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	9A54	pdb_00009a54
PDB-Dev ID	PDBDEV_00000261
Structure Title	Integrative model of GYRB-GYRA 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.
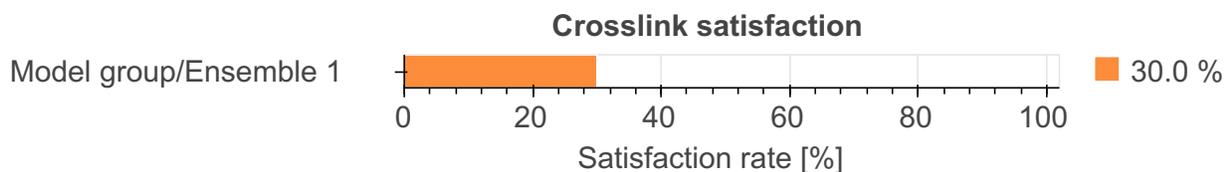
Model Quality: MolProbity Analysis ?
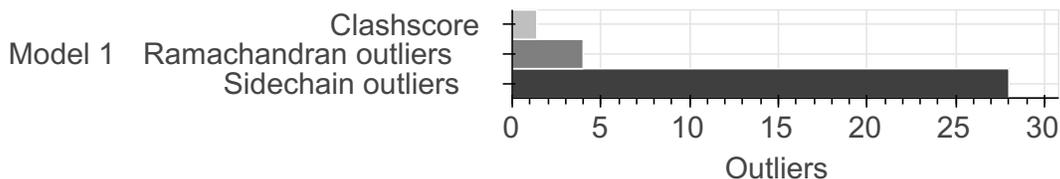
Data Quality ?
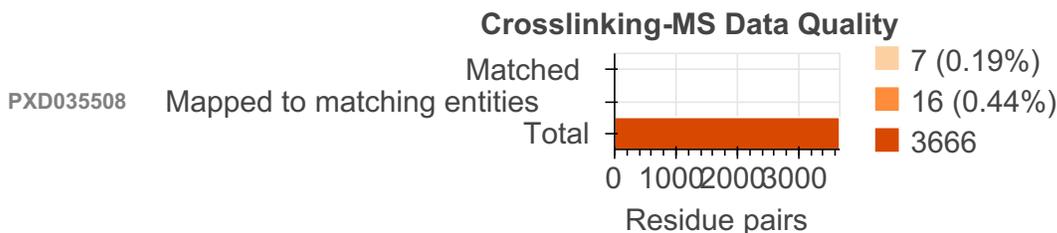
Fit to Data Used for Modeling ?

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	GYRB_BACSU	A	638	-	1-638	100.00 / 0.00	Atomic
		2	GYRA_BACSU	B	821	-	1-821	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	GYRB_BACSU	dbseq_P05652_target	0.00	True
2	GYRA_BACSU	dbseq_P05653_target	0.00	True

Residue pairs stats:

Source	Total	In matched entities	Total matched
9A54	10	10 (100.00%)	7 (70.00%)
PXD035508	3666	16 (0.44%)	7 (0.19%)

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 40 bond angle outliers in this entry (0.25% of 15729 assessed bonds). A summary is provided below.

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
B	333	ASN	OD1-CG-ND2	6.57	116.03	122.60	1	1
B	187	GLN	OE1-CD-NE2	5.43	117.17	122.60	1	1
B	413	GLN	OE1-CD-NE2	5.38	117.22	122.60	1	1
B	599	ASN	OD1-CG-ND2	5.24	117.36	122.60	1	1
B	749	GLN	OE1-CD-NE2	5.13	117.47	122.60	1	1
A	485	ASP	CA-CB-CG	5.10	117.70	112.60	1	1
A	107	ASP	CA-CB-CG	4.99	117.59	112.60	1	1
A	339	GLN	OE1-CD-NE2	4.90	117.70	122.60	1	1
B	818	GLN	OE1-CD-NE2	4.77	117.83	122.60	1	1
A	598	GLN	OE1-CD-NE2	4.76	117.84	122.60	1	1
B	814	ASN	OD1-CG-ND2	4.75	117.85	122.60	1	1
B	529	ARG	CD-NE-CZ	4.70	130.98	124.40	1	1
B	481	ASP	CA-CB-CG	4.56	117.16	112.60	1	1
B	163	ARG	NE-CZ-NH1	4.53	126.03	121.50	1	1
A	440	ARG	NE-CZ-NH2	4.50	123.25	119.20	1	1
B	430	GLN	OE1-CD-NE2	4.45	118.15	122.60	1	1
B	95	GLN	OE1-CD-NE2	4.43	118.17	122.60	1	1
A	3	GLN	OE1-CD-NE2	4.42	118.18	122.60	1	1
B	527	GLN	OE1-CD-NE2	4.40	118.20	122.60	1	1
A	4	GLN	OE1-CD-NE2	4.38	118.22	122.60	1	1
A	523	GLN	OE1-CD-NE2	4.38	118.22	122.60	1	1
A	503	ASP	CA-CB-CG	4.33	116.93	112.60	1	1
B	267	GLN	OE1-CD-NE2	4.31	118.29	122.60	1	1
B	4	GLN	OE1-CD-NE2	4.30	118.30	122.60	1	1
A	5	GLN	OE1-CD-NE2	4.28	118.32	122.60	1	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	12	GLN	OE1-CD-NE2	4.27	118.33	122.60	1	1
A	438	GLN	OE1-CD-NE2	4.27	118.33	122.60	1	1
A	216	ARG	NE-CZ-NH2	4.25	123.02	119.20	1	1
A	9	ASP	CA-CB-CG	4.19	116.79	112.60	1	1
B	528	LYS	C-N-CA	4.18	129.22	121.70	1	1
B	540	ASN	OD1-CG-ND2	4.16	118.44	122.60	1	1
B	629	GLN	OE1-CD-NE2	4.16	118.44	122.60	1	1
B	401	GLN	OE1-CD-NE2	4.15	118.45	122.60	1	1
B	787	GLN	OE1-CD-NE2	4.09	118.51	122.60	1	1
B	81	HIS	CB-CG-CD2	4.08	125.89	131.20	1	1
B	445	GLN	OE1-CD-NE2	4.02	118.58	122.60	1	1
B	357	HIS	CB-CG-CD2	4.01	125.98	131.20	1	1
B	16	GLN	OE1-CD-NE2	4.01	118.59	122.60	1	1
B	542	ASP	CA-CB-CG	4.01	116.61	112.60	1	1
A	51	ASP	CA-CB-CG	4.00	116.60	112.60	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	1.39	32

