Integrative Structure Validation Report

October 09, 2025 - 04:43 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
PrISM Version dbe5a41
PyMOL Version 2.5.0

PDB ID	9A3F | pdb_00009a3f
PDB-Dev ID	PDBDEV_00000200
Structure Title	Modeling of the GLP-1 receptor-Gs complex
Structure Authors	Shijia Yuan; Lisha Xia; Chenxi Wang; Fan Wu; Bingjie Zhang; Chen Pan; Xiaoguang Lei; Raymond Stevens; Andrej Sali; Liping Sun; Wenqing Shui
Deposited on	2023-01-09

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 7 model(s). A total of 12 dataset(s) were used to build this entry.

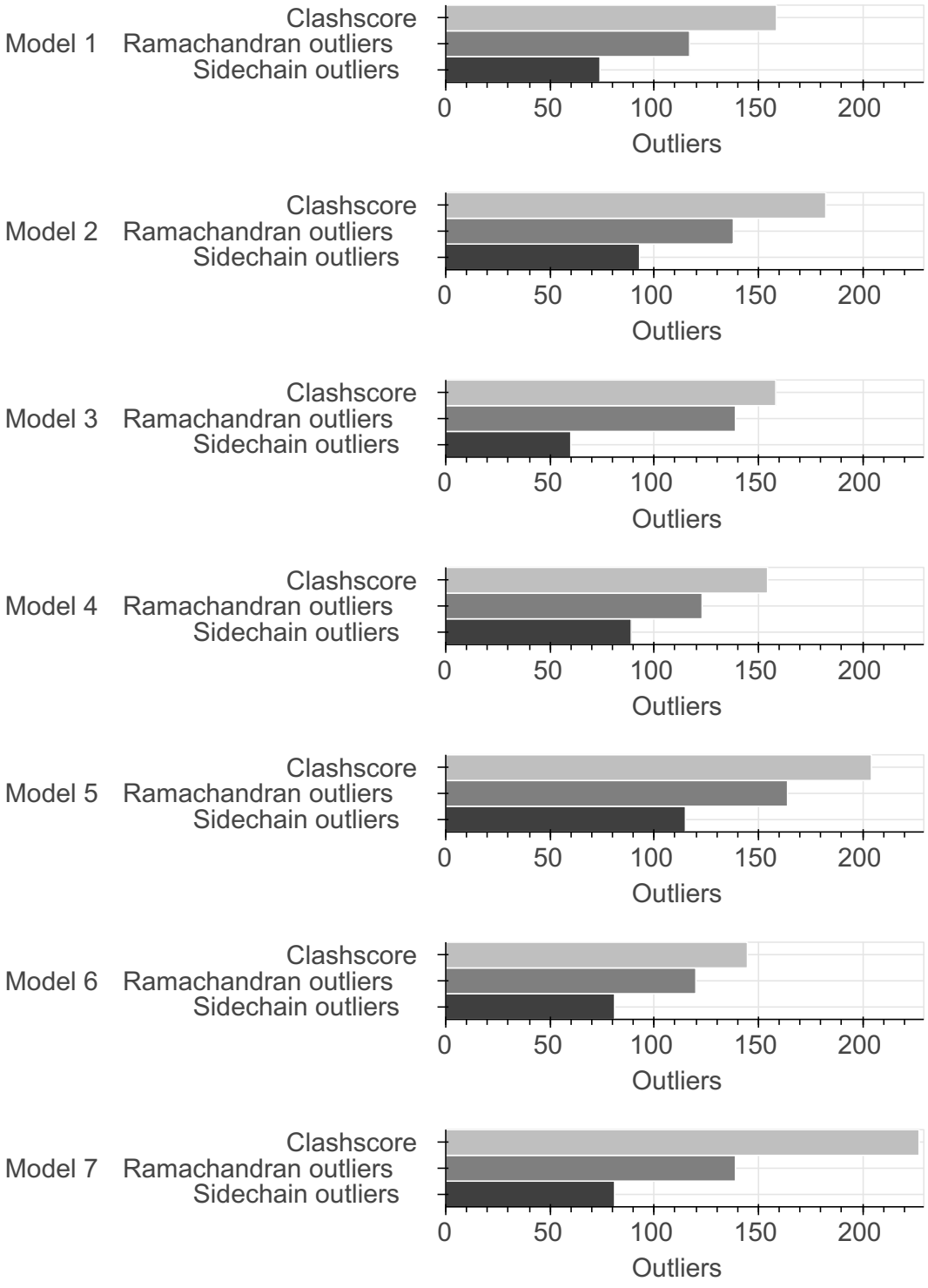
Name	Type	Count
Crosslinking-MS data	Experimental data	1
Other	Experimental data	4

Name	Type	Count
Experimental model	Starting model	7

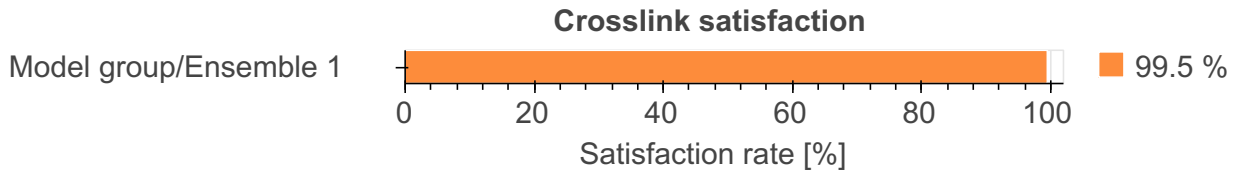
1.2. Overall quality

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 1 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	1-7	1	GLP1R	A	440	-	1-400, 401-440	100.00 / 90.91	Atomic
		2	GLP1	B	31	-	1-31	100.00 / 100.00	Atomic
		3	GA	C	394	-	1-10, 11-391, 392-394	100.00 / 96.70	Atomic
		4	GB	D	340	-	1-2, 3-340	100.00 / 99.41	Atomic
		5	GG	E	71	-	1-6, 7-45, 46-71	100.00 / 54.93	Atomic
		6	NB	F	129	-	1, 2-127, 128-129	100.00 / 97.67	Atomic

