Integrative Structure Validation Report

March 27, 2025 - 10:13 AM PDT

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

Integrative Modeling Validation Version 2.0
Python-IHM Version 1.8

PDB ID	9A8T
Structure Title	Integrative spatiotemporal modeling of biomolecular processes: application to the assembly of the Nuclear Pore Complex
Structure Authors	Latham, A.P.; Tempkin, J.O.B.; Otsuka, S.; Zhang, W.; Ellenberg, J.; Sali, A.
Deposited on	2024-08-09

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 6 distinct ensemble(s).

Summary

This entry consists of 6 model(s). A total of 11 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	1-6	1	Nup133_spoke_1_yc_inner_cr	A	639	1-639	-	100.00 / 100.00	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
				J					
				S					
				AB					
				BK					
				BT					
				CC					
				CL					
				DU					
				ED					
				EM					
				EV					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
				GE					
				GN					
				GW					
				HF					
				IO					
				IX					
				JG					
				JP					
				KY					
				LH					
				LQ					
				LZ					
				NI					
				NR					
				OA					
				OJ					
				PS					
				QB					
				QK					
				QT					
		2	Nup107_spoke_1_yc_inner_cr	B	775	1-775	-	100.00 / 100.00	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
				K					
				T					
				AC					
				BL					
				BU					
				CD					
				CM					
				DV					
				EE					
				EN					
				EW					
				GF					
				GO					
				GX					
				HG					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
				IP					
				IY					
				JH					
				JQ					
				KZ					
				LI					
				LR					
				MA					
				NJ					
				NS					
				OB					
				OK					
				PT					
				QC					
				QL					
				QU					
		3	Nup96_spoke_1_yc_inner_cr	C	475	1-475	-	100.00 / 100.00	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
				L					
				U					
				AD					
				BM					
				BV					
				CE					
				CN					
				DW					
				EF					
				EO					
				EX					
				GG					
				GP					
				GY					
				HH					
				IQ					
				IZ					
				JI					
				JR					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
				LA					
				LJ					
				LS					
				MB					
				NK					
				NT					
				OC					
				OL					
				PU					
				QD					
				QM					
				QV					
		4	SEC13_spoke_1_yc_inner_cr	D	291	1-291	-	100.00 / 100.00	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
				M					
				V					
				AE					
				BN					
				BW					
				CF					
				CO					
				DX					
				EG					
				EP					
				EY					
				GH					
				GQ					
				GZ					
				HI					
				IR					
				JA					
				JJ					
				JS					
				LB					
				LK					
				LT					
				MC					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
				NL					
				NU					
				OD					
				OM					
				PV					
				QE					
				QN					
				QW					
		5	SEH1_spoke_1_yc_inner_cr	E	322	1-322	-	100.00 / 100.00	Multiscale: Coarse-grained: 2 - 10 residue(s) per bead
				N					
				W					
				AF					
				BO					
				BX					
				CG					
				CP					
				DY					
				EH					
				EQ					
				EZ					
				GI					
				GR					
				HA					
				HJ					
				IS					
				JB					
				JK					
				JT					
				LC					
				LL					
				LU					
				MD					
				NM					
				NV					
				OE					
				ON					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
				PW					
				QF					
				QO					
				QX					
		6	Nup85_spoke_1_yc_inner_cr	F	467	1-467	-	100.00 / 100.00	Multiscale: Coarse-grained: 3 - 10 residue(s) per bead
				O					
				X					
				AG					
				BP					
				BY					
				CH					
				CQ					
				DZ					
				EI					
				ER					
				FA					
				GJ					
				GS					
				HB					
				HK					
				IT					
				JC					
				JL					
				JU					
				LD					
				LM					
				LV					
				ME					
				NN					
				NW					
				OF					
				OO					
				PX					
				QG					
				QP					
				QY					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
		7	Nup43_spoke_1_yc_inner_cr	G	377	1-377	-	100.00 / 100.00	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
				P					
				Y					
				AH					
				BQ					
				BZ					
				CI					
				CR					
				EA					
				EJ					
				ES					
				FB					
				GK					
				GT					
				HC					
				HL					
				IU					
				JD					
				JM					
				JV					
				LE					
				LN					
				LW					
				MF					
				NO					
				NX					
				OG					
				OP					
				PY					
				QH					
				QQ					
				QZ					
		8	Nup160_spoke_1_yc_inner_cr	H	1161	1-1161	-	100.00 / 100.00	Multiscale: Coarse-grained: 1 - 10 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
