Integrative Structure Validation Report

October 09, 2025 - 04:40 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
ATSAS Version 3.2.1 (r14885)

PDB ID	9A1T | pdb_00009a1t
PDB-Dev ID	PDBDEV_00000119
Structure Title	SARS-CoV-2 nsp7-8 polyprotein
Structure Authors	Ruchi Yadav; Valentine V. Courouble; Sanjay K. Dey; Francesc X. Ruiz; Patrick R. Griffin; Eddy Arnold
Deposited on	2022-05-20

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 5 dataset(s) were used to build this entry.

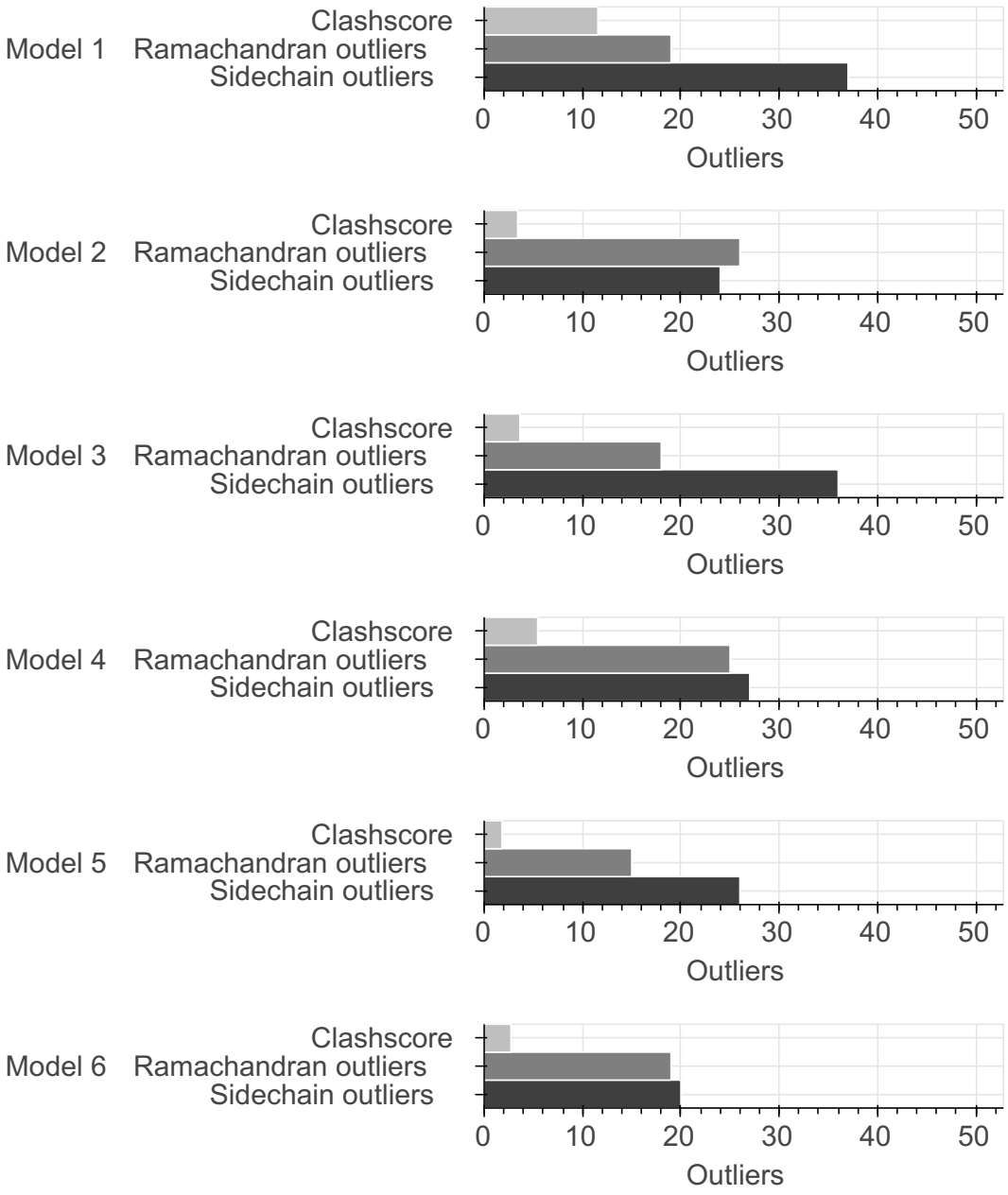
Name	Type	Count
Crosslinking-MS data	Experimental data	1
H/D exchange data	Experimental data	1

