

Integrative Structure Validation Report ?

September 24, 2026 - 12:54 PM PDT

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

IHMValidation Version 3.2

Python-IHM Version 2.9

MolProbity Version 4.5.2

PDB ID	9AAW pdb_00009aaw
Structure Title	Integrative Structural Modeling of the HDAC2/MIER1/MHAP1 Complex
Structure Authors	Nde, J.; Kempf, C.; Zimmermann, R.C.; Cesare, J.; Zhang, Y.; Workman, J.L.; Florens, L.; Washburn, M.
Deposited on	2026-08-04

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 ?

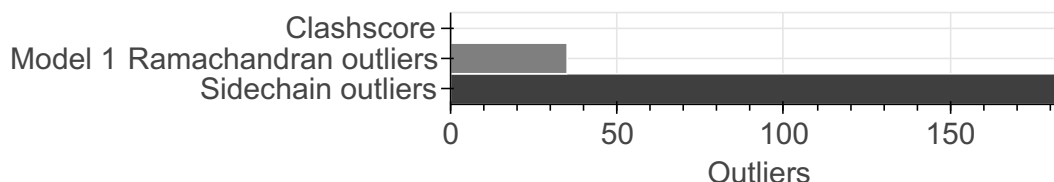
This entry consists of 1 model(s). A total of 5 dataset(s) were used to build this entry.

Name	Type	Count
Crosslinking-MS data	Experimental data	1
Mass Spectrometry data	Experimental data	1
De Novo model	Starting model	3

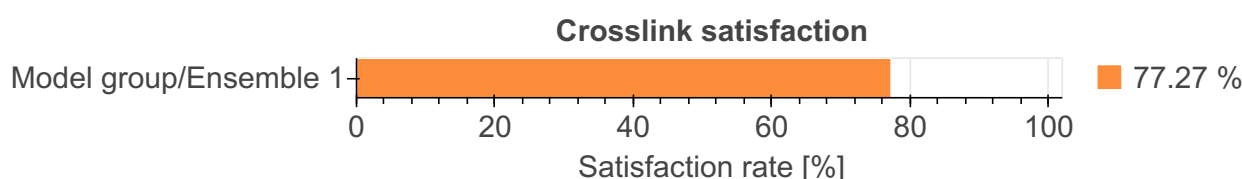
1.2. Overall quality ?

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: MolProbity Analysis ?



Fit to Data Used for Modeling ?



2. Model Details ?

2.1. Ensemble information ?

This entry consists of 0 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	Histone deacetylase 2	A	488	1-488	-	100.00 / 100.00	Atomic
		2	Mesoderm induction early response protein 1	B	512	1-512	-	100.00 / 100.00	Atomic
		3	UPF0547 protein C16orf87	C	154	1-154	-	100.00 / 100.00	Atomic

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	Zenodo	10.5281/zenodo.21810246
2	De Novo model	Zenodo	10.5281/zenodo.21810246

ID	Dataset type	Database name	Data access code
3	De Novo model	Zenodo	10.5281/zenodo.2181024
4	Crosslinking-MS data	Zenodo	10.5281/zenodo.2181024
5	Mass Spectrometry data	MASSIVE	MSV000092927

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	Not available	Prediction of individual protein structures using AlphaFold and i-TASSER	Not available	Not available	False	False
2	1	Not available	XL-MS-guided assembly using HADDOCK	Not available	Not available	False	False

There are 3 software packages reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	AlphaFold	2.00	model prediction	https://github.com/google-deepmind/alphafold
3	HADDOCK	2.50	model building	https://wenmr.science.uu.nl/haddock2.4/
2	i-TASSER	5.20	model building	http://zhanglab.ccmb.med.umich.edu/I-TASSER/download/

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.

3.4. Mass Spectrometry ?

Validation for this section is under development.

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.

4.1b. MolProbity Analysis ?

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.

Standard geometry: bond outliers ?

There are no bond length outliers.

Standard geometry: angle outliers ?

There are no bond angle outliers.

Too-close contacts ?

