Integrative Structure Validation Report ?

October 09, 2025 - 04:41 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	9A2I	pdb_00009a2i
PDB-Dev ID	PDBDEV_00000167
Structure Title	Model of E. coli DnaK by in-cell photo-crosslinking MS and deep learning
Structure Authors	Stahl, K.; Graziadei, A.; Dau, T.; Brock, O.; Rappsilber, J.
Deposited on	2023-02-03

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.
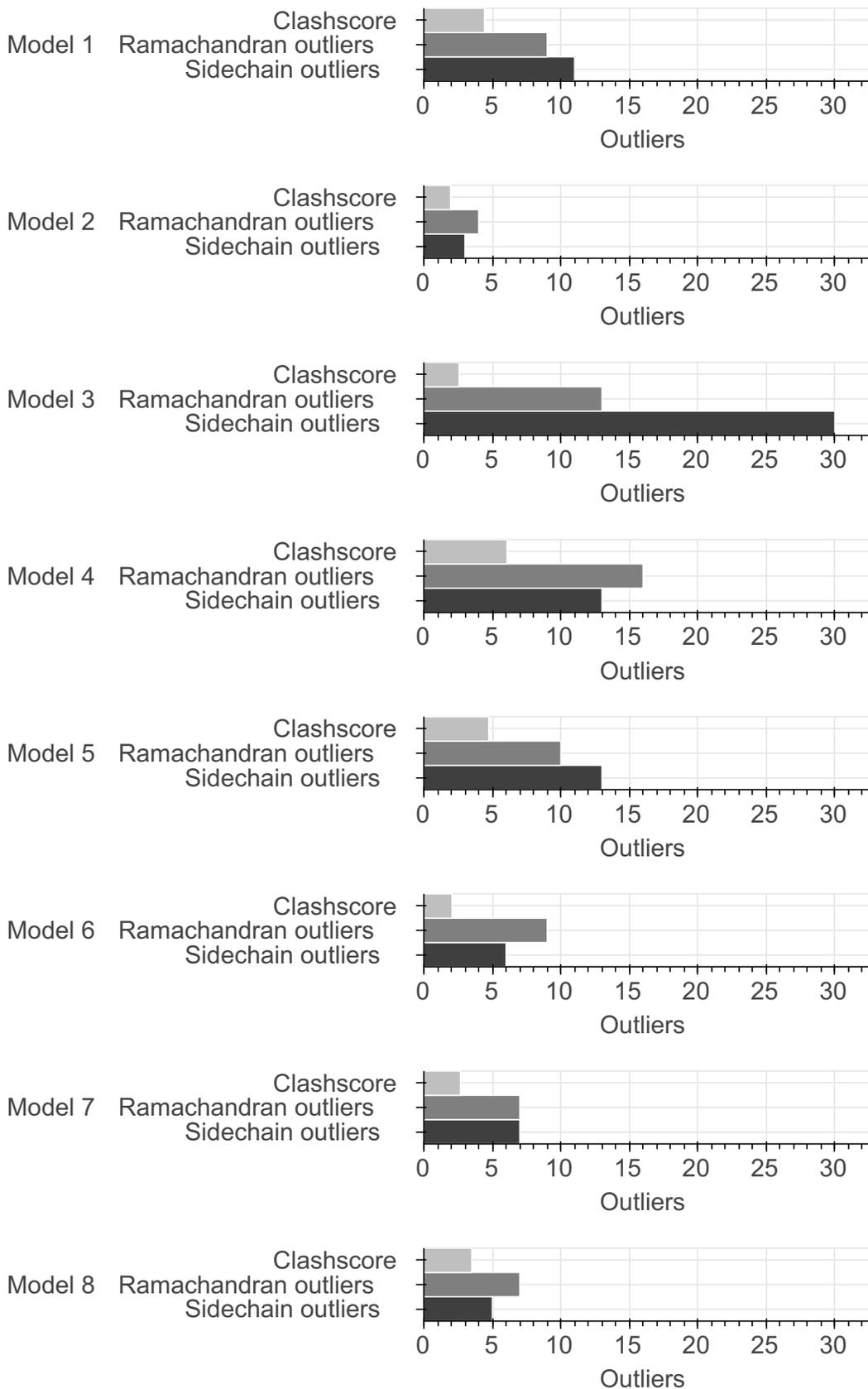
1. Overview ?