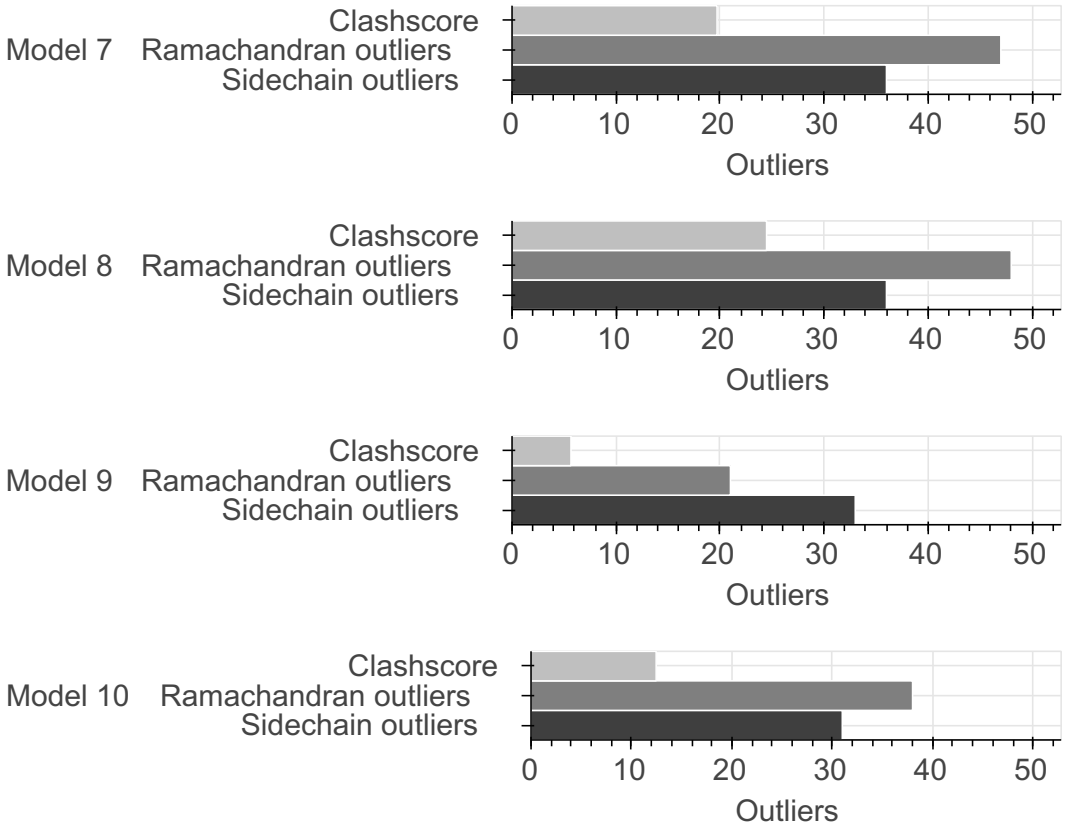
Name	Type	Count
SAS data	Experimental data	1
De Novo model	Starting model	1
Experimental model	Starting model	1

1.2. Overall quality

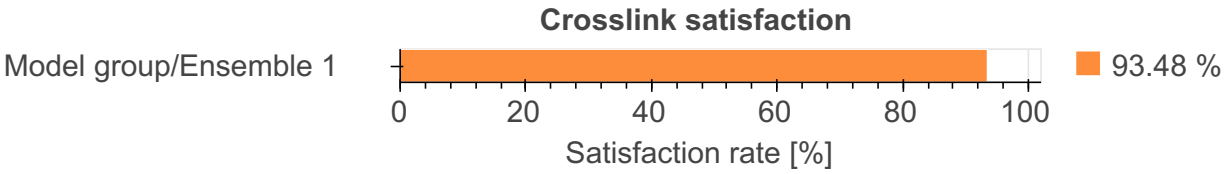
This validation report contains model quality assessments for all structures, data quality and fit to model assessments for SAS and crosslinking-MS datasets. Data quality and fit to model assessments for other datasets and model uncertainty are under development. Number of plots is limited to 256.

