Integrative Structure Validation Report ?

February 18, 2025 - 08:34 AM PST

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

Integrative Modeling Validation Version 2.0
Python-IHM Version 1.8
MolProbity Version 4.5.2
pyHMMER Version 0.11.0

PDB ID	9A2K
PDB-Dev ID	PDBDEV_00000169
Structure Title	Model of E. coli AceA 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 IM Structure Validation Report for a publicly released PDB-IHM entry.

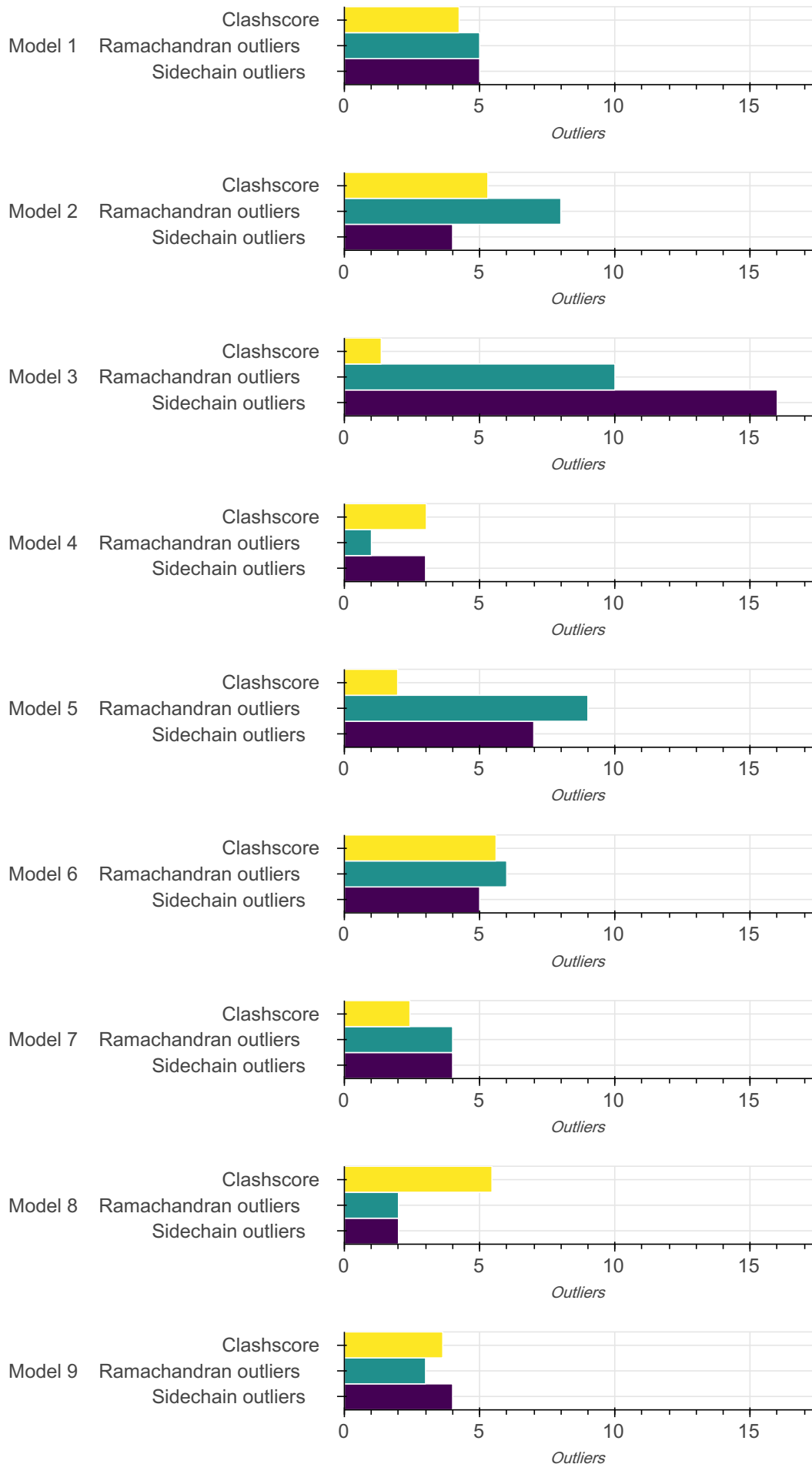*We welcome your comments at helpdesk@pdb-ihm.org*

