

Integrative Structure Validation Report

October 09, 2025 - 04:37 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)

PrISM Version dbe5a41

PyMOL Version 2.5.0

PDB ID	8ZZZ pdb_00008zzz
PDB-Dev ID	PDBDEV_00000035
Structure Title	Rosetta model of human LRH-1 nuclear receptor generated with XL-MS, HDX-MS, and SAXS data
Structure Authors	Seacrist CD; Kuenze G; Hoffmann R; Burke J; Meiler J; Blind RD
Deposited on	2019-09-13

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 11 model(s). A total of 9 dataset(s) were used to build this entry.

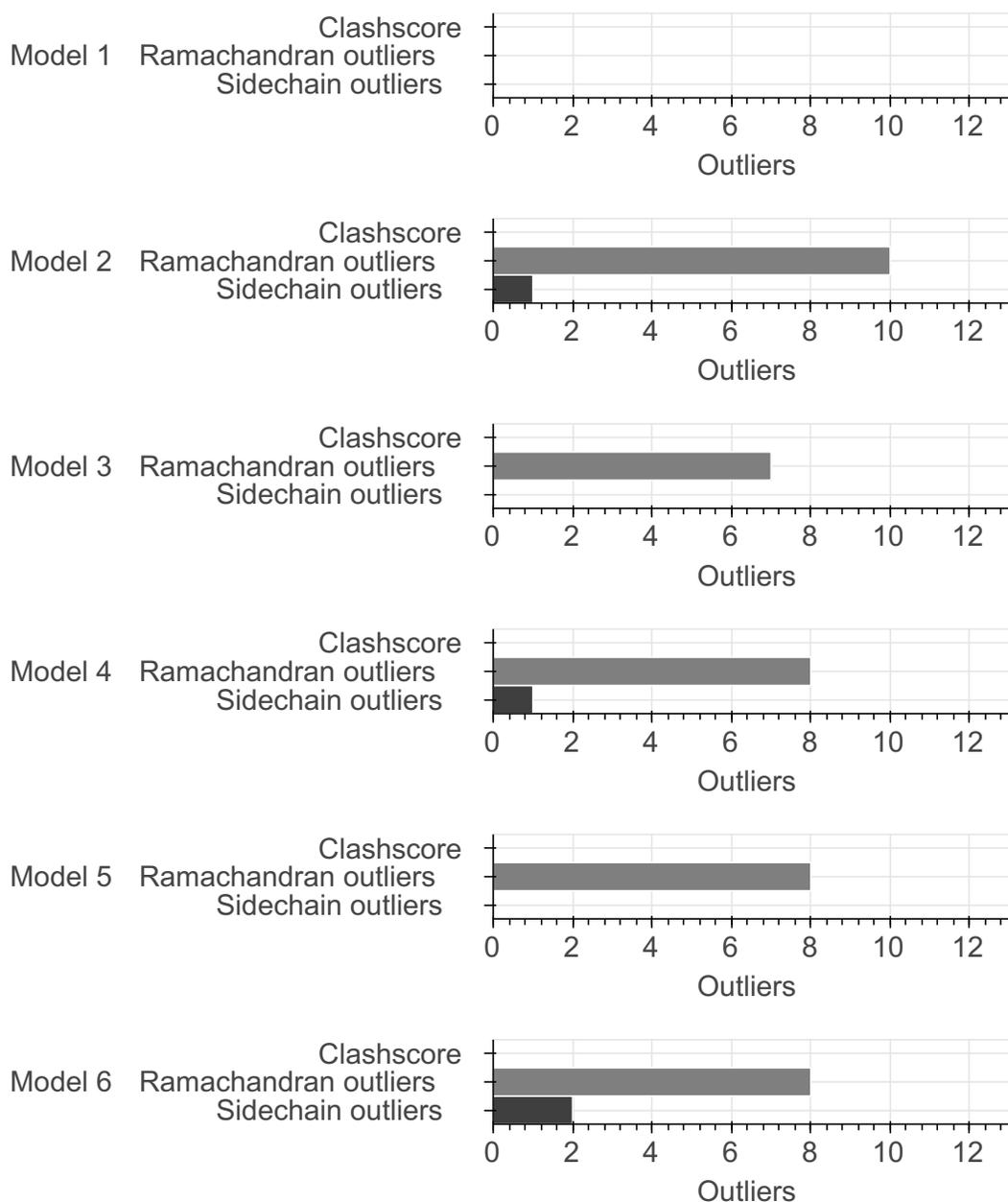
Name	Type	Count
Crosslinking-MS data	Experimental data	2