1.1. Summary ?

This entry consists of 10 model(s). A total of 1 dataset(s) were used to build this entry.
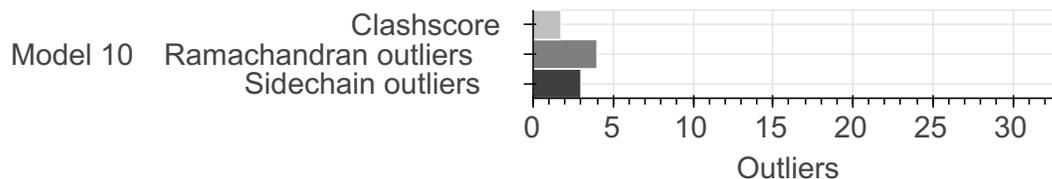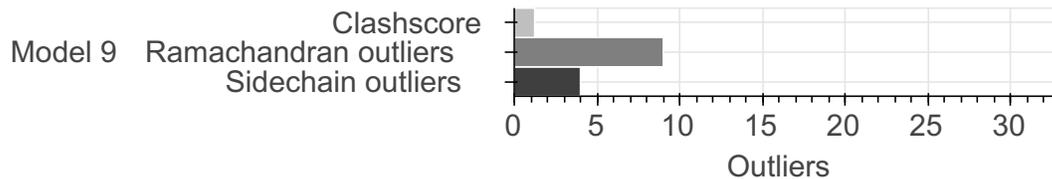
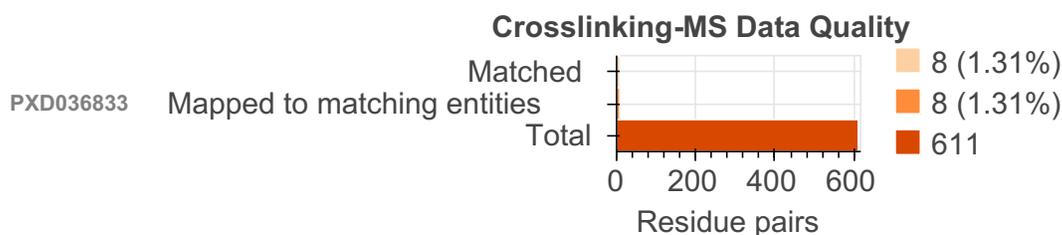| Name	Type	Count
Crosslinking-MS data	Experimental data	1

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 ?

Model 1
Clashscore
Ramachandran outliers
Sidechain outliers
Outliers

Model 2
Clashscore
Ramachandran outliers
Sidechain outliers
Outliers

Model 3
Clashscore
Ramachandran outliers
Sidechain outliers
Outliers

Model 4
Clashscore
Ramachandran outliers
Sidechain outliers
Outliers

Model 5
Clashscore
Ramachandran outliers
Sidechain outliers
Outliers

Model 6
Clashscore
Ramachandran outliers
Sidechain outliers
Outliers

Model 7
Clashscore
Ramachandran outliers
Sidechain outliers
Outliers

Model 8
Clashscore
Ramachandran outliers
Sidechain outliers
Outliers

Model 9 — Clashscore, Ramachandran outliers, Sidechain outliers (Outliers)

Model 10 — Clashscore, Ramachandran outliers, Sidechain outliers (Outliers)

Data Quality ?

Crosslinking-MS Data Quality

PXD036833

- Matched — 8 (1.31%)
- Mapped to matching entities — 8 (1.31%)
- Total — 611

(Residue pairs)

Fit to Data Used for Modeling ?

Crosslink satisfaction

Model group/Ensemble 1

55.56 %

(Satisfaction rate [%])

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	1-10	1	P0A6Y8	A	638	-	1-638	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	jPOSTrepo	JPST001851

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	AlphaLink	AlphaLink with 10 msa subsamples	Not available	10	False	False

There is 1 software package reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	AlphaLink	1.00	model building	https://github.com/lhatsk/AlphaLink

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).