A user guide is available at https://pdb-ihm.org/validation_help.html with specific help available everywhere you see the ? symbol.
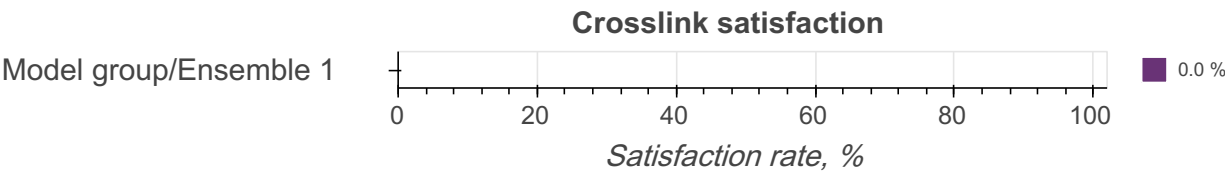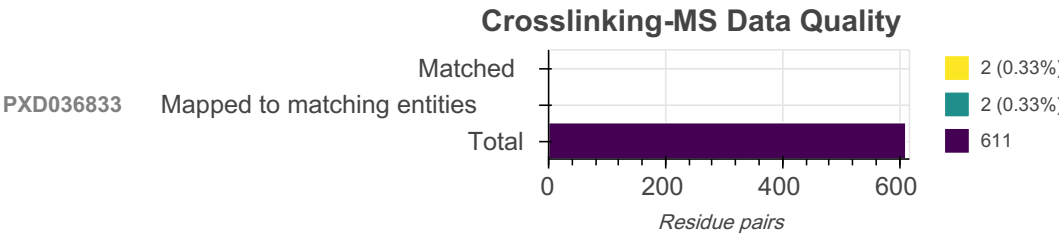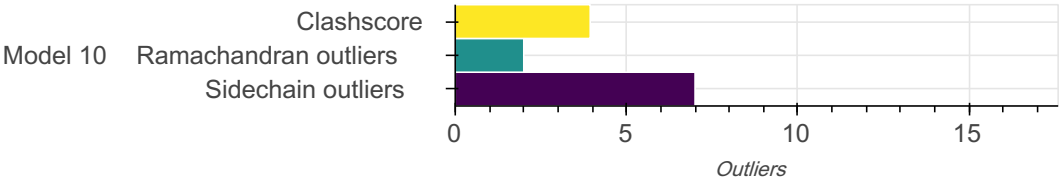
List of references used to build this report is available here.

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

Crosslinking-MS Data Quality

2 (0.33%)	Matched
2 (0.33%)	Mapped to matching entities
611	Total

PXD036833

Crosslink satisfaction

Model group/Ensemble 1 — 0.0 %

Ensemble information ?

This entry consists of 0 distinct ensemble(s).

Summary ?

This entry consists of 10 model(s). A total of 1 datasets were used to build this entry.

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	P0A9G6	A	434	-	1-434	100.00 / 0.00	Atomic

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

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	None	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

Data quality ?

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).

Crosslinking-MS dataset (PRIDE ID)	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	P0A9G6	dbseq_P0A9G6_target	0.00	True

Residue pairs stats:

Source	Total	In matched entities	Total matched
9A2K	2	2 (100.00%)	2 (100.00%)
PXD036833	611	2 (0.33%)	2 (0.33%)

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.

Standard geometry: bond outliers ?

There are no bond length outliers.

Standard geometry: angle outliers ?

There are 240 bond angle outliers in this entry (0.52% of 46490 assessed bonds). A summary is provided below.

