

# 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	9A57   pdb_00009a57
PDB-Dev ID	PDBDEV_00000264
Structure Title	Integrative model of RPOC-NUSG 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](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](#).*

## 1. Overview

### 1.1. Summary

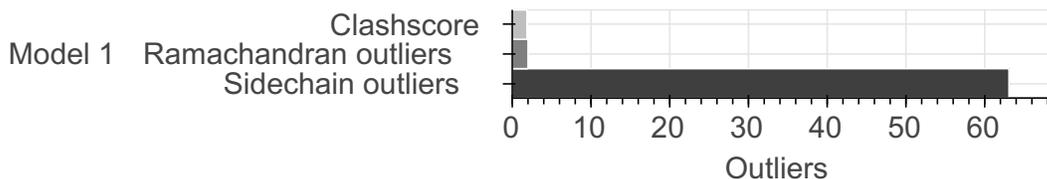
*This entry consists of 1 model(s). A total of 1 dataset(s) were used to build this entry.*

Name	Type	Count
Crosslinking-MS data	Experimental data	1

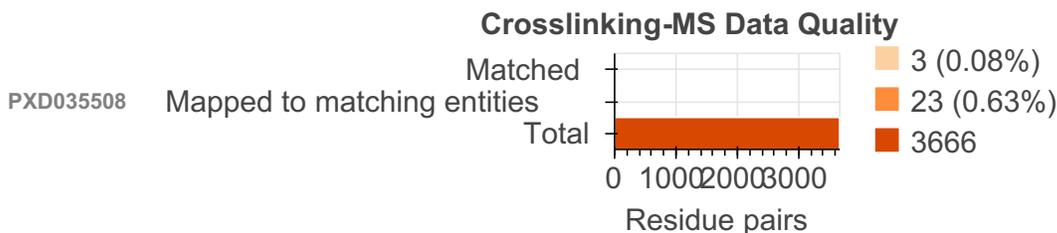
### 1.2. Overall quality

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.

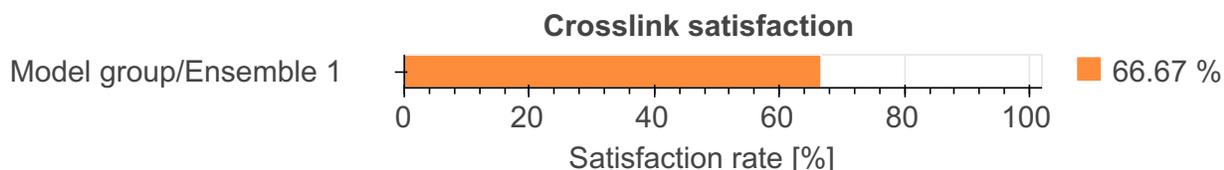
### Model Quality: MolProbity Analysis ?



### 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	RPOC_BACSU	A	1199	-	1-1199	100.00 / 0.00	Atomic
		2	NUSG_BACSU	B	177	-	1-177	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	<a href="#">PXD035508</a>

## 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	<a href="#">AlphaLink2</a>	1.00	model building	<a href="https://github.com/Rappsilber-Laboratory/AlphaLink2">https://github.com/Rappsilber-Laboratory/AlphaLink2</a>

## 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	RPOC_BACSU	dbseq_P37871_target	0.00	True
2	NUSG_BACSU	dbseq_Q06795_target	0.00	True

Residue pairs stats:

Source	Total	In matched entities	Total matched
9A57	3	3 (100.00%)	3 (100.00%)
PXD035508	3666	23 (0.63%)	3 (0.08%)

## 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 1 bond length outliers in this entry (0.01% of 11023 assessed bonds). A summary is provided below.

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	803	ARG	CZ-NH1	4.52	1.26	1.32	1	1

#### Standard geometry: angle outliers

There are 44 bond angle outliers in this entry (0.30% of 14870 assessed bonds). A summary is provided below.

