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	9A1U | pdb_00009a1u
PDB-Dev ID	PDBDEV_00000120
Structure Title	SARS-CoV-2 nsp7-11 polyprotein
Structure Authors	Ruchi Yadav; Valentine V. Courouble; Sanjay K. Dey; Francesc X. Ruiz; Patrick R. Griffin; Eddy Arnold
Deposited on	2022-05-25

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.

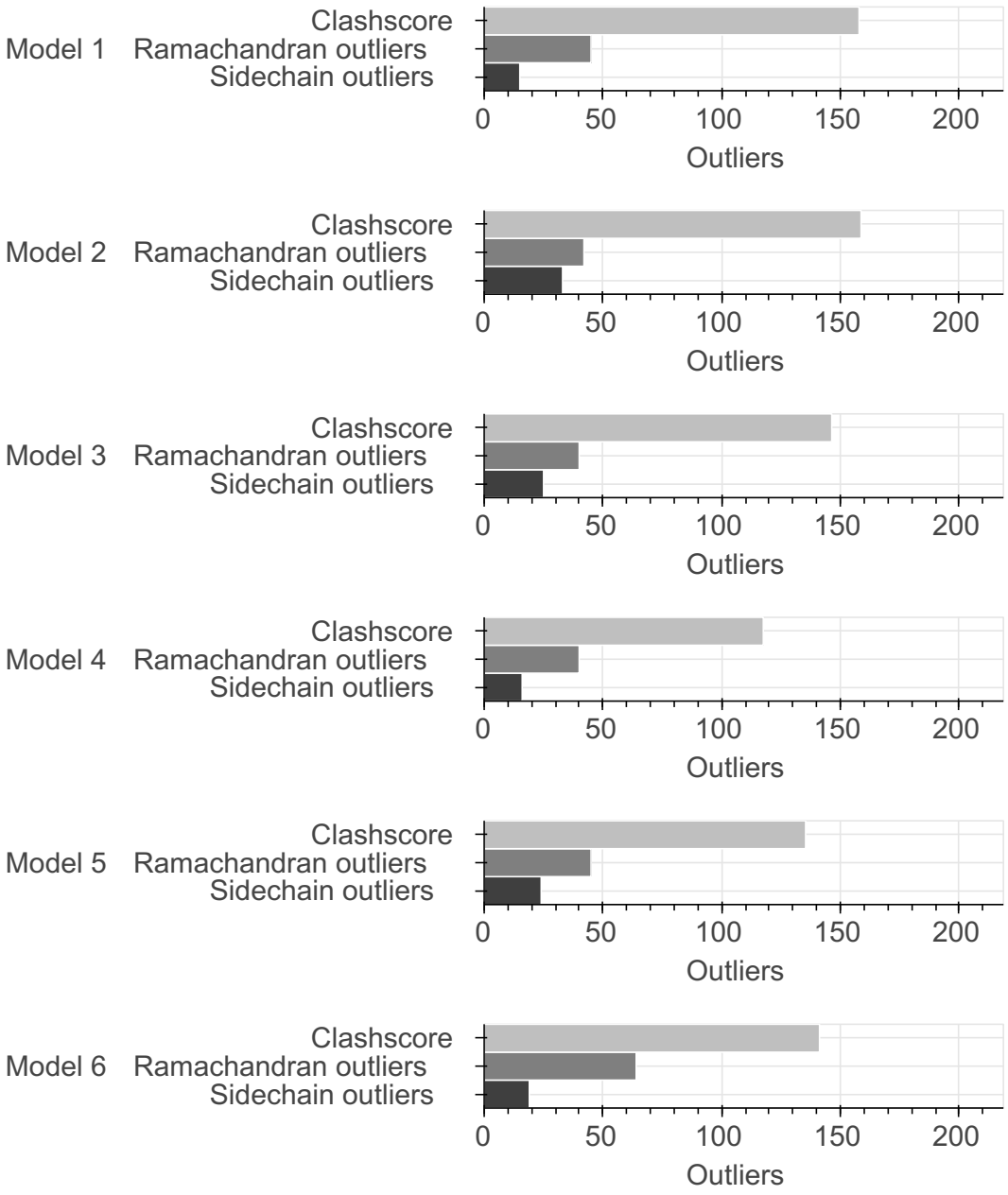
Name	Type	Count
Crosslinking-MS data	Experimental data	1
H/D exchange data	Experimental data	1