PXD036833

Number of entities in the crosslinking-MS dataset:	1102

Number of entities in the entry:	1

Matching entities:

Entity ID	Molecule name	Crosslinking-MS Entity ID	E-value	Exact match
1	P0A6Y8	dbseq_P0A6Y8_target	0.00	True

Residue pairs stats:

Source	Total	In matched entities	Total matched
9A2I	9	9 (100.00%)	8 (88.89%)
PXD036833	611	8 (1.31%)	8 (1.31%)

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 7 bond length outliers in this entry (0.01% of 49020 assessed bonds). A summary is provided below.

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	485	HIS	CD2-NE2	169.52	3.24	1.37	3	1
A	485	HIS	CE1-NE2	121.54	2.54	1.32	3	1
A	485	HIS	ND1-CE1	31.07	1.01	1.32	3	1
A	46	GLY	CA-C	4.68	1.60	1.52	4	1
A	24	PRO	N-CD	4.59	1.41	1.47	4	1
A	467	ARG	CZ-NH2	4.26	1.27	1.33	7	1
A	605	GLN	CA-C	4.08	1.61	1.52	3	1

Standard geometry: angle outliers ?

There are 419 bond angle outliers in this entry (0.63% of 66240 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)
A	485	HIS	ND1-CE1-NE2	51.85	56.55	108.40	3	1
A	485	HIS	CD2-NE2-CE1	43.28	65.72	109.00	3	1
A	485	HIS	CG-CD2-NE2	29.17	78.03	107.20	3	1
A	86	VAL	C-N-CA	19.55	156.89	121.70	3	1
A	605	GLN	C-N-CA	15.92	150.36	121.70	3	1
A	605	GLN	CA-C-N	12.64	141.48	116.20	3	1
A	86	VAL	CA-C-N	12.37	140.95	116.20	3	1
A	86	VAL	O-C-N	12.04	103.74	123.00	3	1
A	605	GLN	O-C-N	11.07	105.29	123.00	3	1
A	485	HIS	CG-ND1-CE1	10.47	127.10	109.30	3	1
A	620	ASN	C-N-CA	10.24	140.13	121.70	7	1
A	24	PRO	CA-N-CD	8.34	100.33	112.00	4	1
A	606	HIS	CA-CB-CG	8.00	121.80	113.80	3	1
A	86	VAL	N-CA-C	7.73	132.65	111.00	3	1
A	375	ALA	C-CA-CB	7.67	122.01	110.50	5	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	64	ASN	CA-CB-CG	7.66	120.26	112.60	5	1
A	201	ASP	CA-CB-CG	7.64	120.24	112.60	5	1
A	637	LYS	C-N-CA	7.50	135.20	121.70	9	5
A	636	ASP	C-N-CA	7.41	135.03	121.70	6	4
A	94	ILE	CA-CB-CG1	7.39	122.96	110.40	3	1
A	87	SER	N-CA-C	7.20	131.16	111.00	3	1
A	624	ASP	C-N-CA	6.82	133.98	121.70	9	3
A	544	HIS	CA-CB-CG	6.74	120.54	113.80	2	1
A	544	HIS	CB-CG-CD2	6.73	122.45	131.20	2	4
A	622	LYS	C-N-CA	6.67	133.70	121.70	4	4
A	471	GLN	OE1-CD-NE2	6.56	116.04	122.60	9	2
A	619	ASN	C-N-CA	6.36	133.15	121.70	7	2
A	23	THR	CA-C-N	6.35	126.43	116.90	4	1
A	103	VAL	C-CA-CB	6.35	123.46	111.40	3	1
A	429	ALA	C-CA-CB	6.24	119.85	110.50	4	1
A	43	THR	CA-CB-OG1	6.23	118.94	109.60	4	1
A	625	ASP	N-CA-CB	6.20	99.96	110.50	9	1
A	23	THR	OG1-CB-CG2	6.18	121.66	109.30	4	1
A	626	VAL	C-N-CA	6.09	132.67	121.70	8	2
A	623	ASP	C-N-CA	6.08	132.65	121.70	4	2
A	606	HIS	CB-CG-CD2	6.01	123.39	131.20	3	1
A	29	ASN	OD1-CG-ND2	5.98	116.62	122.60	4	2
A	260	GLN	OE1-CD-NE2	5.91	116.69	122.60	9	3
A	71	ARG	CD-NE-CZ	5.86	116.19	124.40	6	1
A	606	HIS	C-CA-CB	5.84	121.20	110.10	3	1
A	152	GLN	OE1-CD-NE2	5.81	116.79	122.60	4	6
A	621	ALA	C-N-CA	5.73	132.02	121.70	2	7
A	184	GLY	C-N-CA	5.68	131.93	121.70	4	2
A	605	GLN	N-CA-C	5.63	126.76	111.00	3	1
A	407	VAL	C-N-CA	5.54	131.68	121.70	4	1
A	627	VAL	C-N-CA	5.52	131.64	121.70	8	2
A	625	ASP	C-N-CA	5.49	131.58	121.70	4	1
A	609	GLN	OE1-CD-NE2	5.48	117.12	122.60	3	6
A	86	VAL	CA-CB-CG1	5.47	119.70	110.40	3	1
A	424	GLN	OE1-CD-NE2	5.47	117.13	122.60	10	6