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	354	GLY	C-N-CA	8.75	137.45	121.70	1	2
A	194	LYS	C-N-CA	8.48	136.97	121.70	1	1
A	94	GLN	OE1-CD-NE2	7.26	115.34	122.60	1	6
A	275	ARG	NE-CZ-NH2	7.25	125.72	119.20	10	2
A	430	GLU	C-N-CA	6.80	133.95	121.70	5	1
A	322	PHE	CA-CB-CG	6.56	107.24	113.80	3	2
A	191	SER	C-N-CA	6.49	133.39	121.70	6	1
A	100	ASN	OD1-CG-ND2	6.14	116.46	122.60	5	3
A	271	GLN	OE1-CD-NE2	6.14	116.46	122.60	1	7
A	109	GLN	OE1-CD-NE2	6.14	116.46	122.60	8	6
A	188	GLN	OE1-CD-NE2	6.02	116.58	122.60	4	8
A	351	THR	N-CA-CB	6.02	121.73	111.50	8	1
A	351	THR	CA-CB-CG2	5.73	120.24	110.50	8	2
A	400	GLN	OE1-CD-NE2	5.50	117.10	122.60	4	8
A	79	GLN	OE1-CD-NE2	5.48	117.12	122.60	8	8
A	238	ALA	N-CA-CB	5.45	102.22	110.40	9	1
A	76	GLN	OE1-CD-NE2	5.43	117.17	122.60	6	1
A	167	ASN	OD1-CG-ND2	5.41	117.19	122.60	2	3
A	196	GLY	C-N-CA	5.36	131.35	121.70	1	1
A	207	GLN	OE1-CD-NE2	5.26	117.34	122.60	5	9