2.3. Datasets used for modeling

There are 12 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	Experimental model	PDB	pdb_00006x18
2	Crosslinking-MS data	iProX	IPX0004576001
3	Experimental model	Zenodo	10.5281/zenodo.13325078
4	Experimental model	Zenodo	10.5281/zenodo.13325078
5	Experimental model	Zenodo	10.5281/zenodo.13325078
6	Experimental model	Zenodo	10.5281/zenodo.13325078

ID	Dataset type	Database name	Data access code
7	Experimental model	Zenodo	10.5281/zenodo.13325078
8	Experimental model	Zenodo	10.5281/zenodo.13325078
9	Other	Zenodo	10.5281/zenodo.13325078
10	Other	Zenodo	10.5281/zenodo.13325078
11	Other	Zenodo	10.5281/zenodo.13325078
12	Other	Zenodo	10.5281/zenodo.13325078

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	Sampling	Replica exchange monte carlo	Not available	844020	False	False

There are 3 software packages reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	IMP PMI module	Not available	integrative model building	https://integrativemodeling.org
2	Integrative Modeling Platform (IMP)	Not available	integrative model building	https://integrativemodeling.org
3	MODELLER	Not available	model building	https://salilab.org/modeller/

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).

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

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 1442 bond length outliers in this entry (1.80% of 79975 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)
C	187	GLN	C-N	33.65	1.80	1.33	7	4
C	187	GLN	CA-C	32.11	2.20	1.52	7	4
C	280	LYS	C-N	30.70	1.76	1.33	2	7
C	188	ALA	C-N	28.47	1.73	1.33	7	6
C	305	LYS	C-N	27.40	1.71	1.33	5	6
C	305	LYS	CA-C	27.25	2.10	1.52	5	4
C	188	ALA	N-CA	27.07	1.97	1.46	7	4
C	282	LEU	C-N	27.02	1.71	1.33	2	5
C	188	ALA	CA-C	26.97	2.09	1.52	7	6
C	248	VAL	C-N	26.88	1.71	1.33	3	6
C	58	LYS	C-N	26.53	1.70	1.33	4	7
D	33	ILE	C-N	26.36	1.70	1.33	7	7
D	34	THR	C-N	24.94	1.68	1.33	7	7
C	286	SER	C-N	24.94	1.68	1.33	5	7
D	32	GLN	C-N	24.81	1.68	1.33	7	6
D	35	ASN	C-N	24.41	1.67	1.33	7	5
C	50	GLU	C-N	24.40	1.67	1.33	2	7
D	36	ASN	C-N	24.26	1.67	1.33	7	4
C	286	SER	CA-C	23.88	2.03	1.52	5	7
C	186	LYS	C-N	23.64	1.66	1.33	5	5
C	189	ASP	N-CA	23.46	1.90	1.46	7	6
D	33	ILE	CA-C	23.40	2.02	1.52	7	6
C	281	TRP	N-CA	23.33	1.90	1.46	2	6
C	248	VAL	CA-C	22.96	2.01	1.52	3	5
C	292	ASN	C-N	22.33	1.64	1.33	2	7
C	50	GLU	CA-C	22.18	1.99	1.52	5	7
D	32	GLN	CA-C	22.18	1.99	1.52	7	5
C	186	LYS	CA-C	22.02	1.99	1.52	5	5
D	35	ASN	CA-C	22.00	1.99	1.52	7	4
C	249	ALA	C-N	21.80	1.63	1.33	3	5
C	249	ALA	CA-C	21.75	1.98	1.52	3	5
D	34	THR	CA-C	21.74	1.98	1.52	7	3
D	37	ILE	C-N	21.60	1.63	1.33	7	5

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
C	249	ALA	N-CA	21.59	1.87	1.46	3	5
C	59	GLN	N-CA	21.46	1.87	1.46	4	6
C	283	ARG	C-N	21.28	1.63	1.33	6	6
C	51	SER	C-N	21.22	1.63	1.33	1	4
D	36	ASN	CA-C	20.94	1.96	1.52	7	2
C	51	SER	N-CA	20.91	1.86	1.46	2	6
C	292	ASN	CA-C	20.84	1.96	1.52	2	7
D	34	THR	N-CA	20.76	1.85	1.46	7	5
C	51	SER	CA-C	20.50	1.96	1.52	1	3
C	306	SER	N-CA	20.34	1.84	1.46	5	5
C	287	VAL	N-CA	20.24	1.84	1.46	5	7
C	283	ARG	CA-C	19.57	1.94	1.52	6	6
D	35	ASN	N-CA	19.35	1.83	1.46	7	3
D	33	ILE	N-CA	19.27	1.82	1.46	7	5
C	283	ARG	N-CA	19.20	1.82	1.46	2	2
C	304	GLY	CA-C	19.11	1.86	1.52	7	2
D	36	ASN	N-CA	19.02	1.82	1.46	7	3
D	37	ILE	CA-C	18.88	1.92	1.52	7	2
C	187	GLN	N-CA	18.63	1.81	1.46	1	5
D	37	ILE	N-CA	18.47	1.81	1.46	7	2
C	52	GLY	N-CA	18.41	1.75	1.45	1	3
C	260	ASP	C-N	18.26	1.58	1.33	4	2
C	58	LYS	CA-C	18.24	1.91	1.52	4	7
C	307	LYS	C-N	18.13	1.58	1.33	4	6
C	87	GLU	C-N	17.85	1.58	1.33	7	7
C	392	GLU	C-N	17.84	1.58	1.33	5	5
C	306	SER	C-N	17.78	1.58	1.33	4	5
D	38	ASP	N-CA	17.55	1.79	1.46	7	2
C	295	ASP	C-N	17.50	1.57	1.33	4	4
C	293	LYS	N-CA	17.32	1.79	1.46	2	7
C	284	ASP	C-N	17.23	1.57	1.33	6	5
C	250	SER	N-CA	17.22	1.79	1.46	3	5
C	87	GLU	CA-C	17.19	1.89	1.52	7	6
C	304	GLY	C-N	16.99	1.57	1.33	7	6
D	47	THR	C-N	16.95	1.57	1.33	7	5

