

Integrative Structure Validation Report ?

February 18, 2025 - 08:31 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	9A19
PDB-Dev ID	PDBDEV_00000081
Structure Title	Integrative structure of Smc5/6 complex
Structure Authors	Yu Y; Li S; Ser Z; Sanyal T; Choi K; Wan B; Kuang H; Sali A; Kentsis A; Patel DJ; Zhao X
Deposited on	2021-04-13

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

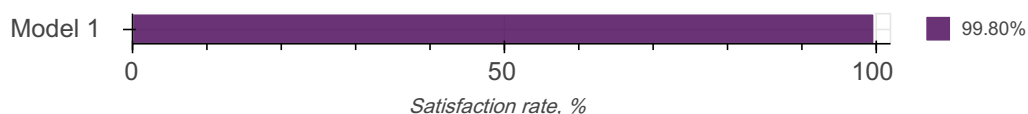
A user guide is available at https://pdb-ihm.org/validation_help.html with specific help available everywhere you see the ? symbol.

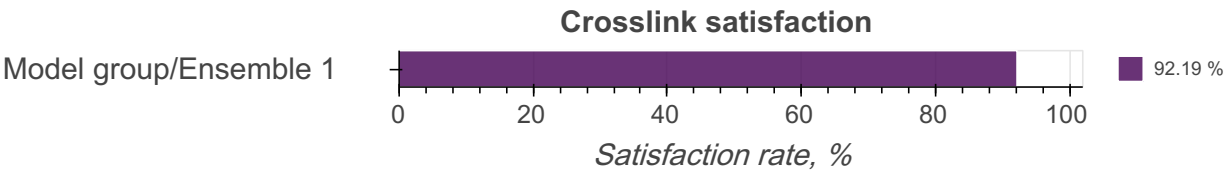
List of references used to build this report is available [here](#).

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 1 distinct ensemble(s).

Summary ?

This entry consists of 1 model(s). A total of 21 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	1	smc5	A	1093	42-204, 208-266, 272-302, 304-363, 365-388, 398-459, 485-633, 653-676, 696-714, 719-737, 739-811, 851-880, 885-946, 950-1093	1-41, 205-207, 267-271, 303, 364, 389-397, 460-484, 634-652, 677-695, 715-718, 738, 812-850, 881-884, 947-949	100.00 / 84.08	Multiscale: Coarse-grained: 1 - 41 residue(s) per bead
		2	smc6	B	1114	80-232, 236-294, 298-310, 314-430, 436-501, 506-692, 699-760, 770-882, 897-916, 923-939, 951-984, 988-1114	1-79, 233-235, 295-297, 311-313, 431-435, 502-505, 693-698, 761-769, 883-896, 917-922, 940-950, 985-987	100.00 / 86.89	Multiscale: Coarse-grained: 1 - 79 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	nse2	C	267	5-258	1-4, 259-267	100.00 / 95.13	Multiscale: Coarse-grained: 1 - 9 residue(s) per bead
		4	nse5	D	556	22-142, 185-257, 274-441	1-21, 143-184, 258-273, 442-556	100.00 / 65.11	Multiscale: Coarse-grained: 1 - 115 residue(s) per bead
		5	nse6	E	464	195-226, 238-379, 387-427, 431-441, 447-464	1-194, 227-237, 380-386, 428-430, 442-446	100.00 / 52.59	Multiscale: Coarse-grained: 1 - 20 residue(s) per bead

Datasets used for modeling ?

There are 21 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	Experimental model	PDB	6QPW
2	Experimental model	PDB	6ZZ6
3	Comparative model	Zenodo	10.5281/zenodo.4685414
4	Comparative model	Zenodo	10.5281/zenodo.4685414
5	Comparative model	Zenodo	10.5281/zenodo.4685414
6	Experimental model	PDB	3HTK
7	Comparative model	Zenodo	10.5281/zenodo.4685414
8	Comparative model	Zenodo	10.5281/zenodo.4685414
9	Experimental model	PDB	5MG8
10	Comparative model	Zenodo	10.5281/zenodo.4685414
11	Comparative model	Zenodo	10.5281/zenodo.4685414
12	Comparative model	Zenodo	10.5281/zenodo.4685414

