

# Integrative Structure Validation Report ?

February 18, 2025 - 08:30 AM PST

*The following software was used in the production of this report:*

*Integrative Modeling Validation* Version 2.0

*Python-IHM* Version 1.8

*MolProbity* Version 4.5.2

PDB ID	9A0L
PDB-Dev ID	PDBDEV_00000057
Structure Title	Integrative structure of Apo-GAFab
Structure Authors	Gupta R; Liu Y; Wang H; Nordyke CT; Puterbaugh RZ; Cui W; Varga K; Chu F; Ke H; Vashisth H; Cote RH
Deposited on	2020-08-07

*This is a PDB-IHM IM Structure Validation Report for a publicly released PDB-IHM entry.*

*We welcome your comments at [helpdesk@pdb-ihm.org](mailto:helpdesk@pdb-ihm.org)*

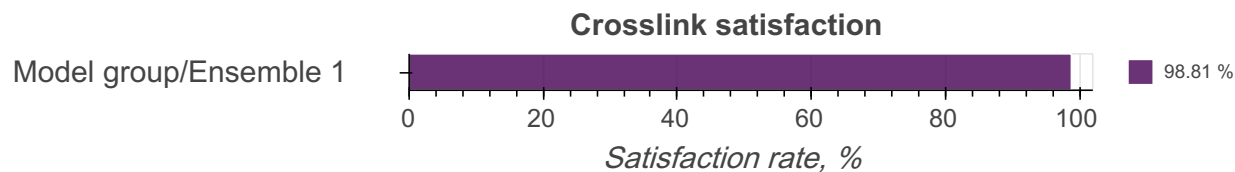
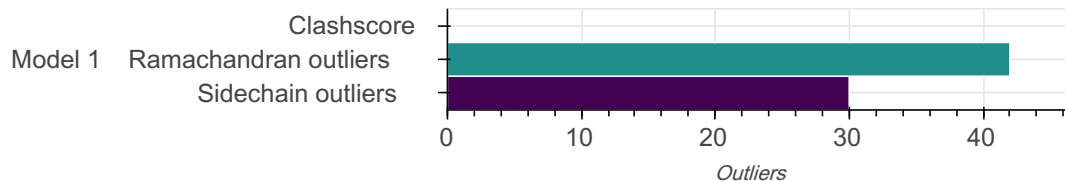
*A user guide is available at [https://pdb-ihm.org/validation\\_help.html](https://pdb-ihm.org/validation_help.html) with specific help available everywhere you see the ? symbol.*

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

## Overall quality ?

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

Model Quality: MolProbity Analysis



### Ensemble information ?

This entry consists of 0 distinct ensemble(s).

### Summary ?

This entry consists of 1 model(s). A total of 2 datasets were used to build this entry.

### Representation ?

This entry has 1 representation(s).

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
1	1	1	GAFab	A	399	-	1-399	100.00 / 100.00	Atomic
				B					

### Datasets used for modeling ?

There are 2 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	Crosslinking-MS data	PRIDE	PXD020817
2	Experimental model	PDB	6X88

## 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	None	None	None	None	False	False

There are 2 software packages reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	<a href="#">Integrative Modeling Platform (IMP)</a>	Not available	integrative model building	<a href="https://integrativemodeling.org">https://integrativemodeling.org</a>
2	<a href="#">Modeller</a>	Not available	model building	<a href="https://salilab.org/modeller/">https://salilab.org/modeller/</a>

## Data quality ?

### Crosslinking-MS

At the moment, data validation is only available for crosslinking-MS data deposited as a fully [compliant](#) dataset in the [PRIDE Crosslinking](#) database. Correspondence between crosslinking-MS and entry entities is established using [pyHMMER](#). Only residue pairs that passed the reported threshold are used for the analysis. The values in the report have to be interpreted in the context of the experiment (i.e. only a minor fraction of in-situ or in-vivo dataset can be used for modeling).