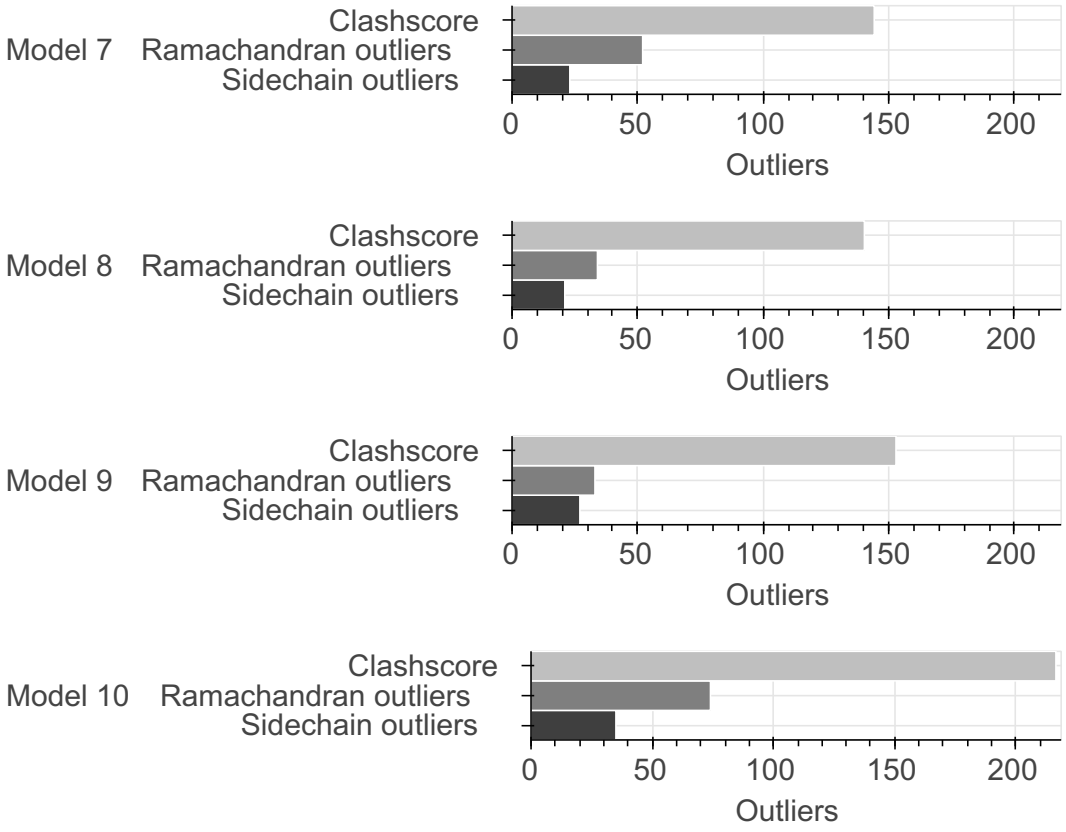
Name	Type	Count
SAS data	Experimental data	1
De Novo model	Starting model	1
Integrative model	Starting model	1

1.2. Overall quality

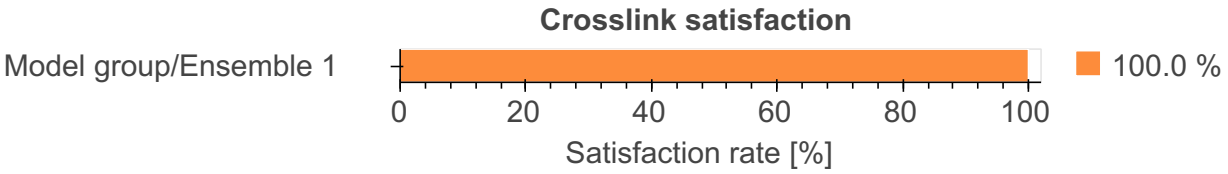
This validation report contains model quality assessments for all structures, data quality and fit to model assessments for SAS and crosslinking-MS datasets. Data quality and fit to model assessments for other datasets and model uncertainty are under development. Number of plots is limited to 256.