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
C	260	ASP	CA-C	16.94	1.88	1.52	4	1
C	284	ASP	N-CA	16.77	1.78	1.46	6	6
A	400	GLU	C-N	16.71	1.56	1.33	7	2
C	285	THR	C-N	16.31	1.56	1.33	5	6
C	279	ASN	CA-C	16.26	1.87	1.52	4	4
C	279	ASN	C-N	16.23	1.56	1.33	4	5
C	49	GLY	CA-C	16.22	1.81	1.52	4	3
C	49	GLY	C-N	15.79	1.55	1.33	4	5
C	46	LEU	CA-C	15.54	1.85	1.52	2	5
C	48	ALA	C-N	15.33	1.54	1.33	1	4
C	307	LYS	CA-C	15.18	1.84	1.52	4	1
D	46	ARG	C-N	15.15	1.54	1.33	6	5
C	282	LEU	CA-C	15.11	1.84	1.52	2	5
C	82	SER	C-N	15.06	1.54	1.33	7	5
C	84	SER	C-N	15.02	1.54	1.33	7	5
C	306	SER	CA-C	14.96	1.84	1.52	4	3
A	149	LEU	C-N	14.80	1.54	1.33	5	6
C	46	LEU	C-N	14.72	1.54	1.33	2	7
C	49	GLY	N-CA	14.72	1.69	1.45	1	2
C	391	TYR	C-N	14.66	1.53	1.33	5	7
C	393	LEU	N-CA	14.64	1.74	1.46	5	4
C	295	ASP	CA-C	14.60	1.83	1.52	4	4
C	284	ASP	CA-C	14.46	1.83	1.52	6	5
C	261	ASN	N-CA	14.37	1.73	1.46	4	1
C	307	LYS	N-CA	14.37	1.73	1.46	5	3
C	261	ASN	C-N	14.34	1.53	1.33	4	2
C	294	GLN	C-N	14.12	1.53	1.33	3	4
D	46	ARG	CA-C	14.12	1.82	1.52	6	3
A	149	LEU	CA-C	14.06	1.82	1.52	5	4
C	88	LYS	N-CA	14.05	1.72	1.46	7	6
C	308	ILE	N-CA	13.96	1.72	1.46	4	1
E	57	SER	C-N	13.84	1.52	1.33	3	3

Standard geometry: angle outliers

There are 9201 bond angle outliers in this entry (8.49% of 108374 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)
C	286	SER	C-CA-CB	22.31	67.72	110.10	2	3
C	286	SER	N-CA-CB	21.09	146.35	110.50	2	4
C	280	LYS	CA-C-O	20.51	85.93	120.80	2	5
C	188	ALA	C-CA-CB	20.13	80.30	110.50	2	4
C	208	PHE	CA-CB-CG	20.04	93.76	113.80	3	3
C	188	ALA	C-N-CA	19.74	157.24	121.70	7	5
C	305	LYS	C-N-CA	19.10	156.07	121.70	5	5
C	50	GLU	C-N-CA	18.94	155.79	121.70	2	6
C	184	VAL	C-CA-CB	18.49	146.54	111.40	4	4
C	51	SER	C-N-CA	18.48	154.96	121.70	1	4
C	238	PHE	CA-CB-CG	18.34	132.14	113.80	7	5
C	178	PHE	CA-CB-CG	18.33	95.47	113.80	4	5
C	282	LEU	CA-C-O	17.80	90.54	120.80	2	2
C	47	GLY	C-N-CA	17.64	153.46	121.70	5	4
F	109	PHE	CA-CB-CG	17.59	131.39	113.80	5	7
C	282	LEU	C-N-CA	17.47	153.14	121.70	2	5
C	200	CYS	N-CA-CB	17.11	139.58	110.50	1	5
C	286	SER	C-N-CA	17.09	152.47	121.70	7	6
C	187	GLN	CA-C-N	16.91	150.02	116.20	7	4
C	249	ALA	C-N-CA	16.83	152.00	121.70	3	5
C	392	GLU	C-N-CA	16.53	151.45	121.70	5	3
C	280	LYS	C-N-CA	16.51	151.41	121.70	2	5
C	46	LEU	C-CA-CB	16.34	141.15	110.10	2	4
D	292	PHE	CA-CB-CG	16.22	97.58	113.80	5	7
C	187	GLN	C-N-CA	15.75	150.05	121.70	2	5
C	58	LYS	C-N-CA	15.68	149.93	121.70	7	5
C	248	VAL	C-N-CA	15.58	149.74	121.70	3	5
C	186	LYS	C-N-CA	15.53	149.65	121.70	1	5
C	306	SER	C-N-CA	15.47	149.54	121.70	5	3
C	94	ASP	N-CA-CB	15.31	136.53	110.50	7	4
D	37	ILE	C-N-CA	15.27	149.19	121.70	7	2
C	295	ASP	C-N-CA	15.24	149.12	121.70	4	4
C	208	PHE	C-CA-CB	15.21	81.19	110.10	2	3
C	279	ASN	CA-C-N	15.11	146.43	116.20	4	4
F	107	ASP	C-CA-CB	15.07	81.47	110.10	5	5