Crosslinking-MS dataset is not available in the [PRIDE Crosslinking](#) database.

## Model quality ?

For models with atomic structures, MolProbity analysis is performed. For models with coarse-grained or multi-scale structures, excluded volume analysis is performed.

### Standard geometry: bond outliers ?

There are 49 bond length outliers in this entry (0.74% of 6663 assessed bonds). A summary is provided below.

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
B	397	ILE	CA-C	12.16	1.78	1.52	1	1
A	314	GLU	C-N	10.30	1.47	1.33	1	1
B	397	ILE	C-N	10.27	1.47	1.33	1	1
B	314	GLU	C-N	10.13	1.47	1.33	1	1
A	209	LYS	CA-C	9.00	1.71	1.52	1	1

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
B	209	LYS	CA-C	8.92	1.71	1.52	1	1
A	397	ILE	C-N	8.90	1.45	1.33	1	1
A	314	GLU	CA-C	8.37	1.70	1.52	1	1
B	314	GLU	CA-C	8.18	1.70	1.52	1	1
A	397	ILE	CA-C	8.08	1.69	1.52	1	1
B	395	LYS	C-N	6.72	1.42	1.33	1	1
A	396	ASP	CA-C	6.54	1.39	1.52	1	1
A	315	TYR	N-CA	6.08	1.57	1.46	1	1
B	388	MET	C-N	6.04	1.41	1.33	1	1
B	342	LYS	N-CA	5.95	1.57	1.46	1	1
B	315	TYR	N-CA	5.78	1.57	1.46	1	1
B	393	ASN	CA-C	5.72	1.40	1.52	1	1
A	397	ILE	N-CA	5.59	1.35	1.46	1	1
A	76	PRO	CA-C	5.59	1.64	1.52	1	1
B	208	HIS	C-N	5.54	1.41	1.33	1	1
B	76	PRO	CA-C	5.31	1.64	1.52	1	1
A	208	HIS	C-N	5.30	1.40	1.33	1	1
B	397	ILE	N-CA	5.23	1.36	1.46	1	1
A	315	TYR	C-N	5.20	1.40	1.33	1	1
B	160	LYS	C-N	4.92	1.40	1.33	1	1
B	232	LYS	C-N	4.88	1.40	1.33	1	1
A	342	LYS	N-CA	4.86	1.55	1.46	1	1
A	395	LYS	C-N	4.85	1.40	1.33	1	1
B	313	ASP	C-N	4.85	1.40	1.33	1	1
A	321	GLY	C-N	4.67	1.27	1.34	1	1
A	209	LYS	C-N	4.65	1.39	1.33	1	1
B	209	LYS	C-N	4.57	1.39	1.33	1	1
B	106	LYS	C-N	4.55	1.39	1.33	1	1
A	341	LYS	CA-C	4.47	1.62	1.52	1	1
B	398	ALA	N-CA	4.43	1.54	1.46	1	1
A	76	PRO	C-N	4.42	1.39	1.33	1	1
B	341	LYS	C-N	4.40	1.39	1.33	1	1
B	315	TYR	C-N	4.39	1.39	1.33	1	1
A	388	MET	C-N	4.37	1.39	1.33	1	1
A	341	LYS	C-N	4.24	1.39	1.33	1	1

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	232	LYS	C-N	4.24	1.39	1.33	1	1
A	160	LYS	C-N	4.22	1.39	1.33	1	1
B	394	ARG	N-CA	4.21	1.38	1.46	1	1
A	105	LYS	C-N	4.16	1.39	1.33	1	1
B	159	PHE	CA-C	4.14	1.44	1.52	1	1
A	391	LEU	C-N	4.14	1.39	1.33	1	1
A	160	LYS	CA-C	4.13	1.61	1.52	1	1
B	76	PRO	C-N	4.13	1.39	1.33	1	1
A	159	PHE	CA-C	4.03	1.44	1.52	1	1

