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	8ZZ2
PDB-Dev ID	PDBDEV_00000002
Structure Title	Structure of Saccharomyces cerevisiae exosome determined with CX-MS
Structure Authors	Shi Y; Pellarin R; Fridy PC; Fernandez-Martinez J; Thompson MK; Li Y; Wang QJ; Sali A; Rout MP; Chait BT
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

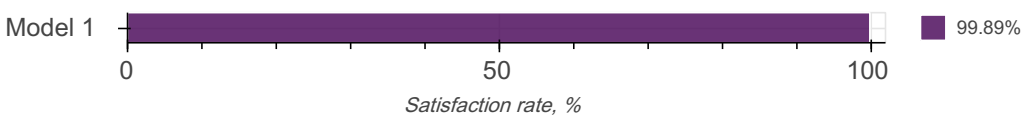
A user guide is available at https://pdb-ihm.org/validation_help.html with specific help available everywhere you see the ? symbol.

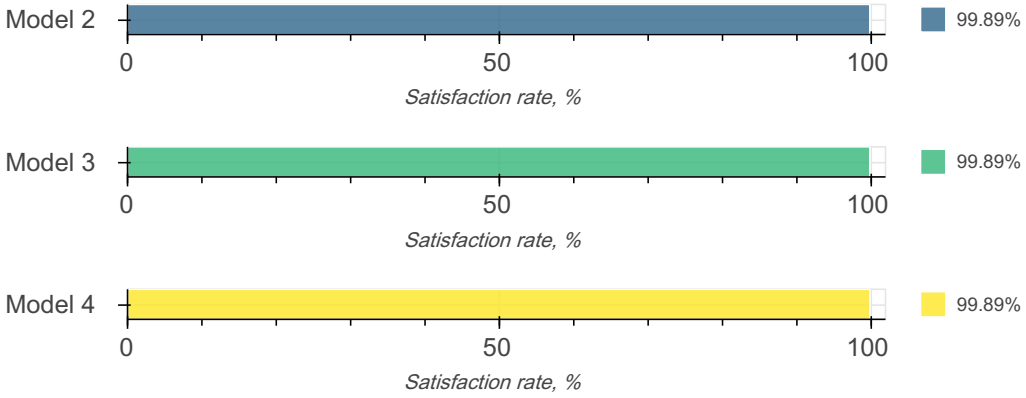
List of references used to build this report is available here.

