

Integrative Structure Validation Report

March 13, 2025 - 10:15 AM PDT

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	8ZZI
PDB-Dev ID	PDBDEV_00000018
Structure Title	The molecular architecture of the BBSome and its implications for facilitated transition zone crossing
Structure Authors	Chou H; Apelt L; Farrell DP; White SR; Woodsmith J; Svetlov V; Goldstein JS; Nager AR; Li Z; Muller J; Dollfus H; Nudler E; Stelzl U; DiMaio F; Nachury MV; Walz T
Deposited on	2018-05-04

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

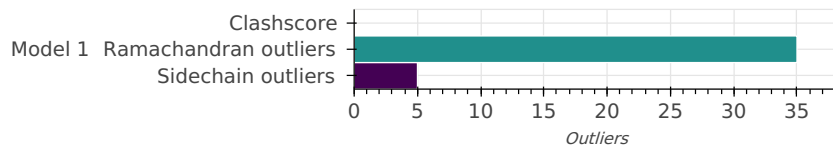
A user guide is available at https://pdb-ihm.org/validation_help.html with specific help available everywhere you see the  symbol.

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

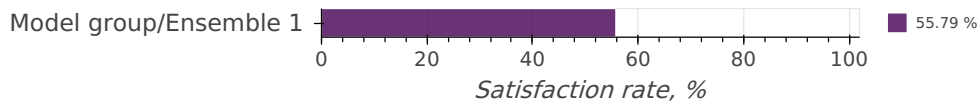
Overall quality

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



Crosslink satisfaction



Ensemble information

This entry consists of 1 distinct ensemble(s).

Summary ⓘ

This entry consists of 1 model(s). A total of 33 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	BBS1	1	593	1-18, 19-409, 116-194, 410-481, 482-593	-	100.00 / 84.82	Atomic
		2	BBS2	2	721	1-318, 319-331, 332-365, 366-368, 369-467, 468-587, 588-721	-	100.00 / 97.78	Atomic
		3	BBS4	4	519	1-461, 462-519	-	100.00 / 88.82	Atomic
		4	BBS5	5	341	1-143, 144-268, 269-341	-	100.00 / 78.59	Atomic
		5	BBS7	7	712	1-5, 6-315, 316-325, 326-365, 366-372, 373-486, 487-494, 495-592, 593-597, 598-708, 709-712	-	100.00 / 94.52	Atomic
		6	BBS8	8	506	1-7, 8-56, 57-174, 175-506	-	100.00 / 75.30	Atomic
		7	BBS9	9	887	1-8, 9-369, 368-409, 410-525, 526-641, 642-742, 743-824, 825-887	-	100.00 / 87.49	Atomic
		8	BBS18	IP	96	1-47, 48-96	-	100.00 / 51.04	Atomic

Datasets used for modeling ⓘ

There are 33 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	Comparative model	Zenodo	10.5281/zenodo.1255360
2	De Novo model	Zenodo	10.5281/zenodo.1255360
3	De Novo model	Zenodo	10.5281/zenodo.1255360
4	Comparative model	Zenodo	10.5281/zenodo.1255360
5	De Novo model	Zenodo	10.5281/zenodo.1255360
6	De Novo model	Zenodo	10.5281/zenodo.1255360
7	De Novo model	Zenodo	10.5281/zenodo.1255360
8	De Novo model	Zenodo	10.5281/zenodo.1255360
9	Comparative model	Zenodo	10.5281/zenodo.1255360
10	Comparative model	Zenodo	10.5281/zenodo.1255360
11	Comparative model	Zenodo	10.5281/zenodo.1255360
12	Comparative model	Zenodo	10.5281/zenodo.1255360
13	De Novo model	Zenodo	10.5281/zenodo.1255360
14	De Novo model	Zenodo	10.5281/zenodo.1255360

