

# Integrative Structure Validation Report ?

February 18, 2025 - 08:35 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	9A2S
PDB-Dev ID	PDBDEV_00000177
Structure Title	Model of E. coli AtpD 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](mailto:helpdesk@pdb-ihm.org)*

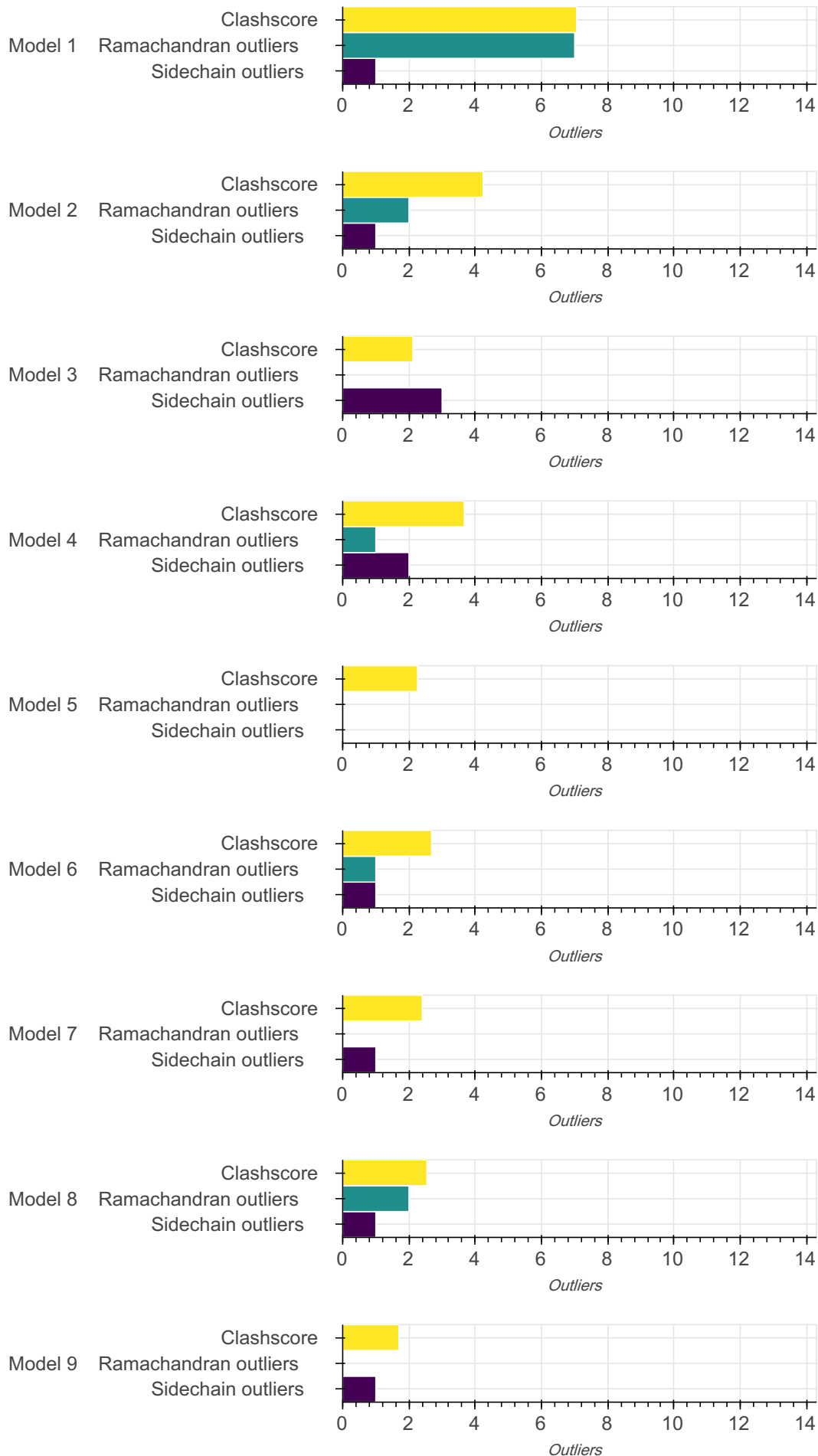
*A user guide is available at [https://pdb-ihm.org/validation\\_help.html](https://pdb-ihm.org/validation_help.html) with specific help available everywhere you see the ? symbol.*