### Standard geometry: angle outliers ?

There are 262 bond angle outliers in this entry (2.90% of 9020 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	106	LYS	C-CA-CB	19.23	73.57	110.10	1	1
B	106	LYS	C-CA-CB	17.63	76.60	110.10	1	1
A	104	VAL	C-N-CA	14.78	148.30	121.70	1	1
B	104	VAL	C-N-CA	14.57	147.92	121.70	1	1
A	107	ASN	CA-CB-CG	12.69	99.91	112.60	1	1
A	396	ASP	C-N-CA	12.32	99.52	121.70	1	1
B	303	PHE	CA-CB-CG	12.23	101.57	113.80	1	1
A	106	LYS	N-CA-C	11.22	79.60	111.00	1	1
B	316	PHE	CA-CB-CG	11.02	102.78	113.80	1	1
B	106	LYS	N-CA-CB	10.68	128.66	110.50	1	1
B	107	ASN	CA-CB-CG	10.34	102.26	112.60	1	1
B	106	LYS	N-CA-C	9.74	83.72	111.00	1	1
A	381	ASN	CA-CB-CG	9.72	122.32	112.60	1	1
B	51	ASN	CA-CB-CG	9.65	102.95	112.60	1	1
B	106	LYS	C-N-CA	9.54	138.87	121.70	1	1
A	106	LYS	N-CA-CB	9.47	126.60	110.50	1	1
A	106	LYS	C-N-CA	9.33	138.50	121.70	1	1
B	388	MET	CA-C-O	9.30	104.98	120.80	1	1
B	315	TYR	C-N-CA	9.29	138.42	121.70	1	1
B	356	ASP	C-N-CA	9.21	138.28	121.70	1	1
B	9	PHE	CA-CB-CG	8.80	105.00	113.80	1	1

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
B	103	ASP	C-N-CA	8.65	137.28	121.70	1	1
B	109	HIS	N-CA-CB	8.63	125.17	110.50	1	1
B	388	MET	C-N-CA	8.08	136.24	121.70	1	1
B	396	ASP	C-N-CA	8.05	107.21	121.70	1	1
A	322	PRO	N-CA-C	8.00	132.10	112.10	1	1
A	316	PHE	CA-CB-CG	7.83	105.97	113.80	1	1
B	393	ASN	C-N-CA	7.79	107.68	121.70	1	1
A	197	PHE	CA-CB-CG	7.77	106.03	113.80	1	1
B	397	ILE	C-N-CA	7.77	135.68	121.70	1	1
A	315	TYR	C-CA-CB	7.67	95.53	110.10	1	1
B	109	HIS	CA-CB-CG	7.58	121.38	113.80	1	1
A	396	ASP	CA-CB-CG	7.53	105.07	112.60	1	1
A	110	PHE	CA-CB-CG	7.52	106.28	113.80	1	1
B	351	PHE	CA-CB-CG	7.43	106.37	113.80	1	1
A	105	LYS	N-CA-C	7.39	131.68	111.00	1	1
A	351	PHE	CA-CB-CG	7.36	106.44	113.80	1	1
B	98	PHE	CA-CB-CG	7.24	106.56	113.80	1	1
A	395	LYS	C-N-CA	7.24	134.74	121.70	1	1
B	341	LYS	C-N-CA	7.23	134.71	121.70	1	1
B	115	ASP	CA-CB-CG	7.20	105.40	112.60	1	1
A	301	ASN	CA-CB-CG	7.18	105.42	112.60	1	1
B	236	PHE	CA-CB-CG	7.16	106.64	113.80	1	1
B	396	ASP	CA-CB-CG	7.15	105.45	112.60	1	1
A	190	LEU	C-CA-CB	7.09	96.62	110.10	1	1
A	109	HIS	N-CA-CB	6.90	122.24	110.50	1	1
B	363	TYR	CA-CB-CG	6.85	101.57	113.90	1	1
B	389	ASN	CA-CB-CG	6.83	105.77	112.60	1	1
B	398	ALA	N-CA-CB	6.71	120.46	110.40	1	1
B	314	GLU	CA-C-N	6.68	129.57	116.20	1	1
A	76	PRO	N-CA-C	6.64	95.50	112.10	1	1
B	76	PRO	N-CA-C	6.62	95.56	112.10	1	1
B	313	ASP	N-CA-CB	6.60	121.72	110.50	1	1
B	180	ASN	CA-CB-CG	6.60	106.00	112.60	1	1
B	398	ALA	C-N-CA	6.59	109.84	121.70	1	1
A	238	ASP	C-N-CA	6.58	133.54	121.70	1	1