The following all-atom clashscore is based on a MolProbity analysis. All-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The table below contains clashscores for all atomic models in this entry.

Model ID	Clash score	Number of clashes
1	0.00	0

There are no too-close contacts.

Torsion angles: Protein backbone ?

In the following table, Ramachandran outliers are listed. The Analysed column shows the number of residues for which the backbone conformation was analysed.

Model ID	Analysed	Favored	Allowed	Outliers
1	1148	980	133	35

There are 35 unique backbone outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	206	PHE	1
A	262	CYS	1
A	378	PRO	1
A	381	GLN	1
A	406	ILE	1
A	408	ILE	1
A	436	ALA	1
A	456	ASP	1
A	457	LYS	1
A	470	ASN	1
A	478	LYS	1
A	479	GLY	1
B	5	SER	1
B	66	ILE	1
B	73	TYR	1
B	82	PRO	1
B	107	SER	1
B	110	SER	1
B	122	SER	1

Chain	Res	Type	Models (Total)
B	159	ASN	1
B	171	ASP	1
B	194	VAL	1
B	212	LEU	1
B	245	PRO	1
B	287	VAL	1
B	334	GLU	1
B	355	GLY	1
B	359	TYR	1
B	410	PRO	1
B	412	GLU	1
B	486	LYS	1
B	489	PRO	1
B	491	SER	1
C	24	PRO	1
C	49	LYS	1

Torsion angles : Protein sidechains ?

In the following table, sidechain rotameric outliers are listed. The Analysed column shows the number of residues for which the sidechain conformation was analysed.

Model ID	Analysed	Favored	Allowed	Outliers
1	1008	686	139	183

There are 183 unique sidechain outliers. Detailed list of outliers are tabulated below. The output is limited to 100 rows.

Chain	Res	Type	Models (Total)
A	1	MET	1
A	35	ARG	1
A	56	ARG	1
A	63	GLU	1
A	66	THR	1
A	71	ASP	1
A	78	ARG	1
A	86	SER	1
A	90	LYS	1
A	96	ASN	1
A	97	VAL	1

Chain	Res	Type	Models (Total)
A	99	GLU	1
A	109	GLU	1
A	130	THR	1
A	141	HIS	1
A	149	SER	1
A	170	ARG	1
A	185	GLU	1
A	190	THR	1
A	200	HIS	1
A	209	THR	1
A	231	ASP	1
A	241	ILE	1
A	262	CYS	1
A	265	ASP	1
A	266	SER	1
A	293	ASN	1
A	307	ARG	1
A	327	ASN	1
A	332	ASN	1
A	352	THR	1
A	358	GLU	1
A	374	LEU	1
A	397	GLU	1
A	407	SER	1
A	413	LYS	1
A	414	ARG	1
A	424	SER	1
A	429	GLU	1
A	437	ASP	1
A	448	GLU	1
A	453	GLU	1
A	455	GLU	1
A	457	LYS	1
A	464	GLU	1
A	467	SER	1

Chain	Res	Type	Models (Total)
A	471	SER	1
A	475	THR	1
A	477	THR	1
A	480	THR	1
A	483	GLU	1
B	7	GLU	1
B	8	SER	1
B	10	SER	1
B	16	THR	1
B	23	ASP	1
B	29	LEU	1
B	32	ASP	1
B	34	ASP	1
B	36	GLU	1
B	49	THR	1
B	57	ASP	1
B	58	LEU	1
B	61	GLU	1
B	64	MET	1
B	68	GLU	1
B	70	LEU	1
B	71	SER	1
B	77	SER	1
B	78	THR	1
B	83	GLU	1
B	96	GLU	1
B	99	GLU	1
B	103	ASN	1
B	104	ASP	1
B	112	GLU	1
B	116	GLU	1
B	119	LYS	1
B	124	GLN	1
B	126	ASP	1
B	127	GLU	1

Chain	Res	Type	Models (Total)
B	128	THR	1
B	129	GLN	1
B	132	ASN	1
B	138	SER	1
B	139	VAL	1
B	146	GLU	1
B	155	TYR	1
B	160	SER	1
B	161	GLU	1
B	163	GLU	1
B	167	GLU	1
B	171	ASP	1
B	172	TYR	1
B	175	SER	1
B	176	GLU	1
B	180	LYS	1
B	194	VAL	1
B	198	ARG	1
B	204	LYS	1

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.