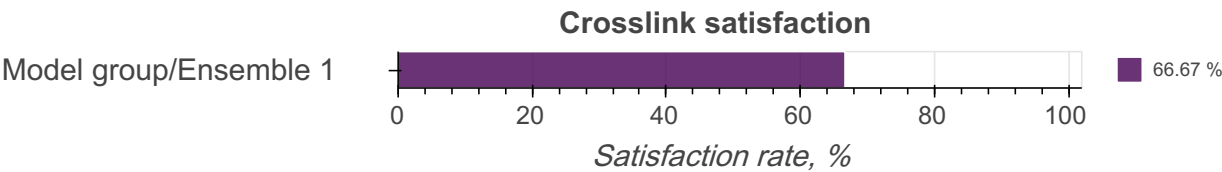
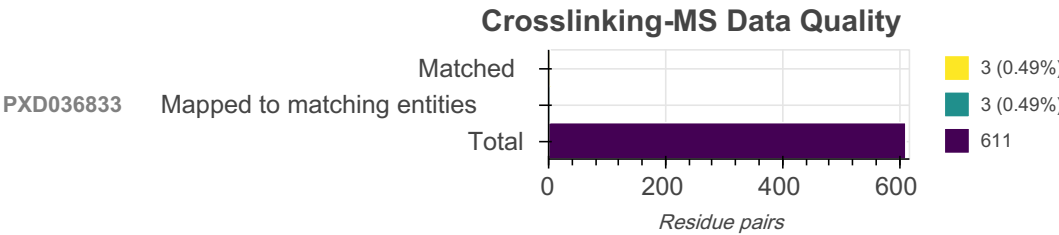
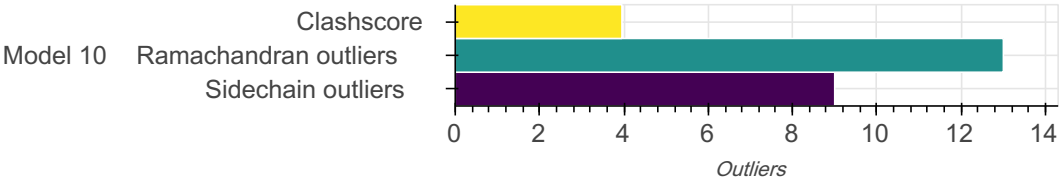
*List of references used to build this report is available [here](#).*

## 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





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	P0ABB4	A	460	-	1-460	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	<a href="#">AlphaLink</a>	1.00	model building	<a href="https://github.com/lhatsk/AlphaLink">https://github.com/lhatsk/AlphaLink</a>

## 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	P0ABB4	dbseq_P0ABB4_target	0.00	True

Residue pairs stats:

Source	Total	In matched entities	Total matched
9A2S	3	3 (100.00%)	3 (100.00%)
PXD036833	611	3 (0.49%)	3 (0.49%)

### 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 1 bond length outliers in this entry (0.00% of 35920 assessed bonds). A summary is provided below.

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	219	ARG	CZ-NH2	4.05	1.28	1.33	1	1

#### Standard geometry: angle outliers ?

There are 179 bond angle outliers in this entry (0.37% of 48610 assessed bonds). A summary is provided below.

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	331	ILE	C-N-CA	9.49	138.78	121.70	10	1
A	344	GLN	OE1-CD-NE2	6.83	115.77	122.60	6	8
A	331	ILE	O-C-N	6.44	112.69	123.00	10	1
A	328	SER	C-N-CA	6.44	133.29	121.70	10	1
A	8	GLN	OE1-CD-NE2	6.32	116.28	122.60	4	10
A	369	GLN	OE1-CD-NE2	6.25	116.35	122.60	9	6
A	243	ASP	C-N-CA	6.18	132.82	121.70	2	4
A	209	GLN	OE1-CD-NE2	6.13	116.47	122.60	9	9
A	403	GLN	OE1-CD-NE2	6.01	116.59	122.60	10	2
A	328	SER	O-C-N	5.75	113.79	123.00	10	1
A	325	GLN	OE1-CD-NE2	5.67	116.93	122.60	3	8
A	189	ASP	CA-CB-CG	5.60	118.20	112.60	1	6
A	405	PHE	CA-CB-CG	5.54	119.34	113.80	1	1
A	366	GLN	OE1-CD-NE2	5.45	117.15	122.60	10	3
A	332	TYR	CA-C-N	5.43	125.04	116.90	4	1
A	331	ILE	CA-C-N	5.40	127.01	116.20	10	1
A	269	GLN	OE1-CD-NE2	5.33	117.27	122.60	10	4