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
C	50	GLU	CA-C-O	15.06	95.19	120.80	2	5
C	286	SER	CA-C-N	15.06	146.32	116.20	5	7
C	97	ASN	N-CA-CB	14.99	135.99	110.50	7	3
C	280	LYS	CA-C-N	14.82	145.85	116.20	2	6
C	247	VAL	C-N-CA	14.79	148.33	121.70	4	5
C	292	ASN	C-N-CA	14.79	148.31	121.70	5	7
C	71	GLU	N-CA-CB	14.73	135.54	110.50	5	3
C	94	ASP	CA-CB-CG	14.56	98.04	112.60	5	3
A	149	LEU	C-N-CA	14.51	147.82	121.70	5	5
D	39	PRO	N-CA-C	14.48	148.29	112.10	7	5
F	116	TYR	C-CA-CB	14.43	137.51	110.10	5	1
A	412	PRO	N-CA-C	14.33	147.92	112.10	2	2
C	273	PHE	CA-CB-CG	14.31	99.49	113.80	1	4
C	84	SER	N-CA-CB	14.23	134.69	110.50	5	3
F	3	VAL	CA-CB-CG2	14.21	86.24	110.40	5	1
C	208	PHE	N-CA-CB	14.16	134.57	110.50	2	5
C	57	VAL	CA-C-N	14.12	87.96	116.20	3	2
C	297	LEU	C-N-CA	14.11	147.09	121.70	4	2
C	182	ILE	CG1-CB-CG2	14.07	152.90	110.70	2	1
C	82	SER	N-CA-C	14.06	71.63	111.00	4	4
C	50	GLU	CA-C-N	13.74	143.68	116.20	2	7
C	186	LYS	CA-C-N	13.71	143.62	116.20	7	5
C	380	ARG	C-CA-CB	13.66	84.15	110.10	7	2
A	161	PHE	CA-CB-CG	13.64	100.16	113.80	7	3
C	292	ASN	CA-C-N	13.52	143.24	116.20	7	7
C	187	GLN	CA-C-O	13.48	97.89	120.80	2	3
C	246	PHE	CA-CB-CG	13.43	127.23	113.80	7	3
C	249	ALA	CA-C-N	13.37	142.95	116.20	3	4
D	32	GLN	CA-C-N	13.33	142.85	116.20	7	5
C	248	VAL	CA-C-N	13.31	142.83	116.20	3	5
C	178	PHE	C-CA-CB	13.29	84.84	110.10	3	1
C	294	GLN	N-CA-C	13.16	74.14	111.00	2	3
A	289	PRO	C-CA-CB	13.16	135.11	110.10	1	1
C	138	PRO	C-CA-CB	13.13	135.04	110.10	5	2
E	61	PHE	CA-CB-CG	13.13	100.67	113.80	2	6

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
C	292	ASN	O-C-N	13.08	102.07	123.00	7	5
C	202	VAL	N-CA-CB	13.05	133.69	111.50	1	6
C	90	THR	N-CA-CB	13.01	133.62	111.50	7	5
C	313	PRO	N-CA-C	12.96	144.50	112.10	4	6
D	33	ILE	C-N-CA	12.94	144.99	121.70	7	6
C	86	GLY	C-N-CA	12.91	144.93	121.70	7	3
C	281	TRP	N-CA-CB	12.87	132.38	110.50	2	3
C	283	ARG	CA-C-N	12.84	141.88	116.20	6	6
C	226	ALA	N-CA-CB	12.81	129.62	110.40	2	7
C	74	GLU	C-CA-CB	12.76	134.34	110.10	7	3
C	82	SER	C-N-CA	12.74	144.64	121.70	5	5
C	129	ASP	C-CA-CB	12.69	85.98	110.10	5	3
D	46	ARG	C-CA-CB	12.64	134.11	110.10	6	3
C	58	LYS	CA-C-N	12.59	141.38	116.20	4	6
D	46	ARG	CA-C-N	12.49	141.17	116.20	6	4
C	49	GLY	C-N-CA	12.48	144.17	121.70	4	3
C	188	ALA	CA-C-N	12.42	141.05	116.20	7	4
C	185	ILE	C-N-CA	12.41	99.37	121.70	1	1
C	307	LYS	C-N-CA	12.38	143.98	121.70	4	4
C	284	ASP	C-N-CA	12.37	143.96	121.70	6	5
D	33	ILE	CA-C-N	12.33	140.85	116.20	7	4
C	252	SER	N-CA-CB	12.32	131.45	110.50	3	3
C	87	GLU	C-N-CA	12.32	143.88	121.70	5	7
D	54	HIS	CA-CB-CG	12.32	101.48	113.80	7	1
C	280	LYS	N-CA-C	12.29	76.60	111.00	7	4
C	282	LEU	CA-C-N	12.28	140.76	116.20	2	2
C	309	GLU	N-CA-CB	12.27	131.36	110.50	7	2
A	401	HIS	C-N-CA	12.26	143.76	121.70	7	2
C	175	ALA	N-CA-CB	12.23	128.74	110.40	2	3
C	292	ASN	C-CA-CB	12.18	86.96	110.10	5	4

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	158.69	3513

Model ID	Clash score	Number of clashes
2	182.34	4036
3	158.34	3505
4	154.41	3418
5	204.24	4521
6	144.74	3204
7	227.11	5027

There are 27224 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)
C:284:ASP:C	C:284:ASP:CA	1.60	7	4
C:293:LYS:C	C:293:LYS:CA	1.59	4	3
E:1:MET:C	E:1:MET:CA	1.58	7	1
C:46:LEU:CB	C:46:LEU:CG	1.58	2	1
C:292:ASN:C	C:292:ASN:CA	1.58	1	6
C:391:TYR:C	C:391:TYR:CA	1.58	4	4
C:305:LYS:CA	C:305:LYS:N	1.57	5	2
D:38:ASP:C	D:38:ASP:CA	1.57	7	1
C:246:PHE:C	C:246:PHE:CA	1.57	2	1
C:250:SER:C	C:250:SER:CA	1.57	3	1
D:48:ARG:C	D:48:ARG:CA	1.57	6	1
A:401:HIS:CA	A:401:HIS:N	1.56	7	1
C:295:ASP:C	C:295:ASP:CA	1.56	3	2
C:278:ASN:C	C:278:ASN:CA	1.56	7	1
C:294:GLN:C	C:294:GLN:CA	1.56	3	1
C:280:LYS:C	C:280:LYS:CA	1.56	5	4
D:32:GLN:C	D:32:GLN:CA	1.55	6	2
C:58:LYS:C	C:58:LYS:CA	1.55	7	4
C:283:ARG:CA	C:283:ARG:N	1.55	1	2
C:279:ASN:C	C:279:ASN:CA	1.55	2	2
E:57:SER:C	E:57:SER:CA	1.55	3	1
C:285:THR:CA	C:285:THR:N	1.54	5	3
D:47:THR:C	D:47:THR:CA	1.54	6	2
C:83:ASN:CA	C:83:ASN:N	1.54	7	1
C:281:TRP:CA	C:281:TRP:N	1.54	4	2