There are 32 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)
A:233:LEU:HD21	A:605:MET:HE3	0.97	1	1
A:229:TYR:CE1	A:605:MET:HE1	0.78	1	1
B:390:LEU:HD11	B:432:LEU:HD22	0.66	1	1
B:75:ILE:HD11	B:83:ASP:HA	0.62	1	1
A:97:MET:HE3	A:134:VAL:HG21	0.62	1	1
B:776:ILE:HD11	B:805:LEU:CD2	0.61	1	1
B:713:ILE:HD11	B:721:LYS:HE2	0.61	1	1
B:359:LYS:HE2	B:465:LEU:HD21	0.58	1	1
B:713:ILE:HB	B:750:LEU:HD21	0.55	1	1
A:329:ILE:HD11	A:344:LEU:HD22	0.55	1	1
B:760:GLU:HB2	B:776:ILE:HD12	0.55	1	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
B:722:ARG:HD3	B:776:ILE:HG23	0.55	1	1
A:427:VAL:HG22	A:499:MET:HE2	0.53	1	1
A:285:HIS:HB3	A:329:ILE:HD13	0.51	1	1
A:633:VAL:HG21	B:27:MET:HG3	0.51	1	1
A:229:TYR:CZ	A:605:MET:HE1	0.50	1	1
B:776:ILE:HD11	B:805:LEU:HD22	0.49	1	1
B:58:MET:HE1	B:70:ILE:HD11	0.49	1	1
A:280:TYR:CE2	A:342:THR:HG21	0.48	1	1
B:666:LYS:HE2	B:749:GLN:OE1	0.47	1	1
B:53:MET:HE3	B:70:ILE:HD13	0.46	1	1
B:390:LEU:HD11	B:432:LEU:CD2	0.46	1	1
A:100:LEU:HD22	A:120:VAL:CG1	0.45	1	1
B:666:LYS:HE2	B:749:GLN:CD	0.45	1	1
B:712:LEU:HD21	B:762:LEU:HD13	0.44	1	1
B:114:ASP:OD1	B:270:LYS:HE3	0.43	1	1
B:510:LEU:HD22	B:601:ILE:HD11	0.42	1	1
A:289:PHE:CD2	A:327:ILE:HD11	0.42	1	1
B:348:LYS:HE2	B:352:GLU:OE2	0.41	1	1
B:58:MET:HE3	B:65:LYS:HB2	0.41	1	1
A:329:ILE:HD11	A:344:LEU:CD2	0.40	1	1
A:237:LYS:HE2	A:612:GLU:OE1	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	1455	1410	41	4

There are 4 unique backbone outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	4	GLN	1
A	6	ASN	1
B	5	ASN	1
B	537	MET	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	1251	1190	33	28

There are 28 unique sidechain outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	120	VAL	1
A	130	THR	1
A	133	ASP	1
A	188	LEU	1
A	200	THR	1
A	206	THR	1
A	326	ILE	1
A	395	LEU	1
A	511	THR	1
A	537	LEU	1
A	565	THR	1
B	6	THR	1
B	32	SER	1
B	84	SER	1
B	290	THR	1
B	435	LEU	1
B	490	SER	1
B	492	LEU	1
B	592	VAL	1
B	605	THR	1
B	638	LEU	1
B	641	LEU	1
B	750	LEU	1
B	756	THR	1
B	785	VAL	1
B	806	VAL	1
B	819	GLU	1
B	820	GLU	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 10 crosslinking restraints combined in 10 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	6
SDA	LYS	CA	SER	CA	upper bound	25.00	4

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
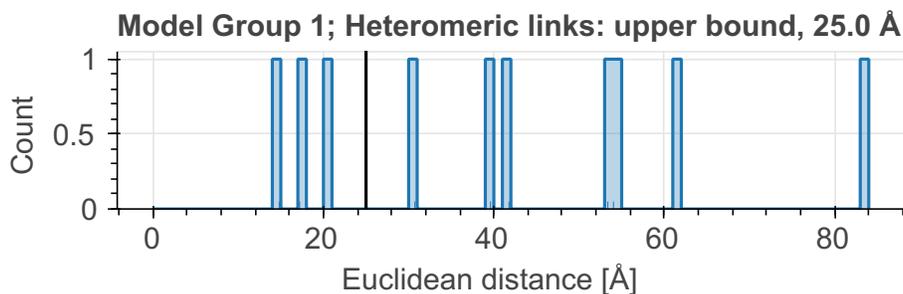
5.2.2. Satisfaction of restraints ⊚

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=10)
1	1	1	1/1	All	30.00	70.00	10
				Heteromeric links/ Intermolecular	30.00	70.00	10

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.