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	44	GLN	OE1-CD-NE2	5.23	117.37	122.60	3	1
A	135	ASP	C-N-CA	5.13	130.93	121.70	10	1
A	137	MET	C-N-CA	5.09	112.53	121.70	10	1
A	280	GLN	OE1-CD-NE2	5.06	117.54	122.60	2	2
A	124	GLN	OE1-CD-NE2	5.05	117.55	122.60	8	8
A	45	GLN	OE1-CD-NE2	5.03	117.57	122.60	10	7
A	243	ASP	CA-C-N	4.90	126.00	116.20	2	1
A	330	GLY	C-N-CA	4.87	130.46	121.70	10	1
A	308	SER	CA-C-N	4.86	124.18	116.90	2	1
A	139	PRO	N-CA-CB	4.85	108.34	103.00	1	2
A	20	GLN	OE1-CD-NE2	4.77	117.83	122.60	10	9
A	171	HIS	CB-CG-CD2	4.67	125.12	131.20	10	1
A	362	GLN	OE1-CD-NE2	4.64	117.96	122.60	10	1
A	408	ALA	N-CA-CB	4.61	103.49	110.40	10	1
A	301	ALA	N-CA-CB	4.59	103.52	110.40	3	1
A	217	ARG	NE-CZ-NH2	4.58	123.32	119.20	2	1
A	1	MET	CA-CB-CG	4.56	104.99	114.10	10	6
A	331	ILE	CA-CB-CG1	4.51	118.07	110.40	10	1
A	244	ASN	OD1-CG-ND2	4.50	118.10	122.60	10	2
A	217	ARG	NE-CZ-NH1	4.48	125.98	121.50	5	6
A	313	PHE	CA-CB-CG	4.46	118.26	113.80	4	1
A	12	ALA	C-CA-CB	4.46	103.81	110.50	2	1
A	111	HIS	CB-CG-CD2	4.43	125.44	131.20	3	9
A	154	VAL	CA-CB-CG2	4.34	117.78	110.40	2	1
A	315	HIS	CB-CG-CD2	4.32	125.58	131.20	2	2
A	426	ARG	NE-CZ-NH2	4.30	123.07	119.20	4	1
A	243	ASP	O-C-N	4.29	116.13	123.00	5	5
A	398	GLN	OE1-CD-NE2	4.26	118.34	122.60	9	2
A	140	PHE	CA-CB-CG	4.25	118.05	113.80	4	1
A	269	GLN	CA-C-N	4.24	123.26	116.90	5	1
A	438	HIS	CB-CG-CD2	4.24	125.69	131.20	9	10
A	442	GLN	OE1-CD-NE2	4.20	118.40	122.60	2	3
A	194	MET	CG-SD-CE	4.20	91.67	100.90	6	1
A	327	ALA	C-N-CA	4.17	129.21	121.70	10	1
A	334	ALA	C-N-CA	4.15	129.17	121.70	8	1

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	188	ASN	OD1-CG-ND2	4.15	118.45	122.60	1	3
A	213	PRO	N-CA-CB	4.12	107.53	103.00	8	1
A	333	PRO	CA-N-CD	4.12	106.24	112.00	2	1
A	305	THR	CA-CB-OG1	4.11	115.77	109.60	2	1
A	159	ASN	OD1-CG-ND2	4.11	118.49	122.60	1	1
A	408	ALA	C-CA-CB	4.10	116.66	110.50	10	1
A	265	ALA	C-N-CA	4.09	129.06	121.70	1	1
A	217	ARG	CD-NE-CZ	4.09	130.12	124.40	3	1
A	402	SER	CA-CB-OG	4.06	119.21	111.10	4	1
A	261	ARG	NH1-CZ-NH2	4.05	114.03	119.30	10	1
A	334	ALA	C-CA-CB	4.05	116.57	110.50	2	1
A	122	ASN	OD1-CG-ND2	4.05	118.55	122.60	1	2
A	328	SER	CA-C-N	4.03	124.26	116.20	10	1
A	233	ASP	CA-CB-CG	4.02	116.62	112.60	2	1
A	287	LYS	CG-CD-CE	4.01	102.08	111.30	10	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	7.06	50
2	4.24	30
3	2.12	15
4	3.67	26
5	2.26	16
6	2.68	19
7	2.40	17
8	2.54	18
9	1.69	12
10	3.95	28