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	78	GLN	OE1-CD-NE2	5.39	117.21	122.60	3	10
A	396	PRO	C-CA-CB	5.39	120.34	110.10	5	1
A	46	GLY	C-N-CA	5.39	131.40	121.70	4	1
A	13	ASN	OD1-CG-ND2	5.38	117.22	122.60	5	3
A	544	HIS	CB-CG-ND1	5.32	130.69	122.70	2	1
A	23	THR	N-CA-CB	5.30	102.49	111.50	4	1
A	57	GLN	OE1-CD-NE2	5.30	117.30	122.60	6	5
A	616	ALA	C-N-CA	5.30	131.24	121.70	7	1
A	617	SER	C-N-CA	5.29	131.22	121.70	7	1
A	47	GLU	N-CA-C	5.28	125.78	111.00	4	1
A	362	ARG	CD-NE-CZ	5.26	131.77	124.40	6	1
A	534	GLN	OE1-CD-NE2	5.26	117.34	122.60	3	3
A	378	GLN	OE1-CD-NE2	5.26	117.34	122.60	2	8
A	396	PRO	C-N-CA	5.24	131.12	121.70	4	1
A	480	ALA	N-CA-CB	5.20	102.60	110.40	8	1
A	69	ILE	CA-CB-CG2	5.20	119.33	110.50	5	1
A	46	GLY	CA-C-N	5.19	126.57	116.20	4	1
A	408	MET	C-N-CA	5.17	131.00	121.70	4	1
A	396	PRO	N-CA-CB	5.16	108.67	103.00	1	1
A	57	GLN	C-N-CA	5.09	130.87	121.70	4	3
A	504	SER	C-N-CA	5.08	130.84	121.70	10	1
A	610	GLN	OE1-CD-NE2	5.07	117.53	122.60	3	5
A	44	GLN	OE1-CD-NE2	5.07	117.53	122.60	5	10
A	63	GLN	OE1-CD-NE2	5.04	117.56	122.60	6	7
A	65	THR	CA-CB-CG2	5.03	119.06	110.50	4	1
A	605	GLN	OE1-CD-NE2	5.03	117.57	122.60	10	10
A	248	GLN	OE1-CD-NE2	5.01	117.59	122.60	2	1
A	46	GLY	N-CA-C	5.00	127.81	113.30	4	1
A	364	ASP	CA-CB-CG	4.97	117.57	112.60	8	2
A	52	GLN	OE1-CD-NE2	4.96	117.64	122.60	9	5
A	549	GLN	OE1-CD-NE2	4.95	117.65	122.60	1	7
A	544	HIS	C-CA-CB	4.94	119.49	110.10	2	1
A	216	PHE	CA-CB-CG	4.94	108.86	113.80	1	2
A	83	GLN	OE1-CD-NE2	4.93	117.67	122.60	10	10
A	34	ARG	NE-CZ-NH2	4.93	123.63	119.20	6	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	477	ASP	CA-CB-CG	4.92	117.52	112.60	10	1
A	625	ASP	CA-CB-CG	4.92	117.52	112.60	9	1
A	456	GLN	OE1-CD-NE2	4.91	117.69	122.60	9	6
A	350	GLN	OE1-CD-NE2	4.90	117.70	122.60	1	1
A	45	ASP	CA-CB-CG	4.89	117.49	112.60	4	1
A	297	ASN	OD1-CG-ND2	4.88	117.72	122.60	8	2
A	620	ASN	CA-C-N	4.87	125.95	116.20	7	1
A	47	GLU	O-C-N	4.86	115.23	123.00	4	1
A	14	SER	C-CA-CB	4.85	119.32	110.10	4	1
A	451	ASN	OD1-CG-ND2	4.82	117.78	122.60	4	1
A	366	ASN	OD1-CG-ND2	4.81	117.79	122.60	7	3
A	108	GLN	OE1-CD-NE2	4.78	117.82	122.60	9	10
A	94	ILE	N-CA-CB	4.76	119.59	111.50	3	1
A	451	ASN	CA-CB-CG	4.76	117.36	112.60	4	1
A	181	LEU	C-N-CA	4.75	130.25	121.70	9	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	4.42	43
2	1.95	19
3	2.57	25
4	6.07	59
5	4.73	46
6	2.06	20
7	2.67	26
8	3.50	34
9	1.23	12
10	1.75	17

