Integrative Structure Validation Report

October 09, 2025 - 04:45 PM PDT

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

IHMValidation Version 3.0
Python-IHM Version 2.5
MolProbity Version 4.5.2
pyHMMER Version 0.11.1

PDB ID	9A4W | pdb_00009a4w
PDB-Dev ID	PDBDEV_00000253
Structure Title	Integrative model of RPOA-RPOC by crosslinking MS and deep learning
Structure Authors	Stahl, K.; Brock, O.; Rappsilber, J.
Deposited on	2024-01-23

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 1 dataset(s) were used to build this entry.

Name	Type	Count
Crosslinking-MS data	Experimental data	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: MolProbity Analysis

Data Quality

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	RPOA_BACSU	A	314	-	1-314	100.00 / 0.00	Atomic
		2	RPOC_BACSU	B	1199	-	1-1199	100.00 / 0.00	Atomic

2.3. Datasets used for modeling

There is 1 unique dataset used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	Crosslinking-MS data	PRIDE	PXD035508

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	AlphaLink2	AlphaLink2	Not available	1	False	False

There is 1 software package reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	AlphaLink2	1.00	model building	https://github.com/Rappsilber-Laboratory/AlphaLink2

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).

PXD035508

Number of entities in the crosslinking-MS dataset:	810
Number of entities in the entry:	2

Matching entities:

Entity ID	Molecule name	Crosslinking-MS Entity ID	E-value	Exact match
1	RPOA_BACSU	dbseq_P20429_target	0.00	True
2	RPOC_BACSU	dbseq_P37871_target	0.00	True

Residue pairs stats:

Source	Total	In matched entities	Total matched
9A4W	4	4 (100.00%)	4 (100.00%)
PXD035508	3666	24 (0.65%)	4 (0.11%)

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-partice 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 60 bond angle outliers in this entry (0.37% of 16271 assessed bonds). A summary is provided below.