Model Quality: MolProbity Analysis





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	1-10	1	SARS-CoV-2 nsp7-8 polyprotein	A	283	-	1-2, 3-82, 83-160, 161-277, 278-283	100.00 / 97.17	Atomic

2.3. Datasets used for modeling

There are 5 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	Crosslinking-MS data	PRIDE	PXD033748
2	H/D exchange data	PRIDE	PXD033698
3	SAS data	SASBDB	SASDPY2
4	Experimental model	PDB	pdb_00006yhu
5	De Novo model	Not available	Not available

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	Integrative modeling	Not available	Not available	Not available	False	False

There is 1 software package reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	I-TASSER	Not available	Integrative modeling	https://zhanggroup.org/I-TASSER/

3. Data quality

3.1.1. Scattering profile

SAS data used in this integrative model could not be validated as the sasCIF file is currently unavailable or incomplete.

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).

Crosslinking-MS dataset is not available in the PRIDE Crosslinking database.

3.4. H/D exchange

Validation for this section is under development.

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 21 bond length outliers in this entry (0.10% of 22050 assessed bonds). A summary is provided below.

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	234	TYR	CE1-CZ	48.48	2.54	1.38	1	1
A	232	PHE	CE2-CZ	38.98	2.55	1.38	1	1
A	232	PHE	CG-CD2	16.16	1.72	1.38	1	1
A	234	TYR	CE2-CZ	15.50	1.75	1.38	1	1
A	234	TYR	CG-CD2	15.49	1.71	1.39	1	1
A	232	PHE	CG-CD1	14.48	1.69	1.38	1	1
A	232	PHE	CD2-CE2	11.85	1.74	1.38	1	1
A	234	TYR	CG-CD1	11.84	1.64	1.39	1	1
A	234	TYR	CD2-CE2	11.31	1.72	1.38	1	1
A	232	PHE	CE1-CZ	10.13	1.69	1.38	1	1
A	232	PHE	CD1-CE1	9.57	1.67	1.38	1	1
A	234	TYR	CD1-CE1	8.55	1.64	1.38	1	1
A	267	TRP	C-N	5.51	1.43	1.34	1	1
A	250	LYS	CA-C	4.78	1.63	1.52	7	1
A	205	ILE	N-CA	4.59	1.55	1.46	3	1
A	268	PRO	N-CA	4.38	1.53	1.47	1	1
A	200	VAL	CA-C	4.16	1.61	1.52	10	1
A	268	PRO	CA-C	4.14	1.61	1.52	8	1
A	204	ILE	C-N	4.13	1.39	1.33	3	1
A	232	PHE	CA-CB	4.07	1.61	1.53	1	1
A	218	PRO	N-CA	4.00	1.53	1.47	7	1