There are 301 clashes. The table below contains the detailed list of all clashes based on a MolProbity analysis. Bad clashes are >= 0.4 Angstrom. The output is limited to 100 rows.

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:485:HIS:CE1	A:500:THR:OG1	1.18	3	1
A:181:LEU:HD13	A:216:PHE:CE2	1.00	1	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:485:HIS:ND1	A:485:HIS:NE2	0.94	3	1
A:485:HIS:HE1	A:500:THR:OG1	0.91	3	1
A:181:LEU:HD22	A:216:PHE:CE1	0.89	1	1
A:14:SER:HB2	A:123:MET:HE1	0.84	4	1
A:252:LEU:HD21	A:296:MET:HE1	0.81	10	2
A:232:PHE:CE2	A:346:MET:HE1	0.79	5	2
A:69:ILE:HG22	A:397:LEU:HD22	0.78	5	1
A:69:ILE:CG2	A:397:LEU:HD22	0.78	5	1
A:181:LEU:HD22	A:216:PHE:CZ	0.77	1	1
A:23:THR:HG22	A:24:PRO:HG3	0.77	4	1
A:485:HIS:CE1	A:485:HIS:NE2	0.76	3	1
A:14:SER:CB	A:123:MET:HE1	0.75	4	1
A:395:THR:C	A:416:THR:HG21	0.70	7	1
A:66:LEU:HD22	A:115:ILE:HG23	0.70	4	1
A:244:PHE:HB2	A:296:MET:HE2	0.69	8	4
A:25:ARG:HE	A:49:LEU:HD13	0.69	4	1
A:396:PRO:HB3	A:414:LYS:HA	0.68	7	1
A:407:VAL:HG23	A:448:ALA:HB2	0.68	7	1
A:5:ILE:HG22	A:18:ILE:HG22	0.67	1	5
A:339:LEU:HB3	A:344:THR:HG22	0.67	1	1
A:116:SER:OG	A:397:LEU:HD21	0.66	5	1
A:394:VAL:C	A:416:THR:HB	0.66	7	1
A:396:PRO:C	A:397:LEU:HG	0.65	5	1
A:286:ILE:HG13	A:296:MET:HE2	0.65	2	1
A:557:LEU:HD11	A:598:LEU:HD21	0.64	8	1
A:399:LEU:HD12	A:412:ILE:HD12	0.64	7	1
A:23:THR:HG21	A:46:GLY:C	0.63	4	1
A:252:LEU:HD21	A:296:MET:CE	0.63	2	1
A:73:ILE:HG23	A:153:ALA:HB3	0.61	6	1
A:103:VAL:HG13	A:104:GLU:H	0.61	3	1
A:181:LEU:CD1	A:377:VAL:HG21	0.60	5	2
A:24:PRO:HA	A:43:THR:H	0.60	4	1
A:394:VAL:CG1	A:415:ASN:HA	0.59	6	2
A:258:ALA:HB2	A:285:TYR:CE1	0.59	8	1
A:426:PHE:CD2	A:474:VAL:HB	0.59	4	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:69:ILE:HG22	A:397:LEU:HB3	0.59	5	1
A:145:TYR:CE1	A:225:THR:HG22	0.58	10	1
A:66:LEU:CD2	A:115:ILE:HG23	0.58	4	1
A:29:ASN:HD21	A:36:THR:CB	0.57	8	1
A:39:ILE:HD11	A:55:LYS:HA	0.57	8	1
A:119:VAL:HG12	A:123:MET:HE3	0.57	4	1
A:85:ASP:HB3	A:89:MET:HE3	0.57	4	1
A:173:THR:HG21	A:391:LEU:HD21	0.57	4	1
A:145:TYR:CE2	A:225:THR:HG23	0.57	5	1
A:442:GLN:HE22	A:518:ASP:HB2	0.57	4	1
A:436:VAL:HG21	A:472:ILE:HG13	0.56	5	1