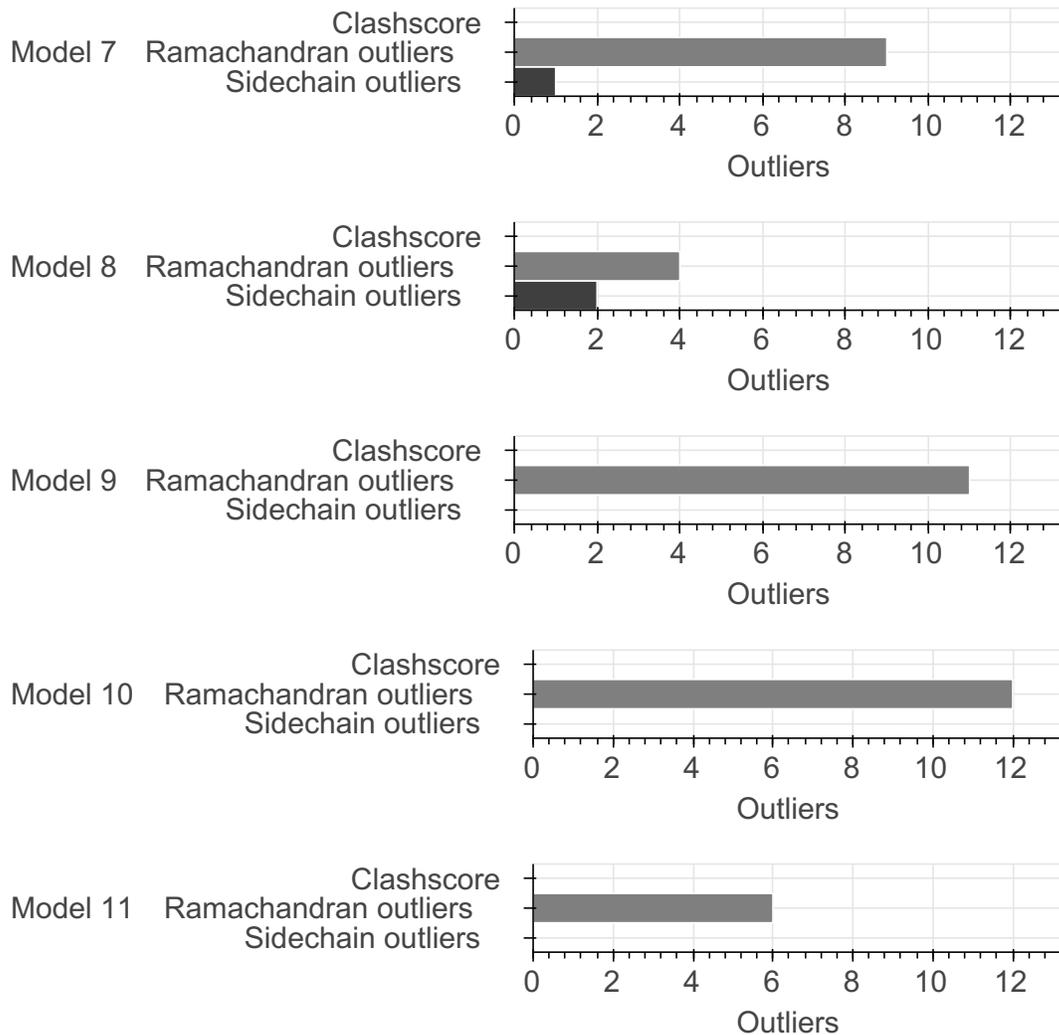
Name	Type	Count
H/D exchange data	Experimental data	1
SAS data	Experimental data	1
Comparative model	Starting model	2
Integrative model	Starting model	1
Experimental model	Starting model	2

1.2. Overall quality ?

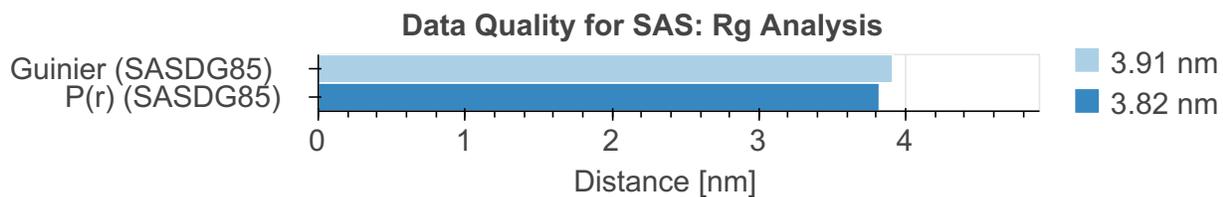
This validation report contains model quality assessments for all structures, data quality and fit to model assessments for SAS and crosslinking-MS datasets. Data quality and fit to model assessments for other datasets and model uncertainty are under development. Number of plots is limited to 256.

Model Quality: MolProbity Analysis ?

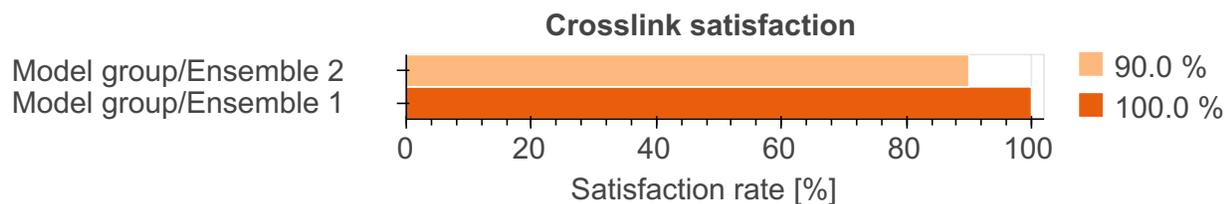




Data Quality ?



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 2 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	1	LRH-1 DNA-binding domain	A	102	-	1-102	100.00 / 100.00	Atomic
		2	LRH-1 Ligand-binding domain	B	240	-	1-240	100.00 / 100.00	Atomic
		4	PGC1-alpha coactivator peptide	D	10	-	1-10	100.00 / 100.00	Atomic
		5	DNA strand 1	E	12	1-12	-	100.00 / 100.00	Atomic
		6	DNA strand 2	F	12	1-12	-	100.00 / 100.00	Atomic
		7	Phospholipid ligand	X	Non-polymeric	-	-	Not available / Not available	Atomic
		8	Zinc ion	Y	Non-polymeric	-	-	Not available / Not available	Atomic
				Z					
2	2-11	3	LRH-1 full length polypeptide	C	538	-	1-538	100.00 / 100.00	Atomic
		4	PGC1-alpha coactivator peptide	D	10	1-10	-	100.00 / 100.00	Atomic
		5	DNA strand 1	E	12	1-12	-	100.00 / 100.00	Atomic
		6	DNA strand 2	F	12	1-12	-	100.00 / 100.00	Atomic
		7	Phospholipid ligand	X	Non-polymeric	-	-	Not available / Not available	Atomic
		8	Zinc ion	Y	Non-polymeric	-	-	Not available / Not available	Atomic
Z									

2.3. Datasets used for modeling

There are 9 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	Experimental model	PDB	pdb_00002a66
2	Experimental model	PDB	pdb_00001yok

ID	Dataset type	Database name	Data access code
3	Comparative model	Zenodo	10.5281/zenodo.3405545
4	Comparative model	Zenodo	10.5281/zenodo.3405545
5	Crosslinking-MS data	Zenodo	10.5281/zenodo.3405545
6	H/D exchange data	Zenodo	10.5281/zenodo.3405545
7	Integrative model	Zenodo	10.5281/zenodo.3405545
8	Crosslinking-MS data	Zenodo	10.5281/zenodo.3405545
9	SAS data	SASBDB	SASDG85

