# Integrative Structure Validation Report ?

## October 09, 2025 - 04:49 PM PDT

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

*IHMValidation Version 3.0*
*Python-IHM Version 2.5*

| | |
|---|---|
| PDB ID | 9A8I \| pdb_00009a8i |
| PDB-Dev ID | PDBDEV_00000382 |
| Structure Title | Integrative structure of the human WDR76-SPIN1-Nucleosome complex |
| Structure Authors | Xingyu Liu; Ying Zhang; Zhihui Wen; Yan Hao; Charles A.S. Banks; Joseph Cesare; Saikat Bhattacharya; Shreyas Arvindekar; 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 Structure Validation Report.*

*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.*

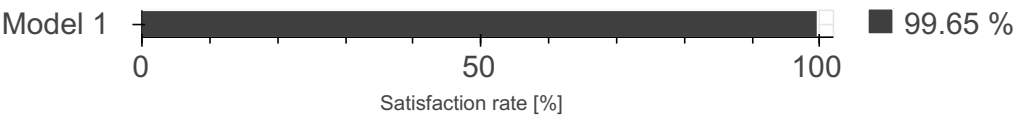*List of references used to build this report is available here.*

## 1. Overview ?

### 1.1. Summary ?

*This entry consists of 1 model(s). A total of 5 dataset(s) were used to build this entry.*

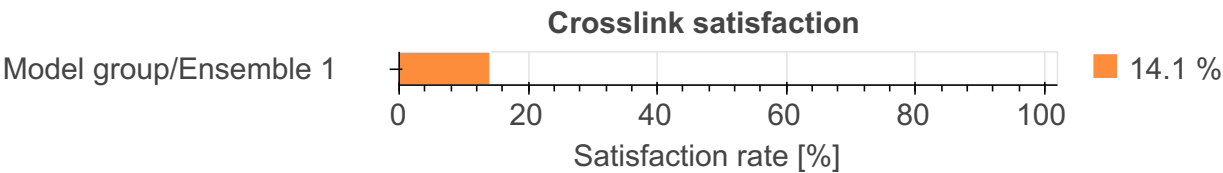| Name | Type | Count |
|---|---|---|
| Crosslinking-MS data | Experimental data | 2 |
| Experimental model | Starting model | 2 |
| De Novo model | Starting model | 1 |

## 1.2. 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 1 ████████████████████ ■ 99.65 %

0        50        100

Satisfaction rate [%]

### Fit to Data Used for Modeling ?

**Crosslink satisfaction**

Model group/Ensemble 1 ███ ■ 14.1 %

0   20   40   60   80   100

Satisfaction rate [%]

# 2. Model Details ?

## 2.1. Ensemble information ?

*This entry consists of 1 distinct ensemble(s).*

## 2.2. 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 |

| ID | Model(s) | Entity ID | Molecule name | Chain(s) [auth] | Total residues | Rigid segments | Flexible segments | Model coverage/ Starting model coverage (%) | Scale |
|---|---|---|---|---|---|---|---|---|---|
| | | 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 |
| | | 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 |

## 2.3. 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 | De Novo model | AlphaFoldDB | AF-Q9H967-F1 |
| 2 | Experimental model | PDB | pdb_00004h75 |
| 3 | Experimental model | PDB | pdb_00005gt0 |
| 4 | Crosslinking-MS data | Zenodo | 10.5281/zenodo.11044599 |
| 5 | Crosslinking-MS data | Zenodo | 10.5281/zenodo.11044599 |

## 2.4. 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 | Not available | 750000 | False | True |

*There are 2 software packages reported in this entry.*

| ID | Software name | Software version | Software classification | Software location |
|---|---|---|---|---|
| 1 | IMP PMI module | 2.16.0 | integrative model building | https://integrativemodeling.org |
| 2 | Integrative Modeling Platform (IMP) | 2.16.0 | integrative model building | https://integrativemodeling.org |

## 3. Data quality ❓

### 3.2. 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.

# 4. 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.

## 4.1a. Excluded Volume Analysis ?

*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 | 966745 | 3427 | 99.65 |

# 5. Fit to Data Used for Modeling Assessment ?

## 5.2. Crosslinking-MS ?

### 5.2.1. Restraint types ?

*This table summarizes information about crosslinker(s) used for data generation, and how crosslinking information was translated into actual modeling restraints. 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 | 25.00 | 69 |
| DSSO | ALA | coarse-grained | LYS | coarse-grained | upper bound | 25.00 | 8 |
| DSSO | LYS | coarse-grained | LYS | coarse-grained | upper bound | 25.00 | 164 |

Distograms of individual restraints

Distograms (i.e., histogram plots of distances) provide an overview of distributions of distances between residues for which chemical crosslinks were identified. The shift of the distogram relative to the threshold value may indicate a poor model. 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.