Standard geometry: angle outliers

There are 560 bond angle outliers in this entry (1.88% of 29830 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	234	TYR	CB-CG-CD1	12.74	101.69	120.80	1	1
A	232	PHE	C-CA-CB	11.89	132.69	110.10	1	1
A	204	ILE	C-N-CA	11.16	141.78	121.70	2	4
A	250	LYS	C-N-CA	10.91	141.34	121.70	7	2
A	246	ASP	CA-CB-CG	10.53	123.13	112.60	3	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	219	ASP	CA-CB-CG	10.16	122.76	112.60	10	3
A	200	VAL	CA-C-N	10.11	132.07	116.90	10	1
A	216	VAL	C-N-CA	9.55	138.89	121.70	7	1
A	232	PHE	CB-CG-CD1	9.26	104.97	120.70	1	1
A	100	PHE	CA-CB-CG	9.20	123.00	113.80	3	3
A	91	PHE	CA-CB-CG	9.16	122.96	113.80	9	4
A	267	TRP	CA-C-N	8.97	130.36	116.90	1	4
A	132	SER	C-CA-CB	8.89	126.98	110.10	8	1
A	232	PHE	CA-CB-CG	8.88	122.68	113.80	4	5
A	209	THR	C-N-CA	8.70	137.36	121.70	10	2
A	155	MET	C-CA-CB	8.47	126.20	110.10	7	1
A	151	ALA	C-CA-CB	8.45	123.18	110.50	7	1
A	241	ILE	C-CA-CB	8.26	128.11	111.60	7	1
A	204	ILE	CA-C-N	8.20	132.61	116.20	3	2
A	250	LYS	CA-C-N	8.18	132.56	116.20	7	1
A	234	TYR	N-CA-CB	8.12	124.30	110.50	1	1
A	242	GLN	C-N-CA	8.07	136.22	121.70	8	1
A	127	LEU	C-CA-CB	7.93	125.17	110.10	8	1
A	79	ASP	CA-CB-CG	7.85	120.45	112.60	6	2
A	125	LYS	C-CA-CB	7.75	124.83	110.10	8	1
A	155	MET	CA-CB-CG	7.73	129.56	114.10	7	1
A	213	LEU	C-CA-CB	7.60	124.54	110.10	7	3
A	115	ASP	CA-CB-CG	7.56	120.16	112.60	10	3
A	216	VAL	CA-C-N	7.51	131.23	116.20	7	1
A	149	ASP	CA-CB-CG	7.44	120.04	112.60	7	4
A	218	PRO	C-N-CA	7.35	134.94	121.70	4	2
A	205	ILE	CA-C-N	7.31	127.86	116.90	7	2
A	135	ASP	CA-CB-CG	7.22	119.82	112.60	8	3
A	208	THR	CA-CB-CG2	7.21	122.75	110.50	7	2
A	173	GLN	OE1-CD-NE2	7.20	115.40	122.60	3	6
A	164	LYS	C-N-CA	7.08	134.45	121.70	7	2
A	245	VAL	CA-CB-CG2	7.05	122.38	110.40	7	1
A	273	ALA	C-CA-CB	6.98	100.03	110.50	8	1
A	282	LEU	C-N-CA	6.92	134.15	121.70	8	1
A	202	LEU	N-CA-CB	6.87	122.18	110.50	7	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	77	MET	C-N-CA	6.82	133.97	121.70	2	1
A	231	THR	C-N-CA	6.80	133.95	121.70	7	1
A	260	ASP	C-N-CA	6.80	133.95	121.70	8	1
A	225	ASN	C-N-CA	6.80	133.94	121.70	7	1
A	203	ASN	CA-CB-CG	6.78	119.38	112.60	10	3
A	248	ASP	CA-CB-CG	6.77	105.83	112.60	9	1
A	201	PRO	CA-N-CD	6.75	102.55	112.00	10	1
A	199	CYS	C-N-CA	6.70	133.75	121.70	7	1
A	250	LYS	O-C-N	6.69	112.30	123.00	7	1