				Q					
				Z					
				AI					
				BR					
				CA					
				CJ					
				CS					
				EB					
				EK					
				ET					
				FC					
				GL					
				GU					
				HD					
				HM					
				IV					
				JE					
				JN					
				JW					
				LF					
				LO					
				LX					
				MG					
				NP					
				NY					
				OH					
				OQ					
				PZ					
				QI					
				QR					
				RA					
		9	Nup37_spoke_1_yc_inner_cr	I	318	1-318	-	100.00 / 100.00	Multiscale: Coarse-grained: 8 - 10 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
				R					
				AA					
				AJ					
				BS					
				CB					
				CK					
				CT					
				EC					
				EL					
				EU					
				FD					
				GM					
				GV					
				HE					
				HN					
				IW					
				JF					
				JO					
				JX					
				LG					
				LP					
				LY					
				MH					
				NQ					
				NZ					
				OI					
				OR					
				QA					
				QJ					
				QS					
				RB					
		10	Nup93_spoke_1_ir_core_1	AK	815	1-815	-	100.00 / 100.00	Multiscale: Coarse-grained: 1 - 10 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
				AN					
				AQ					
				AT					
				CU					
				CX					
				DA					
				DD					
				FE					
				FH					
				FK					
				FN					
				HO					
				HR					
				HU					
				HX					
				JY					
				KB					
				KE					
				KH					
				MI					
				ML					
				MO					
				MR					
				OS					
				OV					
				OY					
				PB					
				RC					
				RF					
				RI					
				RL					
		11	Nup205_spoke_1_ir_core_1	AL	1684	1-1684	-	100.00 / 100.00	Multiscale: Coarse-grained: 1 - 10 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
				AR					
				CV					
				DB					
				FF					
				FL					
				HP					
				HV					
				JZ					
				KF					
				MJ					
				MP					
				OT					
				OZ					
				RD					
				RJ					
		12	Nup155_spoke_1_ir_core_1	AM	1356	1-1356	-	100.00 / 100.00	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
				AP					
				AS					
				AV					
				BI					
				BJ					
				CW					
				CZ					
				DC					
				DF					
				DS					
				DT					
				FG					
				FJ					
				FM					
				FP					
				GC					
				GD					
				HQ					
				HT					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
				HW					
				HZ					
				IM					
				IN					
				KA					
				KD					
				KG					
				KJ					
				KW					
				KX					
				MK					
				MN					
				MQ					
				MT					
				NG					
				NH					
				OU					
				OX					
				PA					
				PD					
				PQ					
				PR					
				RE					
				RH					
				RK					
				RN					
				SA					
				SB					
		13	Nup188_spoke_1_ir_core_2	AO	1564	1-1564	-	100.00 / 100.00	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
				AU					
				CY					
				DE					
				FI					
				FO					
				HS					
				HY					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
				KC					
				KI					
				MM					
				MS					
				OW					
				PC					
				RG					
				RM					
		14	p54_spoke_1_ir_chan_1	AW	366	1-366	-	100.00 / 100.00	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
				AZ					
				BC					
				BF					
				DG					
				DJ					
				DM					
				DP					
				FQ					
				FT					
				FW					
				FZ					
				IA					
				ID					
				IG					
				IJ					
				KK					
				KN					
				KQ					
				KT					
				MU					
				MX					
				NA					
				ND					
				PE					
				PH					
				PK					
				PN					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
				RO					
				RR					
				RU					
				RX					
		15	p58_spoke_1_ir_chan_1	AX	171	1-171	-	100.00 / 100.00	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
				BA					
				BD					
				BG					
				DH					
				DK					
				DN					
				DQ					
				FR					
				FU					
				FX					
				GA					
				IB					
				IE					
				IH					
				IK					
				KL					
				KO					
				KR					
				KU					
				MV					
				MY					
				NB					
				NE					
				PF					
				PI					
				PL					
				PO					
				RP					
				RS					
				RV					
				RY					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
		16	p62_spoke_1_ir_chan_1	AY	169	1-169	-	100.00 / 100.00	Multiscale: Coarse-grained: 9 - 10 residue(s) per bead
				BB					
				BE					
				BH					
				DI					
				DL					
				DO					
				DR					
				FS					
				FV					
				FY					
				GB					
				IC					
				IF					
				II					
				IL					
				KM					
				KP					
				KS					
				KV					
				MW					
				MZ					
				NC					
				NF					
				PG					
				PJ					
				PM					
				PP					
				RQ					
				RT					
				RW					
				RZ					