ID	Dataset type	Database name	Data access code
13	Comparative model	Zenodo	10.5281/zenodo.4685414
14	Comparative model	Zenodo	10.5281/zenodo.4685414
15	Comparative model	Zenodo	10.5281/zenodo.4685414
16	Experimental model	PDB	7LTO
17	Experimental model	Zenodo	10.5281/zenodo.4685414
18	Mass Spectrometry data	PRIDE	PXD023164
19	Crosslinking-MS data	Zenodo	10.5281/zenodo.4685414
20	Crosslinking-MS data	Zenodo	10.5281/zenodo.4685414
21	Crosslinking-MS data	Zenodo	10.5281/zenodo.4685414

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
2	1	Sampling	Replica exchange monte carlo	None	100000	False	True

There are 4 software packages reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	IMP PMI module	develop-31a0ad09b4	integrative model building	https://integrativemodeling.org
2	Integrative Modeling Platform (IMP)	develop-31a0ad09b4	integrative model building	https://integrativemodeling.org
3	MODELLER	9.25	comparative modeling	https://salilab.org/modeller/
4	SWISS-MODEL	3.0.0	protein homology modeling	https://swissmodel.expasy.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.

Mass Spectrometry

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	3932610	7753	99.80

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.

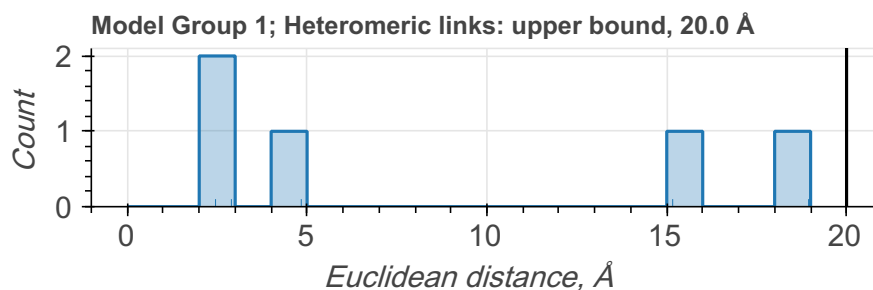
There are 335 crosslinking restraints combined in 335 restraint groups.

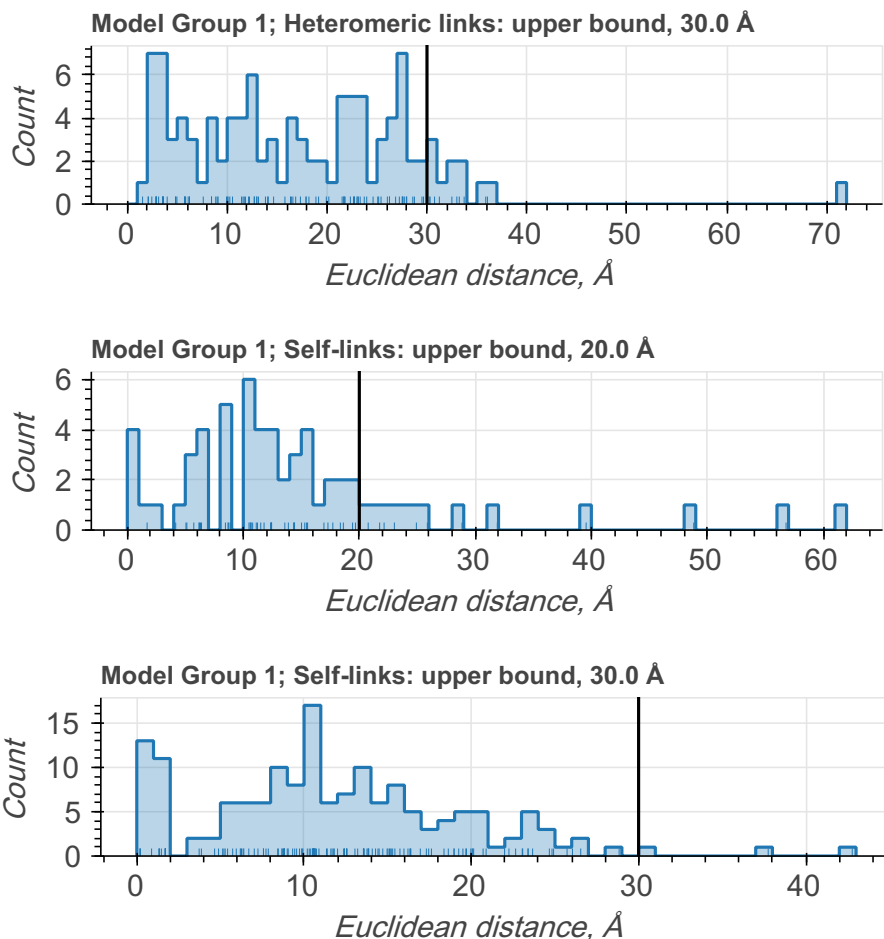
Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSSO	LYS	CA	LYS	CA	upper bound	30.0	110
DSSO	LYS	CA	SER	CA	upper bound	30.0	20
DSSO	LYS	coarse-grained	LYS	coarse-grained	upper bound	30.0	75
DSSO	LYS	CA	THR	CA	upper bound	30.0	14
DSSO	LYS	coarse-grained	SER	coarse-grained	upper bound	30.0	12
DSSO	LYS	coarse-grained	THR	coarse-grained	upper bound	30.0	13
DSSO	LYS	coarse-grained	TYR	coarse-grained	upper bound	30.0	10
DSSO	THR	coarse-grained	THR	coarse-grained	upper bound	30.0	1
DSSO	LYS	CA	TYR	CA	upper bound	30.0	7