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 nsp7-11 polyprotein	A	548	-	1-283, 1-283, 284-548	100.00 / 100.00	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	SASDP23
4	Integrative model	PDB	pdb_00009a1t
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 9 bond length outliers in this entry (0.02% of 42334 assessed bonds). A summary is provided below.

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	472	TYR	N-CA	5.35	1.36	1.46	7	1
A	257	ILE	CA-CB	4.60	1.42	1.54	10	1
A	372	TYR	CB-CG	4.59	1.41	1.51	5	1
A	267	TRP	CD2-CE3	4.56	1.33	1.40	9	1
A	31	TRP	CD2-CE3	4.18	1.33	1.40	2	1
A	201	PRO	N-CD	4.16	1.41	1.47	2	1
A	135	ASP	C-N	4.14	1.39	1.33	10	1
A	492	TYR	CA-CB	4.10	1.61	1.53	6	1
A	423	TYR	CA-C	4.06	1.44	1.52	7	1

Standard geometry: angle outliers

There are 1971 bond angle outliers in this entry (3.43% of 57418 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	412	PHE	CA-CB-CG	19.13	94.67	113.80	7	6
A	100	PHE	CA-CB-CG	19.07	94.73	113.80	8	8
A	39	ASN	CA-CB-CG	19.03	93.57	112.60	10	7
A	464	PHE	CA-CB-CG	17.57	96.23	113.80	2	4
A	373	PHE	CA-CB-CG	17.02	96.78	113.80	9	7
A	546	PHE	CA-CB-CG	16.20	97.60	113.80	3	6
A	485	PHE	CA-CB-CG	14.95	98.85	113.80	2	5
A	510	ASN	C-N-CA	14.62	148.01	121.70	1	3
A	107	TYR	CA-CB-CG	13.82	89.02	113.90	10	4
A	358	PHE	CA-CB-CG	13.70	100.10	113.80	8	8
A	31	TRP	CB-CG-CD2	13.62	107.74	126.80	2	4
A	296	SER	C-N-CA	13.41	145.83	121.70	4	1
A	373	PHE	C-CA-CB	13.29	84.86	110.10	6	2
A	372	TYR	CA-CB-CG	13.09	90.34	113.90	5	5
A	423	TYR	N-CA-CB	13.04	132.67	110.50	7	2
A	266	ALA	C-N-CA	12.94	145.00	121.70	8	2
A	267	TRP	C-CA-CB	12.82	85.74	110.10	9	2
A	239	TRP	C-CA-CB	12.60	86.17	110.10	3	4