A	137	ASP	CA-CB-CG	6.67	119.27	112.60	6	1
A	256	GLU	C-N-CA	6.63	133.64	121.70	10	1
A	251	ILE	O-C-N	6.56	112.50	123.00	1	1
A	234	TYR	CD1-CG-CD2	6.55	127.93	118.10	1	1
A	268	PRO	N-CA-C	6.49	128.34	112.10	7	1
A	93	SER	N-CA-CB	6.41	121.39	110.50	8	1
A	7	ASP	CA-CB-CG	6.41	119.01	112.60	7	5
A	125	LYS	C-N-CA	6.36	133.15	121.70	8	1
A	205	ILE	C-CA-CB	6.35	124.31	111.60	2	1
A	271	VAL	C-CA-CB	6.33	123.44	111.40	10	1
A	126	SER	N-CA-CB	6.33	121.27	110.50	8	2
A	178	THR	CA-CB-CG2	6.31	121.22	110.50	8	1
A	173	GLN	C-CA-CB	6.30	122.08	110.10	10	2
A	267	TRP	C-CA-CB	6.22	121.92	110.10	7	1
A	214	MET	C-N-CA	6.19	132.85	121.70	10	2
A	142	ARG	C-CA-CB	6.19	121.86	110.10	4	2
A	218	PRO	CA-N-CD	6.18	103.34	112.00	7	2
A	95	PRO	CA-N-CD	6.18	103.35	112.00	8	1
A	204	ILE	O-C-N	6.17	113.12	123.00	3	2
A	184	ASP	CA-CB-CG	6.16	118.76	112.60	7	3
A	197	ASP	CA-CB-CG	6.15	118.75	112.60	1	4
A	79	ASP	C-CA-CB	6.13	121.74	110.10	8	3
A	84	LEU	N-CA-CB	6.13	120.91	110.50	1	1
A	157	LYS	C-N-CA	6.12	132.72	121.70	1	1
A	251	ILE	C-CA-CB	6.10	123.79	111.60	1	2
A	137	ASP	C-N-CA	6.08	132.64	121.70	8	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	55	VAL	CA-CB-CG1	6.08	120.73	110.40	1	1
A	168	VAL	C-CA-CB	6.06	99.89	111.40	10	1
A	173	GLN	CB-CG-CD	6.01	122.82	112.60	4	1
A	104	GLN	C-CA-CB	6.00	121.51	110.10	3	1
A	206	PRO	N-CA-C	5.97	127.03	112.10	10	1
A	114	GLY	C-N-CA	5.96	132.42	121.70	10	1
A	15	LEU	C-CA-CB	5.95	121.40	110.10	5	1
A	268	PRO	O-C-N	5.95	113.49	123.00	8	1
A	143	LYS	CA-CB-CG	5.93	125.95	114.10	8	1
A	234	TYR	CE1-CZ-OH	5.91	137.62	119.90	1	1
A	158	GLN	C-N-CA	5.89	132.31	121.70	8	2
A	262	SER	C-CA-CB	5.89	121.29	110.10	7	1
A	257	ILE	CA-CB-CG2	5.87	120.48	110.50	8	1
A	207	LEU	C-CA-CB	5.86	121.22	110.10	10	1
A	81	ARG	CD-NE-CZ	5.83	132.56	124.40	6	3
A	243	GLN	N-CA-C	5.82	127.29	111.00	8	1
A	216	VAL	O-C-N	5.81	113.71	123.00	7	1
A	263	PRO	N-CA-C	5.80	126.61	112.10	7	1
A	92	SER	N-CA-CB	5.80	100.64	110.50	7	1
A	202	LEU	N-CA-C	5.80	94.77	111.00	7	1
A	241	ILE	CG1-CB-CG2	5.76	127.98	110.70	10	1
A	109	GLN	OE1-CD-NE2	5.76	116.84	122.60	10	4
A	252	VAL	CA-CB-CG2	5.75	120.17	110.40	7	2
A	268	PRO	CA-C-N	5.75	127.69	116.20	8	1
A	274	LEU	C-CA-CB	5.74	121.01	110.10	7	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	11.56	51
2	3.40	15
3	3.63	16
4	5.44	24
5	1.81	8
6	2.72	12

