Integrative Structure Validation Report

February 18, 2025 - 08:27 AM PST

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

Integrative Modeling Validation Version 2.0
Python-IHM Version 1.8
ATSAS Version 3.2.1 (r14885)

PDB ID	8ZZK
PDB-Dev ID	PDBDEV_00000020
Structure Title	Structure of the S. cerevisiae nuclear pore complex cytoplasmic mRNA export platform, Nup82
Structure Authors	Fernandez-Martinez J; Kim SJ; Shi Y; Upla P; Pellarin R; Gagnon M; Chemmama IE; Wang J; Nudelman I; Zhang W; Williams R; Rice WJ; Stokes DL; Zenklusen D; Chait BT; Sali A; Rout MP
Deposited on	2018-06-06

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

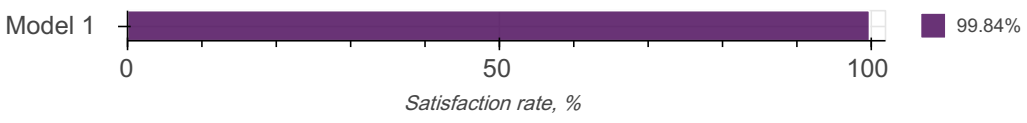
A user guide is available at https://pdb-ihm.org/validation_help.html with specific help available everywhere you see the ? symbol.

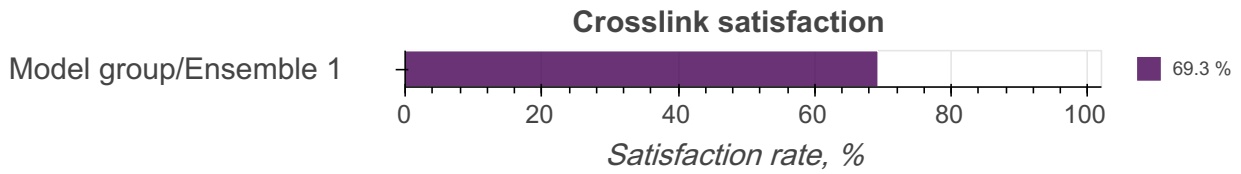
List of references used to build this report is available here.

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 37 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	Dyn2	A B	92	7-92	1-6	100.00 / 93.48	Multiscale: Coarse-grained: 1 - 6 residue(s) per bead
		2	Nup82	C D	713	7-16, 23-120, 123-452, 522-612, 625-669, 678-713	1-6, 17-22, 121-122, 453-521, 613-624, 670-677	100.00 / 85.55	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
		3	Nup159	E F	1460	2-347, 362-381, 1117-1126, 1211-1239, 1266-1321, 1332-1372, 1382-1412, 1429-1456	1, 348-361, 382-1116, 1127-1210, 1240-1265, 1322-1331, 1373-1381, 1413-1428, 1457-1460	100.00 / 38.42	Multiscale: Coarse-grained: 1 - 100 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
		4	Nsp1	G H	823	637-727, 742-778, 788-823	1-636, 728-741, 779-787	100.00 / 19.93	Multiscale: Coarse-grained: 1 - 100 residue(s) per bead
		5	Nup116	I J	1113	966-1111	1-965, 1112-1113	100.00 / 13.12	Multiscale: Coarse-grained: 1 - 100 residue(s) per bead

Datasets used for modeling

There are 37 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	Experimental model	PDB	4DS1
2	Experimental model	PDB	3PBP
3	Experimental model	PDB	5CWS
4	Comparative model	Zenodo	10.5281/zenodo.1256259
5	Comparative model	Zenodo	10.5281/zenodo.1256259
6	Comparative model	Zenodo	10.5281/zenodo.1256259
7	Comparative model	Zenodo	10.5281/zenodo.1256259
8	Comparative model	Zenodo	10.5281/zenodo.1256259
9	Comparative model	Zenodo	10.5281/zenodo.1256259
10	Experimental model	PDB	1XIP
11	Crosslinking-MS data	Zenodo	10.5281/zenodo.1256259
12	Crosslinking-MS data	Zenodo	10.5281/zenodo.1256259
13	Crosslinking-MS data	Zenodo	10.5281/zenodo.1256259
14	2DEM class average	Zenodo	10.5281/zenodo.1256259
15	2DEM class average	Zenodo	10.5281/zenodo.1256259
16	2DEM class average	Zenodo	10.5281/zenodo.1256259