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	472	TYR	C-CA-CB	12.55	133.94	110.10	7	3
A	279	ALA	C-N-CA	12.48	144.16	121.70	9	2
A	134	PHE	C-CA-CB	12.47	86.40	110.10	9	1
A	134	PHE	CA-CB-CG	12.24	101.56	113.80	8	8
A	223	TYR	C-CA-CB	12.09	87.13	110.10	10	1
A	271	VAL	C-N-CA	12.07	143.43	121.70	10	2
A	406	ASN	CA-CB-CG	11.94	100.66	112.60	5	5
A	251	ILE	C-CA-CB	11.90	87.80	111.60	6	1
A	226	THR	C-N-CA	11.85	100.37	121.70	10	2
A	484	GLY	C-N-CA	11.83	100.41	121.70	7	1
A	366	PRO	C-N-CA	11.80	142.95	121.70	1	4
A	67	ALA	C-N-CA	11.74	142.82	121.70	2	2
A	113	ASN	C-CA-CB	11.66	87.94	110.10	10	2
A	509	LYS	C-N-CA	11.54	142.48	121.70	1	2
A	225	ASN	CA-CB-CG	11.51	101.09	112.60	10	9
A	290	VAL	C-N-CA	11.49	142.39	121.70	7	1
A	323	PHE	CA-CB-CG	11.48	102.32	113.80	3	7
A	382	ARG	C-CA-CB	11.26	88.71	110.10	6	2
A	30	LEU	C-N-CA	11.24	101.47	121.70	2	4
A	31	TRP	CB-CG-CD1	11.18	143.67	126.90	2	4
A	444	HIS	CA-CB-CG	11.14	124.94	113.80	9	4
A	415	PHE	CA-CB-CG	11.05	102.75	113.80	7	6
A	336	TRP	C-CA-CB	11.00	89.20	110.10	1	2
A	381	ASN	CA-CB-CG	10.98	101.62	112.60	7	6
A	314	TYR	C-N-CA	10.95	141.40	121.70	1	1
A	184	ASP	CA-CB-CG	10.89	101.71	112.60	7	8
A	148	ALA	C-CA-CB	10.87	94.20	110.50	9	2
A	239	TRP	N-CA-CB	10.86	128.96	110.50	10	5
A	323	PHE	C-N-CA	10.82	141.17	121.70	7	2
A	189	ASN	CA-CB-CG	10.77	101.83	112.60	9	7
A	51	PHE	CA-CB-CG	10.67	103.13	113.80	9	9
A	472	TYR	CA-CB-CG	10.61	94.80	113.90	6	3
A	213	LEU	C-CA-CB	10.60	89.95	110.10	3	3
A	285	ASN	CA-CB-CG	10.52	102.08	112.60	2	6
A	145	GLU	C-CA-CB	10.51	90.13	110.10	10	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	177	PHE	CA-CB-CG	10.48	103.32	113.80	6	7
A	452	THR	C-N-CA	10.42	140.46	121.70	6	3
A	27	SER	C-N-CA	10.37	103.04	121.70	10	2
A	277	ASN	CA-CB-CG	10.35	102.25	112.60	2	7
A	492	TYR	N-CA-CB	10.25	127.93	110.50	6	4
A	272	THR	C-N-CA	10.24	140.13	121.70	9	1
A	209	THR	C-N-CA	10.20	140.06	121.70	9	2
A	418	ASP	CA-CB-CG	10.19	102.41	112.60	6	6
A	506	PHE	CA-CB-CG	10.10	103.70	113.80	6	8
A	520	LYS	C-N-CA	9.91	139.53	121.70	6	1
A	87	ILE	C-N-CA	9.90	139.52	121.70	10	1
A	88	ALA	N-CA-CB	9.89	125.23	110.40	10	1
A	156	TYR	CA-CB-CG	9.87	96.13	113.90	8	5
A	148	ALA	N-CA-CB	9.79	125.09	110.40	8	2
A	168	VAL	C-N-CA	9.74	104.17	121.70	10	2
A	265	LEU	C-N-CA	9.72	104.20	121.70	10	2
A	66	GLY	C-N-CA	9.69	139.13	121.70	10	3
A	91	PHE	CA-CB-CG	9.62	104.18	113.80	1	3
A	213	LEU	N-CA-CB	9.54	126.72	110.50	1	5
A	270	ILE	N-CA-CB	9.53	127.71	111.50	10	1
A	527	ASP	C-CA-CB	9.52	128.19	110.10	8	1
A	372	TYR	C-CA-CB	9.49	92.06	110.10	8	4
A	145	GLU	N-CA-CB	9.49	126.64	110.50	10	2
A	31	TRP	N-CA-CB	9.45	126.56	110.50	2	3
A	323	PHE	C-CA-CB	9.44	92.16	110.10	4	2
A	61	LEU	C-N-CA	9.36	104.84	121.70	2	2
A	205	ILE	C-CA-CB	9.29	93.02	111.60	3	3
A	47	THR	N-CA-CB	9.28	95.72	111.50	8	1
A	329	SER	C-N-CA	9.24	138.34	121.70	7	3
A	423	TYR	C-CA-CB	9.19	92.64	110.10	7	3
A	317	THR	C-N-CA	9.17	138.21	121.70	6	2
A	455	PRO	CA-N-CD	9.11	99.24	112.00	6	2
A	419	ALA	N-CA-CB	9.10	96.75	110.40	6	2
A	429	SER	C-N-CA	9.10	105.32	121.70	5	3
A	343	ASP	CA-CB-CG	9.08	103.52	112.60	8	4

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	419	ALA	C-CA-CB	9.01	124.02	110.50	6	1
A	62	LEU	C-N-CA	9.00	105.50	121.70	10	2
A	526	CYS	C-N-CA	8.97	105.55	121.70	1	4
A	475	CYS	C-CA-CB	8.91	93.18	110.10	1	1
A	273	ALA	C-CA-CB	8.86	97.21	110.50	9	1
A	234	TYR	CA-CB-CG	8.82	98.02	113.90	9	1
A	269	LEU	C-CA-CB	8.77	93.44	110.10	8	3
A	69	ASP	C-N-CA	8.75	137.45	121.70	8	3
A	64	MET	N-CA-C	8.75	86.51	111.00	4	1
A	448	GLY	C-N-CA	8.73	105.98	121.70	1	1
A	59	SER	C-N-CA	8.67	137.30	121.70	3	1
A	95	PRO	CA-N-CD	8.63	99.92	112.00	10	3