A:179:TYR:CE1	A:338:ILE:HG23	0.56	7	2
A:73:ILE:HG23	A:153:ALA:CB	0.56	6	1
A:411:LEU:HD23	A:429:ALA:HB1	0.56	4	1
A:467:ARG:HH22	A:540:ASP:CG	0.55	7	1
A:339:LEU:HD21	A:353:VAL:HG11	0.55	1	1
A:167:ARG:HH11	A:382:LEU:HD21	0.55	5	1
A:23:THR:HG21	A:46:GLY:O	0.55	4	1
A:258:ALA:HB2	A:285:TYR:HE1	0.55	8	1
A:68:ALA:HB1	A:71:ARG:HH21	0.55	6	1
A:401:ILE:HD11	A:474:VAL:HG11	0.55	7	1
A:407:VAL:HG21	A:441:LEU:CD1	0.55	7	1
A:23:THR:C	A:43:THR:H	0.54	4	1
A:219:LEU:HA	A:480:ALA:HB1	0.54	8	1
A:549:GLN:HB3	A:598:LEU:HD21	0.54	1	1
A:232:PHE:CD2	A:346:MET:HE1	0.54	7	2
A:69:ILE:HD13	A:115:ILE:HD13	0.54	4	1
A:395:THR:O	A:416:THR:HG21	0.54	7	1
A:39:ILE:HG22	A:130:TYR:CE1	0.54	5	1
A:547:ARG:HH22	A:573:GLU:CD	0.54	8	7
A:103:VAL:HG22	A:115:ILE:HD11	0.54	7	2
A:181:LEU:HD13	A:216:PHE:CZ	0.53	1	1
A:239:TYR:CE1	A:304:LYS:HE3	0.53	2	1
A:442:GLN:HB2	A:515:MET:HE3	0.53	4	1
A:24:PRO:HA	A:42:TYR:HA	0.53	4	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:181:LEU:CD2	A:216:PHE:CE1	0.53	1	1
A:181:LEU:HD11	A:377:VAL:HG11	0.52	5	1
A:5:ILE:HD11	A:130:TYR:CD2	0.52	5	1
A:462:ILE:HG21	A:472:ILE:HD11	0.52	9	1
A:23:THR:HG22	A:24:PRO:CG	0.52	4	1
A:64:ASN:HD21	A:403:THR:HG23	0.52	5	1
A:440:VAL:HG12	A:515:MET:HE1	0.52	9	1
A:19:MET:SD	A:380:GLY:HA3	0.52	4	1
A:173:THR:HG21	A:404:MET:CE	0.52	1	1
A:67:PHE:CG	A:89:MET:SD	0.52	5	1
A:73:ILE:HD11	A:397:LEU:HD11	0.52	5	1
A:320:LEU:HD22	A:356:PHE:CE2	0.51	10	1
A:286:ILE:HD12	A:296:MET:HE1	0.51	4	1
A:396:PRO:CB	A:414:LYS:HA	0.51	7	1
A:141:THR:HG21	A:371:VAL:HG12	0.51	6	2
A:115:ILE:HG21	A:399:LEU:HD11	0.51	5	1
A:150:GLN:HE22	A:225:THR:HG21	0.51	5	1
A:70:LYS:HA	A:397:LEU:HD12	0.51	5	1
A:139:VAL:HB	A:390:LEU:HD12	0.51	5	1
A:268:LYS:HE2	A:272:GLU:OE2	0.50	10	6
A:81:GLU:HB3	A:226:HIS:CE1	0.50	1	1
A:51:GLY:HA3	A:122:LYS:HZ2	0.50	8	1
A:427:SER:HB3	A:468:GLY:HA2	0.50	1	1
A:375:ALA:HB1	A:392:LEU:CD2	0.50	5	1
A:177:LEU:HD23	A:181:LEU:HD11	0.50	1	1
A:167:ARG:NH1	A:382:LEU:HD21	0.50	5	1
A:410:THR:HG21	A:422:HIS:CD2	0.50	7	1
A:378:GLN:OE1	A:391:LEU:HD13	0.49	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	636	609	18	9
2	636	619	13	4
3	636	565	58	13