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
B	449	ASP	CA-CB-CG	6.15	118.75	112.60	1	1
B	768	ARG	NE-CZ-NH2	5.42	124.07	119.20	1	1
B	439	HIS	CB-CG-CD2	5.20	124.43	131.20	1	1
B	180	GLN	OE1-CD-NE2	4.94	117.66	122.60	1	1
A	187	GLN	OE1-CD-NE2	4.93	117.67	122.60	1	1
A	232	GLN	OE1-CD-NE2	4.92	117.68	122.60	1	1
B	668	ASP	CA-CB-CG	4.92	117.52	112.60	1	1
B	821	ASP	CA-CB-CG	4.90	117.50	112.60	1	1
B	957	GLN	OE1-CD-NE2	4.90	117.70	122.60	1	1
A	182	ASN	OD1-CG-ND2	4.83	117.77	122.60	1	1
B	713	GLN	OE1-CD-NE2	4.83	117.77	122.60	1	1
A	179	GLN	OE1-CD-NE2	4.81	117.79	122.60	1	1
B	329	GLN	OE1-CD-NE2	4.78	117.82	122.60	1	1
B	994	GLN	OE1-CD-NE2	4.73	117.87	122.60	1	1
B	968	GLN	OE1-CD-NE2	4.71	117.89	122.60	1	1
B	925	GLN	OE1-CD-NE2	4.71	117.89	122.60	1	1
B	553	GLN	OE1-CD-NE2	4.70	117.90	122.60	1	1
B	670	GLN	OE1-CD-NE2	4.70	117.90	122.60	1	1
B	933	GLN	OE1-CD-NE2	4.69	117.91	122.60	1	1
B	1190	GLN	OE1-CD-NE2	4.69	117.91	122.60	1	1
B	803	ARG	NE-CZ-NH2	4.62	123.35	119.20	1	1
B	753	ASN	OD1-CG-ND2	4.59	118.01	122.60	1	1
B	1140	GLN	OE1-CD-NE2	4.55	118.05	122.60	1	1
B	532	HIS	CB-CG-CD2	4.52	125.32	131.20	1	1
B	1044	GLN	OE1-CD-NE2	4.51	118.09	122.60	1	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	192	ASP	CA-CB-CG	4.46	117.06	112.60	1	1
B	686	GLN	OE1-CD-NE2	4.44	118.16	122.60	1	1
B	394	GLN	OE1-CD-NE2	4.42	118.18	122.60	1	1
B	1086	THR	CA-CB-CG2	4.42	118.01	110.50	1	1
B	243	GLN	OE1-CD-NE2	4.40	118.20	122.60	1	1
B	887	ARG	NE-CZ-NH2	4.35	123.12	119.20	1	1
A	159	GLN	OE1-CD-NE2	4.35	118.25	122.60	1	1
B	1058	GLN	OE1-CD-NE2	4.31	118.29	122.60	1	1
A	184	ARG	NE-CZ-NH1	4.30	125.80	121.50	1	1
B	737	ASN	OD1-CG-ND2	4.26	118.34	122.60	1	1
B	988	GLN	OE1-CD-NE2	4.26	118.34	122.60	1	1
B	466	GLN	OE1-CD-NE2	4.24	118.36	122.60	1	1
B	324	GLN	OE1-CD-NE2	4.24	118.36	122.60	1	1
A	143	GLN	OE1-CD-NE2	4.21	118.39	122.60	1	1
A	275	GLN	OE1-CD-NE2	4.20	118.40	122.60	1	1
B	608	GLN	OE1-CD-NE2	4.19	118.41	122.60	1	1
B	585	ASN	OD1-CG-ND2	4.18	118.42	122.60	1	1
B	413	ASN	OD1-CG-ND2	4.16	118.44	122.60	1	1
B	609	GLN	OE1-CD-NE2	4.15	118.45	122.60	1	1
A	124	ASP	CA-CB-CG	4.15	116.75	112.60	1	1
B	253	ASP	CA-CB-CG	4.14	116.74	112.60	1	1
B	167	GLN	OE1-CD-NE2	4.12	118.48	122.60	1	1
B	1022	GLN	OE1-CD-NE2	4.12	118.48	122.60	1	1
B	753	ASN	CA-CB-CG	4.12	116.72	112.60	1	1
B	493	GLN	OE1-CD-NE2	4.12	118.48	122.60	1	1
B	201	GLN	OE1-CD-NE2	4.11	118.49	122.60	1	1
B	499	ASN	OD1-CG-ND2	4.10	118.50	122.60	1	1
B	113	ARG	NH1-CZ-NH2	4.10	113.98	119.30	1	1
A	63	HIS	CB-CG-CD2	4.06	125.93	131.20	1	1
A	244	GLN	OE1-CD-NE2	4.06	118.54	122.60	1	1
B	1066	HIS	CB-CG-CD2	4.05	125.93	131.20	1	1
B	424	GLN	OE1-CD-NE2	4.05	118.55	122.60	1	1
A	314	ASP	CA-CB-CG	4.05	116.65	112.60	1	1
A	122	ASN	CA-CB-CG	4.02	116.62	112.60	1	1
B	478	ASN	OD1-CG-ND2	4.01	118.59	122.60	1	1

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	1.83	44

There are 44 clashes. The table below contains the detailed list of all clashes based on a MolProbity analysis. Bad clashes are >= 0.4 Angstrom.