Model ID	Clash score	Number of clashes
7	19.71	87
8	24.47	108
9	5.67	25
10	12.46	55

There are 401 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:234:TYR:CE2	A:234:TYR:CZ	1.66	1	1
A:232:PHE:CE2	A:234:TYR:CE1	1.37	1	1
A:232:PHE:CZ	A:234:TYR:CE1	1.36	1	1
A:232:PHE:CZ	A:234:TYR:CZ	1.36	1	1
A:232:PHE:CE2	A:234:TYR:CZ	1.36	1	1
A:232:PHE:CZ	A:234:TYR:CE2	1.19	1	1
A:232:PHE:CD2	A:234:TYR:CE1	1.19	1	1
A:232:PHE:CG	A:234:TYR:CE1	1.19	1	1
A:232:PHE:CE2	A:234:TYR:CD1	1.19	1	1
A:232:PHE:CE2	A:234:TYR:CE2	1.18	1	1
A:232:PHE:CD2	A:234:TYR:CZ	1.17	1	1
A:232:PHE:CE1	A:234:TYR:CZ	1.16	1	1
A:232:PHE:CZ	A:234:TYR:CD2	1.16	1	1
A:232:PHE:CE1	A:234:TYR:CE1	1.15	1	1
A:232:PHE:CZ	A:234:TYR:CG	1.15	1	1
A:232:PHE:CZ	A:234:TYR:CD1	1.14	1	1
A:232:PHE:CD1	A:234:TYR:CE1	1.14	1	1
A:232:PHE:CD1	A:234:TYR:CZ	1.12	1	1
A:232:PHE:CE2	A:234:TYR:CD2	1.12	1	1
A:232:PHE:CE2	A:234:TYR:CG	1.10	1	1
A:232:PHE:CG	A:234:TYR:CZ	1.06	1	1
A:232:PHE:CE2	A:232:PHE:CZ	0.95	1	1
A:234:TYR:CE1	A:234:TYR:CZ	0.94	1	1
A:252:VAL:H	A:270:ILE:HA	0.90	1	1
A:199:CYS:HA	A:273:ALA:HB3	0.88	8	1
A:200:VAL:HA	A:216:VAL:HG21	0.86	7	1
A:239:TRP:H	A:271:VAL:HG12	0.86	10	1
A:213:LEU:HD22	A:241:ILE:HG23	0.78	10	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:257:ILE:HG21	A:269:LEU:C	0.78	7	1
A:54:MET:HA	A:54:MET:HE2	0.78	4	1
A:217:ILE:HD12	A:218:PRO:HD2	0.77	10	1
A:58:LEU:HG	A:84:LEU:HD12	0.76	1	1
A:54:MET:SD	A:77:MET:HE2	0.76	8	1
A:148:ALA:HA	A:276:ALA:HB2	0.74	7	1
A:207:LEU:HD13	A:211:ALA:HB2	0.74	7	1
A:262:SER:HA	A:265:LEU:HB3	0.71	8	1
A:244:VAL:HG21	A:249:SER:C	0.71	8	1
A:120:LEU:HA	A:124:LYS:H	0.71	8	1
A:254:LEU:HA	A:257:ILE:HB	0.71	8	1
A:213:LEU:HD21	A:272:THR:HA	0.70	10	1
A:205:ILE:HA	A:208:THR:HB	0.70	8	1
A:254:LEU:H	A:254:LEU:HD13	0.70	5	1
A:213:LEU:HD13	A:241:ILE:HG12	0.69	10	1
A:239:TRP:O	A:273:ALA:HB3	0.68	1	1
A:252:VAL:HG23	A:253:GLN:HA	0.67	8	1
A:59:SER:HA	A:87:ILE:HG23	0.67	1	1
A:143:LYS:HE2	A:180:LEU:HB3	0.67	8	1
A:142:ARG:HA	A:177:PHE:CD2	0.65	8	1
A:207:LEU:H	A:274:LEU:HD23	0.65	10	1
A:9:LYS:HA	A:41:ILE:HG22	0.65	4	1
A:213:LEU:HD13	A:241:ILE:CG1	0.65	10	1
A:174:THR:O	A:178:THR:HG23	0.65	8	1
A:244:VAL:HG22	A:245:VAL:N	0.65	8	1
A:19:LEU:HD13	A:34:CYS:SG	0.65	5	1
A:152:MET:HA	A:155:MET:HG3	0.65	7	1
A:110:ALA:HA	A:113:ASN:HB3	0.64	8	1
A:201:PRO:HB3	A:273:ALA:HA	0.63	8	1
A:92:SER:HB3	A:98:ALA:HB2	0.63	7	1
A:149:ASP:HB2	A:273:ALA:HB1	0.63	7	1
A:199:CYS:HA	A:273:ALA:CB	0.63	8	1
A:168:VAL:H	A:217:ILE:HB	0.63	10	1
A:234:TYR:CD1	A:234:TYR:N	0.62	1	1
A:240:GLU:C	A:257:ILE:HG23	0.62	8	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:183:LEU:HB2	A:191:ILE:HG21	0.62	7	1
A:215:VAL:C	A:216:VAL:HG22	0.62	7	1
A:59:SER:HB2	A:87:ILE:HD12	0.62	1	1
A:213:LEU:HD21	A:272:THR:CA	0.62	10	1
A:77:MET:O	A:84:LEU:HD13	0.62	1	1
A:92:SER:HA	A:95:PRO:CD	0.62	8	1
A:13:VAL:HG22	A:38:HIS:HE2	0.62	10	1
A:82:ALA:HA	A:96:SER:HB3	0.61	8	1
A:257:ILE:HG22	A:258:SER:N	0.61	7	2
A:122:LYS:HA	A:126:SER:HB3	0.61	8	1
A:121:LYS:HA	A:125:LYS:HB2	0.60	8	1
A:202:LEU:N	A:216:VAL:HG23	0.60	7	1
A:93:SER:HB2	A:178:THR:HB	0.60	8	1
A:174:THR:HA	A:177:PHE:CD2	0.60	8	1
A:252:VAL:HB	A:257:ILE:HD11	0.60	8	1
A:191:ILE:O	A:195:ALA:HB3	0.60	10	1
A:91:PHE:CD1	A:147:MET:HB2	0.59	9	1
A:241:ILE:H	A:257:ILE:HA	0.59	8	1
A:87:ILE:HD11	A:152:MET:H	0.59	9	1
A:152:MET:HA	A:155:MET:HB3	0.59	8	1
A:204:ILE:O	A:274:LEU:HD21	0.59	10	1
A:243:GLN:C	A:252:VAL:HG21	0.58	8	1
A:243:GLN:HA	A:252:VAL:HG21	0.58	8	1
A:58:LEU:HB2	A:77:MET:HE1	0.58	8	1
A:58:LEU:HD13	A:77:MET:HB2	0.58	4	1
A:251:ILE:HA	A:257:ILE:HD11	0.58	7	1
A:95:PRO:CA	A:99:ALA:HB3	0.58	8	1
A:142:ARG:HB3	A:148:ALA:HA	0.58	8	1
A:171:ALA:HB2	A:234:TYR:H	0.58	7	1
A:241:ILE:H	A:282:LEU:HA	0.57	7	1
A:244:VAL:HG23	A:252:VAL:HG23	0.57	10	1
A:8:VAL:HG22	A:41:ILE:HD11	0.57	8	1
A:151:ALA:HB3	A:276:ALA:CB	0.57	7	1
A:203:ASN:O	A:204:ILE:HG23	0.57	9	1
A:232:PHE:CG	A:234:TYR:HE1	0.57	1	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:184:ASP:HA	A:187:ALA:HB3	0.57	3	1
A:214:MET:HA	A:271:VAL:CG1	0.57	10	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	281	217	45	19
2	281	205	50	26
3	281	208	55	18
4	281	223	33	25
5	281	232	34	15
6	281	218	44	19
7	281	183	51	47
8	281	188	45	48
9	281	213	47	21
10	281	204	39	38

