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	9A4O	pdb_00009a4o
PDB-Dev ID	PDBDEV_00000245
Structure Title	Integrative model of IF2-TAGH 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.
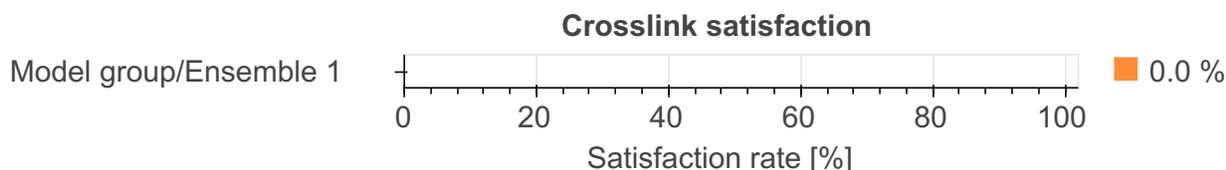
Model Quality: MolProbity Analysis ?

Data Quality ?
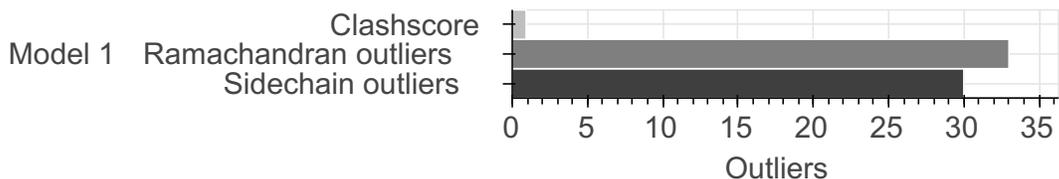
Crosslinking-MS Data Quality

Fit to Data Used for Modeling ?

Crosslink satisfaction
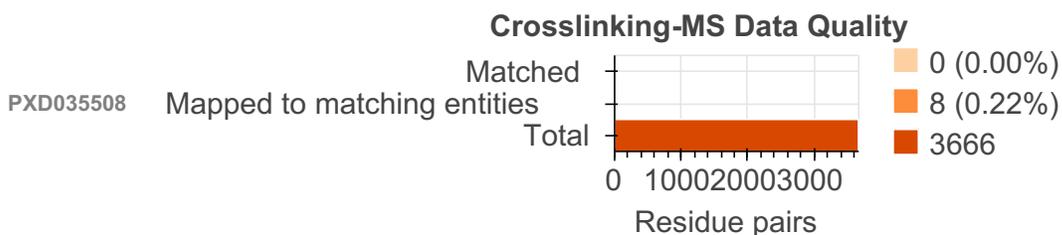
2. Model Details ?

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	IF2_BACSU	A	716	-	1-716	100.00 / 0.00	Atomic
		2	TAGH_BACSU	B	527	-	1-527	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	IF2_BACSU	dbseq_P17889_target	0.00	True
2	TAGH_BACSU	dbseq_P42954_target	0.00	True

Residue pairs stats:

Source	Total	In matched entities	Total matched
9A4O	1	1 (100.00%)	0 (0.00%)
PXD035508	3666	8 (0.22%)	0 (0.00%)

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 71 bond length outliers in this entry (0.72% of 9830 assessed bonds). A summary is provided below.