2.4. Methodology and software ?

This entry is a result of 2 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	Loop modeling	RosettaRemodel	Not available	100	False	False
2	1	Docking	RosettaDock	Not available	40000	False	False
1	2	Linker modeling	Ranch and RosettaMinimize	Not available	45000	False	False

There are 3 software packages reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	Rosetta	3.10	model building, model validation	https://www.rosettacommons.org/
2	ATSAS	2.8.4	model building, model validation	https://www.embl-hamburg.de/biosaxs/software.html
3	REMO	1	model building	https://zhanglab.ccmb.med.umich.edu/REMO/

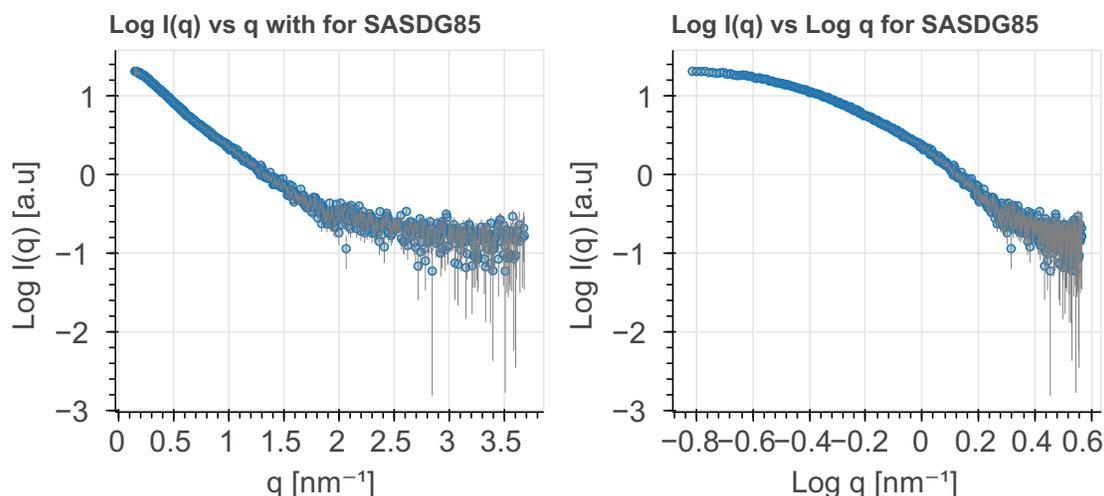
3. Data quality ?

3.1. SAS ?

3.1.1. Scattering profile ?

SAS data used in this integrative model was obtained from 1 deposited SASBDB entry (entries).

Scattering profile for [SASDG85](#): data from solutions of biological macromolecules are presented as both $\log I(q)$ vs q and $\log I(q)$ vs $\log(q)$ based on [SAS validation task force \(SASvtf\) recommendations](#). $I(q)$ is the intensity (in arbitrary units) and q is the modulus of the scattering vector.



3.1.2. Key experimental estimates ?

Molecular weight (MW) estimates from experiments and analysis: Theoretical MW can be compared to SAS-derived values using the forward scatter ($I(0)$) and the known concentration and partial specific volume of the scattering particle, or as estimated from the Porod volume and partial specific volume (Trehella et al., 2017, Trehella et al., 2023).

SASDB ID	Chemical composition MW	Standard MW	Porod Volume/MW
SASDG85	67.0 kDa	Not available	Not available

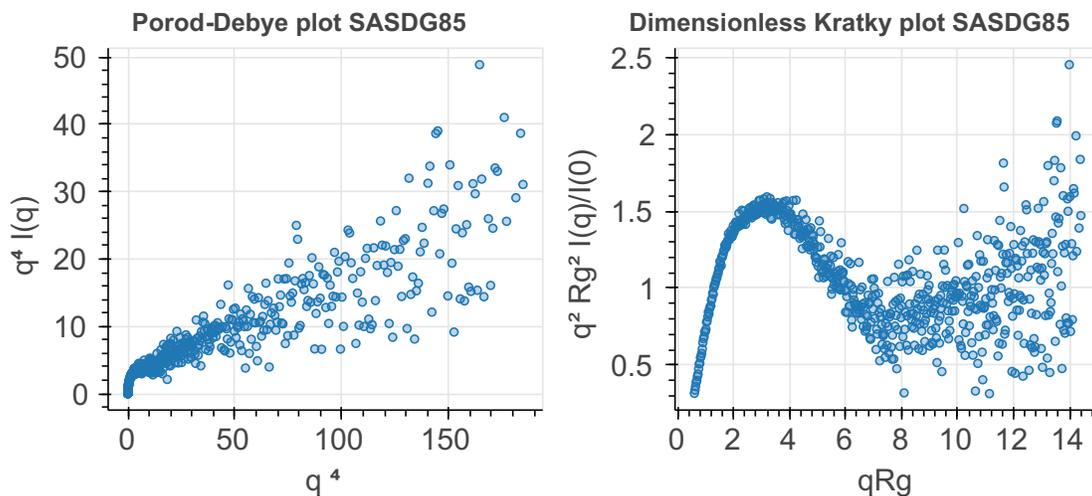
Volume estimates from experiments and analysis: estimated volume can be compared to Porod volume obtained from scattering profiles.

SASDB ID	Estimated Volume	Porod Volume	Specific Volume	Sample Contrast	Sample Concentration
SASDG85	Not available	76.00 nm ³	Not available	Not available	7.50 mg/mL

3.1.3. Flexibility analysis ?

In a Porod-Debye plot, a clear plateau is observed for globular (partial or fully folded) domains, whereas flexible-modular, fully unfolded domains or extended/stiff rodshaped domains lack a discernible plateau (Rambo and Tainer 2013). A bell-shaped Kratky plot ($q^2I(q)$ vs. q) with a well-defined maximum is observed for compact/folded structures. For partially flexible/modular or extended structures the Kratky plot can show multiple maxima and/or an increase in intensity at higher q -values depending on the degree of flexibility and extension. Fully intrinsically disordered structures yield a Kratky plot that systematically increases with increasing q values and will be near linear for highly extended molecules. The dimensionless Kratky plot ($(qR_g)^2I(q)$ vs. qR_g) is useful for quantifying differences in shape and foldedness among scattering objects of different sizes (Trehella et al., 2023).