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	341	LYS	C-N-CA	6.54	133.46	121.70	1	1
B	197	PHE	CA-CB-CG	6.52	107.28	113.80	1	1
B	356	ASP	CA-CB-CG	6.51	106.09	112.60	1	1
A	314	GLU	CA-C-N	6.49	129.19	116.20	1	1
A	341	LYS	CA-C-N	6.49	129.19	116.20	1	1
B	396	ASP	CA-C-N	6.49	103.23	116.20	1	1
B	108	ASN	CA-CB-CG	6.38	106.22	112.60	1	1
A	314	GLU	N-CA-C	6.36	93.19	111.00	1	1
A	207	PHE	CA-CB-CG	6.34	107.46	113.80	1	1
B	314	GLU	N-CA-C	6.31	93.32	111.00	1	1
B	341	LYS	CA-C-N	6.31	128.82	116.20	1	1
A	381	ASN	N-CA-CB	6.28	121.18	110.50	1	1
A	45	PHE	CA-CB-CG	6.20	107.60	113.80	1	1
A	356	ASP	CA-CB-CG	6.19	106.41	112.60	1	1
A	75	ASN	CA-CB-CG	6.17	106.43	112.60	1	1
A	105	LYS	CA-C-O	6.12	110.40	120.80	1	1
A	217	TYR	CA-CB-CG	6.11	102.91	113.90	1	1
B	238	ASP	C-N-CA	6.10	132.68	121.70	1	1
A	94	HIS	C-CA-CB	6.10	98.52	110.10	1	1
B	105	LYS	N-CA-C	6.09	128.05	111.00	1	1
A	340	ASN	CA-CB-CG	6.08	106.52	112.60	1	1
B	398	ALA	N-CA-C	6.05	127.93	111.00	1	1
A	71	ASP	CA-CB-CG	6.03	106.57	112.60	1	1
B	106	LYS	CA-CB-CG	6.01	126.13	114.10	1	1
A	105	LYS	CA-C-N	6.00	128.20	116.20	1	1
A	398	ALA	N-CA-C	5.99	127.77	111.00	1	1
B	194	ASN	CA-CB-CG	5.95	106.65	112.60	1	1
B	389	ASN	C-N-CA	5.94	111.00	121.70	1	1
B	209	LYS	N-CA-C	5.91	94.46	111.00	1	1
B	108	ASN	C-N-CA	5.91	132.33	121.70	1	1
B	217	TYR	CA-CB-CG	5.89	103.29	113.90	1	1
A	397	ILE	CA-C-N	5.89	127.97	116.20	1	1
B	388	MET	CA-C-N	5.87	127.93	116.20	1	1
A	262	ASN	CA-CB-CG	5.84	106.76	112.60	1	1
A	209	LYS	N-CA-C	5.83	94.68	111.00	1	1

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	309	ASN	CA-CB-CG	5.81	106.79	112.60	1	1
B	85	ASP	CA-CB-CG	5.80	106.80	112.60	1	1
A	6	ASN	CA-CB-CG	5.78	106.82	112.60	1	1
A	160	LYS	O-C-N	5.73	113.84	123.00	1	1
B	79	GLU	C-CA-CB	5.68	99.31	110.10	1	1
A	398	ALA	C-N-CA	5.67	111.50	121.70	1	1
B	209	LYS	O-C-N	5.65	113.96	123.00	1	1
A	338	ILE	C-N-CA	5.62	111.58	121.70	1	1
B	381	ASN	CA-CB-CG	5.62	106.98	112.60	1	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