There are 231 clashes. The table below contains the detailed list of all clashes based on a MolProbity analysis. Bad clashes are  $\geq 0.4$  Angstrom. The output is limited to 100 rows.

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:250:LEU:HD13	A:334:ALA:HB1	0.92	1	1
A:253:THR:HG23	A:269:GLN:HE22	0.91	9	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:279:LEU:HD22	A:313:PHE:CZ	0.91	1	1
A:167:ILE:HD12	A:239:LEU:HD11	0.87	7	1
A:136:LEU:HD22	A:146:VAL:HG13	0.86	10	1
A:218:LEU:HD13	A:276:MET:CE	0.82	1	1
A:137:MET:HE1	A:362:GLN:HG3	0.76	5	3
A:408:ALA:HB1	A:412:THR:OG1	0.75	10	1
A:88:ILE:HG23	A:207:TYR:CE1	0.74	8	1
A:83:ALA:HB3	A:104:GLU:HG3	0.73	6	1
A:338:LEU:HD22	A:366:GLN:HE21	0.72	1	1
A:221:ALA:CB	A:245:ILE:HD11	0.70	2	1
A:137:MET:HE2	A:167:ILE:HD11	0.69	8	1
A:132:LYS:HE3	A:419:VAL:HG11	0.69	1	1
A:218:LEU:HD13	A:276:MET:HE1	0.67	1	1
A:167:ILE:CD1	A:239:LEU:HD11	0.66	7	1
A:333:PRO:CG	A:402:SER:HB3	0.66	4	1
A:167:ILE:HD12	A:239:LEU:HD22	0.66	2	1
A:190:PHE:CZ	A:194:MET:HE3	0.64	6	1
A:279:LEU:HD22	A:313:PHE:CE2	0.64	1	1
A:8:GLN:NE2	A:10:ILE:HD11	0.64	4	7
A:246:TYR:CD1	A:276:MET:HE1	0.64	4	1
A:137:MET:HE1	A:362:GLN:CG	0.62	6	2
A:322:LEU:HD23	A:334:ALA:O	0.62	2	1
A:320:VAL:HG22	A:337:PRO:HB2	0.61	1	1
A:333:PRO:HD3	A:405:PHE:CZ	0.61	6	2
A:248:TYR:CE2	A:272:LEU:HD23	0.61	4	1
A:333:PRO:HD3	A:405:PHE:CE1	0.61	3	1
A:118:GLU:OE1	A:287:LYS:HE3	0.60	10	1
A:217:ARG:HB3	A:248:TYR:CE1	0.60	8	1
A:127:LEU:HD11	A:167:ILE:CD1	0.60	3	1
A:328:SER:HA	A:329:LEU:HG	0.60	10	1
A:177:PHE:HB3	A:205:LEU:HD23	0.60	9	2
A:215:GLY:HA3	A:268:TYR:CE1	0.58	4	1
A:127:LEU:HD11	A:167:ILE:HD11	0.58	3	1
A:160:MET:HE3	A:241:PHE:CD1	0.58	10	2
A:297:VAL:HG21	A:312:THR:CG2	0.57	1	1



Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:149:PHE:CE2	A:319:THR:HG23	0.57	2	1
A:83:ALA:HB3	A:104:GLU:CG	0.57	6	1
A:221:ALA:HB2	A:245:ILE:HD11	0.57	2	1
A:215:GLY:HA3	A:268:TYR:CZ	0.56	4	1
A:331:ILE:HG22	A:332:TYR:CG	0.56	10	1
A:249:THR:HG23	A:272:LEU:HD11	0.56	7	1
A:320:VAL:HA	A:337:PRO:HB3	0.56	1	1
A:180:VAL:HG23	A:210:MET:CE	0.55	4	1
A:329:LEU:HD23	A:331:ILE:HD11	0.55	10	1
A:248:TYR:CD1	A:308:SER:HB2	0.55	1	1
A:136:LEU:CD2	A:146:VAL:HG13	0.55	10	1
A:167:ILE:HD12	A:239:LEU:HG	0.55	10	1
A:297:VAL:HG21	A:312:THR:HG21	0.54	1	2
A:324:ARG:NH2	A:365:LEU:HD13	0.54	1	1
A:219:ARG:HH12	A:269:GLN:NE2	0.54	2	1
A:357:THR:HA	A:432:MET:HE1	0.54	2	4
A:126:LEU:HD11	A:350:VAL:HG12	0.54	3	4
A:324:ARG:HH22	A:365:LEU:CD1	0.54	1	1
A:131:ILE:HG21	A:134:ILE:HD12	0.54	1	1
A:177:PHE:CE2	A:241:PHE:CD1	0.54	1	1
A:149:PHE:CE2	A:313:PHE:CZ	0.53	5	1
A:157:THR:HG22	A:161:MET:HE3	0.53	7	2
A:215:GLY:HA3	A:269:GLN:HE22	0.53	1	1
A:8:GLN:HE21	A:10:ILE:HD11	0.53	1	3
A:408:ALA:HB1	A:412:THR:HG1	0.53	10	1
A:217:ARG:NH1	A:248:TYR:CE1	0.53	2	1
A:190:PHE:CE1	A:205:LEU:HD21	0.53	9	1
A:338:LEU:HD22	A:366:GLN:NE2	0.52	1	1
A:320:VAL:HG13	A:335:VAL:HG22	0.52	2	1
A:246:TYR:OH	A:313:PHE:CE2	0.52	1	1
A:324:ARG:HH22	A:365:LEU:HD13	0.52	1	1
A:246:TYR:CE1	A:276:MET:HE1	0.52	4	1
A:12:ALA:HB2	A:213:PRO:CB	0.52	2	1
A:131:ILE:HD13	A:134:ILE:HD12	0.52	1	1
A:160:MET:HE1	A:296:ALA:HB2	0.52	6	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:221:ALA:HB2	A:248:TYR:CZ	0.51	7	1
A:246:TYR:CE1	A:308:SER:HB2	0.51	9	1
A:249:THR:HG23	A:272:LEU:HD21	0.51	6	1
A:244:ASN:ND2	A:298:TYR:H	0.51	10	2
A:333:PRO:CD	A:405:PHE:CE1	0.51	3	1
A:134:ILE:H	A:134:ILE:HD12	0.51	10	1
A:12:ALA:HB2	A:213:PRO:HB3	0.50	2	1
A:239:LEU:HD23	A:241:PHE:CZ	0.50	2	1
A:308:SER:HB2	A:309:PRO:HD3	0.50	5	4
A:297:VAL:CG2	A:312:THR:HG21	0.50	10	1
A:333:PRO:CD	A:405:PHE:CZ	0.50	3	2
A:332:TYR:CZ	A:412:THR:HG22	0.50	4	1
A:149:PHE:CE1	A:319:THR:HG23	0.50	5	1
A:405:PHE:HB2	A:408:ALA:HB3	0.50	6	2
A:88:ILE:HG23	A:207:TYR:CD1	0.49	8	1
A:225:LEU:HD22	A:279:LEU:HD12	0.49	7	1
A:133:VAL:HG21	A:335:VAL:HB	0.49	4	1
A:160:MET:CE	A:296:ALA:HB2	0.49	6	1
A:279:LEU:C	A:279:LEU:HD23	0.49	1	1
A:239:LEU:CD2	A:241:PHE:CZ	0.48	2	1
A:215:GLY:HA3	A:269:GLN:OE1	0.48	1	1
A:320:VAL:HG22	A:337:PRO:CB	0.48	1	1
A:249:THR:HA	A:272:LEU:HD11	0.48	10	1
A:164:ILE:CD1	A:190:PHE:CZ	0.48	8	3
A:248:TYR:CD1	A:308:SER:HB3	0.47	4	1
A:12:ALA:HB3	A:213:PRO:HG3	0.47	8	1
A:133:VAL:HG13	A:136:LEU:CD1	0.47	1	1
A:179:GLY:O	A:220:VAL:HG11	0.47	2	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	458	430	21	7
2	458	443	13	2
3	458	447	11	0