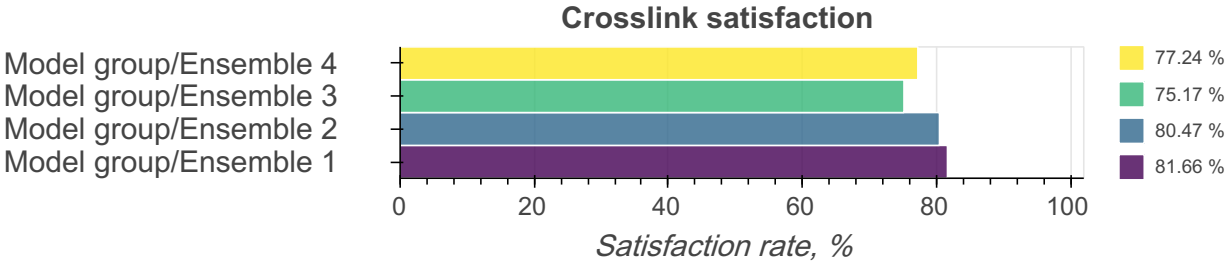
Overall quality

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 5 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	Dis3	A	1001	9-237, 249-329, 364-471, 472-1001	1-8, 238-248, 330-363	100.00 / 94.71	Multiscale: Coarse-grained: 1 - 5 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	Rrp45	B	305	2-301	1, 302-305	100.00 / 98.36	Multiscale: Coarse-grained: 1 - 4 residue(s) per bead
		3	Rrp4	C	359	2-17, 50-102, 103-245, 275-357	1, 18-49, 246-274, 358-359	100.00 / 82.17	Multiscale: Coarse-grained: 1 - 5 residue(s) per bead
		4	Csl4	D	292	1-71, 99-113, 126-162, 185-291	72-98, 114-125, 163-184, 292	100.00 / 78.77	Multiscale: Coarse-grained: 1 - 5 residue(s) per bead
		5	Mtr3	E	250	4-22, 42-149, 163-248	1-3, 23-41, 150-162, 249-250	100.00 / 85.20	Multiscale: Coarse-grained: 1 - 5 residue(s) per bead
		6	Rrp40	F	240	1-60, 61-236	237-240	100.00 / 98.33	Multiscale: Coarse-grained: 1 - 4 residue(s) per bead
		7	Rrp42	G	265	1-265	-	100.00 / 100.00	Coarse-grained: 1 residue(s) per bead
		8	Ski6	H	265	1-242	243-265	100.00 / 91.32	Multiscale: Coarse-grained: 1 - 5 residue(s) per bead
		9	Rrp46_gfp	I	475	1-223, 247-475	224-246	100.00 / 95.16	Multiscale: Coarse-grained: 1 - 5 residue(s) per bead
		10	Rrp43	J	394	7-99, 121-193, 206-309, 327-394	1-6, 100-120, 194-205, 310-326	100.00 / 85.79	Multiscale: Coarse-grained: 1 - 5 residue(s) per bead
		11	Lrp1	K	184	-	1-184	100.00 / 0.00	Multiscale: Coarse-grained: 4 - 5 residue(s) per bead
		12	Rrp6	L	733	127-516, 532-557, 565-619	1-126, 517-531, 558-564, 620-733	100.00 / 64.26	Multiscale: Coarse-grained: 1 - 5 residue(s) per bead
		13	MPP6	M	186	-	1-186	100.00 / 0.00	Multiscale: Coarse-grained: 1 - 5 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
		15	Ski7	N	747	259-747	1-258	100.00 / 65.46	Multiscale: Coarse-grained: 1 - 5 residue(s) per bead

Datasets used for modeling

There are 5 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	Experimental model	PDB	4IFD
2	Experimental model	PDB	1GFL
3	Experimental model	PDB	2HBJ
4	Comparative model	Zenodo	10.5281/zenodo.583313
5	Crosslinking-MS data	Zenodo	10.5281/zenodo.583313

Methodology and software

This entry is a result of 2 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	50000	True	True
1	2	Sampling	Replica exchange monte carlo	None	50000	True	True

There are 3 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	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.

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	8809503	9530	99.89
2	8809503	9561	99.89
3	8563591	9595	99.89
4	8563591	9583	99.89

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.

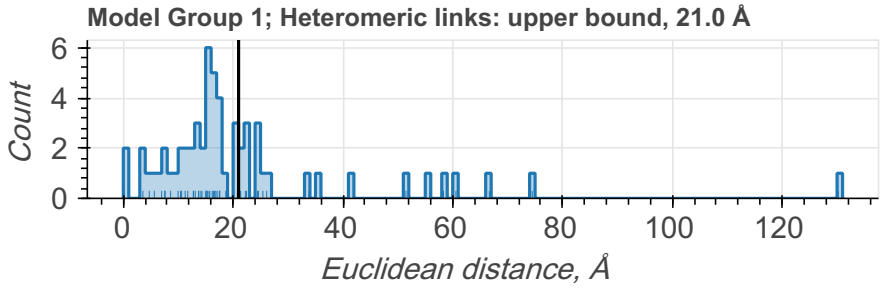
There are 216 crosslinking restraints combined in 216 restraint groups.