**Model Group 1; Heteromeric links: upper bound, 25.0 Å**



**Model Group 1; Self-links: upper bound, 25.0 Å**
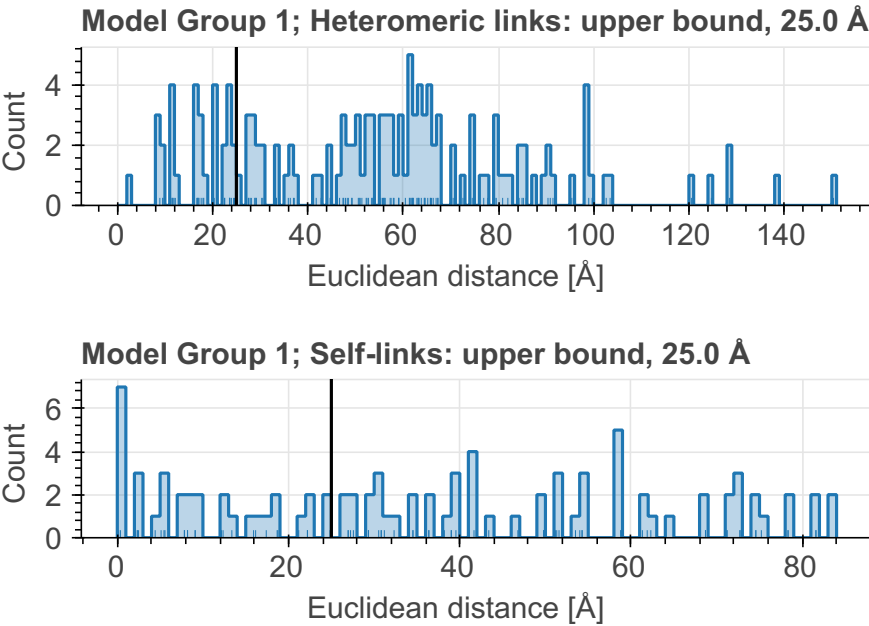


### 5.2.2. Satisfaction of restraints ?

*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.*

**Satisfaction rates in Model Group 1**
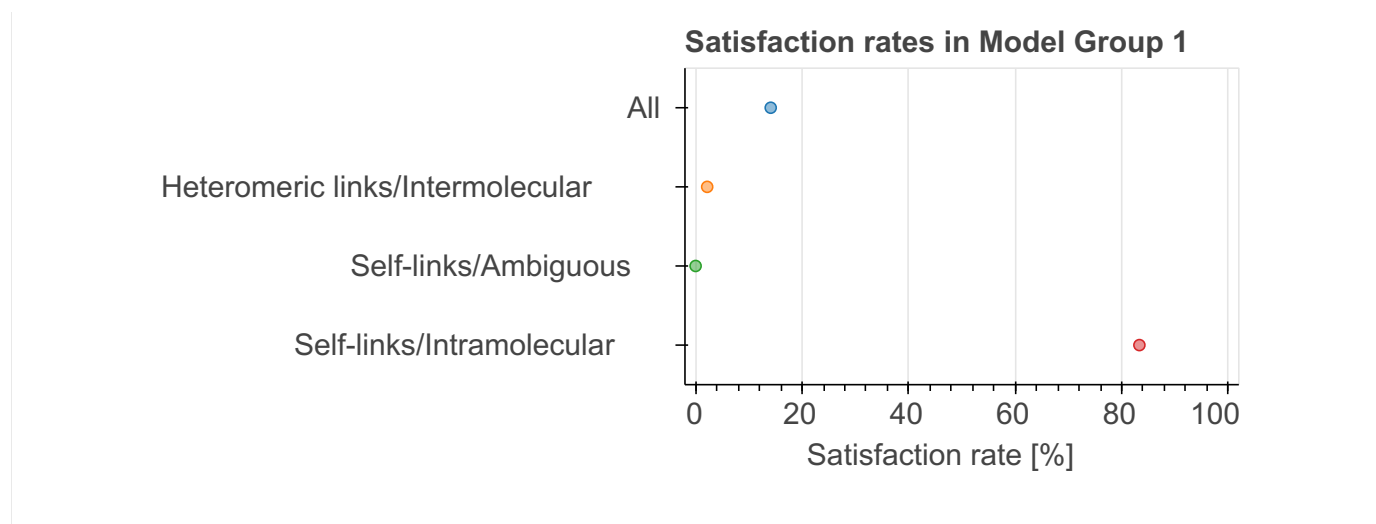
# 6. Fit to Data Used for Validation Assessment ?

Validation for this section is under development.