ID	Dataset type	Database name	Data access code
15	De Novo model	Zenodo	10.5281/zenodo.1255360
16	De Novo model	Zenodo	10.5281/zenodo.1255360
17	Comparative model	Zenodo	10.5281/zenodo.1255360
18	Comparative model	Zenodo	10.5281/zenodo.1255360
19	Comparative model	Zenodo	10.5281/zenodo.1255360
20	De Novo model	Zenodo	10.5281/zenodo.1255360
21	De Novo model	Zenodo	10.5281/zenodo.1255360
22	De Novo model	Zenodo	10.5281/zenodo.1255360
23	De Novo model	Zenodo	10.5281/zenodo.1255360
24	De Novo model	Zenodo	10.5281/zenodo.1255360
25	3DEM volume	EMDB	EMD-7839
26	Crosslinking-MS data	Zenodo	10.5281/zenodo.1255360
27	Experimental model	PDB	4V0N
28	Experimental model	PDB	1VYH
29	Experimental model	PDB	5G05
30	Experimental model	PDB	2CAY
31	Experimental model	PDB	3HSA
32	Experimental model	PDB	1W3B
33	Experimental model	PDB	4YHD

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	Production sampling	Monte Carlo	None	None	False	False
2	1	Rosetta Hybridize	Rosetta Hybridize	None	None	False	False

There are 2 software packages reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	Rosetta	Rosetta version unknown:839226a33c427862a8be7b4ca555493368c1485c 2017-09-18 10:39:53 -0700 from git@github.com:RosettaCommons/main.git	RosettaCM/hybridize, Rosetta Abinitio, and unpublished 'complex assembly'	https://www.rosettacommons.org/
2	HHPred	website	protein homology detection	https://toolkit.tuebingen.mpg.de/hhpred

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.

3DEM volume

Validation for this section is under development.

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 5 bond length outliers in this entry (0.02% of 28381 assessed bonds). A summary is provided below.

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
1	71	LEU	C-N	4.80	1.40	1.33	1	1
1	388	LEU	CB-CG	4.35	1.62	1.53	1	1
9	614	ASN	CB-CG	4.15	1.41	1.52	1	1
9	117	HIS	CB-CG	4.02	1.44	1.50	1	1
7	497	ILE	CB-CG1	4.01	1.45	1.53	1	1

Standard geometry: angle outliers ?

There are 12 bond angle outliers in this entry (0.03% of 38417 assessed bonds). A summary is provided below.

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
9	115	VAL	C-N-CA	19.31	156.46	121.70	1	1
1	71	LEU	C-N-CA	6.94	134.19	121.70	1	1
4	169	LEU	C-N-CA	6.87	134.06	121.70	1	1
1	524	ASN	C-N-CA	6.37	133.16	121.70	1	1
8	19	ARG	C-N-CA	5.52	131.64	121.70	1	1
5	236	LYS	C-N-CA	4.98	130.67	121.70	1	1
4	250	GLY	C-N-CA	4.94	130.59	121.70	1	1
1	62	ASP	CA-CB-CG	4.76	107.84	112.60	1	1
1	47	HIS	CA-CB-CG	4.76	118.56	113.80	1	1
9	476	PHE	CA-CB-CG	4.26	118.06	113.80	1	1
5	235	PHE	CA-CB-CG	4.08	117.88	113.80	1	1
5	113	LEU	C-N-CA	4.08	129.04	121.70	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	3488	3257	196	35

There are 35 unique backbone outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
1	101	THR	1
1	167	LEU	1
1	246	SER	1
1	514	ILE	1
2	4	PRO	1
2	14	ILE	1
2	28	HIS	1
2	131	ILE	1
2	327	ASN	1
2	490	ASP	1
2	499	VAL	1
2	522	GLU	1
2	689	LYS	1
4	14	VAL	1
4	170	THR	1
4	271	PRO	1
5	26	PRO	1
5	171	SER	1
5	203	TYR	1
5	237	ILE	1
5	238	ASP	1
7	9	PHE	1
7	74	THR	1
7	76	GLN	1
7	125	ILE	1
7	183	VAL	1
8	350	ASN	1
9	115	VAL	1
9	462	GLN	1
9	497	PRO	1
9	512	THR	1
9	529	LEU	1
9	638	GLY	1
9	760	GLN	1
IP	63	LEU	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	3068	3007	56	5

There are 5 unique sidechain outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
1	47	HIS	1

Chain	Res	Type	Models (Total)
1	72	LYS	1
1	251	LEU	1
4	229	LEU	1
4	306	ILE	1

Fit of model to data used for modeling ?