There are 154 unique backbone outliers. Detailed list of outliers are tabulated below. The output is limited to 100 rows.

Chain	Res	Type	Models (Total)
A	210	ALA	5
A	268	PRO	5
A	281	LYS	5
A	96	SER	4
A	197	ASP	4
A	213	LEU	4
A	246	ASP	4
A	258	SER	4
A	270	ILE	4
A	280	VAL	4
A	3	SER	3
A	81	ARG	3
A	95	PRO	3
A	115	ASP	3
A	116	SER	3
A	131	LYS	3

Chain	Res	Type	Models (Total)
A	141	GLN	3
A	143	LYS	3
A	145	GLU	3
A	187	ALA	3
A	200	VAL	3
A	204	ILE	3
A	209	THR	3
A	212	LYS	3
A	218	PRO	3
A	219	ASP	3
A	228	ASP	3
A	232	PHE	3
A	249	SER	3
A	251	ILE	3
A	252	VAL	3
A	254	LEU	3
A	257	ILE	3
A	267	TRP	3
A	277	ASN	3
A	23	ARG	2
A	42	LEU	2
A	76	GLU	2
A	78	LEU	2
A	83	THR	2
A	89	SER	2
A	114	GLY	2
A	117	GLU	2
A	125	LYS	2
A	126	SER	2
A	158	GLN	2
A	159	ALA	2
A	164	LYS	2
A	167	LYS	2
A	180	LEU	2
A	181	ARG	2

Chain	Res	Type	Models (Total)
A	195	ALA	2
A	196	ARG	2
A	202	LEU	2
A	203	ASN	2
A	208	THR	2
A	211	ALA	2
A	214	MET	2
A	215	VAL	2
A	216	VAL	2
A	217	ILE	2
A	229	GLY	2
A	230	THR	2
A	238	LEU	2
A	241	ILE	2
A	243	GLN	2
A	245	VAL	2
A	248	ASP	2
A	260	ASP	2
A	261	ASN	2
A	262	SER	2
A	263	PRO	2
A	269	LEU	2
A	278	SER	2
A	2	PRO	1
A	4	LYS	1
A	25	GLU	1
A	28	SER	1
A	30	LEU	1
A	33	GLN	1
A	43	LEU	1
A	44	ALA	1
A	47	THR	1
A	58	LEU	1
A	59	SER	1
A	64	MET	1

Chain	Res	Type	Models (Total)
A	65	GLN	1
A	67	ALA	1
A	68	VAL	1
A	69	ASP	1
A	70	ILE	1
A	75	GLU	1
A	79	ASP	1
A	82	ALA	1
A	85	GLN	1
A	86	ALA	1
A	87	ILE	1
A	88	ALA	1
A	90	GLU	1
A	91	PHE	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	245	191	17	37
2	245	181	40	24
3	245	187	22	36
4	245	191	27	27
5	245	201	18	26
6	245	196	29	20
7	245	171	38	36
8	245	189	20	36
9	245	190	22	33
10	245	188	26	31

There are 166 unique sidechain outliers. Detailed list of outliers are tabulated below. The output is limited to 100 rows.