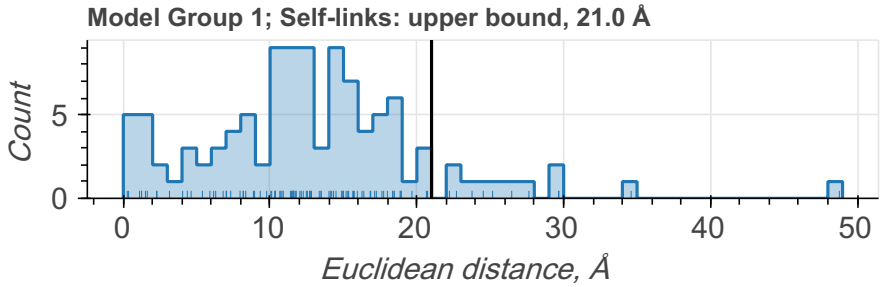
Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSS	LYS	CA	LYS	CA	upper bound	21.0	115

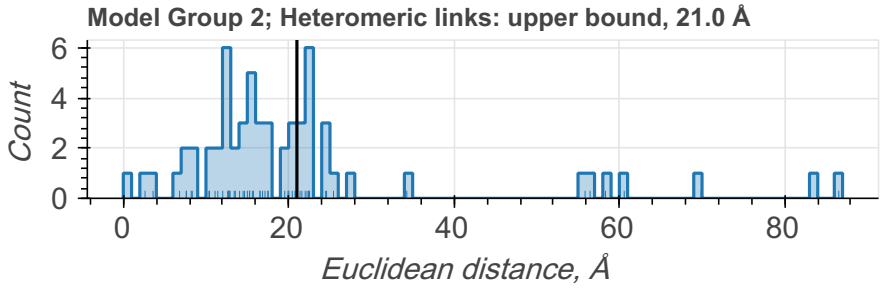
Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSS	LYS	coarse-grained	LYS	coarse-grained	upper bound	21.0	97
DSS	GLY	coarse-grained	LYS	coarse-grained	upper bound	21.0	1
DSS	GLY	CA	LYS	CA	upper bound	21.0	3