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	803	ARG	NE-CZ-NH2	10.16	128.34	119.20	1	1
A	803	ARG	NH1-CZ-NH2	9.29	107.23	119.30	1	1
A	449	ASP	CA-CB-CG	6.27	118.87	112.60	1	1
A	869	ASP	CA-CB-CG	5.76	118.36	112.60	1	1
A	585	ASN	OD1-CG-ND2	5.65	116.95	122.60	1	1
A	668	ASP	CA-CB-CG	5.63	118.23	112.60	1	1
A	180	GLN	OE1-CD-NE2	5.18	117.42	122.60	1	1
A	439	HIS	CB-CG-CD2	4.92	124.80	131.20	1	1
A	821	ASP	CA-CB-CG	4.86	117.46	112.60	1	1
A	394	GLN	OE1-CD-NE2	4.83	117.77	122.60	1	1
A	670	GLN	OE1-CD-NE2	4.79	117.81	122.60	1	1
B	123	ASP	CA-CB-CG	4.79	117.39	112.60	1	1
A	686	GLN	OE1-CD-NE2	4.77	117.83	122.60	1	1
A	527	GLN	OE1-CD-NE2	4.75	117.85	122.60	1	1
A	1044	GLN	OE1-CD-NE2	4.74	117.86	122.60	1	1
A	753	ASN	OD1-CG-ND2	4.74	117.86	122.60	1	1
A	1140	GLN	OE1-CD-NE2	4.71	117.89	122.60	1	1
A	329	GLN	OE1-CD-NE2	4.71	117.89	122.60	1	1
A	933	GLN	OE1-CD-NE2	4.70	117.90	122.60	1	1
A	925	GLN	OE1-CD-NE2	4.60	118.00	122.60	1	1
A	968	GLN	OE1-CD-NE2	4.59	118.01	122.60	1	1
A	994	GLN	OE1-CD-NE2	4.55	118.05	122.60	1	1

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	957	GLN	OE1-CD-NE2	4.50	118.10	122.60	1	1
A	988	GLN	OE1-CD-NE2	4.48	118.12	122.60	1	1
A	324	GLN	OE1-CD-NE2	4.48	118.12	122.60	1	1
A	532	HIS	CB-CG-CD2	4.44	125.43	131.20	1	1
A	768	ARG	NE-CZ-NH2	4.41	123.17	119.20	1	1
A	713	GLN	OE1-CD-NE2	4.38	118.22	122.60	1	1
A	1190	GLN	OE1-CD-NE2	4.28	118.32	122.60	1	1
A	553	GLN	OE1-CD-NE2	4.26	118.34	122.60	1	1
A	201	GLN	OE1-CD-NE2	4.26	118.34	122.60	1	1
B	32	GLN	OE1-CD-NE2	4.19	118.41	122.60	1	1
A	753	ASN	CA-CB-CG	4.13	116.73	112.60	1	1
A	243	GLN	OE1-CD-NE2	4.12	118.48	122.60	1	1
A	318	HIS	CB-CG-CD2	4.12	125.85	131.20	1	1
A	1066	HIS	CB-CG-CD2	4.11	125.85	131.20	1	1
A	951	GLN	OE1-CD-NE2	4.11	118.49	122.60	1	1
A	608	GLN	OE1-CD-NE2	4.10	118.50	122.60	1	1
A	381	HIS	CB-CG-CD2	4.10	125.87	131.20	1	1
A	493	GLN	OE1-CD-NE2	4.10	118.50	122.60	1	1
A	167	GLN	OE1-CD-NE2	4.07	118.53	122.60	1	1
A	349	HIS	CB-CG-CD2	4.07	125.91	131.20	1	1
A	413	ASN	OD1-CG-ND2	4.02	118.58	122.60	1	1
A	419	HIS	CB-CG-CD2	4.02	125.98	131.20	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	1.87	41

There are 41 clashes. The table below contains the detailed list of all clashes based on a MolProbity analysis. Bad clashes are  $\geq 0.4$  Angstrom.