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
B:808:ALA:HB2	B:1070:MET:HE1	0.85	1	1
B:808:ALA:CB	B:1070:MET:HE1	0.79	1	1
B:930:PRO:HB3	B:1060:VAL:HG11	0.70	1	1
B:234:ILE:HD11	B:316:LEU:HD21	0.68	1	1
B:720:LEU:HD11	B:741:PHE:CE2	0.66	1	1
B:371:MET:CE	B:383:ILE:HG23	0.64	1	1
B:124:ALA:HB1	B:149:LEU:HD11	0.61	1	1
B:441:LEU:HD21	B:495:MET:HE1	0.61	1	1
B:1086:THR:HG22	B:1106:VAL:HG21	0.61	1	1
B:726:ILE:HG21	B:741:PHE:CE1	0.60	1	1
B:978:VAL:CG2	B:1018:ILE:HD11	0.58	1	1
B:317:SER:HB3	B:321:LYS:HE2	0.58	1	1
B:317:SER:HB3	B:321:LYS:CE	0.57	1	1
B:84:ARG:O	B:87:VAL:HG22	0.56	1	1
B:503:THR:HG21	B:615:PHE:CD2	0.56	1	1
B:234:ILE:HG22	B:235:PRO:HD2	0.55	1	1
B:978:VAL:HG22	B:1018:ILE:HD11	0.54	1	1
B:214:GLU:OE1	B:1089:LEU:HD21	0.53	1	1
B:234:ILE:HD11	B:316:LEU:CD2	0.52	1	1
A:130:LEU:HD21	A:136:PHE:HB2	0.51	1	1
B:506:ARG:HH21	B:723:LEU:HD22	0.51	1	1
B:726:ILE:HG21	B:741:PHE:CD1	0.50	1	1
B:895:PRO:HA	B:1097:HIS:CD2	0.50	1	1
B:747:MET:HE2	B:749:GLY:HA2	0.50	1	1
B:495:MET:HE3	B:644:LYS:HA	0.49	1	1
B:705:TRP:CD2	B:762:PRO:HG3	0.46	1	1
B:519:THR:HG23	B:565:LEU:HB3	0.46	1	1
A:58:ILE:HG21	A:68:ILE:HD11	0.45	1	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
B:915:VAL:HG12	B:1177:MET:HE1	0.45	1	1
B:539:VAL:HG21	B:565:LEU:HD11	0.45	1	1
B:506:ARG:HE	B:723:LEU:CD2	0.44	1	1
B:1158:LEU:HD22	B:1167:ILE:CD1	0.44	1	1
B:705:TRP:CH2	B:764:LYS:HE2	0.44	1	1
B:978:VAL:HG23	B:1018:ILE:HD11	0.44	1	1
B:806:ASP:CG	B:1162:LYS:HZ1	0.43	1	1
B:808:ALA:HB1	B:1070:MET:HE1	0.42	1	1
B:441:LEU:CD2	B:495:MET:HE1	0.42	1	1
A:22:LYS:HE2	A:197:ASP:OD1	0.41	1	1
B:346:VAL:HG22	B:450:PHE:CZ	0.41	1	1
A:136:PHE:CZ	A:138:VAL:HG21	0.40	1	1
B:1070:MET:HA	B:1070:MET:HE2	0.40	1	1
B:137:THR:O	B:146:LYS:HE3	0.40	1	1
B:969:ALA:HB1	B:1023:VAL:HG13	0.40	1	1
B:113:ARG:HH11	B:212:VAL:CG2	0.40	1	1

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	1509	1471	37	1

There are 1 unique backbone outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	243	ASP	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	1300	1161	81	58

There are 58 unique sidechain outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	43	ILE	1
A	106	THR	1

Chain	Res	Type	Models (Total)
A	113	THR	1
A	124	ASP	1
A	196	LEU	1
A	257	ASP	1
A	286	MET	1
A	291	LEU	1
A	296	LEU	1
B	2	LEU	1
B	55	THR	1
B	76	ASP	1
B	107	PHE	1
B	125	LEU	1
B	159	LEU	1
B	190	MET	1
B	195	LEU	1
B	204	THR	1
B	234	ILE	1
B	260	ARG	1
B	287	MET	1
B	327	PHE	1
B	345	VAL	1
B	357	LEU	1
B	373	GLU	1
B	381	HIS	1
B	393	VAL	1
B	431	VAL	1
B	438	LEU	1
B	444	THR	1
B	449	ASP	1
B	453	ASP	1
B	474	LEU	1
B	488	VAL	1
B	497	LEU	1
B	504	LEU	1
B	634	THR	1

Chain	Res	Type	Models (Total)
B	646	LEU	1
B	666	LEU	1
B	692	ILE	1
B	716	LEU	1
B	761	LEU	1
B	774	LEU	1
B	780	THR	1
B	809	GLN	1
B	813	ILE	1
B	816	THR	1
B	912	ASP	1
B	949	ILE	1
B	953	LEU	1
B	974	ILE	1
B	977	THR	1
B	999	THR	1
B	1003	THR	1
B	1018	ILE	1
B	1041	THR	1
B	1042	THR	1
B	1171	VAL	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 4 crosslinking restraints combined in 4 restraint groups.

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
SDA	LYS	CA	LYS	CA	upper bound	25.00	3
SDA	LYS	CA	MET	CA	upper bound	25.00	1

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=4)
1	1	1	1/1	All	75.00	25.00	4
				Heteromeric links/ Intermolecular	75.00	25.00	4

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.

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.