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	794	712	40	42

There are 42 unique backbone outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	21	SER	1
A	86	ILE	1
A	104	VAL	1
A	107	ASN	1
A	144	ASN	1
A	148	ALA	1
A	232	LYS	1
A	233	GLU	1
A	242	ILE	1
A	255	THR	1
A	256	PRO	1



Chain	Res	Type	Models (Total)
A	258	GLY	1
A	261	VAL	1
A	262	ASN	1
A	302	GLY	1
A	309	ASN	1
A	312	ALA	1
A	321	GLY	1
A	322	PRO	1
A	340	ASN	1
A	343	GLU	1
A	353	ASN	1
A	357	GLY	1
B	51	ASN	1
B	79	GLU	1
B	104	VAL	1
B	107	ASN	1
B	144	ASN	1
B	148	ALA	1
B	233	GLU	1
B	234	LYS	1
B	255	THR	1
B	258	GLY	1
B	261	VAL	1
B	262	ASN	1
B	311	PRO	1
B	313	ASP	1
B	322	PRO	1
B	340	ASN	1
B	343	GLU	1
B	353	ASN	1
B	357	GLY	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	724	634	60	30

*There are 30 unique sidechain outliers. Detailed list of outliers are tabulated below.*

Chain	Res	Type	Models (Total)
A	16	GLN	1
A	28	LYS	1
A	79	GLU	1
A	101	ILE	1
A	103	ASP	1
A	106	LYS	1
A	187	GLN	1
A	206	GLN	1
A	215	ARG	1
A	260	GLU	1
A	274	LYS	1
A	307	MET	1
A	320	LYS	1
A	336	LEU	1
A	342	LYS	1
B	50	ARG	1
B	101	ILE	1
B	103	ASP	1
B	106	LYS	1
B	117	LYS	1
B	124	ASN	1
B	147	ASN	1
B	200	LEU	1
B	232	LYS	1
B	234	LYS	1
B	251	LYS	1
B	260	GLU	1
B	274	LYS	1
B	320	LYS	1
B	336	LEU	1