Chain	Res	Type	Models (Total)
A	137	ASP	5
A	205	ILE	5
A	81	ARG	4
A	89	SER	4

Chain	Res	Type	Models (Total)
A	146	LYS	4
A	193	ASN	4
A	196	ARG	4
A	212	LYS	4
A	215	VAL	4
A	230	THR	4
A	245	VAL	4
A	250	LYS	4
A	251	ILE	4
A	252	VAL	4
A	24	VAL	3
A	45	LYS	3
A	46	ASP	3
A	72	LYS	3
A	87	ILE	3
A	105	GLU	3
A	119	VAL	3
A	152	MET	3
A	155	MET	3
A	180	LEU	3
A	192	ILE	3
A	203	ASN	3
A	213	LEU	3
A	223	TYR	3
A	224	LYS	3
A	231	THR	3
A	234	TYR	3
A	254	LEU	3
A	261	ASN	3
A	270	ILE	3
A	271	VAL	3
A	277	ASN	3
A	281	LYS	3
A	3	SER	2
A	6	SER	2

Chain	Res	Type	Models (Total)
A	7	ASP	2
A	8	VAL	2
A	9	LYS	2
A	38	HIS	2
A	40	ASP	2
A	42	LEU	2
A	47	THR	2
A	54	MET	2
A	55	VAL	2
A	65	GLN	2
A	70	ILE	2
A	71	ASN	2
A	73	LEU	2
A	78	LEU	2
A	79	ASP	2
A	83	THR	2
A	84	LEU	2
A	90	GLU	2
A	95	PRO	2
A	108	GLU	2
A	115	ASP	2
A	123	LEU	2
A	124	LYS	2
A	125	LYS	2
A	128	ASN	2
A	142	ARG	2
A	156	TYR	2
A	162	GLU	2
A	164	LYS	2
A	168	VAL	2
A	175	MET	2
A	179	MET	2
A	185	ASN	2
A	189	ASN	2
A	191	ILE	2

Chain	Res	Type	Models (Total)
A	199	CYS	2
A	208	THR	2
A	209	THR	2
A	217	ILE	2
A	226	THR	2
A	232	PHE	2
A	233	THR	2
A	238	LEU	2
A	240	GLU	2
A	249	SER	2
A	256	GLU	2
A	262	SER	2
A	265	LEU	2
A	4	LYS	1
A	5	MET	1
A	11	THR	1
A	13	VAL	1
A	14	VAL	1
A	23	ARG	1
A	29	LYS	1
A	35	VAL	1
A	43	LEU	1
A	48	THR	1
A	51	PHE	1
A	52	GLU	1
A	56	SER	1

5. Fit to Data Used for Modeling Assessment

SAS data used in this integrative model could not be validated as the sasCIF file is currently unavailable or incomplete.

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

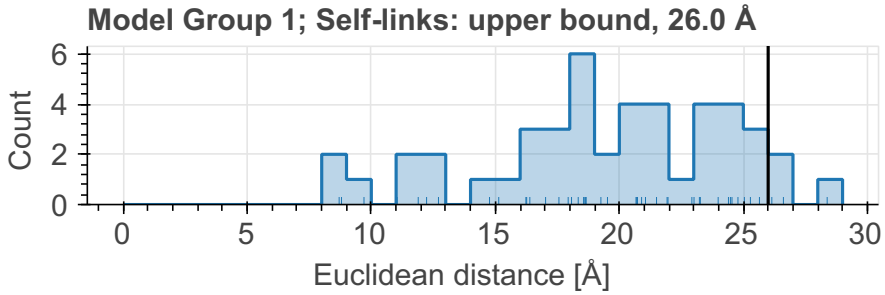
modeling.

There are 46 crosslinking restraints combined in 46 restraint groups.

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSSO	LYS	CA	LYS	CA	upper bound	26.00	33
DSSO	LYS	CA	SER	CA	upper bound	26.00	9
DSSO	THR	CA	THR	CA	upper bound	26.00	1
DSSO	LYS	CA	THR	CA	upper bound	26.00	2
DSSO	LYS	CA	TYR	CA	upper bound	26.00	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.



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=46)
1	1	1	10/10	All	93.48	6.52	46
				Self-links/ Intramolecular	93.48	6.52	46

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.

5.4. H/D exchange

Validation for this section is under development.

6. Fit to Data Used for Validation Assessment

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 Trewhella, 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.