Flexibility analysis for SASDG85

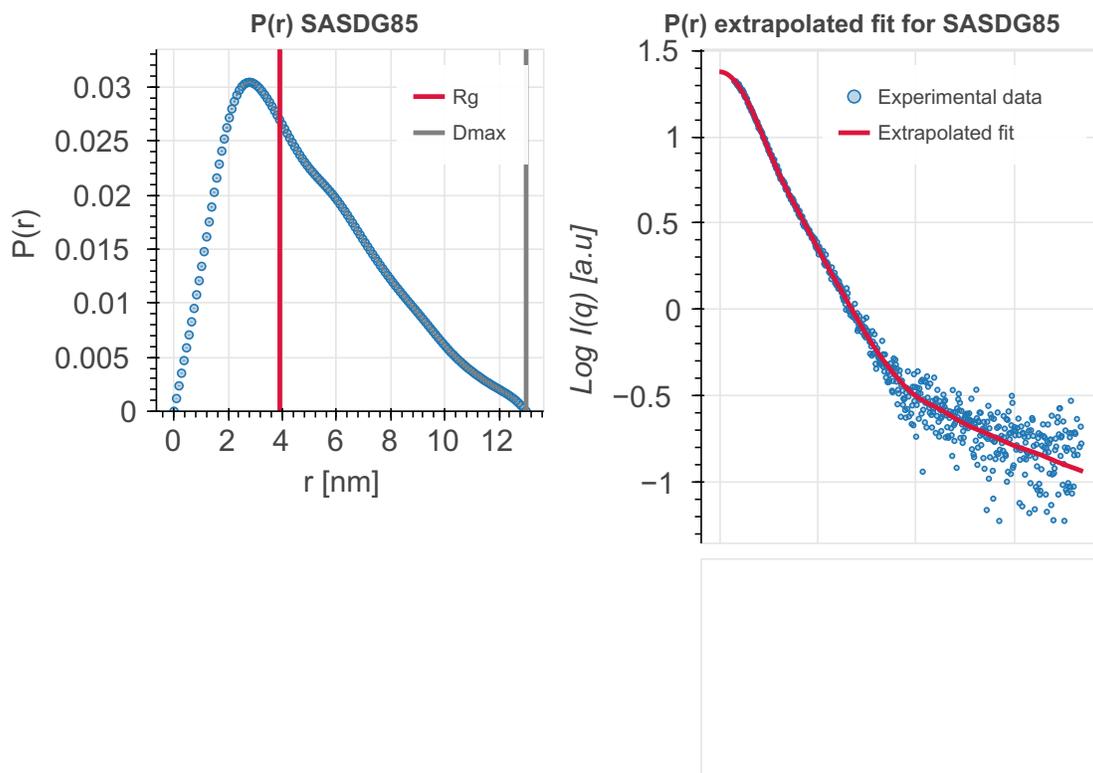


3.1.4. Pair-distance distribution analysis ?

The atom-pair distance distribution function (PDDF) or $P(r)$ represents the distribution of distances between all pairs of atoms within the particle weighted by the respective scattering contrasts (Moore, 1980). The second moment of $P(r)$ yields the radius of gyration (R_g), which is a measure of the overall size and shape of a macromolecule (i.e. the spatial distribution of volume elements). A protein with a smaller R_g is more compact than a protein with a larger R_g , provided both have the same molecular weight.

SASDB ID	Software used	D_{\max}	D_{\max} error	R_g	R_g error
SASDG85	GNOM 4.6	13.000 nm	Not available	3.907 nm	0.020 nm

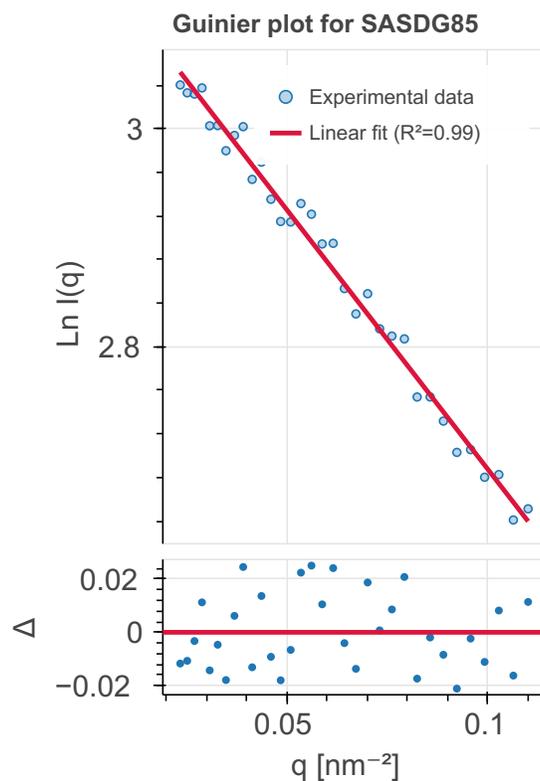
P(r) for SASDG85: The value of $P(r)$ should be zero beyond $r=D_{\max}$.



3.1.5. Guinier analysis ?

The linearity of the Guinier plot ($\ln(q)$ vs. q^2) at very-low angle ($qR_g < 1.3$) is a sensitive indicator of the quality of the sample in relation to its homogeneity; a linear Guinier plot is a necessary but not sufficient demonstration that a solution contains monodisperse particles of the same size. Deviations from linearity can point to strong interference effects from particle attraction or repulsion, polydispersity of the samples, or improper background subtraction (Feigin et al., 2013). Residual difference plots and Pearson correlation coefficient determination (R^2) are measures to assess quality of the linear fit to the Guinier region. A perfect fit has an R^2 value of 1. Residual values should be equally and randomly spaced around the horizontal axis with no evident systematic upward or downward curvature.

SASDB ID	R_g	R_g error	MW	MW error
SASDG85	3.82 nm	0.05 nm	Not available	Not available



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-particle or particle-atom pairs for which excluded volume was analysed.