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
C:305:LYS:C	C:305:LYS:CA	1.54	7	3
C:50:GLU:C	C:50:GLU:CA	1.54	1	4
C:286:SER:CA	C:286:SER:N	1.54	2	2
C:392:GLU:C	C:392:GLU:CA	1.53	4	2
C:187:GLN:C	C:187:GLN:CA	1.53	5	4
C:49:GLY:C	C:49:GLY:CA	1.53	4	1
C:51:SER:C	C:51:SER:CA	1.53	3	2
C:188:ALA:C	C:188:ALA:CA	1.53	5	3
C:84:SER:C	C:84:SER:CA	1.53	7	1
C:283:ARG:C	C:283:ARG:CA	1.53	1	2
C:51:SER:CA	C:51:SER:N	1.53	3	3
C:282:LEU:C	C:282:LEU:CA	1.53	1	2
C:293:LYS:CA	C:293:LYS:N	1.53	7	3
A:149:LEU:C	A:149:LEU:CA	1.53	5	1
C:286:SER:C	C:286:SER:CA	1.52	3	4
C:285:THR:C	C:285:THR:CA	1.52	2	2
C:87:GLU:C	C:87:GLU:CA	1.52	5	2
C:52:GLY:CA	C:52:GLY:N	1.52	3	2
C:261:ASN:C	C:261:ASN:CA	1.52	4	1
A:400:GLU:C	A:400:GLU:CA	1.52	7	1
C:306:SER:CA	C:306:SER:N	1.52	4	2
C:307:LYS:CA	C:307:LYS:N	1.52	4	2
C:49:GLY:CA	C:49:GLY:N	1.52	1	1
D:47:THR:CA	D:47:THR:N	1.51	6	1
C:248:VAL:C	C:248:VAL:CA	1.51	6	4
C:82:SER:C	C:82:SER:CA	1.51	7	1
D:48:ARG:CA	D:48:ARG:N	1.51	7	1
C:280:LYS:CA	C:280:LYS:N	1.51	4	1
C:189:ASP:CA	C:189:ASP:N	1.51	5	3
C:296:LEU:CA	C:296:LEU:N	1.51	4	1
C:306:SER:C	C:306:SER:CA	1.50	5	2
C:48:ALA:C	C:48:ALA:CA	1.50	1	1
C:304:GLY:C	C:304:GLY:CA	1.50	5	2
C:88:LYS:CA	C:88:LYS:N	1.50	7	2
C:58:LYS:C	C:59:GLN:N	1.50	4	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
C:261:ASN:CA	C:261:ASN:N	1.50	4	1
C:46:LEU:C	C:46:LEU:CA	1.50	2	1
C:307:LYS:C	C:307:LYS:CA	1.50	4	1
D:32:GLN:C	D:33:ILE:N	1.49	7	1
D:46:ARG:C	D:46:ARG:CA	1.49	6	1
C:50:GLU:CA	C:50:GLU:N	1.49	4	1
C:287:VAL:CA	C:287:VAL:N	1.49	2	3
C:393:LEU:CA	C:393:LEU:N	1.49	5	1
D:35:ASN:C	D:36:ASN:N	1.49	7	1
C:308:ILE:CA	C:308:ILE:N	1.49	4	1
C:286:SER:C	C:287:VAL:N	1.49	5	1
C:249:ALA:CA	C:249:ALA:N	1.48	2	3
D:34:THR:C	D:35:ASN:N	1.48	7	1
C:248:VAL:C	C:249:ALA:N	1.48	3	1
C:249:ALA:C	C:249:ALA:CA	1.48	4	2
C:284:ASP:CA	C:284:ASP:N	1.47	6	1
C:250:SER:CA	C:250:SER:N	1.47	4	2
C:187:GLN:CA	C:187:GLN:N	1.47	2	4
C:50:GLU:C	C:51:SER:N	1.46	2	1
C:282:LEU:C	C:283:ARG:N	1.45	2	1
D:33:ILE:C	D:34:THR:N	1.45	7	1
C:260:ASP:C	C:260:ASP:CA	1.44	4	1
D:38:ASP:CA	D:38:ASP:N	1.44	7	1
C:305:LYS:C	C:306:SER:N	1.44	5	1
C:188:ALA:C	C:189:ASP:N	1.44	7	1
D:37:ILE:CA	D:37:ILE:N	1.43	7	1
D:36:ASN:CA	D:36:ASN:N	1.43	7	1
D:37:ILE:C	D:37:ILE:CA	1.43	7	1
D:35:ASN:CA	D:35:ASN:N	1.42	7	1
C:186:LYS:C	C:186:LYS:CA	1.42	7	4
C:280:LYS:C	C:281:TRP:N	1.42	2	1
C:187:GLN:C	C:188:ALA:N	1.41	2	2
D:33:ILE:CA	D:33:ILE:N	1.39	7	1
C:286:SER:C	C:286:SER:CB	1.39	2	3
D:36:ASN:C	D:36:ASN:CA	1.39	7	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
C:59:GLN:CA	C:59:GLN:N	1.38	4	1
D:34:THR:CA	D:34:THR:N	1.37	7	1
D:34:THR:C	D:34:THR:CA	1.36	7	1
C:187:GLN:HA	C:187:GLN:N	1.34	1	2
D:35:ASN:C	D:35:ASN:CA	1.33	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	1393	1090	186	117
2	1393	1059	196	138
3	1393	1075	179	139
4	1393	1105	165	123
5	1393	1018	211	164
6	1393	1085	188	120
7	1393	1033	221	139

There are 507 unique backbone outliers. Detailed list of outliers are tabulated below. The output is limited to 100 rows.