Datasets used for modeling

There are 11 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
6	3DEM volume	Zenodo	10.5281/zenodo.11129725
7	3DEM volume	Zenodo	10.5281/zenodo.11129725
8	3DEM volume	Zenodo	10.5281/zenodo.11129725
9	3DEM volume	Zenodo	10.5281/zenodo.11129725
10	3DEM volume	Zenodo	10.5281/zenodo.11129725
11	3DEM volume	Zenodo	10.5281/zenodo.11129725
3	Other	Zenodo	10.5281/zenodo.11129725
1	3DEM volume	EMDB	EMD-3820
2	Other	Other	idr0115
4	Experimental model	PDB	5a9q
5	Experimental model	PDB	5ijo

Methodology and software

This entry is a result of 6 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	Monte Carlo sampling step 1. Repeated for each snapshot model.	MC sampling	None	1600000	False	False
2	1	Choose the lowest scoring structure from each simulation.	Filtering	None	200	False	False
3	1	Monte Carlo sampling step 2. Repeated for each snapshot model.	MC sampling	None	320000	False	False
4	1	Choose the lowest scoring structure from each simulation.	Filtering	None	200	False	False
5	1	Monte Carlo sampling step 3. Repeated for each snapshot model.	MC sampling	None	320000	False	False
6	1	Choose the lowest scoring structure from each replica, and further filter by the median score.	Filtering	None	801	False	False
7	1	Score trajectories based on the scores of snapshot models and the transitions between them.	Trajectory construction	None	None	True	False
1	2	Monte Carlo sampling step 1. Repeated for each snapshot model.	MC sampling	None	1600000	False	False
2	2	Choose the lowest scoring structure from each simulation.	Filtering	None	200	False	False

Step number	Protocol ID	Method name	Method type	Method description	Number of computed models	Multi state modeling	Multi scale modeling
3	2	Monte Carlo sampling step 2. Repeated for each snapshot model.	MC sampling	None	320000	False	False
4	2	Choose the lowest scoring structure from each simulation.	Filtering	None	200	False	False
5	2	Monte Carlo sampling step 3. Repeated for each snapshot model.	MC sampling	None	320000	False	False
6	2	Choose the lowest scoring structure from each replica, and further filter by the median score.	Filtering	None	801	False	False
7	2	Score trajectories based on the scores of snapshot models and the transitions between them.	Trajectory construction	None	None	True	False
1	3	Monte Carlo sampling step 1. Repeated for each snapshot model.	MC sampling	None	1600000	False	False
2	3	Choose the lowest scoring structure from each simulation.	Filtering	None	200	False	False
3	3	Monte Carlo sampling step 2. Repeated for each snapshot model.	MC sampling	None	320000	False	False
4	3	Choose the lowest scoring structure from each simulation.	Filtering	None	200	False	False
5	3	Monte Carlo sampling step 3. Repeated for each snapshot model.	MC sampling	None	320000	False	False
6	3	Choose the lowest scoring structure from each replica, and further filter by the median score.	Filtering	None	801	False	False
7	3	Score trajectories based on the scores of snapshot models and the transitions between them.	Trajectory construction	None	None	True	False
1	4	Monte Carlo sampling step 1. Repeated for each snapshot model.	MC sampling	None	1600000	False	False
2	4	Choose the lowest scoring structure from each simulation.	Filtering	None	200	False	False
3	4	Monte Carlo sampling step 2. Repeated for each snapshot model.	MC sampling	None	320000	False	False
4	4	Choose the lowest scoring structure from each simulation.	Filtering	None	200	False	False

Step number	Protocol ID	Method name	Method type	Method description	Number of computed models	Multi state modeling	Multi scale modeling
5	4	Monte Carlo sampling step 3. Repeated for each snapshot model.	MC sampling	None	320000	False	False
6	4	Choose the lowest scoring structure from each replica, and further filter by the median score.	Filtering	None	801	False	False
7	4	Score trajectories based on the scores of snapshot models and the transitions between them.	Trajectory construction	None	None	True	False
1	5	Monte Carlo sampling step 1. Repeated for each snapshot model.	MC sampling	None	1600000	False	False
2	5	Choose the lowest scoring structure from each simulation.	Filtering	None	200	False	False
3	5	Monte Carlo sampling step 2. Repeated for each snapshot model.	MC sampling	None	320000	False	False
4	5	Choose the lowest scoring structure from each simulation.	Filtering	None	200	False	False
5	5	Monte Carlo sampling step 3. Repeated for each snapshot model.	MC sampling	None	320000	False	False
6	5	Choose the lowest scoring structure from each replica, and further filter by the median score.	Filtering	None	801	False	False
7	5	Score trajectories based on the scores of snapshot models and the transitions between them.	Trajectory construction	None	None	True	False
1	6	Monte Carlo sampling step 1. Repeated for each snapshot model.	MC sampling	None	1600000	False	False
2	6	Choose the lowest scoring structure from each simulation.	Filtering	None	200	False	False
3	6	Monte Carlo sampling step 2. Repeated for each snapshot model.	MC sampling	None	320000	False	False
4	6	Choose the lowest scoring structure from each simulation.	Filtering	None	200	False	False
5	6	Monte Carlo sampling step 3. Repeated for each snapshot model.	MC sampling	None	320000	False	False

Step number	Protocol ID	Method name	Method type	Method description	Number of computed models	Multi state modeling	Multi scale modeling
6	6	Choose the lowest scoring structure from each replica, and further filter by the median score.	Filtering	None	801	False	False
7	6	Score trajectories based on the scores of snapshot models and the transitions between them.	Trajectory construction	None	None	True	False

There are 2 software packages reported in this entry.

ID	Software name	Software version	Software classification	Software location
2	gmconvert	2022/05/09	data processing	https://pdbj.org/gmfit/doc_gmconvert/README_gmconvert.html
1	Integrative Modeling Platform (IMP)	2.20	model building	https://integrativemodeling.org

Data quality

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.

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	12417636	7624	99.94
2	55582696	15809	99.97
3	158268736	26373	99.98
4	307110936	37257	99.99
5	385961436	41753	99.99
6	520208640	48260	99.99

Fit of model to data used for modeling

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 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.