ID	Dataset type	Database name	Data access code
17	2DEM class average	Zenodo	10.5281/zenodo.1256259
18	2DEM class average	Zenodo	10.5281/zenodo.1256259
19	2DEM class average	Zenodo	10.5281/zenodo.1256259
20	2DEM class average	Zenodo	10.5281/zenodo.1256259
21	2DEM class average	Zenodo	10.5281/zenodo.1256259
22	2DEM class average	Zenodo	10.5281/zenodo.1256259
23	2DEM class average	Zenodo	10.5281/zenodo.1256259
24	2DEM class average	Zenodo	10.5281/zenodo.1256259
25	2DEM class average	Zenodo	10.5281/zenodo.1256259
26	2DEM class average	Zenodo	10.5281/zenodo.1256259
27	2DEM class average	Zenodo	10.5281/zenodo.1256259
28	2DEM class average	Zenodo	10.5281/zenodo.1256259
29	2DEM class average	Zenodo	10.5281/zenodo.1256259
30	2DEM class average	Zenodo	10.5281/zenodo.1256259
31	2DEM class average	Zenodo	10.5281/zenodo.1256259
32	2DEM class average	Zenodo	10.5281/zenodo.1256259
33	2DEM class average	Zenodo	10.5281/zenodo.1256259
34	2DEM class average	Zenodo	10.5281/zenodo.1256259
35	SAS data	Zenodo	10.5281/zenodo.1256259
36	SAS data	Zenodo	10.5281/zenodo.1256259
37	SAS data	Zenodo	10.5281/zenodo.1256259

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	1000	False	True
2	1	Sampling	Replica exchange monte carlo	None	1350000	False	True
3	1	Sampling	Replica exchange monte carlo	None	10000	False	True

There are 9 software packages reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	Integrative Modeling Platform (IMP)	develop-0a5706e202	integrative model building	https://integrativemodeling.org
2	IMP PMI module	67456c0	integrative model building	https://integrativemodeling.org
3	HHpred	2.0.16	protein homology detection	https://toolkit.tuebingen.mpg.de/hhpred
4	PSIPRED	4.00	secondary structure prediction	http://bioinf.cs.ucl.ac.uk/psipred/
5	DISOPRED	3	disorder prediction	http://bioinf.cs.ucl.ac.uk/psipred/?disopred=1
6	DomPred	Not available	domain prediction	http://bioinf.cs.ucl.ac.uk/psipred/?dompred=1
7	COILS/PCOILS	Not available	coiled-coil prediction	https://toolkit.tuebingen.mpg.de/pcoils
8	Multicoil2	Not available	coiled-coil prediction	http://groups.csail.mit.edu/cb/multicoil2/cgi-bin/multicoil2.cgi
9	MODELLER	9.15	comparative modeling	https://salilab.org/modeller/

Data quality

SAS:Scattering profile

SAS data used in this integrative model could not be validated as the sascif file is currently unavailable.

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.

2DEM class average

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	5357901	8455	99.84

Fit of model to data used for modeling

SAS data used in this integrative model could not be validated as the sascif file is currently unavailable.

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.

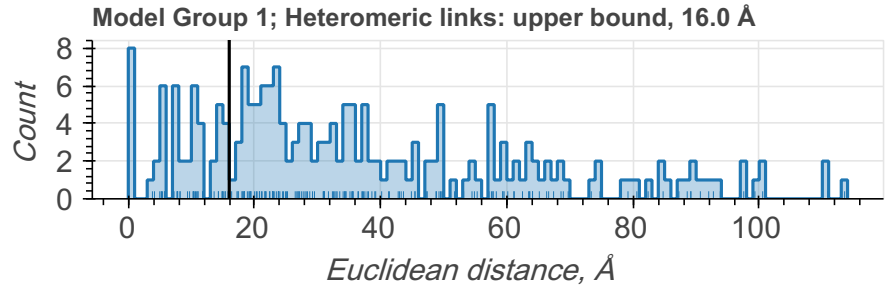
There are 1702 crosslinking restraints combined in 430 restraint groups.

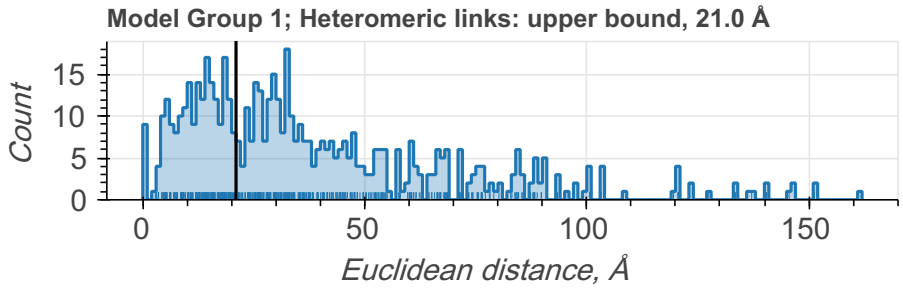
Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSS	LYS	coarse-grained	LYS	coarse-grained	upper bound	21.0	675
DSS	LYS	CA	LYS	CA	upper bound	21.0	487
DSS	LYS	coarse-grained	SER	coarse-grained	upper bound	21.0	36
DSS	LYS	CA	THR	CA	upper bound	21.0	4
DSS	LYS	CA	SER	CA	upper bound	21.0	56
DSS	ARG	coarse-grained	LYS	coarse-grained	upper bound	21.0	12
DSS	ILE	CA	LYS	CA	upper bound	21.0	8
DSS	ASN	coarse-grained	LYS	coarse-grained	upper bound	21.0	20
DSS	ASP	coarse-grained	SER	coarse-grained	upper bound	21.0	4
DSS	GLU	CA	SER	CA	upper bound	21.0	4
DSS	MET	coarse-grained	SER	coarse-grained	upper bound	21.0	8
DSS	GLU	CA	LYS	CA	upper bound	21.0	16
DSS	LYS	coarse-grained	MET	coarse-grained	upper bound	21.0	8
DSS	GLU	coarse-grained	MET	coarse-grained	upper bound	21.0	4
DSS	ASP	coarse-grained	LYS	coarse-grained	upper bound	21.0	12
DSS	ALA	coarse-grained	LYS	coarse-grained	upper bound	21.0	8
DSS	SER	coarse-grained	THR	coarse-grained	upper bound	21.0	4
DSS	ALA	coarse-grained	THR	coarse-grained	upper bound	21.0	4
DSS	SER	coarse-grained	SER	coarse-grained	upper bound	21.0	4