Standard geometry: bond outliers

There are 251 bond length outliers in this entry (0.47% of 53677 assessed bonds). A summary is provided below.

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
X	1	P6L	C39-C40	540.30	12.33	1.52	2	11
X	1	P6L	C2-O3	401.94	9.45	1.41	3	11
X	1	P6L	C14-O4	397.30	9.28	1.33	5	11
X	1	P6L	C14-C18	387.59	9.26	1.50	5	11
X	1	P6L	C16-O8	286.46	7.06	1.33	10	11
X	1	P6L	C34-C35	240.84	6.31	1.50	11	11
X	1	P6L	C26-C37	208.08	5.68	1.52	7	11
X	1	P6L	C27-C28	189.36	5.31	1.52	11	11
X	1	P6L	C5-C7	173.40	4.98	1.52	11	11
X	1	P6L	C5-C6	95.06	3.41	1.51	10	11
C	44	LYS	CB-CG	94.02	4.34	1.52	6	3
X	1	P6L	C36-C38	49.46	2.51	1.52	2	11
C	43	PRO	N-CD	17.06	1.71	1.47	8	3
C	79	ASP	C-N	16.53	1.56	1.33	4	2
C	34	PRO	N-CD	13.22	1.66	1.47	9	1
C	298	PRO	C-N	12.84	1.51	1.33	2	1
C	80	GLU	N-CA	12.47	1.69	1.46	4	2
C	298	PRO	CG-CD	11.97	1.91	1.50	2	1
C	82	LEU	CG-CD2	11.08	1.16	1.52	4	2
X	1	P6L	C6-O8	10.43	1.21	1.42	2	11
C	78	TYR	CB-CG	10.08	1.73	1.51	4	1
X	1	P6L	C22-C23	9.50	1.52	1.33	8	11
C	44	LYS	CG-CD	9.48	1.80	1.52	6	2
C	44	LYS	CA-CB	9.39	1.72	1.53	6	4
C	298	PRO	N-CD	9.27	1.60	1.47	2	1
X	1	P6L	C33-C34	8.98	1.51	1.33	3	11
X	1	P6L	C7-O9	8.50	1.24	1.41	7	11
C	76	TYR	CE2-CZ	8.47	1.17	1.38	11	1

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
C	250	PRO	N-CD	8.34	1.59	1.47	8	1
C	76	TYR	CG-CD2	8.21	1.22	1.39	11	1
C	46	GLU	CB-CG	7.96	1.28	1.52	3	2
C	43	PRO	C-N	7.41	1.43	1.33	2	5
C	82	LEU	CB-CG	7.03	1.67	1.53	4	1
C	231	PRO	N-CD	6.76	1.57	1.47	10	1
C	298	PRO	CB-CG	6.72	1.83	1.49	2	1
C	265	PRO	N-CD	6.68	1.38	1.47	11	1
C	83	GLU	CB-CG	6.27	1.33	1.52	3	2
C	76	TYR	CE1-CZ	6.11	1.23	1.38	11	1
C	78	TYR	CE2-CZ	6.09	1.23	1.38	4	1
C	84	GLU	C-N	5.96	1.41	1.33	8	1
C	76	TYR	CD1-CE1	5.88	1.56	1.38	11	1
C	83	GLU	CG-CD	5.87	1.37	1.52	4	2
C	81	ASP	CB-CG	5.56	1.65	1.52	3	1
C	106	CYS	CB-SG	5.54	1.99	1.81	11	10
C	204	PRO	N-CD	5.51	1.55	1.47	5	2
A	24	CYS	CB-SG	5.51	1.99	1.81	1	1
C	44	LYS	N-CA	5.49	1.56	1.46	2	5
C	83	GLU	CD-OE2	5.42	1.15	1.25	3	1
C	276	PRO	N-CD	5.15	1.55	1.47	2	1
C	78	TYR	CE1-CZ	5.09	1.26	1.38	4	1
C	239	PRO	N-CD	5.05	1.40	1.47	8	1
C	76	TYR	CG-CD1	4.97	1.28	1.39	11	1
C	78	TYR	CG-CD1	4.96	1.28	1.39	4	1
C	43	PRO	N-CA	4.96	1.54	1.47	2	1
C	83	GLU	C-N	4.95	1.40	1.33	3	1
C	244	PRO	N-CD	4.87	1.54	1.47	7	2
C	78	TYR	CG-CD2	4.77	1.49	1.39	4	1
C	44	LYS	CA-C	4.60	1.62	1.52	6	3
C	41	MET	C-N	4.57	1.39	1.33	2	1
C	42	LEU	CA-C	4.49	1.62	1.52	2	1
C	44	LYS	C-N	4.43	1.39	1.33	3	1
C	34	PRO	N-CA	4.36	1.53	1.47	9	1
C	84	GLU	N-CA	4.21	1.54	1.46	3	1

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
C	32	PRO	N-CD	4.17	1.41	1.47	8	1
C	76	TYR	CB-CG	4.05	1.60	1.51	11	1
C	83	GLU	CA-C	4.00	1.61	1.52	3	1

Standard geometry: angle outliers

There are 522 bond angle outliers in this entry (0.71% of 73409 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	79	ASP	C-N-CA	24.23	165.32	121.70	4	3
C	298	PRO	C-CA-CB	23.44	65.56	110.10	2	1
C	18	LEU	C-N-CA	22.58	162.34	121.70	10	1
C	82	LEU	C-N-CA	22.23	161.71	121.70	5	3
C	209	ILE	CA-CB-CG1	22.11	72.82	110.40	2	1
X	1	P6L	C18-C14-O4	21.78	45.83	111.16	4	11
C	81	ASP	C-N-CA	21.09	159.66	121.70	11	1
X	1	P6L	C1-C2-O3	20.99	49.51	112.48	4	11
C	20	PRO	C-N-CA	20.82	159.17	121.70	10	1
C	54	ARG	C-N-CA	19.98	157.67	121.70	9	1
C	298	PRO	N-CA-C	19.94	62.25	112.10	2	1
C	37	GLY	C-N-CA	19.72	157.19	121.70	5	1
X	1	P6L	C27-C16-O8	19.53	52.78	111.38	6	11
X	1	P6L	C27-C28-C29	19.33	54.14	112.12	5	11
C	7	THR	C-N-CA	19.31	156.46	121.70	8	1
C	64	GLU	C-N-CA	19.10	156.08	121.70	5	1
X	1	P6L	C6-C5-C7	19.07	57.83	115.05	10	11
C	21	ILE	C-N-CA	19.03	155.95	121.70	10	1
C	221	LYS	C-N-CA	19.02	155.93	121.70	9	1
C	57	GLY	C-N-CA	18.52	155.04	121.70	11	1
C	298	PRO	N-CA-CB	18.39	123.23	103.00	2	1
C	237	ARG	C-N-CA	17.97	154.04	121.70	4	1
C	277	TYR	C-N-CA	17.83	153.79	121.70	4	1
C	213	ILE	C-N-CA	17.34	152.91	121.70	5	1
C	258	TYR	C-N-CA	17.30	152.84	121.70	3	2
C	209	ILE	CA-CB-CG2	16.77	81.99	110.50	2	1
C	24	GLY	C-N-CA	16.68	151.72	121.70	11	1
C	32	PRO	C-N-CA	16.58	151.54	121.70	7	2