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	157.86	1318
2	158.70	1325
3	146.36	1222
4	117.53	981
5	135.35	1130
6	141.21	1179
7	144.21	1204
8	140.29	1171
9	152.87	1276
10	216.48	1807

There are 12613 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:423:TYR:HB3	A:474:ARG:HB3	1.19	5	1
A:377:LEU:HD11	A:382:ARG:HB2	1.18	2	1
A:141:GLN:HB2	A:180:LEU:HD23	1.18	10	1
A:424:LYS:HE3	A:509:LYS:HD2	1.18	8	1
A:155:MET:HE3	A:160:ARG:HB3	1.17	7	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:336:TRP:HE1	A:338:ARG:HB2	1.17	7	1
A:339:PHE:HD2	A:369:LYS:HD3	1.17	2	1
A:437:CYS:HA	A:468:SER:HB3	1.17	5	1
A:372:TYR:HE1	A:385:VAL:HA	1.17	4	1
A:141:GLN:HG3	A:145:GLU:HG2	1.17	9	1
A:94:LEU:HD12	A:137:ASP:HB3	1.16	2	1
A:240:GLU:HB2	A:268:PRO:HB2	1.16	10	1
A:162:GLU:HG3	A:374:ILE:HG22	1.16	1	1
A:87:ILE:HA	A:90:GLU:HG2	1.16	4	3
A:363:PRO:HA	A:367:LYS:HE2	1.16	4	1
A:226:THR:HG21	A:233:THR:HG22	1.16	7	2
A:292:LEU:HD23	A:317:THR:HG22	1.16	2	1
A:140:MET:HE1	A:188:LEU:HG	1.16	4	1
A:192:ILE:HG23	A:197:ASP:HB3	1.15	8	1
A:62:LEU:HD11	A:70:ILE:HA	1.15	10	1
A:217:ILE:HD11	A:223:TYR:HA	1.15	1	1
A:488:LEU:HD23	A:493:VAL:HG13	1.15	3	1
A:12:SER:HB2	A:54:MET:HE1	1.15	5	2
A:70:ILE:HB	A:157:LYS:HE2	1.15	6	1
A:243:GLN:HA	A:272:THR:HG23	1.15	8	1
A:6:SER:HB2	A:9:LYS:HG2	1.15	10	1
A:164:LYS:HE2	A:281:LYS:HG2	1.15	9	1
A:135:ASP:HB2	A:137:ASP:HA	1.15	10	1
A:361:ASP:HB3	A:362:THR:HA	1.15	1	1
A:438:VAL:HG22	A:507:THR:HG21	1.15	3	1
A:25:GLU:HA	A:31:TRP:HE1	1.15	9	1
A:302:THR:HG21	A:331:LEU:HD22	1.15	2	1
A:77:MET:HE2	A:156:TYR:HB3	1.14	5	1
A:434:ILE:HB	A:475:CYS:HB2	1.14	6	1
A:328:LEU:HB3	A:424:LYS:HE2	1.14	7	1
A:351:GLU:HB3	A:354:PRO:HG3	1.14	1	1
A:322:ARG:HD2	A:386:LEU:HD22	1.14	5	1
A:328:LEU:HD13	A:356:CYS:HA	1.14	1	1
A:244:VAL:HB	A:270:ILE:HB	1.14	9	1
A:62:LEU:HD21	A:70:ILE:HG23	1.14	10	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:491:LYS:HA	A:508:LEU:HD11	1.14	2	1
A:14:VAL:HG21	A:78:LEU:HD11	1.14	8	1
A:350:THR:HG22	A:354:PRO:HB3	1.14	5	1
A:212:LYS:HD3	A:226:THR:HA	1.14	10	1
A:245:VAL:HG22	A:251:ILE:HG22	1.14	1	1
A:142:ARG:HD2	A:204:ILE:HG21	1.14	8	1
A:172:MET:HE3	A:317:THR:HA	1.13	1	1
A:87:ILE:HG12	A:89:SER:H	1.13	2	2
A:201:PRO:HA	A:215:VAL:HG12	1.13	3	1
A:367:LYS:HE3	A:389:LEU:HD23	1.13	3	1
A:124:LYS:HE2	A:482:PRO:HA	1.13	7	1
A:434:ILE:HD11	A:468:SER:HA	1.12	5	1
A:488:LEU:HD11	A:507:THR:HG21	1.12	6	1
A:85:GLN:HB3	A:90:GLU:HB2	1.12	1	1
A:31:TRP:HB3	A:34:CYS:HB3	1.12	10	1
A:191:ILE:HD11	A:215:VAL:HA	1.12	10	1
A:16:LEU:HB2	A:37:LEU:HD13	1.12	1	1
A:204:ILE:HG22	A:205:ILE:HG13	1.12	10	1
A:424:LYS:HE3	A:507:THR:HB	1.12	1	1
A:213:LEU:HD13	A:276:ALA:H	1.12	2	1
A:375:LYS:HB2	A:388:SER:HB3	1.12	6	1
A:243:GLN:H	A:270:ILE:HD13	1.12	10	1
A:27:SER:HB3	A:30:LEU:HD23	1.11	9	1
A:141:GLN:HG2	A:243:GLN:HG3	1.11	9	1
A:341:LYS:HG2	A:367:LYS:HG2	1.11	2	1
A:115:ASP:HB3	A:119:VAL:HG11	1.11	1	1
A:201:PRO:HG2	A:278:SER:H	1.11	2	2
A:377:LEU:HD11	A:383:GLY:H	1.11	5	1
A:408:THR:HG22	A:480:PRO:HG2	1.11	1	1
A:5:MET:HG3	A:8:VAL:HG12	1.11	2	1
A:9:LYS:HG3	A:42:LEU:HG	1.11	5	2
A:375:LYS:HB3	A:392:THR:HB	1.11	6	1
A:225:ASN:HA	A:234:TYR:CD1	1.11	10	1
A:378:ASN:HB3	A:380:LEU:HD13	1.11	2	1
A:451:ILE:HG12	A:488:LEU:HD21	1.11	4	4