## Fit of model to data used for modeling ?

### Fit of model(s) to crosslinking-MS data

#### Restraint types

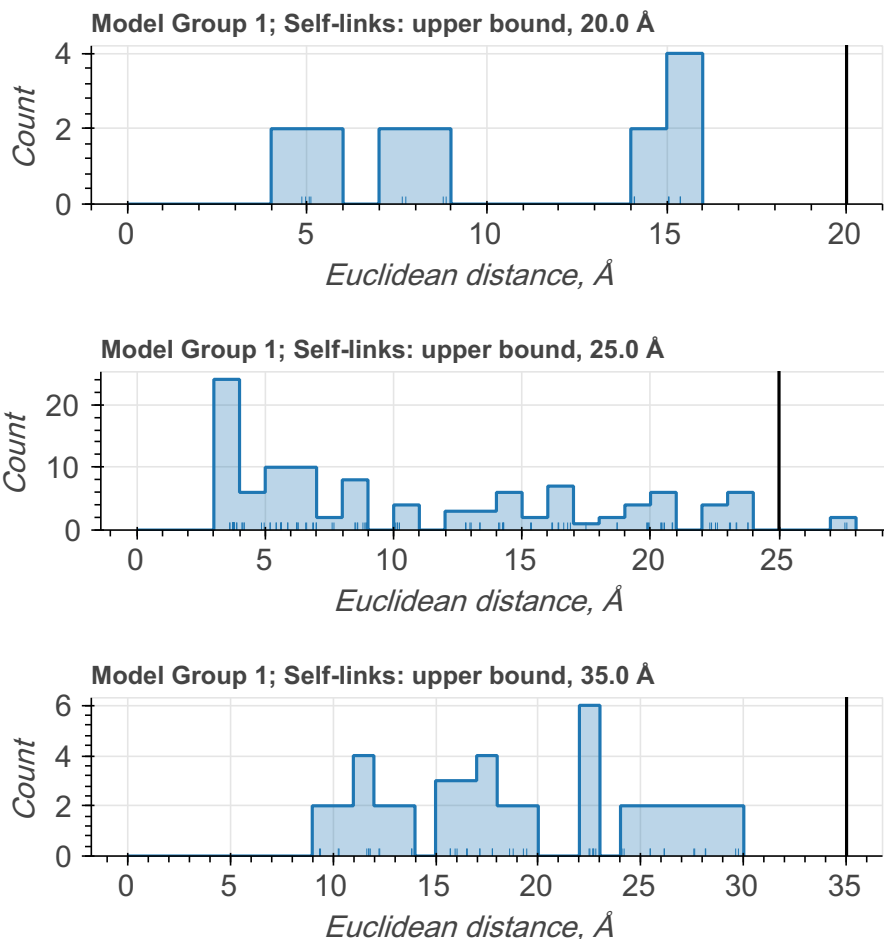
Restraint types are summarized in the table below. Restraints assigned "by-residue" are interpreted as between CA atoms. Restraints between coarse-grained beads are indicated as "coarse-grained". *Restraint group* represents a set of crosslinking restraints applied collectively in the modeling.

There are 168 crosslinking restraints combined in 84 restraint groups.

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
sulfo-SDA	GLU	CA	LYS	CA	upper bound	25.0	38
sulfo-SDA	LYS	CA	VAL	CA	upper bound	25.0	12
sulfo-SDA	ILE	CA	LYS	CA	upper bound	25.0	8
sulfo-SDA	LYS	CA	THR	CA	upper bound	25.0	2
sulfo-SDA	ARG	CA	LYS	CA	upper bound	25.0	4
sulfo-SDA	LYS	CA	PRO	CA	upper bound	25.0	6
sulfo-SDA	TYR	CA	VAL	CA	upper bound	25.0	2
sulfo-SDA	LEU	CA	LYS	CA	upper bound	25.0	6
sulfo-SDA	HIS	CA	LYS	CA	upper bound	25.0	2
sulfo-SDA	LYS	CA	LYS	CA	upper bound	25.0	4
sulfo-SDA	ASP	CA	LYS	CA	upper bound	25.0	6
sulfo-SDA	ALA	CA	LYS	CA	upper bound	25.0	4
sulfo-SDA	LYS	CA	TYR	CA	upper bound	25.0	8
sulfo-SDA	GLY	CA	LYS	CA	upper bound	25.0	6
sulfo-SDA	ASN	CA	LYS	CA	upper bound	25.0	2
BS3	LYS	CA	LYS	CA	upper bound	35.0	44
EDC	LYS	CA	MET	CA	upper bound	20.0	2
EDC	ASP	CA	LYS	CA	upper bound	20.0	4
EDC	GLU	CA	LYS	CA	upper bound	20.0	8

#### Distograms of individual restraints

Restraints with identical thresholds are grouped into one plot. Only the best distance per restraint per model group/ensemble is plotted. Inter- and intramolecular (including self-links) restraints are also grouped into one plot. Distance for a restraint between coarse-grained beads is calculated as a minimal distance between shells; if beads intersect, the distance will be reported as 0.0. A bead with the highest available resolution for a given residue is used for the assessment.



