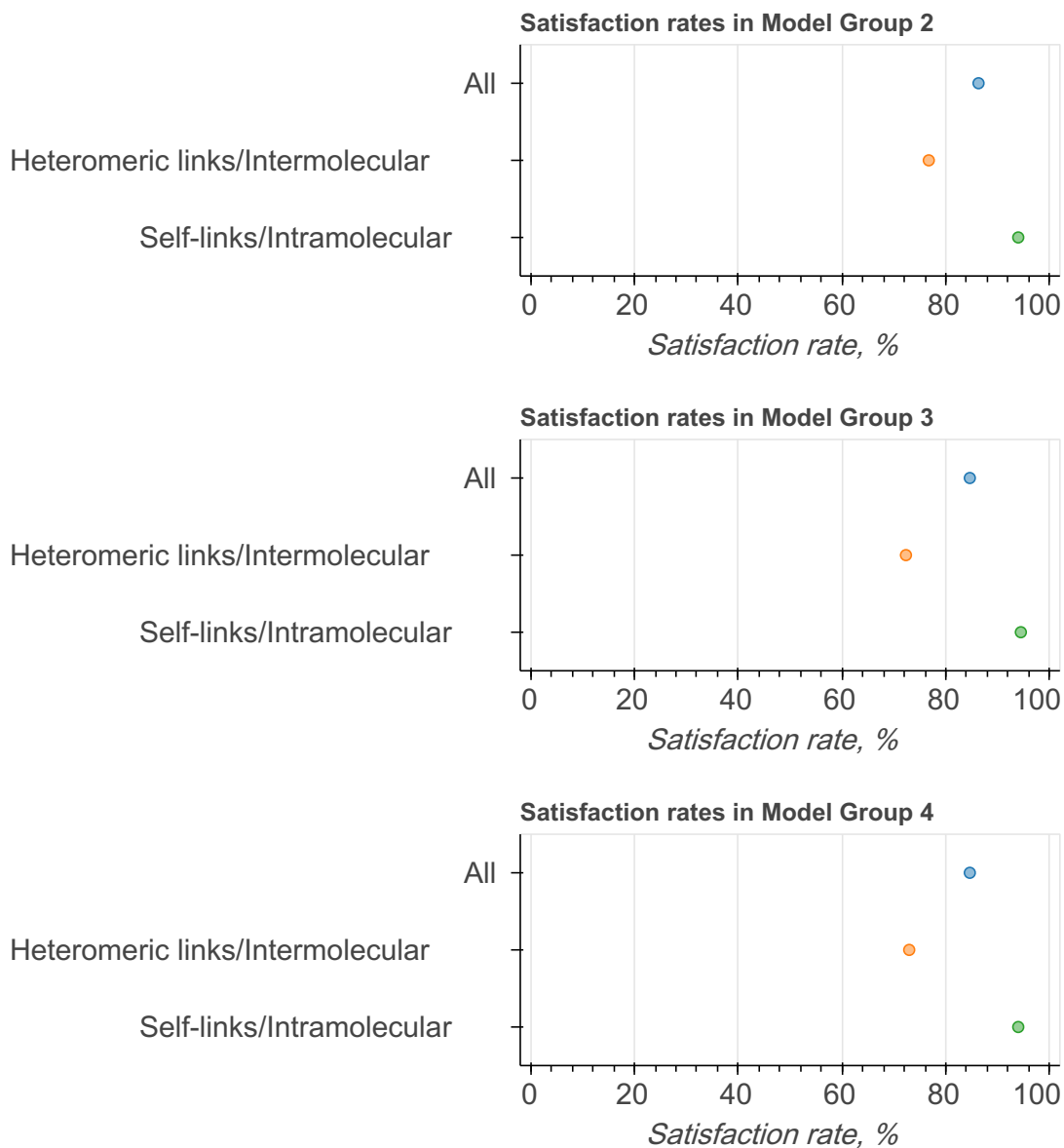
Satisfaction of restraints is calculated on a [restraint group](#) (a set of crosslinking restraints applied collectively in the modeling) level. Satisfaction of a restraint group depends on satisfaction of individual restraints in the group and the conditionality (all/any). A restraint group is considered 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=359)
1	1	1	1/142	All	85.20	14.80	358
				Heteromeric links/ Intermolecular	74.21	25.79	159
				Self-links/ Intramolecular	93.97	6.03	199
1	1	2	1/192	All	86.31	13.69	358
				Heteromeric links/ Intermolecular	76.73	23.27	159
				Self-links/ Intramolecular	93.97	6.03	199
1	1	3	1/39	All	84.64	15.36	358
				Heteromeric links/ Intermolecular	72.33	27.67	159
				Self-links/ Intramolecular	94.47	5.53	199
1	1	4	1/126	All	84.64	15.36	358
				Heteromeric links/ Intermolecular	72.96	27.04	159
				Self-links/ Intramolecular	93.97	6.03	199

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.





Mass Spectrometry

Validation for this section is under development.

3DEM volume

Validation for this section is under development.

Fit of model to data used for validation ?

Validation for this section is under development.

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 Trehwella, 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.