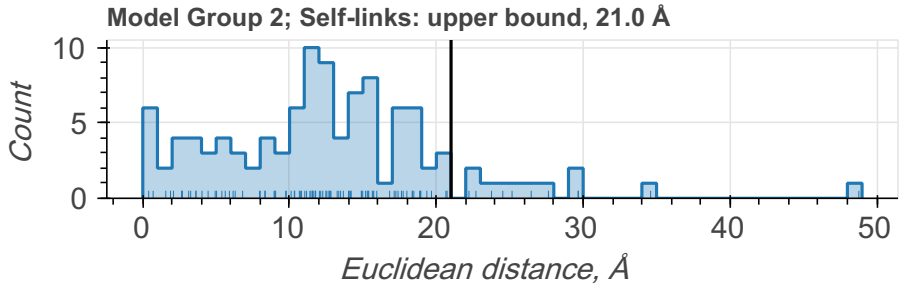
Distograms of individual restraints

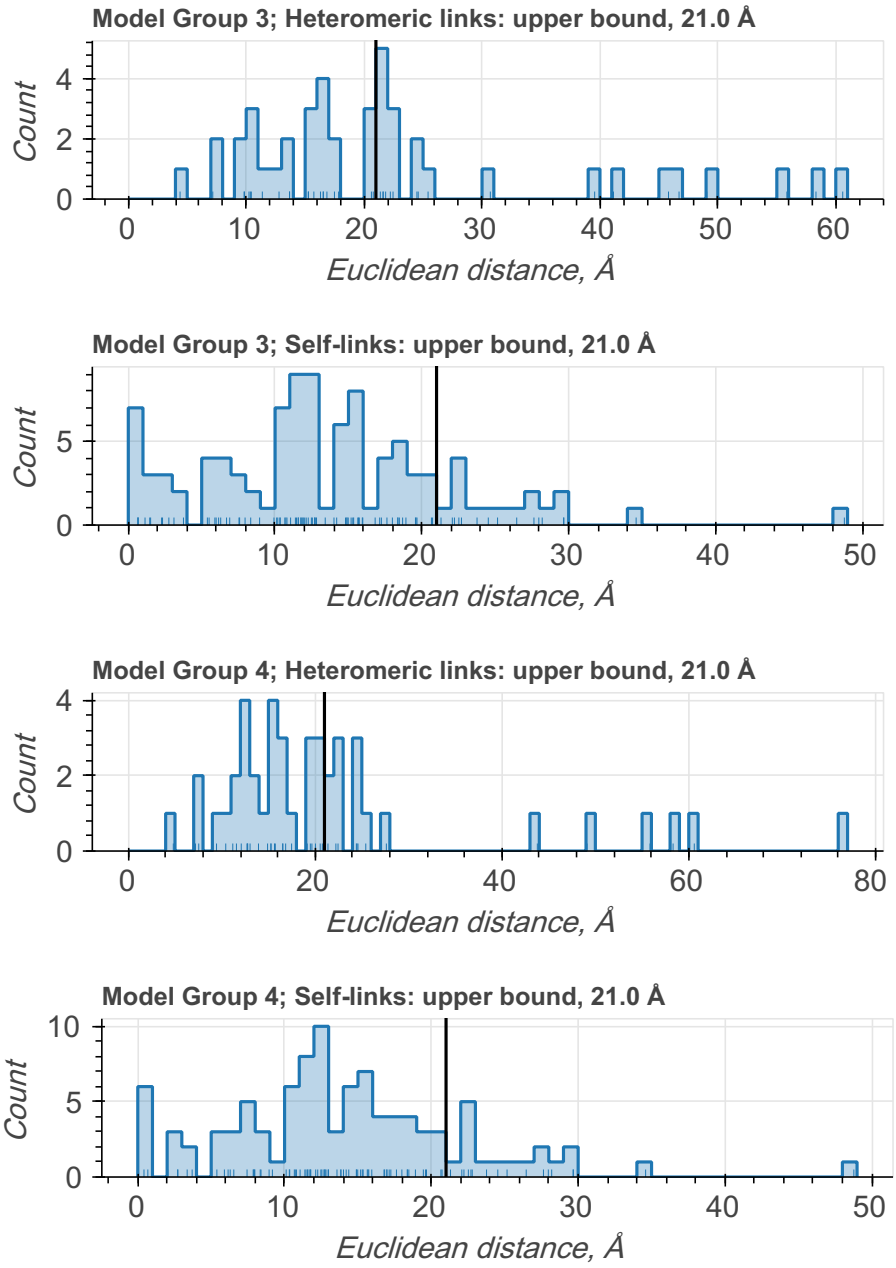
Restraints with identical thresholds are grouped into one plot. Only the best distance per restraint per model group/ensemble is plotted. Inter- and intramolecular (including self-links) restraints are also grouped into one plot. Distance for a restraint between coarse-grained beads is calculated as 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.