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
B	313	GLY	N-CA	8.52	1.59	1.45	1	1
B	313	GLY	CA-C	7.73	1.66	1.52	1	1
B	316	PRO	N-CA	7.57	1.58	1.47	1	1
B	309	ILE	N-CA	7.35	1.60	1.46	1	1
B	308	THR	N-CA	7.27	1.60	1.46	1	1
B	315	ILE	N-CA	7.17	1.59	1.46	1	1
B	307	ARG	N-CA	7.08	1.59	1.46	1	1
B	305	PRO	N-CA	7.05	1.57	1.47	1	1
B	319	GLU	N-CA	6.98	1.59	1.46	1	1
B	317	GLN	N-CA	6.93	1.59	1.46	1	1
B	312	PHE	CA-C	6.88	1.67	1.52	1	1
B	318	ASN	CA-C	6.86	1.67	1.52	1	1
B	315	ILE	CA-C	6.84	1.67	1.52	1	1
B	306	LEU	CA-C	6.76	1.67	1.52	1	1
B	310	ALA	C-N	6.76	1.42	1.33	1	1
B	304	ALA	CA-C	6.70	1.67	1.52	1	1
B	307	ARG	CA-C	6.69	1.67	1.52	1	1
B	317	GLN	CA-C	6.69	1.67	1.52	1	1
B	314	ALA	N-CA	6.68	1.58	1.46	1	1
B	309	ILE	C-N	6.68	1.42	1.33	1	1
B	314	ALA	C-N	6.65	1.42	1.33	1	1
B	308	THR	CA-C	6.64	1.66	1.52	1	1
B	314	ALA	CA-C	6.63	1.66	1.52	1	1
B	311	SER	C-N	6.61	1.42	1.33	1	1
B	316	PRO	C-N	6.61	1.42	1.33	1	1
B	311	SER	N-CA	6.60	1.58	1.46	1	1
B	302	PHE	N-CA	6.59	1.58	1.46	1	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
B	317	GLN	C-N	6.59	1.42	1.33	1	1
B	316	PRO	CA-C	6.58	1.66	1.52	1	1
B	305	PRO	C-N	6.48	1.42	1.33	1	1
B	312	PHE	C-N	6.47	1.42	1.33	1	1
B	313	GLY	C-N	6.44	1.42	1.33	1	1
B	310	ALA	N-CA	6.44	1.58	1.46	1	1
B	318	ASN	N-CA	6.42	1.58	1.46	1	1
B	306	LEU	N-CA	6.42	1.58	1.46	1	1
B	303	ASN	CA-C	6.41	1.66	1.52	1	1
B	302	PHE	C-N	6.39	1.42	1.33	1	1
B	308	THR	C-N	6.35	1.42	1.33	1	1
B	300	MET	N-CA	6.33	1.58	1.46	1	1
B	310	ALA	CA-C	6.30	1.66	1.52	1	1
B	306	LEU	C-N	6.27	1.42	1.33	1	1
B	319	GLU	CA-C	6.26	1.66	1.52	1	1
B	305	PRO	CA-C	6.19	1.66	1.52	1	1
B	311	SER	CA-C	6.17	1.65	1.52	1	1
B	304	ALA	N-CA	6.14	1.57	1.46	1	1
B	312	PHE	N-CA	6.11	1.57	1.46	1	1
B	319	GLU	C-N	6.05	1.41	1.33	1	1
B	318	ASN	C-N	6.03	1.41	1.33	1	1
B	303	ASN	C-N	5.95	1.41	1.33	1	1
B	320	VAL	N-CA	5.93	1.57	1.46	1	1
B	298	GLY	N-CA	5.92	1.55	1.45	1	1
B	301	PHE	C-N	5.88	1.41	1.33	1	1
B	301	PHE	N-CA	5.84	1.57	1.46	1	1
B	299	THR	CA-C	5.83	1.65	1.52	1	1
B	300	MET	C-N	5.73	1.41	1.33	1	1
B	303	ASN	N-CA	5.73	1.57	1.46	1	1
B	309	ILE	CA-C	5.66	1.64	1.52	1	1
B	315	ILE	C-N	5.65	1.43	1.34	1	1
B	307	ARG	C-N	5.64	1.41	1.33	1	1
B	301	PHE	CA-C	5.60	1.64	1.52	1	1
B	298	GLY	CA-C	5.57	1.62	1.52	1	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
B	302	PHE	CA-C	5.45	1.64	1.52	1	1
B	299	THR	N-CA	5.40	1.56	1.46	1	1
B	298	GLY	C-N	5.38	1.40	1.33	1	1
B	304	ALA	C-N	5.32	1.43	1.34	1	1
B	300	MET	CA-C	5.23	1.63	1.52	1	1
B	320	VAL	C-N	4.61	1.39	1.33	1	1
B	321	LYS	N-CA	4.53	1.54	1.46	1	1
B	299	THR	C-N	4.48	1.39	1.33	1	1
B	320	VAL	CA-C	4.47	1.62	1.52	1	1
B	297	ALA	CA-C	4.09	1.61	1.52	1	1

Standard geometry: angle outliers ⓘ

There are 251 bond angle outliers in this entry (1.90% of 13231 assessed bonds). A summary is provided below. The output is limited to 100 rows.