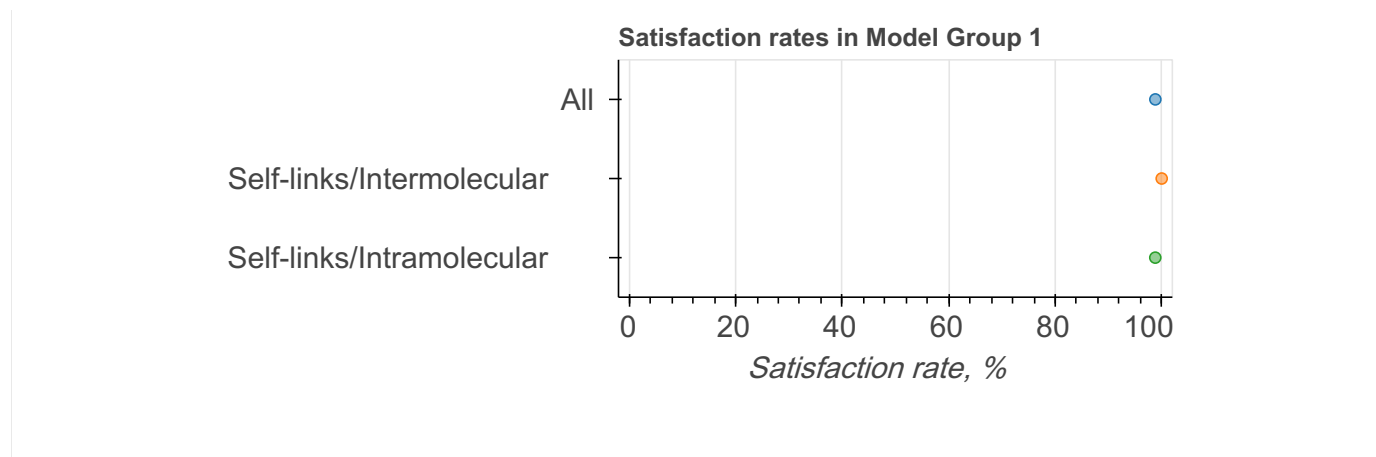
### Satisfaction of restraints

Satisfaction of restraints is calculated on a [restraint group](#) (a set of crosslinking restraints applied collectively in the modeling) level. Satisfaction of a restraint group depends on satisfaction of individual restraints in the group and the conditionality (all/any). A restraint group is considered satisfied, if the condition was met in at least one model of the model group/ensemble. The number of measured restraints can be smaller than the total number of restraint groups if crosslinks involve non-modeled residues. Only deposited models are used for validation right now.

State group	State	Model group	# of Deposited models/Total	Restraint group type	Satisfied (%)	Violated (%)	Count (Total=84)
1	1	1	1/1	All	98.81	1.19	84
				Self-links/ Intermolecular	100.00	0.00	1
				Self-links/ Intramolecular	98.80	1.20	83

### Per-model satisfaction rates in ensembles

Every point represents one model in a model group/ensemble. Where possible, boxplots with quartile marks are also plotted.



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

Validation for this section is under development.

#### Acknowledgments

*The development of integrative model validation metrics, implementation of a model validation pipeline, and creation of a validation report for integrative structures are funded by NSF awards to the [PDB-IHM team](#) (DBI-1756248, DBI-2112966, DBI-2112967, DBI-2112968, and DBI-1756250) and awards from NSF, NIH, and DOE to the [RCSB PDB](#) (DBI-2321666, R01GM157729, and DE-SC0019749). The PDB-IHM team and members of the [Sali lab](#) contributed model validation metrics and software packages.*

*Dr. Jill Trehwella, Dr. Dina Schneidman, and members of the [SASBDB](#) repository are acknowledged for their advice and support in implementing SAS validation methods. Team members from the labs of Dr. Juri Rappsilber, Dr. Alexander Leitner, Dr. Andrea Graziadei, and members of [PRIDE](#) database are acknowledged for their advice and support in implementing crosslinking-MS validation methods. We are grateful to Dr. Shruthi Viswanath for discussions about uncertainty assessment of integrative structural models.*

*Members of the [wwPDB Integrative/Hybrid Methods Task Force](#) provided recommendations and community support for the project.*