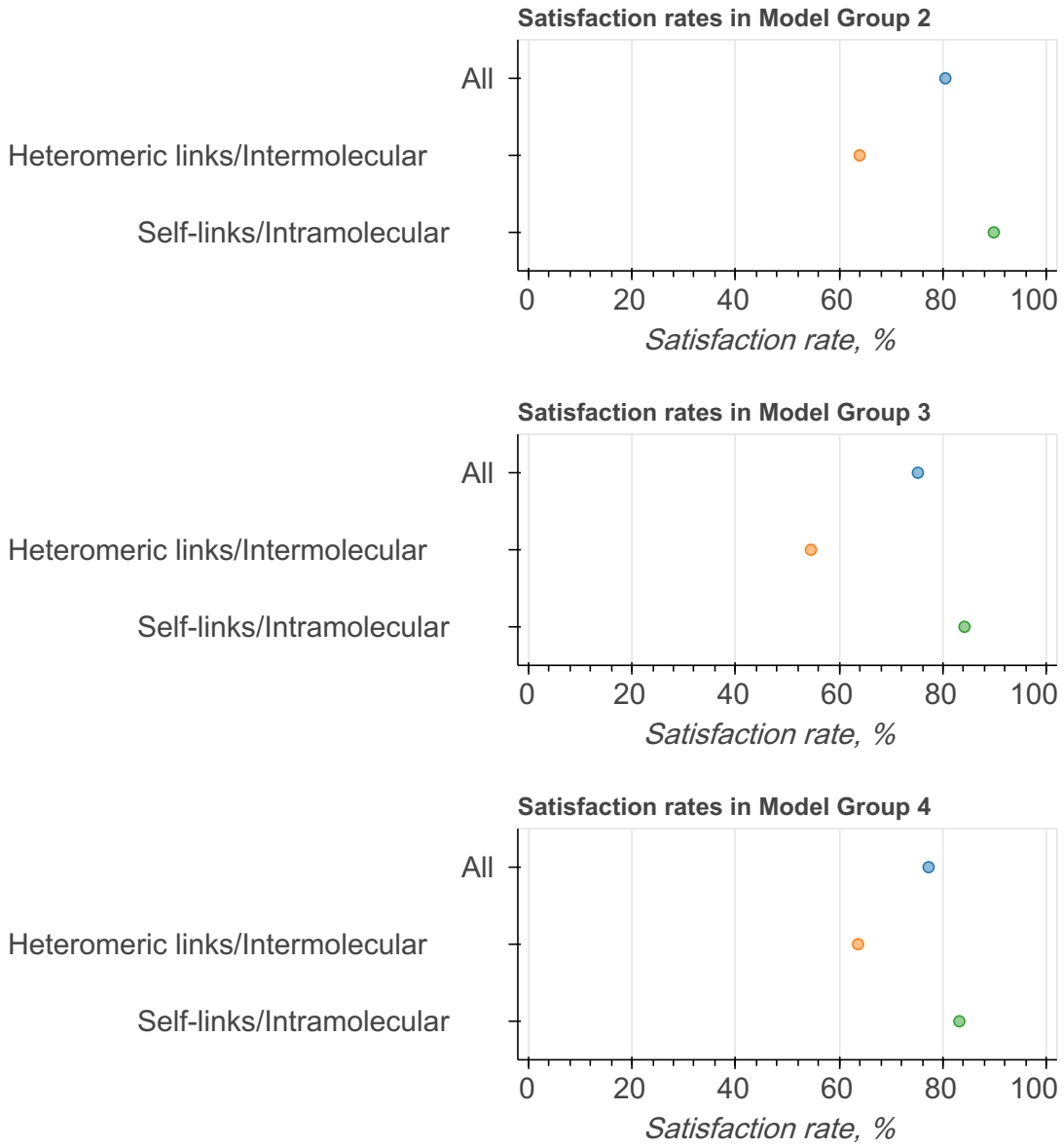
Satisfaction of restraints

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=216)
1	1	1	1/69	All	81.66	18.34	169
				Heteromeric links/ Intermolecular	67.21	32.79	61
				Self-links/ Intramolecular	89.81	10.19	108
1	1	2	1/131	All	80.47	19.53	169
				Heteromeric links/ Intermolecular	63.93	36.07	61
				Self-links/ Intramolecular	89.81	10.19	108
1	2	3	1/159	All	75.17	24.83	145
				Heteromeric links/ Intermolecular	54.55	45.45	44
				Self-links/ Intramolecular	84.16	15.84	101
1	2	4	1/41	All	77.24	22.76	145
				Heteromeric links/ Intermolecular	63.64	36.36	44
				Self-links/ Intramolecular	83.17	16.83	101

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.



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.