

# Integrative Structure Validation Report ?

February 18, 2025 - 08:46 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	9A8I
PDB-Dev ID	PDBDEV_00000382
Structure Title	Integrative structure of the human WDR76-SPIN1-Nucleosome complex
Structure Authors	Xingyu Liu; Ying Zhang; Zihui Wen; Yan Hao; Charles A.S. Banks; Joseph Cesare; Saikat Bhattacharya; Shreyas Arvindkar; Jeffrey J. Lange; Brian D. Slaughter; Jay R. Unruh; Shruthi Viswanath; Laurence Florens; Jerry L. Workman; Michael P. Washburn
Deposited on	2024-05-10

*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](mailto:helpdesk@pdb-ihm.org)*

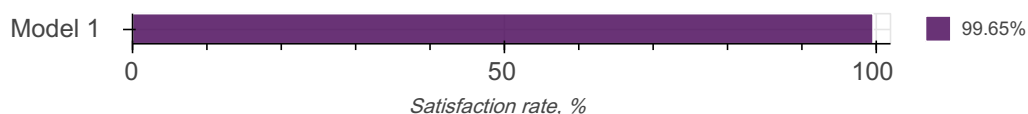
*A user guide is available at [https://pdb-ihm.org/validation\\_help.html](https://pdb-ihm.org/validation_help.html) with specific help available everywhere you see the ? symbol.*

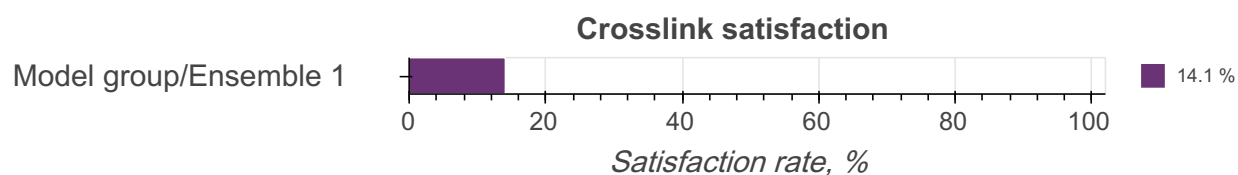
*List of references used to build this report is available [here](#).*

## 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 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	1	WD repeat-containing protein 76	A	626	245-276, 288-626	1-244, 277-287	100.00 / 59.27	Multiscale: Coarse-grained: 1 - 20 residue(s) per bead
		2	Spindlin-1	B	262	45-259	1-44, 195-210, 260-262	100.00 / 82.06	Multiscale: Coarse-grained: 1 - 20 residue(s) per bead
		3	Histone H2A type 1	C	130	13-118	1-12, 119-130	100.00 / 81.54	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
		3	Histone H2A type 1	D	130	14-119	1-13, 120-130	100.00 / 81.54	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
		4	Histone H2B type 1-C/E/F/G/I	E	126	32-125	1-31, 126	100.00 / 74.60	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
		4	Histone H2B type 1-C/E/F/G/I	F	126	28-125	1-27, 126	100.00 / 77.78	Multiscale: Coarse-grained: 1 - 10 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
		5	Histone H3.1	G	136	37-135	1-36, 136	100.00 / 72.79	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
				H					
		6	Histone H4	I	103	17-102	1-16, 103	100.00 / 83.50	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
		6	Histone H4	J	103	24-102	1-23, 103	100.00 / 76.70	Multiscale: Coarse-grained: 1 - 10 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
4	Crosslinking-MS data	Zenodo	<a href="https://zenodo.org/record/11044599">10.5281/zenodo.11044599</a>
5	Crosslinking-MS data	Zenodo	<a href="https://zenodo.org/record/11044599">10.5281/zenodo.11044599</a>
1	De Novo model	AlphaFoldDB	<a href="https://alphafolddb.org/entry/AF-Q9H967-F1">AF-Q9H967-F1</a>
2	Experimental model	PDB	<a href="https://www.rcsb.org/structure/4H75">4H75</a>
3	Experimental model	PDB	<a href="https://www.rcsb.org/structure/5GT0">5GT0</a>

### 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	750000	False	True

There are 2 software packages reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	<a href="https://integrativemodeling.org">IMP PMI module</a>	2.16.0	integrative model building	<a href="https://integrativemodeling.org">https://integrativemodeling.org</a>

ID	Software name	Software version	Software classification	Software location
2	<a href="#">Integrative Modeling Platform (IMP)</a>	2.16.0	integrative model building	<a href="https://integrativemodeling.org">https://integrativemodeling.org</a>