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	319	SER	CA-C-N	5.26	124.79	116.90	6	1
A	408	PHE	CA-CB-CG	5.25	108.55	113.80	1	1
A	396	PHE	CA-CB-CG	5.22	108.58	113.80	10	1
A	184	HIS	CB-CG-CD2	5.20	124.43	131.20	6	7
A	112	TYR	CA-C-N	5.19	124.69	116.90	6	1
A	299	ARG	CD-NE-CZ	5.16	131.62	124.40	3	1
A	194	LYS	CA-C-N	5.13	126.46	116.20	1	1
A	317	ASN	OD1-CG-ND2	5.12	117.48	122.60	1	1
A	326	LYS	CA-C-N	5.09	126.39	116.20	6	1
A	383	GLN	OE1-CD-NE2	5.08	117.52	122.60	9	2
A	324	TRP	CE2-CD2-CE3	5.03	113.77	118.80	8	1
A	433	GLN	OE1-CD-NE2	4.96	117.64	122.60	6	10
A	243	SER	C-N-CA	4.91	130.53	121.70	2	1
A	293	PRO	CA-N-CD	4.90	105.14	112.00	3	1
A	67	ASN	CA-CB-CG	4.88	117.48	112.60	6	2
A	371	GLN	OE1-CD-NE2	4.88	117.72	122.60	6	6
A	133	GLN	OE1-CD-NE2	4.87	117.73	122.60	8	6
A	19	ARG	NE-CZ-NH2	4.87	123.58	119.20	2	3
A	433	GLN	C-N-CA	4.87	130.46	121.70	5	1
A	7	GLN	OE1-CD-NE2	4.85	117.75	122.60	6	6
A	401	GLN	OE1-CD-NE2	4.81	117.79	122.60	3	10
A	267	ALA	N-CA-CB	4.80	103.20	110.40	10	1
A	384	GLN	OE1-CD-NE2	4.80	117.80	122.60	5	4
A	303	GLN	OE1-CD-NE2	4.78	117.82	122.60	6	10
A	326	LYS	CA-C-O	4.77	112.69	120.80	6	1
A	190	ALA	C-CA-CB	4.76	117.64	110.50	6	1
A	337	GLN	OE1-CD-NE2	4.75	117.85	122.60	3	3
A	232	ARG	NH1-CZ-NH2	4.73	113.15	119.30	5	2
A	187	ASP	CA-CB-CG	4.68	117.28	112.60	5	1
A	17	GLN	OE1-CD-NE2	4.66	117.94	122.60	8	2
A	159	GLU	C-N-CA	4.64	130.04	121.70	9	1
A	320	PRO	CA-N-CD	4.58	105.59	112.00	6	1
A	117	VAL	CA-CB-CG2	4.57	118.16	110.40	1	1
A	416	GLN	OE1-CD-NE2	4.55	118.05	122.60	8	9
A	429	THR	C-N-CA	4.55	129.89	121.70	1	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	12	GLN	OE1-CD-NE2	4.54	118.06	122.60	3	5
A	327	ASN	OD1-CG-ND2	4.54	118.06	122.60	5	4
A	275	ARG	NE-CZ-NH1	4.54	116.96	121.50	2	2
A	299	ARG	NE-CZ-NH2	4.52	123.27	119.20	8	1
A	47	GLN	OE1-CD-NE2	4.51	118.09	122.60	7	4
A	80	GLN	OE1-CD-NE2	4.50	118.10	122.60	8	4
A	67	ASN	OD1-CG-ND2	4.50	118.10	122.60	9	3
A	197	HIS	N-CA-C	4.44	123.43	111.00	1	1
A	194	LYS	O-C-N	4.43	115.91	123.00	1	1
A	211	GLN	OE1-CD-NE2	4.43	118.17	122.60	7	10
A	197	HIS	CB-CG-CD2	4.37	125.52	131.20	7	2
A	428	SER	CA-C-N	4.28	124.76	116.20	5	1
A	192	VAL	CA-CB-CG2	4.25	117.62	110.40	3	1
A	267	ALA	C-CA-CB	4.22	116.83	110.50	10	1
A	195	CYS	C-N-CA	4.21	129.28	121.70	5	1
A	221	ASP	CA-CB-CG	4.18	116.78	112.60	9	1
A	377	HIS	CB-CG-CD2	4.17	125.78	131.20	3	1
A	67	ASN	C-CA-CB	4.17	118.03	110.10	5	1
A	184	HIS	ND1-CG-CD2	4.17	110.27	106.10	2	1
A	167	ASN	CA-CB-CG	4.16	108.44	112.60	10	2
A	356	HIS	CB-CG-CD2	4.15	125.80	131.20	5	4
A	427	GLY	C-N-CA	4.15	129.17	121.70	9	1
A	325	GLN	OE1-CD-NE2	4.13	118.47	122.60	3	3
A	409	ASP	CA-CB-CG	4.11	108.49	112.60	4	1
A	320	PRO	N-CA-CB	4.11	107.52	103.00	4	1
A	387	PHE	CA-CB-CG	4.10	117.90	113.80	3	1
A	25	ARG	NE-CZ-NH2	4.09	115.52	119.20	7	1
A	329	ASP	CA-CB-CG	4.07	116.67	112.60	7	1
A	399	HIS	CB-CG-CD2	4.05	125.93	131.20	6	1
A	40	ASN	OD1-CG-ND2	4.04	118.56	122.60	9	1
A	333	ILE	CA-CB-CG2	4.02	103.66	110.50	6	1
A	73	THR	CA-CB-CG2	4.01	117.31	110.50	2	1
A	266	HIS	CB-CG-CD2	4.01	125.99	131.20	10	1

Too-close contacts ⑦

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	4.24	28
2	5.30	35
3	1.36	9
4	3.03	20
5	1.97	13
6	5.61	37
7	2.42	16
8	5.45	36
9	3.64	24
10	3.94	26