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
B	310	ALA	C-N-CA	16.83	151.99	121.70	1	1
B	309	ILE	C-N-CA	15.99	150.49	121.70	1	1
B	311	SER	C-N-CA	15.28	149.20	121.70	1	1
B	302	PHE	C-N-CA	14.47	147.74	121.70	1	1
B	310	ALA	CA-C-N	11.56	139.31	116.20	1	1
B	311	SER	CA-C-N	11.50	139.21	116.20	1	1
B	309	ILE	CA-C-N	10.39	136.98	116.20	1	1
B	302	PHE	CA-C-N	10.16	136.53	116.20	1	1
A	76	GLN	C-N-CA	9.78	139.30	121.70	1	1
A	26	ASN	C-N-CA	9.76	139.27	121.70	1	1
A	38	LEU	C-N-CA	9.39	138.61	121.70	1	1
A	58	LYS	C-N-CA	9.36	138.54	121.70	1	1
B	299	THR	N-CA-C	8.67	135.27	111.00	1	1
A	77	GLN	C-N-CA	8.57	137.13	121.70	1	1
A	6	VAL	C-N-CA	8.51	137.01	121.70	1	1
A	85	LYS	C-N-CA	8.45	136.91	121.70	1	1
B	310	ALA	O-C-N	8.38	109.60	123.00	1	1
B	307	ARG	N-CA-C	8.36	134.42	111.00	1	1
A	65	THR	C-N-CA	8.31	136.65	121.70	1	1
A	7	TYR	C-N-CA	8.26	136.57	121.70	1	1
A	86	ILE	C-N-CA	8.09	136.27	121.70	1	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
B	299	THR	CA-C-N	8.07	132.33	116.20	1	1
B	311	SER	O-C-N	8.04	110.13	123.00	1	1
A	92	ASN	CA-CB-CG	8.03	120.63	112.60	1	1
A	101	LYS	C-N-CA	7.96	136.03	121.70	1	1
B	308	THR	N-CA-C	7.81	132.88	111.00	1	1
A	23	ALA	C-N-CA	7.80	135.74	121.70	1	1
A	79	ALA	C-N-CA	7.68	135.52	121.70	1	1
A	114	ASN	C-N-CA	7.37	134.97	121.70	1	1
B	309	ILE	O-C-N	7.30	111.32	123.00	1	1
B	302	PHE	O-C-N	7.22	111.45	123.00	1	1
B	321	LYS	N-CA-C	7.21	131.20	111.00	1	1
B	306	LEU	N-CA-C	7.20	131.17	111.00	1	1
A	11	LYS	C-N-CA	7.14	134.55	121.70	1	1
A	37	MET	C-N-CA	7.10	134.48	121.70	1	1
A	112	LYS	C-N-CA	6.99	134.28	121.70	1	1
B	318	ASN	N-CA-C	6.99	130.57	111.00	1	1
A	15	VAL	C-N-CA	6.92	134.15	121.70	1	1
A	38	LEU	CA-C-N	6.90	130.00	116.20	1	1
B	316	PRO	CA-N-CD	6.86	102.40	112.00	1	1
B	308	THR	C-N-CA	6.81	133.96	121.70	1	1
A	77	GLN	CA-C-N	6.78	129.76	116.20	1	1
A	96	ASN	C-N-CA	6.75	133.84	121.70	1	1
A	71	PRO	C-N-CA	6.72	133.80	121.70	1	1
A	4	MET	C-N-CA	6.69	133.74	121.70	1	1
A	31	VAL	CA-CB-CG1	6.68	121.76	110.40	1	1
B	315	ILE	CA-C-N	6.63	126.85	116.90	1	1
A	94	VAL	C-N-CA	6.62	133.62	121.70	1	1
B	297	ALA	C-N-CA	6.46	133.34	121.70	1	1
B	301	PHE	C-N-CA	6.44	133.29	121.70	1	1
A	66	ASN	CA-CB-CG	6.43	119.03	112.60	1	1
B	312	PHE	CA-C-N	6.35	128.91	116.20	1	1
B	298	GLY	N-CA-C	6.34	131.67	113.30	1	1
A	26	ASN	CA-C-N	6.29	128.79	116.20	1	1
A	59	SER	C-N-CA	6.28	133.01	121.70	1	1
B	305	PRO	CA-N-CD	6.22	103.29	112.00	1	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	63	ALA	C-N-CA	6.21	132.88	121.70	1	1
B	304	ALA	CA-C-N	6.21	126.22	116.90	1	1
A	1	MET	C-N-CA	6.21	132.88	121.70	1	1
B	318	ASN	C-N-CA	6.20	132.87	121.70	1	1
B	319	GLU	C-N-CA	6.18	132.83	121.70	1	1
A	77	GLN	O-C-N	6.18	113.12	123.00	1	1
B	307	ARG	C-N-CA	6.10	132.68	121.70	1	1
A	7	TYR	CA-C-N	6.07	128.34	116.20	1	1
A	47	ASP	C-N-CA	6.02	132.54	121.70	1	1
A	98	GLN	C-N-CA	5.99	132.48	121.70	1	1
B	311	SER	CA-C-O	5.96	110.66	120.80	1	1
A	526	GLN	OE1-CD-NE2	5.96	116.64	122.60	1	1
B	316	PRO	N-CA-C	5.88	126.80	112.10	1	1
A	604	ASP	CA-CB-CG	5.83	118.43	112.60	1	1
B	305	PRO	N-CA-C	5.83	126.67	112.10	1	1
B	314	ALA	C-N-CA	5.81	132.15	121.70	1	1
A	49	LYS	C-N-CA	5.80	132.14	121.70	1	1
B	319	GLU	N-CA-C	5.78	127.19	111.00	1	1
B	306	LEU	CA-C-N	5.78	127.76	116.20	1	1
A	94	VAL	CA-CB-CG2	5.77	120.20	110.40	1	1
A	59	SER	CA-C-N	5.76	127.72	116.20	1	1
B	300	MET	N-CA-C	5.76	127.13	111.00	1	1
B	310	ALA	CA-C-O	5.72	111.08	120.80	1	1
B	303	ASN	CA-C-N	5.72	127.63	116.20	1	1
A	76	GLN	CA-C-N	5.71	127.63	116.20	1	1
A	7	TYR	O-C-N	5.70	113.89	123.00	1	1
B	307	ARG	CA-C-N	5.68	127.57	116.20	1	1
A	109	LYS	C-N-CA	5.60	131.78	121.70	1	1
B	316	PRO	C-N-CA	5.60	131.77	121.70	1	1
A	16	SER	C-N-CA	5.59	131.77	121.70	1	1
A	92	ASN	CA-C-N	5.58	127.37	116.20	1	1
A	27	MET	C-N-CA	5.52	131.63	121.70	1	1
B	304	ALA	N-CA-C	5.50	126.40	111.00	1	1
B	317	GLN	N-CA-C	5.45	126.26	111.00	1	1
A	69	LYS	C-N-CA	5.43	131.47	121.70	1	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
B	317	GLN	CA-C-N	5.42	127.04	116.20	1	1
B	318	ASN	CA-C-N	5.41	127.01	116.20	1	1
A	6	VAL	CA-C-N	5.41	127.01	116.20	1	1
B	299	THR	O-C-N	5.40	114.37	123.00	1	1
A	516	ASP	CA-CB-CG	5.38	117.98	112.60	1	1
B	309	ILE	CA-C-O	5.36	111.69	120.80	1	1
A	96	ASN	CA-C-N	5.35	126.89	116.20	1	1
A	660	ARG	CD-NE-CZ	5.31	131.84	124.40	1	1
A	103	LYS	C-N-CA	5.30	131.24	121.70	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	0.92	18