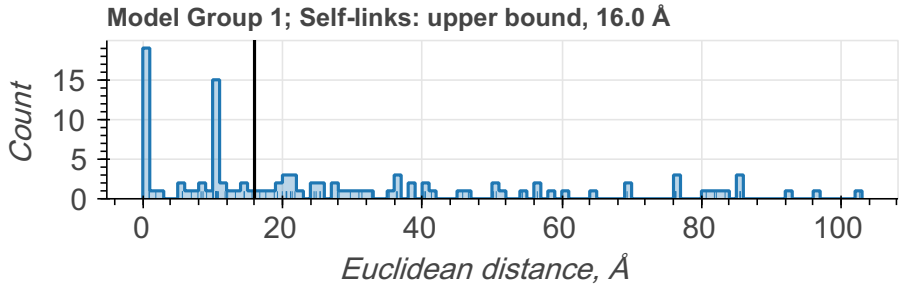
Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSS	ASP	coarse-grained	MET	coarse-grained	upper bound	21.0	4
EDC	ASP	coarse-grained	LYS	coarse-grained	upper bound	16.0	96
EDC	GLU	CA	LYS	CA	upper bound	16.0	60
EDC	ASP	CA	LYS	CA	upper bound	16.0	40
EDC	GLU	coarse-grained	LYS	coarse-grained	upper bound	16.0	128

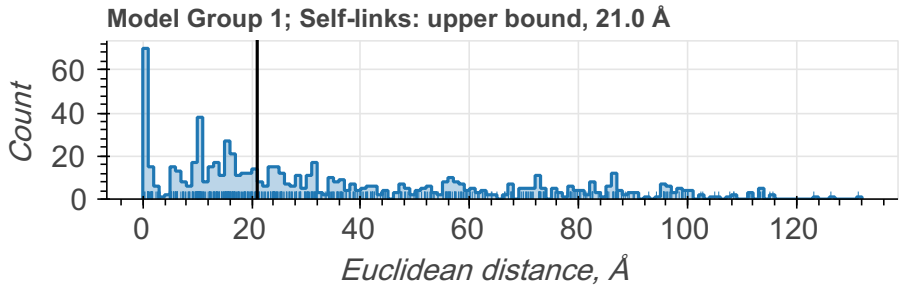
Distograms of individual restraints

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.









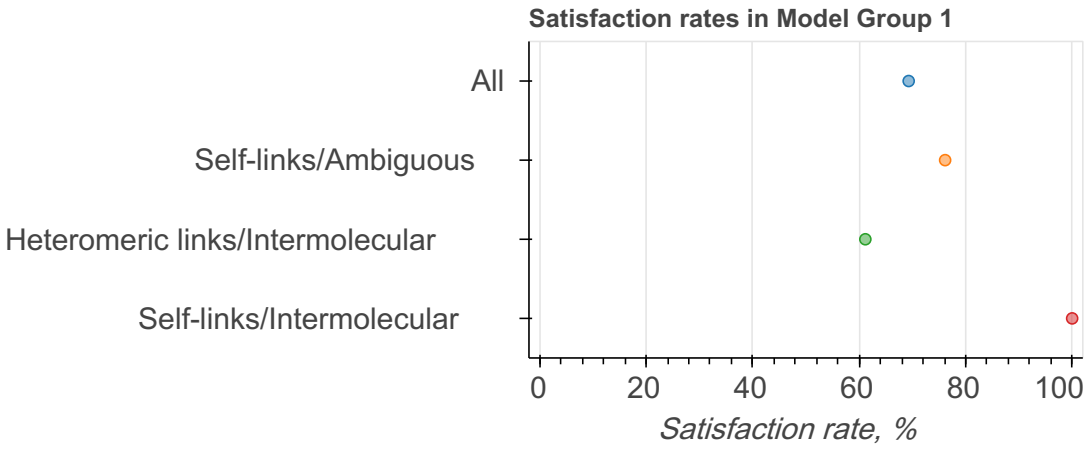
Satisfaction of restraints

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=430)
1	1	1	1/370	All	69.30	30.70	430
				Self-links/ Ambiguous	76.15	23.85	218
				Heteromeric links/ Intermolecular	61.17	38.83	206
				Self-links/ Intermolecular	100.00	0.00	6

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.



2DEM class average

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.