There are 244 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:95:VAL:HG21	A:117:VAL:HG22	0.90	1	2
A:53:MET:HE1	A:347:PHE:CE1	0.85	2	1
A:291:SER:HB3	A:429:THR:HG23	0.82	9	1
A:240:LEU:HD11	A:425:LEU:HD21	0.78	4	1
A:352:LEU:HD13	A:362:MET:HE1	0.78	8	3
A:267:ALA:HB1	A:271:GLN:HB2	0.78	10	1
A:195:CYS:H	A:198:MET:HE2	0.77	8	1
A:267:ALA:HB1	A:271:GLN:CB	0.76	10	1
A:349:PHE:CE2	A:351:THR:HG22	0.73	8	1
A:72:LEU:HD21	A:94:GLN:HB3	0.70	10	3
A:333:ILE:HD13	A:379:VAL:HG21	0.70	7	1
A:161:GLY:HA2	A:190:ALA:HB1	0.70	8	1
A:203:LEU:HD13	A:275:ARG:HH22	0.69	1	1
A:356:HIS:CD2	A:382:VAL:HG11	0.68	1	1
A:105:MET:HE1	A:403:VAL:O	0.67	4	1
A:196:GLY:HA2	A:201:LYS:HE3	0.66	2	1
A:352:LEU:CD1	A:362:MET:HE1	0.66	8	3
A:53:MET:HE1	A:347:PHE:CD1	0.66	2	2
A:362:MET:HA	A:362:MET:HE2	0.66	8	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:352:LEU:HD11	A:365:LEU:CD2	0.66	8	1
A:328:LEU:HD22	A:332:THR:HG21	0.65	6	2
A:199:GLY:HA2	A:424:ALA:HB2	0.65	9	1
A:333:ILE:HG21	A:375:MET:HB2	0.64	2	2
A:413:THR:HG23	A:418:GLY:O	0.64	7	1
A:357:SER:HA	A:362:MET:HE3	0.64	8	1
A:72:LEU:HD21	A:94:GLN:C	0.64	5	2
A:333:ILE:HG21	A:375:MET:CB	0.64	6	2
A:293:PRO:HD2	A:322:PHE:CZ	0.64	3	1
A:349:PHE:CZ	A:351:THR:HG22	0.63	8	1
A:320:PRO:CB	A:378:TYR:CE1	0.62	2	1
A:203:LEU:HD13	A:275:ARG:NH2	0.62	1	1
A:110:SER:HB2	A:162:PHE:CZ	0.61	1	1
A:352:LEU:HD11	A:365:LEU:HD22	0.61	8	1
A:324:TRP:CD1	A:328:LEU:HD12	0.61	3	1
A:245:CYS:SG	A:253:ILE:HD13	0.60	4	2
A:387:PHE:CD1	A:395:THR:HG21	0.60	2	1
A:293:PRO:HD2	A:322:PHE:CE2	0.60	3	1
A:199:GLY:CA	A:424:ALA:HB2	0.60	9	1
A:90:LEU:HD21	A:124:ILE:HD12	0.59	5	1
A:320:PRO:HB2	A:378:TYR:CE1	0.59	2	1
A:355:ILE:HD11	A:359:TRP:CZ2	0.59	7	1
A:320:PRO:HD3	A:350:ILE:HG23	0.59	1	1
A:293:PRO:HD3	A:322:PHE:HB3	0.59	4	2
A:162:PHE:CE2	A:190:ALA:HB1	0.58	6	1
A:114:ALA:HA	A:162:PHE:CE2	0.57	8	1
A:94:GLN:HE21	A:354:GLY:HA2	0.56	10	1
A:291:SER:C	A:322:PHE:CE2	0.56	6	1
A:291:SER:CB	A:429:THR:HG23	0.56	9	1
A:275:ARG:CG	A:275:ARG:HH11	0.56	10	1
A:97:ALA:HB1	A:408:PHE:CE2	0.55	7	1
A:158:ALA:HB1	A:171:LEU:HD21	0.55	6	1
A:72:LEU:HD21	A:94:GLN:CB	0.55	10	2