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSSO	THR	CA	TYR	CA	upper bound	30.0	2
DSSO	SER	CA	THR	CA	upper bound	30.0	1
DSSO	SER	CA	SER	CA	upper bound	30.0	1
DSSO	THR	coarse-grained	TYR	coarse-grained	upper bound	30.0	1
DSSO	SER	coarse-grained	TYR	coarse-grained	upper bound	30.0	1
CDI	LYS	CA	THR	CA	upper bound	20.0	4
CDI	SER	coarse-grained	SER	coarse-grained	upper bound	20.0	2
CDI	LYS	coarse-grained	TYR	coarse-grained	upper bound	20.0	2
CDI	LYS	CA	LYS	CA	upper bound	20.0	29
CDI	LYS	coarse-grained	LYS	coarse-grained	upper bound	20.0	9
CDI	LYS	CA	TYR	CA	upper bound	20.0	4
CDI	LYS	CA	SER	CA	upper bound	20.0	6
CDI	SER	CA	TYR	CA	upper bound	20.0	1
CDI	SER	coarse-grained	THR	coarse-grained	upper bound	20.0	2
CDI	LYS	coarse-grained	THR	coarse-grained	upper bound	20.0	2
CDI	SER	CA	THR	CA	upper bound	20.0	3
CDI	THR	CA	THR	CA	upper bound	20.0	1
CDI	LYS	coarse-grained	SER	coarse-grained	upper bound	20.0	2

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.





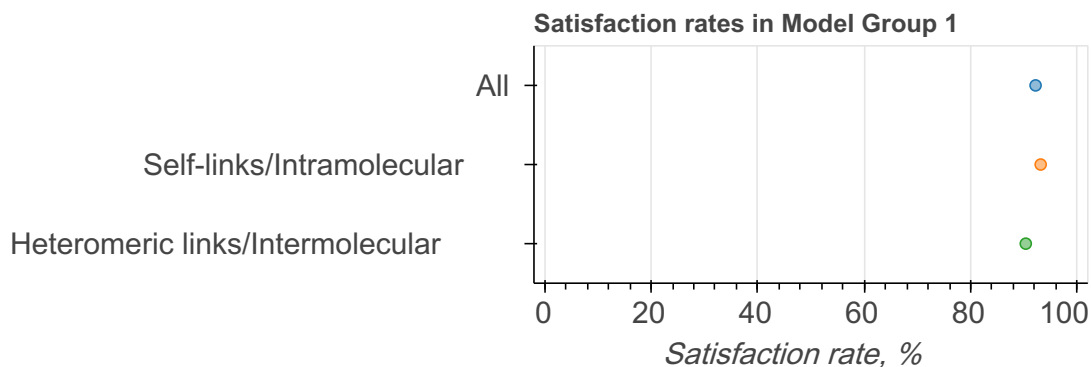
Satisfaction of restraints

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=335)
1	1	1	1/29975	All	92.19	7.81	333
				Self-links/ Intramolecular	93.15	6.85	219
				Heteromeric links/ Intermolecular	90.35	9.65	114

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.

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 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.