There are 18 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:15:VAL:HG21	A:22:THR:H	0.67	1	1
A:86:ILE:H	A:87:ARG:HB2	0.58	1	1
A:6:VAL:HG13	A:30:GLU:CD	0.53	1	1
A:11:LYS:HZ3	A:30:GLU:CD	0.50	1	1
A:223:THR:HB	A:290:THR:HG21	0.50	1	1
B:71:LYS:HE3	B:82:MET:SD	0.47	1	1
B:240:PHE:O	B:248:LYS:HE3	0.46	1	1
A:86:ILE:H	A:87:ARG:CB	0.45	1	1
B:202:ILE:HG21	B:234:LYS:HE3	0.45	1	1
A:61:LYS:HB3	A:62:PRO:HD3	0.44	1	1
A:65:THR:HB	A:66:ASN:HB3	0.44	1	1
A:7:TYR:H	A:8:GLU:HB2	0.43	1	1
A:93:ASP:HB3	A:94:VAL:HG22	0.43	1	1
A:65:THR:HB	A:66:ASN:CB	0.42	1	1
A:294:ILE:HD11	A:376:ILE:HD11	0.42	1	1
B:62:LYS:HE2	B:198:VAL:CG1	0.42	1	1
B:112:MET:HA	B:112:MET:HE2	0.41	1	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:31:VAL:CG1	A:32:ASN:HA	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	1239	1110	96	33