Fit of model(s) to crosslinking-MS data

Restraint types

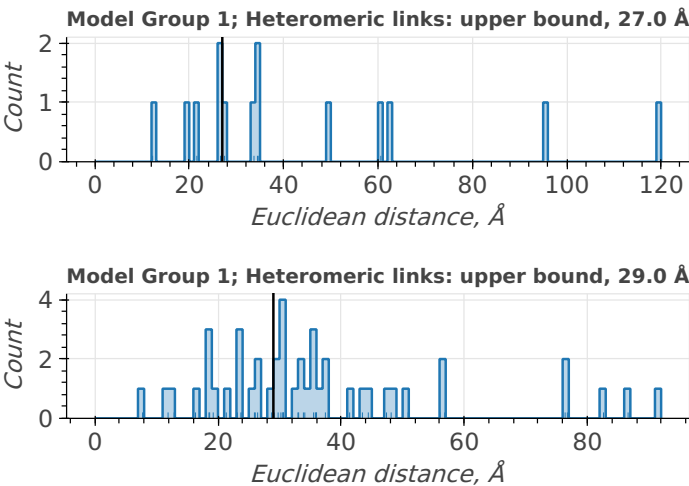
Restraint types are summarized in the table below. Restraints assigned "by-residue" are interpreted as between CA atoms. Restraints between coarse-grained beads are indicated as "coarse-grained". *Restraint group* represents a set of crosslinking restraints applied collectively in the modeling.

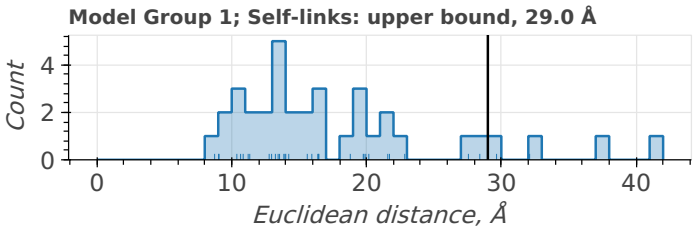
There are 122 crosslinking restraints combined in 122 restraint groups.

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSS	LYS	CA	LYS	CA	upper bound	29.0	99
DSS	LYS	CA	SER	CA	upper bound	29.0	1
DSS	ALA	CA	LYS	CA	upper bound	29.0	2
DSS	ALA	CA	MET	CA	upper bound	29.0	1
BS3	LYS	CA	LYS	CA	upper bound	27.0	17
BS3	LYS	CA	MET	CA	upper bound	27.0	1
BS3	ALA	CA	LEU	CA	upper bound	27.0	1

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.





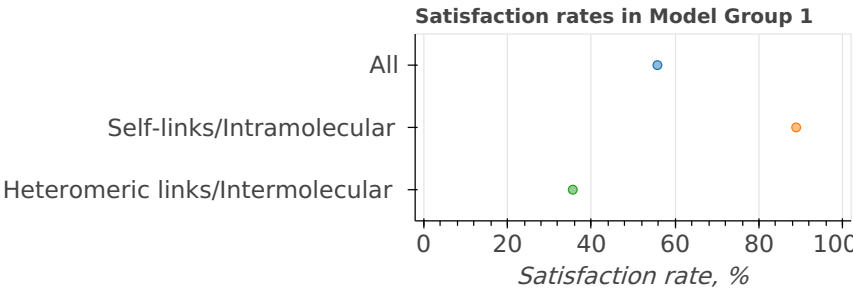
Satisfaction of restraints

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=122)
1	1	1	1/1	All	55.79	44.21	95
				Self-links/ Intramolecular	88.89	11.11	36
				Heteromeric links/ Intermolecular	35.59	64.41	59

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.



3DEM volume

Validation for this section is under development.

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.