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:808:ALA:HB2	A:1070:MET:HE1	0.86	1	1
A:808:ALA:CB	A:1070:MET:HE1	0.76	1	1
A:978:VAL:CG2	A:1018:ILE:HD11	0.71	1	1
A:234:ILE:HD11	A:316:LEU:HD21	0.69	1	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:317:SER:HB3	A:321:LYS:HE2	0.69	1	1
A:720:LEU:HD11	A:741:PHE:CE2	0.66	1	1
B:31:MET:HE1	B:74:ASP:HB3	0.64	1	1
A:726:ILE:HG21	A:741:PHE:CE1	0.61	1	1
A:371:MET:CE	A:383:ILE:HG23	0.60	1	1
A:978:VAL:HG22	A:1018:ILE:HD11	0.57	1	1
B:29:MET:CE	B:78:VAL:HG22	0.56	1	1
A:441:LEU:HD21	A:495:MET:HE1	0.56	1	1
B:29:MET:HE1	B:78:VAL:HG22	0.55	1	1
A:234:ILE:HD11	A:316:LEU:CD2	0.53	1	1
A:495:MET:HE3	A:644:LYS:HA	0.53	1	1
A:506:ARG:HH21	A:723:LEU:HD22	0.51	1	1
A:559:ILE:HD11	A:596:LEU:CD1	0.51	1	1
A:503:THR:HG21	A:615:PHE:CD2	0.51	1	1
A:978:VAL:HG23	A:1018:ILE:HD11	0.51	1	1
A:84:ARG:O	A:87:VAL:HG22	0.49	1	1
A:506:ARG:HE	A:723:LEU:CD2	0.49	1	1
A:234:ILE:HG22	A:235:PRO:HD2	0.48	1	1
A:726:ILE:HG21	A:741:PHE:CD1	0.48	1	1
A:663:ILE:HB	A:768:ARG:HH11	0.47	1	1
A:539:VAL:HG21	A:565:LEU:HD11	0.47	1	1
A:1086:THR:HG22	A:1106:VAL:HG21	0.46	1	1
B:19:LYS:HE3	B:23:GLU:OE2	0.46	1	1
A:317:SER:HB3	A:321:LYS:CE	0.45	1	1
A:705:TRP:CD2	A:762:PRO:HG3	0.44	1	1
A:895:PRO:HA	A:1097:HIS:CD2	0.44	1	1
A:559:ILE:HD11	A:596:LEU:HD11	0.44	1	1
A:1070:MET:HA	A:1070:MET:HE2	0.44	1	1
A:991:ILE:HD11	A:1010:LEU:CD2	0.43	1	1
A:1088:VAL:HG13	A:1094:LEU:HD11	0.43	1	1
A:585:ASN:HD22	A:591:PRO:CD	0.43	1	1
B:25:ARG:HH12	B:81:ASN:C	0.42	1	1
A:500:TYR:CE2	A:614:PRO:HG2	0.42	1	1
A:799:TYR:HE2	A:803:ARG:HH21	0.41	1	1
A:808:ALA:HB1	A:1070:MET:HE1	0.41	1	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:585:ASN:HD22	A:591:PRO:HD3	0.41	1	1
A:803:ARG:NH1	A:1124:THR:OG1	0.41	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	1372	1330	40	2

There are 2 unique backbone outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	581	PRO	1
B	118	ARG	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	1186	1057	66	63

There are 63 unique sidechain outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	2	LEU	1
A	76	ASP	1
A	107	PHE	1
A	125	LEU	1
A	135	VAL	1
A	159	LEU	1
A	187	GLU	1
A	195	LEU	1
A	204	THR	1
A	327	PHE	1
A	345	VAL	1
A	357	LEU	1
A	363	LEU	1
A	373	GLU	1
A	431	VAL	1