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:58:LEU:HD21	A:77:MET:HG2	1.11	9	1
A:120:LEU:HB2	A:144:LEU:HD21	1.11	10	1
A:510:ASN:HB2	A:511:THR:HG22	1.11	1	1
A:201:PRO:HG3	A:215:VAL:HG22	1.11	2	1
A:173:GLN:HG2	A:382:ARG:HG2	1.11	5	1
A:454:THR:HG21	A:493:VAL:HA	1.11	6	1
A:12:SER:HB3	A:41:ILE:HD13	1.11	4	1
A:450:ALA:HA	A:519:TRP:HE1	1.11	1	1
A:54:MET:H	A:163:ASP:HB3	1.11	3	1
A:453:VAL:HG12	A:491:LYS:HE2	1.11	4	1
A:113:ASN:HB3	A:120:LEU:HB2	1.11	9	1
A:91:PHE:HA	A:165:ARG:HB2	1.11	10	1
A:216:VAL:HG22	A:270:ILE:HG12	1.10	2	1
A:315:TYR:HD2	A:390:ALA:HA	1.10	1	1
A:241:ILE:HA	A:265:LEU:HD22	1.10	10	1
A:21:GLN:HA	A:24:VAL:HG23	1.10	1	1
A:400:ALA:HB3	A:403:VAL:HG23	1.10	10	1
A:88:ALA:HB2	A:157:LYS:HG3	1.10	4	1
A:526:CYS:HA	A:529:LEU:HD11	1.10	4	1
A:55:VAL:HG11	A:177:PHE:HE1	1.10	8	1
A:131:LYS:HA	A:134:PHE:HD2	1.10	6	1
A:372:TYR:H	A:385:VAL:HG13	1.10	7	1
A:19:LEU:HD13	A:63:SER:HB2	1.10	2	1
A:374:ILE:HA	A:377:LEU:HD21	1.10	6	1
A:341:LYS:HD2	A:368:VAL:HB	1.10	7	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	546	434	67	45
2	546	443	61	42
3	546	447	59	40
4	546	433	73	40
5	546	434	67	45
6	546	430	52	64

Model ID	Analysed	Favored	Allowed	Outliers
7	546	437	57	52
8	546	464	48	34
9	546	473	40	33
10	546	392	80	74

There are 261 unique backbone outliers. Detailed list of outliers are tabulated below. The output is limited to 100 rows.