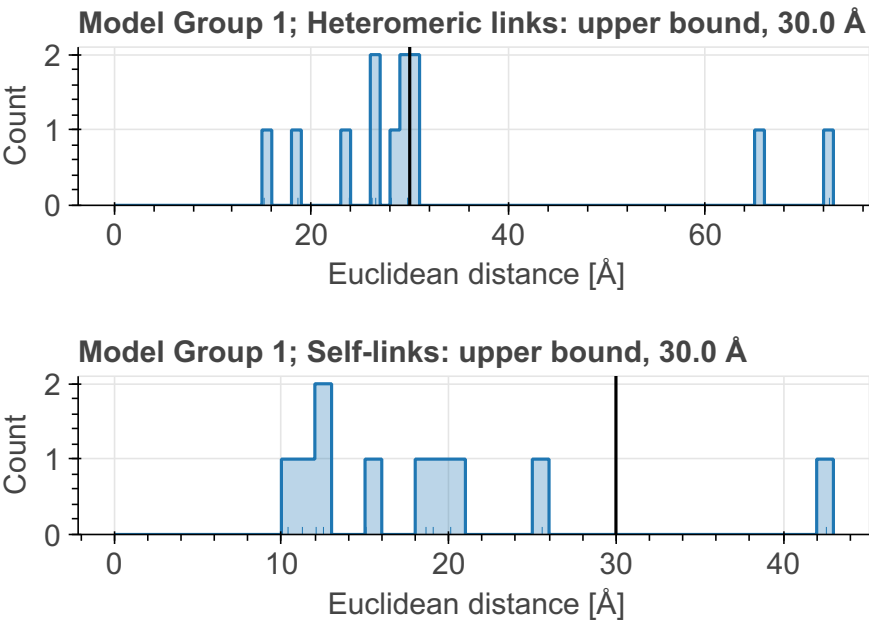
There are 22 crosslinking restraints combined in 22 restraint groups.

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSSO	LYS	CA	LYS	CA	upper bound	30.00	22

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.



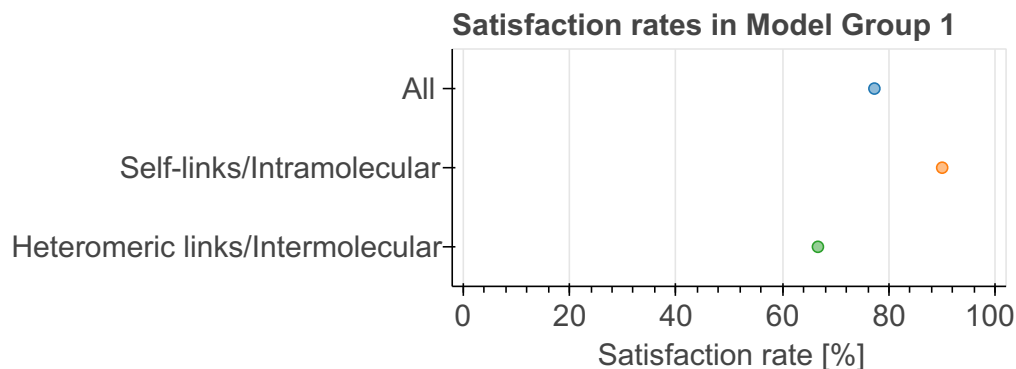
5.2.2. Satisfaction of restraints ?

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=22)
1	1	1	1/1	All	77.27	22.73	22
				Self-links/ Intramolecular	90.00	10.00	10
				Heteromeric links/ Intermolecular	66.67	33.33	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.



5.4. Mass Spectrometry ?

Validation for this section is under development.

6. Fit to Data Used for Validation Assessment ?

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.