Integrative Structure Validation Report

February 18, 2025 - 08:32 AM PST

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

Integrative Modeling Validation Version 2.0
Python-IHM Version 1.8

PDB ID	9A1J
PDB-Dev ID	PDBDEV_00000091
Structure Title	Integrative structure determination of the A3G-CRL5-Vif complex (flexible)
Structure Authors	Robyn M Kaake; Ignacia Echeverria; Seung Joong Kim; John Von Dollen; Nicholas M Chesarino; Yuqing Feng; Clinton Yu; Hai Ta; Linda Chelico; Lan Huang; John Gross; Andrej Sali; Nevan J Krogan
Deposited on	2021-08-11

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

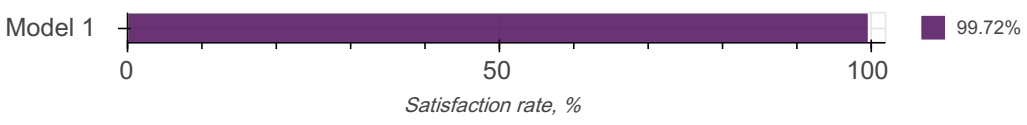
A user guide is available at https://pdb-ihm.org/validation_help.html with specific help available everywhere you see the ? symbol.

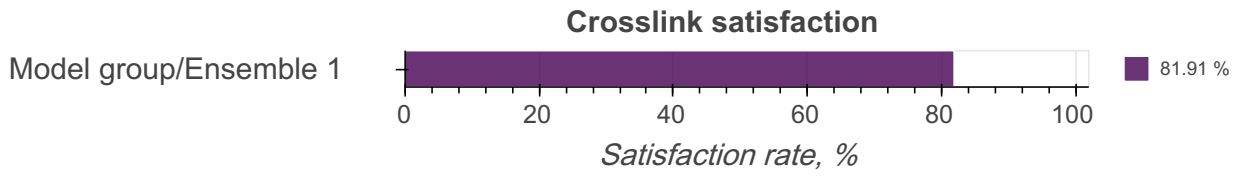
List of references used to build this report is available here.

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: Excluded Volume Analysis





Ensemble information

This entry consists of 1 distinct ensemble(s).

Summary

This entry consists of 1 model(s). A total of 12 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	CBFB	A	182	1-156	157-182	100.00 / 85.71	Multiscale: Coarse-grained: 1 - 5 residue(s) per bead
		2	Vif	B	175	6-154, 166-175	1-5, 155-165	100.00 / 90.86	Coarse-grained: 1 residue(s) per bead
		3	EloB	C	161	1-105	106-161	100.00 / 65.22	Multiscale: Coarse-grained: 1 - 5 residue(s) per bead
		4	EloC	D	112	17-112	1-16	100.00 / 85.71	Multiscale: Coarse-grained: 1 - 5 residue(s) per bead

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
		5	CUL5	E	780	11-302, 308-382, 405-515, 521-568, 574-687, 695-780	1-10, 303-307, 383-404, 516-520, 569-573, 688-694	100.00 / 93.08	Coarse-grained: 1 residue(s) per bead
		6	Rbx2	F	113	27-113	1-26	100.00 / 76.99	Coarse-grained: 1 residue(s) per bead
		7	A3G	G	384	6-194, 200-243, 258-380	1-5, 195-199, 244-257, 381-384	100.00 / 92.71	Coarse-grained: 1 residue(s) per bead

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	4N9F
2	Experimental model	PDB	1LDJ
3	Experimental model	PDB	2ECL
4	Experimental model	PDB	2MA9
5	Comparative model	Zenodo	10.5281/zenodo.5176959
6	Experimental model	PDB	5K81
7	Comparative model	Zenodo	10.5281/zenodo.5176959
8	Experimental model	PDB	3V4K
9	Comparative model	Zenodo	10.5281/zenodo.5176959
10	Mass Spectrometry data	PRIDE	PXD025391
11	Crosslinking-MS data	Zenodo	10.5281/zenodo.5176959
12	Crosslinking-MS data	Zenodo	10.5281/zenodo.5176959

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	None	203100	False	True

There are 4 software packages reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	IMP PMI module	develop-548de65454	integrative model building	https://integrativemodeling.org
2	Integrative Modeling Platform (IMP)	develop-548de65454	integrative model building	https://integrativemodeling.org
3	MODELLER	9.20	comparative modeling	https://salilab.org/modeller/
4	MODELLER	9.19	comparative modeling	https://salilab.org/modeller/

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.

Mass Spectrometry

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.

Excluded volume satisfaction

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.

Model ID	Analysed	Number of violations	Excluded Volume Satisfaction (%)
1	1675365	4692	99.72

Fit of model to data used for modeling

Fit of model(s) to crosslinking-MS data

Restraint types

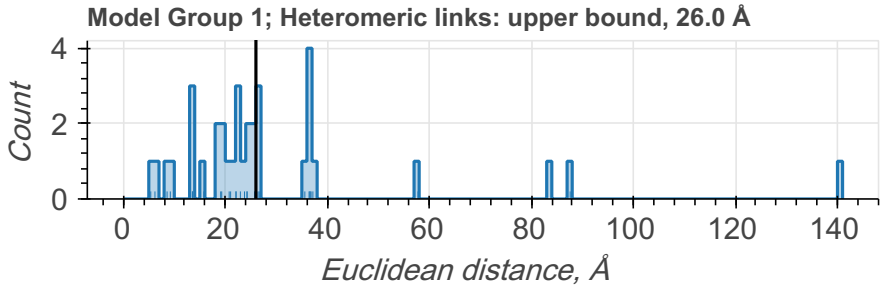
Restraint types are summarized in the table below. Restraints assigned "by-residue" are interpreted as between CA atoms. Restraints between coarse-grained beads are indicated as "coarse-grained". Restraint group represents a set of crosslinking restraints applied collectively in the modeling.

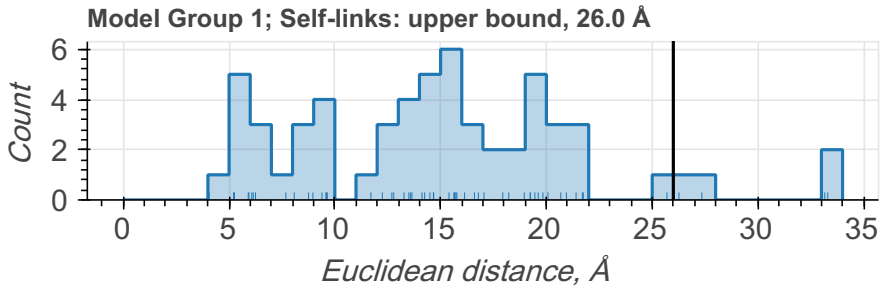
There are 132 crosslinking restraints combined in 132 restraint groups.

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSSO	LYS	CA	LYS	CA	upper bound	26.0	125
DSSO	LYS	CA	MET	CA	upper bound	26.0	7

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=132)
1	1	1	1/198632	All	81.91	18.09	94
				Heteromeric links/ Intermolecular	62.86	37.14	35
				Self-links/ Intramolecular	93.22	6.78	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.

Mass Spectrometry

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 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.