Model ID	Analysed	Favored	Allowed	Outliers
4	458	444	13	1
5	458	446	12	0
6	458	450	7	1
7	458	450	8	0
8	458	436	20	2
9	458	448	10	0
10	458	414	31	13

There are 22 unique backbone outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	133	VAL	2
A	140	PHE	2
A	184	THR	2
A	263	PRO	2
A	104	GLU	1
A	119	GLU	1
A	126	LEU	1
A	128	GLU	1
A	135	ASP	1
A	136	LEU	1
A	141	ALA	1
A	144	GLY	1
A	183	ARG	1
A	185	ARG	1
A	248	TYR	1
A	268	TYR	1
A	318	ALA	1
A	331	ILE	1
A	333	PRO	1
A	342	SER	1
A	405	PHE	1
A	407	VAL	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	381	366	14	1
2	381	376	4	1
3	381	375	3	3
4	381	376	3	2
5	381	376	5	0
6	381	378	2	1
7	381	375	5	1
8	381	373	7	1
9	381	376	4	1
10	381	363	9	9

There are 18 unique sidechain outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	266	VAL	2
A	319	THR	2
A	126	LEU	1
A	137	MET	1
A	138	CYS	1
A	154	VAL	1
A	167	ILE	1
A	182	GLU	1
A	184	THR	1
A	249	THR	1
A	253	THR	1
A	256	SER	1
A	271	THR	1
A	272	LEU	1
A	311	THR	1
A	322	LEU	1
A	329	LEU	1
A	331	ILE	1

Fit of model to data used for modeling ?

Fit of model(s) to crosslinking-MS data

## Restraint types

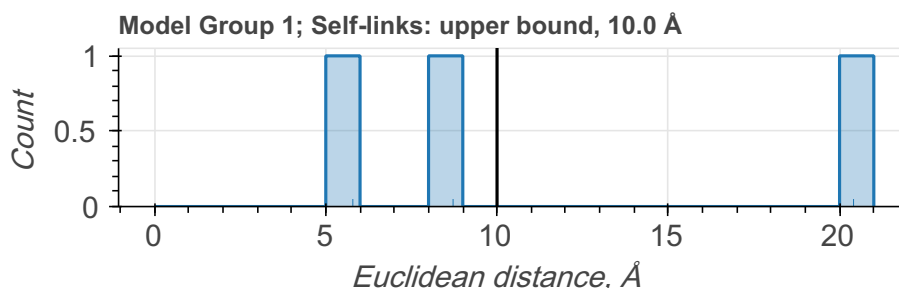
Restraint types are summarized in the table below. Restraints assigned "*by-residue*" are interpreted as between CA atoms. Restraints between coarse-grained beads are indicated as "*coarse-grained*". *Restraint group* represents a set of crosslinking restraints applied collectively in the modeling.

There are 3 crosslinking restraints combined in 3 restraint groups.

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



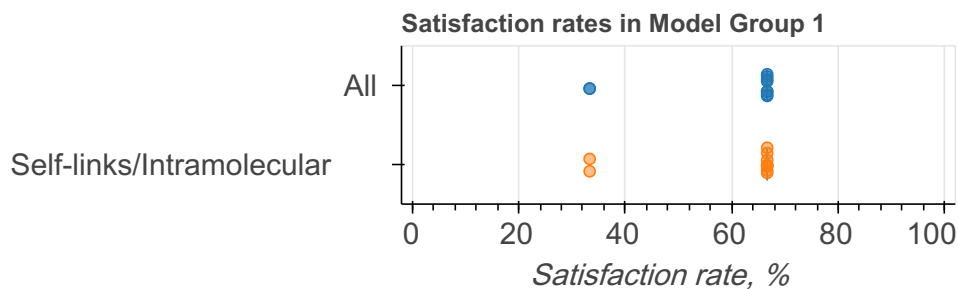
## Satisfaction of restraints

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	10/10	All	66.67	33.33	3
				Self-links/ Intramolecular	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.



### Fit of model to data used for validation ?

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 Trehwella, 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.*