There are 33 unique backbone outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	5	ARG	1
A	7	TYR	1
A	8	GLU	1
A	12	ALA	1
A	15	VAL	1
A	17	SER	1
A	21	LEU	1
A	26	ASN	1
A	28	ASP	1
A	30	GLU	1
A	33	ASN	1
A	48	ALA	1
A	50	TYR	1
A	57	ALA	1
A	60	GLN	1
A	62	PRO	1
A	64	GLU	1
A	66	ASN	1
A	70	GLN	1
A	71	PRO	1
A	74	VAL	1
A	78	SER	1
A	83	PRO	1
A	84	ASN	1
A	92	ASN	1
A	102	ASN	1

Chain	Res	Type	Models (Total)
A	107	ASN	1
A	113	ARG	1
A	575	ASN	1
A	676	PHE	1
B	298	GLY	1
B	305	PRO	1
B	316	PRO	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	1055	963	62	30

There are 30 unique sidechain outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	11	LYS	1
A	13	LEU	1
A	15	VAL	1
A	25	LYS	1
A	35	MET	1
A	38	LEU	1
A	39	GLU	1
A	40	GLU	1
A	94	VAL	1
A	99	PHE	1
A	105	ASN	1
A	122	HIS	1
A	143	LEU	1
A	477	THR	1
A	600	TYR	1
A	680	VAL	1
B	176	THR	1
B	267	PHE	1
B	270	PHE	1
B	286	ILE	1

Chain	Res	Type	Models (Total)
B	292	LEU	1
B	293	THR	1
B	305	PRO	1
B	308	THR	1
B	312	PHE	1
B	316	PRO	1
B	318	ASN	1
B	319	GLU	1
B	336	THR	1
B	479	ASN	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 1 crosslinking restraints combined in 1 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	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.


Model Group 1; Heteromeric links: upper bound, 25.0 Å
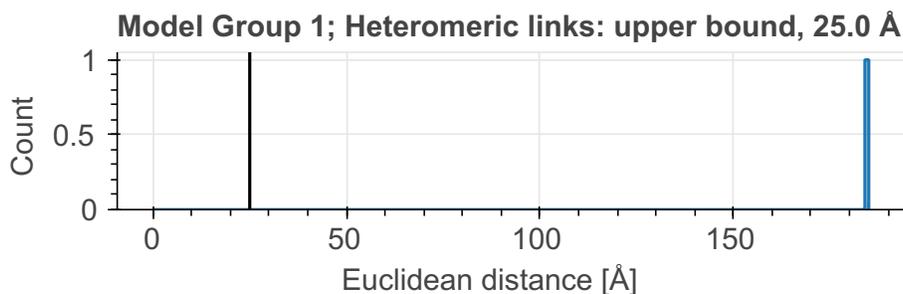
5.2.2. Satisfaction of restraints ❓

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=1)
1	1	1	1/1	All	0.00	100.00	1
				Heteromeric links/ Intermolecular	0.00	100.00	1

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

---
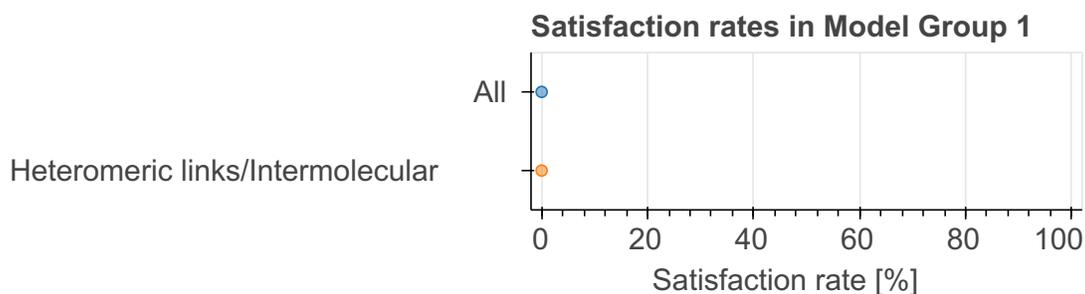
6. Fit to Data Used for Validation Assessment ❓

Validation for this section is under development.