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
C	185	LEU	C-N-CA	16.56	151.52	121.70	7	1
X	1	P6L	C16-C27-C28	16.24	64.84	113.57	1	11
C	80	GLU	C-CA-CB	16.08	79.55	110.10	3	2
C	209	ILE	CG1-CB-CG2	14.65	154.64	110.70	2	1
C	81	ASP	CA-CB-CG	14.41	127.01	112.60	3	2
C	34	PRO	CA-N-CD	13.86	92.60	112.00	9	1
C	80	GLU	N-CA-C	13.84	149.76	111.00	4	3
X	1	P6L	C25-C26-C37	13.03	152.46	113.38	5	11
X	1	P6L	C36-C38-C39	12.46	75.64	113.03	10	11
C	83	GLU	C-N-CA	12.40	144.03	121.70	8	1
C	298	PRO	C-N-CA	11.84	100.39	121.70	9	2
X	1	P6L	C7-C5-O4	11.68	147.31	112.28	10	11
C	53	ALA	C-CA-CB	11.58	93.14	110.50	9	1
C	79	ASP	CA-C-N	11.29	138.78	116.20	4	1
X	1	P6L	C35-C36-C38	11.01	146.18	113.14	7	11
C	33	ILE	C-N-CD	10.51	81.90	125.00	9	1
C	43	PRO	CA-N-CD	10.50	97.30	112.00	8	1
C	44	LYS	CA-CB-CG	10.48	135.05	114.10	8	5
C	33	ILE	C-N-CA	10.36	174.42	122.60	9	1
X	1	P6L	C33-C34-C35	10.28	94.60	125.45	7	11
C	44	LYS	N-CA-CB	10.14	93.27	110.50	10	2
C	82	LEU	CD1-CG-CD2	10.08	132.97	110.80	7	2
C	298	PRO	CA-CB-CG	9.77	85.94	104.50	2	1
X	1	P6L	C16-O8-C6	9.51	89.53	118.05	8	11
C	76	TYR	CB-CG-CD1	9.33	134.79	120.80	11	1
C	242	THR	OG1-CB-CG2	9.00	127.30	109.30	8	1
C	44	LYS	CB-CG-CD	8.87	131.70	111.30	6	2
C	260	THR	OG1-CB-CG2	8.78	126.85	109.30	11	1
C	44	LYS	C-CA-CB	8.71	93.54	110.10	8	4
C	78	TYR	C-N-CA	8.71	137.37	121.70	4	1
C	44	LYS	C-N-CA	8.66	137.28	121.70	4	5
C	199	VAL	CA-CB-CG1	8.63	125.07	110.40	10	1
C	82	LEU	N-CA-CB	8.41	96.21	110.50	4	2
C	84	GLU	C-N-CA	8.36	106.65	121.70	8	2
C	83	GLU	N-CA-C	8.32	134.31	111.00	4	2

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
C	79	ASP	CA-C-O	8.25	106.77	120.80	4	1
C	41	MET	C-N-CA	8.07	136.23	121.70	3	2
C	68	VAL	CA-CB-CG2	7.89	123.82	110.40	6	1
C	82	LEU	CB-CG-CD2	7.88	87.05	110.70	3	1
C	233	THR	OG1-CB-CG2	7.87	125.03	109.30	8	1
C	84	GLU	CA-C-N	7.76	131.71	116.20	8	1
C	42	LEU	N-CA-C	7.58	132.23	111.00	2	2
C	287	TYR	C-N-CA	7.45	135.11	121.70	5	1
C	82	LEU	CA-C-O	7.36	108.29	120.80	4	1
C	83	GLU	CB-CG-CD	7.28	100.22	112.60	3	2
C	80	GLU	N-CA-CB	7.21	122.76	110.50	4	1
C	84	GLU	CA-C-O	7.18	108.60	120.80	8	1
C	84	GLU	N-CA-C	7.07	130.81	111.00	8	2
C	191	LEU	C-N-CA	7.07	134.42	121.70	2	1
C	6	ASP	CA-CB-CG	7.04	119.64	112.60	3	1
C	76	TYR	CB-CG-CD2	6.91	110.44	120.80	11	1
C	67	GLN	C-N-CA	6.89	134.09	121.70	6	1
C	298	PRO	CA-N-CD	6.87	102.39	112.00	2	1
C	45	VAL	C-CA-CB	6.83	98.42	111.40	3	1
X	1	P6L	C34-C35-C36	6.81	132.71	112.27	7	11
C	202	ALA	C-N-CA	6.73	133.82	121.70	9	1
C	240	PHE	CA-CB-CG	6.62	120.42	113.80	6	1
C	79	ASP	CA-CB-CG	6.59	119.19	112.60	4	1
C	294	GLN	C-N-CA	6.57	133.53	121.70	4	1
C	242	THR	C-N-CA	6.56	133.51	121.70	10	1
C	83	GLU	N-CA-CB	6.51	99.43	110.50	3	1
C	208	THR	C-N-CA	6.51	133.42	121.70	6	1
C	34	PRO	N-CD-CG	6.49	112.94	103.20	9	1
C	210	SER	C-N-CA	6.42	133.26	121.70	11	4
C	46	GLU	C-N-CA	6.15	132.77	121.70	3	2
C	82	LEU	N-CA-C	6.13	128.16	111.00	4	2
C	62	MET	CA-C-N	6.10	126.05	116.90	5	1
C	43	PRO	N-CA-C	6.02	127.14	112.10	6	3
C	76	TYR	CA-CB-CG	5.95	124.61	113.90	11	1
X	1	P6L	C38-C39-C40	5.92	95.57	113.33	2	11