Chain	Res	Type	Models (Total)
A	479	HIS	8
A	480	PRO	7
A	450	ALA	6
A	95	PRO	5
A	227	CYS	5
A	370	TYR	5
A	431	GLY	5
A	2	PRO	4
A	24	VAL	4
A	198	GLY	4
A	243	GLN	4
A	267	TRP	4
A	275	ARG	4
A	324	VAL	4
A	335	LYS	4
A	345	THR	4
A	371	LEU	4
A	438	VAL	4
A	547	ALA	4
A	65	GLN	3
A	70	ILE	3
A	87	ILE	3
A	116	SER	3
A	168	VAL	3
A	172	MET	3
A	180	LEU	3
A	205	ILE	3
A	226	THR	3

Chain	Res	Type	Models (Total)
A	251	ILE	3
A	268	PRO	3
A	287	LEU	3
A	289	PRO	3
A	290	VAL	3
A	293	ARG	3
A	301	THR	3
A	304	THR	3
A	327	LEU	3
A	336	TRP	3
A	351	GLU	3
A	354	PRO	3
A	364	LYS	3
A	366	PRO	3
A	395	LEU	3
A	402	GLU	3
A	403	VAL	3
A	404	PRO	3
A	451	ILE	3
A	455	PRO	3
A	477	ILE	3
A	481	ASN	3
A	482	PRO	3
A	502	ASP	3
A	531	GLU	3
A	4	LYS	2
A	7	ASP	2
A	23	ARG	2
A	68	VAL	2
A	69	ASP	2
A	71	ASN	2
A	74	CYS	2
A	75	GLU	2
A	76	GLU	2
A	79	ASP	2

Chain	Res	Type	Models (Total)
A	113	ASN	2
A	114	GLY	2
A	117	GLU	2
A	140	MET	2
A	157	LYS	2
A	164	LYS	2
A	169	THR	2
A	178	THR	2
A	194	ASN	2
A	199	CYS	2
A	211	ALA	2
A	216	VAL	2
A	229	GLY	2
A	237	ALA	2
A	242	GLN	2
A	252	VAL	2
A	262	SER	2
A	272	THR	2
A	274	LEU	2
A	276	ALA	2
A	283	GLN	2
A	288	SER	2
A	294	GLN	2
A	300	GLY	2
A	308	ASP	2
A	312	LEU	2
A	316	ASN	2
A	323	PHE	2
A	325	LEU	2
A	329	SER	2
A	330	ASP	2
A	331	LEU	2
A	332	GLN	2
A	340	PRO	2
A	341	LYS	2

Chain	Res	Type	Models (Total)
A	343	ASP	2
A	346	GLY	2

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	461	408	38	15
2	461	369	59	33
3	461	393	43	25
4	461	395	50	16
5	461	392	45	24
6	461	397	45	19
7	461	389	49	23
8	461	392	48	21
9	461	389	45	27
10	461	372	54	35

There are 158 unique sidechain outliers. Detailed list of outliers are tabulated below. The output is limited to 100 rows.

Chain	Res	Type	Models (Total)
A	519	TRP	6
A	87	ILE	5
A	444	HIS	5
A	141	GLN	4
A	4	LYS	3
A	31	TRP	3
A	77	MET	3
A	123	LEU	3
A	128	ASN	3
A	146	LYS	3
A	212	LYS	3
A	293	ARG	3
A	327	LEU	3
A	328	LEU	3
A	367	LYS	3
A	372	TYR	3