A:19:ARG:HA	A:54:TRP:CZ2	0.55	6	1
A:105:MET:HB2	A:408:PHE:CZ	0.55	1	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:352:LEU:HD21	A:365:LEU:HD21	0.54	2	1
A:275:ARG:HB3	A:275:ARG:HH11	0.54	10	1
A:105:MET:HE2	A:109:GLN:O	0.54	2	1
A:291:SER:HA	A:322:PHE:CE2	0.53	6	1
A:355:ILE:HG21	A:396:PHE:CE2	0.53	10	1
A:355:ILE:HG23	A:411:VAL:CG2	0.53	2	1
A:328:LEU:HD22	A:332:THR:CG2	0.53	6	2
A:157:ASP:HA	A:184:HIS:CD2	0.53	6	1
A:324:TRP:CZ3	A:333:ILE:HG12	0.53	6	1
A:324:TRP:CZ3	A:375:MET:HG3	0.53	1	1
A:159:GLU:OE2	A:193:LYS:HE2	0.52	2	1
A:355:ILE:HG23	A:359:TRP:HE1	0.52	10	1
A:232:ARG:HD2	A:287:TRP:CZ3	0.52	5	3
A:320:PRO:HG2	A:375:MET:HE2	0.52	8	1
A:109:GLN:HB2	A:160:ALA:HB1	0.52	4	1
A:165:VAL:HG13	A:215:ALA:HB2	0.51	1	1
A:319:SER:HB3	A:322:PHE:CD2	0.51	10	1
A:399:HIS:O	A:403:VAL:HG23	0.51	10	1
A:296:GLU:CD	A:299:ARG:HH12	0.51	1	3
A:355:ILE:HG21	A:396:PHE:CD2	0.51	10	1
A:394:TYR:HB3	A:396:PHE:CE2	0.51	9	1
A:114:ALA:HA	A:162:PHE:CZ	0.51	8	1
A:245:CYS:CB	A:253:ILE:HD11	0.50	1	1
A:337:GLN:HE22	A:350:ILE:HD11	0.50	8	1
A:198:MET:SD	A:404:GLY:HA2	0.50	8	1
A:117:VAL:HB	A:175:MET:HE1	0.50	4	4
A:14:GLU:CD	A:19:ARG:HH21	0.49	2	1
A:357:SER:CB	A:362:MET:HE2	0.49	9	1
A:241:ILE:CD1	A:263:PHE:HB2	0.49	10	1
A:203:LEU:HD11	A:238:ALA:CB	0.49	1	1
A:97:ALA:HB2	A:105:MET:HE3	0.49	8	1
A:54:TRP:CZ2	A:58:HIS:CE1	0.49	6	1
A:275:ARG:CB	A:275:ARG:HH11	0.49	10	1
A:325:GLN:HB2	A:379:VAL:HG11	0.49	1	1
A:293:PRO:HD2	A:322:PHE:CE1	0.49	3	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:196:GLY:HA2	A:201:LYS:CE	0.49	2	1
A:105:MET:CE	A:405:THR:HB	0.49	5	1
A:70:GLY:C	A:94:GLN:HE22	0.49	7	1
A:97:ALA:CB	A:105:MET:HE3	0.49	8	1
A:349:PHE:CE2	A:351:THR:CG2	0.48	8	1
A:357:SER:OG	A:362:MET:HE2	0.48	9	1
A:100:ASN:OD1	A:105:MET:HG2	0.48	4	1
A:355:ILE:HG23	A:411:VAL:HG22	0.48	2	1
A:161:GLY:CA	A:190:ALA:HB1	0.48	8	1
A:120:VAL:HG22	A:123:ARG:HH21	0.48	9	1
A:324:TRP:CE3	A:328:LEU:CD1	0.48	8	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	432	404	23	5
2	432	405	19	8
3	432	408	14	10
4	432	421	10	1
5	432	407	16	9
6	432	407	19	6
7	432	414	14	4
8	432	410	20	2
9	432	406	23	3
10	432	413	17	2