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
C	72	LYS	C-N-CA	5.88	132.29	121.70	7	1
C	34	PRO	N-CA-CB	5.88	109.47	103.00	9	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	0.00	0
2	0.00	0
3	0.00	0
4	0.00	0
5	0.00	0
6	0.00	0
7	0.00	0
8	0.00	0
9	0.00	0
10	0.00	0
11	0.00	0

There are no too-close contacts.

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	346	340	6	0
2	544	500	34	10
3	544	511	26	7
4	544	507	29	8
5	544	505	31	8
6	544	509	27	8
7	544	500	35	9
8	544	511	29	4
9	544	497	36	11
10	544	494	38	12
11	544	508	30	6

There are 54 unique backbone outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
C	44	LYS	4
C	249	MET	4
C	19	THR	3
C	45	VAL	3
C	200	ILE	3
C	203	MET	3
C	31	SER	2
C	41	MET	2
C	43	PRO	2
C	47	THR	2
C	48	GLU	2
C	62	MET	2
C	68	VAL	2
C	81	ASP	2
C	82	LEU	2
C	186	ILE	2
C	216	ILE	2
C	241	VAL	2
C	254	SER	2
C	265	PRO	2
C	287	TYR	2
C	7	THR	1
C	14	LEU	1
C	15	LYS	1
C	21	ILE	1
C	34	PRO	1
C	40	VAL	1
C	42	LEU	1
C	55	SER	1
C	60	GLY	1
C	74	VAL	1
C	83	GLU	1
C	84	GLU	1
C	199	VAL	1
C	201	GLN	1

Chain	Res	Type	Models (Total)
C	205	SER	1
C	208	THR	1
C	209	ILE	1
C	210	SER	1
C	213	ILE	1
C	217	HIS	1
C	230	LEU	1
C	231	PRO	1
C	242	THR	1
C	262	GLY	1
C	264	PHE	1
C	267	ARG	1
C	269	ILE	1
C	271	SER	1
C	275	ASP	1
C	283	SER	1
C	284	ILE	1
C	290	MET	1
C	298	PRO	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	315	312	3	0
2	484	479	4	1
3	484	480	4	0
4	484	478	5	1
5	484	481	3	0
6	484	477	5	2
7	484	478	5	1
8	484	478	4	2
9	484	480	4	0
10	484	478	6	0
11	484	479	5	0

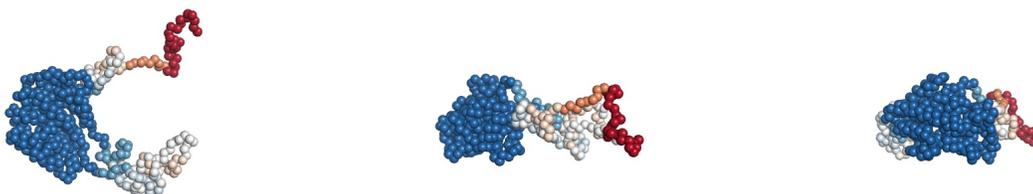
There are 7 unique sidechain outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
C	43	PRO	1
C	68	VAL	1
C	82	LEU	1
C	239	PRO	1
C	250	PRO	1
C	278	THR	1
C	298	PRO	1

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: 10/10).



5. Fit to Data Used for Modeling Assessment ?

5.1. SAS ?

Model and fits displayed below were obtained from SASBDB.

Model(s) and/or fit for this entry have not been deposited or sasCIF file is incomplete.

5.2. Crosslinking-MS ?

5.2.1. Restraint types ?

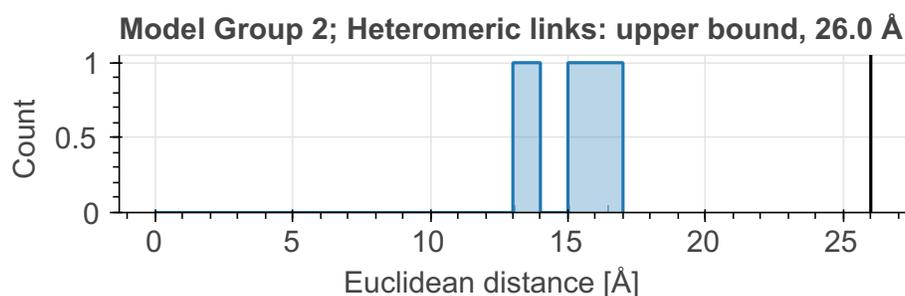
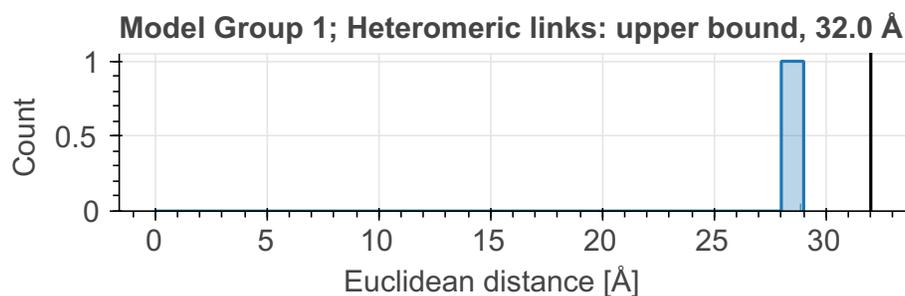
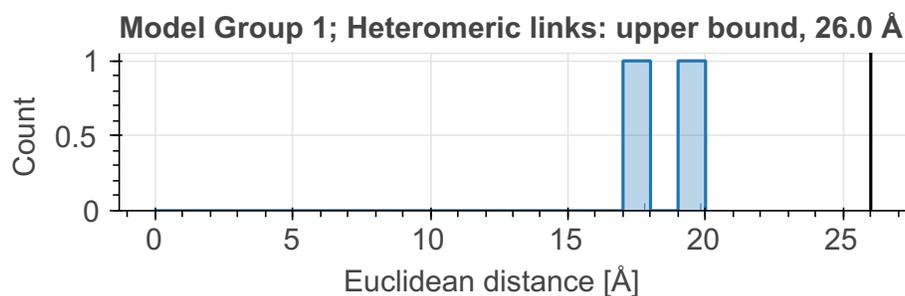
This table summarizes information about crosslinker(s) used for data generation, and how crosslinking information was translated into actual modeling restraints. Restraints assigned "by-residue" are interpreted as between CA atoms. Restraints between coarse-grained beads are indicated as "coarse-grained". *Restraint group* represents a set of crosslinking restraints applied collectively in the modeling.