Model ID	Analysed	Favored	Allowed	Outliers
4	636	600	20	16
5	636	598	28	10
6	636	610	17	9
7	636	610	19	7
8	636	607	22	7
9	636	606	21	9
10	636	613	19	4

There are 57 unique backbone outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	619	ASN	4
A	622	LYS	4
A	56	ARG	3
A	59	VAL	3
A	397	LEU	3
A	404	MET	3
A	620	ASN	3
A	623	ASP	3
A	637	LYS	3
A	34	ARG	2
A	54	ALA	2
A	58	ALA	2
A	60	THR	2
A	183	LYS	2
A	185	THR	2
A	367	PRO	2
A	415	ASN	2
A	618	ALA	2
A	625	ASP	2
A	628	ASP	2
A	24	PRO	1
A	31	GLU	1
A	32	GLY	1
A	35	THR	1
A	46	GLY	1

Chain	Res	Type	Models (Total)
A	53	PRO	1
A	87	SER	1
A	92	LYS	1
A	181	LEU	1
A	211	ASP	1
A	285	TYR	1
A	393	ASP	1
A	394	VAL	1
A	400	GLY	1
A	406	GLY	1
A	407	VAL	1
A	408	MET	1
A	409	THR	1
A	410	THR	1
A	411	LEU	1
A	413	ALA	1
A	414	LYS	1
A	417	THR	1
A	419	PRO	1
A	431	ASP	1
A	434	SER	1
A	448	ALA	1
A	503	ALA	1
A	504	SER	1
A	505	SER	1
A	506	GLY	1
A	617	SER	1
A	621	ALA	1
A	624	ASP	1
A	627	VAL	1
A	629	ALA	1
A	636	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	521	498	12	11
2	521	507	11	3
3	521	451	40	30
4	521	492	16	13
5	521	495	13	13
6	521	502	13	6
7	521	504	10	7
8	521	499	17	5
9	521	495	22	4
10	521	505	13	3

There are 68 unique sidechain outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	611	THR	5
A	417	THR	4
A	420	THR	4
A	1	MET	3
A	394	VAL	3
A	395	THR	3
A	60	THR	2
A	65	THR	2
A	340	VAL	2
A	367	PRO	2
A	399	LEU	2
A	409	THR	2
A	410	THR	2
A	411	LEU	2
A	416	THR	2
A	625	ASP	2
A	636	ASP	2
A	12	THR	1
A	14	SER	1
A	23	THR	1
A	24	PRO	1
A	33	ASP	1