There are 36 unique backbone outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	161	GLY	3
A	323	ASN	3
A	103	ALA	2
A	105	MET	2
A	159	GLU	2
A	160	ALA	2
A	162	PHE	2

Chain	Res	Type	Models (Total)
A	195	CYS	2
A	196	GLY	2
A	244	ASP	2
A	327	ASN	2
A	417	GLY	2
A	41	PRO	1
A	102	ALA	1
A	106	TYR	1
A	107	PRO	1
A	108	ASP	1
A	192	VAL	1
A	193	LYS	1
A	197	HIS	1
A	200	GLY	1
A	242	THR	1
A	245	CYS	1
A	293	PRO	1
A	322	PHE	1
A	325	GLN	1
A	352	LEU	1
A	354	GLY	1
A	355	ILE	1
A	356	HIS	1
A	358	MET	1
A	390	ALA	1
A	418	GLY	1
A	421	SER	1
A	428	SER	1
A	431	GLU	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	343	332	6	5

Model ID	Analysed	Favored	Allowed	Outliers
2	343	335	4	4
3	343	323	4	16
4	343	333	7	3
5	343	331	5	7
6	343	328	10	5
7	343	332	7	4
8	343	333	8	2
9	343	330	9	4
10	343	331	5	7

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

Chain	Res	Type	Models (Total)
A	259	SER	9
A	419	THR	6
A	184	HIS	5
A	429	THR	4
A	192	VAL	3
A	423	THR	3
A	93	TRP	2
A	420	SER	2
A	422	VAL	2
A	72	LEU	1
A	73	THR	1
A	104	SER	1
A	110	SER	1
A	159	GLU	1
A	165	VAL	1
A	197	HIS	1
A	234	ASP	1
A	265	THR	1
A	266	HIS	1
A	275	ARG	1
A	287	TRP	1
A	299	ARG	1
A	321	SER	1

Chain	Res	Type	Models (Total)
A	324	TRP	1
A	352	LEU	1
A	355	ILE	1
A	382	VAL	1
A	408	PHE	1
A	413	THR	1
A	428	SER	1

Fit of model to data used for modeling ?

## Fit of model(s) to crosslinking-MS data

Restraint types

Restraint types are summarized in the table below. 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 2 crosslinking restraints combined in 2 restraint groups.

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
L-Photo-Leucine	LEU	CA	TYR	CA	upper bound	10.0	1
L-Photo-Leucine	LEU	CA	PHE	CA	upper bound	10.0	1

Distograms of individual restraints

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**
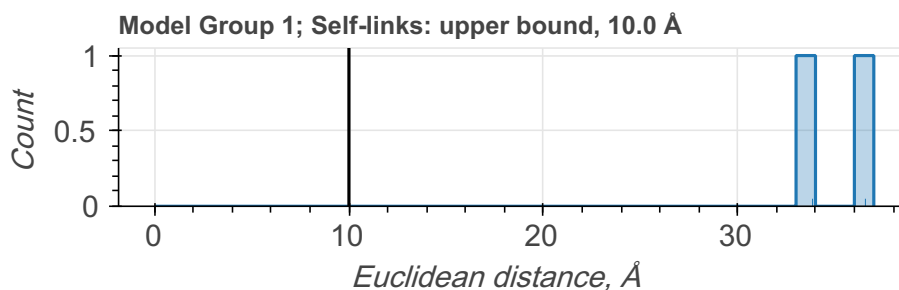
Satisfaction of restraints

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=2)
1	1	1	10/10	All	0.00	100.00	2
				Self-links/ Intramolecular	0.00	100.00	2

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
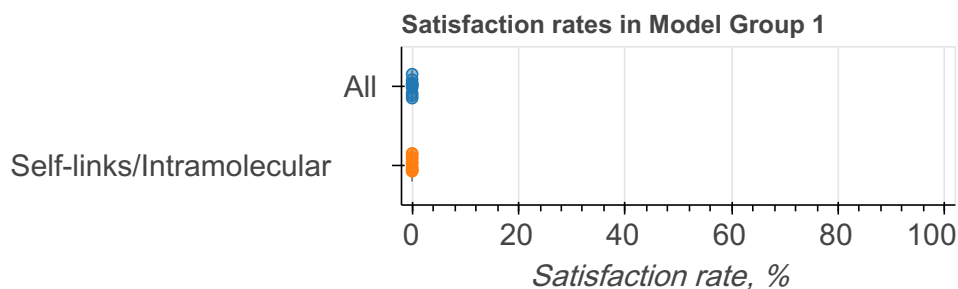
Fit of model to data used for validation ⍰

Validation for this section is under development.