Chain	Res	Type	Models (Total)
D	164	THR	7
A	84	GLU	6
B	19	ALA	6
C	283	ARG	6
D	189	SER	6
F	103	PRO	6
A	35	THR	5
A	105	GLU	5
A	189	HIS	5
A	411	LYS	5
C	183	ASP	5
C	214	VAL	5
C	251	SER	5
C	287	VAL	5
C	288	ILE	5
C	373	ARG	5

Chain	Res	Type	Models (Total)
D	153	ASP	5
F	102	ALA	5
A	47	ALA	4
A	103	CYS	4
A	183	SER	4
A	290	ILE	4
A	347	VAL	4
A	400	GLU	4
C	3	CYS	4
C	74	GLU	4
C	87	GLU	4
C	115	PRO	4
C	152	ALA	4
C	188	ALA	4
C	189	ASP	4
C	207	ILE	4
C	208	PHE	4
C	239	ASN	4
C	247	VAL	4
C	249	ALA	4
C	256	VAL	4
D	28	ALA	4
D	31	SER	4
D	34	THR	4
D	40	VAL	4
D	60	ALA	4
D	95	LEU	4
D	111	TYR	4
D	172	GLU	4
A	74	GLN	3
A	108	ARG	3
A	149	LEU	3
A	152	THR	3
A	162	ALA	3
A	177	ALA	3

Chain	Res	Type	Models (Total)
A	237	LEU	3
A	239	GLU	3
A	253	VAL	3
A	291	LEU	3
A	321	ASP	3
A	322	ILE	3
A	323	LYS	3
A	337	LEU	3
A	352	ALA	3
A	382	VAL	3
A	402	LEU	3
A	413	LEU	3
A	415	CYS	3
A	428	SER	3
C	52	GLY	3
C	59	GLN	3
C	60	MET	3
C	76	ASP	3
C	108	ALA	3
C	123	GLU	3
C	129	ASP	3
C	138	PRO	3
C	147	TYR	3
C	187	GLN	3
C	192	PRO	3
C	209	GLU	3
C	224	VAL	3
C	225	GLY	3
C	246	PHE	3
C	250	SER	3
C	252	SER	3
C	260	ASP	3
C	281	TRP	3
C	285	THR	3
C	289	LEU	3

Chain	Res	Type	Models (Total)
C	295	ASP	3
C	307	LYS	3
C	310	ASP	3
C	322	GLU	3
C	327	GLU	3
C	333	ARG	3
C	341	ILE	3
C	357	HIS	3
C	384	GLN	3
D	25	CYS	3
D	37	ILE	3
D	44	GLN	3
D	48	ARG	3
D	87	THR	3

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	1212	1040	98	74
2	1212	1032	87	93
3	1212	1037	115	60
4	1212	1024	99	89
5	1212	993	104	115
6	1212	1035	96	81
7	1212	1025	106	81

There are 345 unique sidechain outliers. Detailed list of outliers are tabulated below. The output is limited to 100 rows.

Chain	Res	Type	Models (Total)
A	66	LEU	7
A	222	LEU	6
A	373	LEU	5
C	91	LYS	5
C	99	LEU	5
C	181	LYS	5
C	270	LEU	5

Chain	Res	Type	Models (Total)
C	282	LEU	5
D	69	LEU	5
D	190	LEU	5
D	252	LEU	5
D	284	LEU	5
A	195	LEU	4
A	221	LEU	4
C	171	LEU	4
C	247	VAL	4
C	272	LEU	4
D	139	LEU	4
D	198	LEU	4
D	285	LEU	4
F	73	ARG	4
F	80	LEU	4
A	100	LEU	3
A	160	LEU	3
A	194	LEU	3
A	205	LEU	3
A	209	LEU	3
A	290	ILE	3
A	291	LEU	3
A	299	LEU	3
A	311	LYS	3
A	326	LEU	3
A	333	LEU	3
A	336	LEU	3
A	399	LEU	3
B	28	LYS	3
C	43	LEU	3
C	44	LEU	3
C	56	ILE	3
C	95	ILE	3
C	182	ILE	3
C	184	VAL	3

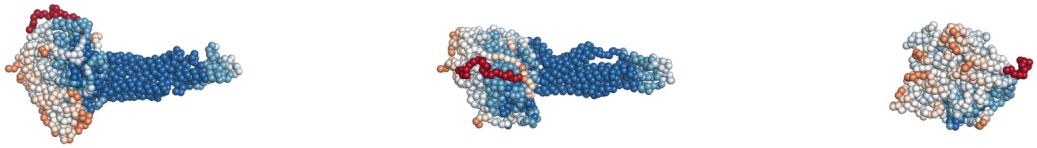
Chain	Res	Type	Models (Total)
C	197	LEU	3
C	198	LEU	3
C	201	ARG	3
C	207	ILE	3
C	232	ARG	3
C	235	ILE	3
C	244	ILE	3
C	265	ARG	3
C	288	ILE	3
C	291	LEU	3
C	293	LYS	3
C	297	LEU	3
C	300	LYS	3
C	394	LEU	3
D	79	LEU	3
D	126	LEU	3
D	146	LEU	3
D	192	LEU	3
D	261	LEU	3
E	65	LYS	3
F	5	LEU	3
F	19	LEU	3
F	59	ILE	3
F	71	ILE	3
F	87	LEU	3
F	99	ARG	3
A	27	LEU	2
A	37	LEU	2
A	41	ARG	2
A	86	LEU	2
A	95	LEU	2
A	117	GLN	2
A	121	LEU	2
A	156	ILE	2
A	166	LEU	2

Chain	Res	Type	Models (Total)
A	201	LEU	2
A	208	LEU	2
A	313	LYS	2
A	319	LYS	2
A	331	LEU	2
A	343	ILE	2
A	356	LEU	2
A	365	LEU	2
A	392	LYS	2
A	406	ARG	2
A	420	LEU	2
B	3	GLU	2
B	14	LEU	2
C	30	LEU	2
C	46	LEU	2
C	50	GLU	2
C	57	VAL	2
C	58	LYS	2
C	59	GLN	2
C	60	MET	2
C	82	SER	2
C	88	LYS	2
C	94	ASP	2

4.2. PrISM Precision Analysis

*Regions of **low** **high** precision, defined as the variability among the models that satisfy the input data and calculated as the density-weighted root mean-square fluctuation (RMSF) from the bead/atom center of density, annotated and visualized using PrISM. The per-bead precision is computed from the deposited ensemble of superposed integrative models. High- and low-precision regions are then determined by clustering beads of similar precision based on their proximity in the structure. Only coarse-grained beads (or CA atoms for atomic models) of deposited models are used for assessment and visualization, and three projections for each representative model are generated.*

PrISM analysis for Ensemble 1 (models deposited/total: 7/29999).



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.