Chain	Res	Type	Models (Total)
A	384	MET	3
A	483	LYS	3
A	509	LYS	3
A	20	GLN	2
A	30	LEU	2
A	38	HIS	2
A	78	LEU	2
A	81	ARG	2
A	97	TYR	2
A	100	PHE	2
A	113	ASN	2
A	202	LEU	2
A	203	ASN	2
A	204	ILE	2
A	217	ILE	2
A	221	ASN	2
A	239	TRP	2
A	257	ILE	2
A	269	LEU	2
A	270	ILE	2
A	275	ARG	2
A	287	LEU	2
A	324	VAL	2
A	331	LEU	2
A	334	LEU	2
A	341	LYS	2
A	378	ASN	2
A	380	LEU	2
A	385	VAL	2
A	421	LYS	2
A	424	LYS	2
A	432	GLN	2
A	440	MET	2
A	472	TYR	2
A	477	ILE	2

Chain	Res	Type	Models (Total)
A	479	HIS	2
A	529	LEU	2
A	9	LYS	1
A	15	LEU	1
A	29	LYS	1
A	36	GLN	1
A	37	LEU	1
A	42	LEU	1
A	43	LEU	1
A	54	MET	1
A	57	LEU	1
A	62	LEU	1
A	65	GLN	1
A	72	LYS	1
A	76	GLU	1
A	84	LEU	1
A	85	GLN	1
A	90	GLU	1
A	94	LEU	1
A	95	PRO	1
A	104	GLN	1
A	119	VAL	1
A	120	LEU	1
A	131	LYS	1
A	144	LEU	1
A	157	LYS	1
A	164	LYS	1
A	165	ARG	1
A	167	LYS	1
A	172	MET	1
A	173	GLN	1
A	179	MET	1
A	182	LYS	1
A	183	LEU	1
A	193	ASN	1

Chain	Res	Type	Models (Total)
A	196	ARG	1
A	214	MET	1
A	222	THR	1
A	224	LYS	1
A	227	CYS	1
A	231	THR	1
A	232	PHE	1
A	244	VAL	1
A	245	VAL	1
A	248	ASP	1
A	251	ILE	1
A	252	VAL	1
A	256	GLU	1
A	267	TRP	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.

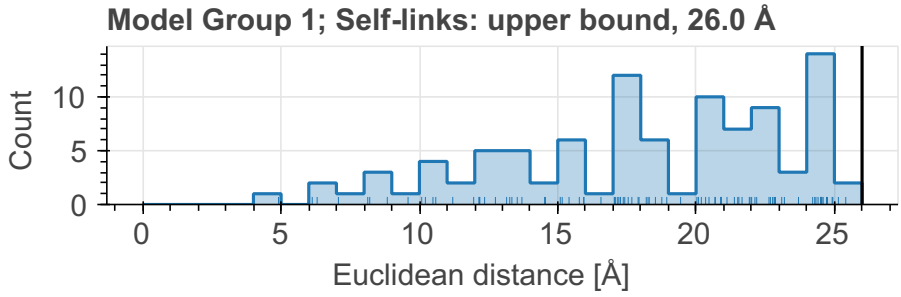
There are 97 crosslinking restraints combined in 97 restraint groups.

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSSO	THR	CA	THR	CA	upper bound	26.00	2
DSSO	LYS	CA	LYS	CA	upper bound	26.00	72
DSSO	LYS	CA	SER	CA	upper bound	26.00	14
DSSO	LYS	CA	THR	CA	upper bound	26.00	6
DSSO	LYS	CA	TYR	CA	upper bound	26.00	3

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.



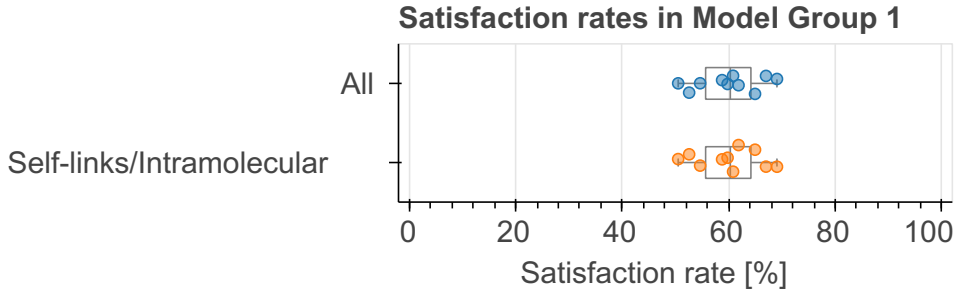
5.2.2. Satisfaction of restraints

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=97)
1	1	1	10/10	All	100.00	0.00	97
				Self-links/ Intramolecular	100.00	0.00	97

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.