Chain	Res	Type	Models (Total)
A	36	THR	1
A	38	SER	1
A	39	ILE	1
A	42	TYR	1
A	43	THR	1
A	45	ASP	1
A	50	VAL	1
A	57	GLN	1
A	78	GLN	1
A	85	ASP	1
A	86	VAL	1
A	87	SER	1
A	94	ILE	1
A	97	ASP	1
A	104	GLU	1
A	123	MET	1
A	189	THR	1
A	215	THR	1
A	225	THR	1
A	283	LEU	1
A	287	THR	1
A	298	ILE	1
A	304	LYS	1
A	332	SER	1
A	386	VAL	1
A	388	ASP	1
A	393	ASP	1
A	397	LEU	1
A	402	GLU	1
A	412	ILE	1
A	428	THR	1
A	436	VAL	1
A	467	ARG	1
A	469	MET	1
A	508	ASN	1

Chain	Res	Type	Models (Total)
A	543	LEU	1
A	544	HIS	1
A	548	LYS	1
A	557	LEU	1
A	598	LEU	1
A	605	GLN	1
A	606	HIS	1
A	615	ASP	1
A	620	ASN	1
A	623	ASP	1
A	631	PHE	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 9 crosslinking restraints combined in 9 restraint groups.

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
L-Photo-Leucine	ASP	CA	LEU	CA	upper bound	10.00	1
L-Photo-Leucine	LEU	CA	SER	CA	upper bound	10.00	1
L-Photo-Leucine	LEU	CA	LYS	CA	upper bound	10.00	1
L-Photo-Leucine	LEU	CA	PRO	CA	upper bound	10.00	1
L-Photo-Leucine	LEU	CA	PHE	CA	upper bound	10.00	1
L-Photo-Leucine	LEU	CA	TYR	CA	upper bound	10.00	1
L-Photo-Leucine	GLU	CA	LEU	CA	upper bound	10.00	1
L-Photo-Leucine	LEU	CA	LEU	CA	upper bound	10.00	2

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; Self-links: upper bound, 10.0 Å
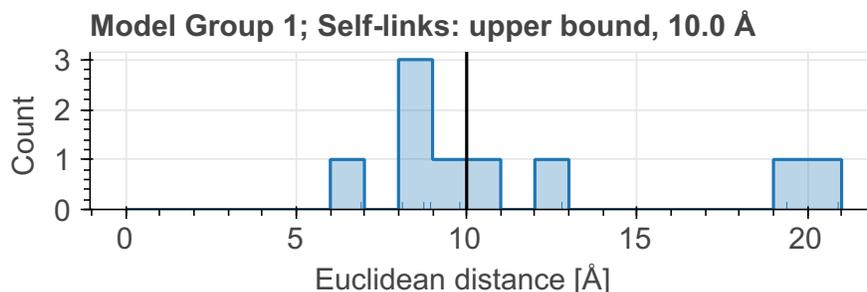
5.2.2. Satisfaction of restraints ⍰

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=9)
1	1	1	10/10	All	55.56	44.44	9
				Self-links/ Intramolecular	55.56	44.44	9

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

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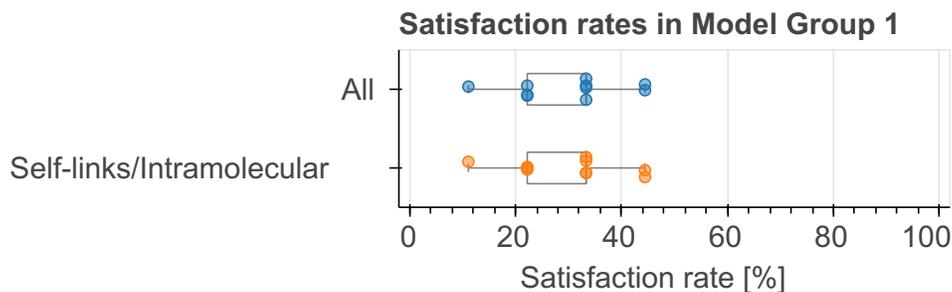
6. Fit to Data Used for Validation Assessment ⍰

Validation for this section is under development.