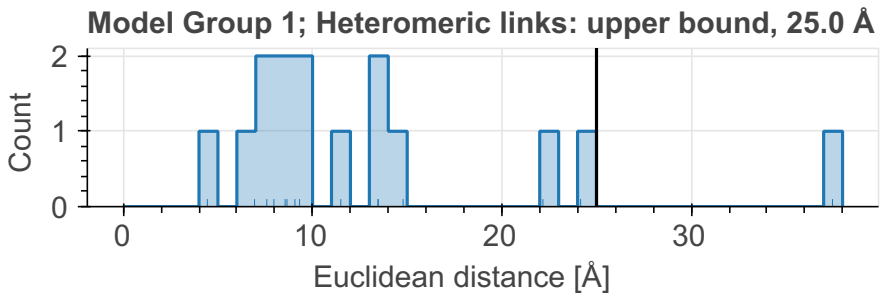
There are 597 crosslinking restraints combined in 597 restraint groups.

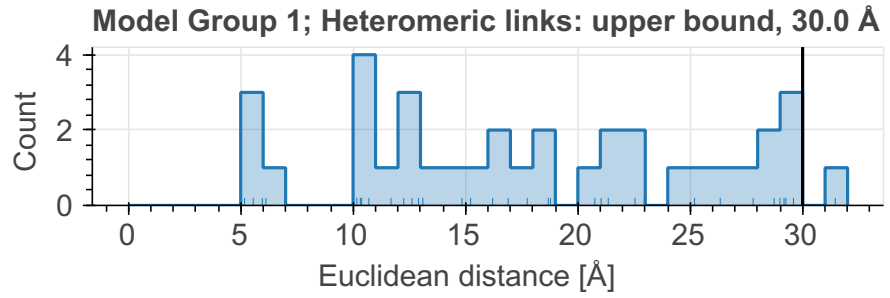
Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
BS3	LYS	CA	SER	CA	upper bound	35.00	67
BS3	LYS	CA	LYS	CA	upper bound	35.00	75
BS3	LYS	CA	TYR	CA	upper bound	35.00	7
BS3	THR	CA	TYR	CA	upper bound	35.00	1
BS3	GLY	CA	LYS	CA	upper bound	35.00	6
BS3	LYS	CA	THR	CA	upper bound	35.00	27
BS3	SER	CA	SER	CA	upper bound	35.00	10
BS3	SER	CA	THR	CA	upper bound	35.00	13
BS3	GLY	CA	SER	CA	upper bound	35.00	2
BS3	THR	CA	THR	CA	upper bound	35.00	4
BS3	ALA	CA	LYS	CA	upper bound	35.00	4
BS3	SER	CA	TYR	CA	upper bound	35.00	2
BS3	MET	CA	SER	CA	upper bound	35.00	2
BS3	LYS	CA	MET	CA	upper bound	35.00	4
BS3	MET	CA	THR	CA	upper bound	35.00	1
BS3	GLN	CA	MET	CA	upper bound	35.00	1
BS3	ALA	CA	MET	CA	upper bound	35.00	1
BS3	HIS	CA	THR	CA	upper bound	35.00	1
DSG	LYS	CA	SER	CA	upper bound	30.00	54

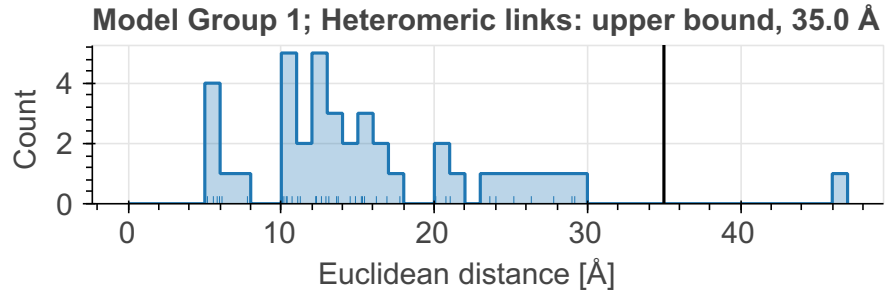
Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSG	LYS	CA	LYS	CA	upper bound	30.00	64
DSG	LYS	CA	THR	CA	upper bound	30.00	37
DSG	LYS	CA	TYR	CA	upper bound	30.00	9
DSG	SER	CA	THR	CA	upper bound	30.00	12
DSG	SER	CA	TYR	CA	upper bound	30.00	1
DSG	THR	CA	THR	CA	upper bound	30.00	4
DSG	ALA	CA	LYS	CA	upper bound	30.00	3
DSG	SER	CA	SER	CA	upper bound	30.00	5
DSG	MET	CA	SER	CA	upper bound	30.00	3
DSG	ALA	CA	SER	CA	upper bound	30.00	1
DSG	LYS	CA	MET	CA	upper bound	30.00	4
DSG	MET	CA	THR	CA	upper bound	30.00	1
DSG	GLY	CA	LYS	CA	upper bound	30.00	4
DSG	THR	CA	TYR	CA	upper bound	30.00	1
EDC	ASP	CA	LYS	CA	upper bound	25.00	49
EDC	GLU	CA	LYS	CA	upper bound	25.00	73
EDC	ASP	CA	MET	CA	upper bound	25.00	1
EDC	GLU	CA	HIS	CA	upper bound	25.00	1
EDC	GLU	CA	GLY	CA	upper bound	25.00	2
KArGO	ARG	CA	GLY	CA	upper bound	40.00	6
KArGO	ARG	CA	LYS	CA	upper bound	40.00	34