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

Model ID	Analysed	Number of violations	Excluded Volume Satisfaction (%)
1	966745	3427	99.65

## Fit of model to data used for modeling ?

### Fit of model(s) to crosslinking-MS data

#### Restraint types

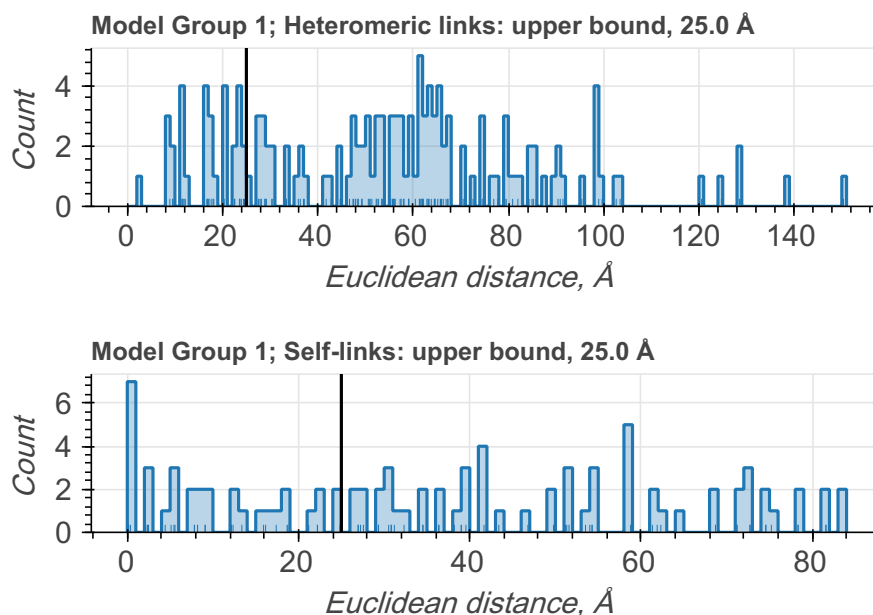
*Restraint types are summarized in the table below. Restraints assigned "[by-residue](#)" are interpreted as between CA atoms. Restraints between coarse-grained beads are indicated as "[coarse-grained](#)". [Restraint group](#) represents a set of crosslinking restraints applied collectively in the modeling.*

There are 241 crosslinking restraints combined in 79 restraint groups.

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSSO	LYS	CA	LYS	CA	upper bound	25.0	69
DSSO	ALA	coarse-grained	LYS	coarse-grained	upper bound	25.0	8
DSSO	LYS	coarse-grained	LYS	coarse-grained	upper bound	25.0	164

#### 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.



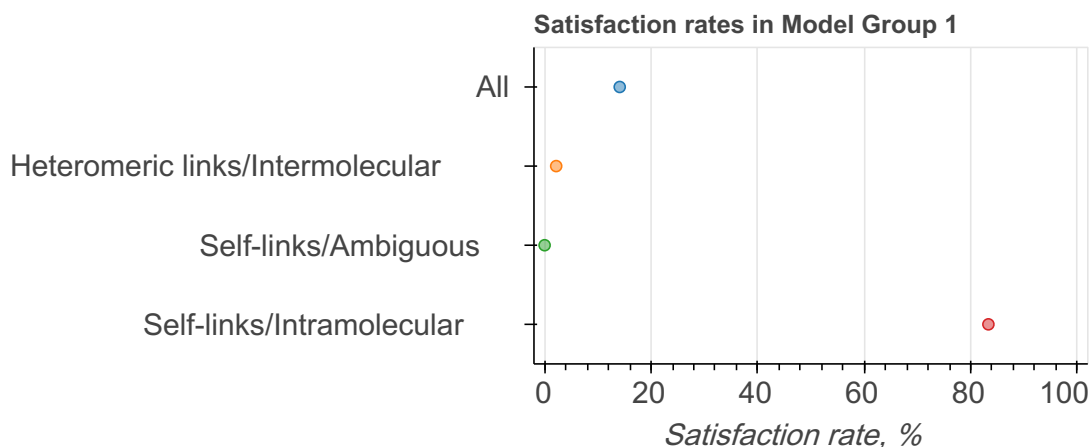
#### Satisfaction of restraints

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=79)
1	1	1	1/27314	All	14.10	85.90	78
				Heteromeric links/ Intermolecular	2.17	97.83	46
				Self-links/ Ambiguous	0.00	100.00	20
				Self-links/ Intramolecular	83.33	16.67	12

### 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 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.