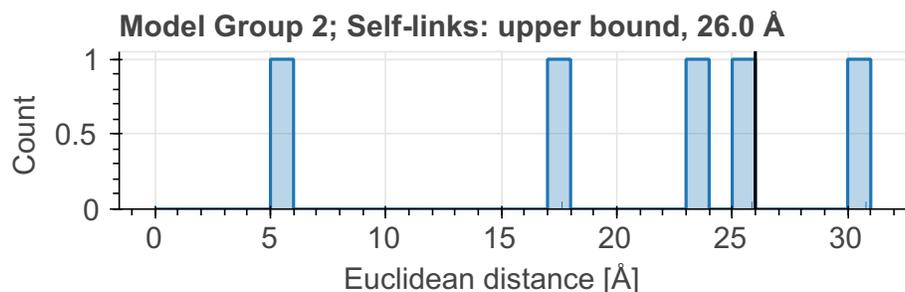
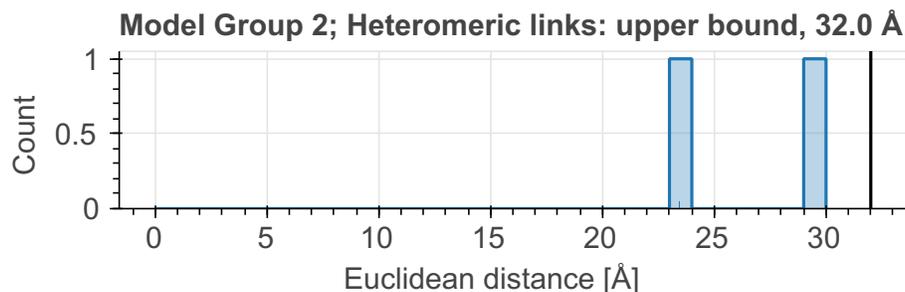
There are 13 crosslinking restraints combined in 13 restraint groups.

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
BS3	LYS	CA	LYS	CA	upper bound	26.00	7
BS3	LYS	CA	PRO	CA	upper bound	32.00	1
BS3	LEU	CA	LYS	CA	upper bound	32.00	2
BS3	LEU	CA	LYS	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.





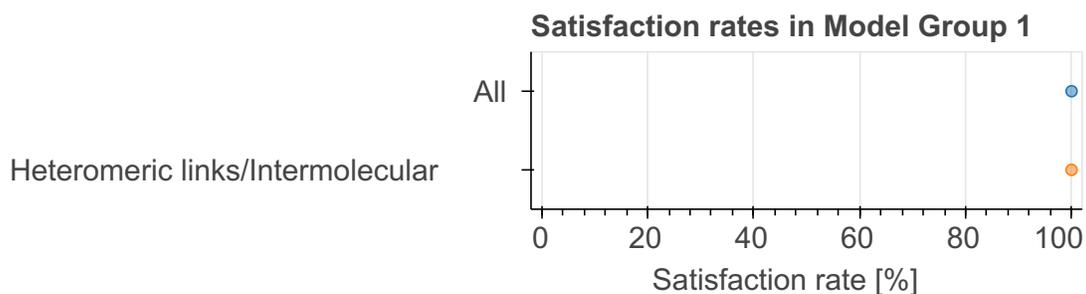
5.2.2. Satisfaction of restraints ?

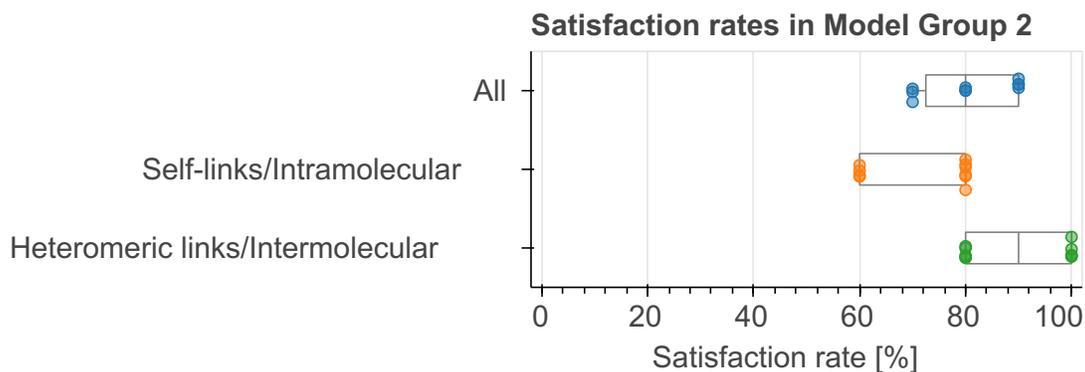
Satisfaction of restraints is calculated on a *restraint group* (a set of crosslinking restraints applied collectively in the modeling) level. Satisfaction of a restraint group depends on satisfaction of individual restraints in the group and the conditionality (all/any). A restraint group is considered 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=13)
1	1	1	1/1	All	100.00	0.00	3
				Heteromeric links/Intermolecular	100.00	0.00	3
1	1	2	10/10	All	90.00	10.00	10
				Self-links/Intramolecular	80.00	20.00	5
				Heteromeric links/Intermolecular	100.00	0.00	5

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 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.