Chain	Res	Type	Models (Total)
A	438	LEU	1
A	444	THR	1
A	449	ASP	1
A	453	ASP	1
A	474	LEU	1
A	488	VAL	1
A	497	LEU	1
A	504	LEU	1
A	519	THR	1
A	524	LEU	1
A	531	VAL	1
A	551	GLU	1
A	580	GLU	1
A	583	LYS	1
A	634	THR	1
A	641	ASP	1
A	646	LEU	1
A	692	ILE	1
A	716	LEU	1
A	761	LEU	1
A	765	SER	1
A	780	THR	1
A	809	GLN	1
A	816	THR	1
A	869	ASP	1
A	912	ASP	1
A	932	THR	1
A	938	THR	1
A	974	ILE	1
A	977	THR	1
A	999	THR	1
A	1003	THR	1
A	1006	TYR	1
A	1041	THR	1
A	1042	THR	1

Chain	Res	Type	Models (Total)
A	1171	VAL	1
A	1177	MET	1
A	1192	THR	1
B	10	THR	1
B	31	MET	1
B	35	ILE	1
B	43	GLU	1
B	45	GLU	1
B	46	THR	1
B	69	ILE	1
B	120	THR	1
B	133	ILE	1
B	174	ILE	1

## 5. Fit to Data Used for Modeling Assessment ?

### 5.2. Crosslinking-MS ?

#### 5.2.1. Restraint types ?

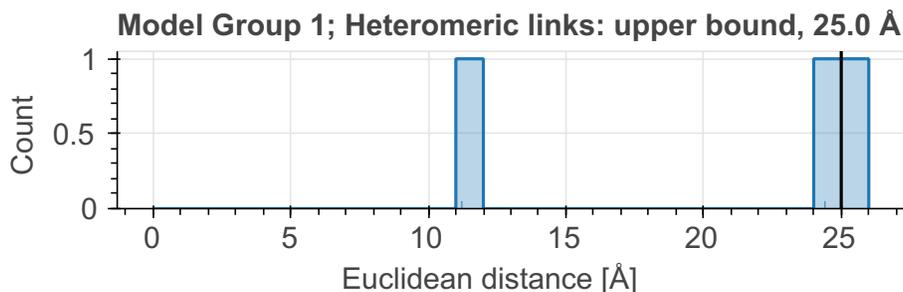
This table summarizes information about crosslinker(s) used for data generation, and how crosslinking information was translated into actual modeling restraints. Restraints assigned "by-residue" are interpreted as between CA atoms. Restraints between coarse-grained beads are indicated as "coarse-grained". *Restraint group* represents a set of crosslinking restraints applied collectively in the modeling.

There are 3 crosslinking restraints combined in 3 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

#### 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.



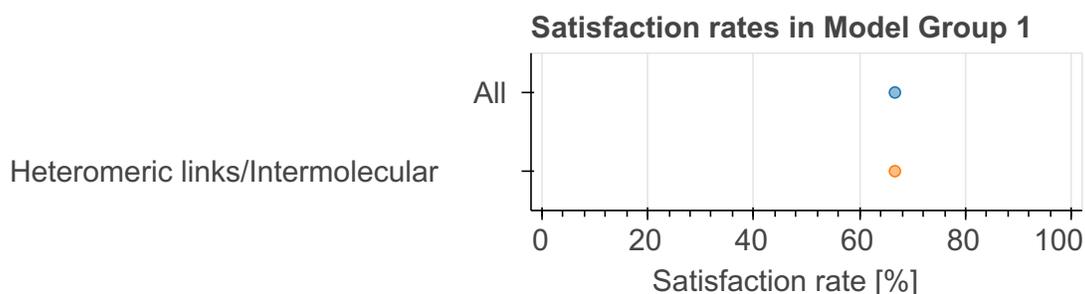
### 5.2.2. Satisfaction of restraints ?

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=3)
1	1	1	1/1	All	66.67	33.33	3
				Heteromeric links/Intermolecular	66.67	33.33	3

#### 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.*