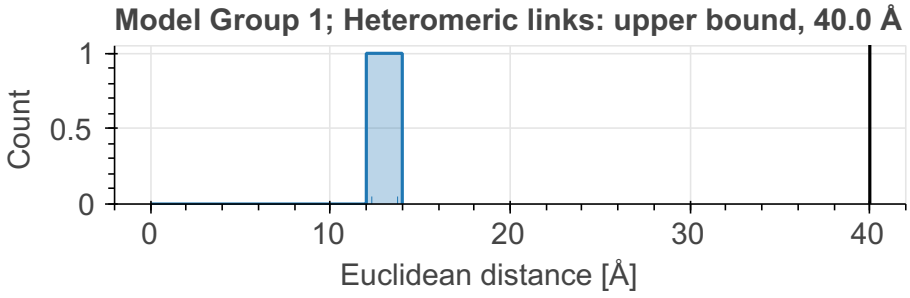
Distograms of individual restraints

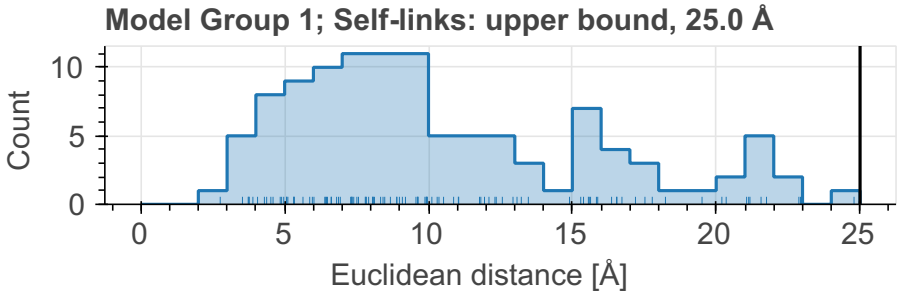
Distograms (i.e., histogram plots of distances) provide an overview of distributions of distances between residues for which chemical crosslinks were identified. The shift of the distogram relative to the threshold value may indicate a poor model. Restraints with identical thresholds are grouped into 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.

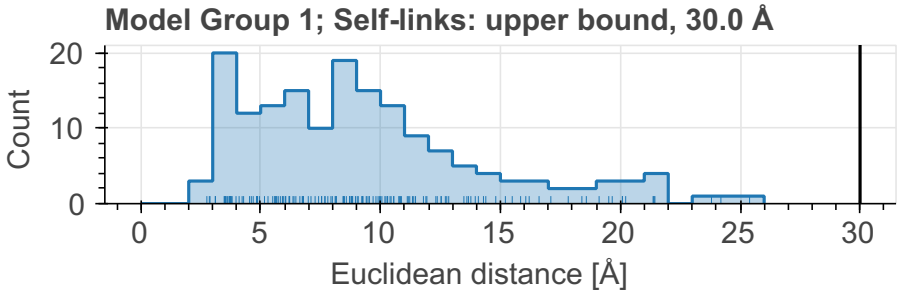


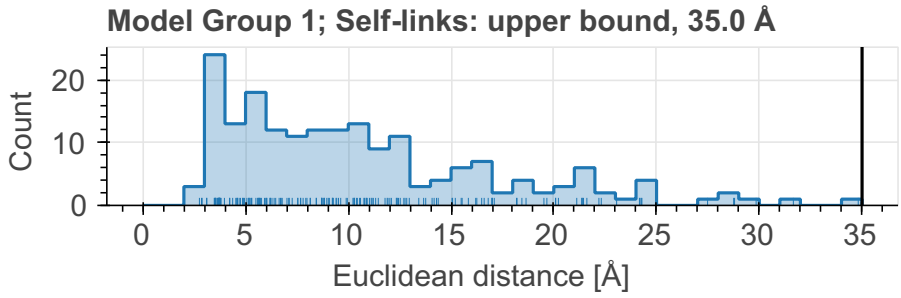

Model Group 1; Heteromeric links: upper bound, 30.0 Å
Count
Euclidean distance [Å]

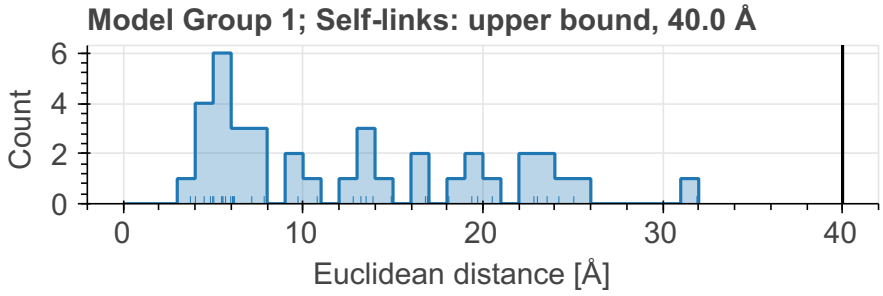

Model Group 1; Heteromeric links: upper bound, 35.0 Å
Count
Euclidean distance [Å]


Model Group 1; Heteromeric links: upper bound, 40.0 Å
Count
Euclidean distance [Å]


Model Group 1; Self-links: upper bound, 25.0 Å
Count
Euclidean distance [Å]


Model Group 1; Self-links: upper bound, 30.0 Å
Count
Euclidean distance [Å]





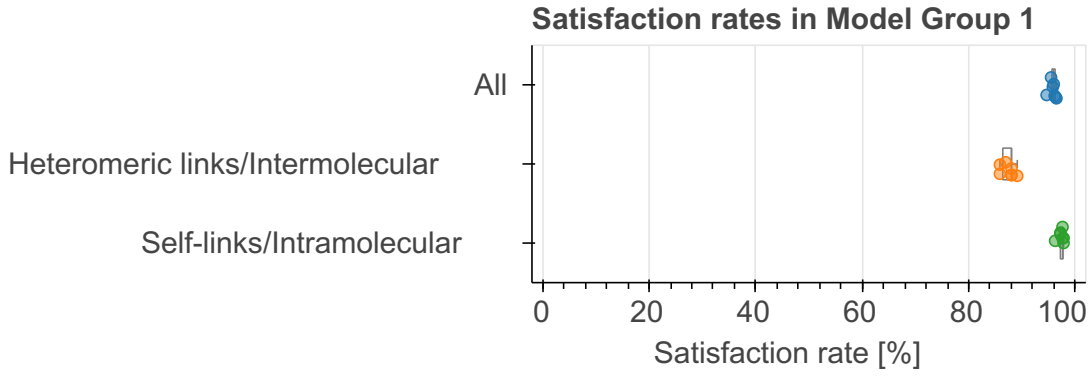
5.2.2. Satisfaction of restraints

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=597)
1	1	1	7/29999	All	99.50	0.50	597
				Heteromeric links/ Intermolecular	96.74	3.26	92
				Self-links/ Intramolecular	100.00	0.00	505

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.



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.